Published in final edited form as:
Proteins. 2004 September 1; 56(4): 828–838. doi:10.1002/prot.20131.

Evaluation of the Relative Stability of Liganded vs. Ligand-Free Protein Conformations Using Simplicial Neighborhood Analysis of Protein Packing (SNAPP) Method

Douglas B. Sherman[1,≠], **Shuxing Zhang**[2,≠], **J. Bruce Pitner**[1], and **Alexander Tropsha**[2,*]

[1]BD Technologies, 21 Davis Dr., Research Triangle Park, NC 27709

[2]The Laboratory for Molecular Modeling, Division of Medicinal Chemistry and Natural Products, School of Pharmacy, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7360

Abstract

Many proteins change their conformation upon ligand binding. For instance, bacterial periplasmic binding proteins (bPBPs) that transport nutrients into the cytoplasm generally consist of two globular domains connected by strands forming a hinge. During ligand binding, hinge motion changes the conformation from the open to the closed form. Both forms can be crystallized without a ligand, suggesting that the energy difference between them is small. We applied Simplicial Neighborhood Analysis of Protein Packing (SNAPP) as a method to evaluate the relative stability of open and closed forms in bPBPs. Using united residue representation of amino acids, SNAPP performs Delaunay tessellation of the protein, producing an aggregate of space-filling, irregular tetrahedra with nearest neighbor residues at the vertices. The SNAPP statistical scoring function is derived from log-likelihood scores for all possible quadruplet compositions of amino acids found in a representative subset of the Protein Data Bank, and the sum of scores for a given protein provides the total SNAPP score. Results of scoring for bPBPs suggest that in most cases, the unliganded form is more stable than the liganded form, and this conclusion is corroborated by similar observations on other proteins undergoing conformation changes upon binding their ligands. The results of these studies suggest that the SNAPP method can be used to predict relative stability of accessible protein conformations. Furthermore, the SNAPP method allows for delineation of the role of individual residues in protein stabilization, thereby providing new testable hypotheses for rational site-directed mutagenesis in the context of protein engineering.

Keywords

Delaunay tessellation; periplasmic binding proteins; conformational stability change; differential SNAPP profile analysis; ligand binding

Introduction

The soluble periplasmic binding proteins of Gram negative bacteria play an important role in the delivery of nutrients from the periplasmic space into the cytoplasm. These bacterial periplasmic binding proteins (bPBPs) bind small molecules, such as sugars, amino acids, and small peptides, and then transport the ligands to permeases bound to the cytoplasmic

*Corresponding author, School of Pharmacy, Campus Box 7360, 327 Beard Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7360. Telephone (919) 966-2955, FAX: (919) 966-0204, Alex_Tropsha@unc.edu.
≠These authors equally contributed to this paper.

membrane. The permeases interact with the bPBPs to release the ligands for translocation into the cytoplasm.[1,2] Although bPBPs vary considerably in size and share little homology,[3] they generally adopt the same structural motif, which consists of two domains connected by two or three peptide strands. The strands act as a hinge that enables the domains to open and close in a manner analogous to a Venus flytrap. When the ligand enters the cleft between the two domains in the open conformation, torsional angles in the hinge region change to bring the domains together to entrap the ligand in the closed conformation.[4] Binding constants for bPBPs are generally in the micromolar range, and the ligands are held in place by hydrogen bonds and van der Waals interactions.[5]

The existence of open and closed conformations for bPBPs has been demonstrated by X-ray crystallography for several proteins, including arabinose-binding protein (ABP),[6] leucine/isoleucine/valine-binding protein (LIVBP),[7] leucine-binding protein (LBP),[8] histidine-binding protein (HBP),[9,10] molybdate-binding protein (ModA),[11] sulphate-binding protein (SBP),[12,13] phosphate-binding protein (PBP),[14–16] galactose/glucose-binding protein (GGBP),[17–20] ribose-binding protein (RBP),[21–24] glutamine binding protein (GlnBP),[25–27] lysine/arginine/ornithine-binding protein (LAOBP),[28,29] allose-binding protein (ALBP),[30,31] *Haemophilus influenzae* Fe^{3+}-binding protein (hFBP),[32,33] vitamin B-12-binding protein (BtuF),[34] *Treponema pallidum* Zn^{2+}-binding protein (TroA),[35,36] dipeptide-binding protein (DppA),[37–39] oligopeptide-binding protein (OppA),[40–45] and maltose-binding protein (MBP).[46–51] In addition, small-angle X-ray scattering (SAXS) experiments for ABP,[52] RBP, MBP, and GGBP[53] have shown that the radius of gyration for these proteins decreases upon ligand binding, which is consistent with a change in overall protein conformation.

Two different models have been proposed to explain the mechanism of action for bPBPs. In the first model,[52] the protein exists in a stable open form, and interaction of the ligand with the protein then triggers a conformational change to bring about closure. In the second model,[29] the protein is in a dynamic equilibrium between open and closed forms in the absence of ligand, which allows for the existence of a closed unliganded conformation. Flocco and Mowbray[17] obtained an X-ray crystal structure of GGBP in a closed unliganded conformation, which supports this hypothesis. Disulfide-trapping experiments[54] between the two domains of GGBP in the absence of ligand and NMR experiments with 5-fluorotryptophan or 3-fluorophenylalanine mutations in GGBP[55] also support the model of dynamic motion of the protein.

The ability to produce crystal structures of open and closed conformations of bPBPs in the absence of ligands and the results from the disulfide trapping studies suggested that the difference in energy between the open and closed forms is small. Calculations of adiabatic bending energy ($E = E_H$ (hinge region) + E_L (remaining lobe + lobe-hinge interactions)) and net free energy on ABP by Mao *et al.*[56] implied that the open conformation should be more stable than the closed conformation by approximately 40 kcal/mol in the absence of ligand. Binding of arabinose would decrease the energy difference by 60 kcal/mol through displacement of waters and would be sufficient to overcome the energy difference between the open and closed forms.

Recently, Simplicial Neighborhood Analysis of Protein Packing (SNAPP)[57–62] was introduced as a method to analyze protein packing. SNAPP is built upon Delaunay tessellation[63] of protein structures using united residue (i.e., single point) representations of amino acids. The tessellation process transforms this representation into an aggregate of space filling, irregular tetrahedra with nearest neighbor residues at the vertices. Log-likelihood scores for all observed quadruplet combinations of amino acids (which is a number close to theoretically possible 8855 compositions) have been calculated using a representative subset of the Protein Data Bank (PDB), and the total SNAPP score can be calculated by summing the scores for all

tetrahedra in the protein.[62] Studies on point mutations in several proteins showed good correlation between the change in SNAPP scores (ΔSNAPP) for wild-type and mutant proteins and $\Delta(\Delta G_{unfold})$.[58] Recent studies in our group[59] have also demonstrated that the SNAPP potential was capable of discriminating between pre-transition state, post transition state, and native conformations of the chymotrypsin inhibitor 2 (CI2) protein (PDB code 2CI2) obtained in the course of folding simulations by Li and Shakhnovich.[64] Thus, we showed that the SNAPP score of pre-transition state was lower than that of the post-transition state, and that the native conformation of CI2 had the highest SNAPP score. We have also demonstrated for a number of proteins that SNAPP potentials successfully ranked native structures higher than their multiple conformational decoys.[59] These recent calculations suggest that the SNAPP score not only correlates with the stability change caused by mutations but it can be used to characterize conformational stability of proteins as well.

In this paper, we have applied the SNAPP method to evaluate the relative stabilities of the open and closed forms of several bPBPs. We find that in the majority of cases, the results of calculations are in a good agreement with experimental observations suggesting that the open (unliganded) conformation is more stable than the closed (liganded) one without taking protein-ligand interactions into account. In addition, we have extended the study to other binding proteins that undergo conformational change to demonstrate the generality of our conclusions based on the results for bPBPs.

Materials and Methods

X-ray crystallographic coordinates of the proteins in this study were taken from the Protein Data Bank.[65] For GlnBP, the coordinates for residues 4, 225, and 226 in 1WDN were removed so that the protein would contain the same residues as 1GGG. A similar modification was made for calmodulin structures 1CFD (calmodulin), 1A29, and 1CLL to contain residues 4 through 146, which all three have in common. The selenomethionine residues in 1N4A and 1N4D were replaced with methionine prior to calculations of the SNAPP scores. Finally, the coordinates for the lysine ligand in 1LST were removed, since the ligand was assigned to the A chain in the PDB file and would be considered by the SNAPP program as part of the protein structure. For 1URP, 1GGG, 1TOA, 1N4A, 1N4D, 1DPP, and 1EZ9, more than one X-ray structure for the same conformation was available in the PDB files. Scores were examined for all such structures, and no significant differences were noted within a given PDB file. Consequently, only the results for the A chains have been reported.

SNAPP scores were calculated as previously described.[58] Briefly, Delaunay tessellation was performed on a set of ca. 1200 high-resolution protein structures from the PDB that had been selected for structural diversity (culled PDB database by R. Dunbrack: http://www.fccc.edu/research/labs/dunbrack/culledpdb.html).[57] The resulting tessellation generated tetrahedra with nearest neighbor residues at the vertices. The log-likelihood of four residues i, j, k, and l being in a tetrahedron was calculated as $q_{ijkl} = \log(f_{ijkl}/ p_{ijkl})$, where f_{ijkl} is the frequency of occurrence of a quadruplet in the set of proteins, and p_{ijkl} is the expected frequency of occurrence of a given quadruplet based on the amino acid frequencies in the set of proteins. In this manner, scores were assigned for nearly all of the 8855 possible amino acid quadruplets.

Proteins in this study were analyzed using a reduced, residue-based representation by replacing the residues with side chain centroids. Delaunay tessellation was performed to generate tetrahedra of nearest neighbors, and the total SNAPP score for a given protein was calculated by summing the log-likelihood scores for all tetrahedra involving contacts between residues in the protein that are not immediately consecutive in the primary sequence. The complete SNAPP method has been implemented in an interactive web interface available at

http://mmlsun4.pha.unc.edu/3dworkbench.html. Total SNAPP scores were calculated using the MuSE module (Mutation with SNAPP Evaluation), which reports the value as the native score. SNAPP scores per residue for each protein were obtained from the ProCAM module, which also provided a list of all tetrahedra in which a given residue participated. Differences in score per residue were then calculated for the open and closed forms and were plotted against residue number to obtain the differential SNAPP score plots. An arbitrary cutoff threshold of ±2 units was employed to distinguish residues that contribute most to the total SNAPP score. Molecular models for the figures were created using Sybyl version 6.7 (Tripos, Inc., St. Louis, MO).

