doi:10.1016/j.jmb.2008.05.042 J. Mol. Biol. (2008) 381, 1068–1087

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Conformer Selection and Induced Fit in Flexible Backbone Protein–Protein Docking Using Computational and NMR Ensembles

Sidhartha Chaudhury[1] and Jeffrey J. Gray[1,2]*

[1]*Program in Molecular and Computational Biophysics, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA*

[2]*Department of Chemical and Biomolecular Engineering, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA*

Received 22 January 2008; received in revised form 15 May 2008; accepted 19 May 2008
Available online 24 May 2008

Accommodating backbone flexibility continues to be the most difficult challenge in computational docking of protein–protein complexes. Towards that end, we simulate four distinct biophysical models of protein binding in RosettaDock, a multiscale Monte-Carlo-based algorithm that uses a quasi-kinetic search process to emulate the diffusional encounter of two proteins and to identify low-energy complexes. The four binding models are as follows: (1) key-lock (KL) model, using rigid-backbone docking; (2) conformer selection (CS) model, using a novel ensemble docking algorithm; (3) induced fit (IF) model, using energy-gradient-based backbone minimization; and (4) combined conformer selection/induced fit (CS/IF) model. Backbone flexibility was limited to the smaller partner of the complex, structural ensembles were generated using Rosetta refinement methods, and docking consisted of local perturbations around the complexed conformation using unbound component crystal structures for a set of 21 target complexes. The lowest-energy structure contained >30% of the native residue–residue contacts for 9, 13, 13, and 14 targets for KL, CS, IF, and CS/IF docking, respectively. When applied to 15 targets using nuclear magnetic resonance ensembles of the smaller protein, the lowest-energy structure recovered at least 30% native residue contacts in 3, 8, 4, and 8 targets for KL, CS, IF, and CS/IF docking, respectively. CS/IF docking of the nuclear magnetic resonance ensemble performed equally well or better than KL docking with the unbound crystal structure in 10 of 15 cases. The marked success of CS and CS/IF docking shows that ensemble docking can be a versatile and effective method for accommodating conformational plasticity in docking and serves as a demonstration for the CS theory—that binding-competent conformers exist in the unbound ensemble and can be selected based on their favorable binding energies.

Keywords: protein–protein docking; flexible docking; ensemble docking; conformer selection; NMR ensembles

Edited by M. Sternberg

Introduction

The formation of highly specific protein complexes is a fundamental process in biology, and the structures of these complexes provide detailed insight into the mechanisms of protein function. Given the enormous number of undetermined complex structures and the difficulties in determining structures through X-ray crystallography or nuclear magnetic resonance (NMR), there is a need for accurate computational methods to predict protein complex structures. Complex structure prediction, also known as 'protein docking,' consists of determining the structure of the bound complex using the unbound structures of its partners (for a review, see Gray[1]). Although protein docking could be viewed as primarily an engineering problem—without the need to follow

**Corresponding author.* E-mail address: jgray@jhu.edu.

Abbreviations used: KL, key lock; CS, conformer selection; IF, induced fit; CS/IF, conformer selection/induced fit; FFT, fast Fourier transform; MC, Monte Carlo; MD, molecular dynamics; CAPRI, Critical Assessment of Protein Interactions; AChe, acetylcholinesterase; FAS2, fasciculin II; PDB, Protein Data Bank; MSF, mean square fluctuation.

the physical basis of protein binding—physical models of biomolecular interactions have long served as starting points for development. Furthermore, the success or failure of docking algorithms based on these models may provide insight into the biophysical models themselves.

A comparison of bound and unbound structures often reveals significant changes in backbone conformation upon binding,[2,3] which confound current docking methods and represent the single greatest challenge to predictive protein docking. Accurate modeling of backbone conformational changes in docking is difficult because of the enormous complexity of backbone conformational space both in size and in degrees of freedom (for a review, see Bonvin[4]), making exhaustive sampling impossible. Furthermore, since changes in backbone conformation affect both the intramolecular and the intermolecular energies of putative complexes, they create additional challenges in the discrimination of near-native structures. Effective strategies for both backbone conformational sampling and discrimination are needed before flexible docking is feasible for predictive applications.

Four biophysical models of protein binding suggest distinct conformational sampling strategies in flexible protein docking. First, the key-lock (KL) model of protein interactions, proposed by Fischer in 1894, states that proteins interact purely via surface complementarity of their rigid unbound structures.[5] The KL model is the most influential model in the development of protein docking algorithms, underlying the original grid-based[6] and fast Fourier transform[7] (FFT) techniques. Since most modern docking methods include side-chain motions in some capacity, for the purposes of this article, we define the KL model in terms of protein backbone flexibility. After allowing side-chain motions, KL docking strategies include, among others, ZDOCK/RDOCK,[8,9] ClusPro,[10] and RosettaDock.[11] These methods are moderately successful at predicting complexes for proteins that undergo minimal backbone conformational change upon binding, but perform progressively worse as the magnitude of backbone conformation changes increases.[12] Furthermore, these methods are far more successful when starting from bound structures than when starting from unbound structures,[11,13,14] even in cases with minor conformational changes, indicating that backbone flexibility is an important component to protein binding in general and that accurate modeling of backbone flexibility may improve docking for all complexes, not just those that exhibit greater flexibility.

Second, the conformational selection (CS) model proposed by Monod *et al.* for protein allostery and later adapted to protein interactions by Kumar *et al.* is a statistical mechanical view of protein binding.[15,16] The unbound state of a protein is represented by an ensemble of low-energy conformations, or conformers, one of which is the bound conformation. During the binding process, the bound-like conformers are selected over the other conformers in the ensemble as a result of their favorable binding energies. Thus, for docking, backbone flexibility is modeled implicitly as a pregenerated ensemble of rigid structures generated from the unbound structure. Previous CS-docking examples include both FFT-based[17–19] and Monte Carlo (MC)-based ensemble docking.[20] Both Smith *et al.* and Grunberg *et al.* used molecular dynamics (MD) to create an unbound ensemble that contained conformations that, in some ways, resemble the bound conformation, but both groups were unable to recover the bound structure in its entirety.[18,19] Subsequent FFT-based cross-docking of an ensemble of structures from the MD simulations showed substantially improved sampling near the bound structure in the former study and marginal improvements in docking accuracy in the latter study. Bastard *et al.* carried out ensemble docking with an ensemble of loop conformations and successfully discriminated near-native structures when the bound loop conformation was deliberately added to the ensemble.[20]

Third, in Koshland's induced fit (IF) model,[21] two proteins recognize each other to form an encounter complex, and then mutually alter their structures to form the intricate surface complementarity observed in bound structures. This model dictates that the bound conformation of a protein exists in response to the presence of the partner in complex, so the backbone conformational space must be sampled explicitly during docking in response to local energetics of the interface. Explicit backbone flexibility has been modeled primarily using MD,[22–24] energy minimization,[23,25] or gradient-based methods in MC minimization,[26,27] but not FFT-based methods. Wang *et al.* have shown impressive results in docking proteins in which a loop undergoes moderate to large conformational changes upon binding using explicit backbone flexibility, but their methods are extremely computationally intensive and require prior knowledge of the flexible regions, limiting their use in blind structure prediction.[26,28] Krol *et al.* carried out both MD relaxation and energy minimization of docking poses and showed significant increases in the fraction of native residue contacts recovered.[23]

The fourth model is a hybrid conformer selection/induced fit (CS/IF) model proposed by Grunberg *et al.* where binding is a two-stage process that begins with conformational selection to form an encounter complex, followed by an IF or 'refolding' step that leads to the final bound conformation.[19] The conformations that are selected to form the encounter complex need not resemble either the bound or the unbound structure. A docking algorithm based on this model would combine both ensemble docking and explicit backbone flexibility during docking. Krol *et al.* employed a docking strategy that combines FFT-based cross-docking of MD-generated ensembles with MD refinement of top-ranked decoys on a limited set of targets.[17] Their approach improved docking accuracy by measure of the number of native residue contacts recovered, but decreased accuracy by other measures such as the root mean square deviation (RMSD) of interface residues. HADDOCK combines both implicit and explicit backbone flexibilities while incorporating

biochemical information,[22] and it is capable of using ensembles from a wide variety of sources including MD, homology modeling, or NMR structures. When HADDOCK was used to dock MD-derived ensembles for two sets of Critical Assessment of Protein Interactions (CAPRI) targets, significant improvements over rigid-body docking were observed.[25,29] In several specific applications, docking of ensembles of NMR models using HADDOCK successfully generated structures of unknown protein complexes that have been subsequently validated by biochemical experiments.[30–32]

Our goal in this work is to implement the four binding models as conformational sampling strategies in a common algorithm (RosettaDock[11]) and to evaluate the differing abilities and limitations of each approach. We aim to create the optimal docking algorithm that allows backbone flexibility. As this is an ambitious goal, we limit this study to docking targets with small to moderate amounts of conformational change upon binding, and we restrict backbone motion to the smaller of the two proteins (ligand). At a minimum, a flexible backbone algorithm must be able to recover correct docked complex structures (near-native) for these simpler targets. Insights gained in this study can be later extended toward the goal of capturing large-scale conformational change in docking. Furthermore, we test our new ensemble-based methods using the sets of structural models provided by NMR solution-state studies (hereinafter referred to as NMR ensembles). Compared to computational ensembles, NMR ensembles are typically more diverse and arguably provide better representation of the unbound state.[33] Although NMR structures make up roughly 15% of known structures, a systematic study showing successful docking using NMR ensembles, to our knowledge, has never been performed.

Results

We have implemented and tested flexible backbone docking approaches using RosettaDock, a multiscale MC-based algorithm that samples docking orientations by emulating the diffusional encounter of two proteins in solution and by identifying low-energy complexes using an approximate energy function. Although RosettaDock is not a rigorous physical simulation of protein binding, its multiscale approach, quasi-kinetic sampling method, and previous success in the CAPRI[27,28,34,35] make it well suited for applying the four kinetic models of binding to protein docking. The low-resolution phase of RosettaDock simulates the formation of an encounter complex between two proteins, while the high-resolution phase models the transition from the encounter complex to a fully complexed structure. Physically realistic conformational sampling strategies may more effectively locate the correct complex structure[36] and, likewise, an efficient sampling technique may provide insight into the underlying theories that inspired its design.

The methods are tested on a set of docking targets using both crystal structures and NMR ensembles of the ligand, and backbone flexibility was limited to the ligand (Table 1). Docking is restricted to local perturbations around the native complexed orientation (local docking) that resembles blind predictive docking in cases where biochemical information on the interaction is available. Global docking, where all of docking search space is sampled, while compatible with the presented methods, is outside of the scope of this study due to the computational cost of evaluating a target benchmark of this size (~100 times the computational cost of local docking).[11]

Ensemble generation

To dock crystal structures, we created computational ensembles using RosettaRelax,[37,38] a multiscale MC-based structural refinement algorithm that samples the local conformation space for alternate low-energy structures using small backbone torsion angle perturbations, side-chain packing, and energy-gradient-based minimization in torsion space. Each structure required 15–30 min to generate on a 1.5-GHz processor, and 10 structures were generated for each target. Fig. 1a illustrates the ensemble generation technique for crystal structures, including idealization of bond lengths and angles, low-resolution relaxation, and high-resolution structural refinement. Fig. 1b illustrates the preparation of NMR ensembles, which consists of idealization followed by high-resolution refinement.

For the Rosetta-generated ensemble, diversity is achieved primarily through the low-resolution relaxation step. Conformers are typically within 1.0 Å C^{α} RMSD of each other, while NMR ensembles are generally much more diverse, as can be seen by comparing Fig. 1a and b. To show the conformational range of each type of ensemble and their potential for containing a binding-competent conformer, the C^{α} RMSD of the interface residues towards the bound ligand conformation after superposition of the entire ligand (BB_rmsd) is shown in Table 1 for the closest and the farthest conformers in each ensemble. In the Rosetta-generated ensembles, the closest conformer is often farther from the bound structure than the unbound crystal structure, suggesting that, within this target set, Rosetta has a limited ability to generate conformers that are closer to the bound state. In the NMR ensemble, the closest conformer is often significantly closer to the bound conformation than the first conformer (Model 1) in the NMR structure coordinate file, which is typically the model that satisfies the most experimental constraints.

