MINIREVIEW

Accommodating Protein Flexibility in Computational Drug Design[1,2]

HEATHER A. CARLSON and J. ANDREW MCCAMMON

Departments of Pharmacology and Chemistry and Biochemistry, University of California, San Diego, La Jolla, California

Received November 24, 1999; accepted December 14, 1999   This paper is available online at http://www.molpharm.org

The need to account for the dynamic behavior of a receptor has long been recognized as a complicating factor in computational drug design. The use of a single, rigid protein structure—usually from a high-quality X-ray crystal structure—still is the standard in most applications (Zheng and Kyle, 1997; Walters et al., 1998). The choice to use a single protein structure is usually based on speed. For example, if a large database of compounds is to be screened for binding affinity, several conformers of each compound will be compared with each protein configuration. Although it is more accurate to use many representative protein configurations, it quickly becomes a very slow process to screen each conformer of each ligand against each protein configuration. It is usually impractical to attempt such prohibitively slow calculations because there is an unfortunate but necessary trade-off between speed and accuracy in computer modeling.

Only recently have advanced methods been introduced to aid in a more accurate description of protein flexibility and its influence on ligand recognition. This has been aided by the exponential growth in the speed of computer processors, available RAM, and disk capacity—all of which are rapidly becoming less expensive. Here, we explain the need for accommodating an ensemble of protein configurations in drug design and the computational methods available for generating and manipulating that dynamic information. Most of the applications discussed below are improvements in ligand-docking or in the generation of pharmacophore models for database searching.

Protein Flexibility and Its Influence on Ligand Binding

For a single, fixed conformation to be an adequate representation of the protein, the system would have to be very rigid. Such a system would correspond with the "lock-and-key" theory of binding in which the protein exists in a single well-defined state with only one optimal complementary ligand. However, the energy landscape of most proteins is frequently described in terms of a folding funnel in which there are many highly unfavorable states that collapse via multiple routes into possibly several favorable folded states (Rejto and Freer, 1996; Onuchic, 1997; Shoemaker et al., 1997; Freire, 1998; Ma et al., 1999). Figure 1 demonstrates the theoretic implications of using a single conformation to represent the behavior of a flexible and rigid protein. In Fig. 1, both proteins are shown to have only one favorable folded state. However, it is implied by the width of the minima that the folded state of the flexible system has higher entropy and, as such, will have a larger subensemble of occupied states. The subensemble is a collection of structurally similar and nearly energetically equivalent conformations of the protein that, together, make up the folded state (Fig. 2). A single structure, even the weighted average provided by a crystal structure, may not describe these substates adequately.

It is important to note that a conformational funnel and the states that it represents are condition-dependent. By altering the conditions (ionic strength, pH, temperature, etc.), the minima can shift in favorability, changing which state is most populated. For example, the funnel in Fig. 2 could correspond to a protein with three conformational states, perhaps identifiable through the solution of different crystal forms. In their review of the use of X-ray crystallography to probe the conformational states of a protein, Rejto and Freer (1996) note that 25 to 30% of a protein surface is in contact with symmetry-related molecules. By changing the

---

ABBREVIATIONS: MC, Monte Carlo; MD, molecular dynamics.

symmetry of the unit cell or other crystallization conditions, the environment is altered. The resulting shift in the most populated configuration would be attributed to crystal packing forces.

The most important point to consider is that introducing a ligand into the system also changes the environment. It too may effect the most populated state of the protein; such a case would correspond to an "induced-fit" system. Recently, Freire (1998) and Ma et al. (1999) have made compelling and subtle arguments address this issue. Rather than viewing the protein in a fixed state that the ligand actively deforms, they argue that the protein most likely exists in a full complement of conformations—most in the native state, some in the induced-fit state, and some in other states. If the ligand binds preferentially to the induced-fit state with sufficiently

favorable free energy (meaning greater than the free energy difference between the two protein states), the *average structure* of the protein will change. If the ligand is capable of multiple binding modes to a protein (binding to different folded forms or multiple binding sites of a single form), it may be necessary to account for these additional states to predict affinity properly (Brem and Dill, 1999).

