Computational Alanine Scanning of the 1:1 Human Growth Hormone–Receptor Complex

SHUANGHONG HUO, IRINA MASSOVA, PETER A. KOLLMAN*

Department of Pharmaceutical Chemistry, University of California, San Francisco, 513 Parnassus Ave., San Francisco, California 94143-0446

Received 15 January 2001; Accepted 27 June 2001

Abstract: The MM-PBSA (Molecular Mechanics–Poisson–Boltzmann surface area) method was applied to the human Growth Hormone (hGH) complexed with its receptor to assess both the validity and the limitations of the computational alanine scanning approach. A 400-ps dynamical trajectory of the fully solvated complex was simulated at 300 K in a 101 Å × 81 Å × 107 Å water box using periodic boundary conditions. Long-range electrostatic interactions were treated with the particle mesh Ewald (PME) summation method. Equally spaced snapshots along the trajectory were chosen to compute the binding free energy using a continuum solvation model to calculate the electrostatic desolvation free energy and a solvent-accessible surface area approach to treat the nonpolar solvation free energy. Computational alanine scanning was performed on the same set of snapshots by mutating the residues in the structural epitope of the hormone and the receptor to alanine and recomputing the $\Delta G_{\text{binding}}$. To further investigate a particular structure, a 200-ps dynamical trajectory of an R43A hormone–receptor complex was simulated. By postprocessing a single trajectory of the wild-type complex, the average unsigned error of our calculated $\Delta\Delta G_{\text{binding}}$ is ~1 kcal/mol for the alanine mutations of hydrophobic residues and polar/charged residues without buried salt bridges. When residues involved in buried salt bridges are mutated to alanine, it is demonstrated that a separate trajectory of the alanine mutant complex can lead to reasonable agreement with experimental results. Our approach can be extended to rapid screening of a variety of possible modifications to binding sites.

 J Comput Chem 23: 15–27, 2002

Key words: protein–protein interaction; site-directed mutagenesis; binding free energy; molecular dynamics; continuum model

Introduction

Both structure and free energy are essential features of macromolecular systems. Although it is relatively straightforward to simulate the structural features of macromolecular systems, provided sufficient sampling is performed, or, if starting near an experimental structure, to characterize the structural features near its initial point, the simulation of free energies is more difficult. Recently, a new computational approach, termed molecular mechanics–Poisson–Boltzmann surface area (MM–PBSA), has shown great promise in its ability to describe the free energy of macromolecular systems.[1] This algorithm combines explicit solvent molecular dynamics (MD) simulations with implicit solvation models, Poisson–Boltzmann analysis[2, 3] and nonpolar solvation free energy calculations,[4] to estimate free energies. A set of "snapshots" along an MD trajectory, which for the highest accuracy is carried out using a periodic box of water molecules and the PME[5] method to represent long-range electrostatics, is saved as representative conformations of a macromolecular complex. When this set of structures is postprocessed, the water molecules are removed and replaced by a continuum solvent model. The free energy of the macromolecular complex consists of the molecular mechanics potential energy of the solute, the solvation free energy, and a solute entropy term. The solvation free energy is further composed of an electrostatic or polar portion, obtained by solving the Poisson–Boltzmann (PB) equation, and a nonpolar solvation contribution associated with cavity formation in the solvent as well as van der Waals (vdW) interactions between the solute and solvent. The entropy of the macromolecule can be estimated by either normal mode or quasi-harmonic analysis.[6] Related free energy approaches have been described by Jayaram[7] and Vorobjev.[8]

The MM-PBSA approach has been successfully used to calculate the relative stability of A- and B-DNA and RNA helices in water, as well as other nucleic acid systems.[1, 9] As an extension of the MM-PBSA approach, computational alanine scanning[10] was developed to estimate the contribution of an individual side chain to the overall protein–protein binding free energies. Massova

*P. A. Kollman died on May 25, 2001.

Correspondence to: S. Huo, Gustaf H. Carlson School of Chemistry & Biochemistry, Clark University, Worcester, MA 01610; e-mail: shuo@clarku.edu

Contract/grant sponsor: NIH; contract/grant number: GM-29072

and Kollman performed an explicit solvent MD simulation of a truncated p53-Mdm2 complex, and then mutated the resulting trajectory by replacing each side chain of p53 one at a time with alanine, and then recalculating the binding free energy. They identified three hydrophobic residues as the critical binding points, consistent with the results of site-directed mutagenesis.[11] A comparison of PB analysis and the generalized Born (GB) model was also performed on the same system,[12] and it was found that the free energy changes upon alanine mutation determined by the GB model are larger and less accurate than those of the PB model. Because GB is significantly faster than PB, however, it may still be an attractive approach for initial screening. The results of computational alanine scanning of p53-Mdm2 are particularly encouraging because the computational data does an excellent job of reproducing the qualitative trends in the experimental data for all 12 amino acids of the p53 peptide.[10] However, the p53-Mdm2 interface is relatively rigid, and there are no completely buried charged or polar residues involved at the binding interface. Moreover, the experimental data reported only the qualitative trend of the contribution to the binding of each residue, instead of the quantitative free energy differences.[11]