Docking algorithms

The general docking algorithm is illustrated in Fig. 2. The ligand in the starting structure is first randomly translated and rotated to generate an initial starting position that approximates a collisional encounter between the two partners in which the

Table 1. Crystal structure and NMR docking targets

Crystal structure docking targets

Complex PDB	Receptor PDB	Receptor	Ligand PDB_X	Ligand	Size	Ub_{xtal}	Ens_{min}	Ens_{max}
2PTC(E:I)	2PTN	B-Trypsin	6PTI	Pancreatic trypsin inhibitor	10	0.28	0.82	1.3
2JEL*(LH:P)	2JEL(LH)	Jel42 Fab fragment	1POH	A06 phosphotransferase	10	0.31	0.39	1.0
1BVK(DE:F)	1BVL(LH)	Antibody Hulys11 Fv	3LZT	Lysozyme	10	0.38	0.71	1.8
1BQL*(LH:Y)	1BQL(LH)	Hyhel-5 Fab	1DKJ	Lysozyme	10	0.40	1.7	2.4
1DFJ(E:I)	2BNH	Ribonuclease inhibitor	7RSA	Ribonuclease A	10	0.40	1.7	2.4
2KAI(AB:I)	2PKA(XY)	Kallikrein A	6PTI	Pancreatic trypsin inhibitor	10	0.40	0.67	1.1
2BTF*(A:P)	2BTF(A)	B-actin	1PNE	Profilin	10	0.44	0.91	1.5
1BRS(A:D)	1A2P(B)	Barnase	1A19(A)	Barstar	10	0.47	0.77	1.7
1BRC(E:I)	1BRA	Trypsin	1AAP(A)	APPI	10	0.57	0.99	1.6
2SIC(E:I)	1SUP	Subtilisin BPN	3SSI	Subtilisin inhibitor	10	0.65	0.97	2.2
1MLC(AB:E)	1MLB(AB)	IgG1 D44.1 Fab fragment	1LZA	Lysozyme	10	0.73	0.69	1.7
2SNI(E:I)	1SUP	Subtilisin novo	2CI2(I)	Chymotrypsin inhibitor 2	10	0.82	0.96	1.9
1WQ1(G:R)	1WER	RAS-activating domain	5P21	RAS	10	0.85	1.0	1.5
1UGH(E:I)	1AKZ	Uracil DNA glycosylase	1UGI(A)	UDG inhibitor	10	0.90	0.92	1.3
1CHO(E:I)	5CHA(A)	α-Chymotrypsin	2OVO	Ovomucoid third domain	10	0.95	1.1	1.8
1ACB(E:I)	5CHA(A)	α-Chymotrypsin	1CSE(I)	Eglin-C	10	1.0	0.79	1.8
1CSE(E:I)	1SCD	Subtilisin Carlsberg	1ACB(I)	Eglin-C	10	1.0	0.79	1.8
1MAH(A:F)	1MAA(B)	AChe	1FSC	FAS2	10	1.0	0.87	1.4
1FSS(A:B)	2ACE(E)	AChe	1FSC	FAS2	10	1.2	0.90	1.6
1TGS(Z:I)	2PTN	Trypsinogen	1HPT	Pancreatic trypsin inhibitor	10	1.6	1.8	2.5
1CGI(E:I)	1CHG	α-Chymotrypsinogen	1HPT	Pancreatic trypsin inhibitor	10	1.8	1.8	2.3

NMR structure docking targets

Complex PDB	Receptor PDB	Receptor	Ligand PDB_X	Ligand PDB_N	Ligand	Size	Ub_{xtal}	Ub_{NMR}	Ens_{min}	Ens_{max}
1AY7	1RGH(B)	Ribonuclease SA	1A19	1BTB	Barstar	30	0.62	0.86	0.57	1.4
1BRS	1A2P(B)	Barnase	1A19	1BTB	Barstar	30	0.52	0.94	0.72	1.3
1EAW	1EAX(A)	Matripase	9PTI	1PIT	Pancreatic trypsin inhibitor	20	0.62	0.97	0.76	1.7
1AK4	2CPL	Cyclophilin	1E6J	1OCA	HIV capsid	20	0.50	0.97	0.78	2.0
2KAI	2PKA(XY)	Kallikrein A	6PTI	1PIT	Pancreatic trypsin inhibitor	20	0.40	1.1	0.74	2.0
2PTC	2PTN	B-Trypsin	6PTI	1PIT	Pancreatic trypsin inhibitor	20	0.28	1.2	0.84	2.2
1KTZ	1TGK	Transforming Growth Factor - Beta	1M9Z	1PLO	Transforming Growth Factor - Beta receptor II	10	0.39	1.2	0.96	1.6
2PCC	1CCP	Cytochrome *C* peroxidase	1YCC	2HV4	Cytochrome *C*	35	0.35	1.3	0.88	2.6
2BTF*	2BTF(A)	Actin	1PNE	1PFL	Profilin	20	0.44	1.8	0.92	2.6
1BVK	1BVL(LH)	Antibody Hulys11 Fv	3LZT	1E8L	Lysozyme	50	0.38	2.1	0.86	6.4
1MLC	1MLB(AB)	IgG1 D44.1 Fab fragment	3LZT	1E8L	Lysozyme	50	0.73	2.2	1.6	4.0
1CSE	1SCD	Subtilisin Carlsberg	1ACB(I)	1EGL	Eglin-C	25	1.0	2.5	1.3	4.4
1CHO	5CHA(A)	α-Chymotrypsin	2OVO	1OMT	Ovomucoid third domain	50	1.0	2.5	1.2	2.7
1B6C	1D6O(A)	FK506 Binding Protein	1IAS	1FKR	Transforming Growth Factor - Beta receptor I	20	0.66	3.3	1.3	3.8
1ACB	5CHA(A)	α-Chymotrypsin	1CSE(I)	1EGL	Eglin-C	25	1.0	3.9	1.9	4.8

PDB codes include chain identifiers. Size is the number of conformers in each ensemble, Ub_{xtal} is the BB_rmsd of the unbound structure crystal structure to the bound crystal structure, Ub_{NMR} is the BB_rmsd of the first model in the NMR structure for NMR targets, and Ens_{min} and Ens_{max} are the minimum and maximum BB_rmsd of conformers in the ensemble, respectively. All RMSD measurements are listed in angstroms.

appropriate interacting surfaces are proximal to each other.[11] Typically <3.5% of initial starting positions are <10 Å from the native structure by C^{α} RMSD of ligand residues after superposition of the receptor (L_rmsd) with the native structure. The initial complex then undergoes 500 cycles of low-resolution docking, which consists of randomized rigid-body moves of ~1 Å each, followed by a Metropolis acceptance criterion. During this phase, side chains are represented rigidly as centroid pseudo atoms.

To capture conformer selection (CS) in binding, we added a step in the low-resolution stage (Fig. 2, green box) of the algorithm immediately following the rigid-body move and preceding the Metropolis step. Following a rigid-body move, the entire ensemble of conformers is superposed along the interface residues of the current conformer (Fig. 1c). The centroid-mode binding energy is calculated for each conformer and used to generate a partition function. A conformer is selected from the ensemble to replace the current conformer based on its Boltzmann-weighted probability within the partition function. Once a conformer is selected, the Metropolis criterion is applied on the combined rigid-body/CS move. The use of a partition function in this manner allows for robustness of this method towards both ensemble size and heterogeneity, compared to a simple random selection. CS is restricted exclusively

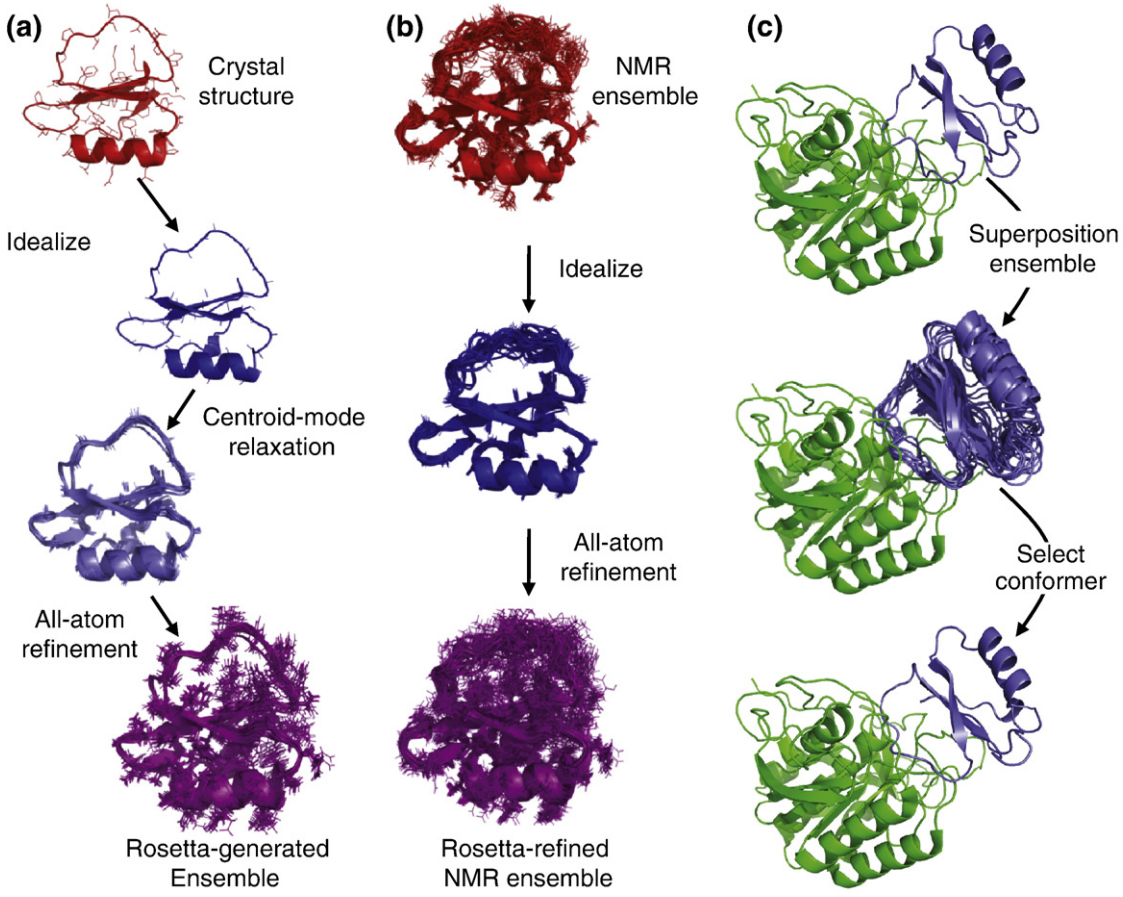

Fig. 1. Ligand ensemble generation and CS. (a) Crystal structures are idealized and then relaxed at low and high resolutions to generate an ensemble of 10 structures. (b) Conformers in an NMR ensemble are idealized and refined at high resolution only. (c) During low-resolution docking, conformers are superposed along the current conformer's interface residues, and a conformer is selected from a partition function using Boltzmann-weighted energies.

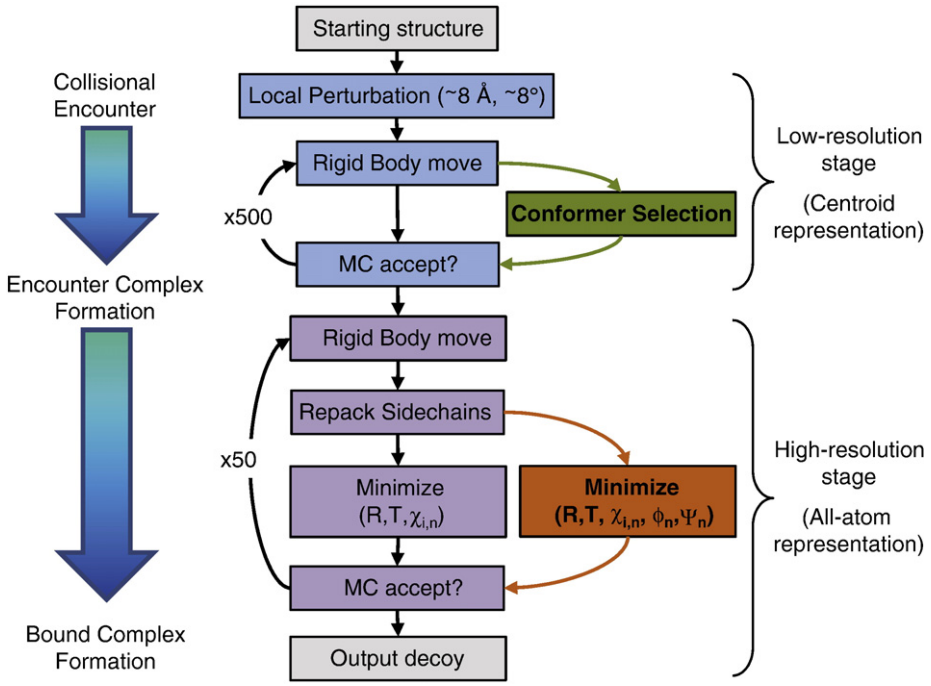

Fig. 2. The flexible docking algorithm. The low-resolution phase models the formation of an encounter complex, and the high-resolution phase models its transition to a bound complex. CS and CS/IF docking include the CS step (green box). IF and CS/IF docking include backbone minimization (orange box).

to the low-resolution phase to remain consistent with the theory that the conformers in the unbound ensemble represent distinct low-energy structures that are the result of thermal fluctuations at time scales much greater than those of a binding event.