The implications for drug discovery are clear; a single protein structure is only useful to identify ligands for that particular narrow state of the subensemble. To obtain new leads and properly predict activity of existing inhibitors, multiple structures are the best option. The available computational methods range in the degree of protein flexibility that they accommodate (focusing on a narrow set of states, a broader subensemble, or the whole conformational funnel in Fig. 2). Here, we present the continuum of methods in the order of increasing flexibility and complexity. Although it often is a temptation to "rank" methods in a review, such an attempt is premature in this area. Most of the advancements have only appeared in the literature over the last five years, and there are few systems to compare to experimental data.

Soft-Docking

The earliest attempts to accommodate small changes in conformation were through the use of implicit methods (Jiang and Kim, 1991). The protein is held fixed, but a "soft"-scoring function is used to evaluate the fit of the ligand to the receptor. Often, scoring functions are derivatives of force fields from molecular mechanics, modified for use in a new application. Soft functions allow for some overlap between the ligand and the protein, giving a small estimate of the plasticity of the receptor. This still restricts the successful ligands so that they frequently will be limited to a certain size and conformation. However, soft-docking has the benefit
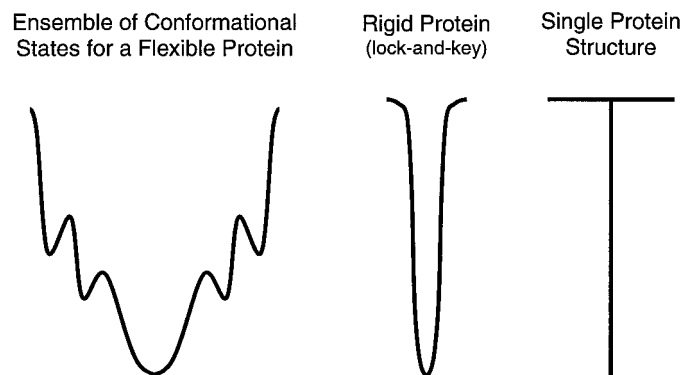
Fig. 1. A single structure of a protein implies an all-or-nothing folding funnel, perhaps best described with the mathematic singularity shown on the right. The folding funnels on the left and in the center demonstrate the conformational flexibility of a "standard" flexible protein and a rigid protein, respectively. Although the rigid system may be described adequately by a single structure (depending on the degree of rigidity as reflected by the narrowness of the funnel), a typical system is not.
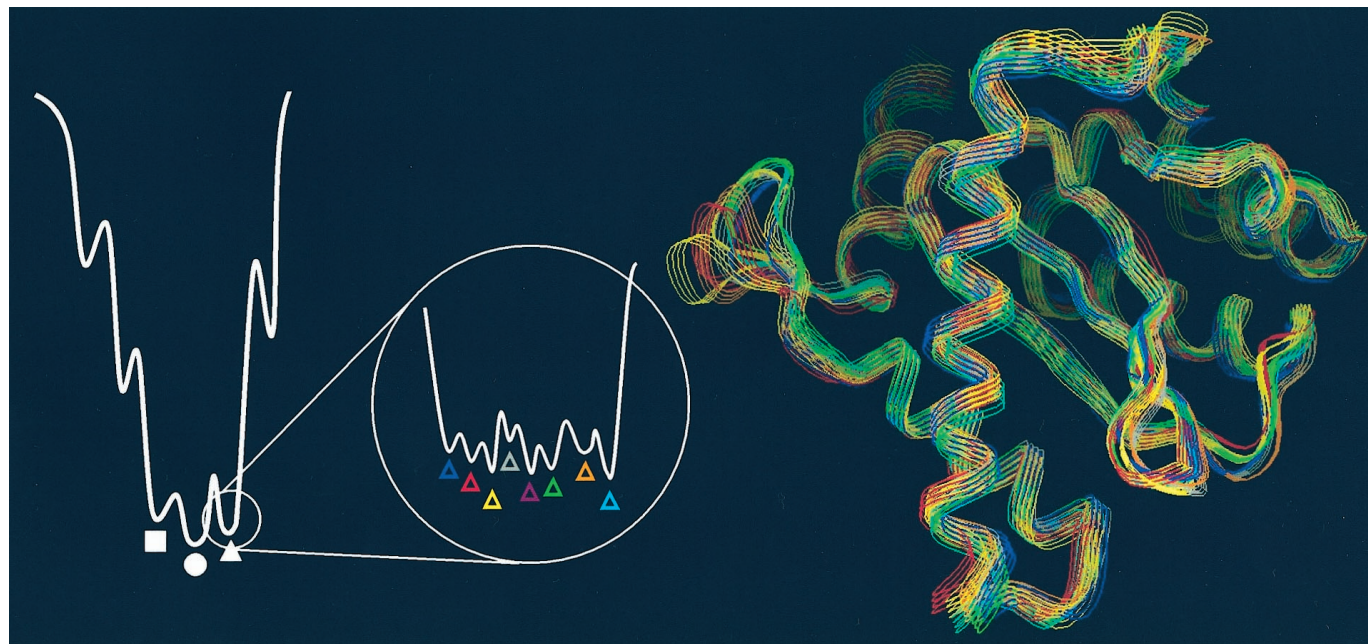
Fig. 2. This conformational funnel demonstrates multiple states for a protein (represented by their weighted-average structures: ■, ●, and ▲). The flexibility inherent in a folded state (▲) is described by a subensemble of conformations (shown here as a collection of colored △). We have used a collection of structures from an MD simulation of HIV-1 integrase (Lins et al., 1999) to demonstrate a subensemble, showing a range of flexibility with modest sampling of the backbone and wide sampling of a small flexible loop on the left. Expanding the minimum of a single state in the subensemble (any of the colored △) would reveal an additional series of subminima that arise from variations in the orientations of the side chains.

