

PROTEIN FLEXIBILITY IN *IN SILICO* SCREENING

HOLGER GOHLKE
HANNES KOPITZ
Department of Mathematics
and Natural Sciences,
Heinrich-Heine-University,
Düsseldorf, Germany

1. INTRODUCTION

Flexibility and mobility are essential for the function of biological macromolecules. As a prominent example, adenylate kinase undergoes large conformational changes of its domains during a catalytic cycle [1,2]. These movements are coupled to small-amplitude fluctuations on the picosecond timescale of backbone atoms, such that local backbone conformational changes are required for the lid closure to occur [3]. As another example, the ribosome behaves in a highly dynamic fashion and so acts as a molecular “machine,” with the large-scale, low-frequency, ratchet-like movements of both subunits as the most dominant motions [4–10]. However, motions on a smaller scale are no less important, as demonstrated by the active role of the ribosomal exit tunnel in cotranslational processes [11]. Overall, these examples demonstrate that biological macromolecules have an intrinsic ability to switch between conformationally distinct states under native conditions and that conformational transitions occur on a wide range of scales, both in time and in space.

The ability to undergo conformational transitions becomes particularly pronounced in the case of ligand binding to several pharmacologically important proteins, for example, HIV-1 protease [12] (Fig. 1), aldose reductase [13], FK506 binding protein [14], renin [15], and DHFR [14]. The mutual conformational adaptation of binding partners is referred to as *plasticity*. Receptor plasticity is also a hallmark of DNA [16] or RNA [17–20] targets. Frequently, hydrophobic interactions between receptor and ligand lead to the observed plasticity, giving rise to the notion of a “hydrophobic collapse” of a receptor around a

ligand [21]. Likewise, binding pockets that show large conformational changes have been found to be dominated by hydrophobic–hydrophobic or aromatic–aromatic residue pair interactions, whereas pockets that do not undergo conformational changes are dominated by mostly polar interactions [22]. Disorder-to-order transitions of intrinsically unfolded proteins upon complex formation are found as an extreme case of collapse [23–25].

Motion (or mobility) of the biomacromolecule is a prerequisite for receptor plasticity. Mobility describes actual movements in terms of directions and amplitudes. Flexibility (and the opposite, *rigidity*), however, is a static property that only determines the possibility of motion, whereas nothing actually moves [26]. As such, flexibility is not necessarily a prerequisite for mobility, as rigid parts of a biomolecule, for example, domains, may well move as a whole, for example, when connected by hinges. Still, knowing which parts of a biomolecule are flexible or rigid is valuable because it considerably simplifies the task of modeling biomacromolecular mobility.

The above examples of receptor plasticity demonstrate that the “rigid receptor hypothesis” [21], which is based on the “lock and key” model [27] of molecular recognition and has served as an underlying principle in structure-based ligand design (SBLD) and *in silico* screening, is no longer tenable. Instead, the ability to understand and predict receptor plasticity, and hence, to deal with protein flexibility and mobility, becomes central for a more in-depth understanding of molecular recognition processes [28] and success in SBLD and *in silico* screening. In fact, if protein–ligand docking is performed with the assumption of a rigid active site in those cases where actual protein movements have been observed, a dramatic drop of the docking accuracy is observed [29,30]. Whereas a docking success rate of 76% was reported for docking a ligand back to the protein structure derived from the ligand’s cocrystal structure (“redocking”), this rate dropped to only 49% when the ligands were docked against protein structures derived from other ligand’s cocrystal structures (“cross-docking”) [30]. Similar drop-offs have also been reported by others [31,32]. The drop in docking accuracy

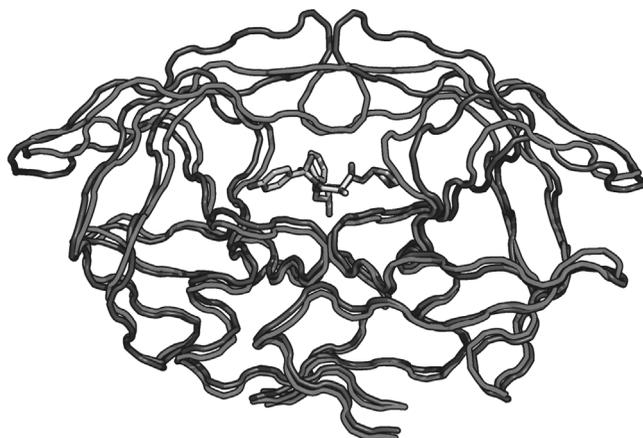


Figure 1. HIV-1 protease as an example for conformational changes between *apo* (dark grey/blue) and *holo* (light grey/green) structures. The *holo* structure is bound to amprenavir.

was often found to be mirrored by the degree to which the protein moves upon ligand binding [31,33], so that docking to an *apo* form usually shows the largest deterioration [29].

Notably, biomacromolecular flexibility and mobility is not limited to influencing only steric complementarity and, hence, direct interactions between the binding partners. Rather, flexibility and mobility and their changes upon complex formation can give rise to pronounced additional, indirect energetic and entropic contributions to the binding affinity [28]. At present, these contributions are almost completely neglected in current SBLD approaches. Finally, protein flexibility and mobility provide a basis for *allosteric regulation*, which is considered to be advantageous compared to competitive regulation [34]. In contrast to the “classical view” [35,36], where allostery is explained by the occurrence of conformational changes, the “modern view of allostery” [37] considers the allosteric modulator to change the population of conformational substates that are already accessible in the global ensemble of the system [38,39]. This leads to differences in binding or signal transduction properties of the system [34,36]. Thus, this view emphasizes the role of (changes in) the protein flexibility as an entropic carrier of allosteric information [37,40–42]. Several opportunities exist to exploit the “modern view of allostery” for SBLD [43], which require a

thorough understanding and accurate modeling of protein flexibility and mobility, however.

In this chapter, we will focus on the first two aspects of biomacromolecular flexibility and mobility in SBLD: modeling receptor plasticity and the energetic and entropic contributions to the binding affinity. With respect to the first, we particularly consider how protein plasticity can be incorporated into docking approaches. We feel that, from a computational point of view, these fields have seen the most progress over recent years.

2. MODELING RECEPTOR PLASTICITY IN SBLD

For incorporating protein plasticity into docking approaches, two requirements need to be met. First, one needs to detect what can move and how. Second, this knowledge needs to be transformed into a docking algorithm. As for the former, knowledge about moving protein parts can be gained from experimental information as well as established techniques such as molecular dynamics (MD) simulations, graph-theoretical approaches, or normal mode analysis. As for the latter, a multitude of approaches has been presented that range from considering protein plasticity only implicitly to modeling side-chain movements to also including backbone changes. In all cases,

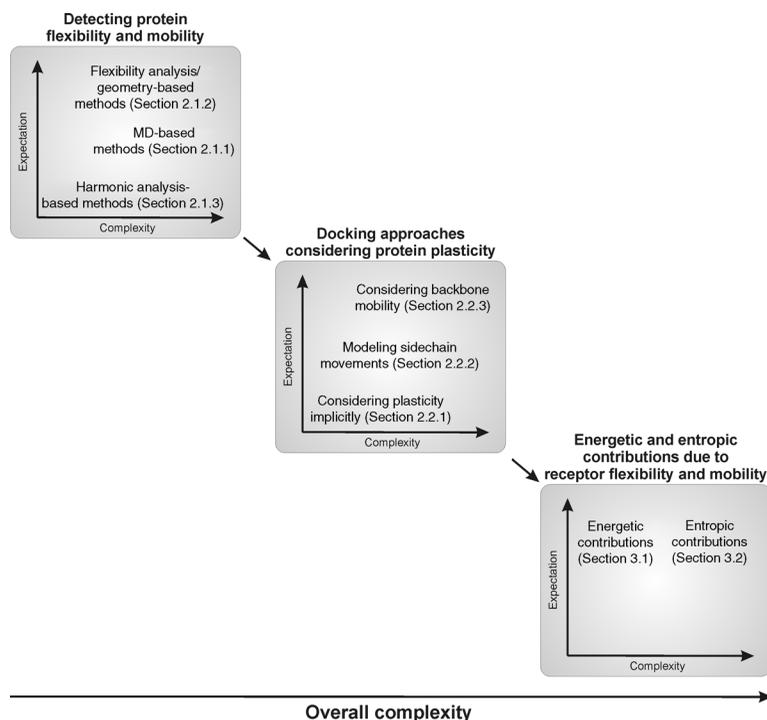


Figure 2. Flow chart showing the different approaches described in this chapter. For each subgroup, the level that we expect each method to impact the modeling of protein flexibility in *in silico* screening is related to the method's complexity. In addition, the overall complexity of the different tasks is depicted. Numbers in parentheses denote the section where each approach is described.

making a tradeoff between accuracy and efficiency is required. Figure 2 shows a flow chart summary of the different approaches described in this chapter. Table 1 lists Web site addresses of programs or servers implementing these algorithms.

2.1. Determining Protein Flexibility and Mobility

Protein plasticity comprises a range of possible movements, from single side chains to drastic structural rearrangements as seen in calmodulin [44]. Recent studies indicate, however, that with little effort much should be gained. Najmanovich et al. [45] indicated that many conformational changes can be captured by side-chain motions only, based on the finding that rotations in side-chains of up to three residues account for $\approx 85\%$ of all the cases where there is a conformational change upon ligand binding. These findings

were corroborated by a study by Zavodszky and Kuhn [46], according to which side-chain rotations were proven to be necessary for correctly docking about half of the protein–ligand complexes considered. However, 95% of the rotations were smaller than 45° . Regarding backbone ϕ/ψ changes, Gunasekaran and Nussinov [22] concluded that such dihedral angle changes upon ligand binding are in general minimal, with the most frequent large changes between the right-handed α -helical and the extended regions. On the contrary, changes between the α_R and the α_L regions were almost never seen, as they would require crossing a high-energy barrier. Apparently, these studies suggest that a protein follows a minimal energy penalty pathway to achieve required conformational changes. Considering such findings will help in developing efficient fully flexible protein–ligand docking approaches (see Section 2.2).