The system of human growth hormone complexed with human growth hormone binding protein (hGHbp) has been extensively examined by Wells and coworkers.[13–17] Based on the combination of X-ray crystallography and systematic alanine scanning of hGH and hGHbp, they revealed the "hot spots" of binding free energy.[15] For the 1:1 hGH:hGHbp complex, the binding interface is 1350 Å^2, with 31 residues forming the structural epitope on the receptor side. A central patch of 11 residues out of these 31 residues contributes more than three-quarters of the total binding free energy, with two tryptophans contributing more than 4.5 kcal/mol each.[15] Given the structural plasticity of the protein–protein interface[14] and the burying of charged or polar residues inside the binding interface (in contrast to the relatively rigid p53-Mdm2 where we tested the method), the hGH:hGHbp complex is a particularly challenging system to further test our computational alanine scanning approach.

In this study, we address the following questions: How well can the computational alanine scanning method reproduce the quantitative free energy differences obtained from experiments? How will this method work on a large and flexible protein–protein binding interface? Will this method identify the hydrophobic residues as "hot spots" of binding free energy among some buried charged or polar groups? Is MD simulation necessary for the accurate estimation of free energy difference between the mutant and wild-type complexes, or will simple minimization suffice?

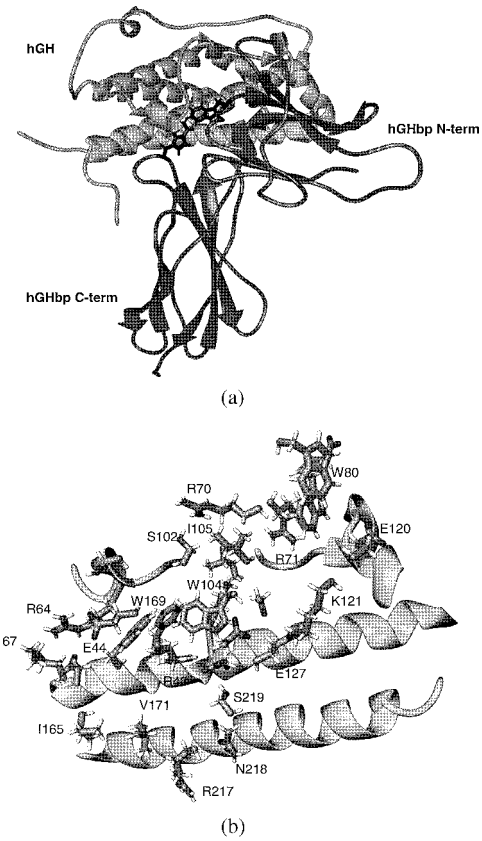


Figure 1. (a) The average structure over the trajectory of the wild-type complex, 1:1 human growth hormone and its receptor (pdb code:1a22). The "hot spots" of binding free energy, hGHbp W104 and hGHbp W169, are shown explicitly. (b) The alanine scanning sites of hGH and hGHbp. For perspicuity, only the interface on the hormone side is shown. The residues subject to alanine mutations in hGH and hGHbp are explicitly shown.

Methodology

The Protocols of MD Simulation

The starting configuration of our simulation is the X-ray crystal structure of human growth hormone bound to the extracellular domain of a single growth hormone receptor.[15] The following missing residues in the X-ray coordinates were built with SYBYL (Tripos Association, St. Louis, MO): 130–135 and 149–153 of the hormone, 212–220, 252–260, and 344–347 of the receptor. None of these residues is close to the binding interface. The structure of the complex is shown in Figure 1(a). The hGH comprises a four-helix bundle core and three minihelices in the connecting loop regions. The hGHbp contains two domains, each of which consists of seven β-strands. The receptor contacts the hormone with some loop regions connecting the β-strands (for details see ref. 15).

A modified shared-memory parallel version of Amber5[18] was employed using the all-atom force field parameters of parm94.[19] Thirteen crystallographic water molecules near at the edge of the binding interface were kept. To keep the whole system neutral, nine Na^+ ions were added. TIP3P three-site rigid water model[20] was used to solvate the protein–protein complex and counterions. Water molecules were removed if they were closer than 1 Å to any protein atoms. The distance between the wall of the water box and the closest atom of the solute was 10 Å. In summary, the system was constructed using periodic boundary conditions with a rectangular box of dimension 101 Å × 81 Å × 107 Å, consisting of an 1:1 hGH:hGHbp complex, 13 crystallographic water molecules, 9 counterions, and 22384 TIP3P water molecules. The system was minimized by 30 steps of steepest decent followed by 970 steps of conjugate gradient to release the bad contacts in the crystal

structure. During the minimization, an 8-Å nonbonded cutoff was applied.

