

УДК 577(075.8)  
ББК 28.071я73  
Ф 59

**Финкельштейн А. В., Птицын О. Б.**  
Ф59 Физика белка: Курс лекций с цветными и стереоскопическими иллюстрациями и задачами. — 3-е изд., испр. и доп. — М.: КДУ, 2012. — 456 с., [32] с. ил.: ил.

ISBN 5-98227-065-2

Физика белка простирается от классификации и принципов устройства белков (мембранных, фибриллярных и, в особенности, водорастворимых глобулярных белков) — до принципов их работы, устройства их активных центров; от элементарных взаимодействий в белках и вокруг них — до структурных переходов в белках, полипептидах и синтетических полимерах.

Этот курс лекций читался студентам МФТИ, а затем, в Учебном Центре Института белка РАН, — студентам и аспирантам Пушкинского университета и Пушкинского филиала МГУ. В настоящее время он, кроме того, читается студентам биологического факультета и факультета биоинженерии и биоинформатики МГУ. Отдельные главы этого курса читались также на множестве отечественных и международных научных школ.

Книга предназначена для широкого круга студентов, аспирантов и научных сотрудников — биологов, медиков, биохимиков, биофизиков, — а также для всех, кто работает с белками или просто размышляет о них.

УДК 577(075.8)  
ББК 28.071я73

ISBN 5-98227-065-2  
© Финкельштейн А. В., Птицын О. Б., 2012  
© Финкельштейн А. В., стереоскопические рисунки, 2005  
© Издательство «КДУ», 2012

## ОГЛАВЛЕНИЕ И КРАТКОЕ СОДЕРЖАНИЕ

Предисловие к четвертому изданию .....	8
Предисловие к третьему изданию .....	9
Предисловие к первому изданию .....	10
ВВЕДЕНИЕ	
Лекция 1 .....	14
Основные функции белков. Аминокислотная последовательность определяет пространственную структуру, пространственная структура — функцию. Обратное — неверно. Глобулярные, фибриллярные и мембранные белки. Первичная, вторичная, третичная, четвертичная структура белка. Биосинтез белка; сворачивание белка <i>in vivo</i> и <i>in vitro</i> . Пост-трансляционные модификации.	
ЭЛЕМЕНТАРНЫЕ ВЗАИМОДЕЙСТВИЯ В БЕЛКАХ И ВОКРУГ	
Лекция 2 .....	25
Стереохимия L-аминокислотных остатков. Валентные связи и углы между ними. Их колебания. Вращение вокруг валентных связей. Пептидная группа. Транс- и цис-пролины.	
Лекция 3 .....	33
Ван-дер-Ваальсово взаимодействие: притяжение на больших расстояниях, отталкивание на малых. Разрешенные конформации аминокислотного остатка (карты Рамачандрана для глицина, аланина, валина, пролина).	
Лекция 4 .....	42
Влияние водного окружения. Водородные связи. Их электрическая природа. Их энергия. Их геометрия в кристаллах. Разболтанность водородных связей в воде. Понятие об энтропии и свободной энергии. Водородные связи в белковой цепи замещают такие же связи этой цепи с водой; в результате водородные связи в белковой цепи приобретают — в водном окружении — энтропийную природу.	
Лекция 5 (сдвоенная) .....	54
Элементы термодинамики. Свободная энергия и химический потенциал. Гидрофобные взаимодействия. Их связь с необходимостью насыщения водородных связей в воде. Доступная воде неполярная поверхность аминокислот и их гидрофобность.	

Лекция 6 (сдвоенная) .....	69	Лекция 13 .....	162
Влияние водного окружения на электростатические взаимодействия. Электрическое поле у поверхности и внутри белка. Диэлектрическая проницаемость. Экранировка зарядов в солевых растворах. Измерение электрических полей в белках при помощи белковой инженерии. Дисульфидные связи. Координационные связи.		Глобулярные белки. Упрощенное представление структур белковых глобул; структурные классы. Строение $\beta$ -белков: $\beta$ -слои, их продольная и перпендикулярная упаковка. Преимущественная антипараллельность $\beta$ -структуры в $\beta$ -белках. Правопропеллерная скрученность $\beta$ -листов. Топология $\beta$ -белков.	
<b>ВТОРИЧНЫЕ СТРУКТУРЫ ПОЛИПЕПТИДНЫХ ЦЕПЕЙ</b>		Лекция 14 .....	177
Лекция 7 .....	87	Строение $\alpha$ -белков. Пучки и слои спиралей. Модель квазисферической глобулы из $\alpha$ -спиралей. Плотная упаковка при контакте $\alpha$ -спиралей. Строение $\alpha/\beta$ -белков: параллельный $\beta$ -слой, прикрытый $\alpha$ -спиралами, и $\alpha/\beta$ -цилиндр. Топология $\beta$ - $\alpha$ - $\beta$ -субъединиц. Строение $\alpha$ + $\beta$ -белков. Отсутствие прямой связи архитектуры белка с его функцией.	
Вторичная структура полипептидов. Спирали: $2_7$ , $3_{10}$ , $\alpha$ , $\pi$ , poly(Pro)II. Антипараллельная и параллельная $\beta$ -структура. $\beta$ -изгибы. Методы экспериментального обнаружения вторичной структуры.		Лекция 15 .....	192
Лекция 8 (сдвоенная) .....	97	Классификация структур белков. Отсутствие наблюдаемой «макроэволюции» укладок белковых цепей — при наблюдаемой «микрэволюции» их структур. Дупликация гена и специализация. Эволюция путем перемешивания доменов. «Стандартные» третичные структуры. Типичность «квазислучайного» чередования аминокислот в первичных структурах глобулярных белков, контраст с периодическими первичными структурами фибриллярных белков и блочными — мембранными белков. Физические принципы строения белковой глобулы. Основные закономерности, наблюдаемые в структурах белковых глобул: наличие отдельно $\alpha$ - и отдельно $\beta$ -слоев; редкость перекрывания петель; редкость параллельности соседних по цепи структурных сегментов; редкость левых $\beta$ - $\alpha$ - $\beta$ -суперспиралей. «Энергетические» и «энтропийные» дефекты редко встречающихся структур и связь этих «дефектов» с относительной редкостью аминокислотных последовательностей, стабилизирующих «дефектные» структуры. «Принцип множественности».	
Элементы статистической физики. Связь температуры с изменением энергии и энтропии. Вероятности состояний с различной энергией (распределение Больцмана–Гиббса). Статистическая сумма и ее связь со свободной энергией. Конформационные превращения. Понятие о фазовом переходе первого рода (переходе «все-или-ничего») и о не-фазовых переходах. Кинетика преодоления свободно-энергетического барьера при конформационных превращениях. Понятие о теории скоростей реакций. Параллельные и последовательные процессы. Характерные скорости диффузионных процессов.		Лекция 16 .....	208
Лекция 9 (сдвоенная) .....	117	Какую вторичную структуру можно ожидать для случайных и квазислучайных аминокислотных последовательностей? Доменное строение стабильных пространственных структур длинных квазислучайных последовательностей наиболее вероятно. Квази-Больцмановская статистика мелких деталей белковых структур. Она возникает из физического отбора стабильных белковых структур. Влияние стабильности структурного элемента на строгость отбора первичных структур, не разрушающих пространственную структуру глобулярного белка, или: почему одни белковые структуры встречаются часто, а другие — редко? Какую структуру — $\alpha$ или $\beta$ — следует чаще ожидать в центре большой глобулы? Связь «энтропийных» дефектов с «энергетическими». Глобулярные белки возникли как «отобранные» случайные полипептиды? Отбор «белкоподобных» случайных последовательностей в белковой инженерии.	
Свободная энергия инициации и элонгации $\alpha$ -спирали. Теорема Ландау и не-фазовость перехода спираль-клубок. Размер кооперативного участка при переходе спираль-клубок. Стабильность $\alpha$ -спирали в воде. Стабильность $\beta$ -структуры в воде. Скорость образования $\beta$ -структуры и $\alpha$ -спирали. Что такое «клубок»?		<b>КООПЕРАТИВНЫЕ ПЕРЕХОДЫ В БЕЛКОВЫХ МОЛЕКУЛАХ</b>	
Лекция 10 .....	132	Лекция 17 (сдвоенная) .....	229
Свойства боковых групп аминокислотных остатков. Включение аминокислотных остатков во вторичную структуру. Аланин, глицин, пролин, валин. Неполярные, короткие полярные и длинные полярные боковые группы. Заряженные боковые группы. Гидрофобные поверхности на вторичных структурах в белках.		Денатурация белка. Нативно-развернутые белки. Кооперативные переходы. Обратимость денатурации белков. Денатурация глобулярного	
<b>ПРОСТРАНСТВЕННОЕ СТРОЕНИЕ БЕЛКОВ</b>			
Лекция 11 .....	140		
Фибриллярные белки, их функции и их периодичные первичные и вторичные структуры; $\alpha$ -кератин, $\beta$ -фибрин шелка, коллаген. Упаковка длинных $\alpha$ -спиралей и обширных $\beta$ -листов. Белки, образующие матрикс; эластин. Генетические дефекты белков и болезни. Амилоиды.			
Лекция 12 .....	150		
Мембранные белки, особенности их строения и функции. Бактериородопсин, рецепторы и G-белки, порин, фотосинтетический центр. Селективная проницаемость мембранных пор. Работа фотосинтетического центра. Понятие о туннельном эффекте. Понятие об электронно-конформационном взаимодействии.			

белка — переход типа « <i>все-или-ничего</i> ». Критерий Вант-Гоффа для перехода « <i>все-или-ничего</i> ». Тепловая и холодная денатурация. Диаграмма фазовых состояний белковой молекулы. Как выглядит денатурированный белок? Клубок и расплавленная глобула. Неоднородность расплавленной глобулы. Отсутствие фазового перехода типа « <i>все-или-ничего</i> » при набухании «обычных» полимерных глобул.	
Лекция 18	250
Почему денатурация глобулярного белка — переход типа « <i>все-или-ничего</i> »? Распад плотной упаковки ядра белка и раскрепощение боковых групп. Проникновение растворителя в денатурированный белок, разрушение расплавленной глобулы, постепенное разворачивание цепи денатурированного белка по мере увеличения силы растворителя. Энергетическая щель между нативной укладкой белковой цепи и прочими ее глобулярными укладками: основное физическое отличие белковой цепи от случайного сополимера. Различия в плавлении «отобранного» гетерополимера (с энергетической щелью) и случайного сополимера.	
Лекция 19	263
Образование структуры белка <i>in vivo</i> и <i>in vitro</i> . Вспомогательные механизмы при самоорганизации <i>in vivo</i> : котрансляционное сворачивание, шапероны, и т. д. Спонтанная самоорганизация возможна <i>in vitro</i> . «Парадокс Левинталя». Опыты по сворачиванию белка в бесклеточных системах, а также — о разном понимании слов <i>in vitro</i> . Стадийный механизм самоорганизации белков. Обнаружение метастабильных (накапливающихся) интермедиатов сворачивания многих белков. Расплавленная глобула — обычно (но не обязательно) наблюдаемый интермедиат сворачивания белка в нативных условиях. Простейшее (одностадийное) сворачивание некоторых белков — без каких-либо накапливающихся метастабильных интермедиатов. Самоорганизация мембранных белков.	
Лекция 20	279
Одностадийное сворачивание малых белков. Теория переходных состояний. Экспериментальный поиск и изучение нестабильных переходных состояний в сворачивании белка. Ядро сворачивания нативной структуры белка. Его экспериментальное обнаружение <i>in vitro</i> методами белковой инженерии. Нуклеационный механизм сворачивания белка.	
Лекция 21	293
Решение «парадокса Левинталя»: к стабильной структуре цепи автоматически ведет сеть быстрых путей сворачивания. Для этого необходимо только, чтобы между нативной укладкой цепи и прочими ее глобулярными укладками существовала бы заметная энергетическая щель. Обсуждение аномально медленного образования стабильной структуры в некоторых белках (серпины, прионы). Представление об «энергетических ландшафтах» белковых цепей. Белковые структуры: физика самоорганизации и естественный отбор самоорганизующихся цепей.	
ПРЕДСКАЗАНИЕ И ДИЗАЙН БЕЛКОВЫХ СТРУКТУР	
Лекция 22	315
Потребность в предсказании структур белков по их аминокислотным последовательностям. «Опознавание» белковых структур и функций	
по гомологии последовательностей. Профили первичных структур семейств белков. Ключевые районы и функциональные сайты белковых структур. Выделение стабильных пространственных структур белковой цепи. «Шаблоны» белковых структур. Мы всегда вынуждены судить о предсказываемой структуре белка только по части взаимодействий, действующих в его цепи. Результат: вероятностные предсказания. Взаимодействия, стабилизирующие и разрушающие вторичную структуру полипептидов. Расчет вторичных структур не-глобулярных полипептидов. Предсказание вторичной структуры белков.	
Лекция 23	332
Представление о подходах к предсказанию пространственных структур белков по их аминокислотным последовательностям. Базы данных по структурам белков. Предсказание общей укладки цепи отдаленных гомологов понижает неопределенности в опознавании структур белков. Структурная геномика и протеомика. Биоинформатика. Белковая инженерия и дизайн. Первые успехи в конструировании белков.	
ФИЗИЧЕСКИЕ ОСНОВЫ ФУНКЦИОНИРОВАНИЯ БЕЛКОВ	
Лекция 24 (сдвоенная)	351
Функция белка и его структура. Элементарные функции. Связывающие белки: ДНК-связывающие белки, иммуноглобулины. Ферменты. Активный центр — «дефект» глобулярной структуры. Твердость белка важна для элементарной ферментативной функции. Каталитический и субстрат-связывающий центры. Ингибирование. Кофакторы. Многовалентные ионы. Механизм ферментативного катализа. Пример: сериновые протеазы. Теория переходного состояния в катализе и ее подтверждение методами белковой инженерии. Абзимы. Специфичность катализа. Узнавание «ключ-замок».	
Лекция 25	371
Сочетание функций. Переход субстрата с одного на другой активный центр. «Двойное сито» повышает специфичность функции. Относительная независимость структуры белка от его элементарной ферментативной активности. Заметная связь структуры с окружением белка. Сопряжение элементарных функций белка и гибкость его структуры. Индуцированное соответствие. Подвижность доменов белка. Перемещение доменов при эволюции белков. Доменная структура: киназы, дегидрогеназы. Аллостерия — взаимодействие активных центров. Аллостерическая регулировка функции белка. Аллостерия и четвертичная структура белка. Гемоглобин и миоглобин. Механизм мышечного сокращения. Шагающий кинезин	
Послесловие	391
Рекомендованная литература	392
Задачи, решения, комментарии	394
Предметный указатель	482

## ПРЕДИСЛОВИЕ К ЧЕТВЕРТОМУ ИЗДАНИЮ

Заново редактируя курс этих лекций, я постарался включить в них все самые яркие результаты, появившиеся в физике белка за последние 6 лет — со времени последнего русского издания, — а также существенно расширил число предлагающихся для самостоятельного решения задач.

*А. В. Финкельштейн, 2011*

## ПРЕДИСЛОВИЕ К ТРЕТЬЕМУ ИЗДАНИЮ

Новое издание «Физики белка» готовится к печати в год 75-летия со дня рождения О. Б. Птицына и 5-летия со дня его безвременной кончины.

Подготавливая к печати это издание, я, во-первых, исправил выявленные опечатки, неточности и неясности в тексте, в некоторых таблицах и рисунках. Во-вторых, я ввел новые данные в лекции 11, 18–22. В-третьих, с помощью А. К. Цатуряна и С. Ю. Бершицкого, которым я глубоко благодарен, я заново переписал тот раздел лекции 25, что касается работы мышц. И, наконец, я ввел в новое издание книги некоторое количество задач для самостоятельного решения.

Я благодарю читателей русского и английского изданий «Физики белка» и слушателей моих лекций, высказавших или приславших мне свои замечания, инициировавшие исправление этих лекций. В особенности я благодарен за замечания и за предоставление мне дополнительных материалов Д. Н. Иванкову, Е. С. Надеждиной, В. Е. Бычковой, О. В. Галзитской, С. О. Гарбузинскому, Н. С. Богатыревой, Г. В. Семисотнову, Н. Ю. Марченко, Е. Н. Барышниковой, М. Г. Шарапову, Л. Б. Переяславцу, Ф. К. Гиеовой, А. Я. Бадретдинову, А. А. Климову, Д. А. Климову, В. Е. Финкельштейну, А. Александреску, А. Кистеру, Д. Бейкеру и К. Плакско.

Я благодарен за обсуждение затронутых в лекциях вопросов А. А. Веденову, А. М. Гутину, Е. И. Шахновичу, Б. А. Реве, М. Д. Франк-Каменецкому, А. М. Дыхне, С. Чотиа, М. С. Гельфанду, И. М. Гельфанду, К. Добсону, А. Фершту, Г. Фринду, А. Трамонтано и А. Леску.

Я благодарен М. Г. Никитиной за помощь, оказанную мне в ходе работы над новым изданием книги, а также моей жене М. С. Вильнер-Мармер, без постоянной поддержки, внимания и содействия которой подготовка этой книги никогда не была бы доведена до конца.

И, наконец, я благодарен Грантам Москвы, фонду ISSEP, Учебному центру Института белка РАН и фирме «Алгодэйн» за поддержку моей преподавательской деятельности, а Российской академии наук, Программе по молекулярной и клеточной биологии РАН, Программе «Ведущие научные школы», Российскому фонду фундаментальных исследований и Медицинскому институту Ховарда Хьюза за постоянную поддержку научной работы нашей лаборатории, основанной О. Б. Птицыным.

*А. В. Финкельштейн, 2004*

## ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

Предлагаемый вашему вниманию курс лекций посвящен **физике** белка, т. е. самым общим проблемам структуры, самоорганизации и функционирования белковых молекул.

Он основан на лекциях, которые читались нами (сначала — О.Б.П., а позже — А.В.Ф.) первоначально в московском Физтехе, а затем — в Пушкинском университете и в Пушкинском филиале МГУ. Первоначально лекции читались для физиков, а потом — в основном для биологов (и отчасти — для химиков). Поэтому за последнее время эти лекции пришлось не только сильно обновить (наука не стоит на месте), но и кардинально переработать — для биологической аудитории.

Поскольку то, что вы видите перед собой — запись курса лекций, а не монография, здесь неизбежны повторы (и в частности — повторы рисунков). Ведь в реальных лекциях невозможны отсылки типа «смотри рисунок 2 и формулу 3 в пред-предыдущей лекции»; однако мы приложили все старания к тому, чтобы такие повторы были минимальны.

Мы приносим свою глубокую благодарность всем соавторам наших работ, использованных в предлагаемых вашему вниманию лекциях, а также — А. Б. Четверину, Д. И. Харакозу, В. Е. Бычковой, Ю. В. Митину, А. С. Спирину, В. А. Колбу, М. А. Ройтбергу, В. В. Велькову, А. В. Ефимову, И. Г. Птициной, Г. И. Гительзону, Д. С. Рыкунову, Г. В. Семисотнову, А. В. Скугареву и Д. Н. Иванкову за чтение рукописи, дискуссии и многочисленные полезные замечания.

Мы благодарны всем сотрудникам нашего Института белка РАН, и особенно — всем, кто работает и работал в нашей Лаборатории физики. Без постоянного общения с ними, научного и просто человеческого, эта книга никогда бы не была написана.

Мы признательны всем слушателям этих лекций — как нашим студентам, так и участникам международных и российских школ — чьи вопросы и замечания позволили авторам уточнить, прежде всего для самих себя, многие из затрагиваемых в этих лекциях вопросов.

Наша особая благодарность — Г. А. Морозову, Д. С. Рыкунову, А. В. Скугареву, Д. Н. Иванкову, М. Ю. Лобанову, Н. Ю. Марченко, И. В. Соколовскому, А. А. Шумилину, М. Г. Дашкевичу, Т. Ю. Сальниковой, Н. С. Богатыревой, А. Э. Жумаеву и А. А. Финкельштейну за помощь в оформлении рисунков и текста лекций, а также — Российскому фонду фундаментальных исследований за финансовую поддержку издания этой книги.

### Предварительные и полуфилософские замечания

*О «лекторе».* Все лекции написаны так, как они читаются в аудитории: от первого лица — от лица «лектора».

*О «внутреннем голосе».* Личность «лектора» объединяет в одном лице обоих соавторов — А.В.Ф. и О.Б.П. Однако это не значит, что соавторы были неизменно согласны друг с другом по всем излагаемым пунктам. Более того, каждый из соавторов порой чувствовал, что лекция касается спорных, далеко не до конца выясненных вопросов. Мы не стали заглушать все эти противоречия и споры, и потому плавный рассказ «лектора» порой прерывает или оспаривает его «внутренний голос». А порой «внутренний голос» просто озвучивает часто задававшиеся нам вопросы и уточняет или развивает рассказ «лектора».

*Почему — «физика белка»? Потому, что нас поражает картина того, как биологическая эволюция многократно усиливает, закрепляет и предъявляет нам последствия тех физических принципов, на которых основываются молекулярные взаимодействия в белке. Нас поражает также то, как сильно наше понимание основных биологических, и, в частности, белковых систем было продвинуто использованием физических методик. Мы можем проследить это от масс-спектрометрии до электронной микроскопии, до рентгена и ядерного магнитного резонанса белков. Наверно, нет ни одной другой области современной науки, где традиционные границы между науками и их философиями были бы взломаны с такой большой пользой.*

*О физике и о биологии в этих лекциях.* По ходу курса мы не упустим возможности изложить те физические идеи и, в частности, те элементы статистической физики и квантовой механики, которые, с нашей точки зрения, абсолютно необходимы не только для понимания строения и функционирования белков, но и для общей научной культуры — и которые «нормальный» студент-биолог либо уже прочно забыл ко времени начала специализации, либо не знал никогда.

А вот из моря информации о функциях белков мы зачерпнем только то, что абсолютно необходимо для иллюстрации роли, играемой пространственными структурами белков в их биологической — точнее, биохимической — деятельности.

*Об эксперименте, физической теории и о расчетах.* Эксперимент дает как основные факты, на которых основываются все наши представления о явлениях, так и массу уточняющих их деталей. Теории позволяют понять суть явлений и связь между ними, помогают планировать содержательные эксперименты. Расчеты позволяют связать теорию с экспериментом и подвергнуть ее проверке в ее ключевых точках. Надо заметить, однако, что не все, что можно рассчитать, следует рассчитывать: так, плотность белка —



или воды — проще просто померить. И не только проще — но и надежнее: детальный расчет зависит от массы параметров, едва ли поддающихся очень точной оценке.

Рассказ о нескольких базовых физических теориях — естественно, в самой упрощенной форме — также включен в наши лекции. И не только потому, что эти теории позволяют с единой точки зрения охватить, упорядочить и понять огромный экспериментальный материал. Не менее важно то, что эти теории красивы. Кроме того, мы считаем, что знание базовых физических теорий и моделей необходимо просто для общей культуры.

*Об эксперименте “in vivo” и “in vitro”.* Говоря об эксперименте “in vivo” и “in vitro”, физики и биологи часто имеют в виду разные вещи. И, вообще говоря, не без оснований: между чистым “in vivo” и чистым “in vitro” есть много допускающих двоякое толкование ступеней. Например, сворачивание белка в бесклеточной системе (со всеми ее рибосомами, факторами инициации и терминации, шаперонами и т. д.) — это столь же явно эксперимент “in vivo” с точки зрения физика (для которого “in vitro” — это отдельный белок в растворе; а бесклеточная система... — уж больно много здесь жизненных реалий!), сколь явно он является экспериментом “in vitro” с точки зрения биолога (для которого “in vivo” — это в живом и желателно неповрежденном организме). Однако структурное, например, исследование отдельного белка в организме практически невозможно. Так что на практике всегда идут на разумный компромисс: на приближение биологически интересного явления *in vivo* доступным экспериментальному наблюдению явлением *in vitro*.

*О физических моделях, грубых оценках и о компьютерном (in silicio) эксперименте.* По ходу лекций мы часто будем рассматривать простые, т. е. резко упрощенные по сравнению с реальностью модели, и прибегать к прикидочным оценкам. И мы хотели бы, чтобы, проштудировав эти лекции, вы тоже научились и делать прикидочные оценки, и использовать простые модели интересующих вас событий. Казалось бы, зачем все это вообще нужно? Многие думают, что, имея мощный компьютер, можно ввести в него «все как есть»: молекулы воды, соль, координаты атомов белка, ДНК, и так далее, задать температуру, — и «все посчитать точно». На самом деле эта картина утопична. Расчет — имеется в виду детальный расчет (обычно он ведется при помощи так называемой «молекулярной динамики»; квантово-механический расчет целого белка и его окружения вообще невозможен) — займет дни и охватит при этом лишь какую-то наносекунду из жизни белка: ведь вам надо будет рассмотреть и тепловые движения, и взаимное влияние десятков тысяч атомов. И, кстати, совсем точным этот расчет все равно не будет: все элементарные взаимодействия мы знаем лишь приблизительно.

А чем подробнее описывается система, чем больше элементарных взаимодействий принимается во внимание, тем больше мелких погрешностей проникает в наши расчеты (не говоря уже о росте затраченного времени). Так что все равно получится не окончательная «истина в последней инстанции», а какая-то более или менее приближенная оценка, — оценка, на которую уйдут дни работы суперкомпьютера. А ведь вас, скорее всего, будет интересовать простая, но быстрая прикидка типа: можно ли ввести в такое-то место белка заряд, или белок при этом взорвется? Так что одна из наших целей — научить вас делать такие прикидки.

Однако это вовсе не означает, что мы будем игнорировать компьютерные эксперименты. Такие эксперименты дают массу полезной информации. Но компьютерный эксперимент — это именно эксперимент (пусть *in silicio*, а не *in vitro* или *in vivo*), причем с очень сложными системами: он поставляет факты, нуждающиеся в дальнейшей интерпретации, и вот тут на помощь приходят упрощенные, но прозрачные теории.

*О том, как относиться к формулам.* Понимая, что математические формулы сложны для биологов, мы постарались по возможности избегать их — и оставили их лишь там, где они действительно нужны. И хотим дать совет: при чтении «поверяйте слова формулами!» Конечно, проще читать только слова, но они часто допускают неоднозначное толкование; так что, поверяя слова формулами и наоборот, вы сами проверите свое понимание.

*О мелком шрифте.* Им даны полезные, но не необходимые отступления, добавления и пояснения.

*О вкусах.* Эти лекции, безусловно, отражают наши личные вкусы и пристрастия. Они направлены более на понимание сути явлений, чем на скрупулезное описание их деталей; в них много теории и физических проблем — и лишь необходимый минимум экспериментальных фактов (и почти нет экспериментальной техники; в частности, мы практически ничего не расскажем о технике рентгеновской кристаллографии и спектроскопии ЯМР — методов, которым мы обязаны львиной долей наших знаний о структурах белков). Поэтому наши лекции никак не заменяют обычные, богатые множеством фактов биофизические и биохимические «белковые» курсы и не должны использоваться как справочник. Говоря о конкретных белках, мы даем лишь важнейшие (с нашей точки зрения!) примеры; таблицы включают лишь абсолютно необходимые данные; рисунки намеренно схематичны; все цифры округлены, и т. д.

*О личном элементе в этих лекциях.* Мы не упустим возможности отметить то, что сделали по физике белка мы и наши коллеги в Институте белка РАН. Это, конечно, усилит «личный элемент» лекций — и, мы надеемся, немного оживит их...

## ВВЕДЕНИЕ

### Лекция 1

Основные функции белков. Аминокислотная последовательность определяет пространственную структуру, пространственная структура — функцию. Обратное — неверно. Глобулярные, фибриллярные и мембранные белки. Первичная, вторичная, третичная, четвертичная структура белка. Биосинтез белка; сворачивание белка *in vivo* и *in vitro*. Посттрансляционные модификации.

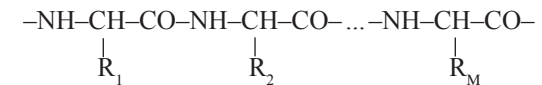
Белки — это и молекулярные машины, и строительные блоки, и оружие живой клетки. Важнейшая и почти монополярная функция белков — ферментативный катализ химических превращений в клетке и вокруг нее. Кроме того, белки-регуляторы контролируют экспрессию генов, а рецепторные белки (они сидят в липидной мембране) обеспечивают восприятие межклеточных сигналов, часто передаваемых гормонами (тоже белками). Иммунные белки и сходные с ними белки гистосовместимости опознают и связывают «чужие» молекулы, а также «свои» клетки, что и позволяет им (клеткам) правильно размещаться в организме. Структурные белки образуют микрофиламенты, микротрубочки, а также фибриллы, волосы, шелк и другие защитные покровы; они армируют мембраны и поддерживают структуру клеток и тканей. Транспортные белки переносят (а запасные — запасают) другие молекулы. Белки, переносящие протоны и электроны через мембрану, обеспечивают всю биоэнергетику: поглощение света, дыхание, выработку АТФ. Другие белки, «сжигая» АТФ, обеспечивают механохимическую деятельность — они работают в мышцах или перемещают элементы клетки.

При всем разнообразии, работа белков всегда базируется на высоко специфическом — как у ключа с замком (точнее: как у гибкого ключа с гибким замком) — взаимодействии белка с обрабатываемой им молекулой. Для этого взаимодействия необходима достаточно «твердая» (во вся-

ком случае, у «работающего белка») пространственная структура. Поэтому биологическая функция белков (как и других важнейших для жизни макромолекул — ДНК и РНК) тесно связана с определенностью их трехмерных структур. Не только разрушение — даже небольшие изменения этих структур часто ведут к утере или резкому изменению активности белков.

Знание молекулярной трехмерной структуры белка необходимо для понимания функционирования белковой молекулы. Поэтому я буду говорить о физике функционирования белков лишь в конце, сосредоточившись вначале на их структуре, на природе ее стабильности и способности к самоорганизации.

- В основе белка лежит полимер — состоящая из аминокислотных остатков полипептидная цепь; это было установлено Э. Фишером в начале XX века. В начале 1950-х годов Сэнгер показал, что белковая цепь имеет уникальную последовательность звеньев — аминокислотных остатков («остаток» — это то, что осталось от свободной аминокислоты после ее встраивания в белковую цепь). Эта цепь имеет химически регулярный остов («главную цепь»), от которого отходят разнообразные боковые группы аминокислот — радикалы  $R_1, R_2, \dots, R_M$



Число звеньев белковой цепи (M) кодируется геном и составляет порой всего несколько десятков, а порой — многие тысячи.

Есть двадцать основных аминокислотных остатков. Их положение в белковой цепи кодируется генами. Однако последующая модификация белка иногда увеличивает разнообразие аминокислот. Кроме того, в некоторые белки включаются разные другие молекулы (*кофакторы*).

В «работающем» белке его цепь свернута строго определенным образом. В конце 1950-х Перутц и Кендрью расшифровали пространственное строение первых белков и показали высокую сложность и уникальность их строения. Впервые же (как потом выяснилось!) строгую определенность строения белковых молекул фактически доказал еще в 60-х годах XIX века Хоппе-Зейлер, получив кристаллы гемоглобина: ведь в кристаллах каждый атом каждой молекулы знает свое место.

Вопрос о том, одинакова ли структура белка в кристалле и в растворе, обсуждался долгие годы, пока об этом свидетельствовали только косвенные данные, однако в конце концов ЯМР (ядерный магнитный резонанс) показал, что обычно она практически одна и та же.

Белки «живут» в разном окружении, и их структура несет на себе явную печать этого окружения. Чем меньше воды вокруг белка, тем невосполнимее разрыв мощных водородных связей, стягивающих его цепь, тем ценнее для белка эти связи (а именно они крепят структуру остова белковой молекулы), тем регулярнее вынуждена быть стабильная структура белка.

Грубо говоря, по «жизненным условиям» и общему типу строения белки можно разбить на три класса.

1) Фибриллярные белки образуют огромные агрегаты, в которых воды довольно мало; их структура обычно высоко регулярна и держится в основном взаимодействиями между разными цепями.

2) Мембранные белки «живут» в мембране, где нет воды, но части их выступают из мембраны в воду. Внутримембранные части таких белков — как и фибриллярные белки — высоко регулярны и прошиты водородными связями, но размер этих регулярных частей ограничен толщиной мембраны.

3) Водорастворимые, живущие в воде глобулярные белки наименее регулярны (особенно небольшие); их структура держится взаимодействиями белковой цепи с самой собой, причем особенно важны взаимодействия далеких по цепи, но сблизившихся в пространстве углеводородных («гидрофобных», т. е. «боящихся воды») групп, а порой также и взаимодействиями белковой цепи с кофакторами.

Наконец, небольшие или бедные углеводородными группами полипептиды могут сами по себе не иметь фиксированной структуры, но приобретать ее при взаимодействии с другими макромолекулами.

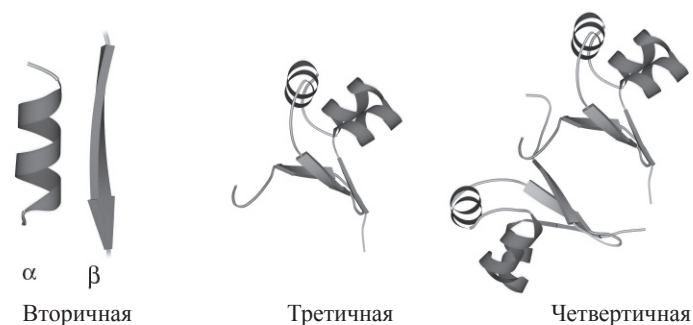
Конечно, намеченное выше деление очень грубо. Порой белок может состоять из фибриллярного «хвоста» и глобулярной «головки» (так устроен, например, миозин), и так далее.

Уже (на 2011 г.) известны многие миллионы белковых аминокислотных последовательностей (для их хранения созданы специальные компьютерные банки — например, SwissProt) и порядка сотни тысяч пространственных структур белков (они хранятся в компьютерном Банке Белковых Структур: Protein Data Bank, или просто PDB). Подавляющая часть того, что мы знаем о трехмерных белковых структурах, относится к водорастворимым глобулярным белкам. Для мембранных же и фибриллярных белков расшифровано лишь относительно немного пространственных структур или их отдельные фрагменты. Причина проста: водорастворимые белки легче выделять в виде отдельных молекул и их структуру легче изучать и рентгеном — в кристаллах, и спектроскопией ЯМР — в растворах. Поэтому, говоря о «структуре белка», «формировании структуры белка» и т. д. — часто, на самом деле, имеют в виду закономерности, доказанные лишь для водорастворимых глобулярных белков. Это надо иметь в виду,

читая книги и статьи, да и эти лекции тоже. Более того, надо иметь в виду, что, по тем же экспериментальным причинам, вся физика белков сейчас является, в основном, физикой небольших белков, а физика крупных белковых молекул и их комплексов только начинает развиваться.

■ Нековалентные взаимодействия, поддерживающие пространственное строение белка, значительно слабее химических связей, фиксирующих последовательность мономеров — аминокислот — в белковой цепи. Эта последовательность — она называется «первичной структурой белка» (рис. 1-1) — создается в ходе матричного биохимического синтеза согласно «инструкции», записанной в гене.

Первичная ...-Gly-Val-Tyr-Gln-Ser-Ala-Ile-Asn-Lys-Ala-...



**Рис. 1-1.** Уровни организации белковой структуры: первичная структура (аминокислотная последовательность); вторичная структура (показаны: α-спираль и один тяж β-структуры); третичная структура глобулы, сложенной одной цепью; и четвертичная структура олигомерного, сложенного из нескольких цепей белка (в данном случае — димерного *cro*-репрессора)

Архитектуры белков, особенно глобулярных водорастворимых белков, сложны и разнообразны — в противоположность универсальности двойной спирали ДНК (кстати, одноцепочечные РНК здесь как бы занимают промежуточное положение). Однако и в белках прослеживается набор «стандартных» структур, о которых мы будем говорить весьма подробно.

Прежде всего здесь речь идет о регулярных вторичных структурах белка — об α-спирали и β-структуре; α-спирали часто изображаются спиральными лентами (см. рис. 1-1) или цилиндрами, а вытянутые β-структурные участки (слипаясь, они образуют листы) — стрелками (см. рис. 1-1). Вторичные структуры отличаются регулярной, периодической формой (или, как говорят, «конформацией») главной цепи — при разнообразии конформаций боковых групп.



Укладка вторичных структур одной полипептидной цепи в глобулу называется третичной структурой (см. рис. 1-1). Объединение же нескольких белковых цепей в «суперглобулу» называется четвертичной структурой белка (см. рис. 1-1). Пример (в дополнение к приведенному на рис. 1-1 димеру белка *cro*-репрессора): гемоглобин состоит из двух  $\beta$ - и двух  $\alpha$ -цепей (к  $\alpha$ - и  $\beta$ -структурам это *никакого* отношения не имеет!). Если четвертичная структура белка сложена из одинаковых цепей, то она обычно симметрична (и *cro*-репрессор, и гемоглобин здесь не исключение). Иногда четвертичная структура объединяет десятки цепей. В частности, вирусные оболочки могут рассматриваться как такие «суперчетвертичные» структуры.

Среди третичных структур есть наиболее типичные, с которыми нам предстоит еще познакомиться. Часто такие типичные структуры охватывают не весь глобулярный белок целиком, а лишь его компактную субглобулу — «домен» (рис. 1-2). Домен (подобно малому глобулярному белку) обычно состоит из 100–200 аминокислотных остатков, т. е. из ~2000 атомов. Его диаметр — 30–40 Å.



**Рис. 1-2.** Доменная структура большого белка сходна с четвертичной структурой, сложенной из малых белков. Разница лишь в том, что все компактные субглобулы (домены) большого белка в данном случае образованы одной цепью, а четвертичная структура складывается из нескольких цепей

- Образование трехмерной структуры белка *in vivo* происходит при биосинтезе или сразу после него. Замечательно, однако, что оно может происходить не только при биосинтезе: около 1960 г. Anfinsen показал, что оно может идти и при «ренатурации» — сворачивании чем-то развернутой белковой цепи *in vitro*; причем идти совершенно спонтанно.

Это значит, что аминокислотная последовательность сама (при подходящей температуре и pH воды!) определяет пространственную структуру белка — т. е. белок способен к *самоорганизации*.

**Внутренний голос:** Строго говоря, это показано для, в основном, относительно небольших (включающих до 200–300 аминокислотных остатков)

водорастворимых глобулярных белков. С более крупными глобулярными белками сложнее, особенно с белками высших организмов: далеко не все они ренатурируют спонтанно...

**Лектор:** Благодарю за уточнение. Да, с такими белками сложнее. Отчасти здесь дело в агрегации; отчасти — в посттрансляционной модификации, особенно у эукариот. Еще меньше известно о спонтанной самоорганизации мембранных и фибриллярных белков. В некоторых таких белках она происходит; но обычно ренатурировать их не удастся. Поэтому надо сразу оговориться: говоря о физике белка, о структуре белка и о ее формировании, я, на самом деле (если не сделаю оговорки), обычно буду иметь в виду относительно небольшие глобулярные белки. Это — обычная в биофизической литературе условность, но то, что она недостаточно артикулирована, часто ведет к недоразумениям.

- Опыты Anfinsen принципиально отделили физический процесс — самоорганизацию пространственной структуры — от биохимического синтеза белковой цепи. Стало ясно, что структура белка детерминирована самой его аминокислотной последовательностью, а не навязывается белку клеточной машинерией. По-видимому, основная задача этой машинерии — оберегать сворачивающийся белок от нежелательных контактов (к коим относятся и контакты между отдаленными, образующими разные домены кусками цепи очень больших белков) — ведь *in vivo* белок сворачивается в клеточном бульоне, где так много разных молекул, так много возможностей прилипнуть к чему-нибудь! Но в разбавленном растворе *in vitro* белок — по крайней мере, не очень большой — складывается спонтанно, сам по себе.
- Строго говоря, белок способен к спонтанной самоорганизации и ренатурации, только если он не подвергся сильной посттрансляционной модификации — т. е. если его химическая структура не была сильно нарушена после биосинтеза и первоначальной самоорганизации. Например, инсулин (где половина цепи вырезается уже после того, как он свернулся) не способен к ренатурации.

Посттрансляционные модификации бывают самые разные. Как правило, химические модификации контролируются специальными ферментами, а не «самоорганизуются» в самом белке. Кроме расщепления белковой цепи (протеолиза: он часто способствует превращению зимогена, неактивного «белка-предшественника», в активный энзим; и он же разделяет полибелок на множество отдельных глобул), наблюдается вырезание кусков белковой цепи (например, при изготовлении инсулина из проинсулина; кстати, вырезанные фрагменты порой служат отдельными гормонами).

Наблюдается и модификация концов цепи, ацетилирование, гликозилирование, пришивание липидов в определенные точки цепи, фосфорилирование определенных боковых групп, и т. д., и т. п. Был обнаружен даже «сплайсинг» белковых цепей (спонтанное вырезание куска цепи и склейка образовавшихся при этом концов). Изредка наблюдается и спонтанная циклизация белковой цепи или каких-то ее кусочков.

Особо надо отметить завязывание S-S-связей между серосодержащими остатками цистеина белковой цепи: «правильные» S-S-связи способны (в подходящих условиях *in vitro*) к спонтанной самоорганизации, хотя *in vivo* их образование катализируется специальным ферментом, дисульфидизомеразой. Как правило, эти связи завязываются только в секретированных белках (внутри клетки нет кислорода, и потому — неподходящий окислительный потенциал). При этом S-S-связи, если они завязаны «правильно», отнюдь не мешают, а скорее способствуют ренатурации белка.

А вот «неправильно» завязанные S-S-связи лишают белок способности к ренатурации. Именно поэтому, кстати, сваренное яйцо не «разваривается» при охлаждении. Дело в том, что белки яйца не только денатурируют при нагревании, но еще и перешиваются между собой дополнительными S-S-связями (как резина). При охлаждении эти химические связи сохраняются и не дают белкам яйца вернуться к первоначальному («нативному») состоянию.

■ Итак, аминокислотная последовательность определяет пространственную структуру белка. А структура эта уже определяет его функцию — т. е. с кем этот белок взаимодействует и что при этом делает.

■ Здесь необходимо сделать несколько замечаний.

Первое. Глядя на рис. 1-1, можно подумать, что внутри белка много пустого пространства и, следовательно, что он «мягок». На самом деле это вовсе не так! Белок тверд — белковая цепь упакована плотно, атомы упираются в атомы (см. левую часть рис. 1-3). Но вот только анатомию белка, его скелет, его внутреннее строение неудобно изучать на полной атомной модели — внутреннее устройство белка можно увидеть, лишь сделав атомы «прозрачными» и обратив основное внимание на ход белковой цепи (см. правую часть рис. 1-3 и, в особенности, рис. 1-1, 1-2, где боковые группы «сбиты», а элементы вторичной структуры выделены). На полноатомной же модели белка (см. рис. 1-3) даже не видно, что белок образован полимерной цепью, видна лишь поверхность глобулы, похожая на картофелину. Однако полноатомная модель нужна для изучения функции белка — специфику его работы определяют именно физико-химические и геометрические свойства поверхности глобулы, этой «картофелины», а дело скелета белка — создать и фиксировать эту поверхность.



Рис. 1-3. Атомная модель белковой глобулы (слева) и ее скелетная модель (справа). В скелетной модели боковые группы — тонкие линии, а главная цепь — толстая

Второе. В состав белка помимо самой белковой цепи часто входят *кофакторы* (рис. 1-4), — малые молекулы, ионы, сахара, нуклеотиды, фрагменты нуклеиновых кислот и т. д. Эти не-пептидные молекулы принимают участие в функционировании, а порой и в формировании структуры белковой молекулы. Кофакторы могут присоединяться к белку химическими связями, а могут просто прилипнуть к определенным местам белковой глобулы. Кроме того, к поверхности белка обычно прочно прилипает множество молекул воды.

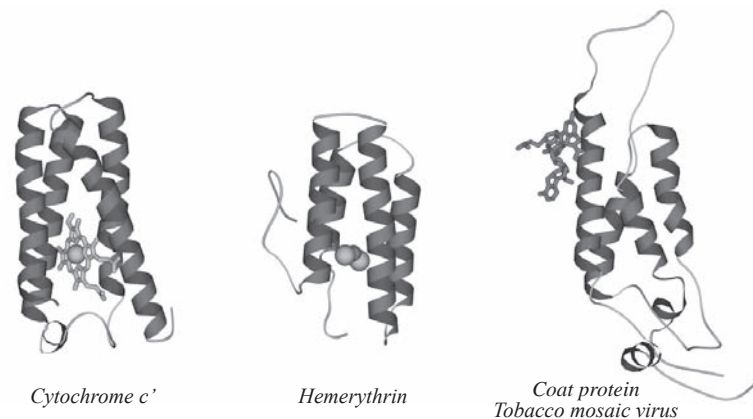


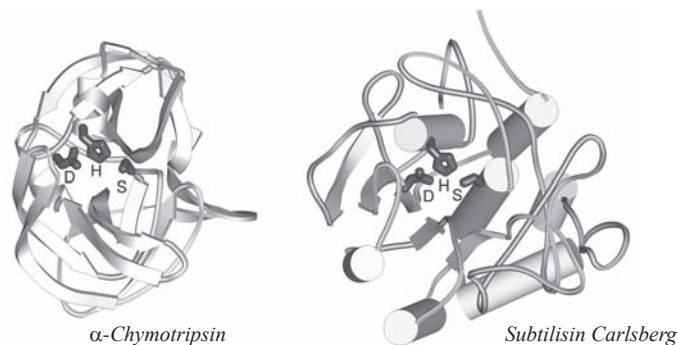
Рис. 1-4. Три сходных по форме (состоящие из четырех  $\alpha$ -спиралей каждый), но разных по функции белка: цитохром с', гемэритрин и белок оболочки вируса табачной мозаики. Показана как белковая цепь, так и кофакторы: скелетные модели — гем (в цитохроме) и фрагмент РНК (в белке оболочки вируса); шарики — ионы железа (в геме цитохрома и в гемэритрине) и связанный железом кислород (в гемэритрине)

Третье. Твердый белок («аперриодический кристалл», говоря словами Шредингера) ведет себя именно как кристалл при изменении внешних условий (например, при повышении температуры) — т. е. он «терпит» до каких-то пор, а потом разом плавится, — а не теряет своей формы и твердости постепенно, как стекло. Это фундаментальное свойство белков тесно связано с надежностью их работы: как электрическая лампочка, белки ломаются по принципу «все или ничего», а не постепенно (последнее привело бы к ненадежности их действия — к расплыванию специфичности и т. д.; мы об этом еще поговорим подробно).

И последнее. Говоря о твердости, надо различать сравнительно небольшие и действительно твердые однодоменные белки (они состоят из одной компактной глобулы) — и более крупные белки, имеющие либо доменную (см. рис. 1-2), либо четвертичную (см. рис. 1-1) структуру: слагающие их субглобулы могут сдвигаться и раздвигаться.

И, конечно, кроме того, — глобулы могут слегка, как твердое тело, деформироваться (но не полностью перестраиваться!) при функционировании белка.

- Обычно сходные по внутренней структуре (по «анатомии») белки делают одно и то же. Например, многие (хотя и не все) цитохромы похожи на тот, что изображен на рис. 1-4; а многие сериновые протеазы (хотя опять-таки не все), происходящие как из позвоночных, так и из бактерий, похожи на химотрипсин (рис. 1-5).



**Рис. 1-5.** Два разных по форме, но почти идентичных по функции белка (сериновые протеазы): химотрипсин, сложенный из  $\beta$ -структуры, и субтилизин, включающий, наряду с  $\beta$ -структурой, много  $\alpha$ -спиралей (некоторые из них входят в активный центр). Несмотря на кардинальное различие в укладке цепи, их каталитические центры состоят из одинаковых и одинаково в пространстве (но не в цепи!) расположенных остатков: Ser195 (S), His57 (H) и Asp102 (D) в химотрипсине, и Ser221 (S), His64 (H) и Asp32 (D) в субтилизине

Однако порой практически одна и та же пространственная структура может обеспечивать совсем разные функции. Так, из трех, изображенных на рис. 1-4, сходных по «грубой» пространственной структуре белков — один (цитохром) связывает электрон, другой (гемэритрин) связывает кислород (эти функции немного похожи — обе они включены в цепь окислительных реакций), а третий (белок оболочки вируса) связывает гораздо более крупные молекулы — РНК, а также другие оболочечные белки.

- Как уже говорилось, структура белка определяет его функцию.

Верно ли обратное, т. е. определяет ли функция белка его структуру? Здесь тоже прослеживаются некие частные корреляции, но, в общем, такое определяющее влияние функции прослеживается только на очень грубом уровне — уровне, определяемом «средой обитания» белковых функций. Так, белки, несущие структурную функцию — например, те, из которых слагаются волосы или фибриллы — как правило, фибриллярны; рецепторы — это мембранные белки, и т. д. Но часто это совсем неверно на уровне архитектуры белка, т. е. на уровне его внутреннего строения. Так, две сериновые протеазы (химотрипсин и субтилизин) имеют одну и ту же каталитическую функцию и даже почти одинаковую специфичность — при совершенно разной (см. рис. 1-5) внутренней структуре глобул [эти белки сходны не более чем тюлень и жук-плавунец: лишь их (белков) «плавники» — т. е. их активные центры, включающие по полудюжине аминокислотных остатков из пары сотен звеньев цепи, — устроены одинаково, а все остальное — по-разному]. Более того: существуют по-разному устроенные активные центры, исполняющие одну и ту же работу (например, у сериновых протеаз, с одной стороны, и у металлопротеаз, с другой).

Итак, ряд примеров показывает, что функция белка не определяет его пространственную структуру. Но, сказав это, надо иметь в виду масштабы. Если обрабатываемая молекула велика — с ней может взаимодействовать почти весь белок, и тогда почти вся его форма важна для его функции. Если же обрабатываемая белком молекула мала (что бывает чаще), то именно мелкомасштабные детали небольшого участка поверхности белка определяют его функцию, а остальное его тело лишь фиксирует эти критически важные детали. Последнее означает, что основная задача большей части белковой цепи — соорудить твердый фундамент для активного центра. Иными словами, специфичность функции белка подразумевает его твердость.

**Внутренний голос:** Подчеркивая нетривиальные и пикантные факты исполнения одной функции белками совершенно разной архитектуры, и разных функций — архитектурно-сходными белками, не следует забывать

о том тривиальном, но от этого не менее верном правиле, что похожие по архитектуре белки очень часто гомологичны (генетически родственны) и имеют идентичные или сходные функции...

**Лектор:** Верно — но тривиально. А я хотел подчеркнуть ту важную для понимания физики белка мысль, что активные центры могут мало зависеть от устройства всего остального тела белка.

Общим же у этого «всего остального» тела является то, что тело белка, во всяком случае «работающего» (а не ожидающего начала работы) белка — твердое: иначе нельзя обеспечить специфичность активного центра.

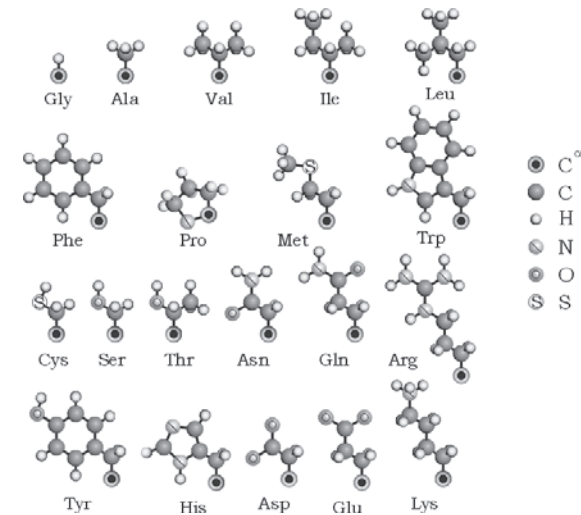
■ В свое время мы подробно рассмотрим структуры белков, их способность к самоорганизации и причину их твердости, обсудим функции и прочие интересные для биолога вопросы — но сначала необходимо рассмотреть аминокислотные остатки и их элементарные взаимодействия между собой и с окружающей средой.

## ЭЛЕМЕНТАРНЫЕ ВЗАИМОДЕЙСТВИЯ В БЕЛКАХ И ВОКРУГ

### Лекция 2

Стереохимия L-аминокислотных остатков. Валентные связи и углы между ними. Их колебания. Вращение вокруг валентных связей. Пептидная группа. Транс- и цис-пролины.

Аминокислоты, из которых сложена полипептидная цепь (рис. 2-1), могут находиться в двух стерических формах: L и D.



**Рис. 2-1.** Боковые цепи двадцати стандартных аминокислотных остатков, исходящие из  $C^\alpha$  атомов главной цепи. Справа показаны входящие в аминокислоты атомы



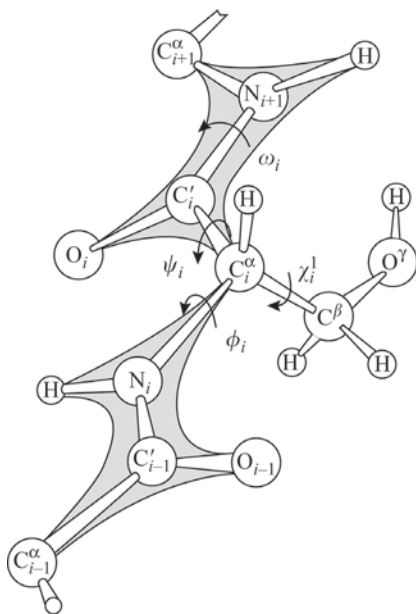
Эти формы, L и D, зеркально-симметричны: в них массивный боковой радикал R (изображенный на рис. 2-1) и H-атом, стоящие при  $\alpha$ -углероде ( $C^\alpha$ ) аминокислоты, меняются местами (стрелки, а также уменьшение размеров и яркости значков указывают углубление атомов в рисунок):



L- и D-форм нет только у глицина Gly, «боковая цепь» которого состоит лишь из H-атома.

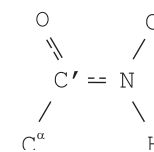
Белковые цепи сложены только из остатков L-аминокислот. Только они кодируются генами. D-аминокислотные остатки — они встречаются в пептидах — не кодируются при матричном синтезе белка, а синтезируются специальными ферментами. Рацемизация (L  $\leftrightarrow$  D переход) спонтанно в белках практически не происходит. Не происходит она и при биосинтезе, хотя часто встречается при чисто химическом синтезе пептидов, где на борьбу с рацемизацией уходит много сил.

Аминокислоты в белковой цепи связаны между собой пептидными связями  $C'$  и N атомов (рис. 2-2).



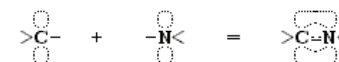
**Рис. 2-2.** Полипептид: главная цепь и боковая группа (Ser) на ней. Пептидные группы заштрихованы. Показаны углы внутреннего вращения в главной ( $\phi$ ,  $\psi$ ,  $\omega$ ) и боковой ( $\chi^1$ ) цепях. Индексы  $i-1$ ,  $i$ ,  $i+1$  показывают последовательность аминокислотных остатков в цепи. Стрелки указывают направление вращения ближней к нам части цепи относительно более отдаленной ее части, ведущее к росту угла поворота. Картинка взята из [3] и адаптирована

■ Важную роль в структуре белка играет как жесткость, так и плоская форма всей пептидной группировки.



Ее плоская форма обеспечивается так называемой  $sp^2$ -гибридизацией электронов N и  $C'$  атомов. «Гибридизация» электронных орбит — это чисто квантовый эффект.  $Sp^2$ -гибридизация преобразует одну сферическую  $s$ - и две «восьмеркообразные»  $p$ -орбиты электронов атома в три вытянутые прочь от ядра  $sp^2$ -орбиты. Эти три орбиты вовлекают атом в три ковалентные связи, лежащие в одной плоскости ( $\rightarrow \leftarrow$ ). Ковалентная связь создается «делокализованным» электронным облаком, охватывающим оба связанных атома.

Жесткость пептидной группы возникает из-за дополнительной связи  $p$ -электронов, не задействованных в  $sp^2$ -орбитах. Эти  $p$ -электроны N-,  $C'$ - и O-атомов также вступают в связь и «делокализируются», создавая электронное облако, охватывающее атомы N,  $C'$  и O (именно поэтому связи  $C=N$  и  $C=O$  рисуются как равноправные «полупортные»,  $=$ ). И так как  $p$ -орбита имеет форму «восьмерки», перпендикулярной всем трем  $sp^2$ -орбитам ( $\rightarrow \leftarrow$ ), то дополнительная ковалентная связь этих «перпендикулярных»  $p$ -орбит препятствует вращению вокруг  $C'-N$ -связи.



Хочу напомнить, что делокализация электронов, их переход с атома на атом — основная причина образования химических связей. Это следует из принципа неопределенности Гейзенберга:

$$\Delta p \Delta x \sim \hbar. \quad (2.1)$$

Здесь  $\Delta p$  — неопределенность в импульсе частицы,  $\Delta x$  — неопределенность в ее координате, а  $\hbar \equiv h/2\pi$ , где  $h$  — постоянная Планка. Так как направление движения электрона в атоме не может быть известно, то  $\Delta p \approx |p| = mv$ , где  $v$  — скорость, а  $m$  — масса частицы. Следовательно,

$$v \sim \hbar/m\Delta x.$$

В то же время, кинетическая энергия частицы  $E = mv^2/2$ , т. е.

$$E \sim \hbar^2/m\Delta x^2. \quad (2.2)$$

Следовательно, при делокализации, когда  $\Delta x$  возрастает, энергия частицы падает, и она переходит в более стабильное состояние. Видно, что этот эффект наиболее силен для легких частиц, т. е. для электронов. Так делокализация электрона приводит к химической связи.

Длина химической связи близка к Ван-дер-Ваальсову радиусу атома, т. е. составляет 1–2 Å (точнее: около 1 Å — для связи C–H, N–H или O–H; около 1,2–1,3 Å — для C=O, C=O, C=N и C=C; 1,5 Å — для C–C и около 1,8 Å — для S–S).

Характерные величины валентных углов составляют около 120 и 109°. Углы в 120° находятся при  $sp^2$ -гибридизованных атомах типа  $-C'<$ ,  $-N<$ , где три валентные связи направлены из центра к вершинам плоского треугольника, углы в 109° — при  $sp^3$ -гибридизованных атомах типа  $>C^a<$ , где четыре связи направлены из центра к вершинам тетраэдра, а также при зазывающих по две связи O< или S< атомах.

- Остановимся теперь на характерных величинах флуктуаций, т. е. тепловых колебаний валентных связей и углов. Они могут вносить вклад в гибкость белковой цепи.

Частоты колебаний проявляются в инфракрасных (ИК) спектрах белков. Характерные частоты колебаний таковы:  $\nu \sim 7 \cdot 10^{13} \text{ с}^{-1}$  для колебания атома H, например, в C–H-связи (при этом длина волны ИК света  $\lambda = c/\nu \sim 5 \text{ мкм}$ ); и  $\nu \sim 2 \cdot 10^{13} \text{ с}^{-1}$  для колебания «тяжелых» групп, например, в  $\text{CH}_3\text{--CH}_3$  связи (при этом  $\lambda = c/\nu \sim 15 \text{ мкм}$ ; здесь  $c$  — скорость света, 300 000 км/с).

Возбуждены ли эти колебания теплом при комнатной температуре?

Для ответа на этот вопрос мы должны сравнить тепловую энергию, приходящуюся на одну степень свободы («тепловой квант»  $kT$ ) с энергией колебания. Оценим величину  $kT$  при «нормальной» температуре. Здесь  $T$  — абсолютная температура в градусах Кельвина ( $T = 300 \text{ К}$  при  $27^\circ\text{C}$ , т. е. при примерно «комнатной» температуре; К — обозначение градуса Кельвина), а  $k$  (иногда еще пишут  $k_B$ ) — постоянная Больцмана [равная 2 (кал/моль)/К, или  $0,33 \cdot 10^{-23}$  кал/К в расчете на одну частицу, так как в одном моле —  $6 \times 10^{23}$  частиц]. Значит, при комнатной температуре «тепловой квант»  $kT = 600$  кал/моль, или  $(600 \text{ кал}) / (6 \cdot 10^{23} \text{ частиц})$ , т. е.  $10^{-21}$  калорий на частицу.

Соответствующую такому тепловому кванту частоту  $\nu_T$  можно найти из соотношения  $kT = h\nu_T$  [где постоянная Планка  $h \equiv 2\pi\hbar = 6,6 \cdot 10^{-34} \text{ Дж}\cdot\text{с} = 1,6 \cdot 10^{-34} \text{ кал}\cdot\text{с}$ ; напомним, что 1 калория (кал) = 4,2 джоуля (Дж)]. Итак, частота  $\nu_T = 7 \cdot 10^{12} \text{ с}^{-1}$  при  $T = 300 \text{ К}$ , т. е. при  $27^\circ\text{C}$ .

«Тепловой квант» не может возбудить более жесткие, т. е. более частые, чем он сам, колебания.

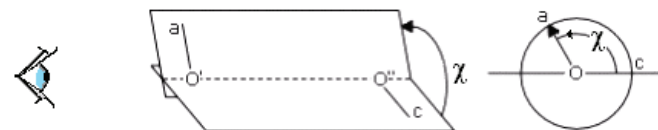
Значит, при комнатной температуре валентные связи «тверды» и не флуктуируют: для этих связей частота колебаний  $\nu \sim 2 \cdot 10^{13} \text{ — } 7 \cdot 10^{13} \text{ с}^{-1}$ , т. е. она на порядок больше, чем  $\nu_T = 7 \cdot 10^{12} \text{ с}^{-1}$ .

Однако валентные связи могут раскачиваться падающим инфракрасным светом, и на этом основана ИК-спектроскопия белков. Собственно, именно ИК-спектроскопия и дает сведения о колебаниях атомов, валентных связей и валентных углов. Конечно, экспериментально их исследуют сначала на малых молекулах, а затем полученную информацию используют при изучении белков.

В отличие от колебаний длин связей, тепловые колебания валентных углов, на порядок более мягкие — их частоты лежат в пределах  $10^{12}\text{—}10^{13} \text{ с}^{-1}$ , — возбуждены при комнатной температуре. Однако их характерная амплитуда невелика — около  $5^\circ$ .

Итак, колебания длин валентных связей вовсе не вносят вклада в гибкость белковой цепи, а колебания валентных углов вносят весьма скромный вклад. На самом деле гибкость белковой цепи, а значит и ее способность складываться и в  $\alpha$ -спирали, и в глобулы обеспечивается вращением (хотя и не вполне свободным вращением, см. ниже) *вокруг* валентных связей. Поэтому и структура цепи в целом часто описывается просто в терминах углов вращения вокруг валентных связей — тогда она называется «конформацией». Впрочем, очень часто термины «структура» и «конформация» употребляются просто как синонимы.

Основные сведения о вращении атомных группировок вокруг валентных связей тоже поставляет ИК-спектроскопия; как и раньше, я остановлюсь только на результатах.

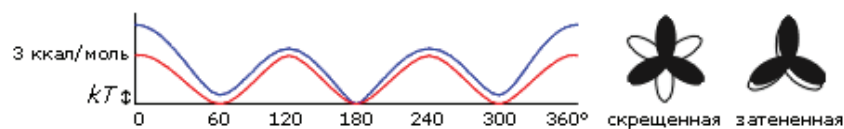


**Рис. 2-3.** Отсчет величины двугранного угла (угла поворота): вид поперек (слева) и вдоль (справа) оси вращения. Дальняя от нас валентная связь  $O''\text{—}c$  — «ось координатного круга» в тригонометрии, центральная связь  $O'\text{—}O''$  — «ось вращения», ближняя к нам связь  $O'\text{—}a$  — «стрелка на координатном круге»

- Взаимное расположение атомных группировок, связанных валентной связью, описывается двугранным углом (двугранным — так как это угол между двумя плоскостями — «гранями»: на рис. 2-3 в одной из этих плоскостей лежат точки  $a, O', O''$ , а в другой —  $O', O''$ ,  $c$ ).

Рисунки 2-2, 2-3 иллюстрируют отсчет этого угла. Отсчет — такой же, как в школьной тригонометрии, если рассматривать ближнюю к нам валентную связь как «вращающуюся стрелку», дальнюю от нас — как «координатную ось», а центральную валентную связь — как «ось вращения». То есть вращение «стрелки» против хода часов увеличивает угол поворота, а по ходу — уменьшает.

■ Рисунок 2-4 показывает характерное поведение энергии при вращении вокруг связи, соединяющей два  $sp^3$ -гибридизованных атома (на примере  $H_3C-CH_3$  и  $CH_2C-CH_2C$ -связей). Такие связи характерны для алифатических боковых групп. Углы вращения боковых групп называются углами  $\chi$ , «хи» (см. рис. 2-2). Максимумы такого трехкратного (соответственно симметрии поворота вокруг  $sp^3-sp^3$ -связи) потенциала (т. е. потенциала, имеющего три максимума и три минимума при повороте на  $360^\circ$ ) соответствуют трем «затененным» конформациям ( $0, 120$  и  $240^\circ$ ), приводящим к сближению (и отталкиванию) электронных облаков. Эти облака отталкиваются, так как образующие их электроны уже задействованы в ковалентных связях. Величина возникающих при этом потенциальных барьеров вращения вокруг  $H_3C-CH_3$ -связи — около 3 ккал/моль, а характерный размах тепловых флуктуаций вокруг этих минимумов (т. е. те отклонения от них, при которых энергия возрастает на  $kT$ ) —  $15-20^\circ$ .

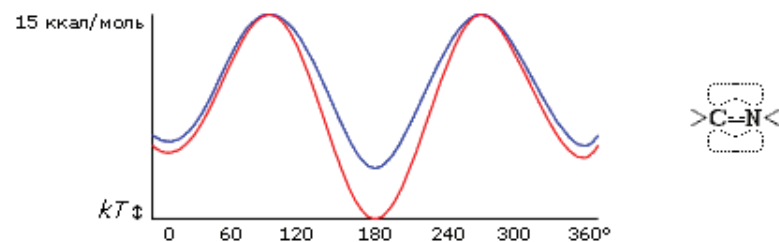


**Рис. 2-4.** Характерный потенциал вращения вокруг одинарной связи между двумя  $sp^3$ -гибридизованными атомами: вокруг  $H_3C-CH_3$  связи (нижняя кривая) и вокруг  $(C-CH_2)-(CH_2-C)$  связи (верхняя кривая). Главным энергетический эффект дает отталкивание электронных облаков: оно максимально в «затененной» конформации (при  $0, 120, 240^\circ$ ) и минимально в «скрещенной» (при  $60, 180, 300^\circ$ ). Отталкивание маленьких H-атомов несущественно. Однако отталкивание тяжелых, окруженных большими электронными облаками C-атомов в окрестности  $0^\circ$  (в цепи  $C-C-C$ ) создает добавочный энергетический эффект, отличающий поворот вокруг  $(C-CH_2)-(CH_2-C)$ -связи от поворота вокруг  $(H_3C)-(CH_3)$  связи. Для сравнения стрелкой  $\updownarrow$  отмечена величина  $kT$

Когда некоторые из валентностей при  $sp^3$ -гибридизованных атомах заняты не водородом, а более массивными атомами — их отталкивание повышает соответствующую слишком тесному сближению этих массив-

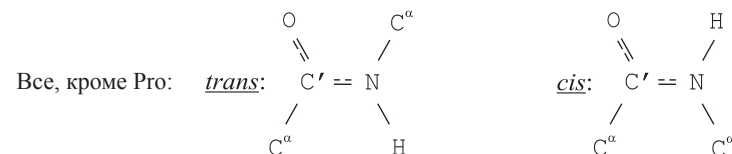
ных атомов часть барьера. Это показано на рис. 2-4 на примере вращения вокруг центральной связи в  $(C-CH_2)-(CH_2-C)$ .

Рисунок 2-5 показывает характерное поведение энергии при вращении вокруг пептидной связи, соединяющей два  $sp^2$ -гибридизованных атома (C' и N). Угол вращения вокруг этой связи называется углом  $\omega$  (см. рис. 2-2). Потенциал двукратен (т. е. имеет два максимума и два минимума при повороте на  $360^\circ$ ), соответственно симметрии поворота вокруг  $sp^2-sp^2$ -связи. Потенциальные барьеры велики из-за участия в пептидной связи р-электронов (как о том говорилось в начале лекции), минимумы потенциала приходятся на  $0$  и  $180^\circ$  (где р-орбиты, стягивающие C'- и N-атомы, наиболее сближены), а его максимумы — на  $90^\circ$  и  $270^\circ$  (где эти стягивающие р-орбиты наименее сближены и, следовательно, хуже всего связаны друг с другом). Из-за высоких барьеров характерный размах тепловых флуктуаций угла вращения вокруг таких связей невелик — около  $10^\circ$ .

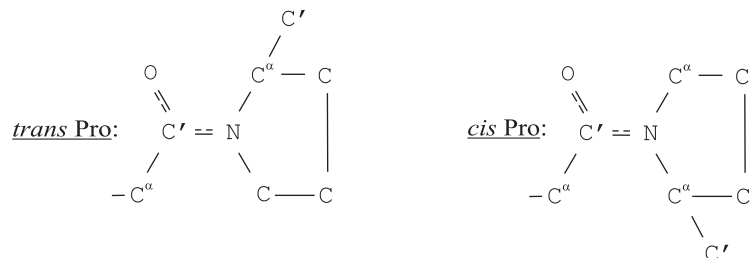


**Рис. 2-5.** Характерный потенциал вращения вокруг кратной связи, соединяющей два  $sp^2$ -гибридизованных атома (C' и N) в пептидной связи. Справа показана связь электронных р-орбит при угле, равном  $0^\circ$  (или  $180^\circ$ ). Для всех (за исключением предпролиновых) пептидных связей их энергия (нижняя кривая) при  $0^\circ$  выше, чем при  $180^\circ$  из-за отталкивания массивных  $C^\alpha$ -атомов. Эта разница энергий мала для пептидной связи, находящейся перед Pro (верхняя кривая), имеющих не один, а два C-атома при N-атоме. См. ниже текст и рисунок, где приведена структурная формула пролина

Стоит отметить, что отталкивание массивных  $C^\alpha$ -атомов делает *cis*-конформацию ( $\omega=0^\circ$ ) энергетически довольно невыгодной; поэтому почти все пептидные группы в белках находятся в *trans*-конформации ( $\omega=180^\circ$ ).

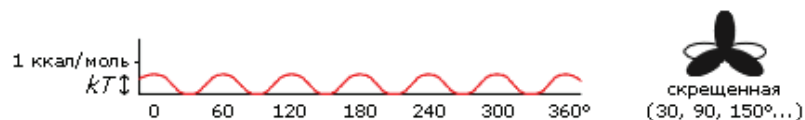


Исключение составляет пептидная связь, предшествующая пролину (*имино-*, а не *аминокислоте*): у N-атома Pro — два (помимо предшествующего C' атома) почти симметричных массивных радикала ( $-C^{\alpha}HC_2$  и  $-CH_2C$ , см. рис. 2-1), так что для этой связи *trans*-конформация лишь ненамного лучше, чем *cis*-.



И в глобулярных белках, и в развернутых (т. е. флуктуирующих, не имеющих фиксированной структуры) пептидах наблюдается около 90 % *trans*- и 10 % *cis*-пролинов. Хочу обратить ваше внимание на эту закономерность — что выгодно само по себе, то чаще встречается и в глобулярных белках. Мы столкнемся с ней еще не раз.

Наконец, рассмотрим потенциал вращения вокруг связи, соединяющей  $sp^3$ -гибридизованный и  $sp^2$ -гибридизованный атомы. Углы вращения вокруг такой связи называются углами  $\phi$  (вращение вокруг N—C $^{\alpha}$ -связи) и  $\psi$  (вращение вокруг C $^{\alpha}$ —C'-связи) (см. рис. 2-2). При вращении по каждому из них возникает шестикратный (6 минимумов и 6 максимумов при повороте на 360°) потенциал с довольно низкими барьерами (рис. 2-6): их высота приблизительно равна 1 ккал/моль — порядка энергии тепловых колебаний (составляющей, как мы помним, 0,6 ккал/моль при комнатной температуре). Именно почти свободное вращение по углам  $\phi$  и  $\psi$  вокруг такого рода связей (между N и C $^{\alpha}$  и между C $^{\alpha}$  и C') в главной цепи полипептида и обеспечивает основную гибкость белковой цепи.



**Рис. 2-6.** Характерный потенциал вращения вокруг одинарной связи между  $sp^3$ - и  $sp^2$ -гибридизованными атомами (на примере вращения вокруг  $H_3C-C_6H_5$  связи)

## Лекция 3

Ван-дер-Ваальсово взаимодействие: притяжение на больших расстояниях, отталкивание на малых. Разрешенные конформации аминокислотного остатка (карты Рамачандрана для глицина, аланина, валина, пролина).

Рассмотрим теперь элементарные взаимодействия атомов, не связанных ковалентно друг с другом.

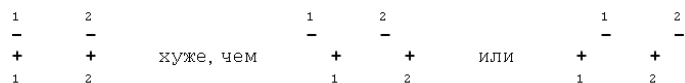
Но сначала вспомним, когда между сближающимися атомами не образуется ковалентной связи. Напомню, что основная причина образования химических связей — делокализация электронов, их «переход» с атома на атом и обратно — или, как говорят, пребывание электронов на совместной, «обобществленной» орбите, охватывающей два атома или больше. При этом, согласно *запрету Паули*, в одном месте, на одной орбите может находиться не более двух электронов — причем тогда их «спины» (т. е. моменты вращения электронов) должны быть противоположно направлены (или, как говорят, «спарены»). Спаривание двух электронов на одной совместной, охватывающей два атома орбите ведет к прочной ковалентной связи.

Если же у сближающихся атомов уже на каждой орбите по два спаренных электрона, то эти атомы не могут образовать ковалентную связь. Иначе на каждой обобществленной орбите электронов было бы слишком много — четыре. А, согласно *запрету Паули*, одна орбита может вынести не более двух! Значит, уже насыщенные, заполненные — каждая — электронной парой орбиты не могут далее обобществляться. И потому несущие такие «насыщенные» орбиты атомы отталкиваются на малых расстояниях, когда их электронные облака начинают перекрываться. Такие атомы непроницаемы друг для друга при обычных, не сверхвысоких температурах.

То же происходит, если в сближающихся молекулах все их валентности уже насыщены: они отталкиваются, когда их атомы сближаются до 2–3 Å.



Однако на больших расстояниях (когда нет перекрытия электронных облаков) все атомы и молекулы притягиваются — если только у них нет электрического заряда, о чем мы поговорим позже. Притяжение это имеет чисто квантовую основу. Оно связано с согласованными колебаниями электронов в обоих атомах. Дело в том, что согласованное («в одну и ту же сторону») смещение электронов приводит к притяжению атомов (в то время как атомы с несмещенными электронами не притягиваются и не отталкиваются). Это ясно из следующей схемы (здесь два атома, «1» и «2», и у каждого — электрон «-» и ядро «+»):



Глядя на схему, нужно помнить, что электрическая энергия растет с уменьшением расстояния  $r$  как  $1/r$ . В результате смещения электронов взаимодействия электрон-электрон и ядро-ядро не меняются, но (рассмотрим самый правый рисунок) усиливается притяжение электрона 1 к ядру 2 — причем оно усиливается больше, чем ослабевает притяжение электрона 2 к ядру 1. Вспомним теперь, что электрон в атоме всегда колеблется («крутится вокруг ядра», не смея упасть на него из-за квантовой неопределенности Гейзенберга). Описанный выше эффект способствует появлению согласованности колебаний электронов сблизившихся атомов и, тем самым, — притяжению этих атомов.

При отдалении атомов их взаимодействие ослабляется. В результате энергия взаимодействия спадает с расстоянием  $r$  между ядрами атомов; можно показать, что она спадает пропорционально  $(1/r)^6$ .

Общая энергия взаимодействия атомов — она еще называется энергией Ван-дер-Ваальсова взаимодействия — охватывается рис. 3-1 и приближенно описывается формой Леннарда-Джонса:

$$U_{LJ}(r) = E_0[(r_0/r)^{12} - 2(r_0/r)^6]. \quad (3.1)$$

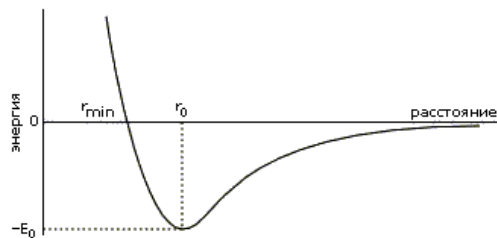


Рис. 3-1. Характерный вид потенциала Ван-дер-Ваальсова взаимодействия

Здесь  $r_0$  — как легко проверить, взяв производную  $U_{LJ}$  по расстоянию  $r$ , — есть то расстояние, на котором энергия  $U_{LJ}$  проходит через минимум, а  $-E_0$  — глубина этого минимума. Последний член формулы (3.1) — тот, что падает с расстоянием, как  $(\text{const}/\text{расстояние})_{\text{в шестой степени}}$  — соответствует притяжению (знак минус перед ним указывает, что соответствующая энергия понижается при сближении); а тот, где значится двенадцатая степень, соответствует отталкиванию (этот член положителен, т. е. соответствующая ему энергия растет при сближении).

Таблица 3-1. Характерные параметры потенциалов Ван-дер-Ваальсовых взаимодействий

Взаимодействие	$E_0$ , ккал/моль	$r_0$ , Å	$r_{\min}$ , Å	Ван-дер-Ваальсов радиус атома, Å
H...H	0,12	2,4	2,0	H: 1,0
H...C	0,11	2,9	2,4	C: 1,5
C...C	0,12	3,4	3,0	O: 1,35
O...O	0,23	3,0	2,7	N: 1,35
N...N	0,20	3,1	2,7	CH <sub>2</sub> : ≈ 1,5
CH <sub>2</sub> ...CH <sub>2</sub>	≈ 0,5	≈ 4,0	≈ 3,0	

*Примечания.* Величины  $E_0$ ,  $r_0$  для межатомных взаимодействий взяты из R. A. Scott, H. A. Scheraga, *J. Chem. Phys.* (1965) **45**:2091, а величины  $r_{\min}$  — из G. N. Ramachandran, V. Sasisekharan, *Adv. Prot. Chem.* (1968) **23**:283. Исходя из этих величин, оценены параметры взаимодействия CH<sub>2</sub>...CH<sub>2</sub>. Взаимодействие CH<sub>2</sub>...CH<sub>2</sub> зависит от взаимной ориентации этих групп; поэтому приведенные в таблице числа — приближительные; их, однако, часто приходится использовать, рассчитывая взаимодействия в белках, так как рентген «не видит» водородных атомов в белках.

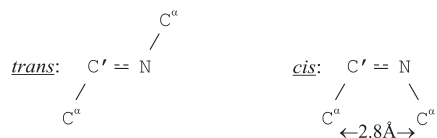
Формула (3.1) точно описывает только притяжение на больших (когда  $r \gg \gg r_0$ ) расстояниях. Отталкивание на малых расстояниях она описывает только качественно, как «очень сильное отталкивание, превосходящее любое притяжение, когда  $r$  стремится к нулю». То, что формула (3.1) приближенная, видно хотя бы по тому, что она подразумевает округлость атомов (так как описываемое ею взаимодействие не зависит от направления), — а электронное облако у атома, вообще говоря, не круглое — вспомните хотя бы о торчащих р-электронах. Вообще, правильно и совсем корректно взаимодействие атомов описывает только квантовая механика — но точно рассчитать она может только очень простые системы, вроде атома He, иона H<sub>2</sub><sup>+</sup> или молекулы H<sub>2</sub>. Для всего остального приходится брать приближенные, «полуэмпирические» формулы вроде (3.1), где вид формулы берется из качественных физических соображений, а параметры (у нас —  $E_0$  и  $r_0$ ) подбираются из опыта. Некоторые основные данные сведены в табл. 3-1.

Обратите также внимание на то, что в таблице дано не только оптимальное расстояние  $r_0$ , но и  $r_{\min}$  — минимальное расстояние, на котором соответствующие атомы встречаются в кристаллах. Расстояние  $r_{\min}$  примерно соответствует той точке рис. 3-1, где энергия при сильном сближении проходит через 0. Величины  $r_{\min}$  удобно применять тогда, когда прикидываешь: возможна ли данная конформация цепи, или нет.

**Внутренний голос:** Существует масса работ по выводу потенциалов атом-атомных взаимодействий. Не означает ли это, что точности всех этих потенциалов доверять не следует?

**Лектор:** В вопросе о *форме* потенциала (см. рис. 3-1) разногласий нет. Нет особых разногласий и в оценках  $r_{\min}$  — они прямо берутся из кристаллических структур. Есть разногласия в вопросе о величинах параметров  $r_0$  и особенно  $E_0$ ... Расстояния  $r_0$  и энергии  $E_0$  в основном берутся тоже из кристаллов — из их структур (расстояния) и теплот сублимации (энергии). Но кристаллы обычно состоят не из атомов, а из молекул. Например:  $\text{CH}_4$ ,  $\text{C}_2\text{H}_6$ ... Так что при расчете энергии всегда возникают вопросы типа: каков вклад  $\text{C}\dots\text{C}$ , каков —  $\text{H}\dots\text{H}$ , каков —  $\text{C}\dots\text{H}$ -взаимодействий? Разные авторы решают его по-разному, поэтому потенциалы у них порой заметно расходятся. Однако эти расхождения обычно нивелируются, когда переходят от атомов обратно к молекулам. Но: если вам придется рассчитывать молекулярные структуры, надо помнить, что нельзя самостоятельно брать одни параметры (например, энергию  $\text{C}\dots\text{C}$ -взаимодействия) у одного автора, а другие (скажем, энергию  $\text{H}\dots\text{H}$ -взаимодействия) — у другого: тут действует принцип «все или ничего», иначе будут ошибки.

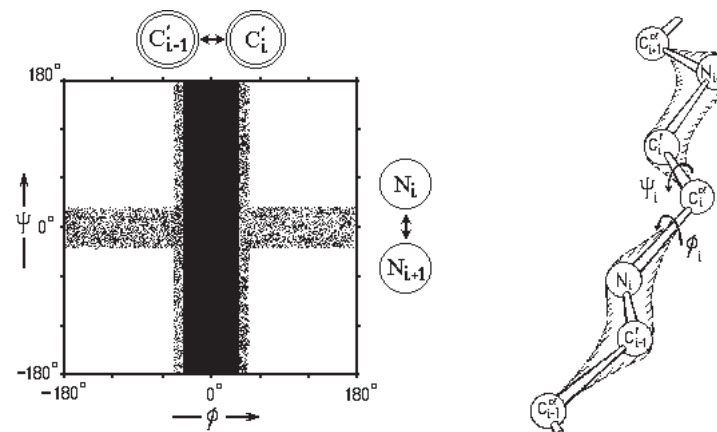
Приведенные в таблице величины помогают понять, почему *trans*-конформация  $\text{C}'\text{--N}$ -связи допустима, а ее *cis*-конформация — нет (для всех аминокислотных остатков, кроме пролина, о чем мы уже говорили): в *trans*-конформации  $\text{C}'\text{--N}$ -связи расстояние между  $\text{C}^\alpha$ -атомами равно  $3,8\text{Å}$ , а в *cis*- (когда эти атомы наиболее сближены) — всего  $2,8\text{Å}$ , что меньше минимально допустимого расстояния  $r_{\min}$ , равного  $3,0\text{Å}$  для пары  $\text{C}\dots\text{C}$ :



■ Из-за жесткости *trans*-формы  $\text{C}'\text{--N}$ -связи и из-за того, что в ней  $\text{C}^\alpha$ -атомы соседних по цепи аминокислот довольно далеко отстоят друг от друга, соседние по цепи остатки могут менять свою конформацию почти независимо друг от друга. Но внутри остатка вращения по углам  $\phi$  и  $\psi$  взаимосвязаны. Изображенные в координатах  $(\phi, \psi)$  «разрешенные» и «запрещенные» конформации остатка называются *картами Рамачандрана* (или, более корректно, — *картами Сасисекхарана-Рамачандрана*).

Прежде чем рисовать их, посмотрим, какие конформации разрешены (и какие — нет) при вращении по углу  $\phi$  (вокруг  $\text{N--C}^\alpha$ -связи) и по углу  $\psi$  (вокруг  $\text{C}^\alpha\text{--C}'$ -связи) в отдельности.

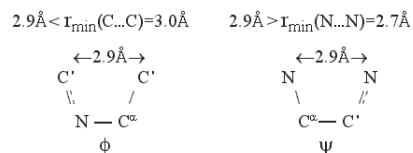
Мы уже знаем, что вращение вокруг этих связей (соединяющих  $\text{sp}^3$ -гибридизованный  $\text{C}^\alpha$ -атом с  $\text{sp}^2$ -гибридизованными  $\text{N}$  или  $\text{C}'$ ) практически свободно. Однако в *cis*-конформациях (при  $\phi=0^\circ$  или  $\psi=0^\circ$ ) вращающиеся вокруг таких связей атомы ( $\text{C}'_{i-1}$  и  $\text{C}'_i$  при вращении по углу  $\phi$  вокруг  $\text{N--C}^\alpha$ -связи, и атомы  $\text{N}_i$  и  $\text{N}_{i+1}$  при вращении по углу  $\psi$  вокруг  $\text{C}^\alpha\text{--C}'$  связи, см. рис. 3-2;  $i-1, i, i+1$  — номера последовательных остатков в цепи) сближаются особенно сильно, и здесь возможен запрет конформации из-за отталкивания этих атомов — как говорят, стерический запрет.



**Рис. 3-2.** Так выглядела бы карта запрещенных (■), напряженных (▨) и полностью разрешенных (□) конформаций при вращении по углам  $\phi, \psi$  во фрагменте  $\text{C}^\alpha\text{C}'\text{N--C}^\alpha\text{--C}'\text{NC}^\alpha$ , если бы ко всем этим атомам не было прикреплено больше никаких других (см. справа), и атомы остатков  $i-1$  и  $i+1$  не взаимодействовали бы между собой

Почему по углу  $\phi$  вращаться труднее, чем по углу  $\psi$ ?

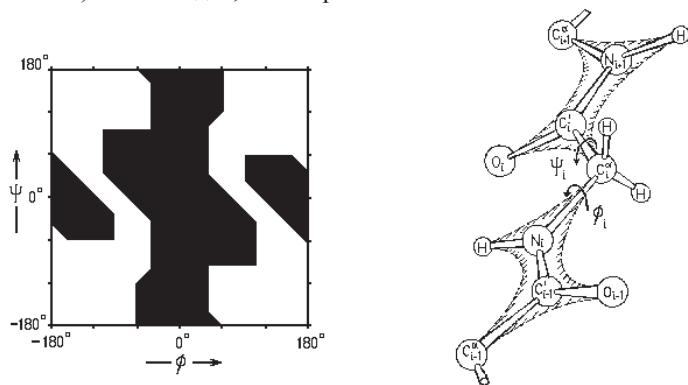
Следующая схема показывает, что минимальное расстояние между атомами  $\text{C}'_{i-1}$  и  $\text{C}'_i$  (при вращении по  $\phi$ ) и между атомами  $\text{N}_i$  и  $\text{N}_{i+1}$  (при вращении по  $\psi$ ) — одно и то же,  $2,9\text{Å}$ : это чуть *меньше*, чем  $r_{\min}=3,0\text{Å}$  для  $\text{C}\dots\text{C}$ -взаимодействия (так что здесь, при вращении по углу  $\phi$ , *cis*-конформация запрещена), и чуть *больше*, чем  $r_{\min}=2,7\text{Å}$  (хоть и меньше, чем оптимальное  $r_0=3,1\text{Å}$ ) для  $\text{N}\dots\text{N}$ -взаимодействия (так что здесь *cis*-конформация не запрещена, но напряжена).



Впрочем, расчеты показывают, что даже та небольшая гибкость, что есть у валентного угла N–C<sup>α</sup>–C', сильно снижает энергию этого напряжения.

Если бы дело ограничивалось только этими C' и N-атомами, карта запрещенных, напряженных и разрешенных конформаций выглядела бы так, как на рис. 3-2. При этом вращения по углам φ и ψ не были бы взаимосвязаны.

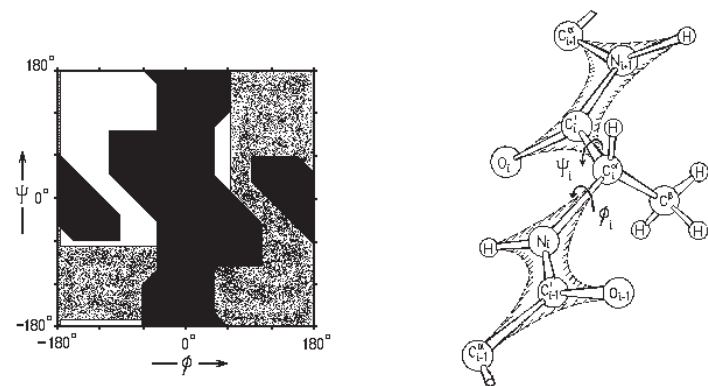
Однако при атоме C' есть еще O- и C<sup>α</sup>-атомы, а при атоме N есть также H- и C'-атомы (причем в воде, как мы вскоре увидим, к H-атому довольно жестко приторочена водородной связью молекула H<sub>2</sub>O). Так что в результате карта Рамачандрана — карта запрещенных и разрешенных конформаций самого маленького аминокислотного остатка, глицина (боковой радикал: -H) — выглядит, как на рис. 3-3.



**Рис. 3-3.** Карта запрещенных (■) и разрешенных (□) конформаций глицина (Gly) при вращении по углам φ, ψ в белковой цепи

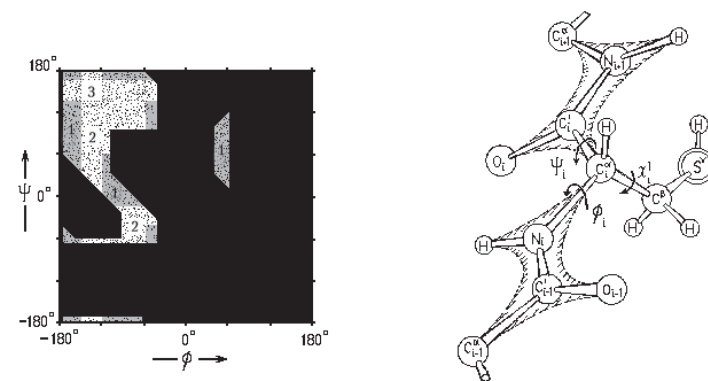
У глицина нет массивного бокового радикала. У всех остальных аминокислотных остатков он есть, и столкновение этого радикала (точнее, его ближайшего к главной цепи C<sup>β</sup>-атома) с C'<sub>i-1</sub>-атомом вырубает запрещенную область по углу φ, а с N<sub>i+1</sub>-атомом — запрещенную область по углу ψ (рис. 3-4).

■ Нарисованная на рис. 3-4 карта относится к аланину, боковой радикал которого мал: C<sup>β</sup>H<sub>3</sub>-группа. У всех остальных аминокислотных остатков радикал крупнее — с одним или двумя тяжелыми γ-атомами при C<sup>β</sup>-атоме.



**Рис. 3-4.** Карта разрешенных (□) конформаций аланина (Ala) при вращении по углам φ, ψ в белковой цепи; (▨) — области, разрешенные лишь для глицина; (■) — области, запрещенные для всех остатков взаимодействиями в главной цепи

Однако от главной цепи эти «новые» γ (и еще более далекие δ, ε и т. д.) атомы находятся далеко, и потому мало влияют на карту Рамачандрана. Точнее: в небольшой (оставленной белой на рис. 3-5) области столкновений этих γ-атомов с главной цепью нет вообще, а в других разрешенных для аланина конформациях такие столкновения могут быть — при некоторых конформациях радикала — и потому здесь разрешены не все конформации радикала; эти области я слегка затемнил на рис. 3-5.

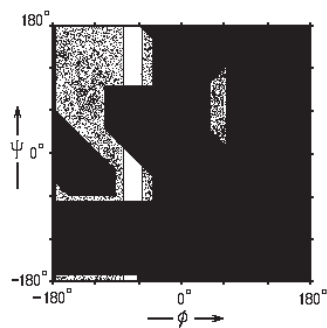


**Рис. 3-5.** Карта запрещенных (■) и разрешенных (□, ▨, ▩) конформаций более крупных остатков при вращении по углам φ, ψ в белковой цепи. В области □ разрешены все 3 ротамера C<sup>γ</sup> атома по χ<sup>1</sup>, в области ▨ — 2 ротамера, в области ▩ — только 1

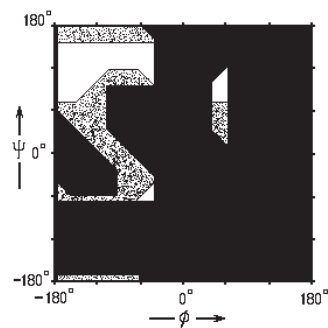
Столкновения тяжелых (не-водородных)  $\gamma$ -атомов с главной цепью особенно существенны для валина, изолейцина и треонина, имеющих по два крупных  $\gamma$ -атома.

- В заключение рассмотрим карту Рамачандрана для аминокислоты пролина. У него угол  $\phi$  практически фиксирован при  $-70^\circ$  кольцом, образуемым боковой группой Pro при участии N-атома его главной цепи; а вращение по углу  $\psi$  — такое же, как у аланина. Так что его разрешенные конформации выглядят, как белая полоса на карте Рамачандрана, рис. 3-6.

Стоит отметить, что кольцо пролина сужает и область разрешенных конформаций остатка, лежащего перед ним в цепи (рис. 3-7).



**Рис. 3-6.** Карта разрешенных (□) для пролина конформаций на фоне конформаций, разрешенных для аланина (▨); ■ — конформации, запрещенные для них обоих



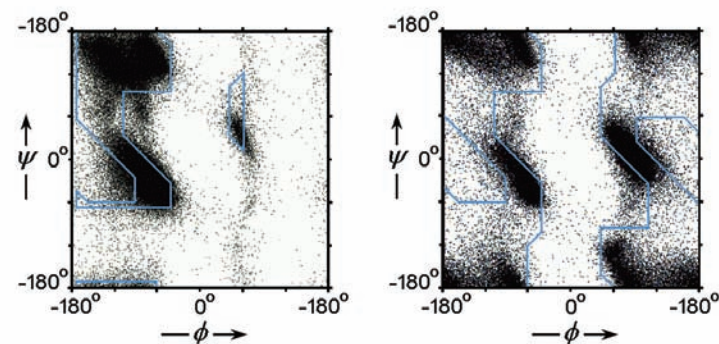
**Рис. 3-7.** Область разрешенных (□) конформаций остатка аланина, лежащего перед пролином в белковой цепи. Область (▨) была бы разрешена, если бы не пролин

- Наконец, посмотрим, как все эти теоретические выкладки согласуются с наблюдаемыми (рентгеном и ЯМР) в белках конформациями аминокислотных остатков. На рис. 3-8 эти конформации отмечены точками на фоне контуров областей, разрешенных теорией для глицина, с одной стороны, и, с другой — для аланина и прочих аминокислотных остатков.

Как видим, согласие это весьма хорошее.

Подавляющее большинство экспериментальных точек попадает в «стерически разрешенные» области. Однако некоторые точки лежат в «стерически запрещенных» областях. Последнее не удивительно, так как мы назвали «стерически запрещенными» те области, где энергия просто высока — не бесконечно велика, что, конечно, привело бы к полному запрету, — а скажем, на пару ккал/моль больше, чем в конформации с минимальной энергией.

То есть белку надо затратить энергию, чтобы загнать аминокислотный остаток в такую область, но он вполне может это сделать. Может сделать, но делает редко.



**Рис. 3-8.** Наблюдаемые в белках конформации аминокислотных остатков (точки) для глицина (справа) и для всех остальных остатков (слева). Контур очерчивает стерически разрешенные области

По ходу лекций мы убедимся, что это — общее правило: напряженные, высокоэнергетичные элементы в белках встречаются редко — но все же встречаются. Что не удивительно: раз белок стабилен, он должен состоять — в массе, но не обязательно целиком — из стабильных деталей.



## Лекция 4

Влияние водного окружения. Водородные связи. Их электрическая природа. Их энергия. Их геометрия в кристаллах. Разболтанность водородных связей в воде. Понятие об энтропии и свободной энергии. Водородные связи в белковой цепи замещают такие же связи этой цепи с водой; в результате водородные связи в белковой цепи приобретают — в водном окружении — энтропийную природу.

До сих пор мы не учитывали водного окружения белков. Пора восполнить этот пробел.

Вода — весьма своеобразный растворитель. Прежде всего, она кипит и затвердевает при аномально высоких для ее малого молекулярного веса температурах. Действительно, вода  $\text{H}_2\text{O}$  кипит при 373 К, а твердеет при 273 К, в то время как  $\text{O}_2$  кипит при 90 К и твердеет при 54 К;  $\text{H}_2$  кипит при всего 20 К и плавится при 4 К;  $\text{CH}_4$  кипит при 114 К, и т. д. То, что структуры воды и льда с трудом разрушаются теплом, показывает, что молекулы воды чем-то очень сильно связаны между собой.

Это «что-то» — связь между О- и Н-атомами молекул  $\text{H}_2\text{O}$  (именно между О и Н, так как и  $\text{O}_2$ , и  $\text{H}_2$  в отдельности легко кипят и плавятся). Эта связь называется *водородной связью*.

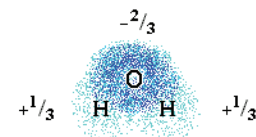
Водородные связи наблюдаются не только в воде. Они наблюдаются всегда, когда водород химически связан с одним электроотрицательным (т. е. притягивающим электрон) атомом и при этом приближается к другому электроотрицательному атому. Примеры:  $\text{O}-\text{H} : : : \text{O}$ ,  $\text{N}-\text{H} : : : \text{O}$ ,  $\text{N}-\text{H} : : : \text{N}$ . Но С-Н-группа, например, водородных связей не образует: атом С недостаточно электроотрицательный атом.

Основные особенности воды как растворителя связаны с наличием в ней мощных водородных связей.

- Водородная связь молекул воды носит электрическую природу. То, что она связана именно с электронами и зарядами, а не с ядрами водородных атомов, прямо следует из того, что температуры и теплоты кипения и

плавления легкой  $\text{H}_2\text{O}$  и тяжелой  $\text{D}_2\text{O}$  воды практически совпадают, несмотря на двукратное различие в массах ядер D и H.

Молекула воды полярна. Это значит, что на ее атомах есть небольшие («парциальные») электрические заряды: на электроотрицательном О — отрицательный, на Н — положительный. Распределение зарядов и электронных облаков на этих атомах выглядит так:



Здесь плотность электронного облака показана плотностью точек, а цифры показывают парциальный заряд атомов этой полярной молекулы. Заряд этот выражен в долях от протонного заряда. У протона, в этих единицах, заряд, естественно, равен +1, а у электрона –1. В целом молекула воды имеет заряд 0.

Заряды на полярных атомах появляются в результате того, что электроотрицательные атомы О оттягивают электронные облака от соседних Н-атомов. В результате на последних возникают небольшие положительные заряды, а на О — отрицательный заряд.

Как вы помните, заряды  $q_1$  и  $q_2$  на расстоянии  $r$  взаимодействуют в вакууме с энергией

$$U = q_1 q_2 / r. \quad (4.1)$$

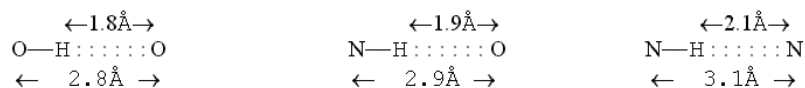
В вакууме заряды взаимодействуют очень сильно. Энергия взаимодействия двух единичных (т. е. протонных или электронных) зарядов на расстоянии 1 Å близка к 330 ккал/моль (запомните эту цифру: она нам пригодится для разных дальнейших оценок), а на более реалистичном (учитывая Ван-дер-Ваальсово отталкивание атомов) расстоянии 3 Å эта энергия близка к 110 ккал/моль. Энергия взаимодействия единичных зарядов — это характерная энергия химической связи; она в сотни раз больше, чем характерная тепловая энергия  $kT$  или характерная энергия Ван-дер-Ваальсова взаимодействия атомов.

Парциальные заряды молекулы воды меньше и потому взаимодействуют слабее — с энергией порядка 10 ккал/моль на расстоянии 3 Å, но и такой энергии хватает на то, чтобы, при притяжении Н к О, «промять» электронные оболочки атомов. Особенно этот эффект «проминания» облаков существен для атомов водорода. Электронная оболочка Н, и так состоявшая всего из одного электрона, оттянута к О-атому, так что смять ее особенно легко (и, вдобавок, Ван-дер-Ваальсов радиус атома Н, участвующего в Н-О- или Н-Н-связи, раза в полтора меньше, чем у Н, участвующего в Н-С-связи).

Снять электронную оболочку, скажем, O-атома много труднее: вокруг ядра O вращается 8 «своих» электронов, да еще какую-то часть электронного облака атом O (в H<sub>2</sub>O) оттягивает у соседних водородов.

Малость электронного облака у атома H и легкость его проминания превращает обычное электростатическое взаимодействие в водородную связь. Это относится ко всем водородным связям, из которых нас будут интересовать следующие, встречающиеся в белках связи: O—H :: O, N—H :: O, N—H :: N.

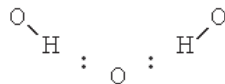
Итак, электронное облако при H-атоме — самое слабое, и оно сильно сминается при притяжении парциального положительного заряда H к парциальному отрицательному заряду O (или N) атома. В результате ядро H подходит к ядру O или N не на 2,35–2,75 Å (как то было бы при Ван-дер-Ваальсовом взаимодействии, о котором мы говорили на прошлой лекции), а заметно ближе — на расстояние 1,8–2,1 Å (эти данные получены из кристаллов малых молекул).



При таком сближении и возникает *водородная связь* (или *H-связь*). H-атом (или, точнее, O—H или N—H-группа) называется *донором* водородной связи, а тот O- или N-атом, к которому этот H приближается, называется *акцептором* водородной связи.

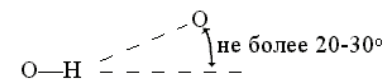
Обратите внимание, что наблюдающееся в кристаллах при образовании H-связи расстояние между O и/или N атомами донора и акцептора близко к 3,0 Å (например, во льду расстояние между ними равно 2,8 Å). Это соответствует оптимальному Ван-дер-Ваальсову расстоянию между O и/или N-атомами, т. е. наличие стоящего между ними H-атома не приводит к увеличению расстояния между донорным O (или N) и акцепторным O (или N) атомами: H-атом их не расталкивает, а связывает.

У каждой H-связи один донор и один акцептор. При этом H почти всегда выступает донором только одной H-связи, а O может быть акцептором двух H-связей. В последнем луче возникает «вилочковая» водородная связь:

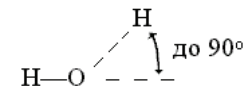


Так как при образовании «вилочки» сильно сближаются два H-атома (с зарядом +1/3 каждый), то ее суммарная энергия меньше, чем удвоенная энергия одинарной (как на предыдущей схеме) H-связи.

- В отличие от Ван-дер-Ваальсова взаимодействия, водородная связь весьма чувствительна к направлению, особенно к направлению донорной группы. Обычно валентная связь донора (N—H или O—H) прямо смотрит на тот акцептор (O- или N-атом), с коим водородная связь образуется:

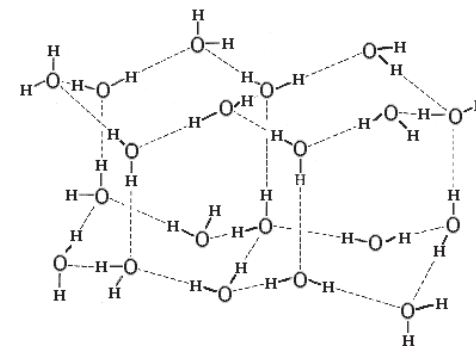


Причина — уже упомянутая малость атома H в O—H группе. К ориентации акцепторной группировки водородная связь менее чувствительна:



- Энергия водородной связи составляет около 5 ккал/моль. Эта оценка базируется на сравнении экспериментальных теплот испарения сходных соединений, из которых одни могут завязывать водородные связи, а другие — нет. Например, теплота испарения диметилэфира — H<sub>3</sub>C—O—CH<sub>3</sub> — составляет около 5 ккал/моль, а теплота испарения спирта — CH<sub>3</sub>—CH<sub>2</sub>—OH — около 10 ккал/моль. Эти вещества состоят из одних и тех же атомов — т. е. Ван-дер-Ваальсовы взаимодействия в них практически одинаковы — но спирт может завязывать водородную связь, а диметилэфир — нет (у него нет O—H-группы!). Каждая OH-группа может быть донором только одной H-связи, а ее O может такую связь акцептировать. Так как у каждой H-связи должен быть один донор и акцептор, в спирте приходится по одной H-связи на молекулу, т. е. цена одной водородной связи — около 10 ккал/моль (*спирт*) — 5 ккал/моль (*эфир*) = 5 ккал/моль.

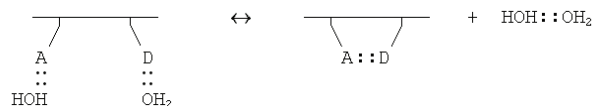
**Рис.4-1.** Нормальный лед («лед I<sub>h</sub>»); есть и другие формы льда, но они стабильны только при очень высоких давлениях). Штриховые линии — H-связи. В ажурной структуре льда видны небольшие полости, окруженные молекулами H<sub>2</sub>O. Картинка, с небольшими изменениями, взята из [6]. На ней каждый H-атом однозначно привязан ковалентной связью к одному O-атому. На самом же деле, как показывает рентгеноструктурный анализ, каждый H-атом может с равной вероятностью занимать два положения: либо он находится у того O-атома, с которым (на рисунке) он ковалентно связан, либо у того, с которым (на рисунке) он связан H-связью. Переход атома H от одного O-атома к другому вызывает перестройку ковалентных и водородных связей у обоих этих O-атомов и в их в окрестности. В результате вся сеть ковалентных и водородных связей во льду сильно флуктуирует, и это приводит к аномально высокой поляризуемости льда: его диэлектрическая проницаемость составляет 97 при 0 °С, превосходя диэлектрическую проницаемость жидкой воды (88 при 0 °С)





И пептидные группы главной цепи, и полярные боковые группы выступают донорами и акцепторами водородных связей. Они могут завязывать водородные связи друг с другом или с молекулами воды — и практически все они завязывают такие связи, так как энергия водородных связей, 5 ккал/моль, раз в десять превосходит энергию теплового движения, т. е. тепловое движение не может разрушить эти связи, разве что время от времени оно случайно нарушает малую часть из них.

Если внутримолекулярная связь между донором D и акцептором A водородной связи в белке завязывается в водном окружении, то эта связь замещает две водородные связи белка с молекулами воды (которые были до того, так как свято место пусто не бывает); и при этом еще создается одна связь между оторвавшимися молекулами воды:



Энергетический баланс этой реакции близок к нулю: две Н-связи было, две стало. Однако *энтропия* воды возрастает, так как вода больше не привязана к белковой цепи и связанные Н-связью молекулы могут плавать где хочется (а энтропия именно и связана с множественностью состояний системы). При этом рост энтропии за счет освобождения димера воды от белка — примерно такой же, как рост энтропии при уходе молекулы H<sub>2</sub>O из льда в жидкую воду (в обоих случаях одна частица обретает свободу плавать).

Энтропийный эффект выхода молекулы H<sub>2</sub>O из льда в воду можно оценить следующим образом. Как известно, в точке плавления (для льда при 0 °С, т. е. при 273 К) энтропийный выигрыш от плавления как раз компенсирует энергетический проигрыш от него. А этот проигрыш нам известен: 80 кал/г, т. е. 80 кал/г × 18 г/моль = 1440 кал/моль (18 г/моль — молекулярный вес H<sub>2</sub>O).

Значит, вследствие роста энтропии воды свободная энергия системы «белок в воде» падает примерно на 1,5 килокалории на каждый моль формирующихся внутримолекулярных водородных связей D::A.

Эта выделенная водой свободная энергия может полностью или частично компенсировать рост свободной энергии, сопровождающий падение энтропии цепи при фиксации групп D и A связью D::A (раньше они болтались как хотели, а теперь, связанные, — не могут). Позже мы увидим, что это падение свободной энергии воды примерно компенсирует энтропию фиксации мономера водородной связью (N–H::O) во вторичных структурах полипептидных цепей. В результате регулярная вторичная структура полипептидов в воде находится как раз на грани стабильности.

Итак, страшные слова — «энтропия» и «свободная энергия» — произнесены. Мой опыт показывает, что «энтропию» нормальный студент-биолог знает на слух (он слышал, что это — мера множественности состояний, мера беспорядка), а свободную энергию представляет совсем смутно... Так как нам понятие «свободная энергия» понадобится часто, я хочу уделить ей немного внимания сейчас, а потом, когда будет необходимо, поговорить о ней и об энтропии поподробнее.

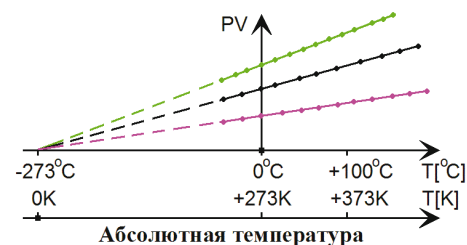
Начнем с простого примера: пусть молекула может находиться в двух состояниях, «а» и «б», например, здесь, в этой комнате, на высоте 200 м от уровня моря («а»), и в монастыре Далай-ламы в Гималаях на высоте 5 км («б»). Как соотносятся вероятности этих двух состояний — при условии, что (1) температура *T* одинакова в обоих помещениях, и (2) мы следим за нашей молекулой так долго, что она успеет побывать и здесь, и там?

Обычно все помнят формулу Больцмана, которая гласит:

$$\text{Вероятность\_быть\_в\_состоянии\_с\_энергией\_}E \text{ пропорциональна } \exp(-E/kT). \quad (4.2)$$

Физический смысл этой формулы заключается в том, что теплота окружающей среды (столкновения с другими молекулами) в определенной степени (в среднем пропорционально температуре среды *T*) возбуждает нашу молекулу, и это позволяет ей забираться в область более или менее высоких энергий. Все это мы подробнее обсудим позже, а пока я позволю себе надеяться, что вы эту формулу помните. Напомню только, что *k* — это постоянная Больцмана, а *T* — абсолютная температура в градусах Кельвина (она отсчитывается от «абсолютного температурного нуля», 0 К = –273,16 °С).

Нет, пожалуй, все-таки не стоит полагаться на то, что вы все это помните. Начнем с абсолютной температуры. Почему она, со времен Клапейрона и Кельвина, отсчитывается от –273 °С? Потому, что эта точка универсальна (в отличие от 0 °С, температуры замерзания одного из миллиардов веществ – воды). Экстраполируя, как то показано на схеме, произведение давления и объема (*P* × *V*) *любого* заданного количества *любого* разреженного газа к низким температурам, мы видим, *P* × *V* обращается в 0 при –273 °С.





Теперь выведем формулу Больцмана, по крайней мере для интересующего нас простейшего случая распределения молекул газа по высоте, когда  $E(h) = mgh$ , где  $m$  — масса молекулы газа,  $g$  — ускорение свободного падения, а  $h$  — высота.

Как известно (помните закон Клапейрона-Менделеева? минуту назад он использовался для определения абсолютной температуры), давление идеального газа  $P$  равно  $nkT$ , где  $n$  — число молекул газа в единичном объеме. Если  $T$  не меняется с высотой  $h$ , то  $dP/dh = (dn/dh)kT$ . С другой стороны, рассматривая столб газа единичного поперечного сечения, мы видим, что  $dP = (mgn)(-dh)$ , так как, поднявшись на  $dh$ , мы уменьшаем на  $(mgn)dh$  суммарный вес молекул газа, давящих на это единичное поперечное сечение. Поэтому  $dP/dh = (dn/dh)kT = -mgn$ . Следовательно,  $dn/dh = -(mg/kT)n$ , или  $d(\ln[n])/dh = -mg/kT$ , то есть  $n \sim \exp(-mgh/kT) = \exp(-E(h)/kT)$ .

Итак, применительно к интересующему нас вопросу — как соотносятся вероятности пребывания молекулы на разных высотах — формула Больцмана сводится к *барометрическому соотношению*

$$\begin{aligned} & \text{[Вероятность\_быть\_на\_высоте\_б]} \text{ относится к} \\ & \text{[вероятности\_быть\_на\_высоте\_а]} \\ & \text{как } \exp(-E_b/kT) \text{ относится к } \exp(-E_a/kT), \end{aligned} \quad (4.3)$$

где  $E_a$  — энергия молекулы в состоянии «а» (т. е. «здесь»),  $E_b$  — в состоянии «б» («на высоте 5 км»), а  $T$  — абсолютная температура (для простоты, как уже сказано, будем считать, что с высотой она не меняется).

Из-за силы тяжести энергия молекулы «здесь» ниже, чем «на высоте 5 км»; так что, согласно Больцману, «на высоте 5 км» молекула будет проводить меньше времени, чем «здесь».

[Раза в полтора-два меньше; подсчитайте сами при  $T=300$  К, вспомнив, что  $E = mgh$ , где  $m$  — средняя масса молекулы воздуха ( $\approx 30$  дальтон = 30 г/моль),  $g \approx 10$  м/с<sup>2</sup> — ускорение свободного падения,  $h \approx 5$  км — разность высот, а постоянная Больцмана  $k$  равна  $1,38 \cdot 10^{-23}$  Джоуля/градус (Дж = кг·м<sup>2</sup>/с<sup>2</sup> = 0,24 кал в расчете на частицу, т. е. 2 кал/град в расчете на моль ( $6 \cdot 10^{23}$ ) частиц). Напомню, что  $R = 2$  (кал/моль)/град. называется «газовой постоянной».]

То есть молекул на высоте в 5 км будет раза в два меньше, чем здесь. Точнее — меньше в *равных единицах объема*, например — внутри легких (в чем легко убедиться, попробовав подышать на разных высотах). Однако в *сумме* число молекул в монастыре Далай-ламы больше, чем в этой комнате: его монастырь куда больше, чем наша комната. То есть там молекула может находиться в большем числе положений — для свободно летающей молекулы воздуха это число пропорционально объему помещения. Физики говорят в этом случае, что в том монастыре возможно куда больше *микросостояний* молекулы, чем в нашей комнате. Поэтому вероятность того, что наша молекула находится *где-то* в монастыре Далай-ламы, относится к вероятности того, что она находится где-то в этой комнате, как

$$\begin{aligned} & \text{[Вероятность\_быть\_где-то\_в\_объеме\_«б»]} : \\ & \text{[вероятность\_быть\_где-то\_в\_объеме\_«а»]} = \\ & = [V_b \exp(-E_b/kT)] : [V_a \exp(-E_a/kT)]. \end{aligned}$$

Здесь  $V_a$  — объем суммы состояний «а» («нашей комнаты»), а  $V_b$  — объем суммы состояний «б» («его монастыря»). Вспомнив элементарную математику, а именно, что  $V$  можно представить как  $\exp(\ln V)$ , эту формулу можно переписать как

$$\begin{aligned} & \text{[Вероятность\_быть\_где-то\_в\_объеме\_«б»]} : \\ & \text{[вероятность\_быть\_где-то\_в\_объеме\_«а»]} = \\ & = [\exp(-E_b/kT + \ln V_b)] : [\exp(-E_a/kT + \ln V_a)] = \\ & = [\exp(-(E_b - T \times k \ln V_b)/kT)] : [\exp(-(E_a - T \times k \ln V_a)/kT)]. \end{aligned}$$

Последняя формула очень напоминает формулу (4.3), формулу Больцмана, но только она относится не к *единице объема*, а ко *всему объему системы*, и — внимание! — вместо энергий  $E$  в ней стоят величины  $E - T \times k \ln V$ . Так вот, величина  $F = E - T \times k \ln V$  называется *свободной энергией* нашей молекулы воздуха при заданном объеме  $V$  и температуре  $T$ . А величина  $S = k \ln V$  называется *энтропией* нашей молекулы в объеме  $V$ .

Это — в данном конкретном случае, когда «число доступных состояний» — это просто доступный для молекулы объем  $V$  ( $V_a$  — «нашей комнаты»,  $V_b$  — «его монастыря»).

В общем случае, энтропия  $S$  просто равна  $k \times$  [логарифм числа доступных состояний]. А свободная энергия  $F$  связана с энергией  $E$ , энтропией  $S$  и температурой  $T$  системы согласно общей формуле

$$F = E - TS. \quad (4.4)$$

Из двух состояний стабильнее, т. е. *вероятнее* то, у которого свободная энергия ниже:

$$\begin{aligned} & \text{[Вероятность\_быть\_где-то\_в\_объеме\_«б»]} : \\ & \text{[вероятность\_быть\_где-то\_в\_объеме\_«а»]} = \\ & = \exp(-F_b/kT) : \exp(-F_a/kT) = \exp[-(F_b - F_a)/kT]. \end{aligned} \quad (4.5)$$

- Иными словами, более вероятно, более стабильному состоянию системы отвечает то, где ниже  $F$ , а *самому* стабильному состоянию системы (при заданной температуре и фиксированном объеме этой системы) отвечает *минимум свободной энергии*  $F$ .

Таким образом, «свободная энергия» — это естественное расширение обычного понятия «энергии» на случай, когда система обменивается теплом с окружающей средой. Напомню: если тело *не* обменивается теплом со средой, его стабильное состояние отвечает минимуму его энергии (или, про-



ще: все, что может упасть, — в конце концов падает). А окружающее тепло возбуждает молекулы системы, они начинают захватывать многочисленные состояния с более высокой энергией (т. е. энтропия системы растет)— и, в результате, молекулы воздуха летают, а не падают на землю.

■ Иначе говоря:

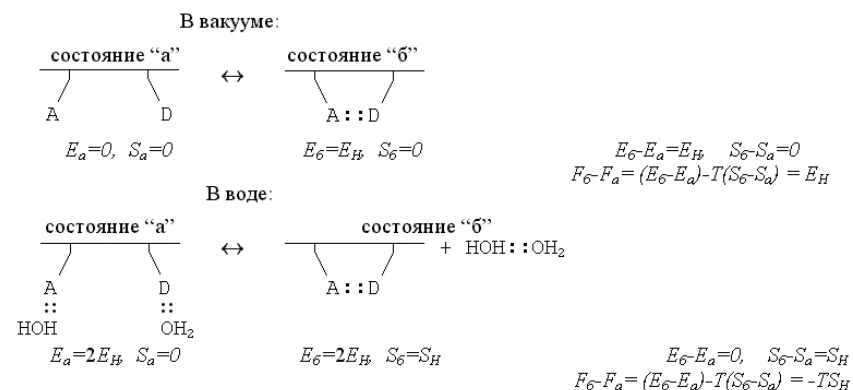
Изменение энергии,  $E_b - E_a$ , есть та работа, которую необходимо вложить для перевода тела из состояния «а» в «б», когда тело **не** обменивается теплом с окружающей средой.

А изменение свободной энергии  $F_b - F_a$  есть та работа, которую необходимо вложить для перевода тела из состояния «а» в состояние «б», когда тело постоянно обменивается теплом с окружающей средой.

Спустимся, однако, с Гималаев к белкам.

Итак — каков же баланс энергий, энтропий и свободных энергий в разобранном нами примере образования водородной связи в белковой цепи?

Для полной наглядности, сравним ход процесса в разных условиях:



Здесь  $E_H < 0$  — энергия Н-связи, а  $S_H > 0$  — энтропия перемещений и вращения свободного тела, т. е. свободной молекулы H<sub>2</sub>O или свободного димера HOH : : OH<sub>2</sub>. Водородные связи — воды с водой, воды с белком — стабильны, когда  $E_H - TS_H < 0$  (а если  $E_H - TS_H > 0$ , т. е. если они нестабильны, — то в этом случае перед нами была бы уже не жидкая вода, а пар).

Сравнение приведенных выше схем показывает, что водородные связи в белковой цепи как бы становятся — за счет изменения энтропии связывающейся воды — менее стабильными в водном окружении, чем в вакууме; действительно, в водном окружении их свободная энергия есть  $F_b - F_a = -TS_H$  — т. е. она меньше по абсолютной величине, чем в вакууме, где  $F_b - F_a = E_H$ .

Еще раз хочу подчеркнуть, что причина этого видимого ослабления — в том, что в водном окружении образовавшаяся внутри белковой цепи Н-связь *замещает* собой связь цепи с водой. И по той же причине водородные связи, стабилизирующие структуру белка *в воде*, носят энтропийную, а не энергетическую природу: энергии двух состояний цепи (с внутрицепочечной связью и без нее) примерно равны, и из этих двух состояний с примерно равной энергией стабильнее та, где выше энтропия, т. е. где больше число микросостояний. А их больше у оторванной молекулы воды, чем у связанной.

Обратите внимание: водородные связи в белковой цепи (*в водном окружении*) носят энтропийную, а не энергетическую природу именно потому, что энергия Н-связей очень высока: раз так, «свободные» от связей *в белке* доноры и акцепторы водородных связей в цепи всегда на деле не свободны от *всех* связей вообще, а связаны с водой. Оторвавшиеся же от белка при образовании внутрибелковой Н-связи молекулы H<sub>2</sub>O тут же связываются друг с другом, так что энергия компенсируется, и свободно-энергетический выигрыш Н-связей в белке идет только от множественности возможных микросостояний оторвавшихся молекул воды. Правда, чтобы связаться друг с другом, молекулы H<sub>2</sub>O должны пожертвовать часть своей обретенной свободы, часть энтропии — но лучше уж потерять небольшую энтропию, чем большую энергию!

То, что (1) молекулы воды сильно связаны друг с другом водородными связями; и (2) то, что эти связи образуются только при определенной взаимной ориентации молекул H<sub>2</sub>O, — определяет специфику воды как растворителя. Это приводит к разным интересным эффектам. Подробнее мы об этом поговорим в двух следующих лекциях.

## Лекция 5

Элементы термодинамики. Свободная энергия и химический потенциал. Гидрофобные взаимодействия. Их связь с необходимостью насыщения водородных связей в воде. Доступная воде неполярная поверхность аминокислот и их гидрофобность.

В первой части этой лекции я хочу продолжить разговор о термодинамике. Это нам пригодится, когда мы будем потом обсуждать специфику воды как растворителя. Свободная энергия погружения разных молекул в воду играет в этом обсуждении ключевую роль.

- Для изучения свободной энергии погружения молекулы в воду берут закупоренную пробирку, наполовину заполненную водой и наполовину — паром, воздухом или органическим растворителем, и смотрят, как изучаемые молекулы распределились между этими двумя фазами.

Мы уже видели, что разность свободных энергий  $F$  определяет предпочтительное состояние рассматриваемой системы (в данном случае — предпочтительное местонахождение рассматриваемой частицы) — согласно формуле

$$\begin{aligned} & \text{[Вероятность\_быть\_где-то\_в\_«б»]} : \\ & \text{: [вероятность\_быть\_где-то\_в\_«а»]} = \\ & = \exp[-(F_b - F_a)/kT]. \end{aligned} \quad (5.1)$$

Свободная энергия  $F = E - TS$  складывается из энергии системы  $E$  и ее энтропии  $S$ . С энергией все, видимо, понятно. С температурой — понятно скорее интуитивно, и к ней мы еще вернемся. С энтропией — сложнее, и хочется повторить еще раз.

В простейшем случае частицы в сосуде, который мы рассмотрели,  $S = k \ln V$ , где  $V$  — доступный для частицы объем.

В чем смысл вхождения в свободную энергию этой энтропии (в виде члена  $-TS$ )? В том, что — если посмотреть на  $\exp[-F/kT]$ , забыв о  $E$ , — то:

$\exp[-(-TS)/kT] = \exp[-(-T \ln V)/kT] = V$  есть просто доступный объем, а он определяет число возможных состояний частицы в пространстве. Чем выше энтропия, тем больше это число состояний, тем вероятнее поэтому нахождение частицы именно в этом объеме.

В общем случае (когда возможные состояния частицы ограничены не только ее столкновениями со стенками, окружающими объем  $V$ , но и, скажем, столкновениями с другими частицами), энтропия  $S$  частицы равна

$$S = k \times [\text{логарифм числа доступных для рассматриваемой частицы состояний}].$$

В молекулярной физике, биологии, химии обычно говорят не об одной частице, а о моле ( $6 \cdot 10^{23}$ ) частиц, и пишут

$$S = R \times [\text{логарифм числа доступных для одной частицы состояний}].$$

Здесь стоит еще раз напомнить, что различие между  $k$  и  $R$  только в том, что  $k$  употребляют, говоря об одной частице, а  $R$  — о моле частиц.

Величина

$$F = E - TS \quad (5.2)$$

называется, строго говоря, свободной энергией Гельмгольца. О ней просто рассказывать, и ее просто и удобно вычислять, так как эта величина относится к системе, которая (как та молекула, о которой мы говорили) находится в замкнутом и неизменном объеме.

Однако обычно на опыте измерения ведутся не при постоянном объеме  $V$ , а при постоянном внешнем давлении  $P$  (например — при атмосферном давлении). В этом случае измеряется не изменение чистой энергии  $E$  изучаемого вещества, а изменение его энтальпии  $H = E + PV$ : в изменение энтальпии  $H$ , помимо собственно изменения энергии тела  $E$ , входит работа, совершенная против постоянного внешнего давления  $P$  при изменении объема  $V$  этого тела.

На практике — той, с которой мы будем иметь дело по ходу этих лекций, — величину  $PV$  можно игнорировать, так как нас будут интересовать жидкости или твердые тела (где на каждую молекулу приходится небольшой объем), причем при невысоком (типа атмосферного) давлении. В этих условиях величина  $PV$  много-много меньше, чем тепловая энергия тела.

Действительно, даже для газов (где приходящийся на молекулу объем особенно велик),  $PV \approx RT \times [\text{число молей}]$  (помните закон Клапейрона-Менделеева?). Следовательно, в расчете на моль вещества, поправка  $PV = 1RT \approx 0,55 \div 0,75$  ккал/моль при температурах от 0 до 100°C, т. е. при  $T = 273 - 373$  К. То есть даже в газах она, как правило, невелика по сравнению с интересующими нас эффектами, составляющими обычно несколько ккал/моль. Для жидких и твердых тел — при давлении порядка атмосферного — роль поправки  $PV$  в величинах  $H$  еще много

меньше: здесь моль вещества занимает не около 25 литров, как газ при давлении в 1 атм. и комнатной температуре, а лишь малую долю литра [ $\approx 1/55$  литра для  $\text{H}_2\text{O}$ ,  $\approx 1/10$  литра для  $(\text{CH}_2)_6$  и т. д.].

Поэтому для нас будет в дальнейшем несущественна разница между  $H$  и  $E$ . Обе эти величины я буду обычно называть просто «энергия».

Столь же несущественным для нас будет различие между *свободной энергией Гельмгольца*  $F = E - TS$  и *свободной энергией Гиббса*

$$G = H - TS = (E + PV) - TS = F + PV. \quad (5.3)$$

И то, и другое (и  $G$ , и  $F$ ) мы, не делая различия, будем именовать просто свободной энергией.

Пожалуй, стоит запомнить, что, обсуждая процессы, протекающие в замкнутом объеме, следует употреблять буквы  $E$  и  $F$ , а обсуждая процессы, протекающие при постоянном внешнем давлении, — буквы  $H$  и  $G$ .

И вот еще что. Работа, потребная на добавление в систему новой частицы, называется *химическим потенциалом*. Мы его рассмотрим чуть позже.

Одно замечание. Если добавлять в систему (например, в жидкость) частицу за частицей при постоянном объеме, то работа, потребная на добавление новой частицы, меняется с ростом числа частиц в системе (в запечатанную бутылку нетрудно впрыснуть каплю; — а вторую? а третью? и т. д.). Но если делать то же самое *при постоянном давлении*, то работа, потребная на добавление новой частицы, **не меняется** с ростом числа частиц в системе. То есть  $E$  и  $F$  при постоянном объеме **не** пропорциональны числу частиц в системе, а  $H$  и  $G$  — при постоянном давлении — пропорциональны. Последнее, при большом числе одинаковых частиц в системе, позволяет определить химический потенциал в очень удобном виде — как «свободную энергию Гиббса в расчете на одну частицу».

Следует запомнить, что равновесному, *стабильному* состоянию любой системы при заданной температуре ( $T = \text{const}$ ) соответствует минимум величины  $F = E - TS$ , *если* равновесие устанавливается при фиксированном объеме, и минимум величины  $G = H - TS$ , *если* оно устанавливается при фиксированном внешнем давлении.

При малом изменении состояния системы ее свободная энергия слегка меняется:

$$F \rightarrow F + dF = F + dE - TdS - SdT$$

или

$$G \rightarrow G + dG = G + dH - TdS - SdT. \quad (5.4)$$

Это значит, в частности, что все возможные перестройки системы в окрестности ее стабильного состояния (а они могут изменять и энергию системы, и ее энтропию) подчиняются следующим уравнениям, характеризующим минимум функции (как мы знаем, точка минимума функции

обладает тем свойством, что небольшое отклонение от нее практически не меняет величину этой функции):

1) при  $V = \text{const}$  стабильному состоянию (при заданном  $T$ ) соответствует минимум  $F$ , т. е.

$$dF|_{V=\text{const}} = dE|_{V=\text{const}} - TdS|_{V=\text{const}} = 0 \quad (5.5)$$

[здесь мы учли, что  $dT=0$  при  $T = \text{const}$ , т. е.  $SdT=0$  в выражении (5.4)];

2) при  $P = \text{const}$ , когда стабильному состоянию (при заданном  $T$ ) соответствует минимум  $G$ ,

$$dG|_{P=\text{const}} = dH|_{P=\text{const}} - TdS|_{P=\text{const}} = 0. \quad (5.6)$$

Из этих уравнений следует термодинамическое определение абсолютной температуры:

$$T = (dE/dS)|_{V=\text{const}} = (dH/dS)|_{P=\text{const}}. \quad (5.7)$$

Я понимаю, что сейчас это определение получено довольно формально, и его физический смысл не очевиден. Поэтому я вернусь к этому вопросу в одной из ближайших лекций.

■ Остановимся теперь на *химическом потенциале* — величине, определяющей термодинамические характеристики не системы в целом, а одной молекулы в этой системе.

Если добавлять в систему молекулу за молекулой при *постоянном давлении*, то на добавление каждой новой частицы надо затратить в точности ту же работу, что на добавление любой предыдущей: объем системы будет расти, а плотность системы — и интенсивность взаимодействий в ней — меняться не будет. Поэтому термодинамическое состояние молекулы в большой однородной системе удобно определять величиной свободной энергии Гиббса  $G$ , деленной на число молекул  $N$ ,

$$\mu = G/N, \quad (5.8)$$

называемой *химическим потенциалом* (а так как в жидкой или твердой фазе и невысоких давлениях  $F \approx G$ , то здесь  $\mu \approx F/N$ ). (Если  $N$  означает не число молекул, а, как обычно в физической химии, число молей молекул, то и  $\mu$  относится не к одной молекуле, а к молю молекул.)

Химический потенциал или, что то же самое, свободная энергия Гиббса в расчете на одну молекулу нам пригодится во второй части сегодняшней лекции, когда речь пойдет о распределении молекул между фазами. Дело в том, что молекулы перетекают из той фазы, где их химический потенциал выше, в ту, где их химический потенциал ниже, это понижает

общую свободную энергию системы и приближает ее к равновесию. А в равновесии химический потенциал молекул в одной фазе равен химическому потенциалу тех же молекул в другой фазе.

■ Для будущего нам еще понадобятся две формулы.

Во-первых, определение теплоемкости, показывающей рост энергии системы с температурой:

$$C_p = dH/dT |_{p=\text{const}} \quad (5.9)$$

(это — при постоянном давлении; можно считать теплоемкость и при постоянном объеме, но эти тонкости нам не нужны).

Во-вторых, — связь энтропии со свободной энергией:

$$S = -dG/dT |_{p=\text{const}} = -dF/dT |_{V=\text{const}} \quad (5.10)$$

Формула (5.10) — важнейшая формула термодинамики. Она прямо следует из того, что малое приращение свободной энергии  $dG = d(H-TS) = dH - TdS - SdT$  (и  $dF = d(E-TS) = dE - TdS - SdT$ ), в то время как — в равновесии системы —  $dH - TdS = 0$  (и  $dE - TdS = 0$ ) в силу *термодинамического определения температуры*:  $T = dH/dS = dE/dS$  [см. (5.7)].

Формула (5.10) показывает, что свободная энергия достигает своего минимального (или максимального) значения при той температуре  $T$ , когда  $S(T) = 0$ .

Полезно знать также, что величина  $G/RT$  достигает своего минимального (или максимального) значения, когда  $H(T) = 0$ , так как

$$d(G/RT)/dT = (dG/dT)/RT - G/RT^2 = -S/RT - (H-TS)/RT^2 = -H/RT^2. \quad (5.11)$$

Замечания.

1. Из уравнений (5.3) и (5.10) следует, что  $H = G + TS = G - T(dG/dT)$ , и поэтому (см. (5.9))

$$C_p = -T(d^2G/dT^2) |_{p=\text{const}} \quad (5.12)$$

2. Нас никогда не интересует величина энергии, энтропии или свободной энергии сама по себе; нас интересуют *изменения* этих величин. В самом деле, гравитационная энергия может отсчитываться от уровня моря, от уровня пола, от центра Земли и т. д. При этом величина гравитационной энергии будет разной, но это не важно: важна только *разность* значений этой энергии при разных положениях частицы. Точно так же, определяя энтропию частицы как  $k \ln V$ , мы можем измерять объем  $V$  в литрах, в кубических аршинах и т. д. Численные значения этой энтропии будут разными, но это не важно: важна только *разность* значений энтропии частицы в двух объемах ( $V$  и  $V'$ ). А эта разность,  $\Delta S = k \ln(V/V')$ , не зависит от единицы измерения объема.

После этого вступления мы готовы поговорить о специфике воды как растворителя. Прежде всего займемся так называемым *гидрофобным эффектом*.

■ Гидро\_фобия — в переводе — водо\_боязнь. Кто же боится воды? Оказывается, все неполярные молекулы: и инертные газы (аргон, ксенон), и водород, и все чисто углеводородные молекулы (метан, этан, бензол, циклогексан и т. д.). Именно поведение углеводородов в воде нас и будет занимать: ведь в белках есть много аминокислот с углеводородными боковыми группами. И именно эти группы, боясь воды и убегая от нее, образуют *гидрофобное ядро белковой глобулы*.

Итак, в чем же — экспериментально — заключается гидро\_фобия? Если мы измерим концентрацию метана  $\text{CH}_4$  в воде и в паре над ней, то в паре концентрация метана будет раз в 10 *больше* (точное число зависит от температуры, но при комнатной температуре, около 20–40°C, — примерно в 10 раз). Приблизительно то же соотношение будет и для  $\text{H}_2$ , и для пропана  $\text{CH}_3\text{CH}_2\text{CH}_3$ : все они боятся (фобия!) смешиваться с водой, и всех их *больше* в паре, чем в воде. А вот этиловый спирт  $\text{CH}_3\text{CH}_2\text{OH}$  и вода хорошо смешиваются, и разделить их довольно трудно — это все знают; но спирт — полярная молекула (ее полярная O–H-группа способна завязывать H-связи). А чисто *неполярным* молекулам лучше уж в вакууме (пар — это почти вакуум), чем в воде... И это — *несмотря* на Ван-дер-Ваальсово притяжение *любых* молекул, в том числе  $\text{H}_2$  или  $\text{CH}_4$ , к  $\text{H}_2\text{O}$ !

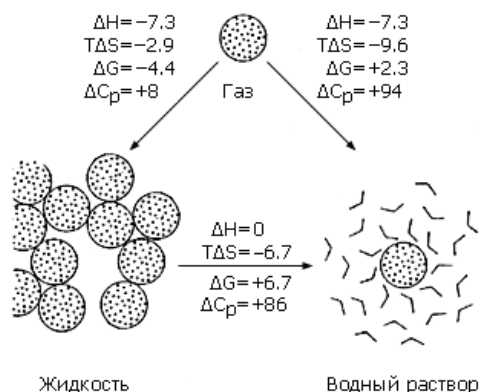
Рассмотрим на одном примере (рис. 5-1) характерные термодинамические эффекты растворения углеводородов в воде. Приведенные на рис. 5-1 термодинамические данные (свободные энергии, энергии и энтропии переноса из фазы в фазу) основаны на изучении равновесного распределения молекул циклогексана,  $(\text{CH}_2)_6$ , между тремя фазами: паром, водным раствором и жидким циклогексаном.

Свободная энергия переноса молекулы из неполярной жидкости в воду определяется как

$$\Delta G_{\text{неполярная жидкость} \rightarrow \text{вода}} = -RT \ln(X_{\text{в воде}}/X_{\text{в непол.}}). \quad (5.13)$$

Здесь  $X_{\text{в воде}}$  и  $X_{\text{в непол.}}$  — равновесные концентрации изучаемых молекул в водном растворе и в контактирующей с ним неполярной жидкости. Величина  $\Delta G_{\text{неполярная жидкость} \rightarrow \text{вода}}$  — разность между «потенциалами средней силы», действующими на нашу молекулу в воде и в неполярной жидкости. Потенциал средней силы создается всеми взаимодействиями нашей молекулы с ее окружением; он содержит как энергетические, так и энтропийные составляющие, возникающие из этих взаимодействий. Так как концентрация есть число молекул в данном объеме, то уравнение (5.13) соответствует Больцмановскому распределению по двум фазам с разными значениями потенциала средней силы.





**Рис. 5-1.** Термодинамика переноса типичной неполярной молекулы, циклогексана ( $\text{CH}_2$ )<sub>6</sub>, из пара (вверху) в воду (справа), а также в жидкий циклогексан (слева), и из него в воду. Цифры примерно соответствуют процессам при 25°C (т.е. при  $T \approx 300\text{ K}$ ; при этом  $RT \approx 0,6$  ккал/моль).  $\Delta H$  (как и энергия, она измеряется в ккал/моль) — изменение энтальпии в рассматриваемом переносе в расчете на моль молекул;  $\Delta S$  — соответствующее изменение энтропии ( $T\Delta S$  — вклад изменения энтропии в свободную энергию; он тоже измеряется в ккал/моль);  $\Delta G = \Delta H - T\Delta S$  (ккал/моль) — это изменение свободной энергии Гиббса в расчете на моль переносимых молекул;  $\Delta C_p$  [кал/(моль·K)] — изменение теплоемкости в расчете на моль переносимых молекул. Картинка взята из [15], но все числа уточнены с использованием справочников и уравнений (5.10)–(5.15)

При 25°C и низком (~1 атм.) давлении,  $X_{\text{в.непол.}} = 9,25$  моль/л для чистого жидкого ( $\text{CH}_2$ )<sub>6</sub>, и  $X_{\text{в.воде}} = 0,0001$  моль/л для его насыщенного раствора в воде. Значит,  $\Delta G_{\text{неполярная жидкость} \rightarrow \text{вода}} = +6,7$  ккал/моль.

Свободная энергия переноса нашей молекулы из газа в неполярную жидкость или в воду определяется таким же образом

$$\Delta G_{\text{газ} \rightarrow \text{неполярная жидкость}} = -RT \ln(X_{\text{в.непол.}}/X_{\text{в.газе}}), \quad (5.14)$$

$$\Delta G_{\text{газ} \rightarrow \text{вода}} = -RT \ln(X_{\text{в.воде}}/X_{\text{в.газе}}), \quad (5.15)$$

где  $X_{\text{в.газе}}$  — равновесная концентрация наших молекул в газе, контактирующем с неполярной жидкостью или с водой. Давление насыщенного пара ( $\text{CH}_2$ )<sub>6</sub> при 25°C — около 0,05 атм., что соответствует  $X_{\text{в.газе}} = 0,002$  моль/л (т.е.  $X_{\text{в.газе}}$  — в 20 раз больше, чем  $X_{\text{в.воде}}$ ). Соответствующие величины  $\Delta G$  представлены на рис. 5-1.

Так как потенциал средней силы, действующий на молекулу в разреженном газе ( $G_{\text{в.газе}}$ ), практически равен нулю, то величина  $\Delta G_{\text{газ} \rightarrow \text{неполярная жидкость}} = G_{\text{в.непол.}} - G_{\text{в.газе}}$  очень близка к  $G_{\text{в.непол.}}$ , т.е. к потенциалу средней силы, действующему на нашу молекулу в неполярной жидкости; а потенциал

средней силы, действующий на нашу молекулу в воде,  $G_{\text{в.воде}}$ , практически равен  $\Delta G_{\text{газ} \rightarrow \text{вода}}$ .

Зная величину  $\Delta G$  и ее изменение с температурой, можно найти величины  $\Delta S$ ,  $\Delta H$ ,  $\Delta C_p$  по формулам (5.10), (5.11) и (5.12).

Итак, приведенные на рис. 5-1 цифры говорят, что *по энергии* молекулы ( $\text{CH}_2$ )<sub>6</sub> притягиваются к воде так же сильно, т.е. с той же энергией  $\Delta H = -7,3$  ккал/моль, как к жидкому циклогексану, но в воду они ни за что не хотят лезть (а в жидкий циклогексан — с большим удовольствием). И виной этой гидрофобии, как показывает рис. 5-1, энтропия: уж очень она падает, когда молекула ( $\text{CH}_2$ )<sub>6</sub> погружается в воду...

Почему же неполярные молекулы, скажем,  $\text{CH}_4$  или ( $\text{CH}_2$ )<sub>6</sub> убегают из воды? Видимо, дело в том, что молекулы  $\text{H}_2\text{O}$  умеют завязывать водородные связи, а неполярные  $\text{H}_2$  или  $\text{Ar}$ ,  $\text{CH}_4$  или ( $\text{CH}_2$ )<sub>6</sub> — нет. Это подтверждается и тем, что, как известно, этиловый спирт  $\text{CH}_3\text{CH}_2\text{OH}$  тоже (как и  $\text{H}_2\text{O}$ ) умеющий завязывать Н-связи, из воды не бежит...

Можно наивно предположить, что все дело в том, что  $\text{CH}_4$  или ( $\text{CH}_2$ )<sub>6</sub>, внедряясь в воду, рвет водородные связи молекул  $\text{H}_2\text{O}$ . Но не все так просто! Если бы ( $\text{CH}_2$ )<sub>6</sub>, внедряясь в воду, просто рвал водородные связи в ней, энергия раствора *сильно росла бы* при внедрении в него ( $\text{CH}_2$ )<sub>6</sub>, — а на самом деле она даже *падает* (см. рис. 5-1) — на восемь килокалорий на моль внедренного ( $\text{CH}_2$ )<sub>6</sub>... А вместо роста энергии *падает энтропия* (и сильно падает:  $T\Delta S = -9,6$  ккал/моль!).

И именно это не дает циклогексану растворяться в воде: все решает свободная энергия  $G = H - TS$ , а она *растет* не только с ростом энергии, но и с *падением* энтропии  $S$  ( $S$  входит в  $G$  со знаком минус!). То есть сильное падение энтропии  $S$  ведет, даже при некотором падении энергии  $H$ , к росту свободной энергии  $G$  и, в результате (см. формулу 5.1), — к *падению вероятности* пребывания молекул в соответствующем состоянии (иными словами — к падению их концентрации в этом состоянии).

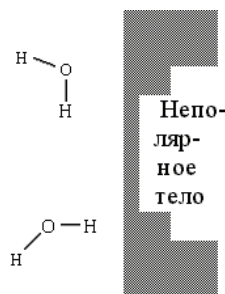
Теперь мы дошли до главного физического вопроса: Почему же энтропия воды падает?

Потому, что неполярная молекула загоразивает своей поверхностью часть заполненного водой пространства для каждой прилегающей к неполярной поверхности молекулы  $\text{H}_2\text{O}$  и тем самым угрожает порвать Н-связи этой молекулы, если она будет смотреть на такую поверхность, скажем, своим Н-атомом (см. рис. 5-2; напомним, что Н-связи — направленные, они завязываются лишь при определенной взаимной ориентации молекул: тогда, когда О–Н-связь утыкается в О-атом другой молекулы воды).



Вспомним также, что в воде завязаны почти все возможные водородные связи, так что жертвовать возможностью их образования без большой потери свободной энергии воды нельзя, а каждая такая связь требует, чтобы О–Н-группа прямо смотрела на тот О-атом, с которым образуется связь. Чтобы избежать потери Н-связей, молекулы  $\text{H}_2\text{O}$  изворачиваются (см. верхнюю молекулу на рис. 5-2) — и частично замораживают на гидрофобной поверхности свободу своих тепловых движений. Они Н-связей не теряют, но платят за это своей энтропией. При этом потеря свободной энергии составляет около 0,2 ккал на моль приповерхностных молекул воды, — на порядок меньше, чем стоила бы потеря их Н-связей.

**Рис. 5-2.** Молекулы  $\text{H}_2\text{O}$  у поверхности неполярного тела. Верхняя молекула способна завязать обе Н-связи, но за ее «подходящее» расположение относительно поверхности тела приходится платить падением энтропии. Нижняя молекула теряет одну Н-связь с водой: ее О-Н-группа смотрит прямо на неполярную преграду



При невысокой температуре частично замороженные у поверхности неполярного тела молекулы  $\text{H}_2\text{O}$  уже легче вступают между собой в Н-связи. При этом, если температура низка (менее  $20^\circ\text{C}$ , а для многих углеводородных гостей — до  $60^\circ\text{C}$ ), приповерхностные водородные связи даже немного выигрывают в энергии (теперь они меньше деформированы движениями полужамороженных приповерхностных молекул воды), но этот выигрыш не может компенсировать потерю энтропии, сопровождающую замораживание этих приповерхностных молекул.

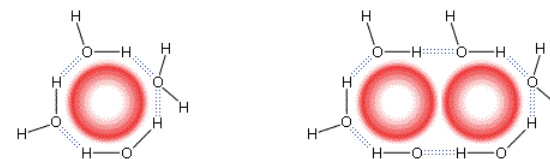
Обратите внимание: снова эффект носит энтропийную, а не энергетическую природу и именно *потому*, что энергия Н-связей очень высока: раз так, молекулам  $\text{H}_2\text{O}$  лучше заморозиться (хотя это все равно термодинамически невыгодно) и пожертвовать частью своей свободы, т. е. часть энтропии, чем потерять большую энергию водородной связи.

Подчеркну, что результирующий энтропийный эффект имеет *тот же знак*, что наивно ожидаемый энергетический эффект, но *меньше* его по величине.

То, что молекулы  $\text{H}_2\text{O}$  у поверхности неполярного тела как бы заморожены, дополнительно подтверждается аномально высокой теплоемкостью циклогексана (и других углеводородов) в воде. В водном окружении у  $(\text{C}_6\text{H}_{12})_6$  в 10 (!) раз (см. рис. 5-1) выше связанная с межмолекулярными

взаимодействиями дополнительная теплоемкость, чем в окружении своих братьев — циклогексанов. Точнее, высока теплоемкость не самого углеводорода, а его льдоподобной гидратной оболочки в воде: этот айсберг как бы подплавляется с ростом температуры, отсюда и его аномально высокая теплоемкость.

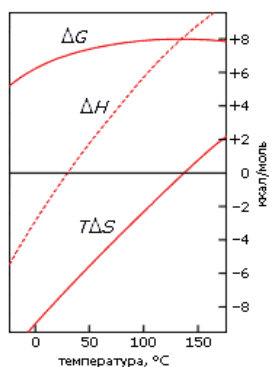
При этом важно помнить, что замороженные и связанные Н-связями поверхностные молекулы  $\text{H}_2\text{O}$  ориентированы друг относительно друга *не совсем так*, как в обычном льду. Во льду молекулы должны стоять регулярно, в узлах кристаллической решетки: им надо сформировать огромный трехмерный кристалл. На поверхности же они могут стоять как угодно — лишь бы завязались водородные связи. Трансляционная симметрия кристаллической решетки этим молекулам не указ, так как образованный ими кристаллик, микроайсберг и не собирается расти до бесконечности, ему лишь бы тонким слоем облепить собой внедренную гидрофобную молекулу (рис. 5-3), или группу таких молекул. В последнем случае между этими молекулами и образуется **гидрофобная связь**.



**Рис. 5-3.** Схема нерегулярной упаковки связанных водородными связями молекул  $\text{H}_2\text{O}$  вокруг одной неполярной молекулы (слева) и вокруг пары таких молекул (справа). В последнем случае образуется гидрофобная связь

Экстремальный случай упорядочения воды при гидрофобном эффекте демонстрируют так называемые *клатраты* — кристаллы из воды и неполярных молекул. Клатраты далеко не столь стабильны, как известные вам из химии кристаллогидраты, сложенные из воды и полярных молекул. Клатраты возникают только при низких (ближе к  $0^\circ\text{C}$ ) температурах и высоких давлениях, когда — из-за давления — в воду внедряется много молекул неполярного газа. В клатратах молекулы  $\text{H}_2\text{O}$  насыщают все свои водородные связи, как во льду, но этот лед имеет другую, чем обычный лед, геометрию. При этом создается как бы кристалл, где квазилед обволакивает и держит в своих ячейках регулярно расположенные молекулы газа. Кстати, говорят, что в клатратах содержится много больше природного газа, чем в обычных его залежах, и, возможно, добыча газа из залежей клатратов (правда, с большой глубины, где высокое давление позволяет им существовать) — дело будущего.

- Гидрофобный эффект довольно сильно зависит от температуры (рис. 5-4). Особенно сильно температура сказывается не на величине самой свободной энергии  $\Delta G$ , а на величине (и даже знаке!) ее компонент  $\Delta H$  и  $\Delta S$ .



**Рис. 5-4.** Термодинамика переноса типичной неполярной молекулы, пентана  $C_5H_{12}$ , из жидкого пентана в воду при разных температурах. Свободная энергия переноса  $\Delta G = \Delta H - T\Delta S$ , ее энтальпийная  $\Delta H$  и энтропийная  $T\Delta S$  составляющие даны в килокалориях в расчете на моль переносимых молекул.  $\Delta G$  достигает своего максимума в точке, где  $\Delta S=0$ , а концентрация насыщенного раствора пентана в воде [она пропорциональна  $\exp(-\Delta G/RT)$ ] минимальна в точке, где  $\Delta H=0$ . Картинка, с небольшими изменениями, взята из P. L. Privalov & S. J. Gill, *Adv. Protein Chem.* (1988) **39**:191–234

При повышении температуры водородные связи в поверхностном слое плавятся, но гидрофобный эффект от этого сначала (до 140°C) только растет — ведь термодинамически невыгодное упорядочение поверхностных молекул воды остается, а их выгодные Н-связи гибнут.

Если при комнатной температуре гидрофобный эффект создается только энтропией, то при повышении температуры все большую и большую роль в этом эффекте начинает играть энергия утерянных (на поверхности неполярной молекулы) водородных связей. Но сам эффект (определяемый величиной  $\Delta G$ ) все возрастает, пока  $T\Delta S < 0$ , т. е. до 140°C.

При еще более высоких температурах в воде (теперь она остается жидкой только под высоким давлением) и так разорвано слишком много Н-связей, гидрофобная поверхность мешает их образованию все меньше и меньше, и гидрофобный эффект начинает спадать.

На рисунке 5-4 я специально привел данные по переносу неполярной молекулы в воду только из неполярного растворителя, а не из пара. Ведь интересующий нас в белках гидрофобный эффект связан с переносом аминокислотных остатков в воду из ядра белка, а оно моделируется скорее неполярным растворителем, чем паром.

- Гидрофобности аминокислот измеряются на опыте по величине констант равновесного распределения аминокислот между водой и слабо полярным растворителем, в качестве которого часто берутся высокомолекулярные спирты (например, октанол) или диоксан. Эти опыты довольно сложны, так как одни аминокислоты практически не растворяются в воде,

а другие — в органике (например, в чисто неполярном циклогексане или бензоле практически не растворяются полярные аминокислоты). Приходится специально подбирать вместо неполярного слабополярный растворитель (удовлетворительный и для более, и для менее полярных аминокислот) и идти на другие ухищрения. При этом, в зависимости от используемых слабополярных растворителей, получаются несколько расходящиеся по величине результаты, особенно для заряженных и сильно полярных аминокислот. Однако качественно они неплохо согласуются.

О гидрофобности аминокислот поговорим чуть позже, а для начала полезно увидеть немного цифр, характеризующих гидрофобный эффект при комнатной температуре для неполярных групп, похожих на те, что встречаются в белках.

**Таблица 5-1. Характерные термодинамические параметры переноса гидрофобных групп из неполярной жидкости в водный раствор при 25°C.**  
Данные взяты из C. Tanford, *The Hydrophobic Effect* (1980)

Вещество	Перенос	$\Delta G$ ккал/моль	$\Delta H$ ккал/моль	$T\Delta S$ ккал/моль	$\Delta C_p$ кал/(моль×К)
Этан $(CH_3)_2$ (ср. с боковой группой Ala: $-CH_3$ )	Бензол → вода	+3.6	-2.2	-5.8	+59
	$CCl_4$ → вода	+3.8	-1.8	-5.4	+59
Бензол $C_6H_6$ (ср. с боковой группой Phe: $-CH_2-C_6H_5$ )	Бензол → вода	+4.6	+0.5	-4.1	+54
	Толуол $C_6H_5CH_3$ (ср. с боковой группой Phe)	+5.4	+0.4	-5.8	+63

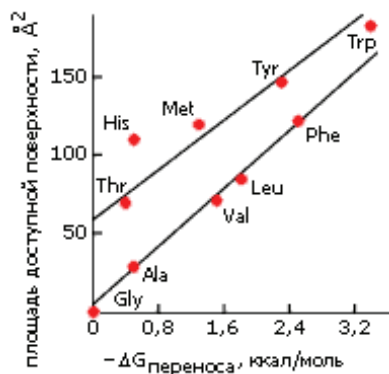
Таблица 5-1 сразу показывает, что  $\Delta G$  (в отличие от  $\Delta H$  и  $\Delta S$ ) растет с размером гидрофобной молекулы. Растет — но как именно? Более подробный анализ гидрофобности разных неполярных молекул показал, что их гидрофобная свободная энергия  $\Delta G$  растет примерно пропорционально *доступной для воды поверхности* гидрофобных молекул. Смысл и способ построения доступной для воды поверхности демонстрируется рис. 5-5.

Свободная энергия гидрофобного эффекта составляет около 0,02–0,025 ккал/моль на каждый квадратный ангстрем «доступной поверхности» неполярной молекулы, перенесенной из неполярного растворителя в воду. В частности, для тех молекул, о которых мы уже говорили: для бензола  $\Delta G \approx 4,6$  ккал/моль при доступной поверхности примерно в  $200 \text{ \AA}^2$ ; для циклогексана  $\Delta G \approx 6,7$  ккал/моль при доступной для воды поверхности примерно в  $300 \text{ \AA}^2$ .



**Рис. 5-5.** Доступная для воды поверхность внедренной в воду молекулы. Точки показывают центры граничащих с водой атомов; сплошная линия — их Ван-дер-Ваальсовы оболочки. Молекула воды представляется как шарик радиусом 1,4 Å. «Доступная для воды поверхность» создается центром этого шарика, когда он обкатывает погруженную в воду молекулу, касаясь Ван-дер-Ваальсовых поверхностей ее внешних атомов. Картинка взята из [3] и адаптирована

Та же закономерность наблюдается и для гидрофобных аминокислотных остатков (рис. 5-6).

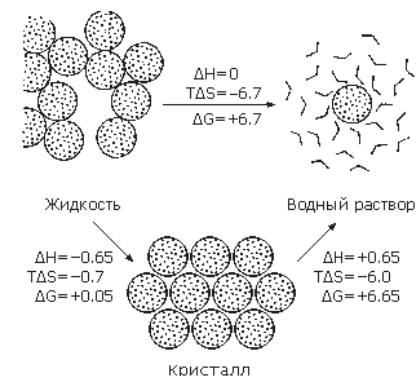


**Рис. 5-6.** Размер доступной для воды поверхности боковых цепей аминокислотных остатков и гидрофобность этих боковых цепей. [Доступная поверхность боковой цепи остатка X равна доступной поверхности аминокислоты X за вычетом доступной поверхности остатка Gly, не имеющего боковой цепи; гидрофобность боковой цепи остатка X равна экспериментально измеренной гидрофобности остатка X минус гидрофобность остатка Gly]. Остатки Ala, Val, Leu, Phe имеют чисто углеводородные боковые цепи. У Thr и Tyr в боковой цепи есть еще по одному O-атому, у Met — S-атом, у Trp — N-атом, у His — два N-атома; поэтому у всех них размер доступной неполярной поверхности меньше полного размера доступной поверхности. Картинка взята из [3] и адаптирована

У боковых групп с полярными атомами гидрофобный эффект (при той же общей доступной для воды поверхности) меньше, чем у чисто неполярных

групп. Однако, если взять только ту доступную для воды поверхность, которая создается неполярными атомами (т. е. вычесть из общей доступной для воды поверхности приблизительно по  $50 \text{\AA}^2$  на каждый полярный атом), мы получим для всех групп примерно одну и ту же зависимость.

■ Гидрофобный эффект играет очень важную роль в поддержании стабильности белковой структуры. Именно он сворачивает белковую цепь в компактную плотную глобулу. Рисунок 5-7 показывает, что свободная энергия перехода гидрофобной группы из воды в гидрофобную же жидкость велика — она составляет несколько килокалорий на моль, в то время как свободная энергия затвердевания этой же неполярной жидкости близка к нулю при физиологических температурах. [Точнее: Для сравнительно низкомолекулярных углеводородов, таких, как циклогексан, показанный на рис. 5-7, свободная энергия отвердевания даже положительна, и это не дает им отвердеть при комнатной температуре. Отвердеванию противится энтропия, в частности, энтропия вращений и перемещений молекул по жидкости, — в жидкости каждая молекула плавает более или менее свободно, а в твердом теле она зажата кристаллической решеткой. Энтропия перемещений молекулы, грубо говоря, не зависит от ее размера в отличие от энтальпии, которая растет с ростом числа контактов одной молекулы с другими, т. е. с ростом ее поверхности. В белковой цепи роль энтропии перемещений меньше, так как аминокислотные остатки связаны цепью, т. е. не могут двигаться независимо друг от друга, и это способствует отвердеванию белка.]



**Рис. 5-7.** Термодинамика переноса типичной неполярной молекулы, циклогексана, из жидкой фазы в твердую и в водный раствор. Цифры соответствуют процессам при 25 °C. Величины  $\Delta H$ ,  $T\Delta S$  и  $\Delta G$  даны в ккал/моль. Картинка взята из [6], но все числа уточнены с использованием справочников и уравнений (5.10)–(5.15)

Если даже пренебречь всей вообще энтропией кристаллизации циклогексана в свободной энергии его отвердевания,  $\Delta G_{\text{жидкий циклогексан-кристалл}}$ , т. е. положить  $\Delta G_{\text{жидкий циклогексан-кристалл}} \approx \Delta H_{\text{жидкий циклогексан-кристалл}} = -0,65$  ккал/моль, — то и в этом случае видно, что для циклогексана термодинамический эффект отвердевания много слабее гидрофобного эффекта, приводящего к конденсации растворенных в воде молекул в жидкую каплю, так как  $\Delta G_{\text{вода} \rightarrow \text{жидкий циклогексан}} = -6,7$  ккал/моль.

■ Таким образом, гидрофобный эффект берет на себя, грубо говоря, 90% работы по созданию белковой глобулы. Но он сам по себе не может создать нативный твердый белок. Он создает лишь расплавленную белковую глобулу, о которой речь впереди. Отвердевает белок, как и все органические жидкости, под действием Ван-дер-Ваальсовых сил, а также водородных и ионных связей — взаимодействий более специфичных и более чувствительных к деталям атомной структуры, чем простая водобоязнь. Но это — окончательная огранка, а большая часть работы, вся черновая работа падает на гидрофобный эффект.

## Лекция 6

Влияние водного окружения на электростатические взаимодействия. Электрическое поле у поверхности и внутри белка. Диэлектрическая проницаемость. Экранировка зарядов в солевых растворах. Измерение электрических полей в белках при помощи белковой инженерии. Дисульфидные связи. Координационные связи.

