THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 264, No. 31, Issue of November 5, pp. 18472-18479, 1989
Printed in U.S.A.

Dependence of Formation of Small Disulfide Loops in Two-Cysteine Peptides on the Number and Types of Intervening Amino Acids*

(Received for publication, June 13, 1989)

Rumin Zhang and Grayson H. Snyder

From the Department of Biological Sciences, State University of New York, Buffalo, New York 14260

Microscopic disulfide-exchange rate constants have been measured for the formation and opening of small disulfide loops in reactions between glutathione and peptides containing 2 cysteines. Twelve cysteine-X_m-cysteine peptides have been studied, where X is an amino acid and m is the number of amino acids between the cysteines. Homopolymers of alanine for m equaling 0–5 are evaluated, as well as X_1 and X_2 series employing glycine, valine, or proline. Equilibrium constants K_c for loop closing are only slightly dependent on the nature of X. Loops with even values of m generally are favored relative to loops with odd values. K_c increases in the rank order X_1, X_3, X_0, X_5, X_4, and X_2. Formation of a disulfide between sequentially adjacent cysteines therefore is not especially difficult. The dependence of K_c on the odd-even nature of m is compared with similar patterns observed both in statistics of disulfide formation in naturally occurring proteins and in theoretical studies of peptide cyclization. The relative equilibrium populations of intramolecular disulfides in peptides containing cysteine-cysteine-cysteine and cysteine-serine-cysteine-serine-cysteine clusters are consistent with predictions based on the values of K_c in the two-cysteine peptides.

Quantitative information about factors influencing the formation of small disulfide loops is important for understanding naturally occurring proteins and for designing new engineered proteins. Examples of native proteins with small loops include acetylcholine receptor and protein disulfide isomerase. A -Cys-Cys- loop formed by 2 cysteines which are adjacent in the primary sequence affects agonist-induced conformational changes at the acetylcholine receptor's binding site (1). Two -Cys-X-X-Cys- loops in the isomerase form the active site of this redox enzyme (2). Proteins with short sequences often employ high densities of disulfides to generate a folded topology. An example is the double-headed trypsin inhibitor from a Chinese medical herb (3), having 41 amino acids including 6 cysteines.

Early systematic studies of air oxidation of peptides of the type Cys-Gly$_m$-Cys (4–6) at low ionic strength revealed no intramolecular loop formation for m = 0 or 1, 15 and 40% monomeric loops for m = 2 and 3, respectively, and 90% or more monomeric loops for m = 4–6. Reduction of monomer loops for the same series by reduced glutathione exhibited monotonically decreasing values for the overall macroscopic rate constant as m increases from 1 to 4 (7). However these overall rate constants are difficult to interpret since they were generated by data requiring a series of several consecutive chemical reactions in an enzyme-coupled spectrophotometric assay system. There have been no published data on details of the microscopic rate constants for loop formation as a function of the number and types of residues in the sequence between the paired cysteines.

Fig. 1 depicts several examples of disulfide exchange reactions pertinent to the data discussed below. *Reaction 1* illustrates attack of the completely reduced peptide (R) by oxidized glutathione to produce either of two species (SI/SII) containing a single peptide-glutathione mixed disulfide. Those species can go on to produce a monomeric loop (L) by intramolecular attack of their mixed disulfide by the free cysteine on the same peptide chain, as shown in *Reaction 2*. The two rate constants are given by k_c, describing loop closing, and k_o, describing loop opening. Alternatively, SI may react again with another GSSG to produce double mixed-disulfides (D), as in *Reaction 3*. It also is possible for one of the mixed disulfide forms, such as SI of *Reaction 4*, to convert directly to the other, SII, by an intramolecular rearrangement. Looped dimers (LD) may be formed by a variety of processes. One such process given by *Reactions 5* and *6* begins with reduction of L by the free sulfhydryl group on SI. The last process in Fig. 1 depicts intramolecular exchange in a 3-cysteine species (LR) containing one loop and 1 reduced cysteine.

We recently presented a detailed study of such reactions in a peptide containing a -Cys-Val-Cys region (8), in which 12 rate constants were determined by kinetic studies beginning with four different monomeric and dimeric forms of the peptide. Data presented below extend those studies to a series of 11 additional synthetic peptides. Here however we focus only on the rate constants k_c and k_o of *Reaction 2* of Fig. 1, determined by a single type of kinetic experiment using L as the starting reagent. Rate constants are evaluated as a function of m for -Cys-Ala$_m$-Cys sequences where m varies from 0 to 5. Rate constants are evaluated as a function of the type of residue between the cysteines in -X_1- and -X_2- series employing Ala, Gly, Val, and Pro variations. These four types of residues have very different effects on formation of reverse turns in polypeptide chains (9). Alanine is the prototype L-amino acid. Glycine provides additional conformations not available to L-amino acids, since its tiny side chain consists of a single hydrogen atom. In contrast, the side chain of valine is branched at the β position adjacent to the backbone, leading to constraints on allowed backbone dihedral angles. Proline provides unique effects by virtue of attachment of its side chain to both the α carbon and the nitrogen atoms of the backbone.

Experiments described below also consider intramolecular disulfide exchange in peptides containing the three-cysteine

* This work was supported by National Institutes of Health Grant GM26715 and by the Center for Applied Molecular Biology and Immunology at the State University of New York at Buffalo. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

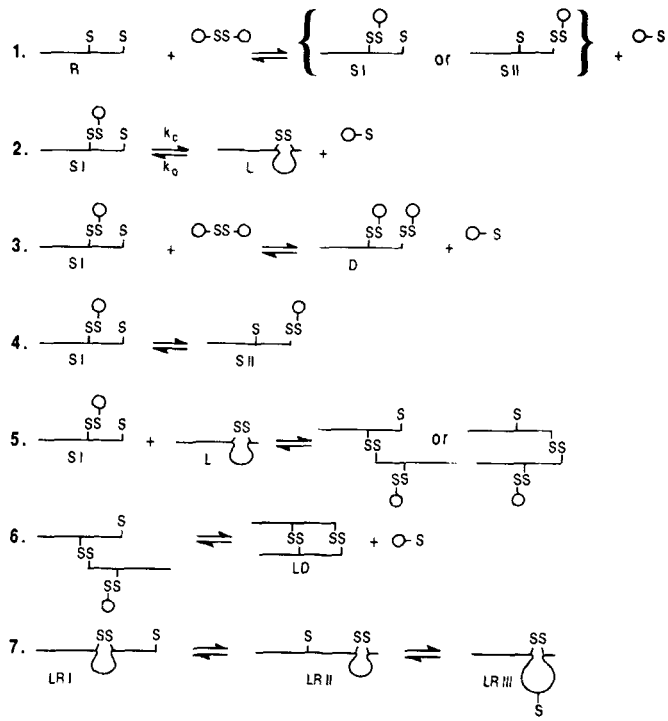

FIG. 1. **Examples of disulfide exchange reactions in mixtures of glutathione and peptides containing 2 or 3 cysteines.** Oxidized and reduced glutathione: *O-SS-O* and *O-S*, respectively.

clusters -Cys-Cys-Cys- or -Cys-Ser-Cys-Ser-Cys-. The observed relative equilibrium populations of the three possible intramolecular disulfide loops in each species are compared with predictions based on the equilibrium constants for loop formation measured in the 2-cysteine peptides.