The MD simulations were performed in the N, P, T ensemble (P = 1 bar, $T = 300$ K). The PME summation method[5] was used to treat the long-range electrostatic interactions. All bond lengths were constrained with the SHAKE algorithm.[21] The time step was 2 fs, and the nonbonded pair list was updated every 25 steps. The short-range nonbonded interactions were truncated with a 9-Å cutoff. The temperature and pressure of the system was regulated using the Berendsen[22] coupling algorithm with a coupling constant of 0.2 ps. The system was gradually heated to 300 K, with a 5 ps interval per 100 K. After 25 ps of equilibration at 300 K, the production MD trajectory with one snapshot per picosecond was collected for 400 ps. The total CPU time for the 440 ps simulation was ca. 985 h, using four processors on an Origin200 located in this laboratory. A second trajectory was generated with the starting point of the X-ray structure of 1:1 hGH:hGHbp with R43 in hGHbp modeled as alanine. Ten Na^+ were added to neutralize this system. Following the same protocol as described above, a 200-ps production trajectory was generated.

In the analysis of the complex structure in solution, we measured the persistence of all possible hydrogen bonds between hGH and hGHbp. We consider hydrogen bond defined by distances of the heavy atoms of donor and acceptor of no more than 3.2 Å, and angles of donor and acceptor diatomic groups of no less than 120°. Furthermore, we consider a hydrogen bond between two oppositely charged groups to be a salt bridge. The occupancy of a hydrogen bond was computed by dividing the number of snapshots showing the hydrogen bond by the total number of snapshots along the MD trajectory. The uncertainty of the hydrogen bond occupancy is estimated by dividing the whole trajectory into blocks with 10 snapshots per block and then computing the standard deviation of the occupancy.

$\Delta\Delta G_{\text{binding}}$ *Calculation Methods*

The binding free energy difference between the mutant and wild type complexes is defined as:

$$\Delta\Delta G_{\text{binding}} = \Delta G_{\text{binding\_mutant}} - \Delta G_{\text{binding\_wild type}} \quad (1)$$

$$\Delta G_{\text{binding}} = G_{\text{(complex)}} - \left[G(\text{hGH}) + G(\text{hGHbp})\right] \quad (2)$$

$$G_{\text{molecule}} = \langle E_{\text{MM}}\rangle + \left\langle G^{\text{polar}}_{\text{solvation}}\right\rangle + \left\langle G^{\text{nonpolar}}_{\text{solvation}}\right\rangle - \text{TS} \quad (3)$$

$$\langle E_{\text{MM}}\rangle = \langle E_{\text{internal}}\rangle + \langle E_{\text{electrostatic}}\rangle + \langle E_{\text{vdW}}\rangle \quad (4)$$

where $\langle\ \rangle$ denotes an average over a set of snapshots along an MD trajectory. E_{internal} includes the bond, angle, and torsional angle energies, while $E_{\text{eletrostatic}}$ and E_{vdw} denote electrostatic and vdW energies, respectively. For the analysis of $\Delta\Delta G_{\text{binding}}$, both the water molecules and counterions were then deleted. As the crystallographic water molecules are at the edge of the binding interface and do not form any water bridges connecting any mutant residues, the effect of crystallographic water on the $\Delta\Delta G_{\text{binding}}$ is insignificant. The internal energy, electrostatic, and vdW interactions were calculated using Amber5[18] with no cutoff. The polar contribution to the solvation-free energy ($\Delta G^{\text{polar}}_{\text{solvation}}$) was calculated by solving the Poisson–Boltzmann equation with the Delphi program.[3] The polar solvation process is equivalent to the transfer of a protein from one medium with dielectric constant equal to that of the interior of the protein to another medium with dielectric constant equal to that of the exterior of the protein. This term yields free energies because it corresponds to the work done to reversibly charge the solute, and it is a polarization free energy because the work goes to the polarization of the solvent. For the Delphi calculation, the grid spacing was set to 0.5 Å, the molecule filled 80% of the grid box, 1000 iterations were performed to ensure the maximum change in potential was less than 0.001 kT/e, and the dielectric constant inside and outside the protein was 1.0 and 80.0, respectively. The ionic strength was zero. PARSE[23] vdW radii and standard parm94[19] charges were used. The nonpolar solvation contribution includes cavity creation in water and vdW interactions between the modeled nonpolar protein and water molecules. This term can be imagined as transferring a nonpolar molecule with the shape of the protein from vacuum to water. This transfer free energy is described as[23]

$$\Delta G^{\text{nonpolar}}_{\text{solvation}} = \gamma A + b,$$

where A is the solvent-accessible surface area calculated by the MSMS program,[4] and γ and b are 0.00542 kcal/mol · Å^2 and 0.92 kcal/mol, respectively, which were derived from experimental transfer energies of hydrocarbons.[23] The probe radius was 1.4 Å.

In eq. (3), S is the entropy of the proteins. The solvent entropy changes caused by polarization and cavity formation are included in the polar and nonpolar solvation-free energy terms. It was found that the entropy changes upon binding for the wild-type and alanine mutant p53–Mdm2 complexes were approximately equal using normal mode analyses.[10] Thus, in this study the entropy changes of the wild type and its mutant complexes upon binding are assumed to be equal.