of being computationally efficient (evaluating the scoring function requires no additional calculation time) and of being relatively easy to implement in existing programs (conformational sampling routines are unchanged). Because of these advantages, there still are improvements being made to scoring functions that use soft-docking techniques (Gschwend et al., 1996; Schnecke et al., 1998).

Conformational Sampling of the Side Chains in the Receptor

The simplest form of conformational sampling accommodates the reorientation of hydrogen atoms. Jones et al. (1995) allowed rotational freedom for hydrogens and lone pairs of hydrogen-bond donor and acceptor atoms in one of the earliest studies using genetic algorithms for ligand-receptor docking. This created more favorable hydrogen bonding networks between the ligand and the receptor. They noted that the method could be extended easily to include complete torsional freedom for the side chains, but the additional computational cost prevented them from including full side chain flexibility in their study of dihydrofolate reductase, L-arabinose binding protein and influenza sialidase. Many of the methods that incorporate flexibility for the entire length of the side chains only allow for torsional sampling—rotations around single and double bonds. Bond lengths and angles are held fixed, but the orientations of the side chains are almost fully accessible through rotations around the bonds. During ligand-docking, these can be minimized through internal coordinate optimization or random thermodynamic sampling techniques such as Metropolis sampling routines (Totrov and Abagyan, 1997; Schnecke et al., 1998).

In one of the earliest incorporations of side chain mobility, Leach (1994) made use of a rotamer library. Alanine, glycine, and proline were given no rotameric states, for obvious reasons. Methionine, lysine, and arginine were given 21, 51, and 55 rotameric states. The 14 additional amino acids ranged from 3 to 10 rotameric states because of their small size and/or rigidity. Although a rotameric library does not allow for complete flexibility of the side chains, it does incorporate the orientations with the lowest energies and provide a good first estimate of the available conformations. Furthermore, searching a rotameric library for side chain orientations that complement a docked ligand is a relatively quick way to search much of the conformational space available to the receptor. The method also may reduce the limitation of conformational barriers; minimization routines can often become trapped in poor local minima that the discrete rotational states in the libraries may "hop" over.