MATERIALS AND METHODS[1]

Iodoacetate, GSH, and GSSG were obtained from Sigma. The two forms of glutathione were further purified by reverse phase HPLC.[2] Synthetic peptides were prepared on a Biosearch model 9500 peptide synthesizer using standard Merrifield chemistry. Crude reduced peptides were purified by preparative HPLC. Their amino acid composition was confirmed by analysis of acid hydrolysates. Their fully reduced nature was confirmed by assaying peptide concentration by UV absorption of their tyrosines (10) and by assaying sulfhydryl concentration with Ellman's reagent (11). Oxidized forms of the peptides were obtained by air-oxidation of sulfhydryl groups in the presence of Cu^{2+}. The monomeric loop form then was separated from oligomer loops by preparative HPLC.

All short synthetic peptides begin with the sequence Tyr-Ser-Arg-. The purpose of the tyrosine is to provide a means of calculating peptide concentrations by integrating HPLC peak areas monitored at 280 nm. The purpose of serine and arginine is to enhance solubility of the peptides. This leader then is followed by the first cysteine of the potential loop-forming region. Twelve such peptides ending in a carboxy-terminal cysteine were studied. One was -Cys-Cys. The others, represented by the amino acids appearing between the two cysteines, were Ala_{1-5},[3] Val_{1or2}, Gly_{1or2}, and Pro_{1or2}. In order to evaluate the effect of the negative backbone carboxyl group, two additional peptides were synthesized. These species begin with the same leader as above, followed by -Cys-Cys-Ala or -Cys-Val-Cys-Ala. Addition of alanine at the carboxyl terminus displaces the ionized carboxyl group 1 residue away from cysteine, eliminating any through bond inductive effect on disulfide exchange. In the 3 M guanidine hydrochloride high salt buffers employed below, additional electrostatic effects are negligible. Two additional short peptides having 3 cysteines were synthesized. These both begin with the leader Tyr-Arg- and are followed by the carboxyl-terminal sequences -Cys-Cys-Cys-amide or -Cys-Ser-Cys-Ser-Cys-amide.

Identities of HPLC peaks were determined by *equilibrium* experiments. Reduced peptide was combined with an excess of reduced and oxidized forms of glutathione in degassed buffers under an argon atmosphere. Buffers contained 3 M guanidine, 30 mM phosphate, 30 mM Tris, and 2 mM EDTA. The purpose of the guanidine is to prevent peptide aggregation and prevent formation of hydrogen bonds and ion pairs. Observation of two identical chromatograms at different long times of incubation confirmed that equilibrium was reached. Ellman's reagent indicated that there was no net loss of sulfhydryl groups by any residual oxygen in the solutions. Monomeric and oligomeric species can be distinguished by varying total peptide chain concentration while holding the GSH/GSSG ratio constant. Decreased collisions between chains give diminished amounts of oligomers at lower peptide chain concentrations. Reduced and oxidized species then can be distinguished by holding peptide chain concentration constant while varying the GSH/GSSG ratio. Final peak assignments may be confirmed quantitatively by verifying that ratios of equilibrium concentrations which should correspond to equilibrium constants are in fact constant under different experimental conditions. Details of these methods were presented previously for Val_1 (8).

A typical *kinetics* experiment begins by mixing 0.1 mM fully oxidized loop monomer with 5 mM reduced GSH at pH 7.0 in the same buffers employed above. At different time points, aliquots are removed and trifluoroacetic acid is added to lower the pH. Disulfide exchange requires the nucleophilic deprotonated thiolate $-S^-$ form of cysteines. Protonation of sulfhydryl groups at low pH is capable of slowing disulfide exchange to the point where it is negligible during the time scale of HPLC separation of components present in the quenched aliquots. Quenched aliquots were frozen and stored at −80 °C until needed for HPLC analysis. Fig. 2 exhibits kinetic data for such an experiment. The starting reagent is monomer loop. This is reduced by excess GSH to produce approximately equal amounts of the SI and SII species. It is not known which peak, S1 or S2, correlates with which structure, SI or SII. At long times, SI and SII are further reduced to generate R with two free sulfhydryl groups. For some peptides, transient population of oligomeric species is observed. Such oligomers in principle may be formed by processes such as *Reactions* 5 and 6 of Fig. 1. No peptide was observed to produce any D under the strongly reducing conditions of these kinetics experiments.

A different protocol was used to monitor intramolecular exchange between LR forms of a 3-cysteine peptide, given by *Reaction* 7 of Fig. 1. The peptide ending in -Cys-Cys-Cys-amide has two stable LR forms at equilibrium. One of these was purified at pH 2 by reverse phase HPLC. In the kinetics experiment, the pH then was jumped up to 7.0, without addition of any GSH or GSSG, to accelerate intramolecular disulfide exchange. As time proceeds, the LRI form converted to an equilibrium mixture of the LRI and LRII forms. At intermediate times during the approach to equilibrium, aliquots were removed and were rejumped back to pH 2 by addition of trifluoroacetic acid to generate quenched samples for HPLC analysis.

Kinetic data were analyzed by Runga-Kutta-Fehlberg numerical integration of the set of differential rate equations describing the contributing reactions (12). Values of rate constants used as input parameters to the simulation were altered to given minimum deviations between observed and simulated concentrations of species. Plots of concentration *versus* elapsed time were constructed. Observed data were found to lie between simulated plots given by the final values for the k values, ± 20% variation. This is the margin of error for all rate constants determined below. Equilibrium constants have a 30% margin of error. Microscopic rate constants were determined for the two separate equilibria describing formation of L by SI or SII. In all cases, rate constants for the SI species were similar to the analogous rate constants for the SII species, differing by no more than 30%. Values reported below are the average of the separate values for the two similar reactions. Since disulfide exchange depends on the thiolate form of sulfhydryl groups, values reported below have been normalized to the high pH limit of 100% thiolate form. Calculations

[1] Portions of this paper (including part of "Materials and Methods," part of "Results," Table III, Refs. 22–91) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

[2] The abbreviations used are: HPLC, high pressure liquid chromatography; -SSG, a glutathione-containing mixed disulfide; CmCys, carboxymethylated cysteine; PTH, phenylthiohydantoin.

[3] Short 2-cysteine peptides sometimes are represented by an abbreviated sequence listing only the amino acids between the cysteines: for example, Ala_1, Tyr-Ser-Arg-Cys-Ala-Cys.

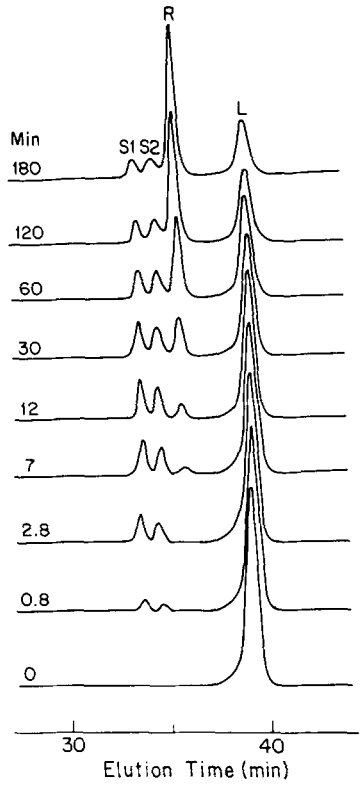

FIG. 2. **Observed kinetics of reduction of monomer loop of Tyr-Ser-Arg-Cys-Cys.** L_0 = 0.16 mM; GSH_0 = 5.2 mM; pH = 6.9, temperature = 23 °C. No dimers or D were observed. Linear 60-min gradiant from 0.1% trifluoroacetic acid in H_2O to 0.1% trifluoroacetic acid in 35% CH_3CN, detected from 13- to 73-min elution time. Flow rate = 1 ml/min; wavelength = 280 nm; column = 0.4 × 30 cm Waters μBondapak C_{18}.

assume a typical value of 8.9 for the pK_a of each cysteine (8). Thus, for example, the Henderson-Hasselbalch relationship permits one to calculate that a typical cysteine studied at pH 7.0 would spend only one-eightieth of its time in the reactive deprotonated state. As such, effective values for observed rate constants at pH 7 are multiplied by 80 to generate the normalized magnitudes.