Results

Although X-ray crystal structures are available for a number of bPBPs, we chose for our studies the following proteins that have been crystallized as wild-type in both the open and closed forms: RBP, GlnBP, ALBP, hFBP, LAOBP, BtuF, TroA, MBP, DppA, and OppA. Multiple conformations of the open unliganded form were available for ALBP and MBP, and multiple closed forms containing different ligands were available for LAOBP, OppA, and MBP. Finally, open forms of MBP were available that had been co-crystallized with a ligand. All of these crystal forms were included in this study, and the PDB codes, ligands, and total SNAPP scores are reported in Table I. Total SNAPP scores were calculated according to the procedure given in Materials and Methods. Ligands were not included in the calculations, since SNAPP operates directly on the protein structure. Consequently, the total SNAPP scores provided a measure of stability of the proteins only in each conformation, and the conformation with the higher total SNAPP score would be expected to be more stable in the context of the protein structure alone.

bPBPs Binding Monosaccharides, Amino Acids, Vitamins, and Ions

Figure 1a illustrates the results for the open and closed conformations of RBP, GlnBP, ALBP, hFBP, LAOBP, BtuF, and TroA, which bind the smallest ligands, namely monosaccharides, amino acids, vitamins, and ions such as Fe^{3+} or Zn^{2+}. The proteins in this group (with the number of residues in the range from 226 to 309) gave total SNAPP scores ranging from approximately 40 to 90 units. Although the magnitude of the total SNAPP score is dependent on the number of residues to a certain extent, the value of the score is also a composite of favorable and unfavorable packing scores for the individual residues (see below). For example, RBP has 271 residues, which is near the middle of this group, but gave the lowest score. GlnBP and hFBP represent the smallest and largest of bPBPs in this group, having 226 and 309 residues respectively, yet both scored approximately 55 SNAPP units. BtuF and TroA have 242 and 276 residues, respectively, but scored the highest. The latter two proteins represent a subclass of bPBPs in which the two domains are connected by a single α-helix of approximately 30 amino acids instead of two or three peptide strands. Thus the magnitude of the scores is more a function of protein packing than of sequence length.

Of the proteins in this group, RBP, GlnBP, ALBP, hFBP, and BtuF all had higher total SNAPP scores in the open unliganded form than in the closed liganded form. The liganded form of TroA scored higher than the unliganded form; however, TroA is unusual, because the liganded form is actually more open than the unliganded form.[36] Consequently, the results are still consistent with the rest of the group, since the open form scored higher than the closed form. A similar result was observed for representative structures of DppA and OppA (Figure 1b), and MBP generally followed this trend as well, as will be discussed below. Examination of the scores in Table I shows that the difference was approximately 3.1 units for RBP, 3.7 units for GlnBP, 6.4 units for hFBP (nearly double that of RBP and GlnBP), 0.4 units for BtuF, and 2.4 units for TroA. The three open unliganded forms of ALBP differed in the angle of opening of the binding cleft (43°, 37°, and 33°), and interestingly, the total SNAPP score decreased in

accordance with the degree of opening, from 10.7 to 9.4 to 6.1. For LAOBP, on the other hand, the three closed forms containing Arg, Lys, or Orn as ligands scored as high or higher than the open form, with differences of 2.6, 1.6, and 0.04, respectively. The closed form containing His scored lower than the open form by 1.1 SNAPP units. Except for LAOBP, the higher scoring of the open forms relative to the closed forms suggested that the packing of the residues in the open form was more stable than that of the closed form when the ligand was not taken into account.

To better understand the difference in the SNAPP scores between the open and closed conformations, differential SNAPP profiles were generated to characterize the role of individual residues in the structures. SNAPP scores per residue were calculated for each conformation as described in Materials and Methods, and the difference in scores per residue was then plotted against residue number. Figure 2a illustrates the differential profile for ALBP, for which the 43° open (1GUD A chain) and closed forms (1RPJ) were compared. The figure shows that a relatively small number of residues (K9, F15, I23, D91, K93, N114, A146, L188, D227, and Q247, with L90 scoring close to the arbitrary threshold) contributed the most to the difference in total SNAPP scores and thus to the relative stability of the protein. Of these, only I23 and L90 scored more favorably in the closed form, meaning that the majority of the significant contributors favored the open conformation, and all residues except I23 and L188 were located in or near the binding pocket.

The tessellation pattern for D91, as depicted in Figure 2b, illustrates how residues in one domain of the binding pocket may gain additional contacts from residues in the other domain upon ligand binding, which may provide favorable or unfavorable contacts. D91 had a score of −1.1 from 5 tetrahedra in the open form and −3.9 from 12 tetrahedra in the closed form, indicating that burial of this hydrophilic residue between the domains had a destabilizing effect as would be expected. Similarly, L90 had a score of 4.5 from 10 tetrahedra in the open form and 6.4 from 12 tetrahedra in the closed form, which indicated that burial of hydrophobic L90 in the binding pocket upon closing contributed favorably to stability, which would also be expected. D91 forms hydrogen bonding contacts with allose in the closed form, which provides offsetting stability for this residue in the binding pocket.

The effect from residues I23 and L188 suggested that repacking of the protein also affects the stability of the structures. I23 had a score of 6.8 in the open form from 11 tetrahedra and 9.3 in the closed form from 13 tetrahedra. The score for L188 changed from 4.7 in 11 tetrahedra in the open form to 2.6 in 9 tetrahedra in the closed form. Thus, even though these residues were located away from the binding site, their change in scoring and number of interactions indicated that the conformational change had effects on packing within the domains themselves.

A differential profile analysis was also performed for LAOBP to understand why the open form scored lower than most of the closed forms. Unlike ALBP, for which the majority of significant residues had more positive scores in the open form than in the closed (Figure 2a), the opposite was true for LAOBP, as seen in Figure 3a for 2LAO (open) and 1LAF (closed). Furthermore, only two of the residues along the rim of the binding pocket (D51 and A141) contributed significantly to the difference in scores. Most of the residues contributing to the difference were located near the hinge (A89, K186, and F191) or near the C-terminal strand of the protein (E162, F231, V235, and G237) in a region where the two domains come in close contact in the open conformation. Thus, the instability of the open form in LAOBP relative to the closed form appears to arise from unfavorable contacts formed between the domains behind the hinge, a phenomenon not seen in the other bPBPs.

Since SNAPP scores are derived from the combination of the number of contacts between residues and the log-likelihood of tetrahedral quadruplet compositions of those residues, we decided to dissect the results further by determining the number of contacts gained or lost per residue upon closure. These results are shown in Figure 3b for 2LAO and 1LAF. Of the three residues indicated in Figure 3a that scored more favorably for the open form, D51 and A141 gained 8 and 5 contacts, respectively, and lost none upon closure, while F231 gained no contacts and lost 2 upon closure yet scored more favorably in the open form. In the case of the remaining six residues in Figure 3b that favor the closed form, A89, E162, K186, and G237 lost more contacts than they gained upon closure. F191 and V235 each lost and gained one contact. The differential SNAPP scores for F191 (−2.2) and V235 (−2.5) were of the same order as K186 (−2.4), which however, lost 7 contacts and gained none. Taken as a whole, these results show that large changes in the number of contacts do not guarantee a large change in SNAPP score, and that even small changes in the number of contacts can have a significant effect on the score.

bPBPs Binding Oligopeptides and Oligosaccharides

DppA and OppA are two of the largest bPBPs, having 507 and 517 residues and molecular weight of 57 and 59 kD, respectively. In addition, DppA and OppA have three domains instead of two found in the other bPBPs,[37] but only domains I and III are directly involved in ligand binding. Unlike most bPBPs that bind monosaccharides or ions, OppA binds to peptides of 2–5 amino acid residues length, and both DppA and OppA can bind peptides with diverse sequences.

The peptide ligands generally adopt an extended conformation, and most of the contacts between protein and ligand are through hydrogen bonds between protein side chains and the peptide backbone. Ordered waters in the binding pocket also play an important role in binding,[45] especially in the region around position 2 of the peptide. These waters serve to fill space depending on the size of the residue, provide hydrogen bonding contacts, and dissipate charge from the ligand.

The results for DppA and OppA are recorded in Table I. A total of 31 closed form crystal structures were available for OppA with a number of different peptides as ligands. As previously mentioned, the results in Figure 1b showed that the open forms in both DppA and OppA had higher SNAPP scores than their respective closed forms, again suggesting that the open forms were more stable than the closed form when the ligand is not taken into account. The difference was 8.1 to 8.8 units for DppA and 2.2 to 15.7 units for OppA. Thus, the OppA closed structures scored across a range of 13.6 SNAPP units, even though the backbone RMSD of the complexes was less than 0.6 Å.

Differential SNAPP analysis for OppA was performed by comparing the differences in scores for a number of closed forms, for example, 1B3G (total score 120.7) and 1B3L (total score 107.3). Here, L348 (2.6) and R237 (2.0) were the only residues that met our cutoff of ±2 units for significant contribution. Several residues scored above ±1.7 units, however, namely A1 (1.8), F74 (1.7), L113 (1.7), F155 (1.8), Y156 (2.0), I218 (1.7), Y236 (−1.9), F353 (1.7), V388 (1.7), and P423 (1.8). Residues Y236 and R237 were in or near the binding pocket, and residues F74, L113, F155, and Y156 were located in the domain that does not directly participate in binding of the ligand. The remaining residues were located in the hydrophobic cores of the two domains that bind to peptides. Similar results were obtained in comparisons between other pairs of closed structures, such as 1JEV (111.8) and 1B3L (107.2) or 1B3G (120.7) and 1QKB (114.6). The latter pairing had a much smaller number of residues with scores above ±1.7 units (F69, W72, L113, L148, F155), and these were all located in the domain not involved in ligand binding. Since OppA can bind a number of different sequences of differing lengths, it is likely that the protein repacks the hydrophobic cores of the domains to accommodate the ligands while retaining binding in nanomolar to micromolar range.

MBP is capable of binding a number of different maltodextrins as demonstrated by crystal structures that include maltose, maltotriose, maltotetraose, reduced maltodextrins maltotriitol and maltotetraitol, and β-cyclodextrin. In addition, MBP is one of the few bPBPs that have been successfully co-crystallized in the open form with a ligand. The data for the MBPs are recorded in Table I, and Figure 1b shows the results for crystal structures with either no ligand or with maltodextrins. In general, the data showed that the majority of open forms had higher total SNAPP scores than the closed liganded forms. This was true whether the proteins were crystallized in the absence of ligand (1OMP and 1JW4), were co-crystallized with a ligand (1FQB, 1EZ9, and 1FQA), or were crystallized in the absence of ligand and the crystal was later saturated with ligand (1JW5). This is consistent with our results for most of the other bPBPs described above.