A more recent method uses larger rotameric libraries and clever search algorithms to save time (Schaffer and Verkhivker, 1998). In a two-step process, a rotamer library is used to generate likely conformations of the side chains, and then the docked system is subjected to unrestricted minimization of the ligand and local side chains. Energy minimization (sometime referred to as optimization) allows for the strongest contacts between the ligand and the receptor, and usually all possible orientations of the side chains and ligand are possible. Most of the applications involving minimizations do not include the entire protein, but instead use a subset of atoms (usually the side chains of the receptor site, but local backbones in the receptor region can be included for

even more local flexibility). The additional atoms outside the receptor region are held fixed. Optimization routines can include simulated annealing, steepest-descent minimization, Monte Carlo (MC) sampling, or other methods. The most interesting optimization methods involve novel algorithms that work around the limitation of becoming trapped in local minima. Apostolakis et al. (1998) have used a conjugate gradient minimization technique that uses a twist from soft-docking. Over the course of the minimization, the van der Waals terms of the potential function are gradually switched on. This allows the ligand to initially adopt a conformation that is a close match to the receptor, but may have some overlap with the side chain residues. As the nonbonded terms are gradually incorporated, the ligand and the receptor can make minor conformational changes to avoid overlap. Another technique for increasing the range of conformations sampled with minimization routines uses multicanonical molecular dynamics (MD) simulations; Nakajima et al. (1997a) used this method in a study of the binding of a peptide to the Src homology 3 domain. The multicanonical algorithm smoothes the energy surface so that unfavorable conformations can exist and barriers can be overcome (Nakajima et al., 1997b). The conformations are evaluated later with a standard force field that reintroduces the full energy landscape, and the binding properties can be assessed with a proper Boltzmann-weighting of the multicanonical sampling.

This last example of minimization methods is an excellent lead into using an ensemble of protein configurations. Sudbeck et al. (1998) recently introduced a simple method that qualitatively incorporates multiple crystal structures into the drug discovery process. They overlaid the structures of nine inhibitor complexes of HIV reverse transcriptase using the $C\alpha$ of the most stable region (residues 97–213). The resulting positions of the nine inhibitors in the overlay were used to create a composite binding site; a map of the surface area that encompasses all the inhibitors was used to described most of the available space in the receptor. Small molecules were docked to a single structure of the reverse transcriptase using conjugate gradient minimization of the ligands and all residues within 5 Å of the ligand. The docked structure was then qualitatively compared with the composite binding site described by the surface area of the nine inhibitors. Although simple and quick, this method leads to two new and highly potent inhibitors ($IC_{50} < 1$ nM).

Generating a Subensemble

The most rigorous methods for generating a subensemble of states and calculating thermodynamic properties are free energy calculations with MC or MD simulations (Lamb and Jorgensen, 1997). These methods are highly reliable and are typically comparable with experimental data within 1 kcal/mol. They are the best choice for accurately assessing chemical modifications that can be made to existing inhibitors to improve their binding affinity. However, comparing two ligands by these methods often can be much slower than screening large libraries of compounds with the more approximate methods described above. In an effort to improve speed and computational efficiency, Luty et al. (1995) used a simplified MD simulation with an implicit solvation model. Much of the protein (trypsin) was held fixed, whereas the active site region and the ligand (benzamidine) were fully

sampled. To dampen the interaction between the rigid and flexible regions of the protein, a small buffer region in a layer between the rigid and flexible atoms was held to their crystallographic coordinates by a small harmonic potential. This method allowed for rapid simulation of the docking process, but it was limited in its conformational sampling (much like the methods discussed above). Using a method that enhances sampling in MD simulations, Mangoni et al. (1999) have studied docking between phosphocholine and the immunoglobulin McPC603 in explicit water. Much of the sampling in the MD simulation is treated in the standard manner, but the center of mass of the protein is given a higher temperature (larger velocity). This encourages faster additional sampling in the system, but the dampening force of the explicit water keeps the system from deforming into less appropriate conformations.