The lowest-energy complex sampled during the low-resolution phase serves as a putative encounter complex and is converted to a high-resolution structure through a side-chain packing step. The complex then undergoes 50 steps of high-resolution refinement consisting of rigid-body moves of ~0.1 Å and periodic combinatorial side-chain packing, followed by energy-gradient-based minimization of rigid-body orientation (**R**,**T**) and side-chain torsion angles ($\chi_{i,n}$). To capture IF in binding, we use a previously developed[26] extended minimization scheme (Fig. 2, orange box) that includes the backbone torsion angles (φ_n,ψ_n), allowing the backbone conformation of the ligand to respond to local energy gradients along the interface created by the docking process. The lowest-energy structure in the high-resolution stage is selected as the final complexed structure for that iteration of the docking algorithm (hereinafter referred to as a 'decoy'), the entire algorithm is repeated to generate 1000 decoys, and all decoys are reranked by all-atom binding energy. To test the four kinetic models (KL, CS, IF, and CS/IF), we simply include or skip the appropriate CS or IF step in the algorithm.

The computational cost of a docking method restricts not only user accessibility but also the extent of conformational sampling, which is often a limiting factor in predictive docking. Using a 1.5-GHz processor, the standard RosettaDock algorithm requires 1.5–5 min to generate a single decoy, depending on the size of the protein partners. The CS method requires time approximately 1.5–2 times longer per decoy for an ensemble of 10 conformers, representing a 5-fold increase in efficiency over exhaustive docking of each conformer in the ensemble to the receptor. The IF method requires time approximately 4–6 times longer per decoy than standard RosettaDock, while the CS-IF method requires time approximately 5–7 times longer. Therefore, some consideration must be placed on the computational costs of each method, along with overall docking performance, when evaluating the methods in this study.

Energy function and discrimination

Energetic discrimination of near-native decoys remains a persistent challenge in flexible docking across FFT-based,[18] MD-based,[17,23] and MC-based methods,[26] due in part to the interplay between intramolecular and intermolecular energies. Previous studies on docking have used either total energy of a complex, intermolecular energies between two partners in a complex, or binding energy of a complex, which measures the difference in energy between the bound complex and the unbound state.[17,18,23,24,26] Although rigid-body docking with standard RosettaDock uses total energy, the variation in the backbone conformation in flexible docking can lead to significant changes in intramolecular energies, which can alter both sampling and discrimination. We therefore compared the use of total energy and binding energy for CS and final decoy discrimination.

Figure 3a shows the observed frequency that a particular conformer was selected in the low-resolution phase of CS docking when using total energy or binding energy in generating the partition function in the CS step, as a function of its unbound energy. Using total energy in the CS step leads to a large bias towards low-energy conformers in the ensemble. While thermodynamically correct, it depends on the accuracy of the relative free energies of the different conformers and prevents adequate sampling of the entire ensemble. By contrast, the use of binding energy leads to a more even distribution between conformers. Unexpectedly, CS is slightly biased towards higher-energy conformers when using binding energy compared to total energy, possibly due to the nonspecific burial of hydrophobic residues.

To illustrate the effects on near-native discrimination, Fig. 3b charts 1000 decoys in docking funnel plots (energy *versus* L_rmsd to the native structure)

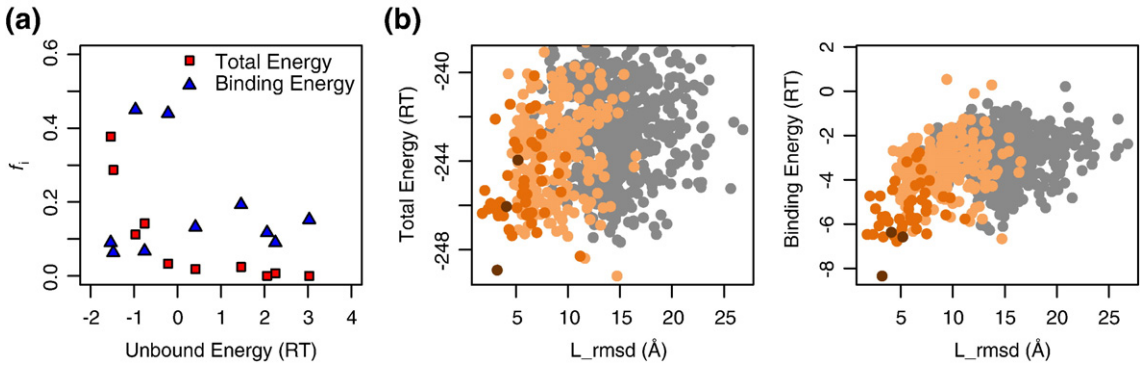

Fig. 3. (a) The frequency of selecting a particular conformer during the CS step *versus* the internal energy for each conformer when using total energy (red) and binding energy (blue) for 1BRC. (b) Total energy *versus* L_rmsd and binding energy *versus* L_rmsd for 1BRC. CAPRI criteria: high-quality decoys are shown in brown, medium-quality decoys are shown in orange, and acceptable-quality decoys are shown in tan; note that high-quality decoys sometimes meet the CAPRI I_rmsd criterion rather than the L_rmsd criterion.

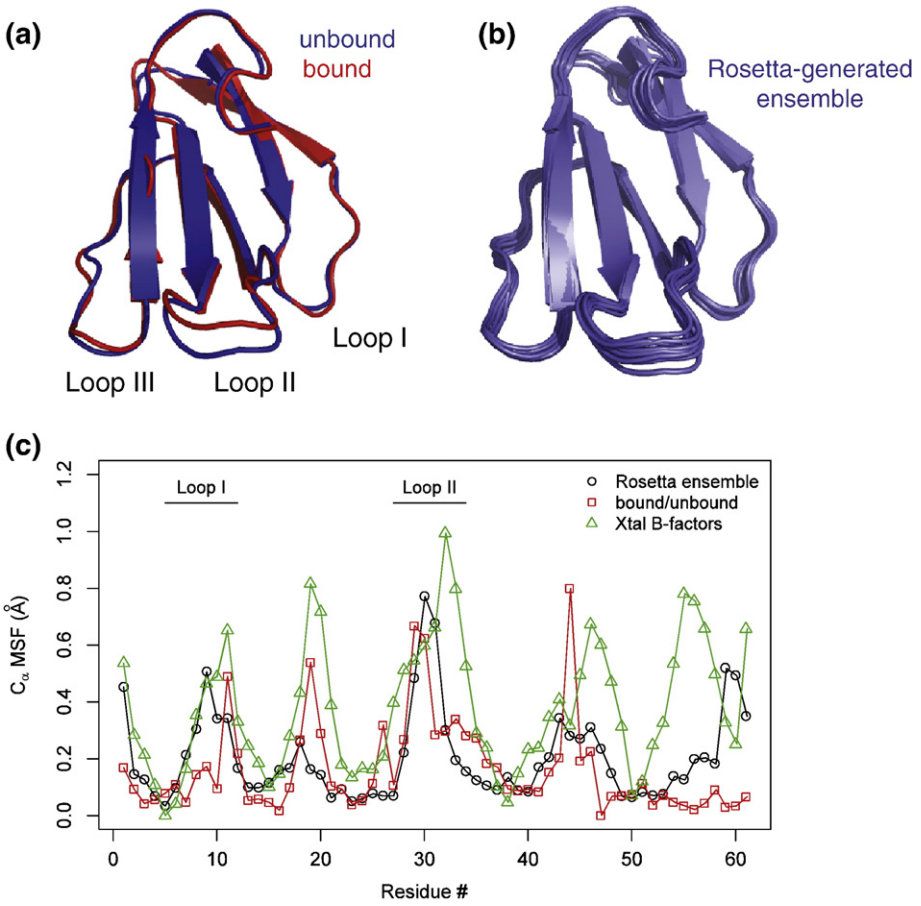

Fig. 4. Backbone variability of FAS2. (a) Bound (red) and unbound (blue) FAS2 conformations. (b) Ensemble generated by Rosetta from the unbound FAS2 conformation. (c) C_α MSF for the Rosetta ensemble (black circles), the C^α MSF between the bound and the unbound FAS2 conformation (red squares), and the C^α MSF calculated from the crystallographic *B*-factors[43] from 1FSC (green triangles).

created by the CS method for the same target. The degree of near-native discrimination in a docking funnel can be quantified by the difference in average energy of near-native decoys and the average energy of non-native decoys, normalized by the standard deviation of the energy of non-native decoys (hereinafter referred to as *Z*-score).[39] *Z*-scores of −1.2 and −2.6 for the total energy and binding energy, respectively, show that the use of binding energy yields significantly improved discrimination. Since binding energy avoids accurate thermodynamic calculation of free energies of the conformers, the starting conformers must be physically realistic and of comparable energies.

Case study: Acetylcholinesterase (AChe)–fasciculin II (FAS2)

To illustrate the results of the four docking algorithms in detail, we analyze the representative case of the AChe–FAS2 complex [Protein Data Bank (PDB)[40] code 1FSS] before describing the results of the entire target set. There is a change in conformation of the interface residues of FAS2 from the unbound state to the bound state that impedes high-quality prediction using current docking methods. In global docking of 1FSS, neither Li *et al.*, using ZDOCK + RDOCK, nor Smith *et al.*, using an FFT-based ensemble docking method, were able to generate a prediction that was within 2.5 Å C^α RMSD of interface residues (I_rmsd) among their 10 top-ranked decoys.[9,18] In local docking, Wang *et al.* used both rigid and flexible backbone docking in RosettaDock, but did not produce a decoy within 2.0 Å I_rmsd in the top three predictions.[26]

FAS2 features a three-finger structural motif common to a number of toxins. A visual comparison of the bound crystal structure[41] with the unbound crystal structure[42] reveals a ~2 Å movement of loop II (Fig. 4a). Computer simulations have provided

Fig. 5. Binding energy *versus* L_rmsd, binding energy *versus* f_{nat}, and the top-ranked decoy (blue) superposed along the receptor (green) with the crystal structure of the bound ligand (red) for AChe binding to FAS2 (1FSS). (a) KL docking using the bound FAS2 structure. (b–e) KL, CS, IF, and CS/IF docking, respectively, using the unbound FAS2 structure. The unbound AChe structure (1FSC) is used in all cases. CAPRI criteria: high-quality decoys are shown in brown, medium-quality decoys are shown in orange, and acceptable-quality decoys are shown in tan.

