Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 22 (2004) 335–348
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Computational analysis of ligand binding dynamics at the intermolecular hot spots with the aid of simulated tempering and binding free energy calculations

Gennady M. Verkhivker*

Pfizer Global Research and Development, La Jolla Laboratories, 10777 Science Center Drive, San Diego, CA 92121-1111, USA

Abstract

Equilibrium binding dynamics is studied for a panel of benzimidazole-containing compounds at the remodeled interface between human growth hormone (hGH) and the extracellular domain of its receptor (hGHbp), engineered by mutating to glycine hot spot residues T175 from the hormone and W104 from the receptor. Binding energetics is predicted in a good agreement with the experimental data for a panel of these small molecules that complement the engineered defect and restore the binding affinity of the wild-type hGH–hGHbp complex. The results of simulated tempering ligand dynamics at the protein–protein interface reveals a diversity of ligand binding modes that is consistent with the structural orientation of the benzimidazole ring which closely mimics the position of the mutated W104 hot spot residue in the wild-type hGH–hGHbp complex. This structural positioning of the benzimidazole core motif is shown to be a critical feature of the low-energy ligand conformations binding in the engineered cavity. The binding free energy analysis provides a plausible rationale behind the experimental dissociation constants and suggests a key role of ligand–protein van der Waals interactions in restoring binding affinity.

Keywords: Computational analysis; Human growth hormone; Ligand binding dynamics

1. Introduction

Understanding mechanisms and fundamental biophysical principles of molecular ecognition continues to present a fundamental experimental and theoretical challenge [1–6]. Alanine scanning mutagenesis of protein–protein interfacial residues, combined with structural and thermodynamic studies, have enabled discovery of energetically important hot spot regions at the intermolecular interfaces that are critical in determining binding affinity, i.e. alanine mutation of a hot spot residue in the binding site results in a pronounced drop in binding affinity of the complex [7]. A comprehensive analysis of protein–protein interfaces [8,9] and a survey of hot spots compiled from various protein binding sites [10] have demonstrated a diversity of interaction patterns and a lack of general rules for hydrophobicity, polarity, or shape, that can be used to unambiguously predict hot spots at the intermolecular interfaces. A recent analysis of conserved residues in 11 clustered interface families comprising a total of 97 crystal structures has shown that the composition of hot spots is typically enriched by certain residues, such as Trp, Tyr, Arg, His, Gln, Asn, and Pro and can be surrounded by a shell of less important residues [10–12]. The discovery of hot spots appears to be broadly relevant in a variety of protein–protein recognition events where, despite a typically large size of intermolecular interface, binding affinity and specificity may be determined by a functional epitope consisting of only a small fraction of the interfacial residues [7,10]. An overlap between flexible consensus binding sites and energetically critical hot spots, discovered through a combination of structural and mutagenesis studies, have triggered experimental and computational studies aiming to understand the nature of structural flexibility and functional diversity in molecular recognition [7–13].

Molecular recognition between proteins and flexible target molecules, including other proteins and small molecules is often accompanied by a considerable flexibility of the protein binding sites and structural rearrangements upon binding between the associated partners [14–18]. Protein dynamics at the intermolecular interfaces can have a profound effect in determining binding thermodynamics, kinetics and consequently in modulating binding affinity and specificity of molecular recognition [19–22]. Protein binding interfaces can be not only structurally flexible, but also functionally adaptive, with a diverse repertoire of protein systems capable of binding with high affinity to ligands, different from their natural binding partners in composition,

* Tel.: +1-858-622-3008; fax: +1-858-678-8244.
E-mail address: gennady.verkhivker@pfizer.com (G.M. Verkhivker).

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2003.12.001

size, and shape [23,24]. Evolution can often find convergent solutions to stable intermolecular interfaces by using structural flexibility and plasticity of the hot spot residues in the binding sites to allow accommodation to different binding partners [12,23,24]. This notion was manifested in discovery of a combinatorially selected high-affinity synthetic peptide that can mimic specific interactions of much larger natural proteins with the hot spot residues of the constant fragment of human immunoglobulin G protein, while using the interacting groups from a different structural scaffold [23]. Computational alanine scanning and binding free energy calculations conducted for this system have indicated that while a specific set of the conserved hot spot protein residues provides thermodynamic stability of the native structure, different hot spot residues contribute decisively to the binding affinity of the peptide–protein complex [25].

Structural and functional studies of human growth hormone–receptor binding between human growth hormone (hGH) and hGHbp have provided a comprehensive view and atomic details of the binding thermodynamics that is determined by a small fraction of residues of a large intermolecular interface [26–28]. In a series of pioneering biochemical studies, homolog-scanning [29] and alanine scanning [30] mutagenesis strategies had allowed an initial mapping of the binding determinants in the hGH and hGHbp molecules even before the high-resolution structures of the complex and individual components became available. These studies determined a relatively small set of residues that resulted in four-time lower binding affinity, when mutated to alanine, with many of these residues being altered in the corresponding nonbinding homologs. Using a variety of biophysical methods, including titration calorimetry, the stoichiometry of the hGH–hGHbp complex was determined, revealing two hGHbp receptor molecules bound to a single hGH hormone molecule [31]. The crystals composition of the hGH–hGHbp complex confirmed that biologically significant dimerization of the growth hormone receptor was indeed mediated through a single hormone molecule [32]. The solution of the X-ray structure of the 1:2 hGH–hGHbp complex [33] has revealed not only the stoichiometry of the complex and the structures of the complex components, but has also identified the critical contact residues of the hormone–receptor interface, thereby permitting a direct structural interpretation of the mutagenesis and biophysical studies. The structure showed that two receptor molecules employed the same binding determinants to interact with two considerably dissimilar sites on the opposite sites of the hormone, that required the receptor binding surfaces to undergo local conformational changes [33,34]. Furthermore, essentially the same set of binding determinants was used by hGH to bind to both hGHbp and prolactin receptor, despite only 28% homology between the complementary binding sites of these receptors [33–35].

A detailed study of hormone binding determinants in the hGHbp receptor was conducted by systematically replacing side chains with alanine and measuring the binding affinities of the alanine hGHbp mutants with the hormone [36]. This mutational strategy probed the importance of 49 charged surface residues, 9 aromatic residues, and 26 neighboring residues. A similar analysis was performed on the hormone side, in which the role of 31 buried hGH residues in binding with the hGHbp receptor was determined by converting each of these residues to alanine and measuring their effect on modulating binding affinity and kinetics of the reaction in the site 1 of hGHbp [37]. Alanine substitutions of only eight residues among the studied set (K41, L45, P61, R64, K172, T175, F176, and R178), forming two small patches on the protein interface, accounted for approximately 85% of the binding energy. Moreover, a significant increase in the off-rates was detected for six mutants P61A, R64A, K172A, T175A, F176A, and R178A, while only subtle effects on the on-rates were observed. These data have indicated that the hormone and the receptor associate by rapid diffusion and the primary role of the intermolecular hot spot residues is to produce a tightly bound complex after the hormone has reached the binding site [37]. In a subsequent widely recognized work [38], the crystal structure of the hGH–hGHbp complex was used to perform alanine scanning of 33 buried side-chains on the receptor. Despite a large hGH:hGHbp intermolecular interface, that covers 1350 $Å^2$ and involves 33 receptor side-chain residues, the discovered hot spot of binding energy comprises of only nine of these residues (R43, E44, I103, W104, I105, P106, I164, D165, and W169), providing virtually all binding affinity. The critical interactions in the hot spot are primarily hydrophobic and the functional epitope is dominated by contributions from two hydrophobic residues W104 and W169 contributing approximately 4.5 kcal/mol, i.e. more than three-quarters of the total binding free energy [38]. Importantly, empirical parameters such as the extent of side-chain burial surface area or van der Waals interactions served as pure predictors of the relative effects on binding affinity, when the alanine scanning was performed. However, these correlations were considerably better when only the burial of well-packed and dominating W104 and W169 hot spot residues was considered.

Mutational studies have shown that more than two-thirds of the contact hGH–hGHbp interfacial residues have practically no effect on binding affinity when converted to alanine [36–38]. Alanine-shaving, which is the process of making multiple simultaneous alanine mutations, has been used to experimentally test the role of affinity-inert residues [39,40]. Strikingly, a less than 10-fold drop in affinity was detected when 16 energetically unimportant residues were simultaneously mutated to alanine [39]. Multiple alanine scanning of peripheral to the functional epitope contact hGH residues F25, Y42, and Q46 have shown only local structural changes near sites of mutation and a little effect on binding kinetics and thermodynamics, resulting from large but compensating changes in the enthalpy and entropy of binding [40] Hence, large cumulative reductions on protein–protein contacts can be tolerated with minimal changes in binding affinity for peripheral, functionally-inert residues in the hGH–hGHbp

interface. These studies have underscored the inherent difficulties of inferring binding free energy changes from local contacts at the flexible regions in the interface and have suggested the possibility for more dynamic regions to be less critical for binding affinity [40]. The 2.6 Å resolution crystal structure of the G120R designed mutant of hGH that binds only a single molecule of hGHbp has guided further a detailed survey of the structural and functional basis for hormone–receptor recognition [41]. This analysis was primarily conducted to distinguish direct binding energy contributions of the hot spot receptor residues from indirect effects that may be mediated through peripheral affinity-inert contacts. The revised alanine scan of the hGHbp residues based on the 1:1 complex showed that only 11 residues affected significantly binding affinity with hGH (R43, E44, I193, W104, I105, P106, D126, E127, D164, I165, and W169) [41]. The true functional epitope is assembled cooperatively and even more localized with only six hGHbp side-chains reflecting major direct contributions to binding affinity [41].