The two exceptions to the general trend were 1FQD, which is a closed form crystallized with maltotetraitol, and 1DMB which is an open form containing β-cyclodextrin. Differential SNAPP analysis between 1FQD and 1FQC (closed form with maltotriitol) showed high changes in scoring for L135 (−2.5), L147 (−2.7), F149 (2.4), L160 (−2.0), L195 (−2.8), and L198 (2.3), which are located in the C-terminal domain, and L285 (2.0) and L304 (2.0), which are located in the N-terminal domain. As can be seen, the distribution between positive and negative scores was nearly even, and the net difference was small. Residues I11, I79, F85, L89, V97, I108, and A264 in the C-terminal domain and D136 and P159 in the N-terminal domain, however, all had differential scores between 1 and 2 that showed more favorable packing in 1FQD than in 1FQC. Only four residues (I132, A223, L262, and L275) scored between −2 and −1 to favor the packing in 1FQC. None of these residues were located in the binding site but were instead distributed throughout the hydrophobic cores of the two domains. A similar comparison between the open form 1OMP and 1DMB showed high differential scores for L198 (3.9), W129 (2.6), L160 (2.6), F169 (2.3), and I368 (2.3) in the C-terminal domain and I79 (3.6) and F279 (2.7) in the N-terminal domain. These residues were also located away from the binding pocket, and all scored more favorably in the unliganded open form (1OMP) than in the open form bound to β-cyclodextrin. In general, these results demonstrate the effects that conformational changes, coupled with the ligand-binding, have on protein packing.

Extension of SNAPP Analysis to Other Ligand Binding Proteins

Based on the results we obtained for SNAPP analysis of bPBPs, we decided to extend our application to non-PBP structures that also display large-scale movement upon ligand binding. As is shown in Table II and Figure 4, we have applied SNAPP analysis to crystal structures of dihydrofolate reductase (DHFR),[66] adenylate kinase,[67–69] and calmodulin.[70–72] The open form of calmodulin was taken from a reported NMR structure, since no crystal structure was available.

Adenylate kinase displays a hinge-based motion that is similar to the bPBPs; however, the two domains that undergo motion are separated by a third domain forming a peptide core. DHFR undergoes the most complex motion and displays three states, namely open, closed, and occluded. The open state acts as a bridge between the closed and occluded conformations.[66] Calmodulin displayed a change of approximately 30 SNAPP units, which was the largest change observed for all proteins studied in this work. The results again support our hypothesis. Except for closed form 1RX3, which scored the highest, all closed or occluded structures had lower SNAPP scores than the open unbound forms.

The analysis of differential SNAPP profiles (data not shown) identified specific residues in the hydrophobic cores of these proteins that were contributing the most to the observed differences in the SNAPP scores. The results for calmodulin were particularly striking. Twenty-eight residues scored over 2 units in the differential profile, and 7 of these scored higher than 6 units. The large changes observed in calmodulin were consistent with the packing of this protein in

the apo and liganded states. In the bound form, calmodulin contains two globular domains that are connected by a central helix covering residues 65–92. Each domain is organized into EF-hand motifs that present a hydrophobic pocket involved in the binding of calmodulin to enzymes. In the apo state, residues 76–81 act as a flexible linker that breaks the central helix into two segments, and the helices within each domain pack more closely.[71] From the point of view of protein packing, the apo state appeared more stable, as reflected by the total SNAPP score, and the liganded state gained stability from the binding of calcium and the interaction of the hydrophobic pocket with target enzymes. Thus, the results of this additional analysis confirm the overall generality of the trend observed on bPBPs, in that the open conformations of proteins appear more stable that the closed (i.e., liganded) ones when the effect of the protein-ligand interactions is not taken into consideration.

Discussion

SNAPP provides a method of analyzing protein packing and stability that is based on unambiguous designation of nearest neighbor residue quadruplets and the associated four-body statistical pseudopotentials. In a previous study,[58] the effect of point mutations in the hydrophobic core of five proteins was examined. Good correlations ($R^2 = 0.86$) were obtained between the change in free energy of unfolding ($\Delta(\Delta G_{unfold})$) for the proteins and the change in SNAPP score (ΔS_{SNAPP}) arising from the mutation. Correlations could be improved further using the average SNAPP score per residue, <SNAPP>, to take into account protein specific effects. As discussed in the Introduction, previous studies have also demonstrated that SNAPP score correlates with protein conformational stability as well.[59] Thus, SNAPP pseudopotentials provide a method to relate protein packing with thermodynamic stability.

Here we have extended SNAPP from the analysis of individual mutations to the evaluation of the entire protein structure. In particular, we have applied SNAPP to examine proteins that undergo significant conformational changes as a result of ligand binding. Total SNAPP scores were calculated for the open and closed conformations of bPBPs and several other proteins that exhibit different conformational states in the liganded and unliganded state. Since the proteins undergo conformational change, the tessellation patterns for the open and closed forms will be different, as certain residues gain or lose spatial proximity as a result of domain movement. Consequently, the SNAPP scores per residue will also change between the two forms, which in turn will change the total SNAPP score for the protein. Given the nature of the scoring process, the conformation for a given protein having the higher total SNAPP score is expected to be more stable.

For RBP, GlnBP, ALBP, hFBP, BtuF, DppA, and OppA, the open unliganded conformation scored higher than the closed conformation, suggesting that for all of these proteins, the open unliganded conformation was more stable than the closed conformation apart from protein-ligand interactions. As previously mentioned, BtuF and TroA are structurally different, because the two domains are connected by a single long helix, whereas the domains in the other proteins are connected by two to three strands and have more flexibility. BtuF opens to a smaller extent than the other bPBPs through a rigid body rotation at residue P105 at the end of the helix in the N-terminal domain.[34] Although TroA is structurally similar to BtuF, the movement of the domains is different. In the unliganded form of TroA, the C-terminal domain tilts 4° with respect to its position in the liganded form, which results in a form that is closed more in the apo form than in the liganded form.[36] The tilting also exposes more of the hydrophobic contacts between the domains to solvent, which would be unfavorable. Thus, the open form of TroA still scored higher than the closed form, even though the open form was liganded and the closed form was unliganded.

Most of the MBP open conformations scored higher than the closed forms. Interestingly, open forms of MBP co-crystallized with ligands generally had higher SNAPP scores than the closed forms that were co-crystallized with ligands. These observations with MBP, along with the results for TroA, would argue against a systematic effect associated with the presence of the ligand that may have caused closed, liganded forms to have a lower score than open, unliganded forms. LAOBP was the one bPBP for which the closed liganded forms generally scored higher than the open unliganded form. SNAPP evaluation of other proteins that undergo conformational change upon ligand binding also suggested that in most cases, the unliganded form is more stable than the liganded form when the ligand is excluded. Overall, these results support the concept that the stability of the closed form is gained from hydrogen bonding and van der Waals interactions between the protein and the ligand, as shown by NMR and small-angle X-ray scattering experiments, but not from intramolecular interactions within the protein itself. This observation is to be expected, since if the opposite were true, then the binding of the ligand with the (more stable) closed protein conformation would be difficult if not impossible.

Differential SNAPP profile analysis provides a means to evaluate the contribution of the individual residues to the change in SNAPP score between open and closed forms. Hypothetically, these residues should then contribute most to the energetics of the conformational change, which could be tested experimentally by mutagenesis at these sites. The analysis of several proteins revealed that only a small number of residues contributed the most to the difference in scores between the two conformations. The differential analysis of L90 and D91 in ALBP also demonstrated the power of SNAPP to correlate with the expectation that burial of these hydrophobic and hydrophilic residues would result in a favorable and unfavorable stability change, respectively, in the closed form without ligands. In fact, D91 is stabilized in the closed form by hydrogen bonding interactions with allose.[30]

At first, it might be expected that most of the residues that contribute to the difference in score between conformations for bPBPs would be located along the surfaces of the domains that come together to form the binding pocket, since these residues would change tessellation patterns and SNAPP score per residue. This expectation was observed in the differential analysis of ALBP. On the other hand, the analysis of LAOBP indicated that residues behind the hinge region were also playing an important role, as shown in Figure 5. Unlike most of the other bPBPs, the large movement of the two lobes in LAOBP is the consequence of a 52° rotation of a single backbone torsion angle (ψ angle of A90) in the first connecting strand as well as distributed smaller changes of three backbone torsion angles of the second connecting strand.[29] It is noteworthy that SNAPP indicated that F231 on the C-terminal strand would score more favorably in the open form, since it is more “buried” as a result of the interdomain contacts that are formed upon opening. Since bPBPs have little sequence homology and can vary in length by almost 300 residues, the domains and folding vary significantly, even though the proteins adopt the same general tertiary form. The manner of hinge opening among the bPBPs also varies because of structure, such that interdomain contacts will not be identical in all situations.

Differential SNAPP analysis also showed how ligand interactions with the protein affected the scoring of some structures, as seen for the closed form of MBP with maltotetraitol, the open form of MBP with β-cyclodextrin, and the closed form of OppA with various peptides. In this case, differential analysis was employed to compare two open or two closed conformations with one another. For both MBP and OppA, the residues that contributed most significantly to the differences were distributed throughout the protein structures. As previously mentioned, the OppA closed structures covered a span of 13.6 SNAPP units among themselves. Attempts in the literature to correlate binding of the tripeptide ligands Lys-X-Lys with ITC thermodynamic data[40,42] and structure generally failed, although recently a QSAR model was

developed using the COMparative BINding Energy (COMBINE) approach.[73] The COMBINE model used ligand and protein desolvation energy, Coulombic interactions, and Lennard-Jones interactions as variables to correlate with ΔG, ΔH, and TΔS and took into account ordered waters found in the binding pocket. Since SNAPP only measures protein packing, it does not directly account for such waters. Rather, the scores indirectly show the influence that the ligand and other variables such as water have on the overall packing of the protein.

Since the bPBPs generally showed higher scores for the open unliganded form, we decided to examine proteins other than bPBPs that also displayed conformational motions to see if similar trends would be obtained. The conformational motion of adenylate kinase, and DHFR has been studied both computationally[74] and by X-ray crystallography, and a variety of crystal structures have been characterized that represent different conformational stages of their enzymatic cycles.[66,69] Calmodulin showed the most dramatic change between open and closed forms, both in terms of the total SNAPP score and the differential SNAPP profile analysis. Notably, the X-ray structure for the open form was not available so we used the NMR structure of the open form instead. It is possible that the dramatic difference between SNAPP scores of the open and close forms of calmodulin is due to the difference in the structural determination technique. The only other ligand-binding protein in this set for which we could find both X-ray (PDB code 1RG7) and NMR (PDB code 1AO8) structures was the closed conformation of DHFR complexed with methotrexate. In this case, the NMR structure indeed had a higher SNAPP score than the corresponding X-ray structure but only by less than three SNAPP units (data not shown). Therefore, the large difference in SNAPP scores for calmodulin is still most likely due to the conformational stability change, and not due to the effects caused by the structure determination technique. As previously mentioned, in all cases except for 1RX3, the open forms of these proteins scored higher than the closed or occluded forms without accounting for protein-ligand interactions. Although by no means exhaustive, the results found with these proteins are in agreement with our general observations for the bPBPs and illustrate the applicability of SNAPP to a variety of protein forms and conformational motions.