Fully reduced 3-cysteine peptides were slowly oxidized at pH 8.3 by addition of oxygen to a buffer containing 0.3 mM peptide. Buffers were similar to those above, but lacked supplementary Cu^{2+} which would have accelerated the slow oxidation process. No glutathione forms were present in the solution. Ellman's reagent was used to determine the time point at which 50% of the total sulfhydryl groups had been oxidized to disulfides. This generally occurred about 3 h after starting the reaction. At that time, the pH was lowered by addition of trifluoroacetic acid. This quenched both net oxidation by oxygen and also disulfide exchange reactions. Different oxidized forms of the 3-cysteine peptide then were separated by preparative HPLC at pH 2.

Some of the partially oxidized forms were LR monomers containing one free cysteine and one intramolecular disulfide. To identify which of the 3 cysteines was free, each LR species purified at pH 2 was combined with 1.0 M iodoacetate at pH 7 to irreversibly alkylate the free sulfhydryl groups. The alkylated derivative then was loaded on an Applied Biosystems model 470A protein sequencer synchronized with an on-line model 120A microbore HPLC analyzer. For each species studied, carboxymethylated cysteine was detected primarily in only 1 of the 3 cysteine positions. Cysteines joined in a disulfide did not give a CmCys signal in the sequencing chromatogram.

RESULTS

Two-Cysteine Peptides—The effect of the negative carboxyl group on disulfide exchange reactions involving carboxyl-terminal cysteines is seen by comparing -Cys-Cys with -Cys-Cys-Ala and by comparing -Cys-Val-Cys with -Cys-Val-Cys-Ala. For -Cys-Cys, addition of an alanine at the carboxyl terminus increases k_c 4.0-fold from 0.30 to 1.2 s^{-1} and increases k_o 5.0-fold from 11 to 55 s^{-1} M^{-1}. Since both rate constants are increased by approximately similar amounts, the equilibrium constant for loop formation given by the ratio $K_c = k_c/k_o$ is essentially unaffected. Similarly, addition of an alanine at the carboxyl end of -Cys-Val-Cys increases k_c and k_o by factors of 4.8 and 5.6, respectively, again not affecting K_c. Thus the negative carboxyl group slows disulfide exchange by a factor of 5, presumably by inhibiting the presence of a neighboring negative $-S^-$ group, both when the carboxyl-terminal cysteine is attacking a disulfide and when it is part of the disulfide being attacked. The carboxyl group does not, however, have a significant impact on the equilibrium constant for forming an adjacent disulfide loop.

Rate constants for carboxyl-terminal loops are arranged in rank order of increasing magnitude of k_o in Table I. The rate constant for loop closing, k_c, varies over a 370-fold range from 0.10 s^{-1} in Pro_2 to 37 s^{-1} in Gly_1. The three slowest loop closing reactions occur in the largest loop, Ala_5, and in the two species -Cys-Cys- and Pro_2. These latter two species have strong limitations on allowed values of backbone dihedral angles. Formation of a disulfide in Cys-Cys requires prior formation of an unfavorable cis peptide bond between the cysteines (13). Prolines have only one favorable value for the dihedral angle ϕ describing rotation about the N–C_α bond. Loop opening processes, given by k_o, vary over a 5000-fold range from 5.0 s^{-1} M^{-1} in Pro_2 to 25 × 10^3 s^{-1} M^{-1} in Gly_1. The five peptides with the smallest values for k_o also have the smallest values for k_c. Gly_1 exhibits the fastest example of both kinetic processes. Peptides with the structure Cys-*X*-Cys have the fastest values for k_o, regardless of the nature of *X*.

The values of equilibrium constants determined by calculating the ratio of rate constants obtained in kinetics experiments are given in Table I. They are identical, within experimental error of ±30%, with values calculated from the ratios of molecular concentrations obtained in equilibrium experiments. Observed numbers vary over a 50-fold range, from 1.5 mM for the most forbidden loop Gly_1 to more than 71 mM for the most favored loop Ala_2. The four smallest K_c values correspond to the four X_1 species, each of which has a very fast k_o. It is important to note that the case of two adjacent cysteines is not the most difficult disulfide to form. In fact, -Cys-Cys- has an intermediate value for K_c. For a given value of *m*, the value of K_c is not strongly dependent on the nature of *X*. Thus in the X_1 series, K_c exhibits a maximum 5-fold difference from the 1.5 mM in Gly_1 to 7.3 mM in Pro_1. For X_2 peptides, K_c again exhibits a maximum 5-fold range, in this case from 16 mM in Pro_2 to 86 mM in Ala_2.

Table II aids in consideration of the effect of varying *m*.

TABLE I

Rate constants and equilibrium constants for loop closing and opening

Peptide	k_c	k_o	$K_c = k_c/k_o$
	s^{-1}	s^{-1} M^{-1}	*mM*
Pro_2	0.10	5.0	20
Cys-Cys	0.30	11	27
Ala_4	0.90	16	55
Ala_5	0.55	18	30
Ala_3	0.90	50	18
Gly_2	4.0	0.13 × 10^3	31
Val_2	10	0.14 × 10^3	71
Ala_2	10	0.14 × 10^3	71
Pro_1	1.5	0.21 × 10^3	7.3
Val_1	2.9	1.8 × 10^3	1.6
Ala_1	5.0	1.6 × 10^3	3.1
Gly_1	37	25 × 10^3	1.5

TABLE II

Dependence of major changes in physical constants on the odd-even nature of m

For each row, the magnitudes of the effects on k_c and k_o are compared, and the greater effect is given in bold print.

Comparison	k_c (forward)	k_o (reverse)	Equilibrium ratio $K_c = k_c/k_o$
	-fold		*-fold*
$Pro_{1\ vs\ 2}$	15 dec[a]	**42 dec**	
$Gly_{1\ vs\ 2}$	9.3 dec	**190 dec**	
$Val_{1\ vs\ 2}$	3.5 inc	**13 dec**	
$Ala_{1\ vs\ 2}$	2.0 inc	**12 dec**	23 inc
$Ala_{2\ vs\ 3}$	**11 dec**	2.8 dec	3.9 dec
$Ala_{3\ vs\ 4}$	No change	**3.1 dec**	3.1 inc
$Ala_{4\ vs\ 5}$	**1.6 dec**	1.1 inc	1.8 dec

[a] dec, decrease; inc, increase.

