Structure *14*, 683–693, April 2006 DOI 10.1016/j.str.2006.01.014

Flexibility and Conformational Entropy in Protein-Protein Binding

Raik Grünberg,[1] Michael Nilges,[1] and Johan Leckner[2,*]
[1] Unité de Bioinformatique Structurale
CNRS URA 2185
Institut Pasteur
25-28 rue du docteur Roux
F-75015 Paris
France
[2] Department of Chemical and Biological Engineering
Molecular Biotechnology
Chalmers University of Technology
P.O. Box 462
SE-405 30 Göteborg
Sweden

Summary

To better understand the interplay between protein-protein binding and protein dynamics, we analyzed molecular dynamics simulations of 17 protein-protein complexes and their unbound components. Complex formation does not restrict the conformational freedom of the partner proteins as a whole, but, rather, it leads to a redistribution of dynamics. We calculate the change in conformational entropy for seven complexes with quasiharmonic analysis. We see significant loss, but also increased or unchanged conformational entropy. Where comparison is possible, the results are consistent with experimental data. However, stringent error estimates based on multiple independent simulations reveal large uncertainties that are usually overlooked. We observe substantial gains of pseudo entropy in individual partner proteins, and we observe that all complexes retain residual stabilizing intermolecular motions. Consequently, protein flexibility has an important influence on the thermodynamics of binding and may disfavor as well as favor association. These results support a recently proposed unified model for flexible protein-protein association.

Introduction

Specific associations between proteins are fundamental to all aspects of cell biology. However, even with their atomic detail structures at hand, we are usually still unable to predict if, how, and with what affinity two proteins interact. Our understanding is obstructed by the complex dynamics of the many thousand atoms that make up the binding partners and the surrounding solvent. Owing to this flexibility, the stability of protein-protein complexes is strongly influenced by entropic contributions (Brady and Sharp, 1997), many of which are notoriously difficult to estimate or measure.

Intuitively, binding is usually assumed to restrict the flexibility of both partner molecules and, as a consequence, to claim a significant cost of conformational entropy. However, the effect of protein flexibility on binding remains elusive, except for the mere fact that structural changes occur between the free and the bound states (Betts and Sternberg, 1999; Lo Conte et al., 1999). Two long-standing models focus on flexibility to describe the binding process. The popular notion of induced fit (Koshland, 1958) assumes a passive mutual adaptation, but it does not explain specific recognition if the two protein structures are not complementary to start with (Bosshard, 2001). The rival idea of conformer selection (Monod et al., 1965; Kumar et al., 2000) suggests a recognition between the two bound conformations that are postulated to occur within the diverse structure ensembles of the two free proteins. Experiments have demonstrated the preferential recognition of distinct subconformations in various binding reactions (Kirschner et al., 1966; Austin et al., 1975; Lancet and Pecht, 1976; Foote and Milstein, 1994; Leder et al., 1995; Berger et al., 1999; James and Tawfik, 2005). Likewise, we showed in a previous simulation study that some conformations among the structure ensembles of unbound proteins are more prone to recognition than others (Grünberg et al., 2004). Nevertheless, recognition seemed not to depend on the presence of the bound conformation. Furthermore, the postulate of bound conformations that have to be visited simultaneously by the two colliding proteins appears to be inconsistent with the fast pace of recognition (Grünberg et al., 2004).

The concept of induced fit as well as the conformer selection model imply a loss of conformational freedom for both binding partners. Experimental studies do not actually support this assumption (Forman-Kay, 1999). Binding of small ligands, peptides, or other proteins globally rigidifies some proteins (Olejniczak et al., 1997; Lee et al., 2000; Mercier et al., 2001; Wang et al., 2001; Kern and Zuiderweg, 2003), whereas others appear to become more flexible (Stivers et al., 1996; Yu et al., 1996; Zidek et al., 1999; Vergani et al., 2000; Yun et al., 2001; Loh et al., 2001; Zhu et al., 2001; Finerty et al., 2002; Arumugam et al., 2003; Fayos et al., 2003; Balog et al., 2004). Changes in dynamics are often detected in regions far from the interacting residues (Yu et al., 1996; Zidek et al., 1999; Loh et al., 2001; Wang et al., 2001; Yun et al., 2001; Arumugam et al., 2003; Fayos et al., 2003). These changes may transfer information to remote parts of the protein (allostery) (Buck and Iyengar, 2003), but they have also been suggested to regulate binding affinity (Steinberg and Scheraga, 1963). Computational studies on three complexes have predicted a loss of overall conformational entropy (Viñals et al., 2002; Gohlke and Case, 2004; Hsu et al., 2004). Conversely, gains were calculated for the dimerization of insulin (Tidor and Karplus, 1994) and were attributed to residual intermolecular motions (Finkelstein and Janin, 1989) and the redistribution of vibrational densities (Steinberg and Scheraga, 1963).

Flexibility may thus have considerable influence on the stability of protein complexes. However, experiments and calculations find no consistent trend even for the sign of this contribution. The available data

*Correspondence: johan.leckner@chalmers.se

Table 1. Protein-Protein Complexes Analyzed in This Study

ID[a]	Receptor/Ligand	PDB Codes (Chain/Model Identifier)			Size[b]		
		Receptor	Ligand	Complex	Receptor	Ligand	ΔRes[c]
c01	Trypsin/Amyloid β-protein precursor inhibitor domain	1BRA	1AAP(A)	1BRC(E:I)	223	56	0
c02	α-chymotrypsinogen/Pancreatic secretory trypsin inhibitor	2CGA(A)	1HPT	1CGI(E:I)	245	56	0
c03	Kallikrein A/Pancreatic trypsin inhibitor	2PKA(AB)	5PTI	2KAI(AB:I)	232	58	−1
c04	Subtilisin BPN/Subtilisin inhibitor	1SUP	3SSI	2SIC(E:I)	275	108	−1
c05	Extracellular domain of tissue factor/Antibody Fab 5G9	1FGN(LH)	1BOY	1AHW(AB:C)	248	211	−11
c06	Humanized anti-lysozyme Fv/Lysozyme	1BVL(AB)	3LTZ	1BVK(AB:C)	224	129	0
c08	Anti-lysozyme antibody Hyhel-63/Lysozyme	1DQQ(AB)	3LTZ	1DQJ(AB:C)	424	129	0
c11	Barnase/Barstar	1A19(A)	1A2P(A)	1BSG(A:E)	108	89	2
c13	Ribonuclease inhibitor/Ribonuclease A	2BNH	7RSA	1DFJ(E:I)	456	124	0
c14	Acetylcholinesterase/Fasciculin-II	1VXR	1FSC(A)	1FSS(A:B)	532	61	0
c15	HIVB-1 NEF/FYN tyrosin kinase SH3 domain	1AVV	1SHF(A)	1AVZ(B:C)	99	59	2
c16	Uracil-DNA glycosylase/Inhibitor	1AKZ	1UGI(A)	1UGH(E:I)	223	83	−1
c17	RAS activating domain/RAS	1WER	5P21	1WQ1(R:G)	324	166	−4
c19	Glycosyltransferase/Tendamistat	1PIF	2AIT(mdl1)	1BVN(P:T)	495	74	−2
c20	CDK2 cyclin-dependant kinase 2/Cyclin A	1HCL	1VIN	1FIN(A:B)	294	252	12
c21	CDK2 cyclin-dependant kinase 2/KAP	1B39(A)	1FPZ(A)	1FQ1(A:B)	290	176	13
c22	Transductin Gt-α/Heteromeric G protein	1TBG(AE)	1TAG	1GOT(A:BG)	408	314	13

[a] Complex identifier used throughout the paper (retained from www.bmm.icnet.uk/docking/systems.html).
[b] Size in residues.
[c] Number of residues resolved in the bound, but not in the free, structure (−, vice versa).

question a general binding-induced loss of conformational entropy and hence also challenge the models from which this assumption was derived. The confusion is fostered by shortcomings of experimental as well as computational methods and the lack of systematic studies that cover more than a single complex. Even though it is possible to make direct measurements with NMR, these experiments focus on fluctuations of selected atoms, traditionally backbone amides, that might not correctly reflect overall flexibility. A recent neutron scattering experiment measured the motion of all hydrogens, but it was limited to harmonic vibrations at very low temperature (Balog et al., 2004). On the other hand, molecular dynamics can be simulated at full atomic detail (Frauenfelder and Leeson, 1998), and entropy estimates can, in theory, be extracted by quasiharmonic analysis (Teeter and Case, 1990). However, when applied to protein-protein complexes (Gohlke and Case, 2004; Hsu et al., 2004), quasiharmonic calculations encounter convergence problems and have thus far yielded unrealistic results (Gohlke and Case, 2004; Hsu et al., 2004) (also see below). Calculations therefore resort to normal mode analysis (Case, 1994; Tidor and Karplus, 1994; Gohlke and Case, 2004), but this approach cannot account for anharmonic motions or solvent effects. Yet, at physiological temperature, overall protein flexibility is dominated by such anharmonic motions (Kitao et al., 1998) and may, to a large extent, be driven by solvent fluctuations (Fenimore et al., 2004). Normal mode analysis of static structures may therefore constitute a rather problematic simplification.