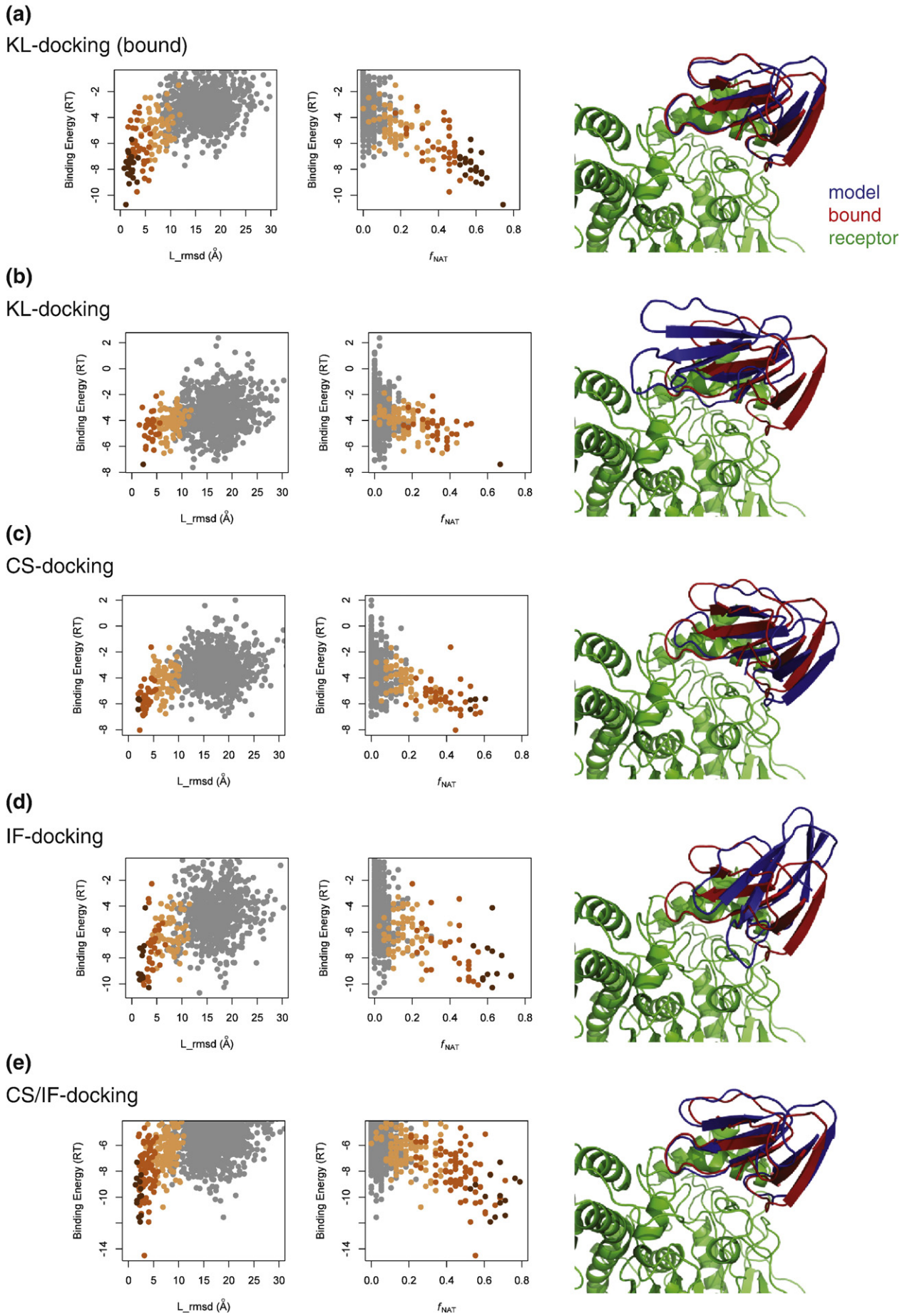

Fig. 5 (*legend on previous page*)

evidence for both conformational gating[44] and an IF mechanism[45] of binding to AChe. The ensemble representing the unbound state of FAS2 that was generated from the unbound crystal structure contained conformational variability in all three loops, but most noticeably in loop II (Fig. 4b). A quantitative analysis of the heterogeneity in the ensemble comparing the mean square fluctuation (MSF) of C^{α} position for the ensemble with the C^{α} MSF between the bound and the unbound conformations and the C^{α} MSF calculated from the crystallographic *B*-factors[43] of the unbound structure (Fig. 4c) reveals good qualitative agreement. The most flexible regions are loop II (residues 27–34), loop I (residues 5–12), and the turn region between loops I and II (residues 15–24). Furthermore, NMR data of the solution-state dynamics of the closely related toxin-α[46] also follow the diversity of the FAS2 ensemble generated using Rosetta. Broad agreement with these diverse measurements of protein flexibility suggests that conformational heterogeneity in the Rosetta ensembles qualitatively reflects inherent flexibility of the protein, validating their use as a representation of the unbound state. In fact, four conformers in the ensemble were closer to the bound state than to the unbound state, with BB_rmsd of 0.90 Å, 0.94 Å, 1.06 Å, and 1.10 Å compared to a BB_rmsd of 1.15 Å for the unbound crystal structure.

To capture the results using the four docking methods, Fig. 5 presents the lowest-energy structure and docking funnel plots for docking FAS2 to the unbound AChe structure as a function of both ligand RMSD (L_rmsd) and the fraction of native contacts recovered (f_{nat}). The upper bound of accuracy is represented by KL docking using the *bound* crystal structure of FAS2 (Fig. 5a), which produces the lowest-energy structure with a CAPRI accuracy rating of high quality, an L_rmsd of 1.1 Å, and an f_{nat} of 0.74. In contrast, KL docking using the *unbound* crystal structure of FAS2 (Fig. 5b) does not produce a docking funnel towards the bound complex and contains numerous false positives, or non-native structures with low energy, including the lowest-energy structure. CS docking (Fig. 5c) produces a lowest-energy structure of medium quality with an L_rmsd of 2.3 Å and an f_{nat} 0.45. IF docking (Fig. 5d) shows a more pronounced docking funnel than the KL method, but also contains a number of false positives, including the lowest-energy structure. Finally, like the CS docking, CS/IF docking (Fig. 5e) produces a lowest-energy structure of medium quality, with an L_rmsd of 2.5 Å and an f_{nat} of 0.61.

The high level of accuracy of the interface in the lowest-energy structure from the CS/IF method is illustrated in Fig. 6, where both the position of FAS2 relative to AChe and the side-chain orientations on both partners are recovered closely. Hydrogen-bond donor atoms from one partner are within 4 Å of hydrogen-bond acceptor atoms from the other partner for 9 of the 15 hydrogen bonds; the unusual hydrophobic stacking interaction between Met33 of FAS2 and Trp279 of AChe[41] is recovered; and, overall, 61% of the native residue–residue contacts are satisfied. The only major interactions not recovered in the model compared to the bound structure are the polar and hydrophobic interactions between the C-terminal Tyr61 of FAS and Lys341, Pro76, and Phe75 of AChe.

The significant difference in docking accuracy achieved by the docking algorithms is a result of their respective treatments of backbone flexibility. Therefore, it is useful to analyze how the different methods affect the backbone conformation of the ligand during docking. Figure 7a demonstrates the breadth and density of backbone conformational

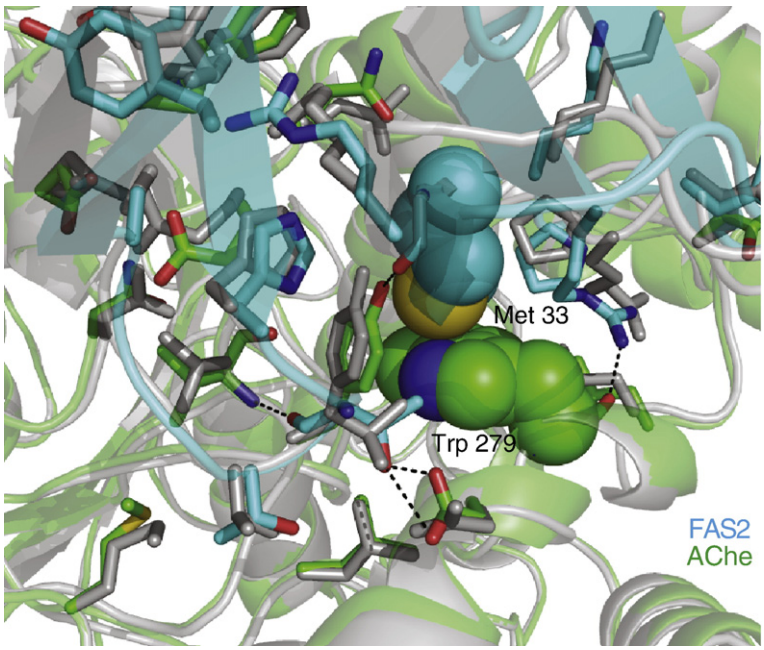

Fig. 6. Details of the lowest-energy structure ligand (cyan) and receptor (green) for the CS/IF method for 1FSS superposed with the native structure (gray) along the receptor. This decoy has an L_rmsd of 2.5 Å, an I_rmsd of 0.94 Å, and an f_{nat} of 0.61. Met33 of FAS2 and Trp239 of AChe are shown as spheres.

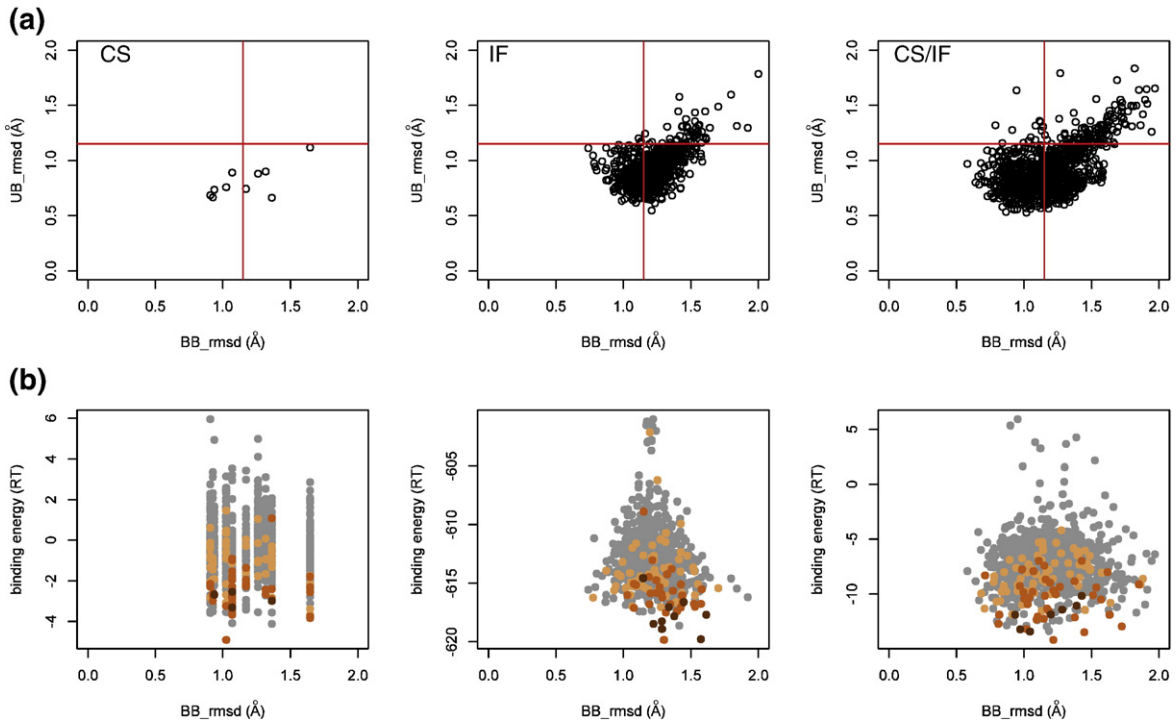

Fig. 7. Backbone sampling and discrimination. (a) The BB_rmsd and UB_rmsd in each decoy output by CS, IF, and CS/IF docking, respectively. Red lines show the BB_rmsd and UB_rmsd of the unbound and the bound conformations, respectively. (b) Binding energy *versus* BB_rmsd for CS, IF, and CS/IF docking, respectively. CAPRI criteria: high-quality decoys are shown in brown, medium-quality decoys are shown in orange, and acceptable-quality decoys are shown in tan.

space being sampled by showing the BB_rmsd and UB_rmsd (C^{α} RMSD of interface residues to the unbound structure) for all 1000 decoys for the flexible docking methods. KL docking samples the single unbound backbone conformation that is 1.15 Å from the bound structure (red line for comparison), while CS docking samples 10 distinct backbone conformations, ranging from 0.90 Å to 1.64 Å from the bound structure. Since both IF and CS/IF docking include explicit backbone flexibility, each backbone conformation generated from docking is a unique result of the stochastic docking process. As a result, these methods sample a wide variety of backbone conformations. Although IF docking samples nearly as wide a range of conformations as CS/IF docking (0.75–2.0 Å compared to 0.63–1.93 Å), a large fraction of these conformations was closer to the unbound state, indicating that it is unable to overcome conformational and energetic barriers to move towards the bound state. Overall, the FAS2 conformations from CS and CS/IF docking were closer to the bound conformation than the unbound crystal structure in 40% and 42% of decoys, respectively, compared to 22% in IF docking.

Ideally, to achieve appropriate energetic discrimination of near-native complexes in flexible docking, an energy funnel should exist not only in the rigid-body conformational space towards the bound docking orientation of the two partners (measured by binding energy *versus* L_rmsd) but also in the backbone conformation space towards the bound backbone conformation of both partners (measured by binding energy *versus* BB_rmsd). Although in this study the receptor was kept fixed in the unbound conformation, an energy funnel towards the bound conformation of the ligand may still be observed. Figure 7b shows the binding energy and the BB_rmsd for all decoys from CS, IF, and CS/IF docking with near-native decoys in tan, orange, and brown for acceptable-quality, medium-quality, and high-quality predictions, respectively (compare to the energy funnel in rigid-body conformation space; Fig. 5c–e). Docking decoys close to the native docked orientation should have lower energies for conformers closer to the bound conformation. Overall, a distinct energy funnel is not observed in the dimension of BB_rmsd for any of the three flexible docking methods, although in both CS and CS/IF docking, the decoy with the lowest binding energy also had a relatively low BB_rmsd (1.0 Å and 1.2 Å, respectively). The lack of an energy funnel in backbone conformation space could be due either to sampling of too few backbone conformations or to a deficiency in energy function in discriminating bound-like conformers during docking. Alternately, differences between the bound receptor conformation and the unbound receptor conformation (which is kept fixed in the unbound form) could force the ligand to adopt a binding-competent conformation slightly away from the bound conformation, obscuring or eliminating an energy funnel in this relatively narrow range of conformation space (BB_rmsd of 0.9–1.3 Å).