For all types of X examined, increasing m from 1 to 2 has a greater effect on k_o than on k_c. For example, changing Ala_1 to Ala_2 increases k_c by a factor of 2.0 but decreases k_o by a factor of 12. The larger effect, 12, is indicated in the table by bold print. Extensive variation in m is given only for the Ala_m series, whose data appear in the lower part of the table. Focusing only on the bold print, one notes that the dominant effect alternates back and forth between k_o and k_c. In other words, the larger effect occurs with opening for the case of changing X_1 to X_2, closing for X_2 *versus* X_3, opening for X_3 *versus* X_4, and closing for X_4 *versus* X_5. All of these dominant effects are decreases in the corresponding rate constant. Furthermore, their absolute magnitudes (12, 11, 3.1, 1.6) decrease monotonically as m increases. Each time a dominant decrease occurs in the forward rate constant, k_c, the equilibrium constant for loop formation exhibits a decrease. Alternatively each time a dominant decrease occurs in the reverse rate constant, k_o, the equilibrium constant increases. The net result on the equilibrium constant, given by the right column in Table II, therefore is an alternating pattern of increases and decreases in K_c with a monotonically declining value in absolute magnitude of the effect as m gets larger. This means that a loop with an even number (m) of alanines between the cysteines is more stable than a loop with $m \pm 1$ intervening alanines, where $m \pm 1$ is an odd number. For example, K_c for Ala_2 is greater than K_c for Ala_1 or Ala_3. Although it also is true that K_c for Cys-Cys ($m = 0$) is greater than K_c for Cys-Ala-Cys ($m = 1$), the case of two sequentially adjacent cysteines requires special considerations to be discussed below.

Three-Cysteine Peptides—Peptides in the 50% oxidized sample of Tyr-Arg-Cys-Ser-Cys-Ser-Cys-amide eluted in two HPLC peaks at pH 2. One of these was fully reduced monomeric R, and the other had an average of one sulfhydryl group and one disulfide per peptide chain. This latter peak tentatively was assigned to the monomeric LR form. To determine if it really was monomer, equilibrium experiments were performed where the concentration of total peptide was varied 10-fold in solutions containing a constant ratio of excess GSH/GSSG. Ten peaks were observed, including the monomeric R peak and the tentatively assigned LR peak observed in the air-oxidation reaction. The relative magnitudes of these 10 peaks were independent of peptide concentration, thereby indicating that all species were monomeric like R and that the tentative LR assignment was correct. A complete analysis of the glutathione-containing equilibrium mixture was not attempted. In reactions between a 3-cysteine species and glutathione, there are 14 possible monomeric forms which could contribute to the HPLC data.

The monomeric LR peak in the slow air oxidation solution was purified by preparative HPLC at pH 2 and then alkylated with 1.0 M iodoacetate at pH 7. Sequencing data for the carboxymethylated derivative are presented in Fig. 3. PTH-derivatives of tyrosine, serine, CmCys, and arginine all elute between 17 and 27 min. Standards containing equimolar amounts of these amino acid derivatives gave peak heights in the rank order Tyr > CmCys > Arg > Ser. The tyrosine ring itself contributes to absorption of 270 nm light in PTH-Tyr. PTH-Arg has a broadened and therefore relatively shorter peak. PTH-Ser is partially degraded by the temperature and timing parameters of the sequencing program used in this particular experiment. Cycle 1 is dominated by the PTH-Tyr peak, and cycle 2 is dominated by PTH-Arg. Similar analysis of subsequent steps gives the complete sequence Tyr-Arg-*X*-Ser-CmCys-Ser-*X*, where *X* corresponds to cycles in which only a small quantity of CmCys is observed. Thus the LR peak predominantly corresponds to the form containing a free cysteine at position 5 and an intramolecular disulfide joining cysteines 3 and 7. The partially oxidized sample at pH 7 therefore contains no significant amount of either of the two possible species of the Cys-*X*-Cys type joining cysteines 3 and 5 or joining cysteines 5 and 7.

A similar set of experiments was performed for the peptide Tyr-Arg-Cys-Cys-Cys-amide. In this case, there are two stable monomeric LR forms giving separate HPLC peaks midway thru the slow O_2-induced oxidation process. Sequencing analysis indicated that they correspond to the LRI and LRII structures given in *Reaction 7* of Fig. 1, where the central cysteine is joined by a disulfide to either of its immediately adjacent neighbors. Thus at pH 7, there is no significant amount of the one possible disulfide of the Cys-*X*-Cys type joining cysteines 3 and 5.

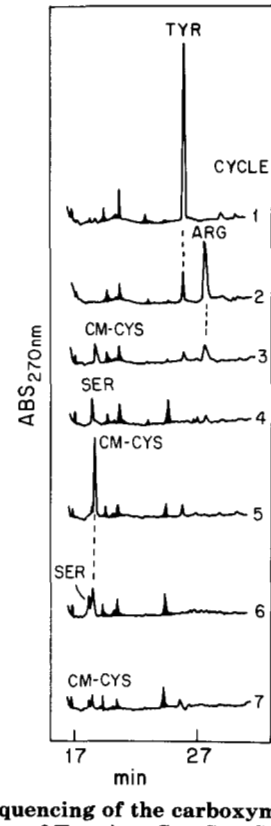

FIG. 3. **Microsequencing of the carboxymethylated derivative of the LR form of Tyr-Arg-Cys-Ser-Cys-Ser-Cys-amide.** Sample load = 1 nmol. *Dashed lines* indicate noticeable occasions where molecules failing to undergo Edman degradation in one cycle react successfully in the subsequent cycle. *Shaded peaks* are impurities or by-products appearing in all cycles.

DISCUSSION

The dependence of the rate constants k_c and k_o on the number of residues between the cysteines reflects several contributing factors. For 2 cysteines separated by many residues in the sequence, the probability of those cysteines colliding with each other in a denaturing solvent decreases as m increases (14, 15). This configurational entropy factor probably is responsible for much of the general trend for non-proline residues in Table I. For such residues, k_c values for different m occur in the rank order ($m = 5$) < ($m = 4$ or 3) < ($m = 2$ or 1). As m gets smaller and smaller, the tendency for cysteines to find each other more rapidly will be offset by unfavorable nonbonding interations. The need to pack backbone groups more tightly together in small loops will lead to distortion of bond lengths and bond angles, preventing formation of a very stable conformation for the looped molecule. This strain factor may contribute to the increase in k_o as m decreases. For non-proline molecules, k_o exhibits the rank order ($m = 5$ or 4) < ($m = 3$) < ($m = 2$) < ($m = 1$).

The case of two sequentially adjacent cysteines requires special consideration. For such cysteines, theoretical studies (13) and crystallographic data (16) indicate that formation of a disulfide requires a cis peptide bond between the cysteines. Formation of a cis peptide bond is very unfavorable unless it precede proline, in which case it is slightly unfavorable (17). In NMR studies of model compounds, 8% cis peptide bonds were detected in *N*-methylformamide but no cis peptide bonds were detected in *N*-methylacetamide, which more closely resembles polypeptide backbones (18). If one estimates that greater than 5% population of cis peptide bonds would have been detectable in the latter compound, then one can estimate that the equilibrium constant for trans/cis interconversion, K_{ISOM}, has a value <1/20 in the direction forming the cis configuration. The overall reaction for forming an intramolecular disulfide in Cys-Cys molecules can be represented by coupling the trans/cis isomerization with disulfide exchange in the cis form.

$$^{HS}\mathbf{Trans}^{SSG} \underset{}{\overset{K_{ISOM}}{\rightleftharpoons}} {}^{HS}\mathbf{Cis}^{SSG} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathbf{Cis}\text{-SS} + GSH$$