We here look at the dynamics of 17 different protein complexes (Table 1) in both their free and bound forms by using molecular dynamics (MD) simulations. We find that binding leads to a redistribution of dynamics and that overall flexibility can equally be lost or gained. We resolve some of the technical issues that thus far prevented the application of quasiharmonic analysis to the study of protein-protein interaction and estimate the entropic contribution to the stability of seven complexes. Our results further question the traditional description of binding, but they blend in with a unified model we recently proposed (Grünberg et al., 2004). Apparently, protein flexibility has a large influence on the thermodynamics of binding and may both favor and disfavor association.

Results and Discussion

Flexibility of Free and Bound States

We selected a set of 17 protein-protein complexes for which three-dimensional structures of both free components and the complex were available (that is 17 × 3 molecular structures). The systems are listed in Table 1 and comprise enzyme-inhibitor pairs, antibody-antigen assemblies, and complexes relaying intracellular signals. They, essentially, all fall into the category of transient interactions between globular proteins. The motion of seven systems (7 × 3 molecules) was sampled by ten independent parallel simulations of 1 ns length. In addition, the complete set of 17 × 3 molecules was subjected to shorter, less elaborate 10 × 50 ps simulations (see the Experimental Procedures). The setup of several short simulations yields better conformal sampling than one long simulation (Caves et al., 1998).

We define flexibility as the average pairwise distance between MD snapshots taken from the second half of mutually independent simulations. The evaluation of pairwise distances eliminates the need for an arbitrary reference structure. Distances were only calculated between snapshots from independent simulations, which diminishes the influence of the sampling interval. Otherwise, the high similarity of neighboring snapshots would create a subpopulation of small distances that would distort the mean as well as width of the overall distribution.

Figure 1 shows the flexibility of the 17 protein pairs before and after binding. There was no common tendency toward restricted motion. We observed both decreased and increased fluctuations, and, as shown in Figure 2A,

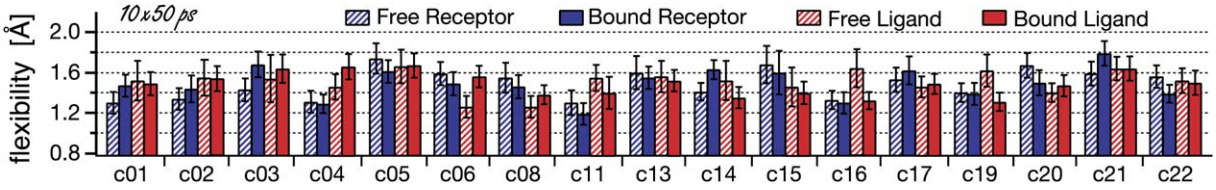

Figure 1. The Flexibility of 34 Proteins before and after Binding

Flexibility is determined for all heavy atoms from the 10 × 50 ps simulations. Flexibility was defined as the average pairwise distance of simulation snapshots. Error bars quantify the standard deviation. There is no trend toward increased or decreased flexibility.

the flexibility remained, on average, unchanged. However, shifts of dynamics spread unevenly through different regions of the proteins. Figure 2A compares the average effect of binding on the flexibility of contact and noncontact surfaces as well as backbone atoms. Not surprisingly, binding sites lost conformational freedom upon formation of the protein complex. However, regions outside the contact area often experienced moderate gains of mobility (this is not an artifact of superpositioning; also, the conformations of the complex were fitted separately for receptor and ligand). The assumption that binding generally restricts the flexibility of proteins is thus not supported by our simulations, which is in line with the picture emerging from experimental studies (Forman-Kay, 1999). A redistribution rather than loss of mobility may be a common feature of protein-protein interactions.

The shorter simulations did not adequately sample the slow, residual intermolecular motions of receptors and ligands in protein complexes and were not sufficient for the calculation of entropies (see below). We therefore performed extended and more elaborate simulations of 10 × 1 ns length on a subset of seven complexes. Representative snapshots of these 21 simulations are shown in Figure 3. The selected set mirrored the diversity of the complete list of 17 complexes but was (owing to computational restraints) somewhat biased to smaller systems. The longer simulations yielded a similar pattern of flexibility changes (Figure 2B). The overall flexibility of receptor and ligand structures correlated between the two very different simulation setups (R = 0.95, P = 10^{-4} for the free ensembles [removing c17 receptor as outlier]; R = 0.77, P = 0.01 for the bound ensembles). The less elaborate 10 × 50 ps simulations, performed on the complete list of complexes, thus appear to be sufficient for the study of flexibilities. In fact, they cover the time window available to recognition (Grünberg et al., 2004).

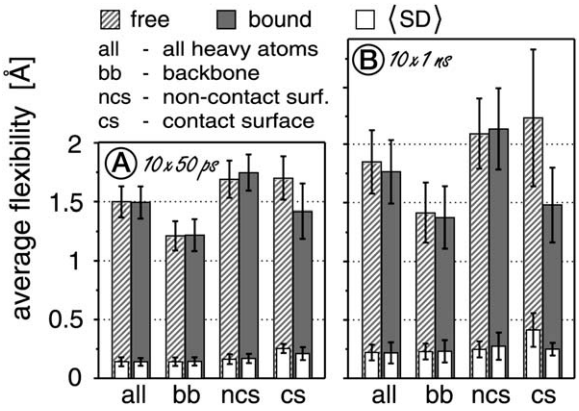

Figure 2. The Average Flexibility before and after Binding

(A) The flexibility before and after binding averaged over 34 proteins (data are derived from the 10 × 50 ps simulations). The white bars are the average of the standard deviations given in Figure 1. Binding sites generally lose flexibility, whereas other regions often gain flexibility. The flexibility of the overall protein appears, on average, to be unaffected by binding. (The measure also depends on the size and shape of the considered region. A fair comparison is thus only possible for the free and bound flexibilities of the same atom selection). (B) The flexibility of 14 proteins, simulated on a longer time scale (10 × 1 ns), before and after binding. The subset is not representative for the 34 proteins in (A).

Our analysis (unlike the simulations) only considered residues that were resolved in both the free and bound experimental structures. Only two complexes among cases 1–14 have fewer unresolved residues than the free state (which may or may not indicate a lower flexibility), whereas 6 complexes have more (see Table 1). Remarkably, and in line with our observations, the unresolved residues are, in the majority of these cases, situated in regions far from the interface, and none of them are located within the contact region. Complexes 20–22 are the exceptions to this rule. These three complexes all have interface loops, unresolved in the free crystallographic structures, that undergo large structural changes upon complex formation (up to 20 Å C_α displacements).

Calculation of Conformational Entropies

Out of the components of free energy, conformational entropy seems currently to be the most difficult to calculate (Gohlke and Case, 2004). So far, two studies have attempted to examine the entropy change of protein complexation by quasiharmonic analysis. Gohlke and Case (2004) discarded the approach owing to insufficient convergence and an unrealistic result. Hsu et al. (2004) only compared entropies of the individual partner proteins in the free and bound states and also reported convergence problems. Their sum of individual (pseudo) entropy changes aligned with an experimental value, but the latter comprised a presumably large share of desolvation entropy. The calculation did not consider desolvation, and the result must thus also be considered unsatisfactory. In the following sections, we describe a protocol that combines quasiharmonic analysis with a careful choice of reference states and yields reasonable, converged entropy differences. The results of our calculations are then presented subsequent to a discussion of possible errors.