Table 2. Results summary for crystal structure targets

	KL(B)					KL					CS					IF					CS/IF				
PDB	N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality	N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality	N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality	N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality	N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality
2SIC	10	0.85	3.9	0.51	***	10	0.72	1.8	0.38	***	10	0.74	3.6	0.59	***	8	0.67	3.4	0.52	***	10	0.74	3.9	0.52	***
1MAH	10	0.64	0.23	0.13	***	3	0.78	2.3	0.66	***	8	0.59	1.6	0.65	***	6	0.64	2.4	1.1	**	6	0.69	3.7	1.4	**
1CHO	10	0.75	0.76	0.31	***	10	0.53	1.8	0.78	***	5	0.81	1.6	0.47	***	6	0.39	6.3	1.5	**	5	0.54	7.7	1.7	**
2PTC	10	0.60	2.7	0.58	***	10	0.48	3.5	0.86	***	10	0.43	4.3	1.0	**	10	0.55	3.0	0.68	***	10	0.62	2.3	0.75	***
2BTF	6	0.86	1.0	0.38	***	8	0.44	2.6	1.2	**	7	0.45	3.0	1.9	**	4	0.41	2.2	1.5	**	9	0.69	1.9	0.67	***
1ACB	10	0.68	0.88	0.40	***	6	0.38	6.1	1.3	**	10	0.51	3.3	0.84	***	10	0.57	2.9	1.1	**	10	0.60	3.8	0.82	***
1BRC	10	0.36	9.2	2.0	**	9	0.43	4.2	1.5	**	9	0.58	3.2	0.81	***	9	0.57	2.8	0.72	***	4	0.36	8.2	1.1	**
2SNI	10	0.76	0.91	0.25	**	8	0.38	5.4	1.5	**	10	0.51	5.7	1.2	**	7	0.40	6.1	1.6	**	10	0.77	2.3	0.82	***
1BQL	7	0.41	3.1	1.1	**	9	0.39	5.1	1.6	**	10	0.74	5.2	1.3	**	6	0.07	22	11		6	0.33	3.6	1.4	**
1UGH	10	0.70	0.43	0.19	***	6	0.37	5.2	2.5	*	3	0.23	5.6	2.8	*	9	0.30	3.4	1.9	**	5	0.31	5.0	2.5	**
1WQ1	4	0.39	4.9	1.5	**	0	0.17	5.8	3.7	*	0	0.26	3.6	2.1	*	0	0.09	6.9	4.1		0	0.07	6.5	4.2	
2KAI	6	0.72	2.5	0.47	***	0	0.04	10	4.5		4	0.03	8.4	3.2	*	6	0.50	5.9	1.2	**	10	0.58	2.9	0.69	***
2JEL	8	0.79	0.74	0.26	***	7	0.10	9.3	4.6	*	7	0.77	1.2	0.43	***	10	0.68	1.3	0.46	***	9	0.56	1.6	0.55	***
1BVK	0	0.11	20	5.9		0	0.25	9.7	4.7	*	0	0.12	30	10		0	0.08	19	9.2		0	0.24	10	4.3	
1CSE	10	0.64	1.3	0.36	***	0	0.00	13	6.0		5	0.32	2.7	1.2	**	2	0.13	18	7.4		2	0.03	13	5.5	
1DFJ	0	0.00	12	7.1		0	0.00	19	7.3		0	0.00	22	11		0	0.00	25	12		0	0.00	18	8.5	
1FSS	10	0.74	1.1	0.43	***	1	0.05	12	7.9		5	0.45	2.3	1.2	**	6	0.00	13	8.1		7	0.55	3.2	1.4	**
1TGS	10	0.74	1.7	0.50	***	5	0.09	15	9.4		10	0.64	2.4	0.87	***	7	0.12	16	9.3		7	0.55	4.9	2.4	**
1CGI	3	0.69	1.2	0.42	***	0	0.06	17	9.5		0	0.08	13	5.0		0	0.16	13	5.8		0	0.28	19	10	
1MLC	4	0.08	19	9.4		2	0.71	23	11		1	0.14	20	9.1		3	0.48	2.3	1.1	**	2	0.14	20	9.2	
1BRS	4	0.09	18	9.5		1	0.03	19	11		8	0.00	18	9.5		9	0.43	3.0	1.5	**	5	0.06	19	11	
Totals					17 (15)					9 (8)					13 (13)					13 (11)					14 (13)

KL(B) is KL docking using the bound ligand, and KL, CS, IF, and CS/IF are the four docking methods using the unbound ligand. N_{10} is the number of structures among the 10 lowest-energy structures that are of at least medium quality. The f_{nat}, L_rmsd, I_rmsd, and CAPRI ratings are for the lowest-energy structure produced. Totals show the number of targets for which the lowest-energy structure was of at least medium quality and, in parenthesis, the number of hits for each method (hit is defined as having a lowest-energy structure of at least medium quality and $N_{10}>4$). Targets are sorted by I_rmsd in the KL(Ub) case. CAPRI ratings are acceptable (*), medium quality (**), and high quality (***).

Flexible docking methods outperform rigid-body docking

We applied the four docking methods to a set of 21 target complexes using crystal structures for both the ligand and the receptor. A docking method was said to produce a 'hit' if it passed two criteria: (1) the lowest-energy structure was of at least medium quality, and (2) at least 5 of the 10 lowest-energy structures were of at least medium quality. This represents a relatively strict criterion in which a docking run is deemed successful only if its lowest-energy structure is of good quality and it has converged on that solution. We also performed KL docking of the bound ligand with the unbound receptor to serve as a control.

Table 2 shows the number of medium-quality or high-quality decoys within the 10 top-scoring structures (N_{10}) and the L_rmsd, I_rmsd, and f_{nat} of the top-scoring decoy. Figure 8a summarizes these results as a histogram of hit quality, and funnel plots for each method for several selected targets are presented in Supplementary Fig. 1. KL docking with the bound ligand structure produced 15 total hits (12 of high quality and 3 of medium quality) and represents the upper bound of results achievable by modeling backbone flexibility in the ligand. KL docking using the unbound crystal structure produced 8 total hits (3 of high quality and 5 of medium quality), demonstrating the potential for improvement in modeling flexibility in the ligand backbone conformation.

As seen in Fig. 8a, CS/IF docking came closest to the quality of KL docking, with the bound ligand conformation producing 13 total hits (7 high-quality predictions and 6 medium-quality predictions). CS docking performed comparably to the CS/IF docking with 13 total hits (7 high-quality predictions and 6 medium-quality predictions, respectively). IF docking produced 11 hits (with 4 high-quality predictions and 7 medium-quality predictions, respectively). At least one of the flexible docking methods produced a hit in 8 of the 13 cases in which KL docking did not; in the 8 cases where KL docking did produce a hit, in 4 cases, at least one of the flexible docking methods produced a higher-quality hit. Overall, there was substantial overlap in the improvement in performance of the flexible docking methods compared to KL docking, with 8 of 21 cases in which two or three of the flexible docking methods produced a top-ranked decoy of a higher CAPRI quality than KL docking (e.g., 2JEL in Table 2; funnel plots in Supplementary Fig. 1).

Assessing ensemble docking and explicit backbone minimization

In a comparison between the flexible docking methods within each target, the two ensemble docking methods (CS and CS/IF) combined to produce 14 top-ranked decoys that were highest in f_{nat}, compared to 5 top-ranked decoys produced by the nonensemble dock methods (KL and IF). Furthermore, the CS and CS/IF methods had, overall, a greater number of hits and a higher accuracy of hits. The larger breadth of conformations sampled by the CS methods is critical to the success of the flexible docking methods, even in complexes with relatively small conformation changes.

In a number of cases, an improvement in docking performance was observed despite the fact that there were no conformers in the ensemble that were closer to the bound structure than to the unbound structure. Although initially counterintuitive, this observation agrees with the findings of both Smith *et al.* and Grunberg *et al.* and suggests that there is inherent value in allowing backbone conformational variability in docking, irrespective of whether the bound conformation is being sampled.[18,19] Even a small relatively homogenous ensemble of 10 structures can make a significant difference in docking compared to a single structure.

The explicit backbone minimization methods (IF and CS/IF) produced greater sampling of higher f_{nat} decoys, as seen in the example of 1FSS (Fig. 4). Similarly, Krol *et al.* observed increased f_{nat} when refining near-native decoys using MD, possibly as the result of the formation of a greater number of energetically favorable contacts along a putative docking interface when minimizing or relaxing.[23] The increase in sampling is not clearly reflected in the overall results, however, because discrimination is more difficult. In the case of 1CSE, the top-ranked decoy is near-native in the CS method, but not in the CS/IF method, despite the fact that both methods use an identical ensemble of ligand conformations. Explicit backbone minimization improves the binding energy across all decoys, and, in this case, the

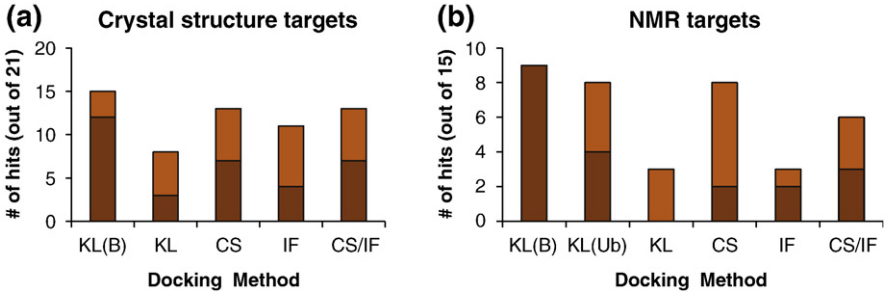

Fig. 8. Histogram of hit quality (quality of the top-ranked decoy for all runs, with at least 5 of the 10 top-scoring decoys being of medium or high quality) for each docking method for (a) crystal structure targets and (b) NMR targets. CAPRI criteria: high-quality decoys are shown in brown, and medium-quality decoys are shown in orange.

binding energy of a non-native docking orientation is improved more than the binding energy of the near-native docked orientation, leading to a false positive. The loss of a near-native lowest-energy structure is also observed in two cases (1BQL and 1WQ1) when comparing KL docking with IF docking. In summary, except in cases where explicit backbone minimization results in a false positive, it generally improves the quality of the top-ranked decoy. Indeed, where CS and CS/IF docking both produced a hit, backbone minimization in CS/IF docking leads to an increase in f_{nat} in 8 of 12 cases.

Ensemble docking with NMR targets

We applied these docking methods to a set of 15 targets in which the unbound ligand structure is an NMR ensemble. A comparison of the BB_rmsd of both the unbound crystal structure and the closest conformer in the NMR ensemble to the bound crystal structure in Table 1 reveals a significant degree of variability in the NMR ensemble. Still, the closest NMR conformer to the bound conformation is often substantially closer than an arbitrary conformer in the ensemble, providing a means for overcoming the structural uncertainties.

Table 3 shows the docking accuracy for the four methods using the NMR structure of the ligand, with rigid-body docking results with the unbound and the bound crystal structures provided for comparison. Figure 8b summarizes these results as a histogram of hit quality, and Supplementary Fig. 2 provides docking funnels for each target across all the methods. For KL and IF docking, the starting structure was the first conformer in the NMR structure coordinate file from the PDB (Model 1). For the CS and CS/IF methods, the entire NMR ensemble was used.

KL docking with the bound ligand conformation produced nine hits, all of high quality. KL docking with the unbound ligand conformation produced seven hits (four of high quality and three of medium quality). In contrast, KL docking with the NMR ensemble Model 1 produced three hits, all of medium quality, demonstrating the difficulties inherent in docking NMR structures compared to crystal structures.

Overall, while none of the flexible docking methods was able to approach the accuracy of docking with the bound ligand conformation, CS docking approached the results of KL docking with the unbound crystal structure, producing eight hits (two of high quality and six of medium quality). CS/IF docking produced six hits (three of high quality and three of medium quality), while IF docking produced three hits (two of high quality and one of medium quality).