Cis-SS is the looped L form. In principle, analysis of kinetic data for these coupled equilibria could be complicated. Since velocities of isomerization are independent of pH but velocities of disulfide exchange depend on the percentage of molecules in the $-S^-$ form, different steps might be rate-limiting at the extremes of very low and very high pH. For the data observed here at pH 7, mathematical analysis is relatively uncomplicated. The low pH quench method will trap $^{HS}\mathbf{Cis}^{SSG}$ and allow it to convert to the more favored trans form via a pH-independent conformational change prior to the HPLC analysis. The HPLC peak for a single mixed-disulfide species (SI or SII) therefore will correspond to the sum of cis and trans S forms at the moment of the pH quench. For the species eluting in the S1 peak, this gives concentration ratio:

$$K_c = \frac{(\text{area of L peak})\ [GSH]}{(\text{area of S1 peak})} = \frac{[\mathbf{Cis}\text{-SS}][GSH]}{[^{HS}\mathbf{Trans}^{SSG}] + [^{HS}\mathbf{Cis}^{SSG}]} \approx \frac{[\mathbf{Cis}\text{-SS}][GSH]}{[^{HS}\mathbf{Trans}^{SSG}]}$$

where concentrations and peak areas are those observed in quenched aliquots of equilibrium mixtures. The symbol representing an approximate equality above is allowed since $^{HS}\mathbf{Trans}^{SSG} \gg {}^{HS}\mathbf{Cis}^{SSG}$ at equilibrium. The final ratio is the exact expression for the overall equilibrium constant for the exact consecutive reactions, such that

$$\frac{[\mathbf{Cis}\text{-SS}][GSH]}{[^{HS}\mathbf{Trans}^{SSG}]} = K_{ISOM}\,\frac{k_1}{k_{-1}}.$$

The Runga-Kutta-Fehlberg numerical simulation ignored separate cis and trans forms. Since the ratio of the simulated rate constants k_c/k_o is empirically observed to be approximately equal to the calculated concentration ratio K_c under the experimental conditions of pH 7, it then follows that

$$\frac{k_c}{k_o} \approx \frac{K_{ISOM}\ k_1}{k_{-1}}.$$

The observed rate constant k_o corresponds to the microscopic rate constant k_{-1} describing reduction of the disulfide loop by GSH. Thus $k_c \approx K_{ISOM} \times k_1$ at pH 7. Since $K_{ISOM} < 1/20$, $k_1 > 20\ k_c$. The value of k_c is 0.30 s^{-1}. Thus once the cis configuration is achieved, the microscopic rate constants for closing the loop, k_1, is > 6.0 s^{-1}. This is at the fast end of the rank order of observed loop closing rate constants in Table I. As mentioned earlier, the equilibrium constant for K_c is generally favorable in -Cys-Cys- sequences. Therefore the unfavorability of having to form a cis peptide bond between the cysteines is more than compensated by 1) the very high probability of collision between the two closely spaced sulfur groups in the cis configuration and 2) the absence of significantly unfavorable nonbonded interactions.

Previous studies of air oxidation of Cys-Cys at low ionic strength (4, 6) detected no monomeric loop. Dimer formation was dominant, in contrast to results observed above. In the previous studies, Cys-Cys was present at a concentration of 45 mM (10 g/liter, cited in Ref. 6), which is significantly higher than the 1.6 mM concentration of Tyr-Ser-Arg-Cys-Cys used here. That higher concentration would favor peptide/peptide collisions and corresponding dimer formation. Moreover at low ionic strength, electrostatic factors may have enhanced formation of antiparallel dimers of (^+H_3N)-Cys-Cys-(COO^-), which would juxtapose the negative carboxyl terminus of one chain with the positive amino terminus of the other chain.

Short polypeptides consist of a chain of rigid links, with each link having a fixed distance between the backbone nitrogen and carboxyl carbon atoms of an amino acid. Given restricted choices of allowed backbone dihedral angles, a chain of such rigid links has a tendency to persist in travelling in a given direction instead of readily reversing its orientation. Monte-Carlo theoretical methods have been applied to the problem of joining the amino- and carboxyl-terminal ends of short polyalanine chains to form a small loop in a denaturing solvent (19). This simulation required not only that the two termini be brought near to each other to create a new peptide bond, but also that they be mutually oriented at the proper angle for peptide bond formation. The probability of forming a cyclic peptide was calculated as a function of m. For small m, that calculated probability alternated between high and low values as m varied from even to odd numbers, with the amplitude of the oscillation decreasing as m increased. This is the same qualitative behavior observed in Table II for equilibrium constants of loop formation in -Cys-Ala$_m$-Cys-peptides in 3 M guanidine. It is possible that the pattern observed for this experimental series originates in the theoretically described persistence length effect for short polymers. Although one might postulate that hydrogen bonds in β turns could enhance loop stability for even-numbered loops such as $m = 2$ or 4, such bonds are not very stable in the 3 M guanidine conditions of these experiments.

This alternating odd-even pattern also appears in statistics

of disulfide loop formation in naturally occurring proteins. A survey of known disulfide positions is presented in the Miniprint. Fig. 4 summarizes data in that survey. The value at the top of each bar gives the absolute number of currently known examples of sequentially unrelated cysteine pairs for each value of m. The *vertical axis* plots the percentage of occasions where the 2 cysteines of those pairs are joined to each other in an intramolecular disulfide. Cysteine pairs with 2 or 4 intervening residues have a greater probability of loop formation than loops with $m = 1$, 3, or 5. This pattern ceases for $m > 5$. Competition between configurational entropy, loop strain, and persistence length effects appears to give the same net dominant influence in the biological and synthetic peptide data. A possible exception is the absence of many cases of Cys-Cys loops in naturally occurring proteins, perhaps reflecting the lack of much biological utility for this tiny but energetically permitted loop.

Given the very different effects of glycine, valine, proline, and alanine on turn formation as mentioned in the introduction, it is not possible to give a definitive explanation of the general lack of dependence of loop stability on the nature of X in X_1 and X_2 series. Each peptide is probably a special case of its own, with different factors compensating to give roughly equivalent values of K_c for a particular value of m. In studies of several X_4 species, the positioning of valine, glycine, and proline residues had a more noticeable effect (20). Loop formation correlated with predicted potentials for β turn formation, yielding no monomer for Val_4 and an 8-fold maximum range in K_c for the heteropolymers examined.

The determination of the predominant disulfide in a given 3-cysteine LR form relies on alkylation of the free cysteine not involved in the disulfide. For a peptide having two or more stable LR forms, it therefore is necessary that alkylation of each purified stable form be faster than intramolecular disulfide exchange between those forms. In the Tyr-Arg-Cys-Cys-Cys-amide species described above, the interconversion of LRI and LRII was depicted in Fig. 1 and is characterized by forward and reverse rate constants which both have the value of 0.024 s^{-1}. Iodoacetate carboxymethylates free cysteines with a rate constant of about 16 s^{-1} M^{-1} (21). Thus in the experiments described above, alkylation by 1.0 M iodoacetate occurs with a pseudo-first order rate constant of 16 s^{-1}, considerably faster than the disulfide exchange. In contrast, for a peptide having only one stable LR form, such as Tyr-Arg-Cys-Ser-Cys-Ser-Cys-amide, it is irrelevant whether alkylation is faster than intramolecular disulfide exchange. If the intramolecular exchange were to be faster, then the solution would establish and maintain an equilibrium population of different LR forms during the slower alkylation. The most stable LR form would predominate among LR forms at all time points. Given equal probabilities of reaction of iodoacetate with any cysteine in the solution, one therefore would obtain the expected single major alkylated product in this fast-exchange limit.