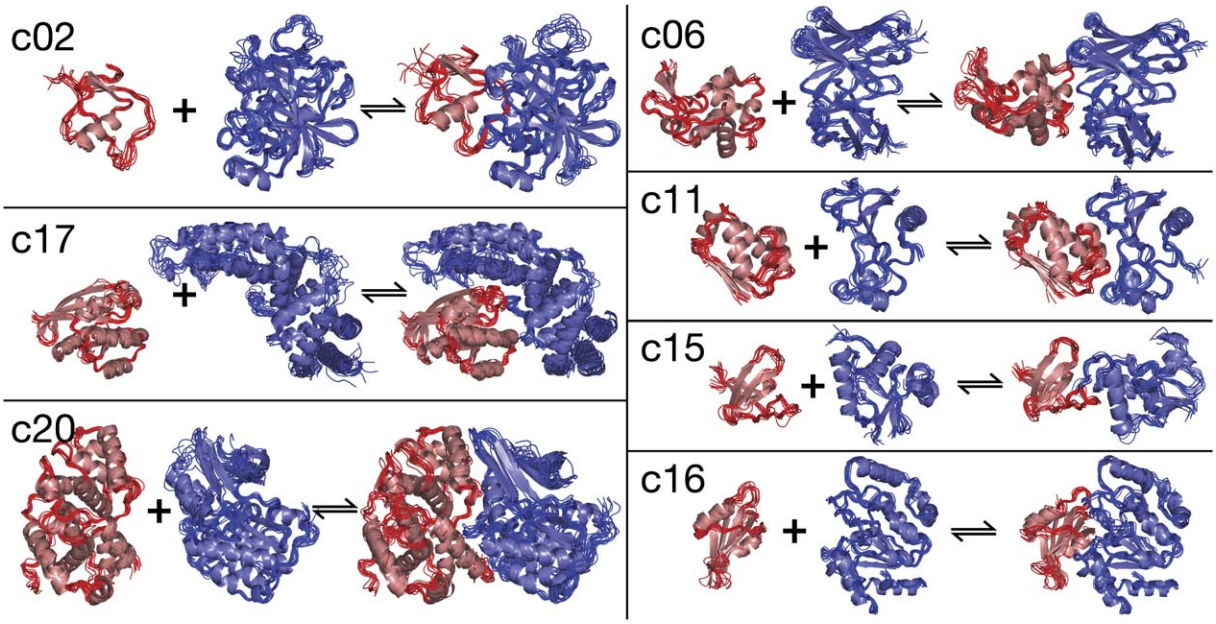

Figure 3. Representative Snapshots from 10 × 1 ns Simulations of the Free and Bound States of Seven Protein-Protein Complexes

For clarity, only the backbone is shown. The ligand and receptor are colored red and blue, respectively. Ten snapshots were selected from each ensemble by c-means fuzzy clustering, which is described earlier (Gordon and Somorjai, 1992; Grünberg et al., 2004). The increased rigidity of interface residues and the increased flexibility of regions distant to the interface are often clearly visible. Note the large changes in c02 and c20, which also experience conformational entropy gains (see Table 3).

Using the same program as Gohlke and Case (2004), we derived absolute entropies from the 10 × 1 ns ensembles of seven free receptors (S_{rec}), ligands (S_{lig}), and their complexes (S_{com}). Subtracting the free from the bound state should yield the change of conformational entropy induced by binding, that is $\Delta S_{conf} = S_{com} - (S_{rec} + S_{lig})$. However, in the one example described by Gohlke and Case (2004), this same strategy overestimated the entropy loss by 790 cal mol^{-1} K^{-1} (compared to a normal mode calculation that was at least compatible with experimental data). Also in our case, the difference turned out to be far too negative for all seven complexes (data not shown).

The effect hence constitutes a systematic error, which we traced to spurious correlations between receptor and ligand. Figure 4 shows the normalized covariance matrix of an artificially constructed trajectory, the combination of a free ligand and a free receptor ensemble. The matrix reveals unphysical correlations between the independently simulated molecules (upper-left quadrant). Such correlations also occur between unrelated noninteracting proteins (data not shown). According to our preliminary analysis, spurious correlations could originate from inaccuracies in calculating the covariance matrix from finite simulation data (combined with the fact that, rather than canceling out, errors add up and always lower the final entropy). A second important component is the random correlations between atomic vibrations that appear to be common and frequent in any deterministic protein simulation. These spurious correlations lower the entropy value of the fake complex ($S_{rec+lig} << S_{rec} + S_{lig}$), but, more importantly, they also compromise the entropy calculated for the real complex. Spurious correlations should cancel out if the bound state is compared to the artificial combination of free receptor and ligand ensembles. We therefore calculated the conformational entropy of binding by using the artificial free "complex" as a reference state: $\Delta S_{conf} = S_{com} - S_{rec+lig}$.

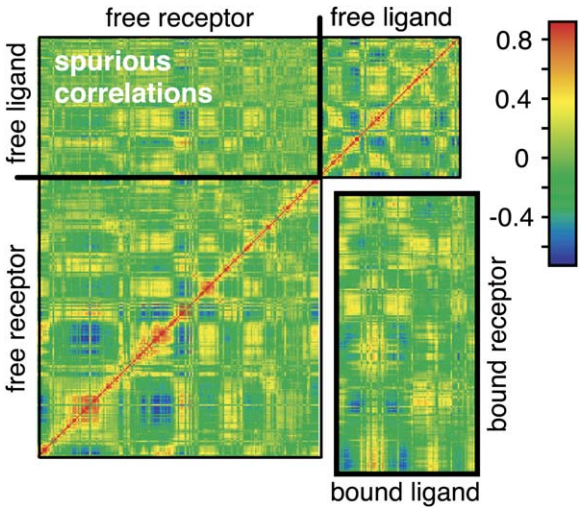

Figure 4. Spurious Correlations between Independent Simulations

The trajectory of a free receptor is artificially combined with the trajectory of a free ligand. The correlation matrix shows the expected intramolecular correlations in the lower-left and upper-right quadrants, but it also reveals unphysical correlations between the independent simulations (upper-left quadrant). For comparison, the lower-right quadrant shows "real" crosscorrelations of the two molecules in the simulation of the protein complex. Data were taken from the last 500 ps of three single 1 ns simulations of c15. For space reasons, the plot only considers the x coordinates of every fourth atom.

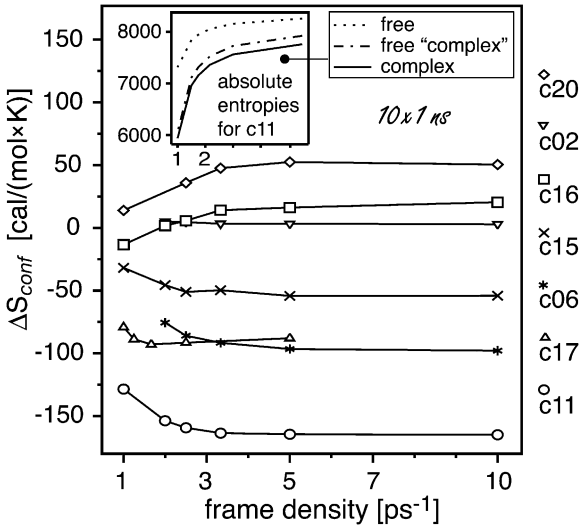

Figure 5. Dependency of Conformational Binding Entropies on the Sampling Interval

Differences of bound and free conformational entropies were calculated for the last 10 or 9 × 200 ps (c17, 300 ps; c20, 100 ps) of simulations covering 10 × 1 ns while using different offsets between the snapshots (from 0.1 to 1 ps).

In line with previous observations (Gohlke and Case, 2004; Hsu et al., 2004), the absolute entropy of the free or bound state did not converge. It generally increased with the addition of additional simulation data — be it additional frames of a constant time segment (inset in Figure 5) or frames that covered a longer time period (inset in Figure 6). By contrast, the entropy *difference* between the bound and free states showed sufficient convergence. This concerns the density of sampling (see Figure 5) as well as the necessary time coverage (shown in Figure 6). For most complexes, sufficient convergence was ensured by 10 or 9 × 300 ps coverage (taken from the simulation end) and a 0.2 ps sampling interval. The largest complex, c20, was only sampled every 0.3 ps. The entropy difference calculated for complex c17 did not converge; the value given below was obtained from the last 9 × 500 ps, with a 0.3 ps sampling interval.