Computational Alanine Scanning Strategy

From the wild-type 400-ps production trajectory, a snapshot was chosen every 4 ps to run alanine scanning. Our experience with MM-PBSA[1, 10, 24–27] indicates that 25–100 snapshots are required for reasonable statistics. One can calculate the $\Delta\Delta G_{\text{binding}}$ using eq. (1) at various levels of approximation. The simplest and most efficient approach is a single trajectory mutation protocol, in which we use the same set of snapshots for the wild-type complex to calculate $\Delta G_{\text{binding\_mutant}}$. We simply truncate a side chain, replacing C_γ with a hydrogen atom, and set the C_β—H bond length and direction to those of the former C_β—C_γ. The underlying approximations of the single trajectory mutation protocol are that the mutant and the wild type undergo similar conformational changes from the unbound to the bound state, and that local side chain reorganizations are small perturbations relative to the alanine mutation itself. However, when the conformations of the wild-type and the mutant are expected to be quite different, as one finds for hGHbp R43A, one can run separate trajectories on the wild-type and mutant species, in which case the evaluation of eqs. (2) and (3) is carried out for two separate trajectories.

Finally, one could, in principle, run separate trajectories of the wild-type and mutant hGH, hGHbp, and hGH:hGHbp, evaluating G with eq. (3) for each of these six separate trajectories. This is the

most computationally demanding and, because the intramolecular terms do not cancel with this approach, the most subject to structural noise. Thus, this model is quite impractical to apply for the hGH and hGHbp system.

Results and Discussion

Evaluating the $\Delta\Delta G_{binding}$ from the Trajectory of the Wild-Type 1:1 Hormone–Receptor Complex (Single Trajectory Mutation Protocol)

For the production (40–440 ps) dynamics, we plotted the time series of root-mean-square (rms) deviation from the X-ray crystal structure of C_α atoms of the complex and of the protein–protein interface. The missing fragments in the crystal structure were not included in the rms deviation analysis. In Figure 2a, it is clear that the interface is very stable during the simulation, while the complex as a whole is relatively flexible. The maximum C_α rms deviation, 1.71 Å, for the complex as a whole occurs at 308 ps; however, the C_α rms deviation of the interface at this point is only 0.66 Å. Superimposing the interface region of the snapshot with the X-ray crystal structure results in two regions with the most deviation: the five loop regions that are unresolved in the X-ray structure and the β-strands in the C-terminal domain of the receptor. To further test the convergence of the simulation, we calculated the effective energy ($\langle E_{MM}\rangle + \langle G^{polar}_{solvation}\rangle + \langle G^{nonpolar}_{solvation}\rangle$) along the molecular dynamics trajectories, as shown in Figure 2b. The relative fluctuation is very small, yet the effective energy does not appear to be well converged until 140 ps; thus, the following analyses are based on the conformations sampled after 140 ps at 4-ps intervals.

As seen in Table 1, the intermolecular vdW interaction and the nonpolar solvation term provide the driving force for binding. The favorable (negative) intermolecular electrostatic interaction is counteracted by the unfavorable (positive) electrostatic desolvation free energy. As a result, the total electrostatic contribution is unfavorable for binding. Because we did not include the solute entropy contribution, the binding energy shown here (−52.0 kcal/mol) is not the total free energy. The entropy change of proteins upon binding can be estimated by either the normal mode analysis or the quasi-harmonic approximation. But it is extremely difficult to handle the Hessian matrix or the covariance fluctuation matrix of an all atom model system with around 400 residues because of the computer resource limitation and the accumulation of round-off errors.[6] Nonetheless, based on previous results,[10, 24, 25] we expect that the magnitude of $T\Delta S$ would be in the range of +30–40 kcal/mol, leading to a qualitatively reasonable free energy of binding.

Figure 1b shows the chosen alanine scanning sites of hGHbp. Because Trp104 and Trp169 of hGHbp were identified as "hot spots" of binding free energy ($\Delta\Delta G_{binding} > 4.5$ kcal/mol) by experimental alanine scanning,[15] we selected all the Trps in the hGHbp structural epitope, which have experimental data for $\Delta\Delta G_{binding}$ as sites for our computational alanine scanning. Other large contributions ($\Delta\Delta G_{binding} > 1.5$ kcal/mol) to the binding free energy found in experiments included hGHbp P106, hGHbp I165, hGHbp R43, hGHbp I105, hGHbp E44, and hGHbp I103.[15] From a crystal structure analysis, Wells and coworkers proposed that hGHbp P106A might cause a large conformational change[15]