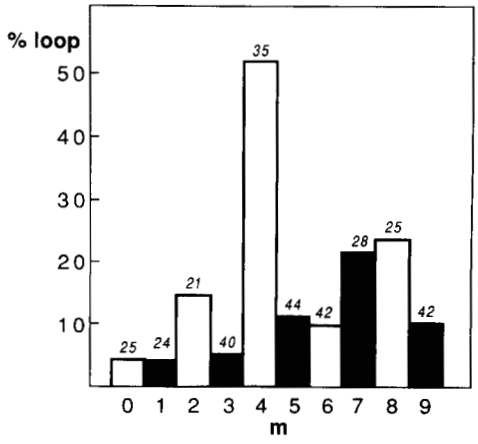

FIG. 4. **Statistics of occurrence of small disulfide loops in naturally occurring proteins.**

A disulfide loop between adjacent cysteines in -Cys-Cys-Cys-amide will have the same bonding interactions and thermodynamic stability as a disulfide loop between adjacent cysteines in -Cys-Cys-Ala. The difference in free energy (ΔG) between the reduced chain and the loop-containing chain should have the same value in both sequences. Values of K_c, which are equal to exp $(-\Delta G/RT)$, therefore should also be independent of the loop's environment. Thus it should be possible to use the relative values of the equilibrium constants K_c in 2-cysteine peptides to predict the relative populations of LR species at equilibrium in a 3-cysteine species. For example, K_c for forming Ala_1 and Ala_3 loops are 3.1 and 18 mM, respectively. It is reasonable to estimate that equilibrium constants for small serine-containing loops will be similar, given the lack of significant dependence of K_c on the type of X observed above for very dissimilar types of amino acids. Thus relative populations of the 3–5, 3–7 and 5–7 disulfides in Tyr-Arg-Cys-Ser-Cys-Ser-Cys-amide should be 3:18:3, respectively. The sequencing data in Fig. 3 exhibit predominance of the 3–7 disulfide, as expected. Similarly, the relative populations of the 3–4, 3–5, and 4–5 disulfides in Tyr-Arg-Cys-Cys-Cys-amide should be 27:3:27, respectively. Experimental observation of equimolar amounts of the 3–4 and 4–5 disulfides, with no significant 3–5 loops, is consistent with the prediction.

Although equilibrium constants are expected to be similar in the 2- and 3-cysteine peptides for the reasons given above, this is not true for the rate constants. Fig. 1 depicts two types of reactions which are unimolecular in both directions. These are interconversion of S forms in *Reaction 4*, and interconversion between LR forms in *Reaction 7*. Interconversion of the S forms was directly measured only for the Val_1 peptide examined previously (8). Those experiments demonstrated that the intramolecular interconversion rate constants and the intramolecular loop-closing rate constant k_c differ only by a factor of 2. It therefore is reasonable to expect that the SI and SII forms of -Cys-Cys-Ala will interconvert with a rate constant similar to the value of 1.2 s^{-1} observed for k_c in this peptide. This estimated value for the 2-cysteine peptide is 50 times faster than the 0.024 s^{-1} interconversion of LRI and LRII forms of the 3-cysteine peptide, -Cys-Cys-Cys-amide. In the former case, a free cysteine is trying to attach to an adjacent cysteine participating in a mobile mixed disulfide with glutathione. In the latter case, a free cysteine is trying to attach to an adjacent cysteine participating in a rigidly constrained **Cis**-SS loop. These are very different processes, exhibiting rate constants which are estimated to differ by more than a factor of 10.

Acknowledgments—We appreciate the contributions by Wenhan Chang, who obtained the data for the Cys-Ala-Cys peptide during a graduate laboratory rotation. We also thank Robert L. Baldwin for directing our attention to the literature on alternating odd-even effects in polymers of amino acids. We are indebted to Yiping Sun in our university's core facility for dependable synthesis of HF-cleaved peptides, development of HPLC separation protocols, and assistance with peptide sequencing.

REFERENCES

1. Kao, P. N., and Karlin, A. (1986) *J. Biol. Chem.* **261**, 8085–8088
2. Edman, J. C., Ellis, L., Blacher, R. W., Roth, R. A., and Rutter,

W. J. (1985) *Nature* **317,** 267–270
3. Tan, F., Zhang, G., Mu, J., Lin, N., and Chi, C. (1984) *Hoppe-Seyler's Z. Physiol. Chem.* **365,** 1211–1217
4. Heaton, G. S., Rydon, H. N., and Schofield, J. A. (1956) *J. Chem. Soc.* 3157–3168
5. Jarvis, D., Rydon, H. N., and Schofield, J. A. (1961) *J. Chem. Soc.* 1752–1756
6. Hardy, P. M., Ridge, B., Rydon, H. N., and Serrao, F. O. S. P. (1971) *J. Chem. Soc. Sect. C Org. Chem.* 1722–1731
7. Weber, U., and Hartter, P. (1974) *Hoppe-Seyler's Z. Physiol. Chem.* **355,** 189–199
8. Zhang, R., and Snyder, G. H. (1988) *Biochemistry* **27,** 3785–3794
9. Rose, G. D., Gierasch, L. M., and Smith, J. A. (1985) *Adv. Protein Chem.* **37,** 1–109
10. Mihalyi, E. (1968) *J. Chem. Eng. Data* **13,** 179–182
11. Riddles, P. W., Blakeley, R. L., and Zerner, B. (1979) *Anal. Biochem.* **94,** 75–81
12. Danby, J. M. A. (1985) in *Computing Applications to Differential Equations*, pp. 1–52 and 245–246, Reston Publishing, Reston, VA
13. Chandrasekaran, R., and Balasubramanian, R. (1969) *Biochim. Biophys. Acta* **188,** 1–9
14. Jacobson, H., and Stockmayer, W. H. (1950) *J. Chem. Phys.* **18,** 1600–1606
15. Kauzmann, W. (1959) in *Sulfur in Proteins* (Benesch, R., Benesch, R. E., Boyer, P. D., Klotz, I. M., Middlebrook, W. R., Szent-Gyorgyi, A. G., and Schwarz, D. R., eds) pp. 93–108, Academic Press, New York
16. Capusso, S., Mattia, C., Mazzarella, L., and Puliti, R. (1977) *Acta Crystallogr. Sect. B Struct. Sci.* **33,** 2080–2083
17. Creighton, T. E. (1983) *Proteins: Structures and Molecular Properties*, pp. 162–165, W. H. Freeman & Co., New York
18. LaPlanche, L. A., and Rogers, M. T. (1964) *J. Am. Chem. Soc.* **86,** 337–341
19. Mutter, M. (1977) *J. Am. Chem. Soc.* **99,** 8307–8314
20. Milburn, P. J., Meinwald, Y. C., Takahashi, S., Ooi, T., and Scheraga, H. A. (1988) *Int. J. Pept. Protein Res.* **31,** 311–321
21. Snyder, G. H. (1987) *Biochemistry* **26,** 688–694

Additional references are found on p. 18479.

Supplementary Material to

Dependence of Formation of Small Disulfide Loops in Two-cysteine Peptides on the Number and Types of Intervening Amino Acids

Rumin Zhang and Grayson H. Snyder

Methods

Before evaluating the tendency of -Cys-X(m)-Cys- sequences to form small disulfide loops in naturally occurring proteins, it is necessary to establish criteria for selecting data which is most relevant to consideration of competitive intramolecular disulfide exchange. In summaries to be described below, the first criterion applied was to exclude proteins like thioredoxin which have only two cysteines, since disulfide formation between those cysteines is the only possibility in a strongly oxidizing buffer. Secondly, proteins like immunoglobulin having interchain disulfides were excluded since their structures are determined by factors not limited to interactions within a single polypeptide chain. Proteins which lack a pair of cysteines separated by 9 or fewer residues in the sequence are not listed below, nor are proteins for which only partial disulfide pairing information is available.