Philippopoulos and Lim (1999) have addressed the issue of conformational sampling in standard MD simulations. A 1.7-ns MD simulation of *Escherichia coli* ribonuclease HI was compared with an ensemble of conformations provided from 15 NMR structures. The conformational sampling for both the NMR and the MD structures is consistent with NMR relaxation data in many regions, implying that the sampling seen in both sets of structures is similar to the dynamic data. The study finds that the MD simulation samples less conformational space than does the NMR structures; the NMR structures show more flexibility in both the side chains and the backbone atoms. The MD and NMR structures also were compared with two crystal structures showing a correlation between the thermal factors in the crystal structures and the sampling in the MD and NMR. When available, NMR structures appear to be an excellent source of protein structures for use in drug discovery. Unfortunately, structure determination with NMR is limited to rather small proteins (~100 residues). Knegtel et al. (1997) have presented the first study using NMR structures; their method is addressed below.

The earliest example of using MD simulations to generate a set of protein conformations for use in docking studies was reported by Pang and Kozikowski (1994). They obtained 69 conformations of *Torpedo* acetylcholinesterase from a 40-ps MD simulation. Because of computational constraints, the MD was relatively short and performed in vacuo with water molecules in the active site gorge. Huperizine A was docked rigidly to the 69 conformations of acetylcholinesterase. It was predicted that huperizine A binds preferentially at the catalytic site at the bottom of the characteristic deep gorge—quite an accomplishment for this early work.

We have recently used MD to generate a subensemble of states for use in pharmacophore modeling (Masukawa et al., 2000). At total of 11 snapshots of an uncomplexed monomer of HIV-1 integrase were taken from an MD simulation. Each snapshot then was held rigid because the flexibility of the system was represented through the use of multiple structures rather than through use of a single, semiflexible structure. Methanol molecules were docked to the 11 receptor conformations to determine complementary positions for hydrogen bond donating groups. The results were overlaid via the essential residues to identify complementary binding regions for the hydrogen-bond donors that were conserved over many protein structures despite the inherent flexibility of the active site. These conserved complementary regions describe a "dynamic" pharmacophore model that compared favorably with known inhibitors of HIV-1 integrase, and the model was used later to search the Available Chemicals Directory to identify successful inhibitors. In a follow-up study (Carlson et al., 1999), we addressed the issue of using multiple crystal structures to create dynamic pharmacophore models. Again, pharmacophore models created with multiple structures compared more favorably with known inhibitors than did models created with a single, rigid crystal structure.

The use of grid representations of receptor sites is similar to the use of pharmacophore models. A grid contains the shape and specificity (hydrophobic regions, hydrogen-bond donor and acceptor regions, charged regions, etc.) of the receptor. Docking a ligand to this type of complementary model of the receptor often is much faster than evaluating each of the many atomic interactions between the ligand and the large protein. MD has been used to generate many configurations of protein-inhibitor complexes and then overlay them with respect to the bound inhibitors (H. B. Broughton, personal communication, 1999). The composite grid is created using a weighted average of the grids that describe the receptor in each of the overlaid protein configurations. This method showed significant improvement in evaluating known ligands of dihydrofolate reductase. Knegtel et al. (1997) have presented a composite grid method incorporating multiple crystal or NMR solution structures of protein-ligand complexes. Different methods for averaging the grid representation of the composite receptor were examined, including geometric averaging of the receptor structures and Boltzmann-weighted averaging of the scoring function grids that describe the separate receptor structures. In this extensive study, they incorporate five crystal structures of inhibitor complexes of HIV-1 protease, five crystal structures of complexed ras p21 protein, 25 NMR structures of uteroglobin complexed with a polychlorobiphenyl metabolite, and five cocrystals of bovine retinol binding protein. The use of a composite grid improves the accuracy of the method, and, surprisingly, it also improves the computational speed of the docking simulations used in database searching.