Table 1. Web site Addresses of Programs or Servers Implementing Algorithms Described in this Chapter

Program/Server	Web Address	Section	Reference(s)
BOSS	http://www.cemcomco.com/BOSS_and_MCPRO_Distribution125.html	2.1.1	[51]
FIRST	http://flexweb.asu.edu/software/first/	2.1.2	[56]
ProFlex	http://www.kuhnlab.bmb.msu.edu/software/proflex/		
ROCK	http://www.bch.msu.edu/~kuhn/software/rock/index.html	2.1.2	[63]
FRODA	http://flexweb.asu.edu/software/first/	2.1.2	[65]
CONCOORD	http://www.mpibpc.mpg.de/groups/de_groot/conCOORD/index.html	2.1.2	[69,70]
Ognm	http://ignm.cccb.pitt.edu/GNM_Online_Calculation.htm	2.1.3	[83,84]
SOFTDOCK	http://bio.iphy.ac.cn/#Software	2.2.1	[111]
DrugScoreONLINE	http://pc1664.pharmazie.uni-marburg.de/drugscore/	2.2.1, 2.2.3	[114]
M-Score	http://sw16.im.med.umich.edu/lab/members/chaoyie/index.html#mscore	2.2.1	[116]
GOLD	http://www.ccdc.cam.ac.uk/products/life_sciences/gold/	2.2.2	[123–125]
AutoDock 4	http://autodock.scripps.edu/	2.2.2	[130]
SLIDE	http://www.bch.msu.edu/~kuhn/software/slide/index.html	2.2.2	[131]
FlexE	http://www.biosolveit.de/FlexX-Ensemble/	2.2.3	[141]
ATTRACT in PTools	http://www.ibpc.fr/chantal/www/ptools/index.html	2.2.3	[142]
FLIPDock	http://flipdock.scripps.edu/	2.2.3	[146]
Fleksy	http://yasara.org/docking.htm#fleksy	2.2.3	[150]
RosettaLigand	http://www.rosettacommons.org/software/index.html	2.2.3	[151,152]
PyRosetta	http://pyrosetta.org/		
Surflex	Available via email request to ajain@jainlab.org	2.2.3	[153]
Discovery Studio	http://accelrys.com/products/discovery-studio/	2.2.3	[154]
M2, ACCENT	http://gilsonlab.umbi.umd.edu/software1a.html	3.2	[36,185,186]

Here, we focus on techniques for the prediction of protein flexibility and mobility that are either characterized by well-established methods that are increasingly applied in SBLD or considerable developments recently. In particular, we emphasize three distinct areas of research: molecular dynamics simulation-based methods, flexibility analysis-based and geometry-based methods, and harmonic analysis-based methods.

2.1.1. Molecular Dynamics Simulations-Based Methods Molecular dynamics simulation is one of the most widely applied and accurate computational techniques currently being used in the field of macromolecular computa-

tion [47–49]. By analyzing the trajectory of a molecular structure through phase space, important information related to SBLD can be derived. This includes the analysis of flexible and mobile regions of the macromolecule and the generation of multiple conformations of a protein.

Regarding receptor plasticity, the potential of MD to enhance SBLD was demonstrated by a recent study, which predicted an intermittent opening of an unknown favorable binding trench adjacent to the catalytic site for HIV-1 integrase [14]. The prediction was experimentally validated later on [50]. Likewise, Carlson et al. [51] developed a new method for generating dynamic pharmacophore models

to compensate for the inherent plasticity of an active site. This method identifies conserved binding regions over multiple configurations of the active site and uses those regions to define a complementary model.

Improvements in MD simulation techniques and increased computational power have recently allowed performing MD simulations on *unbound* protein states that clearly show a potential for generating conformations that mimic *bound* states. Using an artificially low solvent viscosity in a MD simulation [52] of unliganded HIV-1 protease, a comprehensive sampling of the protein's conformational space was achieved, which shed light on the flap dynamics of the protein. Remarkably, multiple conversions between a "closed" and a "semiopen" form observed in crystal structures of inhibitor bound and unliganded protein, respectively, were observed, including reversal of the flap "handedness." In an MD study [53], starting from the unbound form of aldose reductase, a set of distinct conformational substates was identified, which may prove useful as input for flexible docking approaches (see Section 2.2). Similarly encouraging results were reported for the identification of transient pockets in protein-protein binding epitopes by the groups of Helms [54] and Gohlke [55]. Overall, these results are in agreement with the "conformational selection" model [38,39], according to which an appropriate receptor conformation is "picked" during binding from an ensemble of rapidly interconverting conformational species of the unbound macromolecules.

2.1.2. Methods Based on Flexibility Analysis and Geometry-Based Methods Methods based on flexibility analysis of biomacromolecules and geometric simulations are becoming valuable alternatives to force field-based MD simulations for predicting receptor flexibility and mobility due to their lower computational burden. For analyzing protein flexibility, a single, static structure is, first, represented as a so-called bond-bending network or molecular framework, where nodes represent atoms and edges represent covalent and non-covalent interactions. A fast combinatorial algorithm, the "pebble game," then identifies rigid, overrigid, and flexible regions by count-

ing bond-rotational degrees of freedom in the network. This algorithm has been implemented into the FIRST (floppy inclusion and rigid substructure topology) software [56]. FIRST analyses have been used to accurately identify rigid regions as well as collectively and independently moving regions in a series of proteins [57,58]. Recently, the approach has also been extended to RNA structures [59,60]. FIRST results allow investigation of the internal degrees of freedom within a biomacromolecule at different levels of detail: (1) flexibility characteristics at the bond level are instructive for the analysis of binding site regions; (2) flexibility characteristics of larger regions can be related to potential global movements; and (3) rigid cluster decompositions provide hints about movements of structural parts as rigid bodies. The results also allow analysis of the long-range aspects of rigidity percolation, according to which local interactions between a ligand and a receptor may lead to a rigidification of the system at distances beyond the interaction range. Such an effect has been found in the case of a protein-protein complex formation [61], where additional interactions across the interface led to a propagation of rigidity through the binding partners.

As mentioned above, the flexibility analysis provides a natural coarse-graining of macromolecules based on rigid regions [62]. This knowledge can be further exploited for simulating protein mobility using constrained geometric simulation (CGS) [63–65] or as input for elastic network-based approaches (see Section 2.1.3) [66]. CGS-based methods either explore the rigidity-restricted conformational space by satisfying ring closure equations, as given in the program ROCK (rigidity optimized conformational kinetics) [63], or rearrange flexible and rigid parts as "ghost templates," as implemented in FRODA (framework rigidity optimized dynamic algorithm) [65]. From an application point of view, structures resulting from these approaches were used in flexible docking to cyclophilin and the estrogen receptor [67] or in docking studies of the multisubunit protein complex photosystem I [68].

An alternative geometry-based simulation method is CONCOORD (from CONstraint to

COORDinates) [69,70], which applies concepts from distance geometry [71] and generates conformations by satisfying distances and angle constraints. In an application to hyaluronate lyase [72], whose size precludes the use of MD to investigate biologically relevant timescales, flexibility (allosteric) information and functional implications were derived by CONCOORD. Furthermore, Mustard and Ritchie [73] showed that docking to multiple structures, which were obtained by an essential dynamics study following a CONCOORD run, generates better docking predictions than docking only to unbound or modeled structures.

2.1.3. Harmonic Analysis-Based Methods For a long time, normal mode analyses (NMA) have been used to study large-amplitude motions in biomacromolecules [74–76]. About a decade ago, elastic network models (ENMs) were developed, which use simplified force fields [77] and coarse-grained macromolecular models [66,78–82]: the Gaussian network model (GNM) [83,84] and the anisotropic network model (ANM) [85–87]. Since then, ENM have been applied to a vast range of problems concerning the flexibility/mobility of biomacromolecules [88–91].

Notably, conformational changes upon ligand binding are found for most of the proteins to occur along the lowest energy (frequency) modes calculated by NMA or ENM of the *unbound* protein. These modes usually involve large-amplitude and correlated motions [74,78,87,92]. Accordingly, ENMs have been mostly applied as an *a posteriori* analysis in combination with experimental studies, for example, for examining functional dynamics in *Escherichia coli* adenylate kinase, HIV-1 reverse transcriptase, and influenza virus hemagglutinin [93–95], cooperative and allosteric dynamics in tryptophan synthase [96], and binding effects in HIV-1 reverse transcriptase [94]. In agreement with the above findings, ENM have also been successfully used for efficient conformation and pathway generation, which can be exploited for docking studies [97–100] (see Section 2.2.3).

In addition to large-scale dynamical analysis, in particular, GNM has been successfully applied for residue level analyses,

too [84,101–103]. At first sight, this is surprising, considering the simplicity and coarse-graining of the underlying model. Yet, high-frequency modes of GNM have been shown to be important for the identification of binding “hot spot” residues [84], catalytic residues [102], and protein-binding sites [101]. In SBLD, these studies bear the potential for efficiently identifying binding hotspot and catalytic residues. As a hybrid approach, harmonic analyses have been combined with methods that provide atomic detail such as MD. MD/NMA hybrid methods have been proposed [104–106] for amplifying collective motions along normal mode directions in a conventional MD; thus, leading to a better sampling of certain phase space regions. This method was successfully used for docking in the case of HIV-1 protease [105].

Loop motions play an important role in accommodating ligands in binding pockets, but are hard to predict. As such, for highly mobile receptors, the first few low-frequency modes identified by NMA or ENM may not describe movements of the loop region. To overcome this problem, Cavasotto et al. [100] introduced a measure of relevance of normal modes for important loop conformational changes and found that only a few low-frequency modes, usually not among the first ones, are sufficient to represent binding pocket mobility in protein kinases.