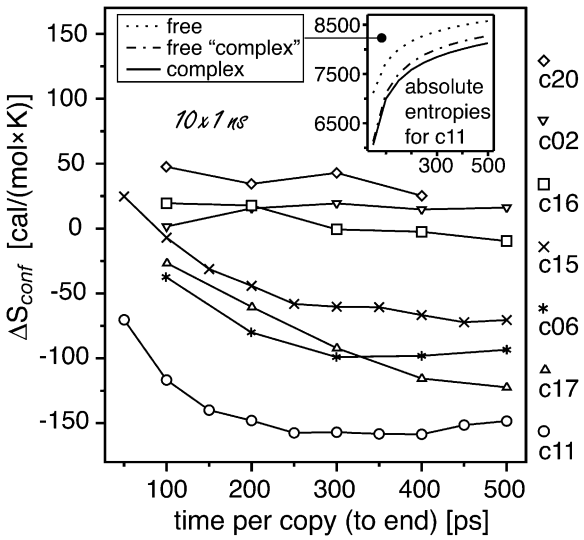

Figure 6. Convergence of Conformational Binding Entropies with Simulation Time

Absolute entropies of the free or bound state do not converge (inset). However, the difference of free and bound entropies converges reasonably well for most protein complexes. Values were calculated for different time segments from the end of 10 × 1 ns or 9 × 1 ns simulations of the free and bound states. For example, 100 ps translates to ten or nine segments covering the range 0.9–1 ns. Snapshots were taken every 0.2 or (c06, c17, c20) 0.3 ps.

Table 2. Protocols Used for Entropy Calculations

Number	Protocol	MD[a]	Analyze[b]	Fit[c]	Pair[d]	Order[e]
1	rec	free	rec	r		
2	lig	free	lig	l		
3	rec+lig	free	rec+lig	r ‖ l	2	intact
4	rec+lig, shuff	free	rec+lig	r ‖ l	2	shuffled
5	rec, bound	com	rec	r		
6	lig, bound	com	lig	l		
7	com	com	rec+lig	r+l	1	intact
8	com, split	com	rec+lig	r ‖ l	1	intact
9	com, swap	com	rec+lig	r ‖ l	2	intact
10	com, shuff	com	rec+lig	r ‖ l	1	shuffled

[a] Simulation of free receptor/ligand (free) or of complex (com).
[b] Consider receptor (rec) and/or ligand (lig).
[c] Fit trajectories on reference rec (r) and lig (l), separately (r ‖ l) or as single molecule (r+l).
[d] Rec and lig taken from the same (1) or two independent (2) simulations.
[e] Time order of ligand coordinates.

Subcomponents of the conformational entropy can be isolated by disrupting correlations. To this end, we developed a number of entropy calculation protocols, outlined in Table 2, in which alternate fitting and reorganization of trajectories destroyed different correlations. The separate superpositioning of receptor and ligand in the complex trajectory (giving $S_{com,split}$) reveals the entropy content of residual intermolecular motions of receptor versus ligand (denoted rec|lig in Table 3). The remaining correlations across the binding interface are uncovered if the complex trajectory is, moreover, divided into receptor and ligand and reassembled from independent bound trajectories ($S_{com,swap}$). The difference between $S_{com,split}$ and $S_{com,swap}$ (denoted rec×lig in Table 3) yields the negative entropy contribution from such correlated motions.

Errors and Limitations of Entropy Calculations

Alternate snapshots or time windows of a single MD simulation are commonly used to estimate errors when quantities are calculated from simulations. However, a deterministic simulation in the nanosecond range cannot possibly give a converged picture of protein motions that, in reality, stretch at least to millisecond timescales (Bouvignies et al., 2005). Variation within a single simulation is hence a poor measure for the variation to be expected across the conformational space accessible to the protein. Indeed, Caves et al. (1998) showed that multiple short simulations sample a larger conformational space than a single simulation of equivalent length. Our multiple simulation approach (10 × 1 ns) thus improves conformal sampling, but, given the large

Table 3. Conformational Binding Entropy and Its Decomposition

	ΔS_{conf} Components				ΔS_{conf} Total	
	Rec[a]	Lig[b]	Rec|Lig[c]	Rec×Lig[d]	Backbone[e]	All Atoms[f]
Protocol[g] →	5-1	6-2	8-7	9-8	7-3	7-3
c02	20 ± 38	−46 ± 19	46 ± 3	−11 ± 1	19 ± 26	19 ± 50
c06	−196 ± 59	31 ± 28	55 ± 3	−8.2 ± 1	12 ± 16	−101 ± 46
c11	−115 ± 49	−123 ± 47	55 ± 5	−8.9 ± 1	−27 ± 22	−157 ± 57
c15	−55 ± 34	−103 ± 29	70 ± 7	−6.6 ± 1	12 ± 14	−60 ± 49
c16	75 ± 46	−144 ± 25	45 ± 2	−13 ± 1	40 ± 15	−0.7 ± 41
c17[h]	−83 ± 71	−153 ± 36	52 ± 11	−11 ± 1	−0.4 ± 29	−147 ± 78
c20	−149 ± 34	116 ± 39	60 ± 9	−14 ± 2	18 ± 16	49 ± 48

All values give the difference of bound to free states (in cal mol^{-1} K^{-1}).
[a] Receptor only.
[b] Ligand only.
[c] Entropy gain from rigid body motions of receptor against ligand.
[d] Entropy loss from motions correlated across the binding interface.
[e] Carbonyl C and O only.
[f] All heavy atoms.
[g] Combination of protocols used to calculate this value (see Table 2).
[h] Not converged.

discrepancy between computationally accessible and theoretically necessary timescales, it should not be expected to yield exhaustive sampling either. However, and more importantly, the availability of several independent simulations allows us to assess the impact of (necessarily) insufficient sampling, which may be largely neglected in errors that are calculated from a single simulation.

We estimated errors with a standard jackknife procedure, that is we recalculated all entropy protocols nine or ten times, leaving out each independent simulation. This yielded errors of around 48 cal mol^{-1} K^{-1} for the difference of free and bound total conformational entropies (see Table 3). Neither Gohlke and Case (2004) nor Hsu et al. (2004) provide errors for the corresponding result of their quasiharmonic calculations. Gohlke and Case (2004) eventually resorted to normal mode analysis and report a standard error of 5 cal mol^{-1} K^{-1} for different snapshots from a single simulation. Hsu et al. (2004) focused on pseudo entropy differences of small protein segments for which they extrapolated errors with a block average procedure from a single simulation. This gave standard errors of 19 and 60 cal mol^{-1} K^{-1} for the two largest segments, comprising 198 and 356 atoms, respectively.

Furthermore, we observed that some of the 10 × 1 ns simulation sets contained singular 1 ns trajectories that (judged by the rmsd to the start or end structure) kept diverging, whereas the remaining nine trajectories appeared "equilibrated" over the last 500 ps. These are not necessarily artifacts, but they serve as another reminder of the globally incomplete sampling (and the issue with single simulation analysis). We sought to stabilize the results against such stochastic events by excluding outlier trajectories with an automatic procedure (see Experimental Procedures).

The large error, outlier trajectories, and spurious correlations discussed above highlight that the sampling procedure still remains the weak point of entropy calculations. Recently advocated replica exchange simulation methods (Gnanakaran et al., 2003; Habeck et al., 2005) may help resolve some of these issues. Inaccuracies may also arise from the conventional covariance matrix, which could give an inappropriate description of the anharmonic multimodal fluctuations in protein structure (Kitao et al., 1998).

Errors of close to ±50 cal mol^{-1} K^{-1} would introduce an uncertainty of 15 kcal mol^{-1} to calculations of binding free energies. Substantial improvements are required before we reach the accuracy needed for predictive or technical applications. Our individual entropy estimates are therefore approximate. Nevertheless, the more global picture emerging from seven different complexes offers a first insight into how flexibility may generally affect—and be affected by—protein-protein association. The range of total and disassembled entropy changes gives a preview of what one can expect from future, more sophisticated, experiments and calculations.

Conformational Entropy of Binding

Table 3 provides the total conformational entropy of binding, ΔS_{conf}, calculated for seven protein complexes. The values span a wide range from stark entropy loss (−157 ± 57 cal mol^{-1} K^{-1}) to substantial gain (49 ± 48 cal mol^{-1} K^{-1}). The two smallest complexes, c11 and c15, exhibited a large entropy loss, as did the antibody antigen system (c06). The largest assembly, c20, yielded the highest gain of conformational entropy. Two other comparatively large systems, c02 and c16, gave a moderately positive or unchanged entropy. Another large complex, c17, seemed to assemble at a high cost of conformational entropy, but the corresponding value did not converge. The assumption that conformational entropy changes may, to some extent, correlate with protein size and class appears attractive, but our data are yet too few and approximate to clearly support such a trend.