Теперь мы обсудим электростатические взаимодействия, точнее — те особенности, которые им придает наличие белковой глобулы и воды вокруг нее. Казалось бы, что здесь обсуждать: вы и так помните, что заряд  $q_1$  на расстоянии  $r$  от себя создает, в среде с диэлектрической проницаемостью  $\epsilon$ , электрическое поле с потенциалом

$$\varphi = q_1 / \epsilon r \quad (6.1)$$

и взаимодействует с находящимся на этом расстоянии зарядом  $q_2$  с энергией

$$U = \varphi q_2 = q_1 q_2 / \epsilon r. \quad (6.2)$$

Помните вы, наверно, и то, что в вакууме (или в воздухе)  $\epsilon = 1$ , в воде  $\epsilon$  близко к 80, а в среде типа пластмассы (и, между прочим, сухого белка)  $\epsilon$  лежит где-то между 2 и 4.

**Внутренний голос:** Это все верно, но только «почти». Во-первых, когда заряды находятся не в вакууме, а в какой-то среде, то  $U$  является, строго говоря, не энергией, а *свободной* энергией [ведь именно она — в присутствии теплообмена нашей системы (зарядов) со средой — стремится к минимуму при стремлении к равновесию, т. е. именно она должна отвечать за притяжение разноименных зарядов и отталкивание одноименных]. Кроме того, энергия элементарного взаимодействия не должна зависеть от температуры, а  $U$  зависит от диэлектрической проницаемости  $\epsilon$ , и, следовательно, меняется с температурой.

**Лектор:** Это замечание абсолютно верно, но все это — постольку, поскольку нас интересует влияние взаимодействия зарядов на стабильность белка — мелочь, пуристика.

Важнее то, что уравнение (6.2) справедливо только тогда, когда «среда» однородна. А когда нас интересует взаимодействие зарядов в белках, мы имеем дело с крайне неоднородной средой. У самого белка, как у любой пластмассы,  $\epsilon$  невелико, где-то около 2 или 4 (напомню: в воздухе  $\epsilon=1$ ), а у воды  $\epsilon=80$ . И заряженные группы белка обычно расположены на его поверхности, у самой воды (позже мы поймем, почему). Какое же  $\epsilon$  мы должны брать для оценок электростатических взаимодействий в белке? Если мы выберем  $\epsilon \approx 80$ , два единичных (протонных) заряда на расстоянии  $3 \text{ \AA}$  будут взаимодействовать с энергией  $\approx 1,5$  ккал/моль; если же мы возьмем  $\epsilon \approx 3$ , то они будут взаимодействовать с энергией  $\approx 40$  ккал/моль. Разница уж очень велика: дополнительные +40 ккал/моль разрушат любую белковую структуру...

Кстати, как понять, — существен или нет тот или иной (свободно-) энергетический эффект для белка? Тут полезно помнить две величины: (а)  $kT \approx 0,6$  ккал/моль; эффект, уступающий  $kT$  (в расчете на одну молекулу белка) всегда не существен — он «замывается» тепловыми флуктуациями; (б) характерный «запас стабильности» белковой структуры (разность свободных энергий между нативной и денатурированной формой белка),  $\approx 10$  ккал/моль; эффект, превышающий 10 ккал/моль, «взрывает» белок.

Второй вопрос связан вот с чем. Формулы (6.1) и (6.2) справедливы, когда  $r$ , расстояние между зарядами, много больше размера молекул среды. Но в белках заряды часто находятся в прямом контакте, на расстоянии  $3\text{--}4 \text{ \AA}$  друг от друга, когда даже молекула воды, не говоря уже о боковой группе, между ними не влезет. Как же оценивать электростатические взаимодействия в этом случае? Брать  $\epsilon=1$ , как в вакууме? Или все же брать  $\epsilon=80$ ? Или...?

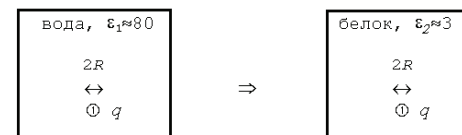
Маленькое философское отступление. Зачем вообще нужны такие прикидочные оценки? Казалось бы, имея мощный компьютер, можно ввести в него «все как есть»: молекулы воды, координаты атомов белка, в том числе координаты зарядов, задать температуру, т. е. энергию теплового движения, и «все посчитать точно». На самом деле эта картина несколько утопична. Расчет (я имею в виду детальный расчет, он ведется при помощи так называемой «молекулярной динамики») — займет часы или дни (и, кстати, совсем точным он не будет, вспомните хотя бы о «некруглых», из-за р-орбит, атомах): ведь вам надо будет рассмотреть и тепловые движения, и поляризацию десятков тысяч атомов. А вас, скорее всего, интересует простая, но быстрая прикидка: можно ли ввести в такое-то место белка заряд, или белок при этом взорвется? И моя цель — научить вас делать такие прикидки.

■ Прежде всего оценим, как меняется энергия заряда при переносе его из воды ( $\epsilon \approx 80$ ) в белок ( $\epsilon \approx 3$ ). Корпускулярностью, т. е. атомным строением и белка, и воды мы пока пренебрежем, точнее, отложим этот вопрос.

Согласно классической электростатике, шарик с зарядом  $q$  и радиусом  $R$  в среде с диэлектрической проницаемостью  $\epsilon$  имеет энергию

$$U = q^2/2\epsilon R. \quad (6.3)$$

Эта формула прямо следует из формулы (6.2): когда мы заряжаем шарик (от заряда 0 до заряда  $q_1$ ), перенося на его поверхность малые заряды  $dq$ , каждый зарядик  $dq$  повышает энергию шарика на  $dU = qdq/\epsilon R$ , согласно (6.2), а интеграл  $q dq$  от 0 до  $q_1$  равен  $q_1^2/2$ .



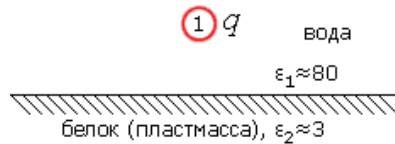
$$\Delta U_{1 \rightarrow 2} = q^2/2\epsilon_2 R - q^2/2\epsilon_1 R$$

Радиус заряженного атома — около  $1,5 \text{ \AA}$ ; значит, его (свободная) энергия близка к  $1,5$  ккал/моль при  $\epsilon \approx 80$  (в воде), и к  $40$  ккал/моль при  $\epsilon \approx 3$  (в белке). Эта большая разница объясняет, почему внутри белка — в отличие от его поверхности — практически нет заряженных групп (нетрудно прикинуть, что даже погружение в среду с  $\epsilon \approx 3$  тесно (на  $\approx 3 \text{ \AA}$ ) сближенной пары противоположных по знаку зарядов повышает свободную энергию примерно на те же  $40$  ккал/моль). Поэтому погружаемая в глубь белка заряженная боковая группа всегда разряжается, т. е. отдает воде свой избыточный  $H^+$  (если она была заряжена положительно) или ловит в воде недостающий ей  $H^+$  (если она была заряжена отрицательно). Правда, разрядка заряженной группы тоже повышает свободную энергию, но «лишь» на несколько (а не на несколько десятков!) ккал/моль. Если говорить точнее, заряды внутри белка иногда все же встречаются, но практически всегда при этом играют очевидную функциональную роль: тогда белку как бы приходится терпеть заряды внутри себя, чтобы работать...

■ Теперь научимся оценивать взаимодействие зарядов с учетом поверхности раздела между белком ( $\epsilon \approx 3$ ) и водой ( $\epsilon \approx 80$ ).

Рассмотрим следующую простую учебную задачу. Пусть вода (среда с  $\epsilon_1=80$ ) занимает одну половину пространства, «белок» (среда с  $\epsilon_2=3$ ) — другую, а поверхность раздела между ними плоская. Пусть в воде, в точке  $\odot$  над плоской поверхностью «белка», находится заряд  $q$  (рис. 6-1).





**Рис. 6-1.** Ситуация, типичная для заряженных боковых групп в белках: находящийся в точке 1 вблизи от поверхности белковой глобулы заряд  $q$  окружен водой, т. е. он сам находится в среде с высокой диэлектрической проницаемостью  $\epsilon_1$ , но вблизи него находится среда с низкой диэлектрической проницаемостью  $\epsilon_2$

Требуется найти электрическое поле  $\phi$ , создаваемое этим зарядом в произвольной точке 2.

Если бы белка не было, или если бы он был очень далеко, это поле в произвольной точке 2 рассчитывалось бы по классической формуле (6.1) с  $\epsilon = 80$  (диэлектрическая постоянная воды) и  $r = r_{12}$  (расстояние между точками 1 и 2).

Что же меняет наличие вблизи от заряда белка — среды с другой диэлектрической проницаемостью?

Для подсказки, вспомним очень похожую задачу, которую вы, должно быть, решали в курсе общей физики. В ней — все то же самое, только теперь «среда 2» — это металл, проводник тока, т. е. его  $\epsilon_2$  равно бесконечности (рис. 6-2). Решение этой задачи выглядит так:

$$\phi = \begin{cases} 0 & \text{внутри металла — там поля нет} \\ q/\epsilon_1 r_{12} + (-q)/\epsilon_1 r_{02} & \text{над поверхностью металла} \end{cases} \quad (6.4)$$



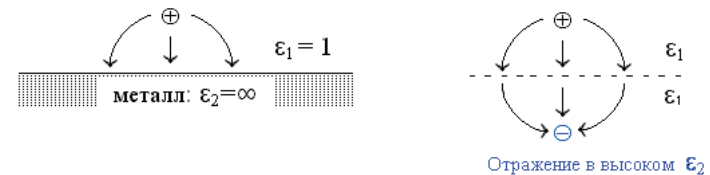
**Рис. 6-2.** Иллюстрация к вспомогательной школьной задаче о поле заряда, находящегося вблизи среды с очень высокой диэлектрической проницаемостью («металла»). На одной схеме (справа) показаны наведенные поляризационные заряды «металла», а на другой (слева) — «отраженный» заряд  $-q$ , суммирующий действие этих поляризационных зарядов за пределами металла. Дальнейшие пояснения — в тексте

Поля *внутри* металла нет, так как иначе свободные заряды в нем двигались бы, пока поле не исчезло.

Результирующее поле *над* поверхностью металла (в точке 2) создается (см. рис. 6-2) зарядом  $q$  и его как бы «отражением» в металле. Это отражение «0» суммирует действие заряда, наведенного в металле сместившимися под действием заряда 1 электронами. Эти электроны распределяются по поверхности металла так, что создают «эффект отражения». Это «отражение» имеет заряд  $-q$  и как бы находится в точке 0, лежащей на той же глубине *под* поверхностью металла, как точка 1 *над* его поверхностью (т. е.  $r_{02}$  — расстояние от отраженного заряда до точки, поле в которой нужно рассчитать).

Строго доказывать это решение (точнее, его вторую, неочевидную часть, относящуюся к «полю над металлом») я не буду, но коротко изложу основную идею доказательства.

Так как внутри металла, если по нему уже не течет ток, электрического поля нет, то силовые линии этого поля (напомню, что они показывают направление силы, действующей на заряд, см. стрелки на рис. 6-3) должны входить в металл строго перпендикулярно (иначе у них была бы составляющая, параллельная поверхности, и она бы гнала ток по металлу). Но и силовые линии поля, созданного (в отсутствие всякого металла) двумя противоположными по знаку, но равными по величине зарядами, тоже перпендикулярны плоскости, проходящей точно посередине между этими зарядами. А раз силовые линии поля выглядят одинаково, то и потенциалы ведут себя одинаково.

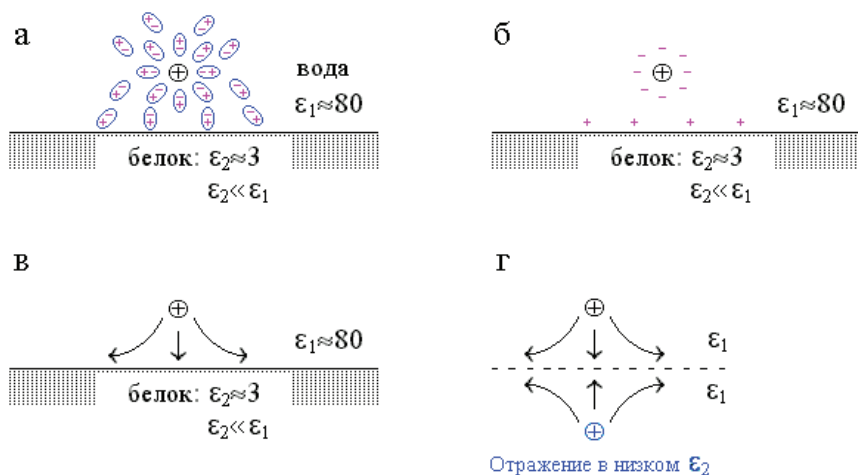


**Рис. 6-3.** Силовые линии поля заряда, находящегося над металлом, и аналогичная (в своей верхней части) картина для силовых линий поля, созданного равными по величине зарядами противоположного знака

Решение «Задачи о поле заряда над металлом» подсказывает, что такие задачи удобно решать при помощи «отраженных» зарядов и анализа хода силовых линий поля. И еще — что эти линии стремятся *побыстрее уйти в среду с более высокой диэлектрической проницаемостью*.

Вернемся теперь к нашей задаче о поле заряда, находящегося у границы двух сред. В этом случае тоже основную роль играет поляризация среды с более высокой диэлектрической проницаемостью — воды (рис. 6-4,

вверху). Небольшими поляризационными зарядами белка (на границе раздела белок/вода) можно пренебречь по сравнению с большими поляризационными зарядами воды. Поляризация среды приводит к тому, что силовые линии стремятся остаться в воде (рис. 6-4, внизу слева) — ведь ее диэлектрическая проницаемость выше, чем белка. Такого же вида поле (если рассматривать только его лежащую *над* поверхностью раздела часть) создается двумя одноименными зарядами (рис. 6-4, справа внизу). Это подсказывает, какой нам нужен отраженный заряд и где его помещать: помещать его надо под поверхность раздела — в то место, где мы видим отражение, — а знак его должен быть тот же, что у нашего (находящегося над белком) заряда «1».

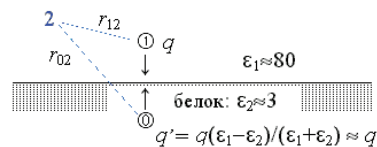


**Рис. 6-4.** Иллюстрация решения задачи о поле заряда, находящегося в воде у поверхности белка (см. рис. 6-1). (а) Диполи (молекулы воды) поворачиваются своим «минусом» к нашему заряду ⊕, а «плюсом» — в противоположную сторону. (б) В результате в воде возникают поляризационные заряды. Они возникают на ее границе, так как внутри толщи воды «плюсы» и «минусы» ее диполей компенсируют друг друга. Те поляризационные заряды, что стоят непосредственно у ⊕, уменьшают поле заряда ⊕, что приводит к появлению высокой диэлектрической проницаемости воды (этот эффект не зависит от соседства белка). Те молекулы воды, что стоят у поверхности раздела сред, создают на этой поверхности поляризационный заряд: обратите внимание, что он — того же знака, что и наш заряд ⊕. (в) Результирующая картина силовых линий электростатического поля у поверхности среды с низкой (по сравнению с той средой, где находится наш заряд ⊕) диэлектрической проницаемостью. (г) Эквивалентная (с точки зрения хода силовых линий над поверхностью раздела сред) картина, вызванная отраженным зарядом того же знака

Я не буду утомлять вас подробным решением всей задачи, а сразу напишу ответ. Итак.

Поле над белком создается и зарядом  $q$ , и отраженным зарядом  $q' = q(\epsilon_1 - \epsilon_2)/(\epsilon_1 + \epsilon_2) \approx +q$  (близким по величине к  $+q$  при  $\epsilon_1 \gg \epsilon_2$ ), лежащим на той же глубине под поверхностью, как  $q$  — над поверхностью белка (рис. 6-5).

$$\varphi_{\text{НАД БЕЛКОМ}} = q/\epsilon_1 r_{12} + q'/\epsilon_1 r_{02}$$

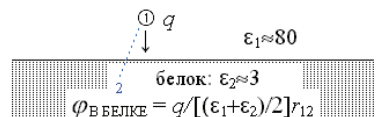


**Рис. 6-5.** Расчет поля над белком, делаемый при помощи отраженного в белке заряда

В результате эффективная диэлектрическая проницаемость  $\epsilon_{\text{эфф}}$  (это та проницаемость, которую надо подставлять в формулу  $\varphi = q/\epsilon_{\text{эфф}} r_{12}$ ) в зоне над белком близка к  $\epsilon_1$ , т. е. к 80 только на самых малых расстояниях от заряда ⊕ (когда  $r_{12}$  много меньше, чем расстояние от ⊕ до поверхности белка; при этом  $r_{12} \ll r_{02} \approx r_{01}$ ).

Но на больших расстояниях от ⊕  $\epsilon_{\text{эфф}}$  всюду близка к 40. В самом деле, когда  $r_{12} \gg r_{10}$ , то  $r_{02} \approx r_{12}$ , и так как  $q' \approx q$  при  $\epsilon_1 \gg \epsilon_2$ , то  $\varphi \approx 2q/[(\epsilon_1/2)r_{12}] \equiv q/[(\epsilon_1/2)r_{12}]$ .

При этом  $\epsilon_{\text{эфф}} \approx 40$  относится и к полю над (а следовательно — и под) самой поверхностью белка (где всегда  $r_{02} \approx r_{12}$ , см. рис. 6-5).



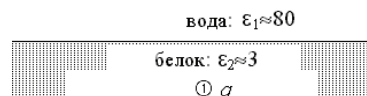
**Рис. 6-6.** Расчет поля, создаваемого внутри белка зарядом, лежащим над поверхностью белка

Итак, поле внутри белка создается зарядом, стоящим в точке ⊕ (рис. 6-6), и для этого поля (я опять позволю себе опустить выкладки) эффективная диэлектрическая проницаемость во всем пространстве под поверхностью равна  $\epsilon_{\text{эфф}} = (\epsilon_1 + \epsilon_2)/2$ , т. е. близка к 40 (т. е. равна  $\epsilon_{\text{эфф}}$  на поверхности белка). И это равенство естественно, так как потенциалы «поля внутри белка» и «поля вне белка» должны совпадать на поверхности раздела сред (воды и белка).

Очень похожий ответ получится, если мы будем рассматривать комбинированную задачу: о поле заряда, находящегося *под* поверхностью белка (рис. 6-7).

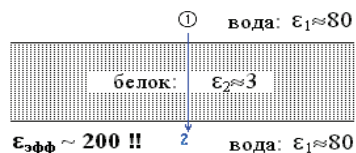
В этом случае эффективная диэлектрическая проницаемость  $\epsilon_{\text{эфф}}$  тоже близка к  $(\epsilon_1 + \epsilon_2)/2$ , т. е. к 40 *всюду* на больших расстояниях от заряда  $q$ , а к  $\epsilon_2$  (т. е. к 3) только когда  $r_{12}$  много меньше, чем расстояние от заряда  $q$  до поверхности белка.

**Рис. 6-7.** Иллюстрация задачи о поле заряда, находящегося внутри белка, но вблизи от его граничащей с водой поверхности



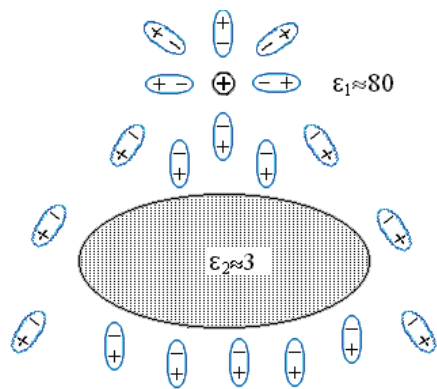
- Еще более любопытный результат дает задача о поле, которое заряд, находящийся на одном краю белка, создает на другом краю белка (рис. 6-8).

**Рис. 6-8.** Иллюстрация задачи об электрическом поле, создаваемом зарядом ① в точке 2 на противоположной поверхности белка



Казалось бы, в этом случае, поскольку взаимодействие идет через белок, через среду с низкой ( $\epsilon_2 \approx 3$ ) диэлектрической проницаемостью, мы должны были бы ожидать, что  $\epsilon_{эфф}$  будет близко к 3, в крайнем случае — ведь вода вокруг — лежать где-то между 3 и 80. А на самом деле —  $\epsilon_{эфф}$  зашкаливает за 200! Как так??

Для объяснения посмотрим, как ориентируются в этом случае полярные молекулы воды вокруг белка и вокруг «нашего» заряда ①, и как это меняет поле заряда (рис. 6-9; он дан, для наглядности, в несколько другом масштабе, чем рис. 6-8, но описывает то же явление).



**Рис. 6-9.** Ориентация молекул воды (они изображены в виде диполей) вокруг белка и заряда ⊕ (он изображен положительным просто для определенности)

Молекулы воды ориентируются по полю: их «-» преимущественно смотрит в сторону «нашего» заряда ⊕, а «+» — в противоположную сторону. В результате, во-первых, заряд ⊕ частично компенсируется примаыкающими к нему «минусами» молекул воды; это — тривиальное явление, оно просто приводит к появлению большой диэлектрической проницаемости воды  $\epsilon_1$ . Во-вторых, возникает то явление, которое нас интересует. А именно, повернутые к белку «плюсы» молекул воды, примаыкающих к белку с той стороны, где находится заряд, создают на этой стороне поляризационный заряд «+»; а «минусы» молекул воды, примаыкающих к белку с другой стороны, создают там противоположный поляризационный заряд «-» (поляризацией самого тела белка можно пренебречь, по сравнению с поляризацией воды она мала, так как его  $\epsilon_2 \ll \epsilon_1$ ).

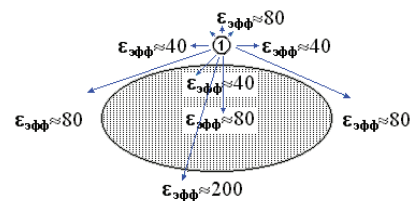
В результате создаются поляризационные заряды, создаются на поверхностях раздела «заряд — вода» и «белок — вода». Только на поверхностях раздела! Внутри однородных (и исходно незаряженных) сред поляризационные заряды «зануляются» — есть такая теорема в электростатике. По той же теореме равна нулю и сумма поляризационных зарядов, созданных внешним зарядом у поверхности тела (нашего белка), окруженного однородной средой (у нас — водой).

Значит, к ⊕ обращен суммарный поляризационный заряд «+» (!!), созданный у поверхности белка повернувшимися молекулами воды, а равный ему поляризационный «-» (созданный молекулами воды у другой поверхности белка) обращен в обратную сторону.

В результате на обращенной к заряду стороне белка потенциал поля возрастает по сравнению с тем, что было бы без белка. Здесь, в дополнение к потенциалу заряда ⊕, потенциал создается и поляризационными «плюсами» у этой поверхности белка (а поляризационные «минусы» от этого места далеко и влияют мало). Поэтому тут  $\epsilon_{эфф} \approx 40$ , как о том уже говорилось.

В то же время, на отвернутой от заряда ⊕ стороне белка потенциал поля падает по сравнению с тем, что было бы без белка; здесь к потенциалу ⊕ заряда добавляется противоположный по знаку потенциал поляризационных «минусов» у ближайшей поверхности белка (а поляризационные «плюсы» от этого места далеко и влияют мало). А раз здесь потенциал поля упал по сравнению с тем, что было бы без белка, — то здесь (на противоположной ⊕'у стороне)  $\epsilon_{эфф}$  больше, чем то  $\epsilon_1$ , что было без белка, т. е. больше, чем  $\epsilon_1 = 80$ .

В целом, распределение значений  $\epsilon_{эфф}$  «в белке и вокруг» для поля, создаваемого зарядом ① и наведенными им же поляризационными зарядами, выглядит так, как на рис. 6-10, если заряд ① находится у поверхности белка, и так, как на рис. 6-11, если заряд ① находится в глубине белка.

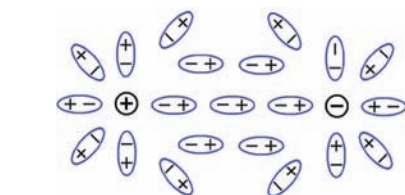


**Рис. 6-10.** Характерные величины  $\epsilon_{эфф}$  для потенциала, создаваемого в разных точках  $r$  зарядом, расположенным у поверхности белка

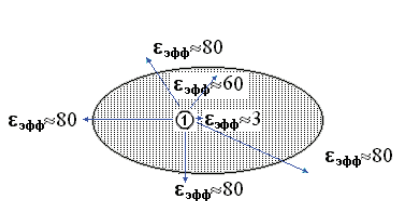
Напоминаю, что  $\epsilon_{эфф}$  — это та эффективная величина диэлектрической проницаемости для точки  $r$ , которую нужно подставлять в формулу расчета потенциала заряда «1» в точке  $r$ , т. е.  $\phi(r) = q_1/\epsilon_{эфф}|r-r_1|$  (где  $r_1$  — точка, в которой стоит заряд).

- Отметим еще одно явление, вытекающее из наличия поверхности раздела и связанных с ней «отраженных зарядов». Оно касается действия заряда на самого себя [см. также обсуждение формулы (6.3)]: находящийся вне белка заряд *отталкивается* от поверхности белка, а заряд, находящийся внутри белка, *сильно притягивается* к его поверхности — т. е., в обоих случаях, среда с более высокой диэлектрической проницаемостью притягивает заряд, а среда с более низкой — его выталкивает. Силу этого притяжения и выталкивания легко оценить: ведь речь здесь идет о взаимодействии заряда с его отражениями (см. рис. 6-5, 6-6); советую попрактиковаться в этом самостоятельно.

[Лишь один численный пример. Можно показать, что энергия проводящего шарика (иона) с зарядом  $q$  и радиусом  $R$ , погруженного наполовину в воду ( $\epsilon_1=80$ ), наполовину в белок ( $\epsilon_2=3$ ), есть  $q^2/(\epsilon_1+\epsilon_2)R \approx 3$  ккал/моль; в воде же она равна  $q^2/2\epsilon_1 R \approx 1,5$  ккал/моль (а в белке —  $q^2/2\epsilon_2 R \approx 40$  ккал/моль, см. (6.3)). Итак, ион, сидящий на белке, теряет (по сравнению с ионом в воде) 1,5 ккал/моль. Но, если ион зажат между двумя белками, — он теряет вовсе не 3 ккал/моль, как то было бы, если бы его взаимодействия с этими двумя белками складывались, а на порядок больше — 37 ккал/моль! Это показывает, что электростатические взаимодействия в неоднородной среде *резко неаддитивны*.]



**Рис. 6-12.** Ориентация молекул воды, окружающих заряды  $\oplus$  и  $\ominus$ , и молекул, находящихся между этими зарядами



**Рис. 6-11.** Характерные величины  $\epsilon_{эфф}$  для потенциала, создаваемого в разных точках  $r$  зарядом, расположенным внутри белка

- Теперь займемся эффектами, связанными с корпускулярностью, т. е. с атомарным строением среды.

Собственно говоря, сама величина диэлектрической проницаемости  $\epsilon$  определяется атомарным строением среды. Если среда состоит из неполярных молекул, электрическое поле лишь смещает электроны в них, что сделать трудно: поэтому электроны смещаются мало, и  $\epsilon$  невелико. Если среда состоит из полярных молекул (пример — вода), электрическое поле разворачивает эти молекулы; это сделать легче, и  $\epsilon$  такой среды велико. В обоих случаях — и при смещении электронов, и при повороте полярных молекул — поляризация среды как бы частично «гасит» внесенные в нее заряды (заряды  $\oplus$  и  $\ominus$ , рис. 6-12), и тем самым уменьшает электрическое поле в среде — по сравнению с тем, что было бы в вакууме.

- Естественно было бы предположить, что связанные с корпускулярностью эффекты должны сильно влиять на взаимодействие зарядов на малых расстояниях, ведь классические формулы (6.1)–(6.3) справедливы, строго говоря, только когда между взаимодействующими зарядами находится много молекул среды. А если заряды (как часто бывает в белках) находятся на расстоянии 3–4 Å, никакой другой атом между ними уже не влезет и не изменит их взаимодействия.

Казалось бы, в случае столь тесного контакта зарядов диэлектрическая проницаемость для их взаимодействия должна приближаться к 1 даже в водном окружении. Эту точку зрения — точнее, это опасение — до сих пор можно встретить в литературе.

Однако, как ни странно, корпускулярность среды не меняет кардинально «макроскопическую» (т. е. выведенную для больших расстояний между зарядами) диэлектрическую проницаемость среды даже на расстоянии порядка 3 Å. То есть даже здесь величина диэлектрической проницаемости воды гораздо ближе к 80 или 40, чем к 1 или 3.

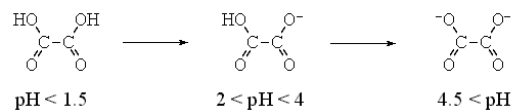
Об этом свидетельствует то, что соль хорошо растворяется (диссоциирует) в воде, что возможно только при слабом притяжении противоположно заряженных ионов даже на самых малых,  $\approx 3$  Å, расстояниях.

В самом деле, ионы  $\text{Na}^+$  и  $\text{Cl}^-$  могут сблизиться до расстояния 3 Å. При этом свободная энергия их притяжения составляла бы  $-1,5$  ккал/моль при  $\epsilon=80$ ,  $-3$  ккал/моль при  $\epsilon=40$ , и  $-6$  ккал/моль при  $\epsilon=20$ . Последнее (6 ккал/моль) превосходит энергию водородной связи. При такой энергии противоположно заряженные ионы слипались бы друг с другом сильнее, чем молекулы воды, и тогда насыщенный раствор соли имел бы концентрацию порядка  $10^{-4}$  моля на литр — как насыщенный водяной пар. Но это явно неправда: растворить 1 моль  $\text{NaCl}$  (58 г) в литре воды — не проблема (это обычный, может



чуть крепкий рассол). Значит, диэлектрическая проницаемость воды существенно больше 20 даже на расстоянии  $\approx 3 \text{ \AA}$ .

- Более точно оценить величину  $\epsilon$  на самых малых расстояниях можно, исходя из величин первой и второй констант диссоциации двухосновных кислот и оснований в воде. Например, диссоциация щавелевой кислоты происходит так:



Вторая диссоциация сдвинута на  $\approx 2,5$  единицы pH относительно первой, т. е. происходит при в  $10^{2,5} = e^{2,3 \times 2,5}$  раз меньшей концентрации  $\text{H}^+$ . Это показывает, что свободная энергия взаимодействия первого заряда со вторым составляет  $2,5 \times 2,3RT \approx 3,5$  ккал/моль. И это — при расстоянии  $\approx 3 \text{ \AA}$  между зарядами. Такая величина энергии взаимодействия соответствует  $\epsilon \approx 40$  на расстоянии  $3 \text{ \AA}$ . К сходному результату —  $\epsilon \approx 30$  на расстоянии  $\approx 2,5 \text{ \AA}$  приводит рассмотрение диссоциации угольной кислоты,  $\text{H}_2\text{CO}_3 \rightarrow \text{HCO}_3^- \rightarrow \text{CO}_3^{2-}$ . А «правильная» свободная энергия водородной связи в воде получается, если  $\epsilon \approx 15-20$  на расстоянии  $\approx 2,0 \text{ \AA}$ .

Конечно,  $\epsilon$  должно приближаться к 1 на расстояниях  $\approx 1 \text{ \AA}$ , иначе страшно поменялась бы энергия всех ковалентных взаимодействий (чего не наблюдается!), но этот «внутриатомный» диапазон расстояний относится уже скорее к химии, чем к физике белка.

Итак, даже *солевая связь* противоположно заряженных боковых групп на поверхности белка должна «стоять» всего пару-тройку ккал/моль. Внутри белка она должна «стоять» больше, но погружение заряженных групп внутрь глобулы обойдется еще дороже — так что не надо удивляться, что такие связи в природных белках наблюдаются нечасто.

Итак, мы пришли к выводу, что корпускулярность не меняет радикально «макроскопическую» (выведенную для больших расстояний между зарядами) диэлектрическую проницаемость воды даже на расстоянии в  $2-3 \text{ \AA}$ , когда между взаимодействующими зарядами не может быть никаких других молекул. Причина этого в том, что заряды достаточно сильно экранируются даже подошедшими «с другой стороны и с боков» (рис. 6-13) молекулами среды, которые поляризуются (в случае воды — просто поворачиваются) так, что «+»'ы этих молекул смещаются к заряду  $\ominus$ , а «-»'ы — к заряду  $\oplus$ .

Здесь мы снова (ср. рис. 6-10, 6-11) видим, что электростатическое взаимодействие между зарядами как бы проходит в основном через среду с более высокой диэлектрической проницаемостью и почти игнорирует среду, которая слабо поляризуется.

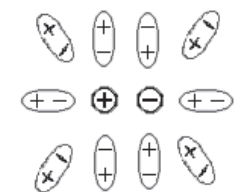


Рис. 6-13. Ориентация молекул воды, окружающих предельно сближенные заряды  $\oplus$  и  $\ominus$

- Все, о чем мы говорили до сих пор, относилось к «мелкомолекулярным» системам. Сохраняются ли все эти выводы для белков? Ведь там эффекты корпускулярности среды (воды прежде всего) накладываются на огромную разницу диэлектрических проницаемостей воды и белка...

- Опыты, проведенные в лаборатории Фершта — зачинателя белковой инженерии, — показывают, что приведенные выше оценки в полной мере относятся и к белкам.

Опыты основывались на следующем. Есть белки, ферменты, активность которых особенно велика при определенном значении pH (у них, как говорится, есть pH-оптимум). Вводя в белок, путем мутации его гена, заряженный остаток, можно сдвинуть этот pH-оптимум (рис. 6-14) — и, по его сдвигу, оценить электрическое поле, созданное мутированным остатком в активном центре.

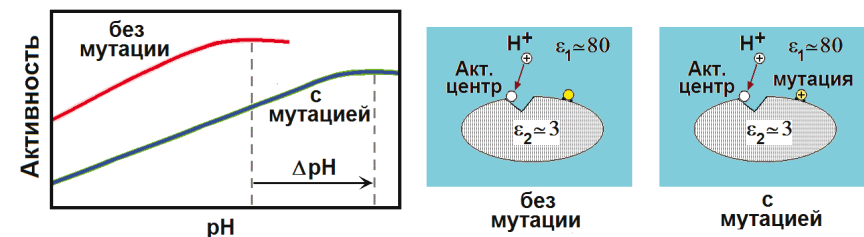


Рис. 6-14. Вызванный мутацией сдвиг pH-оптимума активности белка на величину  $\Delta\text{pH}$ . Из A. J. Russel & A. R. Fersht, *Nature* (1987) **328**:496–500, с упрощениями

Причина pH-оптимума в том, что какая-то группа в активном центре должна, для успешной работы фермента, находиться в определенной зарядовой форме, а заряженность этой группы зависит от концентрации водородных ионов в среде. Концентрация  $\text{H}^+$ , т. е.  $[\text{H}^+]$  (моль/литр) равна  $10^{-\text{pH}}$  по определению, а концентрация  $\text{OH}^-$  в воде есть примерно  $10^{-14+\text{pH}}$ .

Пусть активный центр (АЦ) принимает ион  $\text{H}^+$ :  $\text{АЦ} + \text{H}^+ = \text{АЦH}^+$ . Тогда, по закону действующих масс, соотношение концентраций двух (с  $\text{H}^+$  и без  $\text{H}^+$ ) форм активного центра составляет



$$\begin{aligned} & [\text{АЦН}^+]/[\text{АЦ}] = \exp(-\Delta F_{\text{АЦН}^+}/RT) \times [\text{H}^+] = \\ & = \exp(-\Delta F_{\text{АЦН}^+}/RT) \times 10^{-\text{pH}} = \exp\{-(\Delta F_{\text{АЦН}^+}/RT + 2,3 \times \text{pH})\}, \end{aligned}$$

где  $\Delta F_{\text{АЦН}^+}$  — свободная энергия связывания  $\text{H}^+$  (при  $[\text{H}^+] = 1$  моль/л) с активным центром, а значок [ ] означает концентрацию.

Если мутация вносит в белок заряд, создающий в его активном центре потенциал  $\phi$ , то  $\Delta F_{\text{АЦН}^+}$  меняется:  $\Delta F_{\text{АЦН}^+|_{\text{с мутацией}}} = \Delta F_{\text{АЦН}^+|_{\text{без мутации}}} + \phi e$ , где  $e$  — заряд  $\text{H}^+$ . Так как в рН-оптимуме величина  $[\text{АЦН}^+]/[\text{АЦ}]$  (и, следовательно, величина  $\Delta F_{\text{АЦН}^+}/RT + 2,3 \times \text{pH}$ ) должна быть одной и той же и при, и без мутации, то

$$\begin{aligned} & \Delta F_{\text{АЦН}^+|_{\text{без мутации}}}/RT + 2,3 \times \text{pH}|_{\text{опт. без мутации}} = \\ & = \Delta F_{\text{АЦН}^+|_{\text{с мутацией}}}/RT + 2,3 \times \text{pH}|_{\text{опт. с мутацией}} \end{aligned}$$

То есть

$$\begin{aligned} & \phi e = \Delta F_{\text{АЦН}^+|_{\text{с мутацией}}} - \Delta F_{\text{АЦН}^+|_{\text{без мутации}}} = \\ & = 2,3RT(\text{pH}|_{\text{опт. без мутации}} - \text{pH}|_{\text{опт. с мутацией}}) = 2,3RT \times (-\Delta \text{pH}). \end{aligned}$$

Так, зная сдвиг рН-оптимума, можно оценить потенциал, созданный в активном центре мутированным остатком белка. Теперь, зная трехмерную структуру этого белка и, следовательно, расстояние  $r$  от мутированного остатка до активного центра, можно оценить эффективную диэлектрическую проницаемость  $\epsilon_{\text{эфф}}$  (входящую в формулу  $\phi = q/\epsilon_{\text{эфф}}r$ ) для взаимодействия внесенного мутацией заряда  $q$  с районом активного центра.

Мутации в опытах Фершта вводились на поверхность белка, чтобы не разрушить его структуру (мы уже знаем, что энергия глубоко погруженного в белок заряда очень велика, так что он может буквально взорвать белок).

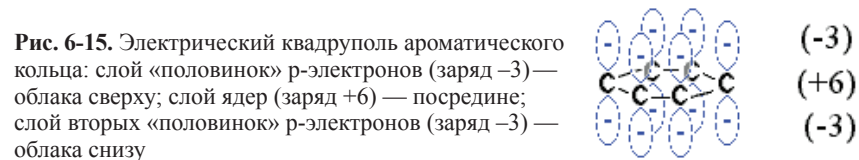
Результат: эффективная диэлектрическая проницаемость  $\epsilon_{\text{эфф}}$  колебалась в этих опытах от 40 до 100, причем первое характерно для близких к активному центру мутаций, а второе — для удаленных от него и «затененных» телом белка. И большая величина  $\epsilon_{\text{эфф}}$  вообще, и, в особенности, то, что  $\epsilon_{\text{эфф}}$  может достигать 100 — вызвало изрядное удивление, так как многие, не знакомые как следует с электростатикой, считали, что  $\epsilon_{\text{эфф}}$  должны лежать где-то между 3 (как внутри белка) и, в самом крайнем случае, 80 (как в воде). Однако нас эти величины удивлять не должны — они хорошо согласуются с тем, что можно ожидать из рис. 6-10.

- Небольшое отступление — о белковой инженерии. Ее главная прелесть заключается в том, что, меняя такой-то кодон в гене белка, мы можем ввести мутацию в точно определенное место белковой глобулы, так как и ген этого белка, и его аминокислотная последовательность, и его трехмерная структура известны. Кроме того, влияние мутации на структуру также можно контролировать — рентгеноструктурным анализом или ЯМР. Таким образом, вся работа ведется с открытыми глазами.

В опытах, о которых шла речь, белок служит микроскопическим (вернее, наноскопическим) электрометром. И белковая инженерия дает возможность использовать такие приборы и при этом прыгать от физической теории к генным манипуляциям и обратно, что чрезвычайно занимательно.

- Теперь я хочу сделать несколько добавлений касательно электростатических взаимодействий.

Первое. До сих пор я говорил только о взаимодействии отдельных зарядов. Однако к электростатике относятся и взаимодействия диполей (например, диполей  $\text{H}^{(+)}-\text{O}^{(-)}$  и  $\text{H}^{(+)}-\text{N}^{(-)}$ , вовлеченных в образование водородных связей), а также квадруполей — последние присутствуют, например, в ароматических кольцах (рис. 6-15).



Я остановился на взаимодействии зарядов потому только, что они наиболее сильны, — даже при прямом контакте они в несколько раз сильнее, чем взаимодействия диполей (и к тому же медленнее спадают с расстоянием), а взаимодействия диполей — сильнее, чем взаимодействия квадруполей.

Второе. При наличии свободных зарядов (например, соли) в воде электростатические взаимодействия ослабевают с расстоянием  $r$  не по закону «энергия пропорциональна  $(1/r)$ », а гораздо быстрее — как  $(1/r)\exp(-r/D)$ . Здесь  $D$  — радиус Дебая-Хюккеля; он соответствует характерному размеру противоионного облака вокруг заряда. Величина  $D$  не зависит от самого заряда, но зависит от зарядов и концентрации ионов соли в среде, от диэлектрической проницаемости среды и от температуры. В воде при комнатной температуре

$$D \approx 3\text{I}^{1/2} \text{ \AA}, \quad (6.5)$$

где

$$\text{I} = 1/2 \sum_i c_i z_i^2 \quad (6.6)$$

— ионная сила раствора в моль/литр. В формуле (6.6) сумма берется по всем сортам ионов в растворе, причем  $z_i$  — заряд (в единицах протонного заряда), а  $c_i$  — концентрация (в моль/л) иона  $i$ . Обычным физиологическим условиям соответствует  $\text{I} \approx 0,1-0,15$ ; при этом  $D \approx 8 \text{ \AA}$ . Однако некоторые микроорганизмы живут при  $\text{I} \sim 1$  моль/л и выше; при этом сохраняются, и то в очень ослабленном виде, только те электростатические взаимодействия, что соответствуют «солевым мостикам», т. е. прямому контакту зарядов.

В целом, при наличии в растворе ионной атмосферы энергия взаимодействия двух зарядов имеет вид

$$U = [q_1 q_2 / \epsilon_{\text{эфф}} r] \exp(-r/D). \quad (6.7)$$

Третье. Электростатическое взаимодействие — яркий пример *не* парного взаимодействия частиц. Оно зависит не только от расстояния  $r$  между зарядами  $q_1$  и  $q_2$ , но и от свойств среды (меняющих и  $\epsilon$ , и  $D$ ), и, в частности, от расстояний от зарядов до других тел и от формы этих тел (все это влияет на  $\epsilon_{\text{эфф}}$ ), а также от концентрации свободных ионов в растворе (она влияет на величину  $D$ ).

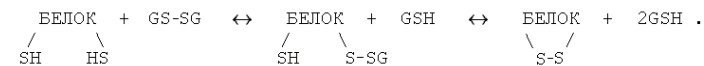
И еще одно добавление. До сих пор я обычно говорил «энергия электростатических взаимодействий». Это говорилось только для простоты слога — как я уже говорил в начале лекции, строго говоря, надо было говорить «свободная энергия». Ведь мы следили только за притяжением и отталкиванием зарядов, но отнюдь не запрещали им (и их окружению) *обмениваться теплом* с окружающей средой. А раз теплообмен возможен, значит, по определению, речь шла о свободной энергии.

Более того. Исследуя температурную зависимость электростатических эффектов в водном окружении, можно показать, что энтропийная составляющая в них доминирует, а собственно энергетическая (энтальпийная) составляющая близка к нулю. Это видно из того, что диэлектрическая проницаемость воды меняется от 88 до 55 (т.е. электростатические взаимодействия растут на 40%) при росте абсолютной температуры  $T$  от 273 до 373 К (тоже примерно на 35%). А у взаимодействия, растущего пропорционально абсолютной температуре, есть только энтропийная, но нет энтальпийной части. Значит, в воде весь электростатический эффект связан не с энергией, а с упорядочением воды вокруг зарядов и с изменением этого упорядочения при сближении или отдалении зарядов друг от друга.

Значит, как ни парадоксально, электростатика в воде имеет энтропийную, а не энергетическую природу, впрочем, так же, как и гидрофобные взаимодействия или образование водородных связей в водном окружении.

- Заклучая раздел «Элементарные взаимодействия в белках и вокруг», я хочу еще упомянуть о дисульфидных связях и о координационных связях. Не столь массовые, как, например, водородные связи, — эти связи, однако, часто играют важную роль в белках.
- *Дисульфидные* (или S–S) связи образуются цистеиновыми (Cys) аминокислотными остатками (боковая группа цистеина:  $-\text{C}^\beta\text{H}_2-\text{SH}$ ). Непосредственное, с выделением водорода (по схеме  $-\text{CH}_2-\text{SH} + \text{HS}-\text{CH}_2- \rightarrow -\text{CH}_2-\text{S}-\text{S}-\text{CH}_2- + \text{H}_2$ ) окисление цистеинов в белках не идет — оно

происходит слишком медленно при комнатной температуре. Однако образование S–S-связей в белках может быстро происходить при помощи тиол-дисульфидного обмена. Полагают, что в клетке это происходит при участии *глутатиона*, существующего и в мономерной тиольной (GSH), и в димерной дисульфидной (GSSG) формах, — и происходит по схеме



Как разрыв, так и образование S–S-связей в клетках катализируется (т.е. ускоряется, но не направляется) специальным ферментом — дисульфидизомеразой.

Образование S–S-связей в клетке обратимо, так как энергетический баланс этой реакции — тиол-дисульфидного обмена — близок к нулю (было две ковалентные S–H-связи и одна S–S-связь, и столько же осталось; не правда ли, это очень похоже на энергетический баланс образования внутрибелковых водородных связей в водном окружении?). Более того, имеющаяся относительно высокая концентрация GSH в клетке сдвигает равновесие в сторону разрыва тех связей, которые могли бы образоваться при «случайном» сближении цистеинов. Поэтому здесь способны образоваться и выжить S–S-связи только между «и без того» (т.е. — другими взаимодействиями) сближенными цистеинами.

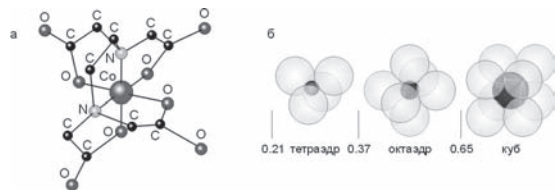
Особенно важны S–S-связи для белков, которым придется жить и работать вне клетки. С одной стороны, там нет дисульфидизомераз и глутатиона, так что уже завязавшиеся (внутри клетки или на выходе из нее) связи «замораживаются», они уже не порвутся и не перестроятся. С другой — белок вне клетки попадает в разные условия и дополнительный запас прочности, даваемый стабильными замороженными S–S-связями, ему не помешает. Поэтому S–S-связи гораздо более типичны для секретируемых белков, чем для внутриклеточных. Обычно в секретируемых белках все имеющиеся цистеины (кроме одного, если их число нечетно) вовлечены в S–S-связи.

- *Координационные* связи образуются N-, O- и S-атомами белка (а также O-атомами воды) с двух- и трехвалентными ионами металлов: Fe, Zn, Co, Ca, Mg и т. д.

У ионов этих металлов есть незаполненные, но низко (по энергии), чуть над заполненными лежащие электронные орбиты. Каждая из них способна связать электронную пару. А у O-, N-, S-атомов («доноров электронов») есть электронные пары, которые могут перетечь на свободные орбиты ионов. Образующаяся при этом связь ничем не отличается от обычной химической связи кроме того, что для обычной связи электроны

поставляются обоими входящими в связь атомами, а для координационной — только одним из них.

При образовании координационных связей ион металла связывается с несколькими донорами электронов. При этом крупные (с радиусом  $\sim 1,5 \text{ \AA}$ ) атомы-доноры со всех сторон окружают маленький (с радиусом  $\sim 0,7 \text{ \AA}$ ) двух- или трехвалентный ион. Наиболее часто шесть атомов-доноров окружают (координируют) ион металла, располагаясь по вершинам правильного октаэдра (рис. 6-16).



**Рис. 6-16.** (а) Структура октаэдрического комплекса, образованного атомом  $\text{Co}^{+++}$  с ЭДТА. (б) Характерная координация центрального иона при различном соотношении его радиуса с радиусами окружающих его доноров электронов. Картинки взяты из [10] и адаптированы

Так как ион может связаться и с электронными донорами белка, и с кислородом воды, он (несмотря на большую энергию каждой связи) переходит из воды на белок и обратно без драматического выигрыша или проигрыша энергии. Важнее, пожалуй, другое: если в белке атомы-доноры уже стоят в «правильных» (для образования координационной связи) положениях — ион может связаться с ними, отпустив ранее окружавшие его молекулы воды, а это приведет к прочной связи из-за выигрыша энтропии движения отпущенных молекул воды. В среднем, каждая координационная связь стоит несколько ккал/моль — несколько больше, чем водородная связь в воде.

Такие связи, образуемые несколькими атомами одной молекулы, способными связать один ион, называются *хелатными* («клешневидными»). Роль таких связей в белках, и в частности в их активных центрах, мы рассмотрим позже. Мы увидим также, что хелатные, полностью обволакивающие ион комплексы могут входить в гидрофобное ядро белка. А пока я хочу снова обратить ваше внимание на рис. 6-16, где изображен часто используемый в лабораторной практике реагент ЭДТА (этилендиаминтетраацетат) в хелатной связи с металлом. У ЭДТА эта связь особенно прочна из-за связи отрицательных зарядов  $\text{COO}^-$  групп ЭДТА с положительно заряженным ионом металла.

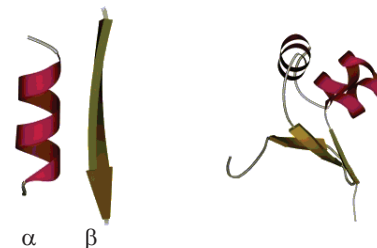
## ВТОРИЧНЫЕ СТРУКТУРЫ ПОЛИПЕПТИДНЫХ ЦЕПЕЙ

### Лекция 7

Вторичная структура полипептидов. Спирали:  $2_7$ ,  $3_{10}$ ,  $\alpha$ ,  $\pi$ , poly(Pro)II. Антипараллельная и параллельная  $\beta$ -структура.  $\beta$ -изгибы. Методы экспериментального обнаружения вторичной структуры.

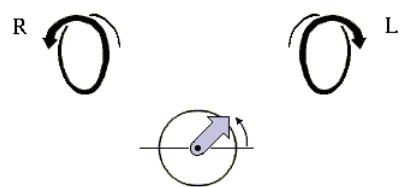
Разобравшись с элементарными взаимодействиями, рассмотрим сегодня вторичную структуру белков. Прежде всего у нас речь пойдет о регулярных вторичных структурах: об  $\alpha$ -спиралях и о  $\beta$ -структуре, но не только о них.

Укладка  $\alpha$  и  $\beta$ -структур в глобулу определяет третичную структуру белка (рис. 7-1). Эти вторичные структуры отличаются определенными, периодическими конформациями главной цепи — при разнообразии конформаций боковых групп.



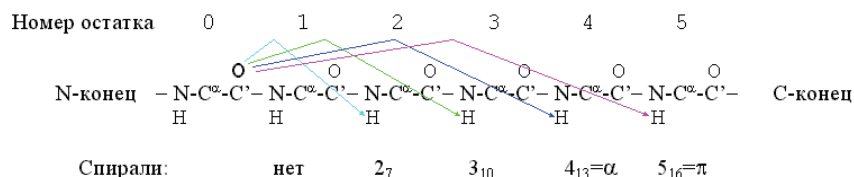
**Рис. 7-1.** Вторичная структура полипептидной цепи ( $\alpha$ -спираль и тяж  $\beta$ -листа) и третичная структура белковой глобулы. Обычно  $\alpha$ - и  $\beta$ -структуры, в сумме, охватывают половину цепи глобулярного белка

Начнем со спиралей. Они могут быть левые и правые (рис. 7-2), у них может быть разный период и шаг. Правые (R) спирали приходят к нам, завиваясь против часовой стрелки (что отвечает положительному отсчету угла в тригонометрии); левые (L) — приходят, вращаясь по стрелке.



**Рис. 7-2.** Правые (R) и левые (L) спирали. Под ними показан отсчет положительного угла в тригонометрии: при этом «близкая к нам» (серая) стрелка вращается *против* хода часов. На спираль надо, естественно, смотреть вдоль ее оси — с любого конца

Важнейшие спирали в полипептидной цепи держатся водородными связями, где C=O группы остова полипептида связаны с лежащими от них в направлении С-конца цепи Н-Н-группами. В принципе, возможны следующие спирали, стянутые Н-связями (рис. 7-3):  $2_7$ ,  $3_{10}$ ,  $4_{13}$  (обычно именуемая  $\alpha$ ) и  $5_{16}$  (она же  $\pi$ ). Здесь в названии « $2_7$ » — «2» означает связь со 2-м по цепи остатком (см. рис. 7-3), а «7» — число атомов в цикле (O...H-N-C'-C-N-C'), замыкаемом этой связью. Тот же смысл имеют цифры и в названии других спиралей.

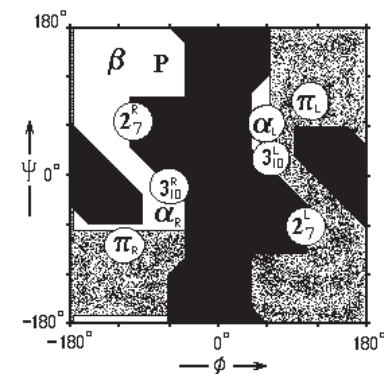


**Рис. 7-3.** Водородные связи (они показаны стрелками), характерные для разных спиралей

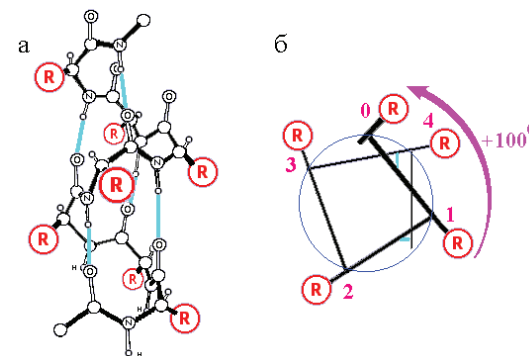
■ Какие из этих спиральных структур преобладают в белках?  $\alpha$ -спирали. Почему? Ответ на этот вопрос дает карта Рамачандрана для типичного аминокислотного остатка — аланина (рис. 7-4), на которой отмечены конформации, периодическое повторение которых приводит к завязыванию изображенных на рис. 7-3 водородных связей.

Видно, что только спираль  $\alpha_R$  ( $\alpha$ -правая) лежит достаточно глубоко внутри области, разрешенной для аланина (и для всех других остатков L-аминокислот). Другие спирали лежат либо на краю этой области (например, левая спираль  $\alpha_L$  или правая спираль  $3_{10}$ ), где конформационные напряжения уже возрастают, либо в области, доступной только глицину.

Поэтому можно ожидать, что именно правая  $\alpha$ -спираль должна быть, как правило, более стабильной и потому преобладать в белках, что и наблюдается. В правой  $\alpha$ -спирали (рис. 7-5) все атомы упакованы оптимально: плотно, но без напряжений; поэтому не удивительно, что в белках таких спиралей много, а в фибриллярных белках они достигают гигантской длины и включают сотни аминокислотных остатков.



**Рис. 7-4.** Конформации различных вторичных структур на фоне карты разрешенных и запрещенных конформаций аминокислотных остатков.  $2_7^R$ ,  $2_7^L$  — правая и левая спираль  $2_7$ ;  $3_{10}^R$ ,  $3_{10}^L$  — правая и левая спираль  $3_{10}$ ;  $\alpha_R$ ,  $\alpha_L$  — правая и левая  $\alpha$ -спираль;  $\pi_R$ ,  $\pi_L$  — правая и левая  $\pi$ -спираль.  $\beta$  —  $\beta$ -структура (подробности см. на рис. 7-8в). P — спираль Poly(Pro)II. □ — конформации, разрешенные для аланина (Ala); ▨ — области, разрешенные лишь для глицина, но не для аланина и других остатков; ■ — области, запрещенные для всех остатков.  $\phi$  и  $\psi$  — углы внутреннего вращения в главной цепи



**Рис. 7-5.** Правая  $\alpha$ -спираль. (а) Атомарная структура. R — боковые группы. Серые линии — водородные связи. (б) Схематическое изображение одного витка той же  $\alpha$ -спирали, вид с торца. Стрелка показывает поворот спирали (в расчете на один остаток) по мере ее приближения к нам (номера остатков уменьшаются по мере приближения к читателю). Картинка (а) взята из [3] и адаптирована

Левых  $\alpha$ -спиралей в белках практически нет. Нет и спиралей  $2_7$ , которые мало того что лежат на самом краю разрешенной области, но еще имеют энергетически невыгодный для Н-связи, почти прямой угол схождения N-H и O=C групп.



Практически нет в белках и спиралей  $\pi$ . Они тоже лежат на самом краю разрешенной области, да еще и витки в них слишком широки, так что  $\pi$ -спирали имеют энергетически невыгодную пустую «дырку» на оси. А вот спирали  $3_{10}$  (в основном правые, левые пригодны практически лишь для глицинов) в белках есть, правда, в виде коротких (из трех-четырех остатков) и деформированных фрагментов (из-за стерических напряжений: слишком тугая спираль! соответствующая спирали  $3_{10}$  конформация лежит на самом краю разрешенной области).

Отметим одно свойство спиралей, хорошо видное на рис. 7-5а: на их N-конце сидят свободные от внутриспиральных водородных связей H-атомы N–H-групп, а на C-конце — свободные от водородных связей O-атомы C=O-групп. Так как электронное облако с H-атома частично стянуто электроотрицательным N-атомом, а электроотрицательный O-атом сам стягивает электрон с C'-атома, — N-конец спирали несет положительный, а C-конец — отрицательный парциальный заряд. То есть спираль представляет собой длинный диполь: величина суммарного (поставляемого тремя NH-группами) парциального «+»-заряда на ее N-конце составляет около половины протонного, а «-»-заряда на C-конце  $\alpha$ -спирали — около половины электронного заряда.

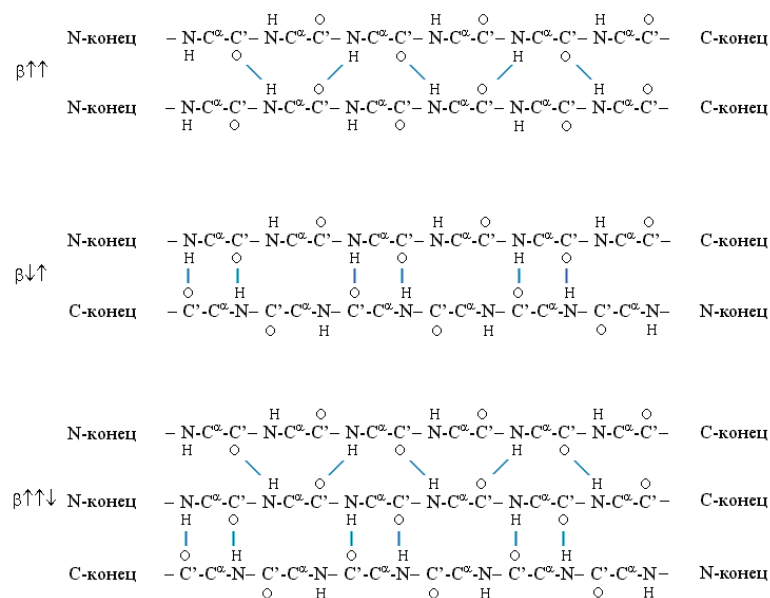
■ Теперь рассмотрим регулярные структуры без водородных связей внутри каждой, но соединенные водородными связями между собой.

Вытянутые (все углы в главной цепи — почти *trans*), слегка скрученные цепи образуют  $\beta$ -структуру. Она бывает (рис. 7-6) параллельной ( $\beta\uparrow\uparrow$ ), антипараллельной ( $\beta\downarrow\uparrow$ ) и смешанной (состоящей из  $\beta\uparrow\uparrow$  и  $\beta\downarrow\uparrow$ ).  $\beta$ -структура стянута водородными связями (символы /, \ и | на схемах рис. 7-6). Она существует в виде более или менее крупных листов. Так как поверхность  $\beta$ -структуры — рифленая, ее еще называют «складчатой  $\beta$ -структурой» (рис. 7-7).

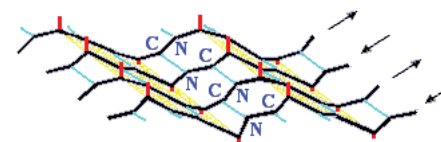
$\beta$ -структурные листы всегда несколько скручены как целое (рис. 7-8а) из-за того, что несколько скручены (рис. 7-8б) отдельные  $\beta$ -тяжи, и потому по ходу тяжа несколько меняется направленность водородных связей. А тяжи, как показал С. Чотиа, скручены из-за того, что наиболее энергетически выгодная конформация остатков с боковыми группами сдвинута к центру стерически разрешенной области (рис. 7-8в). Скрученность отдельного  $\beta$ -тяжа — левая (у L-аминокислот! — у D было бы наоборот): вы видите (рис. 7-8б), что боковые группы тяжа поворачиваются по часовой стрелке (на  $\approx 165^\circ$  на каждый остаток) по мере приближения тяжа к нам.

В результате, поворачиваются и H-связи (на  $\approx 165^\circ$  на каждый остаток, т. е. на  $-330^\circ = +30^\circ$  на пару остатков — повторяющуюся единицу  $\beta$ -структуры), так что угол между соседними тяжами  $\beta$ -структуры (если

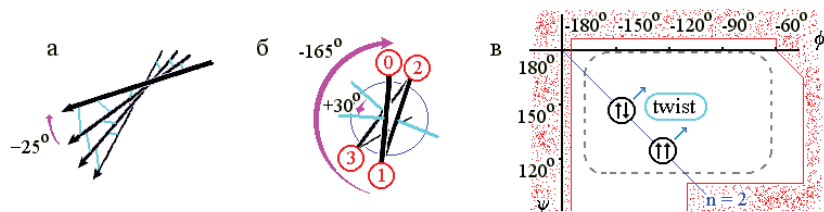
смотреть с кромки листа, рис. 7-8а) обычно составляет около  $-25^\circ$  («минус», как всегда, означает, что ближний к нам  $\beta$ -тяж повернут по часовой стрелке относительно более далекого). Таким образом,  $\beta$ -лист имеет *левое* скручивание, если смотреть с края этого листа (и *правое*, если смотреть вдоль  $\beta$ -тяжей на поворот линии H-связей).



**Рис. 7-6.** Схема хода цепи и расположения водородных связей в параллельной  $\beta\uparrow\uparrow$ , антипараллельной  $\beta\downarrow\uparrow$  и смешанной  $\beta\uparrow\uparrow\downarrow$   $\beta$ -структуре. Видно, что H-связи одного остатка в  $\beta$ -тяже направлены в одну сторону, следующего — в противоположную, и т. д.



**Рис. 7-7.** Лист  $\beta$ -структуры имеет складчатую поверхность. Тонкими светлыми линиями показаны водородные связи  $\beta$ -тяжей. Боковые группы (см. маленькие отростки) расположены на складках; каждая обращена в ту же, что и складка, сторону, т. е. направленные вниз и вверх боковые группы чередуются вдоль  $\beta$ -тяжа. Картинка взята из [3] и адаптирована



**Рис. 7-8.** (а) Скрученность  $\beta$ -листа.  $\beta$ -тяжи изображены стрелками, водородные связи между ними — тонкими светлыми линиями. (б) Схематическое изображение одного витка  $\beta$ -тяжа, вид с торца. Кружки — боковые группы; их номера возрастают по мере удаления от читателя. Светлые линии указывают направление  $C=O$ -групп, завязывающих  $H$ -связи в листе, большая стрелка — поворот  $\beta$ -тяжа при приближении к нам на один остаток, малая стрелка — поворот направленных в одну сторону водородных связей при приближении к нам на два остатка. (в) Конформация идеальной (не скрученной) параллельной ( $\uparrow\uparrow$ ) и антипараллельной ( $\uparrow\downarrow$ )  $\beta$ -структуры для поли(Gly), и усредненная конформация реальной (сложной из L аминокислот) скрученной (twist)  $\beta$ -структуры. Штриховая линия показывает область энергетического минимума для отдельно взятого остатка Ala; тонким контуром с затемнением показаны границы области разрешенных конформаций этого остатка. Диагональ  $\phi\psi$ -карты соответствует плоской периодичной структуре, имеющей 2 остатка на виток. Над диагональю лежат левые (L) спирали, под ней — правые (R). Картинки (а), (в) взяты из [3] и адаптированы

■ Есть спирали и без водородных связей, где плотная (а значит, энергетически выгодная) упаковка держится чисто на Ван-дер-Ваальсовых контактах. Это — полипролиновая спираль. При этом три цепи, каждая из которых скручена в довольно растянутую левую спираль, перевиваясь, образуют правую суперспираль — они плотно закручиваются друг вокруг друга. Из двух возможных типов полипролиновой спирали для нас важна спираль poly(Pro)II: такого рода спираль реализуется в коллагене. В этой спирали пептидные группы пролинов находятся в обычной (*trans*) конформации. Отложим более подробное рассмотрение коллагеновой спирали до соответствующего места курса, а пока ограничимся общим ее видом (рис. 7-9) и отметим на рис. 7-4 область соответствующей ей конформации цепи: видно, что она довольно близка к  $\beta$ -структуре.



**Рис. 7-9.** Общий вид тройной правой суперспирали из левых спиралей Poly(Pro)II

Параметры наиболее важных регулярных вторичных структур белковых цепей суммируются в табл. 7-1:

Таблица 7-1. Основные геометрические параметры наиболее распространенных в белках вторичных структур

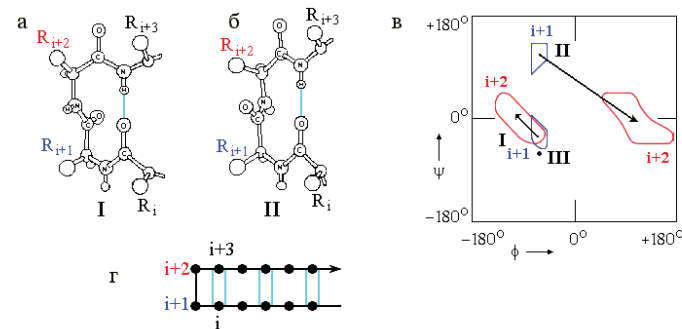
Структура	$H$ -связь	Остаток/виток	Смещение/остаток (Å)	$\phi$	$\psi$
Спираль $\alpha_R$	$CO_0-HN_{+4}$	+3,6	1,5	$-60^\circ$	$-45^\circ$
Спираль $(3_{10})_R$	$CO_0-HN_{+3}$	+3,0	2,0	$-50^\circ$	$-25^\circ$
Лист $\beta\uparrow\downarrow$	меж цепей*	-2,3	3,4	$-135^\circ$	$+150^\circ$
Лист $\beta\uparrow\uparrow$	меж цепей*	-2,3	3,2	$-120^\circ$	$+135^\circ$
Спираль Poly(Pro) II	Нет	-3,0	3,0	$-80^\circ$	$+155^\circ$

\* Расстояние между тяжами в  $\beta$ -листе: 4,8 Å

Примечание. Данные взяты из [3, 6]. Все цифры округлены. «+» в колонке «Остаток/виток» означает правую спираль, «-» — левую.

■ Кроме регулярных, в полипептидных цепях есть еще и нерегулярные вторичные структуры, т. е. стандартные структуры, не образующие длинные периодические системы.

Это — так называемые  $\beta$ -изгибы (« $\beta$ » — потому, что они часто стягивают верхушки соседних  $\beta$ -тяжей в антипараллельных  $\beta$ -шпильках). Характерный вид наиболее важных  $\beta$ -изгибов и конформации входящих в них остатков представлены на рис. 7-10. Сравните рис. 7-10в с рис. 7-4 и табл. 7-1 и обратите внимание на то, что изгибы I (и особенно III) близки по конформации к витку спирали  $3_{10}$ .



**Рис. 7-10.**  $\beta$ -изгибы. (а)  $\beta$ -изгиб типа I ( $\beta$ -изгиб типа III очень на него похож и потому не нарисован отдельно). (б)  $\beta$ -изгиб типа II. Его основное отличие от  $\beta$ -изгиба I — переворот пептидной группы, соединяющей остатки  $i+1$  и  $i+2$ . (в) Конформации фиксируемых водородной связью остатков  $i+1$  и  $i+2$  в  $\beta$ -изгибах. В  $\beta$ -изгибе III оба остатка  $i+1$  и  $i+2$  имеют одинаковую конформацию (отмечена жирной точкой). Конформации остатков  $i$  и  $i+3$  в  $\beta$ -изгибах не фиксированы; они фиксируются  $\beta$ -структурой — если она прорастает из изгиба, как на рисунке (г), где дана схема  $\beta$ -шпильки с  $\beta$ -изгибом в ее вершине. Картинки (а), (б), (в) взяты из [3] и адаптированы

В изгибы обычно входит около половины остатков, не попавших в регулярные вторичные структуры белка.

■ В заключение, несколько слов о том, как экспериментально обнаруживается вторичная структура.

Конечно, если сделан рентген (или точный многомерный ЯМР) белка, вторичная структура берется из атомных координат.

Впрочем, ЯМР (ядерный магнитный резонанс), который хорошо фиксирует сближенность (до 4–5 и менее Å) ядер атомов, позволяет определять вторичную структуру даже тогда, когда полную атомную структуру белка построить еще не удается.

Метод ЯМР основан на возбуждении радиоволнами ориентированных в сильном магнитном поле ядер — тех ядер, которые имеют нечетное число нуклонов (протонов и нейтронов): только они имеют спин и, вследствие этого, — магнитный момент. В белке это — природные «легкие» водороды  $^1\text{H}$ , а также вводимые изотопы  $^{13}\text{C}$ ,  $^{15}\text{N}$  и т. д. Магнитный резонанс наступает на радиочастоте, характерной для данного атома, причем эта частота слегка модифицируется его соседями по химическим связям и по пространству (что и позволяет судить, атом какого остатка возбуждился). Возбужденное ядро может передать свое возбуждение соседнему с ним в пространстве ядру с магнитным моментом, и оно отапортует о полученном возбуждении уже на своей частоте (что и позволит судить о сближенности этих двух магнитных ядер).

Для  $\alpha$ -спиралей особенно характерна сближенность Н-атома группы  $\text{C}^{\alpha}\text{H}$  с Н-атомом NH-группы 4-го от нее (к С-концу цепи) остатка, а для  $\beta$ -структуры — сближенность Н-атомов NH- и  $\text{C}^{\alpha}\text{H}$ -групп у остатков, непосредственно соседствующих по цепи, и у остатков, связанных Н-связями в  $\beta$ -листе (рис. 7-11).

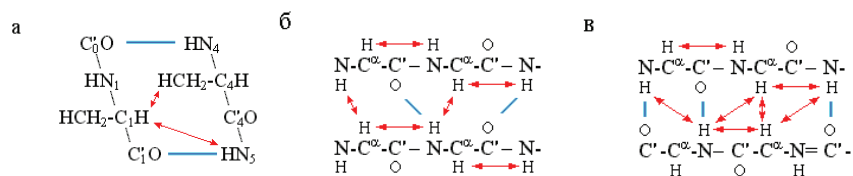


Рис. 7-11. Сближенность ( $\leftrightarrow$ ) ядер водородных атомов, наиболее характерная для  $\alpha$ -спирали (а), параллельной (б) и антипараллельной (в)  $\beta$ -структуры. Индексы при атомах главной цепи в рисунке (а) показывают взаимное расположение остатков в цепи

Однако наиболее важную, пожалуй, роль в определении вторичной структуры играет метод кругового дихроизма (КД). Он не требует знания общей пространственной структуры белка. Наоборот, структурное исследование белка обычно начинается с получения спектров КД. Метод КД основан на различии в поглощении право- и левополяризованного света в спиральных различной закрученности. Из-за этого различия в поглощении плоскополяризованный свет превращается в эллиптически поляризованный.

Характерные спектры эллиптичности в области «дальнего» ультрафиолета (190–240 нм) приведены на рис. 7-12. Показанные спектры зависят от асимметрии окружения пептидных групп и потому рапортуют о том, есть ли в белке вторичная структура, какая и сколько ее.

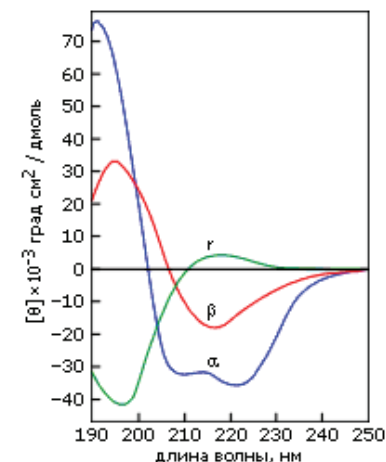
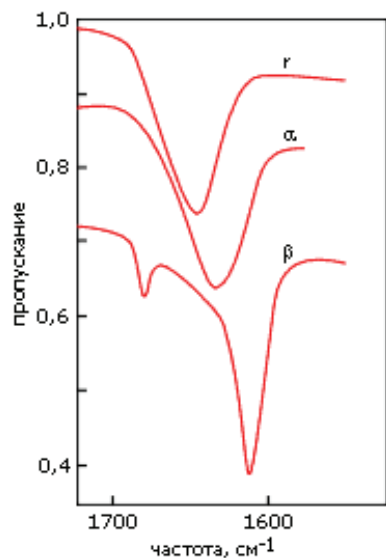


Рис. 7-12. Характерные формы спектров КД для полилизина в форме  $\alpha$ -спирали ( $\alpha$ ),  $\beta$ -структуры ( $\beta$ ) и неупорядоченного клубка ( $\gamma$ ). Картинка взята из N. J. Greenfield et al., *Biochemistry* (1969) 8:4104–4116 и адаптирована

Пептидные группы оптически возбуждаются в «дальнем УФ», при длине волны порядка 200 нм. Это примерно вдвое большая длина волны, чем та, на которой возбуждаются отдельные атомы. Причина того, что пептидная группа возбуждается более длинноволновым (т. е. менее «жестким») светом, — в делокализации электронов пептидной группы по нескольким атомам, о чем мы уже говорили.

Еще больше делокализованы электроны в ароматических группах — там они «размазаны» не по трем, как в пептидной группе, а по шести атомам. Спектры КД ароматических боковых групп приходятся на длину волны ~250–280 нм (хотя «хвост» этих спектров доходит до ~220 нм). В этом диапазоне длин волн, ~250–280 нм (в «ближнем» ультрафиолете), изучают асимметрию окружения ароматических боковых групп, т. е. эффекты, связанные с образованием уже не вторичной, а третичной структуры белка.

В скобках отмечу, что при еще большей делокализации электрона (в более крупных молекулах с кратными связями) он начинает возбуждаться уже не ультрафиолетовым, а видимым светом (400–600 нм): свечение таких молекул видно на глаз, т. е. они являются красителями.



**Рис. 7-13.** Характерные формы инфракрасных спектров пропускания, измеренных в тяжелой воде D<sub>2</sub>O для полилизина в форме α-спирали (α), β-структуры (β) и неупорядоченного клубка (r). Измерения, в данном случае, проводились в области «амид I», отражающей колебания C=O-связи. Картинка взята из H. Susi, *Methods Enzymol.* (1972) **26**:455–472 и адаптирована

Кроме ультрафиолетовых спектров, для регистрации вторичной структуры полипептидов и белков используются инфракрасные (ИК) спектры. Они отражают различия в колебаниях пептидных групп, вовлеченных и не вовлеченных в разные вторичные структуры (рис. 7-13). Эти измерения более сложны, чем измерения УФ-спектров, так как H<sub>2</sub>O поглощает примерно в той же области; поэтому такие измерения обычно проводятся в D<sub>2</sub>O. Кроме того, они требуют больше белка, чем измерения УФ-спектров, и более высоких концентраций белка в растворе.

## Лекция 8

Элементы статистической физики. Связь температуры с изменением энергии и энтропии. Вероятности состояний с различной энергией (распределение Больцмана–Гиббса). Статистическая сумма и ее связь со свободной энергией. Конформационные превращения. Понятие о фазовом переходе первого рода (переходе «*все-или-ничего*») и о не-фазовых переходах. Кинетика преодоления свободно-энергетического барьера при конформационных превращениях. Понятие о теории скоростей реакций. Параллельные и последовательные процессы. Характерные скорости диффузионных процессов.

Теперь можно было бы перейти к рассказу об образовании и распаде вторичной структуры.

Однако прежде я хочу поговорить об основах статистической физики и термодинамики «вообще», так как без этого трудно рассказывать и о стабильности вторичной структуры, и о стабильности белков, и о кооперативных переходах в полипептидах и белках, и о кинетике этих переходов.

Термодинамика дает представление о типах возможных кооперативных переходов в системах, состоящих из очень большого числа частиц. Статистическая физика позволяет указать, когда и какие переходы произойдут в рассматриваемой системе и описать детали этих переходов, исходя из свойств рассматриваемых частиц и взаимодействий между ними.

Прежде всего мы рассмотрим основные понятия статистической физики и термодинамики — энтропию, температуру, свободную энергию и статистическую сумму.

Итак, системы с большим числом степеней свободы (т. е. состоящие из большого числа молекул или даже из одной большой и гибкой молекулы) описываются при помощи статистической физики. «Статистической» — ибо число конфигураций таких больших систем колоссально. Только один



пример: если каждое из  $N$  звеньев цепи может находиться всего в двух возможных конфигурациях (например: «спиральном» и «вытянутом»), то вся  $N$ -звенная цепь имеет  $2^N$  возможных конфигураций. То есть «нормальная» для белка цепь из 100 звеньев может иметь, по крайней мере,  $2^{100}$ , или около 1 000 000 000 000 000 000 000 000 000 000 000 конфигураций. Это очень много. Чтобы просмотреть их все, тратя 1 наносекунду ( $10^{-9}$  с) на просмотр каждой, понадобилось бы  $3 \cdot 10^{13}$  лет — в 1000 раз больше, чем существует наша Вселенная... А ведь на опыте в пробирке находятся миллиарды таких цепей — не говоря о растворителе! И если бы мы захотели инвентаризовать все конфигурации этих цепей, мы бы пропали навсегда. Но нас, конечно, будут интересовать более простые и разумные вещи, например, средняя (т. е. *статистически усредненная*) спиральность цепей и то, как эта спиральность меняется с нагреванием. А замечательным свойством статистического, т. е. пренебрегающего всеми несущественными подробностями усреднения, является кардинальное упрощение ситуации.

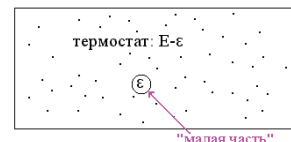
Важнейшую роль в статистическом усреднении играет *энтропия*. Она говорит, сколько конфигураций системы (или, как говорят, сколько ее *микроскопических состояний*) соответствует наблюдаемому нами, т. е. *макроскопическому* (усредненному длительностью нашего наблюдения и большим числом одновременно наблюдаемых молекул) *состоянию* изучаемой системы. Мы уже говорили на эту тему, но тогда мы рассматривали, по существу, только один частный случай: тогда число «микроскопических состояний» определялось объемом, в который была заключена молекула, а ее энтропия была пропорциональна логарифму этого объема (причем нахождение в данном объеме — например, в данной комнате — и было «макроскопическим состоянием» молекулы).

Здесь уместно ответить на естественный вопрос: почему физики предпочитают рассматривать *логарифм* числа микросостояний, а не само это число? Дело в том, что, если рассматривать много отдельных систем (например — отдельных молекул), то мы увидим, что их объемы, энергии, числа степеней свободы при этом складываются, а количества возможных состояний перемножаются. Это различие неудобно, оно сделало бы громоздким все расчеты. Но вот *логарифмы* чисел возможных состояний складываются при перемножении этих чисел [вы помните, что  $\ln(AB) = \ln(A) + \ln(B)$ ] — складываются, как энергии или объемы. Эта однотипность удобна при решении задач, и к тому же операция сложения позволяет легко воспользоваться всей мощью дифференциального исчисления.

■ Теперь — *температура*. Она тесно связана с энтропией: *нет большого числа состояний, нет энтропии — нет и температуры*. Для выяснения связи рассмотрим замкнутую (т. е. не обменивающуюся энергией с окружением) систему. Пусть ее полная энергия равна  $E$ , и пусть эта система нахо-

дится в равновесии, т. е. все ее микросостояния, имеющие энергию  $E$ , равновероятны (а имеющие другую энергию — имеют нулевую вероятность).

Выделим теперь в этой системе «наблюдаемую малую часть» (например, отдельную молекулу в газе или макромолекулу с окружающей ее жидкостью). Вся остальная система тогда может рассматриваться как термостат, в который погружена наша «малая часть».



Замкнутая система,  
 $E = \text{const}$

Разобьем все микроскопические состояния системы, имеющей полную энергию  $E$ , на классы, различающиеся состоянием нашей «малой части». Чем больше микроскопических состояний всей системы в данном классе, тем больше вероятность его (этого класса) наблюдения, т. е. тем больше вероятность наблюдения данного состояния нашей «малой части».

Пусть наша «малая часть» фиксирована в каком-то микроскопическом состоянии (например, пусть наблюдаемая молекула в газе имеет заданное положение в пространстве и заданную скорость). Пусть она обладает при этом энергией  $\varepsilon$ . Поскольку вся система в целом замкнута, то *полная энергия всей системы сохраняется* (закон!), и энергия термостата равна  $E - \varepsilon$ . Пусть этой энергии соответствует  $M_{\text{therm}}(E - \varepsilon)$  возможных микросостояний термостата. Тогда вероятность наблюдения данного состояния нашей «малой части» просто пропорциональна  $M_{\text{therm}}(E - \varepsilon)$ . [Замечание. Мы здесь неявно предположили, что конкретное микросостояние нашей системы не влияет на микросостояния термостата. Строго говоря, это не вполне верно (точнее: это верно только для «термостата», состоящего из разреженного идеального газа), но учет всех явлений на границе «нашей системы» и термостата только затемнил бы все дальнейшее изложение. Для совмещения строгости и наглядности можно представить себе для начала, что наша «наблюдаемая часть» заключена в фиксированный и непроницаемый для молекул термостата объем, а уже потом, если понадобится (нам — не понадобится...), изучить взаимодействие «наблюдаемой части» с термостатом.]

Если число возможных состояний термостата равно  $M_{\text{therm}}(E - \varepsilon)$ , то логарифм этого числа, *по определению*, пропорционален энтропии термостата:

$$S_{\text{therm}}(E - \varepsilon) = k \ln[M_{\text{therm}}(E - \varepsilon)], \quad (8.1)$$

коэффициент  $k$  здесь вводится только для того, чтобы энтропия, как обычно, измерялась в единицах кал/К; как мы увидим позже, он окажется просто константой Больцмана.

Энергия  $\varepsilon$  «малой части» тоже должна быть относительно мала. Поэтому можно воспользоваться обычным дифференциальным разложением величины  $S_{\text{therm}}(E - \varepsilon)$  по малому параметру  $\varepsilon$  [вы, конечно, помните:  $f(x_0 + dx) = f(x_0) + (df/dx)|_{x_0} dx + 1/2(d^2f/dx^2)|_{x_0} (dx)^2 + \dots = f(x_0) + (df/dx)|_{x_0} dx$  при малом  $dx$ , причем  $(df/dx)|_{x_0}$  означает, что производная  $df/dx$  берется в точке  $x_0$ ]. Итак,

$$S_{\text{therm}}(E - \varepsilon) = S_{\text{therm}}(E) - (dS_{\text{therm}}/dE)|_E \varepsilon. \quad (8.2)$$

[Примечание. Так как и  $S$ , и  $E$  пропорциональны числу частиц, то  $dS_{\text{therm}}/dE$  от числа частиц в термостате не зависит, в то время как  $d^2S_{\text{therm}}/dE^2$  *обратно* пропорционально этому числу, т. е.  $d^2S_{\text{therm}}/dE^2 \rightarrow 0$  в очень большом термостате; это и позволяет нам пренебречь членами порядка  $\varepsilon^2$ , а также  $\varepsilon^3$  и т. д. — в формуле 8.2.]

Значит, число возможных микросостояний термостата зависит от энергии  $\varepsilon$  нашей «малой части» как

$$M(E - \varepsilon) = \exp[S_{\text{therm}}(E - \varepsilon)/k] = \exp[S_{\text{therm}}(E)/k] \times \exp[-\varepsilon \times \{(dS_{\text{therm}}/dE)|_E/k\}]. \quad (8.3)$$

При этом ни общий множитель  $\exp[S_{\text{therm}}(E)/k] = M(E)$ , ни число  $(dS_{\text{therm}}/dE)|_E$  не зависит ни от  $\varepsilon$ , ни от конкретного микросостояния нашей «малой части» вообще.

Так как число микросостояний должно расти с энергией (чем больше энергия, тем большим числом способов ее можно разделить), то формула (8.3) означает просто следующее: чем больше энергии отобрала у термостата наша «малая часть», тем меньше энергии осталось у него, и тем меньшим числом способов ее можно разделить. При этом падение числа возможных микросостояний термостата (числа способов раздела его энергии) *экспоненциально* зависит от энергии, содержащейся в нашей «малой части».

Резюмируем. Вероятность наблюдения заданного микросостояния нашей «малой части» (молекулы и т. п.) пропорциональна  $\exp[-\varepsilon \{(dS_{\text{therm}}/dE)|_E/k\}]$ , где  $\varepsilon$  — энергия этой «малой части», а величина  $(dS_{\text{therm}}/dE)|_E/k$  зависит не от «малой части», а только от усредненных свойств окружающей ее среды.

Но, по формуле Больцмана, вероятность пребывания молекулы в заданном состоянии с энергией  $\varepsilon$  пропорциональна  $\exp(-\varepsilon/k_B T)$  [где  $T$  — температура, а  $k_B$  — константа Больцмана!] Сравнивая тождественные выражения  $\exp[-\varepsilon \{(dS_{\text{therm}}/dE)|_E/k\}]$  и  $\exp(-\varepsilon/k_B T)$  — видим, что

$$(dS_{\text{therm}}/dE)|_E = 1/T, \quad (8.4)$$

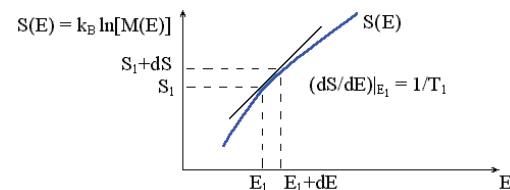
а  $k$  в формулах (8.1) и (8.3) — это та же константа Больцмана  $k_B$  — если, как обычно, энергия измеряется в джоулях (или в калориях), температура в К, а энтропия в Дж/К (или в кал/К).

Формулы (8.1) и (8.4), определяющие температуру как величину, обратную скорости изменения энтропии  $S$  (или логарифма числа микросостояний) с энергией  $E$  системы, — основные формулы статистической физики и термодинамики.

Они показывают, в частности, что  $\ln[M(E + k_B T)] = S(E + k_B T)/k_B = [S(E) + (k_B T)(1/T)]/k_B = \ln[M(E)] + 1$ , т. е. при увеличении энергии на  $k_B T$  число микросостояний растет в  $e = 2,72$  раз, т. е. примерно втрое, *независимо* от размеров системы, действующих в ней сил и т. д.

Они означают также, что, зная число микросостояний любой большой системы («термостата») при разных ее энергиях (или, как говорят, «плотность энергетического спектра» системы), а точнее, зная зависимость логарифма плотности энергетического спектра системы от энергии, мы можем найти температуру, соответствующую каждой энергии этой системы. Соответствующий график показан на рис. 8-1.

Надо только, чтобы микросостояний было очень много — производную  $dS/dE$  можно, естественно, брать только тогда, когда в малом (по сравнению с  $k_B T$ ) интервале энергий находится много микросостояний системы. Именно поэтому температура появляется только в науках, рассматривающих огромное число возможных микросостояний.



**Рис. 8-1.** Определение температуры большой системы. Зависимость энтропии  $S$  этой системы от ее энергии  $E$  показана жирной кривой. Наклон этой кривой,  $dS/dE$ , определяет температуру  $T$ , соответствующую данной энергии  $E$ .  $M(E)$  — число микросостояний с заданной энергией  $E$ , т. е. плотность спектра энергий состояний системы

■ Продолжим рассмотрение системы, заключенной в фиксированный и непроницаемый для молекул термостата объем (через стенки которого она, однако, может обмениваться энергией с этим термостатом, температура которого равна  $T$ ).

Формулы (8.1) и (8.3) показывают, что вероятность пребывания нашей системы в заданном микросостоянии  $i$  с энергией  $\varepsilon_i$  при температуре  $T$  есть

$$w_i(T) = \exp(-\varepsilon_i/k_B T) / Z(T), \quad (8.5)$$

где

$$Z(T) = \sum_j \exp(-\varepsilon_j/k_B T) \quad (8.6)$$

— это нормировочный коэффициент, учитывающий, что сумма вероятностей всех состояний,  $\sum_j w_j$ , должна быть равна 1 (сумма  $\sum_j$  здесь и выше берется по всем возможным микросостояниям  $j$  рассматриваемой «малой системы»).

Величина  $Z$  называется *статистической суммой* для рассматриваемой системы. Зная ее, можно вычислить по формуле (8.5) вероятность каждого из микросостояний этой системы при заданной температуре, — и, далее, ее среднюю энергию при этой температуре

$$E(T) = \sum_j w_j \varepsilon_j \quad (8.7)$$

и ее среднюю энтропию

$$S(T) = k_B \sum_j w_j \ln(1/w_j). \quad (8.8)$$

Обратите внимание, что формула (8.8) усредняет  $\ln(w_j)$  по всем состояниям системы  $j$  с учетом их (состояний  $j$ ) вероятностей  $w_j$ . Она есть прямое обобщение давно знакомого нам (см. формулу (8.1)) определения энтропии  $S = k_B \ln[M(E)]$  на случай, когда все состояния системы имеют самые разные вероятности — а не только две, следовавшие тогда из закона сохранения энергии:  $w_j = 1/M(E)$  для всех  $M(E)$  состояний с  $E_j = E$ , и  $w_j = 0$  при  $E_j \neq E$ .

**Внутренний голос:** Я чувствую, что смысл формулы (8.8) надо объяснить лучше, и доказать, что  $S(T)/k_B$  там является логарифмом среднего числа состояний системы.

**Лектор:** Тогда прошу потерпеть немного математики...

Для доказательства рассмотрим очень большое число  $N$  систем, каждая из которых может быть в состоянии «1» с вероятностью  $w_1$ , в состоянии «2» с вероятностью  $w_2$ , ..., в состоянии « $J$ » с вероятностью  $w_J$ . Значит, в среднем,  $n_1 = w_1 N$  из  $N$  систем находится в состоянии «1»,  $n_2 = w_2 N$  из  $N$  систем находится в состоянии «2», и т. д. (причем  $\sum_j w_j N = N$ ). Сколькими способами можно распределить наши  $N$  систем по этим  $J$  состояниям так, чтобы  $n_1$  из этих систем были бы в состоянии «1»,  $n_2$  в состоянии «2», и т. д.?

Этот подсчет делается так. Занумеруем наши системы номерами 1, 2, ...,  $N$ . Теперь выберем из них те  $n_1$ , что находятся в состоянии 1. Сколькими различными способами это можно сделать? Так как первую из выбранных мы выбираем из  $N$  систем, вторую — из  $N-1$  оставшихся, ...,  $n_1$ -ю — из  $N-(n_1-1)$  оставшихся — то, *при заданном порядке отбора*, мы можем выбрать  $n_1$  из  $N$  систем  $N(N-1) \dots [N-(n_1-1)]$  способами. Но нам не важно, в каком порядке отбирались системы: сначала 5-я, потом 10-я, или наоборот: это — *не разные* отборы, их *результат* будет *один и тот же*. А так как  $n_1$  отобранных систем можно выбрать при  $n_1(n_1-1) \dots 1$  разных порядках действий [сначала можно выбрать первую из всех  $n_1$  штук, потом — вторую из  $(n_1-1)$  оставшихся, и т. д.] — то выбор  $n_1$  из  $N$  систем даст  $\{N(N-1) \dots [N-(n_1-1)]\} / \{n_1(n_1-1) \dots 1\}$  *разных* результатов. Обозначая  $n(n-1) \dots 1$  через  $n!$  (читается: « $n$  факториал») — видим, что выбор  $n_1$  из  $N$  систем даст  $N! / [(N-n_1)! n_1!]$  разных результатов. Далее мы должны выбрать  $n_2$  из  $N-n_1$  оставшихся систем. Согласно описанному выше, это можно сделать  $(N-n_1)! / [(N-n_1-n_2)! n_2!]$  способами при каждом способе выбора  $n_1$  из  $N$  систем на первом шаге. Итого, выбрать  $n_1$

систем в первом состоянии и  $n_2$  систем во втором состоянии из  $N$  систем можно  $\{N! / [(N-n_1)! n_1!]\} \cdot \{(N-n_1)! / [(N-n_1-n_2)! n_2!]\} = N! / [(N-n_1-n_2)! n_2! n_1!]$  разными способами. Продолжая ту же цепь рассуждений, видим, что наши  $N$  систем можно распределить  $N! / [n_1! \dots n_2! \dots n_1!]$  разными способами так, чтобы  $n_1$  из них были бы в состоянии «1»,  $n_2$  в состоянии «2», ...,  $n_J$  в состоянии « $J$ » (напомню, что  $n_j = N - n_1 - n_2 - \dots - n_{j-1}$ ). Предпоследнее усилие: существует формула Стирлинга, согласно которой  $n! \approx (n/e)^n$ . Значит,

$$\begin{aligned} N! / [n_1! \dots n_2! \dots n_1!] &\approx (N/e)^{n_1 + \dots + n_1} / [(n_1/e)^{n_1} \dots (n_1/e)^{n_1}] = \\ &= (N/n_1)^{n_1} \dots (N/n_1)^{n_1} = (1/w_1)^{Nw_1} \dots (1/w_1)^{Nw_1}. \end{aligned} \quad (8.8a)$$

В последнем равенстве мы воспользовались тем, что  $n_1 = w_1 N$  и т. д. И, наконец, интересующее нас число способов распределения наших  $N$  систем по  $J$  состояниям (стем, чтобы  $n_1$  из этих систем были бы в состоянии «1»,  $n_2$  в состоянии «2», и т. д.) получается из (8.8a) в виде  $[1/(w_1^{w_1} \dots w_1^{w_1})]^N$ . А так как это число (для  $N$  независимых систем) — это просто число состояний одной системы, возведенное в  $N$ -ю степень, то число состояний одной системы равно  $1/(w_1^{w_1} \dots w_1^{w_1})$ , и его логарифм есть — как то и должно было оказаться — именно величина  $S(T)/k_B$  из формулы (8.8).

Все!

Статистическая сумма играет очень важную роль в статистической физике, так как величина  $Z(T)$  позволяет прямо вычислить свободную энергию системы, находящейся в фиксированном объеме:

$$F(T) = E(T) - TS(T) = \sum_j \{w_j \varepsilon_j - T[-k_B w_j \ln(w_j)]\} = -k_B T \ln[Z(T)] \quad (8.9)$$

(в предпоследнем равенстве мы воспользовались формулой (8.5)). Затем можно прямо найти энтропию системы:  $S(T) = -dF/dT$  (мы с этой формулой уже встречались, но проверьте ее самостоятельно, исходя из определений, даваемых формулами (8.9), (8.6), (8.5) и (8.8)), а затем — и ее энергию  $E(T) \equiv F(T) + TS(T) = F(T) - T(dF/dT)$ .

■ *Существенные, но второстепенные замечания.*

1) Если «малая система» сама имеет много степеней свободы, то она имеет свою энтропию, и потому — свою внутреннюю температуру. Эта внутренняя температура «малой системы»  $T_{in}$  равна температуре термостата  $T$ . Это следует из того, что, согласно приведенным выше определениям и уравнениям,  $T_{in} \equiv dE(T)/dS(T) = d[F(T) - T(dF/dT)]/d(-dF/dT) = [dF/dT - T(d^2F/dT^2) - dF/dT]dT / (-d^2F/dT^2)dT = T$ .

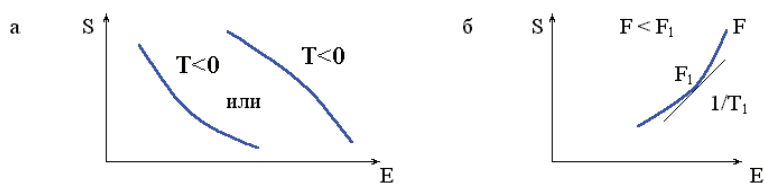
2) Полная энергия складывается из кинетической и потенциальной. Первая зависит только от скоростей частиц, а вторая — только от их места в пространстве, но не от скоростей. «Микросостояние» каждой частицы задается ее координатой в пространстве и ее скоростью. В классической (не квантовой) механике сочетание скоростей и координат может быть любым (как мы знаем, квантовое соотношение Гейзенберга,  $\Delta v \Delta x \sim \hbar/m$ , при комнатной температуре ограничивает сочетание скоростей и координат только у очень легких частиц — электронов и протонов). Это значит, что вероятности координат и скоростей можно «расцепить», т. е.  $w(\varepsilon_{кинет} + \varepsilon_{коорд.}) \sim \exp(-\varepsilon_{кинет}/k_B T) \times \exp(-\varepsilon_{коорд.}/k_B T)$ . Продолжая несложные вычисления (предлагаю их сделать самим!), легко увидеть, что свободные энергии, энергии и

энтропии тоже можно «расцепить» на кинетические и координатные части, т. е.  $F = F_{\text{кин.}} + F_{\text{коорд.}}$  и т. д. Важно, что кинетические части *не зависят* от конфигурации системы (фактически, они дают только постоянный вклад в теплоемкость) и их можно отбросить при изучении конформационных превращений. Поэтому в дальнейшем мы будем говорить только о конфигурационных (или конформационных) энергетических спектрах, энергиях, энтропиях и т. д.

3) Выше мы суммировали по микросостояниям, но в рамках классической механики мы можем с тем же успехом интегрировать по координатам и скоростям, задающим микросостояние каждой частицы.

4) Равновесная температура может быть *только положительной*. Иначе интегрирование вероятностей по скоростям, т. е.  $\int \exp(-mv^2/2kT)dv$ , обращается в бесконечность на больших скоростях и система «взрывается».

Поэтому стабильное состояние никак не может наблюдаться в тех областях конформационного энергетического спектра, где его плотность (и, значит, энтропия системы) падает с ростом ее энергии: этим областям соответствует  $T < 0$  (см. формулы (8.4) и (8.1) и рис. 8-2а).



**Рис. 8-2.** Участки графика  $S(E)$ , которым не может соответствовать никакое стабильное состояние системы. (а) Участки, где энтропия  $S(E)$  падает с ростом энергии  $E$ : здесь  $T = 1/(dS/dE) < 0$ . (б) «Вогнутый» участок зависимости  $S$  от  $E$ : здесь в точке касания (где  $dS/dE = 1/T_1$ ) свободная энергия  $F = E - TS$  не меньше, а больше, чем на соседних с ней участках кривой  $S(E)$ ; ср. с рис. 8-3

5) Величина  $k_B T$  измеряется в единицах типа «энергия на частицу» или «энергия на моль ( $6,02 \cdot 10^{23}$ ) частиц». Если бы люди с самого начала выражали температуру в энергетических единицах, константа Больцмана  $k_B$  не потребовалась бы вообще. Однако история сложилась так, что сначала ввели температурную единицу — «градус» — а потом уже поняли, что «градус» можно всегда пересчитать, конвертировать в «энергию на расчете на частицу»: для этого достаточно умножить его (градус) на некоторую константу — «константу Больцмана». Соответственно, размерность  $k_B$  — «энергия на частицу/градус». Величина  $k_B$  зависит от того, в чем выражается энергия — то ли в «джоуль на частицу», то ли в «калориях на моль частиц». Согласно измерению энергетической величины градуса,  $k_B = 1,38 \cdot 10^{-23}$  (джоуль/частицу)/градус. Однако  $k_B$  можно выразить и по-иному — в расчете не на штуку, а на моль частиц:  $1,38 \cdot 10^{-23}$  (джоуль/частицу)/градус =  $1,38 \cdot 10^{-23}$  (джоуль  $\times$   $[6,02 \cdot 10^{23}] / [6,02 \cdot 10^{23}$  частиц]/градус =  $[1,38 \cdot 10^{-23} \times 6,02 \cdot 10^{23}]$  (джоуль/моль частиц)/градус = 8,31 (Дж/моль)/градус = 1,99 (кал/моль)/градус. Последнюю величину, 1,99 (кал/моль)/К, традиционно именуют «газовой постоянной»  $R$ .

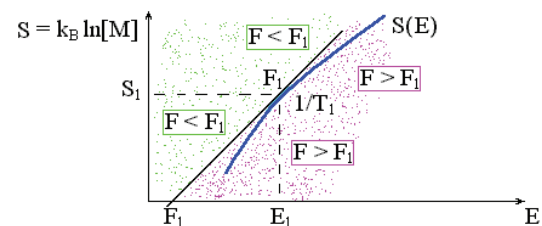
6) Удельная энтропия часто выражается в энтропийных единицах, «э.ед»: э.ед. = (кал/моль)/К  $\approx k_B/2$ . Стало быть, энтропии в 1 э.ед. соответствует  $e^{1/2} \approx 1,65$  состояний системы в расчете на частицу.

■ Теперь, наконец, мы рассмотрим **конформационные превращения**. Именно для их грамотного рассмотрения мы и повторяли азы статистической физики и термодинамики.

Конформационные превращения бывают постепенные («континуальные») и резкие («фазовые»). Попробуем разобраться в их различии, используя графики плотности энергетических состояний — вроде тех, что представлены на рис. 8-1. Для этого прежде всего надо научиться находить на графике стабильное состояние, или состояния, нашей системы при заданной температуре среды.

Пусть нам задано число  $M(E)$  состояний нашей системы («макромолькулы») при каждой ее энергии  $E$ , а также температура  $T_1$  окружающей среды («термостата»). Как, зная график энтропии  $S(E) = k_B \ln[M(E)]$  и  $T_1$ , найти на графике точку, соответствующую стабильному состоянию нашей системы? Температура  $T_1$  показывает, какой наклон  $dS/dE$  должен быть в той точке графика, которая нас интересует: здесь  $dS/dE = 1/T_1$ , ведь температура у нашей макромолькулы та же, что и у среды.

Но, вообще говоря, таких точек — точек, где график имеет заданный наклон  $1/T_1$ , — может быть несколько (см. ниже, рис. 8-5). Какая из них отвечает стабильному состоянию? Рассмотрим касательную к графику в точке  $E_1$ , где  $dS/dE|_{E_1} = 1/T_1$ . Уравнение такой касательной таково:  $S - S(E_1) = (E - E_1)/T_1$ , или:  $E - T_1 S = E_1 - T_1 S(E_1)$ . Величина  $F_1 = E_1 - T_1 S(E_1)$  — это просто свободная энергия при температуре  $T_1$ . Вдоль касательной, как мы видим, величина  $(E - T_1 S)$  неизменна. Всюду слева от касательной величина  $E - T_1 S$  ниже, а всюду справа — выше, чем на касательной (рис. 8-3).

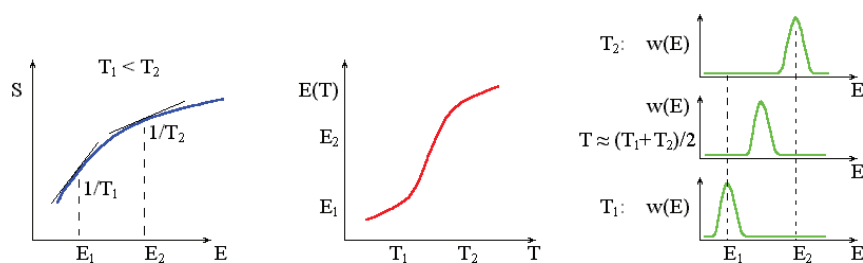


**Рис. 8-3.** Графическое определение температуры и свободной энергии. Зависимость энтропии  $S = k_B \ln(M)$  от энергии  $E$  показана жирной кривой  $S(E)$ .  $M(E)$  — плотность спектра энергий состояний системы. Наклон кривой  $S(E)$ ,  $dS/dE = 1/T$ , определяет температуру  $T$ . Физически возможным температурам  $T > 0$  отвечает подъем кривой  $S(E)$ .  $F_1$  — свободная энергия, соответствующая данному энергетическому спектру при температуре  $T_1$ . Величина  $F = E - T_1 S$  постоянна вдоль касательной и равна  $F_1$ . Левее (и выше) этой касательной  $F = E - T_1 S < F_1$ , правее (и ниже) касательной  $F = E - T_1 S > F_1$ . Так как кривая  $S(E)$  здесь выпукла, т. е. лежит правее и ниже касательной, то точка касания отвечает минимуму свободной энергии при температуре  $T_1$ . Дальнейшие пояснения — в тексте



Последнее означает, в частности, что «вогнутым» участкам зависимости  $S$  от  $E$  (рис. 8-2б) стабильное состояние соответствовать никак не может: здесь в точке касания  $F$  не меньше, а больше, чем на соседних с ней участках кривой  $S(E)$ . Это значит, что, сдвинувшись из точки касания с вогнутой кривой  $S(E)$ , система понизит свою свободную энергию.

Если кривая  $S(E)$  всюду выпукла, то ее наклон все падает по мере увеличения энергии  $E$ . Тогда каждому наклону  $1/T$  отвечает только одна точка кривой (рис. 8-1, 8-4) — т. е. она, эта точка, и отвечает стабильному состоянию при данной температуре  $T$ . С изменением температуры эта точка постепенно движется, а система постепенно («градуально») меняет свое термодинамическое состояние, т. е. энтропию и энергию (рис. 8-4).

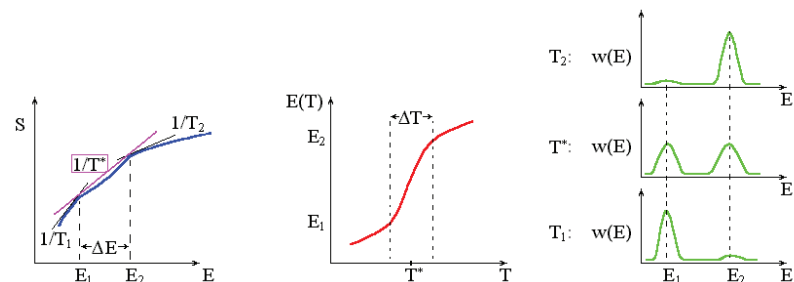


**Рис. 8-4.** Градуальное изменение состояния системы с изменением температуры. Для этого необходимо, чтобы кривая  $S(E)$  была выпукла. Зная зависимость  $S(E)$  (слева), можно найти зависимости  $T(E)$  и, далее,  $E(T)$ ; последняя показана в центре рисунка. Справа:  $w(E)$  — распределение наблюдаемых состояний молекул по энергиям, т. е. вероятность иметь энергию  $E$  при заданной температуре  $T$ ; вероятность  $w(E)$  пропорциональна  $\exp[-(E-TS(E))/k_B T]$

Если наклон кривой  $S(T)$  то падает, то растет по мере увеличения энергии  $E$  (рис. 8-5) — то касательных с одинаковым наклоном к этой кривой может быть несколько — и касание самой левой из них (той, у которой  $F$  меньше) выделяет стабильное состояние. При этом до какой-то температуры  $T^*$  «лучшей» (соответствующей наименьшему значению свободной энергии  $F$ ) будет касательная в районе малых энергий, после  $T^*$  «лучшей» будет касательная в районе больших энергий.

При температуре  $T^*$  высоко- и низкоэнергетичные структуры будут иметь одинаковую свободную энергию и одинаковую вероятность существования. Это значит, что, имея множество одинаковых систем при температуре перехода  $T^*$ , мы будем наблюдать половину из них в высоко-, а половину — в низкоэнергетичном состоянии (см. рис. 8-5, справа). Иными

словами, каждая система проводит половину времени в высоко-, а половину — в низкоэнергетическом состоянии. Такое «сосуществование» двух возможных, и при этом сильно отличающихся по энергии состояний будет охватывать некий, очень узкий (особенно для больших, «макроскопических» систем) диапазон температур в окрестности  $T^*$ ; этот диапазон мы вскоре оценим.



**Рис. 8-5.** Фазовый переход типа «все-или-ничего» (в макроскопических телах он называется «фазовым переходом первого рода») характеризуется резким изменением состояния системы при изменении температуры. Для этого необходимо, чтобы кривая  $S(E)$  имела вогнутый участок: при температуре перехода  $T^*$  внутри этого участка свободная энергия выше, чем на его краях (см. рис. 8-2б). Переход происходит в узком диапазоне  $\Delta T$  температур, отвечающем сосуществованию (т. е. примерно равной вероятности) низко- и высокоэнергетичного состояния. Касательные соответствуют: температуре середины перехода  $T^*$ , а также  $T_1 < T^*$  и  $T_2 > T^*$ . Обратите внимание, что температура  $T^*$  перехода типа «все-или-ничего» (т. е. фазового перехода первого рода) совпадает именно с серединой резкого изменения энергии и с максимальным раздвоением функции распределения молекул по энергиям  $w(E)$

Наиболее важно, однако, то, что состояния со «средними» энергиями *не будут проявляться* в качестве стабильных: из-за «прогиба» кривой  $S(E)$  соответствующие этим «промежуточным» состояниям точки графика лежат *справа* от касательной, соответствующей температуре  $T^*$ , т. е. при температуре  $T^*$  свободная энергия этих «промежуточных» состояний выше, чем у структур, соответствующих точкам касания, а вероятность проявления ничтожна. Говорят тогда, что два стабильных состояния разделены «свободно-энергетическим барьером».

В этих условиях и протекает переход типа «все-или-ничего».

В микроскопических системах (и, в частности, в белковых глобулах) этот переход происходит путем скачка через свободно-энергетический барьер при температуре  $T^*$ ; мы об этом поговорим в конце лекции.

В макроскопических системах (скажем, в пробирке с замерзающей водой) барьер (при температуре  $T^*$ ) столь высок, что скачок через него занял бы практически бесконечное время. Поэтому в макроскопических системах наблюдается *гистерезис*: сохранение состояния системы вплоть до небольшого переохлаждения (при замерзании) или перегрева (при плавлении) системы по сравнению с температурой  $T^*$ , после чего переход протекает путем временного (т. е. нестабильного) расслоения системы на фазы (например, жидкую и твердую). Такой переход в макроскопических системах называется *фазовым переходом первого рода*.

- Теперь я хочу сразу оценить величину температурного интервала сосуществования высоко- и низкоэнергетичных структур. В середине перехода при температуре  $T^*$  свободная энергия низкоэнергетичной фазы  $F_1(T^*) = E_1 - T^*S_1$  равна свободной энергии высокоэнергетичной фазы  $F_2(T^*) = E_2 - T^*S_2$ , т. е. в середине перехода

$$(E_2 - E_1) = T^*(S_2 - S_1). \quad (8.10)$$

При небольшом (на  $\delta T$ ) сдвиге температуры от  $T^*$  свободные энергии фаз немного меняются, так что разность свободных энергий фаз составляет  $\delta F = F_1(T^* + \delta T) - F_2(T^* + \delta T) = (dF_1/dT)\delta T - (dF_2/dT)\delta T = (-S_1\delta T) - (-S_2\delta T) = \delta T(S_2 - S_1)$ . Фазы сосуществуют (т. е. они примерно равновероятны — скажем, соотношение их вероятностей меняется от 10:1 в пользу первой до 1:10 в пользу второй фазы), пока  $\exp(-\delta F/k_B T^*)$  лежит между 10 и 1/10, т. е. пока  $\delta F/k_B T^*$  находится где-то от  $\ln(1/10) \approx -2$  до  $\ln(10) \approx +2$ . В этой области  $|\delta T| < 2k_B T^*/(S_2 - S_1)$ . Значит, температурный интервал сосуществования фаз есть

$$\begin{aligned} \Delta T &\approx [2k_B T^*/(S_2 - S_1)] - [-2k_B T^*/(S_2 - S_1)] = \\ &= 4k_B T^*/(S_2 - S_1) = 4k_B (T^*)^2/(E_2 - E_1). \end{aligned} \quad (8.11)$$

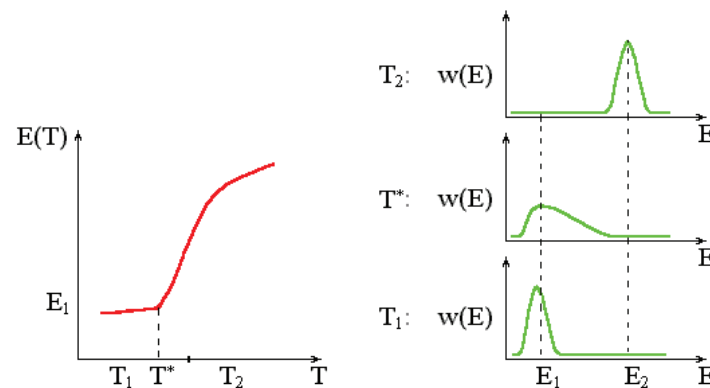
Рассмотрим поучительный численный пример. Если  $T^* \sim 300\text{K}$  (т. е.  $k_B T^* \sim 0,6$  ккал/моль), а  $E_2 - E_1$  составляет, как то характерно для плавления белков, порядка 50 ккал/моль  $\approx 100k_B T^*$  [т. е.  $50/(6 \cdot 10^{23}) \sim 10^{-22}$  килокалорий\_на\_белок], то  $\Delta T$  — порядка десяти градусов. То есть расплавленные и твердые молекулы белка сосуществуют в диапазоне нескольких градусов вокруг середины перехода. Если же  $E_2 - E_1$  составляет порядка 50 килокалорий\_на\_штуку (как при плавлении льдины размером с бутылку), то диапазон сосуществования фаз  $\Delta T$  порядка  $10^{-23}$  градуса.

Иными словами, для фазовых переходов типа «все-или-ничего» в малых системах и — в гораздо большей степени — для фазовых переходов первого рода в макроскопических системах — характерен резкий скачок энергии в узком диапазоне температур. В макроскопических системах этот

диапазон практически бесконечно узок, а для макромолекул он охватывает несколько градусов, т. е. тоже узок по сравнению с «диапазоном наблюдения», обычно охватывающим от 0 до 100°C. А вот в небольших олигопептидах резкого скачка нет — здесь диапазон изменения энергии может охватывать все «экспериментальное окошко» исследуемых температур.

- Несколько слов о фазовых переходах *второго* рода. Если для переходов первого рода характерен *скачок энергии* системы (и ее энтропии, объема и плотности), то для переходов второго рода характерен лишь *излом* зависимости  $E(T)$ , т. е. скачок скорости изменения энергии с температурой, иными словами, *скачок теплоемкости*.

Подчеркиваю, что переход второго рода происходит в *точке излома* (а когда речь идет о маленьких системах типа белка — в узкой области резкого поворота) кривой  $E(T)$  (рис. 8-6), а вовсе не посередине часто идущей за ним более или менее «S-образной» зависимости энергии (или другого наблюдаемого параметра) от температуры. Такой переход возникает, например, когда система находится в каком-то определенном состоянии вплоть до (или начиная с) определенной температуры (и все поглощающееся при нагреве тепло идет лишь на рост тепловых колебаний), а после (или до) этой критической температуры часть тепла идет дополнительно на перестройку системы (например, в ферромагнетике ниже температуры Кюри — на разрушение его спонтанно намагниченного состояния).



**Рис. 8-6.** Характерная зависимость энергии  $E(T)$  и функции распределения вероятности наблюдаемых состояний по энергии,  $w(E)$ , при фазовом переходе *второго* рода.  $T^*$  — температура этого фазового перехода. При этой температуре функция распределения молекул по энергиям  $w(E)$  резко уширяется, и энергия *начинает* (или *кончает* — смотря откуда смотреть) резко изменяться. Обратите внимание, что температура фазового перехода второго рода  $T^*$  совпадает именно с *началом* (а не с серединой) резкого изменения энергии

На рисунках 8-4 – 8-6 я специально представил случаи, когда энергия системы сначала меняется медленно с изменением температуры, потом — резко, а потом — снова медленно. Как вы видите, такое «S-образное» поведение  $E(T)$  совместимо и с фазовым переходом первого рода, и с градуальным изменением системы при отсутствии какого-либо фазового перехода вообще, и с градуальным изменением, начало которому кладет фазовый переход второго рода.

Я прошу обратить на это особое внимание, так как среди не-физиков существуют два непонятно откуда взявшихся предубеждения: что «переход» соответствует именно середине всякого S-образного изменения и что, если этот «переход» не первого рода — значит второго. Это не так!

Обратите внимание: хотя кривые  $E(T)$  на рис. 8-4 – 8-6 выглядят, в общем, одинаково, — поведение кривых распределения по энергиям,  $w(E)$ , при переходе первого рода (т. е. типа «все-или-ничего») выглядит *совсем не так*, как при всех остальных переходах. Для переходов типа «все-или-ничего» — и только для них — характерно наличие двух пиков у  $w(E)$ , т. е. *существование* двух фаз. Поэтому для выяснения того, как — прыжком или постепенно — происходит переход между двумя крайними состояниями, — *недостаточно* увидеть, что энергия или какой-то наблюдаемый параметр резко меняется в узком диапазоне температур. Здесь нужны дополнительные измерения, которые мы рассмотрим в одной из дальнейших лекций.

■ В заключение поговорим о *кинестике* конформационных превращений. Точнее, о том, почему некоторые превращения идут очень медленно. Что здесь значит «медленно»? Предположим, вы знаете время элементарного шага по ходу процесса. Например: один остаток включается во вторичную структуру за 1 наносекунду (нс). Вы знаете также, что в цепи 1000 звеньев. А процесс идет не за 1000 нс, а за 1000 с. Это и есть «медленно»: на порядки медленнее, чем можно ожидать, зная время элементарного шага и число таких шагов. Надо понять — почему...

«Медленность» процессов порой связана с медленностью диффузии при высокой вязкости. Однако часто — подчеркиваю, не всегда, но часто — медленность процесса связана с преодолением высокого *свободно-энергетического барьера*. Особенно это типично для переходов типа «все-или-ничего», где свободно-энергетический барьер разделяет две фазы (см. рис. 8-5); здесь он всегда значителен. Такой барьер очень похож на активационный барьер в химических реакциях, только в данном случае он имеет и энергетическую, и энтропийную составляющие. Напомню, как оценивается скорость такой «барьерной» реакции согласно классической теории переходных состояний.

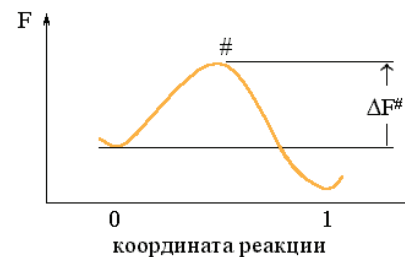


Рис. 8-7. Преодоление свободно-энергетического («активационного») барьера # при переходе из состояния 0 в состояние 1.  $\Delta F^\#$  — свободная энергия барьера (переходного состояния)

Рассмотрим сперва простейший процесс, по ходу которого система переходит из состояния 0 в состояние 1. Пусть на пути процесса 0→1 есть один «барьер» # (рис. 8-7), и ни на этом пути, ни «сбоку» от него нет никаких «ловушек» (рис. 8-8: состояний X, более стабильных, чем исходное 0, но менее стабильных, чем конечное 1; наличие ловушек усложняет процесс, но не меняет его принципиально).

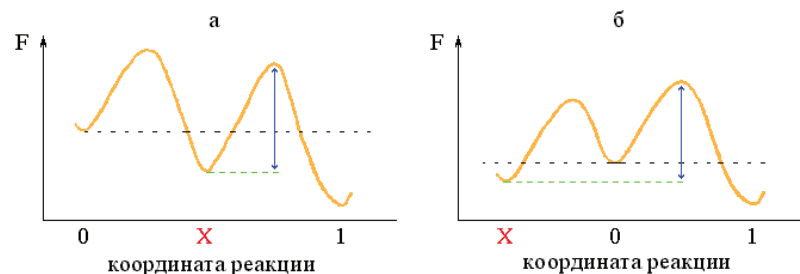


Рис. 8-8. «Ловушка»: ею может быть как интермедиат «X» на пути 0→1 (а), так и побочное (для реакции 0→1) состояние «X» (б). Для наличия кинетической «ловушки» важно только, что она более стабильна, чем исходное состояние «0», но менее стабильна, чем конечное состояние «1», и что на пути из «X» в «1» находится более высокий свободно-энергетический барьер (он отмечен стрелкой), чем на пути из «0» в «X»

Тогда скорость процесса определяется (1) населенностью барьерного («*переходноо*») состояния и (2) скоростью перехода из барьерного в «забарьерное» состояние.

Если  $\Delta F^\#$  — свободная энергия барьера, отсчитанная от свободной энергии начального состояния, причем  $\Delta F^\# \gg k_B T$ , и «ловушек» (см. рис. 8-8) нет, а в начальном состоянии находится  $n$  «частиц» (синонимы: молекул, систем и т. д.), — то на барьере, в каждый момент времени, находится (в результате флуктуаций)  $n^\# \sim n \exp(-\Delta F^\# / k_B T)$  частиц. Пусть каждая из этих на-барьерных частиц переходит за барьер за время  $\tau$  («время

элементарного шага реакции». Тогда за время порядка  $\tau$  за барьер перейдут все  $n^\#$  на-барьерных частиц. На то, чтобы за барьер перевалили бы все  $n$  частиц, понадобится  $n/n^\#$  шагов, т. е. время порядка

$$t_{0 \rightarrow 1} \sim \tau(n/n^\#) \sim \tau \exp(+\Delta F^\# / k_B T). \quad (8.12)$$

Обратная величина,  $k_{0 \rightarrow 1} \equiv 1/t_{0 \rightarrow 1}$ , называется скоростью перехода из 0 в 1.

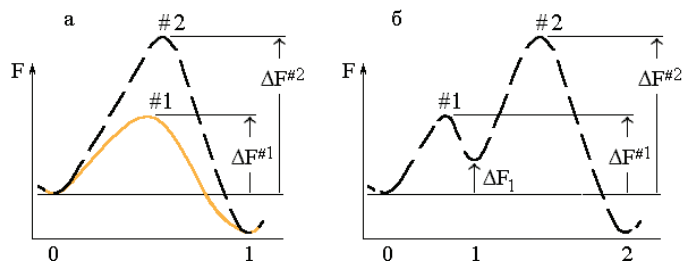
При наличии «ловушки» X таким же образом оценивается время перехода из 0 в X и из X в 1. При этом суммарное время перехода из 0 в 1 складывается из времен переходов  $0 \rightarrow X$  и  $X \rightarrow 1$ .

Добавления.

1) Если переход может осуществляться по нескольким *параллельным* путям (рис. 8-9а), то *скорости* параллельных переходов *складываются*:

$$k_{1+2+\dots} = k'_{0 \rightarrow 1} + k''_{0 \rightarrow 2} + \dots \quad (8.13)$$

Здесь  $k'_{0 \rightarrow 1} \equiv 1/t'_{0 \rightarrow 1} = (1/\tau) \exp(-\Delta F^\# / k_B T)$  — скорость перехода по первому из параллельных путей,  $k''_{0 \rightarrow 2}$  — по второму, и т. д. Так что, если параллельных путей не слишком много, то время процесса диктуется самым быстрым из них (т. е. идущим через **самый низкий** барьер).

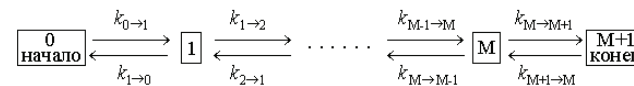


**Рис. 8-9.** Свободно-энергетические барьеры при параллельных (а) и последовательных (б) процессах

2) Если при переходе приходится преодолевать несколько барьеров *последовательно*, то *складываются времена* преодоления этих индивидуальных барьеров. Это утверждение очевидно для изображенного на рис. 8-8а процесса, в котором сначала накапливается стабильный интермедиат X, а затем из него образуется конечное состояние 1. Однако оно далеко не столь очевидно для изображенного на рис. 8-9б процесса, в котором все интермедиаты *нестабильны*. Более того, необходимо уточнить, что под «временем преодоления индивидуального барьера» в этом случае следует понимать время, необходимое для подъема на него *из самого глубокого* из предшествующих, а *не* из непосредственно предшествующего свободно-

энергетического минимума (так, в изображенной на рис. 8-9б ситуации высоту барьера #2 надо отсчитывать от состояния 0, а не от состояния 1).

Для доказательства (я дам только идею и не буду мучить вас выкладками) рассмотрим процесс



(где  $k_{i \rightarrow i+1}$  — скорость перехода из  $i$  в  $i+1$ , а  $k_{i+1 \rightarrow i}$  — скорость обратного перехода из  $i+1$  в  $i$ ), в котором свободные энергии всех промежуточных состояний (1, 2, ..., M) на много  $k_B T$  выше, чем свободные энергии и начального состояния 0, и конечного M+1.

Из-за высокой свободной энергии всех промежуточных состояний в них находится очень мало молекул (по сравнению с их числом в начальном и конечном состояниях в сумме). Поэтому и скорость изменения числа молекул в каждом из промежуточных состояний очень мала по сравнению со скоростью изменения числа молекул в начальном и конечном состояниях. Иными словами, можно положить, что скорость потока постоянна по всему пути реакции:

$$-dn_0/dt \equiv k_{0 \rightarrow 1} n_0 - k_{1 \rightarrow 0} n_1 = k_{1 \rightarrow 2} n_1 - k_{2 \rightarrow 1} n_2 = \dots = k_{M \rightarrow M+1} n_M - k_{M+1 \rightarrow M} n_{M+1} \equiv dn_{M+1}/dt. \quad (8.14)$$

(Такое допущение называется *квазистационарным приближением*; оно широко используется в химической кинетике). Решение этих уравнений приводит к формуле

$$t_{0 \rightarrow \dots \rightarrow M+1} = 1/k_{0 \rightarrow 1} + (1/k_{1 \rightarrow 2}) \exp(\Delta F_1 / k_B T) + \dots + (1/k_{M \rightarrow M+1}) \exp(\Delta F_M / k_B T), \quad (8.15)$$

где  $1/k_{i \rightarrow i+1}$  — время преодоления барьера  $i$  при старте с непосредственно предшествующего состояния  $i$ , а  $\Delta F_i$  — свободная энергия промежуточного состояния  $i$ , отсчитанная от свободной энергии начального состояния. Да, еще надо сказать, что при выводе этой формулы я воспользовался известным соотношением

$$k_{i \rightarrow j} / k_{j \rightarrow i} = \exp[(F_i - F_j) / k_B T]; \quad (8.16)$$

оно следует из того, что равновесные населенности этих состояний,  $n_i^0$  и  $n_j^0$ , должны удовлетворять как кинетическому уравнению  $n_i^0 k_{i \rightarrow j} = n_j^0 k_{j \rightarrow i}$  (отвечающему отсутствию, в равновесии, перетока частиц из одного состояния в другое), так и термодинамическому соотношению  $n_i^0 / n_j^0 = \exp[-(F_i - F_j) / k_B T]$  (связывающему равновесную вероятность нахождения частиц в каждом из состояний с его свободной энергией).

И последнее усилие. Так как, в силу формулы (8.12),  $1/k_{i \rightarrow i} = \tau \exp[(\Delta F^\# - \Delta F_{i-1}) / k_B T]$ , то  $(1/k_{i-1 \rightarrow i}) \exp(\Delta F_{i-1} / k_B T) = \tau \exp(\Delta F^\# / k_B T)$ . Это — как раз время перехода через индивидуальный барьер #  $i$ , если бы он был единственным на пути процесса  $0 \rightarrow M+1$ ; назовем его  $t_{0 \rightarrow \# i \rightarrow}$ . Тогда

$$t_{0 \rightarrow \dots \rightarrow M+1} = t_{0 \rightarrow \# 1 \rightarrow} + t_{0 \rightarrow \# 2 \rightarrow} + \dots + t_{0 \rightarrow \# M+1 \rightarrow}, \quad (8.17)$$

что и требовалось доказать: время протекания *последовательной* реакции *складывается из времен* преодоления индивидуальных барьеров, высоты которых отсчитываются от *самого глубокого* из предшествующих свободно-энергетических минимумов.



Кстати, формула (8.17) показывает, что, если барьеров не слишком много, то время процесса диктуется просто **самым высоким** из них.

■ И самое последнее. Как я уже говорил, мы начинаем подозревать существование свободно-энергетического барьера только тогда, когда видим, что процесс идет много медленнее, чем диффузия. А за какое время происходит диффузия? Чтобы составить себе хотя бы качественное впечатление, поговорим немного о диффузионных процессах.

Но сначала полезно понять, за какое время молекула, из-за трения о среду, забывает, куда она шла, и начинает «диффундировать», т. е. понять, за какое время рассеется, перейдет из-за трения в тепло ее кинетическая энергия.

Можно показать, что это произойдет за пикосекунды.

В самом деле, движение молекулы по инерции в вязкой среде описывается простым уравнением

$$m(dv/dt) = F_{\text{frict}}, \quad (8.18)$$

где  $dv/dt$  — ускорение движения частицы,  $m$  — ее масса,  $F_{\text{frict}}$  — сила ее трения об окружающую среду. Массу можно оценить как  $m = \rho V$ , где  $\rho$  — плотность частицы, а  $V$  — ее объем. Разумно принять, что все линейные размеры частицы примерно равны  $D$ , так что  $V \approx D^3$ . Силу трения  $F_{\text{frict}}$  можно оценить по закону Стокса:  $F_{\text{frict}} = -3\pi D\eta v$ , где  $v$  — скорость движения частицы, а  $\eta$  — вязкость среды. Уравнение

$$m(dv/dt) = -3\pi D\eta v \quad (8.18a)$$

задает характерное время

$$t_{\text{kinet}} \approx m/(3\pi D\eta), \quad (8.19)$$

время затухания движения под воздействием трения о среду. Его легко оценить численно. Так как  $m \approx \rho D^3$ , а  $3\pi \approx 10$ , то  $t_{\text{kinet}} \approx 0,1\rho D^2/\eta$ . И, поскольку плотность  $\rho \approx 1$  г/см<sup>3</sup> у всех молекул, с которыми мы имеем дело, а вязкость воды  $\eta \approx 0,01$  г/(см с) [посмотрите в справочнике], то

$$t_{\text{kinet}} \approx 10^{-13} \text{с} (D/\text{нм})^2, \quad (8.19a)$$

где  $(D/\text{нм})$  — диаметр молекулы, измеренный в нанометрах.

Значит, кинетическая энергия малой молекулы ( $D \approx 0,3$  нм) рассеется за  $\sim 10^{-14}$  с, небольшого белка ( $D \approx 3$  нм) — за  $\sim 10^{-12}$  с, а большого ( $D \approx 10$  нм) — за  $\sim 10^{-11}$  с. Итак: типичное время — пикосекунда. Еще быстрее рассеется кинетическая энергия при трении же о более вязкое, чем вода, окружение (например, о мембрану).

Конечно, без энергии наша частица не останется: молекулы среды, налетая со всех сторон на нашу молекулу, вернут ей ее энергию. Но вот направление своего первоначального движения она полностью забудет — за пикосекунду...

*Примечание.* Вышеприведенная  $t_{\text{kinet}}$  относится к водному окружению. Вязкость клеточного окружения молекул белков, по имеющимся данным, примерно на порядок выше. Поэтому в таком окружении  $t_{\text{kinet}}$  примерно на порядок ниже (см. (8.19)), чем в чистой воде.

■ Теперь составим себе представление о диффузионном смещении молекулы.

Поддерживаемая теплом кинетическая энергия любой молекулы,  $mv^2/2$ , составляет в среднем, как мы знаем, величину порядка  $k_B T$ . Молекула помнит на-

правление своего движения в течение времени  $t_{\text{kinet}}$ . За это время, время одного шага, она пройдет расстояние  $\Delta l \sim vt_{\text{kinet}}$ . Потом направление ее движения поменяется, и она сдвинется примерно на то же расстояние  $\sim \Delta l$  в непредсказуемую сторону. При этом средний квадрат расстояния молекулы от исходной точки растет пропорционально времени. Действительно, если за  $n$  шагов молекула сдвинулась на расстояние  $l_n$  в какую-то сторону, то за  $n+1$  шаг она сдвинется на расстояние  $l_{n+1} = l_n \pm \Delta l$ . При этом  $(l_{n+1})^2 = (l_n \pm \Delta l)^2 = (l_n)^2 \pm 2l_n \Delta l + \Delta l^2$ . То есть, в среднем,  $(l_{n+1})^2 = (l_n)^2 + \Delta l^2$ , так как средняя величина члена  $\pm 2l_n \Delta l$  равна нулю. А так как за время  $t$  молекула делает  $t/t_{\text{kinet}}$  шагов, то средний квадрат ее смещения за это время составляет

$$l_t^2 = (t/t_{\text{kinet}}) \Delta l^2.$$

Так как  $\Delta l \sim vt_{\text{kinet}}$ , то

$$l_t^2 \sim (t/t_{\text{kinet}}) (vt_{\text{kinet}})^2 = t(v^2 t_{\text{kinet}}).$$

А так как  $t_{\text{kinet}} \approx m/(3\pi D\eta)$ , и  $mv^2/2 \approx k_B T$ , то

$$l_t^2 \sim t [k_B T / 1,5\pi D\eta].$$

Полученный нами ответ — приблизительный: мы ведь много раз пользовались знаками « $\approx$ » («приблизительно равно») и « $\sim$ » («равно по порядку величины»). Однако, как всегда бывает в таких случаях, мы получили ответ, близкий к точному (который требует гораздо больше выкладок). Точный ответ — таков:

$$l_t^2 = t [2k_B T / \pi D\eta] \equiv t \cdot [6D_{\text{diff}}]. \quad (8.20)$$

Здесь  $D_{\text{diff}} = k_B T / (3\pi D\eta)$  — вычисленная Эйнштейном величина коэффициента диффузии шара диаметром  $D$  в среде с вязкостью  $\eta$  при температуре  $T$ .

Характерное время диффузии — это время, за которое молекула сдвинется на расстояние порядка собственного размера  $D$ . Зная формулу (8.20), оценить это время очень легко:

$$t_{\text{diff}} = \pi D^3 \eta / 2k_B T. \quad (8.21)$$

И, поскольку вязкость воды  $\eta \approx 0,01$  г/(см с), а  $k_B T \approx 600$  кал/моль при комнатной температуре, то

$$t_{\text{diff}} \approx 0,4 \cdot 10^{-9} \text{с} (D/\text{нм})^3, \quad (8.22)$$

где  $D/\text{нм}$  — диаметр молекулы, измеренный в нанометрах.

Можно показать, что за то же примерно время происходит переворот молекулы (это соответствует смещению полюса молекулы на расстояние  $\approx D$ ).

*Примечание.* Вышеприведенная оценка времени диффузии относится к водному окружению. Вязкость клеточного окружения молекул белков, по имеющимся данным, примерно на порядок выше. Поэтому в таком окружении все характерные диффузионные времена примерно на порядок выше (см. (8.21)), чем в чистой воде.

Итак, мы можем заключить, что характерное время диффузии молекулы в воде — наносекунды: это — время, за которое молекула, испытав

массу столкновений с молекулами среды, сдвинется на расстояние порядка собственного размера и перевернется.

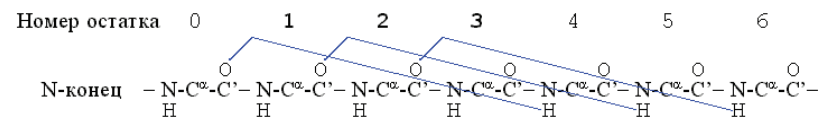
Процессы, происходящие много более медленно, заставляют предполагать существование на их пути свободно-энергетического барьера. В случае межмолекулярных реакций (или реакций между отдаленными кусками цепи) такой барьер может создаваться потерей энтропии при сближении частиц, вступающих в реакцию.

## Лекция 9

Свободная энергия инициации и элонгации  $\alpha$ -спирали. Теорема Ландау и не-фазовость перехода спираль-клубок. Размер кооперативного участка при переходе спираль-клубок. Стабильность  $\alpha$ -спирали в воде. Стабильность  $\beta$ -структуры в воде. Скорость образования  $\beta$ -структуры и  $\alpha$ -спирали. Что такое «клубок»?

Обсудим теперь, зная базовую физику, стабильность вторичной структуры и кинетику ее образования. Сегодня я буду говорить только о *гомо*-полипептидах — о полипептидах, состоящих из одинаковых аминокислотных остатков.

- Начнем с  $\alpha$ -спирали. Первая водородная связь в ней  $(CO)_0 - (HN)_4$  фиксирует конформации трех остатков — 1, 2, 3; следующая водородная связь  $(CO)_1 - (HN)_5$  дополнительно фиксирует конформацию только одного остатка — остатка 4; связь  $(CO)_2 - (HN)_6$  дополнительно фиксирует остаток 5, и т. д.



Значит, если в спирали фиксировано  $n$  остатков, то их фиксирует  $n-2$  водородные связи. Рассмотрим свободную энергию образования такой спирали из клубка в водном окружении («клубок» — это полимер без фиксированной структуры и без взаимодействия дальних по цепи звеньев). Эту свободную энергию можно записать как

$$\Delta F_{\alpha} = F_{\alpha} - F_{\text{клуб.}} = (n-2)f_{\text{H}} - nTS_{\alpha} = -2f_{\text{H}} + n(f_{\text{H}} - TS_{\alpha}). \quad (9.1)$$

Здесь  $f_{\text{H}}$  — свободная энергия образования водородной связи и сопутствующих ей взаимодействий в  $\alpha$ -спирали (вы помните, что  $f_{\text{H}}$  — не просто

энергия, как было бы в вакууме: в нее входит и энергия, и энтропия перестройки водородных связей в водном окружении), а  $S_\alpha$  — потеря энтропии при фиксации одного остатка в спирали.

Вы видите, что в  $\Delta F_\alpha$  есть два члена. Один ( $-2f_H$ ) не зависит от длины спирали; величина

$$f_{\text{init}} = -2f_H \quad (9.2)$$

традиционно называется свободной энергией инициации спирали (на самом деле,  $f_{\text{init}}$  — суммарная свободная энергия обеих границ спирали с клубком; она учитывает и инициацию, и терминацию спирали). Другой член,  $n(f_H - TS_\alpha)$ , прямо пропорционален длине спирали; величина

$$f_{\text{el}} = (f_H - TS_\alpha) \quad (9.3)$$

называется свободной энергией элонгации спирали на один остаток. В общем виде,

$$\Delta F_\alpha = f_{\text{init}} + nf_{\text{el}}. \quad (9.4)$$

При этом отношение вероятности чисто спирального состояния цепи из  $n$  остатков к ее же чисто клубковому (и начисто лишенному спиральных примесей) состоянию, равно

$$\exp(-\Delta F_\alpha/kT) = \exp(-f_{\text{init}}/kT) [\exp(-f_{\text{el}}/kT)]^n = \sigma s^n. \quad (9.5)$$

Здесь я использовал общепринятые обозначения:

- фактор элонгации спирали:  $s = \exp(-f_{\text{el}}/kT)$ ;
- фактор инициации спирали:  $\sigma = \exp(-f_{\text{init}}/kT)$ .

Ясно, что  $\sigma \ll 1$ , так как  $\sigma = \exp(-f_{\text{init}}/kT) = \exp(+2f_H/kT)$ ; а свободная энергия водородной связи — большая отрицательная величина, порядка нескольких  $kT$ .

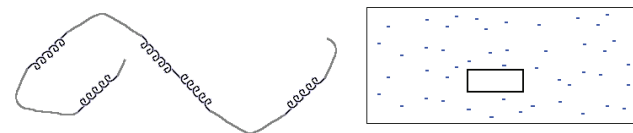
Величина  $\exp(-\Delta F_\alpha/kT) = \sigma s^n$  — это просто **константа равновесия** между двумя состояниями (« $\alpha$ » и «клубок») для куска цепи, состоящего из  $n$  звеньев.

- Прежде чем говорить о том, как экспериментально определяются величины  $\sigma$  и  $s$ , — выясним общий вопрос о том, как образуется спираль при изменении условий среды (температуры, растворителя и т. д.) — переходом все-или-ничего или постепенно?

Поначалу кажется, что такая отличная от клубка структура, как спираль, должна «вымораживаться» из него путем фазового (т. е. все-или-ничего) перехода — как лед из воды...

Однако на этот счет есть теорема Ландау, которая гласит, что в системе, где обе фазы *одномерны*, — фазовый переход первого рода невозможен. Попытаюсь объяснить эту теорему.

Прежде всего — что значит «одномерность». Это значит, что свободная энергия, и прежде всего размер границы, «стыка» фаз не зависит от размера кусков этих фаз. В этом смысле и спираль, и клубок в полимере — одномерны. Рисунок 9-1 поясняет, что граница (стык) спирали и клубка не зависит ни от длины спирали, ни от длины клубка, в то время как поверхность, граница трехмерной фазы (например, льдинки в воде) зависит от ее размера. Соответственно, свободная энергия границы спирали и клубка не зависит от их длин, а свободная энергия поверхности трехмерной фазы растет как  $n^{2/3}$  с ростом числа  $n$  вовлеченных в нее частиц.



**Рис. 9-1.** Сравнение одномерной (клубок со спиралями) и трехмерной (льдинка в воде) систем. Размер стыка спирали и клубка не зависит от их длин; поверхность трехмерной льдинки зависит от ее размера

- Теперь — что значит «происходит как фазовый переход первого рода». Это значит, что при температуре перехода стабильной может быть либо одна, либо другая фаза, но перемешивание фаз (например, льда и воды) ведет к повышению свободной энергии, и потому нестабильно. Здесь вас не должна вводить в иску картину плавающего в воде льда: при любой температуре такое состояние неустойчиво (из-за повышенной свободной энергии на границе воды со льдом), и со временем лед или растает, или охватит всю воду — конечно, если нет тока подземного тепла, течений и других неравновесных вещей...

Выгодно ли сосуществование фаз в трехмерной системе? *Нет*. Почему? Обратимся опять к рис. 9-1. Рассмотрим температуру, где безграничная вода и безграничный лед имеют равную свободную энергию (это и есть условие «точки перехода»). Если в воде плавает льдина из  $n$  молекул, свободная энергия границы пропорциональна  $\xi n^{2/3}$ , где  $n^{2/3}$  — характерное число пограничных молекул, а  $\xi > 0$  — свободная энергия границы в расчете на одну из них. (Замечание. Если  $\xi < 0$ , «перемешивание» — оно при этом всегда термодинамически выгодно — произойдет на молекулярных масштабах, и двух фаз не будет вообще). Значит, поверхность льдинки повышает свободную энергию на  $\xi n^{2/3}$ . Правда, льдинка имеет еще позиционную энтропию, так как может находиться в разных точках сосуда. Но эта энтропия не превосходит величину порядка  $k \ln(N)$ , если в сосуде — всего  $N$  молекул (т. е.  $N$  точек, с которых может начинаться

льдинка). Итого — свободная энергия льдинки — порядка  $[\xi n^{2/3} - kT \ln(N)]$ . Но логарифм — очень «слаборастущая» при больших  $N$  функция. Если льдинка занимает заметную часть сосуда (скажем,  $n \sim N/10$ ), и  $N$  очень велико [скажем, 10 000 000 000], то  $\ln(N)$  [в данном случае — 23] очень мал по сравнению с  $(N/10)^{2/3}$  [в данном случае — с 1 000 000], т.е. в свободной энергии льдинки доминирует граничный член  $\xi n^{2/3}$ , а он противится ее образованию... Поэтому в трехмерной системе макроскопические фазы разделяются (нестабильные «льдинки» из нескольких молекул не в счет: это просто микроскопические, локальные флуктуации), и фазовый переход первого рода в ней возможен.

А выгодно ли сосуществование фаз в одномерной системе? *Оказывается, да.* Рассмотрим опять температуру «середины перехода», где спираль и клубок имеют равную свободную энергию, т.е. где  $f_{el}=0$ . Свободная энергия обеих границ спирали и клубка,  $f_{init}$ , не зависит ни от размера спирали, ни от размера клубка. Позиционная энтропия спирали длины  $n$  в цепи длины  $N$  равна  $k \ln(N-n)$ . Итого, свободная энергия этой плавающей спирали есть  $f_{init} - kT \ln(N-n)$ . При больших  $N$  член  $k \ln(N-n)$  *всегда* доминирует в этом выражении, даже если  $n \sim 0,9 N$ ; а этот логарифмический член понижает свободную энергию и способствует внедрению спирали в клубок (и, точно так же, — клубка в спираль)... Поэтому в одномерной системе фазы *не* разделяются, они стремятся перемешаться, а раз так, то и фазовый переход первого рода (или типа все-или-ничего) *невозможен* при достаточно большой длине цепи. Теорема Ландау доказана.

Примечание. В отличие от плавления  $\alpha$ -спирали, плавление двойной спирали ДНК не подпадает, строго говоря, под действие теоремы Ландау, так как ДНК — одномерная система: при выплавлении куска двойной спирали образуются замкнутые на концах пространственные петли, и замыкание петли дает дополнительный, логарифмически растущий с длиной петли вклад в свободную энергию границы петли и двойной спирали.

Теперь можно поставить вопрос: при каких характерных длинах цепи начинается смешение клубковой и спиральной фаз? Или, иначе: какова будет характерная длина  $n_0$  спирального участка в середине перехода спираль-клубок?

Рассмотрим цепь из  $N$  звеньев при температуре «середины перехода», где спираль и клубок имеют равную свободную энергию, т.е.  $f_{el}=0$ . При этом свободная энергия элонгации спирали (а равно и клубка) — ноль, ее инициации —  $f_{init}$ , а число возможных положений спирали в цепи из  $N$  звеньев — порядка  $N^2/2$  (она может начинаться и кончаться в любом месте при единственном условии, что ее длина — не менее трех остатков); и ни расположение спирали в цепи, ни ее длина (при  $f_{el}=0$ ) не влияют на

ее свободную энергию. Для получения качественной оценки пренебрежем мелочами (цифрами) по сравнению с главным (буквами). Итак: размещений спирали — порядка  $N^2$ , т.е. их энтропия —  $k \cdot 2 \ln(N)$ , а полная свободная энергия внедрения куска новой фазы (спирали с флуктуирующими концами) в цепь длины  $N$  — примерно  $f_{init} - 2kT \ln(N)$ . Если она, эта свободная энергия, больше нуля — новая фаза не внедрится; если она меньше нуля — новая фаза может внедриться, и даже неоднократно. Значит, смешение клубковой и спиральной фаз начинается в кусках длины  $N \sim n_0$ , а величина  $n_0$  получается из уравнения  $f_{init} - 2kT \ln(n_0) = 0$ . Итак: характерная длина кусков спирали и клубка в середине перехода

$$n_0 = \exp(+f_{init}/2kT) = \sigma^{-1/2}. \quad (9.6)$$

Точка (температура) середины перехода на опыте находится как та точка, где спиральность очень длинного полипептида составляет 50% (спиральность полипептида обычно измеряют при помощи КД спектров, как о том говорилось выше; при 50% спиральности спектр КД полипептида выглядит как полусумма спектров клубкообразного полипептида и полипептида, спирального на 100%). В этой точке  $f_{el}=0$ , т.е.  $s = \exp(-f_{el}/kT) = 1$ .

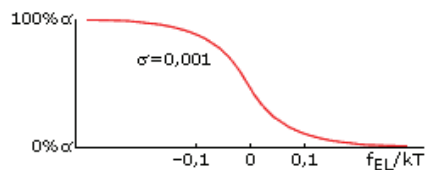
Измеряя в тех же условиях (когда  $\sigma=1$ ) зависимость спиральности от длины полипептида, можно найти  $n_0$  как ту длину цепи, при которой ее средняя степень спиральности в 4 раза меньше, чем у очень длинной цепи (т.е. чем 50%). Доказательство этой численной оценки выходит за рамки нашего курса. Если хотите, попробуйте получить ее самостоятельно.

Наконец, зная  $n_0$ , можно вычислить  $f_{init}$  и  $\sigma$ . Для большинства аминокислот  $n_0 \approx 30$ ,  $f_{init} \approx 4$  ккал/моль, и  $\sigma \approx 0,001$ .

Теперь мы можем найти свободную энергию образования водородной связи (вкуче со всеми сопутствующими образованию Н-связи в  $\alpha$ -спирали взаимодействиями): согласно (9.2),  $f_H = -f_{init}/2 \approx -2$  ккал/моль. Можно найти и конформационную энтропию, теряемую при фиксации одного звена в  $\alpha$ -спирали: согласно формуле (9.3), при  $f_{el}=0$ ,  $TS_\alpha = f_H \approx -2$  ккал/моль.

Оба параметра стабильности спирали — и  $f_{el}$ , и  $f_{init}$  — зависят от температуры, но по-настоящему сильно влияет на стабильность спирали именно отклонение величины  $f_{el}$  от 0. Дело в том, что в спирали, состоящей из  $\sim n_0$  звеньев, это отклонение умножается на большое число  $n_0$  и уже в таком виде входит в свободную энергию спирали. Когда величина  $f_{el} n_0 / kT$  составляет порядка +1 (более точная оценка:  $f_{el} n_0 / kT = +2$ ), спиральность практически исчезает, а когда  $f_{el} n_0 / kT = -2$  — практически исчезает клубок. Значит, переход спираль-клубок в очень длинных ( $N \gg n_0$ ) полипептидных цепях происходит в области  $-2/n_0 < f_{el}/kT < 2/n_0$ , т.е. (при  $n_0 \approx 30$ )  $-0,07 < f_{el}/kT < 0,07$  (рис. 9-2). Это — пример резкого, кооперативного, но *не* фазового (так как ширина его не стремится к нулю по мере роста длины цепи) перехода.





**Рис. 9-2.** Переход спираль-клубок всегда — даже для бесконечно длинных цепей — имеет конечную ширину. Это — пример *не фазового*, но кооперативного перехода: при малой величине фактора инициации спирали ( $\sigma \ll 1$ ) он совершается при совсем небольшом (много меньше, чем на  $kT$ ) изменении величины  $f_{el}$ , что показывает, что «единица перехода» кооперирует много звеньев цепи, но отнюдь не всю эту цепь

- Стабильность  $\alpha$ -спирали обычно падает с ростом температуры и с добавлением полярных денатурантов и растет с добавлением слабополярных растворителей.

Для измерения влияния отдельных аминокислотных остатков на стабильность спиралей сейчас чаще всего используются короткие (длиной  $\sim n_0$  или менее) полипептиды. В них может образоваться только одна спираль, и оценить влияние каждой аминокислотной замены на спиральность здесь наиболее просто. Сейчас известно, что вклад аминокислотного остатка в стабильность спирали лежит в следующих пределах: аланин, самый «спиралеобразующий» остаток:  $s \approx 2$ , т. е.  $f_{el} \approx -0,4$  ккал/моль; глицин, самый «спиралеразрушающий» остаток:  $f_{el} \approx +1$  ккал/моль, т. е.  $s \approx 0,2$ ; правда, у пролина — *аминокислоты* (не имеющей NH-группы, которая должна завязывать структурообразующую водородную связь в  $\alpha$ -спирали) — у него величина  $\sigma$  еще много меньше, порядка 0,01–0,001, но точно она еще не измерена.

Ранее аналогичные оценки делались (в том числе и одним из нас — О.Б.П.) с использованием статистических сополимеров (например, цепи, включающие случайным образом перемешанные 80% Glu и 20% Ala), и именно так были получены первые — а значит, и самые важные оценки, но с появлением возможности синтеза полипептидов с заданной последовательностью такие работы на случайных сополимерах, видимо, отошли в прошлое.

Используя короткие пептиды с заданной первичной структурой, можно даже оценить влияние на спиральность каждой точечной аминокислотной замены в каждом конкретном месте такого пептида — т. е., фактически, оценить вклад аминокислотного остатка в стабильность спирали в зависимости от его расположения относительно N- и C-концов спирали. Боковые группы, в частности, заряженные группы, взаимодействуют с этими концами

по-разному, потому что на N-конце спирали, как уже говорилось, сидят свободные от водородных связей NH-группы главной цепи (и частичный заряд N-конца  $\alpha$ -спирали равен  $+e/2$ ), а на другом ее конце — CO-группы (с суммарным частичным зарядом в половину электронного,  $-e/2$ ).

Кроме того, для измерения стабильности спирального состояния (величины  $f_{el}$ ) в полипептидах, включающих кислотные или основные боковые группы (например, в поли(Glu) или поли(Lys)), используется потенциометрическое титрование. Идея этого подхода в том, что, заряжая спираль, мы разрушаем ее (так как в спирали заряды боковых групп более сближены и потому сильнее расталкиваются), — так что стабильность спирали можно рассчитать, зная, как спиральность и суммарный заряд цепи зависят от pH среды. К сожалению, более подробное рассмотрение этого интересного метода выходит за рамки данного курса.

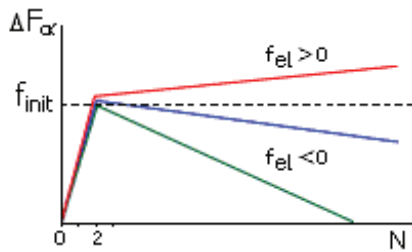
Аналогичные подходы применяются и для измерения стабильности  $\beta$ -структуры в полипептидах. Однако они менее развиты, так как  $\beta$ -структура сильно агрегирует. Сейчас стабильность  $\beta$ -структуры часто измеряют прямо в белке — путем оценки влияния на его стабильность точечных замен отдельных, лежащих на его поверхности  $\beta$ -структурных аминокислотных остатков. Диапазон влияния таких замен на стабильность  $\beta$ -структуры простирается от  $\sim -1$  до  $\sim +1$  ккал/моль. Способность разных остатков стабилизировать  $\alpha$ - и  $\beta$ -структуру мы рассмотрим на следующей лекции.

- Рассмотрим теперь **скорость образования** вторичной структуры в полипептидах.

$\alpha$ -спирали образуются быстро: показано, что за десятую долю микросекунды спираль охватывает пептид из 20–30 остатков (для столь быстрых измерений приходится греть раствор пико- или наносекундным лазерным ударом). Значит, спирали нарастают со скоростью по крайней мере порядка 1 остаток за несколько наносекунд.

Я сказал «по крайней мере», так как скорость образования спирали зависит не только от скорости, с которой она нарастает, но и от скорости, с которой в цепях появляются первые зародыши спиральной структуры. А инициация спирали требует преодоления активационного барьера — и потому образование первого витка спирали требует много больше времени, чем ее последующее удлинение еще на один виток. Так что возможно, что элонгация спирали идет много быстрее, а все наблюдаемое время тратится на ее инициацию. Разберемся в этом вопросе чуть внимательнее.

Характерная зависимость свободной энергии спирали от ее длины показана на рис. 9-3. Даже когда  $f_{el} < 0$  — т. е. когда достаточно длинная спираль стабильна — образование первого витка в ней требует преодоления активационного барьера высотой  $f_{init}$ .



**Рис. 9-3.** Характерная зависимость свободной энергии спирали  $\Delta F_\alpha$  от числа вовлеченных в нее аминокислотных остатков  $N$  при разных свободных энергиях ее элонгации  $f_{el}$ . При  $f_{el} < 0$  длинная спираль стабильна, но ее образование требует инициации — преодоления активационного барьера высотой  $f_{init}$ . При  $f_{el} > 0$  спираль любой длины нестабильна и, следовательно, не образуется

По теории переходных состояний, преодоление такого барьера и образование первого витка в данном месте цепи происходит за время

$$t_{init_\alpha} = \tau \exp(+f_{init}/kT), \quad (9.7)$$

где  $\tau$  — время элементарного шага (в данном случае — элонгации спирали на один остаток), а экспонента учитывает малую населенность барьерного состояния, имеющего (по определению «барьера») наивысшую свободную энергию. В силу уравнения (9.6)

$$t_{init_\alpha} = \tau/\sigma. \quad (9.8)$$

Однако инициация может произойти в любой точке будущей спирали, а ее длина (даже в очень длинных цепях) — порядка  $n_0 = \sigma^{-1/2}$ . Значит, характерное время инициации первого витка в *каком-то месте* будущей спирали составляет в  $n_0$  раз меньше времени, а  $t_{init_\alpha}/n_0 = \tau/(n_0\sigma) = \tau/\sigma^{1/2}$ .

На то, чтобы уже инициированная спираль охватила все свои  $\sim n_0$  звеньев, требуется потом примерно то же (что и на инициацию) время,  $\sim \tau n_0 = \tau/\sigma^{1/2}$ . Значит, все время перехода клубок-спираль в полипептиде должно составлять примерно  $2\tau/\sigma^{1/2}$ , а характерное время полуперехода —

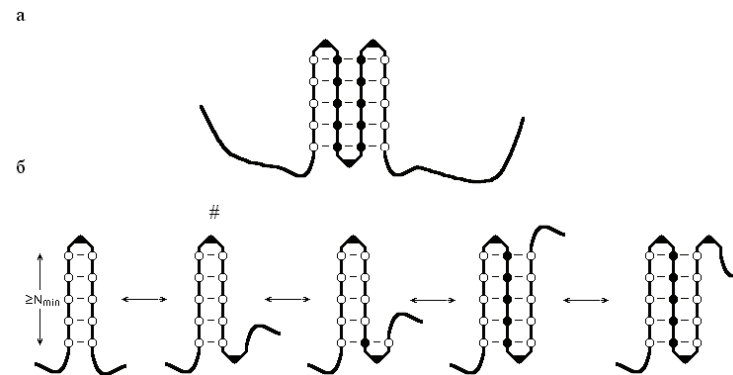
$$t_\alpha \sim \tau/\sigma^{1/2} = \tau \exp(+f_{init}/2kT). \quad (9.9)$$

Половина, грубо говоря, времени идет на инициацию спиралей в разных местах цепи, а половина — на их элонгацию, протекающую, в согласии с предварительной оценкой, со скоростью порядка 1 остаток за наносекунду.

■ Кинетика образования  $\alpha$ -спиралей относительно проста: они всегда образуются быстро. Кинетика образования  $\beta$ -структуры много сложнее и интереснее.

$\beta$ -структура в полипептидах часто образуется очень медленно — за часы, а бывает — за недели (хотя бывает, что и за миллисекунду). Почему? И, в то же время, белки с  $\beta$ -структурой сворачиваются не медленнее (во всяком случае, не намного медленнее)  $\alpha$ -спиральных белков. Как они успевают? И что отвечает за аномалии в образовании  $\beta$ -структуры в полипептидах — медленная инициация или медленная элонгация?

■ «Аномальная» (по сравнению с переходом спираль-клубок) кинетика образования  $\beta$ -структуры связана с тем, что она — двумерный (а не одномерный, как спираль или клубок) объект (рис. 9-4). И, в результате, образование  $\beta$ -структуры может протекать как фазовый переход первого рода. Рассмотрим этот процесс подробнее.



**Рис. 9-4.** (а) Схема  $\beta$ -листа. Аминокислотные остатки внутренних  $\beta$ -участков показаны черными кружками, краевых  $\beta$ -участков — пустыми кружками; изгибы (или петли), соединяющие  $\beta$ -участки, изображены в виде уголков. (б) Описанный в тексте сценарий роста  $\beta$ -листа (для случая, когда отдельные  $\beta$ -шпильки нестабильны). Самая нестабильная структура помечена значком #

Находящиеся на краю листа звенья цепи имеют меньше контактов с другими, чем внутренние звенья листа. Иными словами, край  $\beta$ -листа (как и граница любой другой фазы, например, капли воды, льдинки или  $\alpha$ -спирали) имеет повышенную свободную энергию. Однако, в отличие от  $\alpha$ -спирали (и в сходстве с каплей) —  $\beta$ -лист не одномерен, т. е. размер его края (а значит, и его свободная энергия) растет с числом вовлеченных в этот лист звеньев цепи. Поэтому переход клубок —  $\beta$ -структура становится фазовым переходом первого рода, как образование капли или льдинки. Покажем, что из-за этого на пути образования  $\beta$ -листа возникает высокий (особенно при образовании лишь слегка стабильной  $\beta$ -структуры) свободно-энергетический барьер, который может в миллионы и миллиарды раз замедлить инициацию ее сворачивания.

Край  $\beta$ -листа сложен из (а) краевых  $\beta$ -тяжей и (б) изгибов или петель, соединяющих  $\beta$ -тяжи (рис. 9-4а). Пусть свободная энергия клубка принята за 0 (т. е. за начало отсчета),  $f_\beta$  — свободная энергия остатка в центре  $\beta$ -листа,  $f_\beta + \Delta f_\beta$  — свободная энергия остатка на его краю (т. е.  $\Delta f_\beta$  — краевой эффект), и  $U$  — свободная энергия изгиба. Раз  $\beta$ -лист вообще

образуется — то он стабилен (т.е.  $f_\beta < 0$ ), а краевые эффекты препятствуют его распаду на кусочки (т.е.  $\Delta f_\beta > 0$  и  $U > 0$ ).