Recently, a FIRST analysis-based coarse-graining of protein structures was combined with an ENM analysis, resulting in a rigid cluster normal mode analysis (RCNMA) [66]. With this approach, predicted directions and magnitudes of receptor motions upon ligand binding agree well with experimentally determined ones (Fig. 3a), despite embracing, in extreme cases, >50% of the protein into one rigid cluster. In fact, the results of the method are in general comparable with when no coarse-graining or a uniform coarse-graining is applied; and the results are superior if the movement is dominated by loop or fragment motions. This can be explained by the fact that the appropriate coarse-graining removes irrelevant modes of the system, whereas modes related to flexible regions become more emphasized. Overall, low-frequency modes representing functional motions in flexible

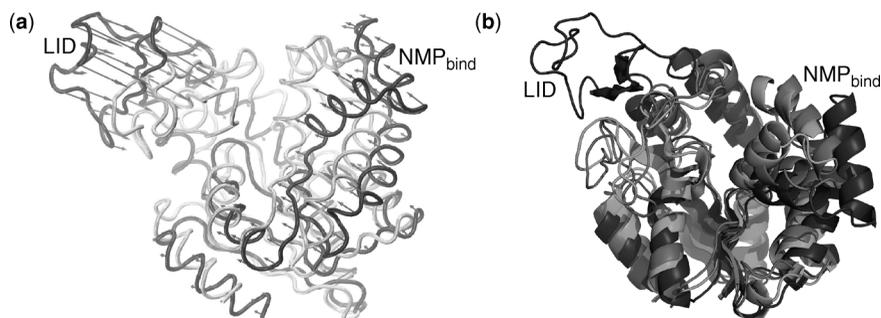


Figure 3. Superimposition of *apo* (dark grey/blue) and *holo* (light grey/green) conformations of adenylate kinase. In (a), the amplitudes and directions of motions as predicted by the mode most involved in the conformational change are depicted as red arrows. The mode was calculated using the RCNMA approach [66] and the *apo* protein conformation. The amplitudes of the motions were scaled for best graphical representation. In (b), additionally, a structure generated by NMSim calculations [107] is given in medium grey/magenta. This structure has moved by about 5 Å RMSD from the *apo* starting structure and comes as close as 2.4 Å RMSD to the *holo* conformation, although no knowledge about the bound ligand was used during simulation. Figure 3a is reproduced from Ref. [66].

regions are obtained. As a further extension of the approach, biomacromolecule conformations were then generated by deforming structures along low-energy normal mode directions predicted by RCNMA plus random direction components. Afterwards, the generated structures were iteratively corrected regarding steric clashes or stereochemical violations. The last step, termed NMSim, is similar in spirit to MD/NMA hybrid methods, but uses constrained geometric simulations instead of more time-consuming MD simulations. In total, when applied repeatedly over all three steps, the procedure efficiently generates series of conformations that lie preferentially in the low energy subspace of normal modes (Fig. 3b) [107].

2.2. Protein–Ligand Docking Approaches that Consider Protein Plasticity

Once mobility information about a protein is known, this knowledge needs to be incorporated in the protein–ligand docking algorithms. Computational efficiency is critical in this step for a thorough sampling of protein conformational space in the available computational time. Further difficulties arise when it comes to recognizing energetically accessible protein conformations among the set of decoys generated (see Section 3.1). Both findings led to the conclusion that docking includ-

ing receptor plasticity is as difficult as the protein folding problem [108]. In addition to algorithmic developments, both the development of computer hardware and a better understanding of biophysical processes on the microscopic level, which lead to new models for the description of movements in biomacromolecules, have aided in considering receptor plasticity. Still, fully flexible docking requires a tradeoff between prediction quality and calculation time.

In the following sections, an overview of current fully flexible protein–ligand docking methods, which have seen a lively development recently, is given (Fig. 4). These methods can be divided into three major groups: (1) Plasticity is considered *implicitly*. In many cases, this is achieved by modifying already existing methods; (2) Side-chain conformational changes in the binding pocket are modeled; (3) Large-scale conformational changes including backbone motions are taken into account.

2.2.1. Considering Plasticity Implicitly

From an algorithmic point of view, soft docking is the simplest approach for considering receptor plasticity. Here, repulsive contributions to the energy function are reduced either by capping van der Waals contributions or by reducing the van der Waals radii of the receptor atoms [109,110]. This allows a ligand atom to

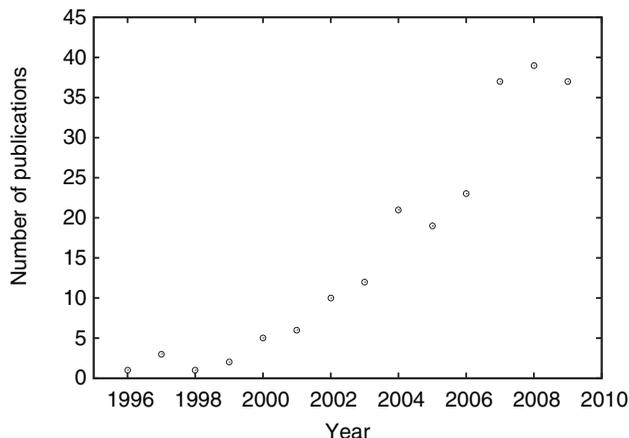


Figure 4. The number of publications dealing with protein plasticity in docking over the past 10 years. The data was derived from the ISI Web of Science database (updated Nov 30, 2009) using the query “docking AND (“flexible protein” OR “protein flexibilit*” OR “protein plasticit*”).”

interpenetrate a receptor surface to a certain degree. In the case of only small side-chain rearrangements in the pocket upon binding, this should be sufficient to find a right binding pose for a ligand even when docking to an *apo* protein conformation. As a main advantage, this approach can be easily implemented into existing docking algorithms [111]. Furthermore, the approach should not be slower than docking considering ligand flexibility only, because no additional degrees of freedom are added to the search problem.

Along these lines, the reduced steepness of knowledge-based potentials has been recognized as an advantage over force field-based or empirical scoring functions [33,112,113], as the former functions are more robust to small changes in a receptor conformation. The potential softness can be amplified during the derivation process by “smearing” observed interactions over a larger distance range, as applied for the DrugScore scoring functions [114,115]. When extending this idea by more explicitly considering atom mobilities of protein atoms based on crystallographic B factors, however, only a slight improvement in the correlation between experimental and calculated binding affinities was found compared to when protein atomic mobility was not considered [116].

Finally, receptor plasticity can also be represented based on multiple structures, which

have been derived from experiment or computational approaches described above (see Section 2.1). Here, we will focus on approaches where properties of several conformations are combined into one protein representation, leading to a mean-field approach, rather than considering the trivial solution to run a parallel docking against every conformation [117,118]. Knegt et al. [119] merged interaction energy contributions of different conformations of the target protein derived from NMR or crystallography into one “average” grid and then, performed a rigid ligand–rigid receptor docking with DOCK [120]. When tested on a database of about 150 small molecules to which known HIV-1 protease ligands were added, all active ligands could be ranked within the top 17% of the database when using the merged grids. This compares favorably to recognizing at most three inhibitors within the top 25% in the case of single structure grids. Similarly, Broughton et al. [121] used combinations of several conformations to obtain merged grids. Here, the conformations were obtained by applying short MD simulations. Österberg et al. [122] pointed out that the way how grids are combined and weighted has a dramatic influence on the outcome of the docking result. They showed that naïve mean, minimum, and maximum grids performed badly, whereas a weighted averaging already applied by

Knegt et al. as well as a Boltzmann-weighting scheme succeeded in carrying over important binding determinants of multiple structures into a single grid representation, resulting in good dockings. After all, while promising, one should note that mean-field representations bear the risk to lead to unphysical protein representations. For example, if a moving side-chain leads to two distinct interaction regions in two different receptor conformations, the combined grid may falsely represent both these regions.

2.2.2. Modeling Side-Chain Conformational Changes

In proteins that undergo only small conformational changes upon ligand binding, side chains move as little as necessary to achieve a collision-free complex, according to the “minimal rotation hypothesis” [46]. This provides the basis for docking approaches in that protein plasticity is modeled by avoiding collisions of side chains with the ligand in the binding pocket. One of the early docking programs taking partial receptor plasticity into account was GOLD [123–125], which optimizes ligand placement and conformations of some terminal bonds of receptor side chains by a genetic algorithm (GA). Thereby, the program fosters the construction of a hydrogen-bond network. GOLD was evaluated on a test set of 100 protein complexes of which it could predict “good” complex configurations in 73% of the cases, albeit with more moderate results for hydrophobic ligands. Systematic problems in ranking very polar ligands and in ranking general ligands in large cavities have been reported, too [126]. Based on earlier developments [127–129], the latest AutoDock (version 4) now also supports side-chain mobility [130]. The user can choose side chains that should be considered explicitly during the docking run. Both the ligand configuration and the side-chain conformations are simultaneously optimized during docking, similar to GOLD. As a drawback, the number of side chains set flexible needs to be restricted to assure suitable calculation times.

As an alternative, the program SPECITOPPE [109] uses side-chain rotations at a late stage of the docking in order to remove clashes between the ligand and the receptor. The side-chain rotations are done in an iterative man-

ner: if a clash is detected, the bond of the side-chain next to the clashing atoms is rotated to remove the clash. If a clash free conformation cannot be obtained that way, another bond next to the ligand will be used. The approach was used to screen a library of 140,000 peptide fragments for the targets serine protease, a DNA repair enzyme, an aspartic proteinase, and a glycosyltransferase, each of which took about an hour. SPECITOPPE was able to narrow down the library of ~140,000 peptides to 10–40 potential ligands for each of the investigated systems. Based on this concept, Schnecke and Kuhn have also proposed the docking algorithm SLIDE [131]. Here, an anchor fragment of the ligand is placed in a first step. Afterwards, the remaining fragments are added according to their database conformation, and clashes of the protein side chains and the ligand are corrected by rotation around the side-chain and ligand bonds in the nonanchor regions. As a disadvantage, only a restricted flexibility of the ligand and receptor is considered that way.