The computed changes in solvent-accessible surface area using the reported alanine shaves [41] resulted in the largest change of only 4.2 $Å^2$ for the shaved mutant of the hGH receptor [10]. Interestingly, truncation to alanine of all 46 affinity-inert side-chains led to only 56% (60.6 $Å^2$) increase in the calculated solvent-accessible surface for the hot spot [10]. The failure of alanine shaving to significantly increase the solvent accessibility of the hot spot was suggested as a rationale for similar binding affinities of alanine shaved mutants compared to those of wild-type proteins [10]. Computational alanine scanning of the hGH–hGHbp complex has shown an excellent agreement between the relative binding free energy changes upon mutation of W104 and W169 residues and experimental data, indicating that it may be energetically costly to significantly change nearby residues to compensate for the loss of critical interactions [42]. However, for charged and polar residues R43, D164, and R217 in hGHbp, which are less important in binding affinity, the agreement with experiment was only qualitative, suggesting that appreciable local side-chain rearrangements of the surrounding residues may take place to restore these defects [42].

The structural and functional data of the hGH–hGHbp binding have suggested that the discovered hot spot could present a template for rational remodeling of the interfaces and design of alternative scaffolds and small molecule mimics [38,41]. Remodeling of the interface between the hGH and the hGHbp receptor was studied by mutating to alanine a critical W104 residue in the receptor and selecting a pentamutant (K168R, D171T, K172Y, E174A, and F176Y) of hGH by phage display that fills the created cavity and largely restores binding affinity [43]. A 2.1 Å resolution X-ray structure of the W104A mutant of the hGHbp receptor bound to the hGH mutant containing five mutations (K168R, D171T, K172Y, E174A, and F176Y) showed that the receptor cavity was filled by selected hydrophobic mutations of hGH and was accompanied by considerable structural rearrangements occurred in the interface at sites distant from the mutations [43]. This pioneering study has shown how structural plasticity at the remodeled hGH–hGHbp interface can rescue large functional changes and largely restore binding affinity of the wild-type hGH–hGHbp complex by a limited number of mutations, thereby providing a mechanism for mutations to be accommodated during coevolution of high affinity binding partners [43].

The hypothesis that functional versatility of hormone–receptor binding may be an evolutionary consequence of selecting a set of structurally diverse binding partners was probed by employing phage display mutagenesis and designing novel hGH variant molecules that are different from the wild-type hormone [43,44]. The structure of the complex between the phage display optimized high-affinity to site 1 hGH bound with hGHbp molecules was determined at 2.6 Å resolution and revealed that 15 mutations introduced in the designed variant of hGH result in significant structural changes in the hormone–receptor interface [45]. A tighter binding in site 1 was accompanied by 15% smaller and 20% more hydrophobic interfacial surface area compared to the wild-type complex. Strikingly, structural plasticity of the hormone–receptor recognition leads to a novel binding solution at the site 2, with a structurally entirely distinct conformation of the receptor molecule obtained in the absence of any selection presure [45]. These results have indicated that structural elements of the hGH–hGHbp system may have an inherent structural flexibility and functional plasticity that enables them to evolve and bind to a variety of binding surfaces.

An elegant protein engineering solution to the specificity problem often implies creating an artificial cavity that can be introduced into the ligand-binding site of the protein by truncating side chains using site-directed mutagenesis. A complementary modification can then be designed on the ligand, generating a new ‘bumped’ compound that no longer binds the wild-type protein but interacts specifically with the mutant. Rational redesign of protein–protein and ligand–protein interfaces and generating modified molecules that interact only with specifically mutated proteins has become a powerful approach for manipulating remodeled protein interfaces and generating ligands with new, orthogonal specificities [46–50]. A recently reported approach for generating synthetic molecules directly modulating specific interactions at the hGH–hGHbp interface involved a cavity formation, introduced at the interface by mutating to glycine the hot spot residues T175 from the hormone and W104 from the receptor [51]. While binding affinity of the W104A mutant of hGHbp is reduced by more than 2500-fold relative to the wild-type complex, T175A mutation results in only 25-fold reduction in binding free energy. The double mutant W104G/T175G reduces binding affinity of the wild-type complex by a factor of 10^6 and results in a formation of a sufficiently large cavity in the interface that can accommodate small molecules complementing this defect and restoring the affinity of the complex. A library of 200 indole analogs and

derivatives of structurally related five- and six-membered fused aromatic heterocycles was then screened for ligands that complement this defect and restore binding affinity of the wild-type complex. A significant recovery of phage was detected only in the presence of ligands containing the benzimidazole core motif; the individually assayed representative molecules that largely restored binding affinity were reported with their respective dissociation constants [51].

A hierarchical computational approach has been recently used to identify the engineered binding site cavity at the remodeled intermolecular interface between the mutants of human growth hormone (hGH) and the extracellular domain of its receptor (hGHbp) [52]. Multiple docking simulations were conducted with the remodeled hGH–hGHbp complex for a panel of potent benzimidazole-containing inhibitors, that can restore the binding affinity of the wild-type complex, and for a set of known non-active small molecules that contain different heterocyclic motifs. Structural clustering of ligand bound conformations and binding free energy calculations, using the AMBER force field and a continuum solvation model, have allowed to rapidly locate and screen numerous ligand binding modes on the protein surface and detect the binding site hot spot at the intermolecular interface. We have found that despite presence of numerous pockets on the protein surface for the mutant hGH–hGHbp complex, the binding site cavity can be distinguished as the energetically most favorable hot spot for the benzimidazole-containing inhibitors, while for a set of non-active molecules the lowest energy ligand conformations do not necessarily bind in the engineered cavity [52].

In this work, we study dynamics and energetics of ligand binding for five compounds with the measured dissociation constants (Fig. 1) which restore the binding affinity of the wild-type complex. We set out to study microscopic details of ligand binding for these small molecules in the engineered cavity created at the hGH–hGHbp remodeled interface to understand the atomic details of ligand binding dynamics and compare the computed binding free energies with the experimental data. Despite the absence of the crystal structures of the bound inhibitors with the remodeled hGH–hGHbp complex, the precise presentation of the ligand elements that could mimic the indole rings of W104 and W169 is known to be the essential prerequisite for the

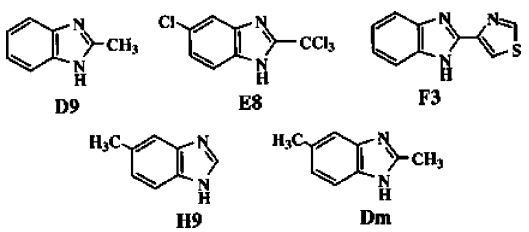

Fig. 1. Chemical structures for a panel of studied inhibitors serving as molecular switches and restoring binding affinity of the hGH–hGHbp remodeled interface. The notations of the compounds from the original work [51] are kept and correspond to the group nomenclature in the initial library screening (see [51] for more details).

molecules to restore the binding affinity of the wild-type hGH–hGHbp complex [51]. As a result, the inhibitors must bind in the engineered cavity of the redesigned interface to fulfill the observed biological function. We show that the low-energy ligand binding modes, generated from equilibrium simulations in the binding site cavity, are compatible with the structural orientation of the benzimidazole core motif that mimics the location of the W104 indole ring in the crystal structure of the wild-type complex. Binding free energy calculations are then performed for all conformations in the low-energy structural clusters featuring this structural positioning in the core motif of the inhibitors. We show that this protocol allows an accurate analysis of binding thermodynamics for studied inhibitors and can provide a plausible rationale of the experimental data by clarifying the role of key energetic contributions in restoring binding affinity of the hGH–hGHbp complex.

2. Materials and methods

2.1. Molecular recognition energy model

Simplified energy models can faithfully describe a multitude of the available binding modes for the complex and reproduce relative thermodynamic stability of the native complex with respect to intermediate complexes and alternative binding modes [53–56]. A multi-stage strategy [57,58] with a hierarchy of different energy functions is pursued in this work to achieve a synergy of robust conformational sampling and accurate estimation of binding energetics. The simplified energy function [58,59] is used in conjunction with evolutionary search and Monte Carlo simulations [59–61] to sample the conformational space and and adequately describe the multitude of the low-energy states available to the system.

The knowledge-based simplified energetic model includes intramolecular energy terms for the ligand, given by torsional and nonbonded contributions of the DREIDING force field [62], and intermolecular ligand–protein steric and hydrogen bond interaction terms calculated from a simplified piecewise linear (PL) potential summed over all protein and ligand heavy atoms (Fig. 2a). The parameters of the pairwise potential depend on the following different atom types: hydrogen-bond donor, hydrogen-bond acceptor, both donor and acceptor, carbon-sized nonpolar, sulfur-sized nonpolar, flourine-sized nonpolar, and large nonpolar. The atomic radius is 1.4 Å for fluorine is 1.4 and 1.8 Å for carbon, oxygen, and nitrogen atoms. The atomic radius of 2.2 Å is assigned to sulfur and phosphorus, chlorine, and bromine atoms, modeled as large nonpolar atom type. Electronegative atoms with an attached hydrogen are defined to be donors, while oxygen and nitrogen atoms with no bound hydrogens are defined to be acceptors.

Sulfur is modeled as being capable of making weak hydrogen bonds which allows for sulfur-donor closer contacts

that are seen in some of the crystal structures. Crystallographic water molecules and hydroxyl groups are defined in this model to be both donor and acceptor, and carbon atoms are defined to be nonpolar. An empirical desolvation correction is applied to the attractive portion of the interactions between nonpolar and polar atoms. This correction is defined as the the ratio between the attractive well depth for nonpolar–polar contacts and the one for nonpolar–nonpolar contacts, and can range between 0 and 1. The parameter is set to 1.0 in this work, thereby imposing a desolvation penalty by disadvantaging the burial of polar groups with the nonpolar atoms.

A hydrogen bond interaction term is assigned to interactions between donor and acceptors, a repulsive interaction contribution is computed for donor–donor and acceptor–acceptor contacts, and a steric intermolecular term is assigned for other contacts. The steric and hydrogen bond-like potentials have the same functional form, with an additional three-body contribution to the hydrogen bond term and the repulsive term for donor–donor and acceptor–acceptor contacts. Both the hydrogen bond interaction energy and the repulsive interaction contribution between donor–donor and acceptor–acceptor close contacts are modulated by an approximate angular dependence (Fig. 2b). These terms are multiplied by the hydrogen bond strength term, which is a function of the angle θ determined by the relative orientation of the protein and ligand atoms (Fig. 2b). The scaling for the repulsive interactions is equivalent to the dependence used for the hydrogen bond interaction term, but in this cases it implies a maximum penalty when the angle θ is 180°, fading to zero at 90° and below. θ is defined to be the angle between two vectors, one of which points from the protein atom to the ligand atom. For protein atoms with a single heavy atom neighbor, the second vector connects the protein atom with its heavy atom neighbor, while for protein atoms with two heavy atom neighbors, it is the bisector of the vectors connecting the protein atom with its two neighbors.