Table 3 also gives the (pseudo) entropy difference between the bound and free states of the individual receptor and ligand proteins (that is, $\Delta S_{rec} = S_{rec}^{bound} - S_{rec}$ and $\Delta S_{lig} = S_{lig}^{bound} - S_{lig}$). Clearly, the individual partners can both gain (up to 116 ± 39 cal mol^{-1} K^{-1}) or lose conformational entropy upon binding. Nevertheless, a loss of conformational entropy seemed to be

more common than perhaps expected from the analysis of flexibilities. On the other hand, our selection of seven complexes was, from the outset, biased toward reduced flexibility of the bound state (see Figure 1). Notably, in four of the seven systems, one binding partner gained (pseudo) entropy. Keeping in mind the generally diminished flexibility of binding interfaces, this is in itself a quite surprising result. Moreover, there appears to be a trend toward entropy compensation between receptor and ligand. The three complexes with positive or unchanged ΔS_{conf} always combined the large entropy gain of one partner with a sizable loss on the other side.

A simple summation of receptor and ligand entropies would ignore important entropy contributions ($S_{com} \neq S_{rec}^{bound} + S_{lig}^{bound}$). In the complex, the two partner molecules still move with respect to each other. This intermolecular motion translated into a substantial entropy gain (rec|lig in Table 3) of between 45 ± 2 and 70 ± 7 cal mol^{-1} K^{-1}. It thus recovers about half of the translational and rotational entropy loss. The value is in remarkable agreement with a "guess" (Janin, 1995) of 50 cal mol^{-1} K^{-1} made by Finkelstein and Janin (1989) over 15 years ago. Minh et al. (2005) very recently devised a different, more specialized quasiharmonic analysis method to derive this particular entropy contribution from MD simulations. They calculated a contribution of ~70 cal mol^{-1} K^{-1} for the Fasciculin-II/Acetylcholinesterase complex (here c14, not simulated on the nanosecond scale), which would fall into the range of values we observed for the different complexes. Contrary to this stabilizing entropy of intermolecular motion, some fluctuations should correlate across the binding interface (rec×lig in Table 3). Compared to other contributions, there was only moderate crosstalk between receptor and ligand fluctuations. The smallest complex lost 6.6 ± 1 cal mol^{-1} K^{-1} and the largest lost 14 ± 2 cal mol^{-1} K^{-1} to such correlations.

Despite the large uncertainties in individual values, our calculations demonstrate that the change of conformational entropy should have a considerable influence on the overall stability of protein complexes. It is perhaps fundamentally impossible to make general statements as to the sign of this contribution. What can be concluded is that protein association may not only deplete, but also boost, conformational entropy. Larger complexes especially seem to be able to compensate for the loss of conformal diversity occurring in the binding region. Partner proteins may experience yet higher changes in their individual dynamics, and this might, in some cases, influence their function.

Attempts are often made at calculating conformational entropy from direct measurements of order parameters of the peptide bond plane in NMR relaxation experiments. For comparison, Table 3 also provides the change of conformational entropy calculated only for fluctuations of backbone carbonyl carbon and oxygen, which can differ substantially from the all-atom entropy. The two values were correlated with R = 0.8 (P = 0.03), and the backbone calculation was strongly biased toward gains of entropy. Backbone atoms make fewer direct contacts with the binding partner and are hence less likely to lose mobility upon binding. They nevertheless benefit just like side chain atoms from flexibility gains outside the contact interface. Entropies calculated from experimental data that are based on subsets of atoms have thus to be interpreted with care. Unfortunately, there are no relaxation data available for any direct comparison with the values in Table 3.

Table 4. Computational Estimates and, Where Available, Experimental Measurements of Overall Binding Entropies

	ΔS					
	Conf[a]	t,r[b]	Solv[c]	Total[d]	Exp[e]	ΔRes[f]
c02	19 ± 50	−100	334	253 ± 50		0
c06	−101 ± 46	−105	135	−71 ± 46		0
c11	−157 ± 57	−101	242	−12 ± 57	(−34[g])[h]	2
c15	−60 ± 49	−98	207	49 ± 49	−1[i]	2
c16	−0.7 ± 41	−103	422	318 ± 41	20[j]	−1
c17[k]	−147 ± 78	−108	587	332 ± 78		−4
c20	49 ± 48	−109	556	496 ± 48		12

All values give the difference of bound to free states (in cal mol^{-1} K^{-1} at 1 M standard state).
[a] Conformational entropy.
[b] Rotational and translational entropy.
[c] Solvent entropy estimated from buried accessible surface.
[d] Conf + t,r + Solv.
[e] Measured entropy change.
[f] Number of residues resolved in the bound, but not in the free, structure (−, vice versa).
[g] Bhat et al., 1994; Sundberg et al., 2000.
[h] Measured for a related molecule.
[i] Frisch et al., 1997.
[j] Arold et al., 1998.
[k] Not converged.

Overall Entropy of Binding

Overall entropy changes due to binding can be reliably measured by calorimetry. Such data are available for the binding of Barnase to Barstar (c11) (Frisch et al., 1997) as well as for the HIV-1 $Nef_{\Delta 1-57}$ to the SH3 domain of Fyn (c15) (Arold et al., 1998). Data have also been published for the interaction between the mouse antibody fragment FvD1.3 and hen egg white lysozyme (Bhat et al., 1994; Sundberg et al., 2000), but not for the variant studied here (c06) that involves an artificial hybrid of FvD1.3 and human antibody segments (Holmes et al., 1998). The comparison of experimental and theoretical entropies of c11 and c15 is complicated by unresolved residues in the structures of both systems. A total of 18 terminal and 30 nonterminal residues of the free $Nef_{\Delta 1-57}$ as well as 15 terminal and 31 nonterminal residues of the bound $Nef_{\Delta 1-57}$ are disordered and are hence not present in the simulations. This putative disorder-order transition of effectively two residues of $Nef_{\Delta 1-57}$ as well as two terminal residues within the Barnase/Barstar complex is also not reflected in the entropy calculations. Salt concentrations, different protonation states, and other details may introduce additional inaccuracies.

In contrast to the calculations, the experimental values also include the entropic contribution from the solvent. The change of solvent free energy is commonly estimated from the accessible surface area, ΔASA, buried upon folding or binding as $\Delta G_{solvent} = \gamma \Delta ASA$ (Brady and Sharp, 1997), where $\gamma \approx 47$ cal mol^{-1} $Å^{-2}$ (Sharp et al., 1991; Noskov and Lim, 2001). Around room temperature, it has been shown to be largely of entropic nature (Privalov and Gill, 1988). Table 4 combines the

changes of vibrational (conformational) entropies given in Table 3 with rotational and translational entropies (Gohlke et al., 2003) as well as the estimated gain of solvent entropy and compares this overall value with the three available experimental binding entropies.

All experimental values fall within the broad error range of the calculated entropies. In fact, certain deviations are to be expected: ignoring the putative ordering of residues upon binding may introduce a systematic error. Furthermore, compared to mouse FvD1.3, the humanized antibody studied here requires additional conformational adjustments to bind its target (Holmes et al., 1998). The entropy loss of c06 may therefore indeed be larger than measured for the complex of mouse FvD1.3. Also, the estimate of solvent free energy (and entropy) from solvent-accessible areas is disputed (Jackson and Sternberg, 1995; Kyte, 2003) and is not considered to be accurate (Gohlke and Case, 2004). In spite of the many uncertainties, our calculations are compatible with the experimentally observed binding entropies.

Mechanism of Protein-Protein Recognition

Our results (as well as several of the experimental studies cited in the Introduction) question both the induced fit (Koshland, 1958) and the conformer selection model (Monod et al., 1965; Kumar et al., 2000): according to both models, binding would always need to overcome a loss of conformational entropy. This apparent contradiction between the entropic cost of specific recognition and an eventually possible entropic gain is resolved by our previously proposed model (Grünberg et al., 2004). We suggested to combine the rival ideas of induced fit and conformer selection into a unified model that describes the process in three steps: (1) diffusion, (2) recognition between complementary free conformers, and (3) relaxation into the bound conformation. By uncoupling the process of recognition (1 and 2) from the search for the bound conformation (3), recognition can transiently restrict conformational entropy, which is eventually regained in the bound state.