By combining soft docking with a postdocking optimization, Mizutani et al. [110] recently proposed an approach based on ADAM [132] that allows the consideration of side-chain movements. After a docking with ADAM into a “softened” cavity, clashes with protein atoms are detected and removed with the energy minimization program BLUTO by moving protein atoms locally. For a test case of 18 complexes, this approach performed better than the docking algorithms FlexX [133], Glide [134], and GOLD [124] with respect to ligand root mean square deviation (RMSD) of the top-ranked solutions.

Rotamer libraries that contain information about favorable discrete conformations provide a way to account for side-chain mobility without sampling the full conformational space of the side chain [135–137]. Following this idea, Frimurer et al. [135] used a rotamer library to generate an ensemble of tyrosine phosphate B1 structures by using three selected active site residues. The conformers were then used for flexible-ligand docking with FlexX [133]. Similarly, Kallblad and Dean [136] used a rotamer library to generate numerous conformers from a single MMP-1 structure. Subsequently, a core ensemble of

conformers was extracted by statistical methods. A synthetic inhibitor was then rigidly docked to this core ensemble using GOLD. This resulted in a good docking pose (RMSD 1.25 Å), in contrast to docking to the experimental conformation or an energy-minimized version.

2.2.3. Considering Backbone Mobility The consideration of receptor backbone degrees of freedom for docking results in a very large search space. In addition, when large-scale macromolecular motions are modeled the issue of scoring the generated receptor conformation becomes critical: disfavorable receptor conformations need to be recognized as such, as they will contribute adversely to the overall binding affinity (see Section 3.1). Approaches that consider backbone mobility follow one of four main strategic ideas: (1) (parallel) docking into multiple conformations; (2) structurally combining multiple conformations; (3) modeling protein motions in reduced coordinates; and (4) pose optimization with MD simulations.

Regarding parallel docking, the algorithm of IFREDA (ICM-flexible receptor docking algorithm) [32] consists of three main steps. In the first step, an ensemble of receptor conformations is generated by placing the ligand at arbitrary positions and orientations into a binding pocket, followed by an energy minimization of these conformations. In this step, side-chain and essential backbone movements are taken into account, leading to an “induced-fit” of the binding pocket. As a drawback, a number of binders need to be known already. In the second step, the receptor conformations are used for a fast rigid receptor, flexible ligand docking. This is finally followed by merging the screening scores and keeping the best rank for each compound. IFREDA was tested for seven protein kinases complexes, for which it was able to generate correct conformations (RMSD < 2 Å) for 70% of the ligands in a cross-docking experiment. The dataset included structures with a backbone motion of up to 2 Å RMSD. In an attempt to improve the performance over sequential docking to multiple receptor conformations, novel ensemble docking methods have been introduced [138–140]. Here, the algorithm automa-

tically selects a good protein structure during docking by simultaneously optimizing ligand configuration and protein conformation.

As for structurally combining multiple conformations, FlexE [141] uses experimentally determined or computed conformations as a “basis set” of receptor plasticity. By structural combination, novel adopted conformations of the receptor are generated. The search for receptor conformations is based on a so-called united protein description. This protein description prunes the search space, but at the same time allows combining alternatives to new structures that were not contained in the ensemble. That way, FlexE can account for mobility in terms of loop movements and side-chain movements. Still, the authors propose to combine FlexE with a rotamer library approach to achieve even better side-chain predictions. Likewise, FlexE’s capability to consider large main-chain variations, such as domain movements, is limited.

Vibrational modes, either obtained from normal mode analysis, ENM, or MD-derived principal components (see Section 2.1.3), describe receptor movements in terms of collective motions. Using such modes as independent variables to describe receptor motions, thus, considerably reduces the complexity of the conformational search space. One of the earliest approaches along these lines was presented by Zacharias and Sklenar [142]. In docking ligands to DNA, harmonic modes derived from the energy-minimized free DNA were combined to model receptor plasticity. About 5–40 low-frequency eigenvectors were sufficient to successfully deform the DNA from the unbound structure to the ligand-bound form, sometimes coming as close as 0.5 Å to the latter. In this early approach, high-frequency movements, for example, rotations of binding pocket side chains, were not considered. This drawback has been overcome recently for protein–ligand docking by May and Zacharias by additionally modeling side-chain mobility in terms of discrete rotamers [143]. Notably, the use of modes may not be successful when it comes to modeling curvilinear receptor motions. It is important to point out, however, that with the use of harmonic modes it is possible to calculate approximate energies for the protein deformation during the docking

run. This is an advantage over approaches that are purely geometry-based such as FlexE. In another approach, Cavasotto et al. [100] used relevant low-frequency modes to describe loop mobility in cAMP-dependent protein kinase (see Section 2.1.3).

A novel reduced representation of protein motions for docking has been introduced by Kazemi et al. [144], who developed an accurate representation of intermolecular interactions that makes use of the high efficiency in evaluating protein–ligand interaction energies from lookup tables even in the case of a moving protein. The new lookup table function for potential fields is based on irregular, deformable 3D grids (Fig. 5). The underlying idea is to adapt a 3D grid with precalculated potential field values, which were derived from an initial protein conformation, to another conformation by moving intersection points in space, but while keeping the potential field values constant. The approach was tested by docking into deformed potential grids for which DrugScore [114] potential values were initially calculated based on *apo* protein structures. The grids were then deformed towards *holo* conformations. For data sets of HIV-1 protease and the kinases CAPK, CDK2, and LCK, this resulted in docking

success rates of 67–100%. Another sophisticated data structure for encoding the conformational space of a receptor is termed flexibility tree [145]. This data structure describes the receptor as a nested system of molecular fragments, which can be involved in a variety of movement types. Implemented into the protein–ligand docking software FLIPDock [146], the data structure provides a small number of variables with which conformational subspaces are parameterized and that can be searched during docking; thus, allowing an effective modeling of receptor conformational changes.

Finally, a number of docking approaches have been published recently that are based on a two-step protocol. In the first step, methods with low accuracy are used, which are fast and applicable to large compound databases. The second step then is more accurate and time consuming. The overall goal here is to screen many compounds without compromising the accuracy. In most of the cases, this is accomplished by docking in combination with MD simulations. The docking is used for two purposes: first, as a filter to identify compounds that can actually bind to the considered target and, second, to provide decoy complex structures as starting points for MD. In

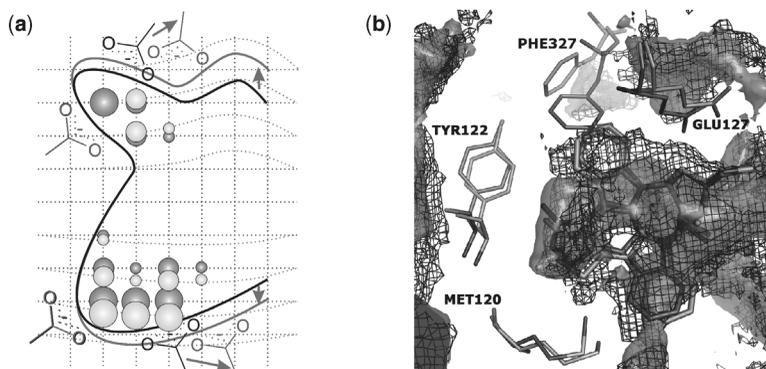


Figure 5. (a) Schematic view of the spatial deformation of potential fields inside a binding pocket according to movements of the surrounding protein. This deformation is modeled using irregular, deformable 3D grids [144]. (b) Docking of staurosporine into potential fields generated from a CAPK *apo* structure but deformed to a *holo* structure. Attractive potential fields for aromatic carbon and side-chain conformations of the *apo* structure are depicted in dark grey/blue. Deformed potential fields and side-chain conformations of the *holo* structure are depicted in light grey/orange. Staurosporine carbon atoms are displayed in light grey/green for the native structure and dark grey/magenta for the solution found for docking into the deformed grids (RMSD to the native structure: 0.64 Å). Note how the most significant movement of Phe327 “drags along” the potential field. Reproduced from Ref. [144].

the subsequent MD step, induced fit effects of the protein will be considered. During the MD step, it is also possible to incorporate explicit water molecules into the complex structure prediction, which is not provided by most of the docking approaches. An intriguing example of such an application is given by the insight gained into the preferential binding of cyclin-dependent kinase inhibitors. Here, Park et al. [147] studied the inhibition of CDK2 and CDK4 by three different selective inhibitors. The inhibitors reduced the mobility of a disordered loop of CDK4 but did not seriously affect CDK2. It was also discovered that the tighter binding of the inhibitor to CDK4 was an effect of a smaller number of water molecules moving into the binding site upon ligand binding. Such insights would not have been possible without the MD step. Furthermore, Cavalli et al. [148] pointed out that the combination of docking with subsequent MD simulations allows identifying those starting structures as being in good agreement with the experimental binding mode that lead to stable trajectories. Similar conclusions were also reported by Zoete et al. [149].

For approaches applying MD simulations, it has to be considered that even short MD runs require considerable computational time and, hence, are limited to only selected cases. Using energy minimization instead of MD simulations helps to overcome this problem, although a more thorough sampling of the protein conformational space is now required in the initial docking step. This philosophy is followed by the Fleksy approach [150], which uses a FlexE-based ensemble docking, followed by an effective Yasara-based complex optimization. Similarly, the previously published RosettaLigand method [151], which only allowed for protein side-chain degrees of freedom so far, has been extended by a stringent gradient-based minimization in the final step [152], now allowing side-chain and backbone torsions of the receptor to vary. Other approaches following the strategy of ensemble docking with subsequent receptor optimization have recently emerged [153,154]. Next, the accuracy of the docking tool used in the first step is critical for the stability of the MD trajectories and, consequently, for the success

of the complex prediction. Finally, appropriate force field parameters for ligands are required that are compatible with biomolecular force fields. Those are available from, for example, the general Amber force field (GAFF) [155], which is compatible with existing Amber biomolecular force fields and has parameters for most organic and pharmaceutical small molecules.