Bouzida et al. (1999) have presented what appears to be the closest reproduction of the process of ligand-protein recognition. We repeat the opening argument that a protein exists in multiple conformations and that a ligand may bind preferentially to any available conformation. Although computationally expensive, Bouzida et al. (1999) actually dock a ligand explicitly to numerous protein conformations. Just as a ligand can bind selectively to one conformation in solution, that same bias can be evaluated computationally by the docking score of the ligand to each protein conformation. Using HIV-1 protease as a test case, they dock two different ligands to 10 different protein conformations obtained from crystal structures of protease complexes. An MC, simulated annealing minimization is used to dock a flexible ligand into the rigid protein structures. Furthermore, a "soft" scoring function is used to allow for additional implicit flexibility of the receptor. They find that the first ligand binds preferentially to only one of the protein conformations, whereas the second can bind more broadly (almost equally to four unique protein conformations). These results point to the cautions of using a single rigid protein structure. Using 9 of the 10 protein structures in this study would not have properly predicted activity for the first ligand, and the second ligand

would only be identified if one of the proper four protein structures were fortuitously chosen.

Predicting Loop Flexibility and Domain Motions

The most challenging task for computational drug design is predicting large domain motions. Sandak et al. (1998) have addressed this issue by limiting their docking routine to hinge-bending motions for both the protein and the ligand. Large stable regions of the protein are identified and modeled as being connected by three-dimensional rotational hinges. Most of the protein is held rigid (all regions between the hinge points). Hinges also are used for flexible points in the ligand. They also note that the hinge could be restricted to a bond rotation if appropriate. Docking is scored through matching appropriate surface areas. A more coarse scoring routine like this must be used, because there is no side chain flexibility or local structure optimization. Additional methods are being developed for the prediction of domain motions and loop fluctuations (Maiorov and Abagyan, 1997; Wriggers and Schulten, 1997; Freire, 1999; Keller et al., 2000). Although they have not been used in drug discovery applications, these methods are quite promising and most likely will be used in the near future.

Zacharias and Sklenar (1999) use harmonic modes to describe protein flexibility. Mobility of large regions of a protein is characterized by low vibrational frequencies. Zacharias and Sklenar have used relaxation of the harmonic modes to improve steric fit between DNA and a minor groove ligand. Hayward et al. (1997) and Bahar et al. (1999) also have developed promising harmonic approximations to estimate global protein flexibility, but they have not yet applied these methods in ligand binding.

Bardi et al. (1997) have used a method to calculate residue stability factors that is a measure of the structural stability of the protein mapped out onto the individual residues. These stability factors can be compared with hydrogen exchange protection factors measured by NMR. This group of researchers compared the residue stability for all residues of HIV-1 protease when bound to 1 of 13 inhibitors. The difference in the inhibitors propagated into the structural and dynamic changes throughout the protein. Stability factors compared well with known escape mutants for the inhibitors (even mutants far from the binding region). The stability factors also were used to describe a dual character of the binding pocket, showing that half of the binding region is part of the most stable region of the protease, whereas the other half is rather plastic before binding. Hilser et al. (1998) have continued to develop this technique; structural information like this is quite useful in understanding the structural control in an enzyme, the design of inhibitors, and the possible escape mutations.

Conclusion

We have covered many of the recent improvements for incorporating protein flexibility in drug discovery and a few of the most promising methods for predicting large-scale motions. It is important to point out that some of these methods have to ignore local conformational flexibility such as side chain reorientation when estimating large protein motions. It is possible that the methods presented above could be combined to generate ensembles of protein structures representing appropriate global and local flexibility. In conclusion, we stress the promise of some of the methods in the last section, which can address both local and global dynamic information (Bardi et al., 1997; Hilser et al., 1998; Bahar et al., 1999; Keller et al., 2000).