Two trends can be observed from the data. First, in contrast to the crystal structure docking targets, with one exception, IF docking only produces a hit if KL docking produces a hit. When IF docking does produce a hit, it generally improves the accuracy of the hit, for example in 2PTC, where IF docking is able to produce a high-quality hit where KL docking produces a medium-quality hit. A potential explanation is that the single backbone methods (KL and IF) only produce hits in cases where the NMR structure is close to the bound conformation: for all three cases in which the IF docking produced a hit, the first conformer in the NMR ensemble Model 1 has a BB_rmsd in the lower range among NMR targets, at 0.94 Å, 1.12 Å, and 1.16 Å for 1BRS, 2KAI, and 2PTC, respectively.

Second, the ensemble docking methods perform significantly better than the single-backbone methods, especially for targets with a BB_rmsd of the first conformer in the NMR ensemble of >1.2 Å. As a representative example, we examine the docking of α-chymotrypsin with the NMR solution structure of

Table 3. Results summary for NMR targets

	KL(B)					KL(Ub)					KL				
PDB	N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality	N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality	N_{10}		L_rmsd	I_rmsd	CAPRI Quality
1KTZ	10	0.80	3.0	0.50	***	9	0.85	1.9	0.47	***	2	0.03	25	9.3	
1EAW	5	0.40	3.4	1.2	**	6	0.50	2.3	0.69	***	2	0.25	8.1	2.6	*
1CHO	10	0.75	0.76	0.31	***	10	0.53	1.8	0.78	***	8	0.22	10	2.2	*
2PTC	10	0.60	2.7	0.58	***	10	0.48	3.5	0.86	***	6	0.43	5.4	1.4	**
1CSE	10	0.64	1.3	0.36	***	8	0.57	3.6	1.0	**	0	0.07	15	7.3	
2BTF*	5	0.86	1.0	0.38	***	8	0.44	2.6	1.2	**	5	0.03	21	13	
1ACB	10	0.68	0.88	0.30	***	6	0.38	6.1	1.3	**	0	0.18	10	2.9	*
1B6C	9	0.70	2.3	0.56	***	7	0.43	8.6	1.4	**	0	0.07	22	13	
1AK4	10	0.70	4.6	0.60	***	1	0.27	19	2.5	*	0	0.00	24	9.3	
2PCC	0	0.00	27	14		1	0.27	8.7	3.7	*	0	0.06	14	5.5	
2KAI	6	0.72	2.5	0.47	***	0	0.04	10	4.5		5	0.47	6.6	1.3	**
1BVK	0	0.11	20	5.9		0	0.25	9.7	4.7	*	1	0.24	16	7.6	
1AY7	6	0.73	3.2	0.86	***	0	0.10	16	7.7		0	0.00	20	9.4	
1MLC	4	0.08	19	9.4		2	0.07	23	11		0	0.08	26	9.7	
2BRS	4	0.09	18	9.5		1	0.03	19	11		7	0.63	2.1	1.0	**
Totals					11 (11)					8 (8)					3 (3)

KL(B) is rigid-body docking using the bound structure, KL(Ub) is rigid-body docking using the unbound crystal structure, and KL, CS, IF, and CS/IF are the four docking methods using the ligand solution-state NMR structure. All column descriptions are the same as in Table 2. Targets are sorted by I_rmsd in the KL(Ub) case. CAPRI ratings are acceptable (*), medium quality (**), and high quality (***).

eglin-C (1ACB). CS docking produces a moderate docking funnel (Fig. 9a) with a *Z*-score of −0.95 and a medium-quality hit. Figure 9b shows the binding energy plotted against the BB_rmsd for each decoy, which, in contrast to the example of 1FSS (Fig. 7b), shows a pronounced energy funnel towards the bound backbone conformation. Out of the considerable diversity in the entire 20-model NMR ensemble of eglin-C (Fig. 9c), CS docking successfully selects the conformer closest to the bound conformation (BB_rmsd of 1.9 Å; Fig. 9d and e). In contrast, the first model in the NMR ensemble used in the (unsuccessful) KL and IF docking is much farther from the bound conformation (BB_rmsd of 3.9 Å; Fig. 9d). Similar results are observed for other cases (e.g., 1EAW and 1CSE) and also for CS/IF docking (funnel plots in Supplementary Fig. 2).

The issues with near-native discrimination observed in docking crystal structures when using IF were present in the results for NMR targets as well. In 1CHO, CS docking produced a medium-quality hit, while CS/IF docking failed to produce even a single medium-quality decoy among the 10 top-scoring decoys. Likewise, CS docking produced a medium-quality hit for 2BTF, while the top-scoring decoy from CS/IF docking was of acceptable quality and only 2 of the top 10 decoys were of medium or better quality. In a surprising number of cases, the top-ranked decoy from CS/IF docking was of acceptable quality. In three cases (1KTZ, 1CHO, and 2PCC), CS/IF docking appears to have converged at the solution, producing at least 5 of the top 10 decoys at acceptable quality. For 1KTZ and 2PCC, this represents an improvement over the other docking methods; for 1CHO, it is a decrease in accuracy compared to CS docking.

Interestingly, in 4 of 15 cases, CS/IF docking produced a top-ranked decoy that was of higher accuracy than KL docking with the unbound crystal structure; in another 5 cases, it performed equally well. In 2 of those 4 cases (2KAI and 2BRS), KL docking using just a single conformer from the NMR ensemble outperformed KL docking with the unbound crystal structure, suggesting that the improvement in docking is a result of differences between all NMR conformers and the crystal structure, and not the use of multiple NMR conformers. Nonetheless, excluding those 2 cases, CS/IF docking performs equal to or better than rigid-backbone docking with the crystal structure in 8 of 13 cases. By using an ensemble docking with minimization method, it is possible for the first time to locally dock NMR structures with comparable accuracy to rigid-body docking with crystal structures.

Discussion

CS and IF are two fundamentally distinct kinetic mechanisms for protein binding, and here we attempt to model them as conformational search strategies within RosettaDock. RosettaDock's MC algorithm samples different thermodynamic configurations of the system irrespective of time, limiting the conclusions we can draw about the physical validity of these two mechanisms. However, the success of these distinct conformational search strategies in recapitulating the binding models that inspired their design can be addressed. The success of CS docking using both computationally generated ensembles and NMR ensembles suggests that binding-competent conformers do exist in the unbound ensemble, as demonstrated by Grunberg *et al.*, and that these conformers can be selected from the ensemble based on their favorable binding energies, both central tenets of the CS binding model.[19] Likewise, success of IF docking through energy-

CS					IF					CS/IF				
N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality	N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality	N_{10}	f_{nat}	L_rmsd	I_rmsd	CAPRI Quality
0	0.29	6.3	4.2	*	1	0.46	4.9	1.8	**	1	0.30	4.8	4.1	**
6	0.38	4.8	1.8	**	7	0.04	10	5.0		8	0.35	3.7	1.5	**
5	0.39	8.9	1.9	**	1	0.44	9.0	2.1	*	0	0.48	9.8	2.2	*
10	0.54	5.5	1.1	**	10	0.35	3.3	1.0	***	9	0.57	4.0	0.75	***
3	0.03	13	5.6		1	0.16	14	6.8		2	0.26	5.6	1.8	**
8	0.48	2.6	1.7	**	0	0.10	15	8.8		2	0.30	6.4	3.1	*
10	0.25	5.1	1.7	**	0	0.00	19	6.8		8	0.31	4.0	1.7	**
0	0.13	13	5.2		0	0.07	16	7.7		0	0.07	14	4.9	
0	0.02	28	7.6		0	0.21	15	3.6	*	0	0.18	15	3.7	*
0	0.00	20	9.3		0	0.38	7.7	3.1	*	0	0.25	8.5	3.4	*
9	0.55	4.5	0.90	***	9	0.55	4.3	0.9	***	8	0.50	5.2	1.1	**
0	0.06	19	10		4	0.00	21	12		0	0.24	9.5	4.2	*
7	0.67	2.1	0.99	***	6	0.06	20	8.8		5	0.72	2.2	0.91	***
0	0.05	19	9.4		0	0.03	17	9.2		0	0.25	8.3	3.4	*
6	0.67	2.5	1.2	**	6	0.58	3.9	1.9	**	10	0.75	2.9	0.87	***
				8 (8)					4 (3)					8 (6)

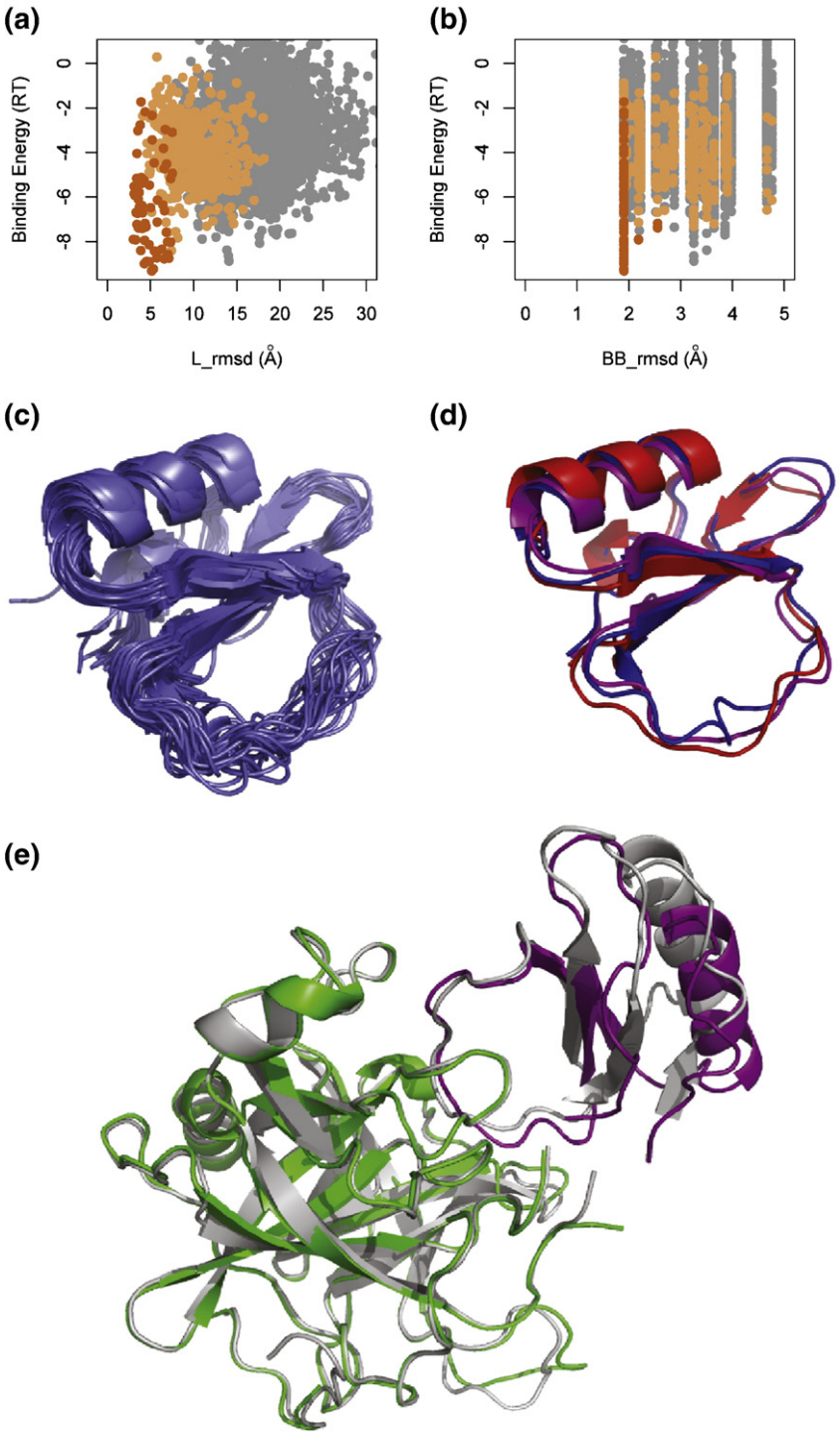

Fig. 9. NMR docking results for CS docking of 1ACB. (a) Binding energy *versus* L_rmsd. (b) Binding energy *versus* BB_rmsd. (c) Rosetta-refined NMR ensemble. (d) Lowest-energy conformer from CS docking (purple) superposed on the bound structure (red) and the first model in the NMR ensemble (blue). (e) Lowest-energy structure from the CS method (receptor, green; ligand, purple) superposed on the native complex (gray).

gradient-based backbone minimization demonstrates that a binding-competent conformation can be reached through a low-energy pathway through conformational space from the unbound structure while in complex with the partner—a necessary component in the IF binding model. Finally, the success of CS/IF docking, especially in the number of top-ranked structures of acceptable quality in the NMR targets, suggests that, for larger conformational changes, both CS over a broad range of conformational space and local IF may be necessary for binding, as proposed by both computational[19] and experimental[47] studies.