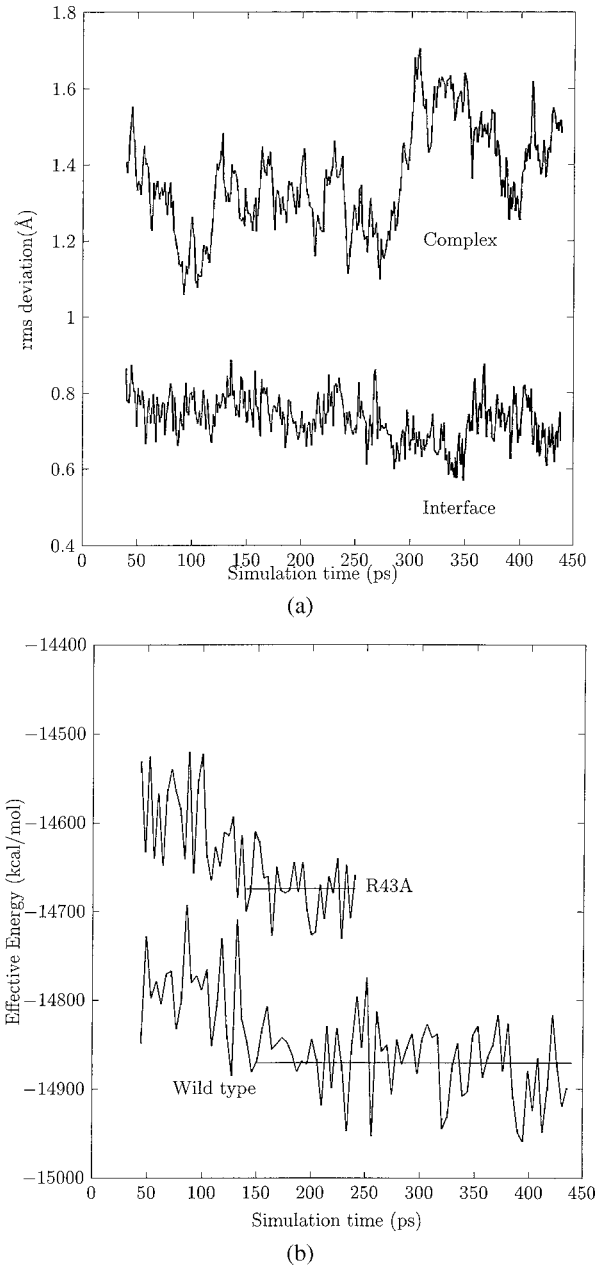


Figure 2. The time series of the rms deviation from the X-ray crystal structure of C_α atoms of the wild-type complex and of the protein–protein interface. The plot describes the 400 ps production dynamics; the data for the first 40 ps of equilibration is not included. (b) The effective energy ($\langle E_{MM}\rangle + \langle G^{polar}_{solvation}\rangle + \langle G^{nonpolar}_{solvation}\rangle$) of the complex along the molecular dynamics trajectories. The horizontal lines are the mean values after 140 ps. Snapshots at 4 ps intervals were used to calculate the effective energy.

and, therefore, we did not include hGHbp P106 in our alanine scanning calculation. We chose hGHbp I105, hGHbp I165, and hGHbp V171 as typical nonpolar residues, and hGHbp R43 and hGHbp E44 as examples of charged residues. As controls, we se-

Table 1. The Components of the Binding Free Energy of the Wild-Type Complex Calculated from the Trajectory.

Contribution	Complex	hGHbp	hGH
$E_{\text{electrostatic}}$	−12635.9 (7.8)	−5894.8 (5.2)	−6489.8 (6.1)
E_{vdW}	−1612.6 (3.9)	−755.2 (2.7)	−690.2 (2.8)
E_{internal}	5672.3 (5.5)	3013.4 (4.9)	2658.9 (3.5)
$G^{\text{polar}}_{\text{solvation}}$	−6400.9 (6.7)	−3472.0 (4.9)	−3314.6 (4.8)
$G^{\text{nonpolar}}_{\text{solvation}}$	106.6 (0.1)	65.4 (0.1)	60.3 (0.1)
G_{subtotal}	−14870.5 (5.1)	−7043.1 (4.4)	−7775.4 (3.5)
	Binding [Complex—(hGHbp + hGH)]		
$\Delta E_{\text{electrostatic}}$	−251.4 (2.6)		
ΔE_{vdW}	−167.1 (0.8)		
$\Delta G^{\text{polar}}_{\text{solvation}}$	385.6 (2.4)		
$\Delta G^{\text{nonpolar}}_{\text{solvation}}$	−19.1 (0.1)		
$\Delta G_{\text{subtotal}}$	−52.0 (1.2)		

The data presented in this table are mean values with the standard deviation of the mean in parentheses. The detailed explanation of the symbols is in the text.

$$G_{\text{subtotal}} = E_{\text{electrostatic}} + E_{\text{vdW}} + E_{\text{internal}} + G^{\text{polar}}_{\text{solvation}} + G^{\text{nonpolar}}_{\text{solvation}}$$

$$\Delta G_{\text{subtotal}} = \Delta E_{\text{electrostatic}} + \Delta E_{\text{vdW}} + \Delta G^{\text{polar}}_{\text{solvation}} + \Delta G^{\text{nonpolar}}_{\text{solvation}}$$

For all the tables, the units of free energies and potential energies are kcal/mol and the standard deviation of the mean is defined as $\sigma/\sqrt{n}$, where n is the number of snapshots.

lected residues exhibiting a loss in side-chain solvent accessibility larger than 10 Å^2 (refer to Table 1 in ref. 15). We also performed the $\Delta\Delta G$ calculations on the hormone side for the alanine mutations causing more than an eightfold reduction in binding affinity reported in ref. 16.