References

Apostolakis J, Plückthun A and Caflisch A (1998) Docking small ligands in flexible binding sites. *J Comput Chem* **19:**21–37.

Bahar I, Erman B, Jernigan RL, Atilgan AR and Covell DG (1999) Collective motions in HIV-1 reverse transcriptase: Examination of flexibility and enzyme function. *J Mol Biol* **285:**1023–1037.

Bardi JS, Luque I and Freire E (1997) Structure-based thermodynamic analysis of HIV-1 protease inhibitors. *Biochemistry* **36:**6588–6596.

Bouzida D, Rejto PA, Arthurs S, Colson AB, Freer ST, Gehlhaar DK, Larson V, Luty BA, Rose PW and Verkhivker GM (1999) Computer simulations of ligand-protein binding with ensembles of protein conformations: A Monte Carlo study of HIV-1 protease binding energy landscapes. *Int J Quantum Chem* **72:**73–84.

Brem R and Dill KA (1999) The effect of multiple binding modes on empirical modeling of ligand docking to proteins. *Protein Sci* **8:**1134–1143.

Carlson HA, Masukawa KM and McCammon JA (1999) Method for including the dynamic fluctuations of a protein in computer-aided drug design. *J Phys Chem A* **103:**10213–10219.

Freire E (1998) Statistical thermodynamic linkage between conformational and binding equilibria. *Adv Protein Chem* **51:**255–279.

Freire E (1999) The propagation of binding interactions to remote sites in proteins: Analysis of the binding of the monoclonal antibody D1.3 to lysozyme. *Proc Natl Acad Sci USA* **96:**10118–10122.

Gschwend DA, Good AC and Kuntz ID (1996) Molecular docking towards drug discovery. *J Mol Rec* **9:**175–186.

Hayward S, Kitao A and Berendsen HJC (1997) Model-free methods of analyzing domain motions in proteins from simulation: A comparison of normal mode analysis and molecular dynamics simulation of lysozyme. *Proteins* **27:**425–437.

Hilser VJ, Dowdy D, Oas TG and Freire E (1998) The structural distribution of cooperative interactions in proteins: Analysis of the native state ensemble. *Proc Natl Acad Sci USA* **95:**9903–9908.

Jiang F and Kim S-H (1991) "Soft docking": Matching of molecular surface cubes. *J Mol Biol* **219:**79–102.

Jones G, Willet P and Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J Mol Biol* **245:**43–53.

Keller PA, Leach SP, Luu TTT, Titmuss SJ and Griffith R (2000) Development of computational and graphical tools for the analysis of movement and flexibility in large molecules. *J Mol Graphics Mod*, in press.

Knegtel RMA, Kuntz ID and Oshiro CM (1997) Molecular docking to ensembles of protein structures. *J Mol Biol* **266:**424–440.

Lamb ML and Jorgensen WL (1997) Computational approaches to molecular recognition. *Curr Opin Chem Biol* **1:**449–457.

Leach AR (1994) Ligand docking to proteins with discrete side-chain flexibility. *J Mol Biol* **235:**345–356.

Lins RD, Briggs JM, Straatsma TP, Carlson HA, Greenwald J, Choe S and McCammon JA (1999) Molecular dynamics studies on the HIV-1 integrase catalytic domain. *Biophys J* **76:**2999–3011.

Luty BA, Wasserman ZR, Stouten PFW, Hodge CN, Zacharias M and McCammon JA (1995) A molecular mechanics/grid method evaluation of ligand-receptor interactions. *J Comput Chem* **16:**454–464.

Ma B, Kumar S, Tsai C-J and Nussinov R (1999) Folding funnels and binding mechanisms. *Protein Eng* **12:**713–720.