3. ENERGETIC AND ENTROPIC CONTRIBUTIONS TO BINDING AFFINITY DUE TO RECEPTOR FLEXIBILITY AND MOBILITY

3.1. Energetic Contributions

Receptor flexibility and mobility are not limited to influencing only steric complementarity and, hence, direct interactions between the binding partners in molecular recognition. Rather, pronounced contributions to the binding affinity arise in a more indirect manner. As such, *conformational variability of the receptor* leads to a disfavorable *reorganization energy* upon binding. This contribution can be large, even for relatively well-preorganized binding sites. In binding free energy calculations of ligands binding to avidin and streptavidin, Lazaridis et al. obtained values between 9 and 36 kcal/mol [156]. Notably, these numbers are of the same order of magnitude as the binding free energies and should clearly be included to correct predicted binding affinities when it comes to flexible receptor docking [157]. In fact, in a recent assessment of 18 available scoring functions for their accuracy to predict binding affinities and to rank-order compounds by their affinity, it was found that most of the assessed scoring functions were much less accurate when a nonnative protein conformation was provided instead of the native one [158], perhaps due to missing reorganization energy contributions. Still, we note that these values are in general difficult to compute because they correspond to the small difference between large conformational energies of the bound and the unbound receptor. Further complications arise from the fact that the protein reorganization energy may not be the same even for slightly different ligands [156].

3.2. Entropic Contributions

Changes in the receptor flexibility upon complex formation, that is, changes in the internal degrees of freedom, lead to *configurational entropy* contributions. Usually, recognition sites that are in direct contact with the binding partner become less flexible, although influences on the flexibility of residues distant from the epitope are also seen. This can be explained in that perturbations at the binding site can propagate to remote locations by altering the dynamic network of interactions in proteins [159–161]. Conversely, it is increasingly recognized that complex formation may also lead to an increase in configurational entropy (which is related to an increase in flexibility of the system), compensating for the loss of translational and rotational entropy upon association [6,162–169]. For some cases, transfer of flexibility to other protein parts (within one binding partner) has been described, leading to a redistribution of protein configurational entropy, also potentially reducing the total entropy loss [38,39,170–173].

Computing configurational entropies is notoriously difficult [174–176], and the methods can be computationally demanding. Methods that belong to this category are harmonic analyses based on static structures [177–180] or covariance matrices obtained from MD simulations [181–183]. These are largely applied in the context of MM/PB(GB)SA calculations or related approaches, with varying success [174]. Other high-level methods for calculating receptor entropies are given by the “mining minima,” [184] “second generation mining minima” [36,185] and “mutual information expansion approximation” [186] methods, which, so far, have been applied to host/guest systems only. Finally, entropic contributions can also be obtained by decomposing free energies computed by free energy perturbation or thermodynamic integration approaches [187,188] into energetic and entropic components. The overall entropy difference then contains changes in the solvent entropy in addition to solute contributions. Dauntingly, when investigating different ways to do so on a test system of liquid water, van Gunsteren and coworkers concluded that “none of the considered techniques seems suitable to

give a perspective for the calculation of the entropy of ligand–protein binding” [189]. As a major obstacle in that respect, the limited sampling of the configurational phase space has been identified.

All of the above methods are inadequate in terms of computational requirements for applications in lead finding and virtual screening. Faster but approximate approaches have been developed for application within scoring functions. In the majority of cases, ligand entropy changes are estimated as sums of constant terms penalizing the restriction of external and internal degrees of freedom of the binding partners, as originally introduced by Böhm [190]. More sophisticated approaches to treat ligand entropy contributions have been introduced in recent years [191–195]. Receptor contributions are still ignored by approximate approaches, however, except for empirical scales considering the restriction of side-chain motions within binding pockets [178]. Clearly, much more work is required in this area, as the above rigorous techniques reveal that receptor contributions should not be neglected.

4. SUMMARY AND OUTLOOK

Computational approaches in structure-based ligand design and *in silico* screening that address issues of protein flexibility and mobility have been reviewed. Motivated by the “conformational selection” model, according to which conformational states adopted by a *holo* protein may be found already by investigating conformational fluctuations of the *apo* structure, in a first step one detects what can move and how. MD simulations, graph theoretical and geometry-based approaches, or harmonic analysis-based methods are used for this. In most of the cases, these methods describe protein motions at a certain scale. Approaches that describe motions on multiple scales are notable exceptions; they usually combine two or more of the established “single-scale” techniques. We expect that further developments along the lines of multi-scale modeling will be particularly fruitful. Subsequently, the knowledge about moving protein parts needs to be included into a

docking strategy. Three main routes of algorithmic development have been identified: (1) implicit consideration of plasticity; (2) modeling of side-chain motions; and (3) accounting for large-scale conformational changes including backbone motions. Only in the latter class, receptor motions at different scales can be accounted for, and we are awaiting further exciting developments here. While the impact of receptor flexibility and mobility on structure is well received by now, energetic and entropic consequences are largely neglected so far. Progress in this area is highly necessary, in particular, as considering receptor plasticity during docking yields ample possibilities for identifying false positive ligand candidates.

ACKNOWLEDGMENTS

This work was supported by funds from Heinrich-Heine-University, Düsseldorf, Germany. We are grateful to Daniel Cashman for critically reading the manuscript.

REFERENCES

- Vonrhein C, Schlauderer GJ, Schulz GE. Movie of the structural-changes during a catalytic cycle of nucleoside monophosphate kinases. *Structure* 1995;3(5): 483–490.
- Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipavlou G, Eisenmesser EZ, Kern D. Linkage between dynamics and catalysis in a thermophilic–mesophilic enzyme pair. *Nat Struct Mol Biol* 2004;11(10): 945–949.
- Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, Kern D. A hierarchy of time-scales in protein dynamics is linked to enzyme catalysis. *Nature* 2007;450(7171): 913–916.
- Frank J, Agrawal RK. A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* 2000;406(6793): 318–322.
- Berisio R, Schlutzen F, Harms J, Bashan A, Auerbach T, Baram D, Yonath A. Structural insight into the role of the ribosomal tunnel in cellular regulation. *Nat Struct Biol* 2003;10(5): 366–370.
- Arumugam S, Gao G, Patton BL, Semenchenko V, Brew K, Van Doren SR. Increased backbone mobility in β -barrel enhances entropy gain driving binding of N-TIMP-1 to MMP-3. *J Mol Biol* 2003;327:719–734.
- Tama F, Valle M, Frank J, Brooks CL 3rd. Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. *Proc Natl Acad Sci* 2003;100(16): 9319–9323.
- Sanbonmatsu KY, Joseph S, Tung CS. Simulating movement of tRNA into the ribosome during decoding. *Proc Natl Acad Sci USA* 2005;102(44): 15854–15859.
- Horan L, Noller H. Intersubunit movement is required for ribosomal translocation. *Proc Natl Acad Sci USA* 2007;104(12): 4881–4885.
- Korostelev A, Trakhanov S, Laurberg M, Noller HF. Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. *Cell* 2006;126(6): 1065–1077.
- Fulle S, Gohlke H. Statics of the ribosomal exit tunnel: implications for cotranslational peptide folding, elongation regulation, and antibiotics binding. *J Mol Biol* 2009;387(2): 502–517.
- Wlodawer A, Vondrasek J. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu Rev Biophys Biomol Struct* 1998;27:249–284.
- Wilson DK, Tarle I, Petrash JM, Quijcho FA. Refined 1.8 Å structure of human aldose reductase complexed with the potent inhibitor zopolrestat. *Proc Natl Acad Sci USA* 1993;90(21): 9847–9851.
- Lin JH, Perryman AL, Schames JR, McCammon JA. Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *J Am Chem Soc* 2002;124(20): 5632–5633.
- Vieira E, Binggeli A, Breu V, Bur D, Fischli W, Guller R, Hirth G, Marki HP, Muller M, Oefner C, Scalone M, Stadler H, Wilhelm M, Wostl W. Substituted piperidines: highly potent renin inhibitors due to induced fit adaptation of the active site. *Bioorg Med Chem Lett* 1999;9(10): 1397–1402.
- Allemann RK, Egli M. DNA recognition and bending. *Chem Biol* 1997;4(9): 643–650.
- Puglisi JD, Chen L, Frankel AD, Williamson JR. Role of RNA structure in arginine recognition of TAR RNA. *Proc Natl Acad Sci USA* 1993;90(8): 3680–3684.
- Puglisi JD, Tan R, Calnan BJ, Frankel AD, Williamson JR. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* 1992;257(5066): 76–80.