To investigate the direct contacts between the chosen mutation sites and the binding interface, we calculated the burial percentage of each mutant residue including backbone atoms upon binding according to:

$$\%\text{Burial} = 100\% - \text{Area}_{\text{complex}}/\text{Area}_{\text{unbound}} \times 100\%.$$

The results, presented in Table 2, demonstrate that hGHbp R43, hGHbp W104, hGHbp S124, hGHbp D126, hGHbp W169, and hGH F176 were completely buried inside the protein–protein interface after binding. In contrast, hGHbp R70, hGHbp W80, and hGH I58 have few contacts with their binding partners, and thus alanine mutation of these residues will have little effect on the binding affinity if the mutations do not cause significant global or local conformational changes. We can, therefore, consider them as controls for the simulation.

Figure 3(a)–(c) shows details of the intermolecular contacts between the chosen mutant residues and their binding partners. hGHbp W104 is surrounded by several charged and polar residues at interface including hGH K168, hGH D171, hGH K172, and hGH T175. While the occupancy of the hydrogen bond between hGHbp W104 (indole HN) and hGH D171 (side chain O) is as high as 96% along the MD trajectory, the hydrogen bond between hGHbp W104 (O) and hGH K168 (side chain HN) is not very stable throughout the simulation. There is a salt bridge between hGHbp R43 and hGH D171 (77% occupancy) and a hydrogen bond between hGHbp R43 (side chain HN) and hGH T175 (side chain O) (95% occupancy). Figure 3(b) reveals that hGHbp W169 is buried in the middle of hGH R64 and hGHbp R43. It is also supported by hGHbp I103. hGHbp E44 forms a salt bridge with hGH R64. Its occupancy is 79% along the time series of the simulation. hGHbp V171 closely contacts with hGH R178 by vdW interactions while it is in loose contact with hGH T175. The partially buried hGHbp R217 forms a salt bridge with hGH E174. Another hydrogen bond connects the HN$^{\varepsilon}$ of hGHbp R217 and the O$^{\varepsilon 1}$ of hGH E174. hGHbp R217 also stacks with the imidazole ring of hGH H18.

We present the $\Delta\Delta G_{\text{binding}}$ data from the computational alanine scanning in Table 2. For the alanine mutations of charged/polar residues involved in buried salt bridges/hydrogen bonds, for example, hGHbp R43A, hGHbp D164A, hGHbp R217A, and hGHbp S219A, the calculated $\Delta\Delta G_{\text{binding}}$ is significantly overestimated using the single trajectory mutation protocol. For hydrophobic residues and polar/charged residues without buried salt bridges, however, the average unsigned error of our calculated $\Delta\Delta G_{\text{binding}}$ is only ~1 kcal/mol. The calculated $\Delta\Delta G_{\text{binding}}$ vs. the experimental data is shown in Figure 4, and the square of the correlation coefficient between the calculated and experimental $\Delta\Delta G_{\text{binding}}$ is 0.5 for 24 out of 29 mutations. We consider hGHbp D164A, hGHbp R217A, hGHbp S219A, hGH R64A, and hGH D171A to be outliers. We find that hGHbp W104A and hGHbp W169A cause the two largest reductions in binding affinity for these 24 mutations.

To gain further insight into the contributions of each component to the binding free energy changes, we present the change of the intermolecular vdW and electrostatic interactions and the polar and nonpolar solvation free energies upon alanine mutation in Figure 5a. In most cases the change of intermolecular electrostatic interactions is seen to strongly anticorrelate with the change of polar solvation free energy. For example, the intermolecular electrostatic interactions between the positively charged R43, R70, R71, and R217 of hGHbp and the negatively charged hGH are strongly favored in the wild-type complex, but large desolvation penalties are paid. On the other hand, the intermolecular electrostatic interactions between the negatively charged E44, E75, E120, and E127 of hGHbp and hGH are unfavorable in the wild-type complex, but polar solvation free energies are favorable for these residues. One exception is hGHbp D164A, which loses favorable intermolecular electrostatic interactions and polar solvation free energy following alanine mutation. As a result, it is important to consider the change of intermolecular electrostatics and polar solvation free energy together. The contributions of $\Delta\Delta E_{\text{vdW}}$ and $\Delta\Delta E_{\text{electrostatic}}$ plus $\Delta\Delta G^{\text{polar}}_{\text{solvation}}$ to $\Delta\Delta G$ are shown in Figure 5b, where most of $\Delta\Delta E_{\text{vdW}}$ are positive, which indicates that the alanine mutation for the residues reduces the vdW contacts at the binding interface. For hGHbp W104A and hGHbp W169A, the magnitude of $\Delta\Delta E_{\text{vdW}}$ is nearly twice the magnitude of the sum of $\Delta\Delta E_{\text{electrostatic}}$ and $\Delta\Delta G^{\text{polar}}_{\text{solvation}}$. As expected, hGHbp W76, hGHbp W104, and hGHbp W169 contribute to the total binding free energy mainly by vdW contacts. The variance of the nonpolar solvation free energy upon alanine mutation is very small compared to the changes of other free energy components. It is encouraging that our calculations reproduce the $\Delta\Delta G_{\text{binding}}$ for all the tryp-

Table 2. $\Delta\Delta G_{\text{binding}}$ of Alanine Scanning of hGH:hGHbp Using the Protocol of Trajectory Mutation.