In order to develop reasonable statistics, it is important to count each situation only once. Thus in families of related proteins like chymotrypsinogen and trypsinogen which have some disulfides in common, no conserved disulfide should be counted more than once. Similarly for proteins like wheat germ agglutinin, which has four internally homologous domains, each type of interdomain or intradomain cysteine pair should be considered only once. Having accounted for such factors it then is necessary to list each and every unrelated pair of cysteines, without regard to whether or not any additional cysteines contribute to the intervening residues between the pair of cysteines under consideration. Thus the -Cys-Cys-X$_3$-Cys- sequence found in proinsulin would contribute three examples to the statistics, namely one case each of m = 0, 3 and 4.

Results

Table III presents a data base of current information about unrelated cysteine pairs separated by 9 or fewer residues in native proteins. The first example, human defensin, is a protein with one example each of a cysteine pair with m = 0,1,4 or 6 and two examples of pairs with m = 9. Since none of these pairs forms a disulfide, each example is listed in the the column for cases "not formed". The six examples (0,1,4,6,9,9) are coded by the string of digits "014699" written in ascending order, one digit for each independent example. For the proinsulin three-cysteine cluster mentioned above, the first and last cysteines pair with each other, generating an example of m = 4 in the "formed" column and examples of m = 0 and 3 in the "not formed" column. A summary of the data base is given by Figure 4 in the text, which shows that pairs with m = 2 or 4 are more likely to form a disulfide than examples with m = 1, 3 or 5.

Table III. Data base of unrelated cysteine pairs with 9 or fewer intervening residues

Protein	Loop Formed	Loop Not Formed	Reference
Human defensin HNP-2		014699	22
Papaya papain		26	23
Bovine prolactin	67		24
Bovine proinsulin	4	038	25
Bovine prothrombin	4	44	26
Ascaris trypsin inhibitor		0124455689	27
Horseradish peroxidase	4	5	28
Bovine carboxypeptidase B		48	29
Bovine neurophysin	45	002555556666799	30
Polypeptide chymotrypsin inhibitor-1		02335689	31
Sea anemone antiviral protein BDS-I		01679	32
Ecballium elaterium trypsin inhibitor		13555679	33
Achromobacter lytic protease I		5	34
Azotobacter lipoamide dehydrogenase	4		35
(Na, K)-ATPase beta subunit		9	36
Wheat germ agglutinin	4	0345556668	37
Influenza virus neuraminidase	448	1113679	38
Porcine pancreatic lipase	45	12	39
Bovine pancreatic phospholipase A2		011456666689	40
Soybean seed trypsin inhibitor	67	01123334567889999	41
Trypanosome variant surface glycoprotein		5678	42
Bovine pancreatic trypsinogen		69	43
Human chorionic gonadotropin beta subunit		122346799	44
Human chorionic gonadotropin alpha subunit	1	0022234	45
Indian cobra toxin B	34	035559	46
Chicken egg-white lysozyme		3	47
Bovine pancreatic ribonuclease	6	6	48
Scorpion insect toxin		013345569	49
Scorpion neurotoxin II		13359999	50
Heteronemertine worm toxin B-IV		2333669	51
Wheat purothionin A-II	8	013357789	52
Porcine pepsin	34		53
Mouse submaxillary gland nerve growth factor		19	54
Ustilago sphaerogena ribonuclease U2		07	55
Heteronemertine worm toxin A-III		359	56
Bovine basic pancreatic trypsin inhibitor		378	57
Bovine acidic pancreatic trypsin inhibitor		267	58
Mouse submaxillary gland epidermal growth factor	8	157	59
Sea anemone toxin II		0169	60
North american rattlesnake myotoxin A		05666	61
Potato carboxypeptidase A inhibitor		23556899	62
Escherichia coli alkaline phosphatase	9		63
Bovine deoxyribonuclease A	2		64
Kunitz soybean trypsin inhibitor	8		65
Rat brain glycoprotein Thy-1		9	66
Actinoxanthin	48		67
Aspergillus oryzae ribonuclease T1	7	33	68
Human platelet beta thromboglobulin		1	69
Honey bee apamine	9	137	70
Human plasma alpha-1 acid glycoprotein		1	71
Stellacyanin		5	72
Human pancreatic stone protein S1		7	73
Acetylcholine receptor alpha subunit	05		74
Human low-Mr kininogen	2		75
Lumber mold proteinase K		9	76
Human free secretory component	6	39	77
Thaumatin I	4589	344579	78
Streptomyces metallo-proteinase inhibitor	45		79
Leech hirudin	7	1557	80
Human mucous proteinase inhibitor		033455667788	81
Porcine pancreatic alpha amylase		35	82
Bovine pancreatic procarboxypeptidase A-III	2	89	83
Sea anemone elastase inhibitor		367	84
Human alpha-2 antiplasmin		8	85
West Nile flavivirus envelope protein		4	86
Escherichia coli enterotoxin STh	477	0022333468	87
Conotoxin MI	49	035	88
Spider hemocyanin	4		89
Human C3a anaphylatoxin		0067	90
Human lactoferrin	4	22788999	91