В кинетике образования  $\beta$ -структуры надо различать два случая:

А)  $f_\beta + \Delta f_\beta < 0$ , т.е. длинная  $\beta$ -шпилька сама по себе стабильнее клубка. Тогда только образование изгиба в вершине этой шпильки требует преодоления активационного барьера (примерно такого же, как при образовании  $\alpha$ -спирали), а дальнейший рост  $\beta$ -структуры идет быстро — примерно как элонгация  $\alpha$ -спирали (см. линию с  $f_{el} < 0$  на рис. 9-3).

Б)  $f_\beta + \Delta f_\beta > 0$ , т.е.  $\beta$ -шпилька сама по себе нестабильна, и только прилипание к инициирующей  $\beta$ -шпильке последующих  $\beta$ -тяжей стабилизирует  $\beta$ -лист. Активационному барьеру тогда соответствует образование «зародыша»: такого  $\beta$ -листа или  $\beta$ -шпильки, что дальнейший рост листа пойдет уже с общим понижением свободной энергии.

■ Образование и последующее «прорастание» зародыша новой фазы — характернейшая особенность фазовых переходов первого рода, к коим принадлежит и образование  $\beta$ -структуры. Однако преодоление связанного с образованием зародыша активационного барьера, как мы скоро убедимся, может быть очень и очень медленным.

Рассмотрим следующий простейший сценарий образования стабильного  $\beta$ -листа в случае, когда  $\beta$ -шпильки нестабильны (рис. 9-4б): образование инициирующей  $\beta$ -шпильки из изгиба и двух  $\beta$ -тяжей по  $N$  звеньев (аминокислот) в каждом; образование последующего изгиба на ее конце; прилипание к листу  $\beta$ -тяжа из  $N$  звеньев; образование последующего изгиба; прилипание  $\beta$ -тяжа; и так далее.

Возникновение из клубка  $\beta$ -шпильки из изгиба и двух  $N$ -звенных  $\beta$ -тяжей повышает свободную энергию цепи на  $U + 2N(f_\beta + \Delta f_\beta)$ ; образование последующего  $\beta$ -изгиба повышает ее еще на  $U$ . Прилипание к  $N$ -звенному краю этой шпильки  $N$ -звенного  $\beta$ -тяжа снижает свободную энергию на  $Nf_\beta$  (ведь при этом число краевых звеньев не меняется, а число внутренних растет на  $N$ ); образование последующего  $\beta$ -изгиба опять повышает свободную энергию на  $U$ ; прилипание нового  $N$ -звенного  $\beta$ -тяжа снова понижает ее на  $Nf_\beta$ , и т.д.

Цикл «прилипание нового  $\beta$ -тяжа и образование нового  $\beta$ -изгиба» изменяет суммарную свободную энергию на  $Nf_\beta + U$ . А так как в результате такого цикла свободная энергия должна снижаться (только тогда пойдет сворачивание цепи), то прилипающий  $\beta$ -тяж (а равно и тот, к которому он прилипает) должен состоять не менее чем из  $N_{\min} = U/(-f_\beta)$  звеньев.

«Переходным», т.е. самым нестабильным состоянием по ходу образования  $\beta$ -структуры является (по нашему сценарию)  $\beta$ -шпилька с последующим изгибом. Так как стабильность шпильки падает с удлинением,

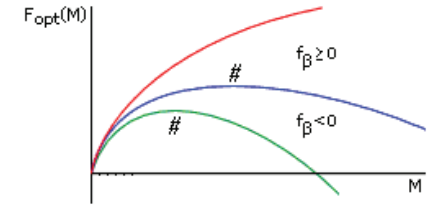
а  $\beta$ -тяж в инициирующей шпильке должен состоять как минимум из  $N_{\min}$  звеньев, то минимальная свободная энергия инициирующей  $\beta$ -шпильки (с последующим изгибом) составляет

$$F^\# = U + 2N_{\min}(f_\beta + \Delta f_\beta) + U = 2(U\Delta f_\beta)/(-f_\beta). \quad (9.10)$$

Это — свободная энергия переходного состояния при образовании  $\beta$ -листа *по нашему сценарию*. Теперь осталось только показать, что, какой сценарий ни возьми, существенно более стабильных переходных состояний быть не может.

Чтобы показать это, оценим, как будет меняться минимальная свободная энергия  $\beta$ -листа с его увеличением. Общий ход ее изменения показан на рис. 9-5. Пока лист мал, доминируют краевые эффекты, и свободная энергия повышается. В большом же листе доминируют внутренние остатки, и его свободная энергия падает.

**Рис. 9-5.** Общий ход изменения  $F_{\text{opt}}$ , минимальной свободной энергии  $\beta$ -листа, с увеличением числа  $M$  звеньев в нем. Кривые относятся к разным свободным энергиям  $f_\beta$  внутренних звеньев  $\beta$ -листа. # — точка максимума величины  $F_{\text{opt}}$  по ходу роста  $\beta$ -листа. При росте  $\beta$ -листа (в отличие от роста  $\alpha$ -спирали, см. рис. 9-3) максимум его свободной энергии не приходится на самое начало этого процесса



Теперь перейдем к расчетам. Сначала оценим минимальную свободную энергию  $M$ -звенного  $\beta$ -листа. Свободная энергия листа, состоящего из  $m$   $\beta$ -тяжей (равной длины, чтобы минимизировать краевую свободную энергию) и  $m-1$  изгиба, равна

$$F(M, m) = Mf_\beta + 2(M/m)\Delta f_\beta + (m-1)U \quad (9.11)$$

(тут я позволил себе пренебречь малым числом звеньев, ушедших в изгибы).

Варьируя число  $\beta$ -участков  $m$  (при заданном  $M$ ), найдем минимум этой свободной энергии из условия

$$dF/dm = -2(M/m^2)\Delta f_\beta + U = 0. \quad (9.12)$$

Отсюда следует оптимальное число  $\beta$ -участков в этом листе,  $m_{\text{opt}} = M^{1/2} \times [2\Delta f_\beta/U]^{1/2}$  и его свободная энергия,  $F_{\text{opt}}(M) = F(M, m_{\text{opt}}) = Mf_\beta - U + 2M^{1/2} \times [2\Delta f_\beta U]^{1/2}$ . Варьируя величину  $F_{\text{opt}}(M)$  по  $M$ , находим ее максимум (см. рис. 9-5) из условия

$$dF_{\text{opt}}/dM = f_\beta + M^{-1/2} [2\Delta f_\beta U]^{1/2} = 0. \quad (9.13)$$

Далее можно определить размер  $\beta$ -листа, соответствующего этому максимуму,  $M^* = 2(\Delta f_\beta U)/(-f_\beta)^2$ , и его свободную энергию

$$F^* = F_{\text{opt}}(M^*) = 2(\Delta f_\beta U)/(-f_\beta) - U. \quad (9.14)$$

Величины  $F^*$  и  $F^\#$  (см. формулу (9.10)) совпадают в главном члене,  $2(\Delta f_\beta U)/(-f_\beta)$ . Именно из-за него свободная энергия переходного состояния  $F^\#$  всегда велика, когда свободная энергия стабилизации  $\beta$ -структуры  $(-f_\beta)$  мала, т. е. когда  $\beta$ -структура находится на грани стабильности. В сущности,  $F^\#$  стремится к бесконечности, когда  $\beta$ -структура приближается к термодинамическому равновесию с клубком, т. е. когда  $(-f_\beta) \rightarrow 0$ .

- Выше мы рассмотрели свободную энергию инициации сворачивания  $\beta$ -листа. Согласно теории переходного состояния, время процесса (в данном случае — инициации сворачивания) зависит от свободной энергии переходного (самого нестабильного промежуточного) состояния  $F^\#$  как

$$t_{\text{init } \beta} \sim \tau_\beta \exp(+F^\#/kT), \quad (9.15)$$

где  $\tau_\beta$  — время элементарного шага (в данном случае — элонгации  $\beta$ -листа на один остаток; и нет причин, чтобы скорость элонгации  $\beta$ -листа качественно отличалась от скорости элонгации спирали).

$t_{\text{init } \beta}$  — характерное время инициации образования  $\beta$ -листа в *одном* месте цепи. Характерное время инициации  $\beta$ -листа в *каком-то* месте цепи из  $M$  остатков равно  $t_{\text{init } \beta}/M$ . Распространение этого листа на всю цепь займет время порядка  $M\tau_\beta$ . Инициация является процессом, лимитирующим скорость сворачивания в коротких цепях (где  $t_{\text{init } \beta}/M > M\tau_\beta$ ). При этом характерное время образования  $\beta$ -структуры составляет  $\sim \tau_\beta \exp(+F^\#/kT)/M$ .

Для  $\beta$ -структуры в длинных цепях (и/или для межмолекулярной  $\beta$ -структуры) характерное время образования составляет  $\sim \tau_\beta \exp(+F^\#/2kT)$ . [Эта задача сходна с уже решавшейся задачей оценки времени образования  $\alpha$ -спиралей в длинных цепях. Основная идея решения — в том, что один  $\beta$ -лист будет охватывать такое число остатков  $M_{\text{эф}}$  что время инициации этого  $\beta$ -листа примерно совпадает со временем его экспансии, т. е.  $\tau_\beta \exp(+F^\#/kT)/M_{\text{эф}} \sim M_{\text{эф}}\tau_\beta$ .]

Таким образом, с учетом того, что  $F^\# = 2(U\Delta f_\beta)/(-f_\beta)$  [см. (9.10)], в обоих случаях зависимость характерного времени образования  $\beta$ -структуры от ее стабильности  $(-f_\beta)$  в расчете на один остаток, при малых  $-f_\beta$ , имеет вид

$$t_\beta \sim \exp[A/(-f_\beta)]. \quad (9.16)$$

Каково бы ни было численное значение константы  $A$ , видно (см. рис. 9-6), что при малых  $-f_\beta$  время образования  $\beta$ -структуры будет огромно (в пределе — бесконечно велико).

Эксперименты показывают именно такую, как на рис. 9-6, резкую зависимость скорости образования  $\beta$ -структуры от ее стабильности.

Итак: малостабильная  $\beta$ -структура должна образовываться очень медленно, и причина здесь не в медленной элонгации, а в медленной инициации, — трудно собрать свободную энергию из теплового движения и преодолеть высокий барьер.

Так объясняется и экспериментально наблюдаемая, очень медленная скорость образования  $\beta$ -структуры в неагрегирующих полипептидах (где  $\beta$ -структура всегда малостабильная), и резкий рост этой скорости с повышением стабильности  $\beta$ -структуры.

В то же время стабильные  $\beta$ -листы и шпильки (а именно такие, стабильные структуры и наблюдаются в белках) должны образовываться довольно быстро — примерно как  $\alpha$ -спирали.

- Крайне медленная инициация — общее свойство фазовых переходов первого рода, когда возникающая фаза находится на грани стабильности. Вспомните: переохлажденная жидкость, переохлажденный пар... Все это связано с большой свободной энергией обширной границы раздела фаз. А  $\beta$ -структура как раз и образуется фазовым переходом первого рода со всеми вытекающими последствиями...

Наоборот,  $\alpha$ -спираль *не* образуется фазовым переходом первого рода [вспомните: граница спирали, в отличие от границы  $\beta$ -структуры (или льдинки), *не* растет с ее размером] — поэтому барьер, который нужно преодолеть при сворачивании спирали, всегда имеет конечную (и небольшую) величину, и она может успеть «проиницироваться» за долю микросекунды.

- Заключительное замечание: что такое «клубок». Этим термином я много раз пользовался, говоря о цепи, не имеющей определенной структуры. Действительно, понятие «клубок» охватывает огромное множество конформаций, не имеющих никакого дальнего (по цепи) порядка, что, однако, не исключает слабый (на масштабе всего нескольких подряд идущих звеньев цепи) ближний порядок.

Самой интересной особенностью клубка (экспериментально наблюдаемой с помощью гидродинамических методов, при помощи рассеяния

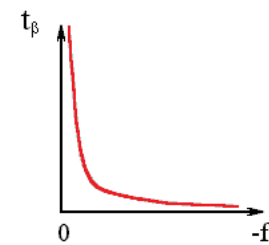


Рис. 9-6. Общий вид зависимости характерного времени образования  $\beta$ -листа от стабильности  $\beta$ -структуры



света и рентгеновских лучей) является его крайне малая плотность, большой объем и совершенно особая зависимость его размеров, радиуса и объема, от длины цепи.

Для выяснения этой особенности рассмотрим простейшую модель клубка, так называемую «свободно-сочлененную цепь» (рис. 9-7). Она состоит из звеньев-«палочек» (если цепь жесткая, такое «звено» может соответствовать нескольким мономерам полимерной цепи); главное, что каждая «палочка» может свободно, как угодно поворачиваться в сочленении относительно своих соседей по цепи. Пусть в цепи  $M$  «палочек», и длина каждой равна  $r$ .

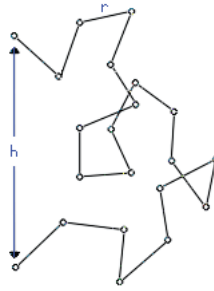


Рис. 9-7. Свободно-сочлененная цепь — модель клубка

Такую цепь можно описать как цепочку из векторов  $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_M$  (напомню: вектор — это отрезок прямой, имеющий направление). Эти векторы имеют одинаковую длину  $r$ , и каждый из них направлен от предыдущего сочленения к последующему. Сумма векторов

$$\mathbf{h} = \sum_{i=1}^M \mathbf{r}_i \quad (9.17)$$

есть просто вектор, идущий от начала к концу цепи. Квадрат длины этого вектора есть

$$h^2 = \left( \sum_{i=1}^M \mathbf{r}_i \right)^2 = \sum_{i=1}^M r_i^2 + \sum_{i \neq j}^M \sum_{j=1}^M \mathbf{r}_i \mathbf{r}_j. \quad (9.18)$$

Найдем теперь среднюю величину  $h^2$ ,  $\langle h^2 \rangle$ , т. е. усредним  $h^2$  по всем возможным тепловым флуктуациям конформации цепи (это усреднение и означает значком  $\langle \rangle$ ). При этом мы должны учесть, что  $\langle r_i^2 \rangle$ , средний квадрат вектора  $\mathbf{r}_i$ , есть просто  $r^2$ , а среднее произведение любых векторов  $\langle \mathbf{r}_i \mathbf{r}_j \rangle$  есть 0 при  $i \neq j$ , так как, свободно вращаясь, эти векторы имеют одинаковую вероятность быть направленными как в одну, так и в разные стороны.

Значит, средний квадрат расстояния между концами свободно-сочлененной цепи есть

$$\langle h^2 \rangle = \left\langle \left( \sum_{i=1}^M \mathbf{r}_i \right)^2 \right\rangle = \left\langle \sum_{i=1}^M r_i^2 + \sum_{i \neq j}^M \sum_{j=1}^M \mathbf{r}_i \mathbf{r}_j \right\rangle = \sum_{i=1}^M \langle r_i^2 \rangle + \sum_{i \neq j}^M \sum_{j=1}^M \langle \mathbf{r}_i \mathbf{r}_j \rangle = Mr^2, \quad (9.19)$$

т. е. линейные размеры (радиус и т. д.) клубка растут с ростом числа  $M$  звеньев в цепи как  $M^{1/2}$ . Отсюда следует, что объем клубка пропорционален  $M^{3/2}$ , в то время как для всех «нормальных» (имеющих фиксированную

плотность) тел объем растет только как число частиц  $M$ , т. е. гораздо медленнее, чем  $M^{3/2}$ . Такая anomalously сильная зависимость объема клубка от размера цепи и есть самое характерное свойство клубка. Из него, в частности, вытекает крайне низкая плотность клубка, образованного длинными цепями, и, как следствие — практическое отсутствие контактов между дальними по цепи звеньями.

Добавление. Так как объем клубка порядка  $M^{3/2}$ , а вероятность встречи концов цепи обратно пропорциональна занимаемому клубком объему, то вероятность такой встречи пропорциональна  $M^{-3/2}$ . Это значит, что свободная энергия замыкания петли клубкообразной цепью растет с ее длиной как  $kT(3/2)\ln(M)$  [закон Флори].

■ Недостаточная реалистичность свободно-сочлененной модели в том, что она допускает, что в любом своем звене цепь может поворачиваться на любой угол, а это, вообще говоря, не так. Однако формулу (9.19) можно переписать в более общем виде,

$$\langle h^2 \rangle = Lr, \quad (9.20)$$

где  $L = Mr$  — полная («контурная») длина цепи, пропорциональная числу звеньев в ней, а  $r$  — эффективное расстояние между «свободными сочленениями» цепи, т. е. характерная длина, на которой цепь «забывает» о своем направлении (для справки: в полипептидной цепи эта характерная длина — или, как говорят, «длина Куновского сегмента» — составляет 30–40 Å, если цепь не содержит ни слишком гибких глицинов, ни слишком жестких пролинов, т. е. Куновский сегмент такой цепи содержит около 10 аминокислотных остатков).

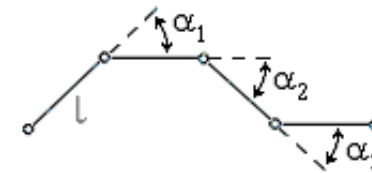


Рис. 9-8. Поворотнo-изомерная модель клубка

В виде (9.20) формула, описывающая размер клубка, уже обретает достаточную общность и может применяться для описания реальных полимеров. При этом, если  $l$  — расстояние между «узлами» — точками, в которых происходит поворот цепи (т. е. длина валентной связи), а  $\langle \cos \alpha \rangle$  — средний косинус угла поворота в узле (рис. 9-8), то

$$r = l(1 + \langle \cos \alpha \rangle) / (1 - \langle \cos \alpha \rangle). \quad (9.21)$$

## Лекция 10

Свойства боковых групп аминокислотных остатков. Включение аминокислотных остатков во вторичную структуру. Аланин, глицин, пролин, валин. Неполарные, короткие полярные и длинные полярные боковые группы. Заряженные боковые группы. Гидрофобные поверхности на вторичных структурах в белках.

Обсудим теперь свойства боковых групп аминокислотных остатков. В особенности я хочу остановиться на том, какие именно структуры стабилизируют те или иные остатки.

- Список 20 «стандартных», т. е. кодируемых ДНК-аминокислотных остатков дан в табл. 10-1; там же дан их молекулярный вес и встречаемость в белках. Структуры этих аминокислотных остатков представлены на рис. 10-1.

Таблица 10-1. Основные свойства аминокислотных остатков

Название	Аминокислотный остаток			$\Delta G_{\text{вода} \rightarrow \text{спирт}}$ бок. группы при 25 °С (ккал/моль)
	Код 3-букв. 1-букв.	% в белках <i>E. coli</i>	Мол. вес при pH7 (дальтон)	
Глицин	Gly G	8	57	0,0
Аланин	Ala A	13	71	-0,4
Пролин	Pro P	5	97	-1,0
Глутаминовая кислота	Glu E	≈6	128	+0,9
Глутамин	Gln Q	≈5	128	+0,3
Аспарагиновая кислота	Asp D	≈5	114	+1,1
Аспарагин	Asn N	≈5	114	+0,8
Серин	Ser S	6	87	+0,1
Гистидин	His H	1	137	-0,2
Лизин	Lys K	7	129	+1,5
Аргинин	Arg R	5	157	+1,5
Треонин	Thr T	5	101	-0,3

Название	Аминокислотный остаток			$\Delta G_{\text{вода} \rightarrow \text{спирт}}$ бок. группы при 25 °С (ккал/моль)
	Код 3-букв. 1-букв.	% в белках <i>E. coli</i>	Мол. вес при pH7 (дальтон)	
Валин	Val V	6	99	-1,6
Изолейцин	Ile I	4	113	-2,4
Лейцин	Leu L	8	113	-2,3
Метионин	Met M	4	131	-1,6
Фенилаланин	Phe F	3	147	-2,4
Тирозин	Tyr Y	2	163	-1,3
Цистеин	Cys C	2	103	-1,3
Триптофан	Trp W	1	186	-3,0

*Примечания.* Все данные взяты из [3] — за исключением данных по гидрофобности ( $\Delta G_{\text{вода} \rightarrow \text{спирт}}$ ) боковых групп, которые взяты из J. L. Fauchère, V. Pliska, *Eur. J. Med. Chem.-Chim. Ther.* (1983) 18:369. Объем (в Å<sup>3</sup>), приходящийся на аминокислотный остаток в белке или в растворе, близок к его молекулярному весу (в дальтонах), умноженному на 1,3. Точнее, процентов на 5 больше, чем (мол. вес) × 1,3, если в остатке много алифатических (-CH<sub>2</sub>-, -CH<sub>3</sub>) групп, и процентов на 5 меньше, чем (мол. вес) × 1,3, если в остатке много полярных (O, N) атомов.

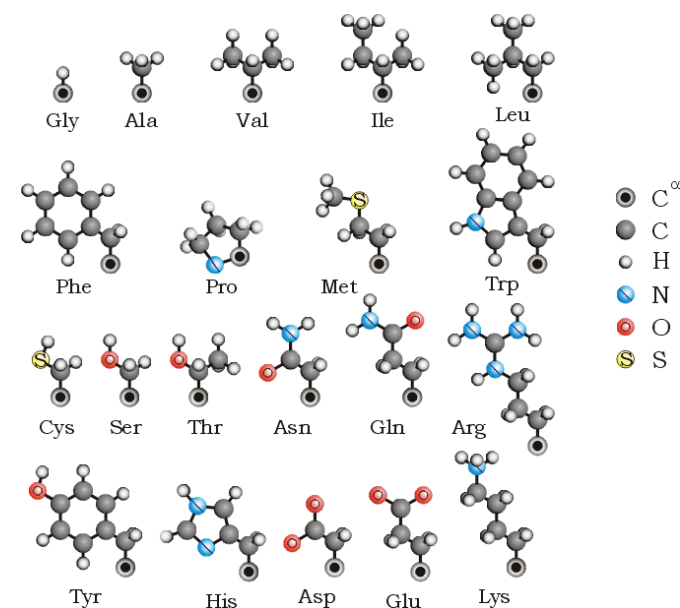


Рис. 10-1. Боковые цепи двадцати стандартных аминокислотных остатков. Справа показаны входящие в аминокислоты атомы

Позже выяснилось, что ДНК иногда кодирует еще одну «нестандартную», 21-ю аминокислоту: *селеноцистеин*, т.е. цистеин, где атом серы заменен на атом селена. Селеноцистеину соответствует кодон UGA матричной РНК, но не всегда (обычно UGA служит терминирующим кодоном), а только в окружении определенного нуклеотидного контекста. И тРНК, заряженная селеноцистеином, опознает только такой кодон.

■ Рассмотрим теперь структурные тенденции аминокислотных остатков; они стали известными после многолетнего статистического исследования белковых структур. Такие исследования отвечают на вопрос: «Что чаще всего бывает, а чего чаще всего не бывает?»

Для систематизации ответов полезной может быть следующая ниже табл. 10-2, куда я вписал, наряду со встречаемостью аминокислотных остатков (А.к. ост.) в разных местах белков, такие свойства остатков, как: наличие NH-группы в главной цепи (ее нет только у иминокислоты пролина); наличие C<sup>β</sup>-атома в боковой цепи (его нет только у глицина); число не-водородных γ-атомов в боковой цепи; наличие и вид полярных группировок в боковой цепи (диполей или зарядов со знаком; жирным выделено то зарядовое состояние, что относится к «нормальному» рН7).

Таблица 10-2. Основные структурные свойства аминокислотных остатков

А.к. ост.	Наличие		Число γ б.гр.	Диполь/заряд	рК	Яркая тенденция быть:					
	NH гл.ц.	C <sup>β</sup> б.гр.				до-α <sub>N</sub>	в-α <sub>N</sub>	за-α <sub>C</sub>	в-α <sub>C</sub>	β	петлях
Gly	+	-				-	-	+			
Ala	+	+				+					
Pro	-	+	1			+	-	-	-		+
Glu	+	+	1	COOH ⇒ CO <sub>2</sub> <sup>-</sup>	4,3	+	+	-	-	-	-
Asp	+	+	1	COOH ⇒ CO <sub>2</sub> <sup>-</sup>	3,9	+	+	-	-	-	+
Gln	+	+	1	OCNH <sub>2</sub>							-
Asn	+	+	1	OCNH <sub>2</sub>		+	-	+	-	+	-
Ser	+	+	1	OH		+					+
His	+	+	1	NH; и N ⇒ NH <sup>+</sup>	6,5	-		+	+		
Lys	+	+	1	NH <sub>2</sub> ⇒ NH <sub>3</sub> <sup>+</sup>	10,5	-	-	+	+	-	-
Arg	+	+	1	HNC(NH <sub>2</sub> ) <sub>2</sub> <sup>+</sup>	12,5	-	-	+	+	-	+
Thr	+	+	2	OH		+				+	
Ile	+	+	2							+	-
Val	+	+	2							+	-
Leu	+	+	1				+			+	-
Met	+	+	1				+			+	-
Phe	+	+	1							+	-
Tyr	+	+	1	OH ⇒ O <sup>-</sup>	10,1					+	+
Cys	+	+	1	SH ⇒ S <sup>-</sup>	9,2					+	+
Trp	+	+	1	NH						+	+

*Примечания.* То зарядовое состояние, что относится к «нейтральному» рН7, выделено жирным шрифтом в столбце «Диполь/заряд»; рК – это то значение рН, при котором группа находится в заряженном или в незаряженном состоянии с равной вероятностью (см. рис.10-5 ниже). К «структурным свойствам» отнесена тенденция быть в α-спирали (α), и особо — в ее N- и C-концевых витках (α<sub>N</sub>, α<sub>C</sub>), а также — непосредственно перед N- и за C-концом спирали (до-α<sub>N</sub>, за-α<sub>C</sub>); тенденция быть в β-структуре (β); тенденция быть в нерегулярных структурах, т.е. «петлях» (включая сюда и β-изгибы цепи); и, наконец, — тенденция быть в гидрофобном ядре глобулы, а не на ее поверхности. «Тенденция» определялась по концентрации остатка в данной структуре, деленной на его среднюю концентрацию в глобулярных белках. Тенденцию «быть» я отмечал значком «+», «не быть» — значком «-». Разница между частотами встречаемости, помеченными значками «+» и «-», примерно двукратна. Более жирным значком отмечалась особо сильная тенденция.

Попробуем понять основные закономерности этой таблицы, исходя из того, что мы уже изучили. При этом мы будем использовать следующую логику: так как белок в целом стабилен, он должен в основном состоять из стабильных элементов, т.е. именно они должны наблюдаться в его структуре чаще всего, а нестабильные должны наблюдаться редко.

Почему пролин не любит вторичной структуры? Потому, что у него нет NH-группы в главной цепи, т.е. у него вдвое уменьшена возможность завязывать водородные связи, а именно на них и держится вторичная структура. Почему он, тем не менее, любит N-конец спирали? Потому, что здесь, на N-конце, NH-группы «торчат» из спирали, т.е. они и так не вовлечены в водородные связи, и здесь пролину нечего терять... С другой стороны, угол φ в пролине фиксирован его кольцом примерно при -60°, т.е. его конформация уже почти «готова» для α-спирали (рис. 10-2а).

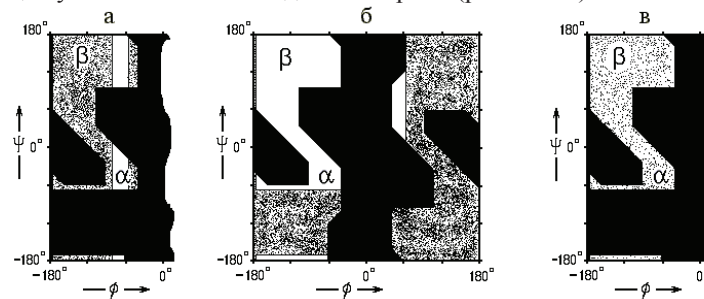


Рис. 10-2. Запрещенные и разрешенные конформации различных аминокислотных остатков и — на их фоне — конформации α- и β-структуры. (а) Разрешенные для пролина конформации (□) на фоне конформаций, разрешенных для аланина (▨); ■ — конформации, запрещенные для них обоих. (б) Разрешенные конформации аланина (□) на фоне конформаций ▨, разрешенных лишь для глицина; ■ — области, запрещенные для всех остатков. (в) Карта запрещенных (■) и разрешенных (□, ▨) конформаций более крупных остатков. В области □ разрешены все конформации боковой группы по углу χ<sup>1</sup>, в области ▨ часть углов χ<sup>1</sup> запрещена

Почему глицин не любит вторичной структуры и предпочитает нерегулярные участки («клубок»)? Потому, что для него допустима очень широкая область углов ( $\phi\psi$ ) на карте Рамачандрана (рис. 10-2б), ему легко принимать самые разнообразные конформации, лежащие вне вторичной структуры.

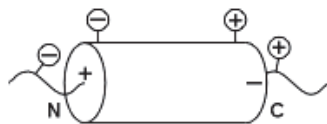
Наоборот, аланин — с более узкой, но включающей и  $\alpha$ - и  $\beta$ -конформацию разрошенной областью на карте Рамачандрана (рис. 10-2б) — предпочитает нерегулярным конформациям  $\alpha$ -спираль (и отчасти  $\beta$ -структуру).

Остальные гидрофобные остатки (т.е. остатки без зарядов и диполей в боковой цепи) предпочитают, как правило,  $\beta$ -структуру. Почему? Потому, что их крупные  $\gamma$ -атомы могут там располагаться более свободно (рис. 10-2в). Особенно это важно для боковых групп с двумя крупными  $\gamma$ -атомами.

А вот аминокислоты с полярными группами в боковых цепях предпочитают поверхностные нерегулярные участки («клубок»), где эти полярные группы могут завязать водородные связи как с полипептидной цепью, так и с водой. Особенно заметна эта тенденция для наиболее полярных, заряженных при «нормальном» рН7 остатков, и для самых коротких (см. рис. 10-1) полярных боковых цепей, с наиболее приближенными к главной цепи полярными группами. Кстати, по той же причине, поскольку у них там есть возможность завязать дополнительную водородную связь, короткие полярные боковые группы любят места у обоих концов спирали.

Некое исключение среди аминокислот с диполями в боковой цепи составляют триптофан и тирозин, имеющие маленький диполь на фоне большой гидрофобной части, и цистеин, у которого (т.е. у SH-группы которого) водородные связи совсем слабые. Они ведут себя, в общем, так же, как гидрофобные остатки.

Мы видим также, что отрицательно заряженные боковые группы предпочитают N-конец спирали (*точнее*: N-концевой виток и один-два остатка перед ним) и не любят C-концевой виток (и пару остатков за ним), а положительно заряженные — предпочитают C-конец спирали и не любят ее N-конец. Почему? Потому, что на N-конце из спирали торчат NH-группы и на нем образуется заметный положительный заряд, и «минусы» боковых цепей притягиваются к нему, а «плюсы» отталкиваются от него (рис. 10-3). А C-конец спирали заряжен, наоборот, отрицательно, и там эффект противоположен: около C-конца любят собираться «плюсы» боковых цепей, а «минусы» его избегают.



**Рис. 10-3.** Предпочтительное расположение заряженных боковых групп у концов  $\alpha$ -спирали

Что касается расположения остатков внутри белка или на его поверхности, здесь общая тенденция заключается в том, что полярные (гидрофильные) боковые группы находятся снаружи, где они могут контактировать с полярной же водой («подобное растворяется в подобном!»). Отрываться от воды полярным группам плохо — теряются водородные связи. Особенно плохо отрываться заряженным группам: переход из среды с высокой диэлектрической проницаемостью (из воды) в среду с низкой (ядро белка) ведет к большому повышению свободной энергии. И действительно, ионизированных групп внутри белка практически нет (а почти все исключения связаны либо с координационными связями с ионом металла, либо с активными центрами, ради которых, собственно, белок и создан...).

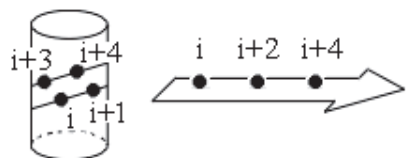
Наоборот, большинство гидрофобных боковых групп находятся внутри белка — они-то и создают здесь гидрофобное ядро (опять: «подобное растворяется в подобном!»). Мы уже говорили, что гидрофобность группы тем больше, чем больше ее неполярная поверхность: именно ее нужно упрятать от воды. Для чисто неполярных групп гидрофобный эффект прямо пропорционален их поверхности, а для групп с полярными вкраплениями — их поверхности за вычетом поверхности этих вкраплений.

Слипание гидрофобных групп — главная движущая сила образования белковой глобулы. Главная, но не единственная — еще есть образование водородных связей во вторичной структуре (о чем мы уже говорили) и образование плотной, квазикристаллической упаковки внутри белка (о чем мы еще поговорим в свое время).

Для создания гидрофобного ядра белковой цепью она должна входить в него с уже насыщенными водородными связями — ведь иначе ее полярным пептидным группам от воды придется оторваться, а разрыв водородной связи дорог. Поэтому в гидрофобное ядро вовлекается цепь, уже образовавшая (или образующая при этом) вторичную структуру и тем самым насытившая водородные связи пептидных групп в главной цепи. Однако при этом в ядро должны увлекаться только гидрофобные остатки вторичной структуры, а входящие в нее полярные остатки должны остаться вне ядра, потому и на  $\alpha$ -спиралях, и на  $\beta$ -структурных участках выделяются гидрофобные и гидрофильные поверхности; для их создания необходимо определенное чередование соответствующих групп в белковой цепи (рис. 10-4).

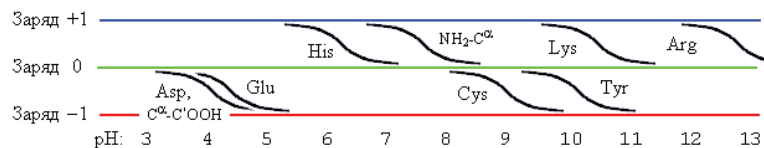
Все закономерности, о которых мы сейчас говорили, используются как для конструирования искусственных белков, так и для предсказания — по аминокислотным последовательностям — вторичной структуры белков, а также для предсказаний тех участков их цепи, что глубоко погружены в белок, или, наоборот, тех участков, что лежат на поверхности белка. К этим вопросам мы еще вернемся.





**Рис. 10-4.** Боковые группы, которые (если все они неполярные) могут формировать единые гидрофобные поверхности на  $\alpha$ -спиралях и на  $\beta$ -структурных участках. Аналогичные сочетания полярных групп в цепи приводят к образованию гидрофильных областей на противоположных поверхностях  $\alpha$ -спиралей и  $\beta$ -тяжей

- В заключение — еще немного о заряженных (или, точнее, ионизируемых) боковых группах. Повышение pH, т. е. падение концентрации  $H^+$ , всегда делает группу «более отрицательно заряженной» — нейтральная группа приобретает отрицательный заряд, а положительно заряженная — разряжается, рис. 10-5. Переход из незаряженного в заряженное или из заряженного в незаряженное состояние происходит у разных групп при разных pH, однако ширина перехода при этом всегда одна и та же — около 2 единиц pH (в этом интервале отношение заряженной и незаряженной форм меняется от 10:1 до 1:10).



**Рис. 10-5.** Заряженность ионизируемых боковых групп, а также N-конца пептидной цепи ( $NH_2-C^\alpha$ ) и ее C-конца ( $C^\alpha-COOH$ ) в водном окружении при разных pH. То pH, что соответствует «полузаряженному» состоянию, и есть величина pK ионизации данной группы (см. табл.10-2). Вероятность заряженного состояния ионизируемой группы относится к вероятности ее незаряженного состояния как  $10^{(pK-pH)}$ :1 для положительно заряжающейся группы, и  $10^{-(pK-pH)}$ :1 для отрицательно заряжающейся группы

- Следует обратить особое внимание на группы, переходящие из незаряженного в заряженное состояние при pH, близком к 7, характерном для жизни белка в клетке: именно такие легко перезаряжаемые группы (и особенно гистидин) часто используются в активных центрах белков.

Как уже говорилось, ионизируемой группе легче погрузиться в неполярную среду (т. е. вглубь белка или мембраны) в разряженном состоянии. Действительно, погружение иона стоит, по сделанной нами оценке, очень дорого — несколько десятков килокалорий на моль. А чего стоит разрядка

ионизируемой группы? Это нетрудно оценить, имея перед глазами рис.10-5. Вероятность незаряженного состояния группы в водном окружении составляет  $W_0 = 1/[1+10^{(pK-pH)}]$  для положительно и  $W_0 = 1/[1+10^{-(pK-pH)}]$  для отрицательно заряжающейся группы. Свободная энергия этого незаряженного состояния  $F_0 = -kT \ln W_0$ ; т. е. для положительно заряжающейся группы  $F_0 \approx 0$ , если  $pH > pK$ , и  $F_0 \approx +2,3kT(pK-pH)$ , если  $pH < pK$ . А для отрицательно заряжающейся группы  $F_0 \approx 0$  при  $pH < pK$ , и  $F_0 \approx -2,3kT(pK-pH)$  при  $pH > pK$ . Таким образом, «свободная энергия разрядки»  $F_0$  не превосходит нескольких килокалорий на моль для всех изображенных на рис.10-5 ионизируемых групп при «нормальных»  $pH \approx 7$ .

## ПРОСТРАНСТВЕННОЕ СТРОЕНИЕ БЕЛКОВ

### Лекция 11

Фибриллярные белки, их функции и их периодичные первичные и вторичные структуры;  $\alpha$ -кератин,  $\beta$ -фиброин шелка, коллаген. Упаковка длинных  $\alpha$ -спиралей и обширных  $\beta$ -листов. Белки, образующие матрикс; эластин. Генетические дефекты белков и болезни. Амилоиды.

Теперь, когда мы знаем свойства вторичных структур полипептидов и свойства аминокислотных остатков, обратимся, наконец, к белкам.

Как уже говорилось, по «жизненным условиям», по стабилизирующим структуре белков взаимодействиям и общему типу строения белки можно разбить на три класса: (1) фибриллярные белки; (2) мембранные белки; (3) водорастворимые глобулярные белки.

Рассмотрим сначала фибриллярные белки.

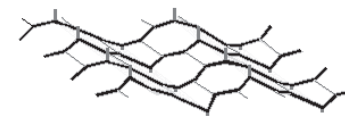
Фибриллярные белки играют в основном структурную роль. Они образуют микрофиламенты, фибриллы, волосы, шелк и другие защитные покровы; они армируют мембраны и поддерживают структуру клеток и тканей. Фибриллярные белки часто очень велики. К ним принадлежит и самый большой из известных белков — мышечный белок титин, состоящий из примерно 30 000 аминокислотных остатков.

Фибриллярные белки часто образуют огромные агрегаты; их пространственная структура высоко регулярна, сложена в основном из очень больших блоков вторичной структуры и держится в значительной степени на взаимодействии между разными полипептидными цепями. Первичная структура фибриллярных белков также высоко регулярна, периодична, потому-то из нее и образуется обширная регулярная вторичная структура.

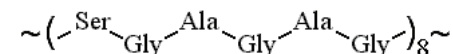
Типичными представителями фибриллярных белков являются:

- а)  $\beta$ -структурный белок фиброин шелка. У  $\beta$ -листа, как мы знаем, периодичность состоит в чередовании остатков, обращенных «вверх» и «вниз» (рис. 11-1).

**Рис. 11-1.** Лист  $\beta$ -структуры. Тенями подчеркнута его складчатость и периодичность. Тонкие линии — водородные связи между цепями, слагающими  $\beta$ -лист



Соответственно, в фиброине шелка основной мотив первичной структуры выглядит как повтор восьми блоков, где в каждом из блоков идет чередование маленьких Gly и более крупных остатков, например:



и этот восьмикратный повтор шести остатков, отороченный менее регулярными кусками цепи, повторяется около 50 раз.

Антипараллельные (такие, как на рис. 11-1)  $\beta$ -слои фиброина шелка уложены друга на друга по принципу «лицом к лицу, спина к спине»: двойной слой глицинов (расстояние между плоскостями 3,5 Å) — двойной слой аланинов/серинов (расстояние между плоскостями 5,7 Å; это хорошо видит рентген) — двойной слой глицинов — и т. д.

В волокне шелка эти квазикристаллы, состоящие из многих  $\beta$ -слоев, погружены в менее упорядоченный матрикс, образованный нерегулярными частями фиброина, а также — специальным неупорядоченным матриксным белком серицином, сшитым S–S-связями в огромный агрегат.

- б)  $\alpha$ -структурные фибриллярные белки, сложенные из длинных перевитых спиралей (coiled coil) (рис. 11-2). В  $\alpha$ -кератине, во всех, обладающих очень разной первичной структурой белках промежуточных филаментов, в тропомиозине такие спирали охватывают всю белковую цепь, да и большая часть миозиновой цепи образует фибриллу такого типа. Такие структуры содержатся также в шелке, но не в обычном шелке тутового шелкопряда, о котором шла речь выше, а шелке пчел и муравьев.

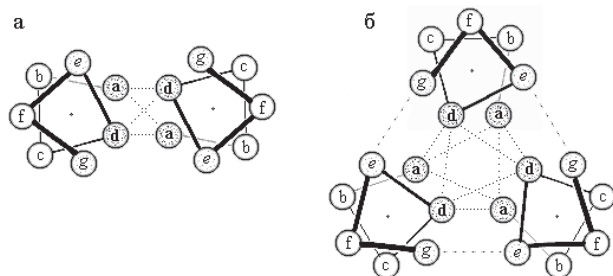
**Рис. 11-2.** Перевитые правые  $\alpha$ -спирали. В комплексе они лежат параллельно друг другу и слегка закручены одна вокруг другой так, что каждая из них образует левую суперспираль с периодом в 140 Å. Контактующие аминокислотные остатки занимают в цепи повторяющиеся позиции **a** и **d** (см. рис. 11-3, 11-4)



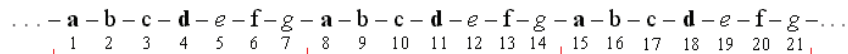
Слипаясь, отдельные цепи образуют суперспираль (рис. 11-2, 11-3).

На следующем структурном уровне изображенные на рис. 11-2 суперспирали часто (но не всегда, например, не в тропомиозине) слипаются друг с другом и образуют фибриллы.

Входящие в суперспираль  $\alpha$ -спирали обычно параллельны, и перевиты они — в разных белках — по две, три или больше. У  $\alpha$ -спирали, как мы уже знаем, период равен 3,6 остатка на виток. В перевитых спиральных периодичность — 7 остатков на два витка  $\alpha$ -спирали, т. е. 3,5 остатка на виток (рис. 11-3, 11-4).



**Рис. 11-3.** Взаимодействие  $\alpha$ -спиралей в двойной (а) и тройной (б) суперспирали (вид с торца спирали). В двойной суперспирали непосредственно контактируют с другой спиралью только остатки **a** и **d**, а в тройной — еще и остатки **e** и **g** (хотя и более слабо). Картинка, с некоторыми изменениями, взята из [6]



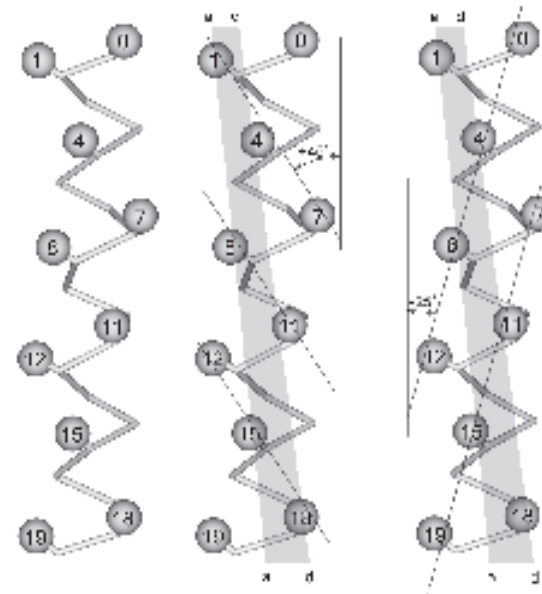
**Рис. 11-4**

Типичная первичная структура при этом имеет ту же периодичность — 7 остатков — и выглядит, как на рис. 11-4. На этом рисунке жирные буквы соответствуют гидрофобным (жирным, слипающимся) аминокислотам, а прочие буквы — аминокислотам гидрофильным.

Интересно, что небольшое увеличение гидрофобности остатков, находящихся в «промежуточных» позициях **e** и **g**, превращает двойную перевитую спираль (рис. 11-3а) в тройную (рис. 11-3б), а еще большее — в четверную и т. д.

Интересно также, что механическое растяжение увлажненного волокна, сложенного из  $\alpha$ -спиралей, может заставить их совершить переход в  $\beta$ -структурную форму, а при снятии напряжения и понижении влажности  $\alpha$ -спиральная структура восстанавливается.

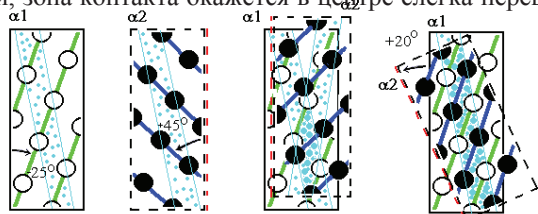
■ Остановимся чуть подробнее на том, как спирали слипаются между собой. На  $\alpha$ -спирали (рис. 11-5, слева) есть несколько спиральных нарезок — «хребтов» из сближенных в пространстве боковых групп. Одни хребты имеют периодичность типа  $1-4-7-...$ , и та их часть, что входит в зону контакта, состоит из пар  $a_1-d_4, a_8-d_{11}, ...$  (рис. 11-5, в центре). Другие хребты имеют периодичность типа  $0-4-8-12-...$ . Та часть этих хребтов, что входит в зону контакта, состоит из пар  $d_4-a_8, d_{11}-a_{15}, ...$  (рис. 11-5, справа).



**Рис. 11-5.**  $\alpha$ -спираль (главная цепь и  $C^\beta$ -атомы) и два сорта хребтов (тонкие линии) из сближенных боковых групп на ее поверхности. В центре: хребты типа  $i, i+3$  (или просто «+3»); справа: хребты типа  $i, i+4$  (или «+4»). Полосой показана контактная поверхность; ее края выделены линии остатков **a** и **d**. Отмечены типичные углы наклона хребтов  $i, i+3$  и  $i, i+4$  относительно оси спирали (на рисунке углы представляются меньшими, так как типичные хребты проходят через массивные боковые группы, а на рисунке — через центры  $C^\beta$ -атомов). Картинка, с некоторыми изменениями, взята из [5]

Первые хребты, хребты типа « $i, i+4$ » (если проводить их через выступы, образованные боковыми группами) идут под углом примерно  $-25^\circ$  к оси спирали, вторые хребты, хребты типа « $i, i+3$ » — под углом примерно  $+45^\circ$  (см. рис. 11-5). Если, перевернув вокруг вертикальной оси, наложить одну поверхность спирали на другую (рис. 11-6), а затем повернуть на  $+20^\circ$  вокруг перпендикулярной странице оси, то хребты

типа «1 – 4 – 7» одной спирали окажутся между хребтами типа «0 – 4 – 8» другой, что обеспечит их плотный контакт (см. рис. 11-6, справа). При этом группы **a** одной спирали окажутся между группами **d** другой, а зона контакта образует на поверхности обеих спиралей (пока они не суперспирализованы) слабо скрученную спиральную полосу. А когда эти спирали суперспирализуются (см. рис. 11-2) и обвиваются вокруг общей оси, зона контакта окажется в центре слегка перевитого пучка.



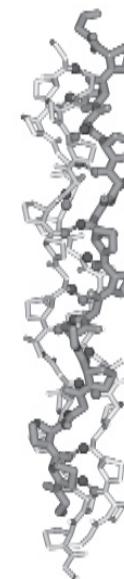
**Рис. 11-6.** Плотная упаковка хребтов боковых групп при контакте спиралей требует разворота на  $+20^\circ$  одной спирали относительно другой. Мы смотрим на зону контакта сквозь одну спираль (сквозь перевернутую вдоль оси  $\alpha 2$ ). Остатки «нижней» спирали  $\alpha 1$  изображены светлыми, а верхней  $\alpha 2$  — темными кружками. Тонкими линиями показаны линии контактирующих остатков **a** и **d**. Картинка, с некоторыми изменениями, взята из Chothia et al., *Proc. Natl. Acad. Sci. USA* (1977) 74:4130–4134

Это — не единственный способ плотного контакта спиралей (с другими мы познакомимся, когда будем говорить о глобулярных белках), но единственный, пригодный для очень длинных спиралей, типичных для фибриллярных белков. Он был предсказан еще Криком в том же 1953 году, когда он вместе с Уотсоном предсказал двойную спираль ДНК.

■ в) Коллаген (по-гречески — «образующий клей»). Главный опорный белок, составляющий около четверти белковой массы позвоночных. Он формирует прочные нерастворимые фибриллы. Молекула коллагена образуется особой, тройной суперспиралью, сложенной из трех полипептидов (рис. 11-7). При этом внутри каждого полипептида, внутри каждой нити этой тройной спирали водородных связей нет — они есть только между нитями.

Конформация всех остатков в каждой цепи коллагена близка к конформации полипролиновой [точнее, poly(Pro)II] спирали; однако, в отличие от poly(Pro)II, коллагеновая спираль имеет межцепные водородные связи. Это — левая спираль, и период ее равен трем. Соответственно, основной мотив первичной структуры в коллагене выглядит как многократный повтор троек (Gly-Pro-Pro)<sub>n</sub> или, точнее, (Gly-нечто-Pro)<sub>n</sub>, причем Gly в такой тройке необходим для образования водородных связей: у него есть NH-группа (в отличие от Pro) и нет боковой группы, а любая боковая группа была бы лишней в центре тугой коллагеновой суперспирали, где сидит образующий H-связь глицин.

**Рис. 11-7.** Модель тройной суперспирали коллагена для последовательности (глицин — пролин — пролин)<sub>n</sub>. Каждая цепь выделена своим оттенком. Водородные связи N-атомов NH-групп глицина с O-атомами CO-групп первых пролинов троек Gly-Pro-Pro показаны в виде маленьких черных гантелек. При этом Gly цепи «1» завязывает связь с цепью «2», а Pro — с цепью «3», и т. д. Завиваясь вокруг двух других, каждая цепь коллагена образует правую суперспираль. «Супер» — потому что на более мелком масштабе, на масштабе конформаций отдельных остатков, коллагеновая цепь уже образует спираль типа poly(Pro)II (причем эта «микроспираль» — левая); ее можно проследить по направлению пролиновых колец. Молекула коллагена имеет около 15 Å в диаметре и около 3000 Å в длину



Интересно, что экзоны, кодирующие коллагеновую цепь, практически всегда начинаются с глицинов и содержат число кодонов, кратное трем. Я надеюсь, вы помните, что гены эукариот содержат экзоны, кодирующие куски белков, и интроны, которые выщепляются из матричной РНК (и потому белки не кодируют).

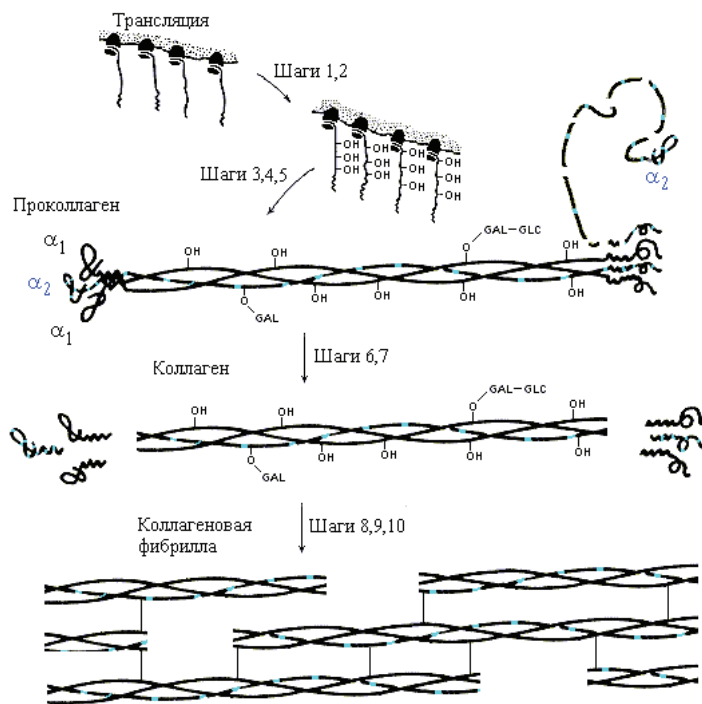
На следующем структурном уровне коллагеновые суперспирали также слипаются друг с другом и образуют фибриллы коллагена.

Биосинтез коллагена, его последующая модификация и образование зрелой структуры коллагеновой фибриллы хорошо изучены (рис. 11-8).

Для самоорганизации коллагена нужен проколлаген, включающий, кроме коллагеновых нитей, глобулярные головки и хвостики. Лишенные головок и хвостиков коллагеновые нити, самоорганизуясь из развернутого состояния *in vitro*, тоже складываются в тройные спирали — но «неправильные», без присущей нативному коллагену гетерогенности тройной спирали (включающей нити двух сортов), без присущего ему регистра (т. е. с неправильным сдвигом нитей относительно друг друга), и т. д. Поэтому сам по себе коллаген *не* способен к правильной спонтанной самоорганизации своей пространственной структуры *in vitro*, так же, как и фиброин шелка, и в отличие от рассмотренных выше  $\alpha$ -структурных суперспиралей и особенно от глобулярных белков, о которых я буду говорить потом.

При повышении температуры коллагеновая спираль плавится (так образуется желатина). Температура плавления коллагена тесно связана с содержанием пролина и оксипролина (чем больше этих жестких остатков, тем выше, естественно, температура плавления), и обычно лишь на несколько градусов превышает температуру тела животного — хозяина этого коллагена. Последний факт возьмем на заметку, через несколько лекций мы к этому еще вернемся.





**Рис. 11-8.** Образование коллагена *in vivo*. *Шаг 1.* Биосинтез про- $\alpha_1$ -цепей и про- $\alpha_2$ -цепей (по 1300 остатков в каждой) в пропорции 2:2. *Шаг 2.* Ферментативное ко-трансляционное гидроксирование некоторых, находящихся перед Gly, остатков Pro и Lys. *Шаг 3.* Присоединение сахаров (GLC-GAL) к части гидроксированных остатков. *Шаг 4.* Образование тетрамера из C-концевых глобул двух про- $\alpha_1$ - и двух про- $\alpha_2$ -цепей; последующая деградация одной про- $\alpha_2$ -цепи; образование проколлагена (про- $\alpha_1$ - $\alpha_2$ - $\alpha_1$ -тримера) и S-S-связей на его глобулярных концах. *Шаг 5.* Образование тройной спирали в середине проколлагена. *Шаг 6.* Секретция проколлагена во внеклеточное пространство. *Шаг 7.* Отщепление глобулярных частей. *Шаги 8-10.* Спонтанное образование фибрилл из тройных суперспиралей, окончательная модификация аминокислотных остатков и образование (под действием особого фермента) ковалентных сшивок модифицированных остатков коллагеновых цепей. Картинка взята из [3] и адаптирована

Тему сворачивания коллагена делает особенно интересным то, что ряд наследственных болезней связан с мутациями, влияющими на сворачивание коллагенов. Наиболее изученной из этих болезней является *osteogenesis imperfecta*, или болезнь «ломких костей». Наиболее частой причиной этого синдрома является замена одного нуклеотида, которая приводит к замене глицина на другую аминокислоту. Это нарушает повто-

ряющуюся последовательность типа (Gly-X-Y)<sub>n</sub> — последовательность, составляющую основу характерной трехспиральной структуры коллагена. По-видимому, по крайней мере одним из следствий такой мутации является замедление сворачивания коллагена — замедление, позволяющее происходить аномальным посттрансляционным модификациям этого белка. Для того чтобы преодолеть трудности, связанные с изучением «цельного» коллагена, эффект мутации, связанной с болезнью, исследовался на упрощенных периодических пептидах. Это позволило провести детальные биофизические исследования, включая ЯМР-ные эксперименты в «реальном времени», с тем, чтобы исследовать шаги сворачивания на уровне индивидуальных остатков. Этот пример, кстати, иллюстрирует как все более ясно прослеживаемую связь между неправильным сворачиванием белка и человеческими болезнями, так и мощь должным образом примененных структурных методов в исследовании молекулярных основ этих болезней.

В заключение, я хочу подчеркнуть, что фибриллярные белки обычно устроены относительно просто в силу периодичности своей первичной и, в силу этого, также и своей вторичной структуры.

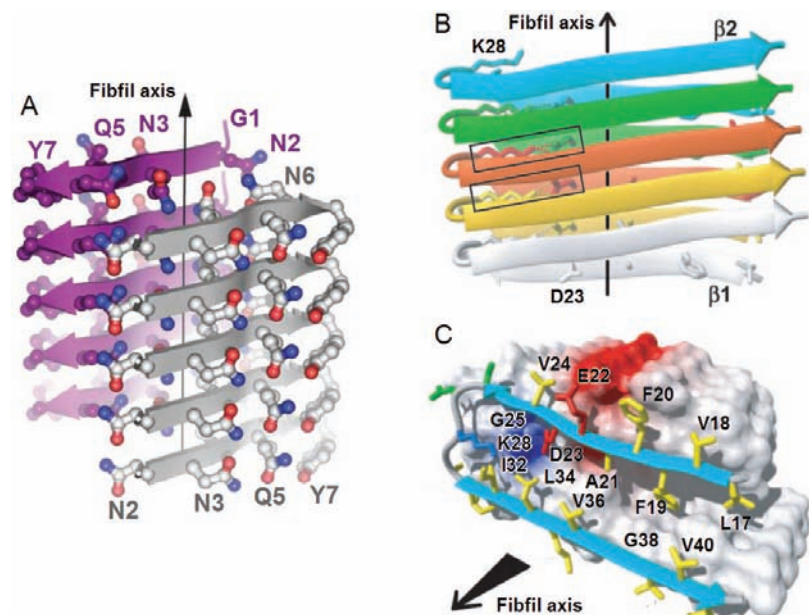
Впрочем, здесь уместно сделать еще одно добавление.

К фибриллярным белкам часто относят и белки, образующие огромные агрегаты, но не имеющие внутри них никакой определенной структуры. Такие белки образуют прошитый теми или иными химическими связями упругий матрикс, в который погружены другие, более структурированные белки.

Типичным матриксным белком является эластин. Он играет важную роль в формировании стенок артерий, легких и связок. Его длинная цепь довольно гидрофобна, и состоит из коротких аминокислотных повторов нескольких типов. Каждая цепь эластина образует неупорядоченный клубок, но эти цепи сшиты в сетку ферментно-модифицированными лизинами (которые сшиваются по четыре в одном узле). В результате образуется нечто вроде резины. Не могу не отметить, что нарушения в работе фермента, модифицирующего лизины в эластине, приводят к утрате эластичности связок и сосудов, а иногда и к разрыву аорты.

И последнее. Недавно обнаружилось, что многие «нормальные» водорастворимые глобулярные белки (о мембранных на сей счет данных пока нет) способны образовывать так называемые амилоидные фибриллы. Среди этих белков — лизоцим, миоглобин и так далее, а в особенности некоторые их мутанты, но, прежде всего — знаменитые прионы. Об амилоидных фибриллах было известно давно, их с белками как-то не

увязывали — пока не увидели, что прионы, эти «инфицирующие белки», образуют такие фибриллы в пораженных ими клетках. Потом выяснилось, что это сходные амилоиды образуют и другие, самые обычные белки, и прионы, на их фоне, выделяются, кажется, только своей инфекционностью. В изученных амилоидных фибриллах белковая цепь уложена в двух- или многослойные сэндвичи, состоящие из параллельной  $\beta$ -структуры.  $\beta$ -тяжи лежат перпендикулярно оси фибриллы, а водородные связи между ними идут вдоль этой оси (рис. 11-9).  $\beta$ -сэндвичи плотно упакованы в



**Рис. 11-9.** (А) Кусок расшифрованной рентгеноструктурным анализом амилоидной фибриллы, образованной коротким (7 остатков) фрагментом дрожжевого белка Sup35 (вызывающим в дрожжах болезнь, сходную с той, что вызывают прионы у млекопитающих); картинка, с некоторыми упрощениями, взята из Nelson et al., *Nature* (2005) **435**:733–778. Обратите внимание, что «гидрофобное ядро» этой фибриллы состоит из полярных (более полярных, чем вода) аминокислот Asn (N) и Gln (Q), боковые группы которых образуют между собой водородные связи. (В), (С): Вид с разных сторон на расшифрованную менее прямыми ЯМР-ными методами амилоидную фибриллу, образованную 42-членным пептидом  $A\beta(1-42)$ , взятым из белка, вызывающего болезнь Альцгеймера; картинки, с некоторыми упрощениями, взяты из Lührs et al., *Proc.Natl. Acad. Sci. USA* (2005) **102**:17342–17347. В пептиде  $A\beta(1-42)$  фрагмент 1-17 неупорядочен (и потому на рисунке от него виден только Leu17), а остальная часть пептида  $A\beta(1-42)$  входит в остов протофиламента, гидрофобное ядро которого состоит в основном из «нормальных» неполярных аминокислот, хотя и включает в себя ионную пару Asp23-Lys28.

протофиламенты, трубчатые или ленточные (их форма зависит от белка и внешних условий), укладываемые в длинные амилоидные нити. Видимо, в  $\beta$ -структуру увлечена не вся белковая цепь, а только ее часть (в то время как остальная остается глобулярной — или развернутой? — и лежит вне остова амилоидного протофиламента). По-видимому, в амилоидных нитях уникальной пространственной структурой обладает только та часть белка, что вовлечена в  $\beta$ -сэндвичи, образующие остов протофиламента. Известно, что образование амилоидных фибрилл имеет вид медленной (часто идущей с «лагом», т.е. с большим запаздыванием перед началом роста) и обычно необратимой  $\beta$ -структурной агрегации. Существуют веские основания полагать, что вред организму наносят не столько сами амилоидные фибриллы, зрелые и крупные, сколько формирующиеся в процессе их образования небольшие, активные олигомерные «зародыши». Так или иначе, процесс и/или результат образования амилоидных фибрилл приводит к тяжелым болезням с огромным порой инкубационным периодом (типа куру, «коровьего бешенства», болезни Альцгеймера и т.д.). Изучение этих и других «конформационных болезней» фибриллярных (и других) белков — одна из «горячих точек» современной молекулярной биологии.

■ И еще. Под маркой фибриллярных белков иногда рассматривают фибриллы, сложенные из «нормальных», не подвергнувшихся перестройке глобул (например, актиновые фибриллы). Мы их не будем здесь рассматривать — о глобулярных белках речь пойдет позже.

## Лекция 12

Мембранные белки, особенности их строения и функции. Бактериородопсин, рецепторы и G-белки, порин, фотосинтетический центр. Селективная проницаемость мембранных пор. Работа фотосинтетического центра. Понятие о туннельном эффекте. Понятие об электронно-конформационном взаимодействии.

Рассмотрим теперь мембранные белки. Они тоже — я имею в виду их трансмембранные части — устроены довольно просто, почти как фибриллярные белки.

Мембраны создают поверхность клетки, а внутри нее — разные замкнутые объемы — или, как говорят, компартменты. Мембраны состоят из жира (липидов) и белков (рис. 12-1). Особая роль мембранных белков (а они составляют до половины веса мембраны) — обеспечивать транспорт через нее различных веществ, а также сигналов. Мембрана — это, так сказать, «изолятор», а белки — точнее, как мы увидим, каналы в них — «проводники». Эти проводники специфичны, каждый из них пропускает через мембрану только определенные молекулы (или — по-видимому, посредством небольшого изменения конформации белка — сигналы от определенных молекул).

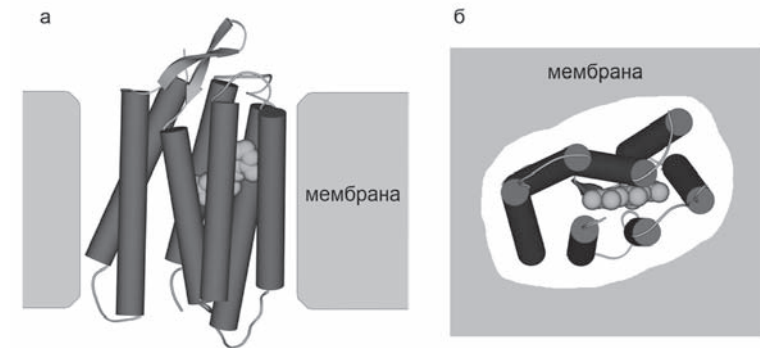


**Рис. 12-1.** Белки в мембране. Серым цветом выделены внемембранные домены. Внутримембранные части белка практически не содержат нерегулярных участков цепи. Выступающие наружу из клетки части эукариотических мембранных белков, как правило, сильно гликозилированы, что усиливает их гидрофильность

Истинно мембранные белки «живут» внутри мембраны — там воды практически нет, так что внутримембранные части таких белков должны состоять (и состоят, как мы увидим) из регулярной вторичной структуры, причем размер этих частей ограничен толщиной мембраны.

Посмотрим на нескольких примерах, как устроены мембранные белки. Вообще надо сказать, что мы еще знаем мало структур мембранных белков — их расшифровано пока всего несколько десятков, — так как эти белки плохо растворимы в воде (приходится использовать детергенты и т. д.), и их очень трудно кристаллизовать, так как они склонны слипаться неупорядоченно.

- Структура бактериородопсина — он проводит протон через мембрану — показана на рис. 12-2; эта структура построена на основе анализа множества электронных микрофотографий очень высокого разрешения, так как получить пригодные для рентгеноструктурного анализа трехмерные кристаллы бактериородопсина удалось только относительно недавно.



**Рис. 12-2.** Бактериородопсин в мембране: (а) — вид вдоль мембраны, (б) — вид на мембрану снизу (изнутри клетки). Цилиндрами показаны семь спиралей этого белка. Показаны и соединяющие эти спирали петли, а также сложенная из шариковатомов модель молекулы ретиналя, прикрепленной внутри бактериородопсина

Мы видим, что трансмембранная часть бактериородопсина сложена из семи регулярных  $\alpha$ -спиралей, идущих от одного до другого края мембраны и образующих слегка наклоненный к ее плоскости пучок, а одинокая  $\beta$ -шпилька и все нерегулярные участки цепи (соединяющие спирали петли) выходят из мембраны.

Высокая регулярность укладки остова трансмембранной белковой цепи закономерна: в «жирном», почти безводном липидном окружении цена каждой водородной связи очень высока, что заставляет белковую цепь, уж если она попадает в мембрану, принимать структуры с полностью

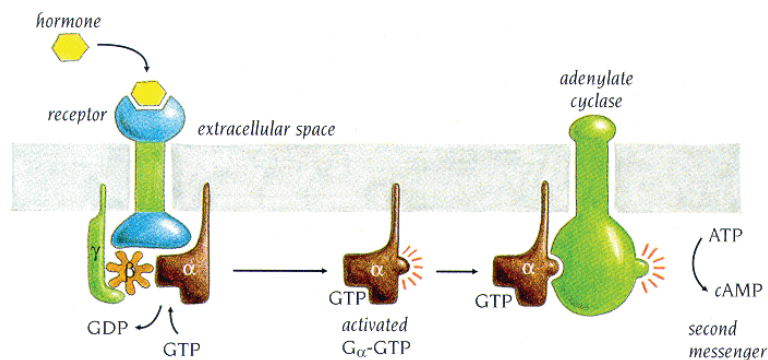
завязанными водородными связями, т. е. либо  $\alpha$ -спираль, либо  $\beta$ -структуру (точнее,  $\beta$ -цилиндр, не оставляющий края  $\beta$ -листа открытыми, см. ниже).

Сидящие на  $\alpha$ -спиралях бактериородопсина гидрофобные группы обращены «наружу», к (тоже гидрофобным) липидам мембраны. Полярные же группы — их немного — обращены внутрь очень узкого канала, по которому идет протон. Протонная проводимость осуществляется при содействии прикрепленной внутри пучка спиралей молекулы кофактора — ретиналя. Он перекрывает центральный канал бактериородопсина.

Поглотив фотон, ретиналь переходит из *trans*- в *cis*-форму. При этом он изгибается (причем конформация белкового тела также несколько меняется) и переносит протон с одной (внутриклеточной, нижней на рис. 12-2) стороны семиспирального бактериородопсинового пучка на противоположную (обращенную к внешней поверхности клетки, где свободных протонов больше, чем внутри). А потом ретиналь разгибается и возвращается назад, но уже без этого протона — за следующим.

- Аналогичные (но обычно более широкие) поры, устроенные как полый пучок спиралей, могут образовываться в других случаях и из отдельных  $\alpha$ -спиральных трансмембранных пептидов.

Кстати, спиральный пучок часто встречается в совсем других мембранных белках. Они не проводят через мембрану никаких молекул. Они проводят сигнал. Речь идет о разных белках-рецепторах, и, в частности, о рецепторах гормонов (рис. 12-3). Работу одного из множества таких рецепторов мы сейчас вкратце рассмотрим.



**Рис. 12-3.** Активация аденилат циклазы (adenylyl cyclase) посредством G-белка, активированного (activated) трансмембранным рецептором (receptor), связавшим гормон (hormone), пришедший из внеклеточного пространства (extracellular space). Превращение АТФ→цикло-АМФ (АТФ→сАМФ) создает внутриклеточный вторичный сигнал (second messenger). Картинка, с разрешения авторов, взята из [5]

При связывании гормона конформация трансмембранного пучка спиралей этого рецептора как-то меняется (еще неясно, как именно). Таким образом, «сигнал» о прибытии гормона передается внутрь клетки: измененная конформация внутримембранной части рецептора как-то содействует тому, что  $\alpha$ -субъединица сидящего на нем *G-белка* (Guanidine-binding protein) отпускает сидящую на ней молекулу ГДФ. Взамен эта субъединица связывает молекулу ГТФ, выловленную из окружающего ее цитозоля, и покидает и прочие свои субъединицы, и рецептор (к которому теперь может подойти другая  $\alpha$ -субъединица G-белка, нагруженная ГДФ, и тоже, в свою очередь, покинуть его, прихватив ГТФ).

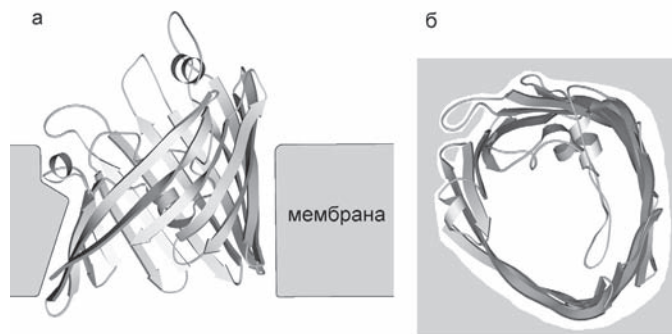
$\alpha$ -субъединица G-белка может расщепить ГТФ (GTP, гуанидин-трифосфат), превратив его в ГДФ (GDP, гуанидин-ди-фосфат) и фосфат (P), но (и это важно!) она делает это очень медленно. А пока, вместе с ГТФ, она дрейфует по мембране (в которую уходит его «хвост»), подплывает к аденилат циклазе, связывается с ней, и, в результате, аденилат циклаза начинает работать, превращая множество молекул АТФ в цикло-АМФ. Это и дает начало физиологическому ответу клетки на связывание гормона. Но действие  $\alpha$ -субъединицы G-белка на аденилат циклазу длится не вечно; в конце концов  $\alpha$ -субъединица превращает свой ГТФ в ГДФ, действие кончается, и  $\alpha$ -субъединица (с ГДФ) возвращается на свой рецептор. Если он все еще связан с гормоном — цикл повторяется; нет — заканчивается.

В результате сигнал одной молекулы гормона многократно усиливается, но длится только определенное время. В этой конечности времени действия и заключается специфика всех G-белков (и не только тех, что активируют аденилат циклазу), работающих, только пока они не успели расщепить ГТФ (при этом медленное расщепление ГТФ служит своеобразным таймером).

- Посмотрим еще на один трансмембранный белок, порин (рис. 12-4). Он также высоко регулярен и имеет вид широкого цилиндра, сложенного из  $\beta$ -структуры. Обратите внимание, что  $\beta$ -лист здесь образует замкнутый  $\beta$ -цилиндр, это позволяет избежать «незавязанных» водородных связей, типичных для краев плоского  $\beta$ -листа. В цилиндре 16 очень длинных  $\beta$ -участков, а диаметр проходящей по его центру довольно широкой поры — около 15 Å. В сторону поры обращены боковые группы полярных остатков, входящих в  $\beta$ -тяжи, а чередующиеся с ними в  $\beta$ -тяжах неполярные остатки обращены своими боковыми группами в мембрану.

Порин служит проводником полярных молекул — кстати, не очень селективным.





**Рис. 12-4.** Порин. (а) Вид вдоль плоскости мембраны (расположение липидов в ней показано очень схематично). (б) Вид поперек плоскости мембраны

■ Селективность проводимости, т. е. специфичность действия мембранных белковых каналов, частично зависит просто от размера пытающейся идти через канал молекулы: слишком большие – не проходят (а порой плохо проходят и слишком малые, отделенные от стенок канала узким, не вмещающим воду вакуумным зазором, – см. Задачу 12.4).

Селективность во многом связана и с тем, что полярные, а тем более заряженные группы «сами по себе» проникают внутрь мембраны с большим трудом.

Вы помните, что свободная энергия заряда  $q$  в среде с диэлектрической постоянной  $\epsilon$  равна  $+q^2/2\epsilon r$ , где  $r$  — Ван-дер-Ваальсов радиус заряда. Легко прикинуть, что при  $q$ , равном электронному заряду, и  $r = 1,5 \text{ \AA}$  (характерный радиус однозарядного иона) величина  $+q^2/2\epsilon r$  близка к 1,5 ккал/моль при  $\epsilon = 80$  (т. е. в воде), а при  $\epsilon_{\text{мембр.}} = 3$  (т. е. внутри «чистой», сложенной лишь из липида мембраны) величина  $+q^2/2\epsilon r$  близка уже к 37 ккал/моль. Итого — повышение свободной энергии на  $\Delta F = +35$  ккал/моль. Вероятность набрать такую свободную энергию составляет, по формуле Больцмана,  $\exp(-\Delta F/kT) = \exp(-35/0,6) = 10^{-25}$ . Значит, успешной будет лишь одна из  $10^{25}$  атак мембраны ионом. А так как «атака» длится не менее  $10^{-13}$  с (это — время теплового колебания, как мы помним), то для проникновения иона через чисто липидную мембрану нужно по крайней мере порядка  $10^{12}$  секунд, или порядка десятка тысяч лет... Все это делает чисто липидную мембрану практически непроницаемой для ионов.

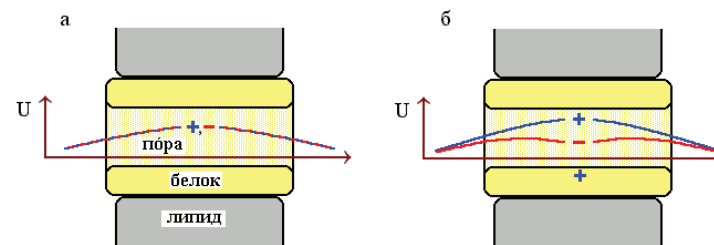
Другое дело, если в мембрану встроен белок, а внутри него образован более или менее широкий, заполненный водой канал, по которому ион может идти, не покидая среду с высокой диэлектрической проницаемостью. Правда, и в канале ион будет чувствовать низкую диэлектрическую прони-

цаемость окружающей канал мембраны, но эффект будет уже много меньше — повышение свободной энергии иона составляет величину порядка  $+q^2/(\epsilon_{\text{мембр.}} \epsilon_{\text{воды}})^{1/2} R$ , где  $R$  — радиус канала,  $\epsilon_{\text{мембр.}} = 3$ , а  $\epsilon_{\text{воды}} = 80$ . Легко прикинуть, что при этом канал радиусом  $R \approx 1,5 \text{ \AA}$  в мембране преодолевается ионом за долю секунды, а  $R \approx 3 \text{ \AA}$  — за малую долю миллисекунды.

Пропускание иона через канал регулируется наличием в нем центров, способных притянуть данный ион и тем уменьшить барьер, который ему приходится преодолевать.

Например, наличие на белке вблизи канала хоть какого-то положительного заряда ускоряет пропускание отрицательно заряженных ионов и резко замедляет пропускание ионов, заряженных положительно (рис. 12-5), а наличие отрицательного заряда — наоборот, ускоряет пропускание положительных ионов и замедляет пропускание отрицательных.

Этот эффект (и, конечно, размер поры) служит основой селективности той проводимости, что обеспечивается мембранными белками.

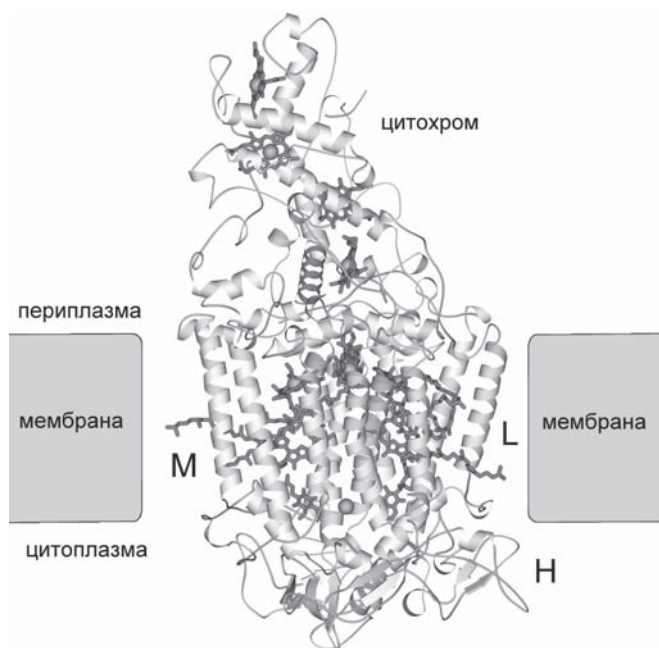


**Рис. 12-5.** Схема трансмембранной поры (мембрана в данном случае расположена вертикально, а пора — горизонтально) и электростатическая свободная энергия  $U$  для положительно (— + —) и отрицательно (— - —) заряженных ионов. (а) На внутренней поверхности поры нет заряда. (б) На внутренней поверхности поры — положительный заряд

■ Теперь давайте более подробно рассмотрим фотосинтетический реакционный центр (рис. 12-6). Его задача — проводить выбитые светом электроны на другую сторону мембраны, и тем самым создавать тот электрический потенциал, на котором и работает фотосинтез.

Фотосинтетический реакционный центр включает в себя цитохром с четырьмя гемами (это, собственно говоря, — не мембранный белок: он находится вне мембраны, в периплазматическом пространстве) и три мембранные субъединицы: L, M и H (от последней, впрочем, в мембране находится только одна  $\alpha$ -спираль). Субъединицы L и M очень похожи.

Все трансмембранные части —  $\alpha$ -спирали. Они, как всегда, длинны (их длина равна толщине мембраны) и регулярны. Внутри мембраны нет



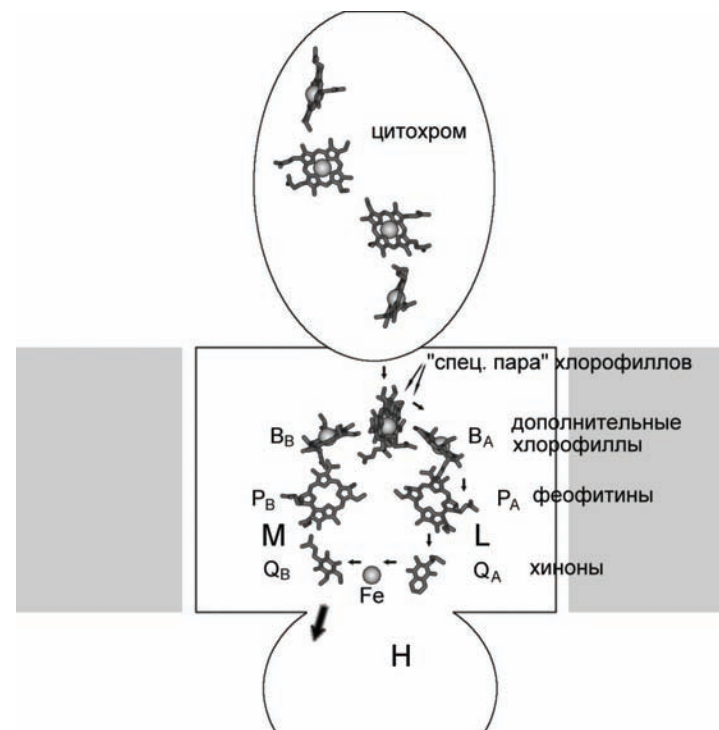
**Рис. 12-6.** Фотосинтетический реакционный центр. Мембрана показана схематично. Трансмембранная часть белка состоит из спиралей субъединиц L и M, есть одна трансмембранная спираль и у субъединицы H. Цитохром целиком находится в периплазме, вне мембраны. Обратите внимание, насколько внутримембранные части белков регулярнее внешне мембранных! Субъединицы L и M связывают фотосинтетические пигменты — хлорофиллы и феофитины (их расположение лучше видно на рис. 12-7). И у тех, и у других пигментов есть длинные гидрофобные «хвосты», выступающие из белка в мембрану. Субъединицы L и M связывают, в своей нижней (на рисунке) части, два хинона (см. также рис. 12-7). Цитохром, лежащий вне мембраны, включает четыре гема (изображенные, как и все кофакторы, скелетными моделями). Картинка, с малыми изменениями, взята из [5]

нерегулярных петель. В то же время лежащая вне мембраны цепь значительно менее регулярна и содержит много петель; в общем, она уложена так же, как в «обычных» водорастворимых глобулярных белках, о которых нам еще предстоит говорить.

Обратите внимание на множество относительно небольших циклических молекул — пигментов, сидящих в белке фотосинтетического реакционного центра: именно по сложным из них дорожкам, «проводникам» идет электрон (это видно по изменяющимся, как только на них попадает «проводимый» электрон, электронным спектрам этих пигментных молекул), а сам полипептид служит только, так сказать, формообразующим изолятором.

Как же работает фотосинтетический реакционный центр?

Сначала квант света выбивает электрон из сидящей в белке «специальной пары» («special pair») хлорофиллов (см. схему на рис. 12-7). Этот электрон мгновенно (за  $\sim 10^{-12}$  с) перескакивает [видимо, через соседний «дополнительный» (accessory) хлорофилл  $B_A$ ] на феофитин  $P_A$  (именно  $P_A$ , а не на  $P_B$ !), а затем, за  $\sim 10^{-10}$  с, — на хинон (quinone)  $Q_A$ . Потом — уже за  $\sim 10^{-4}$  с — электрон переходит на хинон  $Q_B$ . Почему электрон идет столь круглым путем на хинон  $Q_B$ , почему он не регистрируется (по электронным спектрам) на «дополнительном» хлорофилле  $B_A$  — еще не ясно.



**Рис. 12-7.** Схема расположения фотосинтетических пигментов в фотосинтетическом реакционном центре. Ось псевдосимметрии L и M субъединиц проходит через «специальную пару» хлорофиллов и ион Fe. Длинные «хвосты» пигментов «сбиты», чтобы не загромождать рисунок. Электрон преимущественно идет по правой (на рисунке) ветви пигментов, привязанной к субъединице L. Путь электрона отмечен короткими стрелками. Большая стрелка отмечает выход хинона, нагруженного двумя последовательно прибывшими электронами. Левая цепь не используется. Предполагается, что она использовалась в прошлом, а теперь осталась, как аппендикс. Картинка, с малыми изменениями, взята из [5]

Ушедший со «специальной пары» хлорофиллов электрон возмещается электроном, пришедшим с гема цитохрома. На этом завершается первый полуцикл реакции.

В результате второго аналогичного полуцикла хинон  $Q_B$  нагружается вторым электроном и уходит с обоими электронами (дважды заряженный, он легче вырывается из мембраны), чтобы принять участие в дальнейшем фотосинтезе.

■ Итак, фотосинтетический реакционный центр переносит электроны из верхнего (на рис. 12-6, 12-7) в нижний компартмент — против создающейся, в результате, разности потенциалов между компартментами. Коэффициент полезного действия этого реакционного центра около 50% (т. е. 50% поглощенного света превращается в энергию разделенных зарядов: вполне неплохо!).

Отметим два важных физических момента:

1) И хлорофиллы (рис. 12-8), и феофитины, и хиноны, и другие пигменты включают в себя системы «полупроводников» (...-C=C-C=...), т. е. p-электронных связей. Иначе говоря, электронные облака в каждой из этих молекул обобществлены (вследствие Полингова резонанса: ...C=C-...  $\leftrightarrow$  ...C-C=...), и электроны бегают по такой молекуле, как по кусочку металла. Тем самым создается потенциальная яма для электронов, где они «делокализуются», т. е. могут смещаться на расстояния, значительно превышающие диаметр атома (кстати, именно делокализация электронов создает характерную окрашенность пигментов: локализованный в отдельной валентной связи электрон возбуждается коротковолновым ультрафиолетовым светом, а делокализованный электрон — «обычным», видимым, более длинноволновым светом).

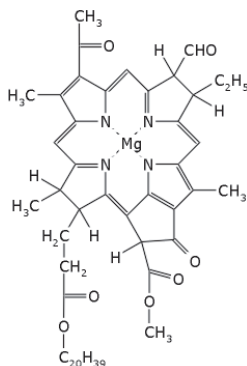


Рис. 12-8. Молекула бактериохлорофилла

2) Передача электрона с одного «кусочка металла» (пигмента) на другой не требует непосредственного контакта этих пигментов. Эта передача осуществляется путем квантового *туннельного перехода* (см. схему на рис. 12-9).

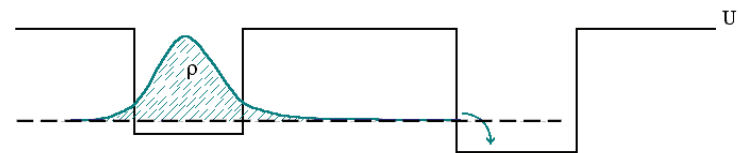


Рис. 12-9. Схема, поясняющая туннельный эффект. Жирная линия показывает профиль потенциальной энергии  $U$  электрона. Пунктирная линия — уровень полной (потенциальной + кинетической) энергии электрона. Колоколообразная линия (со штриховкой под ней) — плотность электронного облака  $\rho$ . Электрон первоначально сидит в левой яме (это его состояние и показано на рисунке), но край его облака (пусть и обладающего здесь очень малой плотностью) достигает правой ямы, так что электрон может со временем весь перетечь туда, если там энергия ниже. При глубине ямы в несколько электрон-вольт (характерной энергии, необходимой для ионизации молекулы, ~100 ккал/моль) — характерное расстояние, на котором плотность его облака спадает на порядок, — около 1 Å

Суть туннельного перехода (или, как говорят, «подбарьерного» перехода: ведь при этом электрон проходит как бы *под* энергетическим барьером) в том, что, по законам квантовой механики, электрон (как и всякая частица, а в особенности — легкая частица) несколько «выступает» за пределы той потенциальной ямы, в которой он находится. «Потенциальной ямой» (областью низкой потенциальной энергии  $U$ ) в данном случае служит та молекула (хлорофилл, феофитин и т. д.), на которой сидит электрон. Вне «ямы», в которой сидит этот электрон (рис. 12-9), его потенциальная энергия выше, чем его полная (потенциальная + кинетическая) энергия в яме. Если бы не квантовый эффект, этот дефицит энергии не давал бы электрону возможности высовываться из ямы. Квантовый же эффект приводит к тому, что волновая функция электрона (или, попросту, плотность электронного облака) «выплескивается» за потенциальную яму, хотя величина этой плотности и очень быстро, экспоненциально спадает по мере удаления от «ямы».

Это — тот же квантовый эффект, что не позволяет электрону упасть на ядро: хоть это падение и уменьшило бы потенциальную энергию электрона, его кинетическая энергия возросла бы при этом еще больше. Дело здесь в том, что, при стремлении к нулю расстояния  $\Delta x$  от электрона до ядра, потенциальная энергия электростатического взаимодействия электрона с ядром стремится к минус бесконечности как  $1/\Delta x$ , а его кинетиче-

ская энергия, в силу принципа неопределенности Гейзенберга стремится (при  $\Delta x \rightarrow 0$ ) к плюс бесконечности куда быстрее, — как  $1/(\Delta x)^2$ .

Действительно, согласно принципу Гейзенберга, неопределенность в скорости  $\Delta v$  и неопределенность в координате  $\Delta x$  частицы связаны соотношением  $m\Delta v\Delta x \sim \hbar$ , где  $\hbar$  — постоянная Планка, а  $m$  — масса частицы. Иными словами, абсолютная величина скорости частицы  $v$  в яме шириной  $\Delta x$  составляет порядка  $\hbar/(m\Delta x)$  при полной неопределенности направления движения частицы в данный момент времени. Следовательно, кинетическая энергия частицы  $E = mv^2/2$  есть величина порядка  $m[\hbar/(m\Delta x)]^2 = (\hbar^2/m)/\Delta x^2$ .

Так же обстоит дело и в случае электрона в потенциальной яме: если бы он ни на йоту не выступал за свою яму, его полная энергия была бы выше.

Поэтому электронное облако слегка «выплескивается» за потенциальную яму. При этом плотность его спадает экспоненциально, как и плотность электронного облака атома. Характерное расстояние, на котором плотность этого облака спадает на порядок (в 10 раз), — около  $1 \text{ \AA}$  (это, как мы знаем, — характерный радиус атома).

Уточнение. Квантово-механический расчет показывает, что характерное расстояние  $\lambda$ , на котором плотность электронного облака спадает в 10 раз, составляет  $\ln 10 \cdot \hbar / \sqrt{8m \cdot \Delta E}$  (где  $m$  — масса частицы, т.е. электрона, а  $\Delta E > 0$  — разность между энергией барьера и уровнем энергии частицы в яме), — что, при типичном значении  $\Delta E \approx 5$  электрон-вольт ( $\approx 120$  ккал/моль), приводит к  $\lambda = 1 \text{ \AA}$ .

Значит, плотность электронного облака спадает в 1000 раз на расстоянии в  $3 \text{ \AA}$  от «своей ямы». Это означает, что вероятность того, что за одно колебание электрон отойдет от «своей ямы» на  $3 \text{ \AA}$ , — порядка  $10^{-3}$  (а на  $5 \text{ \AA}$ , — порядка  $10^{-5}$ , на  $10 \text{ \AA}$ , — порядка  $10^{-10}$  и т. д.). Электрон в «яме» (молекуле пигмента) совершает  $\sim 10^{15}$  колебаний в секунду (в видимом световом диапазоне частот: это хорошо видно по спектрам поглощения таких молекул). Значит, характерное время его перехода в другую «яму» (другую молекулу пигмента), отстоящую на  $3 \text{ \AA}$ , — порядка  $10^{-15} / 10^{-3} = 10^{-12}$  с; в яму, отстоящую на  $5 \text{ \AA}$ , — порядка  $10^{-10}$  с; в яму, отстоящую на  $10 \text{ \AA}$ , —  $10^{-5}$  с; а в яму, отстоящую на  $15 \text{ \AA}$ , — порядка 1 с. Такая зависимость скоростей переходов от расстояний согласуется с тем, что наблюдается в фотосинтетическом реакционном центре.

■ Обратите внимание на ряд существенных моментов.

Первое. Суммарный переход совершается на расстояние около  $40 \text{ \AA}$ . Такое расстояние нельзя покрыть одним туннельным прыжком (такой прыжок занял бы  $\sim 10^{-15} / 10^{-40} \sim 10^{25}$  с  $\sim 10^{17}$  лет — на него не хватило бы

всего времени жизни Вселенной). Однако конструкция белка разбивает этот большой прыжок на четыре малых — от одного притягивающего электрон пигмента к другому — и электрон покрывает эти  $40 \text{ \AA}$  за долю миллисекунды.

Второе. Для того, чтобы электрон не вернулся тотчас со второго пигмента на первый, а пошел дальше, на третий и т. д. — его суммарная (потенциальная + кинетическая) энергия должна падать по ходу процесса — или, иными словами, электрон должен на каждом шаге переходить с высокоэнергетической атомной орбиты на низкоэнергетическую. Конструкция фотосинтетического реакционного центра призвана обеспечить такое понижение энергии электрона на пигментах по ходу процесса.

Третье. Совершая туннельный переход, электрон не расходует энергии на преодоление барьера (здесь нет никакого «трения»). Понижение энергии осуществляется за счет электронно-конформационного взаимодействия (ЭКВ). А именно, когда электрон переходит на следующий пигмент, он застаёт его в конформации, соответствующей энергетическому минимуму без этого нового электрона. Минимуму же энергии в присутствии «нового» электрона отвечает другая, несколько деформированная по отношению к первой, конформация (т.е. расположение ядер атомов) пигмента. Спускаясь в эту новую конформацию, атомы пигмента трутся об окружение, и избыточная энергия диссипирует. В результате электрон на каждом шаге идет от более высокоэнергетической атомной орбиты к более низкоэнергетической, его энергия понижается и она расходуется на то, чтобы сделать туннельный переход «эффективным», т.е. необратимым.

И последнее. Туннельный (или, как говорят, «под-барьерный») переход можно отличить от обычного активационного механизма перехода через энергетический барьер по тому, что на скорость туннельного перехода температура существенным образом не влияет (и потому этот переход не исчезает при низкой температуре), а скорость активационного перехода (пропорциональная  $\exp(-\Delta E^\ddagger/k_B T)$ , где  $\Delta E^\ddagger$  — энергия активационного барьера, а  $T$  — температура) резко уменьшается при падении температуры.



## Лекция 13

Глобулярные белки. Упрощенное представление структур белковых глобул; структурные классы. Строение  $\beta$ -белков:  $\beta$ -слои, их продольная и перпендикулярная упаковка. Преимущественная антипараллельность  $\beta$ -структуры в  $\beta$ -белках. Правопропеллерная скрученность  $\beta$ -листов. Топология  $\beta$ -белков.

Теперь в фокусе нашего внимания будут находиться глобулярные, а точнее — водорастворимые глобулярные белки. Именно они наиболее исследованы: для сотен — изучена спонтанная самоорганизация, для тысяч — расшифрована атомная трехмерная структура. Поэтому именно их обычно имеют в виду, говоря о «типичных белковых структурах», о «закономерностях, наблюдаемых в строении и самоорганизации белков» и т. д. После этой необходимой оговорки обратимся к структурам глобулярных белков.

- Рентгеноструктурные исследования (а позже — двух- и многомерный ЯМР) позволили за более чем полвека интенсивной работы сотен лабораторий установить атомную структуру порядка 10 000 белков (а если считать и мутантные формы, и различные функциональные состояния, то раз в десять больше).

**Внутренний голос.** А совпадает ли та структура, что видит рентген — в кристалле! — со структурой белка в растворе?

**Лектор.** Как правило, практически совпадает. Это подтверждается тремя группами данных. Во-первых, часто можно показать, что белок в кристалле сохраняет свою активность. Во-вторых, иногда можно получить один и тот же белок в разных кристаллических формах (рис. Д-17). При этом его строение остается практически неизменным. И, наконец, ЯМР белка в растворе (правда, он возможен пока только для не очень больших белков) показывает ту же структуру, что и рентген в кристалле (ср. рис. Д-18 с Д-17). Однако надо оговориться: иногда в белке есть гибкие участки (некоторые боковые группы, петли, а также междоменные перемиčky

в больших белках), и вот их структура может меняться при кристаллизации или наводиться ею. Но это относится либо к мелким деталям белковых структур, либо к смещению доменов, но не к устройству практически твердых доменов и небольших белков.

Надо добавить, что рентген видит не только «статическую», усредненную структуру белка, которую мы будем рассматривать сегодня и на ближайших лекциях, но и тепловые флуктуации атомов белка, которые будут вкратце обсуждаться позже.

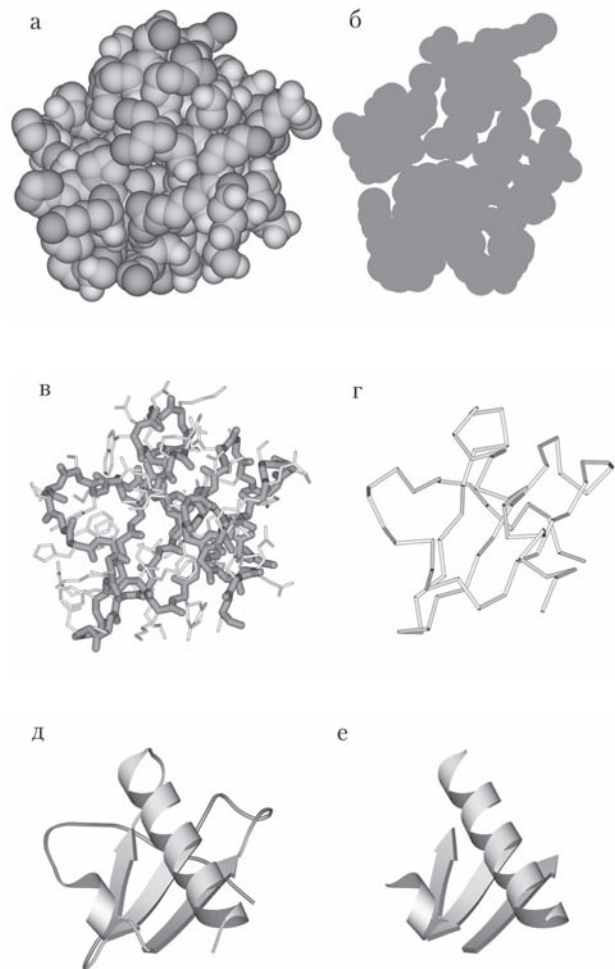
- Итак, что можно увидеть в структурах белков — водорастворимых глобулярных белков — с птичьего полета?

Мы видим, что небольшие (из 50–150, реже из 200–250 остатков) цепи укладываются в компактную глобулу диаметром 25–40 Å (рис. 13-1); и что более крупные белки состоят из нескольких таких субглобул — «доменов» (рис. 13-2). Белковая цепь упаковывается в глобулу так же плотно, как органические молекулы — в кристалл. Это видно и при взгляде на поверхность белка (рис. 13-1а), и на «срезе» белковой глобулы, показанном на рис. 13-1б. Однако, рассматривая белок, нам предстоит рассматривать не плотноупакованные электронные облака (или Ван-дер-Ваальсовы поверхности) атомов — иначе ничего не будет видно внутри белка — а только очищенные от «мяса» атомов (рис. 13-1в) и даже от боковых групп вообще (рис. 13-1г) скелеты белковых молекул. Поэтому не поддавайтесь часто создаваемому рисунками ощущению «рыхлости» белковых глобул!

Каркас пространственной структуры подавляющего большинства глобул (доменов) сложен из уже изученных нами регулярных вторичных структур —  $\alpha$ -спиралей и  $\beta$ -листов (рис. 13-1д), которые стабилизированы регулярными водородными связями в регулярной главной цепи. В глобулярных белках на долю  $\alpha$ - и  $\beta$ -структур в сумме обычно приходится 50–70% остатков цепи. Кстати, эти вторичные структуры были теоретически предсказаны Полингом, Кори и Брэнсоном еще до расшифровки атомного строения белковых молекул. «Штабель», сложенный из этих структур (рис. 13-1е), определяет основные особенности строения белка.

Каркас из  $\alpha$ - и  $\beta$ -структур как бы окружает гидрофобное ядро (или ядра) белка, а нерегулярные петли лежат дальше, на периферии глобулы. Петли практически никогда не входят внутрь белковой глобулы — и неудивительно: их не вовлеченные во вторичную структуру пептидные группы не должны порывать своих водородных связей с водой, это вело бы к нестабильности глобулы. Кстати: водородные связи петель (а также — торцов  $\alpha$ -спиралей и краев  $\beta$ -листов) с молекулами воды часто обнаруживаются рентгеном в белковых молекулах.

В зависимости от строения каркаса глобулярные белки подразделяются на: «чистые»  $\beta$ -белки; «чистые»  $\alpha$ -белки; и «смешанные»  $\alpha/\beta$  и  $\alpha + \beta$ -белки.



**Рис. 13-1.** Картинки, изображающие строение небольшого белка —  $\alpha$ -субъединицы интерлейкина 8 — при разной степени схематизации. (а) Атомная модель (изображены только «тяжелые», не-водородные атомы); в силу того, что цепь в глобуле упакована плотно, мы видим лишь поверхность белка. (б) Срез атомной модели подчеркивает плотность упаковки. (в) Скелетная модель главной цепи (темная линия) и боковых групп (более светлые отростки). (г) Ход главной цепи. (д) Схема строения белка, на которой выделены вторичные структуры (две  $\alpha$ -спирали и один  $\beta$ -лист из трех  $\beta$ -тяжей) в главной цепи белка. (е) Структурный каркас белка, сложенный из вторичных структур. Все рисунки даны в одной и той же проекции и в одинаковом масштабе

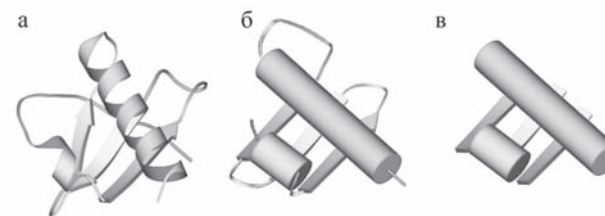
Строго говоря, эта классификация типов белковых структур относится к малым белкам, а также к отдельным доменам (т. е. к компактным субглобулам, из которых сложены большие белки), так как большие белки могут содержать, например, и  $\beta$ -, и  $\alpha$ -домены одновременно.



**Рис. 13-2.** Глобулярные домены в  $\gamma$ -кристаллине

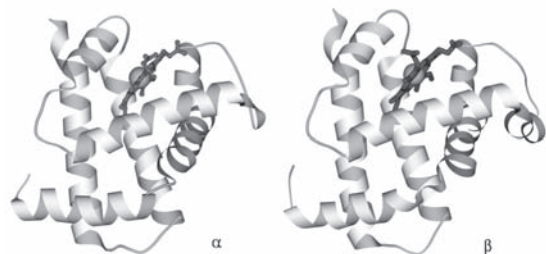
Нас будет прежде всего интересовать: 1) архитектура упаковки  $\alpha$ - и  $\beta$ -структурных сегментов в компактную глобулу (рис. 13-1е); и 2) ход цепи через глобулу (рис. 13-1д) — или, как часто говорят, «топология белковой глобулы».

При этом мы часто будем использовать упрощенные схемы строения белков (рис. 13-3). Такое упрощение достигается не только обращением основного внимания на вторичные структуры (при пренебрежении детальным строением петель), но и пренебрежением различий в размерах этих структур и в деталях их взаимной ориентации [при этом мы переходим от «укладок» (“folds”, рис. 13-3а) к «мотивам укладки» (“folding patterns”, рис. 13-3б) белковых цепей].



**Рис. 13-3.** Упрощенные представления белковых структур. (а) Детальная укладка (“fold”), описывающая размещение вторичных структур в цепи белка и в пространстве (см. также рис. 13-1д). (б) Мотив укладки белковой цепи (“folding pattern”): опущены детали хода петель, размера и точной ориентации  $\alpha$ -спиралей (изображенных в виде параллельных цилиндров) и  $\beta$ -тяжей (изображенных в виде стрелок). (в) Упаковка: штабель (“stack”) структурных сегментов: петли удалены, опущены размер, ориентация и направление  $\alpha$ -спиралей и  $\beta$ -тяжей (изображенных поэтому в виде лент)

Упрощение это оправдывается тем, что детали строения петель и точные размеры и ориентации структурных сегментов (и даже некоторые малые структурные сегменты) не сохраняются, когда мы переходим от рассматриваемого белка к его довольно близкому (имеющему сходную аминокислотную последовательность, т. е. явно общее происхождение) родственнику, например, от гемоглобина  $\alpha$  к гемоглобину  $\beta$  (рис. 13-4).



**Рис. 13-4.** Два близко родственных белка: гемоглобин  $\alpha$  и гемоглобин  $\beta$  лошади (оба — с гемом). Найдите сходства и различия! (Подсказка: при общем высоком сходстве — различия в деталях конформации петель, в деталях ориентации некоторых спиралей, и в одном дополнительном спиральном витке в  $\beta$ -глобине, справа)

На следующем, более высоком уровне упрощения, необходимом для классификации структур белковых глобул, мы иногда вообще будем рассматривать лишь упаковку структурных сегментов в глобулу — т. е. сложенные из вторичных структур упаковки, или «штабели» (“stacks”), — временно забыв о петлях, соединяющих эти вторичные структуры в единую молекулу (рис. 13-3в).

Я буду специально пользоваться такими упрощенными схемами мотивов укладки и упаковки белковой цепи, вырисовывая в то же время, для сравнения, и «истинные» структуры белков так, как их дает компьютер. Казалось бы, зачем нужны упрощенные схемы, если компьютер может нарисовать «все как есть». Дело в том, однако, что «все как есть» содержит очень много деталей, а схемы берут из них главные, повторяющиеся в разных сходных белках. Поэтому они удобны и для классификации белковых структур, и для выделения их главных, типических черт. Рассматривая подробные картинки белков, вы все равно в уме выделяете такие типические черты, а схемы просто позволяют вам систематизировать ваше интуитивное выделение. Кроме того, они дают возможность составить «словесный портрет» белка. На практике такие схемы и такие словесные портреты вам понадобятся сразу же, как только вы захотите выяснить, на кого похож интересующий вас белок: они подчеркивают главное, убирая детали. Конечно, эти «детали»

могут быть самыми важными для функционирования белка, и мы с этим встретимся, но это обстоятельство только подчеркивает относительную независимость функции белка от мотива укладки цепи в нем.

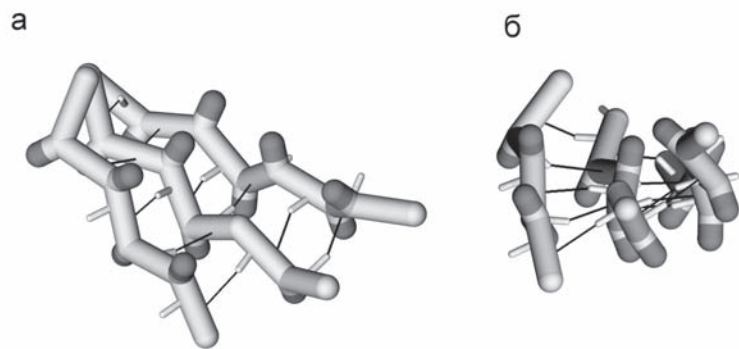
- И «упаковки», и «мотивы укладки» цепи в белковую глобулу сосредотачивают наше внимание не просто на всевозможных (рыхлых, ажурных и т. п.) комплексах из структурных сегментов, а лишь на тех, где они, эти структурные сегменты, уложены плотно. Так мы очерчиваем те области конфигурационного пространства, что соответствуют плотной (но без стерического перекрытия) упаковке белковой цепи в глобулу, т. е. окрестности достаточно глубоких минимумов энергии невалентных взаимодействий. Они дают возможность не только классифицировать уже известные белковые структуры, но и предвидеть новые, еще не найденные в природе.

Здесь уместно сделать пояснение. Когда я буду говорить о классификации белковых структур, об их сходстве и т. д. — я буду иметь в виду не тривиальные вещи типа того, что все глобины похожи друг на друга, независимо от того, работают ли они в человеке или в миноге. Это конечно так, и белки можно классифицировать по филогенетическим семействам, внутри которых функции и, главное, аминокислотные последовательности белков варьируют не слишком сильно. Однако часто сходными пространственными структурами обладают белки, эволюционно никак, по всем тестам, не связанные между собой. И вот это, чисто структурное (а не генетическое) сходство я буду подчеркивать.

- Начнем с изучения строения  $\beta$ -белков.  $\beta$ -структурные домены устроены, пожалуй, проще других: вытянутые участки цепи собраны обычно в два, реже — в несколько  $\beta$ -листов, уложенных друг на друга. Иными словами, «штабеля» вторичной структуры в  $\beta$ -белках выглядят довольно просто.

В  $\beta$ -белках преобладает *антипараллельная*  $\beta$ -структура.

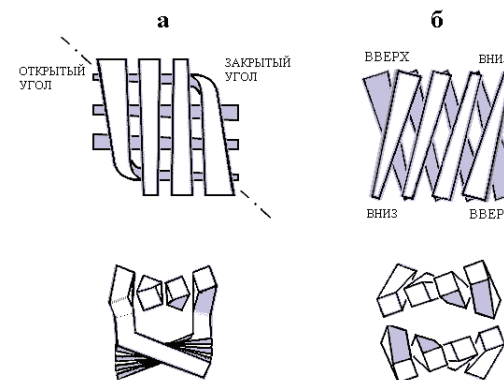
Так как белки собраны из асимметричных L-аминокислот, то вытянутые участки слегка скручены сами по себе: как вы помните, минимум энергии вытянутой конформации лежит на карте Рамачандрана выше диагонали. Витог объединяются, при помощи водородных связей, скрученных  $\beta$ -участков в листы, определенным образом скручены и эти  $\beta$ -листы (рис. 13-5). Их поверхность напоминает пропеллер. Угол между смежными вытянутыми участками  $\beta$ -листа составляет около  $-25^\circ$ . Значит, этот пропеллер выглядит как *левовинтовой*, если смотреть на него поперек хода  $\beta$ -участков (см. рис. 13-5а). И то же скручивание является *правовинтовым*, если смотреть вдоль хода  $\beta$ -участков (см. рис. 13-5б). Так обычно и делают — смотрят *вдоль* хода  $\beta$ -тяжей и говорят, что у  $\beta$ -листа — *правопропеллерная* скрученность.



**Рис. 13-5.** Лист  $\beta$ -структуры, вид поперек (а) и вдоль (б)  $\beta$ -тяжей. Лист имеет складчатую поверхность (она подчеркивается выступающими  $C^\beta$ -атомами) и, обычно, правопропеллерную (если смотреть вдоль тяжей) скрученность. Водородные связи между  $\beta$ -тяжами изображены тонкими линиями

■ Существует два основных типа упаковки  $\beta$ -листов: ортогональная и продольная (рис. 13-6а и 13-6б). В обоих случаях  $\beta$ -листы упакованы «лицом к лицу», и гидрофобное ядро домена заключено между ними, но в первом случае листы повернуты друг относительно друга сильно — на угол около  $90^\circ$  (плюс-минус  $10\text{--}15^\circ$ ), а во втором — слабо, на  $-30^\circ$  (плюс-минус  $10\text{--}15^\circ$ ); углы, лежащие вне этих двух диапазонов (в частности, углы около  $+30^\circ$ ), наблюдаются редко.

При ортогональной упаковке (рис. 13-6а)  $\beta$ -участки скручены и обычно несколько изогнуты, так что общая форма «штабеля» начинает напоминать цилиндр с сильным наклоном  $\beta$ -участков к оси. Такую упаковку часто так и называют:  $\beta$ -цилиндром или  $\beta$ -бочонком, хотя в  $\beta$ -цилиндрах, сложенных из антипараллельной  $\beta$ -структуры (в отличие от  $\beta$ -цилиндров, сложенных из параллельной  $\beta$ -структуры, о которых речь впереди), — сеть водородных связей часто разорвана (или надорвана) на противоположных боках бочонка, что позволяет довольно четко выделить два  $\beta$ -листа. В «закрытом» углу такой упаковки  $\beta$ -участки обоих листов близко сходятся, так что цепь перетекает из листа в лист ценой изгиба и поворота на  $90^\circ$ : можно сказать, что единый  $\beta$ -лист согнут и уложен сам на себя. В противоположных — «открытых» — углах  $\beta$ -листы расходятся, а образующая щель обычно заполняется или  $\alpha$ -спиралью, или нерегулярными петлями, или активным центром, как в ретинол-связывающем белке (рис. 13-7).

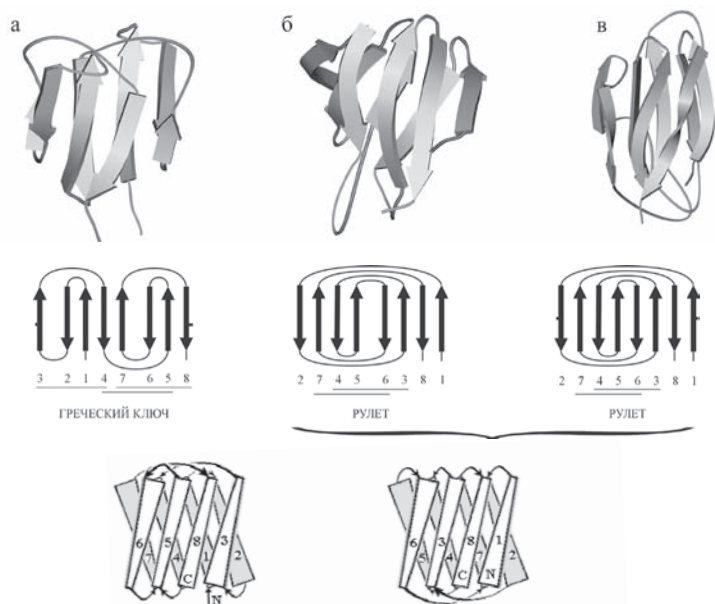


**Рис. 13-6.** Ортогональная (а) и продольная (б) упаковка  $\beta$ -листов. Вид сверху на упаковку  $\beta$ -листов (вверху) и вид с нижнего торца этой упаковки (внизу). На виде сверху — приближение  $\beta$ -тяжа к читателю показано его (тяжа) расширением. Штрих-пунктир — ось ортогонального  $\beta$ -бочонка. На оси находятся оба «открытых» угла этой упаковки. Здесь  $\beta$ -листы наиболее раздвинуты. В двух других, «закрытых» углах листы наиболее сближены; здесь цепь изгибается и переходит из слоя в слой. При ортогональной упаковке гидрофобное ядро имеет приблизительно цилиндрическую форму. В продольной упаковке, наоборот, ядро плоское, расстояние между скрученными листами меняется мало, а поворот одного листа относительно другого позволяет гидрофобным поверхностям лежащих друг над другом скрученных  $\beta$ -тяжей соприкоснуться на большой длине. Картинка, с небольшими изменениями, взята из С. Chothia & А. V. Finkelstein, *Annu. Rev. Biochem.* (1990) 59:1007–1039



**Рис. 13-7.** Ретинол-связывающий белок: пример ортогонально упакованного  $\beta$ -листа. Ход цепи напоминает узор «меандр» (см. топологическую схему — плоскую развертку  $\beta$ -цилиндра — сбоку). На этой схеме  $\beta$ -тяжи изображены стрелками. «Меандр» получается потому, что соседние по цепи  $\beta$ -тяжи являются соседями на поверхности цилиндра; они связаны водородными связями (связь, существующая между крайними  $\beta$ -тяжами 1 и 8, изображена на плоской развертке маленькими черточками). Ретинол-связывающий центр находится у оси цилиндра. Сам ретинол изображен скелетной моделью. Цифры на топологической схеме указывают порядок структурных сегментов в цепи





**Рис. 13-8.** Примеры продольной упаковки  $\beta$ -листов. Мотивы укладки цепи в домене  $\gamma$ -кристаллина (а) (см. также рис. 13-2), в  $\beta$ -домене белка — катаболитического активатора (б) и в белке оболочки сателлитного вируса некроза табака (в). Внизу показаны топологические схемы этих белков. Все эти белки включают характерную топологию «греческого ключа» (или «согнутой пополам  $\beta$ -шпильки»), в которой четыре соседних по цепи  $\beta$ -тяжа антипараллельны, причем первый и четвертый связаны водородными связями. На топологических схемах двухслойный  $\beta$ -сэндвич развернут (как разрезанный сбоку цилиндр) и выложен на плоскость. Место разреза выбрано так, чтобы подчеркнуть симметрию укладки цепи. Так, в  $\gamma$ -кристаллине оно приходится на стык  $\beta$ -тяжей 3 и 8, подчеркивая сходство первой (тяжи 1–4) и второй (тяжи 5–8) половин домена. Сближением  $\beta$ -тяжей на схеме показаны Н-связи между ними; Н-связи между краями развертки (если они есть) показаны маленькими черточками на краях. Увеличенный зазор между  $\beta$ -тяжами отделяет один  $\beta$ -лист сэндвича от другого (если меж ними нет водородных связей). Домен  $\gamma$ -кристаллина содержит повтор, два греческих ключа: из тяжей 1–4 и 5–8; и еще один греческий ключ образован  $\beta$ -тяжами 4–7. Изображенные на рисунках (б, в) белки содержат «ключи» из тяжей 3–6 и 4–7. Более того, они имеют топологию многократно согнутой  $\beta$ -шпильки (она обычно называется «рулет»): здесь Н-связи соединяют также  $\beta$ -тяжи 1 и 8, 2 и 7 — в дополнение к связям тяжей 3 и 6, 4 и 5, типичным для «греческих ключей». Обратите внимание, что топология «греческого ключа» позволяет белковой цепи окружить ядро глобулы лучше, чем позволяет «меандр» (см. рис. 13-7): не только  $\beta$ -структурой с боков, но и петлями сверху и снизу. Обычно продольные упаковки  $\beta$ -листов представляют собой  $\beta$ -сэндвич (а, б), но некоторые из них (например — белка оболочки сателлитного вируса некроза табака и белков оболочки ряда других вирусов) можно представить также в форме  $\beta$ -цилиндров с коллинеарными  $\beta$ -тяжами (в)

■ Продольная упаковка (рис. 13-бб) характерна для листов, которые лишь скручены, как пропеллер, но не изогнуты дугой. Такую упаковку обычно называют  $\beta$ -сэндвичем. Торцы  $\beta$ -сэндвича прикрыты нерегулярными петлями (вы это видите на рис. 13-8). В некоторых  $\beta$ -сэндвичах крайние тяжи  $\beta$ -листов столь сближены (иногда они даже соединены несколькими водородными связями), что эта упаковка принимает форму цилиндра с малым наклоном  $\beta$ -участков к оси.

Необходимо подчеркнуть, что мотивов укладки белковой цепи имеется гораздо меньше, чем известных белковых глобул, а типов «штабелей» — гораздо меньше, чем мотивов укладки белковой цепи. Одному и тому же «штабелю», т. е. одной и той же упаковке структурных сегментов могут соответствовать разные мотивы укладки цепи в глобулу, т. е. эти сегменты могут быть по-разному связаны единой полипептидной цепью. Рисунок 13-8 показывает, как один и тот же  $\beta$ -сэндвич (при различной топологии, т. е. при различном ходе цепи через этот сэндвич) служит основой структуры трех разных белковых доменов — домена  $\gamma$ -кристаллина (а),  $\beta$ -домена белка — катаболитического активатора (б), а также белка вирусной оболочки (в). Причем два последних белка имеют даже одинаковый мотив укладки цепи — т. е. в них цепь одинаково проходит через одинаковый  $\beta$ -структурный штабель (это подчеркивается горизонтальной фигурной скобкой на рисунке 13-8)!

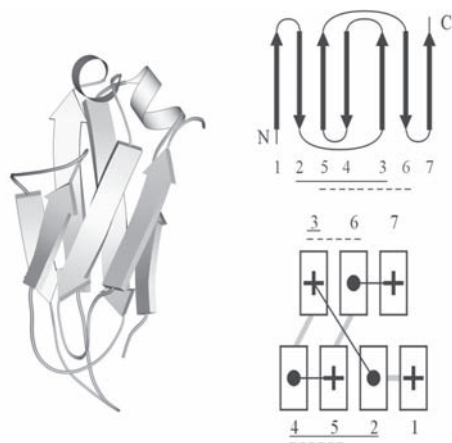
■ Примеры такого сходства структур при отсутствии какой-либо другой видимой связи между белками очень многочисленны.

Здесь я не могу удержаться от искушения показать вам еще один белок, в основе которого лежит  $\beta$ -сэндвич. Это — иммуноглобулин, точнее — так (или примерно так) устроен каждый домен этого большого белка (рис. 13-9).

Подобным же образом уложены цепи белков примерно пятидесяти (!) других суперсемейств (хотя часть из них и отвечает, как иммуноглобулин, за специфическое связывание каких-то объектов, например, при межклеточном узнавании). В некоторых из этих суперсемейств, правда, наблюдаются небольшие отличия от «стандартной» структуры, изображенной на рис. 13-9.

Так, в одних  $\beta$ -тяж 1 образует параллельную  $\beta$ -структуру с тяжом 7 (и даже может потерять связь с тяжом 2); в других —  $\beta$ -тяж 4 так же переползает к тяжу 3 от  $\beta$ -тяжа 5; в третьих — в перемычке между  $\beta$ -тяжами 3 и 4 образуется добавочная  $\beta$ -шпилька. Но остов укладки, состоящий из  $\beta$ -тяжей 2, 3, 5, 6, 7, остается неизменным.

Помимо удовольствия показать вам столь популярный мотив укладки цепи, я преследую еще одну цель: показать, что мотив строения белка с более или менее коллинеарной упаковкой структурных сегментов проще всего отражается схемой, представляющей вид с торца на такой белок.

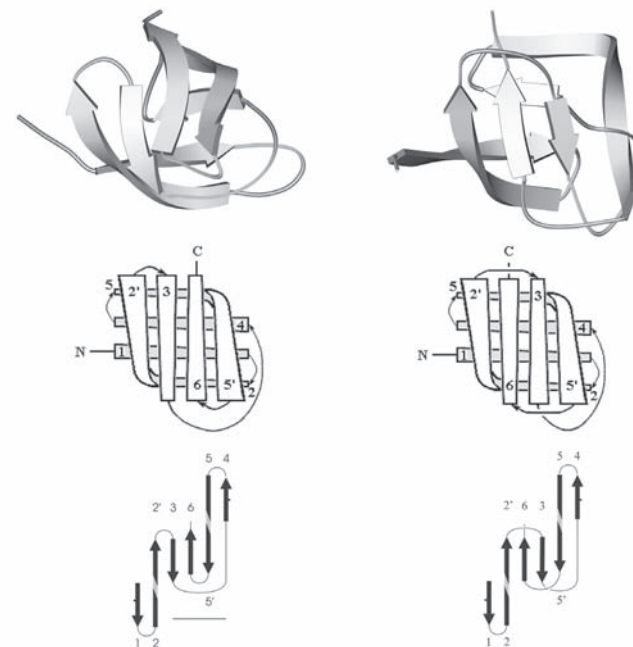


**Рис. 13-9.** Продольная упаковка  $\beta$ -листов в константном домене легкой цепи иммуноглобулина  $\kappa$ . Слева представлена подробная картина белка. На топологической схеме (справа сверху) подчеркнуты «греческие ключи». На рисунке справа внизу дан вид на белок «снизу» (с торцов структурных сегментов). Прямоугольники — торцы  $\beta$ -тяжей. Крестик соответствует N-концу сегмента (т. е. он «идет от нас»), точка — его C-концу (т. е. он «идет к нам»). Ход петель, соединяющих структурные сегменты, показан черной линией, если петля обращена к нам, и светлой, если она находится на противоположной стороне укладки. Обратите внимание, что такая схема позволяет представить коллинеарную упаковку этих сегментов ( $\beta$ -тяжей) наиболее просто. Кроме того, она дает возможность увидеть пространственное строение «греческих ключей» и заметить, что два имеющихся «ключа» (из тяжей 2, 3, 4, 5 и 3, 4, 5, 6, соответственно) в пространстве организованы по-разному. Структура, образованная двумя перекрывающимися (так, как на этом рисунке) греческими ключами, иногда называется «полным греческим ключом»

Вернемся к ортогональной упаковке  $\beta$ -листов. Рисунок 13-10 показывает, как один и тот же  $\beta$ -цилиндр (при различной топологии, при различном ходе цепи через ортогональную укладку  $\beta$ -листов) служит основой и сериновой протеазы типа трипсина или химотрипсина (а), и протеазы кислой, типа пепсина (б).

- Выше мы рассмотрели наиболее показательные «базовые» конструкции  $\beta$ -белков. Есть и другие «базовые конструкции» — например, «многозаходный пропеллер».

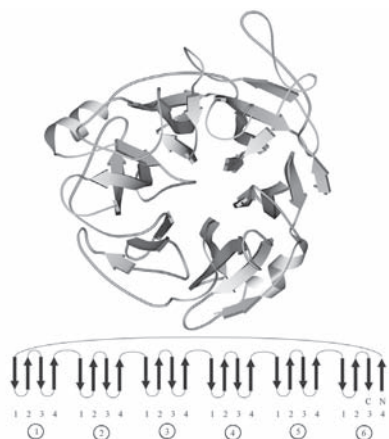
В «пропеллере» нейраминидазы (рис. 13-11) шесть наклонно уложенных  $\beta$ -листов образуют розетку (в других белках бывает до восьми  $\beta$ -листов). Если рассматривать эти листы попарно, они образуют  $\beta$ -сэндвичи, так что «пропеллер» часто описывают как суперцилиндр, сложенный из  $\beta$ -сэндвичей.



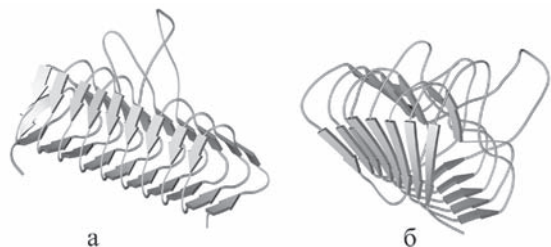
**Рис. 13-10.** Мотив укладки цепи в сериновой протеазе типа химотрипсина (слева) и в кислой протеазе (справа). В последнем белке петли укорочены и изображены довольно схематично. Показаны схемы ортогональной упаковки  $\beta$ -листов в этих белках, а также топологические схемы  $\beta$ -листов. Эти листы согнуты (при изгибе края листов уходят от читателя и смыкаются; возникающая при этом водородная связь краев  $\beta$ -листа изображена черточками); места перегибов  $\beta$ -тяжей соответствуют их более светлым частям. «Греческий ключ» в трипсине подчеркнут

Во «вмятине» на оси этого суперцилиндра — вы видите, что она не прикрыта петлями, — находится активный центр. В ретинол-связывающем белке мы уже встречались с таким расположением активного центра — во вмятине в центре цилиндра, и мы встретимся с ним еще.

- Интересна также, прежде всего своей регулярностью, конструкция типа « $\beta$ -призмы» или (другое название) « $\beta$ -спирали» (рис. 13-12). Три грани такой призмы образованы тремя  $\beta$ -листами — причем параллельными! — а цепь проходит сквозь них по спирали, непрерывно переходя из одного листа в другой. При этом она как бы навивается на ось призмы, образуя либо обычную при соединении параллельных  $\beta$ -тяжей правую спираль, либо (в других призмах) спираль левую, исключительно редкую при соединении  $\beta$ -тяжей.



**Рис. 13-11.**  $\beta$ -структура в форме «шестилопастного пропеллера» в нейраминидазе. Внизу — схема топологии этого белка, состоящего из шести антипараллельных  $\beta$ -листов. Картинка, с некоторыми изменениями, взята из [5]



**Рис. 13-12.**  $\beta$ -призма в ацилтрансферазе (а) и в пектатлиазе С (б). Обратите внимание на разную закрученность цепи вокруг длинной оси призмы: необычную, *левую* на рисунке (а) и обычную, *правую* — на рисунке (б), а также на то, что при левой закрученности цепи отсутствует обычная (правопропеллерная, если смотреть вдоль  $\beta$ -тяжей, см. рис. 13-12б и рис. 13-5–13-11) скрученность  $\beta$ -листа

Достоин внимания, что в некоторых примитивных рыбах (например, миногах) роль иммуноглобулинов исполняют белки, построенные не на классической иммуноглобулиновой укладке (см. рис. 13-9), а на  $\beta$ -призмах.

■ Теперь уместно поговорить о топологии  $\beta$ -белков. Прежде всего обращает на себя внимание то, что  $\beta$ -белки сложены преимущественно из антипараллельной  $\beta$ -структуры. Большинство рассмотренных нами до сих пор  $\beta$ -белков состояло из чистой антипараллельной  $\beta$ -структуры. Порой к ним примешивается небольшое количество параллельной (см. рис. 13-10б). Очень редко, но бывают белки, сложенные из чисто параллельной  $\beta$ -структуры (см. рис. 13-12).

То, что «примесь» параллельной к антипараллельной  $\beta$ -структуре довольно мала, в общем, не удивительно, так как они имеют несколько разные конформации, так что их стык должен быть энергетически невыгодным. В какой мере энергетическая невыгодность структуры сочетается с относительной редкостью ее встречаемости, об этом мы поговорим на одной из ближайших лекций; однако в целом ясно, что стабильная система — а белок стабилен — должна состоять в основном из стабильных же элементов и избегать внутренне нестабильных.

Преимущественно же *антипараллельный* характер  $\beta$ -листов в  $\beta$ -белках тесно связан с тем, что их архитектура обычно основана на  $\beta$ -шпильках (рис. 13-13). Такие шпильки часто бывают согнуты пополам, а иногда — согнуты дважды или даже трижды (как на рис. 13-8б, в).



**Рис. 13-13.** Антипараллельные  $\beta$ -шпильки. Согнутые шпильки характерны для краев  $\beta$ -белков, зацепленные пары — для середин этих белков

Петли, соединяющие  $\beta$ -участки, обычно имеют вход и выход на одном и том же краю укладки (т. е. они не пересекают «штабель», а прикрывают его торец). Это видно почти на всех рисунках. При этом петли, даже если они длинные, чаще всего соединяют близкие между собой в пространстве концы  $\beta$ -участков. Поэтому соседние по цепи  $\beta$ -участки, как правило, *не* параллельны и часто образуют антипараллельные  $\beta$ -шпильки.

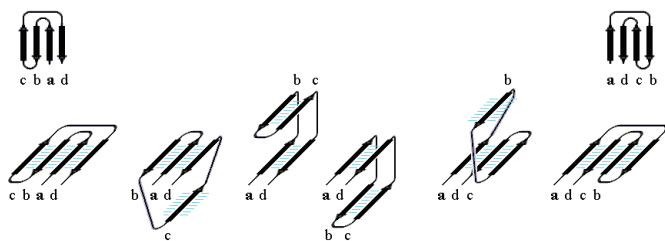


**Рис. 13-14.** Возможные топологии листов из четырех  $\beta$ -участков. Показаны только листы, где каждые два подряд (в цепи) идущих  $\beta$ -участка направлены в противоположные стороны. Из этих топологий часто встречаются только «меандр» (подчеркнут одной чертой) и два (подчеркнуты двойной чертой) «греческих ключа» (они различаются тем, лежит ли тяж 1 внутри «ключа» или на его краю). Пример белка с меандром — ретинол-связывающий белок (см. рис. 13-7); примеры белков с «греческим ключом» — изображенные на рис. 13-8–13-10 (причем в одном и том же белке часто есть «ключи» обоих типов). Обратите внимание, что редко встречаются все топологии со «смешанной» (анти- и параллельной)  $\beta$ -структурой и все топологии с перекрещиванием петель

Отметим также, что «наложение» петли на петлю (или, как говорят, «пересечение петель») встречается редко (такое исключение из правил показано на рис. 13-10б, снизу), видимо, потому что при этом петли должны дополнительно изгибаться (чтобы избежать столкновения или дегидратации), что опять-таки энергетически невыгодно. Избегание наложения петель — общее структурное правило, наблюдающееся в белках.

В результате, из многих изображенных на рис. 13-14 возможных конфигураций  $\beta$ -листа, сложенного из единого куска цепи, по-настоящему часто встречаются лишь три мотива: «меандр» и два «греческих ключа» (они подчеркнуты на рисунке) — именно эти мотивы свободны от всех отмеченных выше недостатков.

«Меандр» — это, кстати, имя очень извилистой реки в Греции — характерен тем, что соседние по цепи  $\beta$ -тяжи являются также ближайшими соседями в пространстве (см. рис. 13-8); обычно (но не всегда) они связаны водородными связями. «Греческий ключ» (такой орнамент можно увидеть на старых вазах и садовых решетках) характерен тем, что четыре соседних по цепи  $\beta$ -тяжа антипараллельны, причем первый и четвертый связаны водородными связями. При этом на самом деле второй и/или третий  $\beta$ -тяжи «ключей» часто лежат не в том же (как может показаться из рис. 13-14), а в другом  $\beta$ -листе. При этом в пространстве образуются различные структуры — так называемые «*abcd*» структуры Ефимова — с одной и той же топологией греческого ключа (рис. 13-15). Найдите их на рисунках 13-8, 13-10.



**Рис. 13-15.** Различные пространственные «*abcd*» структуры Ефимова с топологией греческого ключа. Обратите внимание на *правый ход суперспирали*, состоящей из двух параллельных  $\beta$ -тяжей одного  $\beta$ -листа и лежащего между ними в цепи  $\beta$ -тяжа другого листа (т. е. на сложенную из  $\beta$ -тяжей суперспираль *b-c-d* на второй слева внизу и на суперспираль *a-b-c* на второй справа внизу схемах). Именно такое, «правовинтовое», соединение параллельных  $\beta$ -тяжей одного  $\beta$ -листа типично для белков; обратное, «левовинтовое», встречается очень редко

- Все такие характерные, часто встречающиеся в белках структуры (шпильки, меандр, греческий ключ, *abcd* структуры и т. д.), сложенные из соседних по цепи элементов  $\beta$  (и/или  $\alpha$ ) структуры, часто называют «*супервторичными*» структурами.

## Лекция 14

Строение  $\alpha$ -белков. Пучки и слои спиралей. Модель квазисферической глобулы из  $\alpha$ -спиралей. Плотная упаковка при контакте  $\alpha$ -спиралей. Строение  $\alpha/\beta$ -белков: параллельный  $\beta$ -слой, прикрытый  $\alpha$ -спиральями, и  $\alpha/\beta$ -цилиндр. Топология  $\beta$ - $\alpha$ - $\beta$ -субъединиц. Строение  $\alpha+\beta$ -белков. Отсутствие прямой связи архитектуры белка с его функцией.

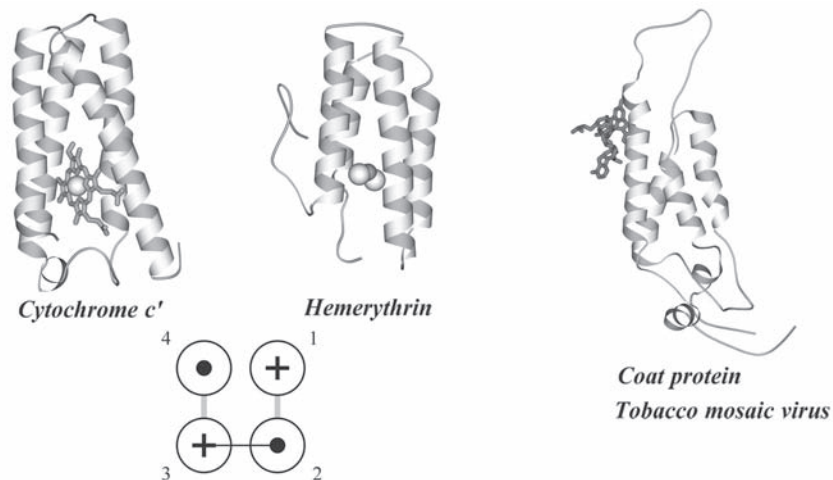
Перейдем теперь к  $\alpha$ -белкам — белкам, сложенным из  $\alpha$ -спиралей. Их труднее классифицировать, чем  $\beta$ -белки. Дело в том, что взаимное расположение  $\beta$ -участков в листах поддерживается водородными связями в главной цепи (а она — повсюду одинакова), в то время как расположение  $\alpha$ -спиралей в глобуле поддерживается плотной упаковкой их боковых групп, которые весьма разнообразны по размеру. Поэтому  $\alpha$ -спирали не уложены в такие более или менее стандартные листы, как  $\beta$ -тяжи.

- Проще других устроены те  $\alpha$ -белки, где  $\alpha$ -спирали длинные. Такие спирали образуют пучки: они лежат (почти) параллельно или антипараллельно (в общем, «коллинеарно») друг к другу. С пучками спиралей мы уже встречались в фибриллярных и мембранных белках.

На рис. 14-1 показаны три четырехспиральных  $\alpha$ -белка. Эти, столь похожие между собой белки исполняют разные функции: один из них, цитохром, связывает электрон; другой (гемэритрин) связывает кислород, а третий (белок оболочки вируса табачной мозаики) связывает гораздо более крупные молекулы — другие оболочечные белки и РНК. В первых двух случаях какую-то общность действия проследить можно, так как оба белка что-то переносят в дыхательной цепи. Правда, сходство здесь весьма отдаленное: в цитохроме полипептид, а именно о его укладке идет речь, связывает гем, гем — ион железа, а уж оно-то и хватает электрон; а в гемэритрине полипептид связывает железо непосредственно, без гема, — и уже два иона железа, а не один, — которые и хватают кислород. Итак, у гемэритрина и цитохрома некую общность функции, пусть с большой натяжкой, но можно найти, но, конечно, у их действия нет ничего общего с действием РНК-связывающего



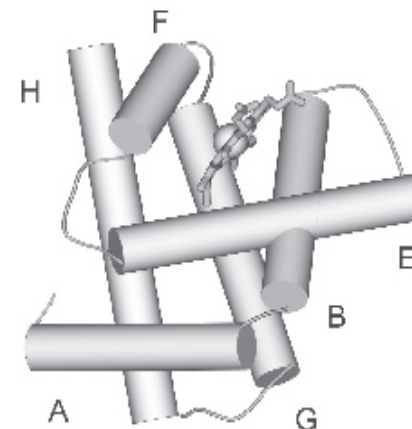
белка оболочки вируса, несмотря на очень сходную архитектуру структурного каркаса, четырехспирального пучка. Причем сходство распространяется не только на архитектуру — четырехспиральный пучок — но и на ход цепи сквозь эту архитектуру, т. е. на мотив укладки цепи. Последнее подчеркивается общей для всех трех белков топологической схемой, приведенной внизу рисунка (схема изображает вид пучка с торца).



**Рис. 14-1.** Три сходных по архитектуре («четырёхспиральный пучок»), но разных по функции  $\alpha$ -спиральных белков: цитохром  $c'$ , гемэритрин и белок оболочки вируса табачной мозаики. Показана как белковая цепь, так и кофакторы: скелетные модели — гем (в цитохроме) и фрагмент РНК (в белке оболочки вируса), шарики — ионы железа (в геме цитохрома и в гемэритрине) и связанный железом кислород (в гемэритрине). Общая архитектура таких «пучков» напоминает коллинеарную упаковку  $\beta$ -листов. На топологической схеме (внизу) дан вид на белок «снизу» (сторцов структурных сегментов). Кружки — торцы  $\alpha$ -спиралей. Крестик соответствует N-концу сегмента (т. е. он «идет от нас»), точка — его С-концу (т. е. он «идет к нам»). Показан ход петель, соединяющих структурные сегменты: черной линией, если петля обращена к нам, и светлой, если она находится на противоположной стороне укладки. Цифры на схеме указывают порядок структурных сегментов в цепи (от N- к С-концу цепи)

Итак, при одинаковом мотиве укладки цепи белки могут работать совсем по-разному. А вот гемэритрин и классический кислород-связывающий белок миоглобин (рис. 14-2) делают одно и то же дело (только первый — в червях, а второй — в хордовых, в том числе и в нас с вами), но их архитектуры совсем непохожи, за исключением того, что оба они —  $\alpha$ -белки. Но

в гемэритрине все  $\alpha$ -спирали лежат параллельно, а в миоглобине спирали организованы в два перпендикулярных слоя. Это — еще один пример того, что сходная функция может осуществляться белками с разной архитектурой, а белки одинакового устройства могут заниматься совсем разным делом.



**Рис. 14-2.** Структура глобина: скрещенные слои по три  $\alpha$ -спирали в каждом. Спираль А, Е и F (они занумерованы буквами по порядку нахождения в цепи) находятся в верхнем слое, спирали H, G и B — в нижнем. Короткие (из 1–2 витков каждая) спирали С и D не изображены, так как они не консервативны в глобинах. В щели верхнего слоя находится гем. Такие «скрещенные слои» напоминают ортогональную упаковку  $\beta$ -листов. [Ортогональный контакт спиралей В и Е особенно плотен, так как в нем образованная в месте нахождения глицина: вмятина на поверхности одной спирали входит в такую же вмятину на поверхности другой.]

Я, как и раньше, хочу привлечь ваше внимание к нетривиальным случаям малой связанности структуры и функции белка, понимая, что вы и так знаете, что родственные белки — например, миоглобин и другие глобины — похожи по строению и занимаются одним и тем же делом.

Сравнивая миоглобин с гемэритрином, я хочу обратить ваше внимание еще на то, что активный центр обоих этих белков (в первом случае этот центр — гем с ионом железа внутри, во втором — два иона железа) — этот центр локализован в «архитектурном дефекте» структуры — в данном случае, в щели между раздвинутыми спиралями.

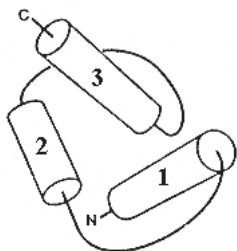
■ Рассмотренные выше «пучки» типичны для упаковки весьма длинных  $\alpha$ -спиралей. Они встречаются как в водорастворимых глобулярных белках, так и в фибриллярных и в мембранных белках. При этом ядро белка, заключенное между  $\alpha$ -спиралями (гидрофобное — в водорастворимых

глобулярных и фибриллярных белках, гидрофильное — в мембранных), имеет вытянутую, квазицилиндрическую форму. «Скрещенные слои» также образуются довольно длинными спиралями; такие упаковки обладают плоским гидрофобным ядром.

■ Однако для относительно коротких спиралей, типичных для глобулярных белков — их длина обычно составляет около 20 Å при диаметре 10 Å — более характерна квазисферическая укладка спиралей вокруг «квазишарового» гидрофобного ядра, сложенного из сидящих на этих спиралях боковых групп.

Типичная упаковка спиралей в глобулярном белке показана на рис. 14-3. Ее нельзя описать в терминах параллельной и перпендикулярной упаковок спиралей — большинство межспиральных контактов имеет угол около 40–60°.

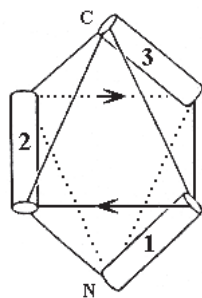
Однако и такие сложные укладки можно неплохо описывать и классифицировать при помощи модели «квазисферических многогранников», которую А. Г. Мурзин и А. В. Ф. предложили более 20 лет назад. Для примера покажу (рис. 14-4), как эта модель описывает только что показанную (на рис. 14-3) α-спиральную глобулу. Эти два последних рисунка, кстати, взяты из опубликованной в *Nature* рецензии на вышеупомянутую работу А. Г. М. и А. В. Ф.



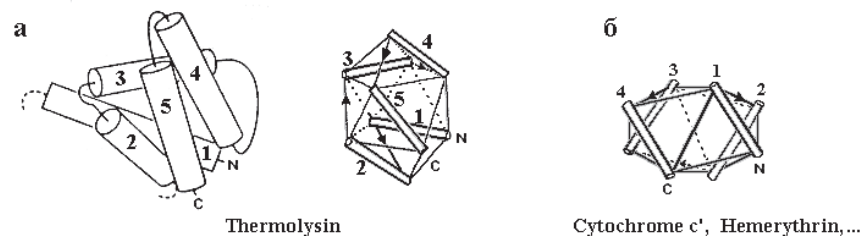
**Рис. 14-3.** Типичная упаковка спиралей в глобулярном белке: N-концевой домен актинидина (ход петель прослежен очень грубо). Обратите внимание, что архитектуру этого домена невозможно описать в терминах коллинеарных и ортогональных упаковок α-спиралей

Не откажу себе в удовольствии показать и еще пару рисунков из той же рецензии (рис. 14-5а).

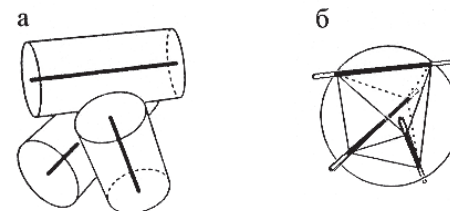
Кстати, модель квазисферических многогранников довольно прилично описывает и довольно длинные спиральные пучки — те, что мы видели на рис. 14-1. Это демонстрируется рисунком 14-5б.



**Рис. 14-4.** Размещение α-спиралей на ребрах квазисферического многогранника, моделирующее N-концевой домен актинидина. Стрелки указывают ход соединяющих α-спиралей петель



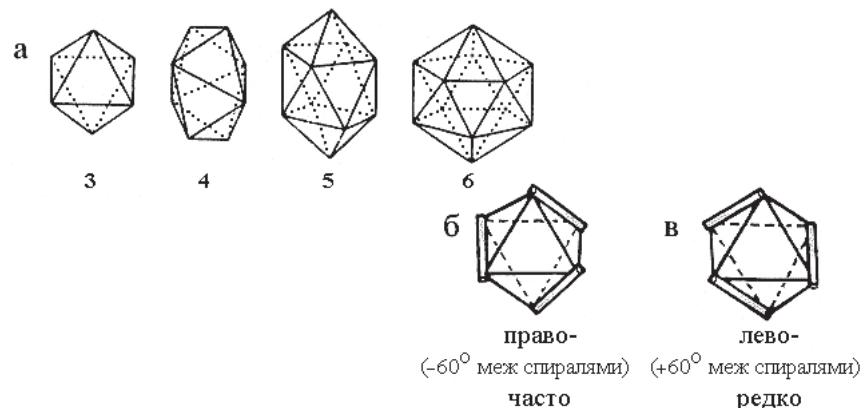
**Рис. 14-5.** Другие примеры того, как упаковка спиралей в глобулярном белке описывается моделью квазисферических многогранников. (а) С-концевой домен термолизина и его модель, основанная на размещении спиралей на ребрах квазисферического многогранника; (б) модель четырехспиральной глобулы, показанной на рис. 14-1. Картинка 14-5а (как и 14-3, 14-4), с небольшими изменениями, взята из вышеупомянутой рецензии (С. Chothia, *Nature* (1989) 337:204–205)



**Рис. 14-6.** Построение многогранника для описания упаковки спиралей. (а) Компактная упаковка трех α-спиралей (цилиндров — 10 Å в диаметре — с осями). (б) Построение многогранника: центр упаковки окружается сферой радиусом в 10 Å; ее пересечение с осями спирали дает вершины многогранника. На трех из его ребер лежат оси спиралей. Те части ребер, что лежат внутри сферы, оставлены темными. Точки пересечения осей со сферой образуют вершины многогранника. Каждой вершине соответствует одна половинка одной спирали. Оси спиралей образуют часть ребер многогранника, а прочие ребра описывают контакты спиралей

Суть модели квазисферических многогранников заключается в том, что она концентрирует наше внимание на расположении спиралей вокруг шарового ядра глобулы. Модель учитывает только, что α-спиралы — твердые вытянутые частицы — плотно окружают ядро; и что полярные концы спиралей должны находиться на поверхности глобулы. Каждая упаковка спиралей может моделироваться многогранником (рис. 14-6), каждая вершина которого соответствует как бы половине спирали. Самые компактные, «квазисферические» многогранники (рис. 14-7) описывают компактные глобулы. Упаковки, близкие к идеальным, и наблюдаются в глобулярных α-белках. При этом каждому данному числу спиралей отвечает один многогранник, а в его рамках существует несколько (от двух до десяти)

типов упаковок, соответствующих различным размещениям осей спиралей на ребрах этого многогранника. Среди этих упаковок есть и рассмотренные выше «пучки спиралей», и «скрещенные слои».



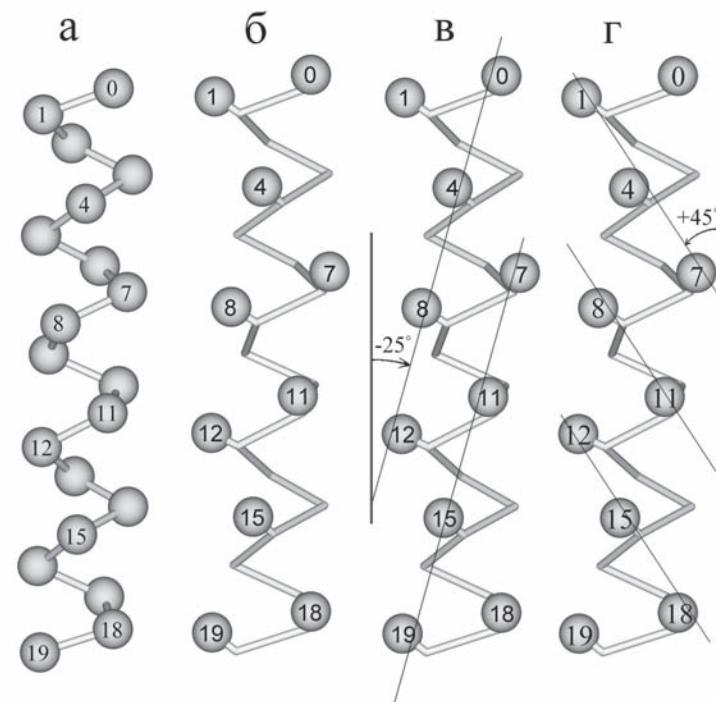
**Рис. 14-7.** Квазисферические многогранники (а), описывающие компактные укладки трех, четырех, пяти и шести спиралей. Большое число спиралей не может уложиться вокруг округлого ядра. Каждый многогранник описывает несколько типов упаковок, т. е. типов «штабелей» спиралей, соответствующих различным размещениям осей спиралей на его ребрах. Таких упаковок — две для трехспирального комплекса [(в) лево-, и (б) правозакрученный (как на рис. 14-3, 14-4) пучок], десять — для четырехспирального, десять — для пятиспирального и восемь — для шестиспирального комплекса («штабеля» для 4–6-спиральных глобул не показаны, но их, при желании, легко построить самостоятельно, разместив спирали — всеми возможными способами — на ребрах многогранника так, чтобы каждая вершина многогранника соответствовала одному концу одной спирали). Те упаковки, где межспиральные углы способствуют плотному контакту спиралей — см. рис. 14-9 — встречаются в белках чаще, чем прочие

■ Интересно, что в наблюдаемых архитектурах  $\alpha$ -спиральных белков вдоль ребер квазисферических многогранников идут не только спирали, но и — как правило — соединяющие их нерегулярные петли (см. рис. 14-4, 14-5). Иными словами, в типичном случае белковая цепь как бы обволакивает свое гидрофобное ядро, следуя по непрерывной цепочке ребер квазисферического многогранника.

■ Обратимся теперь к вопросу о том, как создается плотная упаковка в белковой глобуле. То, что такая упаковка существует, следует из экспериментов, показавших, что белок так же плотен и так же тверд, как органический кристалл. Однако еще предстоит объяснить, как достигается такая упаковка, — слишком уж сложны по форме и разнообразны боковые группы белковой цепи.

Точнее, принцип создания плотной упаковки — и то в самых общих чертах — более или менее ясен только для  $\alpha$ -спиралей, почему и уместно рассмотреть этот вопрос именно сейчас.

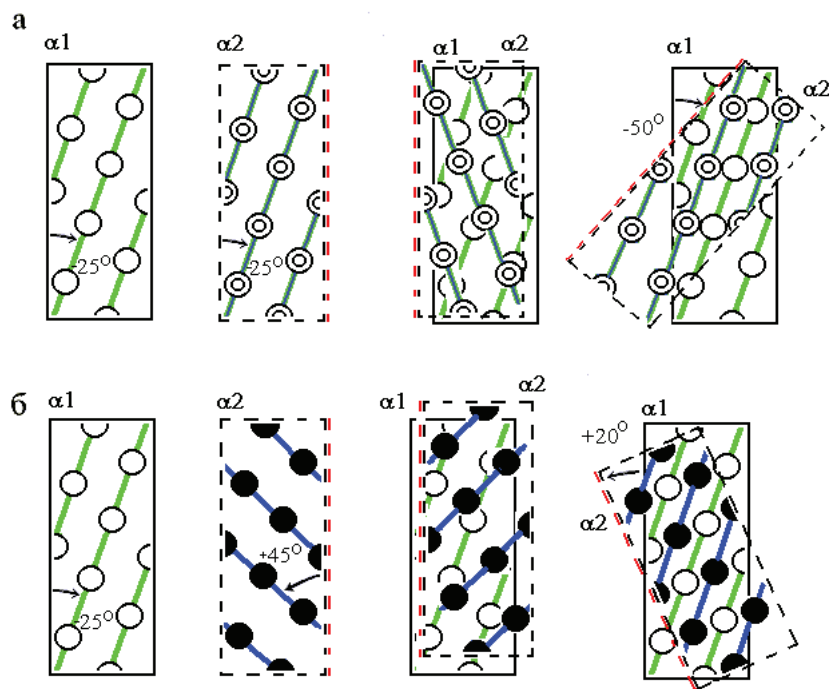
Первая модель плотной упаковки  $\alpha$ -спиралей, упаковка по принципу «выступы (боковые группы) во впадины (между боковыми группами)», была предложена Криком в 1953 г., еще до расшифровки трехмерной структуры первого белка. Затем она была развита Ефимовым и, независимо, группой Чотиа-Левитт-Ричардсон, и к настоящему времени приобрела вид модели «хребты (боковых групп) в лоцины (между таковыми)».



**Рис. 14-8.**  $\alpha$ -спираль; отмечены  $C^\alpha$ -атомы (а) и  $C^\beta$ -атомы (б–г). Нумерованные остатки обращены к читателю. Показаны два сорта хребтов (тонкие линии на лицевой поверхности спирали) из сближенных боковых групп (в, г). Хребты из боковых групп « $i$ »–« $i+4$ »–« $i+8$ »... идут под углом  $-25^\circ$  к оси спирали (в), хребты из групп « $i$ »–« $i+3$ »–« $i+6$ »... — под углом  $+45^\circ$  (г); на рисунке углы представляются меньшими, так как типичные хребты проходят через массивные боковые группы, а на рисунках (в, г) — через центры  $C^\beta$ -атомов. Картинка, с некоторыми изменениями, взята из [5]

Согласно этой модели, боковые группы на поверхности спирали образуют выступы, создающие разделенные ложнами хребты. «Хребты и ложныны» несколько лучше описывают реальность, чем «выступы и впадины», так как разворот одного выступа (одной боковой цепи) в сторону другого (другой боковой цепи) может сделать тот или другой «хребет из выступов» более отчетливым. Хребты (и идущие вдоль них ложныны) бывают двух типов. Хребты типа «+4» создаются боковыми группами остатков, расположенных в цепи под номерами « $i$ », « $i+4$ », « $i+8$ », и т. д. (иными словами — с периодом 4), хребты типа «+3» создаются боковыми группами номер « $i$ », « $i+3$ », « $i+6$ », и т. д. (т. е. с периодом 3). Рисунок 14-8 показывает, что эти хребты образуют разного знака углы с осью спирали.

При плотной упаковке хребты одной спирали входят в ложныны другой. При этом есть две основные возможности (рис. 14-9).



**Рис. 14-9.** Два основных способа плотной упаковки боковых групп при контакте спиралей: под углом  $-50^\circ$  (а) и  $+20^\circ$  (б). Мы смотрим на зону контакта сквозь одну спираль (сквозь перевернутую вдоль своей оси  $\alpha 2$ ). Остатки «нижней» спирали  $\alpha 1$  изображены светлыми, а верхней  $\alpha 2$  — темными кружками. Картинка, с некоторыми изменениями, взята из [5]

А) Хребты «+4» одной спирали входят в ложныны между такими же хребтами «+4» другой (рис. 14-9а; видно, что для получения плотной упаковки спираль  $\alpha 2$  переворачивается, накладывается на спираль  $\alpha 1$  и поворачивается, пока хребты «+4» обеих спиралей не станут параллельно друг другу). При такой упаковке угол между осями спиралей близок к  $-50^\circ$ . Этот угол наиболее характерен для контактов спиралей в  $\alpha$ -спиральных глобулах. Он также типичен для контакта спиралей в  $\alpha/\beta$ - и  $\alpha+\beta$ -белках, о которых речь пойдет ниже. Дело в том, что при этом контакте спиралей скрученность слоя  $\alpha$ -спиралей (угол скручивания в нем близок к  $-50^\circ/10 \text{ \AA}$ , где  $-50^\circ$  — угол между осями соседних спиралей, а  $10 \text{ \AA}$  — ширина  $\alpha$ -спирали) неплохо стыкуется с типичной скрученностью  $\beta$ -листа (имеющем тот же угол скручивания,  $-25^\circ/5 \text{ \AA}$ , где  $-25^\circ$  — угол между осями соседних  $\beta$ -тяжей, а  $5 \text{ \AA}$  — ширина  $\beta$ -тяжа).

Б) Хребты «+3» одной спирали входят в ложныны между хребтами «+4» другой (рис. 14-9б). При такой упаковке угол между осями спиралей близок к  $+20^\circ$ . Этот угол наиболее характерен для контактов спиралей в пучках — и в  $\alpha$ -спиральных глобулах, и в фибриллярных, и в мембранных белках.

Кроме того, хребты «+3» одной спирали могут входить в ложныны между такими же хребтами «+3» другой, образуя очень короткий контакт почти перпендикулярных спиралей. В силу малости размера этого контакта, он на рис. 14-9 не показан, хотя такой «перпендикулярный» контакт спиралей довольно типичен для  $\alpha$ -спиральных глобул.

Заклучая описание плотной упаковки, надо заметить, что реальные отклонения от приведенных выше «идеальных» углов весьма велики, так как боковые группы сильно варьируют по размеру. По той же причине в  $\beta$ -структуре (где боковые группы выступают меньше — поверхность  $\beta$ -листа довольно плоская, а спирали — выпуклая) картина проникновения хребтов в ложныны совсем смазана и наблюдается лишь в отдельных случаях.

■ И последнее. Как согласуется плотная упаковка спиралей с моделью размещения спиралей на ребрах квазисферических многогранников, о которой речь шла выше? Оказывается, довольно любопытным образом. Те «многогранные» упаковки, где углы между спиралами близки к  $-50^\circ$  и/или  $+20^\circ$ , требующимся для плотного контакта спиралей, встречаются часто; а прочие — редко (но тоже встречаются). Так, из двух изображенных на рис. 14-7 трехспиральных упаковок одна, *правозакрученный* пучок, приводит к межспиральным углам в  $-60^\circ$  (что близко к углу  $-50^\circ$ , требующемуся для плотной упаковки, см. рис. 14-9а), и такой трехспиральный пучок встречается часто; а другая, *левозакрученный* пучок, приводит к межспиральным углам в  $+60^\circ$  (что далеко от всех углов —  $-50^\circ$ ,  $+20^\circ$ ,  $90^\circ$ , оптимальных для плотного контакта) — и такой трехспиральный пучок встречается на порядок реже.



- Обратимся теперь к «смешанным» белкам, сложенным из  $\beta$ -листов и  $\alpha$ -спиралей. Для них характерна слоистая структура, причем  $\alpha$ -спирали и  $\beta$ -участки не могут лежать в одном листе — это привело бы к энергетически невыгодной дегидратации водородных связей на краю  $\beta$ -листа (рис. 14-10).



**Рис. 14-10.** Слоистая структура смешанных ( $\alpha/\beta$  и  $\alpha+\beta$ ) белков. Вид с торца  $\alpha$ -спиралей (изображенных — с торца — в виде квадратиков, для подчеркивания плотности упаковки глобулы) и  $\beta$ -тяжей (при виде с торца — прямоугольников).  $\alpha$ -спирали и  $\beta$ -тяжи не могут лежать в одном слое — это привело бы к дегидратации водородных связей на краю  $\beta$ -листа (доноры и акцепторы водородных связей в  $\beta$ -листе изображены черными точками)

Различают  $\alpha/\beta$  (« $\alpha$  дробь  $\beta$ ») и  $\alpha+\beta$  (« $\alpha$  плюс  $\beta$ ») белки (точнее — домены). Иногда их объединяют в общий разряд  $\alpha\&\beta$  (т. е. « $\alpha$  и  $\beta$ »), или  $\alpha-\beta$  белков.