The preexisting equilibrium between high- and low-affinity conformations (Kirschner et al., 1966; Austin et al., 1975; Lancet and Pecht, 1976; Foote and Milstein, 1994; Leder et al., 1995; Berger et al., 1999; James and Tawfik, 2005) should hence not simply be taken for an equilibrium between the free and bound conformations. Unfortunately, kinetic experiments cannot usually monitor association and conformational change in parallel. However, two recent reports hint at further complexity and support our unified model: an unorthodox protein-protein binding study by Becker et al. (2003) revealed a biphasic fluorescence signal composed of a fast concentration-dependent step and a slower concentration-independent process. Very recently, James and Tawfik (2005) studied the binding of a small molecule to an antibody. They determined structures of different free antibody conformations, binding intermediates, and the bound state, and they performed kinetic experiments. They arrived at a mechanistic model that is identical to the one proposed by Grünberg et al. (2004). Protein-protein and protein-small molecule recognition may thus, on the protein side, operate on common principles. It should be noted however that we are here studying transient interactions between globular proteins. The picture may look different for interactions involving unstructured, nonglobular partners (Shoemaker et al., 2000; Levy et al., 2005).

Conclusions

In the majority of cases examined here, binding did not impair a protein's overall flexibility, and, ultimately, conformational entropy could be lost but also gained. In fact, this entropic contribution should constitute a major component (positive or negative) in the balance of forces governing the stability of protein complexes. Substantial shifts in the dynamics of individual partner proteins lead to large, and often opposing, negative or positive entropy contributions. These are overlaid by significant residual intermolecular motions that generally compensate for about half of the loss of translational and rotational entropy. This complex picture is at odds with the simplifying traditional models of the protein-protein binding process and indicates that specific recognition is not necessarily dependent on the search for the bound conformation. Furthermore, the large effect of molecular flexibility may generally hamper the evaluation of interaction energies from static structures.

Structure fluctuations have an important impact on the thermodynamics of protein-protein interaction. The proper treatment of molecular dynamics may well be the key to our understanding of this fundamental process.

Experimental Procedures

We integrated and automated the miscellaneous tasks of data management, analysis, and handling of external programs in a modular object-oriented Python library that will be published elsewhere.

MD Simulations

The X-Plor (Brünger, 1992) simulations performed on 33 free proteins (c06 and c08 share the same ligand) have been described before (Grünberg et al., 2004) and were here extended to the 17 complexes. In short, the proteins were surrounded by a 9 Å layer of water and subjected to 10 parallel, independent simulations of 50 ps length each, summing up to 500 ps total simulation time per molecule.

Extended simulations were performed on 7 complexes and their free components (21 protein structures) with Amber 7.0 (Pearlman et al., 1995) by using the modified all-atom force field parm98 (Wang et al., 2000). Hydrogens and waters were removed, and the proteins were subjected to a WhatIf hydrogen bond network optimization (Hooft et al., 1996). Breaks and premature ends of peptide chains (due to unresolved residues) were capped with a N-methylamine or acetyl group, but no attempts were made to model missing residues. We retained GTP, GDP, or ATP nucleotides (Meagher et al., 2003) and Mg^{2+} ions (this concerns c17 and c20). Hydrogens were added, and each protein was surrounded by a box of TIP3P water (Jorgensen et al., 1983) with at least a 10 Å distance between the protein and the edge of the box. Net charges were neutralized with Na^+ or Cl^-. The solvent was minimized while keeping all protein coordinates restrained.

The following simulation protocol was applied to ten independent copies of each molecule. Bond lengths involving hydrogen atoms were fixed with the SHAKE algorithm (van Gunsteren and Berendsen, 1977). Periodic boundary conditions were applied with a direct-space nonbonded cutoff of 9 Å and particle mesh Ewald (PME) treatment of long-range electrostatic forces (Essmann et al., 1995). The solvent was heated to 300 K over 10 ps NVT MD and was equilibrated with 10 ps NPT MD at 300 K, keeping the protein coordinates harmonically restrained (K = 50 kcal mol^{-1}) and applying an integration time step of 1 fs and temperature control (time constant of 0.5 ps) (Berendsen et al., 1984). Restraints on the solute were then stepwise released during 20 ps NPT MD. During the entire

40 ps equilibration phase, velocities were reassigned every 1 ps from a Maxwell distribution. A production MD of 1 ns was then performed by using an integration time step of 2 fs under NVT conditions at 300 K with the default time constant of 1 ps for heat bath coupling. This protocol of minimization, equilibration, and simulation was automated, parallelized, and applied in identical fashion to all 21 proteins and resulted in 10 independent 1 ns trajectories for each free receptor, free ligand, and complex.

We calculated the trace of mean C_α distance to the last structure for each single 1 ns simulation and determined the gradient of this distance over the last 500 ps by a linear least-squares fit (excluding the last 50 ps). Single trajectories were classified as outliers if their gradient fell 1.5 standard deviations below the average of all 10 simulations (this does not apply to the 10 × 50 ps simulations).

Water molecules, hydrogens, nonprotein atoms, and any atom not present in both the free and bound structures were removed prior to the following analysis.

Flexibility

Flexibility was defined as the average pairwise distance between simulation snapshots. Snapshots were extracted from the last 30 ps of the shorter 10 × 50 ps simulations in an interval of 2 ps. The rmsd was calculated (after individual least-squares fitting) between every pair of structures that did *not* stem from the same 50 ps trajectory. The same procedure was applied to the longer 10 × 1 ns simulations. Here, snapshots were taken in an interval of 5 ps from the last 500 ps of each (nonoutlier) 1 ns trajectory.

The protein surface was defined as any atom participating in the molecular surface of free or bound structures as calculated with FastSurf (Tsodikov et al., 2002) by using a probe radius of 1.4 Å. The binding interface was defined as any surface atom within 6 Å of the other molecule after the superposition of the two free structures on the bound state.

Entropy Calculations

Entropy differences were determined from a combination of several quasiharmonic calculations applied to both free and bound trajectories by using different protocols for the superposition and rearrangement of coordinate frames. Each single protocol consisted of the following steps: (1) certain single trajectories were excluded from the calculation either because they were classified as outliers or in order to estimate errors. Additional trajectories were removed arbitrarily to adjust receptor, ligand, and complex ensembles to the same number of independent trajectories (nine in most cases). (2) We removed all hydrogens and nonprotein atoms, as well as any atom not common to both the free and bound structures. Judging from complex c15, removing hydrogens introduces an error to binding entropies (−7.9 cal mol^{-1} K^{-1} difference) but was essential for scaling up the calculations. (3) Receptor coordinates were in some protocols paired up with ligand coordinates from an independent trajectory, either with or without retaining the relative time order of ligand frames. (4) The single trajectories were iteratively fitted to their respective average until the rms distance between the last and the previous average structure fell below 10^{-6} Å and were then transformed "en bloc" onto the bound state. (5) The modified set of coordinate frames was exported into Amber file format and passed to the ptraj program of Amber 8.0 (Pearlman et al., 1995) for the calculation of the mass-weighted covariance matrix and determination of vibrational, rotational, and translational entropies. We automated steps (1)–(5) and parallelized the calculation of ten different protocols, which are summarized in Table 2.

For each complex, we evaluated the convergence of free versus bound entropy differences by repeating the ten calculations by using different starting frames and different frame offsets. Errors were estimated by repeating all calculations nine or ten times, each time excluding a single trajectory from the sets of free and bound simulations. The standard error was estimated with the jackknife formula $\sigma = (n^{-1}\,[n-1]\,\Sigma[\Delta S_{n-1} - \Delta S_n]^2)^{1/2}$. About 200 individual entropy calculations were hence required per complex to evaluate the convergence and errors of the change in total conformational entropy and entropy components. The individual calculations lasted between several minutes and several hours, depending on the size of the system. Parallelization was therefore crucial.

Acknowledgments

J.L. and R.G. were supported by fellowships from the Knut and Alice Wallenberg Foundation and the Boehringer Ingelheim Fonds, respectively. M.N. acknowledges support from the European Union (QLG2 CT 2000-01313). We thank Wolfgang Rieping and Michael Habeck for discussions and various help and Tru Huynh for the supervision of computer hardware and software. Heather Carlson and Kristin Meagher kindly provided force field parameters for GTP and GDP.