Maiorov V and Abagyan R (1997) A new method for modeling large-scale rearrangements of protein domains. *Proteins* **27:**410–424.

Mangoni M, Roccatano D and Nola AD (1999) Docking of flexible ligands to flexible receptors in solution by molecular dynamics simulation. *Proteins* **35:**153–162.

Masukawa KM, Carlson HA and McCammon JA (2000) Technique for developing a pharmacophore model that accommodates inherent protein flexibility: An application to HIV-1 integrase, in *Pharmacophore Perception, Development, and Use in Drug Design* (Güner OF ed) in press, International University Line, La Jolla, CA.

Nakajima N, Higo J, Kidera A and Nakamura H (1997a) Flexible docking of a ligand peptide to a receptor protein by multicanonical molecular dynamics simulation. *Chem Phys Lett* **278:**297–301.

Nakajima N, Nakamura H and Kidera A (1997b) Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides. *J Phys Chem B* **101:**817–824.

Onuchic JN (1997) Contacting the protein folding funnel with NMR. *Proc Natl Acad Sci USA* **94:**7129–7131.

Pang Y-P and Kozikowski AP (1994) Prediction of the binding sites of huperizine A in acetylcholinesterase by docking studies. *J Comp Aided Mol Design* **8:**669–681.

Philippopoulos M and Lim C (1999) Exploring the dynamic information content of a protein NMR structure: Comparison of a molecular dynamics simulation with the NMR and X-ray structures of *Escherichia coli* ribonuclease HI. *Proteins* **36:**87–110.

Rejto PA and Freer ST (1996) Protein conformational substates from X-ray crystallography. *Prog Biophys Mol Biol* **66:**167–196.

Sandak B, Wolfson HJ and Nussinov R (1998) Flexible docking allowing induced fit in proteins: Insights from an open to closed conformational isomers. *Proteins* **32:**159–174.

Schaffer L and Verkhivker GM (1998) Predicting structural effects in HIV-1 protease mutant complexes with flexible ligand docking and protein side-chain optimization. *Proteins* **33:**295–310.

Schnecke V, Swanson CA, Getzoff ED, Tainer JA and Kuhn LA (1998) Screening a peptidyl database for potential ligands to proteins with side-chain flexibility. *Proteins* **33:**74–87.

Shoemaker BA, Wang J and Wolynes PG (1997) Structural correlations in protein folding funnels. *Proc Natl Acad Sci USA* **94:**777–782.

Sudbeck EA, Mao C, Vig R, Venkatachalam TK, Tuel-Ahlgren L and Uckun FM (1998) Structure-based design of novel dihydroalkoxybenzyloxopyrimidine derivatives as potent nonnucleoside inhibitors of the human immunodeficiency virus reverse transcriptase. *Antimicrob Agents Chemother* **42:**3225–3233.

Totrov M and Abagyan R (1997) Flexible protein-ligand docking by global energy optimization in internal coordinates. *Proteins* **1 (Suppl):**215–220.

Van de Waterbeemd H, Carter RE, Grassy G, Kubinyi H, Martin YC, Tute MS and Willett P (1997) Glossary of terms used in computational drug design. *Pure Appl Chem* **69:**1137–1152.

Walters WP, Stahl MT and Murcko MA (1998) Virtual screening—an overview. *Drug Discov Today* **3:**160–178.

Wriggers W and Schulten K (1997) Protein domain movements: Detection of rigid domains and visualization of hinges in comparisons of atomic coordinates. *Proteins* **29:**1–14.

Zacharias M and Sklenar H (1999) Harmonic modes as variables to approximately account for receptor flexibility in ligand-receptor docking simulations: Application to DNA minor groove ligand complex. *J Comput Chem* **20:**287–300.

Zheng Q and Kyle DJ (1997) Computational screening of combinatorial libraries via multicopy sampling. *Drug Discov Today* **2:**229–234.

Send reprint requests to: Dr. Carlson, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093. E-mail: hcarlson@mccammon.ucsd.edu