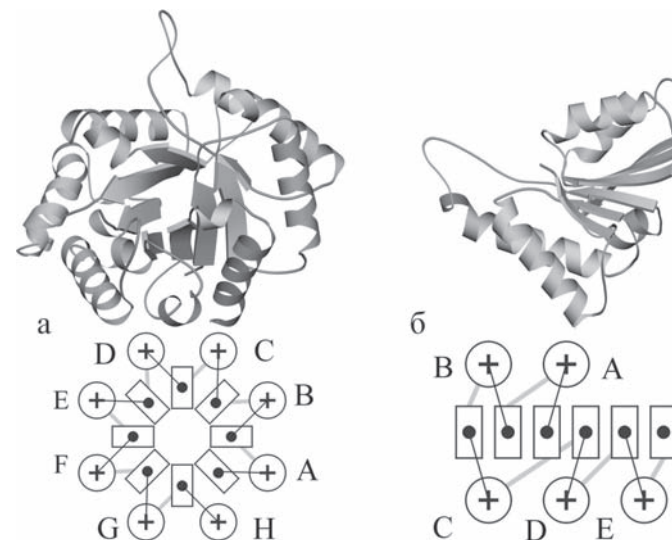
В  $\alpha/\beta$  доменах  $\beta$ -структура — параллельная, и  $\alpha$ -спирали также параллельны друг другу (и антипараллельны  $\beta$ -участкам), а характерное чередование  $\alpha$ - и  $\beta$ -участков в цепи имеет вид  $-\alpha-\beta-\alpha-\beta-\alpha-\dots$

Есть два наиболее характерных мотива строения  $\alpha/\beta$ -белков:  $\alpha/\beta$ -цилиндр, где  $\beta$ -цилиндр лежит внутри цилиндра, сложенного из  $\alpha$ -спиралей (рис. 14-11а), и «укладка Россманна», где более или менее плоский (не считая обычного — правого, если смотреть вдоль  $\beta$ -тяжей — пропеллерного скручивания)  $\beta$ -слой лежит между комплементарно ему скрученными слоями  $\alpha$ -спиралей (рис. 14-11б). В отличие от ранее рассмотренных доменов,  $\alpha/\beta$ -домены обычно имеют два гидрофобных ядра: в укладке Россманна — между  $\beta$ -листом и каждым из слоев спиралей; в  $\alpha/\beta$ -цилindre — внутри  $\beta$ -цилиндра (меньшее ядро) и между  $\beta$ - и  $\alpha$ -цилиндрами (большее).

Кстати, встречаются и  $\alpha/\beta$  почти цилиндры — «недовернутые», не замкнутые в кольцо (и, значит, не имеющие гидрофобного ядра, укрытого внутри  $\beta$ -цилиндра). Их называют « $\alpha/\beta$ -подковами», и в них насчитывается по полтора десятка  $\alpha/\beta$ -повторов.

Однако обычно в  $\alpha/\beta$ -цилindre насчитывается восемь  $\alpha$ - и восемь  $\beta$ -участков, причем почти все  $\alpha/\beta$ -цилиндры имеют одинаковую топологию: все соседние и расположенные через один по цепи  $\beta$ - и  $\alpha$ -участки контакти-

руют друг с другом. По-видимому, такое строение обеспечивает особую стабильность белковой глобулы, так как многочисленные белковые глобулы с такой архитектурой (наблюдаемой в 10% белков) все очень похожи друг на друга по форме — часто и без каких-либо следов общего происхождения в аминокислотной последовательности, и без какой-либо общности функций.

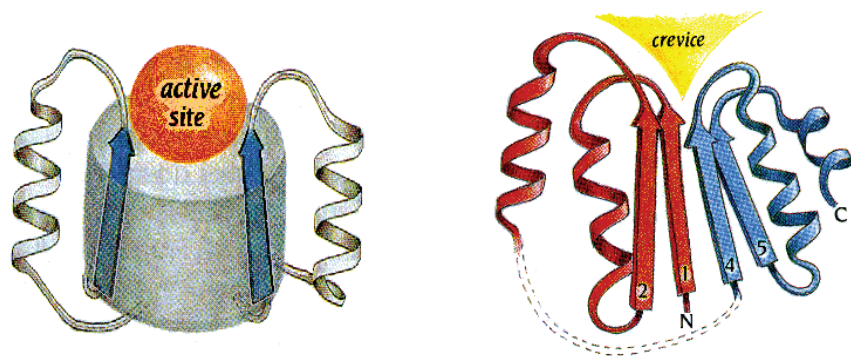


**Рис. 14-11.** Типичные мотивы строения  $\alpha/\beta$ -белков и их упрощенные модели (вид на модели — с торца  $\beta$ -слоя): « $\alpha/\beta$ -цилиндр» в триозофосфатизомеразе (а); «укладка Россманна» в NAD-связывающем домене малатдегидрогеназы (б). На детальной картинке в первом белке видна обращенная к нам впадина, образованная расходящимися в форме розетки петлями; она идет к центру  $\beta$ -цилиндра. Во втором — слева сверху видна щель между расходящимися вниз и вверх от  $\beta$ -листа петлями

Без общности функций — да; без общности в строении активного центра — да; но не без общности места активного центра в архитектуре глобулы — каждая архитектура содержит места (обычно — впадины), как бы специально созданные этой архитектурой для активного центра — что бы он ни делал.

- Я хочу обратить ваше внимание на «воронку» на оси  $\alpha/\beta$ -цилиндра (рис. 14-12а) — вы видите, что эта вмятина, заложенная в общей архитектуре белка, в самом мотиве укладки его цепи, не прикрыта петлями. Здесь находится активный центр. Точнее: из двух таких «воронок», расположенных на противоположных торцах  $\alpha/\beta$ -цилиндра, под активный центр используется только одна — та, куда смотрят С-концы  $\beta$ -участков и N-концы

$\alpha$ -спиралей. Считается, что именно эти концы (возможно, из-за множества открытых NH-групп N-концов спиралей), вместе с соединяющими их короткими петлями, особенно пригодны для связывания разнообразных субстратов. Впрочем, этот вопрос еще недостаточно ясен.



**Рис. 14-12.** Типичное положение активного центра (active site) в  $\alpha/\beta$ -белках: в «воронке» на оси  $\alpha/\beta$ -цилиндра (а), и в щели (crevice), образованной расходящимися петлями в «укладке Россманна» (б). Картинки, с разрешения, взяты из [5]

Аналогичное размещение активного центра — в щели, во вмятине, причем во вмятине, куда тоже смотрят С-концы  $\beta$ -участков и N-концы  $\alpha$ -спиралей, — наблюдается и в «укладках Россманна». Только здесь вмятина образуется не при расхождении петель от центра цилиндра, а при расхождении петель, одна часть которых идет от  $\beta$ -листа к спиральям, лежащим под листом, а другая — над этим листом (рис. 14-12б).

**Внутренний голос:** Только не надо думать, что активный центр всегда помещается в такой очевидной вмятине. Это бывает часто, но далеко не всегда!

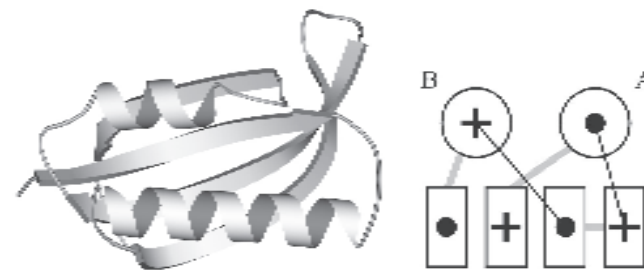
**Лектор:** Верно. Как правило (в ~80% случаев), активный центр находится в самой большой вмятине на глобуле, а она, как правило, определяется архитектурой образованного вторичными структурами скелета этой глобулы. Но много раз, уже зная полную пространственную структуру белка, на нем долго, и не всегда успешно, искали этот самый активный центр...

■ Теперь перейдем к  $\alpha + \beta$ -белкам. В их основе лежит антипараллельная (а не параллельная, как в  $\alpha/\beta$ -белках)  $\beta$ -структура.

В  $\alpha + \beta$ -белках выделяются два класса. Белки одного класса (их порой называют  $\alpha\beta$ -складками, “ $\alpha\beta$ -plaits”) напоминают  $\alpha/\beta$  белки тем, что в них слой  $\alpha$ -спиралей лежит на  $\beta$ -листе. Они напоминают  $\alpha/\beta$ -белки также

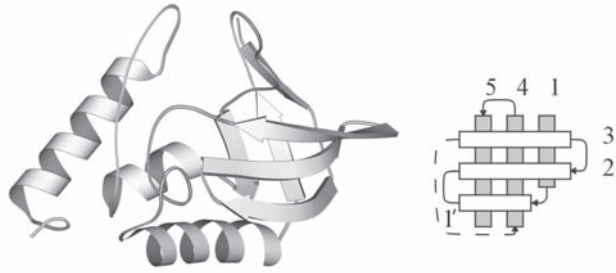
регулярным (но с иным, чем в  $\alpha/\beta$ -белках) чередованием  $\alpha$ - и  $\beta$ -участков в цепи и в пространстве. Белки другого класса («собственно»  $\alpha + \beta$ -белки) не имеют такого чередования, в их цепях  $\alpha$ -структура «отмешана» от  $\beta$ -структуры.

Характерное чередование  $\alpha$ - и  $\beta$ -участков в цепи  $\alpha\beta$ -складки имеет вид ... $\alpha$ - $\beta$ - $\alpha$ - $\beta$ ... или ... $\alpha$ - $\beta$ - $\beta$ - $\beta$ - $\alpha$ - $\beta$ - $\beta$ ... (рис. 14-13). Здесь отдельные  $\alpha$ -спирали лежат между  $\beta$ -шпильками или  $\beta$ -листами из четного числа  $\beta$ -тяжей. Соседние по цепи  $\beta$ -участки в  $\alpha\beta$ -складках образуют антипараллельные  $\beta$ -шпильки; а из-за четного (а не нечетного, как в  $\alpha/\beta$ -белках) числа  $\beta$ -тяжей между  $\alpha$ -спиральями и из-за общей коллинеарности этих тяжей  $\alpha$ -спиральям,  $\alpha$ -спирали также образуют антипараллельные шпильки. Интересно, что такая (или, точнее, сходная  $\alpha\beta\alpha\beta\beta$ ) «складчатая» конструкция белка была сначала теоретически предсказана (точнее, предложена в качестве проекта белково-инженерного дизайна), а затем уже такие архитектуры были массово обнаружены в природе, причем их оказалось особенно много среди ферредоксинов и... РНК-связывающих белков.



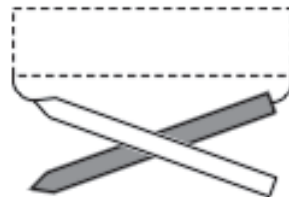
**Рис. 14-13.** Один из типичных мотивов строения  $\alpha + \beta$ -белка: « $\alpha\beta$ -складка» ( $\alpha\beta$ -plait) в рибосомальном белке S6.  $\alpha\beta$ -складка отличается более регулярным, чем в «собственно»  $\alpha + \beta$ -белках, чередованием вторичных структур в цепи (в данном случае  $\beta\alpha\beta\alpha\beta$ ); это — так называемая «ферредоксиновая укладка». Справа помещена схема строения этого белка (вид с торца структурных участков, приблизительно коллинеарных). Спирали занумерованы буквами. «+» на торце  $\alpha$ - или  $\beta$ -участка означает, что он идет от нас (т.е. что к нам обращен его N-конец), точка — что к нам

В «собственно»  $\alpha + \beta$ -доменах (рис. 14-14)  $\alpha$ - и  $\beta$ -участки расположены в цепи нерегулярно, и скорее — как бы блоками. Такие белки обычно выглядят как  $\beta$ -лист (часто — загнутый сам на себя и образующий субдомен), прикрытый отдельными  $\alpha$ -спиральями или  $\alpha$ -спиральным субдоменом;  $\beta$ -структура в  $\alpha + \beta$ -доменах — в основном антипараллельная, как в «чистых»  $\beta$ -белках.



**Рис. 14-14.** Характерный мотив строения  $\alpha + \beta$ -белков: нуклеаза стафилококка. «Собственно»  $\alpha + \beta$ -белки отличаются менее регулярным, чем в  $\alpha/\beta$ -белках или  $\alpha\beta$ -складках, чередованием вторичных структур в цепи (в данном случае  $\beta\beta\beta\alpha\beta\beta\alpha$ ). Мотив укладки цепи, наблюдаемый в  $\beta$ -домене нуклеазы, называется «ОБ-укладка» (“OB-fold”, т. е. “Oligonucleotide-Binding fold”). Справа помещена схема строения этого, часто встречающегося в разных много- и однодоменных белках,  $\beta$ -структурного ОБ-домена (вид сверху на ортогональную упаковку  $\beta$ -участков).  $\beta$ -тяжи занумерованы цифрами. Первый  $\beta$ -тяж согнут (в сущности, сломан); две его половинки носят номера 1 и 1'. Обратите внимание на «эффект матрешки» (“Russian doll effect”): одна типовая структура (ОБ-укладка) составляет часть другой, тоже типовой структуры (нуклеазной укладки)

- Очень характерной чертой  $\alpha/\beta$ - и  $\alpha + \beta$ -белков (а равно и  $\beta$ -белков) является *правовинтовой* (т. е. «к нам — против часовой стрелки») ход перемычек (рис. 14-15) между параллельными  $\beta$ -участками одного и того же  $\beta$ -листа, следующими друг за другом в цепи (но не обязательно притом соседними в этом листе) (см. рис. 14-11 и далее). При этом перемычка между такими параллельными  $\beta$ -участками в  $\alpha/\beta$ - и  $\alpha + \beta$ -белках, как правило, включает в себя  $\alpha$ -спираль. В  $\beta$  (и порой в  $\alpha + \beta$ ) белках в такой правовинтовой перемычке между параллельными  $\beta$ -участками одного листа лежит — вы должны помнить из прошлой лекции («abcd» структуры и т. д.) —  $\beta$ -участок другого листа, а иногда и отдельный  $\beta$ -лист. Иногда (довольно редко) в перемычке между параллельными  $\beta$ -участками нет ни  $\alpha$ -, ни  $\beta$ -структуры, но и тогда эта перемычка обычно *право*-, а не *лево*винтовая.



**Рис. 14-15.** Типичный, правовинтовой ход перемычек между параллельными  $\beta$ -тяжами одного листа. В перемычке обычно находится еще один участок вторичной структуры

На следующей лекции мы увидим, что такой ход перемычки обычно способствует стабилизации белка и, тем самым, позволяет большее разнообразие кодирующих, т. е. стабилизирующих такую структуру последовательностей, а потому этот ход и наблюдается чаще.

- Завершая беглый обзор структур глобулярных белков, хочется еще раз подчеркнуть, что одни и те же — или очень похожие — архитектуры часто встречаются в белках, совсем не сходных функционально или филогенетически. Это открытие заложило основы физической, или рациональной (как это чаще называется в литературе), классификации белков. Об этом мы также поговорим на следующей лекции.

## Лекция 15

Классификация структур белков. Отсутствие наблюдаемой «макроэволюции» укладок белковых цепей — при наблюдаемой «микроразнообразии» их структур. Дупликация гена и специализация. Эволюция путем перемешивания доменов. «Стандартные» третичные структуры. Типичность «квазислучайного» чередования аминокислот в первичных структурах глобулярных белков, контраст с периодическими первичными структурами фибриллярных белков и блочными — мембранных белков. Физические принципы строения белковой глобулы. Основные закономерности, наблюдаемые в структурах белковых глобул: наличие отдельно  $\alpha$ - и отдельно  $\beta$ -слоев; редкость перекрытия петель; редкость параллельности соседних по цепи структурных сегментов; редкость левых  $\beta$ - $\alpha$ - $\beta$ -суперспиралей. «Энергетические» и «энтропийные» дефекты редко встречающихся структур и связь этих «дефектов» с относительной редкостью аминокислотных последовательностей, стабилизирующих «дефектные» структуры. «Принцип множественности».

На этой лекции мы попробуем разобраться в вопросе о том, почему большинство белков вписывается в узкий набор стандартных структур, — из прошлых лекций вы должны были вынести именно такое впечатление.

На самом деле здесь действует правило «80% : 20%». В его исходном виде оно гласит: «80% всего пива выпивается 20% населения». В применении к белкам — «80% всех белков вписывается в 20% наблюдаемых архитектур белковых глобул». И в предыдущих лекциях я позволил себе сосредоточиться именно на типичных структурах.

Итак, почему же большинство белков вписывается в узкий набор стандартных структур? И почему не все (как цепи ДНК)? На каком структурном уровне проявляется это сходство? И что стоит за этими общими структурами: память об общем происхождении? функциональная целесообразность? или необходимость удовлетворять общим принципам сворачивания стабильных белковых структур? И на каком структурном уровне проявляется сходство белков, имеющих разное происхождение и функцию?



На этой лекции мы рассмотрим эти вопросы на качественном уровне, а потом, накопив немного знаний по самоорганизации белков, попробуем ответить на те же вопросы более строго.

- Пока расшифрованных белков было мало — где-то до середины 1970-х гг. — бытовало представление, что каждая третичная структура глубоко уникальна, т. е. что белки разных эволюционных семейств не похожи друг на друга. Однако, по мере роста информации о пространственном строении белковых молекул становилось все яснее, что существуют какие-то « типовые проекты » строения белковых глобул. Архитектуры вновь расшифрованных белков (или, по крайней мере, их доменов) все чаще и чаще оказывались сходными с архитектурами белков уже известных — но при этом совсем других и по функции, и по аминокислотной последовательности. Поэтому начало складываться и крепнуть мнение, не раз уже встречавшееся в этих лекциях, что причина сходства третичных структур белков заключается, видимо, **не только** в эволюционной дивергенции и **не** (или не только) в функциональной конвергенции белков, а просто в ограничении набора укладок какими-то физическими закономерностями.

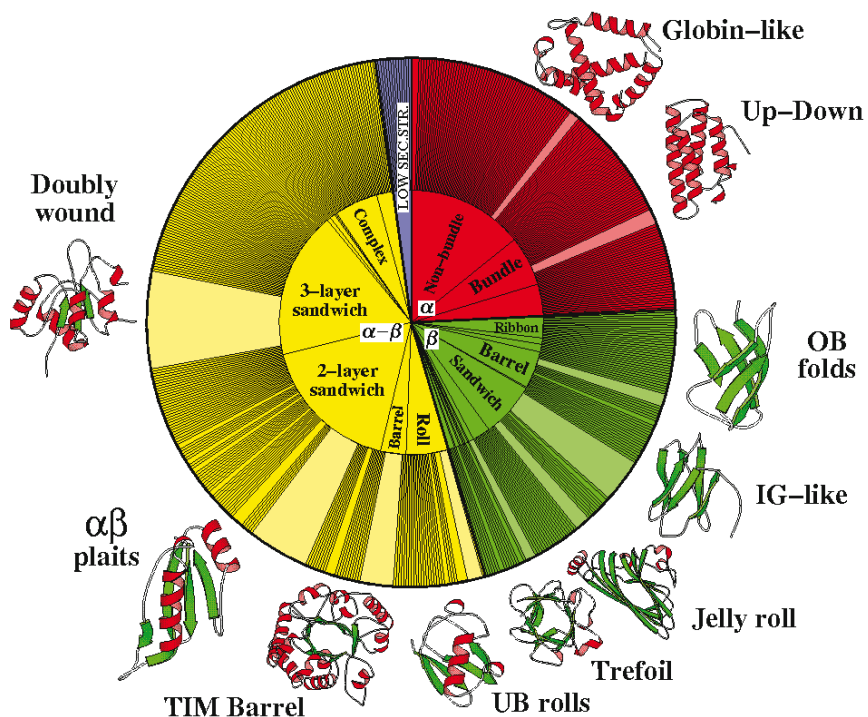
К концу 1970-х гг. стало окончательно ясно, что между двумя « традиционными » структурными уровнями (вторичная структура белка и его детальная атомная трехмерная структура) находится промежуточный уровень — уже известный нам « мотив укладки » белковой цепи (“folding pattern”), что он определяется взаимным расположением  $\alpha$ - и /или  $\beta$ -участков в глобуле и что именно на этом уровне проявляется сходство белков, не связанных ни эволюционно, ни функционально. В отличие от детальной, атомной трехмерной структуры, « мотивы укладки » удивительно просты и даже красивы (рис. 15-1).

Открытие того, что одни и те же — или очень похожие — архитектуры часто встречаются в белках, совсем не сходных функционально или филогенетически, заложило основы физической, или рациональной (как это чаще называется в литературе), классификации белков.

- Наиболее полная классификация структур белков представлена в компьютерных классификаторах “Dali/FSSP”, сделанном Л. Холм и К. Сандером, “CATH” (Class — Architecture — Topology — Homology), сделанном в группе Дж. Торнтон, и в самом, пожалуй, популярном из них — “SCOP” (Structural Classification of Proteins), сделанном А. Г. Мурзиным после его переезда из Пушкино в Кембридж.

Классификация относится, собственно, к строению доменов, т. е. компактных глобул, существующих либо самих по себе, либо в составе многодоменного белка. Она начинается (см. рис. 15-1) со структурного *класса*

домена ( $\alpha$ ,  $\beta$  и т. д.). Классы подразделяются по архитектурам (точнее — архитектурным типам) каркаса, сложенного из  $\alpha$ - и/или  $\beta$ -участков; а архитектуры разделяются по топологиям прохождения цепи через каркас (т. е. по мотивом укладки цепи).



**Рис. 15-1.** Структурные классы белков ( $\alpha$ ,  $\beta$ ,  $\alpha$ - $\beta$  и “Low Secondary Structure”, т. е. «со скудной вторичной структурой»), типичные архитектуры (“Non-bundle”, т. е. «не-пучки», “Bundle”, т. е. «пучки», “Ribbon”, т. е. «ленты», “Barrel”, т. е. «цилиндры», «Sandwich», т. е. «сэндвичи», “Roll”, т. е. «рулоны», “2-layer” и “3-layer sandwich”, т. е. «двухслойные» и «трехслойные сэндвичи», и “Complex”, т. е. «сложные [для описания]»), и типичные мотивы укладки цепи (топологии) по физической классификации белков (CATH): глобиновый мотив; пучок типа «вверх-вниз»; ОБ-укладка; иммуноглобулиновый мотив; рулет; трилистник; УБ-рулон;  $\alpha/\beta$ -цилиндр;  $\alpha/\beta$ -складка; двойная обмотка. Ширина сектора отвечает распространенности структур данного типа в негомологических белках. Обратите внимание на слоистую упаковку  $\alpha$ - и  $\beta$ -структур и на то, что каждый слой сложен либо только из  $\alpha$ -спиралей, либо только из  $\beta$ -тяжей, но не из  $\alpha$ -спиралей и  $\beta$ -тяжей одновременно. Картинка, с внесенными в нее (с любезного согласия С. А. Orengo и J. M. Thornton) небольшими изменениями, взята из С. А. Orengo, A. D. Michie, S. Jones, D. T. Jones, M. B. Swindells & J. M. Thornton, *Structure* (1997) 5, 1093

И уже далее мотивы подразделяются на суперсемейства, где просматривается хоть какая-то гомология (след общего происхождения) последовательностей, суперсемейства — на семейства с явно проявляющейся гомологией, и т. д. вплоть до отдельных белков конкретных организмов.

Физическая классификация белковых структур (класс—архитектура—топология) дает возможность не только систематизировать структуры уже изученные, но и предсказывать будущие открытия белковых структур. Так, долгое время были известны только такие  $\beta$ -белки, которые сложены из антипараллельной  $\beta$ -структуры, а на месте  $\beta$ -белков, сложенных из параллельной  $\beta$ -структуры зияла лакуна. Однако потом она заполнилась (помните  $\beta$ -призмы?).

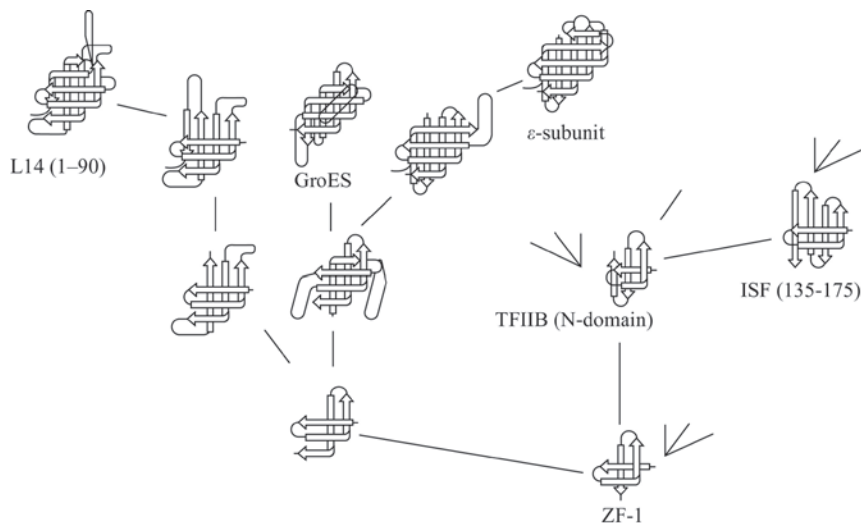
■ Рисунок 15-2 показывает пример такой классификации структур белковых глобул, которая оставляет «вакансии» для возможных, но (еще?) не обнаруженных укладок цепи. Любопытно, что эта классификация («деревья» А. В. Ефимова) основана на постепенном росте и усложнении небольших исходных «заправочных» структур разных типов. Эта схема может интерпретироваться в терминах имитации сворачивания белка (отмечу, что современное представление об этом процессе, полученное из совершенно других экспериментальных и теоретических данных, — об этом речь пойдет через несколько лекций — базируется на нуклеации сворачивания глобулы образованием ее небольшой части). Однако эти же «деревья» могут интерпретироваться в терминах имитации эволюционной истории белков.

■ Вопрос, который всегда волнует биолога, видим ли мы эволюцию структур белков?

В сущности, этот вопрос распадается на два: (1) Видим ли мы микроскопическую эволюцию белков, т. е. не просто «дрейф», изменение белка от организма к организму, но такое его изменение, которое как-то связано с изменением всего организма? (2) Видим ли мы «макроскопическую» эволюцию белков, т. е. усложнение их структуры с усложнением структуры организма?

■ На первый вопрос мы можем смело ответить — да. Хотя далеко не все изменения белка играют ясную функциональную роль (что подчеркивается «нейтральной теорией эволюции» Кимуры), однако в некоторых случаях функциональная роль изменений белка понята и хорошо изучена. Например, гемоглобин ламы (высокогорного животного) сильнее связывает кислород, чем гемоглобины ее родственников, живущих на равнине. Еще явственнее такая подстройка под условия существования наблюдается при сравнении гемоглобинов взрослого животного и плода: последние должны забирать кислород у матери, т. е. связывать его сильнее.

И М. Перутц показал, какие микроскопические изменения в структуре гемоглобина ответственны за это усиление связывания.



**Рис. 15-2.** Часть «структурного дерева» (Efimov A. V., *FEBS Letters* (1997) 407: 37–46), основанного на постепенном росте и усложнении небольших исходных «затравочных» структур. Обратите внимание на «эффект матрешки» — содержание более простой структуры внутри более сложной. Показанный фрагмент ветви относится к той части β-белков, в основе которых лежит β-шпилька, согнутая под прямым углом. На этом фрагменте отмечены уже найденные в природе структуры; остальные пока вакантны

Считается, что эволюция часто протекает путем амплификации гена и последующих мутаций его копий, так что одна копия этого гена до поры до времени поддерживает «старую» функцию (и жизнь организма), а другая или другие получают свободу мутировать в (случайных!) поисках такого изменения, которое бы приспособило функцию белка к биологической потребности. Так, α-лактальбумин молока явно произошел из лизоцима при возникновении млекопитающих. При этом известно, что обычно существует лишь одна копия гена каждого мажорного белка (точнее, две одинаковые — с учетом диплоидности), — но под давлением среды картина меняется: «почти летальная» доза яда может способствовать размножению копий гена, ответственного за элиминацию этого яда. А далее в дело вступают случайные мутации этих копий — и отбор...

Эволюции белков способствует их доменная структура. Известно, что гены доменов, как целое, могут кочевать из белка в белок, порой объ-

единяясь в разных сочетаниях с другими, порой разъединяясь: близко родственные домены часто наблюдаются в разных белках, а также по отдельности (примеры: кальций-связывающий домен калмодулина, парвальбумина и т. д.; разнообразные крендельные (kringle) домены; и т. п.). По-видимому, таким перетасовкам генов способствует их интрон-экзонная структура (в частности, это хорошо видно в иммуноглобулинах); однако гипотеза о том, что «модулем» перетасовки вообще является скорее экзон, чем целый домен, по всей видимости, не подтвердилась.

■ А вот на второй вопрос — видим ли мы «макроскопическую» эволюцию пространственных структур белков, т. е. усложнение структуры белков с усложнением организма — следует, по всей видимости, ответить отрицательно. Обзор белковых структур показывает, что одни и те же мотивы укладки цепи (в частности, изображенные на рис. 15-1) присутствуют и в высших организмах (многоклеточных и одноклеточных эукариотах), и в бактериях-прокариотах (хотя частота встречаемости наиболее «популярных» мотивов несколько отличается у эукариот, с одной стороны, и у прокариот — с другой). То есть мы не наблюдаем у белков того их усложнения с усложнением организма, которое видим, например, на уровне устройства клеток, хроматина и органелл — вплоть до рибосом (даже наоборот: мы видим, что наиболее простые по структуре фибриллярные белки более типичны для высших организмов, чем для прокариот, в особенности археобактерий).

Есть, правда, некие намеки на то, что белки и домены белков у эукариот чуть больше, чем белки прокариот, имеющие тот же мотив укладки цепи, но этот вопрос еще недостаточно выяснен.

Есть, впрочем, и еще одно важное «макроскопическое» структурное отличие — не касающееся, однако, мотивов укладки цепи. Белки эукариот, особенно многоклеточных, гораздо больше подвержены ко- и посттрансляционной химической модификации (типа гликозилирования, иодирования, и т. д., и т. п.). При этом места модификации намечаются первичной структурой белка, а сама модификация производится специальными ферментами; причем часто проводится лишь частично — что ведет к возникновению разнообразия форм тех или иных белков, хотя обычно не меняет их активность. И еще — альтернативный сплайсинг, присущий эукариотам, также повышает разнообразие белков.

**Внутренний голос:** И все же есть данные о том, что белки эукариот не только крупнее, но и «многодоменнее», чем белки прокариот (типичный белок эукариот состоит из четырех-пяти доменов, прокариот — из двух).

**Лектор:** Это так. Однако, возможно, что общий вывод об укрупнении белков эукариот связан более с возникновением у высших организмов больших многодоменных «внешних» белков типа иммуноглобулина,

чем с изменением белков, связанных с внутренним хозяйством клетки. И еще надо сказать, что вопрос о «макроэволюции» мотивов укладок белковых цепей очень затруднен возможностью горизонтального переноса генов, в результате которого «новые» белки могут попадать в «старые» организмы.

**Внутренний голос:** Я бы хотел вернуться к оставленному Вами в стороне «дрейфу» белковых структур, а также к вопросу об их происхождении. Существует гипотеза, что «прыгающими элементами» при эволюции белков могут быть не целые домены, а последовательности из ~10 остатков. Ведь архитектуры белков, обладающих заметным сходством первичных структур, довольно часто отличаются лишь добавленным (или выброшенным, или переложеным в другое место глобулы) кусочком цепи. А так как этот кусочек часто содержит  $\alpha$ - или  $\beta$ -структуру, то явно родственные белковые домены порой относятся к разным мотивам укладки цепи. Так что «дрейф» белков включает и переходы от одного мотива к другому. А раз так — не возникли ли белки из ассоциатов таких мелких модулей?

**Лектор:** Время от времени эти вопросы дискутируются вновь и вновь. Определенно ответить на них до сих пор не удалось. Обсуждаемые структурные модули столь малы, а их последовательности столь широко варьируют, что доказать их родство невозможно. Достаточно сходны только короткие последовательности, связанные с определенными функциями (например — со связыванием гема). Однако сходство функциональных центров порой наблюдается и тогда, когда они (как, например, центры катализа гидролиза) образованы далекими по цепи остатками в белках, абсолютно разных по мотиву укладки цепи. В последнем же случае трудно говорить о «переносе» этих центров из одного белка в другой. Так что, возвращаясь к локальным функциональным и структурным модулям: возможно, они переносятся из белка в белок; возможно, возникают в каждом семействе белков заново.

Точнее, однако, будет сказать, что на эти вопросы не удалось ответить для глобулярных белков. А вот последовательности фибриллярных белков, как о том уже не раз говорилось на лекциях, действительно, выглядят как многократные повторы коротких фрагментов. И гипотеза о «модульном» происхождении фибриллярных белков выглядит вполне обоснованной — тем более, что повторяющиеся структурные модули таких белков часто кодируются отдельными экзонами.

Впрочем, есть все основания надеяться, что геномные и особенно структурно-геномные программы, столь быстро прогрессирующие, дадут ответ на все эти волнующие вопросы.

■ Имея огромное упорядоченное досье белковых структур, мы можем теперь задаться философскими вопросами: (1) В чем физическая причи-

на простоты и регулярности типичных мотивов укладки белковой цепи? и (2) Почему одни и те же мотивы встречаются в самых разных белках, и чем замечательны именно эти мотивы?

В переводе на более научный язык мы хотим выяснить, какие мотивы укладки цепи выводятся из изученной нами физики белковых молекул как наиболее вероятные, много ли их и совпадают ли они с теми мотивами, что наблюдаются в природных белках. В поисках ответа мы исследуем прежде всего стабильность различных структур. Такой подход — изучение в первую голову именно стабильности, а не хода процесса самоорганизации — оправдывается тем, что одни и те же пространственные структуры белков могут быть получены в результате кинетически совсем разных процессов. Они получаются и *in vivo* (как в процессе биосинтеза белка на рибосоме, так в процессе транслокации — в более или менее развернутом виде — белка через мембрану), и *in vitro*, при сворачивании (ренатурации) целой белковой цепи из развернутого состояния. Это значит, что детальная последовательность действий не играет решающей роли при сворачивании белка.

Начнем с простого вопроса — почему существует слоевое строение глобулярных белков, о котором мы говорили на прошлой лекции. Иными словами — посмотрим, почему стабильность плотной глобулы требует, чтобы каркас белковой молекулы выглядел как компактная упаковка  $\alpha$ - и  $\beta$ -слоев, чтобы  $\alpha$ - и  $\beta$ -участки шли от одного края глобулы до другого и чтобы нерегулярные участки *не* лежали внутри глобулы.

В общем, мы об этом уже говорили. Тут все дело в водородных связях, которые стоят дорого, а потому все должны быть насыщены в стабильной структуре. Доноры и акцепторы таких связей есть в пептидной группе каждого аминокислотного остатка. Насытиться они могут или водой, или при образовании вторичной структуры. Поэтому только вторичные структуры — а не нерегулярные петли — стабильной (если она хочет быть таковой) глобулы имеют право не контактировать с водой и лежать внутри глобулы. Элементы же, содержащие свободные (от внутримолекулярных водородных связей) NH- и CO-группы, — петли, края  $\beta$ -листов и концы  $\alpha$ -спиралей — должны быть на поверхности.

Вытянутые  $\alpha$ - и  $\beta$ -структурные участки должны, ради стабильности глобулы, со всех сторон плотно окружать гидрофобное ядро, создаваемое боковыми группами этих участков, и тем самым отделять его от воды. В то же время  $\alpha$ -спираль и  $\beta$ -листы не могут смешиваться в одном слое — пропадут водородные связи края  $\beta$ -листа. Значит, стабильность глобулы требует образования  $\beta$ -слоев и, отдельно,  $\alpha$ -слоев (см. рис. 15-1). Иными словами, отдельные  $\alpha$ - и  $\beta$ -слои являются стабильными элементами глобулярной белковой структуры, а смешение в одном слое  $\alpha$ - и  $\beta$ -структур





предназначенные находиться в мембране, перемежаются с более гидрофильными, предназначенными формировать торчащие из мембраны петли и целые домены), ни *периодичности*, характерной для белков фибриллярных (с их огромными регулярными вторичными структурами).

**Внутренний голос:** Не могу не отметить, что зашифрованное послание может тоже выглядеть как случайная последовательность букв, хотя такой отнюдь не является...

**Лектор:** Конечно, аминокислотные последовательности глобулярных белков не являются истинно случайными (в смысле — какую последовательность ни возьми, из нее получится глобулярный белок). Конечно, эти последовательности отобраны на предмет создания стабильных белковых глобул. Но — самых разных по форме глобул. Так что в наборе их первичных структур присутствует весь спектр закономерностей, присущих каждой из этих глобул (т. е. огромный набор разных шифров). И когда мы называем эти первичные структуры «случайными» (точнее, «квазислучайными»), мы имеем в виду только то, что следы отбора белок-образующих аминокислотных последовательностей во всей совокупности первичных структур глобулярных белков не прослеживаются так явно (и потому не ограничивают множество возможных последовательностей так сильно), как следы отбора на периодичность в фибриллярных белках или отбора на блочность в мембранных. Именно в этом смысле мы и говорим, что аминокислотные последовательности водорастворимых глобулярных белков выглядят примерно так же, как случайные последовательности.

А что такое «выглядеть как случайная последовательность»? Это значит — выглядеть как *большинство* из всех возможных последовательностей... Значит, рассматривая водорастворимые глобулярные белки, вполне осмысленно ставить вопрос о том, какие стабильные пространственные структуры обычно кодируются самыми массовыми, случайными или похожими на них («квазислучайными») последовательностями.

И еще. Если в белковой глобуле есть «структурный дефект» — например, погружение нерегулярной петли или края  $\beta$ -листа в гидрофобное ядро — то обеспечение стабильности такой глобулы требует особо тщательного подбора аминокислотной последовательности (чтобы побольше взаимодействий в ней работали на стабильность). Причем подбор должен быть тем тщательнее, чем сильнее «энергетический дефект» белковой структуры. А если дефектов нет, нет нужды в столь уж тщательном подборе. То есть «бездефектную» глобулу может стабилизировать довольно много последовательностей, глобулу с небольшим дефектом — немного, а с большим — исчезающе малое число последовательностей.

И, если принимать во внимание только физику, те структуры, которые кодируются большим числом последовательностей, должны наблюдаться часто, а те, что малым, — редко. Так может возникнуть «физический отбор» белковых структур.

А так как мы выяснили, что типичные упаковки, «штабеля» вторичных структур в глобулярных белках (см. рис. 15-1) выглядят именно так, как должны выглядеть стабильные упаковки случайных или почти случайных аминокислотных последовательностей, то мы тем самым убедились, что — по крайней мере на уровне *упаковок* — наблюдаемый результат их естественного (биологического) отбора не противоречит отбору физическому.

■ Пойдем дальше и рассмотрим *мотивы укладок* белковых цепей — рассмотрим с той же точки зрения, с точки зрения энергетических дефектов и физического отбора преимущественно «бездефектных» (и потому «годных» для многих последовательностей) структур.

Как мы уже видели, мотивы укладок белковых цепей часто удивительно красивы. Ход белковых цепей часто напоминает линии, орнаментирующие керамику (рис. 15-5). И, по глубокой мысли Джейн Ричардсон, открывшей это сходство, оно не случайно, так как и линия орнамента, и белковая цепь «решает» *одну и ту же* задачу — окружить объем (в белке это центр глобулы, ее гидрофобное ядро), избежав самопересечений этой линии.

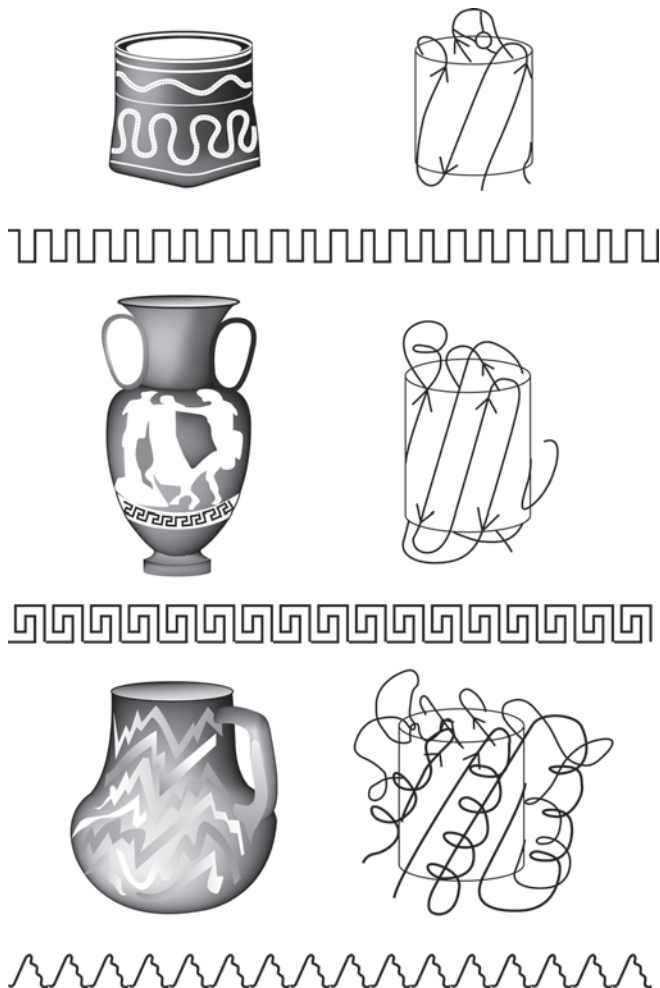
В белках такой эффект достигается тем, что структурные участки уложены вокруг ядра (или двух ядер; последнее типично для  $\alpha/\beta$ -белков), а петли скользят по поверхности ядер, соединяя, как правило, *антипараллельные* (а не параллельные) друг другу соседние участки вторичной структуры (рис. 15-6а) и *не перекрывая* друг друга (рис. 15-6б). Отмечу, что последнее практически исключает и появление узлов в белковой цепи.

Чем же параллельные соединения петлями хуже антипараллельных? Может быть, тем, что это требует слишком длинной и нерегулярной (т. е. не поддерживаемой водородными связями) петли? А, может быть, тем, что довольно жесткую полипептидную цепь приходится сильнее изгибать и тратить на это энергию?

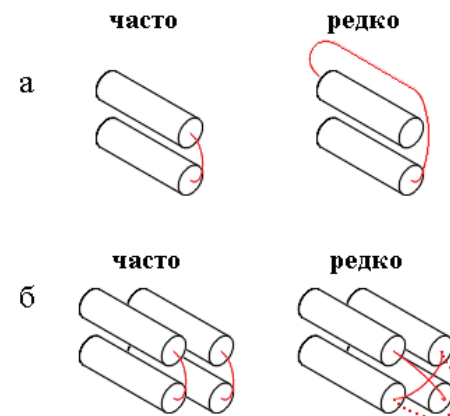
А чем плохи перекрывания петель — ведь при этом цепи не врезаются друг в друга, а просто одна из них проходит над другой? Может быть, тем, что «нижняя», прижатая к ядру петля лишается водородных связей с водой? А чтобы залечить эту потерю (этот «энергетический дефект»), опять нужны «редкие» последовательности...

Надо отметить, что белковые структуры с относительно небольшими дефектами типа перекрывания петель все же встречаются в природе (в отличие от структур с большими дефектами, типа смешения  $\alpha$ - и  $\beta$ -слоев).

Но они встречаются весьма редко — и эта редкость дефектов тем более знаменательна, что уложить цепь в «структуру с дефектом» можно гораздо большим числом способов, чем в «бездефектную» структуру.



**Рис. 15-5.** Мотивы укладки белковой цепи и орнаменты на индейских и греческих вазах: два решения задачи окружения объема несамопересекающейся линией. Вверху: мотив меандра; в середине: мотив греческого ключа; внизу: мотив зигзага-«молнии». Рисунок взят (с разрешения) с обложки Nature, v. 268, No. 5620, 1977 (© 1977, Macmillan Magazines Limited), где была напечатана статья J. Richardson о мотивах укладки белковых цепей



**Рис. 15-6.** Петли, как правило, соединяют *антипараллельные* (а не параллельные) друг другу соседние участки вторичной структуры (а), причем перекрывание петель редко наблюдается в белках, будь то проход одной петли над другой или обход одной петли вокруг другой (б)

Итак, «структурные дефекты» сильно влияют на встречаемость разных белковых структур в природе. Здесь, правда, нас должно смутить то, что при «дефекте» типа петель пропадет лишь одна, максимум две водородные связи, т. е. энергии потеряется немного, килокалорий три или пять. Это не только много меньше, чем полная энергия взаимодействий в белке, измеряемая обычно сотнями килокалорий (судя по опытам по плавлению белков), но и заметно меньше, чем обычный «запас стабильности» белка (т. е. разность свободных энергий его нативного и денатурированного состояния). В нативных условиях этот запас составляет (по тем же опытам) порядка 10 ккал/моль. Почему же «дефект» ценой всего в 5 ккал/моль запрещает, или почти запрещает, перекрывание петель в нативных белковых глобулах?

И еще вопрос: что мешает сделать в петле дополнительный изгиб (пунктир на рис. 15-6) и тем избежать физического перекрывания одной петли другой (т. е. заменить перекрывание петель обходом)? Может быть, здесь дело опять в упругости полимерной цепи — ведь за дополнительный изгиб петле пришлось бы заплатить (как показывает расчет) несколько (все же несколько!) ккал/моль?

Отложим ответ на эти вопросы, а пока рассмотрим еще одну характерную черту белковых архитектур — то, что перемычка между параллельными  $\beta$ -участками почти всегда образует с ними *правозакрученную*, а не *левозакрученную* спираль (рис. 15-7).



**Рис. 15-7.** Левовинтовой ход перемычек между параллельными  $\beta$ -участками очень редко наблюдается в белках, правовинтовой — часто. [Перемычки, изображенные здесь просто веревочками, на самом деле обычно включают в себя участок  $\alpha$ -или  $\beta$ -структуры.] При правовинтовом ходе цепь, приближаясь к нам (смотрящим вдоль оси витка), движется против часовой стрелки; при левовинтовом — по часовой стрелке

Критерий стабильности позволяет и в этом случае выделить «лучшую» из двух зеркально-симметричных белковых архитектур. В основе различия лежит зеркальная асимметрия природных аминокислот. Она приводит, как вы помните, к преимущественно *правому* (показанному на рис. 15-7) скручиванию  $\beta$ -слоев, состоящих из L-аминокислот. При этом угол между осями соседних  $\beta$ -участков близок к  $30^\circ$ , так что полный угол поворота близок к  $330^\circ$  для правовинтовой перемычки и к  $390^\circ$  — для левовинтовой. В результате, из-за жесткости полипептидной цепи, правовинтовая перемычка выгоднее, ее свободная энергия ниже — но опять немного, на пару ккал/моль.

Здесь любой человек, знающий физику полимеров, должен прервать меня и сказать: «Упругость полимера — энтропийный, а вовсе не энергетический эффект! То есть сильно изогнутая цепь не может флуктуировать так свободно, как прямая или слабо изогнутая. Иными словами, с сильно изогнутой формой цепи совместимо гораздо меньше конформаций, чем с ее более или менее вытянутой формой. Однако эффект, о котором Вы говорите, относится к флуктуирующей цепи, т. е. к цепи, *не* фиксированной в глобуле. Но в нативном белке цепь фиксирована, и — так ли она идет, иначе ли — она все равно будет иметь лишь одну какую-то конформацию. Какое отношение имеют описываемые Вами энтропийные потери к нативной структуре белка, где цепь все равно фиксирована, т. е. все равно имеет нулевую энтропию?»

Ответить на этот вопрос — об энтропийных, по видимости, дефектах слишком крутых поворотов цепи — не очень сложно. Дадим его, пока что, на качественном уровне. Здесь дело в том, что «зажатость» конформации цепи лишает ее возможности подстроиться под слагающую ее аминокислотную последовательность. Такая конформация будет иметь низкую энергию

только для строго определенной последовательности (или узкой группы последовательностей) и высокую — для других, т. е. для большинства. Таким образом энтропийный дефект преобразуется в энергетический. А вот если цепь может выбирать из большого набора конформаций, то почти каждая последовательность найдет там что-то для себя приемлемое.

На следующей лекции мы рассмотрим вопрос об энергетических и энтропийных дефектах более строго. Рассмотрим мы и оставшийся у нас вопрос о том, почему «дефект» ценой всего в несколько килокалорий на моль играет столь существенную роль для встречаемости белковых структур.

А пока, вместо ответа на последний вопрос, но в качестве намека, посмотрим, как связаны с энергией другие статистические закономерности, отмеченные в белковых структурах. Например, погруженность в глобулу гидрофобных и гидрофильных боковых групп, разные углы вращения в цепи и т. д. Оказывается, точно так же, как то отмечено для мотивов укладки белковой цепи: «дефекты» встречаются редко, а «хорошие детали» часто! Но здесь мы располагаем большей статистикой и можем получить не только качественные, но и количественные оценки.

Все эти оценки сводятся к следующему. И пониженная встречаемость «дефектов» (будь то «плохой» угол внутреннего вращения или сильная погруженность в глобулу полярной группы), и повышенная встречаемость «хороших деталей» (например, солевых мостиков, образуемых группами с противоположными зарядами) описывается общей феноменологической формулой

$$\text{ВСТРЕЧАЕМОСТЬ} \sim \exp(-\text{ЭНЕРГИЯ\_ДЕТАЛИ}/kT_c), \quad (15.1)$$

где  $T_c$  — какая-то температура, близкая не то к комнатной температуре, не то к характерной температуре плавления белка: ведь 300 К и 370 К совпадают «по порядку величины».

Формула (15.1) для встречаемости структурных деталей в белках удивительно похожа на статистику Больцмана, правда, похожа только *по своей экспоненциальной форме*, а *не* по физическому смыслу. Стоит напомнить, что статистика Больцмана поддерживается тем, что частицы бродят с места на место, и каждая из них больше времени проводит там, где ее энергия ниже. В то же время в наблюдаемых (нативных) белковых глобулах все структурные детали закреплены, они *не* возникают и исчезают, *не* бродят с места на место!

То есть обычная Больцмановская статистика здесь, в применении к встречаемости деталей белковой структуры, вовсе ни при чем. А если так, почему статистика встречаемости этих деталей в белках имеет такую знакомую, «квази-Больцмановскую» форму?



Отложим обстоятельный ответ и на этот вопрос до следующей лекции, а пока примем как феноменологический факт, что дефект ценой в  $kT_c$ , то есть примерно в 1 ккал/моль, в несколько раз снижает наблюдаемость структур, отличающихся этим дефектом. И, как мы уже понимаем, это означает, что такой дефект снижает в несколько раз число аминокислотных последовательностей, придающих стабильность белку с этим дефектом.

Итак, мы приходим к следующим выводам:

1. «Популярные» мотивы укладки белковой цепи выглядят так «стандартно», т. е. так просто и регулярно потому, что каркас белковой структуры представляет собой компактную упаковку слоев вытянутых стандартных твердых тел ( $\alpha$ -спиралей и  $\beta$ -участков), а нерегулярные переемы идут по поверхности глобулы, не пересекая ни друг друга, ни торцов структурных сегментов. Физическая причина такого устройства в том, что оно наиболее благоприятствует стабильности нативной глобулы, позволяя неполярным боковым группам укрываться от воды, а всем пептидным группам главной цепи — насытить свои водородные связи даже при погружении в компактную глобулу.

2. Число таких «стандартных» стабильных мотивов укладки цепи относительно невелико (порядка сотен, а белков — многие десятки тысяч); неудивительно поэтому, что некоторые из этих «стандартных» структур встречаются в разных, со всех остальных точек зрения, белках.

3. В то же время и другие, «дефектные» мотивы укладки белковой цепи не запрещены, они просто маловероятны (т. е. лишь относительно небольшое число последовательностей может придать им стабильность), и тем менее вероятны, чем сильнее дефект.

4. По-видимому, для структур доменов водорастворимых глобулярных белков, с их «квазислучайными» аминокислотными последовательностями, можно сформулировать **«принцип множественности»**: *чем больше аминокислотных последовательностей можно вписать в данную архитектуру без разрушения ее стабильности, тем чаще эта архитектура встречается в природе.*

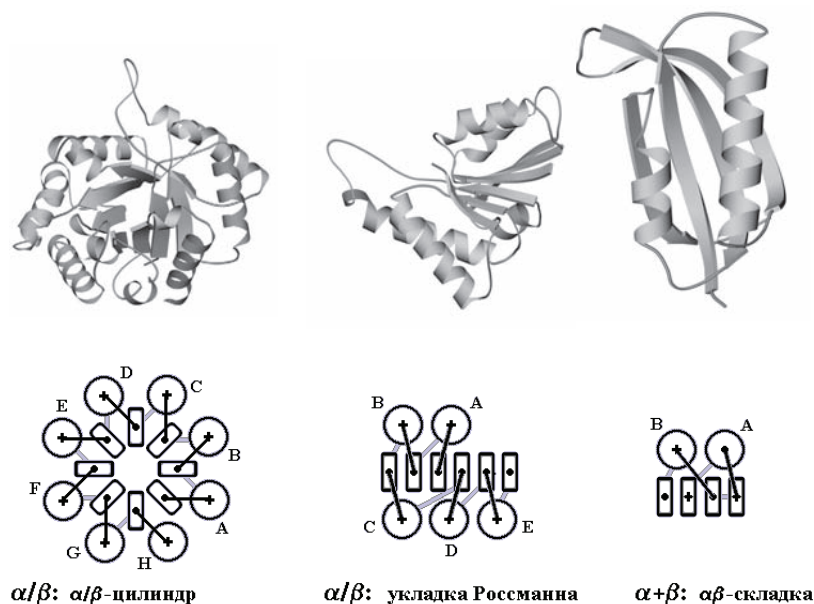
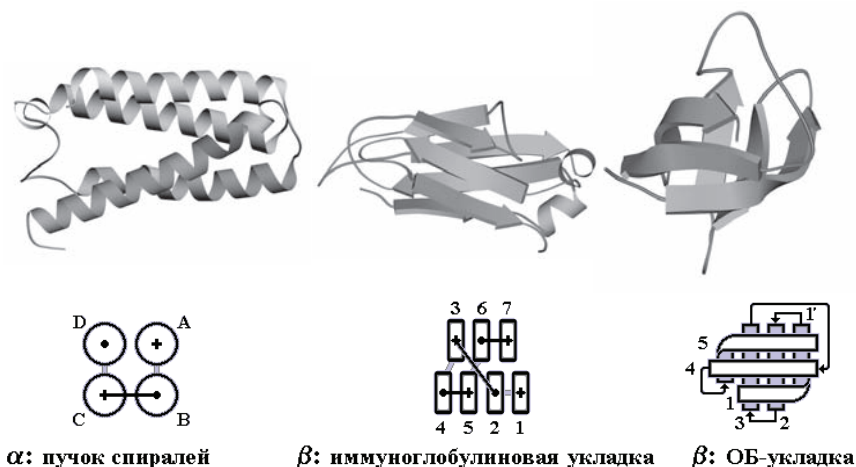
## Лекция 16

Какую вторичную структуру можно ожидать для случайных и квазислучайных аминокислотных последовательностей? Доменное строение стабильных пространственных структур длинных квазислучайных последовательностей наиболее вероятно. Квази-Больцмановская статистика мелких деталей белковых структур. Она возникает из физического отбора стабильных белковых структур. Влияние стабильности структурного элемента на строгость отбора первичных структур, не разрушающих пространственную структуру глобулярного белка, или: почему одни белковые структуры встречаются часто, а другие — редко? Какую структуру —  $\alpha$  или  $\beta$  — следует чаще ожидать в центре большой глобулы? Связь «энтропийных» дефектов с «энергетическими». Глобулярные белки возникли как «отобранные» случайные полипептиды? Отбор «белковоподобных» случайных последовательностей в белковой инженерии.

Теперь мы более подробно поговорим о связи общих закономерностей, наблюдаемых в структурах белков, со стабильностью этих структур и с числом кодирующих их аминокислотных последовательностей.

Посмотрим на рис. 16-1, чтобы вспомнить, что каркас типичной белковой глобулы представляет собой компактную упаковку слоев вытянутых регулярных твердых тел ( $\alpha$ -спиралей и  $\beta$ -участков), а нерегулярные переемы идут по поверхности глобулы, не пересекая ни друг друга, ни торцов структурных сегментов.

Мы говорили, что физическая причина такого устройства белка в том, что оно наиболее благоприятствует стабильности глобулы, позволяя неполярным боковым группам укрыться от воды, а пептидным группам главной цепи — насытить свои водородные связи даже при погружении в компактную глобулу. А это «благоприятствование стабильности», в свою очередь, позволяет большему числу разнообразных аминокислотных последовательностей вписаться в данную архитектуру глобулы без разрушения ее стабильности (что и приводит к более частому наблюдению этой архитектуры в природе — мы назвали это **«принципом множественности»**).

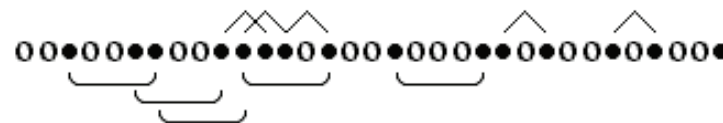


**Рис. 16-1.** Характерные мотивы укладки белковой цепи в  $\alpha$ -,  $\beta$ -,  $\alpha/\beta$ - и  $\alpha + \beta$ -белках. Внизу — их упрощенные схемы. Обратите внимание на слоевую упаковку  $\alpha$ - и  $\beta$ -структур и на то, что каждый слой сложен либо только из  $\alpha$ -спиралей, либо только из  $\beta$ -тяжей, но не из  $\alpha$ -спиралей и  $\beta$ -тяжей одновременно

Как мы уже отмечали, «своеобразие» типичных водорастворимых глобулярных белков состоит в отсутствии каких-либо общих ярких черт в их первичных структурах (и потому-то эти последовательности наиболее разнообразны и многочисленны, что и позволяет рассматривать их с точки зрения «принципа множественности»). Действительно, в их первичных структурах не прослеживается каких-либо явных корреляций: ни периодичности, характерной для фибриллярных белков, ни чередования блоков, характерного для белков мембранных. Наоборот, полярные группы в водорастворимых глобулярных белках довольно равномерно «перемешаны» с неполярными. К тому же эти белки содержат примерно поровну полярных и неполярных аминокислотных остатков. В результате первичные структуры водорастворимых глобулярных белков больше всего напоминают наугад синтезированные «случайные сополимеры» из гидрофобных и гидрофильных аминокислот.

■ Но совместимы ли «случайные» последовательности с компактной укладкой цепи в глобулу? Совместимы ли они с наблюдаемыми в белках (и обычно охватывающими в сумме добрую половину белка) вторичными структурами? Для ответа на эти вопросы рассмотрим гипотетические «белковые цепи», получаемые случайной сополимеризацией равного числа полярных и неполярных групп.

Для того, чтобы  $\alpha$ - или  $\beta$ -структурный сегмент мог уложиться в компактную глобулу, он должен обладать непрерывной гидрофобной поверхностью. Такая поверхность образуется неполярными боковыми группами, занимающими в цепи позиции  $i - i + 4$  (или  $i - i + 3$ ) друг от друга в  $\alpha$ -спирали и  $i - i + 2$  в  $\beta$ -участке. Легко показать, что даже в случайном сополимере имеется достаточно периодических сгущений неполярных групп (рис. 16-2) для образования гидрофобных поверхностей  $\alpha$ - и  $\beta$ -сегментов, способных уложиться в глобулу, характерную для белковых доменов среднего размера.



**Рис. 16-2.** Характерный мотив чередования гидрофобных (•) и полярных (o) аминокислот в «квазислучайной» первичной структуре водорастворимых глобулярных белков. Дуги и уголки указывают на расположение в этой первичной структуре потенциальных гидрофобных поверхностей, годных для  $\alpha$ -спиральных и  $\beta$ -структурных участков, т. е. пар гидрофобных остатков в положениях  $i - i + 4$  и  $i - i + 2$ , соответственно

Пусть  $p$  — доля неполярных, и  $1-p$  — доля полярных групп в сополимере. Вероятность того, что в данной точке цепи начинается периодическая последовательность из ровно  $r$  неполярных групп (и ограниченная, следовательно, с начала и с конца двумя полярными) есть

$$W(r) = (1-p)p^r (1-p). \quad (16.1)$$

О «гидрофобной поверхности»  $\alpha$ - или  $\beta$ -сегмента имеет смысл говорить, когда она охватывает больше одного витка соответствующей структуры. Иначе говоря, в гидрофобной поверхности  $\alpha$ -спирали  $r \geq 2$  гидрофобных групп расположены в периодических позициях  $i - i + 4 - \dots$  (на стоящие между ними остатки мы внимания не обращаем, так как они не входят в рассматриваемую поверхность спирали). В гидрофобной поверхности  $\beta$ -тяжа  $r \geq 2$  гидрофобных групп расположены в позициях  $i - i + 2 - \dots$  (см. рис. 16-2). Применяя формулу (16.1) к таким периодическим позициям, мы видим, что средне-ожидаемое число витков и  $\alpha$ -, и  $\beta$ -сегмента, покрытое гидрофобной поверхностью, есть

$$\langle r \rangle = \sum_{r \geq 2} [W(r)r] / \sum_{r \geq 2} W(r) = \sum_{r \geq 2} [rp^r] / \sum_{r \geq 2} p^r = 2 + p/(1-p) \quad (16.2)$$

(я позволю себе не останавливаться на суммировании рядов  $\sum_{r \geq 2} [rp^r]$  и  $\sum_{r \geq 2} p^r$  — все это можно найти в математическом справочнике). Итак, и «средняя»  $\alpha$ -спираль, и «средний»  $\beta$ -участок включают (при  $p = 1/2$ )  $\langle r \rangle = 3$  регулярно расположенных гидрофобных групп, т. е.  $3 \pm 0,5$  полных периода  $\alpha$ - или  $\beta$ -структуры ( $\pm 0,5$  — так как гидрофобная группа может быть и в начале, и в середине, и в конце витка). Среднеожидаемые числа остатков в  $\alpha$ - и  $\beta$ -участках (периоды этих участков — 3,6 и 2) суть  $\langle n_\alpha \rangle = 11 \pm 2$  и  $\langle n_\beta \rangle = 6 \pm 1$ , соответственно. Это практически совпадает со средними длинами  $\alpha$ - и  $\beta$ -участков в глобулярных белках. Интересно, что в случайных последовательностях — так же, как и в первичных структурах реальных белков, — на одном и том же участке цепи часто перекрываются (см. рис. 16-2) сгущения неполярных групп, способные образовать гидрофобные поверхности и  $\alpha$ -, и  $\beta$ -участков.

Аналогичные оценки показывают, что средняя длина петель между вторичными структурами в случайном сополимере — величина порядка  $3 + 0,5p^{-2}$ , т. е. при  $p \sim 1/2$  петли в среднем короче, чем участки вторичной структуры, что и наблюдается.

Итак, в случайном сополимере при  $p \sim 1/2$  «сами собой» образуются непрерывные гидрофобные поверхности для  $\alpha$ - и  $\beta$ -участков, способных прилипнуть к гидрофобному ядру по крайней мере одной своей стороной, а петли относительно коротки. Поэтому «заурядные», случайные последовательности вполне могут формировать укладки из по крайней мере двух слоев вторичной структуры.

**Внутренний голос:** Не следует, однако, забывать, что в некоторых, пусть немногих, глобулярных белках (например, в гемагглютинине или в лейциновом зиппере) встречаются чрезвычайно длинные спирали, не укладывающиеся в описанные выше принципы. А в некоторых (в супероксиддисмутазе, например) — чрезвычайно длинные неупорядоченные петли...

**Лектор:** Верно. И при этом они выглядят либо как блоки, позаимствованные из фибриллярных спиральных белков, либо (если речь идет о длинных петлях) как аномально гидрофильные блоки. Но в среднем, массово и  $\alpha$ -, и  $\beta$ -, и нерегулярные участки глобулярных белков невелики и имеют как раз такую длину, какую можно ожидать для «случайной» аминокислотной последовательности с равным содержанием гидрофобных и полярных групп.

■ Гармония между «случайными» последовательностями и компактными, потенциально наиболее стабильными формами глобул существует, пока длина цепи не превосходит  $\sim 150$  остатков. Однако, по мере роста размеров глобулы, по мере роста числа слоев вторичной структуры в ней, ее «годность» для случайной аминокислотной последовательности падает. Ведь участки, расположенные внутри белка, должны быть велики (длиной с диаметр глобулы) и почти целиком сложены из одних только гидрофобных остатков — глобула не вынесет наличия многих гидрофильных групп, укрытых от воды, и взорвется. Небольшое число таких длинных и почти целиком гидрофобных участков первичной структуры тоже может «случайно» образоваться в случайном сополимере из  $\sim 50\%$  гидрофобных и  $\sim 50\%$  гидрофильных остатков — но только небольшое число. Поэтому для случайной последовательности мы можем ожидать наличия только двух-трех, изредка — четырех слоев структурных сегментов, не более, — что и наблюдается в не слишком больших водорастворимых глобулярных белках и в доменах более крупных водорастворимых глобулярных белков (см. рис. 16-1). А крупные белки должны состоять из субглобул, доменов — что и наблюдается.

■ Теперь нам предстоит ответить на два поставленных на прошлой лекции (и пока оставленных без должного ответа) вопроса:

1) Почему «энергетический дефект» ценой всего в несколько килокалорий на моль — на фоне гораздо большей полной энергии белка — может практически запрещать многие мотивы белковых архитектур?

2) Какое отношение имеют «энтропийные» эффекты к нативной структуре белка, где цепь все равно фиксирована?

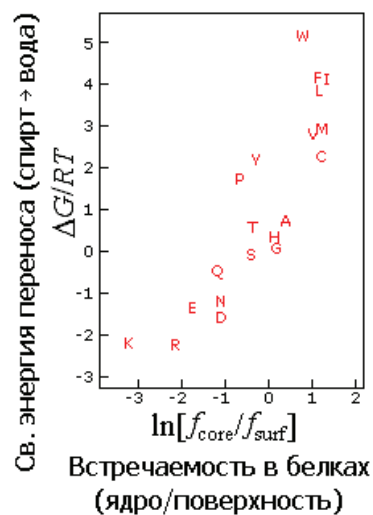
Начнем с первого вопроса — о проявлении энергии «дефекта» в статистике белковых архитектур. Но сначала усугубим его, вспомнив, что отмеченная нами на качественном уровне редкость «структурных дефектов» подкрепляется наблюдаемой квази-Больцмановской (т. е. «похожей

на Больцмановскую») статистикой «малых деталей» белковых структур, которую нам также следует понять. И начнем именно с нее — со статистики «малых деталей».

- Для примера, рассмотрим статистику распределения аминокислотных остатков между внутренностью и поверхностью белковой глобулы и посмотрим, как она связана с гидрофобностью аминокислотных остатков.

Гидрофобность аминокислотных остатков обычно измеряется в свободной энергии их переноса — переноса из октанола, моделирующего гидрофобное ядро белка, в воду. На рис. 16-3 эта свободная энергия переноса (деленная на  $RT$ , где  $T = 300$  К) отложена по вертикальной шкале. По горизонтальной шкале отложен логарифм отношения числа поверхностных и внутренних остатков в белках. Мы видим, что точки более или менее ложатся на прямую линию, и наклон такой прямой более или менее близок к 1–1,5.

**Рис. 16-3.** Экспериментально определенная свободная энергия переноса боковых групп аминокислотных остатков из неполярного растворителя в воду,  $\Delta G$ , выраженная в единицах  $RT$ , и «кажущаяся свободная энергия переноса остатка из ядра на поверхность белка», вычисленная из наблюдаемых частот встречаемости аминокислотных остатков внутри ( $f_{\text{core}}$ ) и на поверхности ( $f_{\text{surf}}$ ) белка по формуле  $\Delta G_i/RT = -\ln [f_{\text{surf}}/f_{\text{core}}]$ . Картинка, с небольшими изменениями, взята из S. Miller, J. Janin, A. M. Lesk, C. Chothia, *J. Mol. Biol.* (1987) **196**:641–656



Таким образом, наблюдаемая статистика распределения остатков между нутром и поверхностью глобулы неплохо описывается формулой

$$\text{ВСТРЕЧАЕМОСТЬ} \sim \exp(-\text{СВОБОДНАЯ\_ЭНЕРГИЯ\_В\_ДАННОЙ\_СРЕДЕ}/kT_c), \quad (16.3)$$

где «конформационная температура»  $T_c$  близка к 300–400 К по порядку величины.

**Внутренний голос:** Я бы рекомендовал быть более осторожным с данными, приведенными на рис. 16-3... Ведь экспериментальная гидрофобность там определяется по переносу аминокислот из воды в высокомолекулярный спирт. Но кто сказал, что гидрофобное ядро белка похоже именно на спирт? Не говоря уже о том, что спирт — жидкий, а ядро — твердое: не является ли, скажем, чисто гидрофобный циклогексан лучшей моделью гидрофобного ядра, чем спирт? А ведь растворимость полярных групп в циклогексане очень низка, что неизбежно сильно *растянет* шкалу гидрофобности рис. 16-3... Верно, корреляция между гидрофобностью и встречаемостью сохранится и в этом случае, но вот наклон проведенной через «циклогексановые» точки интерполяционной прямой окажется много больше, где-то 3–4. И, кстати, именно такой наклон наблюдается в верхней части рис. 16-3, там, где лежат гидрофобные аминокислоты, на которых, к слову, замена спирта на бензол скажется мало...

**Лектор:** Да, это действительно вопрос — чем моделировать гидрофобное ядро белка при экспериментальной оценке гидрофобности... Я бы сказал, что оно все же больше напоминает спирт, чем циклогексан: ведь в ядре белка есть полярные атомы, я имею в виду NH- и CO-группы главной цепи, окружающей ядро [правда, они обычно образуют H-связи в окружающей ядро вторичной структуре, но CO-группа способна образовать две H-связи, «вилочку», и только одна из них используется во вторичной структуре, а вторая остается ненасыщенной]. Так что вряд ли шкалу гидрофобности надо подстраивать под чисто неполярный циклогексан. С другой стороны — если уж речь идет о численных оценках — не надо забывать, что рис. 16-3 сделан при грубом разбиении всех боковых групп на два класса («ядерные» и «поверхностные») по их погруженности в белок; а при таком разбиении экспонированность «ядерных» и «поверхностных» групп отличается не на полную поверхность группы, а примерно на ее половину. Так что экспериментальные гидрофобности надо, соответственно, *уменьшить* раза в два, а это (в противоположность эффекту от замены спирта на циклогексан) уменьшит наклон интерполяционной прямой... В общем, я совершенно согласен, что к всем *численным* оценкам здесь надо относиться со здоровым недоверием, и обращать внимание именно на качественную связь энергетики разных элементов и встречаемости их в белках.

- Итак, статистика встречаемости аминокислотных остатков внутри и на поверхности белка удивительно похожа на статистику Больцмана — *похожа по форме!* Это впервые было замечено Полем в 1971 г. для распределения углов внутреннего вращения в боковых цепях аминокислотных остатков в белках. Потом это было показано и для статистики многих других элементов белковых структур: для встречаемости ионных пар, для встречаемости остатков во вторичных структурах, для встречаемости полостей в белках и т. д., и т. п. К настоящему времени эта аналогия стала столь привычной, что статистика белковых структур часто используется для оценки свободной энергии различных взаимодействий аминокислотных остатков.

Здесь, однако, следует подчеркнуть, что белковая статистика *похожа* на статистику Больцмана именно *по своей экспоненциальной форме*, а *не* по физическому смыслу. Напомню, что статистика Больцмана поддерживается тем, что частицы бродят с места на место, и каждая из них больше



времени проводит там, где ее энергия ниже. В то же время в нативных белках аминокислотные остатки *не* бродят с места на место! Например, Leu72 цепи миоглобина кашалота всегда находится внутри глобулы в нативной структуре этого белка — мы никогда не видим его на поверхности. И если мы видим, по статистике, что 80–85 % всех лейцинов находится внутри белка и 15–20 % на поверхности, то здесь дело не в том, что каждый лейцин проводит 80–85 % времени внутри и 15–20 % времени на поверхности глобулы. Здесь дело в том, что естественный отбор закрепил большинство лейцинов в тех точках цепи, что лежат внутри глобулы.

То есть обычная Больцмановская статистика, статистика флуктуаций в обычном трехмерном пространстве, здесь, в распределении остатков между ядром и поверхностью белка, вовсе ни при чем. В глобуле нет таких флуктуаций, которые бы — в соответствии с гидрофобностью лейцинов — выносили каждый из них на 15–20 % времени на поверхность, и на 80–85 % времени прятали бы его внутрь белка. То есть для каждого **отдельного** лейцина **нет** Больцмановского распределения, диктуемого его (лейцина) гидрофобностью. А если так — чем же объяснить то, что встречаемость **всей совокупности** лейцинов внутри и на поверхности белка соответствует Больцмановскому распределению, диктуемому гидрофобностью лейцина?

Посмотрим с другой стороны. Чем хорошо, скажем, преимущественное расположение лейцинов внутри глобулы? Тем, что оно повышает ее стабильность. Но почему тогда естественный отбор не настоял на том, чтобы все лейцины были внутри белка? Может быть, потому, что уже 80–85 % внутренних лейцинов *достаточно* для стабильности белка, а возня с остальными потребовала бы от него слишком больших усилий?

Оставим вопрос о психологии естественного отбора как бесперспективный и ненаучный и рассмотрим другой вопрос — вопрос о том, как изменяет внутренняя свободная энергия какого-то элемента белковой структуры *число аминокислотных последовательностей*, способных придать стабильность белку с этим структурным элементом.

- Например, посмотрим, как изменится число стабилизирующих структуру белка последовательностей при мутации Leu → Ser во внутренней точке белка.

Нативная структура стабильна, если ее свободная энергия меньше, чем свободная энергия и развернутого белка, и всех других глобулярных укладок его цепи, т. е. если ее свободная энергия лежит ниже какой-то грани, не зависящей от нативной структуры. Будем, для простоты, считать, что нативная структура конкурирует лишь с развернутым состоянием цепи, и что (1) вклад остатка в стабилизацию белка определяется его гидрофобной свободной энергией [что не вполне верно — есть еще, по крайней

мере, водородные связи и необходимость плотной упаковки — но сойдет для «нулевого» приближения...], (2) что внутренние остатки белка целиком укрыты от воды, а поверхностные — целиком ей доступны [что является довольно грубым, но качественно верным допущением], и что денатурированный белок совсем развернут [это бывает часто, но не всегда; поэтому излагаемая ниже теория приближительна, но зато она проста].

Свободная энергия переноса боковой группы серина из гидрофобного окружения в воду — около 0, а лейцина — около +2 ккал/моль. В развернутом денатурированном белке все остатки окружены водой. Значит, если принимать во внимание только гидрофобные взаимодействия, структура белка с серином внутри нативной глобулы на 2 ккал/моль менее стабильна против разворачивания, чем структура с лейцином в той же внутренней точке. Следовательно, грубо говоря, все последовательности, стабилизирующие белок с серином в ядре, будут стабилизировать и белок с лейцином в ядре — но, кроме того, белок с лейцином внутри будут стабилизировать и какие-то последовательности, *не* способные стабилизировать белок с серином в той же точке.

Как упадет число стабилизирующих нативную структуру белка последовательностей при замене более стабильного элемента этой структуры («Leu внутри») на менее стабильный («Ser внутри»)?

Рассмотрим такую задачу формально и попробуем понять, откуда берется наблюдаемая, примерно экспоненциальная зависимость встречаемости разнообразных элементов от их свободной энергии и что за температура  $T_c$  стоит в уравнении (16.3).

Я заранее прошу меня извинить, что все выкладки я дам в самом упрощенном, а значит, не вполне точном виде. Моя цель — дать вам почувствовать, в чем суть дела, не заходя в математические дебри, по которым один из нас (А. В. Ф.) с А. М. Гутиным и А. Я. Бадретдиновым в свое время нагулялся вволю...

- Итак, пусть  $\Delta\varepsilon$  — свободная энергия рассматриваемого элемента в нативном белке, включая его взаимодействия с остальной цепью (например, свободная энергия лейцина или серина в центре нативной глобулы), отсчитанная от его же свободной энергии в денатурированном белке. Пусть  $\Delta F$  — свободная энергия всей остальной цепи в рассматриваемой пространственной структуре, за вычетом ее (этой цепи) свободной энергии в денатурированном белке. Значит,  $\Delta\varepsilon + \Delta F$  — разность свободных энергий нативной и денатурированной форм белка.

Для того чтобы нативный белок был стабилен,  $\Delta F + \Delta\varepsilon$  должно быть меньше 0, т. е. для стабильного белка

$$\Delta F < -\Delta\varepsilon. \quad (16.4)$$

Величины  $\Delta F$  и  $\Delta \epsilon$  зависят от аминокислотной последовательности полипептидной цепи. Рассмотрим множество последовательностей, оставляющих неизменной величину  $\Delta \epsilon$  [в данном случае — все последовательности с инвариантным Leu (или Ser) в данном месте цепи; при этом, если это место лежит в ядре нативного белка,  $\Delta \epsilon = 2$  ккал/моль для Leu и около 0 для Ser, в то время как для всех поверхностных остатков  $\Delta \epsilon$  близко к 0, если денатурированный белок развернут]. Величина же  $\Delta F$  будет меняться от последовательности к последовательности.

При этом вероятность  $P^*$  того, что  $\Delta F < -\Delta \epsilon$ , есть

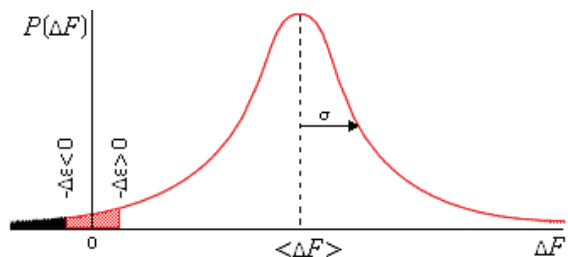
$$P^*(\Delta F < -\Delta \epsilon) = \int_{-\infty}^{-\Delta \epsilon} P(\Delta F) d(\Delta F), \quad (16.5)$$

где  $P(\Delta F)$  — вероятность появления заданной величины  $\Delta F$  в наудачу выбранной случайной последовательности.

Величина  $\Delta F$  складывается из свободных энергий множества разных взаимодействий и из энтропий фиксации множества разных аминокислотных остатков. Так как складываемых «случайных» величин много, то, по «центральной предельной теореме» математической статистики, вероятность  $P$  того, что  $\Delta F$  равна той или иной величине, имеет простую, так называемую Гауссову (см. рис. 16-4) форму:

$$P(\Delta F) = (2\pi\sigma^2)^{-1/2} \exp[-(\langle \Delta F \rangle - \Delta F)^2 / 2\sigma^2]. \quad (16.6)$$

Здесь  $\langle \Delta F \rangle$  — средняя (усредненная по всем последовательностям) величина  $\Delta F$ , а  $\sigma$  — средняя (точнее — среднеквадратичная, усредненная по тем же последовательностям) величина отклонения  $\Delta F$  от среднего  $\langle \Delta F \rangle$ .



**Рис. 16-4.** Типичный (Гауссов) вид распределения величины свободной энергии  $\Delta F$  по случайным аминокислотным последовательностям. В  $\Delta F$  входит вся свободная энергия данной укладки белковой цепи (отсчитанная от свободной энергии денатурированного белка), за исключением фиксированной свободной энергии  $\Delta \epsilon$  интересующего нас элемента структуры. Значения  $\Delta F < -\Delta \epsilon$  (т. е. удовлетворяющие условию  $\Delta F + \Delta \epsilon < 0$ ) отвечают стабильной укладке. Зачерненная область соответствует значениям  $\Delta F < -\Delta \epsilon$  при  $\Delta \epsilon > 0$ , «серая + зачерненная» область — при  $\Delta \epsilon < 0$ . Вторая область больше, т. е. больше случайных последовательностей стабилизирует укладку цепи, если свободная энергия интересующего нас элемента  $\Delta \epsilon < 0$ , чем если  $\Delta \epsilon > 0$

Напомню, что «центральная предельная теорема» описывает ожидаемое распределение суммы *большого числа* случайных величин и отвечает на вопрос: «какова вероятность того, что эта сумма будет иметь то или иное значение?». У нас «сумма большого числа случайных величин» — это  $\Delta F$  (она включает множество взаимодействий в наудачу выбранной, а значит, «случайной» последовательности). Так вот, математика говорит, что для большинства последовательностей величина такой суммы  $\Delta F$  должна лежать между  $\langle \Delta F \rangle - \sigma$  и  $\langle \Delta F \rangle + \sigma$ , и что вероятность того, что сумма всех свободно-энергетических членов равна  $\Delta F$ , круто, экспоненциально падает при удалении  $\Delta F$  от  $\langle \Delta F \rangle$  (см. рис. 16-4). А так как подавляющее большинство случайных последовательностей явно неспособно стабилизировать заданную укладку цепи, то величины  $P$  и  $P^*$  должны быть много, много меньше 1 на интересующей нас «границе стабильности», при  $\Delta F \approx 0$ . Следовательно, величина  $\langle \Delta F \rangle$  должна быть не только положительной, но и *большой* — существенно больше, чем  $\sigma$ .

Величина  $(\langle \Delta F \rangle - \Delta F)^2 = \langle \Delta F \rangle^2 - 2\langle \Delta F \rangle \Delta F + \Delta F^2$ , т. е. при малых (относительно большого  $\langle \Delta F \rangle$ ) значениях  $\Delta F$ , мы имеем  $(\langle \Delta F \rangle - \Delta F)^2 \approx \langle \Delta F \rangle^2 - 2\langle \Delta F \rangle \Delta F$ , и, значит,

$$P(\Delta F) \approx \{(2\pi\sigma^2)^{-1/2} \exp[-\langle \Delta F \rangle^2 / 2\sigma^2]\} \exp[\Delta F(\langle \Delta F \rangle / \sigma^2)]. \quad (16.7)$$

При этом [прошу проверить — или (хуже!) поверить] вероятность того, что  $\Delta F < -\Delta \epsilon$ , есть

$$P^*(\Delta F < -\Delta \epsilon) = \int_{-\infty}^{-\Delta \epsilon} P(\Delta F) d(\Delta F) \approx \text{const} \times \exp[-\Delta \epsilon / (\sigma^2 / \langle \Delta F \rangle)], \quad (16.8)$$

причем константа  $\text{const} = \{(2\pi\sigma^2)^{-1/2} \exp[-\langle \Delta F \rangle^2 / 2\sigma^2]\} (\sigma^2 / \langle \Delta F \rangle)$  нас не интересует, а интересующий нас член  $\exp[-\Delta \epsilon / (\sigma^2 / \langle \Delta F \rangle)]$  демонстрирует, что свободная энергия  $\Delta \epsilon$  любого элемента белковой структуры *экспоненциально* изменяет вероятность того, что наугад выбранная аминокислотная последовательность будет стабилизировать включающую этот элемент укладку цепи. Таким образом, увеличение  $\Delta \epsilon$  *экспоненциально* уменьшает число аминокислотных последовательностей, способных придать стабильность нативной структуре, содержащей элемент с энергией  $\Delta \epsilon$ .

■ **Добавление.** Выше мы считали, что «годными» для белка являются все последовательности, обеспечивающие  $\Delta F + \Delta \epsilon < 0$ , т. е. хотя бы минимальную стабильность его структуры. Через несколько лекций мы убедимся, что белковая структура должна обладать неким минимальным «запасом стабильности», иначе она будет плавиться у нас в руках и не будет способна к быстрой и однозначной самоорганизации. Поэтому более правильно считать, что «годные» для белка последовательности обеспечивают  $\Delta F + \Delta \epsilon < -F_{\min} < 0$ , т. е.  $\Delta F < -\Delta \epsilon - F_{\min}$ . Для обеспечения

стабильности нативной глобулы достаточно, чтобы величина  $F_{\min}$  составляла несколько килокалорий на моль, т. е. можно полагать, что  $F_{\min} \ll \langle \Delta F \rangle$ . При этом условие (16.8) приобретает вид

$$P^* (\Delta F < -\Delta\varepsilon - F_{\min}) \approx \text{const} \times \exp [-(\Delta\varepsilon + F_{\min})/(\sigma^2/\langle \Delta F \rangle)] = \text{const}^* \times \exp [-\Delta\varepsilon/(\sigma^2/\langle \Delta F \rangle)]. \quad (16.8a)$$

То есть (ведь изменение предэкспоненциальной константы нас не волнует) мы приходим все к той же экспоненциальной форме зависимости вероятности того, что данный элемент закрепится «отбором на стабильности» в структуре белка, от энергии этого элемента. Заметьте, что рост стабильности белка, т. е. величины  $F_{\min}$ , экспоненциально уменьшает число «годных» для него последовательностей.

- Полученная зависимость имеет экспоненциальную форму (и тем она похожа на формулу Больцмана), но, *в отличие* от формулы Больцмана,  $\Delta\varepsilon$  здесь делится не на температуру среды (точнее, не на  $kT$ ), а на непонятную пока величину  $\sigma^2/\langle \Delta F \rangle$ .

Что же это за величина? Прежде всего, отметим, что  $\sigma^2/\langle \Delta F \rangle$  не зависит от размера белка. В самом деле, по законам математической статистики, средняя величина  $\langle \Delta F \rangle$  пропорциональна числу суммируемых в  $\Delta F$  членов (которое, в нашем случае, примерно пропорционально размеру белка), а среднеквадратичное отклонение от среднего  $\sigma$  пропорционально квадратному корню из числа этих членов, т. е.  $\sigma^2$  тоже примерно пропорционально размеру белка.

То, что  $\sigma^2/\langle \Delta F \rangle$  *не растет с размером белка*, очень важно: это значит, что величину «дефекта»  $\Delta\varepsilon$  надо сравнивать — в формуле (16.8) — не с полной энергией белка (которая пропорциональна его размеру), а с какой-то характерной энергией  $\sigma^2/\langle \Delta F \rangle$ , т. е. с чем-то вроде средней энергии нековалентных взаимодействий в цепи в расчете на один ее аминокислотный остаток (которая тоже не зависит от размера белка!).

Это, вместе с экспоненциальным видом формулы (16.8), сразу отвечает на вопрос о том, почему  $\Delta\varepsilon$  ценой всего в несколько килокалорий на моль может существенно влиять на встречаемость элементов белковых структур (например, почему лейцина внутри белка на порядок больше, чем серина). Это происходит потому, что число последовательностей, стабилизирующих нативную структуру белка, падает в «е» раз (примерно втрое), когда  $\Delta\varepsilon$  растет на  $\sigma^2/\langle \Delta F \rangle$ .

Итак, величина  $\sigma^2/\langle \Delta F \rangle$  — это что-то вроде энергии цепи в расчете на одно ее звено (или на одну ее степень свободы). Или (напомним, что  $kT$  — это характерная тепловая энергия в расчете на степень свободы) — можно полагать, что  $\sigma^2/\langle \Delta F \rangle = kT_c$ , где  $T_c$  — какая-то температура, характерная для белковой цепи. Какая? В «случайной» глобуле есть только одна характерная температура — температура «вымораживания» самой стабильной укладки такой глобулы. И в белке есть одна характерная тем-

пература — температура его плавления, денатурации,  $\approx 350$  К (которая близка к «температуре жизни» белка,  $\approx 300$  К). Причем можно показать, что, пока мы требуем от белка только минимальной стабильности его структуры (т. е. если отбор его «квазислучайной» первичной структуры определяется только этим, не слишком жестким требованием), то он плавится примерно при той же температуре, при которой замерзает «случайная» глобула (на самом деле — при чуть более высокой). Поэтому можно полагать, что величина «конформационной температуры»  $T_c$  определяется температурой денатурации белка. Иными словами, можно предполагать, что  $kT_c = \sigma^2/\langle \Delta F \rangle$  составляет примерно 0,5–1 ккал/моль.

Это можно не только предполагать, но и показать. За словами «можно показать» стоит довольно сложное теорфизическое доказательство — теорема Шахновича-Гутина — от которого я вас избавлю.

- Итак, мы разобрались с вопросом о том, почему дефект ценой всего в несколько килокалорий на моль — на фоне гораздо большей полной энергии белка — может практически запрещать многие мотивы белковых архитектур. Потому, что «дефект», любой дефект, ценой в 1 ккал/моль снижает число «годных» для белка последовательностей раз в пять, «дефект» ценой в 2 ккал/моль — раз в двадцать, и т. д.

Поэтому, в частности, перекрывание нерегулярных перемычек в белках (а цена этому дефекту — килокалории две-три, максимум пять) наблюдается очень редко (рис. 16-5).

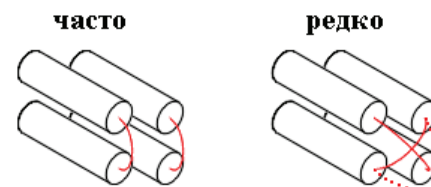


Рис. 16-5. Перекрывание петель редко наблюдается в белках

**Внутренний голос:** Не могу не вмешаться. Вся логика Вашего рассказа основана на том, что последовательности, не способные сложиться в стабильную структуру, отбраковываются. Вполне правдоподобно. Но не отбраковываются ли и те последовательности, что дают *слишком стабильные* структуры? Ведь наблюдаемая стабильность нативных белков *всегда* не слишком высока! К тому же, многократно отмечалась корреляция между температурой денатурации белка и температурой существования этого белка в организме. Но тогда получается, что *слишком* стабильные структуры белков (а, значит, — по Вашей же логике! — и *слишком* стабильные

структурные элементы) должны выбраковываться. А ведь Вы этого совсем не учитываете!

**Лектор:** Сделать стабильный белок трудно. Об этом свидетельствует, в частности, весь накопленный опыт по дизайну искусственных белков. То есть последовательности, способные сложиться в хоть сколько-нибудь стабильную структуру, встречаются редко. Но «суперстабильный» белок сделать еще труднее, уж очень редки пригодные для этого аминокислотные последовательности! О малости доли последовательностей, пригодных для создания «суперстабильных» белковых структур (с очень низкой свободной энергией  $\Delta F$ ), свидетельствует тот же рис. 16-4. Поэтому среди цепей, обеспечивающих хотя бы минимально приемлемую стабильность белков, лишь ничтожная доля будет создавать «суперстабильные» белки... Доля таких белков (если отбор не настаивает на «суперстабильности», но и не препятствует ей) *автоматически* будет ничтожна. Этого вполне достаточно для объяснения всех отмеченных Вами экспериментальных фактов и корреляций...

- Теперь разберемся со вторым вопросом: почему энтропийные эффекты имеют отношение к стабильности нативной структуры белка, где цепь все равно фиксирована.

Здесь все тоже довольно просто.

Сравним два мотива укладки цепи, например, правовинтовой ход перемычки между параллельными  $\beta$ -участками с левовинтовым (рис. 16-6).



**Рис. 16-6.** Левовинтовой (требующий большего изгиба) ход перемычек между параллельными  $\beta$ -участками очень редко наблюдается в белках, правовинтовой — часто

Так как  $\beta$ -лист из  $L$  аминокислот имеет тенденцию к скручиванию с углом около  $-30^\circ$  между соседними  $\beta$ -тяжами, то *правовинтовая* перемычка требует поворота цепи на  $360^\circ - 30^\circ = 330^\circ$ . Другая (*левовинтовая*) перемычка делает больший поворот, на  $360^\circ + 30^\circ = 390^\circ$ . Полимерная физика говорит, что чем сильнее согнута цепь, тем меньше у нее может быть конформаций. Значит, правовинтовой (менее закрученный) ход цепи совместим с большим числом конформаций цепи, а левовинтовой (более закрученный, более напряженный) — с небольшим. Иначе говоря, «левый»

ход перемычки создает «энтропийный дефект» у «левой» перемычки. Это понятно. Однако, как этот «дефект» сказывается на числе последовательностей, способных сделать стабильным тот или иной ход перемычки?

Наугад выбранная аминокислотная последовательность может сделать самой стабильной структурой цепи либо какую-то одну из многих возможных конформаций, соответствующих «правой» перемычке, либо какую-то из немногих, отвечающих «левой» перемычке, либо, наконец, она не способна сделать самой стабильной ни одну из этих конформаций. Последнее, конечно, наиболее вероятно, так как на опыте случайно сваренные сополимеры аминокислот не имеют стабильной трехмерной структуры. Однако, если выбирать из двух: какая перемычка все же имеет больший шанс стать самой стабильной — правая или левая?

Каждая отдельная конформация с «правой» перемычкой ничуть не лучше и не хуже, чем каждая отдельная конформация с «левой» перемычкой, если они имеют равную компактность, одинаковое содержание вторичной структуры и т. д. Однако «правых» конформаций больше... Итак, мы видим, что перед нами — нечто вроде лотереи с малым шансом на выигрыш главного приза («приз» — возможность создать стабильную трехмерную структуру) — лотереи, где «правая» перемычка имеет много «билетов» (возможных конформаций), а левая — мало. Кому же, скорее, достанется приз, если он вообще кому-то достанется? Конечно, «правой» перемычке, той, у которой «билетов» больше: вероятность победы той или иной перемычки прямо пропорциональна числу имеющихся на руках «билетов» — конформаций...

Мы видим, что «энтропийный дефект» конвертируется в энергетический: имея мало конформаций на выбор, трудно рассчитывать на хорошую энергию.

Иными словами, шанс на то, что случайная аминокислотная последовательность сделает стабильной какую-то из *многих* конформаций «правой» перемычки, относится к шансу на то, что она сделает стабильной какую-то из *немногих* «левых» конформаций, как число «правых» конформаций относится к числу «левых». Или: чем шире набор *возможных* конформаций (а их больше у «правой» перемычки), тем больше последовательностей найдет в этом наборе самую стабильную структуру. При этом каждая из таких последовательностей выберет из этого набора одну, наиболее подходящую именно для нее конформацию.

Именно это и наблюдается в глобулярных белках: здесь «правый» ход перемычки — правило, а «левый» — исключение.

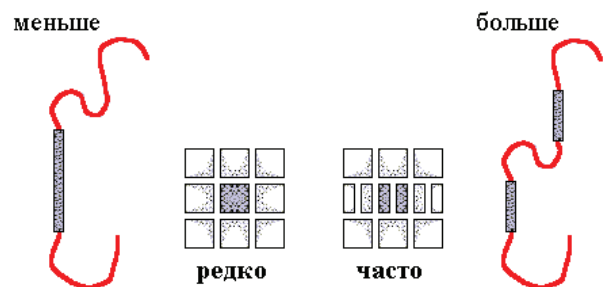
Напомню самое яркое из таких редчайших исключений. Это — левая  $\beta$ -призма; мы о ней говорили пару лекций назад. Ее пространственная структура, действительно, уникальна: сплошь левые перемычки между  $\beta$ -участками. А как насчет первичной структуры? Оказывается, она тоже



уникальна для глобулярных белков: она выглядит вовсе не как «случайная» (что столь типично для глобулярных белков): она очень периодична, она состоит из десятка повторов 18-членного пептида, образующего один виток левой суперспирали (с тремя  $\beta$ -тяжами в каждом витке; а в этих тяжах аномально много глицинов, *не* скручивающих  $\beta$ -лист). Так что, нетипичная пространственная структура сочетается с нетипичной первичной...

- Рассмотрим еще одну задачу, связанную с энтропийными дефектами. Она уже относится не к мотиву укладки цепи, а к слоевой структуре глобулы.

Что должно чаще наблюдаться: белки, через центр которых проходит слой  $\alpha$ -спиралей, или белки, через центр которых проходит  $\beta$ -лист?



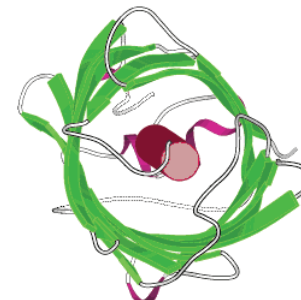
**Рис. 16-7.** Многослойная упаковка, в центре которой лежит  $\alpha$ -спираль, должна быть менее вероятна (и на самом деле встречается гораздо реже), чем многослойная упаковка, в центре которой лежат два  $\beta$ -участка. Дело в том, что лежащий в центре глобулы участок должен состоять из одних гидрофобных остатков, причем большая длина такого структурного участка диктуется диаметром глобулы, а, при равной длине,  $\alpha$ -спираль состоит из вдвое большего числа остатков, чем вытянутый  $\beta$ -тяж; поэтому для создания внутренней  $\alpha$ -спирали (одного большого блока из гидрофобных групп) нужна более «редкая» последовательность, чем для создания двух внутренних  $\beta$ -участков (двух вдвое меньших гидрофобных блоков, *как-то* расположенных в цепи). Вероятность существования одного большого (из  $t$  остатков) гидрофобного блока в данном месте случайной цепи — порядка  $p^t$  (где  $p$  — доля гидрофобных остатков в цепи); и такова же вероятность существования двух вдвое меньших блоков в двух данных местах случайной цепи —  $(p^{m/2})^2 = p^m$ . Однако один блок может размещаться в цепи из  $N$  остатков примерно  $N$  способами, а способов разместить два блока много больше, примерно  $N \times N/2$ , т. е. цепей с двумя короткими гидрофобными блоками много больше, чем цепей с одним длинным блоком

Имея одну и ту же скрытую от воды поверхность, обе эти структуры должны иметь примерно одинаковую гидрофобную свободную энергию, *если* цепи удастся вписаться в такую структуру, не допустив полярных боковых групп внутрь белка. Рассмотрим только «самые трудные» для случайной последовательности структурные сегменты, лежащие в самом

центре глобулы. В первом случае (рис. 16-7) цепи предстоит сотворить внутреннюю спираль; длина этой спирали определяется диаметром глобулы. Для создания спирали, стабильной в центре глобулы, цепь должна обладать одним длинным гидрофобным участком. Во втором случае она должна обладать двумя вдвое более короткими (так как на один остаток цепи приходится около 3 Å длины  $\beta$ -тяжа и 1,5 Å —  $\alpha$ -спирали) гидрофобными участками (они, практически вне зависимости от расстояния между ними по цепи, годятся на роль двух внутренних  $\beta$ -участков). Рисунок 16-7 объясняет, почему для создания внутренней  $\alpha$ -спирали (единого большого блока из гидрофобных групп) нужна более «редкая» последовательность, чем для создания двух внутренних  $\beta$ -участков (двух вдвое более коротких гидрофобных блоков, *как-то*, т. е. любым из большого числа способов, расположенных в цепи).

Итак, ожидаемый ответ (см. рис. 16-7) гласит, что глобулы, через центр которых проходит  $\beta$ -лист, стабилизируются гораздо большим числом первичных структур, чем глобулы с  $\alpha$ -спиралью по центру.

И, действительно, внутренняя  $\alpha$ -спираль в белках очень редка (в качестве известного носителя такого редчайшего признака назову «зеленый флуоресцентный белок», рис. 16-8), а внутренние  $\beta$ -тяжи — типичны (например, в «укладках Россмана»).



**Рис. 16-8.** Необычная архитектура зеленого флуоресцентного белка:  $\alpha$ -спираль окружена  $\beta$ -листом. Центральная спираль на самом деле не сплошная: она изломана хромофором примерно посередине (этот излом не показан на рисунке)

- С той же точки зрения — «насколько часто данная структурная деталь стабилизируется случайными аминокислотными последовательностями?» — можно объяснить и многие другие закономерности, наблюдаемые в глобулярных белках. Например, средний размер домена для цепи с заданным соотношением гидрофобных и гидрофильных групп (этот вопрос исследовался еще ленинградскими учеными Бреслером и Талмудом в 1944 г.,

а затем американским ученым H/Fisher'ом), а также — мы уже об этом говорили — средние длины  $\alpha$ - и  $\beta$ -участков и нерегулярных петель.

- Итак, наш анализ показывает, что вероятность обнаружения заданного структурного элемента в стабильных (т. е. наблюдаемых) глобулах, образованных теми из случайных последовательностей, что вообще способны сделать стабильной хоть какую-то глобулярную структуру, — эта вероятность должна быть тем больше, чем ниже его, этого элемента, энергия и чем большим числом конформаций его можно создать.

А так как энергия и число конформаций объединяются в единое целое в свободной энергии, то статистика наблюдения самых разных структурных элементов в «случайно сделанных» стабильных глобулах должна иметь вид

$$\text{ВСТРЕЧАЕМОСТЬ} \sim \exp(-\text{СВОБОДНАЯ\_ЭНЕРГИЯ}/kT_c), \quad (16.9)$$

где  $T_c$  близка (не в точности равна, как показывает более тщательный анализ, но близка) к температуре затвердевания глобулы, образованной «случайной» аминокислотной последовательностью, которая, в свою очередь, близка к температуре денатурации белка.

И действительно, белковая статистика, в общем, имеет задаваемый формулой (16.9) вид — вид, присущий статистике структур, сложенных из случайных последовательностей, отобранных только на предмет создания стабильной глобулярной структуры.

- Еще раз напомним физическую причину соотношения (16.9). Она заключается в том, что свободная энергия рассматриваемого элемента белковой структуры экспоненциально изменяет число аминокислотных последовательностей, способных придать стабильность белку, нативная структура которого содержит этот элемент. Если он сам по себе стабилен — белок с ним «терпит» много даже неблагоприятных мутаций, т. е. довольно много аминокислотных цепей стабилизируют белок с этим элементом. Если же рассматриваемый элемент нестабилен, белок с ним требует очень тщательного подбора своей первичной структуры (стабильность его пространственной структуры легко разрушается даже немногими мутациями), а таких «тщательно отобранных» последовательностей мало.

Здесь нужно подчеркнуть, что так называемые «запрещенные», т. е. не наблюдаемые (или редко наблюдаемые) в белках структуры не невозможны в принципе — они просто маловероятны, т. е. создаются малым числом аминокислотных последовательностей.

- Похоже, что глобулярные белки могли относительно легко возникнуть из случайных гетерополимеров аминокислот (или, точнее: из кусков

ДНК, кодирующих случайные последовательности аминокислот), нужно было только немного, при помощи нескольких мутаций, стабилизировать самую стабильную пространственную структуру исходного случайного полипептида (и тем самым сделать ее единственной наблюдаемой структурой цепи), да еще «навесить» на ее поверхность активный центр (для «биологически-нужного» взаимодействия с окружающими молекулами). Да еще — убрать с поверхности белка те остатки, что могут вовлечь его в «биологически-вредные» ассоциаты (типа тех, что, склеивая гемоглобины, вызывают серповидно-клеточную анемию).

**Внутренний голос:** Следует ли Вас понимать так, что Вы предполагаете, что белки с «новыми» архитектурами возникали из случайных последовательностей, а не путем сильной мутации белков с какой-то «старой» архитектурой? И что те мотивы укладки цепи, стабильность которых совместима со многими случайными последовательностями, много раз (столько, сколько есть в них гомологических семейств) возникали из случайных последовательностей — в то время как те мотивы, которые охватывают только одно семейство гомологов, смогли возникнуть только единожды?

**Лектор:** Достаточно уверенно можно говорить лишь о том, что не случайно малая доля случайных последовательностей способна породить белки; и о том, что стабильность некоторых мотивов (причем тех, что часто встречаются в белках) совместима с большим числом случайных последовательностей, чем стабильность других, «редких» мотивов. Что касается Ваших вопросов об имевших место быть исторических событиях, я думаю, ответить на них сейчас невозможно. В частности, невозможно сказать, много ли раз возникали представители «популярных» мотивов укладки цепи или нет. Возможно, они возникали многократно, из разных случайных последовательностей. Возможно — однократно (и не из случайной последовательности, а из кусков каких-то других белков), — а потом, в ходе эволюции, имеющие общий корень последовательности так широко разошлись в рамках сохраненного пространственного мотива (благо «популярные» мотивы совместимы со столь разными последовательностями), что все следы гомологии, генетического родства этих белков изгладились, и мы их не можем опознать. Хочу только подчеркнуть, что «популярные», совместимые с большим числом последовательностей структурные мотивы предоставляют куда большее пространство для любого сценария происхождения и дальнейшей эволюции белков, чем «редкие».

- Не могу не отметить в этой связи, что современные работы по дизайну белков *de novo* широко используют как множественные случайные мутации аминокислотных последовательностей, непременно сопровождаемые

отбором «подходящих» (скажем, с чем-то связывающихся) вариантов, так и случайное перемешивание олигопептидов (сопровождающееся таким же отбором). Современные методы отбора «подходящих» случайных последовательностей (упомяну только название самого, пожалуй, мощного из них: фаговый дисплей; читайте о нем в учебниках по молекулярной биологии!) позволяют проанализировать порядка  $10^{10}$  и более случайных последовательностей. То, что среди них порой находят «белковоподобные» продукты, показывает, что такие «белковоподобные» цепи составляют долю порядка  $10^{-10}$  от всех случайных полипептидов, состоящих из нескольких десятков аминокислотных остатков.

Я специально подчеркиваю: возникновение белков из случайных последовательностей можно предположить именно для глобулярных белков, так как именно их последовательности внешне напоминают «случайные» (т. е. самые массовые) сополимеры (см. рис. 16-2). В то же время для первичных (а равно и пространственных) структур фибриллярных и мембранных белков ясно просматривается их «неслучайное» строение (но кстати: строение, также простое с точки зрения его создания, — повтор коротких блоков для фибриллярных, и чередование гидрофобных и гидрофильных блоков — для мембранных белков).

- Проведенный анализ подчеркивает тот факт, что получающиеся в результате эволюции структуры белков выглядят очень «разумно» с физической точки зрения — так же, как и двойная спираль ДНК, как и мембранный бислой. Видимо, и на уровне архитектур белковых доменов эволюция не «изобретает» нечто физически маловероятное, а «выбирает» среди физически разумных (т. е. стабильных и потому способных к быстрой самоорганизации, мы это скоро увидим) структур. В этом и заключается смысл «физического отбора» белковых структур.

## КООПЕРАТИВНЫЕ ПЕРЕХОДЫ В БЕЛКОВЫХ МОЛЕКУЛАХ

### *Лекция 17*

Денатурация белка. Нативно-развернутые белки. Кооперативные переходы. Обратимость денатурации белков. Денатурация глобулярного белка — переход типа «все-или-ничего». Критерий Вант-Гоффа для перехода «все-или-ничего». Тепловая и холодная денатурация. Диаграмма фазовых состояний белковой молекулы. Как выглядит денатурированный белок? Клубок и расплавленная глобула. Неоднородность расплавленной глобулы. Отсутствие фазового перехода типа «все-или-ничего» при набухании «обычных» полимерных глобул.

Ранее мы рассматривали стабильность фиксированных, «твердых» белковых структур. Но, при определенных внешних условиях, самой стабильной может оказаться не твердая, а расплавленная или даже развернутая форма белковой молекулы. Тогда белок «денатурирует», теряет свою нативную, «рабочую» структуру.

- Обычно денатурация белка наблюдается *in vitro*, при воздействии на него аномальной температуры или денатуранта [мочевины,  $H^+$  или  $OH^-$  ионов (т. е. аномального pH) и т. д.]. Однако распад «твердой» структуры белка и затем ее повторная самоорганизация происходит и в живой клетке, что играет важную роль, например, в процессе транспорта белков через мембраны.

Более того. Вообще не все белки «сами по себе» имеют определенную пространственную структуру при физиологических условиях. Не имеющие получили название «нативно-развернутых». Иногда «нативно-развернут» весь белок, иногда развернуты какие-то участки его цепи. Некоторые из нативно-развернутых белков (или нативно-развернутых участков цепи) имеют странный, упрощенный аминокислотный состав и, возможно, никогда не имеют определенной пространственной структуры. Пока непонятно, зачем они нужны клетке. Возможно — как источник каких-то олигопепти-

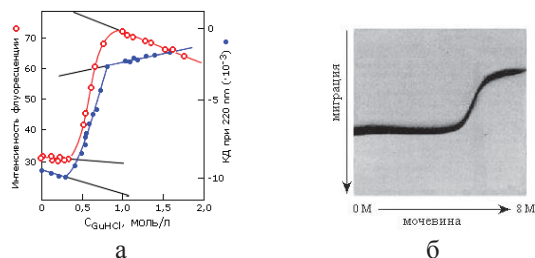
дов. Возможно — «низачем», как «эгоистическая ДНК». Однако многие «нативно-развернутые» белки приобретают свою уникальную пространственную структуру, связав лиганд или сев на другой белок, ДНК или РНК. [Предполагается, что у связывания белка с нестабильной структурой есть то преимущество, что при большой поверхности (а значит, и выборочности) связывания оно оказывается не слишком прочным.] Таким образом, изучение «денатурированного» (неупорядоченного) состояния белковых цепей имеет прямое отношение и к состоянию некоторых белков в клетке.

Денатурация и ренатурация глобулярных белков *in vitro* — объект интенсивных исследований, интерес к которым поддерживается их связью с проблемой самоорганизации белка, т. е. с вопросом о том, как белковая цепь находит свою уникальную структуру среди гигантского числа возможных альтернатив.

В этой лекции я практически не буду касаться кинетических аспектов денатурации и самоорганизации белков и сосредоточусь на термодинамических и структурных аспектах этих явлений. Больше всего изучены водорастворимые глобулярные белки, и именно о них я буду рассказывать.

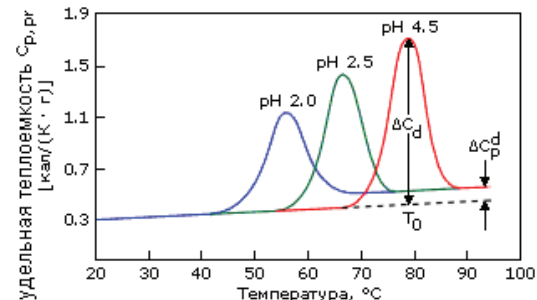
#### ■ Что показывает эксперимент?

Твердо установлено, что денатурация малых белков является кооперативным переходом с одновременным и резким, «S-образным» изменением многих (хотя порой и не всех) характеристик молекулы (рис. 17-1). S-образность экспериментальных кривых показывает, что соответствующие характеристики молекулы меняются от тех, что характерны для нативного белка, до тех, что характерны для белка денатурированного; а узость этих S-образных кривых свидетельствует о кооперативности перехода, т. е. о том, что он охватывает сразу много аминокислотных остатков.



**Рис. 17-1.** Денатурация белка сопровождается резким, «S-образным» изменением многих характеристик молекулы. В данном случае денатурация белка вызывается повышением концентрации денатуранта, которая измеряется в числе молей молекул на литр раствора (обозначение: моль/л, или М). (а) Синхронное изменение КД (длина волны — 220 нм) и флуоресценции при денатурации фосфоглицераткиназы в растворе гуанидингидрохлорида. Картинка, с небольшими изменениями, взята из H. Nojima et al., *J. Mol. Biol.* (1977) **116**:429–442. (б) Электрофорез цитохрома *c* при разных концентрациях мочевины: замедление миграции вызвано разбуханием молекулы при денатурации. Картинка, с разрешения, взята из T. E. Creighton, *J. Mol. Biol.* (1979) **129**:235–264

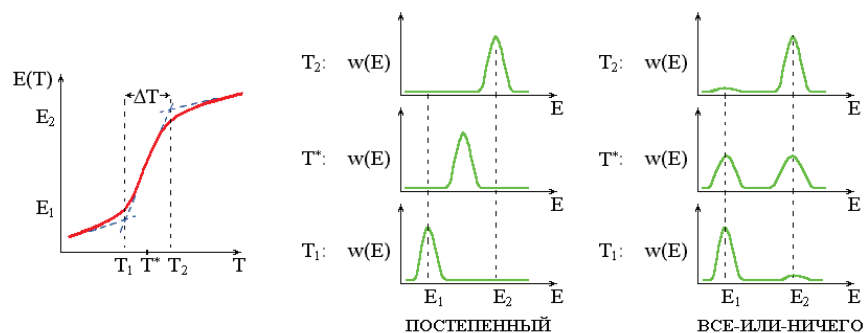
Более того: денатурация белка происходит как переход типа «все-или-ничего» (рис. 17-2).



**Рис. 17-2.** Калориметрическое исследование тепловой денатурации лизоцима при разных pH. Положение пика удельной теплоемкости  $C_p$  определяет температуру перехода  $T_0$ , его ширина — ширину перехода  $\Delta T$ , площадь под пиком — поглощенное при плавлении тепло  $\Delta H$  в расчете на грамм белка. То, что они ( $\Delta T$ ,  $\Delta H$ ,  $T_0$ ) удовлетворяют условиям (17.4), (17.5), свидетельствует, что денатурация происходит как переход типа «все-или-ничего». Увеличенная теплоемкость денатурированного белка — следствие увеличения поверхности контакта его гидрофобных боковых групп с водой при (частичном или полном) разворачивании белка. Картинка, с небольшими изменениями, взята из P. L. Privalov & N. N. Khechinashvili, *J. Mol. Biol.* (1974) **86**:665–684

Последнее (вы должны это помнить) означает, что при таком переходе только начальное (нативное) и конечное (денатурированное) состояния наблюдаются в заметных количествах (рис. 17-3), а «полуденатурированных» молекул практически нет. [Хотя, конечно, они тоже должны быть, пусть в мизерном количестве, ведь не переходит же одно состояние в другое при помощи «нуль-транспортировки»; но наличие «полусвернутых» структур отражается только на кинетике перехода, разговор о которой еще впереди.] Иначе говоря, переход «все-или-ничего» является микроскопическим аналогом фазового перехода первого рода в макроскопических системах (например, плавления кристалла). Однако, в отличие от истинного фазового перехода, S-образность перехода «все-или-ничего» имеет не нулевую, а конечную ширину, так как этот переход охватывает не макроскопическую, а микроскопическую, очень небольшую систему. Каким образом было экспериментально показано то, что денатурация небольших белков подчиняется принципу «все-или-ничего», об этом чуть позже, а пока уместно уточнить, что денатурация типа «все-или-ничего» относится к небольшим глобулярным белкам и отдельным доменам крупных белков, а денатурация больших глобулярных белков, как показывают многочисленные опыты, складывается из плавления входящих в них доменов.





**Рис. 17-3.** При одном и том же виде зависимости энергии  $E$  (или другого наблюдаемого параметра) от температуры, кооперативный (« $S$ -образный») переход может быть и переходом типа «все-или-ничего» (пример: денатурация белка), и постепенным, «безродным» (пример: переход спираль-клубок в полипептидах). Различия проявляются не в виде кривой  $E(T)$ , а в виде функции распределения  $w(E)$  молекул по энергии (или по другому наблюдаемому параметру). Штриховые линии на левом рисунке поясняют графическое определение  $\Delta T$ , температурной ширины перехода

Денатурировать белок можно не только теплом (см. рис. 17-2), но порой и холодом (о чем речь пойдет несколько позже). Кроме того, белки денатурируют под действием слишком высоких или слишком низких pH (вызывающих исчезновение зарядов одного знака и, тем самым, — расталкивание оставшихся зарядов другого знака). Денатурируют белки (см. рис. 17-1) и при добавлении в раствор денатурантов типа мочевины ( $\text{NH}_2 - \text{CO} - \text{NH}_2$ ) или гуанидингидрохлорида ( $[(\text{NH}_2)_3\text{C}]^+\text{Cl}^-$ ) — эти вещества оттягивают к себе водородные связи O-атомов воды, нарушают баланс водородных связей в ней, и тем самым заставляют ее активнее рвать такие связи в белке. Денатурируют белки и при добавлении в раствор молекул додецилсульфата натрия (чья гидрофобные хвосты проникают в ядро белка и ломают его), различных спиртов, солей и т. д., а также — сверхвысоких давлений.

Интересно, что белки существ, обитающих в экстремальных (для нас!) условиях типа горяче-кисло-соленых источников, являются «экстремофилами»: они существуют в нативной форме именно при таких экстремальных условиях, а при «физиологических» (для большинства организмов) условиях денатурируют. Следует отметить, что в архитектурах белков-экстремофилов нет никаких ярких особенностей, отличающих их от «обычных» («мезофильных») белков, а адаптация к экстремальным условиям обеспечивается «неяркой» подгонкой первичных структур, направленной на усиление тех взаимодействий, что могут поддержать стабильность пространственной структуры при данных внешних условиях. Впрочем, рассказ об этом увел бы нас далеко в сторону...

- Может ли денатурированный белок ренатурировать и обрести свою нативную структуру? То есть, обратима ли его денатурация?

Да (это известно со времен опытов Анфинсена в 1960-х гг.; кстати, за них он получил Нобелевскую премию), если белок не слишком велик и не подвергался сильной химической модификации после сворачивания *in vivo*: тогда «мягко» (без разрыва цепи) разрушенная температурой, денатурантом и т. д. архитектура белка спонтанно восстанавливается при «нормализации» среды. Правда, эффективная ренатурация *in vitro* нуждается в тщательном подборе экспериментальных условий, иначе ей может воспрепятствовать выпадение из раствора и/или агрегация (межмолекулярная, а для больших белков, видимо, и внутримолекулярная агрегация отдаленных кусков цепи: известно, что способность к ренатурации и ее «выход» обычно падают, а экспериментальные сложности — растут с увеличением размера белка).

Обратимость денатурации белка чрезвычайно важна: она показывает, что вся нужная для создания уникальной нативной структуры белка информация находится в его аминокислотной последовательности, а сама нативная структура белка (строго говоря: не модифицированного небольшого белка) равновесна и стабильна; и эта обратимость позволяет применять термодинамику для описания и изучения такого перехода.

- Тот фундаментальный факт, что денатурация белка происходит как переход типа «все-или-ничего», был показан П. Л. Приваловым, работавшим в нашем Институте белка. Он изучал тепловую денатурацию белка, которая обычно сопровождается большим тепловым эффектом,  $\sim 1$  ккал/моль аминокислотных остатков.

- Как доказать, что плавление белковой молекулы — переход типа «все-или-ничего»? Для этого недостаточно показать существование пика теплоемкости (или, что есть то же самое, что энергия белка  $S$ -образно зависит от температуры). Все это свидетельствует только о резкости перехода, о том, что он охватывает много аминокислотных остатков одновременно, но ничего не говорит о том, плавится ли белок целиком или по частям. Чтобы доказать, что плавление белка — переход типа «все-или-ничего», нужно сравнить (1) «эффективную теплоту» перехода, вычисляемую из его ширины, с (2) «калориметрической теплотой» этого перехода, т. е. с количеством тепла, поглощаемым *одной молекулой* белка в процессе плавления. Совпадение этих двух независимых величин показывает, что молекула плавится как единое целое.

Это и есть «критерий Вант-Гоффа» для наличия перехода типа «все-или-ничего». Ввиду его важности рассмотрим этот критерий внимательно.

«Эффективная теплота» перехода, следующая из его ширины, есть количество тепла, поглощенного одной независимой «единицей плавления»

ния». Если эффективная теплота перехода меньше калориметрической — «единица плавления» меньше, чем сама молекула, т. е. молекула плавится по частям. Если эффективная теплота перехода больше калориметрической, то «единица плавления» больше молекулы, т. е. плавится не одна молекула белка, а какой-то их агрегат.

Как связана эффективная теплота перехода с его шириной? Рассмотрим «единицу плавления», которая может находиться в двух состояниях: «твердом», с энергией  $E$  и энтропией  $S$ , и «расплавленном», с энергией  $E'$  и энтропией  $S'$ . Пусть, для простоты,  $E$ ,  $E'$ ,  $S$  и  $S'$  не зависят от температуры  $T$  (рассмотрение более общего случая я предоставляю читателю...).

Так как состояний — всего два, то из формулы Больцмана легко видеть, что вероятность пребывания «единицы плавления» в расплавленном состоянии есть

$$P_{\text{распл}} = \exp[-(E' - TS')/kT] / \{\exp[-(E - TS)/kT] + \exp[-(E' - TS')/kT]\} = 1 / \{\exp[(\Delta E - T\Delta S)/kT] + 1\}, \quad (17.1)$$

где  $\Delta E = E' - E$ ,  $\Delta S = S' - S$ . Вероятность пребывания «единицы плавления» в твердом состоянии есть (для перехода «все-или-ничего»)  $P_{\text{ТВ}} = 1 - P_{\text{распл}}$ . Производная  $dP_{\text{распл}}/dT$  показывает, как быстро меняется  $P_{\text{распл}}$  с температурой. Простое вычисление (с использованием уже известной нам формулы  $d(F/T)/dT = d[(E - TS)/T]/dT = -E/T^2$ ) показывает, что

$$dP_{\text{распл}}/dT = P_{\text{распл}} (1 - P_{\text{распл}}) (\Delta E/kT^2). \quad (17.2)$$

Пожалуй, надо расписать здесь все преобразования подробнее, чтобы не принимать на веру, а понимать. Обозначим  $(\Delta E - T\Delta S)/kT$  через  $X$ . Тогда  $dX/dT = -\Delta E/kT^2$ ,  $P_{\text{распл}} = 1/(e^X + 1)$ ,  $P_{\text{ТВ}} = 1 - P_{\text{распл}} = e^X/(e^X + 1)$ , и

$$\begin{aligned} dP_{\text{распл}}/dT &= d[1/(e^X + 1)]/dT = -[1/(e^X + 1)^2] [de^X/dT] = \\ &= -[1/(e^X + 1)^2] e^X [dX/dT] = -[1/(e^X + 1)] [e^X/(e^X + 1)] [dX/dT] = \\ &= P_{\text{распл}} (1 - P_{\text{распл}}) (-dX/dT) = P_{\text{распл}} (1 - P_{\text{распл}}) (\Delta E/kT^2). \end{aligned}$$

Точка середины перехода  $T_0 = \Delta E/\Delta S$ . В этой точке  $P_{\text{распл}} = P_{\text{ТВ}} = 1/2$ , и здесь величина  $P_{\text{распл}} (1 - P_{\text{распл}})$  проходит через максимум, равный  $1/4$ . Точка самого быстрого изменения  $P_{\text{распл}}$ , т. е. максимум производной  $dP_{\text{распл}}/dT$  тоже падает практически на температуру  $T_0$ , если  $\Delta E/kT \gg 1$ , т. е. если переход происходит в узкой области. Здесь наклон кривой  $P_{\text{распл}}(T)$  максимален и равен

$$(dP_{\text{распл}}/dT)|_{T=T_0} = (1/4) \Delta E/kT_0^2. \quad (17.3)$$

Графическое определение ширины перехода показано на рис. 17-3, слева. Оно делается путем линейной экстраполяции максимального наклона

кривой  $P_{\text{распл}}(T)$  до пересечения с базовыми прямыми, соответствующими нативному ( $P_{\text{распл}} = 0$ ) и денатурированному ( $P_{\text{распл}} = 1$ ) состояниям. В зоне перехода экстраполированная (штриховой линией) величина  $P_{\text{распл}}|_{\text{экстрап.}}$  меняется от 0 до 1 (то есть  $\Delta P_{\text{распл}}|_{\text{экстрап.}} = 1$ ), а температура меняется на  $\Delta T$ . Значит

$$(dP_{\text{распл}}/dT)|_{T=T_0} = \Delta P_{\text{распл}}|_{\text{экстрап.}}/\Delta T = 1/\Delta T,$$

— т. е.  $\Delta T$  определяется (см. рис. 17-3) зоной самого крутого подъема величины  $P_{\text{распл}}$ , или, в общем виде, шириной зоны резкого изменения любого экспериментального параметра, определяющегося вероятностью расплавленного состояния (например, шириной зоны резкого изменения спиральности белковой цепи).

В окончательной форме связь эффективной теплоты плавления «плавящейся единицы»  $\Delta E$  с шириной  $\Delta T$  и температурой  $T_0$  перехода имеет вид  $1/\Delta T = (1/4) \Delta E/kT_0^2$ , или

$$\Delta E = 4kT_0^2/\Delta T. \quad (17.4)$$

Рассчитанная таким образом — только из наблюдаемой формы перехода — теплота  $\Delta E$  денатурации «единицы плавления» сравнивается с калориметрической теплотой плавления целого белка (рассчитываемой как  $\Delta H/N$ , где  $\Delta H$  — количество тепла, поглощенного всеми имеющимися в калориметре  $N$  молекулами белка, причем  $N = m/M$ , где  $m$  — масса имеющегося в пробирке белка, а  $M$  — его молекулярный вес). Если

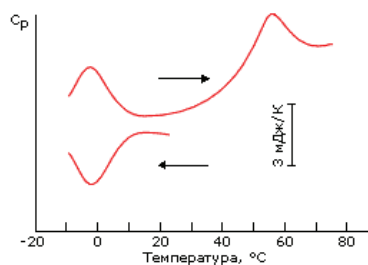
$$\Delta E = \Delta H/N \quad (17.5)$$

— значит плавление целого белка является переходом типа «все-или-ничего». Это и есть «критерий Вант-Гоффа».

Если  $\Delta E < \Delta H/N$  (т. е. если ширина перехода  $\Delta T$  больше, чем того требует критерий Вант-Гоффа) — то «единица плавления» меньше, чем весь белок, т. е. он плавится по частям. Если  $\Delta E > \Delta H/N$  — то «единица плавления» больше белка, т. е. плавится как целое не одна молекула белка, а какой-то белковый агрегат.

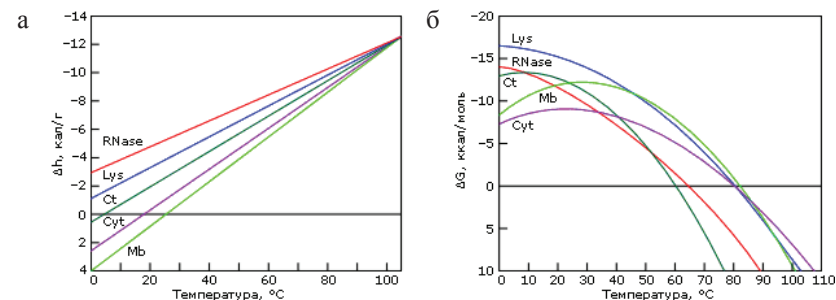
■ Разрушение структуры белка при повышении температуры — плавление — выглядит естественно; однако существует и «холодовая» денатурация белка при аномальном понижении температуры (рис. 17-4). Правда, для многих белков такая денатурация не наблюдается, так как вода в пробире успевает замерзнуть раньше...

Трудами лаборатории Привалова, где «холодовая» денатурация была впервые обнаружена на опыте, было показано также, что она, как и обычная тепловая денатурация, происходит как переход типа «все-или-ничего».



**Рис. 17-4.** Обратимая «холодовая» денатурация белка (апомиоглобина) при аномальном понижении температуры (нижняя кривая; стрелка показывает, что температура в ходе опыта понижается). Верхняя кривая показывает ренатурацию белка (левый пик) при обратном повышении температуры и его дальнейшую тепловую денатурацию (правый пик). Кривые относятся к избыточной теплоемкости раствора белка по сравнению с теплоемкостью растворителя; они не отнормированы на количество белка в растворе. Картинка, с небольшими изменениями, взята из Yu. V. Griko, P. L. Privalov, S. Yu. Venyaminov & V. P. Kutyshenko, *J. Mol. Biol.* (1988) **202**:127–138

Причина «холодовой» денатурации (тут я привожу аргументы Брандтса, предсказавшего ее задолго до опытов Привалова и Грико) кроется в том, что гидрофобные взаимодействия, которые держат белок компактным, сильно растут с температурой, т. е. сильно падают с ее понижением. Рост гидрофобных эффектов — и, в частности, их энергии — с температурой проявляется в относительно высокой теплоемкости денатурированного, менее компактного состояния белка (см.  $\Delta C_p^d$  на рис. 17-2). В результате теплота плавления белка сильно растет с температурой и *падает* с понижением температуры (рис. 17-2, 17-5а). Падает до того, что может изменить знак при низких температурах (рис. 17-5а)! То есть, если обычно более упорядоченная структура (нативный белок) имеет более низкую энергию, чем менее упорядоченная (денатурированный белок) — то при аномально низких температурах ( $\sim 10^\circ\text{C}$  и ниже) энергия нативного белка (плюс энергия окружающей его воды) может быть *выше*, чем денатурированного. То же самое, однако, происходит и с энтропией денатурации: она, как и положено энтропии гидрофобного эффекта, тоже падает с падением температуры. В результате стабильность белка (разность свободных энергий его нативной и денатурированных форм) проходит через максимум при «комнатных» или нулевых температурах и начинает падать при более значительном охлаждении (рис. 17-5б). При этом иногда стабильность белка падает до того, что при аномально низкой (с точки зрения физиологии:  $\sim 0^\circ\text{C}$  и ниже) температуре белок разваливается. Правда, обычно дело до этого не доходит, так как вода с белком замерзает, и все процессы крайне замедляются (что позволяет хранить белок на холоде).



**Рис. 17-5.** Изменение удельной (на грамм белка) разности энергий нативного и денатурированного состояний белка с температурой (а) и изменение свободных энергий в расчете на моль белковых молекул (б). Картинки, с небольшими изменениями, взяты из P. L. Privalov & N. N. Khechinashvili, *J. Mol. Biol.* (1974) **86**:665–684

Парадоксально, но при холодной денатурации дело обстоит в точности так, как при кипении воды. Как известно, кипения можно добиться двумя путями: либо подняв температуру, либо понизив внешнее давление. Гидрофобное давление, сжимающее белковую цепь, сильно падает при уменьшении температуры — и белок может «закипеть».

Кстати — действительно «закипеть» (в смысле — резко понизить свою плотность): если при повышении температуры белок как бы «плавится», т. е. его объем обычно растет не слишком сильно, то при холодной денатурации белковая цепь всегда совершенно разворачивается и занимаемый ею объем растет многократно.

■ Это естественно подводит нас к вопросу о том, как выглядит денатурированный белок.

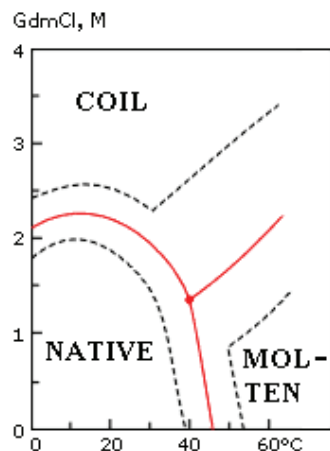
Бурные дискуссии на эту тему велись еще в 1950-х гг., продолжались в 60-х, 70-х, 80-х и только потом привели к некоторому консенсусу.

Проблема была в том, что разные опыты приводили к разным выводам.

Так, многочисленные *термодинамические* опыты показывали, что в области денатурированного состояния белковой молекулы в ней нет никаких кооперативных структурных превращений. Поэтому первоначально полагали, что денатурированный белок всегда является очень рыхлым клубком (каковым он, действительно, является в хороших растворителях типа концентрированных растворов мочевины).

Однако столь же многочисленные *структурные* исследования денатурированных белков упорно свидетельствовали о каких-то крупных перестройках в рамках денатурированного состояния, о каких-то «промежуточных» структурных состояниях между полностью развернутой (клубковой) и нативной формами молекулы (рис. 17-6).

**Рис. 17-6.** Фазовая диаграмма конформационных состояний лизоцима при pH 1,7 в растворе гуанидингидрохлорида при различных температурах. NATIVE — нативное состояние, COIL — клубок, MOLTEN — температурно-денатурированное состояние. Сплошная линия показывает середину перехода, штриховая — зону перехода (предположительно, от соотношения 9:1 в пользу одного состояния до 1:9 в пользу другого; видно, что переход COIL — MOLTEN значительно шире, «менее кооперативен», чем переходы NATIVE — COIL и NATIVE — MOLTEN). Картинка, с небольшими изменениями, взята из С. Tanford, *Adv. Prot. Chem.* (1968) **23**:121–282

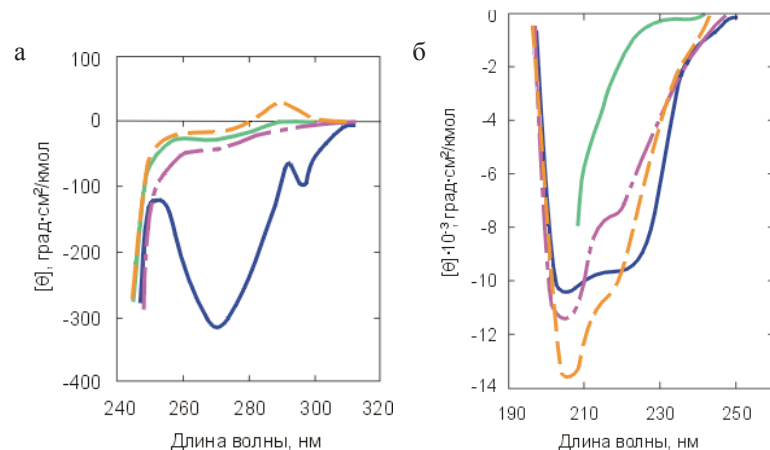


Вообще, экспериментальные данные о состоянии белковой молекулы после денатурации были очень противоречивы. В одних случаях белки казались полностью «развернутыми» и бесструктурными, в других — довольно структурированными и/или компактными, а для их полного разворачивания нужны были сильные растворители («денатуранты»).

Разобраться в этой противоречивой картине удалось только, исследуя структуры денатурированных белков целым букетом методов. В этот букет входили и измерения вязкости (дающие объем молекул белка), и спектры КД в дальнем ультрафиолете (отражающие вторичную структуру белка), и спектры КД в ближнем ультрафиолете (отражающие асимметрию, т. е. упорядоченность окружения ароматических боковых групп), а также и инфракрасная спектроскопия, и ЯМР, и исследование активности белков, и многое другое. Весь этот арсенал и был применен к исследованию денатурации белков группой О. Б. П., к которой А. В. Ф. принадлежал в качестве теоретика.

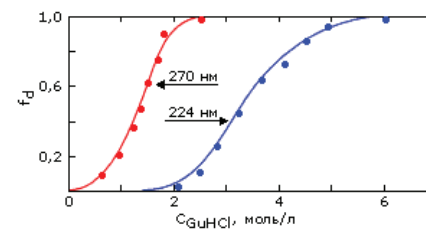
Оказалось, что, кроме активности белка, есть только две вещи, которые всегда резко меняются при денатурации. Это — (1) упорядоченность окружения боковых групп белка, наблюдаемая как по КД в ближнем ультрафиолете (рис. 17-7а), так и при помощи ЯМР, и (2) фиксированность глобулярной структуры, наблюдаемая — посредством ЯМР — по скорости обмена водорода Н полярных групп белка на дейтерий D окружающей воды, а также по ускорению протеолиза белковой цепи.

Степень же упорядоченности главной цепи белка (ее вторичная структура), размеры глобулы и даже плотность ее гидрофобного ядра могут как практически сохраняться, так и меняться очень сильно (рис. 17-7б).



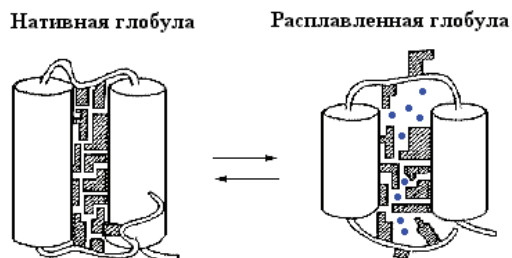
**Рис. 17-7.** Спектры КД в ближнем (а) и дальнем (б) ультрафиолете для нативного  $\alpha$ -лактальбумина (—) и для него же, денатурированного теплом (—), кислотой (—) и сильным денатурантом (—). Спектры КД в ближнем УФ всегда сильно меняются при денатурации. Спектры КД в дальнем УФ существенно меняются только при сильном разворачивании белка денатурантом (см. кривую —). Картинки, с разрешения, взяты из статьи: D. A. Dolgikh, L. V. Abaturov, I. A. Bolotina, E. V. Brazhnikov, V. E. Bychkova, V. N. Bushuev, R. I. Gilmanshin, Yu. O. Lebedev, G. V. Semisotnov, E. I. Tiktopulo & O. B. Ptitsyn. Compact state of a protein molecule with pronounced small-scale mobility: bovine alpha-lactalbumin. (Figs. 3a, 4). *Eur. Biophys. J.* (1985) **13**:109–121, © Springer-Verlag

В ходе этих исследований был обнаружен весьма универсальный интермедиат разворачивания (и сворачивания) белков, позже названный «расплавленной глобулой» (рис. 17-8, 17-9). Важнейшие свойства этого интермедиата суммируются табл. 17-1.



**Рис. 17-8.** Относительное изменение спектров КД  $\alpha$ -лактальбумина в ближнем (270 нм) и в дальнем (224 нм) ультрафиолете с изменением концентрации денатуранта (гуанидингидрохлорида).  $f_d = 0$  отвечает сигналу от нативной структуры,  $f_d = 1$  — от полностью денатурированной молекулы. «Промежуточное состояние» существует между первым и вторым переходом. В области, отвечающей примерно двум молям гуанидингидрохлорида, все молекулы белка существуют (как показало их дальнейшее исследование) в виде «расплавленных глобул». Картинка, с небольшими изменениями, взята из W. Pfeil, V. E. Bychkova & O. B. Ptitsyn, *FEBS Letters* (1986) **198**:287–291





**Рис. 17-9.** Схематическая модель нативной и расплавленной глобулы в молекуле белка. Для простоты, на рисунке показаны только две  $\alpha$ -спирали, соединенные петлей. Каждая вторичная структура покрыта множеством боковых групп; прошитая водородными связями, вторичная структура стабильна, пока проникающий внутрь растворитель не «растворит» глобулу. Вода обычно не способна это сделать без добавления денатуранта. Входящие в гидрофобное ядро белка боковые группы заштрихованы. В расплавленной глобуле они обретают свободу движений, выигрывают энтропию, но теряют энергию плотного контакта. Так в расплавленной глобуле возникают поры, куда вода (•) проникает, не разваливая расплавленную глобулу

Таблица 17-1. Основные свойства расплавленной глобулы и методы их регистрации

«ГЛОБУЛА» (сходство с нативным белком)		«РАСПЛАВЛЕННАЯ» (несходство с нативным белком)	
Гидродинамический объем Мало- и среднеугловое рассеяние рентгеновских лучей	⇒ КОМПАКТНОСТЬ	КД в ближнем УФ $^1\text{H}$ ЯМР спектры	⇒ НЕТ УНИКАЛЬНОЙ УПАКОВКИ БОКОВЫХ ГРУПП
Большеугловое рассеяние рентгеновских лучей	⇒ НАЛИЧИЕ ЯДРА	Водородный обмен, ЯМР Протеолиз	⇒ ФЛУКТУАЦИИ МОЛЕКУЛЫ
ЯМР (спиновое эхо)	⇒ ЗАЖАТОСТЬ ЧАСТИ АРОМАТИЧЕСКИХ БОКОВЫХ ГРУПП	ЯМР (спиновое эхо)	⇒ ПОДВИЖНОСТЬ АЛИФАТИЧЕСКИХ БОКОВЫХ ГРУПП
КД в дальнем УФ ИК спектры ЯМР с $\text{H} \leftrightarrow \text{D}$ обменом	⇒ ВТОРИЧНАЯ СТРУКТУРА	Сканирующая и изотермическая микрокалориметрия	⇒ НЕТ (как правило) ДАЛЬНЕЙШЕГО ПЛАВЛЕНИЯ
Флуоресценция	⇒ ЧАСТЬ ОСТАТКОВ Tr СКРЫТА ОТ ВОДЫ	Флуоресценция	⇒ ЧАСТЬ ОСТАТКОВ Tr ВЫХОДИТ В ВОДУ
Двумерный ЯМР	⇒ СОХРАНЕНИЕ ЧАСТИ ДАЛЬНИХ ПО ЦЕПИ КОНТАКТОВ	Двумерный ЯМР	⇒ РАСПАД БОЛЬШИНСТВА ДАЛЬНИХ ПО ЦЕПИ КОНТАКТОВ
Хроматография (HPLC)	⇒ ВОЗМОЖНОСТЬ ОБРАЗОВАНИЯ «ПРАВИЛЬНЫХ» SS-СВЯЗЕЙ		

**ОТЛИЧИЕ и от нативного, и от развернутого белка:**

Усиленное связывание неполярных молекул  
(флуоресценция связавшихся с белком гидрофобных красителей, АНС)

Свойства расплавленной глобулы на первый взгляд противоречивы. В ней существует развитая вторичная структура и притом обычно практически в том же количестве, что и в нативном белке. Зато в ней (как и в полностью развернутом белке) почти отсутствует упорядоченность боковых групп, столь характерная для нативного белка. В то же время в ней явно остается какая-то доля нативных контактов боковых групп, особенно ароматических (но не алифатических), это показывает ЯМР. Она немного менее компактна, чем нативный белок (как показывает измерение ее гидродинамического объема), а в своем ядре даже почти так же компактна (как показывает «среднеугловое» рассеяние рентгеновских лучей) — но довольно высокая скорость водородного обмена в расплавленной глобуле показывает, что по крайней мере отдельные молекулы растворителя легко проникают вглубь нее. До известного предела эти противоречия удается примирить предположением о неоднородности расплавленной глобулы, о наличии в ней более плотного, более нативоподобного ядра и более распущенных петель — но все же предположение о полностью нативном ядре и полностью развернутых петлях явно не согласуется с опытом, хотя бы с отсутствием в расплавленной глобуле упорядоченности окружения даже ароматических боковых групп. Последнее наблюдается как по КД в ближнем ультрафиолете, так и по спектрам ЯМР.

Для очень многих (хотя и не для всех) белков «расплавленная глобула» появляется при умеренном денатурирующем воздействии на нативный белок и исчезает (превращается в неупорядоченный клубок) лишь с добавлением концентрированного денатуранта. Состояние, похожее на расплавленную глобулу, часто получается и при повышении температуры. При этом переход нативного белка в расплавленную глобулу всегда, когда этот вопрос выяснялся, оказывался переходом «все-или-ничего». «Расплавленная глобула» обычно не плавится кооперативно при дальнейшем нагреве, но ее разворачивание в клубок сильным денатурантом выглядит как кооперативный S-образный переход.

Некоторые нативные белки (особенно небольшие), однако, разворачиваются прямо в клубок, минуя стадию расплавленной глобулы. А некоторые другие белки одними агентами денатурации (например, температурой или кислотой) превращаются в расплавленную глобулу, а другими (например, мочевиной) — в клубок.

■ Противоречивость результатов структурных и термодинамических исследований денатурации белков, а также физическая природа нового фазового состояния белков, состояния «расплавленной глобулы», требовали теоретического осмысления.

Для понимания природы состояния расплавленной глобулы важно, что она получается из нативного состояния путем кооперативного температурного плавления — фазового перехода первого рода. Это означает, что оно обладает значительно большей энтальпией и энтропией, чем нативное состояние, т. е. что внутримолекулярные взаимодействия в нем резко ослаблены, а подвижность белковой цепи резко увеличена. Так как большинство внутренних степеней свободы в белке связано с мелкомасштабными флуктуациями структуры, и прежде всего с движениями боковых групп, именно раскрепощение таких флуктуаций может сделать состояние расплавленной глобулы термодинамически выгодным. Однако раскрепощение мелкомасштабных флуктуаций *не требует* полного разворачивания белковой цепи — достаточно лишь ее небольшого набухания. При этом, однако, в белке резко ослабляется Ван-дер-Ваальсово притяжение: оно сильно зависит от расстояния и даже небольшого увеличения размеров молекулы достаточно для его значительного ослабления (см. рис. 17-9).

В общем, здесь все похоже на плавление кристалла: небольшой рост его объема вчетверо часть Ван-дер-Ваальсовых взаимодействий в нем и раскрепощает движения молекул. Неясно было только, почему столь гетерогенная система, как белок, разрушается фазовым переходом, как кристалл, составленный из одинаковых молекул (или молекул немногих сортов): ведь в белке нет выраженной периодической кристаллической решетки, разрушение которой и плавит кристалл.

Изучение «расплавленной глобулы» послужило и ключом к пониманию кооперативности денатурации белка — процесса, не похожего на то, что происходит ни с синтетическими полимерами, ни с ДНК или РНК.

Что до последних двух — понятно: у них нет единого перехода типа «все-или-ничего». То же касается заодно и переходов спираль-клубок в полипептидах.

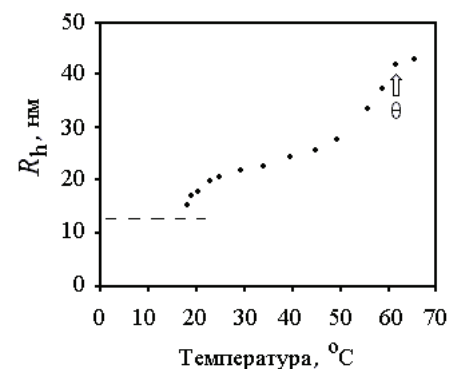
А вот, скажем, уже изученное нами образование  $\beta$ -структуры? Ведь там переход типа «все-или-ничего» существует? Да, существует. Однако при этом «не- $\beta$ -структурная» цепь — типичный клубок: она совсем рыхлая, и в ней нет вторичной структуры. А при денатурации белка часто образуется расплавленная глобула, где практически сохраняется как количество вторичной структуры, так и объем нативного белка.

■ Не могут ли нам помочь синтетические полимеры, не имеющие вторичной структуры? Ведь в них существует переход глобула-клубок... Может быть, этот переход является ближайшим аналогом денатурации белка? Одно время так и думали.

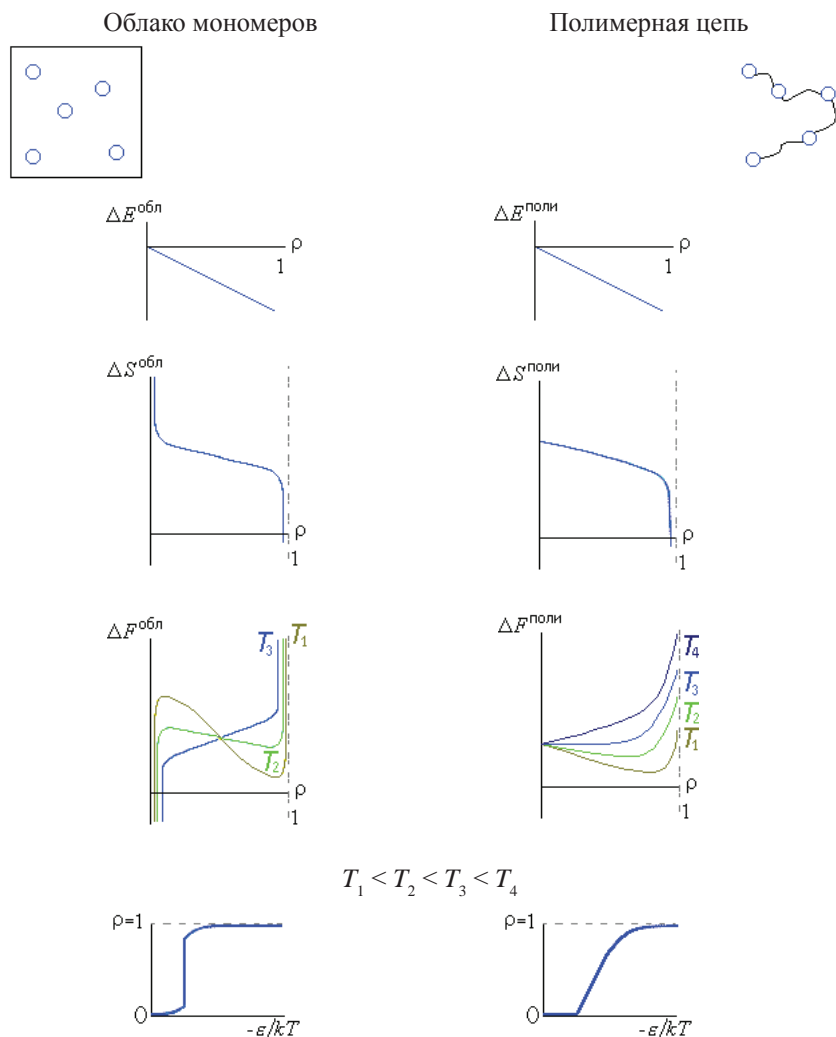
Однако потом оказалось, что процесс денатурации белка качественно отличается от переходов глобула-клубок в «нормальных» полимерах.

Дело в том, что как эксперименты с гомополимерами (крайне трудные: из-за угрозы агрегации их приходится проводить при предельно низких концентрациях полимера), так и физические теории гомополимеров и статистических гетерополимеров (не «отобранных», как белки, а «нормальных», т. е. «случайных») расходились с тем, что наблюдается в белках.

Они все говорили (рис. 17-10, 17-11), что разрушение плотной глобулы в полимерах должно начинаться постепенно и до самого (или почти до самого) конца идти без резкого скачка плотности молекулы, связанного с переходом одной фазы («глобула») в другую («клубок»). Они говорили также, что к моменту превращения глобулы в клубок она уже становится очень рыхлой и что эти две фазы различаются, в основном, масштабом флуктуаций плотности молекулы (малым — в глобуле, и большим — в клубке), т. е., что этот переход совсем *не* похож на фазовый переход первого рода — скажем, на плавление или испарение. А денатурация белка похожа как раз на плавление или испарение твердого тела.



**Рис. 17-10.** Зависимость гидродинамического радиуса  $R_h$  от температуры для разбавленного раствора цепей поли(метил-метакрилата), с молекулярным весом  $6,5 \cdot 10^6$ , в изоамил ацетате [по данным Kayaman N., Guerel E. E., Baysal B. M., Karasz F., *Macromolecules* (1999) **32**, 8399]. Точка перехода глобула-клубок (« $\theta$ -точка») приходится на  $61^\circ\text{C}$  (эта точка отмечена стрелкой). Глобула располагается при температуре ниже  $61^\circ\text{C}$ , клубок — выше  $61^\circ\text{C}$ . Рассчитанный радиус глобулы из твердого поли(метил-метакрилата), с молекулярным весом  $6,5 \cdot 10^6$ , показан пунктиром. Видно, что клубок имеет примерно в четыре раза больший радиус (и, стало быть, почти на два порядка больший объем и меньшую плотность), чем глобула из твердого полимера, и что вслед за переходом клубок-глобула (при понижении температуры) цепь образует набухшую глобулу с почти вдвое большим радиусом (т. е. с почти на порядок меньшей плотностью), чем глобула из твердого полимера (см. область  $50\text{--}30^\circ\text{C}$ ). За этой первой компактизацией следует, при дальнейшем понижении температуры, вторая (см. область  $25\text{--}20^\circ\text{C}$ ), и только она ведет уже к плотной глобуле



**Рис. 17-11.** Качественное сравнение поведения зависимости энергии  $\Delta E$ , энтропии  $\Delta S$  и свободной энергии  $\Delta F = \Delta E - T\Delta S$  добавления одного мономера от плотности  $\rho$  для двух систем: «облако мономеров» и «полимерная цепь». Графики для энергий приведены для притягивающихся ( $\epsilon < 0$ ) мономеров. Зависимости  $\Delta F$  от  $\rho$  показаны для разных температур  $T$ . На дополнительных графиках (внизу) показана качественная зависимость плотности  $\rho$  обеих систем от энергии взаимодействия мономеров, выраженной в единицах  $kT$ ; графики относятся к невысокому внешнему давлению [при очень высоком, «надкритическом» давлении разреженных ( $\epsilon < 0,5$ ) состояний не существует]

■ Несмотря на то, что классическая теория переходов глобула-клубок в гомополимерах *не* применима к денатурации белков, я считаю не лишним рассказать о ней, естественно, в самой упрощенной форме. Такой рассказ, с одной стороны, позволит мне позже отгнать то, что существенно именно для белковых, «отобранных» цепей, а не просто для полимеров. С другой стороны, эта теория вполне применима к поведению уже денатурированных белковых молекул. И, наконец, я считаю, что знание базовых физических моделей необходимо для общей культуры.

■ Я рекомендую вам «поверять слова формулами» при чтении дальнейшего текста. Конечно, проще читать только слова, но они часто допускают неоднозначное толкование; так что, поверяя слова формулами, вы сами проверите свое понимание.

Итак, я рассмотрю простейшую «полимерную» модель — много ( $N$ ) одинаковых и связанных в цепь мономеров («бусинок»). И я сравню ее с простейшей моделью, используемой со времен Ван-дер-Ваальса (XIX век!) для описания переходов газ-жидкость, с «облаком мономеров», т. е. с совокупностью из  $N$  таких же, но не связанных в цепь мономеров (см. рис. 17-11). Мы увидим, что в полимере нет аналога расслоения на «жидкую» и «газообразную» фазы.

*Предварительное замечание.* Ниже я буду считать, что мономеры притягиваются (иначе слишком скучно: плотная фаза не образуется, и все) и что интенсивность этого притяжения не зависит от температуры, т. е. что оно имеет чисто энергетическую природу. Последнее часто неверно, так как мономеры плавают в растворителе (вспомните о гидрофобном эффекте: он растет с температурой!), но так легче найти и рассмотреть все основные сценарии возможных фазовых превращений. А потом уже можно усложнить задачу — учесть зависимость «энергии» притяжения от температуры — и указать, когда, при каких температурах, какой из этих сценариев реализуется.

Сначала я хочу дать качественную картину и объяснить, почему «облако» может претерпевать скачок из плотной в разреженную фазу, а полимер не может, т. е. почему он должен сгущаться постепенно.

■ Чем жизнь мономера в полимере отличается от его жизни в облаке? Тем, что в полимере он не может далеко отойти от своего соседа по цепи; а в облаке он может отлететь от любого другого мономера куда угодно, пока не наткнется на стенку сосуда.

Почему из мономеров образуется плотная фаза — будь то жидкость (если мономеры не связаны цепью) или глобула (если связаны)? Потому, что мономеры притягиваются друг к другу, и, при низкой температуре,

когда роль энтропии падает (вспомните: в свободной энергии  $F = E - TS$  энергия  $E$  сравнивается с энтропией, *умноженной* на температуру), энергия притяжения побьет любое желание мономеров рассредоточиться по пространству и выиграть энтропию. Побьет всегда — *кроме* одного случая: когда такое рассредоточение приведет к практически бесконечно большому энтропийному выигрышу. А последнее возможно тогда, когда облако мономеров помещено в очень большой объем. И вот здесь-то мономерам этого облака приходится выбирать между двумя фазами: либо быть каплей (низкая энергия, *но и* малая энтропия), либо быть разреженным газом (ничтожная энергия, *но зато* очень большая энтропия: летай, где угодно!). А вот промежуточные состояния — с не очень низкой энергией и с невысокой энтропией (когда плотность, скажем, раза в 2 или в 10 меньше, чем у капли) — будут иметь более высокую свободную энергию, и потому не будут стабильны. Так что стабильной — при низкой температуре — будет либо очень плотная фаза облака, либо очень разреженная, но не нечто рыхлое. Иначе говоря, при изменении внешних условий (скажем, при повышении температуры) в «облаке мономеров» может наблюдаться резкое испарение плотной фазы.

Так обстоит дело в «облаке». А в полимере? *Не* так, потому что в самой разреженной глобуле мономер никак не может далеко отойти от своего соседа по цепи, т. е. его энтропия в самом разреженном клубке никогда не превазойдет какого-то предела и не сможет (при низкой температуре) конкурировать с энергией, выигрываемой при сжатии глобулы. Так что полимеру не из чего выбирать: при низкой температуре его мономерам никак не разлететься по всему сосуду — как бы велик тот ни был, а быть ему только сжатым в глобулу, и, значит, в отличие от «облака», он не может претерпеть скачок из плотной в разреженную фазу. То есть при изменении внешних условий (скажем, температуры) возможно только постепенное разрыхление (или уплотнение) полимерной глобулы, но никак не ее резкое «испарение».

Теперь повторим то же самое на языке формул.

Сравним «облако» из  $N$  мономеров, помещенных в объем  $V$ , с глобулой того же объема, образованной цепью из также  $N$  мономеров. Считая, что мономеры и «облака», и полимера более или менее равномерно распределены по объему, добавим к каждой из этих систем еще один мономер и посмотрим, как зависит свободная энергия этого мономера (или, как говорят, его *химический потенциал*) от плотности системы.

- Химический потенциал — удобный инструмент анализа сосуществования фаз. Дело в том, что молекулы перетекают из той фазы, где их химический потенциал выше, в ту, где их химический потенциал ниже, — это понижает общую свободную энергию системы и приближает ее к равновесию.

А в равновесии химический потенциал молекул в одной фазе равен химическому потенциалу тех же молекул в другой фазе. Значит, скачкообразный переход одной фазы в другую возможен, только если разным плотностям, разным фазам системы отвечает один и тот же химический потенциал. А если химический потенциал системы только растет (или только падает) с ее плотностью, никакого скачкообразного фазового перехода быть не может.

Еще надо заметить, что для устойчивости системы надо, чтобы ее химический потенциал *рос* с ростом плотности. Тогда каждая новая частица будет входить в эту систему все с большим и большим трудом. Это — устойчивое состояние: система все более и более (как пружина) сопротивляется усиливающемуся воздействию на нее. В противном случае система не устойчива к спонтанным

Суммирую: конкретное состояние системы определяется конкретными внешними условиями (давлением или объемом), но необходимой предпосылкой наличия фазового перехода в системе является наличие двух подымающихся (с ростом плотности) кусков графика («ветвей») величины химического потенциала, разделенных областью, на которой химический потенциал спадает с ростом плотности..

- Исследуя величину химического потенциала, рассмотрим сначала энергию добавленного мономера в каждой системе («облако», полимер), точнее, ту ее часть, что зависит от плотности, т. е. от взаимодействия ковалентно не связанных звеньев.

В обоих случаях энергия системы (и полимера, и имеющего ту же плотность «облака») меняется — при добавлении мономера — на *одну и ту же* величину. Она меняется на величину энергии взаимодействий добавленного мономера,  $\Delta E = \varepsilon\rho$ , где  $\rho = N\omega/V$  — плотность системы, т. е. доля ее объема  $V$ , занятая  $N$  мономерами (с объемом  $\omega$  у каждого), а  $\varepsilon$  — энергия всех взаимодействий мономера в предельно сжатой системе (при  $\rho = 1$ ). Мы будем считать, что  $\varepsilon < 0$ , т. е. что мономеры притягиваются.

Итак, энергии меняются одинаково, а вот энтропия мономера, добавленного к полимеру и к облаку мономеров, меняется с плотностью *по-разному*.

В «облаке» одному мономеру предоставлен объем  $V/N$ . Точнее — так как прочие мономеры занимают объем  $N\omega$  — свободный объем для мономера в «облаке» составляет

$$V_1^{\text{обл}} = (V - N\omega)/N = (V/N)(1 - \rho) = (\omega/\rho)(1 - \rho). \quad (17.6)$$

Обратите внимание, что этот объем неограниченно велик при стремлении плотности  $\rho$  к 0.

Не то в полимере. Здесь «новому» мономеру доступен лишь ограниченный объем



$$V_1^{\text{полн.}} = \Omega(1 - \rho), \quad (17.7)$$

причем объем  $\Omega$  ограничен связью мономера с предшествующим звеном цепи, а фактор  $(1 - \rho)$ , как и раньше, учитывает, что прочие звенья занимают часть  $\rho$  всего объема.

Результат: изменение энтропии при добавлении одного мономера к «облаку» составляет

$$\Delta S^{\text{обл.}} = k \ln[V_1^{\text{обл.}}] = k \ln[(\omega/\rho)(1 - \rho)], \quad (17.8)$$

а к цепи —

$$\Delta S^{\text{полн.}} = k \ln[V_1^{\text{полн.}}] = k \ln[\Omega(1 - \rho)]. \quad (17.9)$$

Графики зависимости этих функций от плотности  $\rho$  (см. рис. 17-11) ведут себя сходно в районе  $\rho \rightarrow 1$  (где в  $\Delta S$  доминирует член  $\ln(1 - \rho)$ , так что обе  $\Delta S$  падают при  $\rho \rightarrow 1$ ), но *совсем по-разному* при  $\rho \rightarrow 0$ . Если  $\Delta S^{\text{полн.}}$  остается конечной, то  $\Delta S^{\text{обл.}}$  неограниченно растет (из-за члена  $\omega/\rho$ ).

Поэтому для «облака» существует *две растущие с ростом  $\rho$  ветви* величины химического потенциала  $\Delta F = \Delta E - T\Delta S$ , т. е. *две* области потенциально стабильных состояний при низких температурах  $T$  (газ — при малых  $\rho$ , за счет члена  $-T\Delta S$ , и жидкость — при больших  $\rho$ , за счет члена  $\Delta E$ ). А при высоких температурах для «облака» существует только *одна растущая ветвь* величины  $\Delta F = \Delta E - T\Delta S$ , т. е. только одно стабильное состояние (газ — за счет члена  $-T\Delta S$ ).