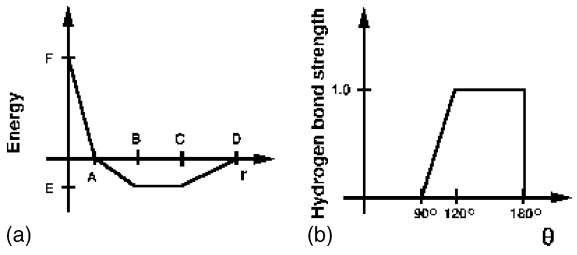

Fig. 2. (a) The functional form of the ligand–protein interaction energy. For steric interactions, $A = 0.93B$, $C = 1.25B$, $D = 1.5B$, $E = -0.4$, $F = 15.0$, and $B = r_{\text{ligand}} + r_{\text{protein}}$ is the sum of the atomic radii for the ligand and protein atoms. For hydrogen bond interactions, $A = 2.3$, $B = 2.6$, $C = 3.1$, $D = 3.4$, $E = -4.0$, $F = 15.0$. For sulfur hydrogen bond interactions, $A = 2.7$, $B = 3.0$, $C = 3.5$, $D = 3.8$, $E = -2.0$, $F = 15.0$. For chelating interactions with the metals $A = 1.5$, $B = 1.7$, $C = 2.5$, $D = 3.0$, $E = -10.0$, $F = 15.0$. For repulsive interactions, $A = 3.2$, $E = 0.1$, $F = 15.0$. The repulsive potential is then linearly scaled from $E = 0.1$ to 0 between 3.2 and 5.0 Å. The units of A, B, C, and D are Å, for E and F the units are kcal/mol. (b) The hydrogen bond interaction energy and the repulsive term are multiplied by the hydrogen bond strength term, which is a function of the angle θ determined by the relative orientation of the protein and ligand atoms.

For molecular docking simulations, the energy landscape must be relatively smooth for robust structure prediction of ligand–protein complexes and softening the potentials is a way to smooth the force field and enhance sampling of the conformational space while retaining adequate description of the binding energy landscape [53–58]. The PL energy function has no singularities at interatomic distances and can effectively explore large conformational spaces. While the PL energy function is proven to be more adequate for sampling non-polar and hydrogen bonds patterns, this simplified energy model does not include a direct electrostatic component and therefore may be less accurate in detecting the exact energetics of the binding modes, especially when extensive networks of electrostatic interactions are present in the crystal structure. This function is less accurate in detecting the exact location and energetics of the native state because of the inaccuracy in quantifying the magnitude of protein–protein and ligand–protein interactions.

The conformational states generated with the PL energy function and docking and equilibrium simulations are evaluated with a more detailed binding free energy model, which includes the molecular mechanics AMBER force field [63] and the solvation energy term based on continuum generalized Born and solvent-accessible surface area (GB/SA) solvation model [64–70]. This procedure is conceptually similar to the MM/PBSA (molecular mechanics Poisson–Boltzmann (PB) surface area) approach [71–78], and replaces time-consuming PB continuum calculations with less demanding GB solvation calculations, correlating well with the PB results [79]. The MM/PBSA approach uses molecular dynamics simulations of the system to generate a thermally averaged ensemble of conformations. Based on this set of structures, the total free energy of the system is evaluated as a sum of the polar solvation energy, which is computed using a finite-difference Poisson–Boltzmann approach, the nonpolar solvation term derived from the solvent-accessible surface area (SA), and solute entropy contribution. The molecular mechanical energy of the molecule includes the electrostatic, van der Waals contributions and internal strain energy. The ensemble of structures for the uncomplexed protein and ligand are generated in the MM/PBSA approach by using the molecular dynamics trajectory of the complex, and simply separating the protein and ligand coordinates, followed by an additional minimization of the unbound protein and unbound ligand. This methodology has been successfully applied in various applications of protein–protein and ligand–protein recognition [72–78] and recently in the computational alanine scanning of the hGH–hGHbp complex [42].

Binding free energy contributions are computed by averaging over conformations from families that are characterized by an overlap of the benzimidazole core motif with the

indole ring of the key W104 residue of the receptor. The ensemble of structures for the uncomplexed protein and ligand are generated by using the equilibrium trajectory of the complex, and separating the protein and ligand coordinates, followed by an additional minimization of the unbound protein and unbound ligand.

The average total free energy of the molecule G is evaluated as follows:

$$G_{\text{molecule}} = G_{\text{solvation}} + E_{\text{MM}} - TS_{\text{solute}} \quad (1)$$

$$G_{\text{solvation}} = G_{\text{cavity}} + G_{\text{vdw}} + G_{\text{pol}} \quad (2)$$

In the GB/SA model, the G_{cavity} and G_{vdw} contributions are combined together via evaluating solvent-accessible surface areas:

$$G_{\text{SA}} = G_{\text{cavity}} + G_{\text{vdw}} = \sum_i \sigma_i \text{SA}_i \quad (3)$$

where G_{SA} is the nonpolar solvation term derived from the solvent-accessible surface area.

$$G_{\text{pol}} = -166.0\left(1 - \frac{1}{\epsilon}\right)\sum_i \sum_j \frac{q_i q_j}{(r_{ij}^2 + \alpha_{ij}^2 \exp(-D_{ij}))^{0.5}} \quad (4)$$

where G_{pol} is the polar solvation energy which is computed using the GB/SA solvation model. S_{solute} is the vibrational entropy of the molecule. E_{MM} is the molecular mechanical energy of the molecule summing up the electrostatic E_{es} interactions, van der Waals contributions E_{vdw}, and the internal strain energy E_{int}:

$$E_{\text{MM}} = E_{\text{es}} + E_{\text{vdw}} + E_{\text{int}} \quad (5)$$

Using these equations, the binding free energy of the ligand–protein complex is computed as follows:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{protein}} - G_{\text{ligand}} \quad (6)$$

From this equation, one can determine contributions of the ligand–protein interaction energy ΔG_{MM}, strain energy ΔG_{strain}, and solvation energy $\Delta G_{\text{GB/SA}}$ to the total binding free energy.

$$\Delta G_{\text{bind}} = \Delta G_{\text{interaction}} + \Delta G_{\text{strain}} + \Delta G_{\text{solvation}} \quad (7)$$

$$\begin{aligned}\Delta G_{\text{interaction}} &= E_{\text{MM}}^{\text{complex}} - E_{\text{MM}}^{\text{bound protein}} - E_{\text{MM}}^{\text{bound ligand}} \\ &= E_{\text{es}}^{\text{complex}} - E_{\text{es}}^{\text{bound protein}} - E_{\text{es}}^{\text{bound ligand}} \\ &\quad + E_{\text{vdw}}^{\text{complex}} - E_{\text{vdw}}^{\text{bound protein}} - E_{\text{vdw}}^{\text{bound ligand}} \\ &\quad + E_{\text{int}}^{\text{complex}} - E_{\text{int}}^{\text{bound protein}} - E_{\text{int}}^{\text{bound ligand}}\end{aligned} \quad (8)$$

$$\begin{aligned}\Delta G_{\text{strain}} &= (E_{\text{MM}}^{\text{bound protein}} - E_{\text{MM}}^{\text{free protein}}) \\ &\quad + (E_{\text{MM}}^{\text{bound ligand}} - E_{\text{MM}}^{\text{free ligand}})\end{aligned} \quad (9)$$

$$\Delta G_{\text{solvation}} = G_{\text{solvation}}^{\text{complex}} - G_{\text{solvation}}^{\text{free protein}} - G_{\text{solvation}}^{\text{free ligand}} \quad (10)$$

The crystal structure of the wild-type hGH–hGHbp hormone–receptor complex was used in computational modeling and was subjected to glycine mutation at the W104 position on the hGHbp and T175 position on the hormone. The mutated chimera structure was minimized using 200 iteration of conjugate gradient minimization to relieve any strain contacts. The energy of each ligand–protein complex is subjected to the conjugate gradient minimization as implemented in the version 7.0 of the MacroModel molecular modeling software package [68]. All protein residues within 3 Å radius sphere from the ligand are treated as flexible during minimization. All protein residues within 2 Å radius from the flexible shell form a first shell of restrained atoms with the force constant 100.0 kJ/mol A^2. A second shell of restrained atoms with the force constant 200 kJ/mol A^2 consists of the residues within 2 Å radius from the first shell and finally, the third shell of restrained atoms is generated by the residues which reside within 2 Å from the second shell and they are restrained with the force constant 300 kJ/mol A^2. The remaining protein atoms are treated as frozen atoms and do not move during the minimization procedure. The interactions between frozen atoms and restrained atoms, and frozen atoms and flexible atoms are included in the total energy value. A residue-based cutoff of 8 Å is set for computing non-bonded van der Waals interactions and 20 Å residue-based cutoff is used for computing electrostatic interactions. The structures of the studied ligands were built and minimized with the MNDO atomic charges calculated using MOPAC program. The protein atoms have been assigned the AMBER force field charges. Because of significant variances in computing solute entropy using the MM/GBSA approach, this term was not included in the total binding free energy value.