Conclusions

In summary, the results presented in this study are rather consistent in indicating that unliganded forms of proteins are more stable that the liganded ones (without considering protein-ligand interactions) with only a few exceptions. Future studies using an all-atom SNAPP scoring function that considers ligand-protein contacts may establish correlations with experimental binding affinities of bPBPs. Another important component of our studies is that the differential profile analysis helps to identify residues mostly responsible for the stability change. This provides suggestions for specific protein sequence modification via site-directed mutagenesis, especially in non-binding regions of the proteins where mutagenesis is unlikely to influence ligand binding directly. These computational and experimental avenues for future research are under investigation in our laboratories.

Abbreviations used

ABP, arabinose-binding protein
ADP, adenosine diphosphate
ALBP, allose-binding protein
AMP, adenosine monophosphate
AMPPNP, adenyl-5'-yl imidodiphosphate
Ap_5A, P^1, P^5-bis(adenosine-5'-)pentaphosphate
bPBP, bacterial periplasmic binding protein
BtuF, Vitamin B-12-binding protein
Chx, cyclohexylalanine

CI2, chymotrypsin inhibitor 2
COMBINE, comparative binding energy
Dab, diaminobutyric acid
Dap, diaminopropionic acid
DDF, 5,10-dideazatetrahydrofolic acid
DHFR, dihydrofolate reductase
DppA, dipeptide-binding protein
FOL, folic acid
GGBP, galactose/glucose-binding protein
GlnBP, glutamine-binding protein
HBP, histidine-binding protein
hFBP, *Haemophilus influenzae* Fe^{3+}-binding protein
Hpe, homophenylalanine
LAOBP, lysine/arginine/ornithine-binding protein
LBP, leucine-binding protein
LIVP, leucine/isoleucine/valine-binding protein
MBP, maltose-binding protein
ModA, molybdate-binding protein
MTX, methotrexate
NAP, nicotinamide adenine dinucleotide phosphate
Nle, norleucine
Npa, naphthylalanine
Nva, norvaline
OppA, oligopeptide-binding protein
Orn, ornithine
PBP, phosphate-binding protein
PDB, protein data bank
RBP, ribose-binding protein
SAXS, small-angle X-ray scattering
SBP, sulfate-binding protein
SNAPP, simplicial neighborhood analysis of protein packing
TFP, trifluoroperazine
TroA, *Treponema pallidum* Zn^{2+}-binding protein

Acknowledgments

We kindly thank Dr. Sherry L. Mowbray of the Swedish Agricultural University (Uppsula) for providing the coordinates for the open forms of ALBP prior to their deposition in the Protein Data Bank. We also thank Dr. Stephen Cammer and Mr. Yuanyuan Qiao, respectively, for development and maintenance of the SNAPP web server. SZ acknowledges an internship from BD Technologies, and AT acknowledges the support from the National Science Foundation (grant number ITR 0112896) and the North Carolina – Israel Research Partnership (NCI-SRP 1999032).

References

1. Ames GF-L. Bacterial periplasmic transport systems: structure, mechanism, and evolution. Annu Rev Biochem 1986;55:397–425. [PubMed: 3527048]
2. Furlong, CE. Osmotic-shock-sensitive transport systems. In: Neidhardt, FC.; Ingraham, JL.; Low, KB.; Magasanik, B.; Schaechter, M.; Umbarger, HE., editors. *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. Washington, D. C.: American Society for Microbiology; 1987. p. 768-796.
3. Tam R, Saier MH Jr. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. Microbiol Rev 1993;57:320–346. [PubMed: 8336670]
4. Gerstein M, Lesk AM, Chothia C. Structural mechanisms for domain movements in proteins. Biochemistry 1994;33:6739–6749. [PubMed: 8204609]

5. Quiocho FA, Ledvina PS. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. Mol Microbiol 1996;20:17–25. [PubMed: 8861200]
6. Quiocho FA, Vyas NK. Novel stereospecificity of the L-arabinose-binding protein. Nature 1984;310:381–386. [PubMed: 6379466]
7. Sack JS, Saper MA, Quiocho FA. Periplasmic binding protein structure and function: refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. J Mol Biol 1989;206:171–191. [PubMed: 2649682]
8. Sack JS, Trakhanov SD, Tsigannik IH, Quiocho FA. Structure of the L-leucine-binding protein refined at 2.4 Å resolution and comparison with the Leu/Ile/Val-binding protein structure. J Mol Biol 1989;206:193–207. [PubMed: 2649683]
9. Oh BH, Kang CH, De Bondt H, Kim SH, Nikaido K, Joshi AK, Ames GF-L. The bacterial periplasmic histidine-binding protein: structure/function analysis of the ligand-binding site and comparison with related proteins. J Biol Chem 1994;269:4135–4143. [PubMed: 8307974]
10. Yao N, Trakhanov S, Quiocho FA. Refined 1.89-Å structure of the histidine-binding protein complexed with histidine and its relationship with many other active transport/chemosensory proteins. Biochemistry 1994;33:4769–4779. [PubMed: 8161536]
11. Hu Y, Rech S, Gunsalus RP, Rees DC. Crystal structure of the molybdate binding protein ModA. Nat Struct Biol 1997;4:703–707. [PubMed: 9302996]
12. Pflugrath JW, Quiocho FA. Sulphate sequestered in the sulphate-binding protein of *Salmonella typhimurium* is bound solely by hydrogen bonds. Nature 1985;314:257–260. [PubMed: 3885043]
13. Pflugrath JW, Quiocho FA. The 2 Å resolution structure of the sulfate-binding protein involved in active transport in *Salmonella typhimurium*. J Mol Biol 1988;200:163–180. [PubMed: 3288756]
14. Luecke H, Quiocho FA. High specificity of a phosphate transport protein determined by hydrogen bonds. Nature 1990;347:402–406. [PubMed: 2215649]
15. Wang Z, Luecke H, Yao N, Quiocho FA. A low energy short hydrogen bond in very high resolution structures of protein receptor-phosphate complexes. Nat Struct Biol 1997;4:519–522. [PubMed: 9228942]
16. Yao N, Ledvina PS, Choudhary A, Quiocho FA. Modulation of a salt link does not affect binding of phosphate to its specific active transport receptor. Biochemistry 1996;35:2079–2085. [PubMed: 8652549]
17. Flocco MM, Mowbray SL. The 1.9 Å X-ray structure of a closed unliganded form of the periplasmic glucose/galactose receptor from *Salmonella typhimurium*. J Biol Chem 1994;269:8931–8936. [PubMed: 8132630]
18. Mowbray SL, Smith RD, Cole LB. Structure of the periplasmic glucose/galactose receptor of *Salmonella typhimurium*. Receptor 1990;1:41–53. [PubMed: 1967096]
19. Vyas NK, Vyas MN, Quiocho FA. Sugar and signal-transducer binding sites of the *Escherichia coli* galactose chemoreceptor protein. Science 1988;242:1290–1295. [PubMed: 3057628]
20. Zou JY, Flocco MM, Mowbray SL. The 1.7 Å refined X-ray structure of the periplasmic glucose/galactose receptor from *Salmonella typhimurium*. J Mol Biol 1993;233:739–752. [PubMed: 8240551]
21. Björkman AJ, Binnie RA, Zhang H, Cole LB, Hermodson MA, Mowbray SL. Probing protein-protein interactions: the ribose-binding protein in bacterial transport and chemotaxis. J Biol Chem 1994;269:30206–30211. [PubMed: 7982928]
22. Björkman AJ, Mowbray SL. Multiple open forms of ribose-binding protein trace the path of its conformational change. J Mol Biol 1998;279:651–664. [PubMed: 9641984]
23. Mahendroo M, Cole LB, Mowbray SL. Preliminary X-ray data for the periplasmic ribose receptor from *Escherichia coli*. J Mol Biol 1990;211:689–690. [PubMed: 2179563]
24. Mowbray SL, Cole LB. 1.7 Å X-ray structure of the periplasmic ribose receptor from *Escherichia coli*. J Mol Biol 1992;225:155–175. [PubMed: 1583688]
25. Hsiao CD, Sun YJ, Rose J, Cottam PF, Ho C, Wang BC. Crystals of glutamine-binding protein in various conformational states. J Mol Biol 1994;240:87–91. [PubMed: 8021944]
26. Hsiao CD, Sun YJ, Rose J, Wang BC. The crystal structure of glutamine-binding protein from *Escherichia coli*. J Mol Biol 1996;262:225–242. [PubMed: 8831790]