- В то же время для полимера, где энтропия *не* становится бесконечно высокой при  $\rho \rightarrow 0$ , существует, при всех температурах, только *одна растущая ветвь* величины  $\Delta F$ . При совсем низких температурах (или, точнее, при больших значениях величины  $-\varepsilon/kT$ ) начало этой ветви приходится на  $\rho \approx 1$ , что соответствует плотному глобулярному состоянию. Потом (с ростом  $T$ , а точнее, с падением величины  $-\varepsilon/kT$ ) начало этой растущей ветви постепенно смещается в сторону низких плотностей (что соответствует все более рыхлым глобулярным состояниям), и, наконец, это начало приходит к плотности  $\rho = 0$ , и все (см. рис. 17-11). Как вы видите, у химического потенциала глобулы никогда нет двух растущих ветвей. Значит, весь процесс набухания глобулы нигде не сопровождается расслоением на две фазы. А раз нет расслоения на фазы, нет и перехода типа «*все-или-ничего*», т. е. нет фазового перехода I рода. Можно показать, что при  $\rho \sim 0$  (точнее: при плотности клубка,  $\sim N^{-1/2}$ ) происходит фазовый переход II рода (что и было сделано И. М. Лифшицем, А. Р. Хохловым и А. Ю. Гроссбергом, а также П. Де Женном), но мы этим заниматься не будем.

■ Уточнение. Строго говоря, общая теория утверждает, что скачкообразный переход более или менее плотной глобулы в клубок возможен только при определенном, трудно выполнимом условии. Он возможен лишь для такого полимера, мономеры которого отталкиваются, взаимодействуя попарно, и притягиваются, взаимодействуя в большом числе одновременно. Такой эффект наблюдается в белках, например, в очень жестких цепях. Однако, в отличие от того, что наблюдается в белках, скачок энергии (в расчете на мономер) при этом очень мал (как и при образовании жидких кристаллов из жестких вытянутых молекул). Большой скачок энергии требует очень специальной конструкции мономеров — конструкции, которой не обладают звенья обычных полимерных цепей. Однако белковые цепи, как мы увидим, обладают именно таким свойством: аминокислотные остатки в них начинают сильно притягиваться, именно взаимодействуя в большом числе. Почему это так — выяснится на следующей лекции.

■ Итак, классическая теория переходов «глобула-клубок» в обычных гомополимерах не может объяснить плавление белка. [Не объясняет плавление белка и теория переходов глобула-клубок в случайных гетерополимерах, но обсуждение ее лежит за рамками этих лекций.]

Такая теория говорит, что глобула расширяется постепенно и что клубок возникает только в конце этого расширения, причем вовсе не путем фазового перехода *первого* рода. А денатурация белков, этих «аперiodических кристаллов» Шредингера, происходит при больших плотностях глобулы, часто вовсе не превращает белок в клубок и напоминает именно разрушение кристалла — фазовый переход первого рода.

■ Объяснением этих особенностей денатурации белков, и особенно фазового характера их плавления, мы займемся на следующей лекции.

## Лекция 18

Почему денатурация глобулярного белка — переход типа «все-или-ничего»? Распад плотной упаковки ядра белка и раскрепощение боковых групп. Проникновение растворителя в денатурированный белок, разрушение расплавленной глобулы, постепенное разворачивание цепи денатурированного белка по мере увеличения силы растворителя. Энергетическая щель между нативной укладкой белковой цепи и прочими ее глобулярными укладками: основное физическое отличие белковой цепи от случайного сополимера. Различия в плавлении «отобранного» гетерополимера (с энергетической щелью) и случайного сополимера.

Итак, как мы выяснили на прошлой лекции, классическая теория переходов «глобула — клубок» не может объяснить плавление белка. Она говорит, что глобула расширяется постепенно и что клубок возникает вовсе не путем фазового перехода *первого* рода, а денатурация белков, этих «аперiodических кристаллов» Шредингера, происходит при больших плотностях глобулы и напоминает именно разрушение кристалла, т. е. фазовый переход I рода.

Не может объяснить денатурацию белка и «обычное» плавление кристаллов — ведь в белках нет периодической кристаллической решетки, которая отличает обычный кристалл от жидкости или стекла (тоже жидкости, но очень вязкой), той решетки, распад которой и составляет сущность плавления кристалла.

Удивительно также, что денатурация белка — гетерогенной системы, где каждый атом, тем не менее, сидит на своем месте, — идет просто как резкий переход «все-или-ничего», аналогичный фазовому переходу первого рода в макроскопических системах: в обычных макроскопических молекулярных системах гетерогенность смазывает переход.

- Здесь уместно обратить внимание на одно обстоятельство, которое отличает «обычную» физику от физики белка и вообще от физики сложных систем. Речь, конечно, идет не о физических законах — они, естественно, одинаковы в обоих случаях — но о том, что приковывает внимание исследователей. «Обычная» физика

привыкла оперировать с более или менее однородными вещами: газами, кристаллами, полями и т. д. Поэтому она обращает особое внимание на возникающие в этих «однородных вещах» неоднородности, например, на квазичастицы. Физика белка, напротив, с самого начала имеет дело с крайне неоднородным объектом. Здесь полно напряжений — если хотите, можно сказать, что белок весь наполнен разнообразными флуктуонами, но это будет не более, чем выраженная высоким штилем констатация неоднородности белка. А вот однородные явления, охватывающие такую неоднородную систему, как целое, например, охватывающая сразу весь белок денатурация, действительно представляют особый интерес для физики белка.

**Внутренний голос:** А действительно ли плавление белка всегда происходит буквально как переход типа «все-или-ничего», т. е. из одного состояния в другое, между которыми не наблюдается промежуточных?

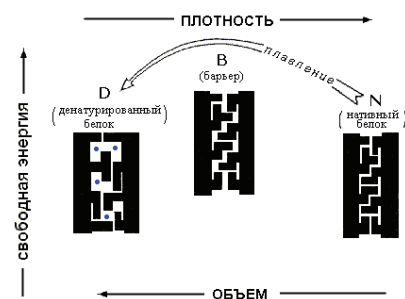
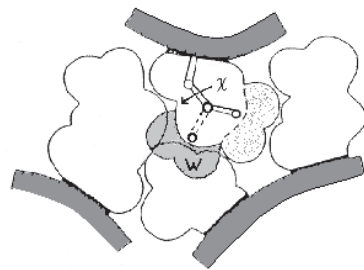
**Лектор:** Во-первых, я сейчас, как обычно, говорю об однодоменных белках (в больших многодоменных есть несколько последовательных — обычно подоменных — переходов типа «все-или-ничего», т. е. такие белки могут быть «полурастворенными»). Во-вторых, в переходе «все-или-ничего», как и в фазовых переходах первого рода, короткоживущие промежуточные состояния наблюдать можно, в основном косвенно, — по их влиянию на кинетику перехода через свободно-энергетический барьер, разделяющий два непосредственно наблюдаемых состояния. Об этом речь пойдет на следующей лекции. В-третьих, было обнаружено, что в некоторых маленьких белках этот свободно-энергетический барьер так мал, что наблюдаемый в них переход сильно напоминает, по заключению В. Муноца, фазовый переход второго рода. Но это — наблюдаемое в очень маленьких белках исключение из общего (присущего белкам среднего размера и большинству маленьких белков) правила.

- Чтобы понять денатурацию белка, надо объяснить, почему существуют два равно стабильных фазовых состояния белковой цепи (как мы уже знаем, в обычных полимерах это невозможно) и почему они разделены свободно-энергетическим барьером (который, как мы знаем, и создает переход типа «все-или-ничего»). То есть надо понять, почему белковая глобула не может разрушиться путем постепенного набухания, как то делают обычные полимерные глобулы.

При таком объяснении надо учесть основные (по сравнению с «просто полимером») особенности белка: то, что он имеет одну укладку цепи, отличающуюся особой стабильностью; то, что в белковой цепи подвижные боковые группы сидят на жесткой главной цепи; и то, что нативный белок упакован плотно, как кристалл (хотя и без кристаллической решетки): в белках, как и в молекулярных кристаллах, Ван-дер-Ваальсовы объемы атомов занимают 70–80 % всего объема, а в жидкостях (в расплавах) — только 60–65 %.

Итак. Боковые группы белковой цепи способны к поворотной изомеризации, т. е. к резким прыжкам из одной разрешенной конформации в другую. Но для такого прыжка необходим определенный свободный объем вблизи «прыгающей» боковой группы. Боковые же группы, слагающие гидрофобное ядро белка, сидят на жесткой (в особенности из-за наличия  $\alpha$ -спиралей и  $\beta$ -листов, необходимых, как мы видели, для создания ядра) главной цепи (рис. 18-1). Эти структуры стабильны, пока в глобулу не попал растворитель (для чего нужен тот же свободный объем). И жесткие эти сегменты перемещаются как целое (по крайней мере в начале расширения глобулы), со всем своим лесом боковых групп. Поэтому расширение глобулы, т. е. расхождение этих сегментов (вспомним, что нативная укладка белковой цепи отличается плотной упаковкой вторичных структур в глобулу, и поэтому особой стабильностью), создает примерно одинаковый свободный объем вблизи каждой боковой группы, и он либо недостаточен для изомеризации каждой из них (пока глобула расширилась мало), либо достаточен для поворотной изомеризации многих. Так что освобождение боковых групп (и проникновение растворителя) наступает лишь тогда, когда расширение глобулы превосходит некий порог, «барьер» (рис. 18-2). Эти два события могут сделать менее плотное состояние белковой цепи столь же стабильным, как ее нативное состояние, но не ранее, чем барьерная плотность останется позади.

**Рис. 18-1.** Схема упаковки боковых групп. Показан лишь маленький фрагмент ядра. Заштрихованная область  $W$  соответствует альтернативному ротамеру боковой группы ( $\chi$  — угол ее вращения); этот ротамер запрещен плотной упаковкой. Для его появления нужен дополнительный объем  $W \approx 30 \text{ \AA}^3$  (т. е.  $\approx 1/5$  объема среднего аминокислотного остатка)

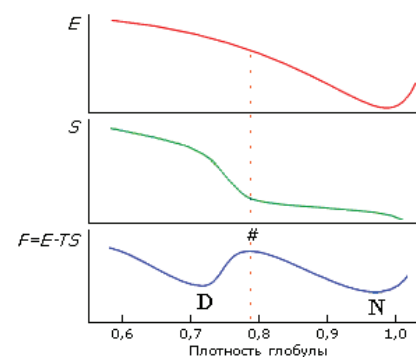


**Рис. 18-2.** Причина существования свободно-энергетического барьера между нативным и любым денатурированным состоянием белка. «Барьерное» состояние возникает при небольшом расширении нативного. В этом состоянии поры в белке уже приводят к значительному повышению Ван-дер-Ваальсовой энергии, но еще не позволяют ни поворотной изомеризации, ни проникновения воды (\*) внутрь белка

Поэтому малое («до-барьерное») расширение нативной глобулы *всегда* невыгодно — оно повышает ее энергию (так как части глобулы уже расходятся), но *не* повышает ее энтропию, так как *еще не* открывает поворотной изомеризации ее боковых групп. То есть свободная энергия глобулы растет при малом расширении. Наоборот, большое («за-барьерное») расширение глобулы открывает поворотную изомеризацию и ведет (при достаточно высокой температуре) к падению ее, глобулы, свободной энергии. В результате денатурация белка при изменении внешних условий происходит не постепенно, а скачком (скачком через барьер), по принципу «все-или-ничего».

То есть белок, не меняясь, терпит изменение внешних условий до некоторого предела, а потом плавится, как микроскопическое твердое тело, весь сразу. Такая устойчивость и твердость белка, в свою очередь, обеспечивает надежность его работы в организме.

Иными словами, фазовый переход между нативным и денатурированным состояниями объясняется скачкообразным ростом энтропии (и прежде всего — энтропии боковых групп) при расширении глобулы (рис. 18-3),



**Рис. 18-3.** Причина наличия перехода «все-или-ничего» при денатурации белка. Энергия  $E$  минимальна при плотной упаковке глобулы (здесь ее плотность  $\rho = 1$ ) и монотонно растет при сжатии или расширении. Энтропия  $S$  растет с падением плотности  $\rho$  сперва медленно (при  $\rho = 1,0-0,8$ , пока боковые группы лишь колеблются, а их поворотная изомеризация еще невозможна), затем быстро (когда поворотная изомеризация начинает раскрепощаться), и затем (когда поворотная изомеризация уже свободна) — опять медленно. Неравномерный рост энтропии связан со следующим эффектом. Для поворотной изомеризации вблизи боковой группы должен быть свободный объем размером по крайней мере с  $\text{CH}_3$ -группу. А так как боковая группа привязана к жесткой главной цепи, такой объем не может появиться у одной лишь группы, он одновременно появится у многих. Поэтому и существует характерная («барьерная») плотность глобулы, где энтропия начинает резко расти с расширением глобулы. Эта плотность довольно высока:  $\approx 80\%$  от нативной плотности, так как объем  $\text{CH}_3$ -группы (характерной поворачивающейся единицы, см. рис. 18-1) составляет  $\approx 20\%$  объема аминокислотного остатка. В результате зависимость суммарной свободной энергии  $F = E - TS$  от плотности глобулы проходит через максимум («барьер» #), отделяющий нативную (плотнупакованную) глобулу  $N$  от любого более рыхлого денатурированного состояния  $D$ . Из-за этого барьера денатурация белка и происходит по типу «все-или-ничего», независимо от дальнейшего состояния денатурированной молекулы.

а его фазовый, кооперативный характер связан с тем, что боковые группы прикреплены к главной цепи и не могут раскрепощаться поодиночке.

■ До открытия расплавленной глобулы денатурация белков обычно трактовалась как полное разрушение структуры и переход глобулы в клубок. После этого открытия стало ясно, что денатурированный белок может быть как довольно плотным, так и рыхлым в зависимости от силы растворителя и гидрофобности цепи. Оценки показывают, что поры расплавленной глобулы (поры, необходимые для свободы движений боковых групп) заполнены растворителем, так как перенос молекулы воды в пору внутри белка термодинамически выгоден по сравнению с существовавшим там до того вакуумом. На опыте об этом свидетельствует ничтожность изменения парциального объема при денатурации белка. Но проникающему внутрь растворителю, если только он не слишком сильно притягивается порами (ведь они окружены в основном неполярными группами, слабо притягивающими воду), удастся только заполнить поры уже существующие (необходимые для свободы движений боковых групп), но не удастся создать новые (уже не необходимые для этих движений и «нужные» лишь растворителю), т. е. не удастся разорвать оставшиеся в белке взаимодействия. Тогда он не сможет расширить глобулу — как вода не может растворить губку, хоть и заполняет ее поры, — и денатурированный белок так и останется компактной расплавленной глобулой (причем «мокрой», насыщенной растворителем расплавленной глобулой).

**Внутренний голос:** Если Вы говорите — «мокрой», то должна быть еще и «сухая» расплавленная глобула?

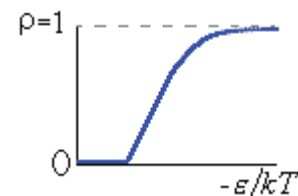
**Лектор:** Да. Это — глобула, поры в которой не заполняются водой. Такая глобула (так же, как и «мокрая») была, из теоретических соображений, предсказана А. В. Ф. и Е. И. Шахновичем в 1980-х годах, но анализ теоретических и экспериментальных данных показал, что сухая расплавленная глобула менее стабильна, чем мокрая, и потому меньше подходит на роль интермедиата плавления белка. Однако совсем недавно сухая расплавленная глобула была обнаружена (Jha & Udgaonkar, *Proc. Natl. Acad. USA* (2009) **106**:12289–12294) при флуктуациях, предшествующих плавлению белка.

■ Компактность расплавленной глобулы обеспечивается остаточными гидрофобными взаимодействиями. Они не очень сильны. Даже в расплавленной глобуле апомиоглобина, обладающей развитой вторичной структурой и почти нативной топологией укладки большей части цепи (Jennings & Wright, *Science* (1993) **262**:892–896), остаточные взаимодействия гидрофобных боковых групп по крайней мере в три раза слабее, чем в нативном белке (Samatova et al., *Protein Sci.*

(2009) **18**:2152–2159), а у некоторых таких групп эти остаточные взаимодействия и вовсе отсутствуют (что подчеркивает неоднородность расплавленной глобулы). Если же остаточные гидрофобные взаимодействия слабы, т. е. если растворитель сильно притягивается к порам в белке, то он начинает расширять поры, глобула начнет разбухать, и тем сильнее, чем сильнее притяжение растворителя к боковой цепи, т. е. чем меньше выгоды находят звенья цепи в контакте друг с другом. По мере роста притяжения растворителя к боковой цепи она разбухнет до неупорядоченного клубка.

При относительно низких температурах, когда роль энтропии раскрепощения боковых групп менее заметна и основная работа падает на растворитель, клубок возникает непосредственно из нативной структуры (вспомните холодовую денатурацию). При высоких же температурах он может, скорее, возникнуть после набухания расплавленной глобулы (вспомните экспериментальную фазовую диаграмму Танфорда; скоро я покажу вам ее опять).

Набухание расплавленной глобулы, возможно, описывается классической теорией переходов глобула-клубок, иллюстрируемой рис. 18-4: если боковая цепь достаточно гидрофобна, денатурированный белок останется в состоянии расплавленной глобулы; если недостаточно — он набухнет или (если растворитель «лучше» или цепь менее гидрофобна) вообще полностью развернется.

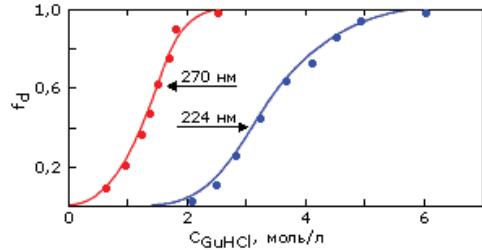


**Рис. 18-4.** Качественная зависимость плотности  $\rho$  обычного полимера от энергии взаимодействия мономеров  $\epsilon$ , выраженной в единицах  $kT$ . Скачка плотности здесь не наблюдается, т. е. это превращение не является переходом типа «все-или-ничего». Однако, начиная с какой-то величины  $-\epsilon/kT$ , скорость роста  $\rho$  с  $-\epsilon/kT$  резко меняется (см. излом кривой, слева), что характерно для фазовых переходов второго рода. В дальнейшем этот рост постепенно затухает, уже без всякого фазового перехода

Так говорит простейшая теория расплавленной глобулы. Фактически, она говорит, что расплавленная глобула не является точно очерченным состоянием со стороны клубка — она не отделена от него переходом «все-или-ничего». Но это — результат простейшей теории, не принимающей во внимание неоднородность белка. Может быть, нужна более изощренная теория?



■ Посмотрим, что говорит эксперимент. Он, прежде всего, говорит, что разворачивание расплавленной глобулы выглядит как кооперативный («S-образный») переход (рис. 18-5). Однако — эта «S-образность» совместима и с рис. 18-4 (разворачиванием, когда плотность меняется быстро, но без скачка), и с переходом типа «все-или-ничего».

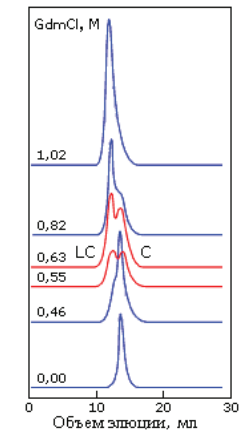


**Рис. 18-5.** Относительное изменение спектров КД  $\alpha$ -лактальбумина в ближнем (270 нм) и в дальнем (224 нм) ультрафиолете с изменением концентрации денатуранта (гуанидингидрохлорида).  $f_d = 0$  отвечает сигналу от нативной структуры,  $f_d = 1$  — от полностью денатурированной молекулы. «Промежуточное состояние» существует между первым и вторым переходом и является (как показало его дальнейшее исследование) «расплавленной глобулой». Картинка, с небольшими изменениями, взята из W. Pfeil, V. E. Bychkova & O. B. Ptitsyn, *FEBS Letters* (1986) **198**:287–291

Итак: является ли разворачивание расплавленной глобулы переходом типа «все-или-ничего» на опыте?

Здесь пока нет полной ясности. Термодинамические исследования многих белков показывают, что, по всей видимости, — нет, не является. Однако Уверским и Птицыным было показано, что состояние расплавленной глобулы может быть, по крайней мере в некоторых белках (например, в карбоксиангидразе), отделено от «менее плотного» состояния переходом типа «все-или-ничего» (рис. 18-6). Правда, карбоксиангидраза —  $\beta$ -структурный белок, так что наблюдаемый переход может относиться к распаду ее  $\beta$ -структуры (в синтетических полимерах он протекает как переход типа «все-или-ничего»). К тому же наблюдающийся в ней переход столь медленен, что засекается хроматографией, а образование расплавленной глобулы в кинетических опытах (о которых будет говориться в следующей лекции) — процесс очень быстрый... Неясности остаются...

В момент перехода это «менее плотное» состояние лишь немногим менее компактно, чем расплавленная глобула, и содержит примерно вдвое меньше вторичной структуры, но оно намного более компактно и структурировано, чем клубок. Поэтому это состояние называют «пред-расплавленной глобулой». Надо отметить, что близость свойств «пред-расплавленной» и «расплавленной» глобул порой ведет к большим недоразумениям...



**Рис. 18-6.** Сосуществование двух фаз, «компактных расплавленных глобул» (C) и «менее плотных глобул» (LC) при хроматографировании белка в области таких концентраций денатуранта (гуанидингидрохлорида), где расплавленная глобула начинает разваливаться. Обратите внимание на раздвоенность пика на хроматограммах, сделанных при 0,55–0,63 М GdmCl. Картинка, с небольшими упрощениями, взята из V. N. Uversky, G. V. Semisotnov, R. H. Pain & O. B. Ptitsyn, *FEBS Letters* (1992) **314**:89–92

Дальнейшее разворачивание «пред-расплавленной» глобулы до клубка является плавным, как того требует показанная на рис. 18-4 теория переходов «глобула — клубок» в «обычных» полимерах, обсуждавшаяся на прошлой лекции.

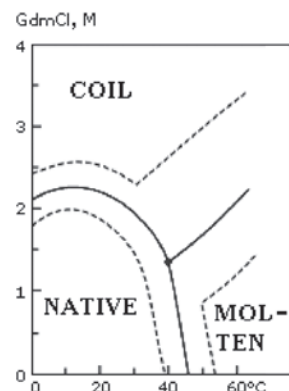
Сосуществование двух фаз — «расплавленных» и «менее плотных» глобул — наблюдается при хроматографировании белка в области таких концентраций денатуранта, где расплавленная глобула разваливается. Наличие раздвоенного пика на таких хроматограммах (рис. 18-6) доказывает, что данный белок существует *либо* в одной форме (расплавленной — это установлено другими измерениями), *либо* в другой (какой-то «менее плотной»), но *не* в форме глобулы какой-то промежуточной между ними плотности.

О наличии переходов типа «все-или-ничего» свидетельствует, в принципе, достаточная узость зоны температурного плавления белка (вспомните критерий Вант-Гоффа). Но критерий Вант-Гоффа не приложишь к плавлению расплавленных белков: как показывает опыт, они, как правило, дальше не плавятся! Однако плавление расплавленных глобул было обнаружено при калориметрическом исследовании ряда искусственно созданных белков, которые, как оказалось, даже при физиологических условиях имеют форму расплавленной глобулы. К тому же дополнительное исследование ширины переходов «расплавленная глобула — клубок» при денатурации природных белков сильными денатурантами показало, что эти ширины для многих однодоменных белков обратно пропорциональны размеру белка, что более всего совместимо с тем, что кооперативной единицей разворачивания является целый такой белок, — то есть с переходом типа «все-или-ничего».

- Похоже, что переход расплавленной глобулы в клубок начинается, по крайней мере в некоторых белках, с ее фазового перехода в «пред-расплавленную» глобулу, а затем последняя разворачивается до клубка уже постепенно. Но выяснение природы этих переходов требует дальнейших экспериментальных и теоретических исследований.

Итак, возвращаясь к фазовой диаграмме Танфорда (рис. 18-7), мы видим, что нативное состояние отделено от всех других (и от клубка COIL, от температурно-денатурированного состояния MOLTEN) переходом типа «все-или-ничего» (что и обуславливает сравнительную узость этих переходов); что денатурированное состояние обязательно отличается от нативного подвижными боковыми группами, но может быть весьма компактным; и, наконец, что переход из температурно-денатурированного состояния в клубок может быть как фазовым переходом типа «все-или-ничего» для одних белков, так и (возможно) не-фазовым переходом — для других.

**Рис. 18-7.** Фазовая диаграмма конформационных состояний лизоцима при pH 1,7 в растворе гуанидинхлорида при различных температурах. NATIVE — нативное состояние, COIL — клубок, MOLTEN — температурно-денатурированное состояние. Жирная линия показывает середину перехода, пунктир — зону перехода (предположительно, от соотношения 9:1 в пользу одного состояния до 1:9 в пользу другого; видно, что переход COIL — MOLTEN значительно шире, «менее кооперативен», чем переходы NATIVE — COIL и NATIVE — MOLTEN). Картинка, с небольшими изменениями, взята из С. Tanford, *Adv. Prot. Chem.* (1968) 23:121–282



Маленькое отступление. Вообще, нас не должно удивлять, что разные белки ведут себя по-разному. И здесь дело не столько в качественных различиях между белками, сколько в том, что мы наблюдаем каждый белок через «экспериментальное окошко», где температурный диапазон ограничен существованием жидкой воды, а диапазон воздействия денатуранта — его нулевой концентрацией. Поэтому, например, менее гидрофобные по своему аминокислотному составу белки — белки, чья глобула легко разрушается денатурантом, — будут демонстрировать в этом «окошке» только переход в клубок (ведь расплавленная глобула держится только остаточными гидрофобными взаимодействиями). И эти же малостабильные белки

(или белки «ослабленные», как, например, апомиоглобин, т. е. миоглобин, ослабленный лишением гема) демонстрируют холодовую денатурацию — денатурацию, при которой всегда образуется клубок. У более стабильных белков она не наблюдается, точнее, она, видимо, просто приходится на температуры, когда вода замерзает и тем прекращает наши опыты...

- Сейчас мы продолжим рассмотрение термодинамики белковых молекул под несколько иным углом зрения, а именно: какие свойства молекулы белка нужны для того, чтобы она разрушалась не постепенно, а переходом типа «все-или-ничего»? То есть, чтобы она, не размягчаясь, долго терпела внешнее воздействие, а потом сразу ломалась бы? (Ясно, что такое поведение способствует надежности работы белка: уж если он работает — то работает, как надо; а не работает — значит, совсем сломался...)

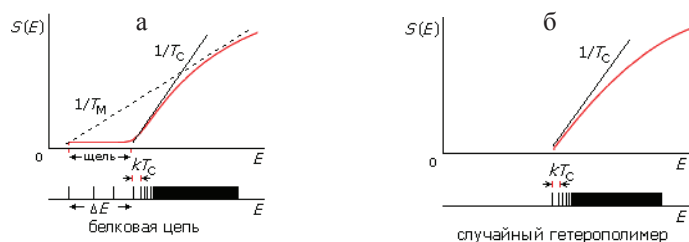
- Имея графики зависимости энтропии  $S$ , и энергии  $E$  от плотности глобулы  $\rho$  (см. рис. 18-3), мы можем построить график зависимости  $S$  от  $E$  и, далее, нарисовать характерный вид энергетического спектра белка (рис. 18-8а). Позвольте напомнить вам, что  $S(E)$  пропорционально логарифму числа структур с энергией  $E$ . Поэтому, зная  $S(E)$ , мы можем оценить плотность энергетического спектра при любой энергии. Позвольте также напомнить вам, что  $dS(E)/dE = 1/T(E)$ , где  $T(E)$  — температура, соответствующая энергии  $E$ .

«Нарисовать характерный вид энергетического спектра» значит показать, как распределены по энергии различные конформации белковой цепи. В спектре нас не интересует, *какая* конформация имеет такую-то энергию; нас интересует только, *сколько* конформаций какую энергию имеет. Каждая линия спектра соответствует одной конформации цепи (уточним: «конформация» здесь соответствует одному локальному минимуму свободной энергии, включающей энергию взаимодействия внутри белка и взаимодействий белка с окружающим растворителем).

Основной чертой полученного спектра белковой глобулы является отрыв энергии самой стабильной, «аномально стабильной», т. е. нативной структуры белка от тех энергий, где начинается спектр подавляющего большинства остальных структур белковой цепи. Именно этот отрыв (или, как говорят, «энергетическая щель между самой стабильной структурой и ее конкурентами») создает барьер свободной энергии между нативной структурой и прочими структурами цепи и заставляет белок плавиться путем перехода «все-или-ничего», а до того «терпеть», не меняясь, изменение внешних условий.

Выше щели (по энергии) лежит практически непрерывный спектр «неправильно сложившихся» структур (непрерывный, так как там очень

много линий, что-то вроде  $\exp(N)$ , где  $N$  — число звеньев в цепи, а ширина спектра энергий пропорциональна всего лишь  $N$ ). Ниже щели лежит нативная структура, а на щель приходится несколько ее «немного разболтанных» вариантов.



**Рис. 18-8.** Предполагаемая зависимость энтропии  $S$  от энергии  $E$  для белковой цепи (а) и для цепи случайного гетерополимера (б). Показаны касательные к нижней части кривых  $S(E)$ , определяющие температуру  $T_C$  стеклования полимера и температуру плавления  $T_M$  белка (температуру сосуществования, т. е. равенства свободной энергии его самой низкоэнергетичной структуры и множества высокоэнергетичных). Наклон второй касательной несколько меньше, т. е.  $T_M$  несколько выше  $T_C$ . То есть при охлаждении белка («цепи со щелью») его нативная, самая низкоэнергетическая структура вымораживается (скачком!) раньше, чем в нем может начаться постепенное стеклование «неверных» структур, лежащих на правом краю щели. Внизу показаны характерные формы энергетических спектров обеих молекул. Каждая линия спектра (здесь показаны лишь немногие из этих линий) соответствует одной конформации цепи. В правой, высокоэнергетичной части обоих спектров этих линий очень много и они сливаются. Основной чертой спектра белка — «отобранного» гетерополимера — является энергетическая щель (ширины  $\Delta E \gg kT_C$ ) между основной, самой стабильной укладкой цепи, и другими, не похожими на нее укладками. Немногие структуры, чьи энергии приходятся на эту щель, представляют собой более или менее «разболтанные» или «недовернутые» варианты основной укладки цепи

Температура плавления белка,  $T_M$ , соответствует наклону касательной (см.  $1/T_M$  на рис. 18-8а), касающейся одновременно как самых малых, практически нулевых, энтропий  $S$  на самом левом («за-щелевом») краю энергетического спектра, так и области высоких энтропий в районе непрерывного энергетического спектра «неверно сложившихся» структур.

Кроме  $T_M$  у спектра белка есть еще одна замечательная температура:  $T_C$ . Она определяется наклоном кривой  $S(E)$  перед самой щелью (см. рис. 18-8а), т. е. скоростью схождения на нет энтропии «неправильно сложившихся» структур цепи. При температуре  $T_C$  «неправильные» структуры должны были бы заморозиться (застекловаться, так как их много, с почти одинаковой энергией) — если бы нативная (отделенная от них

широкой щелью, самая стабильная) структура белка не вымораживалась при еще более высокой температуре  $T_M$ .

А вот у «случайного» сополимера (рис. 18-8б) есть только одна замечательная температура:  $T_C$ , критическая температура его стеклования.

Теорфизические исследования разнообразных моделей гетерополимерных цепей показали, что в случайных гетерополимерах «аномально стабильной» структуры нет, т. е. что щель там не наблюдается (см. рис. 18-8б). Точнее, у случайных цепей она обычно очень мала, порядка  $kT_C$ , так что такие цепи замораживаются (при температуре  $T_C$ ) постепенно, не скачком, как кристалл или белок, а постепенно, как стекло.

■ Значит, необходимая для фазового плавления белка достаточно большая (с шириной  $\Delta E \gg kT_C$ ) энергетическая щель между самой стабильной структурой цепи и ее конкурентами возникает не в любой «случайной» аминокислотной последовательности, а создается отбором «белковоподобных», годных для создания белков последовательностей (последовательностей, допускающих одну-единственную плотную упаковку цепи в глобулу).

Естественный отбор подхватывает такие последовательности, потому что только они способны создавать твердые белковые глобулы со строго определенными структурами, а только среди таких глобул можно найти те, что будут надежно служить организму.

Удалось даже (экстраполируя кривую  $S(E)$ , описывающую левый край непрерывного спектра, на еще более низкие энергии) ориентировочно оценить долю «белковоподобных» цепей (т. е. цепей, имеющих энергетическую щель данного размера), среди всех случайных цепей:

$$\text{Доля } (\Delta E) \sim \exp(-\Delta E/kT_C). \quad (18.1)$$

То есть эта доля очень мала, если щель значительна ( $\Delta E \gg kT_C$ ) — что разумно: белок сделать нелегко! Однако эта доля все же не безумно мала: так, при  $\Delta E \approx 10$  ккал/моль, что достаточно велико по сравнению с  $kT_C \sim 1$  ккал/моль, эта доля составляет  $\sim 10^{-8}$ .

Найти и выделить такую малую, но крепко и специфично к чему-то липнущую, долю молекул уже по силам современным методам, скажем, фаговому дисплею. Поэтому не столь удивительно, что ведущиеся сейчас опыты со случайными полимерами аминокислот позволяют порой получить «белковоподобные» молекулы. Недавно сделанные опыты показывают, что такие молекулы составляют  $\sim 10^{-10}$ – $10^{-12}$  от всей совокупности случайно синтезированных белковых цепей. Эти опыты (через формулу (18.1)) позволяют оценить величину энергетической щели в 10–15 ккал/моль.

Примечательно, что из той же теории, что привела к формуле (18.1), следует, что сделать цепь, стабилизирующую одновременно две разные

нативные структуры (обе отделенные щелью ширины  $\Delta E$  от прочих укладок цепи) совсем трудно: вероятность этого порядка  $\exp(-2\Delta E/kT_c)$ . Иными словами, сделать цепь с одной нативной структурой — чудо, с двумя — чудо в квадрате, с тремя — в кубе... Не удивительно, что белков с двумя и более наблюдаемыми альтернативными структурами почти нет (здесь я говорю, естественно, только о существенно разных структурах, а не о тех небольших деформациях структур, что часто встречаются при функционировании белков; речь о них пойдет в конце этого курса лекций). «Почти нет», потому что несколько таких белков-«хамелеонов» — среди тысяч изученных — все же есть. К этим хамелеонам относится серпин, весьма медленно, но без всякой олигомеризации или агрегации переходящий в свою неактивную форму; белок Ltn40, при димеризации которого  $\alpha$ -спираль преобразуется в часть сильно перестроившегося  $\beta$ -листа; прион, «заразный белок», о котором я уже говорил в 11-й лекции (правда, очень медленное образование его альтернативной структуры связано со слипанием перешедших в «зараженную» форму прионовых цепей); и ряд других белков, причем некоторые из них были специально сконструированы людьми так, чтобы иметь альтернативные укладки, — о них я расскажу позже.

Температура  $T_c$  определяется наклоном кривой  $S(E)$  перед самой щелью (см. рис. 18-8); величина  $T_c$  не зависит от величины  $\Delta E$ , так как  $T_c$  зависит от поведения массы ненативных структур, а щель — только от отрыва от них одной нативной структуры. Иными словами, величина  $T_c$  одинакова и для случайных, и для «белковоподобных» цепей одинакового аминокислотного состава. И, что замечательно: судя по соотношению (18.1),  $T_c$ , характерная температура края непрерывного спектра, — это та самая конформационная температура, что фигурирует в «квази-Больцмановской» белковой статистике, о которой мы говорили на позапрошлой лекции.

Рисунок 18-8 показывает, что, при не слишком больших  $\Delta E$ , температура  $T_c$  лежит чуть ниже температуры плавления белка. А очень большие щели, большие  $\Delta E$ , по-видимому, не нужны для «белковоподобного» поведения молекулы, так что отбор на них настаивать не должен; и в то же время они маловероятны, согласно формуле (18.1). Впрочем, здесь слово за экспериментом...

- Энергетическая щель — фундаментальная вещь в физике белка. Она необходима не только для того, чтобы белок разрушался только фазовым переходом, т. е. чтобы он, не изменяясь, терпел (до известного предела) изменение внешних условий, т. е. для надежности работы белка. Скоро вы увидите, что она необходима и для быстрого и безошибочного сворачивания его нативной структуры. Поэтому экспериментальная оценка этой щели — она еще, увы, не сделана — будет столь интересна.

## Лекция 19

Образование структуры белка *in vivo* и *in vitro*. Вспомогательные механизмы при самоорганизации *in vivo*: котрансляционное сворачивание, шапероны, и т. д. Спонтанная самоорганизация возможна *in vitro*. «Парадокс Левинталя». Опыты по сворачиванию белка в бесклеточных системах, а также — о разном понимании слов *in vitro*. Стадийный механизм самоорганизации белков. Обнаружение метастабильных (накапливающихся) интермедиатов сворачивания многих белков. Расплавленная глобула — обычно (но не обязательно) наблюдаемый интермедиат сворачивания белка в нативных условиях. Простейшее (одностадийное) сворачивание некоторых белков — без каких-либо накапливающихся метастабильных интермедиатов. Самоорганизация мембранных белков.

Сегодня мы посмотрим, как протекает процесс обретения белковой цепью своей нативной структуры во времени, и обсудим интермедиаты, возникающие при самоорганизации белков. Я буду говорить только о водорастворимых глобулярных белках. Самоорганизация мембранных белков изучена еще сравнительно мало, и о них я скажу буквально пару слов. Самоорганизация фибриллярных белков изучена лучше, но скорее не на физическом, а на биохимическом уровне (она уже обсуждалась на примере образования коллагена).

- В живой клетке белок синтезируется на рибосоме. Биохимический синтез белковой цепи занимает порядка минуты, да и все изготовление «готового», свернутого белка, занимает примерно столько же — эксперимент не видит разницы (для справки: весь жизненный цикл бактерии может длиться всего пару десятков минут). Поэтому естественно предположить, что сворачивание белка может начинаться еще на рибосоме, еще до окончания ее полного синтеза всей белковой цепи.

К сожалению, достоверные экспериментальные данные об образовании белковой структуры *in vivo* — в клетке — весьма скудны: очень



трудно увидеть структурные превращения растущего белка на фоне всего клеточного супа. Обычно приходится останавливать синтез, выделять «недоделанный» белок, исследовать его отдельно — на все уходит время, десятки минут, в течение которых пространственная структура исследуемой белковой цепи может кардинально измениться...

Ряд такого рода опытов с большими белками показывает, что их первые, N-концевые домены успевают свернуться (или, точнее, учитывая длительность самого опыта, обладают способностью свернуться) еще до окончания полного синтеза цепи.

Здесь, однако, надо подчеркнуть, что все эти данные относятся к *много*-доменным белкам; это существенная оговорка, так как «единицей самоорганизации», по-видимому, является не белок в целом, а его отдельный домен. То есть «полусвернутый» домен не наблюдается, и мы не можем сказать, что сначала сворачивается его N-концевая половина, потом — C-концевая.

О том, что домен является единицей сворачивания, свидетельствуют два ряда данных, полученных, правда, не *in vivo*, а *in vitro*: во-первых, отдельно взятые домены часто способны к правильной самоорганизации; во-вторых, однодоменный белок, лишенный примерно десятка аминокислот на своем C-конце, к самоорганизации не способен. С другой стороны, А. А. Комаром с соавторами было показано, что цепь глобина, синтезируемая в бесклеточной системе, способна связать гем, когда рибосомой изготовлено только чуть больше половины этой цепи. [Глобин — однодоменный белок, судя и по его структуре, и по характеру его плавления *in vitro* (т. е. его денатурация протекает по типу «все-или-ничего»). Однако его N-концевая половина образует компактный субдомен, в котором находятся связывающие гем участки цепи.]

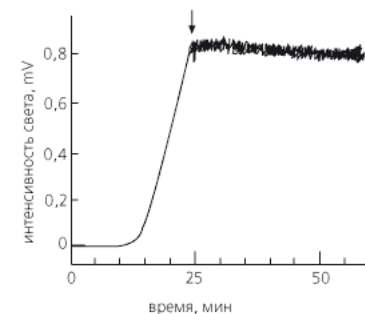
Ряд интересных данных по сворачиванию белков получен в *бесклеточной* белок-синтезирующей системе (т. е. не совсем *in vivo*, но и не совсем *in vitro*). Такая система состоит из рибосом, транспортных и информационных РНК и других факторов, необходимых для матричного синтеза белковой цепи.

Здесь уместно отметить, что, говоря об эксперименте “*in vivo*” и “*in vitro*”, физики и биологи часто имеют в виду разные вещи, так как между чистым “*in vivo*” (в организме) и чистым “*in vitro*” («в стекле») есть много допускающих двоякое толкование ступеней. Например, сворачивание белка в бесклеточной системе (со всеми ее рибосомами, факторами инициации и терминации, шаперонами и т. д.) — это столь же явно эксперимент “*in vivo*” с точки зрения физика (для которого *in vitro* — это отдельный белок в растворе; а бесклеточная система... — уж больно много здесь жизненных реалий!), сколь явно он является экспериментом “*in vitro*” с точки зрения биолога (для которого “*in vivo*” — это в живом и желательном неповрежденном организме). Однако структурное, например, исследование отдельного белка в организме практически невозможно. Так что на практике всегда идут на разумный компромисс — на приближение биологически интересного явления *in vivo* доступным экспериментальному наблюдению явлением *in vitro*.

К сожалению, и в бесклеточной системе очень трудно увидеть строение растущего белка на фоне громадной рибосомы. *Строение*, действительно, увидеть трудно. Но, если повезет, можно увидеть *активность* вновь синтезированного белка.

Счастливым находкой здесь оказалась люцифераза: этот белок, приняв свою нативную конформацию, катализирует реакцию с испусканием света. Так что появление нативной люциферазы легко увидеть — ведь больше в клетке ничего не светится. Исследуя биосинтез и сворачивание этого белка, В. А. Колб, Е. В. Макеев и А. С. Спиринов в нашем Институте белков РАН показали, что первый активный белок появляется через 10 мин после включения его биосинтеза, а выключение, блокирование биосинтеза немедленно прекращает выход активного белка (рис. 19-1). Это значит, что *уже* синтезированных, но *еще не* свернувшихся цепей люциферазы практически нет — т. е. мы видим, что сворачивание этого большого, содержащего свыше 500 остатков в цепи белка идет либо котрансляционно, либо почти мгновенно после конца биосинтеза.

Правда, недавно сделанные эксперименты (Eichmann et al., *Proc. Natl. Acad. Sci. USA* (2010) **107**:9111–9116) показывают, что *не до конца* синтезированная рибосомой белковая цепь не образует определенной пространственной структуры ни висая на рибосоме, ни будучи освобожденной от нее (что, в принципе, не противоречит образованию этой цепью расплавленной глобулы). Таким образом, не до конца синтезированная рибосомой белковая цепь *in vivo* ведет себя так же, как укороченная на несколько C-концевых аминокислот цепь *in vitro*: она не образует определенной трехмерной структуры.



**Рис. 19-1.** Свечение нативной люциферазы, синтезируемой в бесклеточной системе. Время «0» — включение биосинтеза люциферазы. Время выключения биосинтеза указано стрелкой. Сразу после этого свечение прекращает свой рост, т. е. выход нового активного белка больше не наблюдается. Картинка, с небольшими изменениями, взята из V. A. Kolb, E. V. Makeev & A. S. Spirin, *EMBO J.* (1994) **13**:3631–3637

- Однако для выяснения не только результата, но и хода процесса самоорганизации приходится исследовать самоорганизацию белка «совсем *in vitro*» (без всяких рибосом, шаперонов и т. д.), т. е. при денатурации белка в растворе.

Около 1960 г. было сделано замечательное открытие: глобулярный белок способен к спонтанной самоорганизации *in vitro* (ренатурации), если после биосинтеза он не подвергся сильной химической модификации. В этом случае его архитектура, «мягко» (без разрыва цепи) разрушенная температурой, растворителем и т. д., спонтанно восстанавливается при «нормализации» среды. Правда, эффективная ренатурация нуждается в тщательном подборе экспериментальных условий, иначе ей может воспрепятствовать агрегация белка.

**Внутренний голос:** А не сохраняется ли в денатурированном белке каких-то унаследованных от нативного состояния пространственных структур, которые могли бы направить его сворачивание на правильный путь?

**Лектор:** Если белок денатурирован «как следует», то нет. Конечно, если вы не оставите в денатурированном белке его интактные S–S-связи, а разорвете их; и если вы выждете несколько минут, достаточных для исчезновения памяти о нативной форме *cis*- и *trans*-пролинов. И если вы поместите белок в достаточно сильный денатурант, где он превратится в клубок (где его вторичная структура разрушится, а объем увеличится на порядок, т. е. все дальние по цепи контакты в цепи тоже пропадут).

**Внутренний голос:** Однако ведь есть вполне надежные сведения о существовании заметного числа «нативных» контактов между не-соседними по цепи остатками даже во вполне развернутом белке!

**Лектор:** Есть такие данные. Правда (если оставить в стороне белки с S–S-связями), они касаются контактов между довольно близкими остатками, разделенными в цепи всего несколькими другими. И основной вопрос (в свете того, что мы сейчас обсуждаем) заключается вот в чем: эти контакты — они что, унаследованы от нативной структуры белка и заморожены, или же они просто время от времени спонтанно образуются в развернутой цепи? К сожалению, ЯМР (а именно он, в основном, и видит обсуждаемые контакты) не может различить «замороженные» контакты от тех, что образуются время от времени. Дело в том, что сигнал ЯМР спадает как  $R^{-6}$  (где  $R$  — контактное расстояние), так что один и тот же сигнал соответствует контакту с  $R = 4,5 \text{ \AA}$ , существующему 100 % времени, и контакту с  $R = 3,0 \text{ \AA}$ , существующему 10 % времени. Я лично полагаю, что наблюдаемые контакты не заморожены (если это, конечно, не ковалентные S–S-связи; но о них мы сейчас и не говорим). Я не вижу никакой физической причины для этого. Я полагаю, что обсуждаемые контакты просто часто образуются между притягивающимися кусочками цепи (напомню, что именно такие, притягивающиеся кусочки в основном и контактируют в нативном белке). То же относится, по-видимому, и к «остаточной» вторичной структуре развернутых цепей, которую порой видит КД.

■ Явление спонтанной самоорганизации белков было открыто в группе Анфинсена в 1961 г. Открыто на примере спонтанного восстановления биохимической активности и «правильных» S–S-связей в бычьей рибонуклеазе А — после ее полного разворачивания с разрывом всех S–S-связей и последующего помещения в «нативные» условия. Это открытие, впоследствии подтвержденное на множестве других белков, а также возможность чисто химического синтеза белковой цепи, спонтанно сворачивающейся далее в активный белок (опыты Меррифилда и др.), позволяет в первом приближении отделить изучение структурообразования белка от изучения биосинтеза белковой цепи.

■ Самоорганизация белка *in vitro* может рассматриваться как самый простой (и поэтому для меня, физика, наиболее интересный) случай «чистой» самоорганизации, когда белку не помогает никто.

Вообще, самоорганизация пространственных структур белков (а также РНК) — уникальное физическое явление, не имеющее аналогов в «неживой» природе. Эта самоорганизация напоминает образование кристаллов — но кристаллов, во-первых, не имеющих периодической пространственной структуры; во-вторых, чрезвычайно сложно устроенных; и, наконец, очень маленьких.

Самоорганизация белковых структур относится, с физической точки зрения, к классу явлений «возникновение порядка из беспорядка» (по классификации Пригожина): трехмерный «аперриодический кристалл» (говоря словами Шредингера) структуры белка порождается заранее фиксированным порядком звеньев в его цепи. Видимо, следует сразу отметить, что самоорганизация трехмерной структуры белков (и РНК, и кристаллов вообще) возникает из стремления молекул к термодинамическому равновесию, и тем принципиально отличается от той самоорганизации типа «порядок из беспорядка» (будь то самоорганизация в осциллирующей химической реакции Белоусова—Жаботинского или самоорганизация в экологической системе «хищник—жертва»), которую обычно имеют в виду, говоря о самоорганизации в неравновесных, работающих на протоке энергии системах.

■ Однако прежде, чем перейти к рассказу о самоорганизации белка *in vitro*, к рассказу о физике этого спонтанного процесса, я хочу кратко упомянуть ту машинерию, которая используется клеткой для повышения эффективности самоорганизации белка.

Рибосома выдает белковую цепь постепенно и не вполне равномерно — есть паузы, приостановки биосинтеза цепи на «редких» (отвечающих «редким» в клетке тРНК) кодонах. Предполагается, что соответствие пауз границам структурных доменов способствует их спокойному созреванию.

Далее. В клетке белковая цепь сворачивается под опекой специальных белков — шаперонов. Шапероны препятствуют агрегации белков. Многие из них продуцируются клеткой в ответ на тепловой шок, так как при повышении температуры возрастают вызывающие агрегацию гидрофобные силы.

Агрегируя в клетке, белки часто образуют «тела включения». Белки не имеют там нативной структуры (в частности, там их S–S-связи завязаны беспорядочно). Однако, растворяя тела включения, из них можно добыть нативный белок — путем ренатурации *in vitro*.

С неприятностями, причиняемыми агрегацией (столь естественной в концентрированном «клеточном супе»), в клетке борются шапероны. «Малые» шапероны типа hsp70 (heat shock protein, 70 килодальтон) связываются с белком, предохраняя его от агрегации, а потом сбрасываются (на что расходуется АТФ). «Большие» шапероны типа GroEL/GroES или TrjC работают, в основном, с многодоменными белками и, особенно — с белками, чьи домены составлены из отдаленных кусков цепи. Эти шапероны образуют как бы пробирку, сделанную из белка GroEL (диаметром в несколько нанометров — и с крышечкой, сделанной из белка GroES, то прилипающей к GroEL, то отлипающей от него). Внутри этой пробирки (или, скорее, на самый вход в нее, как показано работами группы Г.В. Семисогнова) и садится – *in vitro* – развернутая (или находящаяся в состоянии расплавленной глобулы) белковая цепь. *In vivo* так же, по-видимому, поступает новосинтезированный (и до встречи с GroEL облепленный шаперонами типа hsp70 и/или hsp40) белок или его домен. Согласно распространенной в настоящее время точке зрения, эта «пробирка» (иногда ее называют «ячейкой Анфинсена») защищает новорожденный белок и от агрегации, и от действия концентрированного клеточного «супа» со всеми его протеазами и т. д. При этом, как полагают, эта пробирка время от времени активно «трясется» (при этом в GroEL происходят конформационные превращения, увеличивающие и уменьшающие его гидрофобную поверхность), т. е. она открывается (на что расходуется АТФ) и закрывается, но отпускает она белок только тогда, когда он уже свернется и перестанет липнуть к «пробирке». Впрочем, работами группы Г.В. Семисогнова показано, что, скорее, GroEL просто убирает из раствора «избыточные» неструктурированные белки, что позволяет многочисленным, оставшимся в растворе белкам избежать агрегации и правильно свернуться без всякого прямого взаимодействия с шапероном GroEL. Таким образом, шапероны, так или иначе, служат «инкубаторами» для сворачивающегося белка; однако, как было недавно показано в группе А.С. Спирина, они работают и как «холодильники», задерживая сворачивание белка до тех пор, пока он не сможет включиться в нужную четвертичную структуру или не будет перенесен в нужное место клетки.

Кроме того, самоорганизация белков может ускоряться некоторыми ферментами типа пролил-изомеразы (катализирующей медленно само по себе идущее превращение *trans*-пролинов в *cis*- и обратно; иногда именно оно лимитирует скорость сворачивания) или дисульфид-изомеразы (катализирующей сшивание и расщивание S–S-связей).

■ В живой клетке сворачивание белка происходит в насыщенном макромолекулярном окружении.

При этом, согласно хорошо обоснованной гипотезе К. Добсона, ключевым элементом процесса обычно является расплавленная глобула: именно ее образует белковая цепь при биосинтезе, именно ее оберегают от агрегации шапероны (а если не уберегут – именно из нее образуются амилоиды и другие агрегаты, часто болезнетворные), и именно из нее образуется нативный белок.

За время эволюции возник ряд механизмов, позволяющих уйти от проблем, создаваемых насыщенным окружением белка в клетке. Эти механизмы включают как действие разных семейств молекулярных шаперонов, так и котрансляционное сворачивание, важность которого все возрастает по сравнению с посттрансляционным, особенно в эукариотических клетках с их преимущественно многодоменными белками. Однако при этом, по-видимому, не существует оснований предполагать, что что-либо, кроме самой аминокислотной последовательности, обеспечивает верную конформацию белка (во всяком случае, одно- или двухдоменного) в клеточном окружении — хотя некоторые исследователи все еще предполагают противоположное.

Одной из замечательных идей, следующей из работ недавнего времени, является гипотеза «эволюционного сдвига» от посттрансляционного к котрансляционному сворачиванию по мере того, как эукариотические клетки возникали из комбинаций прокариотических, и по мере того, как развивались более крупные и сложные белки (типичный белок эукариота состоит из четырех-пяти доменов, прокариот — из двух).

Котрансляционное сворачивание, т. е. сворачивание растущей белковой цепи за время, пока она еще прикреплена к рибосоме, вызывает особенно большой интерес из-за последних замечательных достижений в исследовании структуры рибосомы в атомном разрешении. Сворачивание растущего пептида может облегчаться «туннелем» рибосомы, где этот пептид, возможно, первоначально находится — защищенный от агрегации и деградации. Более того, существуют некоторые (правда, не вполне еще достаточные) данные, что сама рибосома и некоторые из компонент рибосомы, в особенности ее большая субъединица и, в частности, ее 23S РНК, могут функционировать как молекулярные шапероны, способные облегчить и ускорить сворачивание предварительно денатурированных белков. Такое действие рибосом дало бы дополнительное (наряду с простой затрудненностью агрегации для сидящих на рибосоме белков) объяснение тому факту, что сво-

рачивание больших денатурированных белков *in vitro* (без рибосом) часто происходит гораздо медленнее, чем сворачивание белков *in vivo*, которое происходит одновременно с их биосинтезом на рибосоме.

Однако опыты по сворачиванию белка *in vitro* показывают, что работа всей этой клеточной машинерии может быть заменена подбором соответствующих внешних условий (малой концентрации белка, нужного окислительно-восстановительного потенциала). Эта замена не меняет результат сворачивания: уж если белок не выпал в осадок, а свернулся *in vitro*, то он свернулся в ту же структуру, что и *in vivo*. Правда, часто это займет больше (а иногда и меньше!) времени, чем самоорганизация *in vivo*, но результат будет тот же.

Более того: известно, что небольшую белковую цепь можно чисто химически синтезировать *in vitro* (в пробирке) — и притом с C-, а не с N-конца — и она, с хорошим выходом, свернется в правильную пространственную структуру.

Все это доказывает, что вся необходимая для построения трехмерной структуры белка информация содержится в химической последовательности аминокислот в его цепи.

Похоже, что — помимо самого синтеза белковой цепи — биосинтетический аппарат клетки (рибосома + шапероны + ...) служит лишь чем-то вроде инкубатора для сворачивания пространственной структуры белка: этот «инкубатор» не определяет структуру белка, но он создает условия для ее созревания, так же, как обычный инкубатор обеспечивает выведение птенца, не предопределяя, кто выведется — цыпленок или утенок.

**Внутренний голос:** Все это, строго говоря, относится в основном к небольшим водорастворимым глобулярным белкам (и доменам больших белков). Сложнее с крупными белками, особенно — с белками высших организмов: далеко не все они ренатурируют спонтанно. Что касается спонтанной самоорганизации мембранных и фибриллярных белков, в некоторых белках она происходит, но обычно полностью ренатурировать их не удается...

**Лектор:** Насчет «трудных» для ренатурации белков: возможно, здесь дело в агрегации (так, некоторые мембранные белки ренатурируют в детергентах, хотя в воде они не ренатурируют); возможно — в посттрансляционной модификации, особенно у эукариот. Договоримся, что пока я буду иметь в виду относительно небольшие глобулярные белки. Пойдем сначала самоорганизацию хотя бы их...

- Загадочность явления спонтанной самоорганизации белков (и РНК) суммируется «парадоксом Левинталь». Загадка состоит вот в чем. У белковой цепи есть бездна возможных конформаций (каждый аминокислотный остаток

имеет около 10 возможных конформаций, то есть цепь из 100 остатков — порядка  $10^{100}$  возможных конформаций). Так что белок должен искать «свою» пространственную структуру среди порядка  $10^{100}$  возможных. При этом белок может «почувствовать» стабильность конформации, только попав прямо в нее, так как отклонение даже на 1 Å может очень сильно повысить энергию цепи в плотной белковой глобуле. И так как переход из одной конформации в другую занимает  $\sim 10^{-13}$  секунды как минимум, перебор всех  $10^{100}$  структур должен был бы занять порядка  $10^{80}$  лет, на фоне которых время жизни нашей Вселенной —  $10^{10}$  лет — величина бесконечно малая... Вопрос: как белок может «найти» свою структуру за минуты?

Парадокс же заключается в следующем. С одной стороны, нативная пространственная структура по *всем* тестам ведет себя как самая стабильная из всех структур цепи: белковая цепь попадает в нее при разных кинетических процессах [и при сворачивании на рибосоме в процессе биосинтеза, и после транслокации сквозь мембрану, и при сворачивании в пробирке (ренатурации) — чем бы и как бы она ни была в этой пробирке развернута (или просто синтезирована химически)]. С другой стороны, не может быть никаких гарантий, что эта структура — самая стабильная из всех возможных: как показывает прикидка Левинталь, у белковой цепи просто нет времени на то, чтобы убедиться в этом!

«Как же белок выбирает свою нативную структуру среди бесчисленного множества возможных? — спросил Левинталь, и ответил: — По-видимому, самоорганизующийся белок следует по какому-то специальному “пути сворачивания”, и та структура, где этот путь заканчивается, и является его нативной структурой». Иными словами, Левинталь предположил, что нативная структура белка определяется не стабильностью, не термодинамикой, а кинетикой, т. е. она соответствует не глобальному, а просто быстро достижимому минимуму свободной энергии цепи.

Вопрос о том, что именно — кинетика или термодинамика — определяет укладку белковой цепи, отнюдь не чисто умозрительный. Он постоянно возникает на пути решения конкретных задач физики белка, идет ли речь о предсказании структуры белка по его аминокислотной последовательности (надо знать, что предсказывать: самую стабильную или самую быстро сворачивающуюся его структуру), или о дизайне новых, не встречающихся в природе белков (надо знать, что делать: максимально усиливать стабильность желаемой структуры или пролагать максимально быстрый путь к ней).

Обсуждение механизма сворачивания белков началось сразу же после расшифровки первых трехмерных структур и открытия явления самоорганизации. Первой, видимо, была гипотеза Филлипса, согласно которой на N-конце растущей цепи (с которого начинается биосинтез белка) воз-



никает зародыш структуры, и остальная цепь наматывается на него. В той или иной форме такая точка зрения присутствует в ряде работ и по сей день. Однако она не подтвердилась. В изящных работах Гольденберга и Крейтона было показано, что N-конец цепи не играет решающей роли в самоорганизации *in vitro* — по крайней мере для однодоменного белка. Опыт показал, что замкнутая в кольцо цепь небольшого белка, ингибитора трипсина, сохраняет способность к самоорганизации; и даже если разрезать это кольцо так, что новым N-концом окажется бывшая середина цепи — самоорганизация ведет к прежней пространственной структуре. В наши дни изготовление циркулярно пермутированных белков стало обыденным.

■ В поисках пути к решению проблемы самоорганизации белков, О. Б. Птицын в 1973 г. предложил концепцию стадийного сворачивания белка (рис. 19-2). В рамках этой гипотезы, получившей позже название «framework model» («каркасной модели»), и проходило, в значительной мере, дальнейшее обсуждение проблемы самоорганизации белка в мировой научной литературе.



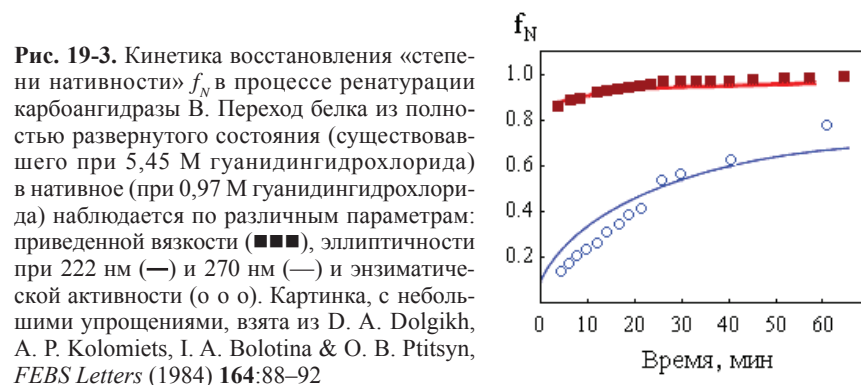
**Рис. 19-2.** Стадийная модель сворачивания белка по Птицыну (1973). Выделены вторичные структуры —  $\alpha$ -спирали (цилиндры) и  $\beta$ -участки (стрелки). Оба предсказанные интермедиата впоследствии найдены на опыте и названы «предрасплавленной» и «расплавленной» глобулами

Эта модель постулировала последовательное вовлечение различных взаимодействий в формирование структуры белка и подчеркивала важность образования зародышевых  $\alpha$ -спиралей и  $\beta$ -шпилек на ранних стадиях сворачивания, слипания этих зародышей в глобулу, грубо напоминающую нативную, и окончательной «доводки» строения глобулы на последнем этапе самоорганизации.

Краеугольным камнем этой концепции было то, тогда чисто гипотетическое, а теперь вошедшее в учебники состояние белковой цепи, которое ныне известно как «расплавленная глобула».

Выведенная буквально «на кончике пера» расплавленная белковая глобула была затем экспериментально обнаружена и детально исследована Долгих, Семисотновым, Гильманшиным, Бычковой и др. в лаборатории Птицына в 1980-х гг. — первоначально как равновесное состояние «слабоденатурированного» белка (об этом я говорил на прошлой лекции), а потом и как кинетический интермедиат сворачивания белка *in vitro*.

Так, расплавленная белковая глобула накапливается в процессе ренатурации карбоангидразы В (рис. 19-3). В ходе такого опыта исходное клубковое состояние денатурированного белка получается инкубированием его в растворе с высокой концентрацией денатуранта, а ренатурация получается быстрым разведением этого раствора большим количеством воды, т. е. наблюдаемый процесс ренатурации стартует с клубка и заканчивается восстановлением нативного белка. Однако разные свойства ренатурирующего белка меняются с разной скоростью, что свидетельствует о накоплении — в начале процесса ренатурации — какого-то «промежуточного» состояния белковой молекулы.



**Рис. 19-3.** Кинетика восстановления «степени нативности»  $f_N$  в процессе ренатурации карбоангидразы В. Переход белка из полностью развернутого состояния (существовавшего при 5,45 М гуанидингидрохлорида) в нативное (при 0,97 М гуанидингидрохлорида) наблюдается по различным параметрам: приведенной вязкости (■ ■ ■), эллиптичности при 222 нм (—) и 270 нм (—) и энзиматической активности (○ ○ ○). Картинка, с небольшими упрощениями, взята из D. A. Dolgikh, A. P. Kolomiets, I. A. Bolotina & O. B. Ptitsyn, *FEBS Letters* (1984) **164**:88–92

В данном случае видно, что «нативные» значения приведенной вязкости (характеризующей объем глобулы) и эллиптичности при 222 нм (характеризующей наличие вторичной структуры) восстанавливаются (точнее, почти восстанавливаются) в процессе ренатурации значительно быстрее, чем эллиптичность при 270 нм (характеризующая упаковку боковых групп в третичной структуре белка) или активность белка (являющаяся мерой нативности его третичной структуры). Это говорит о существовании накапливающегося в значительных количествах (т. е. метастабильного — «как бы стабильного»: ведь он существует в течение заметного времени) кинетического интермедиата. Причем этот интермедиат является расплавленной глобулой: его компактность и вторичная

структура близки к таковым для нативного белка, однако он не имеет ни нативной упаковки боковых групп (судя по эллиптичности в «ближнем» ультрафиолете, при 270 нм), ни энзиматической активности нативного («твердого») белка.

Расплавленная глобула оказывается — подчеркну, в примерно физиологических условиях — ранним промежуточным состоянием на стартующем из клубка пути самоорганизации многих глобулярных белков *in vitro*. Для образования такого интермедиата достаточно миллисекунд, полное же восстановление свойств нативного белка из 100–300 аминокислотных остатков требует от секунд (для одних белков) до часов (для других).

Необходимо подчеркнуть, что самый медленный (rate-limiting) шаг самоорганизации приходится не на раннюю стадию, не на образование расплавленной глобулы, а на образование из нее плотно упакованной, нативной глобулы (см. рис. 19-3).

- Расплавленная глобула — не единственный интермедиат, наблюдаемый при сворачивании белка. Состояние с частично сформированной вторичной структурой и с частичной компактизацией цепи (т. е. напоминающее предрасплавленную глобулу) было обнаружено, когда стало возможным исследовать явления, происходящие в течение первых миллисекунд сворачивания, т. е. еще до образования расплавленной глобулы. В белках с дисульфидными связями были зафиксированы различные интермедиаты, позволяющие проследить последовательность формирования S-S-связей, и т. д.

Надо сказать, что гипотеза о «кинетическом контроле» сворачивания белка чрезвычайно способствовала изучению интермедиатов сворачивания. Главная идея этого направления работ заключалась в том, что, выделив интермедиаты, мы сможем проследить путь сворачивания белка. В химии и биохимии это обычно позволяет понять, как протекает сложная реакция. Таким образом, исследования шли в русле, как теперь говорят, «химической логики», императив которой — «ищи и выделяй промежуточные состояния!».

Но при изучении самоорганизации белков эта логика сработала только частично: интермедиаты (расплавленные глобулы) были определены для очень многих белков, но ключевой вопрос — как может белковая цепь найти свою нативную структуру за очень малое время? — остался без ответа.

- Прогресс в понимании был достигнут при изучении маленьких (из 50–100 остатков) белков. Оказалось, что некоторые из них сворачиваются *in vitro* без наблюдаемых, без «накапливающихся» в эксперименте интермедиатов. Причем это отсутствие наблюдаемых интермедиатов относится не только к специально созданным условиям, близким к точке плавления белка (где одностадийной кинетики сворачивания можно ожидать *a priori*, исходя из того, что *термодинамика* де- и ренатурации соответ-

ствует, как известно, переходу типа «все-или-ничего», т. е. отсутствию сколько-нибудь заметного количества «полусвернутых» форм белка); оно относится — для этих небольших белков — также и к «физиологическим» условиям, при которых сворачивание большинства белков проходит через интермедиаты типа расплавленной глобулы.

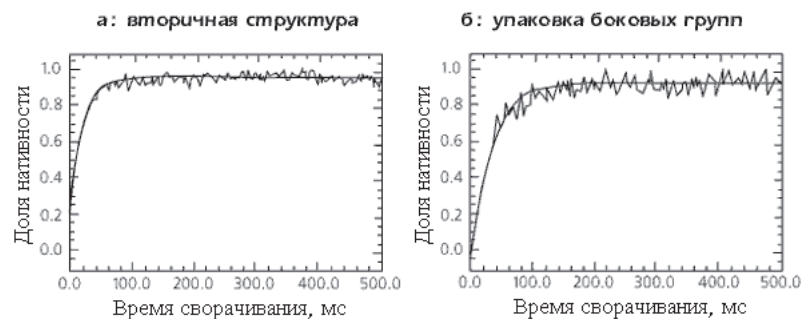
При изучении малых белков, не имеющих ни S-S-связей, ни *cis*-пролинов (т. е. именно тех особенностей, которые ранее широко эксплуатировались, чтобы поймать интермедиаты сворачивания) оказалось, что такие, лишенные «ненужных усложнений» (S-S-связей и *cis*-пролинов) белки порой сворачиваются очень быстро, особенно в чистой воде (или в физиологическом растворе) — не только много быстрее больших белков (что не удивительно), но и не медленнее тех малых, наличие интермедиатов в сворачивании которых должно было, казалось бы, ускорить и облегчить их правильное сворачивание...

Обычно белки «однодоменного» размера сворачиваются в физиологическом растворе (с наблюдаемыми, накапливающимися интермедиатами) за секунды или десятки секунд, но некоторые из малых белков, сворачиваясь без накапливающихся интермедиатов, достигали нативной структуры за миллисекунды (рис. 19-4) или даже за доли миллисекунд (правда, это были самые маленькие белки — до 100 остатков длиной). Отсутствие интермедиатов при сворачивании этих белков регистрировалось по одинаковости скорости восстановления вторичной структуры, упаковки боковых групп, флуоресценции триптофанов (зависящей от погруженности триптофанов в глобулу), по обмениваемости водородов в белке и т. д.

- Однако — что в таких экспериментах можно исследовать, чтобы пролить свет на природу процесса самоорганизации? Ведь интермедиатов, которые можно выделить и изучать, тут нет?!

Ответ: здесь можно изучать *переходное состояние* — это «самое узкое место» на пути сворачивания белка.

Я думаю, надо с самого начала сказать, что переходное состояние, играющее ключевую роль в кинетике процесса сворачивания и определяющее его скорость, является, по определению, абсолютно нестабильным, *самым нестабильным* состоянием на пути сворачивания. Таким образом, это — не какой-то из «интермедиатов сворачивания», о которых шла речь выше (и, в частности, не расплавленная глобула!). Те соответствуют минимумам свободной энергии; поэтому они стабильны хотя бы в течение какого-то времени, т. е. они могут накапливаться в процессе сворачивания, и их можно наблюдать непосредственно. Наоборот, переходное состояние соответствует максимуму свободной энергии; поэтому оно не накапливается и не наблюдается непосредственно. Его можно наблюдать только по его влиянию на скорость сворачивания.



**Рис. 19-4.** Одностадийная ренатурация АСВР (белка, связывающего ацил-коэнзим А), регистрируемая по восстановлению эллиптичности при 225 нм (а) и 286 нм (б); первое отражает восстановление вторичной структуры, второе — третичной (точнее, восстановление упаковки боковых групп). Тонкая, ломаная линия — экспериментальный сигнал, жирная кривая — результат его сглаживания и экстраполяции, с учетом погрешности опыта. Обратите внимание на отсутствие изменений, происходящих за «мертвое время» опыта; оно свидетельствует, что за это время никаких интермедиатов не образуется. Картинки, с разрешения, взяты из Kragelund B. B., Robinson C. V., Knudsen J., Dobson C. M. & Poulsen F. M. *Biochemistry* (1995) **34**: 7217–7224. © 1995, ACS

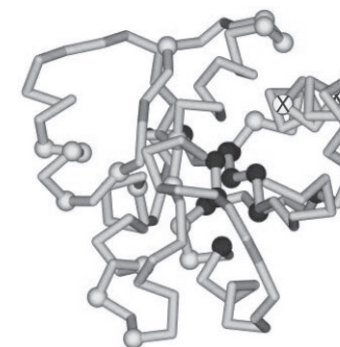
- Отложим пока подробное описание исследования переходных состояний, наблюдаемых при сворачивании белков, и, забежав вперед, суммируем только его основной результат.

В переходном состоянии очерчивается зародыш сворачивания белка, «ядро сворачивания» (рис. 19-5). В этом зародыше относительно небольшое число остатков цепи уже имеет свое нативное (т. е. такое же, как в нативном белке) расположение в пространстве и конформацию. В то же время остальные остатки остаются в денатурированной фазе.

Более подробно мы рассмотрим все это на следующей лекции.

- До сих пор мы говорили только о водорастворимых глобулярных белках.

Действительно, огромное большинство исследований по сворачиванию белка были проведены на таких белках. Но примерно треть белков в геноме каждого организма являются мембранными белками, которые могут достичь своего функционального состояния только в присутствии липидов. В настоящее время существуют данные в пользу того, что сворачивание спиральных мембранных белков проходит через промежуточные состояния, во многом похожие на расплавленную глобулу, наблюдаемую в водорастворимых глобулярных белках, т. е. имеющие нативоподобные вторичные структуры, но лишенные упорядоченности взаимодействий в простран-



**Рис. 19-5.** Экспериментально очерченное ядро сворачивания белка CheY, согласно López-Hernández & Serrano, *Folding & Design* (1996) **1**:43–55. Входящие в ядро сворачивания остатки изображены черными кружками на фоне нативной укладки цепи. Серыми кружками изображены не вовлеченные в ядро сворачивания остатки. Без шариков оставлены те области цепи, где мутации еще не делались. Крестами помечены два сложных для экспериментальной интерпретации остатка цепи

ственной структуре. Присутствие гидрофобных остатков, выступающих из плотных расплавленных глобул, делает их весьма привлекательными в контексте сворачивания в мембране. Ряд опытов заставляет полагать, что разные спирали в мембранных белках сворачиваются более или менее автономно, что совместимо с гибкостью, ожидаемой для расплавленной глобулы. Аналогия между сворачиванием мембранных и водорастворимых белков, возможно, распространяется на существование структурного «ядрышка», которое может служить их зародышем сворачивания. Другие мембранные белки, порины, имеющие структуру типа  $\beta$ -бочонка, сворачиваются, по-видимому, значительно более кооперативно, чем спиральные мембранные белки, т. е. в их сворачивании не просматривается четких стадий и стабильных интермедиатов.

- Прогресс в области исследования сворачивания белков, помимо удовлетворения нашего естественного любопытства, будет, безусловно, способствовать нашей способности эффективно экспрессировать как водорастворимые глобулярные, так и мембранные белки, и повысит наше понимание регуляции внутриклеточных, секреторных и мембранных процессов.

- В заключение этой лекции представляется нелишним провести некую аналогию между самоорганизацией белков и самоорганизацией вирусов. В оболочечных белках многих простых вирусов, по-видимому, имеются все свойства, необходимые для формирования готовых вирусных частиц.

Однако для более сложных вирусов известно, что пути их сборки включают в себя «леса» — белки или другие компоненты, необходимые для формирования вирусных частиц, но не включающиеся в зрелые вирионы. Эти «молекулярные шапероны» поистине замечательны как с точки зрения структуры, так и с точки зрения механизма сборки. Любопытно следующее. Так как простые вирусы собираются без накопления интермедиатов (точно так же, как простые белки сворачиваются без накопления частично свернутых состояний), то мы накопили гораздо больше информации не о фундаментальных путях сборки простых вирусов (и белков), а о более сложных путях сборки сложных вирусов (а равно и белков), так как здесь частично структурированные частицы могут наблюдаться при подходящих условиях.

## Лекция 20

Одностадийное сворачивание малых белков. Теория переходных состояний. Экспериментальный поиск и изучение нестабильных переходных состояний в сворачивании белка. Ядро сворачивания нативной структуры белка. Его экспериментальное обнаружение *in vitro* методами белковой инженерии. Нуклеационный механизм сворачивания белка.

Как уже говорилось на прошлой лекции, важнейшую роль в понимании сворачивания белков сыграло изучение простейшего случая самоорганизации белка — случая, когда по ходу сворачивания не происходит накопления никаких метастабильных промежуточных состояний.

Напомню, что некоторые из малых белков быстро сворачиваются — в «физиологических» условиях среды, моделируемых *in vitro* — без наблюдаемых, «накапливающихся» в эксперименте интермедиатов. При этом переход представляется происходящим в одну стадию, это регистрируется по одинаковой скорости восстановления вторичной структуры, упаковки боковых групп, флюоресценции триптофанов (зависящей от их погруженности в глобулу), по обменяемости водородов в белке и т. д.

Стоит подчеркнуть, что одностадийной кинетики сворачивания близ точки де/ренатурации белка можно *a priori* ожидать для всех однодоменных белков, исходя из того, что *термодинамика* их де- и ренатурации соответствует, как известно, переходу типа «*все-или-ничего*», т. е. отсутствию сколько-нибудь заметного количества «полусвернутых» форм. Не тривиально только то, что это одностадийное сворачивание относится — для некоторых небольших белков — также и к далеким от равновесия «физиологическим» условиям, где скорость сворачивания особенно велика и где сворачивание большинства белков проходит через накапливающиеся интермедиаты типа расплавленной глобулы.

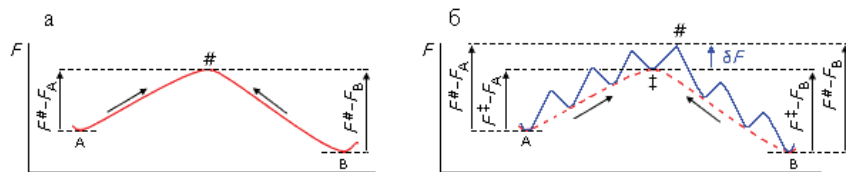
Именно в таком одностадийном сворачивании, происходящем в широком диапазоне внешних условий (и без осложняющих его анализ нака-



плывающих интермедиатов), наиболее удобно исследовать **переходное состояние** — состояние, играющее ключевую роль в кинетике процесса.

■ Вспомним, что переходное состояние соответствует *максимуму* свободной энергии на пути из одного стабильного состояния (*минимума* свободной энергии) к другому.

Напомним также, что кинетика простейшего перехода между двумя непосредственно наблюдаемыми (стабильными) состояниями (А и В на рис. 20-1) успешно описывается *теорией переходных состояний*, которая в химии обычно называется *теорией активированных комплексов*.



**Рис. 20-1.** Преодоление свободно-энергетического барьера при переходе из стабильного состояния А в стабильное состояние В (стрелка слева) и из В в А (стрелка справа).  $F_A$ ,  $F_B$  и  $F^\#$  — свободные энергии состояний А, В и «переходного» (барьерного, имеющего максимальную свободную энергию на пути процесса) состояния #. (а) Элементарная реакция. (б) Процесс, состоящий из нескольких элементарных переходов через устойчивые промежуточные состояния с высокой (выше, чем у А и В) свободной энергией.  $\delta F$  — подъем свободной энергии, который приходится преодолевать на одном шаге. ‡ — промежуточное состояние с максимальной, для устойчивого состояния, свободной энергией  $F^\ddagger$

Суть ее сводится к тому, что скорость процесса лимитируется *низкой населенностью переходного состояния #*, т. е. *наименее* (по термодинамике) *вероятного* (и потому *не* накапливающегося и непосредственно *не* наблюдаемого), *но необходимого промежуточного состояния* пути из А в В, — или, иными словами, что кинетика переходов А→В (и В→А) определяется *высотой* максимума свободной энергии на пути, соединяющем А и В.

При этом скорости перехода из А в В и из В в А, напомним, записываются в виде

$$k_{A \rightarrow B} = k_0 \exp [-(F^\# - F_A)/RT], \quad k_{B \rightarrow A} = k_0 \exp [-(F^\# - F_B)/RT]. \quad (20.1)$$

Здесь  $F_A$ ,  $F_B$  и  $F^\#$  — свободные энергии состояний А, В и «переходного» # (см. рис. 20-1), а  $k_0$  — скорость «элементарного шага» процесса.

Рассматривая элементарные реакции (рис. 20-1а), за  $k_0$  часто принимают величину  $k_{0,el} = RT/h$  (где  $h$  — постоянная Планка), равную (при  $T \sim 300$  К)

примерно  $10^{13} \text{ с}^{-1}$ . (С такой частотой, под воздействием тепловых колебаний, молекула предпринимает попытки преодолеть активационный барьер #.)

Однако нам часто будет удобнее (см. рис. 20-1б) рассматривать только устойчивые состояния (локальные минимумы свободной энергии на пути реакции), предполагая, что их свободная энергия много выше, чем и  $F_A$ , и  $F_B$ . Напомню, что скорость и такого многошагового процесса (если число шагов не слишком велико) также определяется просто самым высоким барьером на его пути.

Рассматривая многошаговой процесс, удобно

1) принимать скорость элементарного шага равной скорости перехода из одного минимума в другой, соседний:

$$k_{0,step} = k_{0,el} \exp (-\delta F/RT)$$

(где  $\delta F$  — подъем свободной энергии, преодолеваемый в ходе элементарного шага);

2) принимать за свободную энергию барьера величину  $F^\ddagger$  — максимальную свободную энергию устойчивого промежуточного состояния в этом процессе.

Тогда

$$k_{A \rightarrow B} = k_{0,el} \exp [-(F^\# - F_A)/RT] = k_{0,el} \exp [-(F^\ddagger + \delta F - F_A)/RT] = k_{0,step} \exp [-(F^\ddagger - F_A)/RT]. \quad (20.2)$$

Эта формула выглядит как (20.1) с заменой  $k_0 \rightarrow k_{0,step}$  и  $F^\# \rightarrow F^\ddagger$ .

При этом характерное время перехода А→В есть

$$t_{A \rightarrow B} = 1/k_{A \rightarrow B} = (1/k_{0,step}) \exp [(F^\ddagger - F_A)/RT]. \quad (20.2a)$$

Эти же пояснения распространяются, естественно, и на  $t_{B \rightarrow A}$  и  $k_{B \rightarrow A}$ .

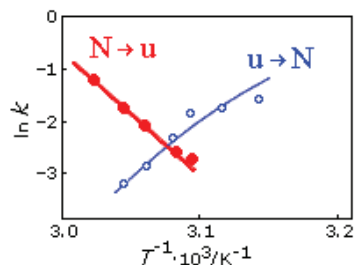
С этим уточнением уравнения (20.1) применимы к процессам, (а) протекающим за время, много большее, чем время элементарного шага, (б) состоящим из не слишком большого числа шагов и (в) не имеющим стабильных (накапливающихся) интермедиатов. И именно с такими процессами мы имеем дело при сворачивании белков.

■ Важно, что температурная зависимость скорости реакции А→В дает возможность судить об *энергии переходного состояния*. Для этого, согласно Аррениусу, вычисляют производную логарифма скорости реакции по обратной температуре ( $1/T$ ):

$$d [\ln (k_{A \rightarrow B})] / d (1/T) = d [\ln (k_{0,el}) - (F^\# - F_A) / RT] / (-T^{-2} dT) \approx \approx -(E^\# - E_A) / R. \quad (20.3)$$

Здесь я воспользовался уже известной нам формулой  $d(F/T)/dT = -E/T^2$  и пренебрег очень слабой зависимостью скорости элементарного шага  $k_{0,el}$  от температуры. Последнее допустимо даже в простых химических реакциях, и тем более допустимо в белках, где энергии  $E^\ddagger$  и  $E_A$  велики, так как определяются взаимодействиями множества частиц.

■ Рисунок 20-2 показывает изменение скорости сворачивания и разворачивания лизоцима при его тепловой ре- и денатурации.



**Рис. 20-2.** Аррениусовы графики для зависимости скорости тепловой денатурации и ренатурации лизоцима от обратной температуры  $T^{-1}$ ; графики извлечены из статьи S. Segava & S. Sugihara, *Biochemistry* (1984) **23**:2473–2488. Константы скорости  $k$  измеряются в  $\text{с}^{-1}$ . Ренатурация, скорость  $k_{u \rightarrow N}$  (эксперимент: точки o, и тонкая интерполяционная кривая); денатурация, скорость  $k_{N \rightarrow u}$  (эксперимент: точки \*, и жирная интерполяционная кривая). Середине плавления отвечает та температура, при которой  $k_{u \rightarrow N} = k_{N \rightarrow u}$ , т. е. где кривые пересекаются (около  $1000/3,08 = 325$  К). При меньших температурах (т. е. при больших  $T^{-1}$ , справа от точки пересечения) превалирует сворачивание, при больших температурах (т. е. при меньших  $T^{-1}$ , слева от точки пересечения) идет разворачивание

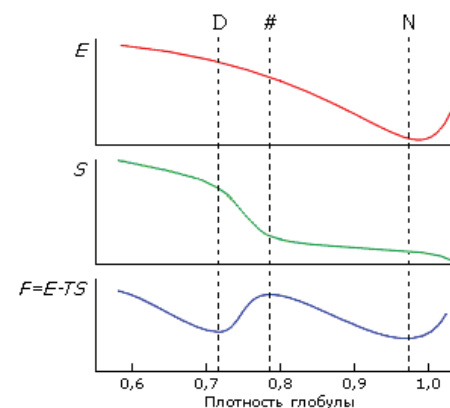
Он показывает, что в самой середине перехода [там, где скорость сворачивания  $k_{u \rightarrow N}$  равна скорости разворачивания  $k_{N \rightarrow u}$ , так что кривые для  $\ln k_{u \rightarrow N}$  и  $\ln k_{N \rightarrow u}$  пересекаются] скорости обоих процессов близки к  $e^{-2.5} \approx 0,1 \text{ с}^{-1}$  (т. е. что здесь времена протекания и де-, и ренатурации — порядка 10 с).

Кроме того, эта картинка показывает, что денатурация ускоряется по мере углубления в «область денатурации», а ренатурация ускоряется по мере углубления в «область ренатурации». И здесь содержится одна очень любопытная вещь.

Она состоит в том, что скорость тепловой денатурации  $k_{N \rightarrow u}$  *падает* с  $T^{-1}$  (т. е. что она *растет* с температурой  $T$ , что обычно для физико-химических реакций), а скорость сворачивания  $k_{u \rightarrow N}$  — *наоборот, растет* с  $T^{-1}$  (т. е. она *падает* с температурой, что *необычно* для физико-химических реакций). По формуле (20.3) это означает, что  $E^\ddagger - E_N > 0$ , а  $E^\ddagger - E_u < 0$ . Иначе говоря,  $E_u > E^\ddagger > E_N$ , т. е. энергия барьера  $E^\ddagger$  лежит выше энергии нативного состоя-

ния  $E_N$ , но *ниже* энергии развернутого состояния  $E_u$  (последнее *необычно* для химических реакций, где барьер по энергии выше, чем *и* начальное, *и* конечное состояния реакции). Дальнейший анализ этого графика [который теперь берется в форме зависимости  $T \ln(k/k_{0,el})$  от  $T$ ] показывает, что такое же соотношение,  $S_u > S^\ddagger > S_N$ , справедливо и для энтропий развернутого, переходного и нативного состояний при сворачивании белка. Это значит, что барьер между нативным и развернутым состояниями белка выглядит «обычным» энергетическим барьером — если глядеть на него со стороны нативного состояния, и что он выглядит *необычным*, энтропийным барьером — если глядеть на него со стороны развернутого состояния.

Впрочем, этого можно было ожидать, исходя из сделанного на одной из прошлых лекций анализа свободно-энергетического барьера и причин фазового, типа «*все-или-ничего*» перехода между нативным и денатурированным состоянием белка (рис. 20-3, который воспроизводит уже известную вам схему изменения энергии, энтропии и свободной энергии по мере расширения глобулы).



**Рис. 20-3.** Изменение энергии  $E$ , энтропии  $S$  и свободной энергии  $F = E - TS$  с изменением плотности глобулы. D — денатурированное состояние (в данном случае — расплавленная глобула, так как плотность его велика), N — нативная глобула, # — «барьер», максимум свободной энергии на пути равномерного расширения глобулы (от N к D).

■ Обратимся еще раз к уравнениям (20.1). Они показывают, что отношение скоростей прямой и обратной реакций,  $k_{A \rightarrow B}/k_{B \rightarrow A}$  есть просто константа равновесия между финальным B и начальным A состояниями

$$K_{B:A} = k_{A \rightarrow B}/k_{B \rightarrow A} = \exp [-(F_B - F_A)/RT]. \quad (20.4)$$

Величина этой константы равновесия, как известно, показывает соотношение между финальным числом  $n_B^\infty$  молекул в состоянии В и финальным числом  $n_A^\infty$  молекул в состоянии А, к которому, в конечном итоге (при стремящемся к бесконечности времени наблюдения), приходит процесс при данных условиях (температуре и т. д.), т. е. когда поток  $A \rightarrow B$  уравнивается потоком  $B \rightarrow A$ , так что  $n_A^\infty k_{A \rightarrow B} = n_B^\infty k_{B \rightarrow A}$ :

$$K_{B:A} = n_B^\infty / n_A^\infty = k_{A \rightarrow B} / k_{B \rightarrow A}. \quad (20.4a)$$

С какой скоростью система приходит к этому равновесному состоянию? Для ответа решим соответствующее дифференциальное уравнение:

$$dn_A/dt = -k_{A \rightarrow B} n_A + k_{B \rightarrow A} n_B, \quad (20.5)$$

где  $t$  — время, а  $n_A$  и  $n_B$  — число молекул в состояниях А и В, соответственно (второе уравнение, для  $dn_B/dt$ , решать не надо, так как  $n_A + n_B \equiv n_0$ , где  $n_0$  — полное число молекул, т. е.  $dn_B/dt \equiv -dn_A/dt$ ).

Подставляя в (20.5)  $n_B = n_0 - n_A$  и используя то, что  $n_0 k_{B \rightarrow A} / (k_{A \rightarrow B} + k_{B \rightarrow A}) = n_A^\infty$ , согласно (20.4a), получаем

$$dn_A/dt = -(k_{A \rightarrow B} + k_{B \rightarrow A}) (n_A - n_A^\infty).$$

Окончательный ответ — позвольте написать его без дальнейших выкладок, а вам рекомендуется сделать их самостоятельно — таков:

$$n_A(t) = [n_A(0) - n_A^\infty] \exp[-(k_{A \rightarrow B} + k_{B \rightarrow A}) t] + n_A^\infty. \quad (20.6)$$

Здесь  $n_A(t)$  — число молекул в состоянии А в момент времени  $t$  от начала процесса, а  $n_A^\infty = n_0 [k_{B \rightarrow A} / (k_{A \rightarrow B} + k_{B \rightarrow A})]$  — конечное, равновесное число молекул в состоянии А.

Значит, видимая (apparent) скорость приближения к равновесию есть

$$k_{app} = k_{A \rightarrow B} + k_{B \rightarrow A}, \quad (20.7)$$

она равна сумме скоростей прямой и обратной реакций.

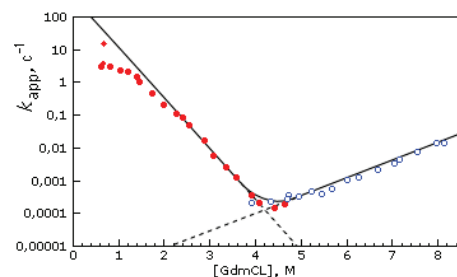
Обратите внимание, что эта скорость  $k_{app}$  зависит только от условий, в которых протекает реакция, и *не* зависит от исходного распределения белковых молекул между нативным и денатурированным состояниями [т. е. от исходной доли  $n_A(0)$ ].

В суммарной скорости  $k_{app}$  доминирует более быстрая реакция. Если условия (температура и т. д.) отвечают большей устойчивости нативного состояния, то в  $k_{app} = k_{u \rightarrow N} + k_{N \rightarrow u}$  доминирует  $k_{u \rightarrow N}$  — скорость сворачивания. Если более устойчиво денатурированное состояние, то в  $k_{u \rightarrow N} + k_{N \rightarrow u}$  доминирует  $k_{N \rightarrow u}$ , скорость разворачивания.

Большим преимуществом измерения скорости прихода к равновесию  $k_{app}$  является то, что она всегда может быть измерена (в отличие от  $k_{u \rightarrow N}$

и  $k_{N \rightarrow u}$  в отдельности), — как тогда, когда более стабильно нативное состояние, так и тогда, когда более стабильно денатурированное.

■ При измерении  $k_{app}$  строятся так называемые «шеvronные графики» (рис. 20-4). Они так называются из-за своей характерной V-образной формы, напоминающей шевроны на военных кителях. Один взгляд на этот график показывает, что скорости сворачивания  $k_{u \rightarrow N}$  и разворачивания  $k_{N \rightarrow u}$  зависят от концентрации денатуранта противоположным образом. С аналогичной ситуацией мы уже сталкивались, когда изучали температурную зависимость скоростей сворачивания и разворачивания белка. Как и тогда, наблюдаемая здесь противоположность наклонов означает «промежуточность» свойств переходного состояния, их среднее положение между свойствами нативного и развернутого белка.



**Рис. 20-4.** «Шевронный график» зависимости  $k_{app} = k_{u \rightarrow N} + k_{N \rightarrow u}$  видимой характерной скорости приближения к равновесию между нативной и развернутой формами белка (лизоцима куриного яйца), от концентрации гуанидинхлорида. Опыт проводился в присутствии нативных S-S-связей. Черные кружки получены при разбавлении крепкого раствора GdmCl, в котором изначально находился денатурированный белок,

т. е. при ренатурации белка. При этом  $k_{app} \approx k_{u \rightarrow N}$  (так как здесь скорость разворачивания  $k_{N \rightarrow u}$  значительно превосходит скорость сворачивания  $k_{u \rightarrow N}$ ). Светлые кружки получены при добавлении GdmCl к раствору изначально нативного белка, т. е. при его денатурации. При этом  $k_{app} \approx k_{N \rightarrow u}$  (так как при денатурации  $k_{u \rightarrow N} \ll k_{N \rightarrow u}$ ). В самой нижней части шеврона  $k_{u \rightarrow N} \approx k_{N \rightarrow u} \approx k_{app}/2$ . Обратите внимание на перекрытие черных и светлых кружков в этой области, в районе точки денатурации: оно показывает, что здесь, действительно, белок сворачивается с той же скоростью, что и разворачивается. Пунктир показывает экстраполяцию величин  $k_{u \rightarrow N}$  и  $k_{N \rightarrow u}$  в области излома шеврона и за нее, сплошная линия — в области низких концентраций GdmCl. Отклоняющиеся от экстраполяционной прямой точки в верхней левой части графика (т. е. *вдали* от точки излома, от точки денатурации белка) свидетельствуют либо о какой-то перестройке переходного состояния, либо о появлении каких-то дополнительных метастабильных интермедиатов (возможно, типа расплавленных глобул), которые могут лежать как на, так и вне основного пути сворачивания. Обратите внимание, что все эти перестройки и/или интермедиаты *не* повышают скорость сворачивания по сравнению с той, что можно было бы ожидать при неизменности переходного состояния: отклоняющиеся точки, хотя и соответствуют наивысшей наблюдаемой скорости сворачивания, лежат *ниже* интерполяционной прямой. Картинка, с небольшими изменениями (добавлены экстраполяционные пунктиры) взята из Т. Kiefhaber, *Proc. Natl. Acad. Sci. USA* (1995) 92:9029–9033

Действительно, то что денатурант вообще разворачивает белок, показывает, что он сильнее притягивается к развернутому белку, чем к нативному. То, что денатурант замедляет сворачивание исходно развернутого белка (см. спад скорости в левой части рис. 20-4) показывает, что он сильнее притягивается к исходному, развернутому белку, чем к переходному состоянию; а то, что денатурант ускоряет разворачивание исходно нативного белка (см. подъем скорости в правой части рис. 20-4), показывает, что он сильнее притягивается к переходному состоянию, чем к исходному, нативному. Это означает, что контакт переходного состояния с денатурантом больше, чем у нативного белка, но меньше, чем у денатурированного. Иначе говоря, приведенный на рис. 20-4 шевронный график показывает, что переходное состояние по степени контакта с растворителем, т. е. по своей компактности, находится где-то на полпути между развернутым и нативным состояниями белка. Этот график дает даже больше: так как наклон для  $k_{N \rightarrow u}$  несколько меньше, чем для  $k_{u \rightarrow N}$ , то компактность переходного состояния несколько ближе (в данном случае) к одной у нативного белка, чем у денатурированного.

Рисунок 20-4 (как и рис. 20-2) относится к лизоциму — белку, денатурация которого гуанидингидрохлоридом приводит прямо к образованию клубка, а не расплавленной глобулы. При умеренных концентрациях денатуранта (при не очень сильном его разбавлении водой) лизоцим сворачивается довольно медленно, за многие минуты. Однако при самоорганизации в почти чистой воде (происходящей при сильном разведении водой крепкого раствора гуанидингидрохлорида, где находились клубкообразные белковые цепи) скорость сворачивания лизоцима максимальна, и уже слабо зависит от остаточной концентрации денатуранта. При этом в сворачивании лизоцима появляется компактный метастабильный интермедиат типа расплавленной глобулы, интермедиат, который не наблюдается при повышенных концентрациях денатуранта.

- Таким образом, вдали от точки равновесия лизоцим демонстрирует двухстадийное сворачивание, а вблизи — одностадийное. То есть нет физической разницы между «двух-» и «одностадийными» белками, если рассматривать их только вблизи точки равновесия.

Однако изучать «одностадийные» белки проще: их переходное состояние в воде — то же, что и при всех других концентрациях денатуранта. Есть еще и техническое преимущество: длинные ветви шевронов проще экстраполировать. Поэтому именно изучение сворачивания «одностадийных» белков сыграло ведущую роль в исследовании переходных состояний.

- Появление компактных метастабильных интермедиатов сворачивания, вообще говоря, всегда «выполаживает» зависимость скорости сворачивания от содержания денатуранта — см. левый край рис. 20-4, или даже меняет знак ее наклона, так как шаг, лимитирующий скорость сворачивания, теперь начинается с более компактного состояния, чьи свойства (например, связанность с денатурантом) ближе к свойствам переходного состояния, чем к свойствам бывшего начального состояния, развернутого. Известно, что ренатурация белка, стартовая не от клубка, а от расплавленной глобулы (например, ренатурация карбоксиангидразы), также демонстрирует довольно слабую зависимость скорости ренатурации от концентрации денатуранта. Последнее означает, что переходное состояние в таком сворачивании мало отличается от исходной расплавленной глобулы по компактности.

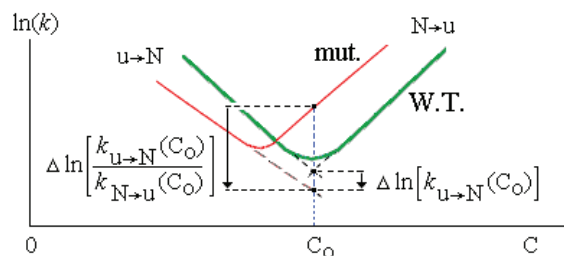
Иначе говоря, переходное состояние — по «промежуточности» некоторых своих свойств между свойствами нативного и развернутого белка — напоминает расплавленную глобулу. Напоминает по «средним» характеристикам, по энергии и компактности, но это еще не означает, что переходное (подчеркнем, нестабильное, в отличие от расплавленной глобулы) состояние действительно похоже на расплавленную глобулу.

Эксперимент (я расскажу о нем через минуту) показывает, что переходное состояние гораздо более неоднородно, чем расплавленная глобула. Его (если речь идет о переходе «клубок → нативный белок») можно представить себе как кусок нативного белка, остальная часть которого все еще находится в денатурированном, клубкообразном состоянии. Вид переходных состояний для переходов типа «расплавленная глобула → нативный белок» (и «клубок → расплавленная глобула») еще не установлен экспериментально, но весьма правдоподобно, что они также состоят из куска более структурированной фазы, в то время как остальная часть цепи остается в менее структурированной фазе.

- Природу переходного состояния удалось выяснить с помощью белковой инженерии. Применяя множество мутаций и анализируя соответствующие им изменения в шевронных графиках (см. рис. 20-4, 20-5), удалось ценой колоссального труда узнать, какие именно остатки вовлечены в «нативоподобную» часть переходного состояния (их мутации существенно влияют на скорость сворачивания), а какие — нет. Этот метод был разработан А. Ферштом в Англии.

Такой метод применим к любым переходам типа «все-или-ничего». Однако, строго говоря, пока что он применяется только к белкам, денатурация которых приводит прямо к образованию клубка, а не расплавленной глобулы.





**Рис. 20-5.** Схема, иллюстрирующая сдвиг «шеvronного графика» при мутации.  $k$  — наблюдаемая скорость установления равновесия;  $C$  — концентрация денатуранта. Жирная линия — исходный белок (W. T.), тонкая — тот же белок с мутацией в одном остатке цепи (mut.). Пунктир показывает экстраполяцию величин  $k_{u \rightarrow N}$  и  $k_{N \rightarrow u}$  в области излома шеврона.  $C_0$  — концентрация денатуранта, соответствующая середине денатурационного перехода в исходном (W. T.) белке (здесь разность свободных энергий нативной и денатурированной форм равна нулю — по определению «середины перехода»). На графике показаны измеряемые в этом опыте величины. Одна из них определяет влияние мутации на высоту свободно-энергетического барьера, стоящего на пути из развернутого состояния в нативное:  $\Delta(F^\# - F_u) = -RT \Delta \ln k_{u \rightarrow N}$ . Вторая — ее влияние на стабильность белка, т. е. на разность свободных энергий нативного и денатурированного состояний:  $\Delta(F_N - F_u) = -RT \Delta \ln [k_{u \rightarrow N} / k_{N \rightarrow u}]$ . Их отношение  $\Delta(F^\# - F_u) / \Delta(F_N - F_u)$  называется величиной  $\Phi_f$  для мутированного остатка. В данном случае  $\Phi_f \approx 1/4$ , т. е. остаток, мутация которого рассматривается, как бы на четверть вовлечен в ядро сворачивания. Показанные на графике величины относятся к концентрации денатуранта  $C_0$ , где требующаяся экстраполяция минимальна (а потому минимальны и погрешности в ней). Однако, ценой несколько большей экстраполяции (до  $C = 0$ ), так же обычно определяются изменения в стабильности нативного белка  $\Delta(F_N - F_u)$  и в высоте барьера  $\Delta(F^\# - F_u)$ , а также величины  $\Phi_f$ , относящиеся к чистой воде

- Для оценки вовлеченности остатка в нативоподобную часть переходного состояния (как говорят, в «зародыш», или в «ядро сворачивания» белка), оценивают, по сдвигу шеврона, влияние мутации данного остатка на (а) скорость сворачивания белка и (б) на стабильность нативной формы белка. Мутируя остаток с целью изучения ядра сворачивания, его обычно меняют на более мелкий (как правило, на аланин или, реже, глицин) с тем, чтобы структура мутантного белка не искажалась бы слишком сильно. Измеряемые величины показаны на рис. 20-5.

Стабильность нативной структуры белка N относительно его развернутого состояния U (при заданном состоянии среды, например, заданной концентрации денатуранта) определяется величиной  $F_N - F_u$ , разностью свободных энергий этих двух состояний белковой молекулы (измеренной при тех же условиях). А  $F_N - F_u$  определяется (см. формулу (20.4)) соотношением между скоростью сворачивания и скоростью разворачивания

белка при заданном состоянии среды. Так как стабильность белка обычно относят к чистой воде, то скорость разворачивания, измеряемую при высоких концентрациях денатуранта, обычно приходится экстраполировать к низким его концентрациям.

Скорость сворачивания белка (перехода  $u \rightarrow N$ ) определяется (см. формулу (20.1)) разностью свободной энергии «зародыша» # и исходного развернутого U состояния белка (т. е. величиной  $F^\# - F_u$ ).

Мутации влияют и на скорость сворачивания, и на стабильность белка. Это дает возможность использовать их для экспериментального определения формы зародыша нативной структуры белка.

Как же можно интерпретировать влияние мутаций на соотношение величин  $F_N - F_u$  и  $F^\# - F_u$  в предположении о **нуклеационном механизме** сворачивания, заключающемся в том, что остатки — уж если они вовлечены в зародыш нативной структуры — стоят там так же, как в нативном белке?

Существование нуклеации при сворачивании белка весьма правдоподобно, так как это сворачивание протекает как переход типа «все-или-ничего» (а это, как мы знаем, — микроскопический аналог фазового перехода первого рода в макроскопических системах). Нуклеация типична для фазовых переходов первого рода (например, для кристаллизации).

Об экспериментальной проверке этого предположения я расскажу вскоре. Сначала я хочу рассказать, как величины  $F_N - F_u$  и  $F^\# - F_u$  используются для того, чтобы очертить зародыш нативной структуры («ядро сворачивания»).

- Если мутация остатка так же влияет на величину  $F^\# - F_u$ , как она влияет на стабильность всего нативного состояния (т. е. на величину  $F_N - F_u$ ), — то это свидетельствует о том, что рассматриваемый остаток так же вовлечен в «зародыш» (образует там те же контакты, имеет ту же конформацию и т. д.), как он вовлечен в нативную глобулу.

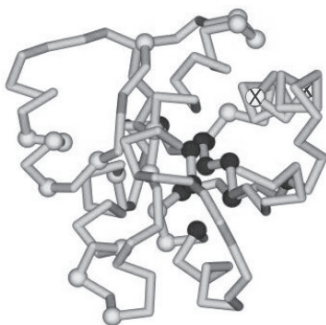
Если, наоборот, мутация остатка оказывает влияние только на стабильность белка (т. е. на величину  $F_N - F_u$ ), но не на скорость сворачивания (т. е. она не влияет на величину  $F^\# - F_u$ ), — значит, этот остаток не вовлекается в глобулярный «зародыш» сворачивания, т. е. он входит в нативную белковую глобулу только уже после образования зародыша.

И, наконец, если мутация остатка сильно влияет на стабильность белка и слабее (но с тем же знаком) — на стабильность зародыша — значит, этот остаток либо входит в один из нескольких альтернативных зародышей, либо он образует в зародыше только часть тех контактов, которые имеет в нативном белке, т. е. лежит на поверхности зародыша

Так очерчивается зародыш сворачивания белка (рис. 20-6): для каждого из мутированных остатков цепи вычисляется величина

$$\Phi_f = \Delta(F^\# - F_u) / \Delta(F_N - F_u) \quad (20.8)$$

— и, если для данного остатка  $\Phi_f$  близка к 1, то говорят, что он входит в зародыш структуры; а если  $\Phi_f$  близка к 0 — то нет.



**Рис. 20-6.** Структура переходного состояния белка CheY, согласно Т. López-Hernández & L. Serrano, *Folding & Design* (1996) 1:43–55. На фоне нативной укладки цепи CheY черными шариками выделены остатки, вовлеченные в структурированную часть переходного состояния (имеющие  $\Phi_f > 0,3$ , т. е. образующие там более 30 % своих нативных контактов), более светлыми шариками — не вовлеченные в структурированную часть переходного состояния остатки с  $\Phi_f < 0,3$ . Без шариков оставлены те области цепи, где мутации еще не делались. Крестами помечены два остатка, сложных для интерпретации, — у них столь малы ( $\ll kT$ ) величины  $\Delta(F^\# - F_u)$  и  $\Delta(F_N - F_u)$  (последнее важнее, так как эта величина стоит в знаменателе в формуле (20.8)), что погрешности в их определении превышают сами эти величины

Замечательно, что лишь очень малое число остатков в белке не удастся проинтерпретировать с такой точки зрения. (А с нее *не* удалось бы проинтерпретировать те остатки, что не вписывались в «нуклеационный механизм» сворачивания, т. е. те остатки, мутации которых влияли бы только на скорость сворачивания, но не на стабильность нативной структуры; или те, что стабилизировали бы только зародыш, но дестабилизировали нативный белок, и т. д. Для этих остатков получались бы величины  $\Phi_f > 1$  и  $\Phi_f < 0$ , что наблюдается крайне редко, и то, как правило, только тогда, когда погрешности в измерении  $\Delta(F^\# - F_u)$  и особенно  $\Delta(F_N - F_u)$  велики.

Это вселяет уверенность в том, что базовая картина, согласно которой остатки — уж если они вовлечены в зародыш — стоят там так же, как в нативном белке, — эта картина в основном справедлива. Подчеркиваю — в основном, но вряд ли полностью. Появляются данные, что находящаяся в «переходном состоянии» белковая цепь более компактна, чем можно было бы ожидать для «нативоподобного кристаллика», окруженного клубком. Можно осторожно предположить, что «нативоподобный кристаллик»

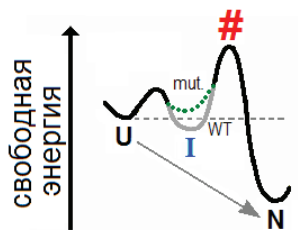
переходного состояния окружен, по крайней мере в некоторых белках, чем-то вроде расплавленной глобулы.

■ Рисунок 20-6 показывает, что найденные описанным выше способом в белке CheY остатки, наиболее существенные для сворачивания этого белка, группируются в компактный «доменчик», компактное ядрышко сворачивания, причем оно лежит не в геометрическом центре белка, а сдвинуто к его поверхности. Аналогичная картина часто наблюдается и в других белках, исследованных на предмет местоположения их ядер сворачивания. Это показывает, что ядро сворачивания белка (folding nucleus) и его гидрофобное ядро (hydrophobic core) — вещи разные. В ряде других белков наблюдается очень «диффузное» — охватывающее большую часть белка, но слабо выраженное (характеризующееся низкими  $\Phi_f$  величинами) и не имеющее четких границ ядро сворачивания. По-видимому, сворачивание этих белков идет многими параллельными путями через разные ядра сворачивания — что в эксперименте, который одновременно наблюдает за множеством сворачивающихся белков, и проявляется как слабо очерченное ядро. Такая точка зрения подкрепляется тем, что некоторые введенные в эти белки мутации могут более четко локализовать ядро сворачивания, причем одни мутации локализуют его в одном, а другие — в другом месте.

**Внутренний голос:** Из того, что Вы сказали, следует, что любая мутация, влияющая на скорость сворачивания белка, должна влиять и на его стабильность. Но есть эксперименты, показывающие обратное!

**Лектор:** То, что я говорил, относится к переходам типа «все-или-ничего», при которых нет накопления метастабильных «полусвернутых» состояний белковой цепи. Если такие промежуточные состояния присутствуют (что возможно только вдали от точки равновесия нативной и денатурированной форм белка, см. загиб кривой в самой левой части рис. 20-4, а также профиль свободной энергии на рис. 20-7), то мутация, не затрагивая стабильности нативного белка, может влиять на его метастабильное промежуточное состояние, и упомянутый Вами эффект — кстати, он и наблюдался в таких условиях — становится возможным.

■ В заключение я хочу вернуться к рис. 20-4 и подчеркнуть, что даже для совершенно конкретного белка не существует четко определенного «характерного времени сворачивания». Действительно, этот рисунок показывает, что сворачивание лизоцима занимает порядка 0,1 с в нативных условиях и порядка 10 000 с в условиях созданного гуанидингидрохлоридом (при той же температуре раствора) равновесия нативной и денату-



**Рис. 20-7.** Схема профиля свободной энергии для перехода из развернутого (U) состояния в нативное (N) через интермедиат (I), который стабилен (по сравнению с U) в белке дикого типа (WT) и не стабилен в белке с мутацией, mut (что показано, в районе I, серой частью сплошной кривой и точечной кривой, соответственно); здесь предполагается, что мутация влияет только на интермедиат I. В данном случае, сворачивание белка с мутацией происходит быстрее, так как этому белку не надо выкарабкиваться из стабильного (по сравнению с U) интермедиата перед попаданием в лимитирующее скорость всего процесса переходное состояние #. Обратите внимание, что переход из U в N происходит в одну стадию (U→N) для белка с мутацией, где интермедиат I неустойчив по сравнению с исходным состоянием U (и, следовательно, не наблюдается), и в две стадии (U→I→N) для белка без мутации, где интермедиат I стабильнее, чем U. Однако повышение свободной энергии состояния N (сближающее стабильность N со стабильностью U) дестабилизирует состояние I и для белка без мутации (мысленно приподнимите немного кривую за ее правый край), и тогда переход в нем также протекает в одну стадию, U→N

рированной форм. А в других условиях, при повышенной температуре, но без гуанидингидрохлорида, области равновесия этих форм отвечает (см. рис. 20-2) время сворачивания порядка 10 с. Так что, обсуждая скорость сворачивания белка, мы должны иметь в виду либо весь наблюдаемый диапазон времен сворачивания данного белка, либо совершенно конкретные экспериментальные условия, например, те условия, при которых белок сворачивается в клетке.

## Лекция 21

Решение «парадокса Левинталь»: к стабильной структуре цепи автоматически ведет сеть быстрых путей сворачивания. Для этого необходимо только, чтобы между нативной укладкой цепи и прочими ее глобулярными укладками существовала бы заметная энергетическая щель. Обсуждение аномально медленного образования стабильной структуры в некоторых белках (серпины, прионы). Представление об «энергетических ландшафтах» белковых цепей. Белковые структуры: физика самоорганизации и естественный отбор самоорганизующихся цепей.

Мы продолжаем разговор о самоорганизации белков.

- Все экспериментальные данные, о которых я рассказывал, — сколь бы интересны они ни были сами по себе — не дают ответа на вопрос, как белок ухитряется найти свою нативную структуру среди астрономического числа возможных за те немногие секунды или доли секунды, что отпущены на его сворачивание.

А число это — как я говорил, его оценил еще в 1968 г. Сайрус Левинталь — действительно огромно:  $\sim 10^{100}$  возможных конформаций для цепи из 100 остатков; их «тупой» перебор занял бы  $\sim 10^{80}$  лет, кладя всего  $10^{-13}$  секунды на переход из одной конформации в другую. А достаточно полный перебор конформаций, считал Левинталь, может быть только «тупым» — ведь белок может «почувствовать» стабильность конформации, только попав прямо в нее, так как отклонение даже на 1 Å может очень сильно повысить энергию цепи в плотной белковой глобуле.

«Как же белок выбирает свою нативную структуру среди бесчисленного множества возможных? — спросил Левинталь, и ответил: — По-видимому, самоорганизующийся белок следует по какому-то специальному «пути сворачивания», и та структура, где этот путь заканчивается, и является его нативной структурой, вне зависимости от того, есть ли еще более стабильная укладка цепи, или нет». Иными словами, Левинталь предположил,

что нативная структура белка определяется *не* стабильностью, *не* термодинамикой, а кинетикой, т. е. она соответствует не глобальному, а просто быстро достижимому минимуму свободной энергии цепи.

Сложность проблемы заключается в том, что поднятый вопрос нельзя решить чисто экспериментально. Действительно: предположим, у белковой цепи есть другая, «ненативная», но еще более стабильная укладка. Как ее найти, если сам белок ее не находит? Ждать результата в течение  $10^{80}$  лет?!

С другой стороны, вопрос о том, что — кинетика или термодинамика — определяет укладку белковой цепи, постоянно возникает на пути решения разных прикладных задач. Он возникает, когда речь идет о предсказании структуры белка по его аминокислотной последовательности (надо знать, что предсказывать: самую стабильную или самую быстро сворачивающуюся его структуру). Он возникает и тогда, когда речь идет о дизайне новых, не встречающихся в природе белков (надо знать, что делать: максимально усиливать стабильность желаемой структуры или пролагать максимально быстрый путь к ней).

■ Однако действительно ли существует противоречие между «структурой стабильной» и «структурой быстро сворачивающейся»? Может быть, стабильная структура *автоматически* является фокусом «быстрых» путей сворачивания и потому *автоматически* обладает свойством сворачиваться быстро?

■ Прежде чем приступить к исследованию этих вопросов, т. е. прежде, чем рассматривать *кинетические* аспекты сворачивания белков, вспомним ряд фундаментальных фактов из области их *термодинамики* (здесь всюду речь идет об относительно небольших, однодоменных белках, т. е. о белках из 50–200 аминокислотных остатков). Эти факты помогут нам понять, какие условия протекания процесса сворачивания мы должны рассматривать. Термодинамические факты таковы:

1) Разворачивание белка обратимо, причем оно происходит как переход «*все-или-ничего*». Последнее означает, что в точке денатурации белка только две формы белковой молекулы — «нативная» и «денатурированная» — присутствуют в заметных количествах, а все прочие («полусвернутые» и «неверно свернутые» формы) практически отсутствуют. Такой переход возможен только в цепях, аминокислотная последовательность которых обеспечивает достаточно большую энергетическую щель между основной массой структур и самой стабильной структурой цепи.

2) Денатурированная форма белков — во всяком случае, небольших белков, развернутых сильным и концентрированным денатурантом, — часто является неупорядоченным клубком.

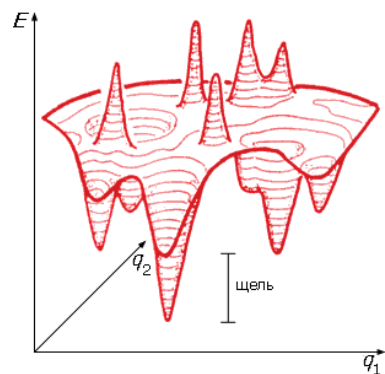
3) В нормальных физиологических условиях нативная форма белка лишь немногим стабильнее его денатурированной формы (а в средней точке перехода обе эти формы имеют, естественно, одинаковую стабильность). При этом нативная структура белка стабильна благодаря своей низкой энергии, т. е. благодаря сильным взаимодействиям в нативной структуре, а развернутая — благодаря своей высокой конформационной энтропии, т. е. благодаря огромному числу разных развернутых конформаций.

*Необходимое пояснение:* как принято в литературе, термин «энергия» здесь означает, строго говоря, *всю свободную энергию взаимодействий*, в том числе взаимодействий цепи с растворителем (например, «энергия» гидрофобных взаимодействий определяется, как вы должны помнить, энтропией растворителя); термин же «энтропия» здесь охватывает лишь конформационную энтропию цепи, но *не* энтропию растворителя. Такая терминология принята, чтобы, оставив растворитель за скобками, сосредоточиться на главной проблеме — как белковая цепь находит «свою» пространственную структуру среди гигантского числа возможных.

Итак, чтобы разрешить «парадокс Левинталя» и показать, что самую стабильную структуру белковой цепи можно найти за разумное время, мы можем, в первом приближении, рассматривать только скорость перехода «*все-или-ничего*» из клубка в самую стабильную структуру цепи, причем нам достаточно рассмотреть случай, когда самая стабильная укладка цепи лишь чуть-чуть более стабильна, чем клубок, а все прочие формы белковой цепи термодинамически нестабильны. Здесь рассматривать сворачивание белка наиболее просто, так как здесь нет стабильных интермедиатов сворачивания: они, как о том говорилось на прошлой и особенно позапрошлой лекции, появляются лишь тогда, когда нативная структура становится много стабильнее клубка. При этом сворачивание белка достигает максимальной скорости, но его анализ усложняется. Поэтому мы, для начала, ограничимся окрестностью точки равновесия нативной структуры с клубком, где белок сворачивается пусть не максимально быстро, но максимально просто. Причем мы должны рассматривать только такие аминокислотные последовательности, которые обеспечивают наличие большой «щели» между энергией самой стабильной структуры цепи и энергиями всех прочих ее укладок (рис. 21-1): как вы помните, статистическая физика гетерополимеров показывает, что фазовый, типа «*все-или-ничего*» распад нативной глобулы требует наличия такой щели, резко отделяющей нативную структуру от всех прочих.

■ Я прежде всего собираюсь показать, что при таких условиях самая *стабильная структура* небольшого белка или домена *автоматически* становится центром «быстрых» путей сворачивания и потому должна сворачиваться за биологически разумное время: секунды или минуты.





**Рис. 21-1.** Грубое схематическое изображение энергетического ландшафта белковой цепи: на рисунке мы можем изобразить только две координаты ( $q_1$  и  $q_2$ ), описывающих конформацию цепи, тогда как реальная конформация белковой цепи описывается сотнями координат. Широкая щель между глобальным энергетическим минимумом и прочими энергетическими минимумами необходима для того, чтобы стабильная укладка цепи разрушалась только путем термодинамического перехода типа «все-или-ничего»; это, в свою очередь, обеспечивает надежность функционирования белка — по принципу «все-или-ничего», как у электрической лампочки

Для того чтобы доказать, что самая стабильная белковая структура должна сворачиваться быстро, достаточно показать, что к ней всегда ведет по крайней мере один «быстрый» путь сворачивания. Наличие не одного, а многих путей сворачивания нативной структуры может только ускорить процесс.

А вот путей сворачивания не нативной, а прочих (и, значит, при наличии «щели» — высокоэнергетичных) структур мы можем не рассматривать вовсе! Они не могут сбить нашу цепь с пути истинного. Ведь вблизи точки перехода типа «все-или-ничего» между самой стабильной структурой цепи и клубком — *никакие* «полу-» и «неверно свернутые» состояния не могут служить ловушками: они не могут «впитать» сворачивающиеся цепи просто в силу малости своей суммарной стабильности. Полезная аналогия здесь — просачивание воды через трещины в стенке, разделяющей два бассейна: если «емкость» трещин мала, т. е. они не способны впитать всю воду, то каждая новая трещина в стенке может только ускорить наполнение второго бассейна, так что, рассмотрев просачивание воды через одну трещину, мы оценим минимальную скорость его наполнения.

Чтобы путь был быстр, каждый шаг на этом пути должен проходиться быстро, таких шагов должно быть не слишком много, и — главное! — этот

путь не должен преграждаться «барьером» в виде очень высокой свободной энергии на одной из стадий сворачивания.

Так как время фиксации одного звена мало ( $\sim 1-10$  нс, судя по измеренной скорости роста  $\alpha$ -спиралей в белковых цепях), то белок, фиксируя одно свое звено за другим, сворачивался бы мгновенно (100-звенная цепь — за  $\sim 100-1000$  нс), *если бы* при этом он не должен был преодолевать свободно-энергетический барьер.

■ Итак: главный вопрос, на который надо ответить, — высок ли барьер на пути, ведущем к самому стабильному состоянию белковой цепи?

Сворачивание белковой цепи ведет к падению ее энтропии (из-за роста упорядоченности цепи) и энергии (из-за образования контактов между сближающимися звеньями цепи). Падение энтропии повышает, а падение энергии понижает свободную энергию цепи. Если, по ходу сворачивания, цепь должна очень близко подойти к своей финальной структуре перед тем, как начнут возникать стабилизирующие эту структуру контакты (т. е. цепь должна потерять почти всю свою энтропию *перед* тем, как начнет выигрываться энергия), то повышение свободной энергии на первом этапе сворачивания будет пропорциональным числу звеньев в цепи, т. е. очень большим, а сворачивание цепи — страшно медленным (как вы помните, согласно химической кинетике, время протекания процесса экспоненциально зависит от достигаемого по его ходу максимального повышения свободной энергии). Именно такая картина (проигрыш всей энтропии *до* начала выигрыша энергии) лежит в основе «парадокса Левинтала», утверждающего, что белковая цепь никак не может, даже за время жизни Вселенной, найти свою самую стабильную структуру.

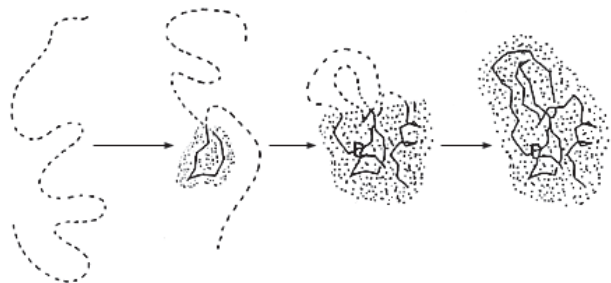
Напротив, если путь сворачивания таков, что по ходу его падение энтропии практически тут же компенсируется падением энергии, то он не перекрыт высоким свободно-энергетическим барьером, и сворачивание идет быстро. Именно такая картина, как мы убедимся, и имеет место.

При наличии барьера характерное время протекания процесса оценивается, исходя из классической теории переходного состояния, как

$$\text{ВРЕМЯ} \sim \tau / \exp(-\Delta F^\ddagger / RT). \quad (21.1)$$

Здесь  $\tau \sim 1$  нс (точнее,  $\sim 10$  нс) — характерное время одного шага процесса сворачивания белка (его можно оценить по скорости сворачивания  $\alpha$ -спиралей),  $T$  — абсолютная температура,  $R$  — газовая постоянная, а  $\Delta F^\ddagger$  — высота максимума свободной энергии на пути сворачивания (иными словами, свободной энергии переходного состояния) относительно свободной энергии исходного (клубкового) состояния цепи.

Рассмотрим изменение энергии  $\Delta E$ , энтропии  $\Delta S$  и результирующей свободной энергии  $\Delta F = \Delta E - T\Delta S$  по ходу изображенного на рис. 21-2 последовательного сворачивания белка. На каждом шаге этого пути одно звено цепи извлекается из клубка и занимает то положение, в котором оно находится в финальной (самой стабильной) структуре глобулы. Такой процесс может показаться несколько искусственным (откуда звено знает свое место в финальной структуре?). Однако это впечатление исчезает, если заметить, что мы таким образом просто просматриваем «кино» распада стабильной структуры белка в обратном направлении.



**Рис. 21-2.** Один из возможных путей последовательного сворачивания белка. Штриховой линией показана цепь, остающаяся в неупорядоченном состоянии — «клубке». Выделенная точками область соответствует уже обретшей свою финальную конформацию части белковой глобулы. Жирной линией выделена главная цепь уже свернувшейся глобулы, а ее боковые группы не показаны для упрощения рисунка. Все промежуточные состояния имеют высокую свободную энергию и потому не накапливаются при сворачивании и не могут наблюдаться непосредственно

**Внутренний голос:** А не может ли белок сворачиваться по одному пути, а разворачиваться — совсем по другому?!

**Лектор:** Не может, если мы рассматриваем оба процесса — и сворачивание, и разворачивание — *при одних и тех же условиях* (температуре, рН, и т. д.). А мы уже договорились рассматривать сворачивание белка вблизи точки термодинамического равновесия его нативной и денатурированной формы. При этом белок сворачивается с той же скоростью, что и разворачивается — по определению точки равновесия. Вспомните, что именно это действительно наблюдается около нижней точки шевронного графика. Причем и сворачивание, и разворачивание должны идти по одному и тому же пути: если бы это было не так, если бы белок сворачивался по одному пути, а разворачивался по другому, то возник бы — в равновесных условиях! — вечный кольцевой поток, т. е. возник бы (если поставить на этот вечный поток турбинку) *вечный двигатель второго рода*. А, как вы знаете (должны знать!), термодинамика утверждает,

что вечный двигатель второго рода невозможен: он противоречил бы общему принципу неубывания энтропии (беспорядка).

Во избежание недоразумений уточню: *если* сворачивание и разворачивание рассматриваются при *разных* условиях (например, сворачивание — в воде, а разворачивание — в крепком растворе мочевины), то эти процессы могут идти по разным путям. Но в одних и тех же условиях оба они идут по одной дорожке. Это — общий физический *принцип детального равновесия*.

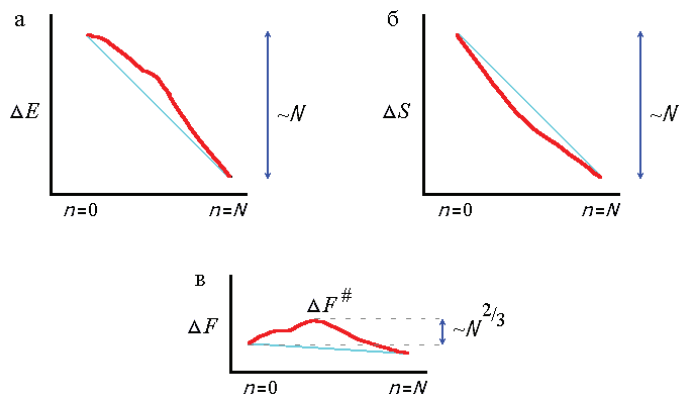
**Внутренний голос:** И все же... Кино о взрыве здания, даже прокрученное в обратную сторону, не похоже на кино о его постройке!

**Лектор:** Да, но и строительство, и взрыв идут за счет громадного необратимого расхода энергии: горючее, взрывчатка, сила мускулов... А сворачивание белка *in vitro*, как и его разворачивание, не потребляют никакого «горючего», и, как показывают шевронные графики, оба могут при этом идти в одних и тех же условиях, вблизи и даже в самой точке равновесия свернутой и развернутой форм. Я много раз подчеркивал этот факт, потому что он очень важен для понимания сворачивания белка: раз оба процесса идут в одних и тех же условиях, они идут по одному пути. То есть, если какой-то путь хорош для разворачивания — то он столь же хорош (если идти в другую сторону) для сворачивания белка.

Итак, изображенный на рис. 21-2 путь сворачивания существует — среди множества других путей сворачивания. И как мы уже договорились, для минимальной оценки скорости сворачивания достаточно оценить скорость сворачивания по *одному* пути.

По мере роста куска финальной глобулы при ее последовательном сворачивании, в ней одно за другим восстанавливаются взаимодействия, стабилизирующие финальную структуру. Если растущая структура все время остается более или менее компактной (а именно такого типа пути сворачивания, с изображенными на рис. 21-2 компактными промежуточными состояниями, нас и должны интересовать), то число этих взаимодействий будет расти (а их энергия, соответственно, будет падать) почти пропорционально числу  $n$  фиксируемых в глобуле звеньев (рис. 21-3а).

Правда, в начале сворачивания падение энергии несколько замедлено, так как прилипание звена к поверхности маленькой глобулы дает, в среднем, меньше контактов, чем прилипание к поверхности большой. В результате возникает нелинейный поверхностный (т. е. пропорциональный  $n^{2/3}$ ) член в энергии  $\Delta E$  растущей глобулы. Таким образом, максимальное отклонение от линейного падения энергии составляет величину порядка  $N^{2/3}$ , где  $N$  — число звеньев в белковой цепи. Это отклонение, очевидно, мало по сравнению с полным падением энергии при сворачивании цепи, которое составляет величину порядка  $N$  (точнее — порядка  $N\epsilon$ , где  $\epsilon$  — среднее изменение энергии остатка при его переходе из клубка в нативную глобулу).

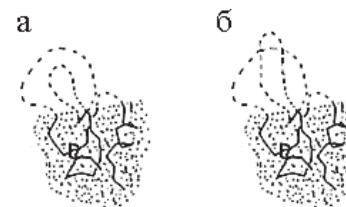


**Рис. 21-3.** Изменение энергии  $\Delta E$  (а), энтропии  $\Delta S$  (б) и свободной энергии  $\Delta F = \Delta E - T\Delta S$  (в) цепи по мере последовательного сворачивания белка ( $n = 0$ : клубок;  $n = N$ : финальная нативная структура) вблизи точки термодинамического равновесия между финальной структурой и клубком. Тонкой линией показана линейная (по  $n$ ) часть изменения  $\Delta E(n)$ ,  $\Delta S(n)$  и  $\Delta F(n)$ . Полное изменение  $\Delta E(N)$  и  $\Delta S(N)$  примерно пропорционально  $N$ , т. е. числу звеньев в цепи, а максимальное отклонение показанных жирной линией величин  $\Delta E(n)$  и  $\Delta S(n)$  от линейного (по  $n$ ) хода относительно мало — оно пропорционально всего  $N^{2/3}$ . В результате  $\Delta F(n) = \Delta E(n) - T\Delta S(n)$  тоже мало (на  $\sim N^{2/3}$ ) отклоняется от линейного хода, и, в точке равновесия [при  $\Delta F(0) = \Delta F(N)$ ], это отклонение  $\Delta F^\#$  (т. е. максимальная величина свободной энергии на пути сворачивания) также пропорционально всего  $N^{2/3}$

Во-вторых, по ходу роста глобулы падает энтропия цепи — примерно пропорционально числу фиксированных в глобуле звеньев (рис. 21-3б). Правда, в начале сворачивания энтропия может падать несколько быстрее из-за образования замкнутых петель, свободно торчащих из растущей глобулы (рис. 21-4). В результате возникает нелинейный (поверхностный) член в энтропии  $\Delta S$  этой растущей глобулы, составляющий (так же, как поверхностный член в  $\Delta E$ ) величину порядка  $N^{2/3}$ , что существенно меньше, чем само падение энтропии, составляющее (как и падение энергии) величину порядка  $N$ .

Входящие в энтропию  $\Delta S$  и энергию  $\Delta E$  линейные и нелинейные члены должны войти и в свободную энергию  $\Delta F = \Delta E - T\Delta S$  растущей глобулы. Однако, когда мы рассматриваем условия, в которых финальная глобула находится в термодинамическом равновесии с клубком (или почти в равновесии — скажем, когда она всего на пару  $RT$  стабильнее клубка), то в разности  $\Delta E - T\Delta S$  большие (пропорциональные  $n$ ) линейные члены взаимно аннигилируют. В самом деле, в самой точке равновесия  $\Delta F = 0$  и в клубке (т. е. при  $n = 0$ ), и в финальной глобуле (т. е. при  $n = N$ ). Значит,

рост энтропии и падение энергии при последовательном сворачивании компенсируют друг друга в главном (линейном по длине свернутой цепи  $n$ ) члене и, *не будь поверхностных* эффектов,  $\Delta F$  было бы равно 0 на всем рассматриваемом пути сворачивания!

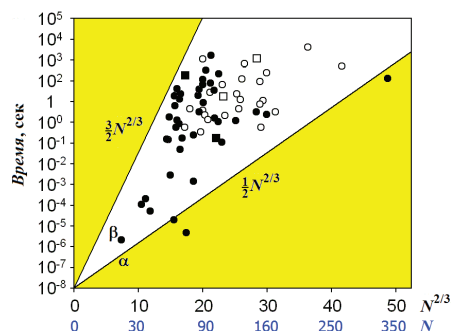


**Рис. 21-4.** (а) Компактный интермедиат сворачивания белка с торчащими из него неупорядоченными петлями. Рост интермедиата соответствует смещению границы между фиксированной (глобулярной) и неупорядоченной (клубковой) частями белковой цепи. Успешное сворачивание интермедиата требует правильной заузленности петель в нем: полусвернутая структура с неправильным заузлением (б) не может прорасти до правильно свернутого белка, ей сперва надо развернуться. Однако в цепи из  $\sim 100$  звеньев может сформироваться лишь один-два узла, так что перебор интермедиатов с разной заузленностью петель не должен лимитировать скорость сворачивания белков

Следовательно, свободно-энергетический барьер связан только с относительно малыми нелинейными поверхностными эффектами. Значит, высота барьера на любом пути последовательного сворачивания (типа того, что изображен на рис. 21-2) пропорциональна *не* числу звеньев  $N$  (как то, фактически, полагал Левинталь), а всего лишь  $N^{2/3}$  (рис. 21-3в).

В результате время достижения самой стабильной структуры растет с числом звеньев цепи  $N$  не «по Левинталю» (т. е. не как  $10^N$  и вообще не как экспонента от  $N$ ), а всего лишь как  $\exp(\lambda N^{2/3})$  — причем здесь наиболее важно как то, что величина  $N^{2/3}$  возникает за счет «расслоения» нативной и развернутой фаз цепи в ходе сворачивания, так и то, что  $N^{2/3}$  существенно меньше, чем  $N$ . В то же время тщательная оценка коэффициента  $\lambda$  показывает, что  $\lambda = 1 \pm 0,5$ , причем конкретная величина  $\lambda$  зависит как от формы укладки цепи в глобулу (от того, много или мало неупорядоченных петель торчит из нативоподобного ядрышка), так и от распределения сильно и слабо притягивающихся аминокислотных остатков в ней (от того, сгруппированы ли в ядрышке сильно притягивающиеся остатки).

Наблюдаемые времена сворачивания (перехода «клубок  $\rightarrow$  нативная глобула» в точке термодинамического равновесия этих двух состояний) лежат (рис. 21-5) в диапазоне от  $10\text{нс} \times \exp[0,5N^{2/3}]$  до  $10\text{нс} \times \exp[1,5N^{2/3}]$ , в соответствии с нашей оценкой.



**Рис. 21-5.** Экспериментальные данные по скорости сворачивания белков на фоне разрешенного формулой (21.2) диапазона. Данные относятся к точке термодинамического равновесия клубка с нативной глобулой, т. е. к самому низу кинетического шеврона.  $N$  — число аминокислот в цепи. Кружки относятся к белкам без S–S-связей, квадратики — к белкам с S–S-связями. Черные значки относятся к белкам, сворачивающимся без видимых интермедиатов во всем диапазоне экспериментальных условий; светлые — к тем, для которых интермедиаты наблюдаются (но вне зоны термодинамического равновесия клубка с нативной глобулой, ближе к «физиологическим» условиям). Значки  $\alpha$  и  $\beta$  относятся к сворачиванию отдельной  $\alpha$ -спирали [Thompson P. A., Eaton W. A., Hofrichter J., *Biochemistry* (1997) **36**, 9200] и  $\beta$ -шпильки [Mucoz V., Thompson P. A., Hofrichter J., Eaton W. A., *Nature* (1997) **390**, 196] соответственно. Картинка взята из работы О. V. Galzitskaya, D. N. Ivankov, A. V. Finkelstein, *FEBS Lett.* (2001) **489**: 113–118 (основанной на списке белков Jackson S. E., *Folding & Design* (1998) **3**: R81–R91) и дополнена данными за последующие пять лет

Причина того, что самая стабильная структура белка достигается всего за

$$\text{Время} \sim \exp [(1 \pm 0,5) N^{2/3}] \times 10 \text{ нс} \quad (21.2)$$

— в том, что падение энтропии по ходу последовательного сворачивания тут же (с точностью до поверхностных эффектов) компенсируется энергией возникающих взаимодействий.

Отметим, что на рассмотренном нами пути глобулярный зародыш белковой структуры *не* перестраивается по ходу сворачивания (на что потребовалось бы гигантское время), а все перестройки происходят только в рыхлом клубке (и потому быстро).

Как бы то ни было, полученная оценка (21.2) и рис. 21-5 показывают, что цепь из 80 и менее аминокислотных остатков должна за минуты (или менее) находить свою самую стабильную структуру всегда, цепь из 100–200 остатков — весьма часто, а цепь из многих сотен остатков найти свою самую стабильную однодоменную структуру за разумное время вообще не может. Она объясняет также, почему большие белки состоят (согласно

старому принципу «разделяй и властвуй») из отдельно сворачивающихся стабильных доменов — или, как теперь говорят в данном случае, «фолдонов» (от слова “fold” — «укладка»): иначе их сворачивание было бы слишком медленным.

Добавление. Формула (21.2) оценивает скорее диапазон времен сворачивания, чем скорость сворачивания каждого конкретного белка. А из рис. 21-5 видно, что скорости сворачивания белков одного и того же размера могут порой различаться в миллионы раз.

Плакско и Бейкер показали, что различия в логарифмах скоростей сворачивания небольших белков, сворачивающихся без наблюдаемых интермедиатов, хорошо коррелируют с введенным ими «контактным порядком»

$$CO = \frac{1}{M \times N} \sum_{(i,j) \in M} \Delta N_{ij},$$

где  $M$  — число межатомных контактов в белке,  $\Delta N_{ij}$  — число остатков цепи между контактирующими (на расстоянии  $\leq 6 \text{ \AA}$ ) атомами  $i$  и  $j$ , а  $N$  — число аминокислотных остатков в цепи.  $CO$  отражает устройство белка: он мал для тех белков, где массово контактируют между собой близкие по цепи остатки (это, в основном,  $\alpha$ -белки), и велик для тех, где контактируют в основном далекие по цепи остатки.

Впоследствии оказалось, что, хорошо работая для небольших белков,  $CO$  (учитывающий только форму укладки белковой цепи, но не ее размер) совсем не работает ни для больших, ни для очень маленьких белков. Поэтому было естественно «скрестить»  $CO$  с учитывающим размер белка фактором  $N^{2/3}$  — тем более, что  $CO$  близок по смыслу к оцененному нами только снизу и сверху коэффициенту  $1 \pm 0,5$  в формуле (21.2):  $CO$  связан с петлями, образованными белковой цепью в нативной глобуле, а « $1 \pm 0,5$ » — с петлями, торчащими из «полусвернутой» глобулы. В результате совместной работы двух групп, нашей и Бейкера, было показано, что скорости сворачивания *всех* белков коррелируют с величиной  $CO \times N$ , которая растет с ростом длины белковой цепи  $N$ , в среднем, как  $N^{2/3 \pm 0,3}$ . Так была получена теория, способная неплохо предсказывать скорости сворачивания белков по их пространственной структуре.

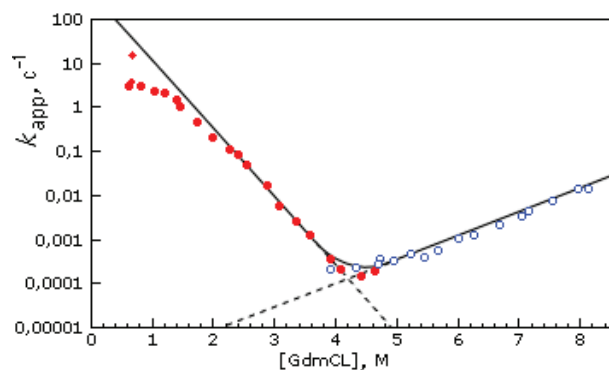
Более того: оказалось, что скорости сворачивания еще лучше коррелируют просто с числом аминокислотных остатков цепи, из которого вычтено суммарное число водородных связей в  $\alpha$ -спиральных участках (что согласуется гипотезой о том, что  $\alpha$ -спирали могут служить блоками самоорганизации: и действительно, они очень стабильны в развернутых цепях некоторых быстро сворачивающихся белков). А так как спиральные участки можно неплохо (например, программой Psipred Д. Джонса) предсказывать по первичной структуре белка — теперь можно предсказывать (с точностью до порядка величины) скорости сворачивания белков (и пептидов) *только* по их аминокислотной последовательности.

#### ■ Примечания.

(1) Найдя максимум свободной энергии на пути сворачивания, мы можем, далее, найти ту точку пути, которой он соответствует, т. е. оценить размер ядра сворачивания белковой структуры. Расчет показывает, что это ядро довольно велико, оно охватывает около половины белковой глобу-



лы. Об этом же говорит эксперимент: примерно одинаковая (но с разным знаком!) скорость изменения  $u$  времени сворачивания,  $u$  времени разворачивания белка с концентрацией денатуранта (рис. 21-6) показывает, что доступная растворителю поверхность цепи в переходном состоянии лежит примерно посередине между доступными растворителю поверхностями цепи в нативном и в развернутом состояниях. Столь большой размер ядра показывает, что белок не может сворачиваться по очень многим параллельным путям — путям с принципиально разными переходными состояниями (и ядрами!): большие компактные ядра, если их много, должны сильно перекрываться, так что число независимых друг от друга путей сворачивания должно быть все равно невелико. Следовательно, мы не сделали большой погрешности, рассматривая только один из возможных путей сворачивания.



**Рис. 21-6.** Зависимость скорости ренатурации и денатурации белка (лизоцима куриного яйца) от концентрации гуанидинхлорида. Экспериментальные точки изображают величину  $k_{app} = k_{u \rightarrow N} + k_{N \rightarrow u}$ , видимой (apparent) характерной скорости приближения к равновесию между нативной и развернутой формами белка. Черные кружки получены при разбавлении раствора GdmCl, в котором изначально находился денатурированный белок, т. е. при полной или частичной ренатурации белка. При этом  $k_{app} \approx k_{u \rightarrow N}$ . Светлые кружки получены при добавлении GdmCl к изначально нативному белку, т. е. при полной или частичной его денатурации. При этом  $k_{app} \approx k_{N \rightarrow u}$ . Пунктир показывает экстраполяцию величин  $k_{u \rightarrow N}$  и  $k_{N \rightarrow u}$  в области излома шеврона. Загиб обеих линий экспериментальных точек в нижней части графика (при  $\approx 4,5$  М GdmCl) соответствует области термодинамического равновесия нативной и развернутой форм белка (где  $k_{u \rightarrow N} \approx k_{N \rightarrow u} \approx k_{app}/2$ ). В этих условиях время ренатурации составляет  $\sim 10^4$  с — при том, что в «почти чистой» воде, при 0,6 М GdmCl (верхняя левая часть кривой), оно приближается к 0,1 с. Картинка, с небольшими изменениями (добавлены экстраполяционные пунктиры) взята из Т. Kiefhaber, *Proc. Natl. Acad. Sci. USA* (1995) **92**:9029–9033

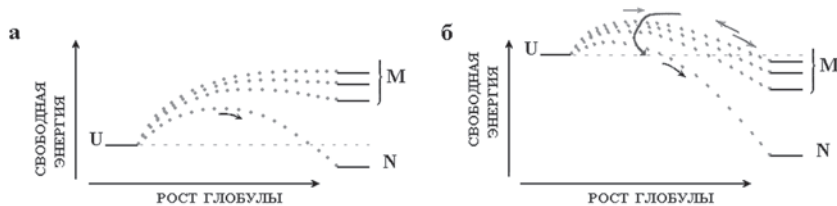
(2) Правда, наша оценка (21.2) относится к точке термодинамического равновесия клубка и нативной структуры белка, где экспериментально наблюдаемое время сворачивания этой структуры максимально, оно может превосходить время сворачивания в нативных условиях на несколько порядков (рис. 21-6). Поэтому оценка (21.2) носит не столько количественный, сколько принципиальный характер — она решает парадокс Левинтала и объясняет, почему белковая цепь может найти свою самую стабильную структуру не за астрономически огромное, а за биологически разумное время.

(3) Для очень длинных цепей, состоящих из многих тысяч звеньев, лимитирующим фактором мог бы стать «квази-Левинталевский» перебор по-разному заузленных интермедиатов — ведь узел в петле (ср. рис. 21-4а и 21-4б) нельзя распустить, не разрушив глобулярную часть интермедиата. Однако столь большие домены и так (по формуле (21.2)) свернуться не могут, а в доменах из сотни-другой звеньев много узлов быть не может (компьютерные эксперименты показали, что для образования одного узла нужно порядка сотни звеньев цепи), так что перебор по-разному заузленных конформаций увеличит время достижения стабильной структуры домена всего в несколько раз.

■ До сих пор мы рассматривали сворачивание белковой цепи вблизи точки фазового перехода, когда только одна («нативная») ее структура сравнима по стабильности с клубком, а все прочие, даже взятые вместе, нестабильны (рис. 21-7а).

Что произойдет со скоростью самоорганизации цепи при удалении от точки фазового перехода? Нас, естественно, интересует случай, когда стабильность нативной структуры растет (когда ее стабильность падает, эта структура просто не образуется).

Рост стабильности нативной глобулы (например, при понижении температуры или при разбавлении раствора денатуранта водой) приводит сперва к увеличению скорости достижения нативной структуры, так как свободная энергия интермедиатов ее сворачивания падает, а конкурирующие с нативной структурой укладки цепи все еще остаются — из-за «энергетической щели» — нестабильными (рис. 21-7а). При этом, если разность свободных энергий нативного и развернутого состояний меняется от нуля до  $\Delta G < 0$ , то разность свободных энергий переходного и развернутого состояний меняется примерно на  $\Delta G/2$ , так как ядро сворачивания охватывает примерно половину белка. Значит, его сворачивание ускоряется примерно в  $\exp(-\Delta G/2RT)$  раз. Следовательно, время достижения  $N$ -звенной белковой цепью своей самой стабильной структуры (свободная энергия которой относительно развернутого состояния составляет  $\Delta G \leq 0$ ) можно оценить как



**Рис. 21-7.** Сворачивание белка при разных условиях. Тонкие линии показывают уровни свободной энергии клубка  $U$ , самой стабильной («нативной») укладки цепи  $N$ , и прочих («misfolded», конкурирующих с самой стабильной) глобулярных структур  $M$ . Точечные линии показывают изменение свободной энергии по ходу сворачивания. Относительно небольшая их неровность (см. график  $\Delta F$  на рис. 21-3) на данном рисунке не показана. Высшие точки этих пунктирных линий соответствуют свободным энергиям переходных состояний на соответствующих путях сворачивания. Основные сценарии процесса сворачивания: (а) самая стабильная структура цепи  $N$  стабильнее клубка  $U$ , а все конкурирующие с ней структуры  $M$  в сумме менее стабильны, чем клубок. Идет быстрое и безошибочное сворачивание самой стабильной структуры; пути сворачивания «неверных» структур  $M$  практически не обследуются молекулой, так как на них свободная энергия только возрастает. (б) Помимо структуры  $N$ , многие свернутые структуры  $M$  более стабильны, чем клубок  $U$ . Идет быстрое сворачивание многих «неправильных» структур с их последующим, очень медленным частичным разворачиванием и переходом в нативную структуру  $N$ . Стрелки показывают главное русло процесса самоорганизации стабильной пространственной структуры белка

$$\text{ВРЕМЯ} \sim \exp[(1 \pm 0,5)N^{2/3} + \Delta G/2RT] \times 10 \text{ нс.} \quad (21.3)$$

Примерно такое ускорение, действительно, наблюдается (см. подъем графика на рис. 21-6, слева) до определенного предела (см. плечо графика на левом краю рис. 21-6). По-видимому, максимальная скорость сворачивания должна достигаться, когда конкурирующие с нативной структурой «неправильные» компактные укладки цепи (а их стабильность тоже растет при понижении температуры и разбавлении денатуранта) как раз сравниваются по стабильности с клубком. При дальнейшем росте стабильности глобул идет быстрое сворачивание многих «неправильных» структур с их последующим медленным частичным разворачиванием и переходом в нативную структуру (рис. 21-7б).

Все это очень напоминает кристаллизацию. При температуре отвердевания можно получить большой хороший монокристалл, но очень-очень медленно. Небольшое понижение температуры ускоряет процесс. А большое приводит к появлению маленьких кристалликов, а не одного большого хорошего монокристалла.

Хочу отметить, что сходство термодинамики самоорганизации белков с фазовыми переходами первого рода экспериментально установлено в работах П. Л. Привалова, а сходство кинетических аспектов этих явлений — в компьютерных экспериментах Е. И. Шахновича и А. М. Гутина и в экспериментальных работах А. Фершта и его группы.

■ Область наиболее быстрого сворачивания исследовалась Шахновичем и его сотрудниками на простых компьютерных моделях белковых цепей. Эти работы как бы моделировали сворачивание белка в «нативных» (или даже еще более благоприятствующих сворачиванию) условиях. Они показали, что характерное время максимально быстрого сворачивания растет с длиной цепи много медленнее, чем то следует из формулы (21.2), оценивающей время сворачивания в точке термодинамического равновесия нативного и развернутого белка. А именно, компьютерный эксперимент показал, что время максимально быстрого сворачивания растет с числом звеньев цепи  $N$  вовсе не как  $\exp(N^{2/3})$ , как то должно быть в точке равновесия, а медленнее — как  $N^6$  для «случайных» цепей и даже как  $N^4$  для цепей, усиленно отобранных на предмет максимально быстрого сворачивания (и обладающих поэтому очень большой щелью между энергией самой стабильной структуры и энергиями прочих компактных укладок цепи).

Все это еще раз подчеркивает обсуждавшуюся на прошлой лекции сильную зависимость времени сворачивания белка от условий опыта, и прежде всего от стабильности нативной и конкурирующих с ней структур белка.

■ Здесь у вас естественно возникает вопрос: «Что будет, если у цепи не одна структура отделена от прочих большей энергетической щелью, а, скажем, две?»

Ответ: «Если обе эти структуры стабильны по сравнению с клубком, первой из них свернется та, к которой ведет лучший (с немного более низким барьером) путь сворачивания». Однако, если эта структура хоть немного менее стабильна, чем другая, последует очень медленный (так как по пути придется разворачивать метастабильную структуру, см. рис. 21-7б) переход в наиболее стабильную форму. Этот переход похож на полиморфные переходы в кристаллах (вспомним «оловянную болезнь» — переход белого олова в серое: эта «болезнь» порой уничтожала, при наступлении холодов, целые склады оловянных пуговиц). По-видимому, таких «полиморфных» белков должно быть мало (теоретические оценки показывают, что аминокислотные последовательности, кодирующие одну выделенную стабильную укладку цепи, редки, а кодирующие сразу две такие укладки — редки в квадрате).

Однако есть основания полагать, что некоторые белки, в частности, серпины (есть такие ингибиторы сериновых протеаз) и прионы (они вызывают «бешенство коров») ведут себя именно так. В последнем случае

полиморфизм осложнен агрегацией (слипанием), которой способствует (или которой наводится) «новая»,  $\beta$ -структурная форма белка. В ней, как показал недавно К. Добсон и его сотрудники, белковая цепь уложена в  $\beta$ -сэндвичи, упакованные в протофиламенты, укладываемые в длинные амилоидные нити (кстати, сходные амилоиды, трубчатые или ленточные, образуют и самые обычные белки, например, лизоцим и миоглобин; на их фоне прионы выделяются, кажется, только своей инфекционностью). В любом случае, возникает  $\beta$ -структурный агрегат.

При агрегации растет эффективная длина цепи фолдона  $N$ , что (согласно сделанным выше оценкам) делает возникновение новой фазы очень медленным. Но зато раз возникший зародыш новой формы агрегированного белка вытягивает в агрегат и переводит в эту новую форму все остальные белковые молекулы (по существу — так же, как при «оловянной чуме»). Эта-то агрегация прионов в клетках мозга и приводит в конце концов к летальному исходу при «коровьем бешенстве», почесухе и родственных им болезнях с огромным инкубационным периодом. Последние — на 2011 г. — данные как будто показывают, что болезнь вызывает не результат агрегации прионов, т.е. не огромные амилоидные нити, а активные «зародыши», возникающие по ходу этой агрегации. Говорят, к таким болезням — болезням с огромным инкубационным периодом — относится и старческий маразм... И вообще, разные амилоиды, которые склонны накапливаться в пожилых людях (или интермедиаты образования этих амилоидов), сейчас рассматриваются как возможная причина многих старческих болезней — хотя до сих пор неясно, действительно ли их огромный инкубационный период связан с медленностью образования амилоидных фибрилл *in vitro*. Однако уже более или менее ясно, что такие болезни являются одной из разновидностей «конформационных болезней», существование которых было давно предположено В. Е. Бычковой и О. Б. П.

Впрочем, есть и более простой, «мирный» и хорошо изученный лабораторный пример того же рода: водорастворимый аминокислотный полимер поли (лизин) в пробирке, при  $\text{pH} > 10$  и температуре 20–50 °С, быстро (за миллисекунды) переходит из клубка в  $\alpha$ -спираль, а потом — гораздо медленнее — из  $\alpha$ -спирали в  $\beta$ -структуру. Последний переход сопровождается агрегацией и может занимать часы, недели и более, его время экспоненциально растет с падением стабильности  $\beta$ -структуры; об этом мы уже говорили, обсуждая кинетику образования  $\beta$ -листов.

- В заключение обратимся еще раз к энергетическому ландшафту белковой цепи (см. рис. 21-1) и посмотрим, как самая стабильная структура автоматически пролагает к себе быстрые, преодолевающие парадокс Левинтала, пути сворачивания.

Через очень холмистый (скорее — даже скалистый) ландшафт (см. рис. 21-1) («скалы» соответствуют мощному, но короткодействующему отталкиванию при столкновении кусков цепи) от каждого энергетического минимума широко расходится «воронка» (или, точнее, сеть) более или менее гладких «дорожек». Эти дорожки соответствуют, в частности, разным путям последовательного (рис. 21-2) сворачивания структуры, соответствующей рассматриваемому минимуму. На таких дорожках происходит фазовое расслоение: из клубка выделяется компактная глобула, составляющая часть нативной, и потому относительно стабильная. Именно это делает дорожку гладкой, так как потеря энтропии одной фазой (клубком) тут же почти полностью (*почти* — из-за малой, но энергетически напряженной области раздела фаз) компенсируется энергией другой фазы (глобулы). На всех этих дорожках не происходит перестройки уже свернувшейся части глобулы, и потому на них нет больших ухабов (а маленькие — не помеха, если температура не слишком низка). Двигаясь по этим дорожкам к энергетическому минимуму, молекула выигрывает энергию, но, одновременно, фиксируя свою укладку, теряет энтропию (см. рис. 21-3). Чем глубже энергетический минимум — тем круче ведущие к нему дорожки, тем легче преодолевать, идя по ним, «энтропийное сопротивление», возникающее при фиксации укладки скатывающейся в энергетический минимум цепи. Сила этого сопротивления, поскольку оно энтропийное, пропорциональна температуре. При очень высокой температуре оно выталкивает цепь из всех энергетических минимумов, так что цепь не сворачивается вообще. При очень низкой температуре энтропийное сопротивление мало — цепь скатывается в любой ближайший минимум, долго из него выбирается, снова куда-то скатывается и никак не может дойти до глобального энергетического минимума (см. рис. 21-7б). И, наконец, при оптимальной для сворачивания температуре энтропийное сопротивление преодолевается на путях только к одному, самому глубокому минимуму — куда цепь и попадает достаточно быстро (см. рис. 21-7а).

- Все изложенные соображения о компенсации энергией энтропии по пути самоорганизации и о том, что это решает парадокс Левинтала, применимы и к образованию нативной структуры белка не только из клубка, но и по другим сценариям — из расплавленной глобулы или через напоминающие ее интермедиаты. Но там сделать какие-либо количественные оценки гораздо сложнее. К тому же хочу подчеркнуть, что эксперимент *не* показывает принципиального, многопорядкового ускорения самоорганизации белка при образовании расплавленных глобул. Поэтому я и ограничился только более простым случаем образования нативной структуры прямо из клубка.

**Внутренний голос:** А не решается ли парадокс Левинталя просто тем, что, продвигаясь к энергетическому минимуму, молекула, теряя энтропию, одновременно обретает энергию? Зачем нужно фазовое расслоение, о котором Вы столько говорите? Может быть, достаточно простой, равномерной, без расслоения на фазы, конденсации клубка?

**Лектор:** Простота такого объяснения, если рассматривать его как общее объяснение, коварна. Прежде всего, равномерная конденсация клубка не обеспечивает перехода «все-или-ничего». Об этом говорилось в конце 17-й лекции. Поэтому гипотеза о равномерной конденсации клубка не может объяснить наблюдаемое на опыте сворачивание и разворачивание белка в окрестности точки сосуществования его денатурированной и нативной форм. Значит, она не может решить проблему сворачивания белка в общем случае. Однако такая гипотеза, возможно, применима при экстремальных условиях, в которых нативная глобула намного (на величину порядка энтальпии плавления белка) стабильнее, чем денатурированный белок.

**Внутренний голос:** А не может ли решать проблему сворачивания белка тот стадийный механизм, о котором Вы говорили в одной из предыдущих лекций?

**Лектор:** Простота такого объяснения, как общего, также коварна. Для того чтобы стадийный механизм помогал сворачиванию белка (для иллюстрации, см. Задачу 19.1), необходимо, чтобы каждая стадия его сворачивания была бы необратима – иначе все опять сведется к перебору всех вариантов укладки цепи... Значит, он может помочь сворачиванию лишь тогда, когда нативная глобула много стабильнее, чем клубок. Но он не может помочь наблюдаемому на опыте сворачиванию белка в окрестности точки сосуществования его денатурированной и нативной форм. Значит, стадийный механизм тоже не может решить проблему сворачивания белка в общем случае, хотя он может помочь сворачиванию белка тогда, когда нативная глобула на много  $kT$  стабильнее, чем развернутая цепь.

- Рассказанное подчеркивает значимость развития «нового взгляда» на сворачивание белка — взгляда, базирующегося не на дискретных интермедиатах сворачивания, а на природе энергетических поверхностей или «ландшафтов», определяющих ход самоорганизации. Этот процесс отличается от более знакомых реакций, происходящих с малыми молекулами, (1) тем, что структура белка возникает как результат множества слабых взаимодействий; (2) тем, что ход сворачивания сопровождается большими изменениями конфигурационной энтропии; (3) тем, что по ходу него возникает расслоение «нативной» и «денатурированной» фаз.

В последнее время возникает понимание того, что внешне совсем разные сценарии экспериментального поведения сворачивающегося белка возникают из-за того, что различные взаимодействия дают разный вклад в общую свободную энергию молекулы при разных условиях. Это объединение идей дает тот скелет, в рамках которого экспериментаторы и теоретики могут работать вместе для получения более детального понимания особенностей процессов сворачивания.

В частности, сейчас интенсивно исследуется связь между ходом процесса сворачивания и множественностью состояний сворачивающихся белков в клеточной системе, что приводит к возможности и необходимости рассматривать более сложные картины сворачивания, включающие не только внутри-, но и межмолекулярные взаимодействия, типа взаимодействия зародышей сворачивания с шаперонами или другими белками, ДНК и РНК.

- Понимание явления самоорганизации белка — это гораздо больше, чем вызов таланту и интеллекту исследователей. Самоорганизация или, как говорят, сворачивание белка представляет собой мост между законами физики и результатами эволюционного отбора, составляющего сущность биологических процессов. В течение долгих лет мы могли обсуждать этот вопрос только в очень общих терминах, подчеркивая биологическую важность такого сложного физического процесса. Теперь мы начинаем понимать, что сворачивание — это физический процесс, происходящий только в системах, которые были биологически отобраны (вспомните «энергетическую щель») именно для того, чтобы обеспечить эффективное прохождение этого процесса.

Физика белка основывается на двух фундаментальных экспериментальных фактах: 1) что белковые цепи (по крайней мере, не слишком длинные и не слишком модифицированные против биосинтеза) способны спонтанно формировать свою трехмерную нативную структуру в подходящем окружении, и 2) что нативные состояния белков отделены от их развернутого состояния термодинамическим переходом типа «все-или-ничего». Последнее, в частности, обеспечивает надежность функционирования белков и минимальное содержание частично свернутых и потому склонных к агрегации молекул.