As observed in CAPRI, accommodating backbone flexibility is currently the biggest obstacle to accurate protein–protein docking. Although most of the crystal structure docking targets we tested were classified as 'rigid body' by Mintseris *et al.*, modest backbone conformational changes between the bound and the unbound states can still make high-resolution docking challenging.[48] In rigid-body global docking, Li *et al.* produced at least one

medium-quality prediction among their 10 top-ranked predictions in 9 of 21 (43%) targets that overlapped with ours.[9] In FFT-based global docking using MD-derived ensembles, Smith *et al.* produced at least one medium-quality prediction in their 10 top-ranked predictions for 4 of 12 (33%) targets that overlapped with ours.[18] Using backbone minimization in local docking with Rosetta (very similar to our IF docking), Wang *et al.* produced at least medium-quality models in their 3 top-ranked predictions in 9 of 17 (53%) targets that overlapped with ours.[26] In our study, CS and CS/IF docking produced a top-ranked medium quality prediction in 13 and 14 of 21 crystal structure targets (62% and 67%), respectively. Although direct comparisons with other studies are difficult both to differences in docking strategies (global *versus* local docking) and to the different target sets used, the general improvement in performance observed in our study may be attributed to better discrimination between conformers due to a more physically realistic scoring function and the greater breadth of conformational sampling due to the use of an ensemble of unbound structures.

Although no flexible docking method clearly outperformed the other in cases where the single unbound conformation was relatively close to the bound conformation (BB_rmsd<1.1 Å), in cases where the single conformation was relatively far from the bound conformation (BB_rmsd>1.3 Å), particularly among the NMR targets, the ensemble docking methods (CS and CS/IF) significantly outperformed the backbone minimization method (IF). Furthermore, CS docking, which was the most successful method in both target sets by some measures, is the least computationally expensive. However, the CS/IF method did show moderate improvements in some cases, particularly in increasing the number of native contacts. Therefore, one potential strategy for predictive docking would be to use CS docking first for global searches where near-native discrimination and computational efficiency are paramount. Potential near-native orientations and their respective binding-competent conformers can then be used as starting structures for further local docking and refinement with IF docking.

Our work included the first systematic study of docking using NMR solution-state structures. NMR ensembles contain conformational variation as a result of both methodological reasons, such as underdetermination of protein structure, and biophysical reasons, such as solution-state protein dynamics.[49] Our results suggest that, while a single conformer may not be sufficient for docking, the use of the entire NMR ensemble can overcome structural uncertainties in the model to a degree that is comparable to rigid-body docking of unbound crystal structure. Conformational heterogeneity in NMR structural ensembles has demonstrated qualitative agreement with a number of quantitative measures thought to be related to protein dynamics, from experimental data such as S^2 order parameters[46] and crystallographic *B*-factors,[43] to computational simulations using MD,[33,43] essential dynamics,[50] or normal-mode analysis.[51] Although, to our knowledge, a direct relationship between this heterogeneity and binding-induced conformation changes has not been studied, our results show that the conformational variation found in NMR ensembles provides meaningful backbone conformational sampling in ensemble docking and may thus, to some degree, be relevant to the conformational changes associated with binding. Still, the difference in performance between docking with the NMR ensembles and docking with the unbound crystal structures suggests that structure underdetermination plays a predominant role in the conformational heterogeneity of NMR ensembles.

These flexible docking methods were inspired by biophysical models of protein binding, but they can be applied more broadly to address uncertainty in the bound backbone conformation in docking, whether that uncertainty is the result of true flexibility (i.e., binding-induced conformational changes) or uncertainties in the initial unbound structures. Efficient methods for general sampling of conformational plasticity in proteins during docking could be applied to docking of homology model or low-resolution structures, where docking of an ensemble of multiple structural models instead of a single model may effectively increase the margin of error in structural modeling as applied to docking. We have recently demonstrated this in a separate study where docking of an ensemble of multiple antibody homology models to their respective antigens, using the CS-docking methods presented here, significantly outperforms the docking of a single homology model in recovering near-native docking solutions.[52] Likewise, HADDOCK has shown several specific successes in docking ensembles from a variety of sources from homology models to NMR structures.[30–32] Together, these results suggest a fundamental robustness of ensemble docking for accommodating conformational plasticity.

There are a number of limitations to the methods for use in predictive docking efforts such as CAPRI. First, this study analyzed local docking only, but often global docking is necessary when no biochemical information exists to assist in prediction. A test of global docking for both target sets with all four docking methods was beyond the scope of this study, but successful local docking can improve global docking. Using RosettaDock, Gray *et al.* showed that global docking succeeded in 18 of 24 cases where local docking produced a medium-quality or high-quality prediction, and in none of the 6 cases where local docking produced, at best, an acceptable-quality prediction, suggesting that the increase in medium-quality and high-quality models using local flexible docking in this study may directly translate to improvements in global docking.[11]

Second, effective energetic discrimination of near-native decoys with explicit backbone flexibility in docking remains a challenge in IF and CS/IF docking. Although the use of binding energy in this study yielded improvements in discrimination

compared to total energy, the existence of false positives still confounded accurate docking, perhaps as a result of inaccuracies in the energy function. Deficiencies in the energy function may include backbone torsion angle and other internal energy components, accurate modeling of electrostatics and side-chain protonation states, ordered water molecules at the interface, or the correct balance between scoring terms. Additionally, more sophisticated approaches to discriminating near-native funnels in rigid-body docking[53] can be extended to include variables such as backbone conformation or conformer identity.

Finally, large global conformation changes (e.g., domain hinge motions) or large local conformation changes (e.g., loops) cannot yet be accurately modeled because neither the ensemble generation method nor the backbone minimization used in this study can capture conformational changes of that magnitude. The docking methods presented here were developed to be largely independent of the method of ensemble generation, so the ensemble generation method can be tailored to accommodate specific types of conformational flexibility. Large global motions may be modeled using larger or more diverse ensembles generated from normal-mode analysis,[54] accelerated MD,[55] or essential dynamics.[50,56] Likewise, specific features of the protein structure can be varied between conformers in the ensemble (e.g., the substrate binding-loop conformation of an enzyme, the conformation of a disordered region of an unbound crystal structure, or V_H–V_L orientation in antibody docking). Previous methods addressed specific types of conformation changes such as hinge motions or domain–domain motions, with some predictive success.[57,58] Progress has been made towards the prediction of flexible regions of proteins that change conformation upon binding,[59] as well as the generation of ensembles of flexible regions that contain bound-like conformations,[60] both of which can improve the quality of the ensembles generated for docking.

Conclusion

In this study, we present three different methods for accommodating backbone flexibility in protein docking and show substantial improvements in overall docking accuracy, including sampling and discrimination, without using any *a priori* knowledge of the flexible regions of the ligand protein. These methods have been intentionally developed to be versatile and can be used in conjunction with ensembles derived from a wide variety of sources, including NMR data, MD simulations, loop ensembles, and homology modeling. Local docking using RosettaDock is featured in a number of successful CAPRI strategies;[27,28,53,61] given the substantial improvements in performance over the standard RosettaDock algorithm, the new methods are a significant step forward in the state of the art of protein–protein docking towards general flexible docking. Incorporation of backbone conformational plasticity into docking might ultimately allow us to expand the list of "dockable" components beyond high-resolution crystal structures to lower-resolution electron microscopy structures, NMR structures, and homology models, which will be essential to applying predictive docking towards a structural understanding of protein interactions in conjunction with ongoing genomic and proteomic efforts.

Materials and Methods

Structural data

We assembled two target sets to test our docking methods (Table 1). The crystal structure target set consists of 21 protein complexes from Protein–Protein Docking Benchmark 1.0,[62] with a BB_rmsd of <2.0 Å and no disordered residues in the interface region of the unbound crystal structures. The NMR structure target set consists of 15 complexes from both Protein–Protein Docking Benchmark 1.0 and Protein–Protein Docking Benchmark 2.0[48] that meet the same criteria and also have an unbound NMR structure and an unbound crystal structure in the PDB.[40]

Docking metrics

A number of metrics that are sensitive to the ligand position relative to the receptor, the specificity of the interactions across the interface, and the backbone conformation of the ligand are used to measure docking accuracy. L_rmsd is an overall measure of the ligand position and orientation with respect the receptor and is the RMSD of C^α coordinates of the ligand protein between the decoy and the native structure after superposition of the receptor. Interface RMSD (I_rmsd) is the C^α RMSD of interface residues in the decoy compared to the native structure after superposition of the interface residues, where interface residues are defined as those with intermolecular distances of <4 Å in the native structure. The fraction of native residue–residue contacts (f_{nat}) is a measure of the specificity of the interactions across an interface. For CAPRI, Mendez *et al.* outline three classifications for docking accuracy based on these metrics: high quality (f_{nat}>0.5 AND [L_rmsd < 1.0 OR I_rmsd < 1.0]); medium quality (f_{nat} > 0.3 AND [L_rmsd < 5.0 OR I_rmsd < 2.0]); acceptable quality (f_{nat} > 0.1 AND [L_rsmd <10.0 OR I_rmsd < 4.0]).[63] The backbone RMSD of the flexible ligand is measured as the C^α RMSD of interface residues after superposition of the entire ligand, to the bound ligand conformation (BB_rmsd) and the unbound ligand conformation (UB_rmsd).

Docking methods

Backbone flexibility is always limited to the ligand, which is defined as the smaller of the two proteins in the complex. All starting structures are prepared by replacing the side chains with rotamers from a standard rotamer library expanded to include rotamers from the unbound crystal structures.[11,39] Independent local docking runs generate 1000 decoys for docking crystal structures or 5000 decoys for docking NMR structures for CS or CS/IF docking in order to accommodate the significantly larger ensemble sizes.

KL docking is local docking using standard RosettaDock as described by Gray *et al.*, with rotamer torsion angle minimization as described by Wang *et al.*[11,39] Antibody docking targets use an alignment file that biases low-resolution docking towards complementarity-determining region residues, as described previously.[34,64,65] For IF methods, the gradient-based energy minimization is expanded beyond rigid-body orientation to include backbone torsion angles during the minimization step in the high-resolution MC minimization phase of RosettaDock, as described by Wang *et al.*[26] The unbound ligand structure used in IF docking was relaxed prior to docking to ensure that the structure was at a local energy minimum with the Rosetta energy function, as described by Wang *et al.*[26]

For CS docking's CS moves, all n conformers in the ensemble are superposed along the interface residues of the current conformer (residues with <4 Å atomic distance to the receptor). Binding energy is calculated for each conformer to create a partition function $Z = \sum_i^n \exp(-\Delta G_i^{\text{binding}}/RT)$, and a random conformer i is selected based on its Boltzmann probability $P_i = \frac{\exp(-\Delta G_i^{\text{binding}}/RT)}{Z}$. As in IF docking, each conformer in the NMR ensemble underwent high-resolution refinement prior to docking to ensure that they were in a local energy minimum within the Rosetta energy function.

Scoring methods

Scoring both during the docking simulation and during the final decoy discrimination is based on binding energy, which is defined as the change in total energy from the final bound complex and the total energy of the initial unbound state: $\Delta G^{\text{binding}} = G(\text{complex}) - G(\text{unbound})$. The unbound energies for the receptor and each ligand conformer are calculated by taking the lowest energy from doing 10 independent side-chain packing runs. The total unbound energy is the sum of the unbound energies of the receptor and the ligand, and is specific to the conformer used in the final complex: $G(\text{unbound}) = G(\text{receptor}) + G(\text{conformer } i)$. All unbound reference energies are calculated prior to docking.

The Rosetta centroid-mode energy function used in docking is identical with that used in Gray *et al.* and consists of contact, bump, residue environment, and residue–residue pair potential.[11] The Rosetta all-atom energy function used in docking consists of van der Waals, hydrogen bonding, side-chain probability, solvation, and electrostatic terms,[11] and the weights used for each component are identical with those used by Wang *et al.*[39] At high resolution, an additional component of a secondary-structure-dependent backbone (Ramachandran) torsional potential was also used[66] and assigned a weight of 0.1.

Ensemble generation

Ensembles are generated using RosettaRelax as described by Misura and Baker, except that 'wobble' and 'crank' perturbations are omitted primarily to reduce computational cost.[37]

Algorithm availability

Ensemble generation, NMR ensemble refinement, and all docking methods presented here are freely available for academic and nonprofit use as part of the Rosetta structure prediction suite†. The distribution includes supporting scripts, documentation, and full source code.