27. Sun YJ, Rose J, Wang BC, Hsiao CD. The structure of glutamine-binding protein complexed with glutamine at 1.94 Å resolution: comparisons with other amino acid binding proteins. J Mol Biol 1998;278:219–229. [PubMed: 9571045]
28. Kang CH, Shin WC, Yamagata Y, Gokcen S, Ames GF-L, Kim SH. Crystal structure of the lysine-, arginine-, ornithine-binding protein (LAO) from *Salmonella typhimurium* at 2.7-Å resolution. J Biol Chem 1991;266:23893–23899. [PubMed: 1748660]
29. Oh BH, Pandit J, Kang CH, Nikaido K, Gokcen S, Ames GF-L, Kim SH. Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. J Biol Chem 1993;268:11348–11355. [PubMed: 8496186]
30. Chaudhuri BN, Ko J, Park C, Jones TA, Mowbray SL. Structure of D-allose binding protein from *Escherichia coli* bound to D-allose at 1.8 Å resolution. J Mol Biol 1999;286:1519–1531. [PubMed: 10064713]
31. Magnusson U, Chaudhuri BN, Ko J, Park C, Jones TA, Mowbray SL. Hinge-bending motion of D-allose-binding protein from *Escherichia coli*: three open conformations. J Biol Chem 2002;277:14077–14084. [PubMed: 11825912]
32. Bruns CM, Nowalk AJ, Arvai AS, McTigue MA, Vaughan KG, Mietzner TA, McRee DE. Structure of *Haemophilus influenzae* Fe^{+3}-binding protein reveals convergent evolution within a superfamily. Nat Struct Biol 1997;4:919–924. [PubMed: 9360608]
33. Bruns CM, Anderson DS, Vaughan KG, Williams PA, Nowalk AJ, McRee DE, Mietzner TA. Crystallographic and biochemical analyses of the metal-free *Haemophilus influenzae* Fe^{3+}-binding protein. Biochemistry 2001;40:15631–15637. [PubMed: 11747438]
34. Karpowich NK, Huang HH, Smith PC, Hunt JF. Crystal structures of the BtuF periplasmicbinding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding. J Biol Chem 2003;278:8429–8434. [PubMed: 12468528]
35. Lee YH, Deka RK, Norgard MV, Radolf JD, Hasemann CA. *Treponema pallidum*TroA is a periplasmic zinc-binding protein with a helical backbone. Nat Struct Biol 1999;6:628–633. [PubMed: 10404217]
36. Lee YH, Dorwart MR, Hazlett KR, Deka RKO, Norgard MV, Radolf JD, Hasemann CA. The crystal structure of Zn(II)-free *Treponema pallidum* TroA, a periplasmic metal-binding protein, reveals a closed conformation. J Bacteriol 2002;184:2300–2304. [PubMed: 11914363]
37. Dunten P, Mowbray SL. Crystal structure of the dipeptide binding protein from *Escherichia coli* involved in active transport and chemotaxis. Protein Sci 1995;4:2327–2334. [PubMed: 8563629]
38. Dunten PW, Harris JH, Feiz V, Mowbray SL. Crystallization and preliminary X-ray analysis of the periplasmic dipeptide binding protein from *Escherichia coli*. J Mol Biol 1993;231:145–147. [PubMed: 8496961]
39. Nickitenko AV, Trakhanov S, Quiocho FA. 2 Å resolution structure of DppA, a periplasmic dipeptide transport/chemosensory receptor. Biochemistry 1995;34:16585–16595. [PubMed: 8527431]
40. Davies TG, Hubbard RE, Tame JRH. Relating structure to thermodynamics: the crystal structures and binding affinity of eight OppA-peptide complexes. Protein Sci 1999;8:1432–1444. [PubMed: 10422831]
41. Sleigh SH, Tame JRH, Dodson EJ, Wilkinson AJ. Peptide binding in OppA, the crystal structures of the periplasmic oligopeptide binding protein in the unliganded form and in complex with lysyllysine. Biochemistry 1997;36:9747–9758. [PubMed: 9245406]
42. Sleigh SH, Seavers PR, Wilkinson AJ, Ladbury JE, Tame JRH. Crystallographic and calorimetric analysis of peptide binding to OppA protein. J Mol Biol 1999;291:393–415. [PubMed: 10438628]
43. Tame JRH, Murshudov GN, Dodson EJ, Neil TK, Dodson GG, Higgins CF, Wilkinson AJ. The structural basis of sequence-independent peptide binding by OppA protein. Science 1994;264:1578–1581. [PubMed: 8202710]
44. Tame JRH, Dodson EJ, Murshudov G, Higgins CF, Wilkinson AJ. The crystal structures of the oligopeptide-binding protein OppA complexed with tripeptide and tetrapeptide ligands. Structure 1995;3:1395–1406. [PubMed: 8747465]
45. Tame JRH, Sleigh SH, Wilkinson AJ, Ladbury JE. The role of water in sequenceindependent ligand binding by an oligopeptide transporter protein. Nat Struct Biol 1996;3:998–1001. [PubMed: 8946852]

46. Duan X, Hall JA, Nikaido H, Quiocho FA. Crystal structures of the maltodextrin/maltose-binding protein complexed with reduced oligosaccharides: flexibility of tertiary structure and ligand binding. J Mol Biol 2001;306:1115–1126. [PubMed: 11237621]
47. Duan X, Quiocho FA. Structural evidence for a dominant role of nonpolar interactions in the binding of a transport/chemosensory receptor to its highly polar ligands. Biochemistry 2002;41:706–712. [PubMed: 11790091]
48. Quiocho FA, Spurlino JC, Rodseth LE. Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. Structure 1997;5:997–1015. [PubMed: 9309217]
49. Sharff AJ, Rodseth LE, Spurlino JC, Quiocho FA. Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. Biochemistry 1992;31:10657–10663. [PubMed: 1420181]
50. Sharff AJ, Rodseth LE, Quiocho FA. Refined 1.8-Å structure reveals the mode of binding of β-cyclodextrin to the maltodextrin binding protein. Biochemistry 1993;32:10553–10559. [PubMed: 8399200]
51. Spurlino JC, Lu GY, Quiocho FA. The 2.3-Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. J Biol Chem 1991;266:5202–5209. [PubMed: 2002054]
52. Newcomer ME, Lewis BA, Quiocho FA. The radius of gyration of L-arabinose-binding protein decreases upon binding of ligand. J Biol Chem 1981;256:13218–13222. [PubMed: 7031058]
53. Shilton BH, Flocco MM, Nilsson M, Mowbray SL. Conformational changes of three periplasmic receptors for bacterial chemotaxis and transport: the maltose-, glucose/galactose- and ribose-binding proteins. J Mol Biol 1996;264:350–363. [PubMed: 8951381]
54. Careaga CL, Sutherland J, Sabeti J, Falke JJ. Large amplitude twisting motions of an interdomain hinge: a disulfide trapping study of the galactose-glucose binding protein. Biochemistry 1995;34:3048–3055. [PubMed: 7893717]
55. Luck LA, Falke JJ. ^{19}F NMR studies of the D-galactose chemosensory receptor. 1. Sugar binding yields a global structural change. Biochemistry 1991;30:4248–4256. [PubMed: 1850619]
56. Mao B, Pear MR, McCammon JA, Quiocho FA. Hinge-bending in L-arabinose-binding protein: the "Venus's-flytrap" model. J Biol Chem 1982;257:1131–1133. [PubMed: 7035444]
57. Cammer, SA.; Carter, CW., Jr; Tropsha, A. Identification of sequence-specific tertiary packing motifs in protein structures using Delaunay tessellation. In: Gan, HH.; Schlick, T., editors. Computational Methods for Macromolecules: Challenges and Applications, Proceedings of the 3rd International Workshop on Algorithms for Macromolecular Modeling, New York; Lecture Notes in Computational Science and Engineering (LNCSE); Oct. 12–14, 2000; New York. Berlin: Springer Verlag; 2002. p. 477-494.
58. Carter CW Jr, LeFebvre BC, Cammer SA, Tropsha A, Edgell MH. Four-body potentials reveal protein-specific correlations to stability changes caused by hydrophobic core mutations. J Mol Biol 2001;311:625–638. [PubMed: 11518520]
59. Krishnamoorthy B, Tropsha A. Development of a four-body statistical pseudo-potential to discriminate native from non-native protein conformations. Bioinformatics 2003;19:1540–1548. [PubMed: 12912835]
60. Masso M, Vaisman II. Comprehensive mutagenesis of HIV-1 protease: a computational geometry approach. Biochem Biophys Res Commun 2003;305:322–326. [PubMed: 12745077]
61. Singh RK, Tropsha A, Vaisman II. Delaunay tessellation of proteins: four body nearest-neighbor propensities of amino acid residues. J Comput Biol 1996;3:213–221. [PubMed: 8811483]
62. Zheng, W.; Cho, SJ.; Vaisman, II.; Tropsha, A. A new approach to protein fold recognition based on Delaunay tessellation of protein structure. In: Altman, RB.; Dunker, AK.; Hunter, L.; Klein, TE., editors. Pacific Symposium on Biocomputing ’97; Singapore. World Scientific; 1997. p. 486-497.
63. Aurenhammer F. Voronoi diagrams: A survey of a fundamental data structure. ACM Comput Surveys 1991;23:345–405.
64. Li L, Shakhnovich EI. Constructing, verifying, and dissecting the folding transition state of chymotrypsin inhibitor 2 with all-atom simulations. Proc Natl Acad Sci USA 2001;98:13014–13018. [PubMed: 11606790]

65. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242. [PubMed: 10592235]

66. Sawaya MR, Kraut J. Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: crystallographic evidence. Biochemistry 1997;36:586–603. [PubMed: 9012674]

67. Berry MB, Meador B, Bilderback T, Liang P, Glaser M, Phillips GN Jr. The closed conformation of a highly flexible protein: the structure of *E. coli* adenylate kinase with bound AMP and AMPPNP. Proteins 1994;19:183–198. [PubMed: 7937733]

68. Müller CW, Schulz GE. Structure of the complex between adenylate kinase from *Escherichia coli* and the inhibitor Ap_5A refined at 1.9 Å resolution: a model for a catalytic transition state. J Mol Biol 1992;224:159–177. [PubMed: 1548697]

69. Müller CW, Schlauderer GJ, Reinstein J, Schulz GE. Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. Structure 1996;4:147–156. [PubMed: 8805521]

70. Chattopadhyaya R, Meador WE, Means AR, Quiocho FA. Calmodulin structure refined at 1.7 Å resolution. J Mol Biol 1992;228:1177–1191. [PubMed: 1474585]

71. Kuboniwa H, Tjandra N, Grzesiek S, Ren H, Klee CB, Bax A. Solution structure of calcium-free calmodulin. Nat Struct Biol 1995;2:768–776. [PubMed: 7552748]

72. Vandonselaar M, Hickie RA, Quail JW, Delbaere LT. Trifluoperazine-induced conformational change in Ca^{2+}-calmodulin. Nat Struct Biol 1994;1:795–801. [PubMed: 7634090]

73. Wang T, Wade RC. Comparative binding energy (COMBINE) analysis of OppA-peptide complexes to relate structure to binding thermodynamics. J Med Chem 2002;45:4828–4837. [PubMed: 12383008]

74. Jacobs DJ, Rader AJ, Kuhn LA, Thorpe MF. Protein flexibility predictions using graph theory. Proteins 2001;44:150–165. [PubMed: 11391777]