19. Fourmy D, Yoshizawa S, Puglisi JD. Paromomycin binding induces a local conformational change in the A-site of 16 S rRNA. *J Mol Biol* 1998;277(2): 333–345.
20. Fulle S, Gohlke H. Molecular recognition of RNA: challenges for modelling interactions and plasticity. *J Mol Recognit* 2009. DOI: 10.1002/jmr.1000.
21. Davis AM, Teague SJ. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew Chem Int Ed Engl* 1999;38(6): 736–749.
22. Gunasekaran K, Nussinov R. How different are structurally flexible and rigid binding sites? Sequence and structural features discriminating proteins that do and do not undergo conformational change upon ligand binding. *J Mol Biol* 2007;365(1): 257–273.
23. Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure–function paradigm. *J Mol Biol* 1999;293(2): 321–331.
24. Levy Y, Cho SS, Onuchic JN, Wolynes PG. A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. *J Mol Biol* 2005;346:1121–1145.
25. Liu J, Tan H, Rost B. Loopy proteins appear conserved in evolution. *J Mol Biol* 2002;322: 53–64.
26. Gohlke H, Thorpe MF. A natural coarse graining for simulating large biomolecular motion. *Biophys J* 2006;91(6): 2115–2120.
27. Fischer E. Einfluss der konfiguration auf die wirkung der enzyme. *Ber Dtsch Chem Ges* 1894;27:2985–2993.
28. Gohlke H, Klebe G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew Chem Int Ed* 2002;41: 2644–2676.
29. Erickson JA, Jalaie M, Robertson DH, Lewis RA, Vieth M. Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy. *J Med Chem* 2004;47(1): 45–55.
30. Murray CW, Baxter CA, Frenkel AD. The sensitivity of the results of molecular docking to induced fit effects: application to thrombin, thermolysin and neuraminidase. *J Comput Aided Mol Des* 1999;13(6): 547–562.
31. Verdonk ML, Mortenson PN, Hall RJ, Hartshorn MJ, Murray CW. protein–ligand docking against non-native protein conformers. *J Chem Inf Model* 2008;48(11): 2214–2225.
32. Cavasotto CN, Abagyan RA. Protein flexibility in ligand docking and virtual screening to protein kinases. *J Mol Biol* 2004;337(1): 209–225.
33. Ferrara P, Gohlke H, Price DJ, Klebe G, Brooks CL. Assessing scoring functions for protein–ligand interactions. *J Med Chem* 2004;47: 3032–3047.
34. Christopoulos A. Allosteric binding sites on cell-surface receptors: novel targets for drug discovery. *Nat Rev Drug Discov* 2002;1: 198–210.
35. Monod J, Wyman J, Changeux J-P. On the nature of allosteric transitions: a plausible model. *J Mol Biol* 1965;12:88–118.
36. Changeux J-P, Edelstein SJ. Allosteric mechanisms of signal transduction. *Science* 2005;308:1424–1428.
37. Kern D, Zuiderweg ER. The role of dynamics in allosteric regulation. *Curr Opin Struct Biol* 2003;13(6): 748–757.
38. Tsai C-J, Ma B, Nussinov R. Folding and binding cascades: shifts in energy landscapes. *Proc Natl Acad Sci USA* 1999;96:9970–9972.
39. Tsai C-J, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels, and protein function. *Prot Sci* 1999;8:1181–1190.
40. Volkman BF, Lipson D, Wemmer DE, Kern D. Two-state allosteric behavior in a single-domain signaling protein. *Science* 2001;291(5512): 2429–2433.
41. Rousseau F, Schymkowitz J. A systems biology perspective on protein structural dynamics and signal transduction. *Curr Opin Struct Biol* 2005;15:23–30.
42. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? *Proteins* 2004;57:433–443.
43. Hardy JA, Wells JA. Searching for new allosteric sites in enzymes. *Curr Opin Struct Biol* 2004;14(6): 706–715.
44. Barnard D, Diaz B, Hettich L, Chuang E, Zhang XF, Avruch J, Marshall M. Identification of the sites of interaction between c-Raf-1 and Ras-GTP. *Oncogene* 1995;10(7): 1283–1290.
45. Najmanovich R, Kuttner J, Sobolev V, Edelman M. Side-chain flexibility in proteins upon ligand binding. *Proteins* 2000;39(3): 261–268.
46. Zavodszky MI, Kuhn LA. Side-chain flexibility in protein–ligand binding: the minimal rotation hypothesis. *Protein Sci* 2005;14(4): 1104–1114.

47. Hansson T, Oostenbrink C, van Gunsteren W. Molecular dynamics simulations. *Curr Opin Struct Biol* 2002;12(2): 190–196.
48. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. *Nat Struct Biol* 2002;9(9): 646–652.
49. Adcock SA, McCammon JA. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev* 2006;106(5): 1589–1615.
50. Hazuda DJ, Anthony NJ, Gomez RP, Jolly SM, Wai JS, Zhuang L, Fisher TE, Embrey M, Guare JP Jr, Egbertson MS, Vacca JP, Huff JR, Felock PJ, Witmer MV, Stillmock KA, Danovich R, Grobler J, Miller MD, Espeseth AS, Jin L, Chen IW, Lin JH, Kassahun K, Ellis JD, Wong BK, Xu W, Pearson PG, Schleif WA, Cortese R, Emini E, Summa V, Holloway MK, Young SD. A naphthyridine carboxamide provides evidence for discordant resistance between mechanistically identical inhibitors of HIV-1 integrase. *Proc Natl Acad Sci USA* 2004;101(31): 11233–11238.
51. Carlson HA, Masukawa KM, Rubins K, Bushman FD, Jorgensen WL, Lins RD, Briggs JM, McCammon JA. Developing a dynamic pharmacophore model for HIV-1 Integrase. *J Med Chem* 2000;43(11): 2100–2114.
52. Hornak V, Okur A, Rizzo RC, Simmerling C. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc Natl Acad Sci USA* 2006;103(4): 915–920.
53. Sotriffer CA, Krämer O, Klebe G. Probing flexibility and “induced-fit” phenomena in aldose reductase by comparative crystal structure analysis and molecular dynamics simulations. *Proteins* 2004;56:52–66.
54. Eyrisch S, Helms V. Transient pockets on protein surfaces involved in protein–protein interaction. *J Med Chem* 2007;50(15): 3457–3464.
55. Cozzini P, Kellogg GE, Spyraakis F, Abraham DJ, Costantino G, Emerson A, Fanelli F, Gohlke H, Kuhn LA, Morris GM, Orozco M, Pertinhez TA, Rizzi M, Sotriffer CA. Target flexibility: an emerging consideration in drug discovery and design. *J Med Chem* 2008;51(20): 6237–6255.
56. Jacobs D, Rader AJ, Kuhn L, Thorpe MF. Protein flexibility predictions using graph theory. *Proteins* 2001;44(2): 150–165.
57. Jacobs DJ, Kuhn LA, Thorpe MF. Flexible and rigid regions in proteins. In: Thorpe MF, Duxbury PM, editors. *Rigidity Theory and Applications*. New York: Kluwer Academic/Plenum Publishers; 1999. p 357–384.
58. Jacobs DJ, Rader AJ, Kuhn LA, Thorpe MF. Protein flexibility predictions using graph theory. *Proteins* 2001;44(2): 150–165.
59. Fulle S, Gohlke H. Analyzing the flexibility of RNA structures by constraint counting. *Biophys J* 2008;94(11): 4202–4219.
60. Fulle S, Gohlke H. Constraint counting on RNA structures: Linking flexibility and function. *Methods* 2009;49(2): 181–188.
61. Gohlke H, Kuhn LA, Case DA. Change in protein flexibility upon complex formation: analysis of Ras-Raf using molecular dynamics and a molecular framework approach. *Proteins* 2004;56:322–337.
62. Gohlke H, Thorpe M. A natural coarse graining for simulating large biomolecular motion. *Biophys J* 2006.
63. Lei M, Zavodszky M, Kuhn L, Thorpe MF. Sampling protein conformations and pathways. *J Comput Chem* 2004;25(9): 1133–1148.
64. Thorpe MF, Lei M, Rader AJ, Jacobs DJ, Kuhn LA. Protein flexibility and dynamics using constraint theory. *J Mol Graph Model* 2001;19(1): 60–69.
65. Wells S, Menor S, Hesperheide BM, Thorpe MF. Constrained geometric simulation of diffusive motions in proteins. *Phys Biol* 2005;2:1–10.
66. Ahmed A, Gohlke H. Multi-scale modeling of macromolecular conformational changes combining concepts from rigidity and elastic network theory. *Proteins* 2006;63:1038–1051.
67. Zavodszky M, Lei M, Thorpe MF, Day A, Kuhn L. Modeling correlated main-chain motions in proteins for flexible molecular recognition. *Proteins* 2004;57(2): 243–261.
68. Jolley CC, Wells SA, Hesperheide BM, Thorpe MF, Fromme P. Docking of photosystem I subunit C using a constrained geometric simulation. *J Am Chem Soc* 2006;128(27): 8803–8812.
69. de Groot BL, Vriend G, Berendsen HJ. Conformational changes in the chaperonin GroEL: new insights into the allosteric mechanism. *J Mol Biol* 1999;286(4): 1241–1249.
70. de Groot BL, van Aalten DMF, Scheek RM, Amadei A, Vriend G, Berendsen HJC. Prediction of protein conformational freedom from distance constraints. *Proteins* 1998;29(2): 240–251.
71. Crippen GM. Novel approach to calculation of conformation: distance geometry. *J Comput Phy* 1977;24(1): 96–107.
72. Singh S, Jedrzejewski MJ, Air GM, Luo M, Laver WG, Brouillette WJ. Structure-based inhibi-