Все, о чем уже рассказывалось на этих лекциях, наводит на мысль, что, по-видимому, биологическая эволюция отбирает только те последовательности, которые способны сворачиваться в хорошо определенные трехмерные нативные структуры. Дело в том, что именно это определяет стабильность выполнения белком своей функции. А для этого нужны цепи, у которых энергия одной из структур существенно ниже, чем энергия у альтернативных укладок цепи. И именно цепи с такими последо-



вательностями способны к переходу типа «все-или-ничего» при сворачивании и разворачивании. Усиленная стабильность нативных структур, по-видимому, связана с плотной упаковкой цепи в глобулу (хотя пока еще не вполне ясно, какие конкретные требования накладывает на белковые цепи их способность к плотной упаковке). Примечательно, что «отобранные» по этому признаку цепи удовлетворяют требованию укладки в стабильную пространственную структуру одновременно с требованием сворачиваться быстро. Эта последняя способность, по-видимому, помогает еще не полностью свернутым полипептидным цепям избежать конкурирующих процессов межмолекулярной агрегации. Способность стабильных структур сворачиваться быстро решает давно отмеченное противоречие между кинетическим (по способу осуществления) и термодинамическим (по результату) выбором нативной структуры.

Последовательности глобулярных белков выглядят довольно «случайно», а их вторичная и третичная структура, как мы видели, несет много черт, типичных для всех более или менее стабильных укладок случайных сополимеров. Однако способность полипептидных цепей белков к быстрому и воспроизводимому сворачиванию в определенную пространственную структуру никак не является характерной для случайных последовательностей. Основной чертой «белковоподобных» аминокислотных последовательностей, чертой, определяющей все их физические свойства, является повышенная стабильность нативной структуры этих последовательностей, т. е. существование большой щели между энергией нативной структуры и минимальной энергией неверно свернутых глобулярных структур. Размер этой щели в реальных белках пока не определен; однако существующие оценки показывают, что для «белковоподобного» поведения обычного домена (из 100–200 аминокислотных остатков) достаточно относительно небольшой щели, в 10–20 ккал/моль. И уже довольно ясно, что необходимое повышение стабильности нативной структуры может начинаться с совсем малой, случайно образовавшейся щели и вестись путем постепенного эволюционного отбора подходящих для стабильного (а потому стабильно работающего) белка точечных мутаций.

- Сворачивание белков — это, наверное, самый простой пример биологического морфогенеза, который может происходить только в системе, эволюционно отобранной для того, чтобы позволить ему совершаться. Тем не менее, сворачивание белков имеет все физические характеристики, типичные для сложного процесса, подчиняющимся законам статистической механики. В частности, сворачивание глобулярных белков включает в себя нуклеацию, которая вообще типична для термодинамических переходов типа «все-или-ничего», т. е. для фазовых переходов первого рода. Путь

сворачивания белка, по-видимому, не уникален (и положение ядра сворачивания в белковой глобуле — тоже). Наблюдаемое порой у мутантных белков изменение ядер и интермедиатов сворачивания (при сохранении нативной структуры белков) показывает, что пути самоорганизации находятся под менее строгим контролем селекции, чем нативная пространственная структура белка: ведь разные пути могут вести к одной цели, хоть и с несколько разной скоростью. [Правда, в малой скорости сворачивания может крыться опасность: как показали К. Добсон и Л. Морозова-Роше, «недо-свернутые» и «расплавленные» белки склонны к агрегации, занимающей порой (в отсутствие охраны со стороны шаперонов!) всего миллисекунды и конкурирующей с «правильным» сворачиванием белка, или к образованию амилоидных фибрилл — а все это может вести к «конформационным болезням», о которых уже говорилось.] Сворачивание белка возникает как результат статистической флуктуации в структуре его развернутой цепи, которая (при «физиологической» температуре, pH и т. д.) приводит к переходу цепи в нативное состояние, т. е. в структуру с наиболее низкой энергией. Таким способом единственная, точно определенная структура белка возникает при сворачивании его цепи из статистического ансамбля развернутых или частично свернутых цепей.

- Все сказанное выше еще раз подчеркивает как связь между структурой белка, его способностью к самоорганизации, его стабильностью, так и роль эволюционного отбора в формировании и поддержке этих свойств белка.

Легко представить себе (но куда труднее доказать), что, зацепившись за случайно немного пониженную энергию одной из структур полипептида, считанного со «случайного» куска ДНК (причем структуры, способной, пусть очень плохо, но все же делать что-то хорошее для клетки), — эволюционное давление привело, путем мутаций и отбора все лучше и лучше сворачивающихся цепей, к возникновению последовательности, обладающей этим редким свойством — свойством спонтанной самоорганизации, и что оно же позволило развиваться специфической функции этого «прото-белка» (причем развиваться способом, чувствительным к контролю обратных связей в сложном и переполненном разными молекулами клеточном окружении).

Впрочем, такой способ возникновения новых белков — из случайных последовательностей — не играл, видимо, уже сколько-нибудь заметной роли на наблюдаемых (т. е. не самых ранних) этапах биологической эволюции. Сейчас все «новые» белки имеют своих «предков». Как уже говорилось, считается, что сейчас в основе возникновения новых белков лежат два механизма. Это, во-первых, дупликация гена «старого» белка, которая позволяет одному гену, практически не меняясь, поддерживать «старую»

функцию, а другому — спокойно мутировать, дрейфуя (при помощи селекции) к измененной, «новой» функции. Во-вторых, это — объединение отдельных белков в многодоменные белки, способные осуществлять более сложные реакции и поддающиеся более тонкому (из-за физического взаимодействия доменов) регулированию. Отмечу, что многодоменная структура особенно характерна для многоклеточных, высших (по сравнению с бактериями и одноклеточными эукариотами) организмов.

Однако возникновение новых белков путем дупликации и рекомбинации генов и их (белков) естественный отбор на полезность в организме никак не отменяет того, что некоторые («бездефектные») белковые архитектуры, стабильность (а значит, и способность к самоорганизации) которых совместима с большим разнообразием кодирующих их последовательностей, предоставляют эволюции и отбору гораздо больший простор, чем другие («дефектные») белковые архитектуры. Конечно, естественный отбор может способствовать появлению даже самой маловероятной архитектуры белка — если только именно она нужна для какой-то его (белка) полезной (для организма) функции. Ведь способствовал же отбор появлению таких маловероятных структур, как глаз или мозг... Однако, как мы увидим, функция белка, во всяком случае — водорастворимого глобулярного белка, обычно весьма слабо связана с его общей архитектурой. Поэтому не столь удивительно (хоть и весьма примечательно) то, что структуры таких белков обычно похожи на те, получение которых можно было бы ожидать при выборке стабильных укладок случайных сополимеров аминокислот, т. е. то, что «принцип множественности», по всей видимости, действует и на современном этапе биологической эволюции белковых структур.

## ПРЕДСКАЗАНИЕ И ДИЗАЙН БЕЛКОВЫХ СТРУКТУР

### *Лекция 22*

Потребность в предсказании структур белков по их аминокислотным последовательностям. «Опознавание» белковых структур и функций по гомологии последовательностей. Профили первичных структур семейств белков. Ключевые районы и функциональные сайты белковых структур. Выделение стабильных пространственных структур белковой цепи. «Шаблоны» белковых структур. Мы всегда вынуждены судить о предсказываемой структуре белка только по части взаимодействий, действующих в его цепи. Результат: вероятностные предсказания. Взаимодействия, стабилизирующие и разрушающие вторичную структуру полипептидов. Расчет вторичных структур не-глобулярных полипептидов. Предсказание вторичной структуры белков.

С тех пор как стало понятно, что аминокислотная последовательность белковой цепи определяет ее пространственную структуру, возникла проблема предсказания этой структуры по последовательности аминокислотных остатков в белковой цепи.

Чем вызвана потребность в предсказании белковых структур, кроме чисто интеллектуального интереса: удастся это сделать или нет? Тем, что экспериментально пространственную структуру белка определить куда труднее, чем его аминокислотную последовательность. А поток вновь полученных последовательностей все возрастает — в ходе выполнения проектов по расшифровке полных геномов разных организмов, от бактерий до человека, они появляются сотнями ежедневно. Понимание же механизма действия белка, подбор искусственных ингибиторов или активаторов к нему, и часто даже просто определение того, чем он занимается в клетке, настоятельно требуют знания его пространственной структуры...

И, конечно же, интерес к предсказанию пространственных структур белков подогревается воспоминанием о том, какое решающее значение имело предсказание строения двойной спирали ДНК для понимания всего генетического механизма.

■ Сейчас известно уже порядка сотни тысяч белковых последовательностей. Но «всего» для нескольких тысяч из них определены (рентгеном или ЯМР) их пространственные структуры. При этом многие из недавно определенных последовательностей просто считаны с ДНК или РНК, т. е. никто не определил на опыте, чем занимаются сделанные из них белки (и, если речь идет о последовательностях, считанных с ДНК, экспрессируются ли такие белки в клетке вообще...).

Что же можно сказать о трехмерных структурах тех последовательностей (я их буду называть «новыми»), для которых эксперимент (рентген или ЯМР) еще не сказал своего слова?

■ Прежде всего возникает мысль о предсказании трехмерной структуры «новой» последовательности на основании *родственного сходства* или, как говорят, «гомологии» ее первичной структуры с какими-то из «старых» последовательностей, пространственное строение коих уже расшифровано.

Опыт показывает, что даже не очень сильного сходства последовательностей достаточно для очень хорошего сходства пространственных структур: часто говорят, что пространственная структура более консервативна, чем аминокислотная последовательность (или что код, определяющий пространственную структуру белка по первичной, вырожден).

Установление гомологии последовательностей — действительно, очень мощный метод выяснения сходства структур и функций (причем не только белков, а и фрагментов ДНК, и РНК, но я буду говорить о белках).

Может быть, стоит упомянуть, что термин «гомология последовательностей» не вполне однозначен. Один пример. Предположим, что после иницирующего кодона белка в ген встраивается еще один нуклеотид. Старая и новая ДНК (и РНК) будут почти тождественны, их опознают как «гомологи» (что правильно, в строгом смысле: у них общее происхождение). В то же время, старая и новая аминокислотные последовательности, из-за сдвига рамки считывания, — совсем разные, между ними не углядишь никакой «гомологии» (что, строго говоря, неверно: их происхождение — общее; однако это заключение верно в смысле, что пространственные структуры старого и нового белка ни в чем не сходны).

Ниже я употребляю термин «гомология» в узком смысле сходства аминокислотных последовательностей, так как наша цель — найти сходство пространственных структур.

Установление гомологии и, далее, сходства пространственных структур не представляет труда только для достаточно близких, как на рис. 22-1, последовательностей.

	1	10	20
Человек, шимпанзе	<b>GD</b> VEKGKK <b>IF</b> IMKCSQCHTV...		
Свинья, корова, овца	<b>GD</b> VEKGKK <b>IF</b> VQKCAQCHTV...		
Курица, индюк	<b>GD</b> VEKGKK <b>IF</b> VQKCSQCHTV...		
Морская собака (акула)	<b>GD</b> VEKGKK <b>IF</b> VQKCAQCHTV...		
Личинка мясной мухи	GVPA <b>GD</b> VEKGKK <b>IF</b> VQRCAQCHTV...		
Ржа (грибок)	GFED <b>GD</b> AKK <b>GA</b> R <b>IF</b> KTRCAQCHTV...		
Рапс, цветная капуста	ASFDEAPP <b>GN</b> SKAGE <b>KIF</b> KTKCAQCHTV...		

Рис. 22-1. Гомологичные аминокислотные последовательности N-концевых фрагментов цитохромов *c* различных митохондрий и хлоропластов эукариотов. Жирным шрифтом выделены остатки, идентичные оным в человечесем (human) белке, подчеркнуты — сходные с ними. Выравнивание аминокислотных последовательностей взято из [6]. Столь высокогомологичные белки имеют очень сходную пространственную структуру

Труднее, когда белки в семействе сильно варьируют (рис. 22-2).

	○ ○ ○	◇◇◇ ○	○ ◇	● ○ ○ ○
<i>E. coli</i>	MLKQVE <b>IF</b> TDG <b>S</b> CLGN---PGGGY <b>GA</b> ILRYRGRE <b>KTF</b> SAGYTR <b>T</b> TNNRMELMAAIVALEAL			
Yeast	YNKSMNYVC <b>DG</b> SSFGNGTSSSRAGY <b>GA</b> YFEGAP <b>E</b> ENISPLL <b>S</b> GAQ <b>T</b> NNRAEIEAVSEALKKI			
MMLV	PDADH <b>TW</b> YTDG <b>S</b> LLQ---EGQR <b>K</b> AGA <b>AV</b> T <b>T</b> ET <b>E</b> VIWAKAL <b>P</b> AGT <b>S</b> AQ <b>R</b> AELIAL <b>T</b> QAL <b>K</b> -L			
RSV	PV <b>P</b> GP <b>T</b> VF <b>TD</b> ASS <b>ST</b> H---KGVVV <b>W</b> REG <b>P</b> RW-- <b>E</b> I <b>K</b> E <b>I</b> AD <b>L</b> GA--SVQ <b>Q</b> LE <b>A</b> RA <b>V</b> AM <b>A</b> LL-L			
HIV	IV <b>G</b> A <b>E</b> TF <b>Y</b> VD <b>G</b> AAN <b>R</b> ---TK <b>L</b> G <b>K</b> AG <b>Y</b> VT <b>N</b> -KGR <b>Q</b> K <b>V</b> VL <b>T</b> N <b>T</b> -- <b>T</b> N <b>Q</b> K <b>T</b> EL <b>Q</b> AI <b>Y</b> L <b>A</b> L <b>Q</b> D-			
	= : : = : :	: :	= : :	= = = : : = =
	← β1 →	← β2 →	← β3 →	← α1 →

Рис. 22-2. Аминокислотные последовательности N-концевых фрагментов рибонуклеаз H бактерии (*E. coli*), эукариота (дрожжи, yeast), и трех разных вирусов. В данном случае выравнивание последовательностей делалось «вручную», без компьютера, так, чтобы не допустить их разрывов (см. «← →») внутри α- и β-структурных участков. Жирным шрифтом выделены остатки, идентичные в трех и более из этих пяти последовательностей. Черными точками (над последовательностями) отмечены остатки активного центра, светлыми кружками и ромбами — остатки, вовлеченные в два гидрофобных ядра этого белка. Внизу отмечены остатки, совпадающие (=) и сходные (: ) у последовательностей из RSV и HIV (помещенных в двух нижних строках выравнивания), а также указана вторичная структура рассматриваемых белков. Картинка, с небольшими изменениями, взята из [7]

В этом случае на помощь приходит компьютер. Разработано множество программ, ищущих гомологии; с ними можно работать по Интернету. Назову только самые популярные из этих программ: BLAST, PSI-BLAST, HMMer, а также программа Смита-Ватермана. Все они строят выравнивание (alignment) последовательностей, добиваясь наибольшего сходства между ними. При этом за повышение сходства часто приходится платить «разрывом» последовательностей (см. знаки «←→» на рис. 22-2). Потом программа оценивает сходство выровненных последовательностей, и сообща-

ет: (1) гомологичны ли они (т. е. связаны ли они генетическим родством) и (2) как выглядит наилучшее выравнивание этих последовательностей. Правильно установить родство последовательностей можно и при невысоком их сходстве, но — внимание! — при этом часто не удается установить их правильное (вытекающее из сопоставления их пространственных структур) выравнивание; я еще к этому вернусь.

■ Разные программы по-разному оценивают, чего стоит совпадение остатков, чего — сходство, чего — несовпадение, чего — начало разрыва, чего — каждый дополнительный остаток в разрыве. Цены совпадений, сходства и несовпадений извлекаются из статистики аминокислотных замен в очевидных гомологах, а штрафы за разрывы подбираются авторами так, чтобы удовлетворительно выделять белки, сходство которых уже известно из других данных (и отбраковывать заведомо не-гомологичные белки). Оптимизированные параметры потом «зашиваются» в программу. Пользуясь программой, обычные люди («пользователи») обычно смутно знают, что «хорошо» согласно этой программе, что «плохо», а просто говорят: «установлено, что гомология последовательностей составляет 25 %», имея в виду, что 25 % выровненных остатков совпали друг с другом.

Встает вопрос — свидетельствуют ли эти 25% о сходстве последовательностей? Вообще говоря, это зависит от длины выравнивания (чем длиннее, тем скорее) и от числа и длины разрывов, сделанных при выравнивании (чем меньше, тем скорее). Хорошие программы обычно указывают вероятность того, что наблюдаемое сходство могло быть получено случайно. Однако всегда поучительно сравнить, *пользуясь той же программой*, и заведомо *несходные* последовательности. И тут выясняется, что «гомология» заведомо *несходных* белков (рис. 22-3) часто составляет 10–15 %, иногда 20 %, и порой (если — с разрывами в последовательностях) даже 25 %!

```
rop    --MTKQEKALNMARFIRSQTLTLEKLNELDADEQADICESLHDHDADELYRSCLARFGDDGENL--
mjc    SGKMTGIVKWFNADKGFGITPDDGSKDVFVHFSAIQNDGYKSLDEQGQKVSFTIESGAKGPAAGNVTSL
```

**Рис. 22-3.** Выравнивание аминокислотных последовательностей *непохожих, не-гомологичных* белков [в данном случае —  $\alpha$ -спирального РНК-связывающего белка (*rop*) и  $\beta$ -структурного белка холодового шока (*mjc*)] часто дает 10–15 % совпадающих аминокислотных остатков [в данном примере — 10 остатков (см. жирный шрифт) из 69, т. е. 14,5 %]. Выравнивание сделано программой BLAST; оно не потребовало разрывов в последовательностях. Таким образом, выравнивание заведомо негомологичных белков *rop* и *mjc* дает, даже и без разрывов в последовательностях, примерно ту же долю совпадающих остатков, что и выравнивание N-концевых половин RSV и HIV рибонуклеаз (см. рис. 22-2)

Эти цифры меняются от программы к программе. Однако накопленный опыт показывает, что тогда, когда «хорошая» (по общему мнению) программа выравнивает последовательности так, что совпадают свыше 30–35 % остатков, выявленной гомологии, родству последовательностей можно смело доверять (с оговоркой: при длине сравниваемых последовательностей свыше 50, а лучше — 100 остатков). Правда, надо учитывать что 30–35 % гомологии между последовательностями, верно (как правило) свидетельствуя об их родстве, позволяют правильно наложить друг на друга только 70–80% их пространственных структур, давая неверное предсказание о наложении остальных 20–30 %. А для того, чтобы верно проследить структуру 95 % главной цепи «нового» белка, нужно, чтобы его гомология с белком с известной структурой достигала 40–50 %.

Если же сходство пары последовательностей не превышает 10–15 %, то их родство обычно нельзя обнаружить: такое сходство находится на уровне «случайного шума» (что, однако, *не* является доказательством, что белки *не* похожи, *не* гомологичны — я к этому еще вернусь). А от 15 до 25 и даже до 30 % простирается «сумеречная зона»: *кажется*, что белки гомологичны, но кто поручится?.. (Да и сходство между выравниванием последовательностей и наложением пространственных структур при гомологии в 15–25 % невелико.)

К сожалению, как я уже сказал, все эти пороги не вполне одинаковы у разных программ (и у разных режимов их работы), а к программе они, эти характеризующие «порог доверия» цифры, обычно не прилагаются (кроме программ типа BLAST, где приводится оценка статистической достоверности полученного сходства; правда, все эти цифры есть в исходных статьях, но кто их читает...). Поэтому я бы рекомендовал, прежде чем доверяться любой такой программе, проверить ее (именно *ее*, и именно в используемом Вами режиме) на известных вам белках примерно той же длины (и сходных, и *несходных*) и понять, «что такое хорошо и что такое плохо» (другой вариант: прочесть исходную статью...).

Больше всего все эти оценки достоверности и *недостоверности* найденного сходства «плавают» от программы к программе из-за того, что разные авторы по-разному оценивают «штраф» за разрыв последовательности. Если штраф столь высок, что запрещает все разрывы, то случайно выбранные 20-буквенные (белковые) последовательности сходны на уровне всего 5 %. Если же этот штраф положить *нулевым*, т. е. позволить делать любые разрывы «бесплатно», то случайно выбранные белковые (и вообще 20-буквенные) последовательности дают сходство на уровне 30–35 % (а ДНКовые и РНКовые, 4-буквенные — на уровне 65 %)! Опыт показывает, что оптимальное отделение «похожих» от «непохожих» белковых последовательностей достигается, когда начало разрыва последовательности штрафуются в цену двух или трех дополнительных совпадений аминокислотных остатков, а за удлинение разрыва платится примерно 1/20–1/100 этой цены за каждый дополнительный остаток в разрыве.

Я умышленно не говорю ничего о математике, лежащей в основе алгоритмов поиска гомологий. Это нас увело бы слишком далеко. Хочу, однако, произнести



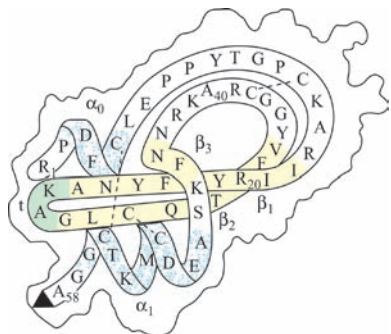


конформаций белковой цепи. Здесь, с весьма переменным пока успехом, приходится прибегать к конформационным расчетам и другим методам *гомологического моделирования*, на которых я останавливаться не буду.

Кроме того, опыт показывает, что даже полная идентичность фрагментов аминокислотных последовательностей не всегда гарантирует одинаковость их пространственных структур. Так, иногда одни и те же последовательности даже из 6–10 остатков входят в  $\alpha$ -спираль в одних белках и в  $\beta$ -структуру — в других.

- Перейдем к методам предсказания пространственной структуры «новых» последовательностей, не имеющих видимой гомологии с уже расшифрованными белками.

К ним сейчас относится около половины «новых» последовательностей. Поэтому о пространственной укладке многих последовательностей, получаемых в ходе выполнения генетических проектов, мы не можем догадаться по их гомологии с белками уже известными — она или слишком слаба для обнаружения, или отсутствует. Тут-то и возникает настоятельная потребность в основанном не на гомологии, а на физике решении задачи предсказания пространственной структуры, а, в перспективе, и функции белка, по его аминокислотной последовательности (рис. 22-5).



**Рис. 22-5.** Схема первичной и пространственной структуры маленького белка (панкреатического ингибитора трипсина). Ход главной цепи изображен на фоне общего контура молекулы; выделены  $\alpha$ -спирали,  $\beta$ -тяжи, резкий поворот цепи (t) и цистеиновые мостики (— —). Так как белок сворачивается сам собой, то, в принципе, все это можно предсказать по одной лишь первичной структуре белка. Боковые группы здесь не показаны, но, в принципе, и их конформации тоже можно предсказывать

- Сразу скажу: сейчас нет надежных и точных методов предсказания структуры белков, не имеющих гомологов с уже известной структурой.

Причин тому (кроме технической: огромного, даже для современных компьютеров, объема вычислений) — видимо, две: (а) ограниченная точ-

ность энергетических оценок и прочих параметров, на которых базируется теоретический расчет белковых структур, и (б) сравнительно малая (и потому уязвимая для всех неточностей расчета) разность энергий между одной «правильной» и многими «неверными» конформациями белковой цепи. Речь идет, собственно, о сравнительно малой ширине уже обсуждавшейся «энергетической щели». Если ширина щели — около 10 ккал/моль, а энергия плавления белка — около 100, то для верного предсказания (я позволю себе опустить доказательство) нужна 90 %-ная корреляция энергетических расчетов с экспериментом — что пока, видимо, недостижимо. Я хочу особо подчеркнуть последнее обстоятельство, т. е. малую ширину щели, — оно резко отличает ситуацию в белках (и в РНК) от ситуации в ДНК — к большому огорчению людей, занявшихся предсказаниями структур белков по их аминокислотным последовательностям. В ДНК комплементарное спаривание нуклеотидов наблюдается по всей длине двойной спирали, а оно обеспечивается тем, что одна цепь по своей первичной структуре строго комплементарна другой. Это и дает огромный свободно-энергетический выигрыш «правильной» структуры ДНК над всеми «неправильными». Такой же «комплементарный код» спаривания нуклеотидов есть и в РНК, но здесь комплементарные участки значительно короче, чем в ДНК, а «код» не всегда строго соблюдается. А в белках **нет** никакого однозначного «кода» слипания остатков: здесь не наблюдается ни какой-либо строгой попарной комплементарности длинных последовательностей контактирующих аминокислотных остатков, ни следующего из нее огромного энергетического преимущества «правильной» укладки цепи...

Однако существующие методы дают все более полную и надежную информацию о возможном строении (или, чаще, о возможных вариантах строения) белка или, особенно, его структурных элементов.

В прагматическом аспекте, вопрос о предсказании трехмерного строения белка сейчас ставится следующим образом: похожа ли пространственная структура рассматриваемой белковой последовательности на какую-либо из уже известных пространственных белковых структур? Если **да**, то **как** рассматриваемая последовательность вписывается в эту структуру? Если **нет**, можем ли мы указать какие-либо характерные детали пространственной организации рассматриваемой последовательности?

Я знаю несколько случаев, в том числе и из собственной практики, когда ответы на эти вопросы существенно способствовали экспериментальному определению структуры и /или функции изучаемой аминокислотной последовательности, помогли планировать белково-инженерные эксперименты и так далее. Впрочем, должен сказать, что я бы с большим интересом прочел хороший обзор о практическом применении предсказаний белковых структур...

■ Перейдем теперь к *методам* предсказания белковых структур. Рассматривая эти методы, я буду уделять особое внимание тем идеям, и в основном физическим идеям, что лежат в их основе.

Есть две стратегии предсказания белковых структур. Согласно первой стратегии, белковая структура ищется как результат кинетического процесса сворачивания.

Согласно второй, она ищется как структура с минимально возможной для данной цепи свободной энергией (или, что то же самое, как наиболее вероятная структура).

В принципе, по-видимому, обе эти стратегии могут привести к правильному результату, так как самая стабильная структура белковой цепи, несмотря на опасения Левинтала, образуется достаточно быстро (этому вопросу была посвящена отдельная лекция, и сейчас я позволю себе его не касаться).

Важно, однако, что, рассматривая предсказание белковых структур, мы можем отвлечься от путей сворачивания и рассматривать только результат — стабильную структуру белка. Тем более, что первая стратегия — стратегия имитационного моделирования процесса сворачивания белка — пока систематического прогресса не дала: уж очень сложны все расчеты, и к тому же — все же весьма приблизительно все потенциалы взаимодействий. А все спекуляции о специфических (например, иерархически) принципах сворачивания белка (которые выглядели так привлекательно и сулили существенно упростить поиск нативной структуры) оказались неверными. Поэтому моделирование сворачивания белков мы рассмотрим позже.

Вторая стратегия оказалась более успешной.

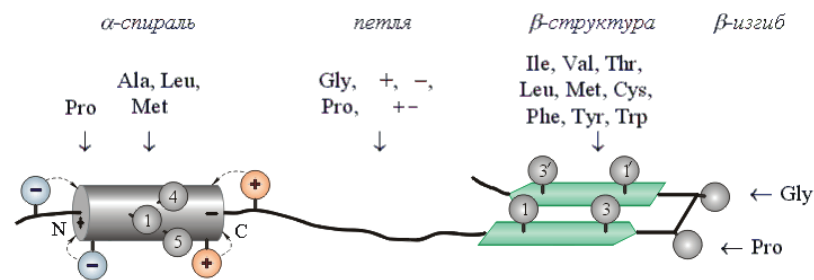
Оценка стабильности (или, что то же, вероятности образования) разных структур базируется либо на потенциалах взаимодействий, либо (чаще) на частотах встречаемости разных структурных элементов и разных взаимодействий в белках. Здесь нет принципиальной разницы: как мы уже знаем, квази-Больцмановская статистика белковых структур определяется потенциалами взаимодействий. Поэтому я буду говорить о «потенциалах», даже если их оценка восходит к белковой статистике, а не к физическому эксперименту. Взаимодействия с молекулами воды обычно не рассматриваются в явном виде; вместо этого используются потенциалы взаимодействий, опосредованных водой (например, гидрофобных взаимодействий).

■ Начнем с определения стабильной вторичной структуры белка по его аминокислотной последовательности;  $\alpha$ -спирали и вытянутые  $\beta$ -участки — важнейшие элементы белковых глобул, во многом определяющие, как вы помните, их общую архитектуру (см. рис. 22-5).

Забудем пока о том, что белковая цепь уложена в твердую глобулу: слишком сложно рассчитывать вторичную структуру одновременно с третичной. Можем ли мы предсказать вторичную структуру цепи по ее аминокислотной последовательности *до* предсказания третичной структуры? Оказывается — да, в основном да, пусть и не абсолютно точно.

Прежде всего, какие аминокислотные остатки будут стабилизировать отдельно взятую вторичную структуру, например,  $\alpha$ -спираль, а какие будут разрушать ее?

Эксперименты дают прямой ответ на этот вопрос. Я имею в виду огромную работу по оценке  $\alpha$ - и  $\beta$ -образующих способностей аминокислотных остатков, проведенную в группах Шераги, Фасмана, Болдвина, Фершта, Серрано, ДеГрадо, Кима и в ряде других групп (и, в частности, у нас — В. Е. Бычковой и О. Б. П.). Кроме того, богатый (хорошо совпадающий с физико-химическим экспериментом и дополняющий его) материал дает статистика аминокислотных последовательностей  $\alpha$ - и  $\beta$ -структур в белках. Рисунок 22-6 суммирует важнейшие (те, которые стоит запомнить) из оценок, полученных всеми этими методами.



**Рис. 22-6.** «Шаблоны» для опознавания  $\alpha$ -спирали, петли,  $\beta$ -структуры и  $\beta$ -изгиба; выделены те остатки, которые стабилизируют их или их отдельные части. «+» означает все положительно заряженные аминокислоты, «-» — все отрицательно заряженные, «+ -» — все аминокислоты с диполем в боковой цепи. Показан также стабилизирующий  $\alpha$ - и  $\beta$ -структуру порядок чередования гидрофобных групп в цепи (см. нумерованные группы). Такого типа чередование приводит также к образованию гидрофобных и полярных поверхностей  $\alpha$ -спиралей и  $\beta$ -тяжей

Здесь надо, однако, сразу сказать, что все закономерности, наблюдающиеся в структурах белков, носят вероятностный, *частотный* характер. Несмотря на многочисленные попытки, не удалось выделить никакого четкого «кода» белковых структур, т. е. ничего похожего на то четкое А-Т и Г-С-спаривание нуклеотидов, которое свойственно двойной спирали ДНК. Впрочем, надо признать, что уже в РНК — с их разнообразным,

в отличие от ДНК, репертуаром пространственных структур — спаривание нуклеотидов происходит далеко не столь однозначно...

Подавляющее большинство из полученных экспериментальных и статистических оценок можно легко понять, исходя из физики и стереохимии аминокислот. Мы уже говорили об этом на одной из прошлых лекций.

Так, Pro не любит входить ни в  $\alpha$ -спираль (кроме ее N-концевого витка), ни в  $\beta$ -структуру. Почему? Потому, что у него нет NH-группы, и он не может завязывать соответствующие  $\alpha$ - и  $\beta$ -структуре водородные связи; а на N-конце спирали NH-группа таких связей и не должна завязывать, — и Pro там встречается часто, тем более что у Pro один из двух углов вращения в главной цепи (угол  $\phi$ ) уже фиксирован пролиновым кольцом в подходящем положении. По тем же причинам часто встречается Pro и на N-конце изгибов.

А вот, например, Ala стабилизирует  $\alpha$ -спираль, а Gly разрушает и ее, и  $\beta$ -структуру, и способствует образованию нерегулярных участков. С чем связана эта разница? С тем, что область конформаций, т. е. область допустимых углов  $\phi$ ,  $\psi$  для Gly в клубке гораздо больше, чем для Ala, в то время как допустимые конформации для обоих этих остатков в  $\alpha$ -спирали примерно совпадают.

По аналогичным причинам  $C^\beta$ -разветвленные остатки — Val, Ile и Thr — больше стабилизируют  $\beta$ -структуру, где их боковые группы имеют 3 разрешенных поворотных изомера, чем  $\alpha$ -спираль или клубок, где они имеют только 1 изомер (при каждом значении  $\phi$  и  $\psi$ ).

Вообще же гидрофобные группы склонны несколько чаще входить в  $\alpha$ - и  $\beta$ -структуру, где они могут «бесплатно» (все уже оплачено водородными связями!) слипаться в гидрофобные кластеры (см. рис. 22-6), чем в клубок, где они этого делать не могут. А вот боковые группы с диполями, особенно в короткой боковой цепи, больше склонны входить в нерегулярные участки цепи, где они могут завязывать дополнительные нерегулярные водородные связи, чем в  $\alpha$ - и  $\beta$ -структуры, где доноры и акцепторы водородных связей уже насыщены внутрицепными связями.

Влияние аминокислотных остатков на вторичную структуру можно не только объяснить, но и *a priori* оценить теоретически. Так, еще до получения экспериментальных оценок, в начале и середине 1970-х гг., А. В. Ф. и О. Б. П. предсказали, что отрицательно заряженные остатки должны стабилизировать N-конец спирали, притягиваясь к его положительному парциальному заряду, и дестабилизировать ее C-конец, отталкиваясь от отрицательного парциального заряда спирального диполя. Положительно же заряженные остатки должны действовать в прямо противоположном направлении. При этом потенциал каждого такого взаимодействия должен, по теоретической оценке, составлять 1/4 или 1/3 килокалории. Что и было подтверждено экспериментально.

■ Зная, какие аминокислотные остатки стабилизируют середину спирали, какие — ее N-конец, а какие — C-конец, мы получаем нечто вроде «шаблона» спирали.

Шаблон  $\alpha$ -спирали можно описать так: если начало фрагмента белковой цепи обогащено отрицательно заряженными группами, да еще там стоит Pro; если середина этого фрагмента обогащена остатками Ala, Leu и Met, а Pro и Gly там нет; и если его C-конец содержит положительно заряженные боковые группы — перед нами  $\alpha$ -спиральный участок. Кроме того, для стабильности  $\alpha$ -спирали полезен, а для ее включения в глобулу — просто необходим определенный (см. рис. 22-6) порядок чередования гидрофобных групп в цепи: этот порядок способствует слипанию этих групп и, кроме того, приводит к образованию на спирали сплошной гидрофобной поверхности, необходимой для ее прилипания к глобуле. Важность правильного чередования боковых групп для формирования вторичной структуры была показана В. И. Лимом.

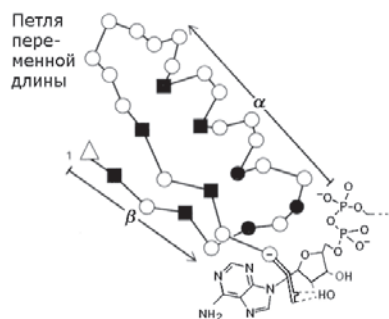
То есть «шаблон» качественно описывает аминокислотную последовательность, подходящую для образования  $\alpha$ -спирали. Чем лучше аминокислотная последовательность удовлетворяет этому шаблону, тем вероятнее спираль в данном месте цепи. Такое же описание — «шаблон» — можно составить и для участков, пригодных для образования  $\beta$ -структуры. А также для участков, особо пригодных для образования  $\beta$ -изгибов и петель.

Более того, «шаблоны» можно применять и для описания кусков цепи, образующих более сложные структуры, например, для описания  $\beta$ - $\alpha$ - $\beta$ -суперспиралей, состоящих из двух параллельных  $\beta$ -участков и  $\alpha$ -спирали между ними (рис. 22-7). Такие структуры типичны для доменов, связывающих нуклеотиды. Особо важную роль в «шаблоне» играют так называемые «ключевые позиции», которые могут быть заняты только строго определенными аминокислотными остатками, например, Gly: только этот остаток может находиться в конформации с  $\phi \approx 60^\circ$ , недоступной всем остальным остаткам. Отметим, что «шаблоны» могут содержать, помимо структурной, и функциональную информацию (см. последний остаток в нуклеотид-связывающей  $\beta$ - $\alpha$ - $\beta$ -субъединице на рис. 22-7).

■ Но вернемся к расчету вторичной структуры. Для начала, рассмотрим вторичные структуры в «развернутой» цепи, т. е. не учитывая взаимодействия между ними.

Зная вклады отдельных взаимодействий в стабильность  $\alpha$ -спиралей и  $\beta$ -шпилек, мы можем рассчитать их свободную энергию в любом месте цепи. Вот уже 30 лет мы используем для этого нашу программу ALB. В одном из режимов (“unfolded chain” — «развернутая цепь») она позволя-

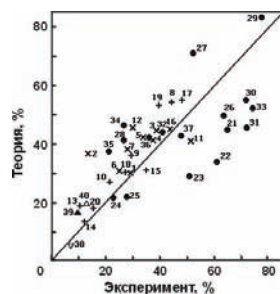




**Рис. 22-7.** «Шаблон» суперспирали  $\beta$ - $\alpha$ - $\beta$ , связывающей нуклеотиды. Квадратики отмечают позиции, обычно занимаемые сравнительно небольшими гидрофобными остатками (Ala, Ile, Leu, Val, Met, Cys): это — гидрофобное ядро суперспирали  $\beta$ - $\alpha$ - $\beta$ . Черные кружки — ключевые позиции, занимаемые только Glu: здесь находятся резкие повороты цепи. Светлый треугольник отмечает первую позицию мотива  $\beta$ - $\alpha$ - $\beta$ , где обычно находится положительно заряженная или дипольная боковая цепь. В последней (-) позиции мотива  $\beta$ - $\alpha$ - $\beta$  находится аминокислота Asp или Glu, связывающая лиганд (нуклеотид). Рисунок, с небольшими упрощениями, взят из R. K. Wierenga et al., *J. Mol. Biol.* (1986) **187**:101–107

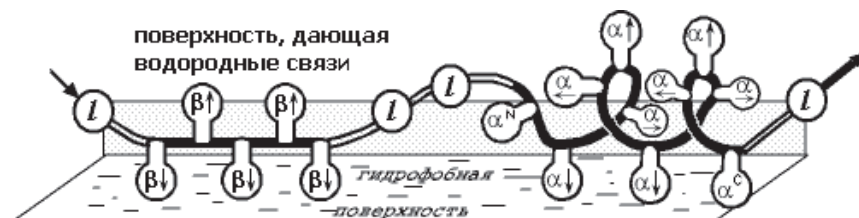
ет рассчитывать вероятность образования  $\alpha$ - и  $\beta$ -структур в каждом месте цепи и среднее их содержание в полипептидах и в неглобулярных (еще) белковых цепях, причем при разной температуре, ионной силе и pH раствора. Потом результат можно сравнить с опытными данными, например, с содержанием вторичной структуры в тех же полипептидах, судя по КД спектрам. Рисунок 22-8 показывает, что теоретический расчет неплохо совпадает с опытом.

**Рис. 22-8.** Теоретическая (вычисленная программой ALB — unfolded chain) и экспериментально найденная спиральность нескольких десятков пептидов при температуре 0–5 °C и разной ионной силе и pH раствора. Рисунок, с небольшими изменениями, взят из A. V. Finkelstein, A. Y. Badretdinov & O. B. Ptitsyn, *Proteins* (1990) **10**:287–299



- Переходя к расчету и предсказанию вторичной структуры глобулярных белков, необходимо учесть, что здесь к взаимодействиям, существующим в несвернутых цепях, добавляется взаимодействие каждого участка цепи с глобулой, *строения которой мы не знаем*. Точнее, мы не знаем ее де-

тального строения, но знаем, что участки цепи как-то примыкают к гидрофобному ядру белка. В простейшем приближении взаимодействие с ядром можно аппроксимировать взаимодействием с «гидрофобным озером», на котором плавает белковая цепь (рис. 22-9).



**Рис. 22-9.** Флуктуирующая вторичная структура белковой цепи (здесь:  $\beta$  —  $\beta$ -тяж,  $l$  — петля,  $\alpha$  —  $\alpha$ -спираль) на поверхности, имитирующей белковую глобулу («модель плавающих бревен»). Эта поверхность состоит из гидрофобного «озера», «берег» которого способен завязывать водородные связи с  $\beta$ -участками. Модель учитывает разное чередование обращенных к поверхности «озера» ( $\downarrow$ ), от поверхности ( $\uparrow$ ) и вдоль нее ( $\rightarrow$ ,  $\leftarrow$ ) боковых групп в разных вторичных структурах, а также эффекты на N- и C-концах спирали, объединенные и приписанные ее, соответственно, N- и C-концевым остаткам ( $\alpha^N$ ,  $\alpha^C$ ). Рисунок, с небольшими изменениями, взят из O. B. Ptitsyn & A. V. Finkelstein, *Biopolymers* (1983) **22**:15–25

Зная из опыта силу гидрофобных взаимодействий, а из стереохимии  $\alpha$ - и  $\beta$ -структуры — мотивы чередования в цепи боковых групп, глядящих в одну и ту же сторону и способных, следовательно, одновременно взаимодействовать с гидрофобной поверхностью, мы можем сосчитать вероятность образования  $\alpha$ -спирали и  $\beta$ -структуры в каждом месте белковой цепи. Этим также — но уже в режиме «глобулярная цепь» (“globular chain”) — занимается программа ALB (кстати, к ней можно обращаться по Интернету через <http://phys.protres.ru/resources.html>).

Вероятности рассчитываются при комнатной температуре. Почему при комнатной? Не лучше ли выделять только *самое лучшее* по энергии расположение  $\alpha$ - и  $\beta$ -структур в цепи, т. е. рассчитывать все вероятности при 0 K?

Можно сказать, конечно, что нас интересует стабильная структура белка при (примерно) комнатной температуре, а она поддерживается, в том числе, и температурно-зависимыми гидрофобными взаимодействиями.

Но еще важнее то, что вероятности, рассчитанные именно при такой температуре (300 K), лежащей чуть ниже нормальной температуры плавления белка (350 K), наилучшим образом позволяют отделить более вероятные  $\alpha$ - и  $\beta$ -участки, в которых мы можем быть более или менее уверены, от тех «менее вероятных», в которых мы уверенными быть не можем.

Ведь, в сущности, мы стараемся предсказать вторичную структуру белка *только по части* тех взаимодействий, что ее держат в действительности: мы знаем (и то приблизительно) внутренние взаимодействия в этой структуре, но *не* знаем (или можем лишь крайне грубо оценить) те, что действуют между рассматриваемым куском цепи и остальной глобулой. А они очень мощны.

То есть мы находимся в том же положении, как если бы пытались предсказать, будет ли данный остаток внутри белка или на его поверхности, зная только его собственную гидрофобность — и больше ничего. И мы знаем, что такая задача имеет ответ, но ответ не точный, а *вероятностный*. Этот ответ содержится в наблюдаемой статистике распределения остатков между нутром и поверхностью глобулы. Мы уже знаем, что она выглядит так:

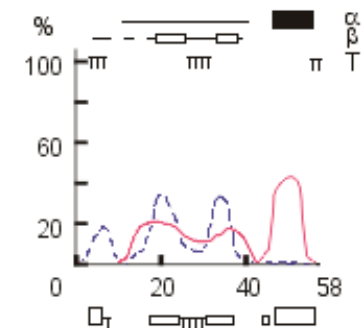
$$\frac{\text{ВЕРОЯТНОСТЬ\_ВНУТРИ}}{\text{ВЕРОЯТНОСТЬ\_НА\_ПОВЕРХНОСТИ}} \sim \exp(-\text{ГИДРОФОБНАЯ\_СВОБОДНАЯ\_ЭНЕРГИЯ}/kT_c), \quad (22.1)$$

где  $T_c$  — характерная температура белковой статистики, лежащая — как мы тоже помним — несколько ниже температуры плавления белка.

И — возвращаясь к предсказаниям белковых структур — с той же температурой  $T_c$  мы можем *вероятностно* предсказывать существование рассматриваемой вторичной структуры, зная только ее собственную свободную энергию.

И если при этом мы получим на выходе уверенное, лишенное или почти лишенное флуктуаций предсказание — в добрый час! — значит, принятые нами в расчет взаимодействия сами по себе обеспечивают достаточно стабильную структуру белка...

- Рисунок 22-10 иллюстрирует расчет вторичной структуры в панкреатическом ингибиторе трипсина при помощи программы ALB-globular chain. Так как такой расчет принимает во внимание взаимодействия внутри вторичных структур и их гидрофобное взаимодействие с подложкой, моделирующей остальную глобулу, но игнорирует плотную упаковку, то он моделирует скорее вторичную структуру расплавленной глобулы. Но, как мы знаем, вторичные структуры нативной и расплавленной глобулы довольно близки. Результирующий график показывает, что рассчитанная вероятность флуктуирующей вторичной структуры белка, даже в самых структурированных местах, далека от 100 %. Однако, в согласии с опытом,  $\alpha$ -спирали преобладают в С-концевой части его цепи, а  $\beta$ -структура — в центре. При этом пик вероятности спирального состояния стоит над большей спиралью, а пики  $\beta$ -структурного — над большими  $\beta$ -структурными участками. По этим пикам (если они превышают определенный, эмпирически устанавливаемый уровень) и идет предсказание вторичной структуры глобулярного белка.



**Рис. 22-10.** Расчет вторичной структуры панкреатического ингибитора трипсина. Абсцисса — номер остатка в белковой цепи, ордината — вычисленная вероятность его  $\alpha$ -спирального (—) и  $\beta$ -структурного (---) состояния. Сверху приведены предсказания наиболее вероятных  $\alpha$ -спиралей ( $\alpha$ ),  $\beta$ -тяжей ( $\beta$ ) и изгибов (T). Черные прямоугольники — уверенное предсказание вторичной структуры, светлые — не столь уверенное предсказание, линии — данная вторичная структура в принципе не исключена, но маловероятна. Внизу изображено расположение в цепи нативных  $\alpha$ -спиралей (высокие прямоугольники),  $\beta$ -тяжей (низкие прямоугольники) и изгибов T

Предсказание вторичной структуры носит вероятностный характер не только в программе ALB, но и во всех остальных «предсказывающих» программах, даже когда эти программы нацелены, как то обычно бывает, на выделение одного, «лучшего» распределения  $\alpha$ - и  $\beta$ -участков в белковой цепи.

Точность распознавания  $\alpha$ -,  $\beta$ -структуры и нерегулярных петель белка по *одной* его аминокислотной последовательности составляет около 65 %.

Можно ли улучшить предсказание вторичной структуры? Да. В методе PHD Роста и Сандера или PsiPred Д. Джонса (очень их рекомендую; они доступны по Интернету) вторичная структура предсказывается не для *одной* аминокислотной последовательности, а (всегда, когда это возможно) для набора гомологов. В результате такого подхода случайные погрешности как бы сглаживаются и предсказание становится намного более точным (достигая, в среднем,  $\approx 75$  % вместо 63–65 %).

Итак — несмотря на свою ограниченную точность, предсказание вторичной структуры стало уже, по сути, рутинной процедурой исследования первичной структуры белка.

Хуже обстоит дело с гораздо более сложной проблемой, проблемой предсказания пространственной структуры белка по его аминокислотной последовательности. Об этом мы поговорим на следующей лекции.

## Лекция 23

Представление о подходах к предсказанию пространственных структур белков по их аминокислотным последовательностям. Базы данных по структурам белков. Предсказание общей укладки цепи отдаленных гомологов понижает неопределенности в опознавании структур белков. Структурная геномика и протеомика. Биоинформатика. Белковая инженерия и дизайн. Первые успехи в конструировании белков.

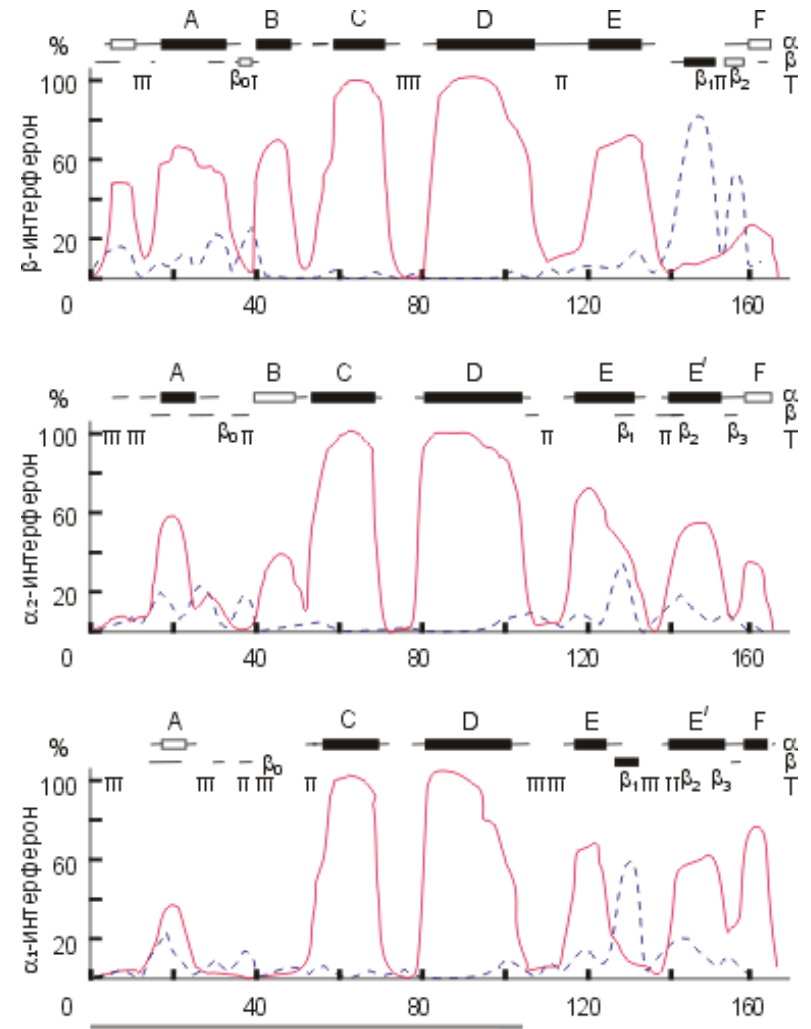
Перейдем теперь к новой, быстро развивающейся области — к работам по предсказанию пространственных укладок белковых цепей.

- Расчет пространственных укладок часто базируется на предварительном предсказании вторичной структуры белковой цепи. Хотя такой подход не является строго обоснованным, так как третичная структура может влиять на вторичную, иногда он приводит к успеху.

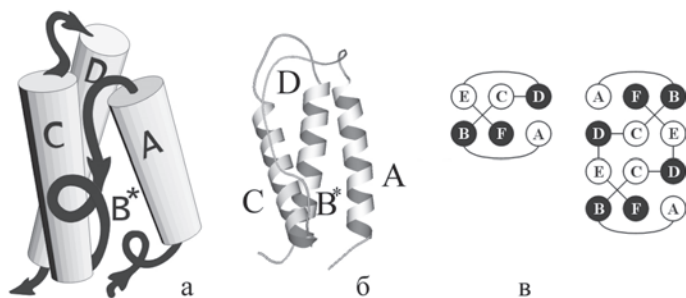
Рисунок 23-1 иллюстрирует расчет вторичной структуры в глобуле. Он был сделан в 1985 г. для интерферонов. Видно, что в них преобладают  $\alpha$ -спирали, особенно в N-концевой части цепи, где (как было известно из опыта) находится функциональный домен интерферона. В данном случае спирали эти предсказывались столь уверенно (а это бывает отнюдь не всегда) и столь одинаково в разных, не так уж и высоко гомологичных цепях, что мы попробовали в том же 1985 г. слепить домен из этих спиралей.

Полученный комплекс, пучок из трех больших спиралей и одной крошечной, показан на следующем рисунке (23-2а). Через пять лет после того, как было сделано это предсказание, была, наконец, получена рентгеновская структура интерферона  $\beta$  (рис. 23-2б), и рентгеновская структура N-концевого его домена довольно точно совпала с той, что была предсказана нами.

Структура N-концевого домена интерферона — одна из первых, более или менее успешно предсказанных до опыта структур белковых молекул.



**Рис. 23-1.** Расчет вторичной структуры нескольких интерферонов. Абсцисса — номер остатка в белковой цепи, ордината — вычисленная вероятность его  $\alpha$ -спирального (—) и  $\beta$ -структурного (---) состояния. Сверху приведены предсказания наиболее вероятных  $\alpha$ -спиралей ( $\alpha$ ),  $\beta$ -тяжей ( $\beta$ ) и изгибов (T). Черные прямоугольники — уверенное предсказание вторичной структуры, светлые — не столь уверенное предсказание, линии — данная вторичная структура в принципе не исключена, но маловероятна. Участок цепи, образующий N-концевой домен, подчеркнут внизу рисунка. Картинка взята из О. В. Ptitsyn, A. V. Finkelstein & A. G. Murzin, *FEBS Letters* (1985) **186**:143–148



**Рис. 23-2.** (а) Предсказание укладки цепи в N-концевом домене интерферонов, сделанное нами в 1985 г. [Ptitsyn, Finkelstein, Murzin, *FEBS Letters* (1985) v. 186, p. 143]. Три большие  $\alpha$ -спирали (A, C, D) изображены цилиндрами; кроме того, была предсказана возможность существования отдельного спирального витка В. (б) Рентгенографическая структура N-концевого домена (С-концевой домен — не изображен)  $\beta$ -интерферона, расшифрованного в 1990 г. [Senda et al., *Proc. Jpn. Acad. Sci., ser. B, Biophys. Biol. Sci.* (1990) v. 66, p. 77]. Рисунок (б) дан в той же ориентации, что и предсказанная модель (а), и на нем даны те же имена больших спиралей (A, C, D), под которыми они значатся на рисунке (а). Район В\* спиралеподобен, но не  $\alpha$ -спирален в интерфероне  $\beta$ , однако в родственном ему интерфероне  $\gamma$  там находится небольшая  $\alpha$ -спираль [Ealick et al., *Science* (1991) v. 252, p. 698]. (в) Топологии  $\beta$  (слева) и  $\gamma$  (справа) интерферонов по Ealick et al. Интерферон  $\gamma$  состоит из двух субъединиц. Обратите внимание на «обмен» С-концевыми кусками [спиралями (E, F)] между двумя субъединицами  $\gamma$  интерферонов

Имеет смысл подчеркнуть факторы, способствовавшие успеху предсказания пространственного строения интерферонов: сами по себе  $\alpha$ -спирали предсказаны здесь очень уверенно; и они одинаково расположены в отдаленно-гомологичных цепях, что позволяет доверять полученным предсказаниям. Однако столь уверенное и повторяющееся в гомологах предсказание вторичной структуры встречается на практике не часто. Более типичны некоторые разночтения, как в С-концевом домене интерферонов (см. рис. 23-1), для которого — именно вследствие этого, очень типичного небольшого разночтения в предсказаниях вторичных структур — однозначного предсказания укладки цепи получить не удалось.

И еще очень важная вещь для успеха предсказания строения интерферона. Когда мы искали наилучшую упаковку предсказанных  $\alpha$ -спиралей, мы могли рационально организовать перебор всех возможных упаковок, так как располагали априорной классификацией  $\alpha$ -спиральных комплексов (я об этом уже говорил) — она как раз была разработана к тому времени Мурзиным и А. В. Ф.

■ Проблему предсказания белковой структуры можно представить себе как проблему выбора — из множества возможных — той пространственной структуры, которая наиболее стабильна для данной аминокислотной последовательности. Откуда же можно взять «возможные пространственные структуры»?

Самый простой, прагматичный ответ — взять из Банка белковых структур, PDB. Тогда, в сущности, речь пойдет не о «предсказании», а об «опознавании» пространственной структуры белка. Но в этом есть свой плюс: «опознав» белок, мы можем надеяться одновременно опознать и его функцию, а ведь именно она обычно и интересует биолога.

В Банке белковых структур, конечно, нет еще всех возможных структур, но есть надежда, что там есть более половины всех существующих в природе мотивов укладки белковых цепей в домены. Эта надежда — ее обосновал Сайрус Чотиа — основана на том, что вновь расшифровываемые белки (точнее, их домены) все чаще и чаще оказываются похожими на уже расшифрованные. Согласно оценкам Чотиа, основанным на экстраполяции числа расшифрованных структур и числа разных мотивов укладки цепи, существует примерно 1500 мотивов укладки, большая часть которых (причем самых «популярных») уже известна.

■ Итак, предсказывая структуру белка, не имеющего видимой гомологии с белками уже расшифрованными, можно попробовать взять, одну за другой, все пространственные структуры из Банка, наложить (возможно, с некоторыми выпетливаниями, рис. 23-3, или даже с изменением порядка прохождения цепи через структурный каркас белка) цепь этого белка на каждую из них, и посмотреть, какая из этих пространственных структур даст — для нашей цепи — наибольший энергетический выигрыш. При этом мы должны разрешать цепи то идти по скелету структуры, то выпетливаться или «сокращать» имеющиеся в скелете выпетливания, если это увеличивает энергетический выигрыш.

Такой подход называют «методом протягивания» (threading method). Он был предложен А. В. Ф. и Б. А. Ревой в 1990 г. и — независимо, в более простом и более удобном, хотя и менее корректном физически, варианте — Д. Айзенбергом и его группой в 1991 г. Сейчас метод протягивания стал весьма популярным методом опознавания структур «новых» белков по их аналогии со «старыми».

В общем, работа по протягиванию напоминает поиск гомологии, только на этот раз «выравниваются» не две первичные структуры, «новая» и «старая», а «новая» первичная структура со «старой» пространственной структурой белкового скелета.





**Рис. 23-3.** Схема, иллюстрирующая идею «протягивания» исследуемой последовательности по Банку белковых структур. Жирная линия показывает те участки, где последовательность идет по скелету, пунктир — те, где она выпетливается

■ Прежде чем перейти к результатам, хочу подчеркнуть две принципиальные проблемы метода «протягивания». В сущности, аналогичные проблемы, в том или ином виде, возникают в любом «предсказательном» методе.

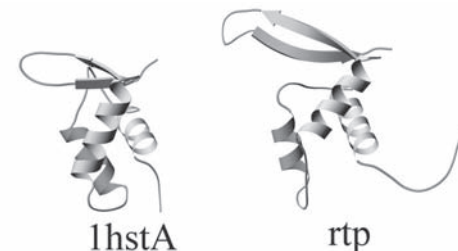
Во-первых, конформацию даже тех кусков цепи, что наложены на скелет, мы рассматриваем с большой погрешностью: ведь мы не рассматриваем конформации боковых групп (разных у «старой» и «новой» цепи), а именно они, боковые группы, в основном, и взаимодействуют. Далее, мы не рассматриваем конформации всех выпетливаний. Оценка показывает, что при протягивании мы рассматриваем примерно половину взаимодействий в белковой цепи, а вторую не рассматриваем вообще. Значит, опять мы вынуждены судить о структуре белка по части взаимодействий, действующих в его цепи. Значит, опять наши предсказания могут носить только вероятный характер.

Во-вторых, как перебрать все наложения и найти лучшее, или лучшие, среди них? Ведь их, возможных наложений, страшно много... Скажу коротко: мощные математические вычислительные методы для этого уже развиты, но рассказ об этом занял бы слишком много времени. Наверно, имеет все же смысл назвать эти методы по имени — может быть, кому-то из вас пригодится (но большинству, конечно, нет). Итак: динамическое программирование и его вариант — статистическая механика одномерных систем (цепных молекул) для расчета протягивания цепи через скелет; теория самосогласованного поля — для расчета действующего на цепь молекулярного поля в каждой точке скелета; стохастическая минимизация энергии методом Монте-Карло; а также — разные варианты метода ветвей и границ, и т. д.

И еще: интуитивный перебор вариантов! Он кажется старомодным в компьютерный век, но именно так А. Г. Мурзин предсказал структуру многих «новых» белков, очень часто выигрывая в эти своеобразные белковые шахматы у самых лучших компьютерных программ.

**Внутренний голос:** Предыдущая фраза была написана к первому изданию «Физики белка», в 2001 г. Забавно, что за прошедшие годы, несмотря на огромный прогресс компьютеров, ситуация с «белковыми шахматами» вовсе не изменилась кардинально (см. статью “Predicting protein structures with a multiplayer online game” в *Nature* (2010) **466**:756–770 и компьютерную игру Foldit, <http://fold.it/portal/node/>).

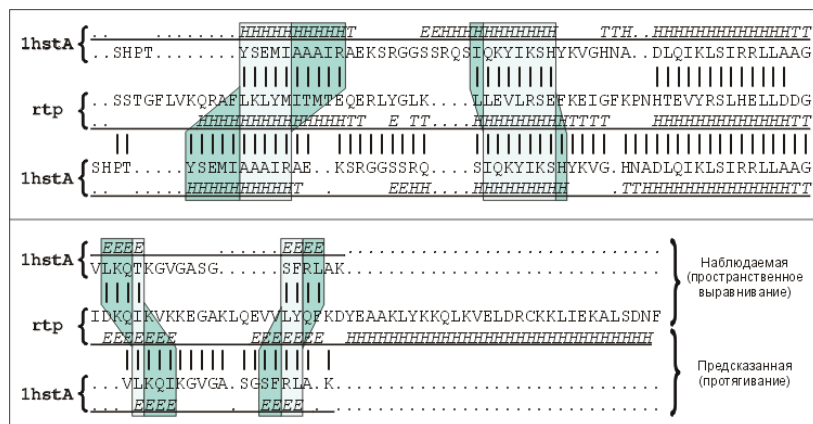
■ Как бы то ни было — какие же получаются результаты? Для примера я покажу предсказанную методом протягивания укладку цепи в белке, терминирующем репликацию. Протягивая цепь этого «нового» белка по всем известным трехмерным белковым структурам, Зиппл и его коллеги из Зальцбурга показали, что укладка гистона H5 наиболее «родственна» этой цепи (рис. 23-4).



**Рис. 23-4.** Гистон H5 цыпленка (1hstA, слева) и белок, терминирующий репликацию (rtp, справа); в последнем не показана С-концевая спираль, не имеющая аналога в гистоне H5. Предсказание сходства укладки цепи в этих двух белках сделано методом «протягивания» в работе Н. Flöckner, М. Braxenthaler, Р. Lackner, М. Jaritz, М. Ortner & М. J. Sippl, *Proteins* (1995) **23**:376–386; это предсказание сделано в ходе «слепого» тестирования предсказательных методов (CASP-1994). Средне-квадратичное отклонение между 65-ю эквивалентными С<sup>α</sup>-атомами в изображенных структурах — 2,4 Å

Таким образом они правильно (несмотря на очень низкое сходство аминокислотных последовательностей) опознали мотив укладки цепи белка — терминатора репликации. Правда, предсказанное (при помощи теоретического протягивания) наложение цепи этого белка на структуру гистона H5 заметно отличается (рис. 23-5) от того, что получается при прямом наложении трехмерных структур этих молекул. Это еще раз показы-

вает, что все погрешности, о которых я говорил по ходу лекций — погрешности в параметрах стабильности, незнание точной конформации петель и т. д. — не позволяют выделить *одну* уникальную наилучшую структуру, а позволяют в действительности только найти более или менее узкий набор лучших структур. Такой набор «лучших» пространственных структур можно выделить довольно надежно; а вот какая из них окажется, согласно расчету, «наилучшей» — это дело случая. Нативная структура найдется *где-то* в наборе «лучших» структур, она приблизительно соответствует предсказанной («наилучшей») — но это все, что можно реально сказать.



**Рис. 23-5.** Последовательность гистона Н5 (1hstA), выровненная относительно последовательности белка, терминирующего репликацию (rtp) согласно рентгеноструктурным данным («наблюдаемая», верхняя пара последовательностей), и она же, выровненная согласно предсказанию методом протягивания («предсказанная», нижняя пара). Серыми зонами показаны сдвиги, расхождения этих двух выравниваний. Они хорошо видны по разным сдвигам вторичных структур гистона Н5, приведенных в самой верхней и самой нижней строках в этих двух выравниваниях. Вторичные структуры записаны в подчеркнутых строках (обозначения: *H* —  $\alpha$ -спираль, *E* —  $\beta$ -тяж, *T* — изгиб цепи). Точками отмечены делеции, т. е. пропуски в первичных и вторичных структурах, внесенные при их выравнивании. Картинка, с минимальными изменениями ее формы, взята из статьи “Protein structure prediction by threading methods: evaluation of current techniques”. С. М. -R. Lemer, V. J. Rooman & S. J. Wodak, *Proteins* (1995) 23:337–355, в которой обсуждаются итоги CASP-1994

Я, конечно, привел только «хорошие» примеры — примеры удавшихся предсказаний. Примеров неудавшихся много больше, и это понятно — ведь, предсказывая, надо выбрать один или пусть несколько вариантов из бездны оцениваемых. Так что попасть рядом с такой бесконечно малой

целью даже в половине случаев (так сейчас работают наиболее успешные группы) — огромный успех.

■ Методы «протягивания» становятся сейчас рабочим инструментом для опознавания трехмерных структур белковых последовательностей. Чрезвычайно важно, что уже сформирован рецепт — делай так, так и так, и в результате ты получишь одну или несколько структур, среди которых, с довольно высокой вероятностью, будет структура, сходная с трехмерной укладкой рассматриваемой последовательности.

■ На этом фоне и на фоне прогресса в опознавании структур белков «по гомологии» (в особенности с привлечением биоинформатических «мегасерверов», опрашивающих серверы, работающие по разным программам, и суммирующих результаты их работы) — успехи в предсказании белковых структур путем имитации их сворачивания — пока выглядят, вообще говоря, несколько более скромно, хотя и здесь делается немало ярких и интересных работ.

Хочу особо отметить подход Д. Бейкера, в котором своего рода имитация сворачивания используется для поиска стабильной укладки цепи. Этот подход воплощен в комплексе программ Rosetta, успешно применявшемся и для предсказания структур белков, и для белкового дизайна, о котором речь пойдет в конце этой лекции. Вначале производится подбор гомологов (с известной пространственной структурой) к каждому короткому фрагменту цепи белка, структура которого предсказывается. Потом из пространственных структур этих фрагментов-гомологов складывается масса возможных пространственных структур рассматриваемого белка, а затем негодные структуры цепи отбрасываются энергетическими расчетами, а лучшие выделяются. Действуя таким образом, Бейкер смог не только «опознать» в PDB, но и *ab initio* предсказать несколько «новых», еще не наблюдавшихся укладок белковых цепей в ходе «слепого» тестирования предсказательных методов («CASP»). И еще необходимо отметить недавние работы групп В. Панде (*J. Am. Chem. Soc.* (2010) 132:1526–1528) и особенно Д. Шоу (*Science* (2010) 330:341–346), которые смогли проследить сворачивание развернутых цепей нескольких, пусть пока только крошечных белков в течение десятков микросекунд, и при этом не только показать, что они достигают своих нативных структур, но и верно оценить скорость их сворачивания.

■ Можно ли как-то устранить зловерное действие приближенности наших оценок взаимодействий в белках (для иллюстрации, см. Задачу 23.1) и улучшить опознавание и предсказание белковых структур? Да, если привлечь гомологи, т. е. предсказывать структуру белка уже не только по *его* аминокислотной последовательности, а по *набору* гомологичных последовательностей. Использование гомологов улучшило предсказания вторичных

структур белков. Этот же метод широко и успешно применяется при предсказании вторичных структур РНК. При опознавании пространственных укладок белковых цепей его тоже начали применять, при этом отдается предпочтение той укладке, что более или менее хороша для всех гомологов. Причем наиболее полезными оказываются гомологи с идентичностью ~ 40 %: они еще сохраняют почти тождественную пространственную структуру, но уже кодируют ее по-разному, что и позволяет опознать «сигнал» (структуру), несмотря на «шум» всех погрешностей предсказания. Правда, при этом мы получим не точную структуру данной цепи, а некоторую обобщенную, т. е. приближенную структуру, общую для всех гомологов.

■ Мой сегодняшний рассказ касался глобулярных белков. А что можно сказать о фибриллярных и мембранных?

В последовательностях фибриллярных белков — без всякого компьютера — просматриваются периодичности, характерные для их вторичных структур, причем они настолько очевидны, что опознание вторичных структур этих белков обычно не составляет труда. А на этой основе удается предложить несколько возможных вариантов их упаковки в фибриллу.

Предсказание строения мембранных белков менее развито, чем предсказание строения белков глобулярных: слишком мало структур мембранных белков мы знаем. Опознать внутримембранные части таких белков довольно легко: это — сплошные (точнее — почти сплошные) гидрофобные блоки, длина которых диктуется толщиной мембраны. Однако принципы формирования укладки этих блоков в белке еще недостаточно выяснены (хотя известно главное; напомним: у мембранных белков «внутри» структуры — не гидрофобное, как у глобулярных, а полярное; гидрофобные же группы контактируют с жиром мембраны). В общем, работы по предсказанию пространственного строения мембранных белков только начинаются.

■ Суммируем.

Предсказание трехмерных укладок белковых цепей, с точки зрения физики, является поиском самой стабильной укладки рассматриваемой цепи. Такое предсказание возможно в принципе, и порой оно приводит к вполне удовлетворительному результату. Однако оно всегда имеет сложности, связанные с недостатком точной информации о репертуаре возможных укладок белковых цепей и о многих взаимодействиях в белковой цепи, с погрешностями в энергетических параметрах, и т. д.

В результате однозначное и надежное предсказание укладки цепи, не имеющей гомолога с уже известной структурой, исключительно по ее последовательности пока что вряд ли возможно, но зато можно выделить несколько наиболее вероятных мотивов укладки цепи и — что очень важно — отбраковать огромное множество других укладок. И такие

предсказания могут быть уточнены с учетом дополнительной информации, в частности, о последовательностях цепей гомологичных белков и об их возможных укладках, об активных центрах и т. д.

■ В прагматическом аспекте, вопрос о предсказании трехмерного строения белка сейчас ставится скорее как задача *опознания*, чем предсказания. Она ставится так: похожа ли пространственная структура рассматриваемой белковой последовательности на какую-либо из уже известных пространственных белковых структур? Если **да**, то **как именно** рассматриваемая последовательность располагается в этой структуре?

Стратегия *опознания* ограничена (ведь не все возможные пространственные структуры белков уже известны, а опознать можно только известную), она обладает одним несомненным преимуществом перед априорным (и не ограниченным набором известных структур) *предсказанием*. Опознав структуру белка, мы часто можем догадаться, чем он занимается; а предсказание структуры — само по себе, без опознания сходства с чем-то уже известным — не дает никаких намеков на этот счет.

■ Надо отметить, что на физике базируется только часть методов, направленных на предсказание (точнее — на опознание) структур белков. Большинство же методов пытается опознать сходство данной последовательности (и, часто, ее окружения в геноме) с последовательностями (и геномным окружением) уже известных белков.

Все «опознавательные» подходы базируются на банках данных по структурам и последовательностям биополимеров. Они содержат такой гигантский объем информации, что сложилась и окрепла тенденция брать всю необходимую информацию именно из этих Банков: и набор возможных структур, и энергии самых разных взаимодействий в них, и семейства гомологов, и набор возможных функций, и годных для того активных центров. Это называется «**биоинформатикой**»; эта ветвь науки быстро развивается, и на это даются гигантские средства. Не меньшие средства расходуются на пополнение баз данных. Вы все знаете о проекте «геном человека», призванном установить первичную структуру всех человеческих генов, а также генов ряда других организмов (всего — порядка сотен тысяч генов). Этот проект уже завершен в той части, что касается секвенирования, т. е. определения нуклеотидной последовательности, человеческих хромосом. Несколько сложнее дело обстоит с выделением и аннотированием генов на этих хромосомах. Выделить ген на ДНК эукариота — задача нетривиальная и сложная: ведь (вспомните сплайсинг, и особенно альтернативный сплайсинг) приходится находить все те кусочки, из которых складывается ген, кодирующий белок. Однако люди надеются, что и эта задача будет решена через несколько лет.

Уже запущены следующие масштабные биоинформационные проекты:

Проект «**Протеомика**» призван опознать все белок-кодирующие гены и, далее, узнать, чем занимается каждый белок-кодирующий ген в контексте генома: опознать биохимическую функцию этого белка (кодирует ли ген оксиредуктазу? или нуклеазу? или?), его роль в клетке (вовлечен ли он в нуклеотидный метаболизм? или в регуляцию? или?), а также его связи с продуктами других генов.

Опознавание укладок белковых цепей, возможно, послужит в этом проекте одним из инструментов. Ведь, в конечном счете, именно функция интересует биолога, а структуры и последовательности обслуживают этот интерес.

Другой проект, «**Структурная геномика**», призван расшифровать (рентгеном или ЯМР) хотя бы по одному представителю из каждого семейства белков — всего этот проект должен дать, за время порядка 10 лет, порядка десятка тысяч разнообразных белковых структур (что обойдется в ~10 миллиардов долларов). Когда этот проект будет завершен, предсказание белковых структур *ab initio* окажется излишним — пространственные структуры интересующих вас последовательностей можно будет просто опознавать по гомологии.

И, наконец, гигантский проект «**Индивидуальные геномы**» — от геномов индивидуальных людей до геномов поразивших именно их опухолей, бактерий и вирусов...

■ В заключительной части этой лекции я хочу кратко рассказать о *белковой инженерии*, то есть о направленной модификации белков уже существующих, и о *белковом дизайне*, т. е. о конструировании новых белков.

Олигонуклеотидный синтез и техника рекомбинантных ДНК дали возможность получения генов белков, не существовавших в природе; рентген и ЯМР позволили увидеть трехмерные структуры белков; а мощные ЭВМ и компьютерная графика позволили вступить в интерактивный диалог с этими пространственными структурами — менять в них что-то и оценивать последствия этих изменений. Объединившись, эти методы стали, соответственно, «руками», «глазами» и «мозгом» новой области молекулярной биологии — белковой инженерии. Ее стратегическая задача — создание знаний и методов, позволяющих получать белки с наперед заданной функцией и структурой. Трудно переоценить перспективность этих исследований для конструирования новых лекарств, катализаторов, для схематехники и т. д.

Прицельный белково-инженерный эксперимент уже дал ответ на ряд фундаментальных вопросов. Показано, что за выбор пространственной структуры отвечают далеко не все детали плотной упаковки — структура белка выдерживает массу точечных мутаций; и вряд ли отвечают за этот

выбор петли: если заменить или вырезать из белковой цепи участок, кодирующий петлю, то «рана» на теле глобулы обычно затягивается. О том, что целенаправленное введение точечных мутаций дает массу информации об энергетике белковой глобулы и промежуточных структур на пути ее самоорганизации, мы уже говорили; а о том, как они применяются для исследования активности белка, — поговорим на следующей лекции.

Следует отметить и промышленное применение белковой инженерии в фармакологии и химии для получения, путем введения направленных мутаций, белков с повышенной (или пониженной — смотря, что требуется спросом) стабильностью и модифицированной каталитической активностью.

■ Сущность белково-инженерного подхода к модификации белка демонстрируется рисунком 23–6. Предположим, мы хотим ввести в структуру белка какой-то сайт или просто стабилизировать эту структуру (на рис. 23-6, в качестве простого примера, изображено введение сайта Asp — His, часто встречающегося в активных центрах белков). Тогда, во-первых, мы ищем в банке «старых» белков (структура которых уже установлена рентгеном или ЯМР) тот белок и те аминокислотные остатки в нем (на рисунке — Z и Y), которые, после замен Z→Asp и Y→His, могут дать искомый сайт с искомой конфигурацией Asp и His. Затем в гене выбранного «старого» белка мы заменяем кодоны, кодирующие Z и Y, на кодоны, кодирующие Asp и His, соответственно, — и получаем ген «нового» белка, содержащего искомый сайт (а после, конечно, надо еще проверить, работает ли этот сайт, и узнать — при помощи рентгена или ЯМР — получился ли он в нужной конфигурации).

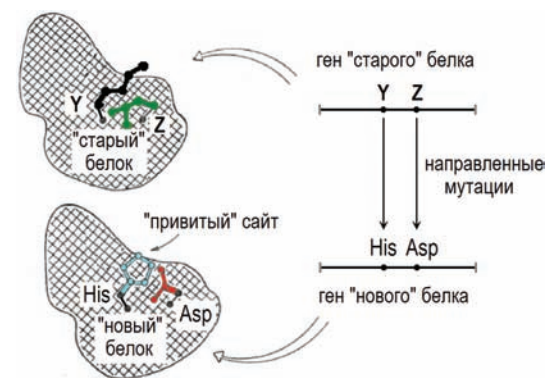


Рис. 23-6. Схема применения белковой инженерии для «прививки» к белку нового функционального сайта. Картинка, с некоторыми упрощениями, взята из: Финкельштейн, *Биополимеры и клетка* (1989) 5:89–93

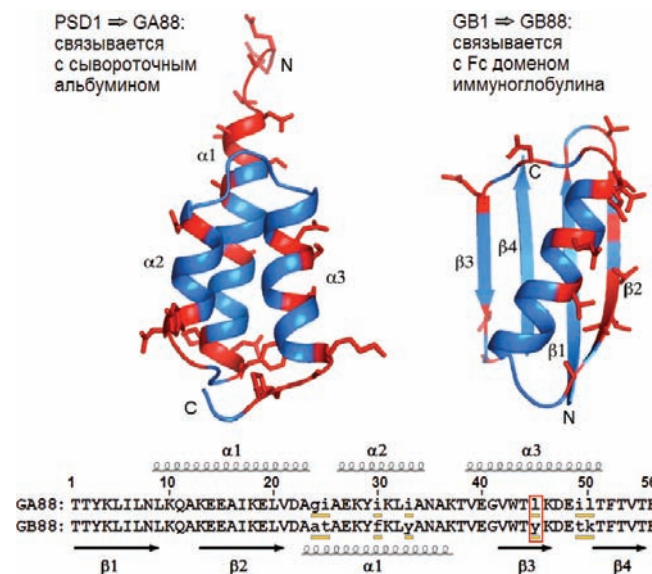


■ То, что я показал на рис. 26-6, — не более, чем простой учебный пример. А вот как была построена замечательная работа группы Дэвида Бейкера, в которой им удалось создать (Röthlisberger et al., *Nature* (2008) **453**:190–195) не существовавший в природе активный центр для не существовавшей в живых организмах химической реакции. В качестве таковой была взята «элиминация Кемпа» — перенос протона с одного отростка ароматической молекулы на другой. Выбор типа и конфигурация функциональных групп, способных катализировать такую реакцию, а значит — стабилизировать переходное состояние на ее пути, был сделан, исходя из физики и химии процесса элиминации (кстати, для этого катализа оказалась пригодной, наряду с рядом других, и определенная конфигурация рассмотренной выше диады Asp — His, лежащей в кармане из ароматических боковых групп). Далее, были выбраны (при помощи уже упомянутого комплекса программ Rosetta) те белки, к которым можно привить выбранную диаду в требуемом окружении и требуемой конфигурации. Затем эта прививка была сделана методами генной инженерии, белок был экспрессирован, и оказалось, что полученный искусственный фермент ускоряет элиминацию Кемпа в миллион раз по сравнению со скоростью ее спонтанного протекания (что очень и очень неплохо для начала, хотя некоторые природные ферменты ускоряют реакции в миллион миллионов раз); а рентгеноструктурный анализ показал, что привитой активный центр выглядит практически так, как того требовал теоретический дизайн. Итак, дело было сделано — был получен искусственный фермент с запланированной, но не существовавшей до того в природе активностью.

■ Белковая инженерия — это метод решения разных задач. Он использовался и для выяснения вопроса о том, может ли одна и та же (или почти одна и та же) аминокислотная последовательность кодировать две существенно разные укладки белковой цепи. Оказалось — может, если тщательно поработать над аминокислотной последовательностью.

Группа Орбана и Брайана стартовала от двух белков, PSD1 и GB1, имеющих равную длину цепи, но разную структуру и функцию (рис. 23-7), и разные последовательности, в которых совпадали всего 9 остатков из 58 (идентичность 15 % — это отвечает практически случайному сходству). Далее, пошагово вводя в эти последовательности сближающие их друг с другом мутации и на каждом шаге контролируя сохранение белками их первоначальных функций, Орбан и Брайан добились того, что сохранившие и свои функции, и свои структуры (Alexander et al., *Proc. Natl. Acad. USA* (2007) **104**:11963–11968; He et al., *Proc. Natl. Acad. USA* (2008) **105**:14412–14417) последовательности совпадали уже на 88 %, т. е. у них было 51

идентичных остатков из 58 (см. рис. 23-7). Дальнейшая работа над этими последовательностями уменьшила их различие — при сохранении различия их структур и функций — всего до 1 остатка (Alexander, He, Chen, Orban, Bryan, *Proc. Natl. Acad. USA* (2009) **106**:21149–21154)!



**Рис. 23-7.** Белки, сохранившие после введения множества мутаций свои функции и пространственные структуры, но приобретшие 88 %-ю идентичность последовательностей. Измененные остатки выделены цветом и скелетными моделями боковых групп на показанных в верхней части рисунка укладках белковых цепей. На показанных в нижней части рисунка последовательностях совпадающие аминокислотные остатки изображены заглавными буквами, а различающиеся — маленькими (и подчеркнуты). Рамкой обведен тот единственный остаток, который, в пределе, различает эти белки, сохранившие свои структуры и функции. Картинка, с некоторыми упрощениями и дополнениями, взята из He et al., *Proc. Natl. Acad. USA* (2008) **105**:14412–14417

В другой интересной работе (Gambin, Schug, Lemke, Lavinder, Ferreón, Magliery, Onuchic, Deniz, *Proc. Natl. Acad. USA* (2009) **106**:10153–10158) показано, что после целенаправленного введения нескольких мутаций белок, имеющий вид четырехспирального пучка, приобретает возможность находиться в другой конформации — также имеющей вид четырехспирального пучка, но с другой укладкой спиралей. При этом одни условия среды стабилизируют одну структуру, другие — другую, а в промежуточных условиях обе структуры сосуществуют одновременно.

Итак, почти одна и та же или даже совсем одна и та же (но очень хорошо отобранная!) аминокислотная последовательность может, в принципе, кодировать две существенно разные укладки белковой цепи.

■ Едва освоив введение мутаций в «природные» белки, белковая инженерия обратилась к дизайну, конструированию белковых молекул.

■ Задача дизайна — обратная по отношению к задаче предсказания структуры. Если при предсказании мы должны найти пространственную структуру, наиболее пригодную для рассматриваемой последовательности, — то при дизайне мы должны найти, сконструировать последовательность, годную для создания желаемой пространственной структуры.

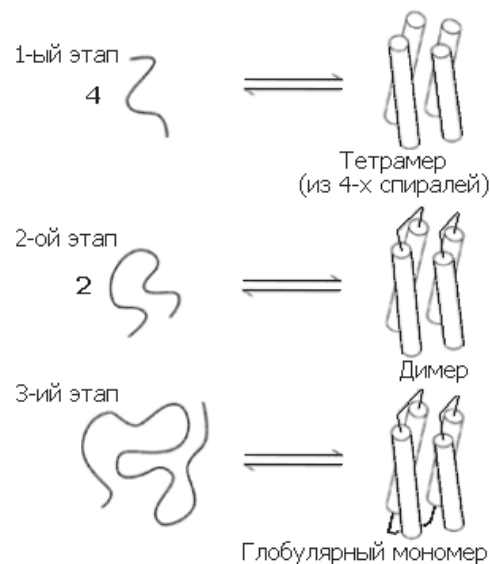
Вообще говоря, расчет искусственных конструкций может быть проще, чем предсказание структуры «натуральных» белков (расчет прочности «искусственной» башни проще, чем «натурального» дерева: она спроектирована так, чтоб поддаваться расчету!). А конструирование новых белков опирается на теорию белковых структур; их «строительными блоками» обычно служат последовательности, кодирующие мощные, внутренне стабильные и способные к эффективному слипанию  $\alpha$ - и  $\beta$ -участки.

Белковый дизайн был поставлен на повестку дня в конце 1970-х — начале 1980-х гг., когда возникла техника создания искусственных генов и техника точного химического синтеза белковых цепей. В конце 1980-х — начале 1990-х гг. при помощи прикидок, проб и ошибок были созданы простейшие белковые молекулы. Их архитектуры брались из природных белков, а их аминокислотные последовательности подбирались так, чтобы, не будучи гомологичными природным, стабилизировать эти архитектуры.

■ Первым был сделан — в группе ДеГрадо — четырехспиральный пучок (рис. 23-8). Дизайн проводился в тесном диалоге с экспериментом. Исследование окончательного варианта белка показало, что он спирален и глобулярен; и структура этого искусственного белка оказалась гораздо стабильнее, устойчивее к нагреву, чем структура любого «естественного» белка! Потом, правда, оказалось, что этот белок не плавился при нагревании потому, что был расплавленной глобулой с самого начала... Пришлось усилить твердую структуру этого искусственного белка, введя в него гистидины, связывающие ион. И вот этот-то белок-комплексон оказался уже твердым, как природные белки.

Долгое время все искусственные белки (кроме тех, что усиливались образованием ионных комплексов) представляли собой коллекцию отличных расплавленных глобул. Они обладали прекрасной вторичной структурой, они были очень компактны, но им не хватало твердости.

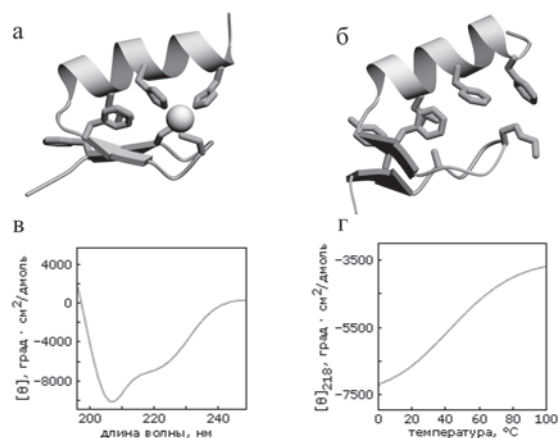
■ Вопрос: почему хотят сделать нормальный, «твердый» белок — а получается расплавленная глобула?



**Рис. 23-8.** Основные этапы дизайна четырехспирального пучка, сделанного в группе ДеГрадо. 1. Подбор коротких спиральных пептидов, способных слипаться в тетрамер. 2. Дизайн петель, сшивающих эти пептиды попарно, и отбор тех вариантов сшитых пептидов, что димеризуются. 3. Дизайн последней петли и отбор мономерного тетраспирального искусственного белка. Рисунок W. F. DeGrado взят, с его любезного разрешения, из [6]

Видимо, потому, что все знают, как сделать стабильные вторичные структуры (лейцины и аланины, отороченные глутаминовыми кислотами с N-конца и лизинами с C-конца — получится  $\alpha$ -спираль; побольше валинов, изолейцинов и треонинов — получится  $\beta$ -структура); и все знают, как заставить эти  $\alpha$ - и  $\beta$ -структуры слипаться: надо сделать на них сплошные поверхности из гидрофобных групп (рис. 23-9); и как заставить эти белки не агрегировать: надо сделать противоположные поверхности этих  $\alpha$ - и  $\beta$ -структур из гидрофильных групп. Но вот для чего рецептов пока никто не сделал — это для создания плотной упаковки боковых групп в гидрофобном ядре белка. Вот эта-то упаковка и не получается. А без нее при наличии слипающихся, но не плотно слипающихся вторичных структур получается лишь расплавленная глобула...

■ Чтобы справиться с недостаточной точностью методов расчета и дизайна, часто отходят от «рационального» конструирования и прибегают к введению в цепь множества случайных мутаций (точечных или блочных, но — массовых!) с последующим отбором имеющих «белковоподобную» активность вариантов (скажем, специфически с чем-нибудь связывающихся). Таким путем группе Н. Райтона (N. Wrighton) удалось сделать мини-белок (димер из двух  $\beta$ -шпилек), обладающий активностью гормона эритропоэтина. (Интересно, что природный гормон состоит из 166 остатков, а его искусственный аналог — всего из 20!).



**Рис. 23-9.** (а) Структура исходного «цинкового пальца» (второго модуля белка Zif268); ион Zn показан шариком. (б) Структура искусственного белка FSD-1. (в) Спектр КД для FSD-1 при 1 °С. (г) Изменение спектра КД искусственного FSD-1 с температурой. Картинки, с небольшими упрощениями, взяты из В. I. Dahiya и S. L. Mayo, *Science* (1997) 278:82–86

Однако, по-видимому, и задача рационального дизайна плотной упаковки может быть решена. Имея хороший алгоритм компьютерного перебора астрономического числа возможных упаковок боковых групп и отсека бесперспективных вариантов, расчет плотной упаковки (для маленького белка) может быть проведен.

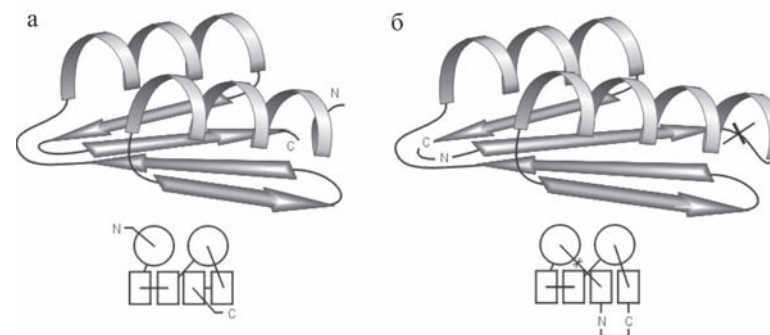
По крайней мере, Дахийат и Майо его провели и создали в 1997 г. твердый (без всякого дополнительного комплексообразования с ионом) маленький белок (см. рис. 23-9). Этот белок был создан на основе архитектуры «цинкового пальца» (широко распространенного ДНК-связывающего мотива) — но без иона Zn, держащего архитектуру природного цинкового пальца. При этом искусственный белок FSD-1 Дахийата и Майо имеет очень невысокую (20 %) гомологию с природным цинковым пальцем и лишен цинк-связывающего центра. И тем не менее он тверд при низких температурах (его структура исследована при помощи ЯМР). Однако он плавится в довольно широком температурном диапазоне, т. е. заметно менее кооперативно (см. рис. 23-9г), чем аналогичные ему природные белки.

Кроме белков, основанных на природных архитектурах, конструируются искусственные белки, чьи архитектуры не имеют природных аналогов.

В основу искусственного белка альбегетина (рис. 23-10а), сконструированного в нашей группе и полученного А. Н. Федоровым, Д. А. Долгих

и М. П. Кирпичниковым, положена структура, не обнаруженная еще в природе. Она состоит из двух повторов типа  $\alpha$ - $\beta$ - $\beta$  (поэтому белок и назван альбегетин), и она не противоречит общим принципам формирования структуры глобулярных белков.

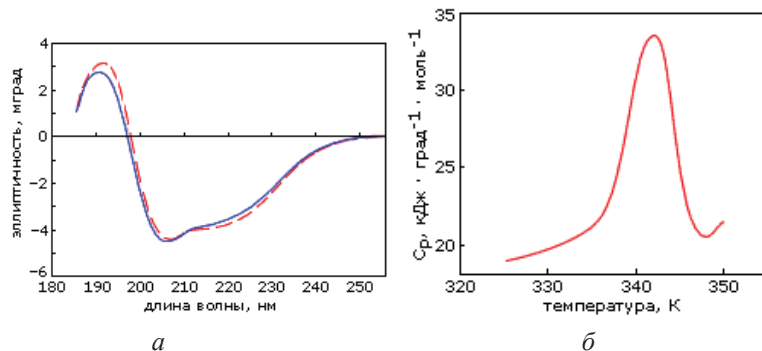
Структурное исследование альбегетина показало, что он обладает заданной вторичной структурой (рис. 23-11а) и компактной, весьма стабильной к разворачиванию мочевиной и к протеолизу пространственной структурой. Однако он кооперативно не плавится и находится скорее в состоянии расплавленной глобулы, чем в твердом.



**Рис. 23-10.** Конструируемая пространственная структура альбегетина (а) и схема экспериментально определенной пространственной структуры рибосомального белка S6 (б). Внизу показаны схемы топологий этих белков. Для S6 показана также искусственно введенная пермутация — разрезание петли (X) и соединение исходных N- и C-концов (N–C). Такая пермутация придает ему архитектуру, сконструированную для альбегетина

Еще один белок со структурой, запланированной для альбегетина, был получен другим методом: при помощи циркулярной пермутации природного белка S6 (он, как и ряд других белков с недавно расшифрованной пространственной структурой, имеет ту же укладку структурных сегментов, что была разработана для альбегетина, но эти сегменты по-иному соединены белковой цепью, см. рис. 23-10б). Полученный белок обладает твердой, кооперативно плавящейся пространственной структурой (рис. 23-11б).

Недавно альбегетин был использован в качестве носителя функциональной активности. В него — точнее, в новый белок, названный альбегероном, — был включен фрагмент 131–138 цепи интерферона  $\alpha 2$  человека, способный активировать бласт-трансформацию тимоцитов. Собственно говоря, активацию делает именно этот фрагмент интерферона, а остальная, уложенная в глобулу цепь этого белка служит как бы ножнами, она



**Рис. 23-11.** (а) Спектр КД альбелбетина (---) и альбелферона (—); (б) кривая микрокалориметрического плавления пермутированного, с целью придания ему топологии альбелбетина, белка S6. Картинки, с небольшими изменениями, взяты из D. A. Dolgikh, A. E. Gabrielian, V. N. Uversky & M. P. Kirpichnikov, *Appl. Biochem. & Biotech.* (1996) **61**:85–96 и из Z. Kh. Abdullaev, R. F. Latypov, A. Ya. Badretdinov, D. A. Dolgikh, A. V. Finkelstein, V. N. Uversky & M. P. Kirpichnikov, *FEBS Letters* (1997) **414**:243–246, соответственно

защищает фрагмент 131–138 от раскусывания и не дает ему действовать слишком уж активно. Опыты показали, что так же работает и альбелферон.

- Работы по функционально-активным искусственным белкам ведутся во многих группах. Создаются модели фибриллярных белков на основе длинных спиральных пучков. Создана первая «работающая» модель мембранного белка из амфифильных (гидрофобно-гидрофильных) спиралей: эти спирали образуют ионные каналы, причем точечные мутации способны резко менять их селективность. Дали функциональный «белок» и  $\alpha$ -спиральные полипептиды, химически пришитые к гему.
- В общем, в настоящее время белки стремительно превращаются из объекта почтительного и изумленного наблюдения в предмет активной инженерной деятельности (но изумление остается...).

## ФИЗИЧЕСКИЕ ОСНОВЫ ФУНКЦИОНИРОВАНИЯ БЕЛКОВ

### Лекция 24

Функция белка и его структура. Элементарные функции. Связывающие белки: ДНК-связывающие белки, иммуноглобулины. Ферменты. Активный центр — «дефект» глобулярной структуры. Твердость белка важна для элементарной ферментативной функции. Каталитический и субстрат-связывающий центры. Ингибирование. Кофакторы. Многовалентные ионы. Механизм ферментативного катализа. Пример: сериновые протеазы. Теория переходного состояния в катализе и ее подтверждение методами белковой инженерии. Абзимы. Специфичность катализа. Узнавание «ключ-замок».

Эту лекцию я посвящаю функционированию белков.

Об этом можно говорить очень много. Я, однако, буду говорить только о физических основах работы белков. Мой рассказ даст вам несколько картинок из жизни работающих белков — картинок, подчеркивающих значение пространственной организации белков для их функции. Некоторые картинки такого рода уже встречались в предыдущих лекциях, в частности, когда я говорил о мембранных белках. Сегодня я буду говорить только о белках глобулярных, водорастворимых.

- Очень грубо, общую схему функционирования белка можно представить в следующем виде:

*СВЯЗАТЬ* → *ТРАНСФОРМИРОВАТЬ* → *ОТПУСТИТЬ*.

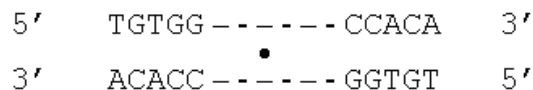
Уточню: некоторые белки делают только часть из этих действий; слова «СВЯЗАТЬ» и «ОТПУСТИТЬ» могут подразумевать связывание и отпуск нескольких разных молекул; а слово «ТРАНСФОРМИРОВАТЬ» может означать и химическую трансформацию, и изменение конформации



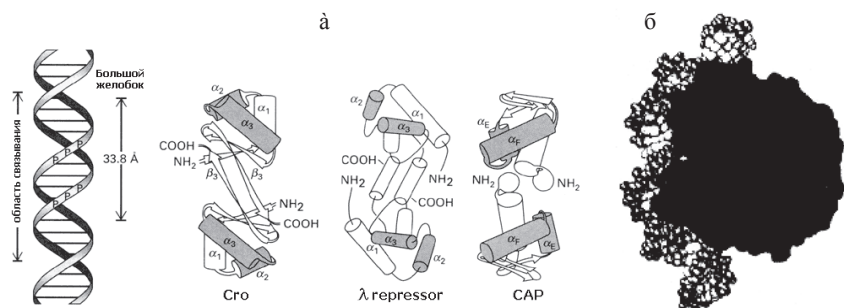
(как самого белка, так и субстрата), и/или перемещение (белка или субстрата) в пространстве.

■ Начнем с белков, чья основная функция — «СВЯЗАТЬ». К ним относятся, например, ДНК-связывающие белки.

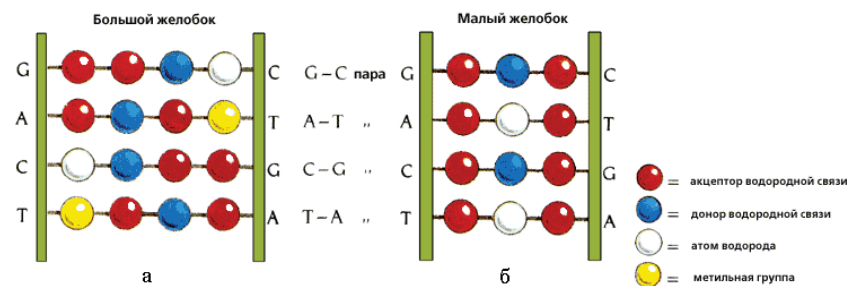
Для связывания с ДНК поверхность белка — причем на большом протяжении — должна быть приблизительно комплементарна поверхности двойной спирали (рис. 24-1а). Тогда выступы белковой поверхности смогут глубоко внедриться в желобок ДНК, и уже там его боковые группы смогут провести тонкое опознавание конкретной ДНКовой последовательности (рис. 24-2) и связаться с той, для которой этот белок предназначен. Все показанные белки на рис. 24-1 — димеры, и именно в такой форме они комплементарны ДНКовому дуплексу. При этом две одинаковые ДНК-опознающие  $\alpha$ -спирали такого димера узнают *палиндром* в двойной спирали ДНК, т. е. такую пару участков в ДНКовом дуплексе, которая одинаково выглядит при перевороте на  $180^\circ$  относительно перпендикулярной к ДНК оси, например



Здесь «—» означают произвольную последовательность ДНК между двумя половинами палиндрома, а «•» — ось переворота.



**Рис. 24-1.** (а) Структура ДНК (слева) и ряда белков, обладающих характерным ДНК-связывающим мотивом «спираль-изгиб-спираль» (он выделен серым цветом). Для белка-активатора катаболитического гена (CAP — catabolite gene activator protein) показан только его С-концевой домен. Все эти белки димерны, и все они опознают большой желобок ДНК своими спиралями  $\alpha_3$  ( $\alpha_F$  у CAP), расстояние между которыми в димере близко к периоду двойной спирали ДНК (33,8 Å). Картинки В. W. Matthews взяты, с разрешения, из [6]. (б) Изгибание ДНК (более светлая спираль слева) связавшимся с ней димерным CAP-белком (справа, черный). Связывание его с ДНК требует присутствия циклоАМФ



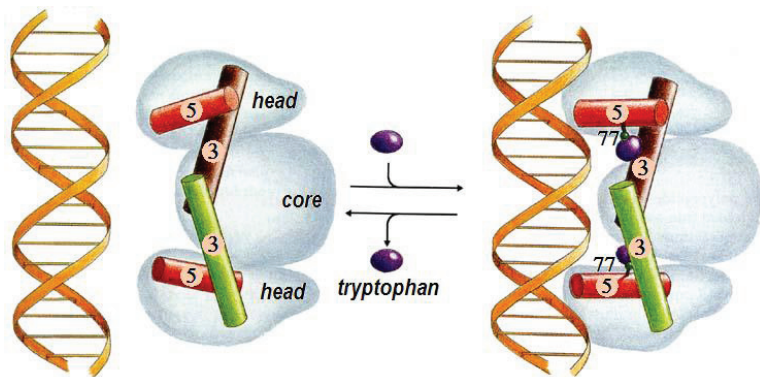
**Рис. 24-2.** Характерные узоры, создаваемые разными функциональными группами А-Т и С-Г пар в большом и малом желобках ДНК. Картинка, с разрешения, взята из [5]

ДНК-связывающие  $\alpha$ -спирали в таком белковом димере *антипаралельны* друг другу, а расстояние между ними близко к периоду двойной спирали ДНК, так что димер садится на один бок двойной спирали ДНК. Однако разный угол наклона этих  $\alpha$ -спиралей к соединяющей их центры оси приводит к тому, что все эти белки по-разному изгибают ДНК при связывании. Порой такой изгиб, *индуцированный* связыванием, довольно велик (рис. 24-1б).

Иногда связывание белка с ДНК требует наличия *кофакторов*, которые изменяют — точнее, слегка деформируют — структуру белка и переводят его из неактивной формы в активную.

Так, в *trp*-репрессоре (он — в *E. coli* — репрессирует оперон, отвечающий за синтез РНК, кодирующей белки, необходимые для синтеза триптофана) — таким *кофактором* (точнее, *корепрессором*) является сам триптофан (рис. 24-3). Пока триптофан не связался с белком, расстояние между ДНК-связывающими спиралями в димере *trp*-репрессора слишком мало (около 28 Å вместо необходимых 34 Å), так что тот не может связать ДНК. Триптофан же, связавшись с белком, отодвигает спирали так, что они становятся комплементарными к желобку в двойной спирали и связываются с ней. Так что, когда триптофана в клетке много, он, связываясь с белком, блокирует дальнейший синтез триптофан-синтезирующих белков, и тем самым, свой (триптофана) дальнейший синтез. Этот способ регуляции называется *отрицательной обратной связью*.

Триптофан в данном случае выступает *стимулятором* ДНК-связывающей активности *trp*-репрессора и *ингибитором* синтеза белков, необходимых для синтеза триптофана. Причем это стимулирование *trp*-репрессора является «аллостерическим», так как Тгр связывается с белком «в другом месте» — не в том, где с белком должна связываться ДНК.



**Рис. 24-3.** Схема действия триптофанового (trp) репрессора. На фоне общего контура димера, обладающего общим сплавленным ядром (core) и двумя идентичными головками (head), показаны только те две спирали (3 и 5), между которыми садится корепрессор — аминокислота Трп. Она отталкивает 77-й остаток цепи, а та —  $\alpha$ -спираль 5. Только при этом обе спирали 5 могут связаться с ДНК. Картинка, с разрешения, взята из [5]

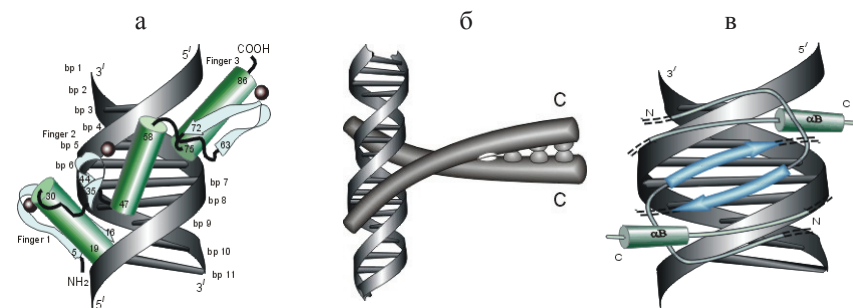
Мотив «спираль-изгиб-спираль», фигурирующий на рис. 24-1 и рис. 24-3, — характерный, но отнюдь не единственный структурный мотив, связывающий ДНК. Чтобы подчеркнуть это, я вынес три других характерных мотива на рис. 24-4. Я хотел подчеркнуть, что ДНК-связывающие белки могут принадлежать к разным структурным классам (на представленных рисунках есть и  $\alpha$ -, и  $\alpha + \beta$ -белки) и что даже само связывание с ДНК может осуществляться как  $\alpha$ -, так и  $\beta$ -структурой.

До сих пор мы говорили о тех грубых (характерный размер:  $\sim 10\text{--}30 \text{ \AA}$ ) чертах белковой структуры, которые позволяют ей внедриться в желобок ДНК. За опознание же той конкретной ДНКовой последовательности, с которой белок должен связаться, отвечают более мелкие черты белковой поверхности (характерный масштаб — размер атома,  $\sim 3 \text{ \AA}$ ).

К сожалению, «общий код» выборочного опознавания белками фрагментов ДНК пока не выяснен (и не ясно, есть ли такой сколько-нибудь четкий «код»), хотя, глядя на детали каждого расшифрованного ДНК-белкового контакта, можно увидеть, какие водородные связи между боковыми группами белка и нуклеотидами и какие другие плотные контакты способствовали образованию ДНК-белкового контакта именно в этом месте.

■ Тонкое, выборочное опознавание белками других молекул удобно рассмотреть на примере *иммуноглобулинов*, или *антител* — белков, предназначенных (в организмах позвоночных) для тонкого распознавания мелкоаппаратных (размером в атом или несколько атомов) *антигенных детерминант*

у самых разных молекул самого разного размера. А аналогичные иммуноглобулинам *рецепторы Т-клеток* опознают таким же образом маленькие антигенные детерминанты у клеток, например, у клеток, зараженных вирусами.

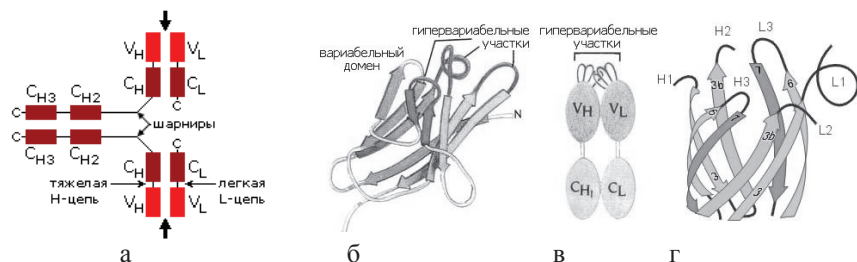


**Рис. 24-4.** Еще три характерных ДНК-связывающих белковых мотива. В двух из них ключевая роль принадлежит  $\alpha$ -спиралям: (а) «цинковые пальцы» (шарики — ионы Zn); (б) «лейциновый zipper». В третьем, met-репрессоре, — ключевая роль принадлежит  $\beta$ -шпильке (v): она специфически связывается с большим желобком ДНК, в то время как  $\alpha$ -спирали  $\alpha\text{V}$  связываются неспецифически с сахаро-фосфатным остовом ДНК. Цинковый палец, “finger” (этот домен можно отрезать от целого белка и выделить отдельно) — один из самых маленьких из известных глобулярных белков, а лейциновый zipper — самый простой из них по структуре. Когда лейциновый zipper не связан с ДНК, он представляет собой просто димер из параллельных  $\alpha$ -спиралей, слепившихся по всей длине своими узкими гидрофобными поверхностями. Образующие эти поверхности боковые группы показаны в виде выступов. Однако одна (на каждой спирали) из этих групп не гидрофобна. У места скрещения спиралей находится вкрапленный в гидрофобную поверхность полярный Asn (см. маленькое пятнышко). Он необходим для образования именно димера, так как его замена на более гидрофобный остаток приводит к тому, что спирали слипаются не по две, а по три и более. Картинки, с небольшими изменениями, взяты из [6] (а) и [5] (б, в). У каждой цепи помечен ее N- и/или C-конец

Иммуноглобулины состоят из многих  $\beta$ -структурных доменов и относительно небольших гибких шарниров между ними (рис. 24-5а). Разнообразие сочетаний переменных (антиген-связывающих) доменов обеспечивает иммуноглобулину широкий спектр сортов и, соответственно, широкий спектр действия, а твердость этих доменов — высокую селективность действия иммуноглобулина каждого сорта. Я не буду пересказывать основы клонально-селекционной теории происхождения огромного разнообразия иммуноглобулинов.

[Вы должны помнить из других курсов, что в зародышевых клетках представлены не целые гены легких и тяжелых цепей иммуноглобулинов, а части, куски этих генов. Там эти части собраны в кассеты — отдельно много сортов для каждого из трех

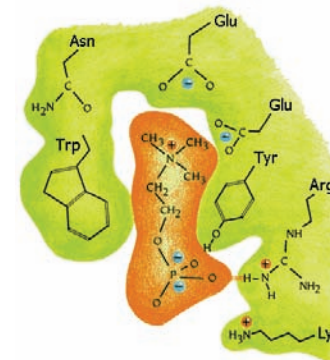
кусков варибельного домена тяжелой цепи, отдельно — легкой цепи; отдельно — константные домены каждой цепи, отдельно — шарниры. При образовании соматических иммунных клеток эти куски всячески тасуются и еще каким-то образом мутируют в своих гиперварибельных участках — и соединяются в целые гены легких и тяжелых иммуноглобулиновых цепей. Как эти «соматические мутации» вводятся, как направляются в особые части генов — все это еще мало известно. И почему такие мутации не взрывают структуру белка, а их может быть до нескольких десятков в домене: слишком много, чтобы сначала ввести их все чохом, а потом отбирать «выжившие» белки! — тоже загадка.]



**Рис. 24-5.** (а) Общее строение одного из иммуноглобулинов (IgG). Отмечены варибельные V и константные C домены двух легких (L: из двух доменов каждая) и двух тяжелых (H: из четырех доменов каждая) цепей, а также C-концы всех этих цепей. Домены  $C_{H2}$  гликозилированы. Жирные стрелки указывают на антиген-связывающие «карманы» между доменами  $V_L$  и  $V_H$ . (б) Строение домена  $V_H$ ; почти так же выглядят домены  $V_L$ ,  $C_L$ ,  $C_H$  и прочие. В показанном домене  $V_H$  выделены гиперварибельные петли; гиперварибельные петли есть и в домене  $V_L$ . Вместе эти два домена образуют антиген-связывающий карман (в). (г) Антипараллельный  $\beta$ -цилиндр, образованный  $\beta$ -листом домена  $V_H$  и  $\beta$ -листом домена  $V_L$ . Антиген-связывающий карман образуют исходящие из этих листов гиперварибельные петли тяжелой и легкой цепи ( $H_1$ ,  $H_2$ ,  $H_3$  и  $L_1$ ,  $L_2$ ,  $L_3$ , соответственно). Картинки (б, в) взяты, с разрешения, из [5]; картинка (г) взята из [5] с небольшими изменениями

Для нас важно сейчас, что антиген опознается варибельными доменами легкой и тяжелой цепей ( $V_H$  и  $V_L$ ) совместно — точнее, их ( $V_H$  и  $V_L$ ) гиперварибельными петлями, оторачивающими антиген-связывающий карман, находящийся на стыке этих двух доменов (рис. 24-5б-г). Первичная структура этих петель варьирует от одного сорта молекул иммуноглобулина к другому (что и создает огромное разнообразие этих сортов). Однако, для каждого данного сорта — не только аминокислотная последовательность, но и конформация всех петель строго фиксирована, а сам антиген-связывающий карман покоится на твердом  $\beta$ -цилиндре, образованном соединившимися в рукопожатии антипараллельными  $\beta$ -листами варибельных доменов. Поэтому каждая молекула иммуноглобулина может сильно связать только определенный антигенный детерминант и равнодушна к другим.

Рисунок 24-6 показывает, что селективность связывания антигенных детерминант определяется не устройством белка в целом (оно служит лишь как бы фундаментом), а прежде всего комплементарностью формы, обводов связываемой молекулы к форме относительно небольшой вмятины, к форме только самого антиген-связывающего кармана.



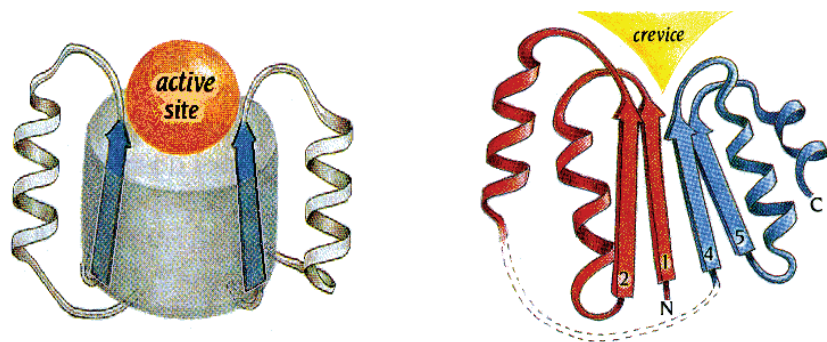
**Рис. 24-6.** Специфическое взаимодействие антигена и связывающего его «кармана» антитела. Показаны сближающиеся заряды и образующиеся водородные связи. Картинка, с разрешения, взята из [5]

Кроме того, гидрофобные части связываемой молекулы контактируют с гидрофобными частями кармана, его заряды комплементарны зарядам, вкрапленным в карман, а находящиеся на антигене доноры и акцепторы водородных связей комплементарны вкрапленным в карман антитела акцепторам и донорам этих связей. Все это делает связывание — но только связывание строго определенного антигена — крепким.

■ Такое же, как в антителах, расположение активного центра — в воронке на торце  $\beta$ -цилиндра — наблюдается и во многих других белках, никак с иммуноглобулинами не связанных. Например, это — стандартное место активного центра в  $\alpha/\beta$ -цилиндрах (где, в отличие от иммуноглобулинов,  $\beta$ -структура параллельна). Вообще, изучая структуры белков, легко заметить, что очень часто активный центр помещается в их «стандартном дефekte», в стандартно расположенной (т. е. определяемой мотивом укладки цепи, а не боковыми группами) вмятине в архитектуре белковой глобулы (рис. 24-7): такая вмятина автоматически способствует окружению субстрата одновременно многими боковыми цепями белка.

Столь же обычным местом активного центра является место стыка доменов. Рисунок 24-8 показывает активный центр сериновых протеаз типа трипсина, он находится на стыке двух  $\beta$ -структурных доменов.





**Рис. 24-7.** Стандартные вмятины в архитектурах белковых глобул часто определяют местоположение (не функцию!) активного центра. Слева: активный центр (active site) в воронке при верхушке  $\beta/\alpha$ -бочонка с параллельным  $\beta$ -цилиндром; сходное расположение активного центра в воронке при верхушке *анти*параллельного  $\beta$ -цилиндра смотри на рис. 24-5г. Справа: активный центр в щели (crevice), образующейся в укладке Россманна в месте расхождения правых  $\beta$ - $\alpha$ - $\beta$ -суперспиралей, идущих в разные стороны (в суперспирали  $\beta_1$ - $\alpha$ - $\beta_2$  цепь идет от нас, в суперспирали  $\beta_4$ - $\alpha$ - $\beta_5$  — к нам). Картинки, с разрешения, взяты из [5]



**Рис. 24-8.** Положение активного центра в сериновых протеазах типа трипсина. Показаны части активного центра: *каталитического центра*, где серым цветом выделены боковые группы «триады переноса заряда» — Ser195 (S), His57 (H) и Asp102 (D) и *субстрат-связывающего центра*, серыми шариками показаны NH-группы, образующие оксианионовую дыру Ox, черным — неспецифическая субстрат-связывающая площадка, и совсем светлыми — группы, выстилающие специфический субстрат-связывающий карман. Стык доменов находится между парой (His57, Asp102) и остальной частью активного центра

- Заключая рассказ о функции *СВЯЗАТЬ*, надо сказать несколько слов о симметричной ей функции *ОТПУСТИТЬ*.

Уместно напомнить, что многие так называемые «нативно-развернутые» белки приобретают свою уникальную пространственную структуру, только связав лиганд (или сев на другой белок, ДНК или РНК): ведь обеспечивающая специфичность связывания твердость структуры белка востребована только тогда, когда наличествует связь белка с лигандом. Поэтому неудивительно, что, выполняя функцию *ОТПУСТИТЬ*, некоторые белки теряют свою уникальную пространственную структуру. Так, например, ретинол-связывающий белок, отпустив ретинол, превращается в расплавленную глобулу.

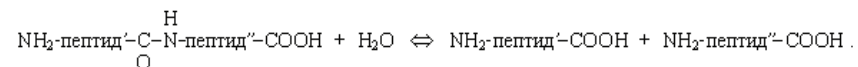
- Теперь естественно рассказать о ферментах — белках, чья главная функция — химически *ТРАНСФОРМИРОВАТЬ* связавшиеся с ними молекулы.

Ферменты не создают новые реакции и не меняют их направление, они «только» ускоряют самопроизвольно протекающие процессы. Но ускоряют их порой в миллионы миллионов раз, в миллионы раз сильнее, чем очень сильные химические катализаторы.

Я позволю себе не касаться классификации ферментов (ее вы найдете в любом биохимическом учебнике) и сосредоточусь на вопросе о том, как ферментам удается столь ускорить химическую реакцию.

Сериновые протеазы — классический лекционный объект, используемый для рассказа о простой ферментативной реакции, и я не буду отступать от этой традиции.

Сериновые протеазы разрезают полипептидные цепи, т. е. проводят реакцию



Реакция гидролиза пептидной цепи идет, когда в среде достаточно много воды, и сама по себе, но очень *медленно*, за многие годы, т. е. гидролиз (при наличии свободной воды) термодинамически выгоден, но требует преодоления очень высокого активационного барьера. Если же свободной воды в среде нет — реакция идет в другую сторону, в сторону синтеза полипептида и выделения воды, но тоже очень медленно.

В присутствии же фермента реакция гидролиза пептида (или, при отсутствии свободной воды, — обратная реакция его синтеза из более мелких фрагментов) занимает доли секунды, то есть фермент резко снижает ее активационный барьер. Посмотрим, как он это делает.



Прежде всего рассмотрим основные компоненты *активного центра* фермента.

Он состоит из *каталитического центра*, ответственного за проведение химической трансформации, и *субстрат-связывающего центра*, призванного правильно подставить субстрат под удар каталитического резака (или, точнее, сварочно-режущего аппарата, так как фермент равно ускоряет и прямую, и обратную реакцию).

Катализ — в сериновых протеазах — непосредственно осуществляется боковой цепью одного определенного серина (называемого «Ser195» — по его расположению в цепи химотрипсина — во всех белках семейства трипсина; именно он, трипсин, изображен на рис. 24-8). Напомню химическую формулу боковой группы серина:  $-\text{CH}_2-\text{OH}$ . Однако для того, чтобы серин мог катализировать гидролиз, его (серин) нужно подготовить. Пока кислород находится в форме  $-\text{OH}$ -группы, он не активен, а активным он становится после утери  $\text{H}^+$  и перехода в форму  $-\text{O}^-$ . Этим отрывом атома  $\text{H}$  от Ser195 занимаются два остальных члена «триады переноса заряда» — His57 (он-то и принимает оторванный атом  $\text{H}^+$ ) и «вспомогательный» Asp102. Мутации обоих этих остатков — не говоря уже о мутации или химической модификации каталитического серина — практически губят каталитическую активность трипсиновых протеаз.

Активированный атом  $\text{O}^-$  боковой группы Ser195 совершает основное каталитическое действие. Он атакует  $\text{C}$ -атом расщепляемой пептидной группы (рис. 24-9), завязывая с ним временную ковалентную связь (правда, более длинную и менее прочную, чем обычные  $\text{C}-\text{O}$ -связи), переводя тем самым этот  $\text{C}$ -атом в тетраэдрическую форму (где он ковалентно связан уже не с тремя, а с четырьмя другими атомами; «дополнительный» атом — это  $\text{O}^-$  боковой группы Ser195). А последующий распад этого комплекса ведет к разрыву  $\text{C}-\text{N}$  связи и последующему разрыву обрабатываемой цепи.

Субстрат-связывающий центр состоит (см. рис. 24-8, 24-9) из *оксианионной дыры*, связывающей кислород расщепляемой пептидной группы, из *неспецифической* пептид-связывающей *площадки* (отвечающей — вместе с оксианионной дырой — за то, что расщепляемая пептидная группа займет правильное положение относительно активированного  $\text{O}$ -атома боковой группы Ser195), и из *специфического* субстрат-связывающего *кармана*, отвечающего за распознавание той аминокислоты, после карбоксила которой производится расщепление пептида.

Общий ход реакции иллюстрируется рис. 24-9. Эта схема была получена ценой многолетних исследований многих групп. Для ее построения привлекались данные и по катализу расщепления разных субстратов, и по химическим модификациям фермента, а равно и белково-инженерные

исследования (которые показали, какие точки фермента вовлечены в катализ), и исследования ферментов в комплексе с нерасщепляемыми аналогами субстратов, и рентгенструктурные работы (показавшие, где образуются водородные связи), и так далее.

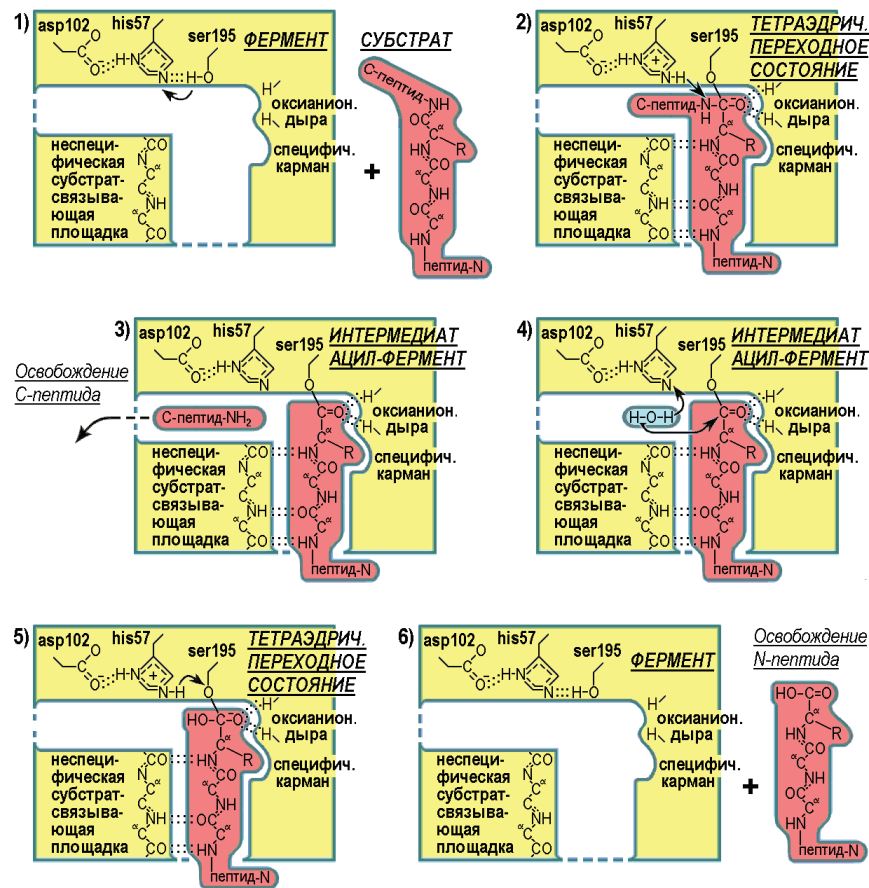
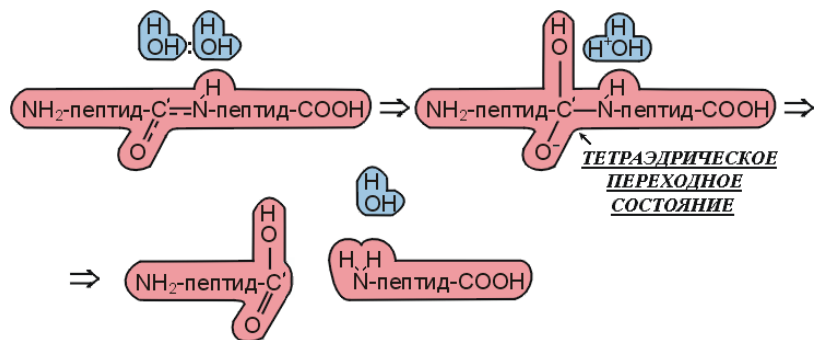


Рис. 24-9. Схема ферментативного гидролиза пептида

■ Что же делает фермент, как он ускоряет химическую реакцию? Для ответа на этот вопрос давайте сравним ферментативную реакцию (см. рис. 24-9) с аналогичной реакцией, протекающей спонтанно — в присутствии воды, но без фермента (рис. 24-10).



**Рис. 24-10.** Схема спонтанного, не ферментативного гидролиза пептида в воде по тому же, что на рис. 24-9, химическому механизму

Мы видим, что ферментативная реакция идет в две стадии: сначала отщепляется C-концевой пептид и образуется комплекс N-концевого пептида с O-атомом серина, а затем этот N-пептид заменяет свою связь с O-атомом серина на связь с O-атомом воды. Без фермента же реакция идет в одну стадию, в которой отщепляется C-концевой пептид и образуется комплекс N-концевого пептида с O-атомом воды. Мы видим также, что все эти три реакции идут через тетраэдрический активированный комплекс, точнее, активированный комплекс, где рядом с тетраэдрическим C-атомом (завязавшим ковалентную связь с O-атомом) находится донор протона ( $\text{His}^+$  в ферменте или, в воде, —  $\text{OH}_3^+$ ).

То, что на пути ферментативной реакции находится два активированных комплекса (а не один, как в не-ферментативной) — не может, само по себе, ускорить реакцию. Однако ускоряет ее то, что активированный комплекс в связи с ферментом *менее* нестабилен (а именно его нестабильность и лимитирует скорость реакции), чем активированный комплекс в воде.

Активированный комплекс (или, что то же самое, «переходное состояние») стабилизируется на ферменте следующим образом.

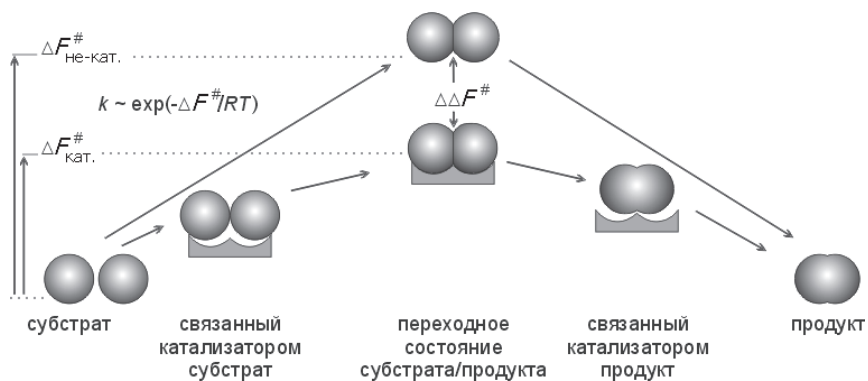
Во-первых, отрицательный заряд C'-O<sup>-</sup> группы пептида, образующийся при тетраэдризации C'-атома, на ферменте втягивается в оксианионную дыру, где стоят два уже нацеленных в эту дыру протона. Водородная и электростатическая связь с этими положительно заряженными протонами понижает энергию отрицательного заряда O<sup>-</sup> и, следовательно, энергию тетраэдризации C'-атома — по сравнению с не-ферментативной реакцией, где такой заранее устроенной «дыры» нет, т. е. приводит к понижению энергии (энтальпии) переходного состояния, или к так называемому *энтальпийному катализу*.

Во-вторых, рядом с тетраэдрическим C'-атомом стоит, с протоном наготове,  $\text{His}^+$  — а он примерно столь же стабилен, как «безпротонный»  $\text{His}^0$ . В воде роль донора протона должен играть ион  $\text{OH}_3^+$  — ион, концентрация которого, вследствие его нестабильности, в воде крайне мала — порядка  $10^{-7}$  моль/л. Его «вылавливание» из воды понижает энтропию и соответственно повышает свободную энергию не-энзиматического активированного комплекса. Поэтому энтропия препятствует сбору «до кучи» всех компонентов активированного комплекса в воде и не мешает этому сбору на ферменте, где все эти компоненты уже собраны вместе. При этом энтропия «вылавливания» фермента субстратом оплачивается его прилипанием к субстрат-связывающему карману. Этот облегченный сбор всех компонентов активированного комплекса понижает его свободную энергию на ферменте по сравнению с одной в не-ферментативной реакции и приводит к так называемому *энтропийному катализу*.

В сумме, энтропийная и энтальпийная компоненты катализа ускоряют ферментативную реакцию примерно в  $10^{10}$  раз по сравнению с не-ферментативной.

Если прилипание субстрата к ферменту достаточно сильно, а концентрация субстрата не слишком низка, то энтропийный катализ понижает наблюдаемый порядок реакции по концентрации субстрата. Так, если скорость неэнзиматической реакции синтеза типа  $A + A \rightarrow A_2$  определяется квадратом концентрации субстрата A, то скорость аналогичной энзиматической реакции, на первой стадии которой субстрат A связывается ферментом, зависит лишь от концентрации фермента, а от концентрации субстрата (или субстратов) в растворе не зависит. Этот эффект особенно силен тогда, когда речь идет о реакции, в которую вовлечено несколько молекул-субстратов, например — при синтезе пептида из двух фрагментов в отсутствие воды.

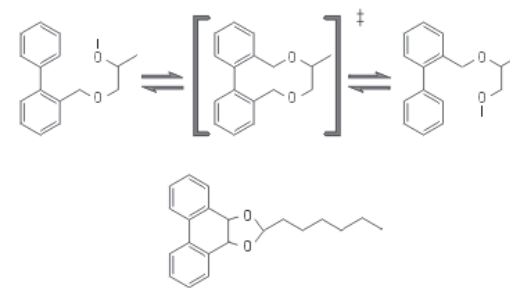
■ Основное и в химическом, и в ферментативном катализе — понижение свободной энергии переходного состояния, т. е. снижение максимума свободной энергии, преодолеваемого по ходу реакции (рис. 24-11). Снижения этой свободной энергии можно достигнуть за счет энтропии, за счет сбора на ферменте всех необходимых компонентов реакции. И его же можно достигнуть, или усилить, за счет энтальпии преимущественного связывания именно переходного состояния обрабатываемой молекулы, а не ее начального состояния — «субстрата» и/или состояния конечного — «продукта». Кстати, если субстрат или продукт будут связываться очень уж сильно (сильнее, чем переходное состояние), фермент будет ими просто ингибироваться, отравляться (он потеряет свою функцию «ОТПУСТИТЬ» и выйдет из игры).



**Рис. 24-11.** Схема, поясняющая важность стабилизации катализатором именно переходного (самого нестабильного по ходу реакции) состояния субстрата/продукта и важность жесткости катализатора для снижения свободно-энергетического барьера реакции  $\Delta F^\#$  и увеличения ее скорости  $k$ . Рисунок, для наглядности, изображает просто комплементарность обводов переходного состояния к обводам активного центра фермента, в то время как в действительности основную роль обычно играют временные ковалентные связи переходного состояния субстрата/продукта с каталитическим центром фермента. Свободные энергии начального и конечного состояний не меняются катализатором. Рисунок подчеркивает, что ускорение реакции зависит от величины  $\Delta\Delta F^\#$ , т. е. от силы связывания ферментом именно переходного состояния

При преимущественном связывании переходного состояния ведущую роль может играть как связывание его деформированной электронной системы (вспомним «оксианионовую дыру» сериновых протеаз), так и преимущественное связывание деформированной — опять-таки в переходном состоянии — конформации всей молекулы.

■ Последнее особенно ярко проявляется в каталитической функции искусственных «абзимов» (*antibody enzyme*) — антител, отобранных так, чтобы, связывая переходное состояние субстрата, они катализировали бы его химическое превращение (рис. 24-12). Некоторые абзимы способны направить на снижение активационного барьера почти всю энергию, выгаданную при связывании субстрата, т. е. почти вся она прилагается к деформации именно обрабатываемой химической связи. И хотя имеющиеся абзимы — не очень мощные ферменты (они ускоряют спонтанную реакцию максимум на 5 порядков, а не на 10–15, как естественные ферменты — ведь в них еще не умеют встраивать доноры и акцепторы электронов) — возможность их создания подтверждает важность преимущественного связывания именно переходного состояния.

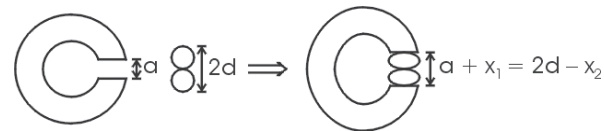


**Рис. 24-12.** Принципиальная схема химического превращения, катализируемого абзимом — искусственно выработанным антителом к переходному состоянию реакции, которая спонтанно идет медленно. В переходном ( $\ddagger$ ) состоянии изображенной реакции изомеризации имеется — в отличие и от начального, и от конечного ее состояний — плоская система из трех колец. Для выработки антител к переходному состоянию такой формы используется включающая аналогичные кольца, но стабильная молекула, изображенная внизу: эта молекула впрыскивается в кровь животного, и на нее вырабатывается иммунитет

Хочу подчеркнуть, что для преимущественного связывания именно переходного состояния, в пике отличающимся от него на какой-то ангстрем начальному и конечному состояниям химической реакции (см. рис. 24-11), белок должен быть как можно более твердым. Мягкий белок — точнее, мягкий активный центр — столь же эффективен при катализе, как резиновое лезвие при бритье...

Может быть, будет полезно рассмотреть следующую простую модель.

Предположим, что две молекулы (на следующем ниже рисунке — два шарика диаметром  $d$  каждый) втягиваются, за счет энергии сорбции, в имеющуюся на белке щель («активный центр»), первоначально имеющую ширину  $a$  (меньшую, чем  $2d$ ). При этом шарики сжимаются, а щель расширяется. И белок, и шарики считаются упругими (последние — до определенной степени сжатия, за которой следует образование химической связи между ними). Пусть коэффициент упругости белка есть  $k_{pr}$ , а коэффициент упругости столбика из двух молекул —  $k_{mol}$ . Спрашивается, какая энергия пойдет на деформацию шариков (только ее можно использовать для преодоления энергетического барьера, стоящего — см. рис. 24-11 — на пути реакции спонтанного слипания двух молекул в одну).



Решаем. Ширина щели под давлением шариков возрастает с  $a$  до  $a + x_1$ , а длина столбика из двух шариков падает с  $2d$  до  $2d - x_2$ , причем  $2d - x_2 = a + x_1$ . Тогда, по закону Гука, на деформацию белка идет энергия  $E_{pr} = \frac{1}{2} k_{pr} x_1^2$ , а на деформацию

молекул —  $E_{\text{mol}} = \frac{1}{2} k_{\text{mol}} x_2^2$ . При этом  $k_{\text{pr}} x_1 = k_{\text{mol}} x_2$ , так как сила давления молекул на белок равна (третий закон Ньютона!) силе давления белка на молекулы. Значит,

$$\begin{aligned} E_{\text{mol}}/E_{\text{pr}} &= (\frac{1}{2} k_{\text{mol}} x_2^2)/(\frac{1}{2} k_{\text{pr}} x_1^2) = \\ &= [(1/k_{\text{mol}}) (k_{\text{mol}} x_2)^2]/[(1/k_{\text{pr}}) (k_{\text{pr}} x_1)^2] = k_{\text{pr}}/k_{\text{mol}}, \end{aligned} \quad (24.1)$$

и (после некоторых вычислений)

$$E_{\text{mol}} = \frac{1}{2} k_{\text{mol}} (2d - a)^2 [k_{\text{pr}}/(k_{\text{pr}} + k_{\text{mol}})]^2. \quad (24.2)$$

Итак, мы видим что при «мягком» белке ( $k_{\text{pr}} \ll k_{\text{mol}}$ ) основная энергия идет на бесполезную для катализируемой реакции деформацию белка и лишь «твердый» белок ( $k_{\text{pr}} \gg k_{\text{mol}}$ ) способен сконцентрировать основную часть энергии на деформации обрабатываемых молекул.

Что и требовалось доказать.

**Внутренний голос:** Значит, «твердым» должен быть не столько сам белок, сколько его активный центр? И, говоря о «твердости», — Вы имеете в виду просто механическую твердость белка, или что-то другое?

**Лектор:** Я имею в виду твердость именно активного центра; твердость остального белка важна — для катализа — только постольку, поскольку она обеспечивает твердость активного центра. И я имею в виду не столько механический, сколько энергетический аспект твердости: то, что белок не подстраивается под изменение субстрата, так что небольшое изменение состояния субстрата вызывает большое изменение энергии его взаимодействия с белком. Этот эффект может вызываться механическим взаимодействием обводов субстрата с белком; здесь картина особенно наглядна, см. рис. 24-11 и 24-12. Но такой же эффект может быть связан и с взаимодействием валентных электронов субстрата и белка; энергия взаимодействия электронных облаков резко меняется при очень «мелкомасштабных» изменениях их формы.

Деформация жестких ковалентных связей субстрата, перестройка которых и составляет суть химической реакции, более эффективно осуществляется жесткими же временными ковалентными связями субстрата с каталитическим центром фермента, чем деформацией субстрата при помощи более мягких Ван-дер-Ваальсовых взаимодействий. Это объясняет, почему абзимы, катализирующие реакцию только за счет Ван-дер-Ваальсовых взаимодействий с субстратом, являются менее эффективными катализаторами, чем природные ферменты с их химически активными каталитическими центрами.

**Внутренний голос:** Значит, гибкость белка не играет никакой роли в катализе?

**Лектор:** Здесь надо, во-первых, четко отличать сам акт катализа, которому гибкость белка не способствует, от этапа проникновения субстрата

в активный центр, для чего гибкость бывает нужна; с этим эффектом мы еще встретимся. Во-вторых, надо иметь в виду следующее. Если активный центр белка не подогнан к переходному состоянию субстрата идеально (никто не чужд недостатков), то некоторая гибкость позволит этому неидеальному активному центру принять должную форму хоть на время. Важно, однако, чтобы активный центр допускал подгонку именно под переходное, а не начальное или конечное состояние субстрата.

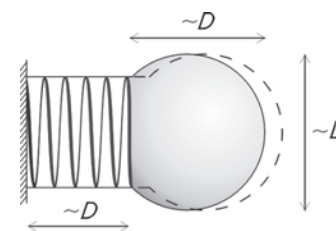
**Внутренний голос:** Все Ваши рассуждения до сих пор касались статического или почти статического фермента. А не может ли энергия, необходимая для катализа, запасаться как кинетическая энергия движущихся частей фермента, скажем, как энергия его колебаний?

**Лектор:** Нет, не может. Можно показать, что кинетическая энергия даже большой молекулы белка рассеивается за  $\sim 10^{-11}$  с максимум (а маленькой — и того быстрее), в то время как один акт ферментативного катализа занимает не менее  $\sim 10^{-6}$  с. То есть выделившаяся при посадке субстрата на фермент кинетическая энергия рассеется, перейдет в тепло задолго до того, как она сможет хоть как-то пригодиться в каталитическом акте.

В самом деле, движение одной части белка относительно другой (скажем, при колебании) происходит в вязком окружении (воде, мембране и т. д.). Такое движение описывается обычным простым уравнением

$$m (d^2x/dt^2) = F_{\text{frict}} + F_{\text{elast}}, \quad (24.3)$$

Здесь  $d^2x/dt^2$  — ускорение движения частицы,  $m$  — ее масса,  $F_{\text{frict}}$  — сила ее трения об окружающую среду, и  $F_{\text{elast}}$  — действующая на нее упругая сила. Массу можно оценить как  $m = \rho V$ , где  $\rho$  — плотность частицы, а  $V$  — ее объем.  $F_{\text{frict}}$  можно оценить по закону Стокса:  $F_{\text{frict}} = -3\pi D\eta (dx/dt) \approx -10D\eta (dx/dt)$ , где  $dx/dt$  — скорость частицы,  $D$  — ее диаметр, а  $\eta$  — вязкость среды.  $F_{\text{elast}}$  можно оценить по закону Гука:  $F_{\text{elast}} = -(ES/L)x$ , где  $x$  — смещение от точки равновесия,  $S$  — площадь сечения сочленения двух частей белка,  $L$  — длина эластичной части, а  $E$  — модуль ее упругости. В ходе дальнейших оценок разумно принять, что все линейные размеры примерно одинаковы (взгляните на прилагаемую схему), так что  $L \approx D$ ,  $S \approx D^2$ ,  $V \approx D^3$ .



Уравнение (24.3) задает два характерных времени.

Одно, самое важное, мы уже обсуждали. Это —  $t_{\text{kinet}}$  время затухания движения белка под воздействием трения об окружающую среду. Оценка этого времени



следует из «кинетической» части этого уравнения:  $m(dx/dt^2) \approx -3\pi D\eta(dx/dt) \approx -10D\eta(dx/dt)$ . Получается, что  $t_{\text{kinet}} \approx m/(10D\eta)$ ; его легко оценить численно. Так как  $m \approx \rho D^3$ , то  $t_{\text{kinet}} \approx 0,1\rho D^2/\eta$ . И, поскольку плотность белка  $\rho \approx 1 \text{ г/см}^3$ , а вязкость воды  $\eta \approx 0,01 \text{ г/(см с)}$ , то  $t_{\text{kinet}} \approx 10^{-13} \text{ с} (D/\text{нм})^2$ , где  $(D/\text{нм})$  — диаметр белка, измеренный в нанометрах. Значит, кинетическая энергия небольшой части белка размером  $D = 1 \text{ нм}$  рассеется в водном окружении за  $10^{-13} \text{ с}$ , а у большого белка, размером  $D = 10 \text{ нм}$ , она рассеется за  $10^{-11} \text{ с}$ . Еще быстрее рассеется кинетическая энергия при трении о более вязкое, чем вода, окружение (о мембрану или о другую часть белка).

Полезно еще оценить, каким может быть характерное время колебания белка  $t_{\text{vibr}}$ . Его величина следует из «колебательной» части уравнения (24.3),  $m(dx/dt^2) \approx -EDx$ :  $t_{\text{vibr}} \approx (m/ED)^{1/2} \approx D(\rho/E)^{1/2}$ . При характерном для белков модуле упругости  $E \sim 10^{10} \text{ г/(см с}^2)$ ,  $t_{\text{vibr}} \approx 10^{-12} \text{ с} (D/\text{нм})$ . Сравнивая  $t_{\text{vibr}}$  с  $t_{\text{kinet}}$  мы видим, что в белках размером  $D < 10 \text{ нм}$  упругие колебания вообще невозможны (сразу затухают) из-за трения.

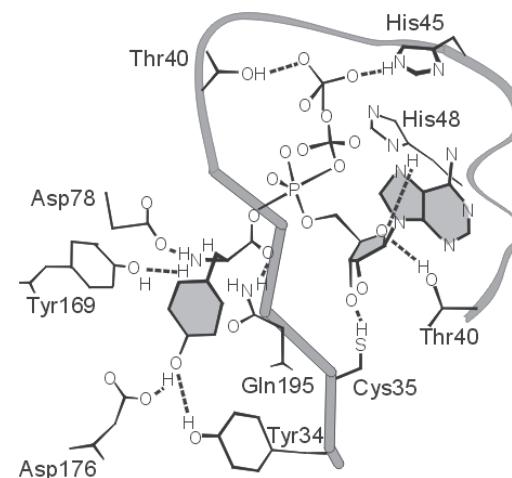
Значит, (1) потенциальная энергия может продуктивно вкладываться только в субстрат (чтобы подвести его к переходному состоянию), а не в белок, и поэтому активный центр белка должен быть тверд; и (2) использование кинетической энергии белка в катализе невозможно.

■ Теория катализа за счет преимущественного связывания переходных состояний была обоснована Холденом и Полингом еще в 1930–40-х гг. В последние годы переходные состояния для ряда энзиматических реакций были «прошупаны» методами белковой инженерии, т. е. путем целенаправленной замены аминокислот активного центра фермента. При этом удалось выяснить (и увидеть, при помощи рентгена), какие из аминокислотных остатков фермента отвечают за слипание разных компонентов, потребных для создания активированного комплекса, и какие — за преимущественное связывание именно и только переходного состояния реакции, дотоле гипотетического, так как его устройство не было видно структурному эксперименту.

Так, в тирозил-тРНК синтетазе, исследованной А. Ферштом, были найдены те остатки (Thr40 и His45, рис. 24-13), что связываются с  $\gamma$ -фосфатом АТФ *только* в переходном состоянии, что и отличает это переходное состояние от просто связанных субстратов или продуктов реакции.

Переходное состояние очень нестабильно (по определению: максимально нестабильно!), и прямо увидеть его нельзя. Так что рентген, видящий лишь стабильное расположение субстрата на ферменте, не видит его (субстрата) связывания с Thr40 и His45 (здесь расстояние слишком велико), между тем как замены Thr40 и His45 на маленькие аланины снижают активность мутантного белка более чем в миллион раз — хотя эти замены почти совсем не влияют на связывание «мутантом» субстра-

тов реакции. Это свидетельствует, что их (Thr40 и His45) связывание с  $\gamma$ -фосфатом АТФ происходит только в переходном состоянии — при напряженном (из-за того, что он вступил в пять ковалентных связей) состоянии  $\alpha$ -фосфора, атакующего СО-группу тирозина.



**Рис. 24-13.** Предполагаемое строение переходного состояния в реакции образования тирозил-АМФ из тирозина и АТФ на ферменте тирозил-тРНК синтетазе. В целях наглядности кольцо тирозина заштриховано и кольца аденозина — тоже. Атакующий тирозин активированный пяти-координатный  $\alpha$ -фосфор (P) молекулы АТФ находится в центре рисунка, а  $\gamma$ -фосфат АТФ — в его верхней части. Этот  $\gamma$ -фосфат образует водородные связи с Thr 40 и His 45 только в переходном состоянии реакции; при простом, непродуктивном связывании субстратов этих Н-связей нет. Картинка, с небольшими изменениями, взята из [9]

■ Говоря о ферментативной реакции, мы не можем упускать из виду также ее очень высокую специфичность.

Так, сериновые протеазы щепят полипептид только после определенных аминокислот: химотрипсин — после ароматических, трипсин — после положительно заряженных, эластаза — только после самых маленьких. Этот эффект [он называется опознаванием субстрата по принципу «ключ (субстрат) — замок (фермент)»] достигается четкой структурой той *специфической* части субстрат-связывающего «кармана» (рис. 24-14), куда должна улечься последняя (как на трипсине) или предпоследняя (как на папаине) перед точкой щепления боковая группа пептида. В термоллизине же ключевым для специфичности является аминокислотный остаток, стоящий сразу *после* расщепляемой С–N-связи.

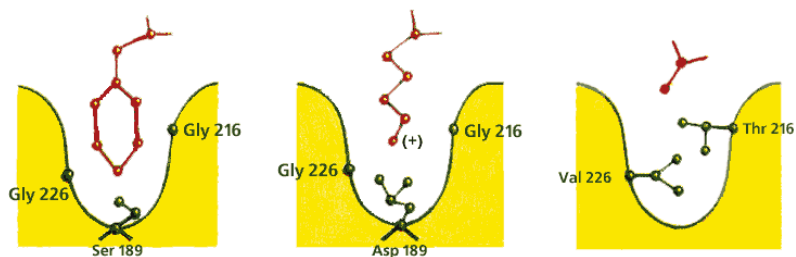


Рис. 24-14. Устройство специфического сайта субстрат-связывающего кармана в разных сериновых протеазах. Картинка, с разрешения, взята из [5]

Ключевые остатки достаточно точно намечают возможную точку щепления, но скорость щепления полипептида в намеченной точке зависит также и от того, какие другие остатки эту точку окружают (и это естественно: ведь они тоже как-то взаимодействуют с протеазой).

Однако еще более важной для эффективности и самой возможности щепления является конформация полипептида в месте щепления. Если эта конформация не годится для щепления (т. е. если ключевая боковая группа и подлежащая щеплению связь C–N не могут занять своего «правильного» положения на ферменте одновременно), то эта связь C–N не будет разорвана.

Поэтому протеазы плохо расщепляют белки с жесткой пространственной структурой — но, как только эти белки денатурируют, легко с ними расправляются.

Имея определенную жесткую конформацию, полипептидная цепь может даже просто прилипнуть к протеазе, закрыть ее активный центр, да так, нерасщепленной, на ней и остаться. На этом принципе «непродуктивного связывания» работают многие ингибиторы.

Протеазы, о которых мы говорили сегодня, являются эффективными, но не абсолютно аккуратными ферментами. По-видимому, организм еще может позволить им изредка ошибаться в выборе точки щепления. Но есть белки, от которых требуется абсолютная надежность. Она достигается работой нескольких активных центров одного белка.

Рассмотрев сегодня элементарные функции белков, мы рассмотрим их более сложные функции на следующей лекции.

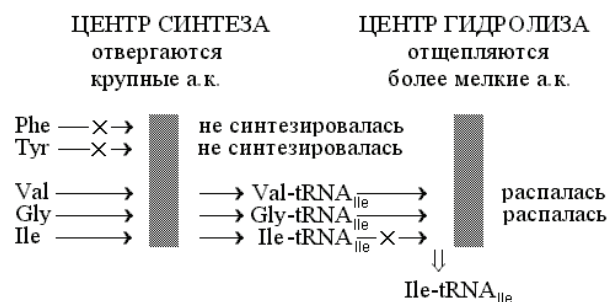
## Лекция 25

Сочетание функций. Переход субстрата с одного на другой активный центр. «Двойное сито» повышает специфичность функции. Относительная независимость структуры белка от его элементарной ферментативной активности. Заметная связь структуры с окружением белка. Сопряжение элементарных функций белка и гибкость его структуры. Индуцированное соответствие. Подвижность доменов белка. Перемешивание доменов при эволюции белков. Доменная структура: киназы, дегидрогеназы. Аллостерия — взаимодействие активных центров. Аллостерическая регулировка функции белка. Аллостерия и четвертичная структура белка. Гемоглобин и миоглобин. Механизм мышечного сокращения. Шагающий кинезин.

Рассмотрев на прошлой лекции элементарные функции белков, мы рассмотрим теперь их более сложные функции.

- Как уже говорилось, есть белки, от которых требуется колоссальная точность работы, и эта точность достигается сопряженной работой нескольких активных центров одного белка. Так обстоит дело, например, в аминоксил-тРНК синтетазах: они «заряжают» тРНК нужными аминокислотами, и от точности их работы зависит точность синтеза белков. Поэтому они не могут позволить себе ошибаться чаще, чем однажды на много тысяч синтезов аминоксил-тРНК, а опознавание аминокислоты специфическим субстрат-связывающим карманом допускает порядка 1% ошибок. Для устранения этих ошибок используется эффект «двойного сита» (рис. 25-1). Этот эффект обеспечивается специальной конструкцией фермента.

У фермента аминоксил-тРНК синтетазы *два* активных центра: центр синтеза и центр гидролиза. Причем гидролиз аминоксил-тРНК на аминоксил-тРНК синтетазе *не* есть просто реакция, в точности обратная синтезу: обе эти реакции идут с выделением свободных фосфатов, т. е. *обе* — с понижением свободной энергии.



**Рис. 25-1.** Схема «двойного сита» для механизма редактирования работы изолейцил-тРНК синтетазы, заряжающей изолейцином (Ile) изолейциновую тРНК (tRNA<sub>Ile</sub>)

При этом субстрат-связывающий карман центра гидролиза *меньше*, чем карман центра синтеза аминоксил-тРНК.

Принцип сита — *не* пропускать частицы, которые *больше*, чем ячейки сита.

На первом сите — при синтезе аминоксил-тРНК — тРНК заряжаются «правильной» аминокислотой и некоторым числом «неправильных», более мелких (более крупные аминокислоты отвергаются все, а большая часть «неправильных» мелких тоже отвергается по причине различий в гидрофобности/гидрофильности, но все же отбор по гидрофобности/гидрофильности не столь жесток, как отбор по возможности втиснуться в карман данного размера). Итак, на выходе этапа синтеза — много аминоксил-тРНК с «правильной» аминокислотой и некоторое количество аминоксил-тРНК с «неправильными» аминокислотами — но только с теми, что *меньше*, чем «правильная»: более крупные аминокислоты отвергаются «карманом» синтетического центра, этим первым «ситом».

На втором активном центре, на втором «сита» — у него «ячейка» (субстрат-связывающий карман) поменьше — идет гидролиз образовавшихся аминоксил-тРНК, но *только* тех, где аминокислота *меньше*, чем «правильная». То есть «правильная» аминокислота отвергается карманом гидролизного центра, этим «вторым ситом», а все остальные аминоксил-тРНК гидролизуются и распадаются.

В сумме два эти сита и обеспечивают выход только «правильно заряженных» аминоксил-тРНК.

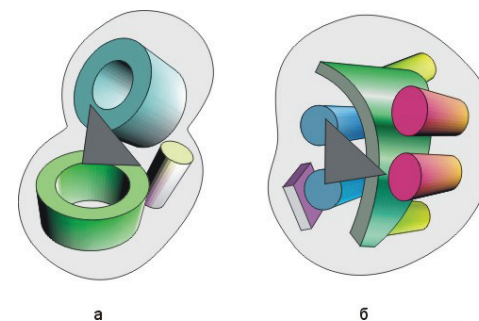
- В завершение рассказа о специфичности ферментов стоит остановиться на двух прагматических аспектах. Во-первых, сейчас масса людей занята изучением субстрат-связывающих карманов и подбором — с открытыми глазами, не наощупь — связывающихся с ними ингибиторов. Особый интерес, по понятным причинам, вызывают протеазы и другие белки вируса

СПИДа. Во-вторых, большое применение находят себе точечные мутации в районе активного (и прежде всего субстрат-связывающего) центра: они позволяют модифицировать «природную» субстратную специфичность (например, менять специфичность сериновых протеаз) и даже создавать новые, искусственные специфичности для нужд промышленности и медицины. Значительный интерес также вызывают попытки — пока это лишь попытки — создания (на «твердом фундаменте» белковой глобулы) «нового» активного центра с новой, «привитой» функцией путем направленных точечных замен аминокислотных остатков исходного белка.

- Изучение ферментативной деятельности белка показывает (вспомните прошлую лекцию), что в эту функцию активно вовлечена совсем небольшая часть белковой глобулы — в то время как остальная ее часть служит лишь как бы твердым каркасом, обеспечивающим «правильное» строение закрепленного на нем активного центра.

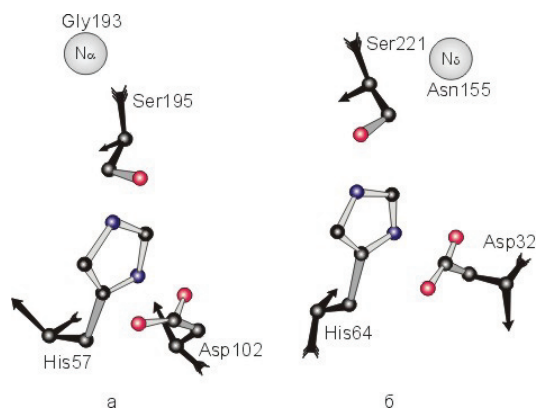
Поэтому нас не должно удивлять, что белки с совсем разной первичной структурой и даже с совсем разной пространственной структурой могут иметь одинаковые или очень сходные биохимические функции.

Классическим примером снова являются сериновые протеазы. Есть два класса таких протеаз: типа трипсина и типа субтилизина. Они совсем *не* похожи ни по аминокислотной последовательности, ни по общей форме (рис. 25-2) и даже принадлежат к разным структурным классам (трипсин — двухдоменный β-белок, субтилизин — однодоменный α/β-белок). Они имеют одинаковую конфигурацию лишь ключевых аминокислотных остатков в каталитическом центре (рис. 25-3), причем не всех остатков целиком, а лишь их функционально важных «кончиков», но не имеют одинаковой конфигурации остатков даже в субстрат-связывающем кармане.



**Рис. 25-2.** Схема строения сериновых протеаз типа трипсина (а) и субтилизина (б). На фоне общего контура глобулы показаны α-спирали (цилиндры) и β-листы (более или менее плоские или скрученные в бочонки). Район активного центра показан черным треугольником

Более того: существуют протеазы, совершенно не похожие на сериновые протеазы даже в зоне каталитического центра. Я имею в виду, в частности, металлопротеазы типа карбоксипептидазы или термолитина. У них ключевую роль в катализе играет плотно встроенный в белок (и удерживаемый там координационными связями) ион  $Zn^{++}$ . Потеряв внешние электроны и имея поэтому маленькие радиусы, многозарядные металлические ионы создают вокруг себя очень сильное электрическое поле. Это позволяет такому иону, встроенному в металлопротеазу, активировать молекулу воды (разломать ее на  $OH^-$  и  $H^+$ ) с тем, чтобы та осуществила далее разрыв пептидной связи. При этом активированная вода «работает» примерно так же, как в не-катализируемой реакции, о которой мы говорили на прошлой лекции. Но существенная разница заключается в том, что фермент активирует именно ту молекулу воды, что сидит рядом с атакуемой пептидной связью, а в не-каталитической реакции ее приходится долго разыскивать среди массы не-активированных молекул  $H_2O$ . Одновременно — в металлопротеазах — заряд  $Zn^{++}$  стабилизирует отрицательный заряд на тетраэдрическом переходном состоянии разрываемого пептида (т. е. имитирует действие оксианионовой дыры сериновых протеаз).



**Рис. 25-3.** Схема строения каталитических центров сериновых протеаз типа трипсина (а) и субтилизина (б). [Кстати, такие же каталитические центры есть не только в протеазах: они есть в липазах и в некоторых других ферментах]. Черным выделены фрагменты главной цепи (вплоть до  $C^{\beta}$ -атомов); направление хода цепи в этих фрагментах показано стрелками. Кругом (с N) показано место оксианионовой дыры. В нее смотрят  $N_{\alpha}$ -H-группы главной цепи в трипсине, и  $N_{\delta}$ -H-группа боковой цепи в субтилизине. Видно, что только расположение концов боковых групп каталитической триады переноса заряда (Ser, His, Asp) и, с натяжкой, положение оксианионовой дыры инвариантно в этих двух классах сериновых протеаз, в то время как кардинальные отличия наблюдаются даже в ходе главной цепи у ключевых каталитических остатков His и Asp

■ Итак, одна и та же химическая реакция может осуществляться совсем разными белками. При этом одну и ту же ключевую роль — роль мощных электрических резаков — в одних белках играют многозарядные металлические ионы, в других — легко принимающие электроны или протоны органические кофакторы, в третьих — активированные своим окружением (как в трипсине) боковые группы...

С другой стороны, как мы помним, разные белки с одной и той же характернейшей формой  $\alpha/\beta$ -цилиндра (а равно и белки, имеющие форму укладки Россманна) могут катализировать совершенно разные реакции.

Классическим примером белков с общей структурой, общим происхождением, но разными функциями является пара лизоцим —  $\alpha$ -лактальбумин. Эти белки не только обладают почти идентичной (с точностью до 1–2 Å) пространственной структурой, но и гомологичны — они имеют 35 % общих идентичных аминокислотных остатков. Тем не менее, один из них (лизоцим) является ферментом (он режет олигосахара), а другой ( $\alpha$ -лактальбумин), недавно — сотню миллионов лет назад, с появлением молока — произошедший от первого, утратил активный центр лизоцима и теперь сам по себе ферментом вообще не является, а модифицирует ферментативную активность другого белка.

■ Анализ корреляции между ферментативными свойствами белков (эти свойства фиксируются классификацией ферментов, «ЕС») и их структурной классификацией (CATH) показал, что, в целом, не существует яркого предпочтения какой-то биохимической функции белками какого-то структурного класса.

Правда, более «крупномасштабные» свойства белков (зависящие не от маленького каталитического центра, а от большей поверхности глобулы), свойства связывать тот или иной крупный лиганд, — эти свойства отчасти коррелируют с общей архитектурой белка. Отчасти коррелирует с общей архитектурой белка и его вовлеченность в те или иные биологические процессы.

Наиболее очевидно это проявляется в белках, ответственных за транспорт — самый разный транспорт — через мембрану. Будучи мембранными белками, все они имеют очень регулярную трансмембранную  $\alpha$ - или  $\beta$ -структуру; мы об этом уже говорили. Говорили мы и о том, что фибриллярные белки, с их огромными блоками вторичной структуры, играют в основном структурную роль. Однако, пусть в меньшей степени, эта связь проявляется и в водорастворимых глобулярных белках.

Так, половина гем- и ДНК-связывающих белков являются  $\alpha$ -белками, а доля  $\alpha$ -белков среди углеводов- и нуклеотид-связывающих белков состав-



ляет не более 20 %. С другой стороны, почти все белки, отвечающие за иммунитет и за опознание одними клетками других, являются  $\beta$ -белками.  $\beta$ -белками является и почти половина углевод-связывающих белков — в то время как среди гем-, нуклеотид- и ДНК-связывающих белков доля  $\beta$ -белков составляет около 10 %. Почти все нуклеотид-связывающие белки принадлежат к  $\alpha/\beta$ -классу. К этому классу принадлежат, например, все 11 белков гликолитического пути (использующие нуклеотиды в качестве кофакторов и субстратов). Вообще,  $\alpha/\beta$ -белки часто — чаще, чем остальные — являются ферментами, и особенно теми, что обеспечивают «внутреннее хозяйство» клетки.