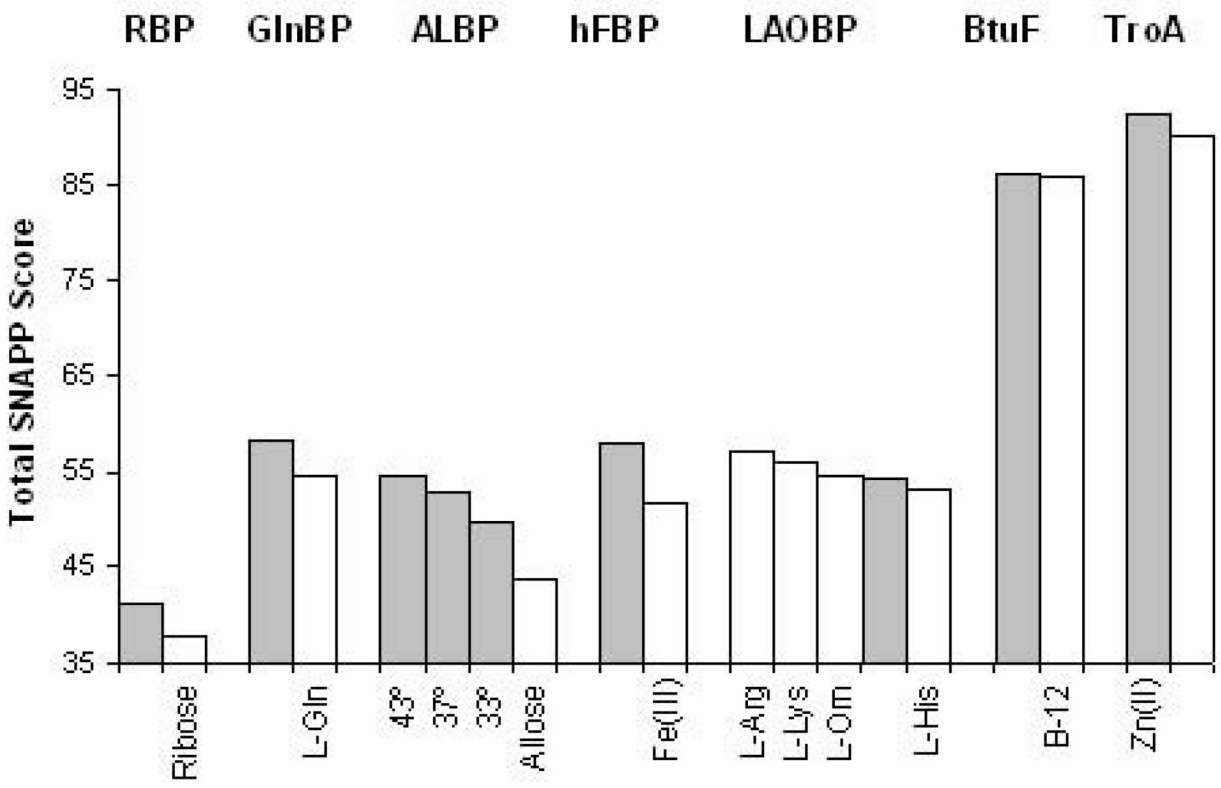

Figure 1a

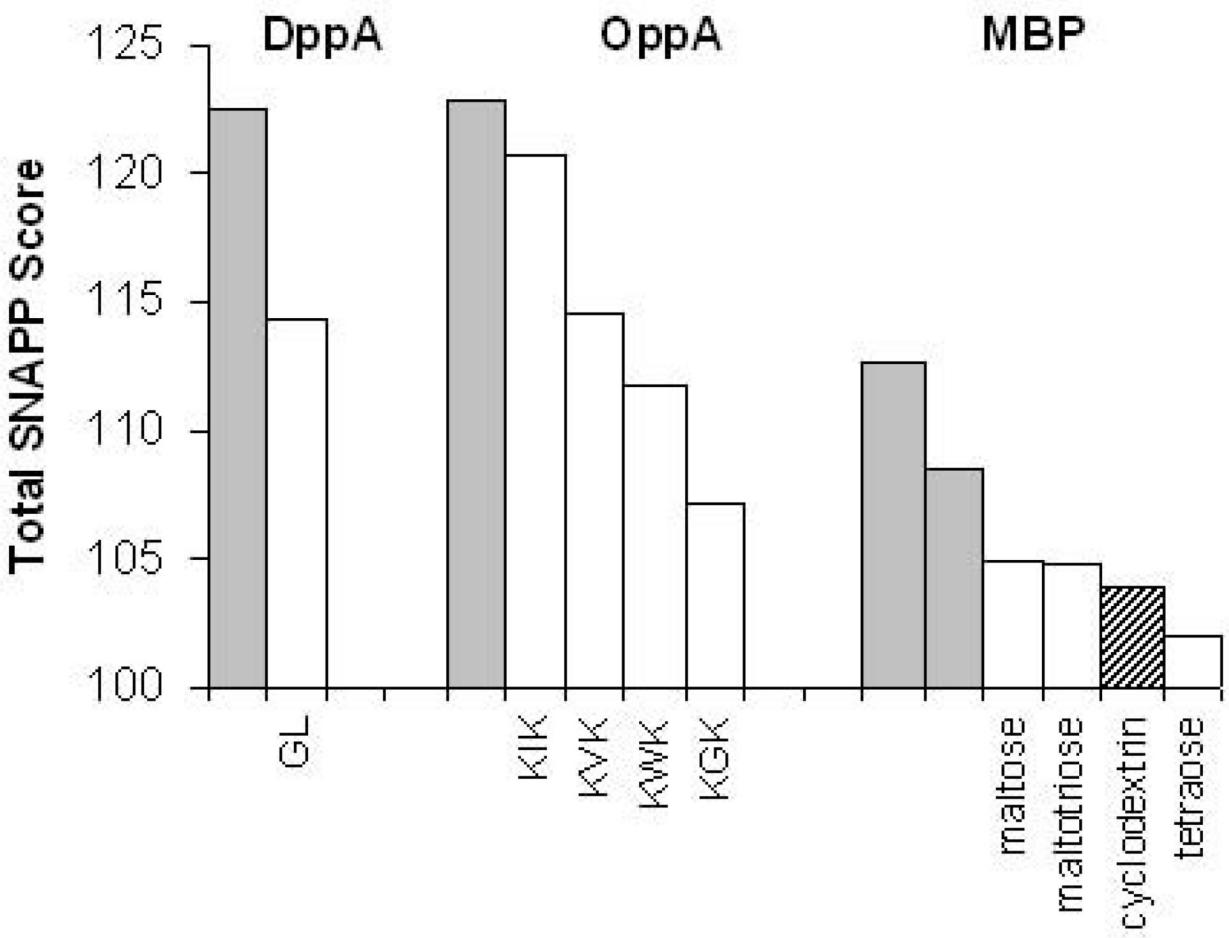

Figure 1b

Figure 1.

Figure 1a. Total SNAPP scores for the opened and closed forms of bPBPs that bind monosaccharides, amino acids, vitamins, and ions. Gray denotes the open form, and white denotes the closed form. TroA is unusual in that the liganded form is more open than the unliganded form.

Figure 1b. Total SNAPP scores for bPBPs that bind polypeptides and oligosaccharides. MBP remains in the open form when bound to β-cyclodextin (hatched), but the protein structure scores in the range of the closed forms.

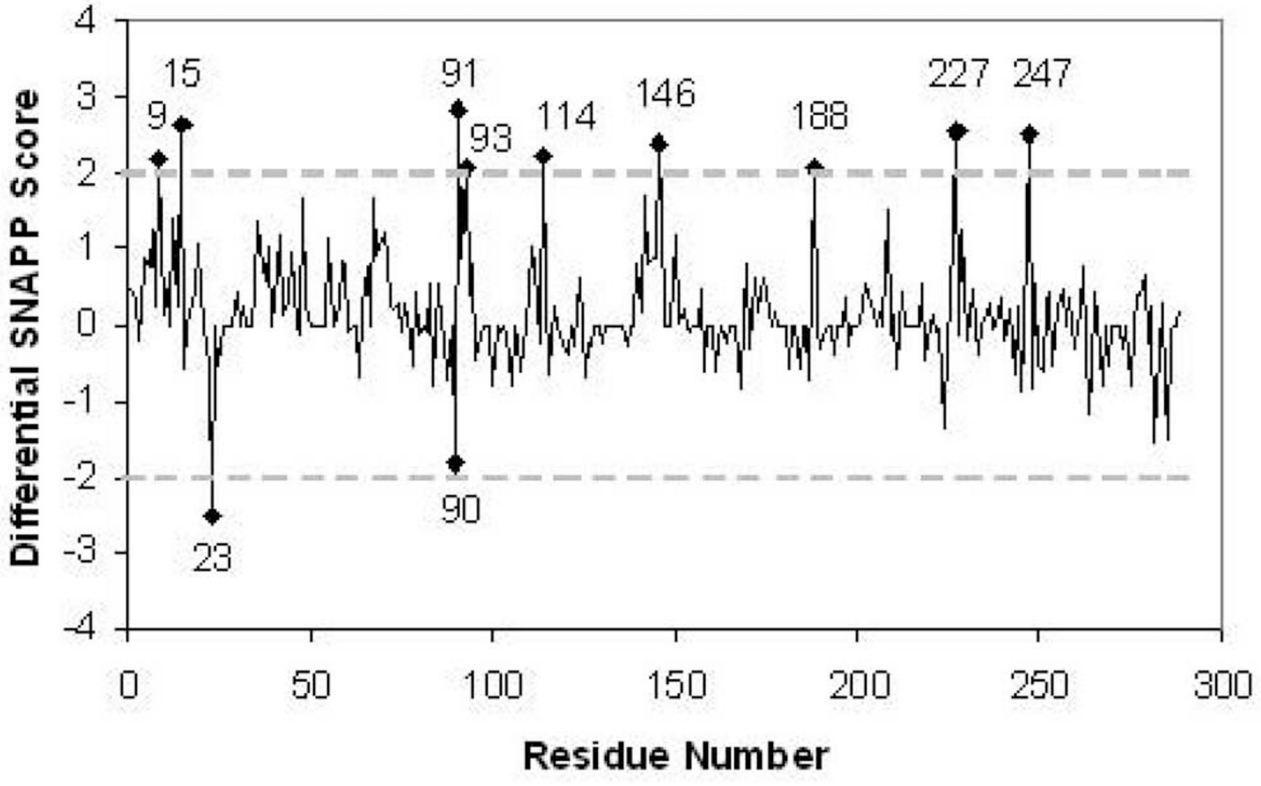

Figure 2a

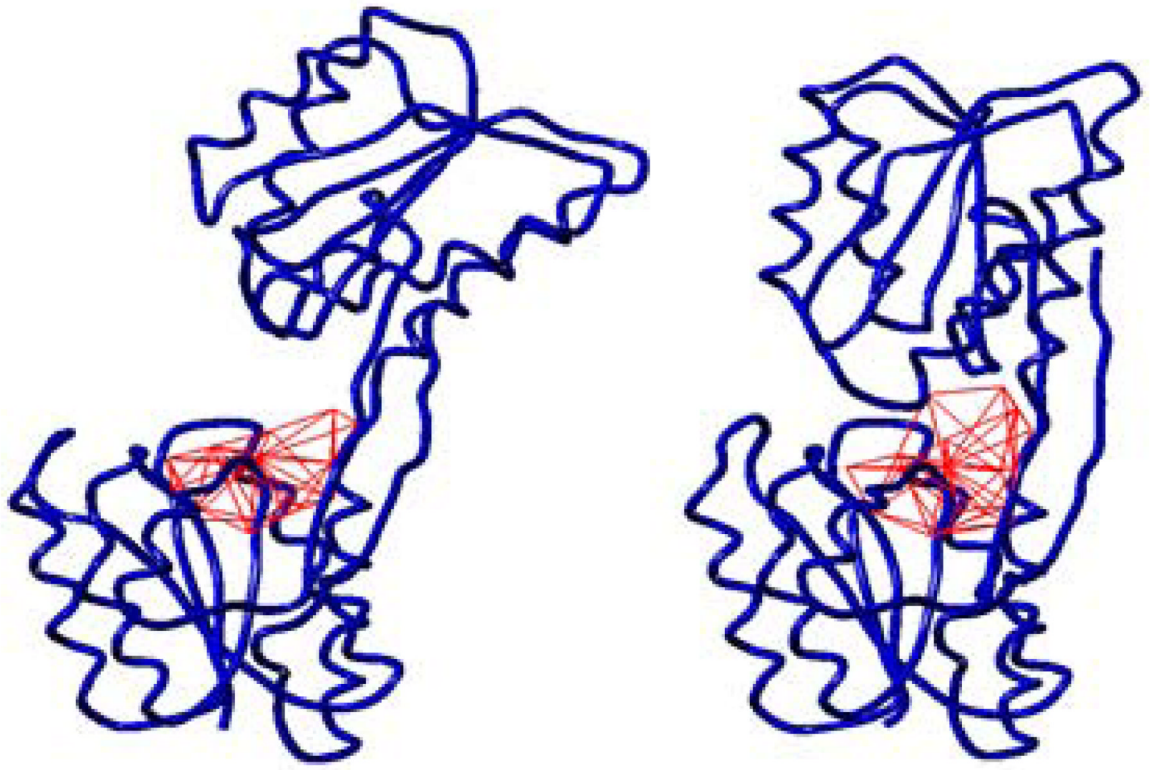

Figure 2b

Figure 2.