2.2. Monte Carlo equilibrium simulations of ligand–protein binding

Equilibrium simulations of ligand binding with the remodeled hGH–hGHbp mutant complex are performed in a parallelepiped that encompasses the engineered cavity with a 5.0 Å cushion added to every side of this box to accurately reproduce the equilibrium distribution between ligand binding modes in the binding site cavity. The protein structure of the remodeled hGH–hGHbp mutant complex is held fixed in its minimized conformation, while rigid body degrees of freedom and rotatable angles of the ligands are treated as independent variables. Bonds allowed to rotate include those linking sp^3 hybridized atoms to either sp^3 or sp^2 hybridized atoms and single bonds linking two sp^2 hybridized atoms. We have carried out equilibrium simulations using parallel simulated tempering dynamics [80–87] with 50 replicas of the ligand–protein system attributed, respectively, to 50 different temperature levels that are uniformly distributed in the range between 5300 and 300 K. Independent local Monte Carlo moves are performed independently for each replica at the corresponding temperature level, but after a simula-

tion cycle is completed for all replicas, configuration exchanges for every pair of adjacent replicas are introduced. The mth and nth replicas, described by a common Hamiltonian $H(X)$, are associated with the inverse temperatures β_m and β_n, and the corresponding conformations X_m and X_n. The exchange of conformations between adjacent replicas m and n is accepted or rejected according to Metropolis criterion with the probability $p = \min(1, \exp[-\delta])$, where $\delta = [\beta_n - \beta_m][H(X_m) - H(X_n)]$. Starting with the highest temperature, every pair of adjacent temperature configurations is tested for swapping until the final lowest value of temperature is reached. This process of swapping configurations is repeated 50 times after each simulation cycle for all replicas whereby the exchange of conformations presents an improved global update which increases thermalization of the system and overcomes slow dynamics at low temperatures on rough energy landscapes, thereby permitting regions with a small density of states to be sampled accurately. During simulation, each replica has a non-negligible probability of moving through the entire temperature range and the detailed balance is never violated which guarantee each replica of the system to be equilibrated in the canonical distribution with its own temperature [80–87].

Monte Carlo simulations allow to dynamically optimize the step sizes at each temperature by taking into account the inhomogeneity of the molecular system [88]. The acceptance ratio method is used to update the step sizes every cycle of 1000 sweeps. For all these simulations, we equilibrated the system for 1000 cycles (or one million sweeps), and collected data during 5000 cycles (or five million sweeps) resulting in 5000 samples at each temperature. A sweep is defined as a single trial move for each degree of freedom of the system. A key parameter is the acceptance ratio which is the ratio of accepted conformations to the total number of trial conformations. At a given cycle of the simulation, each degree of freedom can change randomly throughout some prespecified range determined by the acceptance ratio obtained during the previous cycle. This range varies from one degree of freedom to another because of the complex nature of the energy landscape. At the end of each cycle, the maximum step size is updated and used during the next cycle.

Simulations are arranged in cycles, and after a given cycle i, where the average acceptance ratio for each degree of freedom j is $\langle P_j \rangle^i$, the step sizes σ_j^i for each degree of freedom are updated for cycle $i+1$ according to the formula

$$\sigma_j^{i+1} = \sigma_j^i \frac{\ln[a\langle P_{\text{ideal}}\rangle + b]}{\ln[a\langle P_j\rangle^i + b]} \tag{11}$$

where $\langle P_{\text{ideal}}\rangle$ is the desired acceptance ratio, chosen to be 0.5. The parameters a and b are used to ensure that the step sizes remain well-behaved when the acceptance ratio approaches 0 or 1. They are assigned so that the ratio σ^{i+1}/σ^i is scaled up by a constant value s for $\langle P_j\rangle^i = 0$, and down by the same constant for $\langle P_j\rangle^i = 1$. Solving the equations

$$s^{-1} = \frac{\ln[a\langle P_{\text{ideal}}\rangle + b]}{\ln[b]} \tag{12}$$

$$s = \frac{\ln[a\langle P_{\text{ideal}}\rangle + b]}{\ln[a + b]} \tag{13}$$

with $s = 3$ yields $a = 0.673$ and $b = 0.065$.

2.3. Similarity clustering

The 3D-similarity calculations are based on the spatial proximity of atoms in a binding site and the atom type. Four types of atoms are distinguished: hydrogen bond donors, hydrogen bond acceptors, hydrogen bond donors and acceptors and nonpolar atoms. The atom type compatibility $a(i, j)$ is assigned a value between 0.0 and 1.0, with the compatibility between two atoms of the same type defined to be 1.0, that between donor and acceptor atom is 0.0, and other combinations of atoms have compatibilities between 0.0 and 1.0.

The spatial proximity between two atoms i and j is evaluated with a Gaussian function $p(i, j) = 10^{(-r_{i,j}^2/\sigma^2)}$, where r_{ij} is the distance between atoms i and j, and $\sigma = -c^2/\log_{10}(p)$, where c and p denote the cutoff distance and proximity threshold, respectively. Both the cutoff distance and the proximity threshold determine the shape of the Gaussian function to evaluate spatial proximity of two atoms, with $c = 3.0$ Å and $p = 0.000032$.

A descriptor $d(i, j)$ is calculated from the spatial proximity and the atom type compatibility:

$$d(i, j) = p(i, j)a(i, j) \quad \text{if } r(i, j) \le c \tag{14}$$

$$d(i, j) = 0 \quad \text{if } r(i, j) > c \tag{15}$$

An atom descriptor $D_m^n(i)$ for atom i in molecule m is then calculated by summation over all N atoms in molecule n, $D_m^n(i) = \sum_{j=1}^{N} d_m^n(i, j)$. The intermolecular similarity between molecules m and n is given by the Tanimoto coefficient [89–91]:

$$S(m, n) = \frac{\sum_{i=1}^{M} D_m^m(i) D_m^n(i) + \sum_{j=1}^{N} D_n^m(j) D_n^n(j)}{\sum_{i=1}^{M} D_m^m(i)^2 + \sum_{j=1}^{N} D_n^n(j)^2 - \sum_{i=1}^{M} D_m^m(i) D_m^n(i) - \sum_{j=1}^{N} D_n^m(j) D_n^n(j)} \tag{16}$$

Molecules are grouped into clusters by comparing the intermolecular similarity coefficient. The first molecule is assigned to the first cluster. The next molecule is assigned to the cluster in which a cluster member has the highest similarity with the next molecule, if the similarity is above a threshold, chosen to be 0.85. Otherwise, the next molecule is assigned to a new cluster. The first member of the a cluster is called the cluster center. After all molecules are assigned to clusters, the molecules are arranged in new order, starting with the largest cluster and proceeding to the smallest cluster. The reordered set of molecules is subjected to the same clustering procedure. This procedure is iterated until the information entropy converges to a minimum. The clusters

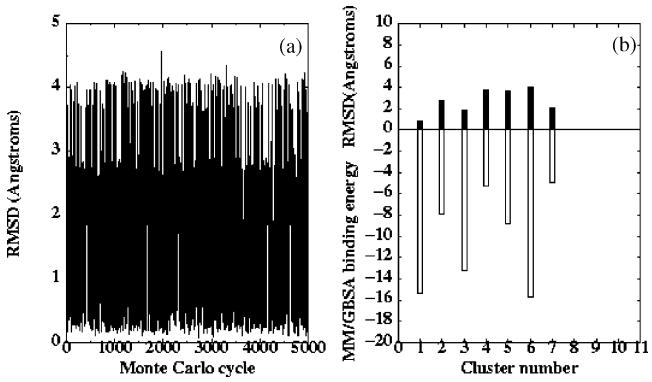

Fig. 3. (a) Time-dependent equilibrium history of D9 binding in the binding site cavity at $T = 300$ K. (b) MM/GBSA binding free energies of conformational clusters generated in equilibrium simulations at $T = 300$ K for D9 and the rmsd differences of the corresponding cluster centers from the reference lowest energy conformation of D9 determined in docking simulations.

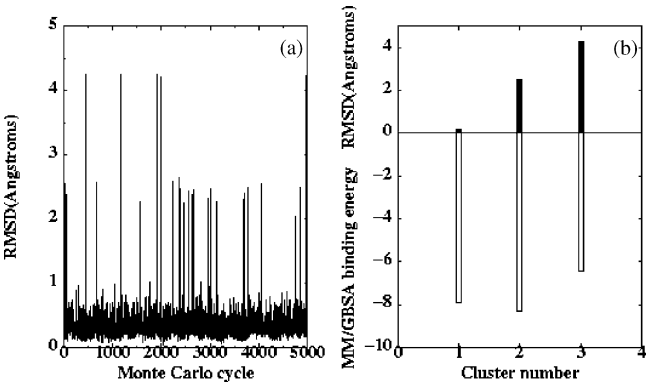

Fig. 4. (a) Time-dependent equilibrium history of Dm binding in the binding site cavity at $T = 300$ K. (b) MM/GBSA binding free energies of conformational clusters generated in equilibrium simulations at $T = 300$ K for Dm and the rmsd differences of the corresponding cluster centers from the reference lowest energy conformation of Dm determined in docking simulations.

with at least 100 members are analyzed. Since conformations which belong to the same cluster are equivalent with 85% structural similarity, different clusters are compared by analyzing cluster centers.

3. Results and discussion

In the absence of the crystal structures for studied inhibitors bound at the remodeled hGH–hGHbp interface, a direct comparison with the experimental data is possible only with the binding affinities of the ligand–protein complexes. Structural and mutagenesis data have indicated that the precise presentation of elements mimicking the indole rings of W104 and W169 is an essential feature for ligands to restore the binding affinity of the wild-type complex, and therefore, the panel of studied inhibitors must bind in the hot spot of the intermolecular interface. To further understand a structural diversity of ligand bound conformations in the binding site cavity that could fulfill an important functional requirement of mimicking critical mutated residues in the hot spot, equilibrium simulations are performed for studied ligands using simulated tempering in conjunction with the simplified PL energy function. During equilibrium simulations in the binding site cavity, we monitor whether putative ligand binding modes result in a unique structural orientation of the benzimidazole core motif that emulates the indole ring of W104 and whether the requirement of structural compatibility with the mutated hot spot residues is a unique structural attribute for the low-energy ligand conformations to bind in the engineered cavity.