Все это показывает довольно широкую независимость общего строения белка от его ферментативной активности, при некоторой связи его архитектуры с «крупномасштабной» субстрат-связывающей активностью белка, с его окружением («средой обитания») и вовлеченностью белков в те или иные биологические процессы.

- До сих пор я почти ничего не говорил о конформационных изменениях белковых молекул в ходе их функционирования и не случайно: пока речь идет — как у нас до сих пор — о катализе одной, отдельно взятой энзиматической реакции — конформационная гибкость белка может только ухудшить его каталитические свойства. В самом деле, для эффективного катализа нужно преимущественное связывание именно переходного состояния, в пику отличающимся от него лишь на какой-то ангстрем начальному и конечному состояниям, и это преимущественное связывание нацелено на синтез или разрыв очень жестких ковалентных связей. Так что рвать их при помощи «гибкого» белка — все равно, что резать проволоку резинкой или подушкой...

Впрочем, здесь надо добавить следующее. Когда активный центр белка не идеально подогнан под свою функцию — а идеальность подгонки требует сверхтщательной работы естественного отбора — этот центр может требовать небольшой деформации для перехода в активную форму. И тогда такая деформация будет, действительно, функционально необходима — она будет, время от времени, доводить несовершенный белок до кондиции. Однако здесь надо четко отличать нужду (деформацию, искупающую несовершенство геометрии активного центра) от добродетели (от совершенства белка).

Конечно, даже самый лучший белок несколько деформируется в процессе своей каталитической работы — ведь он не алмазный — но это играет примерно ту же функциональную роль, что деформация бумаги под пером: чем меньше, тем лучше.

**Внутренний голос:** Значит, в каталитическом центре фермента вообще не должно быть гибких деталей?

**Лектор:** Нет, почему же. Их не должно быть, только если фермент должен работать быстро. А вот в каталитическом центре G-белка такие детали есть. Но — в чем задача G-белка? Его задача расщеплять ГТФ, но — *очень медленно!* Ведь G-белок (мы о нем уже говорили) — это таймер. И вот для *медленной* работы гибкий каталитический центр подходит прекрасно... А для «быстрого» катализа нужен твердый активный центр.

- Другое дело, когда белок должен перейти от одного действия к другому: тогда его деформация, обеспечивающая переход от одной роли к другой, действительно, играет важную функциональную роль.

Так, во многих белках проникновение субстрата в активный центр требует небольшого расхождения близких к этому центру боковых групп. Однако, после того, как субстрат сел на этот центр, его каталитическая работа движения этих групп уже не требует.

Можно сказать, что подвижность нужна при подготовке к реакции, а жесткость — в момент ее проведения. Это похоже на военное «правило Клаузевица»: «на марше — отдельно, в бою — вместе».

Правда, иногда переход от одной функции белка к другой осуществляется просто переходом субстрата с одного на другой активный центр (вспомните аминоксил-тРНК синтетазы: в них вся подвижность сосредоточена в субстрате).

Однако часто переход от одной элементарной функции к другой осуществляется за счет более или менее крупномасштабного преобразования структуры белка, и необходимость в таком преобразовании отражается на строении белковой молекулы. Эта тема и будет занимать нас в дальнейшем ходе лекции.

- Говоря о подвижности белков, надо различать мелкие и более крупные движения.

Мелкие движения — это, по существу, тепловые флуктуации. Рентген-структурный анализ высокого разрешения способен видеть такие флуктуации на фоне усредненной, «статичной» структуры белка. Он видит их как некую размытость положения атомов, причем величина этой размытости, как и положено для тепловых флуктуаций, растет с температурой.

Флуктуации больше у поверхностных групп белка (здесь их амплитуда, в среднем, около половины ангстрема), и в несколько раз меньше у групп, входящих в ядро белка. Аминокислотные остатки в петлях колеблются сильнее, чем во вторичных структурах. Боковые группы колеблются заметно сильнее, чем главная цепь. Особенно сильны колебания длинных поверх-

ностных боковых групп, концы которых не зажаты другими группами и часто столь подвижны, что рентген их вообще не фиксирует. Поэтому порой говорят, что внутренность белка — твердая, а его поверхность напоминает жидкость, которая затвердевает только около  $-100\text{ }^{\circ}\text{C}$ . (В скобках отмечу, что полужидкий поверхностный слой вообще характерен для кристаллов.)

К выводу о сочетании «твердых» и «жидких» составляющих в динамике белковой глобулы приводит и гамма-резонансная («мёссбауэровская») спектроскопия. Она следит за поглощением монохроматических  $\gamma$ -квантов «мёссбауэровскими» ядрами некоторых изотопов тяжелых металлов (например,  $\text{Fe}^{57}$ ) и позволяет оценить, в частности, размер твердой, крепко спаянной с этими мёссбауэровскими ядрами части белка.

Более крупномасштабные изменения связаны с деформацией белковой структуры при связывании лигандов. Динамика релаксации таких деформаций лучше всего изучена в миоглобине. С его гема можно лазерной вспышкой практически мгновенно (за  $\sim 10^{-13}$  с) сорвать лиганд  $\text{CO}$ , а затем следить (по оптическим спектрам) за возвращением к исходному состоянию.

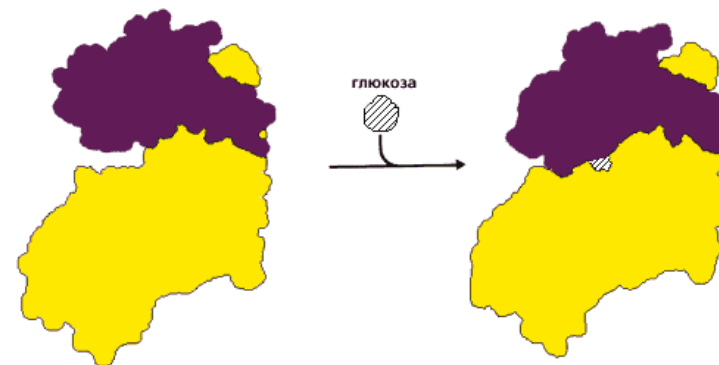
Сложная кинетика такой релаксации, изученная в широком диапазоне температур, указывает на то, что путь к нативной конформации лежит через ряд энергетических барьеров, а сама нативная конформация белка охватывает набор конформаций, различающихся в микроскопических деталях.

Такие и еще более крупномасштабные деформации белков при исполнении их функций изучают также, кристаллизуя белки в разных функциональных состояниях. Богатую информацию дают также разные спектральные методы, химические модификации и т. д.

■ Мы уже познакомились, на прошлой лекции, с одной функционально важной деформацией на примере регуляции ДНК-связывающей активности  $\text{trp}$ -репрессора. Сейчас мы рассмотрим вопрос о функционально важных деформациях более внимательно.

Прежде всего посмотрим, как деформация белка помогает сочетать стадии цикла **СВЯЗАТЬ**  $\rightarrow$  **ТРАНСФОРМИРОВАТЬ**  $\rightarrow$  **ОТПУСТИТЬ**. Рисунок 25-4 демонстрирует *индуцированное соответствие* (введенный Д. Кошландом термин) фосфорилирующего белка — гексокиназы — его субстратам. Этот белок переносит фосфатную группу с АТФ на глюкозу. Но — в принципе, по химии реакции — эта же фосфатная группа может быть перенесена и на воду; однако этого не происходит. В попытке ответить на вопрос, почему это не происходит, Кошланд постулировал следующее. 1) До связывания с субстратом фермент находится в «открытой» форме (в которой он может захватить субстрат из воды, но не способен провести его фосфорилирование). 2) После связывания с субстратом он — фермент — переходит в «закрытую», каталитически-активную форму, где все части активного центра

собраны воедино и способны катализировать реакцию фосфорилирования, но вода вытеснена из активного центра и потому не конкурирует с субстратом за фосфорилирование. 3) После каталитического акта фермент снова открывается, и фосфорилированный субстрат уходит.



**Рис. 25-4.** Индуцированное соответствие при функционировании гексокиназы. В открытой форме два домена разделены глубокой щелью, куда может заплывать глюкоза. Когда глюкоза попала в щель, домены поворачиваются, щель закрывается, вода из нее вытесняется, а все компоненты каталитического центра сходятся вместе. Картинка, с небольшими изменениями, взята из [5]

Впоследствии опыт полностью подтвердил эту гипотезу (см. рис. 25-4), но только для тех белков, которым нужно скрыть обрабатываемый субстрат от конкурирующей с ним воды (хотя попытки распространить механизм индуцированного соответствия на все ферментативные реакции предпринимались многократно). Для действия трипсина, например, этого не нужно — и в нем индуцированного соответствия субстрату не наблюдается: трипсин (а также химотрипсин, эластаза, субтилизин и т. д.) не деформируется и опознает субстрат по простейшему принципу «ключ-замок».

Хочу привлечь Ваше внимание к тому, что индуцированное соответствие достигается *смещением* либо крупных блоков (вспомним прошлую лекцию), либо целых белковых доменов (см. рис. 25-4), но *не* полной перестройкой укладки белковой цепи. А смещения эти происходят в основном путем мелких локальных деформаций. [Аналогия: мышцы сокращаются («локальная деформация»), и пальцы («домены») сжимаются в кулак — но эти пальцы не превращаются в зубы или щупальца...]

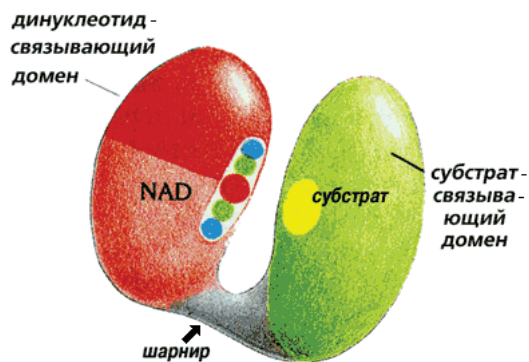
То же самое имеет место и во всех других известных случаях «конформационных перестроек белков» — за немногими исключениями, когда из глобулы иногда вырывается и переходит в нерегулярную конформацию

целая  $\alpha$ -спираль или  $\beta$ -тяж. Одна из самых больших известных мне перестроек происходит в калмодулине. Сам по себе он имеет форму гантели, «головки» которой —  $\alpha$ -спиральные домены — разнесены на большое расстояние «ручкой» — одной длинной  $\alpha$ -спиралью. Однако при связывании калмодулина с рядом других белков «головки» бывшей гантели, не ломаясь, сближаются и слипаются друг с другом и с белком-мишенью, а бывшая «ручка» — длинная  $\alpha$ -спираль — разрушается.

- Домены белков подвижны не только в пространстве, но и в эволюционном времени. Как я уже говорил, гены доменов, как целое, могут кочевать из белка в белок, порой объединяясь в разных сочетаниях с другими, порой разъединяясь. При этом нередки случаи, когда в одном организме наблюдаются несколько мономерных белков, а в другом они сливаются в единый многодоменный белок.

Относительная автономия доменов хорошо видна в большом семействе белков, называемых дегидрогеназами. Эти белки катализируют сходные реакции типа окисления OH-групп при помощи кофакторов  $\text{NAD}^+/\text{NADH}$ , легко принимающих оторванные протоны и отдающих их. Однако обрабатывают они разные вещества: алкогольдегидрогеназа — спирт (этанол), лактатдегидрогеназа — лактат, и т. д.

Дегидрогеназы состоят из двух доменов (рис. 25-5), соединенных перемычкой.



**Рис. 25-5.** Цепи NAD-зависимых дегидрогеназ свернуты в два отдельных домена, соединенных относительно гибкой перемычкой. Один домен более или менее универсален (т. е. очень сходный домен встречается в разных белках) — он связывает кофактор NAD. Второй — не универсальный, а индивидуальный в каждой дегидрогеназе домен — связывает обрабатываемый субстрат. Картинка, с разрешения, взята из [5]

Один, связывающий обрабатываемый субстрат домен, устроен индивидуально в каждой из дегидрогеназ. Например, в алкоголь-дегидрогеназе он содержит  $\beta$ -цилиндр, а в глицеральдегид-3-фосфат-дегидрогеназе — плоский  $\beta$ -лист.

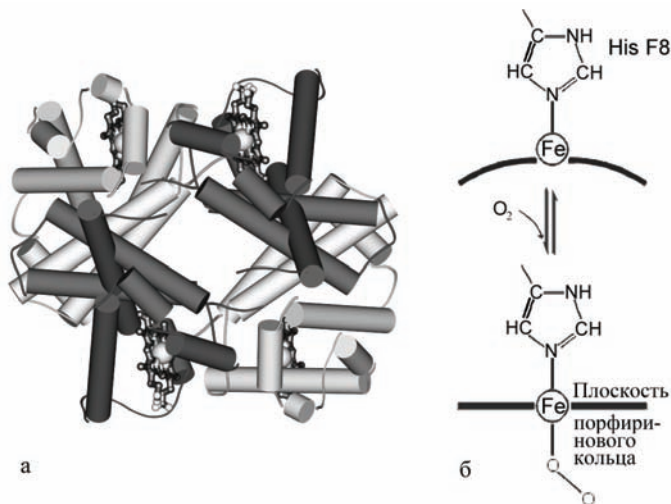
В то же время второй из этих доменов — тот, что связывает кофактор NAD — у всех NAD-зависимых дегидрогеназ устроен практически одинаково, хотя в одних дегидрогеназах он находится в N-, а в других — в C-концевой половине цепи, и несмотря на отсутствие видимой гомологии в первичных структурах у этих NAD-связывающих доменов. В этом домене (типа  $\alpha/\beta$ ) цепь обладает «укладкой Россмана», причем сходство распространяется и на мелкие детали NAD-связывающих доменов разных дегидрогеназ: большую часть этих доменов можно наложить друг на друга с точностью до 2 Å, включая и место связывания NAD [это, по-видимому, показывает, что пространственная структура лучше помнит родство белковых доменов, чем первичная].

Таким образом, разнообразие действия дегидрогеназ обеспечивается — при универсальности NAD-связывающего домена, несущего каталитический центр, — разнообразием субстрат-связывающих центров, расположенных на «субстрат-связывающих» доменах, устроенных по-разному: в структуре целого белка эти две части активного центра соприкасаются (см. рис. 25-5), создается единый активный центр — и он-то и обеспечивает дегидрогеназную функцию.

- Теперь я хочу рассказать о неконтактном — или, как говорят, *аллостерическом* — взаимодействии активных центров. Аллостерические взаимодействия между различными связывающими и активными центрами, особенно в олигомерных белках, играют важнейшую роль в контроле и интеграции биохимических реакций. При этом «сигнал» о состоянии одного активного центра передается на другой через деформацию белковой глобулы — деформации, затрагивающей и этот «другой» центр.

Сейчас я рассмотрю (в самой упрощенной форме) только один, наиболее изученный аллостерический белок — гемоглобин.

Это — тетрамер, точнее — комплекс из двух  $\alpha$ - и двух похожих на них  $\beta$ -цепей (рис. 25-6а). Его дело — *связать* кислород в легких (где его много), *донести* до мышц (где его мало) и *отдать* мышечному миоглобину, мономерному белку, похожему на любую из четырех субъединиц гемоглобина. Активный центр миоглобина и каждой  $\alpha$ - и  $\beta$ -субъединицы гемоглобина — гем: кофактор, содержащий ион железа, связывающий одну молекулу  $\text{O}_2$  (рис. 25-6б). И успешность действия гемоглобина как транспортера  $\text{O}_2$  определяется именно аллостерическим, в данном случае — концертным взаимодействием четырех гемов в нем.



**Рис. 25-6.** (а) Общая структура гемоглобина, состоящего из двух  $\alpha$ - и двух  $\beta$ -цепей; каждая из цепей окрашена в свой цвет;  $\alpha$ -спирали нарисованы как цилиндры; гемы показаны скелетными моделями с ионом железа (шариком) внутри. (б) Схема связывания кислорода гемом в гемоглобине. До связывания  $O_2$  ион железа  $Fe^{2+}$  гема находится в высокоспиновой форме, в которой два внешних электрона находятся на двух разных орбитах. При этом диаметр  $Fe^{2+}$  чуть великоват для того, чтобы войти в центр порфиринового кольца гема (которое при этом находится в слегка изогнутой форме). Связав  $O_2$ , ион  $Fe^{2+}$  приобретает низкоспиновую, чуть более компактную форму, в которой электроны с противоположными спинами занимают одну орбиту. Теперь  $Fe^{2+}$  может войти в центр порфиринового кольца (которое при этом обретает плоскую форму). Показан также гистидин  $\alpha$ -спирали F, координирующий ион железа. Картинка (б), с небольшими изменениями, взята из [9]

Ион железа держится в геме четырьмя координационными связями N– $Fe^{2+}$ . Еще одну такую связь этот ион завязывает в гемо- или миоглобине с ближайшим к нему гистидином, гистидином  $\alpha$ -спирали F. И последнюю связь он может завязать с молекулой  $O_2$  (или CO: вот тут-то и наступает отравление угарным газом...).

Здесь есть, правда, одно препятствие: анализ статической структуры и мио-, и гемоглобинов показал, что молекула  $O_2$  никак не может подойти из окружающей среды к гему! Оказывается, для этого необходима согласованная тепловая флуктуация конформации нескольких боковых групп, прикрывающих гем.

Когда кислород подходит к гему и связывается с его ионом железа, у иона железа гема  $Fe^{2+}$  меняется спиновое состояние и потому — форма

электронной оболочки. Он немного уменьшается в диаметре и входит в центр порфиринового кольца гема. При этом он (ион  $Fe^{2+}$ ) чуть-чуть (на полангстрема) смещается и тянет за собой координирующий его гистидин  $\alpha$ -спирали F.

Так, *электронно-конформационным взаимодействием* (при котором происходит быстрое изменение электронного, а затем и конформационного состояния) начинается цепь конформационной деформации гемоглобина.

Далее смещение гистидина — посредством множества мелких деформаций белка — слегка изменяет обводы связавшей  $O_2$  субъединицы. Ее контакт с другими субъединицами ослабляется. При этом те, еще не связав  $O_2$ , начинают, в свою очередь, релаксировать (приобретая те же обводы, присущие не деформированной взаимодействиями с соседями глобуле, и те же внутренние смещения атомов), так что их гистидины начинают вдавливать соответствующие ионы железа в центры гемов. Теперь эти ионы железа могут уже легче связать кислород. В результате связь одной молекулы  $O_2$  с гемоглобином провоцирует связывание еще трех  $O_2$ . Точно так же утеря одного  $O_2$  ведет к утере их всех.

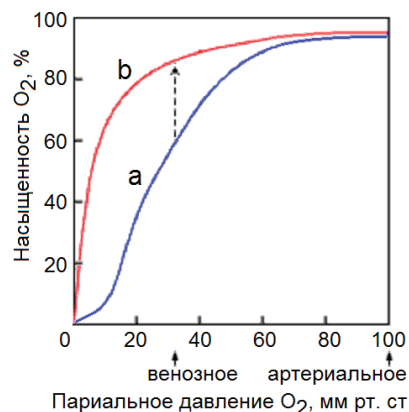
Одновременно со связыванием и утерей кислорода в гемоглобине идут еще два важных процесса, также связанных с его аллостерическими конформационными перестройками. В «дезоксигемоглобине», свободной от  $O_2$ , субъединицы гемоглобина связывают  $CO_2$  (в основном в форме  $HCO_3^-$  — причем не гемами, а N-концами цепей) и ионы  $H^+$  (гистидинами C-концевых спиралей, пространственно близкими к N-концам цепей). А в «окси-форме», связав полученный на вдохе  $O_2$ , гемоглобин теряет и ионы  $H^+$  (это называется «эффект Бора»), и молекулы  $CO_2$ : все это предназначено для выдоха.

Из сказанного выше следует, что гемоглобин ведет себя как белок, связывающий четыре молекулы  $O_2$  *одновременно* — и это отражается в нелинейной, S-образной зависимости насыщения этого (*тетрамерного!*) белка кислородом от его ( $O_2$ ) концентрации (рис. 25-7). А у (*мономерного!*) миоглобина зависимость насыщения кислородом от концентрации  $O_2$  такого S-образного прогиба *не* имеет.

Поэтому в легких, где кислорода много, гемоглобин им насыщается. А в тканях, где (венозное) давление  $O_2$  относительно мало, тетрамерный гемоглобин *уже* отдает  $O_2$ , а мономерный миоглобин *еще* его хватает — и передает  $O_2$  дальше по окислительной цепи мышц — например, тех, что расширяют и сжимают легкие.

Кстати, о мышцах — описанием их работы (или, точнее, описанием *механохимической* функции белка, на которой основана эта работа) я и хочу начать завершение этой лекции.



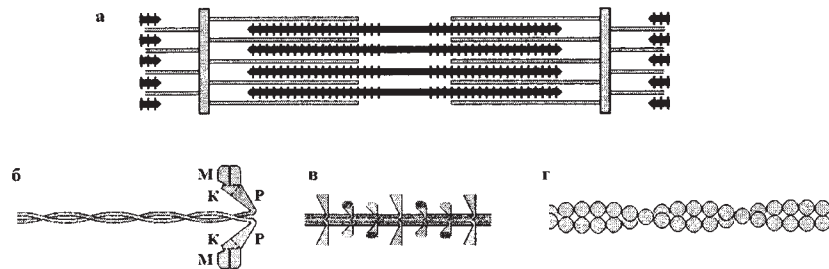


**Рис. 25-7.** Кривые насыщения кислородом тетрамерного белка гемоглобина (а) и мономерного миоглобина (б). Отмечено артериальное давление кислорода, при котором он поглощается гемоглобином крови в легких, и венозное давление, при котором он отдается гемоглобином мышечному миоглобину. Экспериментальная кривая связывания  $O_2$  с миоглобином соответствует реакции первого порядка, а S-образная часть экспериментальной кривой связывания  $O_2$  с гемоглобином — реакции примерно третьего (а не четвертого, как для простоты сказано в тексте) порядка. Стрелка соответствует переходу кислорода с гемоглобина на миоглобин. Картинка, с небольшими изменениями, взята из [1]

Центральную роль в сокращении мышц (рис. 25-8а) играет белок миозин (рис. 25-8б). Я буду рассказывать о работе «миозина II» быстрых поперечно-полосатых скелетных мышц. Миозины других типов работают сходным образом, а похожий механохимический механизм, хотя и с существенными вариациями в деталях, наблюдается в кинезине, динеине и других сократительных белках.

Миозин — это димерный белок. Мономер миозина состоит из фибриллярной части, т. е. «хвоста» (длинной  $\alpha$ -спирали), и из глобулярной части, «головки». Спиральные хвосты скручены в молекуле миозина попарно, и в таком виде они вплетены в толстую миозиновую нить. Продолговатые глобулярные головки периодически, на расстоянии около полутора сотен ангстрем друг от друга, выступают из этого волокна (рис. 25-8в). Они могут связываться с актиновыми нитями (фибриллами, сложенными из глобулярного белка актина, рис. 25-8г). «Сжигая» АТФ, эти головки могут «переступить» по актиновой нити. Последнее и ведет к сокращению мышц.

Как же происходит это «переступание», как химический потенциал АТФ преобразуется в механическое движение?



**Рис. 25-8.** (а) Схема саркомера, структурной единицы мышцы. Толстые миозиновые нити (с торчащими из них головками) размещены между тонкими актиновыми нитями, прикрепленными к ограничивающим саркомер базальным пластинам. Длина актиновой нити — примерно 1 мк. (б) Схема молекулы миозина, состоящей из двух полипептидных цепей. У этой молекулы — две головки и хвост, представляющий собою двойную  $\alpha$ -структурную суперспираль. Центр связывания АТФ/АДФ схематично показан как небольшая щель в миозиновой головке, состоящей на самом деле из трех доменов («моторного» М, «конверторного» К и «рычага» Р). Слипаясь хвостами, молекулы миозина образуют единую миозиновую нить (в), сильно схематизированную на этом рисунке, с периодически торчащими из нее в разные стороны головками; при этом каждая двухголовая молекула миозина сдвинута относительно предыдущей на 143 Å вдоль оси нити. (г) Схема тонкой нити — двойной спирали (с периодом 360 Å), сложенной из глобул актина (другие компоненты актиновой нити (тропонин, тропомиозин, АДФ) не изображены). Актин способен связываться с миозином только при достаточно высокой концентрации кальция

Процесс сокращения мышцы удовлетворительно описывается «моделью шагающих мостиков».

Она основывается на многолетних исследованиях многих лабораторий — исследованиях, включающих и сверхскоростные снимки дифракции рентгеновских лучей сокращающейся мышцей, и электронную микроскопию мышц, и рентгеноструктурный анализ головки миозина в разных ее состояниях, и тонкие эксперименты с единичными молекулами, и многое другое. Эти исследования позволили увидеть большинство основных этапов мышечного сокращения.

Оказалось, что головка миозина может находиться в двух макроскопических конформационных состояниях: в выпрямленном (под углом  $\approx 90^\circ$  к нити), или «атакующем» (А), и в согнутом (R), т. е. наклонном под углом  $\approx 45^\circ$  к миозиновой нити. При этом состояние R обладает двумя подсостояниями: R' (где открыта щель, ведущая от каталитического центра к области связывания «головки» с актином), и R'', где эта щель закрыта. На рис. 25-9 эта щель, служащая и для выхода продуктов расщепления АТФ, и для аллостерической регуляции связывания «головки» с актином, несколько утри-

рована. Переход из состояния R в состояние A осуществляется поворотом двух доменов головки миозина («рычага» и «моторного домена») вокруг  $\alpha$ -спирали лежащего на их стыке «конверторного домена».

Кроме того, головка может быть в двух макроскопических состояниях по связи с актином: либо в связанном ( $a^+$ ), либо в несвязанном ( $a^-$ ).

Эти четыре состояния миозиновой головки образуют «макроскопический цикл» (см. рис. 25-9): ( $R'', a^+ \rightarrow R', a^+ \rightarrow R'', a^- \rightarrow A, a^- \rightarrow A, a^+ \rightarrow \dots$ ). Обратите внимание, что состояний в этом цикле — *больше, чем два*: из двух состояний направленного цикла не построишь (возможны только колебания туда-сюда) и, значит, *направленного* движения не получишь.

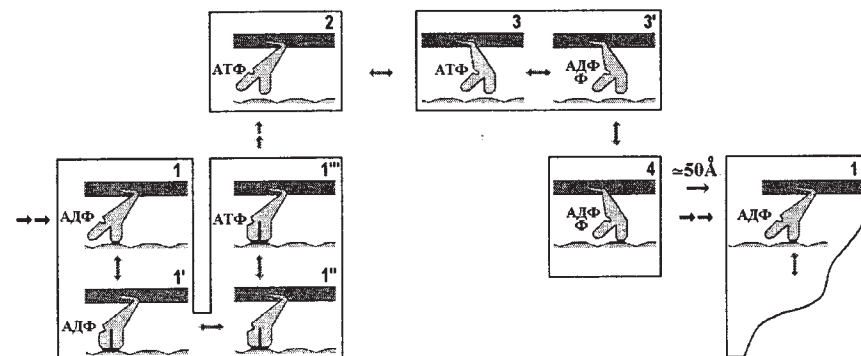
Кроме того, головка может находиться в четырех состояниях по связи с малыми лигандами (АТФ, АДФ и Ф): она может быть свободна от этих лигандов, может связывать АТФ, может связывать оба продукта гидролиза АТФ (АДФ и фосфат Ф) и может связывать один АДФ. Эти четыре состояния образуют «микроскопический цикл».

«Микроскопический цикл» сопряжен с «макроскопическим» (см. рис. 25-9): первый предоставляет энергию, а второй преобразует ее в сокращение мышцы. Для этого сопряжения необходимо, чтобы «несанкционированные» реакции были бы практически исключены, так что *каждый* шаг «микроскопического цикла» мог бы произойти *только* на определенном этапе «макроскопического цикла».

Такое сопряжение может произойти лишь в том случае, если и изменение конформации «головки», и изменение ее связанности с актином изменяет состояние центра, связывающего АТФ или АДФ и Ф, и — в определенном своем состоянии — катализирующего превращение АТФ в АДФ + Ф. А так как этот центр находится в 40 Å от места связывания актина (и примерно там же, где находится точка изгиба миозиновой головки), то связывание с актином вызывает изменение АТФ/АДФ-ного центра аллостерически — путем изменений конформации щели, соединяющей эти центры (примерно так же, как связывание  $O_2$  с одним гемом вызывает изменение константы связи  $O_2$  с другим гемом гемоглобина). Эта связь центров, естественно, действует и в обратную сторону: комплексы миозин-АТФ и миозин-АДФ-Ф слабо связываются с актином, а константа связывания с актином комплекса миозин-АДФ и «пустой» миозиновой головки в 10 000–100 000 раз больше.

В ходе каждого цикла расходуется одна молекула АТФ — тогда, когда цикл идет в сторону, вызывающую сокращение мышцы, а не в противоположную (в последнем случае АТФ, в принципе, должен был бы синтезироваться, хотя бы и чрезвычайно медленно; впрочем, для обращения реакции  $АТФ \rightarrow АДФ + Ф$  надо было бы снизить концентрацию АТФ с «обычной», миллимолярной, до субпиколярной — а это привело бы к распаду актиновых

нитей!). Свободная энергия, получаемая при гидролизе АТФ [ $\approx 15$  ккал/моль АТФ при типичных для мышц концентрациях АТФ ( $\approx 10^{-3}$  М), АДФ ( $\approx 10^{-5}$  М) и Ф ( $\approx 10^{-3}$  М)], расходуется на сокращение мышцы, закрепляя каждый ведущий к нему шаг.



**Рис. 25-9.** Механохимический цикл миозиновой головки. Изображены макроскопические состояния миозина и связанные с ним на разных стадиях этого цикла малые лиганды. Головка миозина сильно связана с актином на стадиях 1', 1'', 1''' (макроскопическое состояние R', a+), слабо — на стадиях 4 (A, a+) и 1 (R'', a+), а на стадиях 2 (R'', a-) и 3, 3' (A, a-) она с актином не связана. На стадиях 1 (R'', a+), 1', 1'', 1''' (R', a+), 2 (R'', a-) головка согнута, а на стадиях 3 (A, a-), 3' (A, a-) и 4 (A, a+) — выпрямлена. При переходе от стадии 1''' к 2 в моторном домене головки открывается щель (т. е. конформация R' переходит в R''), а при переходе от 1 к 1' (от R'' к R') — щель закрывается. В несвязанной головке (a-) щель открыта. При переходе от стадии 2 к 3 и от 4 к 1 головка меняет свой изгиб, причем переход от 4 к 1 (во время которого она связана с актином) сопровождается «силовым ударом», ведущим к сдвигу всего миозинового волокна относительно актинового примерно на 50 Å. Переход от стадии к стадии является либо переходом по макроскопическому циклу (связанному с изменением конформации миозина или его связи с актином), либо переходом по микроскопическому циклу (связанному с изменением связи миозина с малыми лигандами, т. е. АТФ или продуктами его гидролиза, АДФ и Ф). Переход от состояния 4 к состоянию 1 состоит из двух элементарных актов (одного — по макроциклу, и одного — по микроциклу; это изображено двойной стрелкой); порядок этих актов еще не вполне выяснен: отход фосфата (Ф) от миозина может либо предшествовать силовому удару, либо следовать за ним. В переходе 1'''  $\rightarrow$  2, также отмеченному двойной стрелкой, открытые щели в головке может либо предшествовать отрыву от актина, либо следовать за ним. Те переходы, которые практически необратимы при существующих внутри мышцы концентрациях АТФ, АДФ и Ф, показаны односторонними стрелками, а те, что легко обратимы, — двусторонними (обратите внимание, что необратимы «силовой удар» и отход связавшей АТФ головки от актина, а само связывание АТФ и его расщепление — обратимы)

Исследователей мышц издавна занимал вопрос о том, каким устройством энергия химической реакции ( $ATP \rightarrow ADP + P$ ) превращается мышечную тягу. Ведь сам химический акт (сколько бы времени ни затрачивалось на его подготовку и на ожидание флуктуации, преодолевающей активационный барьер) занимает время порядка времени колебания валентной связи, т. е.  $\sim 10^{-14}$  с. При этом участвующие в химическом превращении атомы смещаются примерно на  $1 \text{ \AA}$ . А миозиновая нить смещается на длину «силового удара» ( $\approx 50 \text{ \AA}$ ) за  $\sim 10^{-3}$  с — время, триллион раз большее, чем химический акт! Где же запасается химическая энергия на это время? Существует гипотеза, что она запасается в каких-то эластичных элементах белка. Однако поиск этих элементов пока не дал определенных результатов: все элементы белка обладают некоторой эластичностью — но найти тот, что за счет эластичности обеспечивает смещение на  $50 \text{ \AA}$ , так и не удалось. Однако — нужен ли такой эластичный элемент для работы мышцы? Представляется, что для «силового удара» достаточен сам по себе поворот рычага при переходе из менее стабильной (при наличии связанного АДФ) атакующей (А) конформации в более стабильную (опять-таки, при наличии АДФ) согнутую (R) конформацию. И неважно, будет ли такой переход идти все время «под уклон» (как при релаксации эластичного напряжения), или будет требовать «неэластичного» переваливания через какой-то барьер — при условии, что тот не высок и преодолевается тепловой флуктуацией за биологически-разумное время.

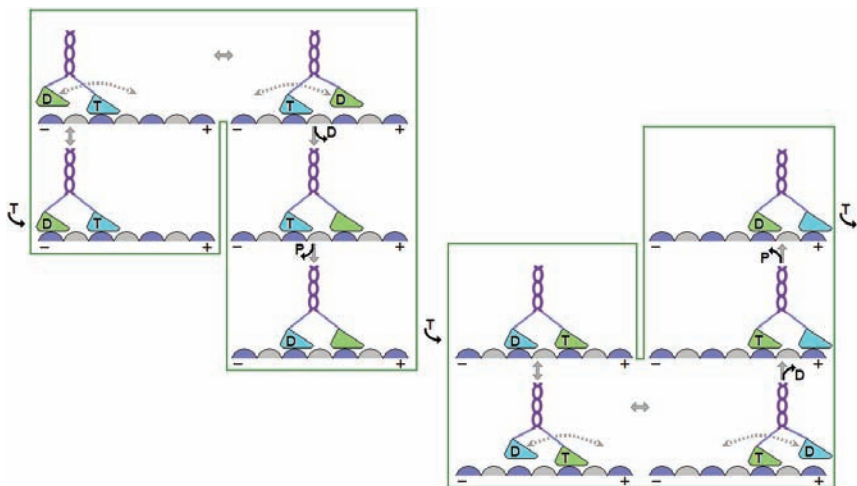
Другая проблема. Рисунок 25-9 может создать впечатление, что работа мышцы требует, по крайней мере на стадии «силового удара», синхронного структурного перехода во всех близко друг от друга расположенных и связанных с актином головках миозиновой нити.

Это впечатление, однако, неверно. «Работающие» головки не расположены в миозиновой нити близко друг от друга. В каждый момент времени «работает», толкая актин, только 1–2 % головок, а остальные ждут своей очереди. Эксперимент показывает, что в работающей мышце каждая головка миозина щепит около 20 молекул АТФ в секунду. Однако за  $1/20$  с головка смещается относительно актиновой нити не на  $50 \text{ \AA}$  (длину шага «силового удара»), а на  $3000\text{--}4000 \text{ \AA}$ ! Значит, 98–99% времени от этой  $1/20$  с головка, не касаясь актина, едет на чужих плечах (и, готовясь к «силовому удару», претерпевает изображенные на рис. 25-9 стадии 2, 3, 3'). Потом, менее чем за тысячную долю секунды, она наносит свой удар по актину — и едет на чужих плечах (точнее, «головах») дальше... В результате число связанных с актином головок, в расчете на миозиновую нить, исчисляется лишь штуками (из сотен в этой нити присутствующих), а расстояние между упирающимися в актин головками составляет

тысячи ангстрем — что, по-видимому, исключает необходимость жесткой синхронизации их действий.

Сочетание крупномасштабного конформационного изменения головки с чередованием ее связанного и не связанного с актином состояний напоминает движение весла при гребле (которое тоже — цикл: «весло позади в воде — весло позади в воздухе — весло впереди в воздухе — весло впереди в воде — весло позади в воде — ...», ср. с рис. 25-9). Итак, миозиновая нить движется по актиновой примерно так же, как лодка по воде (точнее — как многовесельная лодка, где каждый момент гребет, расходуя свою молекулу АТФ, лишь один гребец из ста).

В заключение я хочу показать вам, как работает белок кинезин (я буду говорить об одном из типов кинезина, о кинезине 1). Этот белок, сжигая АТФ, транспортирует мРНК, белковые комплексы и т. д. вдоль микротрубочек, делая один шаг длиной 8 нм от ее «-» к «+» концу на каждую расщепленную молекулу АТФ (рис. 25-10). Кинезин, как и миозин, сопрягает «макроскопический», т. е. механический цикл перемещения своих «головок» (на самом деле, служащих ногами) с «микроскопическим», т. е. химическим циклом расщепления АТФ. В ходе «микроцикла» головка кинезина («моторный домен») может быть связана с АТФ, может быть связана с АДФ и может быть «пустой». В ходе «макроцикла» головка кинезина может иметь 4 положения. Во-первых, она может быть «задней» (т. е. находящейся позади «шеи» димера, ближе к «-» концу микротрубочки), и может быть «передней» (т. е. находящейся впереди «шеи», ближе к «+» концу микротрубочки). Во-вторых, в каждом из этих двух положений головка может быть связанной с тубулином и может быть несвязанной с ним. При этом, из-за взаимодействия с по-разному («впереди» и «сзади») относительно головки расположенным линкером, «задняя» и «передняя» головки имеют чуть разные конформации, и потому — разные свойства. Подтверждаемые экспериментом условия, необходимые для однонаправленного движения кинезина по микротрубочке: (1) АТФ расщепляется *только* «задней» головкой; (2) АДФ может уйти *только* из «передней» головки; (3) головка, на которой сидит АДФ, связывается с тубулином слабо (и в «заднем», и в «переднем» положении), так что такая головка может отсоединиться от турбулина; (4) «пустая» головка, а также головка, связавшая АТФ, сильно связана с тубулином и практически не может отсоединиться от него. При этом кинезин все время связан с микротрубочкой хотя бы одной головкой, а другая может диффузионно перемещаться из того положения, где АДФ не может от нее оторваться, в то, где АДФ может эту головку покинуть.



**Рис. 25–10.** Схема движения димерного кинезина 1 по тубулиновой микротрубочке. Верхушки глобул тубулина (их диаметр — 4 нм) изображены полусферами. Скругленными треугольниками двух разных цветов изображены две головки кинезина, двойной спиралью — их общая шея, линиями — «линкеры», соединяющие шею с головками участки кинезина. Обе головки кинезина одинаковы и прикрепляются к микротрубочке в одинаковой (относительно тубулина) ориентации. Показан полный акт перемещения кинезина, состоящий из двух шагов, каждый из которых сопрягает «макроцикл» с «микроциклом». Каждая светло-серая стрелка отвечает одному элементарному акту движения головки кинезина. Длинная пунктирная серая стрелка показывает диффузию головки на плюс или минус 16 нм (при этом шея кинезина — по ней удобно следить за смещением этого димерного белка, как целого — смещается на 8 нм). Каждая черная стрелка отвечает одному элементарному акту изменения состояния субстрата (Т = АТФ, D = АДФ, P = фосфат). Те переходы, которые практически необратимы при существующих внутри клетки концентрациях АТФ, АДФ и фосфата, показаны однонаправленными стрелками, а те, что легко обратимы, — двунаправленными. Рисунок составлен по статье Yildiz, Tomishige, Gennerich, Vale, *Cell* (2008) 134:1030–1041

В результате кинезин шагает по тубулину, мышцы сокращаются, легкие качают воздух, кислород связывается с гемоглобином, который передает его миоглобину, который передает его мышцам и другим органам, — и так мы дышим и живем.



## *Послесловие*

К моему глубочайшему сожалению, мне приходится дописывать эти страницы в одиночестве: мой учитель, О. Б. Птицын, скончался... Поэтому за все недостатки этого курса несу полную ответственность я один, и я буду благодарен читателям, мне о них сообщившим.

А. В. Финкельштейн  
2000 г.  
afinkel@vega.protres.ru

## РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА

(послужившая также источником многих иллюстраций к этим лекциям)

- [1] Фершт Э. *Структура и механизм действия ферментов*. — М.: Мир, 1980. Гл. 1, 8–12.  
Fersht A. *Enzyme structure and mechanism*. 2nd ed. — N. Y.: W. H. Freeman & Co., 1985. Ch. 1, 8–12.  
Fersht A. *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*. — N. Y.: W. H. Freeman & Co., 1999.
- [2] Волькенштейн М. В. *Биофизика*. — М.: Наука, 1981. Гл. 4, 6.
- [3] Шульц Г. Е., Ширмер Р. Х. *Принципы структурной организации белков*. — М.: Мир, 1982.  
Schulz G. E., Schirmer R. H. *Principles of protein structure*. — N. Y.; Heidelberg; Berlin: Springer-Verlag, 1979.
- [4] Кантор Ч., Шиммель П. *Биофизическая химия*. — М.: Мир, 1982. Т. 1. Гл. 2, 5; Т. 3. Гл. 17, 20, 21.  
Cantor C. R., Schimmel P. R. *Biophysical chemistry*. — San Francisco: W. H. Freeman & Co, 1980. Part 1. Ch. 2, 5; Part 3. Ch. 17, 20, 21.
- [5] Branden C, Tooze J. *Introduction to Protein Structure*. — N. Y.; London: Garland Publ., Inc., 1991; 2nd ed., 1999.
- [6] Creighton T. E. *Proteins*. 2-nd ed. — N. Y.: W. H. Freeman & Co., 1991.
- [7] Perutz M. F. *Protein structure*. — N. Y.: W. H. Freeman & Co., 1992.
- [8] Ленинджер А. *Основы биохимии*. В 3-х т. — М.: Мир, 1985. Гл. 4–8, 23, 29.  
Leninger A. L., Nelson D. L., Cox M. X. *Principles of biochemistry*. 2nd ed. — N. Y.: Worth Publ. Inc., 1993. Ch. 3–8, 21, 26.
- [9] Страйер Л. *Биохимия*. В 3-х т. — М.: Мир, 1984 (т. 1); 1985 (т. 2, 3). Гл. 1–9, 27, 33–34.

Stryer L. *Biochemistry*. 4th ed. — N. Y.: W. H. Freeman & Co., 1995. Ch. 1–3, 7–16, 34–36.

- [10] Полинг Л. *Общая химия*. — М.: Мир, 1974. Гл. 1–6, 9–13, 16, 24.  
Pauling L. *General chemistry*. — N. Y.: W. H. Freeman & Co., 1970. Ch. 1–6, 9–13, 16, 24.
- [11] Степанов В. М. *Молекулярная биология. Структуры и функции белков*. — М.: Высшая Школа, 1996.
- [12] Чернавский Д. С., Чернавская Н. М. *Белок-машина. Биологические макромолекулярные конструкции*. — М.: Изд-во МГУ, 1999.
- [13] Рубин А. Б. *Биофизика*. — М.: Книжный дом «Университет», 1999. Т. 1. Гл. 7–14.
- [14] Эмануэль Н. М., Кнорре Д. Г. *Курс химической кинетики*. 4-е изд. — М.: Высшая школа, 1984.
- [15] Howard J. *Mechanics of motor proteins and the cytoskeleton*. — Sunderland, Massachusetts: Sinauer Associates, Inc., 2001. Part III.
- [16] Сердюк И. Н., Заккаи Н., Заккаи Дж. *Методы в молекулярной биофизике. Структура. функция. динамика. В двух томах*. — М.: Книжный дом «Университет», 2010.

Мы избегали ссылок в тексте для облегчения его чтения, однако ссылки на оригинальные работы и обзоры даны в подписях к рисункам и таблицам.

Рисунки белков извлечены из Банка белковых структур (его создание описано в: Bernstein F. C., Koetzle T. F, Meyer E. F., Jr, Brice M. D., Kennard O., Shimanouchi T., Tasumi T. // *J. Mol. Biol.* 1977. V. 112. P. 535–542) при помощи программ MOLSCRIPT (Kraulis P. J. // *J. Appl. Cryst.* 1991. V. 24. P. 946–950) и WHAT IF (Vriend G. // *J. Mol. Graphics.* 1990. V. 8. P. 52–56), RasMol (Sayle R., Milnerwhite E. J. // *Trends Biochem. Sci.* 1999. V. 20. P. 374–376), Insight II (*Molecular Simulations Inc.*, 1998) и ViewerLite (*Accelrys Inc.*, 2001).

## ЗАДАЧИ, РЕШЕНИЯ, КОММЕНТАРИИ

Изучение точных наук предполагает, в частности, обучение решению задач. Отсутствие задач было существенным недостатком первых редакций этой книги. В третьем русском издании я постарался восполнить этот пробел, а в этом — еще увеличил число задач, особенно сложных.

Так как книга рассчитана на людей с существенно разной профессиональной подготовкой, некоторым одни задачи покажутся излишне сложными, а другие — слишком элементарными. «Слишком элементарные» задачи появились, в основном, как реакция на письма читателей (которые, будучи хорошими специалистами в одной области науки, испытывали затруднения в других). По той же причине я дал довольно подробно решения задач даже там, где большинству читателей они покажутся излишними.

«Слишком сложные» задачи (и комментарии к ним), в сущности, дополняют курс лекций.

А. В. Финкельштейн, 2004–2011

### К лекции 2

#### Задача 2.1

Аминокислоты бывают «левыми» и «правыми». Бывает ли «левая» и «правая» вода? А этиловый спирт? Почему?

#### Решение

«Левыми» и «правыми» могут быть только те молекулы, которые нельзя совместить с их зеркальным отражением путем смещения и поворота в трехмерном пространстве и вращения вокруг валентных связей.

Молекула воды, Н-О-Н, состоит всего из трех атомов. Три точки (ядра Н, О, Н; пространственной конфигурацией электронов мы пренебрегаем)

всегда лежат в одной плоскости, а зеркальное отражение плоского тела всегда можно совместить с исходным телом путем смещения и переворота плоскости в трехмерном пространстве (проверьте на произвольном треугольнике!). Так что «левой» и «правой» воды не бывает никогда.

«Левыми» и «правыми» могут быть, в принципе, только трехмерные тела. Так что мгновенная конфигурация молекулы из четырех и более атомов (например,  $\text{CH}_3\text{CH}_2\text{OH}$ ) может отличаться от своего зеркального отражения. Однако, заменив в такой зеркально-отраженной молекуле все углы внутреннего вращения  $\chi_i$  на  $-\chi_i$ , мы приходим к исходной конфигурации  $\text{CH}_3\text{CH}_2\text{OH}$ . Поэтому зеркальные отражения  $\text{CH}_3\text{CH}_2\text{OH}$  («левый» и «правый» спирт) могут различаться только до тех пор, пока тепловое движение не устранил память об их исходной конфигурации, т. е. в течение  $\sim 10^{-10}$  с при комнатной температуре (и гораздо большее время — при сверхнизких температурах); однако на больших временах наблюдения «левого» и «правого»  $\text{CH}_3\text{CH}_2\text{OH}$  не бывает.

Но вот «тяжелый» этиловый спирт типа  $\text{CH}_3\text{CDH-OH}$  (где D — дейтерий) имеет «правую» и «левую» формы! Здесь у центрального атома С четыре, *не лежащие в одной плоскости* ковалентные связи, причем все — с *разными* атомами (С, Н, D, О), а переход атома с одной связи на другую требует большой активационной энергии и потому занимает практически бесконечное время.

А «тяжелый» этиловый спирт типа  $\text{CDH}_2\text{CH}_2\text{-OH}$  *не* имеет «правой» и «левой» форм: у его центрального атома С две из четырех ковалентных связей (с С, Н, Н, О) одинаковы, и у первого атома С (связанного с С, Н, Н, D) — тоже, так что «зеркально-отраженная» форма этой молекулы превращается в исходную путем поворота вокруг валентных связей (как то было с  $\text{CH}_3\text{CH}_2\text{OH}$ ).

#### Задача 2.2

Тепловые, имеющие энергию  $kT$  колебания имеют частоту  $\nu_T \approx 7 \times 10^{12} \text{ c}^{-1}$  при  $27^\circ\text{C}$ . Какова частота тепловых колебаний при  $-200^\circ\text{C}$ ? При  $+2700^\circ\text{C}$ ?

#### Решение

Частота тепловых колебаний определяется из уравнения  $kT \approx h\nu_T$ , где  $T$  — абсолютная температура.  $27^\circ\text{C}$  соответствует абсолютной температуре  $T = 273 + 27 = 300\text{K}$ ,  $-200^\circ\text{C}$  — абсолютной температуре  $T' = 273 - 200 = 73\text{K}$ , т. е. вчетверо более низкой,  $+2700^\circ\text{C}$  — абсолютной температуре  $T'' = 273 + 2700 = 2973\text{K}$ , т. е. вдесятеро более высокой. Поэтому частота тепловых колебаний при  $-200^\circ\text{C}$  вчетверо ниже, чем при  $27^\circ\text{C}$ , и составляет  $\approx 1.7 \times 10^{12} \text{ c}^{-1}$ , а при  $+2700^\circ\text{C}$  вдесятеро выше, чем при  $27^\circ\text{C}$ , и составляет  $\approx 7 \times 10^{13} \text{ c}^{-1}$ .

### Задача 2.3

Частота колебания С–Н связи,  $\nu_{\text{С–Н}}$ , составляет  $7 \times 10^{13} \text{ с}^{-1}$ . Как зависит это колебание от температуры?

#### Решение

Частота колебания — не зависит от температуры: она зависит только от жесткости связи и масс связанных атомов. Возбужденность колебания — зависит от температуры: колебание практически не возбуждено, пока  $\nu_{\text{С–Н}} > \nu_T$ , и возбуждается, только когда  $\nu_T$  поднимается до  $\nu_{\text{С–Н}}$ , т. е. до  $\approx +2700 \text{ }^\circ\text{C}$  (см. предыдущую задачу).

### Задача 2.4

Найти длину световой волны, соответствующей колебанию с частотой  $\nu = 7 \times 10^{12} \text{ с}^{-1}$ .

#### Решение

Длина световой волны  $\lambda_{\text{свет}} = c/\nu$ , где  $c$  — скорость света ( $300\,000 \text{ км/с}$ ). Значит,  $\lambda = [3 \times 10^8 \text{ м/с}]/[7 \times 10^{12} \text{ с}^{-1}] \approx 0,4 \times 10^{-4} \text{ м} = 40 \text{ мк}$ .

### Задача 2.5

Найти длину звуковой волны, соответствующей колебанию с частотой  $\nu = 7 \times 10^{12} \text{ с}^{-1}$ , считая, что в растворе скорость звука  $u = 1500 \text{ м/с}$ .

#### Решение

Формально,  $\lambda_{\text{звук}} = u/\nu = [1500 \text{ м/с}]/[7 \times 10^{12} \text{ с}^{-1}] = 0,2 \times 10^{-9} \text{ м} = 2 \text{ \AA}$ . Однако такая звуковая волна сильно затухает, поскольку ее длина близка к межатомному расстоянию.

### Задача 2.6

Электрон колеблется вокруг ядра с круговой частотой  $\omega = 2 \times 10^{16} \text{ с}^{-1}$ . Оценить (а) амплитуду и (б) кинетическую энергию его колебания, исходя из принципа Гейзенберга.

#### Решение

Согласно принципу неопределенности Гейзенберга, неопределенность импульса ( $\Delta p = m\Delta v$ ) и неопределенность координаты ( $\Delta x$ ) связаны соотношением  $\Delta p\Delta x \approx \hbar$ . Здесь  $\hbar = 1,05 \times 10^{-34} \text{ Дж} \times \text{с}$  — «квант момента количества движения», связанный с постоянной Планка ( $h = 6,6 \times 10^{-34} \text{ Дж} \times \text{с}$ ) соотношением  $\hbar = h/2\pi$ .

При колебании с амплитудой  $A$  координата меняется от  $+A$  до  $-A$ , т. е. ее неопределенность  $\Delta x = [A - (-A)]/2 = A$ , а скорость, при круговой частоте  $\omega$ , меняется от  $+\omega A$  до  $-\omega A$ , т. е. ее неопределенность  $\Delta v = \omega A$ . Итак,  $\Delta p\Delta x = m\omega A^2 \approx \hbar$ , откуда: амплитуда  $A \approx (\hbar/m\omega)^{1/2}$ , а кинетическая энергия  $E = m(\omega A)^2/2 \approx \hbar\omega/2$ .

При массе электрона  $m = 0,91 \times 10^{-30} \text{ кг}$  и круговой частоте  $\omega = 2 \times 10^{16} \text{ с}^{-1}$ , имеем:  $A \approx 0,75 \times 10^{-10} \text{ м} = 0,75 \text{ \AA}$ ,  $E \approx 1 \times 10^{-18} \text{ Дж} \approx 0,25 \times 10^{-18} \text{ кал} = 150 \text{ ккал/моль}$ .

### Задача 2.7

Протон гармонически колеблется в С–Н связи с частотой  $\nu = 7 \times 10^{13} \text{ с}^{-1}$ . Оценить амплитуду его колебания, исходя из неопределенности Гейзенберга.

#### Решение

Частота  $\nu$  связана с круговой частотой  $\omega$  соотношением  $\nu = \omega/2\pi$ , т. е.  $\omega = 4,4 \times 10^{14} \text{ с}^{-1}$  при  $\nu = 7 \times 10^{13} \text{ с}^{-1}$ . Исходя из полученного в предыдущей задаче соотношения: амплитуда  $A \approx (\hbar/m\omega)^{1/2}$ , имеем, при массе протона  $m = 1,67 \times 10^{-27} \text{ кг}$  и частоте  $\nu = 7 \times 10^{13} \text{ с}^{-1}$ : амплитуда  $A \approx 0,12 \text{ \AA}$ .

### Задача 2.8

Оценить квантовую неопределенность координаты частицы, испытывающей тепловые колебания. Молекулярный вес частицы: (а) — 1 дальтон (Н-атом); (б) — 18 дальтон (молекула воды); (в) — 100 дальтон (аминокислотный остаток).

#### Решение

Согласно принципу Гейзенберга,  $m\Delta v\Delta x \approx \hbar$ , где  $\Delta v$  — квантовая неопределенность скорости частицы, а  $\Delta x$  — неопределенность ее координаты. При колебаниях  $\Delta v \approx \omega\Delta x$  (см. задачу 2.6), где  $\omega$  — круговая частота колебания, так что  $\Delta x \approx (\hbar/m\omega)^{1/2}$ . Характерная круговая частота тепловых колебаний определяется из уравнения  $\hbar\omega_T \approx kT$ . Поэтому квантовая неопределенность координаты частицы, испытывающей тепловые колебания, есть  $\Delta x_T \approx (\hbar/m\omega_T)^{1/2} = \hbar/(mkT)^{1/2}$ .

При  $T = 300 \text{ К}$  [т. е.  $kT = 2,5 \times 10^3 \text{ Дж/моль} = 2,5 \times 10^3 \text{ Дж}/(0,6 \times 10^{24})$ ] и массе частицы  $m = 1$  дальтон =  $1,67 \times 10^{-27} \text{ кг}$ , имеем  $\Delta x_T \approx 0,4 \text{ \AA}$ ; при  $m = 18$  дальтон =  $18 \times (1,67 \times 10^{-27} \text{ кг})$ ,  $\Delta x_T \approx 0,09 \text{ \AA}$ ; и при  $m = 100$  дальтон,  $\Delta x_T \approx 0,04 \text{ \AA}$ .



## К лекции 3

### Задача 3.1 (сложная)

*Предварительная информация.* Простейшая модель Ван-дер-Ваальсового взаимодействия пары атомов выглядит следующим образом:

Даны два атома, «1» и «2», в каждом из которых электрон «гармонично», т. е. согласно уравнению

$$m \left[ \frac{d}{dt} \left( \frac{d\vec{x}_i}{dt} \right) \right] = -k_i \vec{x}_i \quad (i = 1, 2) \quad (3.1.1)$$

колеблется у своего ядра. Здесь  $m$  — масса электрона,  $t$  — время,  $\vec{x}_i$  — трехмерная координата электрона атома  $i$  относительно ядра этого атома, и  $k_i$  — константа жесткости связи электрона с ядром в атоме  $i$ . Потенциальная энергия растяжения связи электрона с ядром, в этом приближении, есть

$$U_i = \frac{1}{2} k_i \vec{x}_i^2. \quad (3.1.2)$$

При этом

$$\omega_i = (k_i/m)^{1/2}, \quad (3.1.3)$$

круговая частота колебаний электрона в атоме  $i$ . Согласно квантовой механике, полная (кинетическая+потенциальная) энергия колебаний невозбужденного осциллятора  $i$  есть

$$E_i = \frac{1}{2} \hbar \omega_i = \frac{1}{2} \hbar (k_i/m)^{1/2}, \quad (3.1.4)$$

где  $\hbar$  — квант момента количества движения, т. е. постоянная Планка  $h$ , деленная на  $2\pi$ . См. задачу 2.6, где соотношение типа 3.1.4 выводится прямо из принципа неопределенности Гейзенберга.

Если электрон атома  $i$  смещен от своего ядра, в атоме  $i$  возникает дипольный момент  $\vec{P}_i = -e\vec{x}_i$ , где  $-e$  — заряд электрона. Между диполями «электрон — ядро в атоме 1» и «электрон — ядро в атоме 2» возникает взаимодействие. Если атомы находятся в вакууме на расстоянии  $r$  друг от друга, и  $r \gg |\vec{x}_1|, |\vec{x}_2|$ , то энергия взаимодействия диполей есть

$$\Delta U = \frac{1}{r^3} [P_1'' P_2'' + P_1''' P_2''' - 2P_1' P_2'] \quad (3.1.5)$$

(где  $P_1', P_2'$  — проекции диполей атомов «1» и «2» на линию, соединяющую эти атомы, а  $P_1'', P_2''$  и  $P_1''', P_2'''$  — перпендикулярные этой линии компоненты диполей; см. задачу 6.2; предполагается, что пара взаимодействующих атомов находится в вакууме). Наличие этого дополнительного взаимодей-

ствия диполей (которое может быть и притягивающим, и отталкивающим, в зависимости от знака скалярного произведения  $\vec{x}_1 \vec{x}_2$ ) как-то меняет колебания электронов в атомах. В результате меняется их суммарная энергия, и...

*Вопросы.*

- (а) будут ли атомы притягиваться или отталкиваться?
- (б) как зависит энергия их взаимодействия (Ван-дер-Ваальсова энергия) от расстояния  $r$ ?
- (в) как зависит эта энергия от жесткости связи электронов с ядрами?
- (г) как зависит эта энергия от массы электрона?
- (д) как зависит эта энергия от частот колебаний электронов во взаимодействующих атомах?

*Решение*

Рассмотрим сначала одномерную задачу, т. е. будем рассматривать только те составляющие колебаний электронов, что коллинеарны соединяющей наши атомы линии. Обозначим одномерную координату электрона  $i$  через  $x_i$ . Тогда

$$U_i = \frac{1}{2} k_i x_i^2, \quad (3.1.6)$$

и 
$$\Delta U = -2 \frac{1}{r^3} P_1' P_2' = -2 \frac{e^2}{r^3} x_1 x_2. \quad (3.1.7)$$

Полную потенциальную энергию взаимодействующих атомов,  $U = U_1 + U_2 + \Delta U$ , можно представить в виде

$$U(x_1, x_2) = \frac{1}{2} k_1 x_1^2 + \frac{1}{2} k_2 x_2^2 + \gamma x_1 x_2, \quad (3.1.8)$$

где  $\gamma = -2e^2/r^3$ . Это — двумерная потенциальная яма центром в точке  $(x_1 = 0, x_2 = 0)$ . В этой яме колеблются два электрона. Их полная кинетическая энергия

$$W = \frac{1}{2} m [(dx_1/dt)^2 + (dx_2/dt)^2]. \quad (3.1.9)$$

Эквипотенциальные линии, подчиняющиеся уравнению  $U(x_1, x_2) = \text{const}$ , представляют собой концентричные эллипсы, причем подобные (т. е. эллипс, соответствующий уравнению  $U(x_1, x_2) = C$ , — это просто в  $C^{1/2}$  раз увеличенный эллипс, соответствующий уравнению  $U(x_1, x_2) = 1$ ). Из геометрии известно, что большая и малая оси эллипса перпендикулярны друг другу.

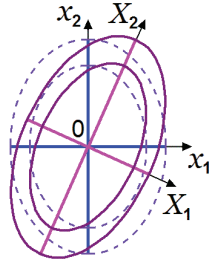
Если  $\gamma = 0$ , то оси эллипса  $\frac{1}{2} k_1 x_1^2 + \frac{1}{2} k_2 x_2^2 = C$  (такие эллипсы изображены на схеме пунктиром) направлены по осям координат. При этом колебание по координате  $x_1$  никак не влияет на колебание по координате  $x_2$ .

Если  $\gamma \neq 0$ , то колебания по координатам  $x_1$  и  $x_2$  влияют друг на друга:

$$m\left[\frac{d}{dt}\left(\frac{dx_1}{dt}\right)\right] = -k_1x_1 - \gamma x_1,$$

$$m\left[\frac{d}{dt}\left(\frac{dx_2}{dt}\right)\right] = -k_2x_2 - \gamma x_1.$$

При  $\gamma \neq 0$  оси эллипса  $\frac{1}{2}k_1x_1^2 + \frac{1}{2}k_2x_2^2 + \gamma x_1x_2 = C$  (такие эллипсы изображены на схеме тонкой линией) наклонены по отношению к осям координат  $(x_1, x_2)$ , длины осей тоже несколько меняются при  $\gamma \neq 0$ , но все равно оси эллипса остаются перпендикулярными друг другу (причем направления осей эллипсов, соответствующих разным  $C$ , естественно, совпадают).



Иначе говоря, повернув систему координат на определенный угол, мы увидим, что эллипс, подчинявшийся уравнению  $\frac{1}{2}k_1x_1^2 + \frac{1}{2}k_2x_2^2 + \gamma x_1x_2 = C$  в «старой» системе координат  $(x_1, x_2)$ , есть эллипс, подчиняющийся уравнению  $\frac{1}{2}K_1X_1^2 + \frac{1}{2}K_2X_2^2 = C$  в «новой» (и тоже ортогональной) системе координат  $(X_1, X_2)$ .

В этой «новой», просто повернутой, системе координат кинетическая энергия электронов сохраняет свой прежний вид (3.1.9):  $W = \frac{1}{2}m[(dX_1/dt)^2 + (dX_2/dt)^2]$ . При этом колебание по координате  $X_1$  никак не влияет на колебание по координате  $X_2$ : каждое из них подчиняется уравнению

$$m\left[\frac{d}{dt}\left(\frac{dX_i}{dt}\right)\right] = -K_iX_i \quad (i = 1, 2),$$

отличающемся от уравнения (3.1.1) для невзаимодействующих атомов только заменой  $x_i$  на  $X_i$  и, главное,  $k_i$  на  $K_i$ . Фактически, мы переходим таким образом от двух взаимосвязанных электронов к двум гармонически и независимо друг от друга колеблющимся квазичастицам. Каждая из них имеет ту же, что электрон, массу  $m$ , но измененную (по сравнению с электроном) константу жесткости  $K_i$  связи с ядром атома.

В результате эти квазичастицы колеблются с измененными круговыми частотами  $\Omega_1 = (K_1/m)^{1/2}$ ,  $\Omega_2 = (K_2/m)^{1/2}$  и, согласно квантовой механике,

полная энергия двух таких взаимодействующих осцилляторов (атомов) есть

$$E = \frac{1}{2}\hbar\Omega_1 + \frac{1}{2}\hbar\Omega_2 = \frac{1}{2}\hbar(K_1/m)^{1/2} + \frac{1}{2}\hbar(K_2/m)^{1/2} = \frac{1}{2}(\hbar/m^{1/2})[K_1^{1/2} + K_2^{1/2}]. \quad (3.1.10)$$

Отличие этой энергии от суммы энергий отдельных атомов,

$$\Delta E = [\frac{1}{2}\hbar\Omega_1 + \frac{1}{2}\hbar\Omega_2] - [\frac{1}{2}\hbar\omega_1 + \frac{1}{2}\hbar\omega_2] = \frac{1}{2}(\hbar/m^{1/2})[K_1^{1/2} - k_1^{1/2} + K_2^{1/2} - k_2^{1/2}], \quad (3.1.11)$$

и есть искомая энергия взаимодействия атомов.

Осталось только вычислить изменения величин  $K_1$  и  $K_2$  по сравнению с  $k_1$  и  $k_2$ .

$k_1$  и  $k_2$  являются собственными числами матрицы  $\begin{vmatrix} k_1 & 0 \\ 0 & k_2 \end{vmatrix}$ : они соот-

ветствуют осям эллипсов, заданных уравнением  $\frac{1}{2}k_1x_1^2 + \frac{1}{2}k_2x_2^2 = \text{const}$ . Соответственно,  $K_1$  и  $K_2$ , соответствующие осям эллипсов, заданных уравнением  $\frac{1}{2}k_1x_1^2 + \frac{1}{2}k_2x_2^2 + \gamma x_1x_2 = \text{const}$ , являются собственными числами матрицы  $\begin{vmatrix} k_1 & \gamma \\ \gamma & k_2 \end{vmatrix}$  (и, одновременно, повернутой матрицы  $\begin{vmatrix} K_1 & 0 \\ 0 & K_2 \end{vmatrix}$ ).

Как известно из математики, собственные числа ( $K$ ) такой матрицы находятся из уравнения

$$\text{Det} \begin{vmatrix} k_1 - K & \gamma \\ \gamma & k_2 - K \end{vmatrix} \equiv (k_1 - K)(k_2 - K) - \gamma^2 = 0.$$

Решая это квадратное уравнение,  $K^2 - K(k_1 + k_2) + k_1k_2 - \gamma^2 = 0$ , находим оба его корня,

$$K_{1,2} = \frac{k_1 + k_2}{2} \pm [(\frac{k_1 - k_2}{2})^2 + \gamma^2]^{1/2},$$

которые (если выбрать атомы так, что  $k_1 \geq k_2$ ) можно представить в виде:

$$K_1 = k_1 + \Delta,$$

$$K_2 = k_2 - \Delta,$$

где

$$\Delta = [(\frac{k_1 - k_2}{2})^2 + \gamma^2]^{1/2} - |\frac{k_1 - k_2}{2}| \equiv \frac{\gamma^2}{[(\frac{k_1 - k_2}{2})^2 + \gamma^2]^{1/2} + |\frac{k_1 - k_2}{2}|} \geq 0 \quad (3.1.12)$$

(в такого рода вычислениях полезно представлять малые разности больших чисел, типа  $(a+1)^{1/2} - a$ , где  $a > 0$ , в эквивалентном виде  $((a+1)^{1/2} - a) \times [(a+1)^{1/2} + a] / [(a+1)^{1/2} + a] = 1 / [(a+1)^{1/2} + a]$ , в котором вычитания больших чисел нет). Величина  $\Delta > 0$ , если  $\gamma \neq 0$ , но она не превышает  $\gamma$ , т. е. она мала относительно  $k_1$  и  $k_2$ : ведь  $k_1$  и  $k_2$  описывают жесткость связи электрона с близким к нему ядром «своего» атома, а  $\gamma$  — жесткость связи между гораздо более отдаленными друг от друга электронами и ядрами разных атомов.

Возвращаясь к уравнению (3.1.11) видим, что его можно преобразовать к виду

$$\begin{aligned} \Delta E &= \frac{1}{2} (\hbar/m^{1/2}) \cdot \{[(k_1 + \Delta)^{1/2} - k_1^{1/2}] + [(k_2 - \Delta)^{1/2} - k_2^{1/2}]\} = \frac{1}{2} (\hbar/m^{1/2}) \times \\ &\times \{\Delta / [(k_1 + \Delta)^{1/2} + k_1^{1/2}] - \Delta / [(k_2 - \Delta)^{1/2} + k_2^{1/2}]\} = \\ &= -(\hbar/2m^{1/2}) \cdot \Delta \cdot [(k_1 + \Delta)^{1/2} + k_1^{1/2} - (k_2 - \Delta)^{1/2} - k_2^{1/2}] / \\ & / \{[(k_1 + \Delta)^{1/2} + k_1^{1/2}][(k_2 - \Delta)^{1/2} + k_2^{1/2}]\}. \end{aligned} \quad (3.1.13)$$

Так как  $\Delta > 0$  при  $\gamma \neq 0$  и  $k_1 \geq k_2$ , то величина  $\Delta E$  отрицательна при  $\gamma \neq 0$ , т. е. рассматриваемые атомы притягиваются друг к другу.

Это — ответ на вопрос (а).

Для ответа на вопросы последующих задач сначала преобразуем величину  $(k_1 + \Delta)^{1/2} + k_1^{1/2} - (k_2 - \Delta)^{1/2} - k_2^{1/2}$ , стоящую в числителе (3.1.13), к виду

$$\begin{aligned} (k_1 + \Delta)^{1/2} - (k_2 - \Delta)^{1/2} + k_1^{1/2} - k_2^{1/2} &= (k_1 - k_2 + 2\Delta) / [(k_1 + \Delta)^{1/2} + \\ & + (k_2 - \Delta)^{1/2}] + (k_1 - k_2) / [k_1^{1/2} + k_2^{1/2}], \end{aligned} \quad (3.1.14)$$

и затем упростим выражение (3.1.13), пожертвовав в знаменателях формул (3.1.13) и (3.1.14), где нет вычитаний, малыми (по сравнению с  $k_1$  и  $k_2$ ) членами  $\Delta$ :

$$\Delta E = -(\hbar/4m^{1/2}) \cdot \Delta \cdot [(k_1 - k_2) + \Delta] / [k_1^{1/2} k_2^{1/2} (k_1^{1/2} + k_2^{1/2})]. \quad (3.1.15)$$

Наконец, используя формулу (3.1.12) для  $\Delta$ , упростим выражение  $\Delta \cdot [(k_1 - k_2) + \Delta]$ . Легко видеть, что, введя  $Z = |k_1 - k_2|/2\gamma$ , величину  $\Delta[(k_1 - k_2) + \Delta]$  можно представить в виде

$$\begin{aligned} &\gamma^2 \cdot \{2Z / [(Z^2 + 1)^{1/2} + Z] + 1 / [(Z^2 + 1)^{1/2} + Z]^2\} = \\ &= \gamma^2 \cdot \{[2Z \cdot [(Z^2 + 1)^{1/2} + Z] + 1] / [(Z^2 + 1)^{1/2} + Z]^2\} = \\ &\gamma^2 \cdot \{[2Z(Z^2 + 1)^{1/2} + 2Z^2 + 1] / [Z^2 + 2Z(Z^2 + 1)^{1/2} + 1 + Z^2]\} = \gamma^2. \end{aligned}$$

Итак, наконец,

$$\Delta E = -(\hbar/4m^{1/2}) \cdot \gamma^2 / [k_1^{1/2} k_2^{1/2} (k_1^{1/2} + k_2^{1/2})], \quad (3.1.16)$$

или, вспомнив, что  $\omega_i = (k_i/m)^{1/2}$  и  $\gamma = -2e^2/r^3$ ,

$$\Delta E = -\hbar (e^4/m^2 r^6) / [\omega_1 \omega_2 (\omega_1 + \omega_2)]. \quad (3.1.17)$$

Отсюда сразу следуют ответы на вопросы (б)–(д):

— энергия Ван-дер-Ваальсова взаимодействия атомов падает с ростом расстояния  $r$  между ними как  $1/r^6$ ;

— эта энергия падает с увеличением жесткости  $k$  связи электронов с ядрами взаимодействующих атомов;

— эта энергия обратно пропорциональна квадратному корню из массы электрона;

— эта энергия падает с частотами колебаний электронов во взаимодействующих атомах.

*Дополнение.* До сих пор мы рассматривали только те составляющие колебаний электронов, что коллинеарны соединяющей атомы линии. Точно такое же решение можно получить для каждой из двух составляющих, перпендикулярных этой линии, причем, согласно уравнению (3.1.5), взаимодействие между взаимно-перпендикулярными составляющими отсутствует. Согласно (3.1.5), диполь-дипольное взаимодействие между составляющими, перпендикулярными соединяющей атомы линии, вдвое меньше, чем диполь-дипольное взаимодействие между составляющими, параллельными этой линии. Поэтому вклад каждой из двух этих перпендикулярных составляющих в Ван-дер-Ваальсово взаимодействие атомов четверо меньше, чем рассмотренный нами вклад параллельных составляющих. Окончательно, энергия Ван-дер-Ваальсова взаимодействия двух атомов равна

$$\Delta E = -3/2 \hbar (e^4/m^2 r^6) / [\omega_1 \omega_2 (\omega_1 + \omega_2)]. \quad (3.1.18)$$

Такое уравнение называется уравнением Лондона. В учебниках по физике и химии (см. [10]) оно обычно приводится в форме

$$\Delta E = -3/2 [(I_1 \alpha_1) (I_2 \alpha_2) / (I_1 + I_2)] \cdot (1/r^6), \quad (3.1.19)$$

где  $I_i = \hbar \omega_i$  — энергия возбуждения электрона в атоме  $i$ , а  $\alpha_i = e^2/m\omega_i^2$  — электронная поляризуемость атома  $i$ . Типичные значения величин  $I$  — порядка десятка электрон-вольт (т. е. сотен ккал/моль), а величин  $\alpha$  — порядка кулического ангстрема.

### Задача 3.2

От каких электронов атома, внешних или внутренних, зависит Ван-дер-Ваальсово взаимодействие?

### Решение

От внешних: они слабее связаны с ядром и потому легче поляризуются, — или, что то же самое, у них меньше частоты собственных колебаний (а удвоение обеих частот,  $\omega_1$  и  $\omega_2$ , согласно (3.1.18), снижает энергию Ван-дер-Ваальсова взаимодействия в восемь раз).

### Задача 3.3

Зависит ли Ван-дер-Ваальсово взаимодействие пары атомов от окружающей их среды?

### Решение

Да. Поскольку Ван-дер-Ваальсово взаимодействие пары атомов определяется квадратом диполь-дипольного взаимодействия поляризованных состояний этих атомов, а диполь-дипольное, как и всякое электростатическое взаимодействие, обратно пропорционально диэлектрической проницаемости среды, то Ван-дер-Ваальсово взаимодействие пары атомов падает с ростом диэлектрической проницаемости среды.

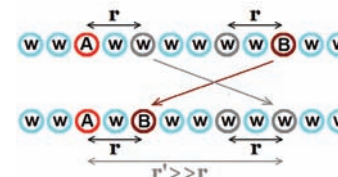
Речь, конечно, идет о той диэлектрической проницаемости, что соответствует частотам собственных колебаний электронов, т. е. ультрафиолетовой области спектра электромагнитных излучений. Эта диэлектрическая проницаемость определяется электронными поляризуемостями атомов, т. е. величинами  $\alpha = e^2/m\omega^2$ , фигурировавшими в уравнении (3.1.18), и числом атомов в единице объема, т. е. плотностью среды. Типичная электронная диэлектрическая проницаемость жидкостей, кристаллов и стекол — около 2. [Отмечу, что ослабление Ван-дер-Ваальсова взаимодействия пары отдельно взятых атомов в плотной среде (по сравнению с их взаимодействием в вакууме) невелико, ~5 % согласно расчетам одноатомных жидкостей и твердых тел, и потому этот эффект обычно не принимается во внимание. Однако можно показать, что описываемый эффект существенно возрастает в присутствии ковалентных связей, т. е. в среде многоатомных молекул.]

### Задача 3.4 (довольно сложная)

Ван-дер-Ваальсово взаимодействие пары любых атомов в вакууме всегда ведет к их притяжению. Почему эта закономерность не сохраняется в плотной среде? Какие атомы там отталкиваются друг от друга? [Подсказка. Воспользоваться тем, что Ван-дер-Ваальсова энергия взаимодействия пары атомов, согласно уравнения 3.1.18–3.1.19, может быть приближена формулой  $E_{AB}(r) \approx -\xi_A \cdot \xi_B / r^6$ , где  $\xi_A$  и  $\xi_B$  зависят от характеристик отдельных атомов, причем  $\xi_A > 0$  и  $\xi_B > 0$ .]

### Решение

В плотной среде практически нет «вакуумных вакансий». Поэтому, рассматривая смещение атома на расстояние порядка ангстрема и более, мы должны принять во внимание, что такое смещение носит характер обмена. При этом атом А, приближающийся к атому В, удаляет от него какой-то атом среды — и наоборот, удаляющийся атом тут же заменяется каким-то приближающимся.



Энергия такого обмена может быть записана в виде

$$\Delta E_{AB}(r) = E_{AB}(r) + E_{ww}(r) - [E_{Aw}(r) + E_{Bw}(r)].$$

Здесь  $E_{AB}(r)$  — собственная (т. е. такая, как в вакууме) энергия взаимодействия наших атомов А и В на расстоянии  $r$ ;  $E_{ww}(r)$  — собственная энергия взаимодействия атомов среды, w, на том же расстоянии; а  $E_{Aw}(r)$  и  $E_{Bw}(r)$  — собственные энергии взаимодействия атомов среды с А и В. Все эти собственные энергии отрицательны, т. е. соответствуют притяжению, но  $\Delta E_{AB}(r)$  может быть и положительной, и отрицательной величиной. Если  $\Delta E_{AB}(r) < 0$  — атомы А и В притягиваются в среде молекул w. Если  $\Delta E_{AB}(r) > 0$  — отталкиваются; это происходит тогда, когда А или В сильно взаимодействует со средой.

Можно показать (впрочем, достаточно взглянуть на уравнения 3.1.18–3.1.19), что величины  $E_{AB}$  и т. д. представимы в виде типа  $E_{AB}(r) \approx -\xi_A \cdot \xi_B / r^6$ , где  $\xi_A$  и  $\xi_B$  зависят от характеристик отдельных атомов, причем  $\xi_A > 0$  и  $\xi_B > 0$ . В этом приближении

$$\Delta E_{AB}(r) = (\xi_A - \xi_w) (\xi_B - \xi_w) / r^6.$$

Теперь видно, что атомы А и В притягиваются, когда  $\xi_A > \xi_w$  и  $\xi_B > \xi_w$  или  $\xi_A < \xi_w$  и  $\xi_B < \xi_w$  (т. е. когда А, и В либо одновременно «сильнее», либо одновременно «слабее», чем атомы среды), — и отталкиваются, когда  $\xi_A > \xi_w > \xi_B$  или  $\xi_A < \xi_w < \xi_B$  (т. е. когда один из атомов А, В «сильнее», а другой — «слабее», чем атомы среды).

### Задача 3.5

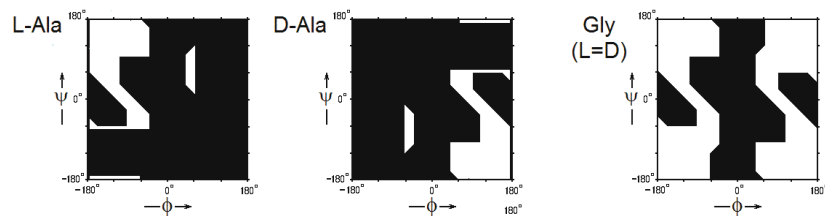
Вспомните карты разрешенных состояний для L-аланина и глицина. Как выглядят карты разрешенных состояний для D-стереоизомеров тех же аминокислот?



**Решение**

D-аланин — зеркальное отражение L-аланина, карта разрешенных состояний которого дана на рис. 3-4. При зеркальном отражении угол поворота вокруг оси (и, в частности, угол  $\phi$  и угол  $\psi$ ) меняет свой знак на противоположный (см. рис. 2-3). Поэтому карта разрешенных состояний D-аланина (см. белую область на схемах ниже) выглядит, как повернутая на  $180^\circ$  карта разрешенных состояний L-аланина.

D- и L-глицин — это одно и то же, так как глицин,  $-\text{NH}-\text{CH}_2-\text{CO}$ , сам себе зеркально-симметричен. Соответственно, карта его разрешенных состояний (см. рис. 3-3), повернутая на  $180^\circ$ , совпадает сама с собой.

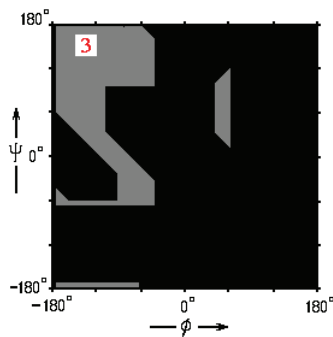


**Задача 3.6**

Предположим, что в полипептидной цепи имеется L-аминокислотный остаток с тремя  $\text{C}^\gamma$  атомами. Как выглядит карта разрешенных состояний для этого остатка?

**Решение**

Рис. 3-5 показывает, что только в одной области углов  $\phi, \psi$  разрешены одновременно все три ротамера  $\text{C}^\gamma$  атома по углу  $\chi_1$ . Только эта область и будет разрешена для остатка с тремя  $\text{C}^\gamma$  атомами (см. белую область на следующей схеме).



**Задача 4.1**

Зная, что расстояние между центрами O атомов в  $\text{O}-\text{H}:::\text{O}$  связи равно  $2,8\text{\AA}$ , длина ковалентной  $\text{O}-\text{H}$  связи равна  $0,96\text{\AA}$ , а минимальное расстояние между атомами O, наблюдаемое в кристаллах, равно  $2,7\text{\AA}$ , — оценить радиус непроницаемой для других атомов «твердой» части атома H, находящегося в молекуле воды.

**Решение**

Радиус «твердой» части атома O равен  $2,7\text{\AA}/2 = 1,35\text{\AA}$ . Поэтому радиус непроницаемой для приближающегося O-атома «твердой» части атома H, находящегося в молекуле воды, равен  $2,8\text{\AA} - 2,7\text{\AA}/2 - 0,96\text{\AA} = 0,49\text{\AA}$ . Он вдвое меньше чем «нормальный» минимально возможный радиус атома H в неполярных молекулах ( $r_{\text{min}}$ ), приведенный в таблице 3-1. Это — результат оттягивания электрона H электроотрицательным атомом O и «проминания» оставшегося на H электронного облака электростатическим взаимодействием  $\text{H}^+$  и  $\text{O}^-$ .

**Задача 4.2**

Существуют рассказы о «собачьих пещерах», где задыхаются мелкие собаки, но вполне может пройти человек. Этому явлению обычно дается следующее объяснение: углекислый газ ( $\text{CO}_2$ ), выделяющийся в пещере, стекает на ее дно, вытесняет оттуда более легкие азот с кислородом, и делает «придонный» слой воздуха — и только его — непригодным для дыхания. Состоятельно ли это объяснение?

**Решение**

Не вполне. Если бы все дело было только в весе отдельных молекул,  $\text{CO}_2$  имел бы практически одинаковую концентрацию по всему объему пещеры. В самом деле, согласно барометрическому соотношению, **вероятность**  $\sim \exp(-mgh/kT)$ , возникающая под действием силы тяжести характерная высота  $h^* = kT/mg$  слоя молекул  $\text{CO}_2$  (чья молекулярная масса 44 дальтон) должна быть всего в  $44/32 \approx 1,4$  раза меньше, чем характерная высота  $h^*$  слоя молекул  $\text{O}_2$  (чья молекулярная масса 32 дальтон). А так как характерная высота  $h^*$  слоя воздуха составляет много километров, то слой  $\text{CO}_2$  тоже охватит километры... Значит, гибельность «собачьей пещеры» вызвана не неравновесным распределением  $\text{CO}_2$  по ее объему, а кинетикой накопления  $\text{CO}_2$  в ней. В «собачьей пещере»  $\text{CO}_2$  либо выделяется со дна, либо, действительно, выделившись из стен, стекает потоком на дно — да там и остается долгое вре-

мя из-за медленности диффузии (по сравнению со скоростью поступления  $\text{CO}_2$  в пещеру) и отсутствия ветра. Однако, если бы  $\text{CO}_2$  перестал прибывать, диффузия (даже в отсутствие ветра) постепенно разнесла бы его по всей пещере, и вся она стала бы непригодной для дыхания.

## К лекции 5

### Задача 5.1

Предположим, что некая молекула тверда (т. е. не имеет возбуждаемых внутренних степеней свободы) и имеет форму шара. Предположим, что она имеет в кристалле колебания с амплитудой  $a = 0,5 \text{ \AA}$ , а ее насыщенный пар над кристаллом имеет давление  $P = 0,1 \text{ атм.}$  при  $27^\circ \text{C}$ .

Оценить свободную энергию  $\Delta G$ , энтропию  $\Delta S$  и энтальпию  $\Delta H$  испарения.

### Решение

$$\Delta G = 0; \Delta S \approx 25,8 \text{ кал/(град}\cdot\text{моль)}; \Delta H \approx 7,7 \text{ ккал/моль.}$$

### Объяснение.

Прежде всего:  $\Delta G = 0$ , так как изменение свободной энергии молекулы при переносе ее из твердого (или жидкого) тела в насыщенный пар равно нулю по определению. Далее, изменение энтропии молекулы при переносе ее из кристалла, где объем ее колебаний есть  $V_{\text{cr}} = (2a)^3$ , в пар, где на нее приходится объем  $V_g = RT/P$ , есть  $\Delta S = R \ln(V_g/V_{\text{cr}})$ . И, наконец, изменение энтальпии  $\Delta H = \Delta G + T\Delta S$ .

Для вычислений надо привести все численные данные в одну систему единиц (например, СИ):

$27^\circ \text{C}$  это  $T = 300 \text{ К}$ ; при этом  $RT = 1,99 \text{ кал/(град}\cdot\text{моль)} \times 300 \text{ К} \approx 600 \text{ кал/моль} \times 4,18 \text{ Дж/кал} \approx 2,5 \times 10^3 \text{ Дж/моль}$ ; а  $1 \text{ атм.} \approx 101 \text{ кН/м}^2$ , так что  $0,1 \text{ атм.} = 0,1 \text{ атм.} \times 101 \text{ (кН/м}^2\text{)/атм.} \approx 1,0 \times 10^4 \text{ Н/м}^2$ .

Теперь:  $V_{\text{cr}} = (2a)^3 = 1 \text{ \AA}^3$ ;

$V_g = RT/P \approx (2,5 \times 10^3 \text{ Дж/моль}) / (1,0 \times 10^4 \text{ Н/м}^2) = 0,25 \text{ м}^3/\text{моль} = 0,25 \text{ м}^3 / (0,602 \times 10^{24}) = 0,42 \times 10^{-24} \text{ м}^3 = 0,42 \times 10^6 \text{ \AA}^3$ ;

$\Delta S = R \ln(V_g/V_{\text{cr}}) = 1,99 \text{ кал/(град}\cdot\text{моль)} \ln(0,42 \times 10^6) \approx 25,8 \text{ кал/(град}\cdot\text{моль)}$ ;

$\Delta H = \Delta G + T\Delta S = 0 + T\Delta S = 300 \text{ К} \times 25,8 \text{ кал/(град}\cdot\text{моль)} \approx 7,7 \text{ ккал/моль}$ .

### Задача 5.2

Какова концентрация молекул в газе, имеющем давление  $0,1 \text{ атм.}$  при  $27^\circ \text{C}$ ?

### Решение

Так как удельный объем этого газа равен  $0,25 \text{ м}^3/\text{моль}$  (см. решение предыдущей задачи), то концентрация молекул в этом газе равна  $1/(0,25 \text{ м}^3/\text{моль}) = 1/(250 \text{ литр/моль}) = 0,004 \text{ моль/литр}$ .

### Задача 5.3

Каково характерное расстояние между ионами  $\text{H}^+$  (точнее,  $\text{H}_3\text{O}^+$ ) при pH7? при pH0?

### Решение

При pH0: концентрация  $1 \text{ моль/литр} \approx [0,6 \times 10^{24} \text{ иона}]/[10^8 \text{ нм}^3] = 0,6 \text{ иона/нм}^3 \approx 1 \text{ ион}/1,7 \text{ нм}^3$ .

Значит, характерное расстояние между ионами  $\text{H}^+$  есть  $(1,7 \text{ нм}^3)^{1/3} \approx 1,2 \text{ нм} = 12 \text{ \AA}$  при pH0.

При pH7: концентрация  $10^{-7} \text{ моль/литр} \approx 1 \text{ ион}/(1,7 \times 10^7 \text{ нм}^3)$ , а характерное расстояние между ионами есть  $(1,7 \times 10^7 \text{ нм}^3)^{1/3} \approx 257 \text{ нм} = 2570 \text{ \AA} \approx 1/4 \text{ микрона}$ .

### Задача 5.4

Предположим, что распределение концентраций некой гидрофобной молекулы между водой и органическим растворителем,  $X_{\text{в воде}} : X_{\text{в неоп. р-рит.}}$ , соотносится как 1:  $(5 \times 10^4)$  при  $0^\circ \text{C}$ , 1:  $(10 \times 10^4)$  при  $50^\circ \text{C}$ , 1:  $(4,4 \times 10^4)$  при  $100^\circ \text{C}$ .

Оценить свободную энергию  $\Delta G$ , энтропию  $\Delta S$ , энтальпию  $\Delta H$  и теплоемкость  $\Delta C_p$  переноса этой гидрофобной молекулы из воды в органический растворитель.

### Решение

$$\Delta G \equiv \Delta G_{\text{вода} \rightarrow \text{неоп. р-рит}} = -RT \ln (X_{\text{в неоп. р-рит}}/X_{\text{в воде}}) = RT \ln (X_{\text{в воде}}/X_{\text{в неоп. р-рит}}) \quad [\text{см. (5.13)}]$$

$\Delta G = RT \cdot \ln [1/(5 \times 10^4)] \approx (0,55 \text{ ккал/моль}) \cdot (-10,8) \approx -6,0 \text{ ккал/моль}$  при  $0^\circ \text{C}$ , т. е. при  $T = 273 \text{ К}$ ;

$\Delta G = RT \cdot \ln [1/(10 \times 10^4)] \approx (0,65 \text{ ккал/моль}) \cdot (-11,5) \approx -7,5 \text{ ккал/моль}$  при  $50^\circ \text{C}$ , т. е. при  $T = 323 \text{ К}$ ;

$\Delta G = RT \cdot \ln [1/(4,4 \times 10^4)] \approx (0,75 \text{ ккал/моль}) \cdot (-10,7) \approx -8,0 \text{ ккал/моль}$  при  $100^\circ \text{C}$ , т. е. при  $T = 373 \text{ К}$ .

Величины  $\Delta S$  и  $\Delta H$  вычисляются по формулам:

$$\Delta S = -d(\Delta G)/dT \quad [\text{см. (5.10)}],$$

$$\Delta H = RT^2 \cdot [-d(\Delta G/RT)/dT] \quad [\text{см. (5.11)}].$$

Величина  $\Delta C_p$  вычисляется по формуле:

$$\Delta C_p = dH/dT \quad [\text{см. (5.9)}].$$

$\Delta S$  в центре интервала 0–50 °С, т. е. при 25 °С:

$$-[-8,5 \text{ ккал/моль} - (-6,0 \text{ ккал/моль})]/50^\circ = 30 \text{ ккал/град}\cdot\text{моль};$$

$\Delta S$  в центре интервала 50–100 °С, т. е. при 75 °С:

$$-[-8,0 \text{ ккал/моль} - (-7,5 \text{ ккал/моль})]/50^\circ = 10 \text{ ккал/град}\cdot\text{моль};$$

$\Delta H$  в центре интервала 0–50 °С, т. е. при 25 °С ( $T = 298\text{K}$ ):

$$\begin{aligned} & (0,6 \text{ ккал/моль} \times 298^\circ) \cdot \{[-(-7,5 \text{ ккал/моль})/(0,65 \text{ ккал/моль}) - \\ & \quad -(-6,0 \text{ ккал/моль})/(0,55 \text{ ккал/моль})]/50^\circ\} = \\ & = (0,6 \text{ ккал/моль} \times 298^\circ) \cdot (11,5 - 10,8)/50^\circ \approx 2,5 \text{ ккал/моль}; \end{aligned}$$

$\Delta H$  в центре интервала 50–100 °С, т. е. при 75 °С ( $T = 348\text{K}$ ):

$$\begin{aligned} & (0,7 \text{ ккал/моль} \times 348^\circ) \cdot \{[-(-8,0 \text{ ккал/моль})/(0,75 \text{ ккал/моль}) - \\ & \quad -(-7,5 \text{ ккал/моль})/(0,65 \text{ ккал/моль})]/50^\circ\} = \\ & = (0,7 \text{ ккал/моль} \times 348^\circ) \cdot (10,7 - 11,5)/50^\circ = -5,5 \text{ ккал/моль}; \end{aligned}$$

$\Delta C_p$  в центре интервала 25–75 °С, т. е. при 50 °С:

$$[-5,5 \text{ ккал/моль} - (2,5 \text{ ккал/моль})]/50^\circ = -160 \text{ ккал/град}\cdot\text{моль}.$$

Посмотрите на рис. 5-4: вы найдете там примерно те же цифры для  $\Delta G$ ,  $\Delta S$ ,  $\Delta H$  — но с обратным знаком, так как рис. 5-4 описывает перенос молекулы из органического растворителя в воду, а в данной задаче рассмотрен обратный процесс: перенос молекулы из воды в органический растворитель.

### Задача 5.5 (довольно сложная)

Известно, что химический потенциал частиц А в растворе, где они находятся в очень малой концентрации, вычисляется по формуле  $\mu_A = kT \ln [C_A] + \mu_{A:B}^0$ , где  $C_A = N_A/V$  — концентрация, т. е. число частиц А в единице объема раствора (см. Задачу 8.2), а не зависящая от  $C_A$  часть потенциала,  $\mu_{A:B}^0$ , зависит от температуры, внешнего давления и от взаимодействия отдельной частицы А с окружающим ее растворителем В.

Найдите химический потенциал частиц В, составляющих основную массу раствора.

### Решение

$$\mu_B = -kT [C_A/C_B] + \mu_{B:}^0,$$

где  $C_A$  и  $C_B$  — концентрации (в моль/литр) частиц А и В в растворе, а  $\mu_{B:}^0$  есть химический потенциал чистого растворителя В (он зависит от температуры и внешнего давления).

*Объяснение.*

Химический потенциал частиц В мало меняется при добавлении малого количества молекул А, и потому его можно представить в виде

$$\mu_B = \mu_{B:}^0 + N_A (\partial \mu_B / \partial N_A),$$

где производная берется при постоянном давлении, температуре и стремящемся к нулю числу частиц ( $N_A$ ) в растворе.

Свободная энергия всего раствора, содержащего много молекул А и В, представима (по определению химического потенциала частиц в большой системе) в виде  $G = \mu_A N_A + \mu_B N_B$ , причем добавление к этому раствору малого количества ( $dN_A$ ) молекул А (при постоянном давлении, температуре и количестве молекул В) меняет эту свободную энергию, по тому же определению, на  $\mu_A \cdot dN_A$ , так что  $\partial G / \partial N_A = \mu_A$ . С другой стороны,

$$\begin{aligned} \mu_A & \equiv \partial G / \partial N_A \equiv \partial (\mu_A N_A + \mu_B N_B) / \partial N_A \equiv \\ & \equiv (\partial \mu_A / \partial N_A) \cdot N_A + \mu_A + (\partial \mu_B / \partial N_A) \cdot N_B + (\partial N_B / \partial N_A) \cdot \mu_B, \end{aligned}$$

причем  $(\partial N_B / \partial N_A) = 0$ , так как количество молекул В не меняется. Поэтому мы получаем уравнение

$$\begin{aligned} \mu_A & = (\partial \mu_A / \partial N_A) \cdot N_A + \mu_A + (\partial \mu_B / \partial N_A) \cdot N_B, \\ \text{или} \quad \partial \mu_B / \partial N_A & = -(\partial \mu_A / \partial N_A) \cdot (N_A / N_B), \end{aligned}$$

а так как  $\mu_A = kT \ln [N_A/V] + \mu_{A:B}^0$ , то  $\partial \mu_A / \partial N_A = kT / N_A$ , и  $\mu_B = \mu_{B:}^0 + N_A (\partial \mu_B / \partial N_A) = \mu_{B:}^0 + N_A (-kT / N_A) \cdot (N_A / N_B) = \mu_{B:}^0 - kT \cdot (N_A / N_B) \equiv -kT [C_A / C_B] + \mu_{B:}^0$ .

Обратите внимание, что химический потенциал растворителя зависит только от *концентрации* молекул примеси относительно молекул растворителя, но вовсе не зависит от силы взаимодействия растворителя с внесенными в него примесями, а вся сила взаимодействия растворителя с примесями входит в химический потенциал этих примесей. [То, что химический потенциал растворителя *падает* пропорционально концентрации молекул примеси, позволяет измерить эту концентрацию по *осмотическому давлению* раствора — т. е. по той разнице давлений, которая устанавливается в равновесии между растворителем с примесью и чистым растворителем, когда они разделены мембраной, проницаемой только для растворителя, но не для примеси.]

### Задача 5.6

Пусть в растворе имеется небольшое число молекул А и В, которые могут объединиться в результате обратимой реакции  $A + B \leftrightarrow AB$ . Доказать, что равновесие молекул А, В и АВ достигается тогда, когда их концентрации [А], [В] и [АВ] подчиняются «закону действующих масс», выражаемому уравнением

$$[A] \cdot [B]/[AB] = K_d.$$

Здесь

$$K_d = \exp[(\mu_{AB}^0 - \mu_A^0 - \mu_B^0)/kT],$$

где  $\mu_I^0$  — не зависящая от концентрации молекулы I составляющая химического потенциала, определяемая (см. задачу 17.2) взаимодействиями внутри молекулы I и взаимодействием молекулы I с окружающим растворителем (I = А, В или АВ).

### Решение

Если равновесие достигнуто, то свободная энергия находится в минимуме, т. е. она не меняется после объединения молекул А и В в АВ. Если концентрации молекул А, В и АВ малы, то их химические потенциалы подчиняются уравнениям  $\mu_I = kT \ln [I] + \mu_I^0$ , где I = А, В или АВ. Значит, происшедшее при реакции  $A + B \rightarrow AB$  изменение свободной энергии  $\mu_{AB} - \mu_A - \mu_B$  равно нулю, то есть

$$(kT \ln [AB] + \mu_{AB}^0) - (kT \ln [A] + \mu_A^0) - (kT \ln [B] + \mu_B^0) = 0$$

и 
$$\mu_{AB}^0 - \mu_A^0 - \mu_B^0 = kT \ln \{[A] \cdot [B]/[AB]\}.$$

Величина  $K_d = \exp[(\mu_{AB}^0 - \mu_A^0 - \mu_B^0)/kT]$  называется «константой диссоциации», а  $\mu_{AB}^0 - \mu_A^0 - \mu_B^0$  — свободной энергией связи частиц А и В. Величина  $K_d$  измеряется в тех же единицах (обычно — моль/литр), что и концентрации молекул. Если молекулы находятся, в основном, в диссоциированном состоянии (А и В) — константа диссоциации велика; если они находятся, в основном, в ассоциированном состоянии (АВ) — константа диссоциации мала.

### Задача 5.7

Пусть концентрации молекул А, В и их ассоциата АВ подчиняются «закону действующих масс», выражаемому уравнением  $[A] \cdot [B]/[AB] = K_d$ , причем  $[A] + [AB]$ , суммарная концентрация молекул А, есть  $\alpha$ , а  $[B] + [AB]$ , суммарная концентрация молекул В, есть  $\beta$ . Найти равновесные концентрации [А], [В] и [АВ]. Особо рассмотреть случай, когда константа диссоциации  $K_d$  мала (по сравнению с  $\alpha$  и  $\beta$ ), причем (а)  $\alpha = \beta$ ; (б)  $\alpha$  существенно больше, чем  $\beta$ .

### Решение

$$[AB] = (\alpha + \beta + K_d)/2 - \{[(\alpha - \beta)/2]^2 + K_d(\alpha + \beta)/2 + (K_d/2)^2\}^{1/2};$$

$$[A] = \alpha - [AB];$$

$$[B] = \beta - [AB].$$

Примечание. Из двух корней квадратного (относительно [АВ]) уравнения  $(\alpha - [AB])(\beta - [AB]) = K_d [AB]$  выбирается, естественно, тот, при котором все концентрации (в том числе [А] и [В]) — положительные числа.

В частном случае (а), когда  $\alpha = \beta \gg K_d$ :

$$[AB] \approx \beta - \{K_d \beta\}^{1/2} \approx \beta;$$

$$[A] = [B] \approx \{K_d \beta\}^{1/2}.$$

В частном случае (б), когда  $\alpha \gg \beta \gg K_d$ :

$$[AB] \approx \beta \{1 - K_d/(\alpha - \beta)\} \approx \beta;$$

$$[A] \approx \alpha - \beta$$

$$[B] = K_d [AB]/[A] \approx \beta K_d/(\alpha - \beta).$$

Эти два решения, (а) и (б), можно объединить в следующем виде, приближенно верном при  $\alpha \geq \beta \gg K_d$ :

$$[AB] \approx \beta;$$

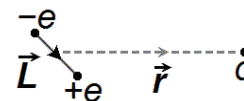
$$[A] \approx \alpha - \beta + \{K_d [AB]\}^{1/2};$$

$$[B] = K_d [AB]/[A].$$

## К лекции 6

### Задача 6.1

Найти свободную энергию  $U$  электростатического взаимодействия диполя длины  $L$ , состоящего из зарядов  $+e$  и  $-e$ , с зарядом  $q$ ; диполь находится на расстоянии  $r \gg L$  от заряда  $q$  (см. схему). Диэлектрическая проницаемость среды равна  $\epsilon$ .





## Решение

$$U = \varphi q,$$

где  $\varphi = \frac{+e}{\varepsilon |\vec{r} - \vec{L}/2|} + \frac{-e}{\varepsilon |\vec{r} + \vec{L}/2|}$  — потенциал, созданный диполем в точке  $q$ .

Воспользуемся разложениями

$$\begin{aligned} \frac{1}{|\vec{r} - \vec{L}/2|} &= \frac{1}{[(\vec{r} - \vec{L}/2)^2]^{1/2}} = \frac{1}{[r^2 - \vec{r}\vec{L} + L^2/4]^{1/2}} = \\ &= \frac{1}{r} [1 - \vec{n}\vec{L}/r + (L/2r)^2]^{-1/2} \approx \frac{1}{r} (1 + \vec{n}\vec{L}/2r) \end{aligned}$$

и 
$$\frac{1}{|\vec{r} + \vec{L}/2|} \approx \frac{1}{r} (1 - \vec{n}\vec{L}/2r);$$

здесь  $\vec{n} = \vec{r}/r$  — единичный вектор, направленный от центра диполя к заряду  $q$ , а в разложениях оставлены только старшие члены по малому параметру  $L/r$ . Тогда

$$\varphi = \frac{e}{\varepsilon} \left[ \frac{1}{|\vec{r} - \vec{L}/2|} - \frac{1}{|\vec{r} + \vec{L}/2|} \right] \approx \frac{e}{\varepsilon \cdot r^2} (\vec{n}\vec{L}), \quad (6.1.1)$$

$$U \approx \frac{qe}{\varepsilon \cdot r^2} (\vec{n}\vec{L}) = \frac{q}{\varepsilon \cdot r^2} (\vec{n}\vec{P}), \quad (6.1.2)$$

где  $\vec{P} = e\vec{L}$  — дипольный момент (он направлен от отрицательного заряда диполя к положительному).

*Добавления.*

1) Дипольный момент системы зарядов обычно вычисляется в виде  $\vec{P} = \sum_i e_i \vec{r}_i$ , где  $e_i$  и  $\vec{r}_i$  — величина заряда  $i$  (со знаком) и его координата в пространстве. Однако такая форма подходит только для такой системы, общий заряд которой,  $\sum_i e_i$ , равен нулю: иначе сдвиг системы, как целого, на расстояние  $\vec{R}$  меняет  $\vec{P}$  (добавляя к нему  $\vec{R}\sum_i e_i$ ), что лишено физического смысла. Поэтому, в общем виде, удобно отсчитывать дипольный момент от «центра тяжести» системы зарядов  $\vec{R}_0 = \sum_i |e_i| \vec{r}_i / \sum_j |e_j|$ , т. е. вычислять

его как  $\vec{P} = \sum_i e_i (\vec{r}_i - \vec{R}_0)$ . В этом случае  $\vec{P}$  остается тем же, что и при вы-

числении по обычной формуле  $\vec{P} = \sum_i e_i \vec{r}_i$  в случае, если  $\sum_i e_i = 0$ , а если все заряды  $e_i$  — одного знака, то  $\vec{P} = \sum_i e_i (\vec{r}_i - \vec{R}_0) = 0$ , что соответствует здравому смыслу.

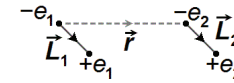
2) Если заряд  $q$  положителен (+), то к нему притягивается  $-e$ , и диполь  $\vec{P}$  стремится ориентироваться против  $\vec{n}$ , так что  $U$  (как и  $\vec{n}\vec{P}$ ) стремится стать отрицательной величиной. Если заряд  $q$  отрицательный (−), то к нему притягивается  $+e$ , т. е. диполь  $\vec{P}$  стремится ориентироваться по  $\vec{n}$ , так что  $\vec{n}\vec{P}$  стремится стать положительной величиной, а энергия  $U$  — опять-таки отрицательной.

3) До сих пор мы смотрели на взаимодействие «со стороны диполя». Теперь посмотрим на него «со стороны заряда». Поскольку напряженность создаваемого зарядом поля есть, как вы знаете,  $\vec{E}_q = q\vec{n}_{q \rightarrow} / \varepsilon r^2$ , где  $\vec{n}_{q \rightarrow}$  — единичный вектор, направленный от заряда  $q$  к точке, где определяется поле (в данном случае — где находится диполь), а  $\vec{n}_{q \rightarrow} = -\vec{n}$  (так как  $\vec{n}$  направлен от диполя к заряду  $q$ ), то ту же энергию  $U$  можно представить в виде

$$U = -\vec{E}_q \vec{P}. \quad (6.1.3)$$

## Задача 6.2

Найти свободную энергию  $U$  электростатического взаимодействия заряда двух диполей: (1) длины  $L_1$ , из зарядов  $+e_1$  и  $-e_1$ , и (2) длины  $L_2$ , из зарядов  $+e_2$  и  $-e_2$ . Диполи находятся друг от друга на расстоянии  $r \gg L_1, L_2$  (см. схему). Диэлектрическая проницаемость среды равна  $\varepsilon$ .



## Решение

$$U = \frac{e_1 e_2}{\varepsilon |\vec{r}|} - \frac{e_1 e_2}{\varepsilon |\vec{r} - \vec{L}_1|} - \frac{e_1 e_2}{\varepsilon |\vec{r} + \vec{L}_2|} + \frac{e_1 e_2}{\varepsilon |\vec{r} + (\vec{L}_2 - \vec{L}_1)|}.$$

Воспользуемся тем же разложением, что в задаче 6.1, но теперь оставим в нем два старших члена по малому параметру  $L/r$ , так как самые старшие члены, как мы увидим, сократятся:

$$\frac{1}{|\vec{r} + \vec{L}|} = \frac{1}{r} [1 + \{2\vec{n}(\vec{L}/r) + (L/r)^2\}]^{-1/2} \approx$$

$$= \frac{1}{r} \left[ 1 + (-1/2) \{2\bar{n}(\bar{L}/r) + (L/r)^2\} + [(-1/2)(-3/2)/2!] \{2\bar{n}(\bar{L}/r) + (L/r)^2\}^2 + \dots \right] = \\ = \frac{1}{r} \left[ 1 - \bar{n}\bar{L}/r - L^2/2r^2 + 3/2 \times (\bar{n}\bar{L}/r)^2 \right],$$

где  $\bar{n} = \bar{r}/r$  — единичный вектор, направленный от диполя 1 к диполю 2. Так же разлагаются и остальные члены. Тогда

$$U = \frac{e_1 e_2}{\varepsilon \cdot r} \left\{ 1 - [1 + \bar{n}\bar{L}_1/r - L_1^2/2r^2 + 3/2 \times (\bar{n}\bar{L}_1/r)^2] - \right. \\ \left. - [1 - \bar{n}\bar{L}_2/r - L_2^2/2r^2 + 3/2 \times (\bar{n}\bar{L}_2/r)^2] + \right. \\ \left. + [1 - \bar{n}(\bar{L}_2 - \bar{L}_1)/r - (\bar{L}_2 - \bar{L}_1)^2/2r^2 + 3/2 \times (\bar{n}(\bar{L}_2 - \bar{L}_1)/r)^2] \right\} = \\ = \frac{e_1 e_2}{\varepsilon \cdot r^3} \left[ \bar{L}_1 \bar{L}_2 - 3(\bar{n}\bar{L}_1)(\bar{n}\bar{L}_2) \right] = \frac{1}{\varepsilon \cdot r^3} \left[ \bar{P}_1 \bar{P}_2 - 3(\bar{n}\bar{P}_1)(\bar{n}\bar{P}_2) \right], \quad (6.2.1)$$

где  $\bar{P}_i = e_i \bar{L}_i$  — дипольный момент диполя  $i$ .

Обозначим через  $P_i'$  ту компоненту дипольного момента  $\bar{P}_i$ , что направлена вдоль  $\bar{n}$ , а через  $P_i''$  и  $P_i'''$  — две его другие, перпендикулярные  $\bar{n}$  компоненты. Тогда  $U$  можно записать в виде

$$U = \frac{1}{\varepsilon \cdot r^3} \left[ P_1'' P_2'' + P_1''' P_2''' - 2 P_1' P_2' \right]. \quad (6.2.2)$$

### Задача 6.3

Найти «энергию Борна», т. е. свободную энергию погружения в воду (из вакуума):

а) для пары далеко друг от друга отстоящих ионов «+» и «-» (с зарядами  $+e$ ,  $-e$  и радиусами  $R^+$ ,  $R^-$ , соответственно) и

б) для пары контактирующих (находящихся на расстоянии  $r' = R^+ + R^-$ ) ионов «+» и «-». Считать, что диэлектрическая проницаемость воды  $\varepsilon = 80$ ; пренебречь зависимостью  $\varepsilon$  от температуры.

в) Решить предыдущую задачу, считая, что те же ионы «+» и «-» образовали диполь с ковалентной связью длиной  $r' = (R^+ + R^-)/2$ .

Во всех случаях полагать  $R^+ = a$ ,  $R^- = 1,5a$ .

### Решение

а) отдельные ионы:

$$\Delta U^+ = (e^2/2R^+) (1/\varepsilon - 1) = -(e^2/2a) (1-1/\varepsilon) = -0,494 (e^2/a);$$

$$\Delta U^- = (e^2/2R^-) (1/\varepsilon - 1) = -(e^2/3a) (1-1/\varepsilon) = -0,329 (e^2/a);$$

$$\Delta U^+ + \Delta U^- = (e^2/2R^+) (1/\varepsilon - 1) + (e^2/2R^-) (1/\varepsilon - 1) = -0,823 (e^2/a);$$

б) Два разноименных иона на расстоянии  $r$ :

$$\delta U = \Delta U^+ + \Delta U^- + [(-e^2)/r\varepsilon - (-e^2)/r] = e^2 (-1/2R^+ - 1/2R^- + 1/r) (1 - 1/\varepsilon) = \\ = 0,9875 (e^2/a) (a/r - 5/6);$$

при  $r = r' = R^+ + R^- = 2,5a$ :  $\delta U'' = 0,9875 (e^2/a)(0,4 - 5/6) = -0,423 (e^2/a)$ : погружение двух контактирующих ионов стоит практически столько же, сколько погружение одного иона;

в) при  $r = r' = (R^+ + R^-)/2 = 1,25a$ :  $\delta U' = 0,9875 (e^2/a) (0,8 - 5/6) = -0,033 (e^2/a)$ : погружение диполя с ковалентной связью стоит довольно дешево.

Если  $e$  — единичный (электронный или протонный) заряд, а  $a = 1 \text{ \AA}$ , то  $e^2/a = 333$  ккал/моль. Тогда выигрыш энергии (точнее — свободной энергии) составляет сто — полтора ккал/моль при погружении в воду отдельного иона (а равно и пары контактирующих ионов), и лишь несколько ккал/моль — при погружении диполя с ковалентной связью ионов.

### Задача 6.4 (довольно сложная)

Кристаллографические (т. е. наблюдаемые в кристаллах) радиусы ионов ( $\text{Na}^+$ : 0,95 Å,  $\text{K}^+$ : 1,33 Å,  $\text{Ca}^{++}$ : 0,99 Å,  $\text{Cl}^-$ : 1,81 Å;  $\text{Br}^-$ : 1,95 Å, ...) отличаются от гидратационных, т. е. тех, что можно вычислить из энергий  $\Delta U$  переноса этих ионов из вакуума в воду ( $\text{Na}^+$ : 1,61 Å,  $\text{K}^+$ : 2,04 Å,  $\text{Ca}^{++}$ : 1,68 Å,  $\text{Cl}^-$ : 1,89 Å,  $\text{Br}^-$ : 2,09 Å, ...). [ $\Delta U$  равно 102 ккал/моль для  $\text{Na}^+$ , 82 ккал/моль для  $\text{K}^+$ , 391 ккал/моль для  $\text{Ca}^{++}$ , 87 ккал/моль для  $\text{Cl}^-$ , 79 ккал/моль для  $\text{Br}^-$ , ...; все численные данные взяты из [10].]

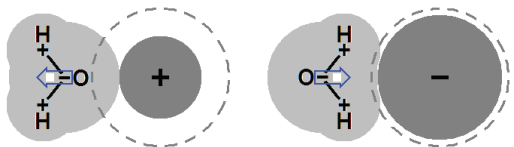
а) Объясните, на качественном уровне, почему гидратационные радиусы ионов больше их кристаллографических радиусов.

б) Вы видите, что различия гидратационных и кристаллографических радиусов существенно больше для положительно заряженных ионов (где они составляют  $\approx 0,7 \text{ \AA}$  и для одно- и для двухзарядных положительных ионов), чем для отрицательно заряженных (где эти различия составляют  $\approx 0,1 \text{ \AA}$ ). Объясните, на полукваликативном уровне, в чем причина этой разницы.

### Решение

а) Кристаллографический радиус атома очерчивает тот объем, куда никакие части других атомов проникнуть не могут. Но электроны и особенно ядра атомов в молекулах, с которыми взаимодействует ион (причем основным взаимодействием является электростатическое), отстоят от его центра на величину большую, чем это минимальное расстояние.

Потому-то гидратационные радиусы, описывающие взаимодействие ионов с водой, больше кристаллографических радиусов тех же ионов. На схеме сферы, образованные кристаллографическими радиусами, показаны темно-серым, а сферы, образованные гидратационными радиусами ионов, показаны пунктиром.

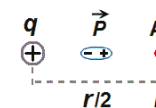


б) Хотя сам по себе гидратационный радиус конкретного иона нельзя оценить из элементарных соображений, мы можем объяснить различие между гидратационными и кристаллографическими радиусами положительно и отрицательно заряженных ионов. Схема показывает, что центр электростатических взаимодействий, т. е. центр диполя, в виде которого можно представить себе молекулу  $\text{H}_2\text{O}$ , не совпадает с центром объема  $\text{H}_2\text{O}$ . Объем  $\text{H}_2\text{O}$  фактически определяется объемом атома О (атомы Н лишь едва выступают за его пределы), так что центр объема  $\text{H}_2\text{O}$  близок к ядру атома О. Центр диполя, отмеченный белой точкой на схеме, находится между несущими парциальные заряды центрами атомов О и Н. Этот центр [его положение можно рассчитать, зная длины связей О–Н ( $0.96\text{Å}$ ) и угол между этими связями ( $104,5^\circ$ )] смещен на  $0,3\text{Å}$  от ядра О в сторону ядер атомов Н. «+» ион притягивает О-атом  $\text{H}_2\text{O}$ , так что, по сравнению с ядром О, центр диполя удален на  $0,3\text{Å}$  от «+» иона; «-» ион притягивает Н-атомы  $\text{H}_2\text{O}$ , так что, по сравнению с ядром О, центр диполя приближен на  $0,3\text{Å}$  к «-» иону (см. схему). Разность этих смещений,  $+0,6\text{Å}$  в пользу «+» иона, близка к наблюдаемой разности различий гидратационных и кристаллографических радиусов «+» и «-» ионов.

### Задача 6.5

Известно, что повышенная (по сравнению с вакуумом) диэлектрическая проницаемость среды создается ориентирующимися по электрическому полю диполями — молекулами этой среды. Известно также, что диэлектрическая проницаемость  $\epsilon > 1$  ослабляет создаваемый зарядом потенциал согласно формуле  $\phi = q/\epsilon r$ , где  $r$  — расстояние от заряда  $q$  до точки, где этот потенциал измеряется.

Предположим, что «среда» состоит из одного-единственного диполя, расположенного на полпути между зарядом  $q$  и точкой А, где измеряется потенциал поля (и что этот диполь, естественно, ориентирован по полю заряда  $q$ ; см. схему).



- Усилит или ослабит этот диполь тот потенциал, что был в точке А в отсутствие диполя?
- Насколько? (Считая, для простоты, что диполь полностью соригентирован по полю заряда  $q$ .)
- Какой потенциал создаст этот полностью соригентированный по полю диполь в точке, где находится заряд  $q$ ?

### Решение

а) Как ни странно, как ни контр-интуитивно — но такой диполь не ослабит, а *усилит* тот потенциал, что создавался зарядом  $q$  в точке А в отсутствие диполя!

К заряду  $q$  направится, естественно, тот конец диполя, что к нему притягивается, т. е. противоположно заряду  $q$  заряженный. При этом одноименный с зарядом  $q$  конец диполя направлен к точке А, т. е. он подходит к А ближе, чем противоположный заряд диполя. Значит, *одноименный* (с зарядом  $q$ ) конец диполя действует на точку А *сильнее*, чем противоположный, — что *усилит* созданный зарядом  $q$  в точке А потенциал.

б) Если диполь (длины  $L$ , с одноименным зарядом  $+q$  зарядом  $+e$  на одном конце и зарядом  $-e$  на другом) полностью соригентирован по полю заряда  $q$ , то исходный потенциал  $\phi = q/r$  повышается на созданный диполем потенциал  $\Delta\phi_A = \frac{+e}{|\vec{r}/2 - \vec{L}/2|} + \frac{-e}{|\vec{r}/2 + \vec{L}/2|} = 4P/r^2$ , где  $P = |eL|$  — величина

момента диполя, внесенного на полпути между зарядом  $q$  и точкой А [ср. с формулами (6.1.1), (6.1.2)].

в) Потенциал, созданный тем же диполем в точке нахождения  $q$ , есть

$$\Delta\phi_q = \frac{+e}{|\vec{r}/2 + \vec{L}/2|} + \frac{-e}{|\vec{r}/2 - \vec{L}/2|} = -4P/r^2 = -\Delta\phi_A.$$

В результате энергия заряда  $q$  *понижается* (как то и должно быть при взаимодействии с поляризуемой средой!) на  $U = q\Delta\phi_q = -4qP/r^2$  [этот расчет сделан в предположении, что тепловые колебания диполя слабы и он полностью соригентирован по полю заряда  $q$ ; ср. с формулами (6.1.1), (6.1.2)].

### Задача 6.6

Решение (а) предыдущей задачи показывает, что расположенный между зарядом и точкой измерения диполь не ослабляет, а *усиливает* тот потен-

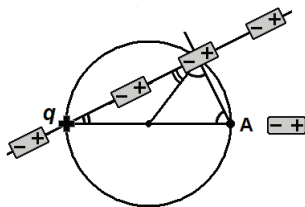
циал, что создавался зарядом  $q$  в точке измерения в отсутствие диполя. Как это может согласоваться с тем, что состоящая из диполей поляризуемая среда *ослабляет* тот потенциал, что создавался зарядом  $q$  в точке измерения в отсутствие поляризуемой среды?

[Подсказка. Этот вопрос можно сформулировать так: где находятся те диполи, что *ослабляют* потенциал, создававшийся зарядом  $q$  в точке измерения в отсутствие поляризуемой среды?]

### Решение

Будем, для определенности, считать, что заряд  $q$  положителен (+). Тогда у тех диполей, что *ослабляют* исходный положительный потенциал  $\varphi = +q/r$ , созданный непосредственно зарядом  $+q$  в точке А, « $\rightarrow$ » должен находиться *ближе* к А, чем « $+$ » (а не так, как у изображенного в задаче 6.5 диполя, размещенного между А и  $q$ , у которого к А обращен « $+$ »).

Так как диполи ориентированы по полю, созданному зарядом  $+q$ , у всех у них « $\rightarrow$ » обращен к заряду  $+q$  (а « $+$ » — наоборот). Схема, изображающая произвольную плоскость, проведенную через точки  $q$  и А, показывает, что те диполи, что находятся на окружности, диаметром которой служит отрезок  $q$ -А, стоят перпендикулярно линии, соединяющей центр диполя с А. (Есть такая теорема в геометрии: вписанный угол, опирающийся на диаметр окружности — прямой.)



Значит, у диполей, стоящих на поверхности сферы, диаметром которой служит отрезок  $q$ -А, « $\rightarrow$ » и « $+$ » находятся на равном расстоянии от А, и они не меняют потенциал в точке А.

Эта же схема показывает, что у тех диполей, что находятся внутри этой сферы, « $+$ » находится *ближе* к А, чем « $\rightarrow$ », и они (как и изображенный в задаче 6.5 диполь) *усиливают* исходный потенциал  $\varphi = +q/r$  в точке А.

Наоборот, у *всех* тех диполей, что находятся *вне* этой сферы, « $\rightarrow$ » находится *ближе* к А, чем « $+$ », и они *ослабляют* исходный потенциал  $\varphi = +q/r$  в точке А (и таких диполей больше, чем находящихся внутри рассматриваемой сферы).

Значит, *ослабляют* исходный потенциал  $\varphi = +q/r$  *те и только те* диполи, что находятся *вне* сферы, диаметром которой служит отрезок  $q$ -А. А на-

ходящаяся внутри этой сферы поляризуемая среда *вовсе не ослабляет*, а, как ни парадоксально, *усиливает* исходный потенциал точки А...

### Задача 6.7

Найти соотношение между энтальпией  $H$  и свободной энергией  $G$  электростатического взаимодействия зарядов в воде. Свободная энергия взаимодействия зарядов есть  $G \equiv U = qe/\epsilon r$ , а диэлектрическая проницаемость воды  $\epsilon$  зависит от температуры  $T$ , причем  $\epsilon(T=273\text{K}) = 88$ , а  $\epsilon(T=373\text{K}) = 55$ .

### Решение

Так как  $G \equiv H - TS$ , то  $H = G + TS$ . При этом, по формуле (5.10), энтропия  $S = -dG/dT$ , так что  $H = G - T(dG/dT) = G [1 - (T/G)(dG/dT)] = G [1 - d(\ln G)/d(\ln T)]$ .

Так как  $G = qe/\epsilon r$ , причем  $q, e, r$  не зависят от  $T$ , то  $H = G [1 + d(\ln \epsilon)/d(\ln T)]$ . Это соотношение не зависит от того, с какой средой мы имеем дело. От среды зависит только величина  $d(\ln \epsilon)/d(\ln T)$ . В частности, при  $\epsilon = \text{const}$  (как, например, в вакууме),  $H = G$ .

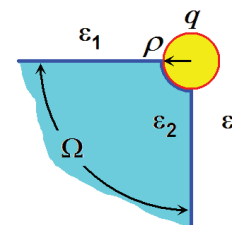
Оцениваем  $d(\ln \epsilon)/d(\ln T)$  у воды как  $\{\ln [\epsilon(T=373\text{K})] - \ln [\epsilon(T=273\text{K})]\} / \{\ln [373\text{K}] - \ln [273\text{K}]\} = \ln [55/88] / \ln [373/273] = -0,47/0,31 = -1,5$ .

Значит, в воде  $1 + d(\ln \epsilon)/d(\ln T) = 1 - 1,5 = -0,5$  и  $H = -0,5G$ .

Итак, *в воде* энтальпия *противоположна* по знаку свободной энергии, т. е. *в воде* энтальпия *падает* (а не растет, как интуитивно кажется!) при сближении одноименных зарядов (а свободная энергия, т. е. работа, конечно, растет).

### Задача 6.8 (довольно сложная)

Заряженный проводящий (металлический) шар радиуса  $\rho$  находится на вершине неограниченно-длинного конуса с диэлектрической проницаемостью  $\epsilon_2$ , окруженного средой с диэлектрической проницаемостью  $\epsilon_1$ . Определить энергию взаимодействия заряда с конусом, заключенный внутри которого телесный угол равен  $\Omega$  (имеется в виду энергия переноса заряда из бесконечно от конуса удаленной точки среды с диэлектрической проницаемостью  $\epsilon_1$  на вершину конуса).





### Решение

Вспомним изученные вами в курсе физики основные уравнения электростатики:

(1)  $\vec{E} = -\text{grad}(\varphi)$ , т. е. напряженность электрического поля (в любой точке пространства) равна градиенту электрического потенциала, взятому с обратным знаком;

(2)  $\oint_S (\varepsilon \vec{E})_{\perp} dS = 4\pi q_S$ , где интеграл берется по произвольной замкнутой поверхности  $S$ ,  $(\varepsilon \vec{E})_{\perp}$  — перпендикулярная этой поверхности составляющая вектора  $\varepsilon \vec{E}$  ( $\varepsilon$  — диэлектрическая проницаемость в данной точке), а  $q_S$  — окруженный этой поверхностью  $S$  «свободный» заряд (т. е. не включающий тех зарядов, что возникли в поляризующемся диэлектрике).

В приложении к шару с радиусом  $\rho$  и зарядом  $q$ , находящемуся в однородной среде с диэлектрической проницаемостью  $\varepsilon_1$ , эти уравнения приводят к хорошо вам известному результату: при  $r \geq \rho$  (где  $\vec{r}$  — ведущий из центра шара радиус-вектор, а  $\rho$  — радиус шара),  $\varphi = q/\varepsilon_1 r$ , откуда  $\vec{E} = q\vec{r}/\varepsilon_1 r^3$ , а энергия шара  $U_0 = q^2/2\varepsilon_1 \rho$ .

Для шара, находящегося на вершине конуса, будем искать решение в сходном виде:  $\varphi = q/Cr$  при  $r \geq \rho$ ; здесь  $C$  — константа, которую нам предстоит определить. Обратите внимание, что такой потенциал соответствует условию проводимости шара, так как он одинаков во всех точках его поверхности (т. е. при  $r = \rho$ ), а его градиент однозначно определен в любой точке вне шара, и он не имеет составляющей, перпендикулярной боковой поверхности конуса, т. е. поверхности контакта сред с  $\varepsilon = \varepsilon_1$  и  $\varepsilon = \varepsilon_2$ .

Если  $\varphi = q/Cr$ , то  $\vec{E} = q\vec{r}/Cr^3$  и, значит,  $U = q^2/2Cr$ . Возьмем интеграл  $\oint_S$  по сфере радиуса  $r \geq \rho$ , концентрически охватывающей наш шар. Этот  $\oint_S$  состоит из двух частей:  $\oint_{S_1}$  берется по той части этой сферы, что лежит вне конуса, а  $\oint_{S_2}$  — по той ее части, что лежит внутри конуса. Так как заключенный внутри конуса телесный угол равен  $\Omega$ , а остающийся вне конуса телесный угол есть, соответственно,  $4\pi - \Omega$ , то  $\oint_{S_2} \varepsilon_2 (qr/Cr^3) dS = \varepsilon_2 (qr/Cr^3) \cdot r^2 \Omega = \varepsilon_2 q\Omega/C$ , а  $\oint_{S_1} \varepsilon_1 (qr/Cr^3) dS = \varepsilon_1 (qr/Cr^3) \cdot r^2 (4\pi - \Omega) = \varepsilon_1 q(4\pi - \Omega)/C$ . При этом сумма  $\varepsilon_2 q\Omega/C + \varepsilon_1 q(4\pi - \Omega)/C = 4\pi q$ , по основному уравнению (2), так что  $C = \varepsilon_1 (1 - \Omega/4\pi) + \varepsilon_2 (\Omega/4\pi)$ . Итак, энергия шара на конусе есть

$$U = q^2 \cdot \frac{1}{2\rho} \cdot \frac{1}{\varepsilon_1 (1 - \Omega/4\pi) + \varepsilon_2 (\Omega/4\pi)}. \quad (6.8.1)$$

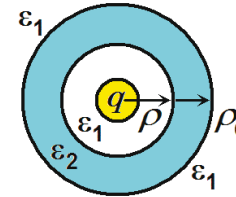
Обратите внимание, что, при  $\varepsilon_1 \gg \varepsilon_2$ , энергия  $U$  остается невысокой, порядка  $q^2/2\rho\varepsilon_1$  при любом  $\varepsilon_2 < \varepsilon_1$ , пока конус не охватит большую часть пространства, а потом быстро растет до  $q^2/2\rho\varepsilon_2$ .

Вычитая из нее энергию того же шара, находящегося в однородной среде с диэлектрической проницаемостью  $\varepsilon_1$ ,  $U_0 = q^2/2\varepsilon_1 \rho$ , получаем искомую энергию взаимодействия шара с конусом:

$$\Delta U = q^2 \cdot \frac{1}{2\rho\varepsilon_1} \cdot \frac{(\varepsilon_1 - \varepsilon_2)(\Omega/4\pi)}{\varepsilon_1 (1 - \Omega/4\pi) + \varepsilon_2 (\Omega/4\pi)}. \quad (6.8.2)$$

### Задача 6.9

Точечный заряд, находящийся в среде с диэлектрической проницаемостью  $\varepsilon_1$ , окружен концентрическим шаровым слоем с диэлектрической проницаемостью  $\varepsilon_2$ . Определить энергию взаимодействия заряда шаровым слоем (имеется в виду энергия переноса заряда из бесконечно от шара удаленной точки среды с диэлектрической проницаемостью  $\varepsilon_1$  в центр шара).



### Решение

В сферически-симметричном случае, изображенном на схеме, заряд  $q$  создает вокруг себя поле с потенциалом:

$$\varphi = q/\varepsilon_1 r \text{ — при } r \geq \rho_0 \text{ (где } \rho_0 \text{ — внешний радиус слоя);}$$

$$\varphi = q/\varepsilon_2 r + \Delta\varphi_1 \text{ — при } \rho \leq r \leq \rho_0;$$

$$\varphi = q/\varepsilon_1 r + \Delta\varphi \text{ — при } r \leq \rho \text{ (где } \rho \text{ — внутренний радиус слоя).}$$

Константы  $\Delta\varphi_1$  и  $\Delta\varphi$  находятся из условия непрерывности потенциала  $\varphi$  на стыках двух сред. (Проверьте, что этот потенциал удовлетворяет обоим основным уравнения электростатики, приведенным в решении задачи 6.8!)  $\Delta\varphi_1$  нас не интересует, а  $\Delta\varphi = q [1/\rho - 1/\rho_0] [1/\varepsilon_2 - 1/\varepsilon_1]$  есть дополнительный, созданный шаровым слоем потенциал в районе заряда  $q$ . В результате существования этого дополнительного потенциала энергия взаимодействия заряда с шаровым слоем равна

$$\Delta U = \frac{1}{2} q \Delta\varphi = q^2 \cdot \left(\frac{1}{\rho} - \frac{1}{\rho_0}\right) \cdot \frac{\varepsilon_1 - \varepsilon_2}{2\varepsilon_1 \varepsilon_2}. \quad (6.9.1)$$

### Задача 6.10 (довольно сложная)

Как, на качественном уровне, зависит энергия взаимодействия точечного заряда  $q$  с маленьким телом, имеющим диэлектрическую проницаемость  $\varepsilon_2$ , в среде, имеющей диэлектрическую проницаемость  $\varepsilon_1$ , от расстояния и от диэлектрической проницаемости обеих сред?

#### Решение

Качественное приближение к ответу можно получить следующим образом.

Известно, что электростатическая энергия в единице объема равна  $\varepsilon E^2/8\pi$ , где  $E$  — напряженность электрического поля, равная, в однородной среде,  $q/\varepsilon r^2$  на расстоянии  $r$  от заряда  $q$ . Итак, электростатическая энергия объема  $dV$  среды с диэлектрической проницаемостью  $\varepsilon_1$  равна  $dV \{q^2/8\pi\varepsilon_1 r^4\}$  на расстоянии  $r$  от заряда  $q$ . [Проверяем вышесказанное «известно...», интегрируя  $q^2/8\pi\varepsilon_1 r^4$  по объему, лежащему вне шара радиуса  $R$ ; получаем  $\iiint_{r \geq R} dV \{q^2/8\pi\varepsilon_1 r^4\} = 1/2 q^2/\varepsilon_1 R$  — т. е. хорошо нам уже известную формулу для энергии заряженного шара в диэлектрике.] Замена  $\varepsilon_1$  на  $\varepsilon_2$  в кусочке пространства не может изменить ни  $q$ , степени  $r$  в выражении  $dV \{q^2/8\pi\varepsilon_1 r^4\}$ . Однако, естественно, она как-то повлияет на множитель, связанный с  $\varepsilon$ . Точно оценить это изменение из элементарных соображений я не могу. Ясно, однако (и мы это уже обсуждали в лекциях), что замена  $\varepsilon_1$  на более высокое  $\varepsilon_2$  приведет к притяжению такого, более поляризующегося кусочка к заряду (т. е. к понижению энергии), а замена  $\varepsilon_1$  на более низкое  $\varepsilon_2$  приведет к отталкиванию этого кусочка. Поэтому знак энергии взаимодействия должен быть тот же, что у  $\varepsilon_1 - \varepsilon_2$ . Из того соображения, что одновременное увеличение и  $\varepsilon_1$ , и  $\varepsilon_2$  в несколько раз должно во столько же раз понизить электростатическую энергию (где какой-то  $\varepsilon$  всегда стоит в знаменателе), можно заключить, что множитель  $1/\varepsilon_1$  в выражении  $dV \{q^2/8\pi\varepsilon_1 r^4\}$  переходит в множитель вида  $(\varepsilon_1 - \varepsilon_2)/\varepsilon_1 \langle \varepsilon \rangle$ , где  $\langle \varepsilon \rangle$ , — какая-то эффективная, усредненная диэлектрическая проницаемость. Рекомендую посмотреть на соответствующие члены в уравнениях (6.8.2), (6.9.1).

Итак, мы можем, на качественном уровне, сказать, что энергия взаимодействия точечного заряда  $q$  с маленьким телом пропорциональна  $(\varepsilon_1 - \varepsilon_2)/\varepsilon_1 \langle \varepsilon \rangle$ , и  $1/r^4$  (а также объему тела и квадрату заряда); можно записать это в виде формулы

$$dU = \frac{1}{8\pi} \cdot dV \cdot \frac{q^2}{r^4} \cdot \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 \langle \varepsilon \rangle}, \quad (6.10.1)$$

где величина  $\langle \varepsilon \rangle$ , нам пока неизвестна; из частных случаев, разобранных в задачах 6.8 и 6.9, можно только предположить, что величина  $\langle \varepsilon \rangle$ , ближе к  $\varepsilon_1$ , т. е. к диэлектрической проницаемости среды, чем к  $\varepsilon_2$ . Для того чтобы сказать больше, нужно точное решение задачи.

*Примечание.* Точное (и сложное!) решение задачи о взаимодействии заряда с шариком показывает, что

$$dU = \frac{1}{8\pi} \cdot dV \cdot \frac{q^2}{r^4} \cdot \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 (\frac{2}{3}\varepsilon_1 + \frac{1}{3}\varepsilon_2)}. \quad (6.10.2)$$

## Лекции 8

### Задача 8.1

Найдите температуру идеального одноатомного газа, энергия которого составляет  $\xi$  в расчете на частицу.

#### Решение

$$T = \frac{2}{3} \xi / k.$$

Расчет базируется на формулах (8.4),  $1/T = dS/dE$ , и (8.1),  $S = k \ln [M(E)]$ . Здесь  $E = N\xi$  — полная энергия газа из  $N$  частиц, а энтропия состоит из кинетической и координатной частей,  $S(E) = S_{\text{кинет.}}(E_{\text{кинет.}}) + S_{\text{кoord.}}(E_{\text{кoord.}})$ , где  $E_{\text{кинет.}} + E_{\text{кoord.}} = E$ . Так как в идеальном газе молекулы не взаимодействуют по определению,  $E_{\text{кoord.}} = 0$ ; значит,  $E_{\text{кинет.}} = E$ . С другой стороны, кинетическая энергия одноатомного газа из  $N$  частиц

$$E_{\text{кинет.}} = \sum_{i=1}^N \frac{1}{2} m [v_{i,1}^2 + v_{i,2}^2 + v_{i,3}^2],$$

где  $m$  — масса молекулы, а  $v_{i,1}$ ,  $v_{i,2}$ ,  $v_{i,3}$  — скорости движения частицы  $i$  по трем координатам трехмерного пространства. Значит, квадрат длины имеющего  $3N$  координат вектора  $(v_{1,1}, v_{1,2}, \dots, v_{N,2}, v_{N,3})$  равен  $2E_{\text{кинет.}}/m$ , а доступные для системы из  $N$  молекул точки в пространстве скоростей лежат на  $3N$ -мерной сфере с радиусом  $R = (2E_{\text{кинет.}}/m)^{1/2}$ . Площадь этой  $3N$ -мерной сферы равна  $C \cdot R^{3N-1}$  (коэффициент  $C$ , также экспоненциально зависящий от  $N$ , нас сейчас не интересует), а соответствующая ей энтропия  $S_{\text{кинет.}} = k[\ln C + (3N-1)\ln(R)]$ . Теперь

$$\begin{aligned} 1/T &= dS_{\text{кинет.}}/dE_{\text{кинет.}} = k \cdot d \{ \ln C + [(3N-1)/2] \ln (2E_{\text{кинет.}}/m) \} / dE_{\text{кинет.}} = \\ &= k \cdot (3N-1) / (2E_{\text{кинет.}}) = k \cdot [(3N-1)/2N] / \xi. \end{aligned} \quad (8.1.1)$$

Итого,  $T = [2N/(3N - 1)](\xi/k) = 2/3 \xi/k$  при  $N \gg 1$ . А так как энергия  $\xi^{(1)}$ , приходящаяся на каждую из трех степеней свободы одноатомной молекулы, равна, в среднем,  $\xi/3$ , то  $kT = 2\xi^{(1)}$ .

*Дополнение.*

Ту же задачу можно решить другим способом — более коротким, но требующим больше математических навыков.

Исходя из распределения Больцмана, найдем среднюю кинетическую энергию, приходящуюся на каждую степень свободы одноатомной молекулы

$$\begin{aligned} \xi^{(1)} &= \int_{-\infty}^{+\infty} \frac{mv^2}{2} \exp\left(-\frac{mv^2}{2kT}\right) dv / \int_{-\infty}^{+\infty} \exp\left(-\frac{mv^2}{2kT}\right) dv = \\ &= \int_{-\infty}^{+\infty} \frac{kT}{2} \cdot \frac{mv^2}{kT} \exp\left(-\frac{mv^2}{2kT}\right) d\left[\frac{vm^{1/2}}{(kT)^{1/2}}\right] / \int_{-\infty}^{+\infty} \exp\left(-\frac{mv^2}{2kT}\right) d\left[\frac{vm^{1/2}}{(kT)^{1/2}}\right] = \\ &= \frac{kT}{2} \cdot \int_{-\infty}^{+\infty} x^2 \exp\left(-\frac{x^2}{2}\right) dx / \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2}\right) dx = \frac{kT}{2}. \end{aligned} \quad (8.1.2)$$

А так как  $\xi^{(1)} = \xi/3$  в трехмерном пространстве, то  $kT = 2/3 \xi$ . Итак, в среднем, кинетическая энергия частицы есть

$$1/2 m (v_1^2 + v_2^2 + v_3^2) = 3/2 kT. \quad (8.1.3)$$

Обратите внимание, что результат  $\xi^{(1)} = kT/2$ , где  $\xi^{(1)}$  — средняя кинетическая энергия в расчете на одну степень свободы, не зависит ни от массы частицы, ни от каких-либо предположений об их взаимодействии, ни даже от размерности системы, и потому он верен для всех частиц в любых системах, подчиняющихся классической механике (при условии, что этих частиц — много).

### Задача 8.2 (есть тонкости...)

Найдите координатную энтропию идеального одноатомного газа, состоящего из (а)  $N$  разных и (б)  $N$  одинаковых частиц, заключенных в объем  $V$ . Найдите также координатную составляющую химического потенциала частицы в каждом из этих двух случаев.

#### Решение

(а)  $S = Nk \ln V$ ,  $\mu = -kT \ln V$ ; (б)  $S = Nk \ln [eV/N]$ ,  $\mu = -kT \ln [V/N]$ .

Начнем со случая (а), когда все частицы — разные. Объем доступного координатного пространства равен  $V$  для любой частицы идеального одноатомного газа (и  $V^N$  — объем пространства конфигураций для всех  $N$  частиц). Следовательно, координатная энтропия одной индивидуальной

частицы есть  $k \ln V$ , и  $S = Nk \ln V$  — координатная энтропия  $N$  разных молекул. При добавлении одной индивидуальной частицы при постоянном объеме  $V$  координатная часть свободной энергии нашей системы увеличивается на  $dF = dE^{\text{коорд.}} - T \cdot k \ln V = -kT \ln V$ , так как в идеальном газе  $E^{\text{коорд.}} = 0$ . Следовательно, координатная часть химического потенциала  $\mu = -kT \ln V$ . [Отмечу, что формула (5.8)  $\mu = G/N$  в данном случае неприменима, так как она справедлива только при большом числе одинаковых частиц в системе.]

Если частицы одинаковы (случай (б)), то мы должны учесть, что обмен местами двух любых частиц не меняет систему. А так как  $N$  частиц можно переставить  $N! = N(N-1) \dots 2 \cdot 1$  способами (первую частицу ставим в любую из  $N$  позиций, вторую — в любую из  $N-1$  оставшихся, и т. д.), то объем пространства нетождественных конфигураций уменьшается  $N!$  раз, и, соответственно, его логарифм падает на  $\ln [N!] \approx N \ln [N/e]$  (последнее преобразование, справедливое при  $N \gg 1$ , называется формулой Стирлинга). Следовательно,  $S = Nk \ln [eV/N]$  для  $N$  одинаковых частиц.

Соответственно, координатная часть свободной энергии Гельмгольца для этой системы есть  $F^{\text{коорд.}} = E^{\text{коорд.}} - T \cdot Nk \ln [eV/N] = -T \cdot Nk \ln [eV/N]$ , так что добавление к этой системе одной частицы увеличивает ее свободную энергию на  $\mu = dF^{\text{коорд.}}/dN = -kT \ln [V/N]$ .

Поскольку в нашей системе много одинаковых частиц, ту же величину  $\mu$  можно вычислить и в виде  $\mu = G/N = (E - TS + PV)/N = (-T \cdot Nk \ln [eV/N] + NkT)/N = -kT \ln [V/N]$ . [Доказательство известной формулы  $PV = NkT$ , где  $P$  — давление газа, см. в следующей задаче.]

При  $N = 1$  этот химический потенциал, естественно, совпадает с тем, что был получен для индивидуальной частицы в случае, когда все частицы — разные.

### Задача 8.3

Докажите, что для заданного количества идеального газа  $PV/T = \text{const}$  (где  $V$  — объем,  $P$  — давление,  $T$  — температура).

#### Решение

Изменение  $dF$  свободной энергии тела при постоянной (обеспечиваемой теплообменом с окружающей средой) температуре равно, по определению, вложенной в тело работе. Давление  $P$ , тоже по определению, — это сила, действующая на единицу площади ограничивающей газ поверхности. На поверхность площадью  $\sigma$  действует сила  $P\sigma$ . Сжатие газа, т. е. изменению его объема на  $-dV < 0$ , соответствует сдвиг поверхности  $\sigma$  на рас-

стояние  $-dV/\sigma$ . Этот сдвиг требует вложения в сжимаемое тело работы  $P\sigma \times (-dV/\sigma) = -PdV$ . Следовательно,  $dF = -PdV$ , и, значит,  $(dF/dV)_{T=\text{const}} = -P$ .

Свободная энергия состоит из кинетической и координатной частей, т. е.  $F = (E^{\text{кинет.}} - TS^{\text{кинет.}}) + (E^{\text{коорд.}} - TS^{\text{коорд.}})$ . При этом  $E^{\text{кинет.}}$  и  $S^{\text{кинет.}}$  зависят только от скоростей, но не координат молекул газа,  $T = \text{const}$  есть температура окружающей среды, а  $E^{\text{коорд.}} = 0$  для идеального газа. Только энтропия  $S^{\text{коорд.}}$  распределения молекул газа в пространстве меняется при сжатии газа. А  $S^{\text{коорд.}} = Nk \ln[eV/N]$ , согласно решению задачи 8.2, и, значит,  $dS^{\text{коорд.}}/dV = Nk/V$ . Следовательно,  $P = -(dF/dV)_{T=\text{const}} = T(dS^{\text{коорд.}}/dV) = TNk/V$ , и  $PV/T = Nk = \text{const}$  для заданного количества идеального газа. Что и требовалось доказать.

#### Задача 8.4 (сложная)

Согласно основному определению энтропии, она пропорциональна логарифму числа доступных для системы состояний. Такое определение согласуется с «Третьим началом термодинамики» (согласно которому энтропия обращается в нуль при нулевой абсолютной температуре), означая, что при нулевой температуре системе доступно только одно (самое низкоэнергетическое) состояние. С другой стороны, энтропия газа (вернее, ее координатная часть) обычно представляется как число, пропорциональное логарифму объема  $V$ . Это второе определение несет в себе противоречие, так как объем может измеряться в разных единицах (литрах,  $\text{Å}^3$ , кубических саженях и т. д.). При этом численная величина объема, измеренного в разных единицах, будет разная — а, следовательно, и численная величина энтропии также будет зависеть от единицы измерения объема. Более того: определяя энтропию через логарифм объема, мы вступаем в противоречие с третьим началом термодинамики, так как если логарифм объема равен нулю, когда объем измеряется в одной из единиц, — то логарифм того же объема, измеренного во всех прочих (столь же почтенных) единицах, не равен нулю.

Попытайтесь разрешить это противоречие. [Подсказка. Это противоречие можно устранить, найдя «естественную» единицу измерения объема, — единицу, определяемую физикой системы, а не человеческой историей (как литры,  $\text{Å}^3$  и т. д.).]

#### Решение

Простой, прагматичный и обычно даваемый (в том числе — в этих лекциях) ответ гласит, что во всех процессах нас интересует не величина энтропии, а ее изменение, — а оно пропорционально логарифму *соотношения* объемов,  $V_1/V_2$ , которое не зависит от того, в какой единице измеряются

объемы  $V_1$  и  $V_2$ . Это справедливо. Более того, то же относится к энергии — нас всегда интересует не абсолютная величина, а лишь изменение энергии (вот почему знаменитая, огромная абсолютная энергия Эйнштейна,  $E = mc^2$ , была впервые получена не из опыта, а из теории, — и была востребована опытом лишь тогда, когда потребовалось понять источник огромной энергии ядерных реакций). Однако этот прагматичный ответ поверхностен: он не устраняет противоречия энтропии, определяемой как логарифм объема, с третьим началом термодинамики (т. е. этот ответ не работает при достаточно низких температурах).

По-настоящему верный и глубокий ответ базируется на принципе неопределенности Гейзенберга: соотношение  $m\Delta v\Delta x \approx \hbar$  задает «квант объема» для рассматриваемой частицы. Однако  $\hbar$  — это «квант объема» пространства «координата — импульс» [точнее, «одномерная координата ( $x$ ) — импульс ( $mv \equiv mdx/dt$ ) по этой одномерной координате»] для этой частицы, а не квант объема обычного (координатного) пространства. [Это соотношение относится к каждой из координат трехмерного пространства; поэтому «квантом объема», т. е. «естественной» единицей измерения объема шестимерного пространства, состоящего из трех обычных пространственных координат  $\{x_1, x_2, x_3\}$  и трех соответствующих импульсов  $\{mv_1, mv_2, mv_3\}$  каждой частицы является  $\hbar^3$ .]

Соотношение неопределенностей показывает, что объем  $mdv \cdot dx$  пространства «координата — импульс» каждой частицы содержит *определенное число*, порядка  $mdv \cdot dx/\hbar$ , *физически различимых* состояний этой частицы, если  $mdv \cdot dx \geq \hbar$  (а такое ограничение объема пространства координата — импульс, которое приводит к  $mdv \cdot dx < \hbar$ , попросту физически невозможно).

В рамках классической механики, в которой кинетическая и потенциальная энергии разделяются, мы можем определить «квант»  $\Delta x$  каждой из координат обычного трехмерного пространства *для каждой данной частицы при заданной температуре*. Приблизительно, это делается так. Поскольку средняя кинетическая энергия  $\langle mv^2/2 \rangle$ , движения частицы по каждой координате при температуре  $T$  равна  $kT/2$  [см. (8.1.2)], то среднеквадратичная величина ее скорости  $\langle v^2 \rangle^{1/2} = (kT/m)^{1/2}$ . При этом неопределенность  $\Delta v$  в величине скорости частицы (которая может быть направлена в любую сторону) составляет также  $\approx (kT/m)^{1/2}$ . Следовательно, по принципу неопределенности Гейзенберга,  $\Delta x \cdot m (kT/m)^{1/2} \approx \hbar$ , и  $\Delta x \approx \hbar/(mkT)^{1/2}$ .

Такие величины мы уже рассматривали, решая задачу 2.8, и нашли, что, при  $T = 300\text{K}$ ,  $\Delta x \approx 0,4\text{Å}$  при  $m = 1$  дальтон,  $\Delta x \approx 0,09\text{Å}$  при  $m = 18$  дальтон и т. д.



Следовательно, объем трехмерного координатного пространства для частицы должен измеряться в «естественных» для нее (при заданной величине ее термического возбуждения) единицах  $[\Delta x]^3 = [\hbar/(mkT)^{1/2}]^3$ . Обратите внимание, что  $\Delta x$  неограниченно растет при  $T \rightarrow 0$ , так что  $\ln(V/[\Delta x]^3)$  падает от огромных величин (при высоких температурах) до 0, соответствующего третьему началу термодинамики, когда  $[\Delta x]^3$  сравнивается с  $V$ , — и дальше упасть уже не может, так как, в силу принципа неопределенности Гейзенберга, величина возбуждения частицы в заданном объеме не может упасть ниже определенного предела. Впрочем, когда  $[\Delta x]^3$  приближается к  $V$ , справедливое для классической механики разделение кинетической и потенциальной энергии теряет свою силу, и мы переходим в область строгой квантовой механики, лежащую за пределами этих лекций.

### Задача 8.5

Найдите  $E^{(1)}$ , среднюю потенциальную энергию частицы, подчиняющейся классической механике, если она колеблется по одной координате, энергия растяжения связи с этой частицей  $U = \frac{1}{2}Kx^2$ , а температура среды равна  $T$ . Найдите также  $\langle x^2 \rangle^{1/2}$ , среднеквадратичное смещение частицы от точки равновесия, возникающее вследствие тепловых флуктуаций.

### Решение

$$E^{(1)} = kT/2, \quad \langle x^2 \rangle^{1/2} = [kT/K]^{1/2}.$$

Исходя из распределения Больцмана,

$$E^{(1)} = \frac{1}{2}Kx^2 = \int_{-\infty}^{+\infty} \frac{Kx^2}{2} \exp\left(-\frac{Kx^2}{2kT}\right) dx / \int_{-\infty}^{+\infty} \exp\left(-\frac{Kx^2}{2kT}\right) dx = \frac{kT}{2},$$

$$\langle x^2 \rangle^{1/2} = [1/2 Kx^2 / (1/2 K)]^{1/2} = [kT/K]^{1/2}$$

(см. решение уравнения 8.1.2).

Обратите внимание, что эта средняя потенциальная энергия одномерной осцилляции,  $E^{(1)}$ , совпадает с  $\xi^{(1)}$ , средней кинетической энергией в расчете на одну степень свободы.

Значит, полная (потенциальная + кинетическая) энергия одномерной осцилляции равна  $kT$ .

### Задача 8.6 (довольно сложная)

Найдите полную статистическую сумму, полную свободную энергию, полную энтропию и полную энергию частицы, подчиняющейся классической механике, если она колеблется по одной координате, энергия растяжения связи с ней есть  $U = \frac{1}{2}Kx^2$ , а температура среды равна  $T$ .

### Решение

Для рассматриваемого «классического осциллятора»

$$Z(T) = Z^{\text{кинет.}}(T) \cdot Z^{\text{коорд.}}(T) = \int_{-\infty}^{+\infty} \exp\left(-\frac{mv^2}{2kT}\right) d(mv) \cdot \int_{-\infty}^{+\infty} \exp\left(-\frac{Kx^2}{2kT}\right) dx =$$

$$= \left\{ (mkT)^{1/2} \int_{-\infty}^{+\infty} \exp\left(-\frac{mv^2}{2kT}\right) d\left[\frac{vm^{1/2}}{(kT)^{1/2}}\right] \right\} \cdot \left\{ (kT/K)^{1/2} \int_{-\infty}^{+\infty} \exp\left(-\frac{Kx^2}{2kT}\right) d\left[\frac{vK^{1/2}}{(kT)^{1/2}}\right] \right\} =$$

$$= \left\{ (mkT)^{1/2} \int_{-\infty}^{+\infty} \exp\left(-\frac{u^2}{2}\right) du \right\} \cdot \left\{ (kT/K)^{1/2} \int_{-\infty}^{+\infty} \exp\left(-\frac{y^2}{2}\right) dy \right\} = kT \cdot (m/K)^{1/2} \cdot 2\pi;$$

$$F(T) = -kT \ln(Z) = -kT \cdot \ln[2\pi kT(m/K)^{1/2}];$$

$$S(T) = -dF/dT = k \cdot \ln[2\pi e kT(m/K)^{1/2}];$$

$$E(T) = E^{\text{кинет.}}(T) + E^{\text{коорд.}}(T) = F + TS = kT;$$

последнее  $[E^{\text{кинет.}}(T) + E^{\text{коорд.}}(T) = kT]$  — в согласии с ответами на вопросы задач 8.1 и 8.5.

Обратите внимание, что «элементарным объемом» интегрирования здесь является (в согласии с принципом неопределенности Гейзенберга) величина  $mdv \cdot dx$  (см. задачу 8.4).

Сейчас уместно вспомнить, что величина  $(K/m)^{1/2}$  — это круговая частота  $\omega$  колебаний нашего осциллятора, что  $\varepsilon_0 = \frac{1}{2}\hbar\omega$  — минимальная, следующая из принципа Гейзенберга, квантовая энергия колебаний в невозбужденном осцилляторе (см. задачу 3.1) и что  $\hbar$  — это «квант объема» пространства «координата — импульс» любой частицы (см. задачу 8.4).

Поэтому полученные выше формулы для  $Z$ ,  $F$  и  $S$  можно переписать в виде

$$Z(T) = \hbar[\pi(kT/\varepsilon_0)];$$

$$F(T) = -kT \cdot \ln[\hbar \cdot \pi(kT/\varepsilon_0)] = -kT \cdot \ln[\hbar] - kT \cdot \ln[\pi(kT/\varepsilon_0)];$$

$$S(T) = k \cdot \ln[\hbar \cdot \pi e(kT/\varepsilon_0)] = k \cdot \ln[\hbar] + k \cdot \ln[\pi e(kT/\varepsilon_0)].$$

Добавление. Рассчитывая величину  $Z(T)$ , мы интегрировали по пространству «координата — импульс», т. е. по  $d(mv) \cdot dx$ , пользуясь обычными величинами измерения массы, длины и времени (например, килограмм — метр — секунда). Поэтому и ответ также получен в этих единицах. Однако его можно получить в безразмерном, независимом от единиц измерения виде, если учесть тот отмеченный в задаче 8.4 факт, что в объеме  $d(mv) \cdot dx$  содержится  $[d(mv) \cdot dx]/\hbar$  физически различных состояний, и интегрировать по  $d(mv) \cdot dx/\hbar$ . Тогда:

$$Z^*(T) = Z(T)/\hbar = \pi(kT/\varepsilon_0);$$

$$F^*(T) = F(T) + kT \cdot \ln[\hbar] = kT \cdot \ln[\pi(kT/\varepsilon_0)];$$

$$S^*(T) = S(T) - k \cdot \ln[\hbar] = k \cdot \ln[\pi e(kT/\varepsilon_0)].$$

Из выражения для  $S^*$  видно, что количество доступных для осциллятора «квантов объема  $\hbar$ » пространства «координата — импульс» составляет порядка  $\pi e(kT/\varepsilon_0)$  *штук*. Это значит, что энтропия (и энергия) осциллятора вполне подчиняется классической механике, пока температура высока, и  $kT > \varepsilon_0$ . При этом число таких «доступных осциллятору квантов объема» велико. Но мы уже никак не можем пользоваться классической механикой и классической статистической физикой для описания этого осциллятора, когда число таких «квантов» падает ниже 1, что происходит при столь низких температурах, что  $kT \ll \varepsilon_0$ .

### Задача 8.7

Во сколько раз замедляет процесс наличие на его пути энергетического барьера высотой в 10 ккал/моль при 0 °С? при 50 °С? при 100 °С?

#### Решение

В  $\exp[-(10 \text{ ккал/моль})/RT]$ , т. е. в  $\approx 8 \times 10^7$  раз при 0 °С, в  $\approx 5 \times 10^6$  при 50 °С, в  $\approx 6 \times 10^5$  при 100 °С.

### Задача 8.8

Согласно уравнению (8.20), квадрат диффузионного смещения  $l$  частицы в жидкости с вязкостью  $\eta$  зависит от времени диффузии  $t$ , величины  $kT$  и диаметра частицы  $D$  как  $l^2 \approx t [2kT/\pi D \eta]$ .

Рассмотрим диффузию в воде с вязкостью  $\eta \approx 0,01$  г/(см·с) при 27 °С. За какое время сместится от своего первоначального положения на 1 микрон, т. е. на величину порядка размера маленькой клетки: (а) молекула воды с диаметром  $D = 3$  Å; (б) маленькая белковая молекула с  $D = 30$  Å; (в) большой белковый комплекс с  $D = 300$  Å. Для сравнения — оценить, за какое время пролетели бы то же расстояние 1 микрон те же частицы в идеальном газе при тех же 27 °С. При всех вычислениях плотность частиц  $\rho$  принять равной 1,3 г/см<sup>3</sup>.

#### Решение

Время диффузии:  $\sim 0,12 \times 10^{-3}$  с для молекулы воды,  $\sim 1,2 \times 10^{-3}$  с для маленькой белковой молекулы,  $\sim 12 \times 10^{-3}$  с для большого белкового комплекса.

Время пролета 1 мкм без столкновений:  $\sim 1,2 \times 10^{-9}$  с для молекулы воды,  $\sim 0,4 \times 10^{-7}$  с для маленькой белковой молекулы,  $\sim 1,2 \times 10^{-6}$  с для большого

белкового комплекса. Скорость полета,  $v = (3kT/m)^{1/2} = (18kT/\pi D^3 \rho)^{1/2}$ , следует из уравнения (8.1.3).

### Задача 8.9

Представим себе, что «сайт связывания» частицы — это объем  $v_1 = 30 \text{ Å}^3$ , куда эта частица должна попасть какой-то своей определенной точкой А. Предположим также, что частица никак не чувствует близость сайта, пока точка А находится вне объема  $v_1$ .

За какое время частица диаметром  $D = 30 \text{ Å}$  найдет свой «сайт связывания» путем диффузии в объеме  $V = 1 \text{ мк}^3$  воды, вязкость которой  $\eta \approx 0,01$  г/(см·с) при 27 °С?

#### Решение

Время поиска сайта:  $\sim 4$  с.

Так как частица никак не чувствует близость сайта, пока точка А находится вне объема  $v_1$ , то ей нужно, путем случайной диффузии, перебрать  $\sim V/v_1$  объемов  $v_1$  прежде, чем она попадет в свой сайт. Переход от одного объема  $v_1$  к соседнему требует смещения на расстояние  $\sim (v_1)^{1/3}$ . Согласно уравнению (8.20), время такого элементарного перехода путем диффузии есть  $t_1 \approx (v_1)^{2/3} / [2kT/\pi D \eta]$ . Полное же время поиска своего «сайта связывания» путем диффузии в объеме  $V$  займет время  $t \sim t_1 [V/v_1] = (v_1)^{2/3} / [2kT/\pi D \eta] \times (V/v_1) = V(\pi \eta / 2kT) (D/v_1)^{1/3}$ .