Figure 2a. (a) Differential SNAPP profile between the open (43°) and closed forms of ALBP reveals residues that contribute significantly to the differences in total SNAPP scores between forms. SNAPP scores for each residue in the closed form 1RPJ were subtracted from the corresponding SNAPP score in 1GUD A chain, and difference was plotted against residue number. The threshold for significant contribution was set arbitrarily at ±2.

Figure 2b. Tube diagrams and tessellation of residue D91 in the 43° open (left) and closed (right) forms of allose binding protein. D91 had the highest differential SNAPP score, as shown in Figure 2a. Closure of the binding pocket buries the residue, thereby increasing the number of tetrahedral contacts. The contacts in the closed form lower the score for D91, which indicates a less stable environment for this residue.

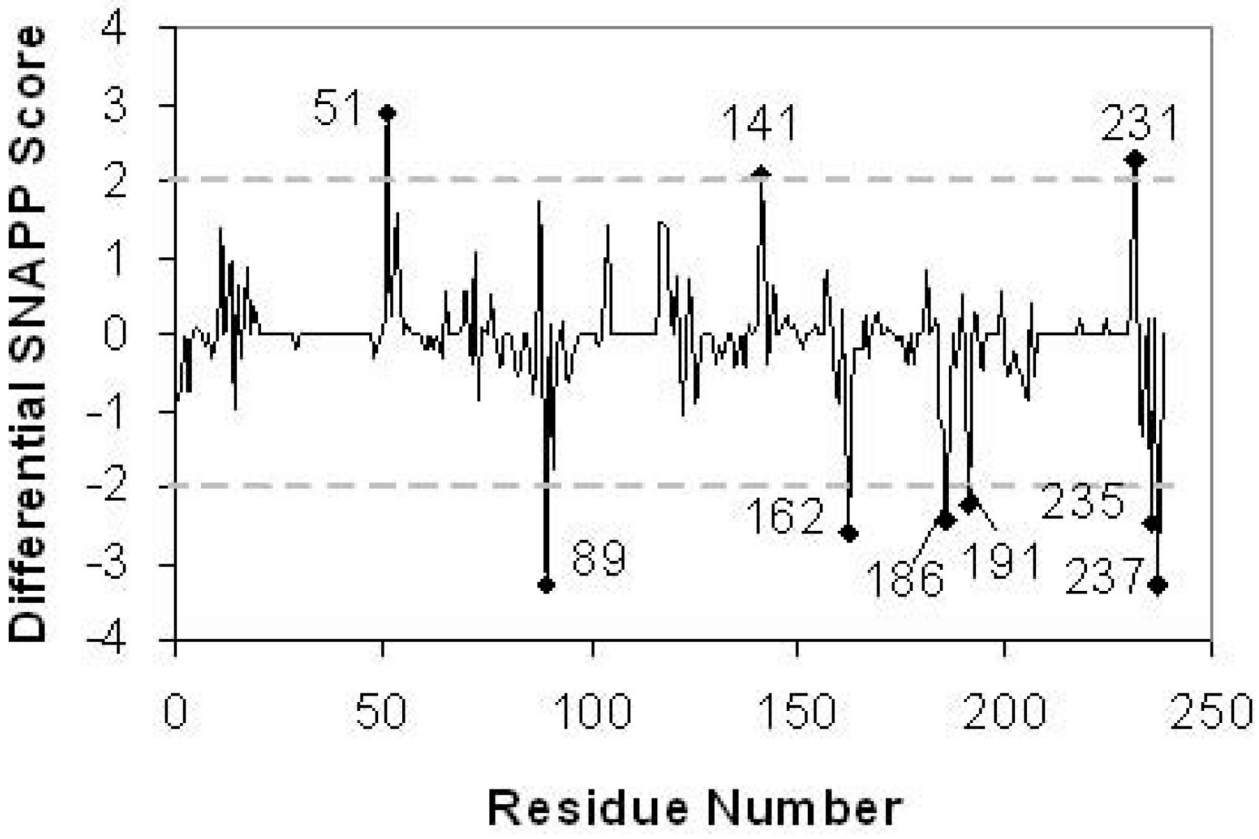

Figure 3a

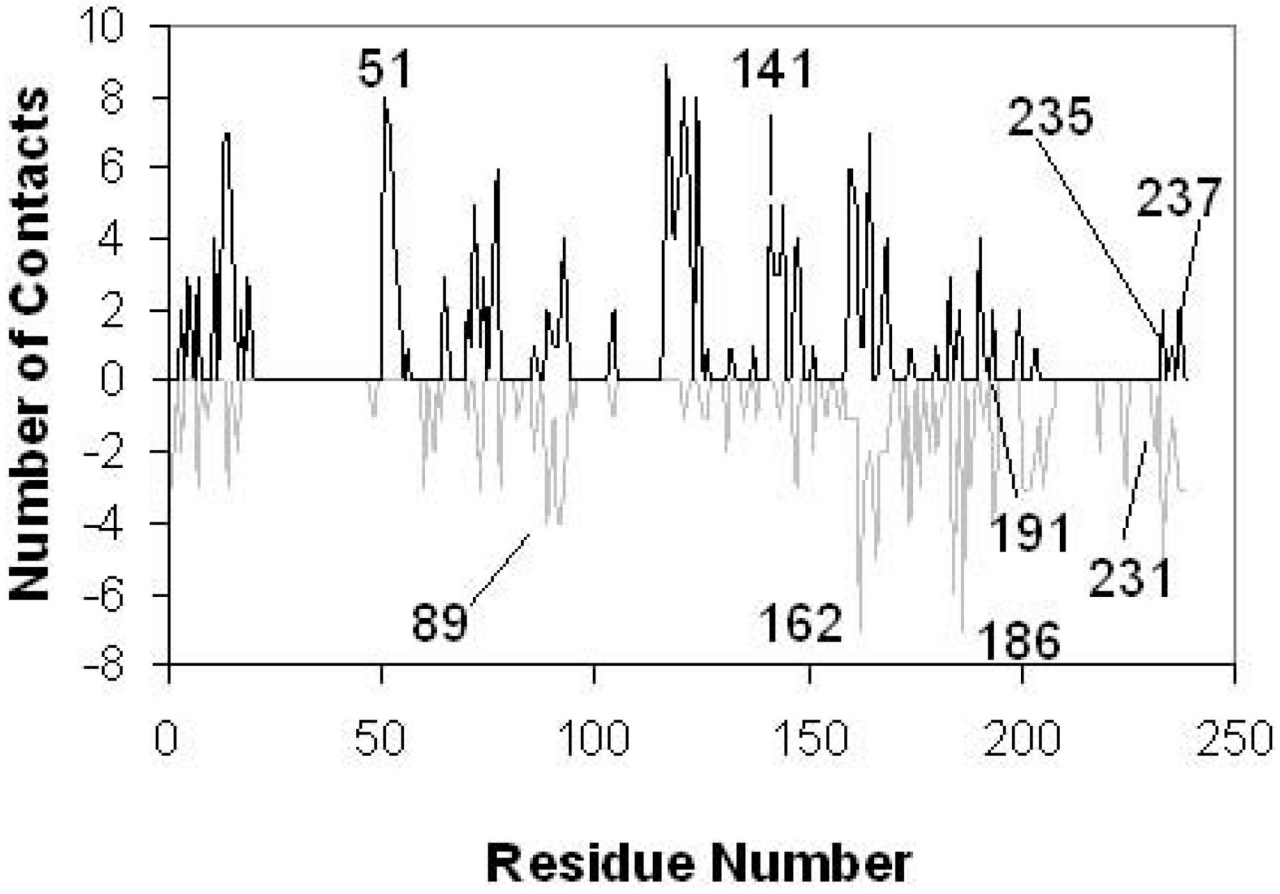

Figure 3b

Figure 3.

Figure 3a. Differential SNAPP profile for LAOBP between the open unliganded form (2LAO) and the closed form (1LAF).

Figure 3b. Number of contacts gained (black) and lost (gray) for each residue upon closure for 2LAO (open) and 1LAF (closed).