- tors of influenza virus sialidase. A benzoic acid lead with novel interaction. *J Med Chem* 1995;38(17): 3217–3225.
73. Mustard D, Ritchie D. Docking essential dynamics eigenstructures. *Proteins* 2005;60(2): 269–274.
 74. Go N, Noguti T, Nishikawa T. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc Natl Acad Sci USA* 1983;80(12): 3696–700.
 75. Brooks B, Karplus M. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc Natl Acad Sci USA* 1983;80(21): 6571–6575.
 76. Case D. Normal mode analysis of protein dynamics. *Curr Opin Struct Biol* 1994;4(2): 285–290.
 77. Tirion MM. Large amplitude elastic motions in proteins from single-parameter atomic analysis. *Phys Rev Lett* 1996;77:1905–1908.
 78. Tama F, Gadea FX, Marques O, Sanejouand Y-H. Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins* 2000;41:1–7.
 79. Kurkuoglu O, Jernigan RL, Doruker P. Mixed levels of coarse-graining of large proteins using elastic network model succeeds in extracting the slowest motion. *Polymer* 2004;45:649–657.
 80. Doruker P, Jernigan RL, Bahar I. Dynamics of large proteins through hierarchical levels of coarse-grained structures. *J Comp Chem* 2002;23(1): 119–127.
 81. Li G, Cui Q. A coarse-grained normal mode approach for macromolecules: an efficient implementation and application to Ca^{2+} -ATPase. *Biophys J* 2002;83(5): 2457–2474.
 82. Durand P, Trinquier G, Sanejouand Y-H. A new approach for determining low-frequency normal modes in macromolecules. *Biopolymers* 1994;34:759–771.
 83. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold Des* 1997;2(3): 173–181.
 84. Bahar I, Erman B, Haliloglu T, Jernigan RL. Efficient characterization of collective motions and interresidue correlations in proteins by low-resolution simulations. *Biochemistry* 1997;36(44): 13512–13523.
 85. Hinsen K. Analysis of domain motions by approximate normal mode calculations. *Proteins* 1998;33(3): 417–429.
 86. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J* 2001;80(1): 505–515.
 87. Tama F, Sanejouand YH. Conformational change of proteins arising from normal mode calculations. *Protein Eng* 2001;14(1): 1–6.
 88. Tozzini V. Coarse-grained models for proteins. *Curr Opin Struct Biol* 2005;15:144–150.
 89. Ma J. Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure* 2005;13(3): 373–380.
 90. Ma J. New advances in normal mode analysis of supermolecular complexes and applications to structural refinement. *Curr Protein Peptide Sci* 2004;5(2): 119–123.
 91. Bahar I, Rader AJ. Coarse-grained normal mode analysis in structural biology. *Curr Opin Struct Biol* 2005;15:586–592.
 92. Brooks B, Karplus M. Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. *Proc Natl Acad Sci USA* 1985;82(15): 4995–4999.
 93. Bahar I, Erman B, Jernigan R, Atilgan A, Covell D. Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. *J Mol Biol* 1999;285(3): 1023–1037.
 94. Temiz NA, Bahar I. Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase. *Proteins* 2002;49:61–70.
 95. Isin B, Doruker P, Bahar I. Functional motions of influenza virus hemagglutinin: a structure-based analytical approach. *Biophys J* 2002;82(2): 569–581.
 96. Bahar I, Jernigan RL. Cooperative fluctuations and subunit communication in tryptophan synthase. *Biochemistry* 1999;38(12): 3478–3490.
 97. Kim MK, Jernigan RL, Chirikjian GS. Efficient generation of feasible pathways for protein conformational transitions. *Biophys J* 2002;83(3): 1620–1630.
 98. Kim MK, Chirikjian GS, Jernigan RL. Elastic models of conformational transitions in macromolecules. *J Mol Graph Model* 2002;21(2): 151–160.
 99. Miyashita O, Onuchic JN, Wolynes PG. Non-linear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proc Natl Acad Sci USA* 2003;100(22): 12570–12575.
 100. Cavasotto CN, Kovacs JA, Abagyan RA. Representing receptor flexibility in ligand dock-

- ing through relevant normal modes. *J Am Chem Soc* 2005;127(26): 9632–9640.
101. Ertekin A, Nussinov R, Haliloglu T. Association of putative concave protein-binding sites with the fluctuation behavior of residues. *Protein Sci* 2006;15(10): 2265–2277.
 102. Yang LW, Bahar I. Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes. *Structure* 2005;13(6): 893–904.
 103. Bahar I, Atilgan AR, Demirel MC, Erman B. Vibrational dynamics of folded proteins: significance of slow and fast modes in relation to function and stability. *Phys Rev Lett* 1998;80: 2733–2736.
 104. He J, Zhang Z, Shi Y, Liu H. Efficiently explore the energy landscape of proteins in molecular dynamics simulations by amplifying collective motions. *J Chem Phys* 2003;119 (7): 4005–4017.
 105. Tatsumi R, Fukunishi Y, Nakamura H. A hybrid method of molecular dynamics and harmonic dynamics for docking of flexible ligand to flexible receptor. *J Comput Chem* 2004;25(16): 1995–2005.
 106. Zhang Z, Shi Y, Liu H. Molecular dynamics simulations of peptides and proteins with amplified collective motions. *Biophys J* 2003;84 (6): 3583–3593.
 107. Ahmed A, Gohlke H. Multiscale modeling of macromolecular conformational changes. In: Nithiarasu P, Löhner R, van Loon R, editors. 1st International Conference on Computational and Mathematical Biomedical Engineering (CMBE09), Swansea, 2009. p 219–222.
 108. Halperin I, Ma B, Wolfson H, Nussinov R. Principles of docking: an overview of search algorithms and a guide to scoring functions. *Proteins* 2002;47(4): 409–443.
 109. Schnecke V, Swanson C, Getzoff E, Tainer J, Kuhn LA. Screening a peptidyl database for potential ligands to proteins with side-chain flexibility. *Proteins* 1998;33:74–87.
 110. Mizutani MY, Takamatsu Y, Ichinose T, Nakamura K, Itai A. Effective handling of induced-fit motion in flexible docking. *Proteins* 2006;63(4): 878–891.
 111. Jiang F, Kim SH. “Soft docking”: matching of molecular surface cubes. *J Mol Biol* 1991;219 (1): 79–102.
 112. Gohlke H, Klebe G. Statistical potentials and scoring functions applied to protein–ligand binding. *Curr Opin Struct Biol* 2001;11: 231–235.
 113. Brylinski M, Skolnick J. Q-Dock: low-resolution flexible ligand docking with pocket-specific threading restraints. *J Comp Chem* 2008;29 (10): 1574–1588.
 114. Gohlke H, Hendlich M, Klebe G. Knowledge-based scoring function to predict protein–ligand interactions. *J Mol Biol* 2000;295(2): 337–356.
 115. Pfeffer P, Gohlke H. DrugScoreRNA-knowledge-based scoring function to predict RNA–ligand interactions. *J Chem Inf Model* 2007;47 (5): 1868–1876.
 116. Yang CY, Wang R, Wang S. M-score: a knowledge-based potential scoring function accounting for protein atom mobility. *J Med Chem* 2006;49(20): 5903–5911.
 117. Carlson HA. Protein flexibility and drug design: how to hit a moving target. *Curr Opin Chem Biol* 2002;6:447–452.
 118. Totrov M, Abagyan R. Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr Opin Struct Biol* 2008;18(2): 178–184.
 119. Knegtel RM, Kuntz ID, Oshiro CM. Molecular docking to ensembles of protein structures. *J Mol Biol* 1997;266(2): 424–440.
 120. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule–ligand interactions. *J Mol Biol* 1982; 161:269–288.
 121. Broughton HB. A method for including protein flexibility in protein–ligand docking: improving tools for database mining and virtual screening. *J Mol Graph Model* 2000;18 (3): 247–257, 302–304.
 122. Österberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins* 2002;46: 34–40.
 123. Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J Mol Biol* 1995;245:43–53.
 124. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 1997;267(3): 727–748.
 125. Sousa SF, Fernandes PA, Ramos MJ. Protein–ligand docking: current status and future challenges. *Proteins* 2006;65(1): 15–26.
 126. Kellenberger E, Rodrigo J, Muller P, Rognan D. Comparative evaluation of eight docking

- tools for docking and virtual screening accuracy. *Proteins* 2004;57(2): 225–242.
127. Goodsell DS, Olson AJ. Automated docking of substrates to proteins by simulated annealing. *Proteins* 1990;8(3): 195–202.
 128. Goodsell DS, Morris GM, Olson AJ. Automated docking of flexible ligands: applications of AutoDock. *J Mol Recognit* 1996;9(1): 1–5.
 129. Morris GM, Goossell DS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 1998;19:1693–1662.
 130. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comp Chem* 2009;30(16): 2785–2791.
 131. Schnecke V, Kuhn LA. Virtual screening with solvation and ligand-induced complementarity. *Persp Drug Discov Des* 2000;20:171–190.
 132. Yamada M, Itai A. Development of an efficient automated docking method. *Chem Pharm Bull* 1993;41:1200–1202.
 133. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol* 1996;261(3): 470–489.
 134. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 2004;47:1739–1749.
 135. Frimurer TM, Peters GH, Iversen LF, Andersen HS, Moller NP, Olsen OH. Ligand-induced conformational changes: improved predictions of ligand binding conformations and affinities. *Biophys J* 2003;84(4): 2273–2281.
 136. Kallblad P, Dean PM. Efficient conformational sampling of local side-chain flexibility. *J Mol Biol* 2003;326(5): 1651–1665.
 137. Leach A. Ligand docking to proteins with discrete side-chain flexibility. *J Mol Biol* 1994; 235:345–356.
 138. Sotriffer CA, Zentgraf M. Addressing selectivity and protein flexibility by simultaneous docking to multiple targets. *Protein Sci* 2004;13:75–75.
 139. Bottegoni G, Kufareva I, Totrov M, Abagyan R. Four-dimensional docking: a fast and accurate account of discrete receptor flexibility in ligand docking. *J Med Chem* 2009;52(2): 397–406.
 140. Huang SY, Zou XQ. Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. *Proteins* 2007;66(2): 399–421.
 141. Claussen H, Buning C, Rarey M, Lengauer T. FlexE: efficient molecular docking considering protein structure variations. *J Mol Biol* 2001;308(2): 377–395.
 142. Zacharias M, Sklenar H. Harmonic modes as variables to approximately account for receptor flexibility in ligand–receptor docking simulations: applications to DNA minor groove ligand complex. *J Comput Chem* 1999;20: 287–300.
 143. May A, Zacharias M. Protein–ligand docking accounting for receptor side chain and global flexibility in normal modes: evaluation on kinase inhibitor cross docking. *J Med Chem* 2008;51(12): 3499–3506.
 144. Kazemi S, Kruger DM, Sirockin F, Gohlke H. Elastic potential grids: accurate and efficient representation of intermolecular interactions for fully flexible docking. *ChemMedChem* 2009;4(8): 1264–1268.
 145. Zhao Y, Stoffler D, Sanner M. Hierarchical and multi-resolution representation of protein flexibility. *Bioinformatics* 2006;22(22): 2768–2774.
 146. Zhao Y, Sanner MF. FLIPDock: docking flexible ligands into flexible receptors. *Proteins* 2007;68(3): 726–737.
 147. Park H, Yeom MS, Lee S. Loop flexibility and solvent dynamics as determinants for the selective inhibition of cyclin-dependent kinase 4: comparative molecular dynamics simulation studies of CDK2 and CDK4. *Chembiochem* 2004;5(12): 1662–1672.
 148. Cavalli A, Bottegoni G, Raco C, De Vivo M, Recanatini M. A computational study of the binding of propidium to the peripheral anionic site of human acetylcholinesterase. *J Med Chem* 2004;47(16): 3991–3999.
 149. Zoete V, Meuwly M, Karplus M. Investigation of glucose binding sites on insulin. *Proteins* 2004;55(3): 568–581.
 150. Nabuurs SB, Wagener M, De Vlieg J. A flexible approach to induced fit docking. *J Med Chem* 2007;50(26): 6507–6518.
 151. Meiler J, Baker D. ROSETTALIGAND: protein-small molecule docking with full side-chain flexibility. *Proteins* 2006;65(3): 538–548.
 152. Davis IW, Baker D. ROSETTALIGAND docking with full ligand and receptor flexibility. *J Mol Biol* 2009;385(2): 381–392.