Protein	Mutation	Loss in Side-Chain SASA	%Burial	$\Delta\Delta G^{\text{Calc}}_{\text{mut-wt}}$	$\Delta\Delta G^{\text{Exp.}}_{\text{mut-wt}}$
hGHbp	R43A(sD171, hT175)	30.9	100.0	13.98 (0.51)	2.12
	R43A[a]	30.9	100.0	2.60	2.12
	E44A(sR64)	34.0	70.5	−0.19 (0.20)	1.69
	R70A	1.1	0.0	0.54 (0.07)	0.69
	R71A(sE56)	80.7	93.4	−1.00 (0.52)	0.54
	E75A	22.8	13.0	−0.81 (0.17)	−0.10
	W76A(hL45)	57.6	61.5	0.70 (0.21)	0.51
	W80A	4.2	7.8	−0.50 (0.11)	−0.02
	S98A	15.3	16.0	−0.19 (0.07)	−0.05
	S102A	22.3	80.3	−0.75 (0.16)	−0.2
	W104A(hD171)	135.3	99.6	5.81 (0.38)	>4.5
	I105A	13.0	57.2	0.64 (0.1)	1.94
	E120A	34.6	46.7	−1.94 (0.22)	−0.19
	K121A	12.0	45.9	−1.42 (0.16)	0.08
	S124 A	12.7	100.0	−2.25 (0.21)	0.28
	D126A	16.6	100.0	2.03 (0.23)	0.99
	E127A(sR167, sK41)	65.9	77.0	−0.99 (0.57)	0.97
	D164A(sR64, hQ68)	16.7	96.7	31.05 (0.94)	1.49
	I165A	18.7	47.3	1.01 (0.22)	2.13
	K167A	37.5	55.8	0.21 (0.23)	−0.02
	W169A	88.7	100.0	3.47 (0.21)	>4.5
	V171A	43.2	86.5	0.68 (0.22)	−0.64
	R217A(sE174)	46.9	75.1	7.32 (0.26)	0.28
	N218A	102.9	95.3	2.7 (0.33)	0.3
	S219A(hD171)	20.9	94.5	6.55 (0.23)	0.03
hGH	I58A	0.0	0.0	0.00 (0.06)	1.7
	R64A(sE44, sD164)	130.3	71.3	5.37 (0.26)	1.8
	D17lA(sR43, hW104, hS219)	62.4	99.0	7.03 (0.63)	1.2
	K172A	54.6	78.6	2.73 (0.16)	1.6
	F176A	6.2	100.0	−1.04 (0.14)	1.6

[a] The experimental $\Delta\Delta G_{\text{mut-wt}}$ and loss in side-chain solvent accessibility for the residues on the receptor side are taken from Table 1 in ref. 15. For the mutations on the hormone side, the experimental $\Delta\Delta G_{\text{mut-wt}}$ is from radioimmunoassay data.[16, 17] The standard deviation of the calculated mean $\Delta\Delta G_{\text{binding}}$ is presented in the bracket. R43A[a] represents the result calculated from the trajectory of the mutant complex R43A. The standard deviation of the mean $\Delta\Delta G_{\text{binding}}$ is not shown in this case because it is not applicable for the separate trajectories of the wild-type and mutant complexes. %Burial, as defined in the text, is calculated from the average structure over the 400-ps production trajectory of the wild-type complex. The salt bridges and the hydrogen bonds between the residues subject to alanine mutation and their binding partners are denoted by s and h, respectively.

tophans (W76, W80, W104, W169) on the receptor side with an average deviation of 0.8 kcal/mol. This is despite the fact that a balance between electrostatic and vdW interactions (Fig. 5a–b), which are an order of magnitude larger than this deviation, must be achieved and that the magnitude of each component except $\Delta\Delta G^{\text{nonpolar}}_{\text{solvation}}$ is much larger than that of the net free energy differences for hGHbp W76A, hGHbp W104A, and hGHbp W169A. This suggests a good balance between the molecular mechanical energies and solvation free energies. The calculated effects of mutating the three hydrophobic residues, I105, I165, and V171 of hGHbp, are calculated in qualitative agreement with experiment (∼1 kcal/mol unsigned error), although hGHbp V171A is predicted to be destabilizing rather than stabilizing.

Finally, calculating $\Delta\Delta G_{\text{binding}}$ for the charged residues is most challenging, because both $|\langle E_{\text{electrostatic}}\rangle|$ and $|\langle \Delta G^{\text{polar}}_{\text{solvation}}\rangle|$ change considerably (10–100 kcal/mol) upon mutating from a charged/polar to a neutral residue. This balance seems to be reasonably achieved for E44A, R70A, R71A, E75A, S98A, S102A, E120A, K121A, S124A, D126A, E127A, K167A, and N218A of hGHbp and hGH K172, where the calculated $\Delta\Delta G_{\text{binding}}$ are within ∼2 kcal/mol of experiment. However $\Delta\Delta G_{\text{binding}}$ is significantly overestimated for R43A, D164A, R217A, and S219A of hGHbp and R64A and D171A of hGH.