Received: November 30, 2005
Revised: January 5, 2006
Accepted: January 6, 2006
Published: April 11, 2006

References

Arold, S., O'Brien, R., Franken, P., Strub, M., Hoh, F., Dumas, C., and Ladbury, J. (1998). RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef. Biochemistry *37*, 14683–14691.

Arumugam, S., Gao, G., Patton, B.L., Semenchenko, V., Brew, K., and Van Doren, S.R. (2003). Increased backbone mobility in β-barrel enhances entropy gain driving binding of n-timp-1 to mmp-3. J. Mol. Biol. *327*, 719–734.

Austin, R., Beeson, K., Eisenstein, L., Frauenfelder, H., and Gunsalus, I. (1975). Dynamics of ligand binding to myoglobin. Biochemistry *14*, 5355–5373.

Balog, E., Becker, T., Oettl, M., Lechner, R., Daniel, R., Finney, J., and Smith, J. (2004). Direct determination of vibrational density of states change on ligand binding to a protein. Phys. Rev. Lett. *93*, 028103.

Becker, C., Hunter, C., Seidel, R., Kent, S., Goody, R., and Engelhard, M. (2003). Total chemical synthesis of a functional interacting protein pair: the protooncogene H-Ras and the Ras-binding domain of its effector c-Raf1. Proc. Natl. Acad. Sci. USA *100*, 5075–5080.

Berendsen, H., Postma, J., Van Gunsteren, W., DiNola, A., and Haak, J. (1984). Molecular dynamics with coupling to an external heat bath. J. Chem. Phys. *81*, 3684–3690.

Berger, C., Weber-Bornhauser, S., Eggenberger, J., Hanes, J., Pluckthun, A., and Bosshard, H. (1999). Antigen recognition by conformational selection. FEBS Lett. *450*, 149–153.

Betts, M., and Sternberg, M. (1999). An analysis of conformational changes on protein-protein association: implications for predictive docking. Protein Eng. *12*, 271–283.

Bhat, T., Bentley, G., Boulot, G., Greene, M., Tello, D., Dall'Acqua, W., Souchon, H., Schwarz, F., Mariuzza, R., and Poljak, R. (1994). Bound water molecules and conformational stabilization help mediate an antigen-antibody association. Proc. Natl. Acad. Sci. USA *91*, 1089–1093.

Bosshard, H. (2001). Molecular recognition by induced fit: how fit is the concept? News Physiol. Sci. *16*, 171–173.

Bouvignies, G., Bernado, P., Meier, S., Cho, K., Grzesiek, S., Bruschweiler, R., and Blackledge, M. (2005). Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bondscalar couplings. Proc. Natl. Acad. Sci. USA *102*, 13885–13890.

Brady, G., and Sharp, K. (1997). Entropy in protein folding and in protein-protein interactions. Curr. Opin. Struct. Biol. *7*, 215–221.

Brünger, A. (1992). X-PLOR. A System for X-Ray Crystallography and NMR (New Haven, CT: Yale University Press).

Buck, E., and Iyengar, R. (2003). Organization and functions of interacting domains for signaling by protein-protein interactions. Sci. STKE *2003*, re14.

Case, D. (1994). Normal mode analysis of protein dynamics. Curr. Opin. Struct. Biol. *4*, 285–290.

Caves, L., Evanseck, J., and Karplus, M. (1998). Locally accessible conformations of proteins: multiple molecular dynamics simulations of crambin. Protein Sci. *7*, 649–666.

Essmann, U., Perera, L., Berkowitz, M., Darden, T., Lee, H., and Pedersen, L. (1995). A smooth particle mesh Ewald method. J. Chem. Phys. B *103*, 6998–7014.

Fayos, R., Melacini, G., Newlon, M., Burns, L., Scott, J., and Jennings, P. (2003). Induction of flexibility through protein-protein interactions. J. Biol. Chem. *278*, 18581–18587.

Fenimore, P., Frauenfelder, H., McMahon, B., and Young, R. (2004). Bulk-solvent and hydration-shell fluctuations, similar to α- and β-fluctuations in glasses, control protein motions and functions. Proc. Natl. Acad. Sci. USA *101*, 14408–14413.

Finerty, P.J., Jr., Muhandiram, R., and Forman-Kay, J.D. (2002). Side-chain dynamics of the SAP SH2 domain correlate with a binding hot spot and a region with conformational plasticity. J. Mol. Biol. *322*, 605–620.

Finkelstein, A., and Janin, J. (1989). The price of lost freedom: entropy of bimolecular complex formation. Protein Eng. *3*, 1–3.

Foote, J., and Milstein, C. (1994). Conformational isomerism and the diversity of antibodies. Proc. Natl. Acad. Sci. USA *91*, 10370–10374.

Forman-Kay, J.D. (1999). The 'dynamics' in the thermodynamics of binding. Nat. Struct. Biol. *6*, 1086–1087.

Frauenfelder, H., and Leeson, D. (1998). The energy landscape in non-biological and biological molecules. Nat. Struct. Biol. *5*, 757–759.

Frisch, C., Schreiber, G., Johnson, C., and Fersht, A. (1997). Thermodynamics of the interaction of barnase and barstar: changes in free energy versus changes in enthalpy on mutation. J. Mol. Biol. *267*, 696–706.

Gnanakaran, S., Nymeyer, H., Portman, J., Sanbonmatsu, K., and Garcia, A. (2003). Peptide folding simulations. Curr. Opin. Struct. Biol. *13*, 168–174.

Gohlke, H., and Case, D. (2004). Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. J. Comput. Chem. *25*, 238–250.

Gohlke, H., Kiel, C., and Case, D. (2003). Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. J. Mol. Biol. *330*, 891–913.

Gordon, H., and Somorjai, R. (1992). Fuzzy cluster analysis of molecular dynamics trajectories. Proteins *14*, 249–264.

Grünberg, R., Leckner, J., and Nilges, M. (2004). Complementarity of structure ensembles in protein-protein binding. Structure (Camb) *12*, 2125–2136.

Habeck, M., Nilges, M., and Rieping, W. (2005). Replica-exchange Monte Carlo scheme for bayesian data analysis. Phys. Rev. Lett. *94*, 018105.

Holmes, M., Buss, T., and Foote, J. (1998). Conformational correction mechanisms aiding antigen recognition by a humanized antibody. J. Exp. Med. *187*, 479–485.

Hooft, R., Sander, C., and Vriend, G. (1996). Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. Proteins *26*, 363–376.

Hsu, S., Peter, C., van Gunsteren, W., and Bonvin, A. (2004). Entropy calculation of HIV-1 Env gp120, its receptor CD4 and their complex: an analysis of configurational entropy changes upon complexation. Biophys. J. *88*, 15–24.

Jackson, R., and Sternberg, M. (1995). A continuum model for protein-protein interactions: application to the docking problem. J. Mol. Biol. *250*, 258–275.

James, L., and Tawfik, D. (2005). Structure and kinetics of a transient antibody binding intermediate reveal a kinetic discrimination mechanism in antigen recognition. Proc. Natl. Acad. Sci. USA *102*, 12730–12735.

Janin, J. (1995). Elusive affinities. Proteins *21*, 30–39.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926–935.

Kern, D., and Zuiderweg, E.R. (2003). The role of dynamics in allosteric regulation. Curr. Opin. Struct. Biol. *13*, 748–757.

Kirschner, K., Eigen, M., Bittman, R., and Voigt, B. (1966). The binding of nicotinamide adenine dinucleotide to yeast glyceraldehyde-3-phosphate dehydrogenase: temperature-jump relaxation studies on the mechanism of an allosteric enzyme. Proc. Natl. Acad. Sci. USA *56*, 1661–1667.

Kitao, A., Hayward, S., and Go, N. (1998). Energy landscape of a native protein: jumping-among-minima model. Proteins *33*, 496–517.

Koshland, D. (1958). Application of a theory of enzyme specifity to protein synthesis. Proc. Natl. Acad. Sci. USA *44*, 98–104.

Kumar, S., Ma, B., Tsai, C., Sinha, N., and Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. Protein Sci. *9*, 10–19.

Kyte, J. (2003). The basis of the hydrophobic effect. Biophys. Chem. *100*, 193–203.