153. Jain AN. Effects of protein conformation in docking: improved pose prediction through protein pocket adaptation. *J Comput Aided Mol Des* 2009;23(6): 355–374.
154. Koska J, Spassov VZ, Maynard AJ, Yan L, Austin N, Flook PK, Venkatachalam CM. Fully automated molecular mechanics based induced fit protein–ligand docking method. *J Chem Inf Model* 2008;48(10): 1965–1973.
155. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem* 2004;25(9): 1157–1174.
156. Lazaridis T, Masunov A, Gandolfo F. Contributions to the binding free energy of ligands to avidin and streptavidin. *Proteins* 2002;47(2): 194–208.
157. Wei BQ, Weaver LH, Ferrari AM, Matthews BW, Shoichet BK. Testing a flexible-receptor docking algorithm in a model binding site. *J Mol Biol* 2004;337(5): 1161–1182.
158. Englebienne P, Moitessier N. Docking ligands into flexible and solvated macromolecules. 4. Are popular scoring functions accurate for this class of proteins? *J Chem Inf Model* 2009;49(6): 1568–1580.
159. Schymkowitz JW, Rousseau F, Wilkinson HR, Friedler A, Itzhaki LS. Observation of signal transduction in three-dimensional domain swapping. *Nat Struct Biol* 2001;8(10): 888–892.
160. Freire E. Can allosteric regulation be predicted from structure? *Proc Natl Acad Sci USA* 2000;97(22): 11680–11682.
161. Pan H, Lee JC, Hilser VJ. Binding sites in *Escherichia coli* dihydrofolate reductase communicate by modulating the conformational ensemble. *Proc Natl Acad Sci USA* 2000;97(22): 12020–12025.
162. Yu L, Zhu CX, Tse-Dinh YC, Fesik SW. Backbone dynamics of the C-terminal domain of *Escherichia coli* topoisomerase I in the absence and presence of single-stranded DNA. *Biochemistry* 1996;35:9661–9666.
163. Olejniczak ET, Zhou MM, Fesik SW. Changes in the NMR-derived motional parameters of the insulin receptor substrate 1 phosphotyrosine binding domain upon binding to an interleukin 4 receptor phosphopeptide. *Biochemistry* 1997;36(14): 4118–4124.
164. Zidek L, Novotny MV, Stone MJ. Increased protein backbone conformational entropy upon hydrophobic ligand binding. *Nat Struct Biol* 1999;6:1118–1121.
165. Loh AP, Pawley N, Nicholson LK, Oswald RE. An increase in side chain entropy facilitates effector binding: NMR characterization of the side chain methyl group dynamics in Cdc42Hs. *Biochemistry* 2001;40(15): 4590–600.
166. Forman-Kay JD. The ‘dynamics’ in the thermodynamics of binding. *Nat Struct Biol* 1999;6:1086–1087.
167. Kay LE, Muhandiram DR, Wolf G, Shoelson SE, Forman-Kay JD. Correlation between binding and dynamics at SH2 domain interfaces. *Nat Struct Biol* 1998;5:156.
168. Kay LE, Muhandiram DR, Farrow NA, Aubin Y, Forman-Kay JD. Correlation between dynamics and high affinity binding in an SH2 domain interaction. *Biochemistry* 1996;35: 361–368.
169. Steinberg IZ, Scheraga HA. Entropy changes accompanying association reactions of proteins. *J Biol Chem* 1963;238:172–181.
170. Stivers JT, Abeygunawardana C, Mildvan AS, Whitman CP. ¹⁵N NMR relaxation studies of free and inhibitor-bound 4-oxalocrotonate tautomerase: backbone dynamics and entropy changes of an enzyme upon inhibitor binding. *Biochemistry* 1996;35:16036–16047.
171. Hodsdon ME, Cistola DP. Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by ¹⁵N NMR relaxation and ¹H exchange. *Biochemistry* 1997;36(8): 2278–2290.
172. Lee AL, Kinnear SA, Wand AJ. Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex. *Nat Struct Biol* 2000;7(1): 72–77.
173. Canino LS, Shen T, McCammon JA. Changes in flexibility upon binding: application of the self-consistent pair contact probability method to protein–protein interactions. *J Chem Phys* 2002;117:9927–9933.
174. Gohlke H, Case DA. Converging free energy estimates: MM-PB(GB)SA studies on the protein–protein complex Ras-Raf. *J Comput Chem* 2004;25:238–250.
175. van Gunsteren WF, Bakowies D, Baron R, Chandrasekhar I, Christen M, Daura X, Gee P, Geerke DP, Glattli A, Hunenberger PH, Kastenholz MA, Oostenbrink C, Schenk M, Trzesniak D, van der Vegt NF, Yu HB. Biomolecular modeling: goals, problems, perspectives. *Angew Chem Int Ed Engl* 2006;45(25): 4064–4092.
176. Meirovitch H. Recent developments in methodologies for calculating the entropy and free energy of biological systems by computer si-

- mulation. *Curr Opin Struct Biol* 2007;17(2): 181–186.
177. Case DA. Normal mode analysis of protein dynamics. *Curr Opin Struct Biol* 1994;4: 285–290.
178. Pickett SD, Sternberg MJ. Empirical scale of side-chain conformational entropy in protein folding. *J Mol Biol* 1993;231(3): 825–839.
179. Go N, Scheraga HA. Analysis of contribution of internal vibrations to statistical weights of equilibrium conformations of macromolecules. *J Chem Phys* 1969;51(11): 4751.
180. Go N, Scheraga HA. Use of classical statistical-mechanics in treatment of polymer-chain conformation. *Macromolecules* 1976;9(4): 535–542.
181. Levy RM, Karplus M, Kushick J, Perahia D. Evaluation of the configurational entropy for proteins: application to molecular dynamics simulations of an α -helix. *Macromolecules* 1984;1984:1370–1374.
182. Schlitter J. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem Phys Lett* 1993;215: 617–621.
183. Karplus M, Kushick JN. Method for estimating the configurational entropy of macromolecules. *Macromolecules* 1981;14(2): 325–332.
184. Head MS, Given JA, Gilson MK. “Mining minima”: direct computation of conformational free energy. *J Phys Chem A* 1997;101(8): 1609–1618.
185. Nam KY, Chang BH, Han CK, Ahn SG, No KT. Investigation of the protonated state of HIV-1 protease active site. *Bull Korean Chem Soc* 2003;24(6): 817–823.
186. Killian BJ, Kravitz JY, Gilson MK. Extraction of configurational entropy from molecular simulations via an expansion approximation. *J Chem Phys* 2007;127 (2): 024107-1–024107-16.
187. Kirkwood JG. Statistical mechanics of fluid mixtures. *J Chem Phys* 1935;3:300–313.
188. Beveridge DL, DiCapua FM. Free energy via molecular simulation: applications to chemical and biomolecular systems. *Annu Rev Biophys Chem* 1989;18:431–492.
189. Peter C, Oostenbrink C, van Dorp A, van Gunsteren WF. Estimating entropies from molecular dynamics simulations. *J Chem Phys* 2004;120(6): 2652–2661.
190. Böhm HJ. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J Comput Aided Mol Des* 1998;12(4): 309–323.
191. Salaniwal S, Manas ES, Alvarez JC, Unwalla RJ. Critical evaluation of methods to incorporate entropy loss upon binding in high-throughput docking. *Proteins* 2007;66(2): 422–435.
192. Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP. Empirical scoring functions. I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J Comput Aided Mol Des* 1997;11(5): 425–445.
193. Mammen M, Shakhnovich EI, Deutch JM, Whitesides GM. Estimating the entropic cost of self-assembly of multiparticle hydrogen-bonded aggregates based on the cyanuric acid center dot melamine lattice. *J Org Chem* 1998;63(12): 3821–3830.
194. Ruvinsky AM. Calculations of protein–ligand binding entropy of relative and overall molecular motions. *J Comput Aided Mol Des* 2007;21(7): 361–370.
195. Salaniwal S, Manas ES, Alvarez JC, Unwalla RJ. Critical evaluation of methods to incorporate entropy loss upon binding in high-throughput docking. *Proteins* 2007;66(2): 422–435.