Acknowledgements

We acknowledge Michael Daily, Monica Berrondo, Arvind Sivasubramanian, and Chu Wang for insightful discussions and for reading and editing of the manuscript. This work was supported by National Institute of Health grant R01 GM078221. S.C. received funding from National Institute of Health training grant T32 GM008403.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.05.042

References

1. Gray, J. J. (2006). High-resolution protein–protein docking. *Curr. Opin. Struct. Biol.* **16**, 183–193.
2. Betts, M. J. & Sternberg, M. J. (1999). An analysis of conformational changes on protein–protein association: implications for predictive docking. *Protein Eng.* **12**, 271–283.
3. Lo Conte, L., Chothia, C. & Janin, J. (1999). The atomic structure of protein–protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198.
4. Bonvin, A. M. (2006). Flexible protein–protein docking. *Curr. Opin. Struct. Biol.* **16**, 194–200.
5. Fischer, E. (1894). Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.* **27**, 2985–2993.
6. Wodak, S. J. & Janin, J. (1978). Computer analysis of protein–protein interaction. *J. Mol. Biol.* **124**, 323–342.
7. Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A. A., Aflalo, C. & Vakser, I. A. (1992). Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc. Natl Acad. Sci. USA*, **89**, 2195–2199.
8. Chen, R., Li, L. & Weng, Z. (2003). ZDOCK: an initial-stage protein-docking algorithm. *Proteins*, **52**, 80–87.
9. Li, L., Chen, R. & Weng, Z. (2003). RDOCK: refinement of rigid-body protein docking predictions. *Proteins*, **53**, 693–707.
10. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. (2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, **20**, 45–50.
11. Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A. & Baker, D. (2003). Protein–protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol.* **331**, 281–299.
12. Vajda, S. (2005). Classification of protein complexes based on docking difficulty. *Proteins*, **60**, 176–180.
13. Goodford, P. J. (1985). A computational procedure for determining energetically favorable binding sites

† www.rosettacommons.org

on biologically important macromolecules. *J. Med. Chem.* **28**, 849–857.
14. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R. & Ferrin, T. E. (1982). A geometric approach to macromolecule–ligand interactions. *J. Mol. Biol.* **161**, 269–288.
15. Monod, J., Wyman, J. & Changeux, J. P. (1965). On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **12**, 88–118.
16. Kumar, S., Ma, B., Tsai, C. J., Sinha, N. & Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci.* **9**, 10–19.
17. Krol, M., Chaleil, R. A., Tournier, A. L. & Bates, P. A. (2007). Implicit flexibility in protein docking: cross-docking and local refinement. *Proteins*, **69**, 7.
18. Smith, G. R., Sternberg, M. J. & Bates, P. A. (2005). The relationship between the flexibility of proteins and their conformational states on forming protein–protein complexes with an application to protein–protein docking. *J. Mol. Biol.* **347**, 1077–1101.
19. Grunberg, R., Leckner, J. & Nilges, M. (2004). Complementarity of structure ensembles in protein–protein binding. *Structure*, **12**, 2125–2136.
20. Bastard, K., Prevost, C. & Zacharias, M. (2006). Accounting for loop flexibility during protein–protein docking. *Proteins*, **62**, 956–969.
21. Koshland, D. E. (1958). Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl Acad. Sci. USA*, **44**, 98–104.
22. Dominguez, C., Boelens, R. & Bonvin, A. M. (2003). HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737.
23. Krol, M., Tournier, A. L. & Bates, P. A. (2007). Flexible relaxation of rigid-body docking solutions. *Proteins*, **68**, 159–169.
24. Smith, G. R., Fitzjohn, P. W., Page, C. S. & Bates, P. A. (2005). Incorporation of flexibility into rigid-body docking: applications in rounds 3–5 of CAPRI. *Proteins*, **60**, 263–268.
25. de Vries, S. J., van Dijk, A. D., Krzeminski, M., van Dijk, M., Thureau, A., Hsu, V. *et al.* (2007). HADDOCK *versus* HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins*, **69**, 726–733.
26. Wang, C., Bradley, P. & Baker, D. (2007). Protein–protein docking with backbone flexibility. *J. Mol. Biol.* **373**, 503–519.
27. Chaudhury, S., Sircar, A., Sivasubramanian, A., Berrondo, M. & Gray, J. J. (2007). Incorporating biochemical information and backbone flexibility in RosettaDock for CAPRI rounds 6–12. *Proteins*, **69**, 793–800.
28. Wang, C., Schueler-Furman, O., Andre, I., London, N., Fleishman, S. J., Bradley, P. *et al.* (2007). RosettaDock in CAPRI rounds 6–12. *Proteins*, **69**, 758–763.
29. van Dijk, A. D., de Vries, S. J., Dominguez, C., Chen, H., Zhou, H. X. & Bonvin, A. M. (2005). Data-driven docking: HADDOCK's adventures in CAPRI. *Proteins*, **60**, 232–238.
30. Dominguez, C., Bonvin, A. M., Winkler, G. S., van Schaik, F. M. A., Timmersm, H. T. M. & Boelens, R. (2004). Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches. *Structure*, **12**, 633–644.
31. Jensen, G. A., Andersern, O. M., Bonvin, A. M., Bjerrum-Bohr, I., Etzerodt, M., Thogersen, H. C. *et al.* (2006). Binding site structure f LRP–RAP complex: implications for a common ligand–receptor binding motif. *J. Mol. Biol.* **362**, 700–716.
32. Drogen-Petit, A., Zwahlen, C., Peter, M. & Bonvin, A. M. (2004). Insight into molecular interactions between two PB1 domains. *J. Mol. Biol.* **336**, 1195–1210.
33. Philippopoulos, M. & Lim, C. (1999). Exploring the dynamic information content of a protein NMR structure: comparison of a molecular dynamics simulation with the NMR and X-ray structures of *Escherichia coli* ribonuclease HI. *Proteins*, **36**, 87–110.
34. Gray, J. J., Moughon, S. E., Kortemme, T., Schueler-Furman, O., Misura, K. M., Morozov, A. V. & Baker, D. (2003). Protein–protein docking predictions for the CAPRI experiment. *Proteins*, **52**, 118–122.
35. Daily, M. D., Masica, D., Sivasubramanian, A., Somarouthu, S. & Gray, J. J. (2005). CAPRI rounds 3–5 reveal promising successes and future challenges for RosettaDock. *Proteins*, **60**, 181–186.
36. Camacho, C. J. & Vajda, S. (2002). Protein–protein association kinetics and protein docking. *Curr. Opin. Struct. Biol.* **12**, 36–40.
37. Misura, K. M. & Baker, D. (2005). Progress and challenges in high-resolution refinement of protein structure models. *Proteins*, **59**, 15–29.
38. Bradley, P., Misura, K. M. & Baker, D. (2005). Toward high-resolution de novo structure prediction for small proteins. *Science*, **309**, 1868–1871.
39. Wang, C., Schueler-Furman, O. & Baker, D. (2005). Improved side-chain modeling for protein–protein docking. *Protein Sci.* **14**, 1328–1339.
40. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. *et al.* (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242.
41. Harel, M., Kleywegt, G. J., Ravelli, R. B., Silman, I. & Sussman, J. L. (1995). Crystal structure of an acetylcholinesterase–fasciculin complex: interaction of a three-fingered toxin from snake venom with its target. *Structure*, **3**, 1355–1366.
42. le Du, M. H., Housset, D., Marchot, P., Bougis, P. E., Navaza, J. & Fontecilla-Camps, J. C. (1996). Structure of fasciculin 2 from green mamba snake venom: evidence for unusual loop flexibility. *Acta Crystallogr. Sect. D*, **52**, 87–92.
43. Brunne, R. M., Berndt, K. D., Guntert, P., Wuthrich, K. & van Gunsteren, W. F. (1995). Structure and internal dynamics of the bovine pancreatic trypsin inhibitor in aqueous solution from long-time molecular dynamics simulations. *Proteins*, **23**, 49–62.
44. Gabdoulline, R. R. & Wade, R. C. (2001). Protein–protein association: investigation of factors influencing association rates by Brownian dynamics simulations. *J. Mol. Biol.* **306**, 1139–1155.
45. Bui, J. M. & McCammon, J. A. (2006). Protein complex formation by acetylcholinesterase and the neurotoxin fasciculin-2 appears to involve an induced-fit mechanism. *Proc. Natl Acad. Sci. USA*, **103**, 15451–15456.
46. Guenneugues, M., Drevet, P., Pinkasfeld, S., Gilquin, B., Menez, A. & Zinn-Justin, S. (1997). Picosecond to hour time scale dynamics of a "three finger" toxin: correlation with its toxic and antigenic properties. *Biochemistry*, **36**, 16097–16108.
47. Tang, C., Schwieters, C. D. & Clore, G. M. (2007). Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature*, **449**, 1078–1082.
48. Mintseris, J., Wiehe, K., Pierce, B., Anderson, R., Chen, R., Janin, J. & Weng, Z. (2005). Protein–Protein Docking Benchmark 2.0: an update. *Proteins*, **60**, 214–216.
49. Andrec, M., Snyder, D. A., Zhou, Z., Young, J., Montelione, G. T. & Levy, R. M. (2007). A large data set comparison of protein structures determined by crystallography and NMR: statistical test for structural

differences and the effect of crystal packing. *Proteins*, **69**, 449–465.
50. de Groot, B. L., van Aalten, D. M., Scheek, R. M., Amadei, A., Vriend, G. & Berendsen, H. J. (1997). Prediction of protein conformational freedom from distance constraints. *Proteins*, **29**, 240–251.
51. Yang, L. W., Eyal, E., Chennubhotla, C., Jee, J., Gronenborn, A. M. & Bahar, I. (2007). Insights into equilibrium dynamics of proteins from comparison of NMR and X-ray data with computational predictions. *Structure*, **15**, 741–749.
52. Sivasubramanian, A., Sircar, A., Chaudhury, S. & Gray, J. J. (2008). Towards high-resolution homology modeling of antibody Fv regions using knowledge-based techniques, de novo loop modeling and docking. Manuscript in preparation.
53. London, N. & Schueler-Furman, O. (2007). Assessing the energy landscape of CAPRI targets by FunHunt. *Proteins*, **69**, 809–815.
54. Cavasotto, C. N., Kovacs, J. A. & Abagyan, R. A. (2005). Representing receptor flexibility in ligand docking through relevant normal modes. *J. Am. Chem. Soc.* **127**, 9632–9640.
55. Markwick, P. R., Bouvignies, G. & Blackledge, M. (2007). Exploring multiple timescale motions in protein GB3 using accelerated molecular dynamics and NMR spectroscopy. *J. Am. Chem. Soc.* **129**, 4724–4730.
56. Seeliger, D., Haas, J. & de Groot, B. L. (2007). Geometry-based sampling of conformational transitions in proteins. *Structure*, **15**, 1482–1492.
57. Ben-Zeev, E., Kowalsman, N., Ben-Shimon, A., Segal, D., Atarot, T., Noivirt, O. *et al.* (2005). Docking to single-domain and multiple-domain proteins: old and new challenges. *Proteins*, **60**, 195–201.
58. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. (2005). Geometry-based flexible and symmetric protein docking. *Proteins*, **60**, 224–231.
59. Tobi, D. & Bahar, I. (2005). Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proc. Natl Acad. Sci. USA.* **102**, 18908–18913.
60. Noy, E., Tabakman, T. & Goldblum, A. (2007). Constructing ensembles of flexible fragments in native proteins by iterative stochastic elimination is relevant to protein–protein interfaces. *Proteins*, **68**, 702–711.
61. Wiehe, K., Pierce, B., Tong, W. W., Hwang, H., Mintseris, J. & Weng, Z. (2007). The performance of ZDOCK and ZRANK in rounds 6–11 of CAPRI. *Proteins*, **69**, 719–725.
62. Chen, R., Mintseris, J., Janin, J. & Weng, Z. (2003). A protein–protein docking benchmark. *Proteins*, **52**, 88–91.
63. Mendez, R., Leplae, R., De Maria, L. & Wodak, S. J. (2003). Assessment of blind predictions of protein–protein interactions: current status of docking methods. *Proteins*, **52**, 51–67.
64. Sivasubramanian, A., Chao, G., Pressler, H. M., Wittrup, K. D. & Gray, J. J. (2006). Structural model of the mAb 806–EGFR complex using computational docking followed by computational and experimental mutagenesis. *Structure*, **14**, 401–414.
65. Sivasubramanian, A., Maynard, J. A. & Gray, J. J. (2008). Modeling the structure of mAb 14B8 bound to the anthrax protective antigen. *Proteins*, **70**, 218–230.
66. Simons, K. T., Kooperberg, C., Huang, E. & Baker, D. (1997). Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *J. Mol. Biol.* **268**, 209–225.