Equilibrium simulations indicate a certain diversity of low-energy ligand binding modes in the binding cavity of the hGH–hGHbp interface (Figs. 3–7). For example, for D9 inhibitor, at least three low-energy binding modes are well populated in equilibrium simulations at room temperature (Fig. 3a). These binding modes deviate by rmsd = 0.8,

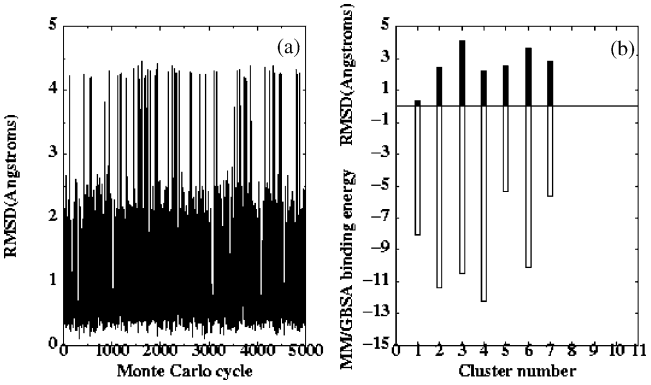

Fig. 5. (a) Time-dependent equilibrium history of H9 binding in the binding site cavity at $T = 300$ K. (b) MM/GBSA binding free energies of conformational clusters generated in equilibrium simulations at $T = 300$ K for H9 and the rmsd differences of the corresponding cluster centers from the reference lowest energy conformation of H9 determined in docking simulations.

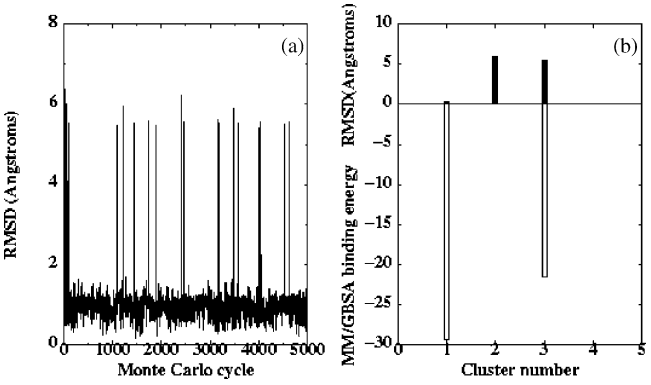

Fig. 6. (a) Time-dependent equilibrium history of E8 binding in the binding site cavity at $T = 300$ K. (b) MM/GBSA binding free energies of conformational clusters generated in equilibrium simulations at $T = 300$ K for E8 and the rmsd differences of the corresponding cluster centers from the reference lowest energy conformation of E8 determined in docking simulations.

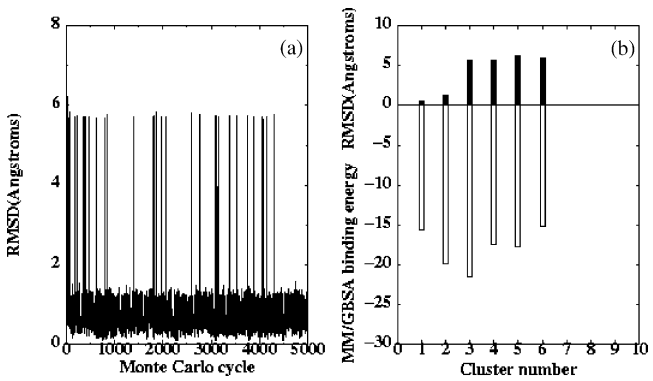

Fig. 7. (a) Time-dependent equilibrium history of F3 binding in the binding site cavity at $T = 300$ K. (b) MM/GBSA binding free energies of conformational clusters generated in equilibrium simulations at $T = 300$ K for F3 and the rmsd differences of the corresponding cluster centers from the reference lowest energy conformation of F3 determined in docking simulations.

1.88, and 4.0 Å from the lowest energy structure (Fig. 3a and b) determined in docking simulations that serves as the reference state in calculations. However, these seemingly alternative binding modes differ by a 180° ligand flip in the binding cavity, whereas still intimately mimicking the position of the W104 hot spot residue from the wild-type complex. The predicted orientation of the benzimidazole core in the first low-energy cluster of D9 closely overlaps W104 side-chain with the 2-methyl group off the benzimidazole protruding further into the cavity and filling the position of the side-chain of T175 in the wild-type complex. Consequently, binding free energy calculations using MM/GBSA energetic model are conducted for all conformations in the structural clusters with an overlap between the benzimidazole core motif and the mutated hot spot residues of the wild-type complex.

Equilibrium simulations and subsequent binding free energy calculations of the low-energy conformational clusters lead to similar results for Dm (Fig. 4a and b) and H9 inhibitors (Fig. 5a and b). Energetically comparable low-energy clusters are observed that differ by 4.25 Å from the reference Dm structure and by 4.12 Å from the H9 reference structure (Figs. 4 and 5). The low-energy ligand binding modes in the engineered cavity are presented for Dm (Fig. 8a and b) and H9 (Fig. 8c and d). While the lowest energy binding modes for Dm and H9 are very similar, they differ by about 180° flip from the lowest energy cluster determined for D9 inhibitor. Nevertheless, the benzimidazole core mimics even more precisely the position of W104 residue in energetically favorable binding modes of Dm and H9 with the 5-methyl group coming off the benzimidazole ring and occu-

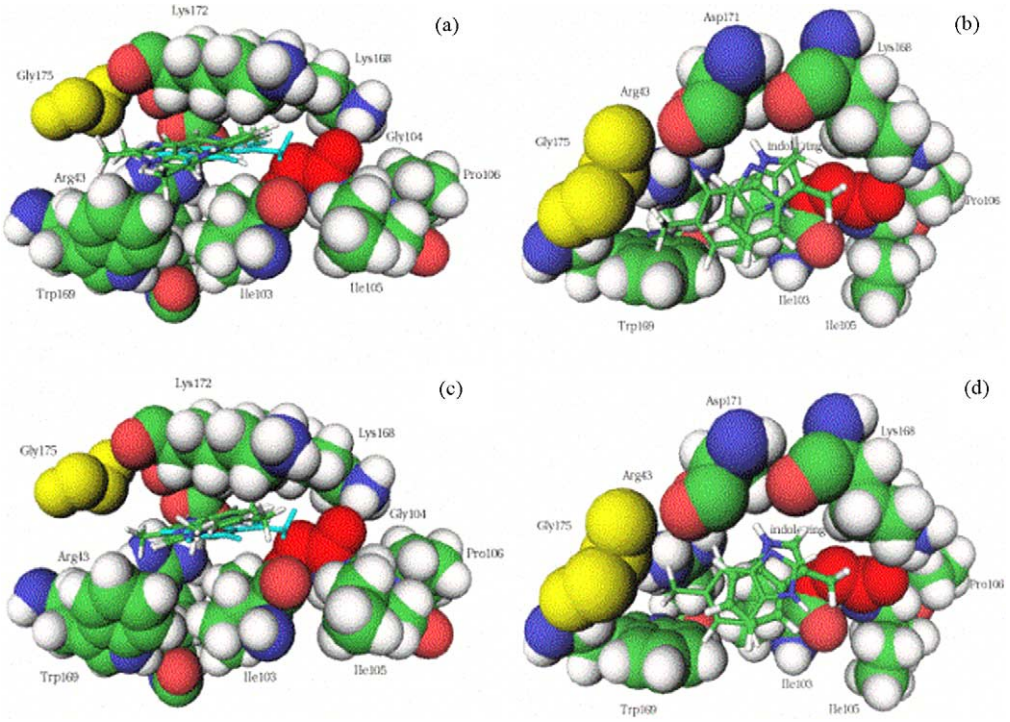

Fig. 8. (a) The low-energy cluster centers (color coded according to their atom type) generated from equilibrium simulations for Dm inhibitor at $T = 300$ K in the binding site cavity. An overlap is seen with the indole ring of W104 residue from the the wild-type complex (colored in light blue). (b) The predicted orientation of the benzimidazole core in the lowest energy cluster obtained for Dm inhibitor from equilibrium simulations. An overlap is seen with the indole ring of W104 in the wild-type complex (all structures are color coded according to their atom type). (c) The low-energy cluster centers (color coded according to their atom type) generated from equilibrium simulations for H9 inhibitor at $T = 300$ K in the binding site cavity. An overlap is seen with the indole ring of W104 residue from the the wild-type complex (colored in grey). (d) The predicted orientation of the benzimidazole core in the lowest energy cluster obtained for H9 inhibitor from equilibrium simulations. An overlap is seen with the indole ring of W104 in the wild-type complex (all structures are color coded according to their atom type).