References

22. Selsted, M. E., and Harwig, S. S. L. (1989) *J. Biol. Chem.* **264**, 4003-4007
23. Dayhoff, M. O. (1972) in *Atlas of Protein Sequence and Structure* Vol. 5, pp. 121, National Biomedical Research Foundation, Silver Spring, MD
24. *Ibid*, pp. 202
25. *Ibid*, pp. 209
26. Dayhoff, M. O. (1976) in *Atlas of Protein Sequence and Structure* Vol.5, Supplement 2, pp.95, National Biomedical Research Foundation, Silver Spring, MD
27. *Ibid*, pp.112
28. Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure* Vol.5, Supplement 3, pp.68. National Biomedical Research Foundation, Silver Spring, MD
29. *Ibid*, pp.123
30. Burman, S., Wellner, D., Chait, B., Chaudhary, T., and Breslow, E. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 429-433
31. Greenblatt, H. M., Ryan, C. A., and James, M. N. G. (1989) *J. Mol. Biol.* **205**, 201-228
32. Driscoll, P. C., Cronenborn, A. M., Beress, L., and Clore, G. M. (1989) *Biochemistry* **28**, 2188-2198
33. Heitz, A., Chiche, L., Le-Nguyen, D., and Castro, B. (1989) *Biochemistry* **28**, 2392-2398
34. Tsunasawa, S., Masaki, T., Hirose, M., Soejima, M., and Sakiyama, F. (1989) *J. Biol. Chem.* **264**, 3832-3839
35. Shierbeek, A. J., Swarte, M. B. A., Dijkstra, B. W., Vriend, G., Read, R. J., Hol, W. G. J., Drenth, J., and Betzel, C. (1989) *J. Mol. Biol.* **206**, 365-379
36. Kirley, T. L. (1989) *J. Biol. Chem.* **264**, 7185-7192
37. Wright, C. S. (1977) *J. Mol. Biol.* **111**, 439-457
38. Varghese, J. N., Laver, W. G., and Colman, P. M. (1983) *Nature* **303**, 35-40
39. Benkouka, F., Guidoni, A. A., de Caro, J. D., Bonicel, J. J., Desnuelle, P. A., and Rovery, M. (1982) *Eur. J. Biochem.* **128**, 331-341
40. Dijkstra, B. W., Darenth, J., Kalk, K. H., and Vanmaelen, P. J. (1978) *J. Mol. Biol.* **124**, 53-60
41. Odani, S., and Ikenaka, T. (1973) *J. Biochem.* **74**, 694-715
42. Allen, G., and Gurnett, L. P. (1983) *Biochem. J.* **209**, 481-487
43. Mikes, O., Holeysovsky, V., Tomasek, V., and Sorm, F. (1966) *Biochem. Biophys. Res. Commun.* **24**, 346-352
44. Mise, T., and Bahl, O. P. (1981) *J. Biol. Chem.* **256**, 6587-6592
45. Mise, T., and Bahl, O. P. (1980) *J. Biol. Chem.* **255**, 8516-8522
46. Ohda, M., and Hayashi, K. (1973) *Biochem. Biophys. Res. Commun.* **55**, 431-438
47. Canfield, R., and Liu, A. K. (1965) *J. Biol. Chem.* **240**, 1997-2002
48. Smyth, D. G., Stein, W. H., and Moore, S. (1963) *J. Biol. Chem.* **238**, 227-234
49. Darbon, H., Zlotkin, E., Kopeyan, C., van Rietschoten, J., and Rochat, H. (1982) *Int. J. Pept. Prot. Res.* **20**, 320-330
50. Kapeyan, C., Martinez, G., Lissitzky, S., Miranda, F., and Rochat, H. (1974) *Eur. J. Biochem.* **47**, 483-489
51. Blumenthal, K. M., Keim, P. S., Heinrikson, R. L., and Kem, W. R. (1981) *J. Biol. Chem.* **256**, 9063-9067
52. Hase, T., Matsubara, H., and Yoshizumi, H. (1978) *J. Biochem.* **83**, 1671-1678
53. Sepulveda, P., Marciniszyn, J. Jr., Liu, D., and Tang, J. (1975) *J. Biol. Chem.* **250**, 5082-5088
54. Angeletti, R. H., Hermodson, M. A., and Bradshaw, R. A. (1973) *Biochemistry* **12**, 100-115
55. Sato, S., and Uchida, T. (1975) *J. Biochem.* **77**, 1171-1176
56. Blumenthal, K. M. (1980) *J. Biol. Chem.* **255**, 8273-8274
57. Anderer, F. A., and Hornle, S. (1966) *J. Biol. Chem.* **241**, 1568-1572
58. Guy, O., Shapanka, R., and Greene, L. J. (1971) *J. Biol. Chem.* **246**, 7740-7747
59. Savage, C. R. Jr., Hash, J. H., and Cohen, S. (1973) *J. Biol. Chem.* **248**, 7669-7672
60. Wunderer, G. (1978) *Hoppe-Seyler's Z. Physiol. Chem.* **359**, 1193-1201
61. Fox, J. W., Elzinga, M., and Tu, A. T. (1979) *Biochemistry* **18**, 678-684
62. Rees, D. C., and Lipscomb, W. N. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 4633-4637
63. Bradshaw, R. A., Cancedda, F., Erickson, L. H., Neumann, P. A., Piccoli, S. P., Schlesinger, M. J., Shriefer, K., and Walsh, K. A. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 3473-3477
64. Liao, T.-H., Salnilow, J., Moore, S., and Stein, W. H. (1973) *J. Biol. Chem.* **248**, 1489-1495
65. Koide, K., Tsunasawa, S., and Ikenaka, T. (1972) *J. Biochem.* **71**, 165-167
66. Campbell, D. G., Gagnon, J., Reid, K. B. M., and Williams, A. F. (1981) *Biochem. J.* **195**, 15-30
67. Pletnev, V. Z., Kuzin, A. P., Trakhanov, S. D., and Kostetsky, P. V. (1982) *Biopolymers* **21**, 287-300
68. Takahashi, K. (1965) *J. Biol. Chem.* **240**, PC4117-4119
69. Begg, G. S., Pepper, D. S., Chesterman, C. N., and Morgan, F. J. (1978) *Biochemistry* **17**, 1739-1744
70. Callewaert, G. L., Shipolini, R., and Vernon, C. A. (1968) *FEBS Lett.* **1**, 111-113
71. Schmit, K., Burgi, W., Collins, J. H., and Nanno, S. (1974) *Biochemistry* **13**, 2694-2697
72. Engeseth, H. R., Hermodson, M. A., and McMillin, D. R. (1984) *FEBS Lett.* **171**, 257-261
73. Rouimi, P., de Caro, J., Bonicel, J., Rovery, M., and de Caro, A. (1988) *FEBS Lett.* **229**, 171-174
74. Mosckovitz, R., and Gershoni, J. M. (1988) *J. Biol. Chem.* **263**, 1017-1022
75. Kellermann, J., Thelen, C., Lottspeich, F., Henschen, A., Vogel, R., and Muller-Esterl, W. (1987) *Biochem. J.* **247**, 15-21
76. Jany, K.-D., Lederer, G., and Mayer, B. (1986) *FEBS Lett.* **199**, 139-144
77. Eiffert, H., Quentin, E., Decker, J., Hillemeir, S., Huftschmidt, M., Klingmuller, D., Weber, M. H., and Hilschmann, N. (1984) *Hoppe-Seyler's Z. Physiol. Chem.* **365**, 1489-1495
78. De Vos, A. H., Hatada, M., van der Wel, H., Krabbendam, H., Peerdeman, A. F., and Kim, S.-H. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 1406-1409
79. Murai, H., Hara, S., Ikenaka, T., Oda, K., and Murao, S. (1985) *J. Biochem.* **97**, 173-180
80. Dodt, J., Seemuller, U., Maschler, R., and Fritz, H. (1985) *Biol. Chem. Hoppe-Seyler* **366**, 379-385
81. Grutter, M. G., Fendrich, G., Huber, R., and Bode, W. (1988) *EMBO J.* **7**, 345-351
82. Pasero, L., Mazzei-Pierron, Y., Abadie, B., Chicheportiche, Y., and Marchis-Mouren, G. (1986) *Biochim. Biophys. Acta* **869**, 147-157
83. Venot, N., Sciaky, M., Puigserver, A., Desnuelle, P., and Laurent, G. (1986) *Eur. J. Biochem.* **157**, 91-99
84. Tschesche, H., Kolkenbrock, H., and Bode, W. (1987) *Biol. Chem. Hoppe-Seyler* **368**, 1297-1304
85. Lijnen, H. R., Holmes, W. E., van Hoef, B., Wiman, B., Rodriguez, H., and Collen, D. (1987) *Eur. J. Biochem.* **166**, 565-574
86. Nowak, T., and Wengler, G. (1987) *Virology* **156**, 127-137
87. Shimonishi, Y., Hidaka, Y., Koizumi, M., Hane, M., Aimoto, S., Takeda, T., Miwatani, T., and Takeda, T. (1987) *FEBS Lett.* **215**, 165-170
88. Gray, W. R., Rivier, J. E., Galyean, R., Cruz, L. J., and Olivera, B. M. (1983) *J. Biol. Chem.* **258**, 12247-12251
89. Eyerle, F., and Schartau, W. (1985) *Biol. Chem. Hoppe-Seyler* **866**, 403-409
90. Chazin, W. J., Hugli, T. E., and Wright, P. E. (1988) *Biochemistry* **27**, 9139-9148
91. Anderson, B. F., Baker, H. M., Dodson, E. J., Norris, G. E., Rumball, S. V., Waters, J. M., and Baker, E. N. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 1768-1773