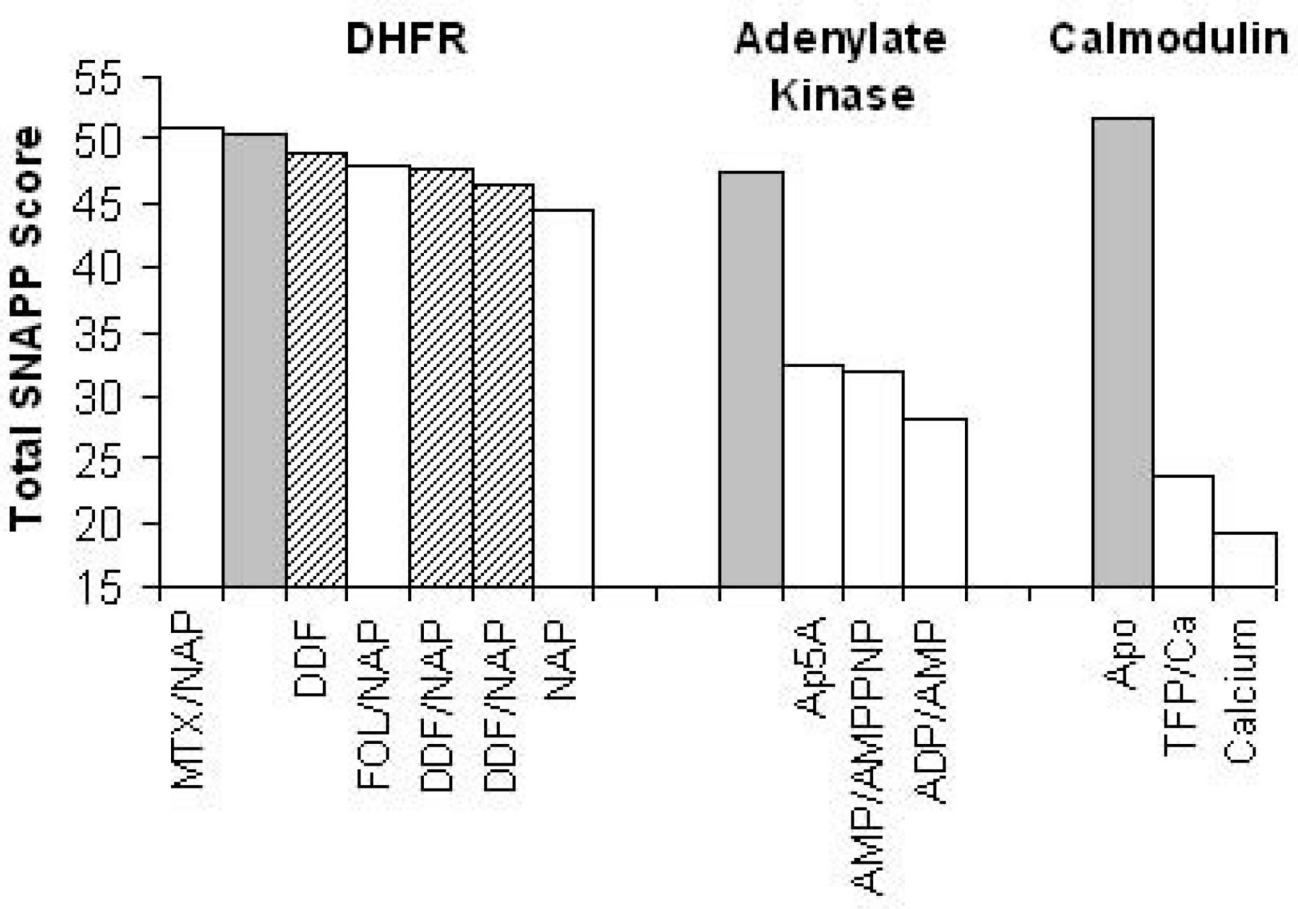

Figure 4.
Total SNAPP scores for unliganded (gray) and liganded (white) forms of the following proteins: Dihydrofolate reductase (DEER), adenylate kinase, and calmodulin. Open forms are in gray, closed forms are white, and occluded forms are hatched.

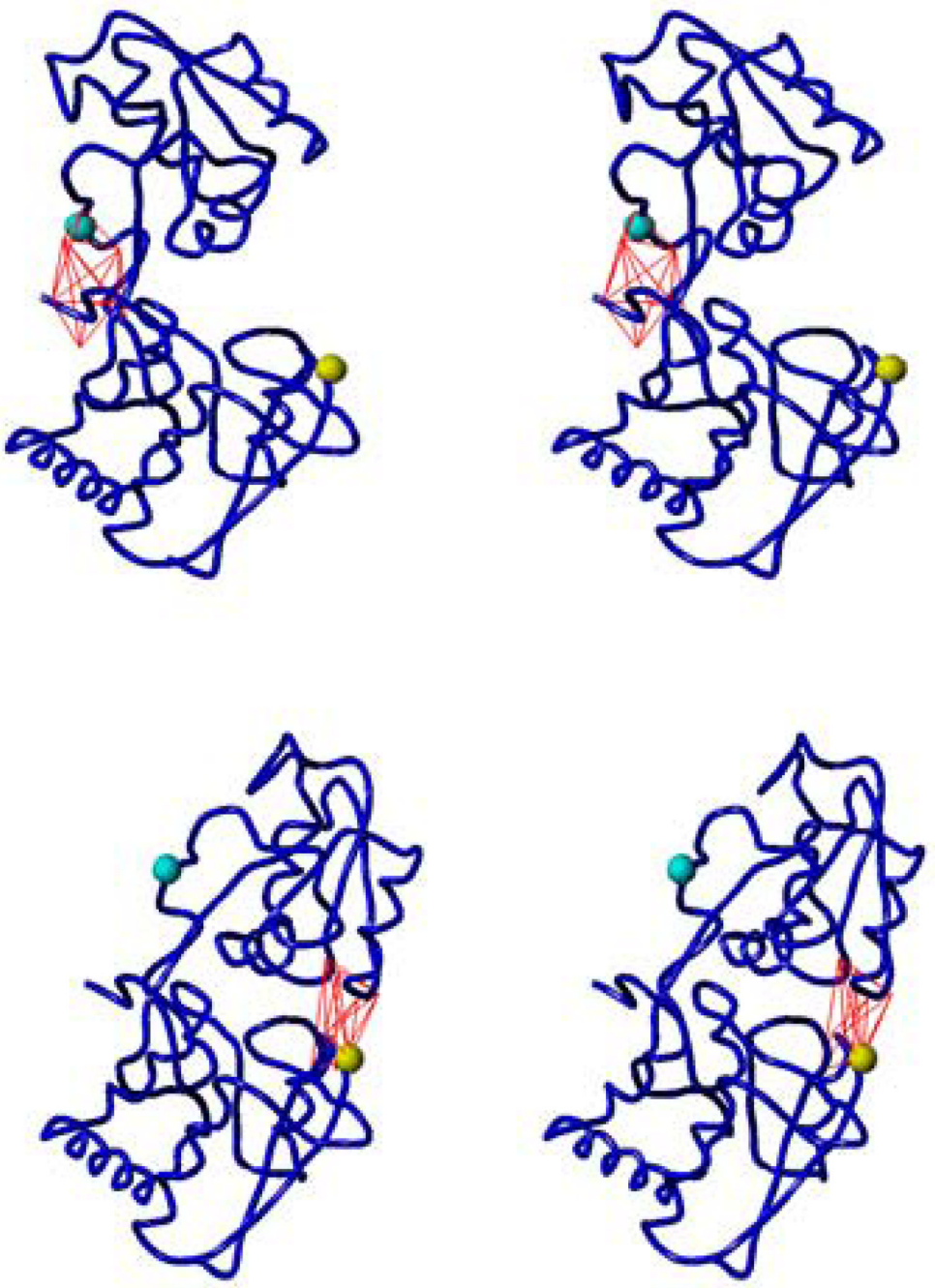

Figure 5.
Contacts determined by tessellation for K186 (cyan) and D51 (yellow) in the open (top) and closed forms of LAOBP (bottom). K186 loses contacts upon closure while D51 gains contacts. The contacts for K186 and D51 are unfavorable, and K186 contributes more positively to the closed form while D51 contributes more positively to the open form.

Table I

Total SNAPP Scores for Periplasmic Binding Proteins

Protein	PDB	Ligand	Conformation[a]	Score
RBP	1URP	None	O	40.925
	2DRI	Ribose	C	37.794
GlnBP	1GGG	None	O	58.078
	1WDN	Glutamine	C	54.399
ALBP	1GUD (A)	None	O	54.292
	1GUB (A)	None	O	52.965
	1GUD (B)	None	O	49.657
	1RPJ	Allose	C	43.543
hFBP	1D9V	None	O	57.964
	1MRP	Fe^{3+}	C	51.586
LAOBP	1LAF	Arginine	C	56.85
	1LST	Lysine	C	55.914
	1LAH	Ornithine	C	54.312
	2LAO	None	O	54.273
	1LAG	Histidine	C	53.196
BtuF	1N4A	None	O	86.160
	1N4D	Vitamin B-12	C	85.715
TroA	1TOA	Zn^{2+}	O	92.330
	1K0F	None	C	89.927
MBP	1FQB	Maltotriitol[b]	O	113.22
	1OMP	None	O	112.584
	1FQD	Maltotetraitol	C	109.637
	1EZ9	Maltotetraitol[b]	O	108.578
	1JW4	None	O	108.472
	1FQA	Maltotetraitol[b]	O	108.119
	1JW5	Maltose[c]	O	106.18
	1FQC	Maltotriitol	C	105.575
	1ANF	Maltose	C	104.937
	3MBP	Maltotriose	C	104.753
	1DMB	β-cyclodextrin	O	103.908
	4MBP	Maltotetraose	C	102.033
DppA	1DPE	None	O	122.445
	1DPP	GL	C	114.336
OppA	1RKM	None	O	122.899
	1B3G	KIK	C	120.729
	1B5I	KNK	C	118.553
	2RKM	KK	C	118.488
	1B3F	KHK	C	118.249
	1JEU	KEK	C	117.078
	1B6H	K-Nva-K	C	116.792
	1B5J	KQK	C	116.121
	1B0H	K-Npa-K	C	116.02
	1B4Z	KDK	C	115.995
	1B46	KPK	C	115.756
	1B5H	K-Dap-K	C	115.424
	1OLC	KKKA	C	114.947
	1B51	KSK	C	114.800
	1B9J	KLK	C	114.585
	1QKB	KVK	C	114.562
	1B3H	K-Chx-K	C	114.058
	1JET	KAK	C	113.741
	1B4H	K-Dab-K	C	113.343
	1OLA	VKPG	C	113.333
	1B58	KYK	C	113.303
	1QKA	KRK	C	113.146
	1B40	KFK	C	112.871
	1B1H	K-Hpe-K	C	112.855
	1B32	KMK	C	112.614
	1B7H	K-Nle-K	C	112.418
	1B2H	K-Orn-K	C	112.191
	1JEV	KWK	C	111.836
	2OLB	KKK	C	111.524
	1B05	KCK	C	109.691
	1B52	KTK	C	108.609
	1B3L	KGK	C	107.169

[a] Open conformation is denoted by O and closed by C

[b] Open form co-crystallized with ligand.

[c] Open form crystal saturated with ligand.

Table II

Total SNAPP Scores for Several Non-periplasmic Binding Proteins

Protein	PDB	Ligand	Conformation	Score
Dihydrofolate Reductase	1RX3	MTX, NAP	C	51.052
	1RA1	NAP	O	50.38
	1RX5	DDF	Occ[a]	48.892
	1RX2	FOL, NAP	C	47.949
	1RX4	DDF, NAP	Occ	47.535
	1RX6	DDF, NAP	Occ	46.302
	1RX1	NAP	C	44.491
Adenylate Kinase	4AKE	None	O	47.338
	1AKE	Ap_5A	C	32.475
	1ANK	AMP/AMPPNP	C	32.050
	2ECK	ADP/AMP	C	28.302
Calmodulin	1CFD	None	O	51.729
	1A29	TFP/Ca^{2+}	C	23.543
	1CLL	Ca^{2+}	C	19.148

[a]Occluded conformation is denoted by Occ