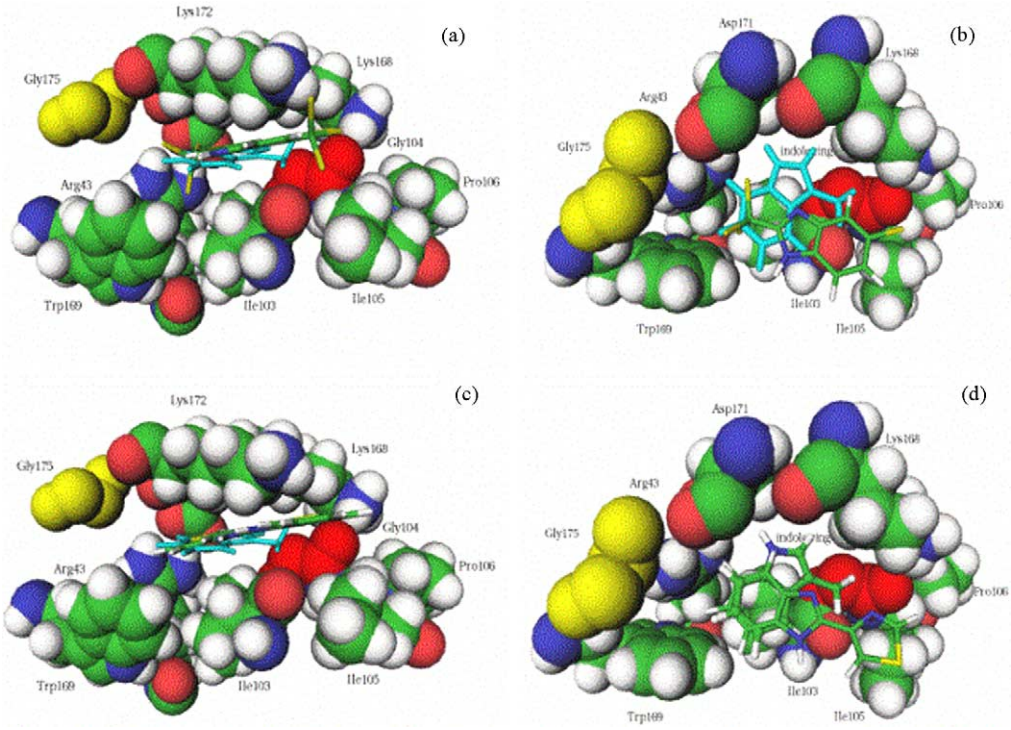

Fig. 9. (a) The low-energy cluster centers (color coded according to their atom type) generated from equilibrium simulations for E8 inhibitor at $T = 300\,\mathrm{K}$ in the binding site cavity. An overlap is seen with the indole ring of W104 residue from the the wild-type complex (colored in light blue). (b) The predicted orientation of the benzimidazole core in the low-energy clusters obateined for E8 inhibitor from equilibrium simulations (color coded according to their atom type). An overlap is seen with the W104 residue in the wild-type complex (colored in light blue). (c) The low-energy cluster centers (color coded according to their atom type) generated from equilibrium simulations for F3 inhibitor at $T = 300\,\mathrm{K}$ in the binding site cavity. An overlap is seen with the indole ring of W104 residue from the the wild-type complex (colored in grey). (d) The predicted orientation of the benzimidazole core in the low-energy clusters of F3 inhibitor obtained from equilibrium simulations (color coded according to their atom type), An overlap is seen with the W104 residue in the wild-type complex (all structures are color coded according to their atom type).

pying a vacancy in the binding cavity left by T175 side-chain in the wild-type complex (Fig. 8b and d). The 2-methyl group in Dm interacts with the alkyl portion of the K168 and K172 side-chains (Fig. 8b), which is consistent with the observed pattern in the wild-type complex where key interactions in the hGH–hGHbp hot spot are alkyl-aromatic stacking interactions, particularly between the alkyl portion of K172 of hGH and W104 of the hGHbp. Alternative binding orientations in the binding site cavity for the E8 and F3 inhibitors (Fig. 9a and c) are also determined by energetically similar ligand conformations that differ by 180° flip but overlap with the indole ring position of W104. In the lowest energy cluster of E8 inhibitor, the trichloromethyl group of the 5-chloro-2-trichloromethyl benzimidazole overlaps considerably with the phenyl portion of the indole ring in W104. This structural arrangement provides more favorable van der Waals contacts in the binding site with W169 and adds to the bulk of critical nonpolar contacts mildly stabilizing electrostatic interactions with R43 residue (Fig. 9b). The lowest energy conformational cluster obtained from equilibrium simulations of F3 is characterized by an overlap between the phenyl portions of the benzimidazole motif in F3 and the indole ring of W104. A bulky thiazole group in 2-position is more directed outwards the cavity to avoid possible short contacts with G175 and interacts favorably with P61, I103, I105 and the aliphatic group of K172 (Fig. 9d). It is obviously possible that local protein conformational changes produced during minimization procedure of complexes in the MM/GBSA model are not sufficient to capture the degree of structural rearrangements that may be necessary to accommodate the thiazole group in the proximity of the G175 residue.

A convincing correlation between the predicted and experimental binding affinities is obtained based on binding free energy calculations of a single lowest energy conformational cluster determined after screening the results of docking simulations (Fig. 10a). By comparing the computed binding free energies determined from equilibrium simulations with the experimentally observed dissociation constants (Fig. 10b), a similar agreement is seen. As a result, we infer that there may be a larger diversity of low-energy conformations in the binding cavity than was originally suggested in docking simulations. However, these ligand binding modes are still compatible with the position of W104 residue. Structural orientations of the ligands that considerably overlap with W104 position, primarily by using the benzimidazole core motif, is a prerequisite for the low-energy ligand conformations binding in the engineered cavity. The results point to a

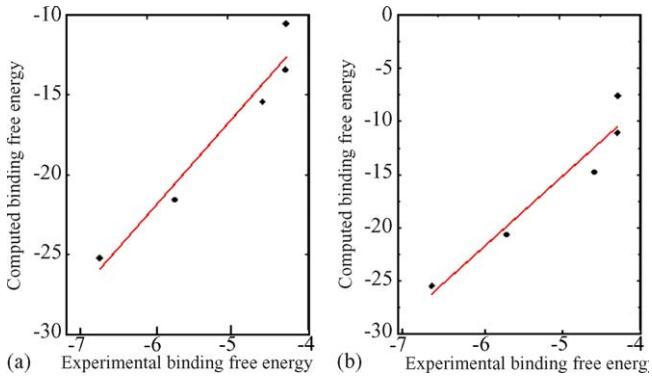

Fig. 10. (a) Correlation between the computed and experimental binding free energies based on binding free energy calculations over a single lowest energy conformational cluster determined after screening the results of docking simulations. The value of correlation coefficient $R = 0.97$ with the regression coefficient (slope) of 5.77. (b) Correlation between the computed and experimental binding free energies determined from equilibrium simulations based on binding free energy contributions averaged over all conformations families of low-energy binding modes that are characterized by an overlap of the benzimidazole core motif with the indole ring of the key W104 residue of the receptor. The value of correlation coefficient $R = 0.96$ with the regression coefficient (slope) of 6.85.

critical role of the intermolecular van der Waals interactions in restoring binding affinity of the remodeled hGH–hGHbp mutant complex. The bulk of the ligand–protein interactions is determined by the appropriate positioning of the benzimidazole core motif in place of mutated W104 residue. The total electrostatics contribution is mildly destabilizing for the majority of the inhibitors, except the most potent E8 ligand, where the electrostatic intermolecular interactions are not fully offset by the unfavorable polar solvation [92–95].

There are a number of limitations of the proposed hierarchical approach in its current implementation, including only a partial account of the protein flexibility in the MM/GBSA model and lack of the solute entropy contribution that was not included in the model due to typical significant variances in its values. The entropy change of proteins upon binding, that can be estimated by either the normal mode analysis or the quasi-harmonic approximation, is extremely difficult to evaluate for the hGH–hGHbp system because of a large size in the all-atom model. In the computational alanine scanning analysis of the hGH–hGHbp system, the results of MM/PBSA simulations without inclusion of the solute entropy term were in a good agreement with the experiment [42]. Another, more serious concern could arise from the assumption of only local conformational changes during binding of the inhibitors. In the framework of the MM/GBSA approach, the ensembles of structures for the uncomplexed protein and ligand are generated by using all conformations from a given cluster. Separation into the protein and ligand structures is followed by minimization of the complexes as well as the unbound protein and unbound ligand for each member of the cluster. This generally allows to capture local protein conformational changes in response to motions of the ligand binding mode. However, in the absence of the crystal structures of the complexes with the inhibitors, this still leaves a considerable ambiguity regarding a possibility of a large global conformational change, observed, for instance, in the remodeled W104A mutant interface in the presence of functional changes in residues K168R, D171T, K172Y, E174A, and F176Y. It is quite clear that there no current modeling techniques that could have predicted or could have adequately simulated the details of these large structural changes. Obviously, only the solution of the crystal structures for the W104G/T175G mutant complex with the panel of studied inhibitors can ultimately validate or disprove some structural suggestions made in our study. Arguably, in the absence of evolutionary pressure to mutate neighboring hot spot residues, as was the case in the phage display remodeling experiments, it may be energetically disadvantageous to rearrange the entire interface to compensate for the loss of interactions, and only small conformational changes near the mutational site may occur. Similar arguments were also presented to explain a good agreement with the experimental changes in binding affinity in computational alanine scanning of W104 and W169 residues [42]. The successful prediction of the relative binding energetics for a panel of studied ligands suggests that the determined ligand binding modes and accompanied local protein conformational changes captured in the MM/GBSA approach may represent a reasonable structural model for ligand binding at the remodeled hGH–hGHbp interface that allows to quantitatively reproduce differences in binding affinity.

4. Conclusions

The presented approach combines equilibrium dynamics, structural clustering and binding free energy evaluations of the low-energy conformational clusters using MM/GBSA model and allows an adequate analysis of binding thermodynamics for studied inhibitors. Binding energetics is predicted in a good agreement with the experimental data for a panel of benzimidazole-containing compounds that complement this defect at the redesigned interface and restore the binding affinity of the wild-type hGH–hGHbp complex. Equilibrium simulations of ligand binding in the binding site cavity of the hGH–hGHbp mutant complex have revealed a certain diversity of binding modes that considerably overlap with the indole ring of W104 residue in the wild-type complex. Structural orientations of the ligands that considerably overlap with W104 position, primarily by using the benzimidazole core motif, is a prerequisite for the low-energy ligand conformations binding in the engineered cavity.

The proposed approach can be also applied to computational mapping of binding sites and finding optimal locations for organic solvents on protein surfaces, that was recently addressed in a conceptually similar hierarchical docking approach [96]. This hierarchical protocol could also present a convenient complementary strategy to computational mapping of potential binding sites in proteins and

locating putative binding points using multiple copy simultaneous search (MCSS) approach [97–100]. A recruitment of unoccupied pockets near the active sites may be another potential application of the presented method, that could be useful in complementing experimental mapping techniques, such as the multiple solvent crystal structures method that can locate and characterize consensus binding sites capable of binding small molecules [101,102]. Understanding structural and energetic characteristics of the hot spots in protein–protein and ligand–protein interfaces by using experimental and computational approaches can ultimately facilitate rational and combinatorial inhibitor design [103].