Lancet, D., and Pecht, I. (1976). Kinetic evidence for hapten-induced conformational transition in immunoglobin MOPC 460. Proc. Natl. Acad. Sci. USA *73*, 3549–3553.

Leder, L., Berger, C., Bornhauser, S., Wendt, H., Ackermann, F., Jelesarov, I., and Bosshard, H. (1995). Spectroscopic, calorimetric, and kinetic demonstration of conformational adaptation in peptide-antibody recognition. Biochemistry *34*, 16509–16518.

Lee, A., Kinnear, S., and Wand, A. (2000). Redistribution and loss of side chain entropy upon formation of a calmodulinpeptide complex. Nat. Struct. Biol. *7*, 72–77.

Levy, Y., Cho, S., Onuchic, J., and Wolynes, P. (2005). A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. J. Mol. Biol. *346*, 1121–1145.

Lo Conte, L., Chothia, C., and Janin, J. (1999). The atomic structure of protein-protein recognition sites. J. Mol. Biol. *285*, 2177–2198.

Loh, A.P., Pawley, N., Nicholson, L.K., and Oswald, R.E. (2001). An increase in side chain entropy facilitates effector binding: nmr characterization of the side chain methyl group dynamics in cdc42hs. Biochemistry *40*, 4590–4600.

Meagher, K., Redman, L., and Carlson, H. (2003). Development of polyphosphate parameters for use with the AMBER force field. J. Comput. Chem. *24*, 1016–1025.

Mercier, P., Spyracopoulos, L., and Sykes, B. (2001). Structure, dynamics, and thermodynamics of the structural domain of troponin C in complex with the regulatory peptide 1–40 of troponin I. Biochemistry *40*, 10063–10077.

Minh, D.D., Bui, J.M., Chang, C.E., Jain, T., Swanson, J.M., and McCammon, J.A. (2005). The entropic cost of protein-protein association: a case study on acetylcholinesterase binding to fasciculin-2. Biophys. J. *89*, L25–L27.

Monod, J., Wyman, J., and Changeux, J. (1965). On the nature of allosteric transitions: a plausible model. J. Mol. Biol. *12*, 88–118.

Noskov, S.Y., and Lim, C. (2001). Free energy decomposition of protein-protein interactions. Biophys. J. *81*, 737–750.

Olejniczak, E.T., Zhou, M.M., and Fesik, S.W. (1997). Changes in the NMR-derived motional parameters of the insulin receptor substrate 1 phosphotyrosine binding domain upon binding to an interleukin 4 receptor phosphopeptide. Biochemistry *36*, 4118–4124.

Pearlman, D., Case, D., Caldwell, J., Ross, W., Cheatham, T., III, DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comp. Phys. Commun. *91*, 1–41.

Privalov, P.L., and Gill, S.J. (1988). Stability of protein structure and hydrophobic interaction. Adv. Protein Chem. *39*, 191–234.

Sharp, K., Nicholls, A., Fine, R., and Honig, B. (1991). Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. Science *252*, 106–109.

Shoemaker, B., Portman, J., and Wolynes, P. (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc. Natl. Acad. Sci. USA *97*, 8868–8873.

Steinberg, I.Z., and Scheraga, H.A. (1963). Entropy changes accompanying association reactions of proteins. J. Biol. Chem. *238*, 172–181.

Stivers, J.T., Abeygunawardana, C., Mildvan, A.S., and Whitman, C.P. (1996). 15N NMR relaxation studies of free and inhibitor-bound

4-oxalocrotonate tautomerase: backbone dynamics and entropy changes of an enzyme upon inhibitor binding. Biochemistry *35*, 16036–16047.

Sundberg, E., Urrutia, M., Braden, B., Isern, J., Tsuchiya, D., Fields, B., Malchiodi, E., Tormo, J., Schwarz, F., and Mariuzza, R. (2000). Estimation of the hydrophobic effect in an antigen-antibody protein-protein interface. Biochemistry *39*, 15375–15387.

Teeter, M., and Case, D. (1990). Harmonic and quasiharmonic descriptions of crambin. J. Phys. Chem. *94*, 8091–8097.

Tidor, B., and Karplus, M. (1994). The contribution of vibrational entropy to molecular association. The dimerization of insulin. J. Mol. Biol. *238*, 405–414.

Tsodikov, O.V., Record, M.T., Jr., and Sergeev, Y.V. (2002). Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. J. Comput. Chem. *23*, 600–609.

van Gunsteren, W., and Berendsen, H. (1977). Algorithms for macromolecular dynamics and constraint dynamics. Mol. Phys. *34*, 1311–1327.

Vergani, B., Kintrup, M., Hillen, W., Lami, H., Piemont, E., Bombarda, E., Alberti, P., Doglia, S.M., and Chabbert, M. (2000). Backbone dynamics of Tet repressor alpha8intersectionalpha9 loop. Biochemistry *39*, 2759–2768.

Viñals, J., Kolinski, A., and Skolnick, J. (2002). Numerical study of the entropy loss of dimerization and the folding thermodynamics of the GCN4 leucine zipper. Biophys. J. *83*, 2801–2811.

Wang, C., Pawley, N., and Nicholson, L. (2001). The role of backbone motions in ligand binding to the c-Src SH3 domain. J. Mol. Biol. *313*, 873–887.

Wang, J., Cieplak, P., and Kollman, P. (2000). How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J. Comput. Chem. *22*, 1048–1074.

Yu, L., Zhu, C.X., Tse-Dinh, Y.C., and Fesik, S.W. (1996). Backbone dynamics of the C-terminal domain of *Escherichia coli* topoisomerase I in the absence and presence of single-stranded DNA. Biochemistry *35*, 9661–9666.

Yun, S., Jang, D.S., Kim, D.H., Choi, K.Y., and Lee, H.C. (2001). 15N NMR relaxation studies of backbone dynamics in free and steroid-bound Delta 5–3-ketosteroid isomerase from *Pseudomonas testosteroni*. Biochemistry *40*, 3967–3973.

Zhu, L., Hu, J., Lin, D., Whitson, R., Itakura, K., and Chen, Y. (2001). Dynamics of the Mrf-2 DNA-binding domain free and in complex with DNA. Biochemistry *40*, 9142–9150.

Zidek, L., Novotny, M.V., and Stone, M.J. (1999). Increased protein backbone conformational entropy upon hydrophobic ligand binding. Nat. Struct. Biol. *6*, 1118–1121.

Structure *14*, 1205, July 2006

Erratum

Flexibility and Conformational Entropy in Protein-Protein Binding

Raik Grünberg,[1] Michael Nilges,[1] and Johan Leckner[2,*]
[1] Unité de Bioinformatique Structurale
CNRS URA 2185
Institut Pasteur
25-28 rue du docteur Roux
F-75015 Paris
France
[2] Department of Chemical and Biological Engineering
Molecular Biotechnology
Chalmers University of Technology
P.O. Box 462
SE-405 30 Göteborg
Sweden

*Correspondence: johan.leckner@chalmers.se

(Structure *14*, 683–693; April 2006)

In Table 4 of this paper, all three values in the sixth column (with the heading "Exp") should be shifted one position upward. The corrected table is below.

Table 4. Computational Estimates and, Where Available, Experimental Measurements of Overall Binding Entropies

	ΔS					
	Conf[a]	t,r[b]	Solv[c]	Total[d]	Exp[e]	ΔRes[f]
c02	19 ± 50	−100	334	253 ± 50		0
c06	−101 ± 46	−105	135	−71 ± 46	(−34[g])[h]	0
c11	−157 ± 57	−101	242	−12 ± 57	−1[i]	2
c15	−60 ± 49	−98	207	49 ± 49	20[j]	2
c16	−0.7 ± 41	−103	422	318 ± 41		−1
c17[k]	−147 ± 78	−108	587	332 ± 78		−4
c20	49 ± 48	−109	556	496 ± 48		12

All values give the difference of bound to free states (in cal mol^{-1} K^{-1} at 1 M standard state).
[a] Conformational entropy.
[b] Rotational and translational entropy.
[c] Solvent entropy estimated from buried accessible surface.
[d] Conf + r,t + Solv.
[e] Measured entropy change.
[f] Number of residues resolved in the bound, but not in the free, structure (−, vice versa).
[g] Bhat et al., 1994; Sundberg et al., 2000.
[h] Measured for a related molecule.
[i] Frisch et al. 1997.
[j] Arold et al. 1998.
[k] Not converged.

DOI: 10.1016/j.str.2006.06.003