Величина  $\pi \eta / 2kT = 0,4 \times 10^{-9} \text{ с/нм}^3 = 0,4 \text{ с/мк}^3$ , если  $\eta \approx 0,01$  г/(см·с) и  $T = 300 \text{ К}$  [см. 8.21 и 8.22].

Значит,  $t \sim 0,4 \text{ с/мк}^3 \times V(D/v_1)^{1/3}$ , и, при  $V = 1 \text{ мк}^3$ ,  $D = 30 \text{ Å}$  и  $v_1 = 30 \text{ Å}^3$ ,  $(D/v_1)^{1/3} \approx 10$ , время поиска сайта  $t \sim 4$  с.

### Задача 8.10

Считаем, что протон мгновенно связывается с центром его связывания после того, как он подойдет к этому центру на расстояние заметного электростатического взаимодействия, где электростатическая энергия порядка  $kT$ .

За какое время центр связывания протона поймает протон в воде при 27 °С и pH7?

#### Решение

Время связывания протона:  $\sim 10^{-3}$  с.

Пояснения.

Формула  $t \sim 0,4 \text{ с/мк}^3 \times V(D/v_1)^{1/3}$  получена в задаче 8.9. Осталось определить  $V$ ,  $D$ ,  $v_1^{1/3}$ .

Так как протон  $\text{H}^+$  путешествует по воде в виде  $(\text{H}_3\text{O})^+$ , можно считать, что  $D \approx 3 \text{ Å}$ .

Так как  $H^+$  «чувствует» отрицательный заряд своего акцептора на расстоянии  $r \approx 8 \text{ \AA}$  от последнего (где энергия взаимодействия  $e^2/\epsilon r = kT$ ), то  $v_1^{1/3} \approx 16 \text{ \AA}$ .

$$V = 1/(10^{-7} \text{ моль/литр}) = 10^7 \text{ литр/моль} = 10^7 \times 10^{15} \text{ мк}^3/6 \times 10^{23} \approx 1/60 \text{ мк}^3.$$

## К лекции 9

### Задача 9.1

Какова константа равновесия  $K$  (спираль/клубок) между спиральным и клубковым состоянием цепи в куске из  $n = 23$  звеньев, если фактор элонгации спирали  $s = 1,1$ , фактор инициации спирали  $\sigma = 0,001$ , а рассматриваемый кусок окружен в своей цепи:

- (а) клубковыми участками и с N-, и с C-конца;
- (б) клубковым участком с N- и спиральным участком с C-конца;
- (в) спиральным участком с N- и клубковым участком с C-конца;
- (г) спиральными участками и с N-, и с C-конца.

Для случая (а) оценить также минимальную длину стабильной спирали, положение концов которой в полипептидной цепи не флуктуирует.

### Решение

$K$  (спираль/клубок) =  $\sigma s^n = 0,009$  в случае (а), когда образование спирали добавляет две границы (в ее начале и в ее конце) спирали и клубка; при этом минимальная длина стабильной спирали, положение концов которой в полипептидной цепи не флуктуирует, составляет  $-\ln(\sigma)/\ln(s) \approx 73$  остатка. Обратите внимание, что длина стабильной спирали с флуктуирующими концами много меньше даже при  $s = 1$ , где она составляет (см. (9.6)) лишь  $\sigma^{-1/2} \approx 32$ .

$K$  (спираль/клубок) =  $s^n = 9$  в случаях (б), (в), когда удлинение спирали не меняет число границ спирали и клубка.

$K$  (спираль/клубок) =  $\sigma^{-1} s^n = 9000$  в случае (г), когда образование спирали ликвидирует две границы спирали и клубка.

### Задача 9.2 (довольно сложная)

Предположим, что в системе имеется очень большое число  $M$  частиц, что время инициации некой структуры вокруг данной частицы равно  $t_{\text{init}}$ , причем  $t_{\text{init}}$  много больше, чем характерное время вовлечения еще одной частицы в уже растущую структуру.

Оцените характерное время, за которое возникающая структура охватит практически всю систему:

(а) если время присоединения новой частицы к растущей структуре равно  $\tau$  вне зависимости от размера этой структуры?

(б) если скорость присоединения новой частицы к растущей структуре, уже охватившей  $m$  частиц, растет пропорционально поверхности этой структуры (т. е. пропорционально  $m^{2/3}$  в трехмерном пространстве), а время присоединения одной новой частицы, соответственно, падает как  $\tau/m^{2/3}$ ?

### Решение

Характерное время перехода всех частиц в новую структуру составляет:

- (а)  $\sim t_{\text{init}}/M$  при  $M < (t_{\text{init}}/\tau)^{1/2}$ ;  
 $\sim \tau(t_{\text{init}}/\tau)^{1/2}$  при  $M \geq (t_{\text{init}}/\tau)^{1/2}$ .
- (б)  $\sim t_{\text{init}}/M$  при  $M < (t_{\text{init}}/\tau)^{3/4}$ ;  
 $\sim \tau(t_{\text{init}}/\tau)^{1/4}$  при  $M \geq (t_{\text{init}}/\tau)^{3/4}$ .

*Объяснение.* Рассмотрим проблему в общем виде. Время присоединения одной новой частицы к растущей структуре, уже охватившей  $m$  частиц, составляет  $\tau/m^L$  [где  $L = 0$  в одномерном случае (а), и  $L = 2/3$  в трехмерном случае (б)]. Тогда время, за которое в растущую структуру (после ее инициации) включится  $m$  частиц, составляет

$$\int_0^m (\tau / m_1^L) dm_1 = \tau \cdot m^{1-L} / (1-L) \quad [\text{т. е.}$$

$\tau \cdot m$  в случае (а) и  $3\tau \cdot m^{1/3}$  в случае (б)].

Предположим, что один центр инициации втягивает в структуру, в конце концов,  $m$  частиц. Тогда на образование этой структуры уйдет время  $t(m) = t_{\text{init}}/m + \tau \cdot m^{1-L}/(1-L)$  — так как то, что центр инициации может находиться в любой из  $m$  частиц финальной структуры, ускорит инициацию в  $m$  раз, а дальнейшее распространение структуры на  $m$  частиц займет  $\tau m^{1-L}/(1-L)$ .

Видно, что инициация занимает основное время при малых  $m$ , а распространение — при больших  $m$ . Видно также, что образование структуры займет меньше всего времени, когда в эту структуру вовлечено такое число частиц  $N$ , что  $d[t(m)]/dm|_{m=N} = 0$ , т. е. когда  $-t_{\text{init}}/N^2 + \tau/N^L = 0$ . Итак,  $N = (t_{\text{init}}/\tau)^{1/(2-L)}$  [т. е.  $N = (t_{\text{init}}/\tau)^{1/2}$  в случае (а) и  $N = (t_{\text{init}}/\tau)^{3/4}$  в случае (б)].

Следовательно: при  $M < N$  время экспансии второстепенно, и весь процесс в целом определяется временем инициации  $M$ -частичной структуры, т. е. он занимает  $\sim t_{\text{init}}/M$  и в случае (а), и в случае (б). — Однако при  $M > N$  система разбивается на  $\sim M/N$  доменов, где структурообразование происходит независимо и — в каждом  $\sim N$ -частичном домене — занимает время порядка

$$\tau \left[ (t_{\text{init}} / \tau) / N + N^{1-L} / (1-L) \right] = \frac{2-L}{1-L} \tau \cdot (t_{\text{init}} / \tau)^{\frac{1-L}{2-L}},$$

т. е.  $\sim \tau \cdot (t_{\text{init}} / \tau)^{\frac{1-L}{2-L}}$  (множителем  $(2-L)/(1-L)$  можно пренебречь: в обоих случаях он  $\sim 1$ ). Значит (ввиду независимости того, что происходит в одном домене, от происходящего в другом) — за то же время  $\sim \tau \cdot (t_{\text{init}} / \tau)^{\frac{1-L}{2-L}}$  возникающие структуры охватят практически всю систему в целом. [В железе и многих других твердых телах хорошо видна стохастическая доменная структура, возникающая вследствие такого процесса. Однако домены в белках имеют другое происхождение, так как они возникают не в «случайных», а в точно определенных местах цепи.]

### Задача 9.3 (довольно сложная)

Свободная энергия клубка равна 0, свободная энергия изгиба  $U = +3$  ккал/моль, добавочная свободная энергия края  $\beta$ -листа  $\Delta f_{\beta} = +0,3$  ккал/моль, а свободная энергия остатка в центре  $\beta$ -листа  $f_{\beta}$  равна: (а)  $-0,54$  ккал/моль; (б)  $-0,18$  ккал/моль; (в)  $-0,06$  ккал/моль.

Оценить минимальный размер цепи, в которой такой  $\beta$ -лист может быть стабильным, и, считая, что время элонгации  $\beta$ -листа на один остаток  $\tau_{\beta} = 10^{-8}$  с, оценить время образования  $\beta$ -листа в цепи из такого числа аминокислотных остатков при температуре  $T=300$  К. Числом аминокислотных остатков в  $\beta$ -изгибе пренебречь. Не  $\beta$ -структурными взаимодействиями и взаимодействием между разными цепями (агрегацией и т. д.) тоже пренебречь.

### Решение

Предварительно отметим, что данные в задаче величины отвечают условиям  $f_{\beta} < 0$  (в противном случае  $\beta$ -лист нестабилен) и  $f_{\beta} > -2\Delta f_{\beta}$  (в противном случае клубок нестабилен против отдельного  $\beta$ -тяжа).

Рассмотрим теперь возможность образования отдельной стабильной  $\beta$ -шпильки.

Условие этого —  $\Delta f_{\beta} + f_{\beta} < 0$ . Ему отвечает только случай (а). Минимальное число  $M$  звеньев в цепи при этом находится из условия  $M(\Delta f_{\beta} + f_{\beta}) + U = 0$ , что, при заданных в случае (а) значениях  $U, \Delta f_{\beta}, f_{\beta}$ , приводит — для шпильки — к  $M_{\text{min}} = 12,5$ .

Максимум свободной энергии (#) при инициации  $\beta$ -шпильки приходится на образование изгиба. Поэтому  $\Delta F^{\#} = U$ , а время инициации  $\beta$ -шпильки в одной точке цепи есть:  $\tau_{\beta} \exp(+U/kT) \approx 1,5 \times 10^{-6}$  с.

Так как после инициации  $\beta$ -шпилька охватывает все  $M_{\text{min}}$  звеньев цепи очень быстро, за  $\tau_{\beta} \cdot M_{\text{min}} \approx 10^{-7}$  с, то  $\tau_{\beta} \exp(+U/kT)$  и есть, фактически, время образования стабильной  $\beta$ -шпильки в столь короткой цепи, что, только перегнутая пополам, она образует стабильную шпильку: этот процесс может начинаться с изгиба только в одной ее точке — в центре цепи.

Это отличает  $\beta$ -шпильку от спирали, которая с равным успехом может начать сворачиваться с любой точки цепи.

Теперь рассмотрим образование  $\beta$ -листов.

Минимальное число  $M_{\text{min}}$  звеньев в цепи, в которой  $\beta$ -лист с заданными значениями  $U, \Delta f_{\beta}, f_{\beta}$  может быть стабильным, находится из уравнения (9.11) при условии равенства нулю свободной энергии  $\beta$ -листа, условия (9.12) и вышеупомянутом условии  $f_{\beta} > -2U\Delta f_{\beta}$ .

Получаем:

$$M_{\text{min}} = [2U\Delta f_{\beta}/(-f_{\beta})^2] \cdot \{1 + [1 + \frac{1}{2}f_{\beta}/\Delta f_{\beta}]\}^{1/2}. \quad (9.3.1)$$

(Второй корень,  $[2U\Delta f_{\beta}/(-f_{\beta})^2] \cdot \{1 - [1 - \frac{1}{2}f_{\beta}/\Delta f_{\beta}]\}^{1/2}$  привел бы к тому, что число  $\beta$ -тяжей в листе  $m = 2/\{1 + [1 + \frac{1}{2}f_{\beta}/\Delta f_{\beta}]\}^{1/2} < 2$ , что невозможно.)

То есть:

(а)  $M_{\text{min}} \approx 11$  звеньев при  $U = +3$  ккал/моль,  $\Delta f_{\beta} = +0,3$  ккал/моль,  $f_{\beta} = -0,54$  ккал/моль;

(б)  $M_{\text{min}} \approx 190$  звеньев при  $U = +3$  ккал/моль,  $\Delta f_{\beta} = +0,3$  ккал/моль,  $f_{\beta} = -0,18$  ккал/моль;

(в)  $M_{\text{min}} \approx 1900$  звеньев при  $U = +3$  ккал/моль,  $\Delta f_{\beta} = +0,3$  ккал/моль,  $f_{\beta} = -0,06$  ккал/моль.

Свободную энергию инициации  $\beta$ -листа ( $\Delta F^{\#}$ ) и время инициации  $\beta$ -листа в одной точке цепи находим по формулам (9.10), (9.15),

$$t_{\text{init } \beta} \sim \tau_{\beta} \exp\{+\Delta F^{\#}/kT\} = \tau_{\beta} \exp\{+[2U\Delta f_{\beta}/(-f_{\beta})]/kT\}, \quad (9.3.2)$$

а время образования  $\beta$ -листа путем его инициации в какой-то точке цепи и его дальнейшего роста есть

$$t_{\beta} \sim t_{\text{init } \beta}/M + M\tau_{\beta}.$$

Инициация образования  $\beta$ -структуры во многих точках, возможная только если  $\tau_{\beta}/M \ll M\tau_{\beta}$ , т. е. если  $\exp([2U\Delta f_{\beta}/(-f_{\beta})]/kT) \ll M^2$ , практически исключена во всех рассматриваемых случаях (проверьте на числах!).

Получаем:

(а)  $\Delta F^{\#} \approx 1,1U = 3,3$  ккал/моль,  $t_{\text{init } \beta} \sim 2,5 \times 10^{-6}$  с,  $t_{\beta} \sim 0,5 \times 10^{-6}$  с (произвольный лист образуется чуть быстрее, чем четко очерченная отдельная шпилька);

(б)  $\Delta F^{\#} \approx 3,3U = 10$  ккал/моль,  $t_{\text{init } \beta} \sim 0,2$  с,  $t_{\beta} \sim 1 \times 10^{-3}$  с;

(в)  $\Delta F^{\#} \approx 10U = 30$  ккал/моль,  $t_{\text{init } \beta} \sim 5 \times 10^{13}$  с,  $t_{\beta} \sim 2 \times 10^{10}$  с.

### Задача 9.4

Предполагая, что  $(-f_{\beta})/(2\Delta f_{\beta}) \ll 1$ , исключить величину  $f_{\beta}$  из уравнений (9.3.1), (9.3.2) и связать между собой полученные в предыдущей задаче величины  $t_{\text{init } \beta}/\tau_{\beta}$  и  $M_{\text{min}}$ .



### Решение

При  $(-f_{\beta})/(2\Delta f_{\beta}) \ll 1$ ,  $M_{\min} \approx [8U\Delta f_{\beta}/(-f_{\beta})^2]$ , т. е.  $(-f_{\beta}) \approx [8U\Delta f_{\beta}/M_{\min}]^{1/2}$ , так что

$$2U\Delta f_{\beta}/(-f_{\beta}) \approx (U\Delta f_{\beta}/2)^{1/2}M_{\min}^{-1/2},$$

$$t_{\text{init}_{\beta}}/\tau_{\beta} \sim \exp\left[\frac{\sqrt{U\Delta f_{\beta}/2}}{kT} \times \sqrt{M_{\min}}\right], \quad (9.4.1)$$

т. е. время образования  $\beta$ -структуры экспоненциально растет с корнем квадратным из числа аминокислотных остатков в минимальном стабильном  $\beta$ -листе — фактически, с потерей свободной энергии на периметре такого  $\beta$ -листа, растущей пропорционально  $M^{1/2}$ .

### Задача 9.5

Вклад аминокислотного остатка в контурную длину цепи составляет около  $3,5 \text{ \AA}$ , а длина Куновского сегмента (т. е. расстояние, на котором цепь «помнит» свое направление) —  $35 \text{ \AA}$ . Каково расстояние между концами клубкообразной цепи (т. е., примерно, диаметр клубка) из 150 аминокислотных остатков?

### Решение

Квадрат расстояния между концами клубкообразной цепи  $L^2 = 3,5 \text{ \AA} \times 150 \times 35 \text{ \AA} = 18375,5 \text{ \AA}^2$ .

Значит,  $L \approx 136 \text{ \AA}$ .

## К лекции 10

### Задача 10.1

Средний молекулярный вес аминокислотного остатка — 110 дальтон.

(а) Оценить средний объем аминокислотного остатка, считая, что плотность белка равна  $1,3 \text{ г/см}^3$ .

(б) Каков диаметр глобулы из 150 аминокислотных остатков?

(в) Найти отношение диаметра глобулы к диаметру клубка из той же цепи, оцененному в задаче 9.5.

(г) Считая, что молекулы начинают интенсивно соприкасаться, когда молекулы занимают  $1/2$  объема — определить, при какой, ориентировочно, концентрации (в мг/мл и в моль/литр) начинают интенсивно соприкасаться рассмотренные выше глобулы и клубки?

### Решение

(а)  $140 \text{ \AA}^3$ ; (б)  $34 \text{ \AA}$ ; (в)  $\approx 1 : 4$ ; (г) глобулы начинают интенсивно соприкасаться при концентрации  $\sim 650 \text{ мг/мл}$  ( $= 0,04 \text{ моль}_{\text{цепей}}/\text{литр}$ ), клубки — при  $\sim 10 \text{ мг/мл}$  ( $= 0,0006 \text{ моль}_{\text{цепей}}/\text{литр}$ ).

### Пояснение:

Если  $\rho$  — плотность белка, а  $m$  — средний молекулярный вес аминокислотного остатка, то его объем

$$V = m/\rho = 110 \text{ г/моль}/1,3 \text{ г/см}^3 =$$

$$= [110 \text{ г}/(0,6 \times 10^{24})]/[1,3 \text{ г}/(10^{24} \text{ \AA}^3)] \approx 140 \text{ \AA}^3.$$

Диаметр глобулы из 150 аминокислотных остатков  $D = (150 \times 140 \text{ \AA}^3 \times 6/\pi)^{1/3} \approx 34 \text{ \AA}$ . Это — в 4 раза меньше, чем диаметр клубка из той же цепи (см. задачу 9.5), т. е. объем клубка примерно в 60 раз больше, чем объем глобулы из 150 аминокислотных остатков.

### Задача 10.2

Оценить соотношение концентраций (в моль/литр) ионов  $\text{H}^+$  при pH7 и заряженных аминокислотных остатков в растворе, в котором, как в клетке, белок составляет  $\sim 10\%$  по весу.

### Решение

1: 2000000.

Если белок составляет в растворе  $10\%$  по весу, а доля заряженных остатков в белке — около  $20\%$ , то таких остатков в 1 литре раствора около  $20 \text{ г}$ , т. е. их концентрация (при молекулярном весе  $\sim 100$  дальтон) — около  $0,2 \text{ моль/литр}$ . А концентрация ионов  $\text{H}^+$  при pH7 —  $10^{-7} \text{ моль/литр}$ .

## К лекции 11

### Задача 11.1

Какая из нижеприведенных последовательностей может кодировать фибриллу из перевитых  $\alpha$ -спиралей, какая — фибриллу коллагена, какая —  $\beta$ -структурный фибриллярный белок:

(а) -Gly-Ala-Gly-Thr-Gly-Ala-Gly-Thr-Gly-Ala-

(б) -Gly-Ala-Pro-Gly-Pro-Pro-Gly-Thr-Pro-Gly-Ala-Pro-Gly-Pro-Pro-

(в) -Gly-Ala-Glu-Ser-Leu-Gly-Asn-Gly-Ala-Glu-Ser-Leu-Gly-Asn-Gly-Ala-

### Решение

(а):  $\beta$ -структурный фибриллярный белок; (б): фибрилла коллагена; (в): перевитые  $\alpha$ -спирали.

## К лекции 12

### Задача 12.1

(а) Почему типичный мембранный белок выглядит либо как пучок  $\alpha$ -спиралей, идущий от одного края мембраны до другого, либо как  $\beta$ -цилиндр, также идущий от одного края мембраны до другого?

(б) Может ли внутри мембраны лежать не  $\beta$ -цилиндр, а  $\beta$ -лист?

### Решение

(а) Потому, что такие структуры обеспечивают отсутствие внутри мембраны не вовлеченных во внутрибелковые взаимодействия доноров и акцепторов водородных связей — а такие «свободные» доноры и акцепторы не могли бы связаться с водой в жирном и потому безводном мембранном окружении (как то они делают в водном окружении), и этот дефицит водородных связей резко снизил бы стабильность структуры белка.

(б) Маловероятно, так как на всех краях  $\beta$ -листа всегда есть не вовлеченные во внутрибелковые взаимодействия доноры и акцепторы водородных связей (в отличие от  $\beta$ -цилиндра, у которого края  $\beta$ -листа спаяны друг с другом), а это резко снижает стабильность  $\beta$ -листа внутри мембраны.

### Задача 12.2

Может ли внутри мембранного белка, в самой середине его пространственной структуры, лежать большой нерегулярный участок?

### Решение

Может — если внутри данный мембранный белок создает широкую, наполненную водой пору, где рассматриваемый нерегулярный участок и помещается. См. рис. 12-4.

### Задача 12.3 (трудная)

Оценить энергию переноса точечного заряда  $q$  из среды с высокой диэлектрической проницаемостью  $\epsilon_1$  в центр бесконечно длинной цилиндрической трубки радиуса  $a$ , имеющей ту же диэлектрическую проницаемость  $\epsilon_1$  и окруженной средой с низкой диэлектрической проницаемостью  $\epsilon_2$ .

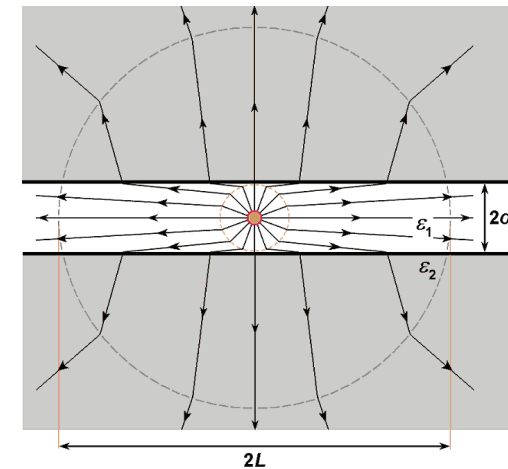
### Решение

Строгое решение этой задачи, приведенное в книге «Электростатика и электродинамика» (В. Смайт, М.: Изд. ИЛ, 1954), требует привлечения специального математического аппарата (бесконечных рядов Бесселевых функций и т. д.) и не дает наглядной физической картины. Поэтому здесь приводится приблизительное, но достаточно точное и наглядное решение, полученное А. В. Ф., Д. Н. Ивановым и А. М. Дыхне [к аналогичному результату пришел S. Teber (частное сообщение) при огрубленном суммировании бесконечных рядов Бесселевых функций].

Картину электрического поля в трубке и вокруг нее можно построить, используя хорошо известную аналогию между ходом силовых линий электрического поля и электрическим током: при этом заряд выступает как электрод, среда с высокой диэлектрической проницаемостью выступает как проводник тока, а среда с низкой диэлектрической проницаемостью — как изолятор.

Непосредственно вокруг заряда линии электрического поля не испытывают влияния среды, окружающей трубку; они расходятся по идущим из центра заряда радиусам, пока, на расстоянии  $a$  от его центра (см. схему), не наткнутся на стенки трубки (так что все точки сферы радиуса  $r = a$  имеют одинаковый потенциал).

Потом они идут в основном вдоль проводящей трубки, лишь постепенно проникая в окружающую ее, плохо проводящую среду. Там они расходятся от трубки примерно по радиусам цилиндра.



И, наконец, на расстояниях  $r > L$  от заряда (величину  $L$  нам предстоит еще определить), силовые линии «забывают» о существовании трубки

и идут сквозь плохо проводящую среду примерно по идущим от заряда радиусам, создавая потенциал  $\varphi(r > L) = q/\varepsilon_2 r$ .

Значит, на расстоянии  $L$ , с которого начинается сферическое распространение электрического поля в среде с диэлектрической проницаемостью  $\varepsilon_2$ , потенциал равен

$$\varphi(L) = q/\varepsilon_2 L. \quad (12.3.1)$$

Изменение потенциала от расстояния  $L$  до расстояния  $a$  от заряда внутри трубки равно

$$\delta\varphi_{a-L, \text{ в трубке}} \approx [(2q/\varepsilon_1 a^2 + q/\varepsilon_2 L^2)/2] (L - a). \quad (12.3.2)$$

Эта оценка получена следующим образом. Длина пути вдоль трубки равна  $L - a$ , а  $(2q/\varepsilon_1 a^2 + q/\varepsilon_2 L^2)/2$  — средняя напряженность электрического поля на этом пути. Средняя — так как напряженность поля составляет  $q/\varepsilon_2 L^2$  на расстояниях  $r \approx L$  (где поле определяется средой с диэлектрической проницаемостью  $\varepsilon_2$ ), а на расстояниях  $r \approx a$  она возрастает до  $\approx 2q/\varepsilon_1 a^2$ . Напряженность равна  $\approx 2q/\varepsilon_1 a^2$  при  $r \approx a$ , потому что здесь почти все силовые линии идут вдоль трубки, через среду с высокой диэлектрической проницаемостью  $\varepsilon_1$ , а сумма площадей двух сечений трубки (справа и слева от заряда), через которые проходят эти линии, равна  $2\pi a^2$ . Последнее вдвое меньше, чем площадь сферы радиуса  $a$ . Значит, поток созданных зарядом силовых линий через эти сечения вдвое плотнее, чем был бы через такую сферу. Соответственно, напряженность поля в этом сечении трубки вдвое выше, чем та напряженность  $q/\varepsilon_1 a^2$ , что была бы на поверхности сферы радиуса  $a$  в однородной среде с той же, что и в трубке, диэлектрической проницаемостью  $\varepsilon_1$ .

Изменение потенциала от расстояния  $L$  до расстояния  $a$  вне трубки, где идет приблизительно цилиндрическое распространение поля, равно

$$\delta\varphi_{a-L, \text{ перн. трубке}} = (q/\varepsilon_2 L) \ln(L/a). \quad (12.3.3)$$

Здесь опять напряженность составляет  $q/\varepsilon_2 L^2$  на расстояниях  $r \approx L$ , а на меньших расстояниях, при цилиндрическом распространении поля, она равна  $q/(\varepsilon_2 L r)$ ; интегрирование этой напряженности по радиусу  $r$  (между  $r = L$  и  $r = a$ ) и приводит к формуле (12.3.3).

При этом, конечно,  $\delta\varphi_{a-L, \text{ в трубке}} = \delta\varphi_{a-L, \text{ перн. трубке}}$  (так как все точки сферы радиуса  $r = a$  имеют одинаковый потенциал, а все точки сферы радиуса  $r = L$  имеют другой, но тоже одинаковый потенциал). Это приводит нас к уравнению

$$[(2q/\varepsilon_1 a^2 + q/\varepsilon_2 L^2)/2] (L - a) = (q/\varepsilon_2 L) \ln(L/a),$$

или

$$[(L/a)^2/(\varepsilon_1/\varepsilon_2) + 1/2] (1 - a/L) = \ln(L/a), \quad (12.3.4)$$

связывающему величину  $L/a$  с величиной  $\varepsilon_1/\varepsilon_2$ .

Уравнение 12.3.4 имеет два решения.

Одно — тривиальное:  $L = a$  при любом соотношении  $\varepsilon_1/\varepsilon_2$ , оно нас не интересует, так как при этом поле вообще не распространяется вдоль трубки, что оно должно делать при  $\varepsilon_1 \gg \varepsilon_2$ .

Другое,

$$L/a \approx (\varepsilon_1/\varepsilon_2)^{1/2} \times [1/2 \ln(\varepsilon_1/\varepsilon_2)]^{1/2}, \quad (12.3.5)$$

получается в интересующем нас случае  $\varepsilon_1 \gg \varepsilon_2$  следующим образом. Так как поле далеко распространяется преимущественно вдоль трубки при  $\varepsilon_1 \gg \varepsilon_2$  (см. схему), то  $L \gg a$ . Поэтому в уравнении (12.3.4) сначала пренебрегают как всеми членами, не содержащими больших величин  $L/a$  и  $\varepsilon_1/\varepsilon_2$ , так и различием величины  $\ln(L/a)$  и 1 (так как логарифм — очень медленно растущая функция). Отсюда получают приближенное соотношение  $(L/a)^2 \sim (\varepsilon_1/\varepsilon_2)$ , или  $2 \ln(L/a) \approx \ln(\varepsilon_1/\varepsilon_2)$ . Затем полученную таким образом величину  $\ln(L/a) = 1/2 \ln(\varepsilon_1/\varepsilon_2)$  подставляют в уравнение (12.3.4), по-прежнему пренебрегая там всеми членами, не содержащими больших величин  $L/a$  и  $\varepsilon_1/\varepsilon_2$ , и получают формулу (12.3.5). Она имеет физический смысл только при  $L > a$ , что имеет место при  $\varepsilon_1 > 2,35\varepsilon_2$ , т. е. в интересующем нас случае  $\varepsilon_1 \gg \varepsilon_2$ .

При этом потенциал на расстоянии  $a$  от центра заряда равен

$$\begin{aligned} \varphi(a) &= \varphi(L) + \delta\varphi_{a-L} = q/\varepsilon_2 L + (q/\varepsilon_2 L) \ln(L/a) = \\ &= (q/\varepsilon_2 L) [1 + \ln(L/a)], \end{aligned} \quad (12.3.6)$$

где величина  $L/a$  определяется уравнением (12.3.5). Значит, он больше, чем потенциал  $\varphi_1(a) = q/\varepsilon_1 a$ , существующий на том же расстоянии  $a$  от заряда в однородной среде с диэлектрической проницаемостью  $\varepsilon_1$ , на величину  $\varphi(a) - \varphi_1(a)$ . На малых расстояниях  $r < a$  поле примерно одно и то же и в однородной среде с диэлектрической проницаемостью  $\varepsilon_1$ , и в трубке радиуса  $a$ , имеющей ту же диэлектрическую проницаемость  $\varepsilon_1$ . Так что потенциал на заряде в трубке выше, чем потенциал на том же заряде в однородной среде, на ту же величину  $\varphi(a) - \varphi_1(a)$ .

Поэтому  $q[\varphi(a) - \varphi_1(a)]/2$  и есть величина искомой энергии переноса заряда  $q$  из среды с диэлектрической проницаемостью  $\varepsilon_1$  в центр длинной трубки радиуса  $a$ , имеющей ту же диэлектрическую проницаемость  $\varepsilon_1$ , и окруженной средой с низкой диэлектрической проницаемостью  $\varepsilon_2$ . Исходя из уравнений (12.3.6), (12.3.5), эту величину можно выразить в виде

$$\Delta U = \frac{q^2}{a\sqrt{\varepsilon_1\varepsilon_2}} \times \left[ \frac{1 + \frac{1}{2} \ln\left(\frac{\varepsilon_1}{2\varepsilon_2} \ln\left(\frac{\varepsilon_1}{\varepsilon_2}\right)\right)}{\sqrt{2 \ln\left(\frac{\varepsilon_1}{\varepsilon_2}\right)}} - \frac{1}{2} \sqrt{\frac{\varepsilon_2}{\varepsilon_1}} \right]. \quad (12.3.7)$$

При этом стоящее в квадратных скобках выражение близко к 1, когда отношение  $\epsilon_1/\epsilon_2$  лежит во всем диапазоне, который может представлять физический интерес (т. е. от  $\sim 10$  до многих сотен), так что, с разумной точностью,

$$\Delta U \approx \frac{q^2}{a\sqrt{\epsilon_1\epsilon_2}}. \quad (12.3.7a)$$

### Задача 12.4 (довольно сложная)

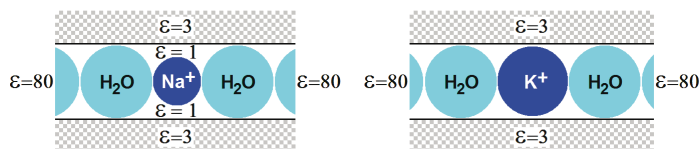
Известно, что мембранные каналы, проводящие  $K^+$ , на два порядка хуже (медленнее) проводят  $Na^+$ ; а те, что проводят  $Na^+$ , на два порядка хуже проводят  $K^+$ . Какие эффекты могут отвечать за такую селективность?

#### Решение

$Na^+$  меньше, чем  $K^+$ : радиус  $Na^+$  равен  $0,95 \text{ \AA}$ , радиус  $K^+$  —  $1,33 \text{ \AA}$  (см. задачу 6.4). Поэтому узкий натриевый канал плохо проводит более крупный  $K^+$ .

Более интересна причина того, что калиевый канал плохо проводит чуть меньший, чем  $K^+$ , ион  $Na^+$ .

В зазор между стенкой калиевого канала и ионом  $Na^+$  не может зайти вода, так как толщина зазора всего около  $1 \text{ \AA}$  (см. схему). Поэтому маленький ион, проникая через несколько более крупное, чем он, отверстие, как бы устраивает вокруг себя вакуумное кольцо, которое, с точки зрения электростатики, сужает этот канал. Это должно несколько повысить свободную энергию этого иона в таком канале.



Точный расчет очень сложен ввиду существенной роли корпускулярности в этой задаче, но ориентировочная оценка может быть сделана. Предположим сначала, что ионный канал в среде с  $\epsilon = \epsilon_{\text{мембраны}}$  сужается с  $R_{K\text{-канал}}$  до  $R_{Na\text{-канал}}$  по всей длине. Тогда энергия иона в канале повышается на  $\Delta U = q^2/(2\epsilon_{\text{эфф}}R_{Na\text{-канал}}) - q^2/(2\epsilon_{\text{эфф}}R_{K\text{-канал}})$ , где  $2\epsilon_{\text{эфф}}R$ , как то сказано в лекциях (и доказано в решении задачи 12.3), примерно равно  $(\epsilon_{\text{воды}}\epsilon_{\text{мембраны}})^{1/2}R \approx (80 \cdot 3)^{1/2}R \approx 15R$ . Поэтому  $\Delta U \approx 6$  ккал/моль при  $R_{K\text{-канал}} = 1,33 \text{ \AA}$ ,  $R_{Na\text{-канал}} = 0,95 \text{ \AA}$ . Однако «электростатическое сужение» канала охватывает только непосредственную окрестность иона, видную из его центра под телесным углом, составляющим примерно половину от полного ( $4\pi$ ). Поэтому результирующий эффект (примерно пропорциональный теле-

сному углу, занимаемому телом с низкой диэлектрической проницаемостью и мало от этой проницаемости зависящий, пока телесный угол не приближается к  $4\pi$ , см. задачу 6.8) должен быть раза в два меньше, чем  $\Delta U$ . Значит, он должен равняться примерно 3 ккал/моль, — что примерно соответствует тому, что калиевые каналы проводят  $Na^+$  на два порядка хуже, чем  $K^+$ .

### Задача 12.5

Предположим, что туннелирующий электрон проникает через барьер шириной  $L$  за  $0,001$  с. За какое время он проникнет через вдвое более широкий барьер той же высоты?

#### Решение

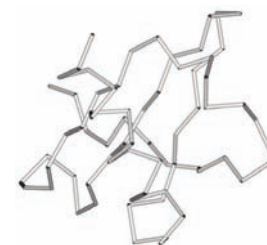
Вероятность туннельного проникновения электрона через потенциальный барьер заданной высоты экспоненциально падает с шириной этого барьера. Если электрон проникает через барьер шириной  $L$  за  $0,001$  с, то вероятность туннелирования электрона сквозь этот барьер за одно колебание электрона (занимающее  $\sim 10^{-15}$  с) составляет  $\sim 10^{-12}$ . Следовательно, электрон проникнет сквозь барьер ширины  $2L$  за  $\sim 10^{-15}/(10^{-12} \times 10^{-12}) \sim 10^9$  с.

Обратите внимание, что время туннельного проникновения через барьер экспоненциально падает с шириной этого барьера, а время проникновения путем теплового возбуждения спадает только пропорционально ширине барьера (см. формулу 8.17).

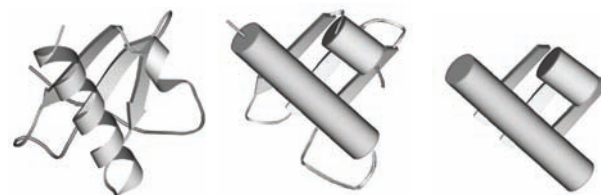
## К лекции 13

### Задача 13.1

На схеме показан ход главной цепи в небольшом белке. Выделить в нем вторичную структуру и затем упростить до «мотива укладки цепи» и, далее, до «штабеля структурных сегментов».



#### Решение

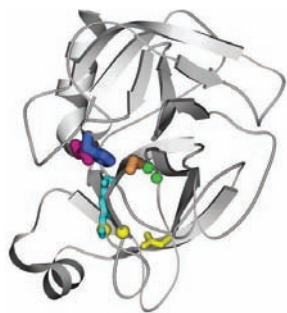


См. Рис. 13–1, 13–3 и стереорисунок Д-15 в перевернутом виде.

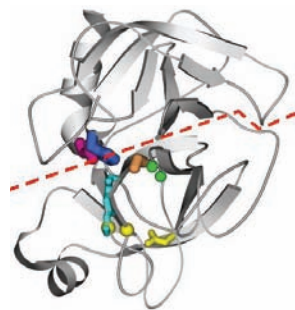


### Задача 13.2

Разделить домены в белке, показанном на схеме.

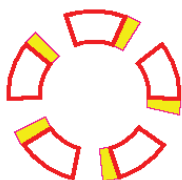


### Решение



### Задача 13.3

Рассмотрим  $\beta$ -цилиндр из пяти  $\beta$ -тяжей (см. схему; на ней тенью показана обычная, правовинтовая, если смотреть вдоль тяжей, скрученность  $\beta$ -листа).



(а) Сколько всего может быть разных мотивов укладки цепи в таком  $\beta$ -цилиндре?

(б) Сколько всего может быть разных мотивов укладки цепи в таком  $\beta$ -цилиндре, если идущие друг за другом в цепи  $\beta$ -тяжи антипараллельны друг другу (т. е. перемычка между ними лежит на торце цилиндра)?

(в) Сколько есть таких  $\beta$ -цилиндров, где идущие друг за другом в цепи  $\beta$ -тяжи антипараллельны друг другу, а перемычки между ними не пересекаются? Нарисуйте те их. Отметьте те, где отсутствует топология «меандр» и «греческий ключ». Отметьте наличие суперспирали из  $\beta$ -тяжей (края которой стянуты Н-связями), ее место в цепи и ее знак.

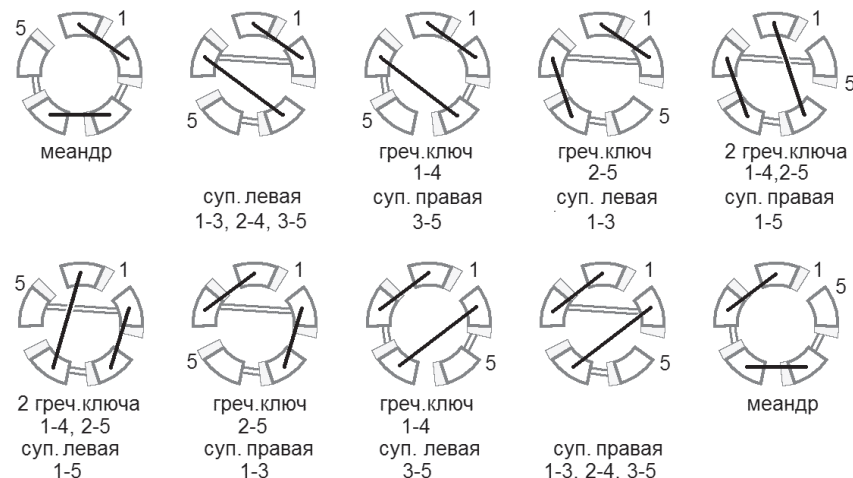
### Решение

(а)  $4! \times 2^4 = 384$  (так как каждый  $\beta$ -тяж белковой цепи может занимать любое из 5 мест в цилиндре и быть направлен в обе стороны, но место

для первого (в цепи)  $\beta$ -тяжа в цилиндре и его направление можно выбрать произвольно).

(б)  $4! = 24$

(в) 10; см следующую схему:



## Классы 14

### Задача 14.1

Последовательность участков вторичной структуры ( $\alpha$  и  $\beta$ ) выглядит как в  $\beta\alpha\beta\alpha\beta\alpha\beta\alpha\beta$  в одном домене и как  $\beta\beta\alpha\beta\beta\beta\alpha\alpha\alpha\alpha$  — в другом. Какой из них принадлежит к классу  $\alpha/\beta$  белков, а какой — к классу  $\alpha+\beta$  белков?

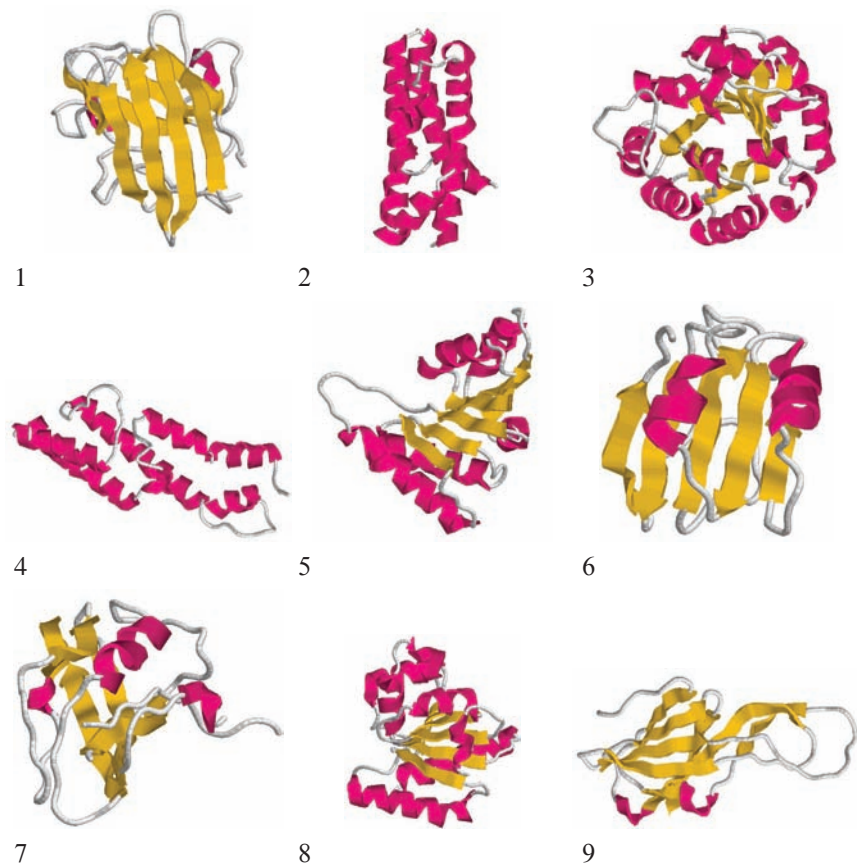
### Решение

К классу  $\alpha/\beta$  белков:  $\beta\alpha\beta\alpha\beta\alpha\beta\alpha\beta$ ; к  $\alpha+\beta$  белков:  $\beta\beta\alpha\beta\beta\beta\alpha\alpha\alpha\alpha$ .

## Классы 15

### Задача 15.1

Изображенные на схеме белковые домены принадлежат к четырем главным для водорастворимых глобулярных белков классам ( $\alpha$ ,  $\beta$ ,  $\alpha/\beta$ ,  $\alpha+\beta$ ).



Они выбраны так, что часть из них принадлежит к наиболее популярным мотивам укладки цепи в данном классе, а часть — к наименее популярным.

- (а) К какому классу принадлежит каждый из доменов?  
 (б) Какие домены имеют популярные мотивы укладки цепи, а какие — редкие?  
 (в) Какой из белков имеет «укладку Россманна»?

### Решение

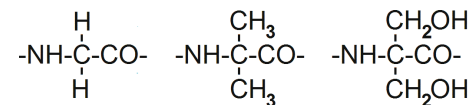
- (а)  $\alpha$ : 2, 4;  $\beta$ : 1, 9;  $\alpha/\beta$ : 3, 5, 8;  $\alpha+\beta$ : 6, 7.  
 (б) Популярные мотивы укладки цепи: 1, 2, 3, 5, 6; остальные (4, 7, 8, 9) — редкие, согласно классификации SCOP. Обычно причины «редкости»

более или менее видны: в структуре 9 перекрываются петли (в правой части этой, не очень компактной структуры); в структуре 4 петли внедряются в центр  $\alpha$ -слоя, и структура не очень компактна; в структуре 7  $\beta$ -тяжи сильно скручены, и  $\beta$ -слой прикрыт не столько спиралями, сколько петлями; однако структура 8 выглядит вполне «нормально»...

(в) 5. Структура 8 похожа на «укладку Россманна», но в ней нет характерного для укладки Россманна порядка тяжей в  $\beta$ -листе и, соответственно, сдвинуто место расхождения (вверх и вниз) исходящих от N-концов  $\beta$ -тяжей петель. В укладке Россманна они расходятся в центре  $\beta$ -листа (состоящего из шести  $\beta$ -тяжей), а в структуре 8, если внимательно присмотреться, — ближе к тому краю  $\beta$ -листа (также состоящего из шести  $\beta$ -тяжей), что ближе к читателю.

### Задача 15.2

Предположим, что «белковая цепь» сложена из зеркально-симметричных аминокислот вроде тех, что показаны на схеме. Предположим также, что выбор нативной пространственной структуры такого «белка» определяется ее максимальной стабильностью для данной первичной структуры.



Может ли такой «белок» иметь одну уникальную по своей стабильности пространственную структуру? Почему?

### Решение

Нет, так как у любой трехмерной укладки цепи есть не совпадающий с ней зеркальный двойник, и обе эти структуры имеют одинаковую энергию, если цепь сложена из зеркально-симметричных звеньев.

Существование такого двойника не угрожает уникальной стабильности пространственной структуры обычного белка, сложенного из L-аминокислот, так как его зеркальный двойник состоит из D-аминокислот, а переход L-аминокислот в D-форму занимает гигантское время.

## К лекции 16

### Задача 16.1

- (а) Вычислить  $\sum_{r \geq R} p^r$  при  $p < 1$ .  
 (б) Вычислить  $\sum_{r \geq R} [rp^r]$  при  $p < 1$ .

### Решение

$$(a) \sum_{r \geq R} p^r = p^R [\sum_{r \geq 0} p^r] = p^R / (1 - p),$$

так как ряд  $p^r$  сходится при  $p < 1$ , стремящемся к бесконечности, и  $\sum_{r \geq 0} p^r = 1 + \sum_{r \geq 1} p^r = 1 + p[\sum_{r \geq 0} p^r]$ , откуда  $(1 - p)[\sum_{r \geq 0} p^r] = 1$ .

$$(b) \sum_{r \geq R} [rp^r] = p d[\sum_{r \geq R} p^r] / dp = p d[p^R / (1 - p)] / dp = [p^R / (1 - p)] [R + p / (1 - p)].$$

### Задача 16.2

В плавающих в воде олигопептидах пролин постоянно переходит из *cis*-формы в *trans*- и обратно. Переход происходит в секундном диапазоне времен, а экспериментально измеренное соотношение *cis*- и *trans*-форм составляет 1:10. В белковых глобулах перехода из *cis*-формы в *trans*- (и из *trans*- в *cis*-) не наблюдается: здесь *cis*-пролины занимают одни положения в цепи белка, а *trans*-пролины — другие. Оцените, какое соотношение между *cis*- и *trans*-формами пролина должно наблюдаться в белковых глобулах?

### Решение

Порядка 1:10 (точнее: белковая статистика показывает, что 1:20), так как статистика элементов нативных белковых глобул сходна с относящейся к 300–400 К Больцмановской статистикой этих элементов, взятых в отдельности.

### Задача 16.3 (сложная)

В теоретических исследованиях белков широко используется так называемая «модель случайных энергий» («Random Energy Model», или просто REM), в простейшей форме описывающая статистические свойства энергетических спектров сложных систем. Приложимость к белкам этой модели, первоначально предложенной Б. Деррида для «неупорядоченных» физических систем типа спиновых стекол и введенной в физику белка Дж. Брингельсоном и П. Волинесом, была обоснована Е. И. Шахновичем и А. М. Гутиним.

Согласно этой модели, энергии разных укладок гетерополимерной цепи не зависят друг от друга.

Это предположение выглядит странно, но Шахнович и Гутин показали, что в переход от одной укладки цепи к другой в плотной глобуле требует столь сильного изменения системы контактов звеньев этой цепи, что, при большом химическом разнообразии звеньев цепи, оно оказывается в достаточной степени справедливым.

Энергия каждой укладки гетерополимера определяется его первичной структурой и системой контактов между звеньями его цепи. Согласно модели случайных энергий, при всевозможных реализациях первичной струк-

туры гетерополимера, каждая его укладка  $i$  (из  $M \gg \gg 1$  возможных глобулярных кладок цепи) обретает энергию  $E$  с вероятностью

$$P_i(E) = (2\pi\sigma_i^2)^{-1/2} \times \exp[-(E - \varepsilon_i)^2 / 2\sigma_i^2], \quad (16.3.1)$$

где  $\varepsilon_i$  — средняя (по всем первичным структурам) энергия этой укладки, а  $\sigma_i$  — среднеквадратичное отклонение  $E$  от  $\varepsilon_i$ . В простейшей версии модели случайных энергий полагают, что все  $\varepsilon_i = 0$  и все  $\sigma_i$  одинаковы.

В рамках этой модели:

(а) Определить (в случае, когда все  $\varepsilon_i = 0$ , а все  $\sigma_i = \sigma$ ) величину энергии  $E_C$ , типичную для нижней границы энергетического спектра.

(б) Определить характерную температуру, соответствующую нижней границе энергетического спектра.

(в) Определить, в том же случае, вероятность того, что одна из  $M$  укладок цепи имеет энергию ниже  $E_C - |\Delta E|$ , а остальные — энергию, превышающую  $E_C$ .

(г) Пусть  $P(E_C)$  соответствует случаю, когда  $\varepsilon_i = 0$ . Как изменится эта вероятность, если  $\varepsilon_i \neq 0$  (при малом  $\varepsilon_i$ ).

### Решение

(а) Пусть все  $\varepsilon = 0$ , все  $\sigma$  одинаковы, а  $p_{E'}$  — вероятность того, что энергия ( $E$ ) укладки ниже уровня  $E'$ . Вероятность того, что энергии всех  $M$  укладок лежат выше  $E'$  есть  $(1 - p_{E'})^M$ . Если  $p_{E'} \ll 1$ , то  $(1 - p_{E'})^M \approx \exp[-Mp_{E'}]$ . Значит, при  $-Mp_{E'} = \ln(1/2)$  все укладки цепи одновременно имеют энергии, превышающие  $E'$ , с шансом  $1/2$ . Итак, характерная величина  $E_C$  может быть найдена из условия  $p_{E_C} = (\ln 2) / M$ . Величина  $p_{E_C}$  очень мала при  $M \gg \gg 1$ . Поэтому

$$\begin{aligned} p_{E_C} &= \int_{-\infty}^{E_C} (2\pi\sigma^2)^{-1/2} \exp[-E^2 / 2\sigma^2] \cdot dE = \\ &= \int_{-\infty}^{E_C} (2\pi\sigma^2)^{-1/2} \exp\{-(E - E_C)^2 / 2\sigma^2\} \cdot dE = \\ &= (2\pi\sigma^2)^{-1/2} \int_{-\infty}^0 \exp[-(x + E_C)^2 / 2\sigma^2] \cdot dx = \\ &= (2\pi\sigma^2)^{-1/2} \exp[-E_C^2 / 2\sigma^2] \int_{-\infty}^0 \exp[x(-E_C) / \sigma^2 - x^2 / 2\sigma^2] \cdot dx \end{aligned}$$

(где  $x = E - E_C$ ). Очевидно, что малой величине  $p_{E_C}$  соответствует  $E_C < 0$  и  $E_C^2 / \sigma^2 \gg 1$  (то есть  $E_C / \sigma \ll -1$ ). Поэтому в том интервале  $x$ , где подынтегральная функция в последнем выражении не мала [и где, следовательно,

$-1 \leq x(-E_C)/\sigma^2 \leq 0$ , т. е.  $(x^2/\sigma^2)(E_C/\sigma)^2 \leq 1$ , и, значит,  $x^2/\sigma^2 \leq (\sigma/E_C)^2 \ll 1$ , величиной  $x^2/\sigma^2$  можно пренебречь, и интеграл равен  $\sigma^2/(-E_C)$ . В результате, получаем уравнение

$$[2\pi(E_C^2/\sigma^2)]^{-1/2} \exp\{-E_C^2/2\sigma^2\} = \ln 2/M.$$

Если  $M$  очень велико, то самыми существенными членами уравнения являются  $\exp\{-E_C^2/2\sigma^2\}$  и  $M$ , так что мы получаем

$$E_C = -\sigma(2 \ln M)^{1/2}. \quad (16.3.2)$$

Это и есть энергия, типичная для нижней границы энергетического спектра (причем середина спектра, по сделанному предположению  $\varepsilon_i = 0$ , приходится на нулевую энергию, а его верхний край — на  $-E_C$ ).

(б) Вблизи этой границы, как видно из предшествующих выкладок, вероятность  $P(E)$  можно записать как  $P(E_C) \times \exp[(E - E_C)(-E_C/\sigma^2)]$ . Это значит, что плотность энергетического спектра,  $m_E = MP(E)$ , достаточно велика уже на расстоянии  $\sim \sigma^2/(-E_C)$  от нижней границы — очень небольшом по сравнению с полушириной этого спектра ( $-E_C$ ). А это значит, что здесь можно вычислить характерную для нижней части спектра температуру:

$$1/T = dS/dE = d[k_B \ln(m_E)]/dE = -k_B E_C/\sigma^2.$$

Окончательно,

$$T_C = \sigma^2/(-k_B E_C) = [\sigma(2 \ln M)^{-1/2}]/k_B \quad (16.3.3)$$

и есть температура, соответствующая нижней границе энергетического спектра.

(в) Вероятность того, что данная укладка цепи имеет энергию  $E_C - |\Delta E|$ , где «энергетическая щель»  $|\Delta E|$  невелика по сравнению с  $|E_C|$ , есть

$$P(E_C - |\Delta E|) \approx P(E_C) \times \exp[(-|\Delta E|)(-E_C/\sigma^2)] = P(E_C) \times \exp[-|\Delta E|/k_B T_C].$$

Вероятность того, что одна из  $M$  укладок цепи имеет энергию  $E_C - |\Delta E|$ , есть

$$\begin{aligned} M \times P(E_C - |\Delta E|) \times (1 - p_{Ec})^{M-1} &= M \times P(E_C) \times \exp[-|\Delta E|/k_B T_C] \times \\ &\times (1 - p_{Ec})^{M-1} \approx \frac{1}{2} \exp[-|\Delta E|/k_B T_C] \\ &\approx \exp[-|\Delta E|/k_B T_C]. \end{aligned} \quad (16.3.4)$$

(г) Если вероятность  $P(E_C)$  соответствует случаю, когда  $\varepsilon_i = 0$ , то, если  $\varepsilon_i \neq 0$  для укладки  $i$  (и  $\varepsilon_i$  мало):

$$\begin{aligned} P_i(E_C) &= (2\pi\sigma^2)^{-1/2} \times \exp[-(E_C - \varepsilon_i)^2/2\sigma_i^2] = P(E_C) \times \exp[E_C \varepsilon_i/\sigma_i^2] \approx \\ &\approx P(E_C) \times \exp[-\varepsilon_i/k_B T_C]. \end{aligned} \quad (16.3.5)$$

### Задача 17.1

Пусть молекула белка имеет лишь два состояния: «0», где и энергия, и энтропия этой молекулы равны нулю, и «1», где энергия этой молекулы равна  $E = const > 0$ , а энтропия равна  $S = const' \gg k$ .

Вслед за Вант-Гоффом, найдите: (а) характерную температуру перехода  $T^*$  из состояния «0» в состояние «1»; (б) характерную ширину области этого перехода по температуре; (в) характерное соотношение концентраций состояния «0» и «1» при температуре, соответствующей началу и концу области перехода; (г) условие, при котором переход будет выглядеть кооперативным.

### Решение

(а)  $T^* = E/S$ ; (б)  $\Delta T = T^*(4k/S)$ ; (в)  $e^2: 1 \approx 7,4: 1$  в начале перехода, и  $1: e^2 \approx 1: 7,4$  в конце; (г) при  $S \gg 4k$ .

### Пояснения.

Больцмановские вероятности состояний «0» и «1» суть

$$P_0 = 1/\{1 + \exp[-(E - TS)/kT]\},$$

$$P_1 = \exp[-(E - TS)/kT]/\{1 + \exp[-(E - TS)/kT]\} = 1 - P_0.$$

Характерная температура перехода  $T^*$  отвечает условию  $P_1 = P_0$ , т. е.

$$T^* = E/S.$$

Величина  $dP_0/d(1/T)$  показывает, как быстро меняется  $P_0$  с обратной температурой  $(1/T)$ .

$$\begin{aligned} dP_0/d(1/T) &= \{\exp[-(E - TS)/kT]/\{1 + \exp[-(E - TS)/kT]\}^2\} \times \\ &\times d[(E - TS)/kT]/d(1/T) = P_0(1 - P_0)E/k. \end{aligned}$$

Эта величина максимальна там, где  $P_0 = 1/2$ , т. е. при температуре  $T^*$ . Здесь  $dP_0/d(1/T) = E/4k$ . С другой стороны, величину  $dP_0/d(1/T)$  можно представить в виде  $\Delta P_0/\Delta(1/T)$ , где  $\Delta P_0$  и  $\Delta(1/T)$  — характерные изменения величин  $P_0$  и  $(1/T)$  при переходе. Поскольку, при  $E > 0$ ,  $P_0(1/T = \infty) \equiv P_0(T = 0) = 1$ , а  $P_0(1/T = 0) \equiv P_0(T = \infty) = \{1 + \exp[S/k]\}^{-1} \approx 0$  при  $S \gg k$ , то  $\Delta P_0 = 1$  при переходе, и его ширина

$$\Delta(1/T) = 1/[E/4k] = 4k/E,$$

или, в более привычных терминах:  $\Delta T \approx 4k(T^*)^2/E = T^*(4k/S)$ .



Таким образом, область перехода простирается от  $1/T' = 1/T'' - 2k/E$  до  $1/T'' = 1/T^* + 2k/E$ . Легко найти, что, при  $T'$ ,  $P_0 = 1/\{1 + e^2\}$  и  $P_1 = e^2/\{1 + e^2\}$ ; а при  $T''$ ,  $P_1 = 1/\{1 + e^2\}$  и  $P_0 = e^2/\{1 + e^2\}$ . Таким образом,

$$P_0/P_1 = e^2/1 \text{ в начале перехода, и } P_0/P_1 = 1/e^2 \text{ в конце.}$$

Переход будет выглядеть кооперативным, если он узок, т. е. при  $\Delta(1/T) \ll 1/T^*$  (или, что то же самое, при  $\Delta T \ll T^*$ ), что происходит при  $S \gg 4k$ .

### Задача 17.2

Какую общую формулу вы можете предложить для описания величины химического потенциала частицы в растворе, где она находится в очень малой концентрации?

#### Решение

$\mu = kT \ln[C] + \mu^0$ , где  $C$  — концентрация наших частиц в растворе, а не зависящий от этой концентрации потенциал  $\mu^0$  зависит от температуры и от взаимодействия отдельной частицы с окружающим ее растворителем.

*Объяснение.* В задаче 8.2 мы получили формулу для координатной составляющей химического потенциала частицы в газе:  $\mu^{\text{коорд}} = -kT \ln [V/N]$ , где  $T$  — температура, а  $N$  — число частиц, заключенных в объем  $V$ . Величина  $C = N/V$  есть, по определению, концентрация частиц.

Кроме координатной составляющей, у химического потенциала частицы в газе есть еще не зависящая от концентрации  $C$  кинетическая составляющая и свободная энергия внутримолекулярных взаимодействий, а для частицы в жидкости, дополнительно, еще и свободная энергия ее взаимодействия с окружающим ее растворителем. Эти члены и объединены в  $\mu^0$ .

### Задача 17.3

Пусть молекула белка имеет лишь два состояния: «0», где ее свободная энергия равна 0, а плавающие в растворе частицы А (скажем, протоны) с ней не связываются, и «1», где с ней связываются  $n_A$  частиц А, так что свободная энергия этой молекулы становится равной  $F = F_1 - n_A \mu_A > 0$  (где член  $F_1$  — собственная свободная энергия белка в состоянии «1» и его связей с  $n_A$  частицами А, а член  $-n_A \mu_A$  учитывает, что  $n_A$  частиц А извлечено из раствора, где их химический потенциал был  $\mu_A$ ).

Переход вызывается изменением величины  $\mu_A$ : он происходит при  $\mu_A^* - \Delta \mu_A/2 \rightarrow \mu_A^* + \Delta \mu_A/2$ , где  $\mu_A^*$  соответствует середине перехода.

Найдите аналог уравнения Вант-Гоффа для такого процесса.

#### Решение

$$n_A \Delta \mu_A = 4kT, \text{ где } T \text{ — температура; } \mu_A^* = F_1/n_A.$$

Решение аналогично решению задачи 17.1. Больцмановские вероятности состояний «0» и «1» суть  $P_0 = 1/\{1 + \exp[-(F_1 - n_A \mu_A)/kT]\}$ ,  $P_1 = 1 - P_0$ .

Величина  $dP_0/d\mu_A = -P_0(1 - P_0)(n_A/kT)$  показывает, как быстро меняется  $P_0$  с химическим потенциалом частиц А. Она максимальна при  $P_0 = 1/2$ . Здесь  $F_1 = n_A \mu_A$ , а  $dP_0/d\mu_A = -n_A/4kT$ . С другой стороны,  $dP_0/d\mu_A = \Delta P_0/\Delta \mu_A$ , где  $\Delta P_0$  и  $\Delta \mu_A$  — характерные изменения величин  $P_0$  и  $\mu_A$  при переходе. Поскольку  $P_0 = 1$  при полном отсутствии частиц А в системе (когда  $C = 0$ , так что  $\mu_A = -\infty$ , согласно задаче 17.2), и  $P_0 \rightarrow 0$  при  $C \rightarrow \infty$  (когда и  $\mu_A \rightarrow \infty$ ), то  $\Delta P_0 = -1$  при переходе от  $C = 0$  до  $C = \infty$ , и  $n_A \Delta \mu_A = 4kT$ .

Таким образом, зная соответствующее переходу изменение химического потенциала частиц А (для  $H^+$ , например,  $\Delta \mu = 2,3kT \Delta pH$ ), мы можем найти число  $n_A = 4kT/\Delta \mu_A$  частиц А, связавшихся с кооперативной единицей перехода, и, далее, если мы знаем, сколько частиц А поглощено белковой молекулой в целом, — можно узнать, совпадает ли эта кооперативная единица перехода с белковой молекулой в целом или нет. Число же поглощенных белком частиц А можно узнать, сравнив числа этих частиц, необходимые для повышения химического потенциала  $\mu_A$  на  $\Delta \mu_A$  в присутствии и в отсутствие белка, и разделив эту разность на число молекул белка в растворе. В случае, когда «частицами А» являются протоны, такой метод называется потенциометрическим титрованием.

### Задача 17.4

Белок плавится (по типу «все-или-ничего») при температуре  $T^* = 350$  К, причем ширина области плавления  $\Delta T = 7^\circ$ .

(а) Какова разность энтальпий  $\Delta H$  нативного и денатурированного состояний белка в точке плавления? Какова разность энтропий  $\Delta S$  нативного и денатурированного состояний белка в точке плавления?

(б) Какова разность свободных энергий  $\Delta G$  нативного и денатурированного состояний белка при температуре плавления  $T^*$ ?

(в) Какова разность свободных энергий  $\Delta G$  нативного и денатурированного состояний белка на  $10^\circ$  ниже температуры плавления  $T^*$ ?

#### Решение

(а) В силу соотношения Вант-Гоффа,  $\Delta T = 4k(T^*)^2/\Delta E$ , где  $\Delta E$  — повышение энтальпии кооперативной единицы при переходе типа «все-или-ничего». Значит, энтальпия денатурированного состояния белка на  $\Delta E$  выше, чем энтальпия нативного, а разность энтальпий нативного и денатурированного состояний  $\Delta H = -\Delta E = -4kT^* \times (T^*/\Delta T) = -4 \times (0,7 \text{ ккал/моль}) \times (50) = -140 \text{ ккал/моль}$ . А так как  $\Delta S = \Delta H/T^*$ , то  $\Delta S = (-140 \text{ ккал/моль})/(350 \text{ К}) = -400 \text{ кал/(моль}\cdot\text{град)}$ .

(б) Ноль — по определению точки плавления!

(в) Так как  $\Delta S = -\partial(\Delta G)/\partial T$ , то  $\Delta G(T^* - 10^\circ) \approx \Delta G(T^*) - 10^\circ \times (-\Delta S) = 0 - 10^\circ \times [400 \text{ кал/(моль}\cdot\text{град)}] = -4 \text{ ккал/моль}$ .

### Задача 17.5 (несложная, но требует вычислений...)

Белок плавится (по типу «все-или-ничего») при температуре  $T^* = 350$  К, причем ширина области плавления  $\Delta T = 7^\circ$ , а рост теплоемкости белка при плавлении равен 3 ккал/(моль·град).

Считая, что разность теплоемкостей нативного и денатурированного белка не зависит от температуры, — определить:

(а) как зависит от температуры разность энтальпий  $\Delta H$  нативного и денатурированного состояний белка?

(б) как зависит от температуры разность энтропий  $\Delta S$  нативного и денатурированного состояний белка?

(в) как зависит от температуры разность свободных энергий  $\Delta G$  нативного и денатурированного состояний белка?

(г) при какой температуре белок максимально стабилен?

(д) каковы  $\Delta S$ ,  $\Delta H$ ,  $\Delta G$  белка при этой температуре?

(е) при какой температуре белок претерпевает холодную денатурацию?

### Решение

В задаче 17.4 рассмотрен тот же белок и определены разности энтальпий  $\Delta H(T^*)$ , энтропий  $\Delta S(T^*)$  и свободных энергий  $\Delta G(T^*)$  его нативного и денатурированного состояний при температуре плавления  $T^*$ . Что касается разности теплоемкостей этих двух состояний, то, поскольку рост теплоемкости белка при плавлении есть  $C_{p\text{денатур.}} - C_{p\text{натив.}}$ , то разность теплоемкостей его нативного и денатурированного состояний есть  $\Delta C_p = C_{p\text{натив.}} - C_{p\text{денатур.}} = -(C_{p\text{денатур.}} - C_{p\text{натив.}}) = -3$  ккал/(моль·К).

(а) Так как  $\Delta C_p = d(\Delta H)/dT$ , то  $\Delta H(T) = \Delta H(T^*) + \Delta C_p(T - T^*)$ .

(б) Так как  $\Delta C_p = d(\Delta H)/dT = d(\Delta G + T\Delta S)/dT = d(\Delta G)/dT + \Delta S + T[d(\Delta S)/dT] = T[d(\Delta S)/dT]/dT$ , то  $d(\Delta S)/dT = \Delta C_p/T$ , и  $\Delta S(T) = \Delta S(T^*) + \Delta C_p \int_{T^*}^T dT/T = \Delta S(T^*) + \Delta C_p \ln(T/T^*)$ .

(в) Так как  $dG/dT = -S$ , то  $\Delta G(T) = \Delta G(T^*) - \int_{T^*}^T \Delta S(T)dT = 0 - \int_{T^*}^T [\Delta S(T^*) + \Delta C_p \cdot \ln(T/T^*)]dT = -\Delta S(T^*)(T - T^*) - \Delta C_p T^* [(T/T^*) \ln(T/T^*) - T/T^* + 1]$ , т. е.  $\Delta G(T) \approx -\Delta S(T^*) \cdot (T - T^*) - \Delta C_p(T - T^*)^2/2T^*$  при  $|T - T^*| \ll T^*$ .

(г) Белок максимально стабилен при той температуре  $T_{st}$ , при которой  $\Delta G(T)$  минимально. В минимуме  $\Delta G$  должно быть:  $d(\Delta G)/dT = 0$  [т. е.  $\Delta S(T) = 0$ ], и  $d^2(\Delta G)/dT^2 > 0$  [последнее всегда выполняется при  $\Delta C_p < 0$ , так как  $d^2(\Delta G)/dT^2 = -d(\Delta S)/dT = -\Delta C_p/T$ ]. Условие  $\Delta S(T) = 0$  означает, что  $\Delta S(T^*) + \Delta C_p \ln(T_{st}/T^*) = 0$ . Тогда  $\ln(T_{st}/T^*) = -\Delta S(T^*)/\Delta C_p$ , и  $T_{st}/T^* = \exp[-\Delta S(T^*)/\Delta C_p] = \exp[0,4/3] = 0,875$ , так что, при  $T^* = 350$  К,  $T_{st} = 306,3$  К  $\approx 33^\circ$ С.

(д) При  $T = T_{st}$ :  $\Delta S(T_{st}) = 0$  (см. выше),  $\Delta G(T_{st}) = \Delta H(T_{st}) = \Delta H(T^*) + \Delta C_p(T_{st} - T^*) = -140$  ккал/моль +  $[-3$  ккал/(моль·К)]  $\times [43,3$  К] =  $-10$  ккал/моль.

(е) Белок претерпевает холодную денатурацию, когда  $\Delta G(T)$  при понижении температуры опять (после  $T^* = 350$  К) проходит через 0. Так как  $\Delta G(T) \approx -\Delta S(T^*) \cdot (T - T^*) - \Delta C_p \cdot (T - T^*)^2/2T^*$ , то температура холодной денатурации  $T^{**}$  находится из условия  $T^{**} - T^* \approx -[\Delta S(T^*)/\Delta C_p] \cdot 2T^*$ , т. е.  $T^{**} - T^* \approx -90^\circ$ , и, значит,  $T^{**} \approx 350$  К  $- 90^\circ$ , т. е. холодная денатурация происходит на  $\approx -13^\circ$ С (так что она может наблюдаться только в переохлажденной воде или при добавлении антифриза).

### Задача 17.6 (сложная)

Часто применяемые денатуранты белков, мочевины ( $\text{NH}_2\text{—CO—NH}_2$ ) и гуанидин гидрохлорид ( $[(\text{NH}_2)_3\text{C}]^+ \text{Cl}^-$ ), имеют избыток  $\text{NH}_2$  групп, в то время как многие «ренатуранты» (вещества, повышающие стабильность белковых структур), например,  $\text{Na}_2\text{SO}_4$  и многие сахара, имеют избыток атомов О или ОН групп.

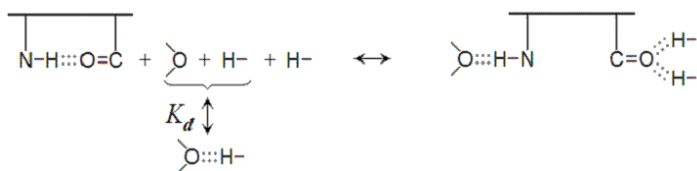
Объясните этот феномен, исследовав влияние добавления в раствор этих веществ на стабильность водородных связей в главной цепи белка. [Совет. Прежде всего, внимательно посмотрите на рисунки в Лекции 4 и на картинки Д4–Д6, Д106, Д12–Д13, Д15]. Считать, что образование и распад водородных связей подчиняется «закону действующих масс», выражаемому уравнением  $[\text{—O}] \cdot [\text{H—}]/[\text{—O::H—}] = K_d$ , причем, так как большинство водородных связей в воде завязано,  $K_d$  много меньше, чем суммарная концентрация и доноров ( $[\text{H}] = [\text{H—}] + [\text{—O::H—}]$ ), и акцепторов ( $[\text{O}] = [\text{—O}] + [\text{—O::H—}]$ ) водородной связи в растворе.

### Решение

Свободная энергия, потребная на разрыв водородной связи во вторичной структуре полипептида, падает при добавлении в раствор «избыточных» доноров и растет при добавлении в раствор «избыточных» акцепторов водородной связи.

Это связано с тем, что каждый атом водорода Н–О (или Н–N) группы способен завязать только одну водородную связь, а каждый атом кислорода — две (эта разница хорошо видно на картине водородных связей во льду, рис. 4-1). В то же время, внутри вторичной структуры О-атом С=О группы (как и Н-атом Н–N группы) имеет только одну водородную связь. При этом атомы С=О::Н–N в ней лежат практически на прямой линии, что затрудняет образование еще одной водородной связи этой С=О группой с еще одним Н-атомом растворителя. К тому же скрепляющие вторичную структуру белка водородные связи часто погребены внутри белковой глобулы, что вообще исключает образование еще одной водородной связи такой С=О группой белка с растворителем.

После разворачивания белка, сопровождаемого разрывом водородных связей внутри него, каждая его N–H группа может завязать *одну* связь с O-атомом растворителя, а каждая его C=O группа может завязать *две* связи с H-атомами растворителя [суммарная энергия такой «вилочковой» связи, видимо, всего в полтора (а не в два) раза больше, чем энергия двух одинарных водородных связей, но нам это сейчас несущественно]. Так как константа диссоциации водородной связи  $K_d$  мала, можно полагать, что при разворачивании белка образуются практически все водородные связи с растворителем, которые становятся возможными после разрыва внутрибелковых водородных связей (см. схему).



Образование тех трех водородных связей, что образуются N–H и C=O группами белка после разрыва связи между ними, требует привлечения одного свободного от водородной связи O-атома и *двух* свободных от водородной связи H-атомов.

Рассмотрим сначала разрыв внутрибелковых водородных связей в водном окружении.

Таких «свободных от связей» H и O атомов в воде мало, в силу малости константы диссоциации  $K_d$  и равенства числа акцепторных и донорных «валентностей для образования водородной связи» у молекул  $H_2O$ , где каждый H служит донором в одной водородной связи, а каждый O (которых вдвое меньше, чем H-атомов) — акцептором в двух связях.

Что же произойдет, если мы добавим к чистой и столь уравновешенной воде «примесь», содержащую избыток «свободных» H или O атомов?

Содержание соответствующего атома в «свободной» (от водородной связи) форме, естественно, возрастет, и это повышает его химический потенциал (см. задачу 17.2). Однако *произведение* концентраций «свободных» H и O атомов мало меняется при добавлении «примеси» — это произведение подчиняется уравнению  $[O] \cdot [H] = K_d [O:H]$ , причем концентрация связанных атомов,  $[O:H]$ , близка к 100 % при малой  $K_d$  и потому практически неизменна. Поэтому и *сумма* химических потенциалов свободных атомов, равная химическому потенциалу их связанной формы (см. задачу 5.6), также практически неизменна, т. е. уменьшение химического потенциала того атома, которого в растворе мало, равно увеличению химического потенциала того атома, который, благодаря добавлению примеси, находится в растворе в избытке.

За вовлечение в водородную связь с белком каждого «свободного» атома надо платить его химическим потенциалом: ведь в растворе становится одним таким «свободным» атомом меньше. Однако образование водородных связей «свободными» (от внутрибелковых связей) N–H и C=O группами белка с растворителем требует *разного* количества «свободных» H и O атомов (N–H группе нужен один O-атом, C=O группе — два H-атома, см. схему). Эта-то разница (при том, что разрыв внутрибелковых водородных связей освобождает *равное* число N–H и C=O групп) и приводит к тому, что избыток H-атомов денатурирует белки, а избыток O-атомов их ренатурирует.

Если в раствор добавлены дополнительные N–H группы, там становится сравнительно много свободных от водородной связи H-атомов, но «свободные» O-атомы становятся редкими (так как эти добавленные N–H группы, если они не связываются с белком сами, отвлекают на себя часть O-атомов воды). При этом образование одной водородной связи «свободной» C=O группой белка облегчается, а образование одной водородной связи его «свободной» N–H группой в той же мере затрудняется.

А так как N–H группа привлекает *лишь один* «свободный» O-атом, а C=O группа — *два* «свободных» H-атома растворителя, то, в сумме, образование водородных связей с растворителем «освобожденными» N–H и C=O группами белка *облегчается* при избытке «свободных» H-атомов (которых нужно много), несмотря на недостаток в растворе «свободных» O-атомов (которых нужно мало).

Это объясняет, почему свободная энергия разрыва внутрибелковой водородной связи *падает* при добавлении в раствор «избыточных» доноров водородной связи, т. е. почему вещества с избыточными N–H группами служат денатурантами белков.

Если же, наоборот, в раствор добавлены дополнительные O-атомы, то они отвлекают на себя часть H-атомов воды, и «свободные» H-атомы становятся редкими.

А так как N–H группа привлекает *лишь один* «свободный» O-атом, а C=O группа — *два* «свободных» H-атома, то, в сумме, образование водородных связей «свободными» N–H и C=O группами белка *затрудняется* при недостатке «свободных» H-атомов (которых нужно много), несмотря даже на избыток в растворе свободных O-атомов (которых нужно мало).

Это объясняет, почему свободная энергия разрыва внутрибелковой водородной связи *растет* при добавлении в раствор «избыточных» акцепторов водородной связи, т. е. почему вещества с избыточными O-атомами служат ренатурантами белков.

### Задача 17.7 (нетрудная, но требует вычислений...)

Опираясь на решение задач 17.6 и 5.7 и считая, что в растворе образуется 85 % от максимально возможного числа водородных связей, — оценить изменение свободной энергии разрыва водородной связи в белке при добавлении в раствор одного моля «избыточных» (а) доноров и (б) акцепторов водородной связи.

#### Решение

Разорванная внутрибелковая водородная связь привлекает к себе два свободных донора (Н-) и один свободный акцептор (О-) водородной связи из растворителя, «заплатив» за это их химическими потенциалами, т. е. из свободной энергии образовавшихся водородных связей с этими донорами и акцепторами надо вычесть  $2\mu_{\text{H}} + \mu_{\text{O}}$  в сумме.

При добавлении Н- или О- в раствор, величина  $-(2\mu_{\text{H}} + \mu_{\text{O}})$  меняется только за счет изменения концентраций [Н-] и [О-], так как сами по себе энергии водородных связи и все остальные взаимодействия, входящие в химические потенциалы, практически не зависят от малой «примеси» свободных Н или О атомов. Не зависит от этой «примеси» и сумма  $\mu_{\text{H}} + \mu_{\text{O}}$ , так как она равна химическому потенциалу связанной формы -Н:О-, а концентрация [-О::Н-] этой формы близка к 100%, и потому меняется мало.

Значит, величину  $-(2\mu_{\text{H}} + \mu_{\text{O}})$  можно представить в виде  $-1,5(\mu_{\text{H}} + \mu_{\text{O}}) - 0,5(\mu_{\text{H}} - \mu_{\text{O}}) = \text{const} - 0,5kT \cdot \ln ([\text{H-}]/[\text{O-}])$ . Величина const не зависит от «примесей» и поэтому нас не интересует, так что изменение свободной энергии разрыва водородной связи при добавлении «избыточных» доноров и акцепторов этой связи представляется в виде

$$\Delta G = -kT \cdot \ln \{([\text{H-}]/[\text{O-}])^{1/2}\}. \quad (17.7.1)$$

Воспользовавшись формулой (5.7.1) для химических потенциалов ассоциирующих молекул, отсюда можно получить:

(а) В случае добавления доноров:

$$\Delta G = -kT \cdot \ln \{1 + ([\text{H}] - [\text{O}])/(K_d[-\text{O}::\text{H-}])^{1/2}\}. \quad (17.7.2)$$

(б) В случае добавления акцепторов:

$$\Delta G = kT \cdot \ln \{1 + ([\text{O}] - [\text{H}])/(K_d[-\text{O}::\text{H-}])^{1/2}\}. \quad (17.7.3)$$

В обеих формулах, [Н] — суммарная концентрация в растворе доноров (связанных и несвязанных водородной связью), [О] — суммарная концентрация в растворе акцепторов, а  $K_d$  — константа диссоциации водородной связи.

Поскольку в 1 литре воды находится  $1000/18 = 55,6$  молей  $\text{H}_2\text{O}$ , и каждая из молекул  $\text{H}_2\text{O}$  содержит 2 акцептора и 2 донора водородной связи, то суммарная концентрация акцепторов водородной связи в воде — 111 моль/литр, и суммарная концентрация доноров — тоже 111 моль/литр. 85 % из них (около 94 моль/литр) связаны водородной связью в форме -О::Н-, а остальные ( $[\text{H-}] \approx 17$  моль/литр доноров и  $[\text{O-}] \approx 17$  моль/литр акцепторов) — свободны. Значит, константа диссоциации водородной связи,  $K_d = [\text{H-}] \cdot [\text{O-}]/[-\text{O}::\text{H-}] \approx 3$  моль/литр, а  $(K_d[-\text{O}::\text{H-}])^{1/2} \approx 17$  моль/литр.

Если  $|\text{[H]} - \text{[O]}| \ll 17$  моль/литр, то (17.7.2) и (17.7.3) приобретают общую форму:

$$\begin{aligned} \Delta G &\approx kT \cdot ([\text{O}] - [\text{H}]) / (K_d[-\text{O}::\text{H-}])^{1/2} \approx \\ &\approx kT \cdot ([\text{O}] - [\text{H}]) / (17 \text{ моль/литр}). \end{aligned} \quad (17.7.4)$$

Из этого следует, что добавлении в раствор одного моля «избыточных» доноров водородной связи изменяет свободную энергию разрыва водородной связи на  $-0,035$  ккал/моль, тем самым облегчая ее разрыв, а добавлении одного моля «избыточных» акцепторов изменяет свободную энергию разрыва водородной связи на  $+0,035$  ккал/моль, что затрудняет ее разрыв.

Цифра  $-0,035$  ккал/моль соответствует  $1/2$  моль/литр мочевины ( $\text{NH}_2\text{-CO-NH}_2$ ), обладающей двумя избыточными акцепторными валентностями (= 4 акцепторные NH — 2 донорные у О). Она же относится к  $1/4$  моль/литр гуанидин гидрохлорида ( $[(\text{NH}_2)_3\text{C}]^+\text{Cl}^-$ ), обладающего четырьмя избыточными акцепторными валентностями (= 6 акцепторные NH — 2 донорные у Cl). А цифра  $+0,035$  ккал/моль соответствует  $1/8$  моль/литр  $\text{Na}_2\text{SO}_4$ , — молекулы-ренатуранта, имеющей восемь акцепторных валентностей.

## К лекции 18

### Задача 18.1

Опыты по денатурации единичных белковых молекул делаются так. Один конец молекулы химически пришивается к подложке, другой — к щупу атомного силового микроскопа, и к этому щупу прикладывается некая сила, разворачивающая белок.

Оценить силу, потребную для разворачивания белка из 100 аминокислотных остатков, стабильность нативного состояния которого, по сравнению с клубком, составляет 10 ккал/моль, если время наблюдения не ограничено.



### Решение

Для разворачивания белка нужно затратить работу  $A = 10$  ккал/моль. Эта работа должна быть затрачена на пути  $L$ , равном разности между линейным размером вытянутой цепи и белковой глобулы, т. е.  $L \approx 350 \text{ \AA} - 40 \text{ \AA} \approx 300 \text{ \AA}$ . Искомая сила  $F = A/L = (10 \text{ ккал/моль})/300 \text{ \AA} \approx (42 \times 10^3 \text{ Дж}/6 \times 10^{23}) / (300 \times 10^{-10} \text{ м}) \approx 2 \times 10^{-12} \text{ Н} = 2$  пиконьютона.

### Задача 18.2

Вопрос о том, какое число сортов звеньев дает возможность цепи однозначно определять пространственную структуру, можно переформулировать в виде следующей задачи:

Предположим, что цепь из  $N$  звеньев имеет  $M$  укладок, одинаковых по такому свойству, как возможность быть оптимальной структурой для какой-то последовательности звеньев (что предполагает также одинаковость этих укладок по таким общим свойствам, как компактность, содержание вторичной структуры и т. д.).

Какое число сортов звеньев  $K$  может обеспечить существование у каждой последовательности звеньев только одной укладки с минимальной для этой последовательности энергией?

### Решение

$$K \geq M^{1/N}$$

Если имеется  $K$  сортов звеньев, то у цепи из  $N$  звеньев есть  $K^N$  возможных последовательностей. Если  $K^N < M$  и каждая последовательность имеет только одну оптимальную укладку, то эти  $K^N$  оптимальных укладок никак не покроют все  $M$  возможных укладок (которые предполагались равноценными с точки зрения возможности быть оптимальными хоть для какой-то последовательности). Это противоречие показывает, что  $K^N$  должно превосходить  $M$ , т. е.  $K \geq M^{1/N}$ .

### Задача 18.3 (сложная)

В задаче 16.3 описана «модель случайных энергий» и, в частности, тот вариант этой модели, что широко используется в теоретических исследованиях белков. В нем «нативная» укладка цепи гетерополимера резко отличается от прочих («ненативных») своей высокой стабильностью.

Математически эта модель формулируется так:  $M-1$  из  $M \gg 1$  укладок цепи гетерополимера обретают свою «случайную» энергию  $E$  с вероятностью

$$P(E) = (2\pi\sigma^2)^{-1/2} \times \exp[-E^2/2\sigma^2], \quad (18.3.1)$$

а оставшаяся одна («нативная») укладка имеет более низкую, чем все прочие, энергию  $E_N = E_C - |\Delta E|$ , где

$$E_C = -\sigma (2\ln M)^{1/2} \quad (18.3.2)$$

— энергия, типичная для нижней границы «случайного» энергетического спектра «ненативных» укладок. [Здесь и ниже речь идет только о координатной части энергии, энтропии и т. д., т. к. их кинетические составляющие не меняются при конформационных переходах]

Эта модель имеет три параметра:  $M$  — число укладок цепи гетерополимера;  $\sigma$  — среднее квадратичное отклонение энергий «ненативных» укладок от средней, принятой за 0 величины;  $|\Delta E|$  — ширину энергетической щели, отделяющей «нативную» укладку от «ненативных».

Установить, как связаны эти параметры с основными наблюдаемыми параметрами плавления белка: температурой плавления  $T_M$ ; теплотой плавления  $\Delta H$ ; скачком теплоемкости при плавлении,  $\Delta C_p$ .

Особо рассмотреть случай «узкой щели», когда  $|\Delta E| \ll |E_C|$ : он, видимо, наиболее характерен для белковых глобул.

### Решение

Рассмотрим сначала энергетический спектр «ненативных» укладок цепи.

Если  $P(E)$  — вероятность того, что данная «ненативная» укладка имеет энергию  $E$ , то  $m_E = (M-1)P(E)dE \approx MP(E)dE$  — среднее число «ненативных» укладок, имеющих энергию в диапазоне  $E - E + dE$ . Следовательно,  $S(E) = k \cdot \ln(m_E) = k(-E^2/2\sigma^2 + \text{const})$  — энтропия «ненативных» укладок, имеющих энергию  $E$ . Постоянная  $\text{const}$  определяется из того условия, что  $S(E_C) = 0$ , так как спектр энергий ненативных укладок начинается с  $E_C$  (и кончается в районе  $-E_C$ ), а  $E_C = -\sigma(2\ln M)^{1/2}$ . Итак, энтропия «ненативных» укладок

$$S(E) = k(E_C^2/2\sigma^2 - E^2/2\sigma^2) \geq 0 \quad (18.3.3)$$

определена в диапазоне энергий  $E_C \leq E \leq -E_C$ , где существует в достаточно плотный энергетический спектр «ненативных» укладок, а вне этого диапазона «ненативных» укладок нет и там  $S(E)$  не определена.

Каждой энергии  $E$  спектра «ненативных» укладок соответствует температура

$$T(E) = [dS/dE]^{-1} = \sigma^2/(-kE) \quad (18.3.4)$$

(при  $E_C \leq E \leq -E_C$ ), и свободная энергия при этой температуре  $T(E)$  есть

$$F(E) = E - T(E) \cdot S(E) = E + (E_C^2 - E^2)/2E. \quad (18.3.5)$$

Теперь вернемся к нативной укладке. Ее энергия  $E_N = E_C - |\Delta E|$ , а энтропия ее равна 0 (так как она одна); значит, ее свободная энергия равна  $E_N$  при любой температуре.

Термодинамическое равновесие нативной и ненативных укладок наступает при той энергии  $E$  равновесных укладок, когда  $F(E) = E_N$ , т. е.

$$E/2 + E_C^2/(2E) = E_N. \quad (18.3.6)$$

Решением возникающего при этом уравнения

$$E^2 - 2E E_N + E_C^2 = 0 \quad (18.3.7)$$

является

$$E_1 = E_N + [E_N^2 - E_C^2]^{1/2} \quad (18.3.8)$$

(второй корень,  $E_2 = E_N - [E_N^2 - E_C^2]^{1/2}$ , не имеет физического смысла, т. к. он меньше  $E_N$ , а при этих энергиях никаких укладок нет вообще).

Следовательно, теплота плавления

$$\Delta H = E_1 - E_N = [E_N^2 - E_C^2]^{1/2} = |\Delta E|^{1/2} \cdot [|\Delta E| + |2E_C|]^{1/2}, \quad (18.3.9)$$

температура плавления

$$T_M = T(E_1) = \sigma^2/(-kE_1) = (\sigma^2/k)/[|E_C| + |\Delta E| - \Delta H], \quad (18.3.10)$$

а скачок теплоемкости при плавлении (от  $C_N = dE_N/dT \equiv 0$  до  $C_{\text{denat}} = (dE/dT)|_{E=E_1} = (dT/dE)^{-1}|_{E=E_1}$ ) есть

$$\Delta C_P = (dT/dE)^{-1}|_{E=E_1} = kE_1^2/\sigma^2,$$

— или, с учетом того, что  $T_M \equiv T(E_1) = \sigma^2/(-kE_1)$ , —

$$\Delta C_P = \sigma^2/(kT_M^2). \quad (18.3.11)$$

Уравнения (18.3.9), (18.3.10), (18.3.11) связывают все три параметра  $|\Delta E|$ ,  $\sigma$  и  $|E_C|$  «модели случайных энергий» с тремя основными наблюдаемыми параметрами плавления белка ( $\Delta H$ ,  $T_M$ ,  $\Delta C_P$ ), причем соотношение  $(2 \ln M)^{1/2} = |E_C|/\sigma$  [см. (18.3.2)] позволяет оценить и  $M$ , число укладок белковой цепи.

Если  $|\Delta E| \ll |E_C|$ , то уравнения (18.3.9), (18.3.10) приобретают совсем простую форму:

$$\Delta H \approx |2\Delta E/E_C|^{1/2} \cdot |E_C|, \quad (18.3.12)$$

$$T_M \approx \{\sigma^2/(k|E_C|)\} / \{1 - |2\Delta E/E_C|^{1/2}\} = \{\sigma^2/k\} / \{|E_C| - \Delta H\}. \quad (18.3.13)$$

Отсюда:

$$\sigma^2 = (kT_M^2) \Delta C_P, \quad (18.3.14)$$

$$|E_C| = \Delta H + \Delta C_P T_M, \quad (18.3.15)$$

$$|\Delta E| = \frac{1}{2} \Delta H^2 / [\Delta H + \Delta C_P T_M]. \quad (18.3.16)$$

*Добавление.*

Возникает вопрос: можно ли использовать выведенные в этой задаче уравнения для получения параметров  $M$ ,  $\sigma$  и  $|\Delta E|$  той «модели случайных энергий», что соответствует белку, основные термодинамические параметры плавления которого ( $T_M$ ,  $\Delta H$ ,  $\Delta C_P$ ) известны из опыта.

*Ответ:*

Формально — да. Однако, к сожалению, это нельзя сделать прямо, без введения дополнительных поправок, так как экспериментальные термодинамические параметры зависят не только от перестроек укладок цепи (принимаемых во внимание моделью случайных энергий), но и от изменения гидрофобных и электростатических взаимодействий с температурой, которое не учитывается моделью случайных энергий, но дает большой вклад в измеряемые величины (в особенности — в величину  $\Delta C_P$ , фигурирующие в каждом из уравнений 18.3.14–18.3.16).

## К лекции 19

### Задача 19.1

Предположим, что выбор одной из  $2^{100}$  структур происходит в 3 кинетических этапа. На первом этапе выбирают себе какие-то структуры из состоящего из  $2^{70}$  структур поднабора, соответствующего первому интермедиату. На втором, не выходя из этого поднабора, цепи выбирают какие-то из  $2^{40}$  структур поднабора, соответствующего второму интермедиату. И на последнем этапе выбирается единственная структура.

Предполагается, что в рамках каждого этапа выбор структуры происходит путем полного перебора, а «ложные» (не соответствующие «правильным» интермедиатам сворачивания «правильной» структуры) энергетические минимумы отсутствуют.

(а) Считая, что переход от структуры к структуре занимает  $10^{-10}$  с, — оценить, сколько времени займет весь процесс.

(б) Какое падение энергии на первом, втором и третьем этапе требуется для того, чтобы цепь не возвращалась на предыдущую стадию? [Считать, что  $kT = 0,6$  ккал/моль.]

### Решение

(а) Перебор на первом этапе требует выбора одной структуры из  $2^{100}/2^{70} = 2^{30} \approx 10^9$  структур; этого достаточно для попадания в 1-й интермедиат. Этот процесс займет  $10^{-10} \text{ с} \times 10^9 \sim 0,1 \text{ с}$ , если цепь не будет «проваливаться» в не соответствующие «правильному» интермедиату энергетические минимумы (а если будет — то больше). Перебор на втором этапе (этапе попадания во 2-й интермедиат) требует выбора одной структуры из  $2^{70}/2^{40} = 2^{30} \approx 10^9$  структур; при том же условии, он тоже займет  $\sim 0,1 \text{ с}$ . Перебор на третьем этапе требует выбора одной структуры из  $2^{40} \approx 10^{12}$  структур: он займет  $10^{-10} \text{ с} \times 10^{12} \sim 100 \text{ с}$  (опять-таки при условии отсутствия «неправильных» энергетических минимумов). Итого — порядка 100 секунд как минимум.

(б) На первом этапе: хотя бы на  $kT$  больше, чем  $kT \ln(2^{30}) \approx 12,5 \text{ ккал/моль}$ ; на втором: столько же; на третьем: хотя бы на  $kT$  больше, чем  $kT \ln(2^{40}) \approx 16,7 \text{ ккал/моль}$ ; итого — не менее  $3kT + kT \ln(2^{30}) \approx 44 \text{ ккал/моль}$ .

### Задача 19.2 (сложная)

Теплоемкость нативной белковой глобулы  $C = 5 \text{ ккал/(К·моль)}$  при комнатной температуре. Каков средний размер тепловых флуктуаций энергии  $E$  этой глобулы?

### Решение

Средний размер тепловых флуктуаций энергии любой системы, по определению, есть  $\delta E = [\sum_i w_i (\varepsilon_i - E)^2]^{1/2}$ , где  $E = \sum_i w_i \varepsilon_i$  — средняя энергия системы,  $\varepsilon_i$  — энергия  $i$ -го микросостояния этой системы, а  $w_i = \exp(-\varepsilon_i/kT) / [\sum_i \exp(-\varepsilon_i/kT)]$  — вероятность  $i$ -го микросостояния при температуре  $T$ ; все суммы здесь берутся по всем микросостояниям рассматриваемой системы. Значит,

$$\delta E^2 = \sum_i w_i \varepsilon_i^2 - (\sum_i w_i \varepsilon_i)^2.$$

В то же время, теплоемкость системы

$$C = dE/dT = \sum_i (dw_i/dT) \varepsilon_i = \sum_i w_i \cdot \varepsilon_i / kT^2 - (\sum_i w_i \varepsilon_i)^2 / kT^2.$$

Следовательно,

$$C = \delta E^2 / kT^2 \text{ и } \delta E = [C \cdot kT^2]^{1/2}.$$

При  $C = 5 \text{ ккал/(К·моль)}$  и  $T = 300 \text{ К}$ ,  $\delta E = 30 \text{ ккал/моль}$ .

### Задача 19.3

Вычисленный в предыдущей задаче размер флуктуаций энергии  $E$  нативной белковой глобулы много больше, чем разность свободных энергий денатурированного и нативного белка, составляющая около  $10 \text{ ккал/моль}$  при комнатной температуре. Почему эти тепловые флуктуации не разваливают нативную глобулу?

### Решение

Энергия тепловых флуктуаций нативного состояния белка действительно может превзойти разность свободных энергий денатурированного и нативного белка, причем энергия нативного состояния, возбужденного тепловыми флуктуациями, может превзойти среднюю энергию денатурированного состояния белка (разность средних энергий этих состояний составляет всего  $\sim 10 \text{ ккал/моль}$  в точке максимальной стабильности белка, см. задачу 17.5, хотя при температуре теплового плавления она на порядок больше).

Однако тепловые флуктуации не разваливают нативную глобулу, так как нативное состояние (включая его флуктуации) занимает одну область конформационного пространства, а денатурированные состояния занимают другую область этого пространства. Переход же из одной области в другую (поскольку в белках это — переход типа «все-или-ничего») требует преодоления активационного барьера, причем это требует не только подъема энергии, но движения не в любую сторону, а лишь в определенном, ведущем к другому состоянию направлении (что ограничивает энтропию переходного состояния).

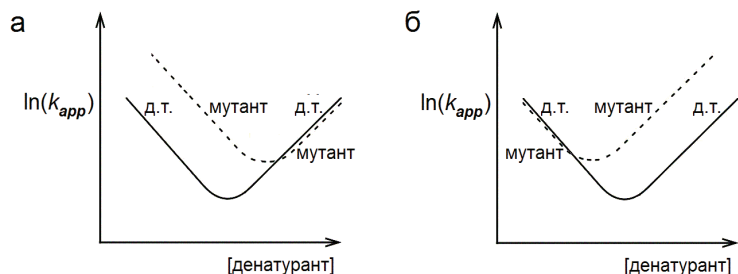
Впрочем, нельзя сказать, что тепловые флуктуации вовсе не могут разрушить стабильную нативную глобулу: изредка и на короткое время им это удается — см. задачу 20.3.

## К лекции 20

### Задача 20.1

Предположим, что в некоем белке мутация одного аминокислотного остатка дала тот сдвиг шевронного графика мутанта относительно шевронного графика белка «дикого типа» (д. т.), что показан на схеме **а**, а мутация другого — тот, что показан на схеме **б**.

Что мы можем сказать о вовлеченности этих остатков в ядро сворачивания белка?



### Решение

Показанный на схеме **а** сдвиг шевронного графика свидетельствует об изменении скорости сворачивания ( $k_f$ ) при практическом отсутствии изменения скорости разворачивания ( $k_u$ ) белка. Значит, мутация влияет на стабильность ядра сворачивания так же, как на стабильность нативной структуры белка, определяемую соотношением  $k_f/k_u$ . Значит, первый рассматриваемый остаток в это ядро вовлечен полностью.

Обратная ситуация, показанная на схеме **б**, свидетельствует, что второй рассматриваемый остаток в это ядро не вовлечен совсем.

### Задача 20.2

В точке термодинамического равновесия нативного и развернутого состояний лизоцима, каждая молекула которого испытывает переход типа «все-или-ничего», скорость приближения раствора белка к равновесию состояний равна примерно  $0,0003 \text{ с}^{-1}$  (см. рис. 20-4).

- Какова при этом скорость сворачивания белка?
- Каково характерное время сворачивания белка?
- Какова скорость и каково характерное время разворачивания белка?
- Как выглядит кинетика процесса приближения раствора белка к равновесию, если белок первоначально находился в 8М растворе мочевины, а в момент времени «0» этот раствор был разбавлен водой до концентрации 4.2 М мочевины, соответствующей точке термодинамического равновесия нативного и развернутого состояний белка?

(д) Как выглядит процесс сворачивания-разворачивания, если, в описанном выше опыте, мы долго следим за отдельной молекулой белка?

(е) Как выглядит кинетика процесса приближения раствора белка к равновесию, если белок первоначально находился в чистой воде, а в момент времени «0» в этот раствор влили мочевины, доведя ее концентрацию до 4,2 М, что соответствует точке термодинамического равновесия нативного и развернутого состояний белка?

(ж) Как выглядит процесс сворачивания-разворачивания отдельной молекулы этого же белка, если, в описанном выше опыте, мы долго следим за отдельной молекулой белка?

### Решение

(а) Скорость сворачивания белка:  $\frac{1}{2} \times$  (скорость приближения к равновесию) =  $0,00015 \text{ с}^{-1}$ .

(б) Характерное время сворачивания белка:  $1/(\text{скорость сворачивания белка}) = 6,7 \times 10^3 \text{ с} \approx 2 \text{ часа}$ .

(в) В точке термодинамического равновесия скорость разворачивания равна скорости сворачивания, т. е. она равна  $0,00015 \text{ с}^{-1}$ , характерное время разворачивания есть  $1/0,00015 \text{ с}^{-1} = 6,7 \times 10^3 \text{ с} \approx 2 \text{ часа}$ .

(г) Доля нативного белка: = 0 до момента «0» разбавления мочевины, а после него (в точке термодинамического равновесия нативного и денатурированного белка) эта доля составляет  $0,5 \times [1 - \exp(-0,00015 \cdot t)]$ , где  $t$  — время (в секундах), прошедшее после момента «0».

(д) После разбавления мочевины отдельная молекула белка остается развернутой порядка двух часов, потом практически мгновенно сворачивается, потом остается свернутой порядка двух часов, потом практически мгновенно разворачивается, потом остается развернутой порядка двух часов, потом практически мгновенно сворачивается... и т. д. [Примечание: «порядка двух часов» означает, что для некоторых молекул это время составляет полчаса, для других — час, для третьих — два, для четвертых — четыре, а в среднем — 2 часа.]

(е) Доля нативного белка: = 1 до момента «0» добавления мочевины, а после него: =  $0,5 \times [\exp(-0,00015 \cdot t) + 1]$ , где  $t$  — время (в секундах), прошедшее после момента «0».

(ж) После добавления мочевины отдельная молекула белка остается свернутой порядка двух часов, потом практически мгновенно разворачивается, потом остается развернутой порядка двух часов, потом практически мгновенно сворачивается... и т. д.

### Задача 20.3

Предположим, что белок испытывает переход типа «все-или-ничего» из развернутого состояния в нативное при комнатной температуре, причем скорость его сворачивания равна  $3 \text{ с}^{-1}$ , а нативное состояние стабильнее развернутого на  $7,5 \text{ ккал/моль}$ .

(а) Каково характерное время пребывания этого белка (после сворачивания) в свернутом состоянии?

(б) Каково характерное время пребывания этого белка (после разворачивания) в развернутом состоянии?



(в) Как выглядит при заданных условиях процесс сворачивания-разворачивания, если мы долго следим за отдельной молекулой белка?

### Решение

(а) Примерно сутки. Если нативное состояние стабильнее развернутого на 7,5 ккал/моль при комнатной температуре, то скорость разворачивания белка в  $\exp[(7,5 \text{ ккал/моль})/(0,6 \text{ ккал/моль})] = 2,5 \times 10^5$  раз меньше, чем скорость его сворачивания. Значит, скорость разворачивания белка составляет  $(3 \text{ с}^{-1})/(2,5 \times 10^5) \approx 1/(83000 \text{ с}) \approx 1 \text{ сутки}^{-1}$ , а его характерное время —  $1/(1 \text{ сутки}^{-1}) = 1 \text{ сутки}$ .

(б) Примерно  $1/(3 \text{ с}^{-1}) = 0,33 \text{ с}$ .

(в)... порядка 0,3 секунды пребывания в развернутом состоянии — практически мгновенное сворачивание — порядка суток пребывания в свернутом состоянии — практически мгновенное разворачивание — порядка 0,3 секунды пребывания в развернутом состоянии — практически мгновенное сворачивание —...

## К лекции 21

### Задача 21.1

Для белка из  $N = 100$  аминокислотных остатков — оценить, очень ориентировочно и только по порядку величины, время сворачивания в районе точки денатурации, если:

(а) мы вовсе не знаем пространственной структуры этого белка?

(б) если мы знаем, что в его  $\alpha$ -спиралях 70 водородных связей?

(в) каково будет время сворачивания, при комнатной температуре, у белка из 100 аминокислотных остатков, с 70 водородными связями в его  $\alpha$ -спиралях в таких условиях, что стабильность его нативной структуры составляет 6 ккал/моль, причем мы знаем, что белок сворачивается без видимых интермедиатов сворачивания?

### Решение

(а)  $\sim 10 \text{ нс} \times \exp(100^{2/3}) \sim 10 \text{ с}$  в среднем; от  $\sim 10 \text{ нс} \times \exp(0,5 \times 100^{2/3}) \sim 10^{-3} \text{ с}$  до  $\sim 10 \text{ нс} \times \exp(1,5 \times 100^{2/3}) \sim 10^5 \text{ с}$  в качестве крайних пределов;

(б) так как в спирали не входит  $100 - 70 = 30$  Н-связей, то, в среднем, время сворачивания  $\sim 10 \text{ нс} \times \exp(30^{2/3}) \sim 0,16 \text{ мс}$ ;

(в) так как ядро сворачивания охватывает, грубо говоря, половину белка, то стабильность ядра повысится примерно на 3 ккал/моль, так что:  $\sim 0,16 \text{ мс} \times \exp[-(3 \text{ ккал/моль})/(0,6 \text{ ккал/моль})] \sim 1 \text{ мкс}$ .

### Задача 21.2 (довольно сложная)

Свободная энергия зарождающегося в растворе кристаллика представлена в виде

$$F = Mf + 6M^{2/3}\psi, \quad (21.2.1)$$

где  $M$  — число молекул в кристаллике,  $f$  — свободная энергия молекулы в кристаллической фазе по сравнению с жидкой,  $\psi$  — дополнительная свободная энергия молекулы кристаллической фазы на ее поверхности.

Считая, только для простоты, что характерное время  $\tau$  встраивания молекулы из жидкой фазы в растущий кристаллик не зависит ни от размера кристаллика, ни от формы его поверхности, —

(а) оценить характерное время  $t_{init}$  инициации кристаллической фазы в заданной точке пространства;

(б) связать  $t_{init}$  с минимальным размером  $M_{st}$  стабильного кристаллика;

(в) оценить характерное время  $t_{trans}$  перехода в кристаллическую фазу в большом объеме;

(г) получить численные оценки при комнатной температуре, если  $f = -0,2$  ккал/моль,  $\psi = 0,3$  ккал/моль,  $\tau = 10^{-11}$  с; если  $f = -0,1$  ккал/моль,  $\psi = 0,3$  ккал/моль,  $\tau = 10^{-11}$  с.

[Подсказка. Воспользоваться решениями задач к главе 9, но учесть, что теперь рассматривается не двухмерный, а трехмерный случай.]

### Решение

Из  $Mf + 6M^{2/3}\psi = 0$  следует минимальный размер стабильного кристаллика:  $M_{st} = [6\psi/(-f)]^3$ .

Из  $d(Mf + 6M^{2/3}\psi)/dM = 0$  следует размер критического зародыша кристалла,  $M_1 = [4\psi/(-f)]^3 = 8/27 M_{st}$ , и его свободная энергия относительно жидкости:  $\Delta F^\# = 2\psi(M_1)^{2/3} = 8/9\psi(M_{st})^{2/3}$ .

Характерное время инициации кристалла в заданной точке пространства:

$$t_{init} = \tau \exp[+\Delta F^\#/kT] = \tau \cdot \exp\left[\frac{4\psi}{9kT} \cdot M^{2/3}\right].$$

Характерное время перехода в кристаллическую фазу в большом объеме:

$$t_{trans} = (t_{init}\tau)^{1/2} = \tau \cdot \exp\left[\frac{4\psi}{9kT} \cdot M_{st}^{2/3}\right].$$

Численные оценки при  $T = 300 \text{ К}$  (т. е.  $kT = 0,6$  ккал/моль),  $\psi = 0,3$  ккал/моль,  $\tau = 10^{-11}$  с:



A:	0	0	$\frac{1}{3}$	0	$\frac{2}{3}$	0	0	0	0	0	$\frac{1}{3}$	0	0	0	0	0	0	0	0
C:	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D:	0	$\frac{2}{3}$	0	0	0	0	0	0	0	0	$\frac{1}{3}$	0	0	0	0	0	0	0	1
E:	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
F:	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
P:	1	0	0	0	0	$\frac{1}{3}$	0	0	1	0	0	0	0	0	0	0	0	0	0
S:	0	0	$\frac{1}{3}$	0	0	0	0	0	0	0	$\frac{2}{3}$	0	0	0	$\frac{1}{3}$	$\frac{2}{3}$	0	0	0
T:	0	0	0	0	0	0	$\frac{2}{3}$	0	0	0	0	0	0	0	0	$\frac{1}{3}$	0	0	$\frac{2}{3}$
V:	0	$\frac{1}{3}$	0	0	0	0	$\frac{1}{3}$	0	0	0	$\frac{1}{3}$	0	1	0	0	0	0	0	0
-:	0	0	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$	0	0	0	$\frac{1}{3}$	0	0	0	$\frac{2}{3}$	0	0	0	0	$\frac{1}{3}$

(г) Опознавание гомологичной последовательности с использованием этого профиля:

**ПРОФИЛЬ:**

А.к. в профиле:	-																			V				
За совпадение:	$+\frac{1}{3}$																			$+\frac{1}{3}$				
А.к. в профиле:	V	S	-														-	V	-	A	-		T	-
За совпадение:	$+\frac{1}{3}+\frac{1}{3}$	$+\frac{1}{3}+\frac{2}{3}+\frac{1}{3}$	$+\frac{1}{3}+\frac{2}{3}+\frac{1}{3}$														$+\frac{1}{3}+\frac{1}{3}$	$+\frac{2}{3}+\frac{1}{3}$	$+\frac{1}{3}+\frac{1}{3}$		$+\frac{1}{3}$			
А.к. в профиле:	P	D	A	C	A	P	T	E	P	S	D	F	V	S	S	D	T							
За совпадение:	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$							

**Последов. 1:** F V S F A V A A A S V - F S S A T  
**Цена:**  $0+\frac{1}{3}+\frac{1}{3}$   $0+\frac{2}{3}$  0 0 0 0  $0+\frac{2}{3}+\frac{1}{3}$  -1  $0+\frac{1}{3}+\frac{2}{3}$   $0+\frac{2}{3}$  = +3  
 не гомологична

**ПРОФИЛЬ:**

А.к. в профиле:	-																			V		
За совпадение:	$+\frac{1}{3}$																			$+\frac{1}{3}$		
А.к. в профиле:	V	S	-														-	A	-		T	-
За совпадение:	$+\frac{1}{3}+\frac{1}{3}$	$+\frac{1}{3}+\frac{2}{3}+\frac{1}{3}$	$+\frac{1}{3}+\frac{2}{3}+\frac{1}{3}$														$+\frac{1}{3}+\frac{1}{3}$	$+\frac{2}{3}+\frac{1}{3}$	$+\frac{1}{3}+\frac{1}{3}$		$+\frac{1}{3}$	
А.к. в профиле:	P	D	A	C	A	P	T	E	P	-	S	D	F	V	S	S	D	T				
За совпадение:	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$	$+1+\frac{2}{3}+\frac{1}{3}$				

**Последов. 2:** P F F C E - E E P T T T F V - A D A  
**Цена:** +1 0 0 +1 0  $-\frac{2}{3}$  0 +1 +1 -1 0 0 +1 +1  $-\frac{2}{3}$  0 +1 0 = +4 $\frac{2}{3}$   
 гомологична

*Пояснение дробных чисел в «Цене»:* -1 за делецию («-») в последовательности, плюс  $(\frac{2}{3}) \times (+\frac{1}{2})$ , где  $+\frac{1}{2}$  — цена совпадения «-» в последовательности и «-» во множественном выравнивании, а  $\frac{2}{3}$  — вес «-» в данном месте множественного выравнивания.

**Задача 22.3.**

Пользуясь приведенной на рис. 22-6 схемой — оценить, какая из нижеприведенных последовательностей содержит  $\alpha$ -спираль, какая — нерегулярную петлю, какая —  $\beta$ -тяж и какая —  $\beta$ -изгибы:

- 1: Gly Thr Ser Val Thr Val Ser Phe Thr Trp Gly Gly
- 2: Ser Gly Gly Pro Gln Glu Ser Asn Gly Tyr Arg Pro
- 3: Val Ala Gly Pro Gly Ser Ser Gly Pro Pro Gly Val
- 4: Gly Pro Ala Glu Gln Leu Leu Gln Lys Met Arg Gly

*Решение*

1:  $\beta$ -тяж, 2: нерегулярная петля, 3:  $\beta$ -изгибы, 4:  $\alpha$ -спираль.

*К лекции 23*

**Задача 23.1**

Предположим, что при работе по предсказанию структуры белка характерная погрешность в определении энергии структуры составляет  $\pm\sigma$ , причем вероятность того, что погрешность имеет величину  $\Delta E$ , подчиняется Гауссовому закону:  $P(\Delta E) = (2\pi\sigma^2)^{-1/2} \times \exp[-\Delta E^2/2\sigma^2]$ .

Такой расчет прекрасно различает две структуры, разность между энергиями которых составляет  $10\sigma$ : он, практически уверенно, может сказать, какая из них стабильнее другой. (Так как разность двух энергий оценивается со стандартной погрешностью  $\pm\sqrt{2}\sigma$ , то в зазоре шириной  $10\sigma$  лежит  $\approx 7$  стандартных погрешностей, т. е. вероятность ошибки в определении существования зазора составляет, согласно таблице вероятностей ошибок,  $\sim 10^{-11} \sim e^{-25}$ . Эту же оценку можно получить иначе: Вероятность того, что погрешность поднимет энергию низкоэнергетичной на  $5\sigma$ , составляет  $\sim e^{-25/2}$ ; вероятность того, что погрешность снизит энергию высокоэнергетичной на  $5\sigma$ , составляет тоже  $\sim e^{-25/2}$ ; значит, вероятность того, что эти два события произойдут одновременно, составляет  $\sim e^{-25/2} \cdot e^{-25/2} = e^{-25}$ ).

Предположим теперь, что мы должны сравнить энергии  $M$  структур (где  $M \gg 1$ ), причем только одна из них стабильна, т. е. имеет низкую энергию, а энергии всех остальных структур больше на  $10\sigma$ .

При каком числе  $M$  результат расчета с описанной выше погрешностью перестанет быть достоверным, т. е. рассчитанная энергия одной из высокоэнергетических структур сравняется с рассчитанной энергией той структуры, что на самом деле имеет на  $10\sigma$  более низкую энергию?

### Решение

Пусть истинная энергия низкоэнергетической структуры равна 0, а высокоэнергетических —  $10\sigma$ . Тогда спектр расчетных энергий  $M$  высокоэнергетических структур простирается, согласно решению задачи 16.3 [см. формулу (16.3.2) и ниже], от  $\approx 10\sigma - \sigma(2\ln M)^{1/2}$  до  $\approx 10\sigma + \sigma(2\ln M)^{1/2}$ . Если  $10\sigma - \sigma(2\ln M)^{1/2} \leq 0$ , то расчетная энергия какой-то или каких-то высокоэнергетических структур перекроет расчетную энергию истинно самой низкоэнергетической структуры. Итак, критическая величина  $M$  находится из уравнения

$$10\sigma \approx \sigma(2\ln M)^{1/2}.$$

Результат:  $M \sim e^{50} \approx 10^{22} \approx 2^{70}$ ; а так как у каждого аминокислотного остатка есть больше двух возможных конформаций, то результат такого расчета перестанет быть достоверным для цепи из более, чем 70 остатков.

## К лекции 24

### Задача 24.1

Иммуноглобулин связывает некую молекулу X с константой связывания  $K_{ass} = \frac{[Ig-X]_{eq}}{[Ig]_{eq}[X]_{eq}} = 10^8$  литр/моль при комнатной температуре. Предполо-

жим, что эта молекула имеет форму шара, что она твердая (т. е. не имеет внутренних степеней свободы) и что она имеет в иммуноглобулине колебания с амплитудой  $a = 0,5 \text{ \AA}$ .

Оценить  $\Delta G_0$ , разность между свободной энергией взаимодействия этой молекулы с иммуноглобулином и свободной энергией взаимодействия с водой.

### Решение

$$\Delta G_0 \approx -15,5 \text{ ккал/моль.}$$

#### Объяснение.

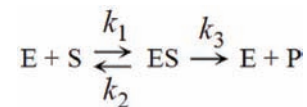
При концентрации  $[X]_0 = 1/K_{ass}$  частиц X в растворе,  $[Ig-X]/[Ig] = 1$ , т. е. иммуноглобулин с равной вероятностью находится в формах Ig-X и Ig. Значит, в этом случае свободные энергии этих двух форм совпадают. Извлекая частицу из раствора, иммуноглобулин затрачивает работу на уменьшение доступного частице объема: в растворе на ее долю приходился объем  $V_0 = 1/[X]_0 = K_{ass}$ , а в иммуноглобулине — только объем ее колебаний,  $V_1 =$

$= (2a)^3$ . Соответствующая работа равна  $kT \ln(V_0/V_1)$ . Ценой этой работы иммуноглобулин приобретает дополнительное взаимодействие с частицей (правда, взамен тех, что у его связывающего кармана и у связываемой поверхности частицы были с водой). Разность  $\Delta G_0$  между свободной энергией взаимодействия частицы с иммуноглобулином и свободной энергией их утраченного взаимодействия с водой должна компенсировать работу, затраченную на уменьшение доступного частице объема: ведь свободные энергии форм Ig-X и Ig совпадают. Значит,  $\Delta G_0 + kT \ln(V_0/V_1) = 0$ , и, окончательно,  $\Delta G_0 = -kT \ln[K_{ass}/(2a)^3]$ .

Вычисляем:  $K_{ass}/(2a)^3 = (10^8 \text{ литр/моль})/1 \text{ \AA}^3 = (10^8 \times 10^3 \text{ см}^3/\text{моль})/(10^{-8} \text{ см})^3 = [10^{11}/(0,6 \times 10^{24})]/10^{-24} = 1,7 \times 10^{11}$ . При  $kT = 0,6 \text{ ккал/моль}$ , имеем  $\Delta G_0 \approx -15,5 \text{ ккал/моль}$ .

### Задача 24.2

Простейшая модель кинетики энзиматической реакции (модель Михаэлиса-Ментен) трактует эту реакцию как обратимое образование интермедиата ES из свободного субстрата S и свободного фермента E, сопровождаемое необратимым превращением находящегося в составе ES субстрата в продукт P. При этом фермент-субстратный комплекс ES образуется из свободного субстрата и свободного фермента скоростью  $[E][S]k_1$  и распадается на свободный субстрат и свободный фермент со скоростью  $[ES]k_2$ , а на свободный продукт и свободный фермент со скоростью  $[ES]k_3$  (см. схему).



Определите зависимость скорости образования продукта P от полного содержания субстрата ( $[S]_T \equiv [S] + [ES]$ ) и фермента ( $[E]_T \equiv [E] + [ES]$ ) в растворе. (Считая, что субстрата в растворе много больше, чем фермента, такой расчет можно делать в стационарном приближении, согласно которому скорость изменения концентрации интермедиата,  $d[ES]/dt$ , столь мала по сравнению с  $d[S]/dt$  (или  $d[P]/dt$ ), что концентрация интермедиата  $[ES]$  может считаться неизменной.)

### Решение

В стационарном приближении  $d[ES]/dt = [E][S]k_1 - [ES]k_2 - [ES]k_3 = 0$ . Следовательно,

$$[ES] = [E][S]/K_M, \quad (24.2.1)$$

где величина



$$K_M = (k_2 + k_3)/k_1 \quad (24.2.2)$$

называется *константой Михаэлиса*.

Так как субстрата в растворе много, а фермента мало, то концентрация свободного субстрата  $[S]$  близка к полной концентрации субстрата  $[S]_T$  (а  $[ES] \ll [S]_T$ ); но концентрация свободного фермента  $[E]$  может быть существенно меньше, чем его полная концентрация  $[E]_T = [E] + [ES]$ . Подставив  $[E] = [E]_T - [ES]$  в (24.2.1), получим

$$K_M [ES] = ([E]_T - [ES])[S]_T,$$

и, далее,

$$[ES] = [E]_T [S]_T / ([S]_T + K_M).$$

Обратите внимание, что в образование фермент-субстратного комплекса вовлечено меньше половины фермента при  $[S] < K_M$  и больше половины при  $[S] > K_M$ .

Наконец — скорость образования продукта:

$$d[P]/dt = [ES]k_3 = k_3[E]_T[S]_T / ([S]_T + K_M). \quad (24.2.3)$$

### Задача 24.3

Какова связь между полученной при решении предыдущей задачи константой Михаэлиса  $K_M = (k_2 + k_3)/k_1$  и константой диссоциации  $K_{diss}$  комплекса  $ES$  на  $E$  и  $S$ ?

#### Решение

Константа диссоциации комплекса  $ES$  на  $E$  и  $S$  есть  $K_{diss} = [E][S]/[ES]_{eq}$ , где  $[ES]_{eq}$  есть та концентрация комплекса  $ES$ , что, не уменьшаясь и не увеличиваясь, находится в равновесии с заданными концентрациями  $([E], [S])$  субстрата и фермента в отсутствие превращения субстрата в продукт (т. е. при  $k_3 = 0$ ). Условие такого равновесия можно представить в виде  $d[ES]/dt = [E][S]k_1 - [ES]_{eq}k_2 = 0$ . Следовательно,

$$k_2/k_1 = [E][S]/[ES]_{eq} \equiv K_{diss}. \quad (24.3.1)$$

В то же время, константу Михаэлиса можно представить в виде

$$K_M = (k_2 + k_3)/k_1 = (k_2/k_1)(1 + k_3/k_2) = K_{diss}(1 + k_3/k_2), \quad (24.3.2)$$

где член  $k_3/k_2$  показывает, в какой пропорции фермент-субстратный комплекс распадается с выделением продукта и с выделением субстрата. При  $k_3 \ll k_2$ ,  $K_M \approx K_{diss}$ .

### Задача 24.4

Предположим, что реакция самопроизвольного (некатализируемого) превращения субстрата в продукт происходит по механизму  $S \leftrightarrow S^* \rightarrow P$  (где  $S^*$  — переходное состояние субстрата), с активационной свободной энергией  $\Delta F_0^\# \equiv F(S^*) - F(S)$ , а реакция превращения субстрата в продукт на ферменте происходит по тому же механизму  $ES \leftrightarrow ES^* \rightarrow EP$ , но с более низкой (из-за взаимодействия переходного состояния субстрата с ферментом) активационной свободной энергией  $\Delta F_{cat}^\# \equiv F(ES^*) - F(ES) = [F(S^*) - F(S)] - \delta$ .

(а) Как соотносится каталитическая константа скорости  $k_3$  (см. задачу 24.2) к скорости  $k_0$  самопроизвольного превращения субстрата в продукт?

(б) Как соотносятся скорости катализируемого ферментом (см. решение задачи 24.2) и самопроизвольного превращения субстрата в продукт?

#### Решение

(а) Так как скорость преодоления активационного барьера можно представить в виде  $k = k_{0,el} \cdot \exp(-\Delta F^\# / k_B T)$ , где  $k_{0,el} = k_B T / h$  — частота тепловых колебаний при температуре  $T$ , а  $\Delta F^\#$  — активационная свободная энергия, то понижение величины  $\Delta F_{cat}^\#$  на  $\delta$  по сравнению с  $\Delta F_0^\#$  увеличивает  $k_3$  в  $\exp[-(-\delta)/k_B T]$  раз по сравнению с  $k_0$ :

$$k_3 = k_0 \cdot \exp(\delta / k_B T).$$

(б) Так как скорость катализируемого ферментом превращения субстрата в продукт есть  $k_3[E]_T[S]_T / ([S]_T + K_M)$ , а скорость самопроизвольного превращения субстрата в продукт есть  $k_0[S]_T$ , то увеличение  $k_3$  в  $\exp(\delta/k_B T)$  раз по сравнению с  $k_0$  делает скорость катализируемого ферментом процесса в  $\exp(\delta/k_B T) \{ [E]_T / ([S]_T + K_M) \}$  раз больше скорости самопроизвольного превращения.

## К лекции 25

### Задача 25.1

В нативной белковой глобуле некая частица находится в одномерной потенциальной яме, создаваемой взаимодействиями с отдаленными участками цепи. При отклонении  $x$  от центра ямы энергия частицы  $E = \gamma x^2/2$ , и частица флуктуирует в яме со средней амплитудой  $\langle x^2 \rangle^{1/2} = 0,25 \text{ \AA}$ . При этом среднее время пересечения ямы частицей составляет порядка  $10^{-9}$  с.

Сколько, по порядку величины, времени надо ждать, пока эта частица отклонится от центра ямы (а) на  $1,0 \text{ \AA}$ ? (б) на  $2,0 \text{ \AA}$ ? (в) на  $3,0 \text{ \AA}$ ? Учтите, что временное полное разворачивание нативного белка происходит примерно раз в сутки (см. задачу 20.3).

### Решение

(а) примерно  $10^{-6}$  секунды; (б) примерно полсутки; (в) примерно одни сутки.

#### Объяснение.

Отклонению частицы в нативном белке на  $\langle x^2 \rangle^{1/2}$  от центра ямы с потенциалом  $E = \gamma x^2/2$  соответствует потенциальная энергия, равная  $kT/2$  (см. задачу 8.5). Значит, отклонению на  $0,25 \text{ \AA}$  от центра этой ямы соответствует энергия  $kT/2$ , отклонению на  $1,0 \text{ \AA}$  от центра ямы соответствует энергия  $8kT$ , отклонению на  $2,0 \text{ \AA}$  соответствует энергия  $32kT$ , а отклонению на  $3,0 \text{ \AA}$  соответствует энергия  $72kT$ . Время ожидания отклонения с энергией  $E$  составляет  $\sim \tau \times \exp(-E/kT)$ . При  $E = kT/2$  оно составляет  $\sim \tau \times \exp(-1/2) \sim 10^{-9}$  с, а при произвольном  $E$  —  $\sim \tau \times \exp(-E/kT) \sim (10^{-9} \text{ с}) \times \exp(\frac{1}{2} - E/kT)$ . Следовательно, оно составляет  $\sim 2 \times 10^{-6}$  с для отклонения в  $1,0 \text{ \AA}$ ,  $\sim 5 \times 10^4$  с  $\approx 13$  часов для отклонения в  $2,0 \text{ \AA}$ , и (формально)  $\sim 10^{22}$  с  $\sim 10^{14}$  лет для отклонения в  $3,0 \text{ \AA}$ . Однако, на самом деле, отклонения в  $3,0 \text{ \AA}$  надо ждать не  $10^{14}$  лет, а всего сутки, так как полное разворачивание нативного белка происходит примерно раз в сутки, и при этом потенциальная яма, создаваемая взаимодействиями с отдаленными участками цепи, исчезает.

### Задача 25.2

Константа диссоциации АТФ в воде  $K_{diss} = \frac{[\text{АДФ}][\text{Ф}]}{[\text{АТФ}]_{eq}} = 5 \cdot 10^5$  моль/литр,

где значком *eq* помечена та концентрация АТФ, при которой он, при данных концентрациях АДФ и фосфата Ф, находится в равновесии с ними.

(а) Какова свободная энергия гидролиза одной молекулы АТФ при физиологических концентрациях  $[\text{АТФ}]_{\text{ф}} \approx 10^{-3}$  моль/литр,  $[\text{АДФ}]_{\text{ф}} \approx 10^{-5}$  моль/литр,  $[\text{Ф}]_{\text{ф}} \approx 10^{-3}$  моль/литр?

(б) При какой концентрации АТФ (и при всё тех же  $[\text{АДФ}]_{\text{ф}}$  и  $[\text{Ф}]_{\text{ф}}$ ) распад АТФ сменится его синтезом?

### Решение

(а) При физиологических концентрациях

$$\frac{[\text{АДФ}]_{\text{ф}}[\text{Ф}]_{\text{ф}}}{[\text{АТФ}]_{\text{ф}}} = \frac{[10^{-5} \quad / \quad ] [10^{-3} \quad / \quad ]}{[10^{-3} \quad / \quad ]} \approx 10^{-5} \text{ моль/литр.}$$

Значит,  $[\text{АТФ}]_{\text{ф}} = [\text{АДФ}]_{\text{ф}} [\text{Ф}]_{\text{ф}} / [10^{-5} \text{ моль/литр}]$  при физиологических условиях, в то время как в равновесии было бы  $[\text{АТФ}]_{eq} = [\text{АДФ}]_{\text{ф}} [\text{Ф}]_{\text{ф}} / [5 \cdot 10^5 \text{ моль/литр}]$ , то есть концентрация молекул АТФ при физиологических условиях в  $5 \cdot 10^5 / 10^{-5} = 5 \cdot 10^{10}$  раз больше, чем было бы при равно-

весии АТФ с физиологическими концентрациями АДФ и Ф. Такому превышению соответствует избыточная свободная энергия, равная  $kT \cdot \ln(5 \cdot 10^{10}) \approx \approx 25 kT \approx 15$  ккал/моль.

(б) Распад АТФ сменится его синтезом, если (при всё тех же  $[\text{АДФ}]_{\text{ф}}$  и  $[\text{Ф}]_{\text{ф}}$ ) концентрация АТФ будет в  $5 \cdot 10^{10}$  (см. выше) раз меньше, чем  $[\text{АТФ}]_{\text{ф}}$ , т. е. будет равна  $\approx 10^{-3}$  моль/литр /  $5 \cdot 10^{10} \approx 0,2 \cdot 10^{-13}$  моль/литр.

### Задача 25.3

Головка миозина расщепляет молекулу АТФ, и 50 % получаемой при этом свободной энергии (50 % от 15 ккал/моль) превращается в производимую мышцей работу. Каждое расщепление приводит к силовому удару, приводящему к сдвигу миозинового волокна относительно актинового на 5 нм.

(а) Оцените, какую силу развивает головка миозина.

(б) С какой скоростью сближаются концы мышцы при ее сокращении, если она имеет 30 000 саркомеров в длину, число одновременно работающих миозиновых головок составляет 1 % от их общего числа, а каждая головка щепит 20 молекул АТФ в секунду?

(в) Какую силу развивает эта мышца, если в ней находится  $10^{-6}$  моля миозиновых головок?

(г) Какую мощность развивает эта мышца?

(д) Как изменится скорость сокращения мышцы, его сила и мощность, если, при тех же прочих параметрах (в частности — числа миозиновых головок), она имеет не 30 000, а 10 000 саркомеров в длину?

### Решение

(а) Каждая головка миозина развивает силу, равную

$$15 \text{ ккал/моль} \times 0,50 / 5 \text{ нм} = (4,2 \times 15 \times 10^3 \text{ Дж/моль} \times 0,50) / (5 \times 10^{-9} \text{ м}) \approx 6 \times 10^{12} \text{ Н/моль} = 10^{-11} \text{ Н.}$$

(б) Так как в каждый момент работает только 1 головка из 100, то за 1/20 секунды, т. е. за время сцепления одной молекулы АТФ в расчете на каждую головку миозина, успевает сработать не одна, а 100 головок в том же контакте актиновой и миозиновой нитей; это приводит к взаимному сдвигу этих нитей на  $5 \text{ нм} \times 100 = 500 \text{ нм}$ . Поскольку вдоль саркомера имеется два контакта актиновой и миозиновой нитей (справа и слева от его центра, см. рис. 25-8а), то саркомер сокращается за 1/20 секунды на вдвое большую длину, 1000 нм, а 30 000 друг за другом идущих саркомеров — на  $30\,000\,000 \text{ нм} = 0,03 \text{ м}$  (за те же 1/20 секунды). Значит, скорость сокращения такой мышцы — 0,6 м/с.

(в) За 1/20 секунды,  $10^{-6}$  моля миозиновых головок расщепляют  $10^{-6}$  моля АТФ (50 % свободной энергии распада которого превращается в работу), и при этом мышца сокращается на 0,03 м. Значит,

сила =  $(15 \text{ ккал/моль} \times 0,50 \times 10^{-6} \text{ моля}) / (0,03 \text{ м}) = 0,25 \text{ кал/м} \approx 1 \text{ Дж/м} = 1 \text{ Н}$ .

(г) Мощность = сила  $\times$  скорость, т. е.  $(1 \text{ Н}) \times (0,6 \text{ м/с}) = 0,6 \text{ Вт}$ .

(д) Мощность (= энергия/время) не изменится, скорость сокращения упадет втрое, сила утроится.

## ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

### А

абзимы 364–366  
агрегация белков 266, 268, 313  
АДФ 385–390, 480–481  
актин 384–389  
активность белка 15, 81, 162, 197, 238, 265, 273–274, 343–368, 375–376  
активные центры 22–24, 81–82, 86, 137–138, 168, 173, 179, 187–188, 227, 317, 320, 341–344, 357–360, 364–368, 370–381  
аланин 38–40, 88–89, 122, 132, 135–136, 141, 288, 347, 368, 405–406  
аллостерические взаимодействия 353, 381  
альбебетин 348–350  
амилоид 147–149, 269, 308  
аминокислотные остатки 15, 18, 23–26, 36, 38, 40–41, 47, 64, 66–67, 88–89, 117, 122–125, 131–135, 140, 146, 199, 201, 211–218, 228, 230, 233, 249–253, 274, 294, 301–303, 312, 315, 318, 320–333, 343–345, 368, 373–377, 436–439, 461, 467, 470–473, 476  
аминокислоты L и/или D-форма 25–26, 90–92, 449  
антигенные детерминанты 354–357  
антитела *см.* иммуноглобулины  
архитектура *см.* структурный каркас  
археобактерии 197  
АТФ 14, 152–153, 268, 368–369, 378, 384–390, 480–482

### Б

бактериородопсин 151–152  
Банк Белковых Структур = Protein Data Bank = PDB 16, 335–336  
банк первичных структур, SwissProt 16  
белки-рецепторы 152

белки со скудной вторичной структурой 194, 200  
белковая глобула 21, 59, 68–69, 72, 82, 87, 107, 137, 163, 165–167, 171, 182, 187, 192–193, 195, 200, 202, 205, 207, 209, 214, 251, 259, 261, 271, 273, 289, 293, 298, 313, 324, 329, 343, 357, 373, 378, 381 *см.* *также* водорастворимые... инженерия 81–83, 287, 342–244, 368  
статистика, квази-Больцмановская 207, 213, 262, 324  
структура 16–17, 67–70, 134, 162, 165–167, 192, 195, 197–199, 202–203, 205, 207–208, 214–216, 219, 220, 222, 228–229, 263, 267, 296, 302–303, 314–315, 323–324, 330, 335–337, 339, 341–342, 346, 354, 378  
белковый дизайн 189, 222, 227, 271, 339, 342, 346  
бесклеточные системы, сворачивание белков 264–265  
биоинформатика 341  
биосинтез *см.* синтез белков биохимический  
боковые группы аминокислот 15, 132, 215  
Больцмана формула 49–50, 100, 154, 220, 234,  
Бора эффект 383

### В

Ван-дер-Ваальсовы взаимодействия 34–35, 43–46, 242, 366, 398, 403–404  
Ван-дер-Ваальсовый радиус атома 28, 35, 43  
Вант-Гоффа критерий для денатурации

белка 233, 235, 257  
валентные  
углы 28–29, 38  
связи 28–30, 44, 131, 158,  
394–395  
вода 12, 16, 18, 21, 38, 42–43, 45–49,  
52–54, 58–86, 96, 108, 114–115,  
118–119, 125, 136–137, 148, 151,  
154, 163, 199–201, 208–209,  
213–215, 217, 224, 231–232, 235–  
238, 240, 252, 254, 258–259, 270,  
273, 275, 286–289, 296, 299, 304,  
324, 359, 361–363, 367–368, 374,  
378–379, 389, 394–395, 397, 407,  
409–410, 416–418, 421, 432–433,  
440, 444, 450, 457–459, 461, 468,  
477, 480  
водорастворимые глобулярные белки  
16–17, 19, 156, 162–163, 179,  
200–201–202, 208, 211, 213, 270,  
276, 314, 351, 375, 447  
водородные связи 16, 38, 42–48, 52–53,  
61–64, 79–80, 84–86, 88–93, 117–  
118, 121–123, 135–137, 144–145,  
148, 151–153, 163, 167–170, 173,  
176–177, 186, 199, 201, 203, 208–  
209, 217, 232, 240, 303, 326, 329,  
354, 357, 369, 440, 457–461, 470  
и гидрофобный эффект 63–68  
направленность 90  
энергия 44  
вращение вокруг ковалентных связей  
26–27, 29, 30–32, 37, 394–395 *см.*  
*также* угол вращения  
время образования структуры 128–129,  
438  
вторичные структуры 17–18, 20, 48, 87,  
89, 92–97, 110, 117, 123, 135–137,  
140, 147, 151, 163–167, 188–190,  
193–194, 199–205, 211–213, 215,  
223, 238, 240–242, 252, 254, 256,  
266, 272–276, 279, 312, 317, 320,  
324–334, 338–340, 346–349, 375,  
377, 445, 447, 457, 462  
предсказание 328–334, 339

выравнивание последовательностей  
317–321, 473  
**Г**  
Гейзенберга принцип неопределенно-  
сти 27, 34, 160, 396–398, 429, 431  
гемоглобин 15, 18, 166, 195–196,  
381–390  
«Геном человека» проект 341  
гетерополимер случайный 226, 243,  
260–261, 295, 451, 462  
гидрация *см.* вода  
гидролиз 198, 359–362, 371–372,  
386–387, 480  
гидрофобная поверхность 64, 137–138,  
169, 211–212, 268, 327, 329, 355  
гидрофобное ядро *см.* ядро белковой  
глобулы  
гидрофобность 64–66, 137, 142,  
214–215, 254, 330, 372  
гистерезис 108  
глицин 26, 38–41, 88–90, 122, 131–136,  
141, 144–146, 179, 288, 405–406  
главная цепь белка 15, 17, 21, 25–26,  
32, 38–40, 48, 87–94, 123, 134–137,  
143, 163–164, 177, 201, 208–209,  
215, 238, 251–254, 298, 319, 322,  
326, 374, 377, 445, 457  
глобула  
нативная 217, 220, 253, 274,  
289, 295, 299, 301–303, 305, 310,  
466–467, 479  
пред-расплавленная 256–258,  
272, 274  
расплавленная 239–242,  
254–258, 265, 268–269, 272–275,  
277, 279, 283–287, 291, 309, 330,  
346–347, 349, 359  
глобулярные белки 16–17, 19, 32, 87,  
135, 140, 144–145, 147, 162–163,  
179–181, 191, 198–203, 208,  
211–213, 223–226, 228, 231, 263,  
270, 276–277, 312, 314, 328, 340,  
349, 351, 355, 375, 447  
глутатион 85  
гомология 195, 227, 316–322, 335, 342,

348, 473  
«греческий ключ» 170, 176, 204  
ГТФ 153, 377  
**Д**  
«двойного сита» эффект 371–372  
двугранный угол *см.* угол вращения  
дегидрогеназа 380–381  
денатурация белка 221, 226, 229–233,  
235–239, 241–243, 245, 249,  
250–251, 253–255, 294–304, 461  
ее обратимость 233  
денатурированный белок 217–218, 232,  
236–238, 254, 266, 269–270, 273,  
285, 304, 310, 456, 467  
детального равновесия принцип 299  
«дефекты» структурные 179, 200, 202,  
205  
диссоциации константа 80, 412, 458,  
460–461, 478, 480  
дисульфидные связи *см.* S-S связи  
дисульфид-изомеразы 269  
диффузия 110, 114–115, 390, 408,  
432–433  
диэлектрическая проницаемость  
45, 69, 71–82, 84, 154, 404, 413,  
415–416, 418, 421–425, 440–444  
ДНК-связывание 348, 352–355,  
375–376, 378  
домены 18–19, 150, 163, 165–172,  
180–181, 186–190, 193–194,  
196–198, 200, 202, 208, 211, 213,  
225, 228, 231, 264, 267–270, 295,  
303, 305–312, 314, 327, 332–335,  
352, 355–358, 379–381, 385–387,  
389, 435–436, 446–448  
доступная поверхность 65–67, 304

**Е**  
Ефимова структурное дерево 195  
**З**  
«зародыш» *см.* ядро сворачивания  
заряд аминокислотного остатка  
138–139

**И**  
иминокислота 40, 122, 134  
иммуноглобулины 171, 174, 197,  
354–357, 476–477  
иммуноглобулиновая укладка цепи 174  
ингибитор 272, 307, 315, 322, 330–331,  
353, 370, 372  
индуцированное соответствие 378–379  
инициация  $\alpha$ -спиралей 118, 120,  
122–125, 128–129, 434  
интермедиаты сворачивания 239,  
273–275, 287, 295, 301, 308, 313,  
465, 470  
интрон-экзонная структура 197  
инфракрасная спектроскопия белков 28

**К**  
калориметрия 231, 233–235, 257  
катализ 14, 344, 360–367, 374, 376–377  
квант тепловой 28  
«квазисферических многогранников»  
модель 180–182, 185  
кинетика 97, 110, 113, 117, 124–126,  
231, 251, 271, 273  
классификация белков 165–166, 191,  
193–195  
классификация белковых структур  
CATH 193, 375  
Dali/FSSP 193  
SCOP 193, 448  
классы (структурные) белков 16, 140,  
188–189, 193–194, 354, 375–376  
клатраты 63  
клубок 95–96, 117–121, 125–126,  
128–131, 136, 237–238, 241–243,  
246, 248–249, 250, 254–259, 266,  
273–274, 286–287, 294–296,  
298–302, 304–310, 326  
ключ-замок 369, 379  
колебания 28–29, 32, 34, 46, 96, 109,  
160, 281, 367–368, 377, 386,  
395–399, 403–404, 408, 431  
коллаген 144–147  
конформационные превращения  
104–105, 110, 268,



конформация 17, 30, 36–40, 87–90, 92–93, 117, 130, 135–136, 144–145, 150, 152–153, 161, 166–167, 175, 206–207, 222–223, 252, 259–260, 265, 270–271, 322–323, 326–327, 336, 338, 345, 364, 370, 388–389  
*см. также* структура белка  
и болезни 149, 308–313  
и изменения функций  
белка 376, 378–379  
кооперативные переходы 97, 121–122, 230, 232, 241, 256 *см. также* де- и ренатурация  
координационные связи 84–86, 137, 374, 382  
корпускулярность 71, 79–81  
котрансляционное сворачивание 265, 269  
кофакторы 15–16, 21, 152, 156, 178, 353, 375–376, 380–381  
кристаллизация 68, 289, 306, 472  
кристаллы 15–16, 22, 36, 44, 63, 151, 162–163, 182, 231, 242, 249, 250–251, 261, 267, 307, 378, 404, 407–408,  
круговой дихроизм (КД) 95, 121, 230, 238–239, 241, 256, 266, 328, 348, 350

## Л

Ландау теорема 118–120  
Левингала парадокс 270, 295, 297, 304, 309–310  
лед, структура 42–46, 48, 63, 119–120  
лейциновый zipper 355

## М

матриксные белки 141, 147  
меандр 169–170, 175–176, 204, 446  
мембранный канал 154–155, 444  
мембранные белки 16, 19, 23, 140, 150–151, 155, 177, 179, 185, 201, 228, 263, 270, 276–277, 340, 350–351, 440  
металлы *см.* координационные связи  
миозин 16, 141, 384–389, 481

многозаходный пропеллер 172  
модель квазисферических многогранников для  $\alpha$ -белка 180–182  
модель случайных энергий 450–451, 462, 464–465  
«модули» 197–198  
мотив «спираль-изгиб-спираль» и связывание ДНК 352, 354  
мотивы укладки цепи 167, 170–171, 173, 178, 190, 194, 197–199, 208, 222, 224, 227, 337, 340, 357, 445–446, 448  
мутации 81–82, 146–147, 196, 216, 226–227, 277, 287–292, 312–313, 342–350, 356, 360, 373, 467–468  
мышцы 14, 379, 381, 383–388, 390, 481–482

## Н

набор стандартных мотивов укладки 193, 208,  
нативно-развернутые белки 229–230, 359  
надежность работы белка 22, 253, 259, 262  
нуклеационный механизм 289–290

## О

«ОБ-укладка» 190, 194  
образование структуры белка  
in vivo 12–13, 18–20, 146, 199, 263, 265–268, 270  
in vitro *см.* самоорганизация  
объем глобулы 273  
объем свободный 247, 252–253  
одностадийное сворачивание белков 279, 286  
отбор 102, 196, 202–203, 216, 220–222, 228, 261–262, 311–314, 347, 372, 376

## П

Паули запрет 33  
пепсин 172  
пептидная связь 26, 31–32, 374

пептидные группы 27, 31, 47–48, 92–96, 137, 163, 199, 208–209, 360  
перевитые спирали 141–142, 439  
переход глобула-клубок 242–243, 245, 249–250, 255, 257, 301  
переход спираль-клубок 120–122, 124–125, 232, 242  
переходное состояние 110–111, 124, 126–128, 275, 280–281, 283, 285–288, 290–292, 297, 304–306, 362–369, 376, 467, 479  
переходы «все-или-ничего» 107–108, 110, 231–235, 241–242, 248, 250–251, 253, 255, 256–259, 275, 279, 283, 287, 289–291, 294–296, 310–312, 455, 467–469  
пермутация 349  
петли 120, 125, 131, 134–135, 151, 156, 162–176, 178, 180, 182, 187–188, 199, 201–203, 205, 212–213, 221, 226, 240–241, 300–305, 321, 325, 327, 329, 331, 338, 343, 347, 349, 356, 377, 449, 475  
плотная упаковка 92, 144, 164, 167, 177, 182–186, 217, 252–253, 261, 312, 330, 342, 347–348  
плотность 11, 42, 57, 101, 104–105, 109, 114, 130–131, 159–160, 237–238, 243–249, 252–259, 283, 367–368, 404, 432, 438, 452  
поверхностный эффект 301–302  
поворотная изомеризация боковых групп 252–253  
полимер 15, 117, 119, 206, 242, 251, 255–257, 260–261, 314  
полиморфные переходы 307  
полипептид 26, 88, 96–97, 117, 121–124, 129, 140, 144, 156, 177, 227–228, 232, 328, 350, 369, 457  
полипролиновая спираль 92, 144  
порин 153–154, 277  
последовательности  
аминокислотные 16–20, 82, 137, 166–167, 187, 201–203, 206, 208–209, 213, 216, 218–219, 222–223, 233, 261, 269, 271, 294–295,

303, 307, 315–327, 331, 335–347, 373 *см. также* структура белка  
случайные 202, 211, 212–213, 218–219, 224–227, 313  
периодические 212

посттрансляционные модификации 19, 147

потенциал

вращения 30–32  
средней силы 59–60

потенциометрическое титрование 455  
право- и левовинтовой ход перемычек 190, 206, 222

предсказание пространственных структур 316

принцип множественности 208–211, 314

прион 147–148, 262, 307–308

проводимость протонная 152

прокариот белки 197, 269

пролин 31–32, 36, 40, 92, 131–132, 135, 145, 266, 269, 275, 450

промежуточное состояние *см.* интермедиаты сворачивания

проникновение растворителя в глобулу 240–241, 252

«пропеллер» 167, 171–172, 174

протеолиз 19, 238, 349 *см. также* гидролиз

«Протеомика» проект 342

протягивания метод 335–339

профили первичной структуры 321, 473

путь сворачивания 271, 274–275, 285, 293–309, 313, 324

пучки  $\alpha$ -спиралей 152, 177–185

## Р

равновесие термодинамическое 128, 267, 298, 300–302, 307, 464, 468–469

разворачивание белков 231, 239, 285, 288, 294, 310, 458, 461–462, 468, 470, 479–480 *см. также* денатурация

Рамачандрана карты разрешенных состояний 37–40, 88, 136, 167  
ренатурация глобулярных белков *см.* самоорганизация  
рентгеноструктурный анализ 45, 82, 148, 151, 344, 385  
репрессор 17–18, 353–355, 378  
рецептор 152–153, 355  
рибосома и синтез белка 197, 199, 263–265, 267, 269–271  
родственность *см.* гомология  
Россмана укладка 186–188, 225, 358, 375, 381, 448–449  
ротамер 39, 252, 406

### С

самоорганизация  
белков 18–20, 193, 199, 230, 263–267, 269–270, 272, 274, 277, 303, 307, 311  
стадийная 272, 310  
вирусов 277  
саркомер 385, 481  
свободная энергия 48–49, 51–52, 54–62, 64–65, 67–71, 79–80, 82, 84, 97, 103–108, 111, 113, 117–121, 123, 125–129, 137, 139, 154–155, 205–206, 214–219, 222, 224, 226, 236–237, 244, 246–247, 253, 259–260, 271, 275, 280–281, 283, 289, 291–292, 294–295, 297–298, 300, 303, 305–306, 311, 324, 327, 363–364, 371, 387, 408–409, 411–413, 415, 417, 421, 427–428, 430, 436–438, 454–455, 457, 459–461, 463–463, 467, 471, 476–477, 479–482  
и гидрофобность 59, 62–68  
и сворачивание белка 300–308  
свободно-энергетический барьер 107, 110–112, 114, 116, 125, 251, 280, 283, 288, 297, 301, 364  
сворачивание белков *см.* самоорганизация  
связывание 82, 153, 171, 188, 196,

198, 230, 321, 351–354, 357, 359, 363–365, 368–370, 376, 378, 381–387, 433  
связывания константа 386, 476  
селективная проводимость мембранных белков 154–155  
селеноцистеин 134  
середина перехода 107–108, 120–121, 234, 238, 258, 282, 288 *см. также* равновесие...  
сериновые протеазы 22–23, 307, 357–360, 364, 369–370, 373–374  
серпин 262, 307,  
синтез белков  
биохимический 17, 19  
химический 267, 346  
скорость *см. также* время...  
перехода 111–113, 161, 280–281, 295  
приближения к равновесию 284–285, 304, 468–469  
разворачивания 282, 284–285, 288–289, 468–470  
сворачивания 269, 275, 279, 282, 284–292, 297, 299, 301–303, 306, 313, 468–470  
слоевая структура 200, 224  
«словесный портрет» белка 166  
солевая связь 80  
случайные последовательности 202–203, 211–213, 218–219, 224–227, 261, 313  
специфичность ферментов 369, 372  
спирали 17, 87–90, 92–96, 117–129, 135–138, 141–147, 151–152, 156, 166, 173, 177–190, 205, 213, 225, 228, 277, 316, 323–334, 337, 345, 350, 353–355, 383, 390, 434, 436, 449, 470  
без водородных связей 92  
спираль как длинный диполь 90  
стабильность глобулярных белков и их сворачивание 199, 202, 205, 208–209, 216, 220–222, 288–289, 305, 306–308, 312, 440, 457

статистика Больцмана 207, 213–216, 450  
статистическая  
механика 11, 312, 336  
сумма 97, 102–103, 430  
физика 11, 97, 101, 103, 295, 432  
структура белка *см. также* конформация и последовательности  
вторичная 17–18, 20, 87–89, 92–97, 110, 117, 123, 135–137, 140, 147, 151, 163–167, 193, 203, 238, 272–276, 324, 330, 339, 457  
первичная 17, 122, 140–142, 197, 201–202, 211–213, 224–226, 303, 316, 321–323, 331, 338, 449–450  
третичная 17–18, 87, 95, 193, 273, 312, 325, 332  
упрощенные схемы  
165–166, 210  
четвертичная 17–18, 22  
«Структурная геномика» проект 342  
структурный каркас 163–164, 178, 208–209, 335  
субтилизин 22–23, 373–374, 379  
суперспираль 92, 141–146, 176, 224, 327–328, 358, 385, 446

### Т

твердость белков 22–24, 253, 359, 366  
температура 12, 18, 22, 28–29, 32–33, 49–51, 54–67, 69–70, 83–85, 97–110, 115, 161, 207, 226, 229–249, 253–262, 268, 281–285, 297, 306–309, 328–330, 348, 377–378, 395–396  
конформационная 214, 221, 262  
перехода 119–121, 453  
плавления 42, 145, 221, 260, 262, 329–330, 455–456, 463–464, 467  
стеклования 260–261  
теплоемкость 58–63, 104, 109, 231, 233, 236, 456, 463–464, 466

«тепловой квант» (кТ) 28  
термодинамика 58, 60, 64, 67, 97, 101, 105, 233, 259, 271, 274, 279, 280, 294, 298, 307, 428–429  
термолизин 181, 374  
тетрамер 146, 347, 381  
топология 165, 170–176, 186, 194, 254, 305, 334, 349–350, 446  
трипсин 172–173, 272, 322, 330–331, 357–358, 360, 369, 373–375, 379  
туннельный эффект 159

### У

угол вращения 26, 29–32, 89, 207, 215, 252, 326, 395  
упаковка 63, 92, 137, 144, 164–168, 177–186, 194, 200, 203, 210, 217, 224, 252, 261, 273–276, 279, 312, 330, 334, 340, 342, 347–348  
ортогональная 168–169, 172–173, 179–180, 190  
продольная 168–172  
компактная 181, 199, 208–209  
упругость полимера 205–206

### Ф

Ф-величины 288–291  
фаговый дисплей 228  
фазовая диаграмма белка 238, 255, 258  
фазового перехода отсутствие 110, 229  
фазовые переходы  
в статистической механике 107–110  
первого рода 107–108, 110, 118120, 125–126, 129, 231, 242–243, 248–251, 289, 307, 312  
и их нуклеация 289–290  
второго рода 109, 248, 251, 255  
ферментативный катализ 361–365  
ферменты 359–364, 366–372  
фибриллярные белки 16, 19, 88, 140–141, 144, 147, 149, 197–198, 201–202, 211, 213, 228, 263, 270, 340, 350, 375, 439  
фиброин 140–141, 145

формы стерические 25  
фотосинтетический реакционный  
  центр 155–158, 161,  
функции белков 14–15, 20–24, 167,  
  177–179, 196, 313–314, 321–323,  
  342, 344–345, 351–352, 359, 363,  
  371–381

## Х

химическая связь *см.* валентная связь  
химический потенциал 56–58, 246–248,  
  384, 410–412, 426–427, 454–455,  
  458–460  
химотрипсин 22, 172–173, 360, 369  
холодовая денатурация глобулярных  
  белков 235–237, 255, 259, 456–457  
хроматография 256–257

## Ц

цикл 386–387  
цинковый палец 348, 355  
цистеин *см.* S-S связи

## Ч

число  
  аминокислотных последователь-  
  ностей 208–209, 216, 219, 226  
  конформаций 206–207

## Ш

Шапероны 12, 264–265, 268–270, 278,  
  311, 313  
шеvronный график 285–288, 299,  
  467–468  
ширина перехода 121–122, 138,  
  231–235  
«штабель» *см.* структурный каркас

## Э

эволюция белка 195–196, 198, 227–228  
электростатические взаимодействия  
  44, 69–70, 78, 80, 83–84, 159, 404,  
  413, 415, 418, 465  
  и корпускулярность среды  
  79–81  
элонгация  $\alpha$ -спиралей 118, 120,

  123–124, 126  
энергетическая щель 259–262, 294–  
  296, 305, 307, 311–312, 323, 463  
энергетический  
  барьер 161, 283, 365, 378,  
  432  
  ландшафт 296, 308–310  
  спектр 101, 104–105,  
  259–262, 450–452, 463  
энергия 27–28, 30–32, 34, 36, 38, 40–41,  
  43, 45, 48–52, 78, 82–86, 98–110,  
  158–161, 167, 205–207, 213, 220–  
  221, 232–234, 236–237, 240, 244–  
  247, 249, 255, 259–261, 281–283,  
  287, 293–295, 299–302, 306–307,  
  310–313, 323, 365–368, 386, 388,  
  395–401, 403–405, 415–417, 419,  
  421–426, 428–433, 440–443,  
  450–453, 462–467, 475–476, 480  
  водородной связи 45, 48,  
  62, 79–80  
  свободная 49, 51–52,  
  54–62, 64–65, 67–71, 79–80, 82,  
  84, 97, 105–108, 111, 113, 117–121,  
  123–129, 131, 137, 139, 154–155,  
  214–219, 222, 224, 226, 236–237,  
  244, 246–247, 253, 259–260, 267,  
  271, 275, 280–281, 288–289, 291–  
  292, 294–295, 297–300, 303–306,  
  311, 324, 327, 330, 363–364, 371,  
  387, 408–409, 411–413, 415, 421,  
  427–428, 436–438, 444, 454–461,  
  467, 471, 476–477, 479–482  
энтальпийный катализ 362  
энтальпия 55, 60, 67, 242, 310, 362–363,  
  408–409, 421, 455–456 *см. также*  
  энергия  
энтропийный катализ 363  
энтропия 48–49, 51–56, 58–62, 64, 67–68,  
  86, 97–106, 109, 116, 118–121, 206,  
  218, 234, 236, 240, 242, 244, 246–  
  248, 253, 255, 259–260, 283, 295,  
  297–302, 309–310, 363, 408–409,  
  425–430, 432, 453–456, 463, 467  
эукариот белки 19, 197

## Я

ядро  
  белковой глобулы 59, 135,  
  170, 181  
  сворачивания 276–277,  
  288–289, 291, 303, 305, 313,  
  467–468, 471  
ЯМР-спектроскопия 13, 15–16, 40, 82,  
  94, 147–148, 162, 238–241, 266,  
  316, 342–343, 348

## а

$\alpha$  +  $\beta$ -белки 163, 185–186, 188–190,  
  209, 354  
 $\alpha/\beta$ -белки 163, 185–186, 187–190  
 $\alpha/\beta$ -бочонок *см.*  $\alpha/\beta$ -цилиндры  
 $\alpha/\beta$ -складка 188–190, 194  
 $\alpha/\beta$ -цилиндры 186–188, 194, 357, 375  
 $\alpha$ -белки 163, 177–178, 181, 303, 375  
 $\alpha$ -кератин 141  
 $\alpha$ -спирали 17, 21–22, 29, 87–90, 94–96,  
  117, 120–129, 135–138, 141–143,  
  151–152, 155, 163–165, 168,  
  177–183, 185–186, 188–190, 194,  
  199, 208–212, 224–226, 240, 252,  
  262, 272, 297, 302–303, 308, 322,  
  324–327, 329–331, 332–334, 338,  
  347, 352–355, 373, 380, 382–184,  
  386, 439–440, 470, 475

## β

$\beta$ -белки 163, 167, 172, 174–175, 177,  
  189–190, 195–196, 376  
 $\beta$ -бочонок *см.*  $\beta$ -цилиндр  
 $\beta$ -изгибы 93, 126, 135, 325, 327, 436,  
  475  
 $\beta$ -листы 87, 91–94, 125–129, 140–141,  
  152–153, 163–164, 167–176,  
  178–179, 185–190, 199, 201–202,  
  222, 224–225, 252, 262, 308, 356,  
  373, 381, 436–438, 440, 446, 449  
 $\beta$ -призмы 173–174, 195, 223  
 $\beta$ -структура 17–18, 22, 87, 89–96,  
  123–126, 128–129, 135–138, 141,  
  148–149, 152–153, 163, 167–168,  
  170–171, 174–175, 185–186, 188–

  190, 194–195, 198–200, 206, 210,  
  212, 242, 256, 308, 322, 325–331,  
  347, 354, 357, 375, 437–438  
  антипараллельная 90–92,  
  94, 167–168, 174–175, 188–189,  
  195  
  параллельная 90–92, 94,  
  148, 168, 171, 174–175, 186  
  смешанная 90–91, 175  
  скрученность 90, 92,  
  167–168, 174, 185, 446  
 $\beta$ -сэндвич 148–149, 170–172, 308  
 $\beta$ -тяж 90–93, 125–127, 138, 148, 153,  
  164–165, 167–174, 176–177,  
  185–186, 189–190, 194, 210, 212,  
  222–224, 322, 325, 329, 331, 333,  
  338, 380, 437, 446–447, 449, 475  
 $\beta$ -цилиндр 152–153, 168–170, 172,  
  186–187, 356–358, 381, 440, 446  
 $\beta$ -шпильки 93, 125–127, 151, 170–171,  
  175, 189, 196, 272, 302, 327, 247,  
  355, 436

## В

BLAST 317–320

## С

cis-конформация 31–32, 36–37

## G

G-белок 152–153, 377

## H

HMMer 317, 320

## P

PSI-BLAST 317, 320

## S

S-S-связи 20, 28, 84–85, 141, 146, 201,  
  266–269, 274–275, 285, 302

## T

trans-конформация 31–32, 36, 92