References

[1] J.A. McCammon, Theory of biomolecular recognition, Curr. Opin. Struct. Biol. 8 (1998) 245–249.
[2] K. Muller-Dethlefs, P. Hobza, Noncovalent interactions: a challenge for experiment an theory, Chem. Rev. 100 (2000) 143–167.
[3] P.G. Gane, P.M. Dean, Recent advances in structure-based rational drug design, Curr. Opin. Struct. Biol. 10 (2000) 401–404.
[4] A.M. Davis, S.J. Teague, Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis, Angew. Chem. Int. Ed. Engl. 38 (1999) 736–749.
[5] M.H. Van Regenmortel, Molecular recognition in the post-reductionist era, J. Mol. Recognit. 12 (1999) 1–2.
[6] H.A. Carlson, J.A. McCammon, Accommodating protein flexibility in computational drug design, Mol. Pharmacol. 57 (2000) 213–218.
[7] W.L. DeLano, Unraveling hot spots in binding interfaces: progress and challenges, Curr. Opin. Struct. Biol. 12 (2002) 14–20.
[8] S. Jones, J.M. Thornton, Principles of protein–protein interactions, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 13–20.
[9] L. Lo Conte, C. Chothia, J. Janin, The atomic structure of protein–protein recognition sites, J. Mol. Biol. 285 (1999) 2177–2198.
[10] A.A. Bogan, K.S. Thorn, Anatomy of hot spots in protein interfaces, J. Mol. Biol. 280 (1998) 1–9.
[11] Z. Hu, B. Ma, H. Wolfson, R. Nussinov, Conservation of polar residues as hot spots at protein interfaces, Proteins 39 (2000) 331–342.
[12] B. Ma, H.J. Wolfson, R. Nussinov, Protein functional epitopes: hot spots, dynamics and combinatorial libraries, Curr. Opin. Struct. Biol. 11 (2001) 364–369.
[13] E.J. Sundberg, R.A. Mariuzza, Luxury accommodations: the expanding role of structural plasticity in protein–protein interactions, Struct. Fold Des. 8 (2000) R137–R142.
[14] H. Frauenfelder, D.T. Leeson, The energy landscape in non-biological and biological molecules, Nat. Struct. Biol. 5 (1998) 757–759.
[15] P.A. Rejto, S.T. Freer, Protein conformational substates from X-ray crystallography, Prog. Biophys. Mol. Biol. 66 (1996) 167–196.
[16] C.J. Tsai, B. Ma, R. Nussinov, Folding and binding cascades: shifts in energy landscapes, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 9970–9972.
[17] S. Kumar, B. Ma, C.J. Tsai, N. Sinha, R. Nussinov, Folding and binding cascades: dynamic landscapes and population shifts, Protein Sci. 9 (2000) 10–19.
[18] A.P. Demchenko, Recognition between flexible protein molecules: induced and assisted folding, J. Mol. Recognit. 14 (2001) 42–61.
[19] J.D. Forman-Kay, The 'dynamics' in the thermodynamics of binding, Nat. Struct. Biol. 6 (1999) 1086–1087.
[20] A.L. Lee, S.A. Kinnear, A.J. Wand, Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex, Nat. Struct. Biol. 7 (2000) 72–77.
[21] L. Zidek, M.V. Novotny, M.J. Stone, Increased protein backbone conformational entropy upon hydrophobic ligand binding, Nat. Struct. Biol. 6 (1999) 1118–1121.
[22] J. Cavanagh, M. Akke, May the driving force be with you—whatever it is, Nat. Struct. Biol. 7 (2000) 11–13.
[23] W.L. DeLano, M.H. Ultsch, A.M. de Vos, J.A. Wells, Convergent solutions to binding at a protein–protein interface, Science 287 (2000) 1279–1283.
[24] D. Tondi, U. Slomczynska, M.P. Costi, D.M. Watterson, S. Ghelli, B.K. Shoichet, Structure-based discovery and in-parallel optimization of novel competitive inhibitors of thymidylate synthase, Chem. Biol. 6 (1999) 319–331.
[25] G.M. Verkhivker, D. Bouzida, D.K. Gehlhaar, P.A. Rejto, S.T. Freer, P.W. Rose, Monte Carlo simulations of the peptide recognition at the consensus binding site of the constant fragment (Fc) of human immunoglobulin G: the energy landscape analysis of a hot spot at the intermolecular interface, Proteins 48 (2002) 539–557.
[26] J.A. Wells, Systematic mutational analyses of protein–protein interfaces, Methods Enzymol 202 (1991) 390–411.
[27] J.A. Wells, A.M. de Vos, Structure and function of human growth hormone: implications for the hematopoietins, Annu. Rev. Biophys. Biomol. Struct. 22 (1993) 329–351.
[28] J.A. Wells, Binding in the growth hormone receptor complex, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 1–6.
[29] B.C. Cunningham, P. Jhurani, P. Ng, J.A. Wells, Receptor and antibody epitopes in human growth hormone identified by homolog-scanning mutagenesis, Science 243 (1989) 1330–1336.
[30] B.C. Cunningham, J.A. Wells, High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis, Science 244 (1989) 1081–1085.
[31] B.C. Cunningham, M. Ultsch, A.M. de Vos, M.G. Mulkerrin, K.R. Clauser, J.A. Wells, Dimerization of the extracellular domain of the human growth hormone receptor by a single hormone molecule, Science 254 (1991) 821–825.
[32] M. Ultsch, A.M. de Vos, A.A. Kossiakoff, Crystals of the complex between human growth hormone and the extracellular domain of its receptor, J. Mol. Biol. 222 (1991) 865–868.
[33] A.M. de Vos, M. Ultsch, A.A. Kossiakoff, Human growth hormone and extracellular domain of its receptor: crystal structure of the complex, Science 255 (1992) 306–312.
[34] A.A. Kossiakoff, W. Somers, M. Ultsch, K. Andow, Y.A. Muller, A.M. de Vos, Comparison of the intermediate complexes of human growth hormone bound to the human growth hormone and prolactin receptors, Protein Sci. 3 (1994) 1697–1705.
[35] W. Somers, M. Ultsch, A.M. de Vos, A.A. Kossiakoff, The X-ray structure of a growth hormone–prolactin receptor complex, Nature 372 (1994) 478–481.
[36] S.H. Bass, M.G. Mulkerrin, J.A. Wells, A systematic mutational analysis of hormone-binding determinants in the human growth hormone receptor, Proc. Natl. Acad. Sci. U.S.A. 88 (1991) 4498–44502.
[37] B.C. Cunningham, J.A. Wells, Comparison of a structural and a functional epitope, J. Mol. Biol. 234 (1993) 554–563.
[38] T. Clackson, J. Wells, A hot spot of binding energy in a hormone–receptor interface, Science 267 (1995) 383–386.
[39] L. Jin, J.A. Wells, Dissecting the energetics of an antibody–antigen interface by alanine shaving and molecular grafting, Protein Sci. 3 (1994) 2351–2357.
[40] K.H. Pearce Jr., M.H. Ultsch, R.F. Kelley, A.M. de Vos, J.A. Wells, Structural and mutational analysis of affinity-inert contact residues at the growth hormone–receptor interface, Biochemistry 35 (1996) 10300–10307.
[41] T. Clackson, M.H. Ultsch, J.A. Wells, A.M. de Vos, Structural and functional analysis of the 1:1 growth hormone:receptor complex reveals the molecular basis for receptor affinity, J. Mol. Biol. 277 (1998) 1111–1128.

[42] S. Huo, I. Massova, P.A. Kollman, Computational alanine scanning of the 1:1 human growth hormone–receptor complex, J. Comput. Chem. 23 (2002) 15–27.
[43] S. Atwell, M. Ultsch, A.M. de Vos, J.A. Wells, Structural plasticity in a remodeled protein–protein interface, Science 278 (1997) 1125–1128.
[44] H.B. Lowman, J.A. Wells, Affinity maturation of human growth hormone by monovalent phage display, J. Mol. Biol. 234 (1993) 564–578.
[45] C. Schiffer, M. Ultsch, S. Walsh, W. Somers, A.M. de Vos, A. Kossiakoff, Structure of a phage display-derived variant of human growth hormone complexed to two copies of the extracellular domain of its receptor: evidence for strong structural coupling between receptor binding sites, J. Mol. Biol. 316 (2002) 277–289.
[46] A. Bishop, O. Buzko, S. Heyeck-Dumas, I. Jung, B. Kraybill, Y. Liu, K. Shah, S. Ulrich, L. Witucki, F. Yang, C. Zhang, K.M. Shokat, Unnatural ligands for engineered proteins: new tools for chemical genetics, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 577–606.
[47] A.C. Bishop, O. Buzko, K.M. Shokat, Magic bullets for protein kinases, Trends Cell. Biol. 11 (2001) 167–172.
[48] M.A. Shogren-Knaak, P.J. Alaimo, K.M. Shokat, Recent advances in chemical approaches to the study of biological systems, Annu. Rev. Cell. Dev. Biol. 17 (2001) 405–433.
[49] P.J. Alaimo, M.A. Shogren-Knaak, K. Shokat, Chemical genetic approaches for the elucidation of signaling pathways, Curr. Opin. Chem. Biol. 5 (2001) 360–367.
[50] K.M. Specht, K.M. Shokat, The emerging power of chemical genetics, Curr. Opin. Cell. Biol. 14 (2002) 155–159.
[51] Z. Guo, D. Zhou, P.G. Schultz, Designing small-molecule switches for protein–protein interactions, Science 288 (2000) 2042–2045.
[52] G.M. Verkhivker, D. Bouzida, D.K. Gehlhaar, P.A. Rejto, S.T. Freer, P.W. Rose, Computational detection of the binding site hot spot at the remodeled human growth hormone–receptor interface. Proteins 53 (2003) 201–219.
[53] P.A. Rejto, G.M. Verkhivker, Unraveling principles of lead discovery: from unfrustrated energy landscapes to novel molecular anchors, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 8945–8950.
[54] G.M. Verkhivker, P.A. Rejto, A mean field model of ligand–protein interactions, implications for the structural assessment of human immunodeficiency virus type 1 protease complexes and receptor-specific binding, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 60–64.
[55] G.M. Verkhivker, P.A. Rejto, D.K. Gehlhaar, S.T. Freer, Exploring energy landscapes of molecular recognition by a genetic algorithm, analysis of the requirements for robust docking of HIV-1 protease and FKBP-12 complexes, Proteins 25 (1996) 342–353.
[56] G.M. Verkhivker, D. Bouzida, D.K. Gehlhaar, P.A. Rejto, L. Schaffer, S. Arthurs, A.B. Colson, S.T. Freer, V. Larson, B.A. Luty, T. Marrone, P.W. Rose, Hierarchy of simulation models in predicting molecular recognition mechanisms from the binding energy landscapes. Structural analysis of the peptide complexes with SH2 domains, Proteins 45 (2001) 456–470.
[57] G.M. Verkhivker, D. Bouzida, D.K. Gehlhaar, P.A. Rejto, S. Arthurs, A.B. Colson, S.T. Freer, V. Larson, B.A. Luty, T. Marrone, P.W. Rose, Deciphering common failures in molecular docking of ligand–protein complexes, J. Comput. Aided Mol. Des. 14 (2000) 731–751.
[58] G.M. Verkhivker, D. Bouzida, D.K. Gehlhaar, P.A. Rejto, L. Schaffer, S. Arthurs, A.B. Colson, S.T. Freer, V. Larson, B.A. Luty, T. Marrone, P.W. Rose, Hierarchy of simulation models in predicting structure and energetics of the Src SH2 domain binding to the tyrosyl phosphopeptides, J. Med. Chem. 45 (2002) 72–89.
[59] D. Bouzida, P.A. Rejto, S. Arthurs, A.B. Colson, S.T. Freer, D.K. Gehlhaar, V. Larson, B.A. Luty, P.W. Rose, G.M. Verkhivker, 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 (1999) 73–84.
[60] D. Bouzida, P.A. Rejto, G.M. Verkhivker, Monte Carlo simulations of ligand–protein binding energy landscapes With the weighted histogram analysis method, Int. J. Quantum Chem. 73 (1999) 113–121.
[61] G.M. Verkhivker, P.A. Rejto, D. Bouzida, S. Arthurs, A.B. Colson, S.T. Freer, D.K. Gehlhaar, V. Larson, B.A. Luty, T. Marrone, P.W. Rose, Towards understanding the mechanisms of molecular recognition by computer simulations of ligand–protein interactions, J. Mol. Recognit. 12 (1999) 371–389.
[62] S.L Mayo, B.D. Olafson, W.A. Goddard III, DREIDING: a generic force field for molecular simulation, J. Phys. Chem. 94 (1990) 8897–8909.
[63] S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Chio, G. Alagona, S. Profeta, P. Weiner, A new force field for molecular mechanical simulation of nucleic acids and proteins, J. Am. Chem. Soc. 106 (1984) 765–784.
[64] W.C. Still, A. Tempczyk, R.C. Hawley, T. Hendrickson, Semianalytical treatment of solvation for molecular mechanics and dynamics, J. Am. Chem. Soc. 112 (1990) 6127–6129.
[65] F. Mohamadi, N.G.J. Richards, W.C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W.C. Still, MacroModel—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics, J. Comput. Chem. 11 (1990) 440–467.
[66] D. Qiu, P.S. Shenkin, F.P. Hollinger, W.C. Still, The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate born radii, J. Phys. Chem. A 101 (1997) 3005–3014.
[67] J. Weiser, A.A. Weiser, P.S. Shenkin, W.C. Still, Neighbor-list reduction: optimization for computation of molecular van der Waals and solvent-accessible surface areas, J. Comput. Chem. 19 (1998) 797–808.
[68] J. Weiser, A.A. Weiser, P.S. Shenkin, W.C. Still, Erratum: neighbor-list reduction: optimization for computation of molecular van der Waals and solvent-accessible surface areas, J. Comput. Chem. 19 (1998) 1110.
[69] J. Weiser, P.S. Shenkin, W.C. Still, Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO), J. Comput. Chem. 20 (1999) 217–230.
[70] J. Weiser, P.S. Shenkin, W.C. Still, Fast, approximate algorithm for detection of solvent-inaccessible atoms, J. Comput. Chem. 20 (1999) 586–596.
[71] J. Srinivasan, T.E. Cheatham, P. Cieplak, P.A. Kollman, D.A. Case, Continuum solvent studies of the stability of DNA, RNA and phosphoramidate-DNA helices, J. Am. Chem. Soc. 120 (1998) 9401–9409.
[72] I. Massova, P.A. Kollman, Computational alanine scanning to probe protein–protein interactions: a novel approach to evaluate binding free energies, J. Am. Chem. Soc. 121 (1999) 8133–8143.
[73] L.T. Chong, Y. Duan, L. Wang, I. Massova, P.A. Kollman, Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 14330–14335.
[74] B. Kuhn, P.A. Kollman, A ligand that is predicted to bind better to avidin than biotin: insights from computational fluorine scanning, J. Am. Chem. Soc. 122 (2000) 3909–3916.
[75] B. Kuhn, P.A. Kollman, Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models, J. Med. Chem. 43 (2000) 3786–3791.
[76] M.R. Lee, Y. Duan, P.A. Kollman, Use of MM-PB/SA in estimating the free energies of proteins: application to native, intermediates, and unfolded villin headpiece, Proteins 39 (2000) 309–316.
[77] W. Wang, P.A. Kollman, Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model, J. Mol. Biol. 303 (2000) 567–582.

[78] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T.E. Cheatham III, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc. Chem. Res. 33 (2000) 889–897.
[79] V. Tsui, D.A. Case, Molecular dynamics simulations of nucleic acids with a generalized Born solvation model, J. Am. Chem. Soc. 122 (2000) 2489–2498.
[80] E. Marinari, G. Parisi, Simulated tempering: a new Monte Carlo scheme, Europhys. Lett. 19 (1992) 451–458.
[81] K. Hukushima, K. Nemoto, Exchange Monte Carlo method and application to spin glass simulations, J. Phys. Soc. Jpn. 65 (1996) 1604–1607.
[82] U.H.E. Hansmann, Y. Okamoto, Monte Carlo simulations in generalized ensemble: multicanonical algorithm versus simulated tempering, Phys. Rev. E 54 (1996) 5863–5865.
[83] U.H.E. Hansmann, Y. Okamoto, Generalized-ensemble Monte Carlo method for systems with rough energy landscape, Phys. Rev. E 56 (1997) 2228–2233.
[84] U.H.E. Hansmann, Y. Okamoto, Numerical comparisons of three recently proposed algorithms in the protein folding problem, J. Comput. Chem. 18 (1997) 920–933.
[85] U.H.E. Hansmann, Parallel tempering algorithm for conformational studies of biological molecules, Chem. Phys. Lett. 281 (1997) 140–150.
[86] Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding, Chem. Phys. Lett. 314 (1999) 141–151.
[87] G.M. Verkhivker, P.A. Rejto, D. Bouzida, S. Arthurs, A.B. Colson, S.T. Freer, D.K. Gehlhaar, V. Larson, B.A. Luty, T. Marrone, P.W. Rose, Navigating ligand–protein binding free energy landscapes: universality and diversity of protein folding and molecular recognition mechanisms, Chem. Phys. Lett. 336 (2001) 495–503.
[88] D. Bouzida, S. Kumar, R.H. Swendsen, Efficient Monte Carlo methods for the computer simulation of biological molecules, Phys. Rev. A 45 (1992) 8894–8901.
[89] P. Willet, V. Winterman, A comparison of some measures for the determination of intermolecular structural similarity, Quant. Struct.-Act. Relat. Pharmacol. Chem. Biol. 5 (1986) 18–25.
[90] P. Willet, V. Winterman, D. Bawden, Implementation of non-hierarchical cluster analysis methods in chemical information systems: selection of compounds for biological testing and clustering of substructure search output, J. Chem. Inf. Comput. Sci. 26 (1986) 109–118.
[91] D. Bawden, Browsing and clustering of chemical structures, in: W.A. Warr (Ed.), The International Language of Chemistry, Springer-Verlag, Berlin, 1988, pp. 145–150.
[92] J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar, S. Subramaniam, Analytical shape computation of macromolecules. I. Molecular area and volume through alpha shape, Proteins 33 (1998) 1–17.
[93] J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar, S. Subramaniam, Analytical shape computation of macromolecules. II. Inaccessible cavities in proteins, Proteins 33 (1998) 18–29.
[94] J. Liang, H. Edelsbrunner, C. Woodward, Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design, Protein Sci. 7 (1998) 1884–1897.
[95] F.B. Sheinerman, B. Honig, On the role of electrostatic interactions in the design of protein–protein interfaces, J. Mol. Biol. 316 (2002) 161–177.
[96] S. Dennis, T. Kortvelyesi, S. Vajda, Computational mapping identifies the binding sites of organic solvents on proteins, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 4290–4295.
[97] A. Caflisch, A. Miranker, M. Karplus, Multiple copy simultaneous search and construction of ligands in binding sites: application to inhibitors of HIV-1 aspartic proteinase, J. Med. Chem. 36 (1993) 2142–2167.
[98] C.M. Stultz, M. Karplus, MCSS functionality maps for a flexible proteins, Proteins 37 (1999) 512–529.
[99] D. Joseph-McCarthy, S.K. Tsang, D.J. Filman, J.M. Hogle, M. Karplus, Use of MCSS to design small targeted libraries: application to picornavirus ligands, J. Am. Chem. Soc. 123 (2001) 12758–12766.
[100] R. Bitetti-Putzer, D. Joseph-McCarthy, J.M. Hogle, M. Karplus, Functional group placement in protein binding sites: a comparison of GRID and MCSS, J. Comput. Aided Mol. Des. 15 (2001) 935–960.
[101] C. Mattos, D. Ringe, Locating and characterizing binding sites on proteins, Nat. Biotechnol. 14 (1996) 595–599.
[102] D. Ringe, C. Mattos, Analysis of the binding surfaces of proteins, Med. Res. Rev. 19 (1999) 321–331.
[103] R.A. Powers, B.K. Shoichet, Structure-based approach for binding site identification on AmpC beta-lactamase, J. Med. Chem. 45 (2002) 3222–3234.