Recalculating $\Delta\Delta G_{\text{binding}}$ of hGHbp R43 Mutated to Alanine from the Trajectory of R43A (Separate Trajectory Protocol)

Our overestimate of the $\Delta\Delta G_{\text{binding}}$ for the polar/charged residues involved in buried salt bridges/hydrogen bonds suggests that our assumption of minimal conformational change upon alanine mutations is less justified in these cases. Because of the large size of

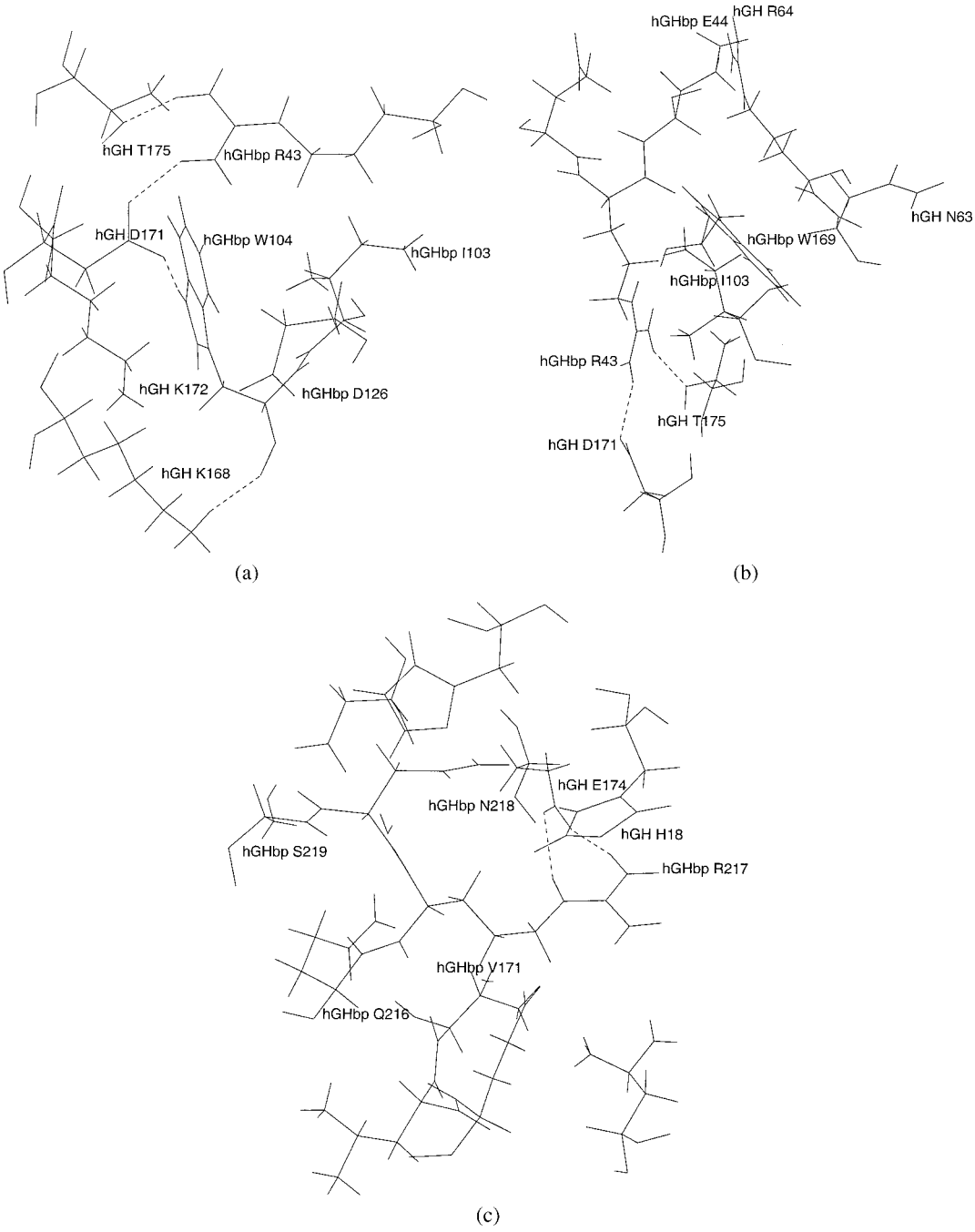


Figure 3. The details of intermolecular contacts between the residues subject to alanine mutation and the hGH interface. The dashed lines denote hydrogen bonds.

the system and the considerable computational cost, we chose to do further calculations only on hGHbp R43A, which has one of the largest errors in calculated $\Delta\Delta G_{\text{binding}}$, and is one of the most fully buried charged residues at the interface. Again, because of the computational cost, we ran a more limited trajectory on this mutant, 240 ps total, with 200 ps for production as noted in the Method section. We collected a new trajectory starting from the same wild-type crystal structure as before, but with the side chain of hGHbp R43 modeled as alanine. The C_α rms deviation from the crystal structure as a function of the simulation time is shown in Figure 6. Although this mutant complex is not as stable as the wild type, the C_α rms deviation of the interface remains very small. As shown in Figure 2b, the effective energy along the MD simulation of the mutant complex reaches reasonable convergence by 140 ps

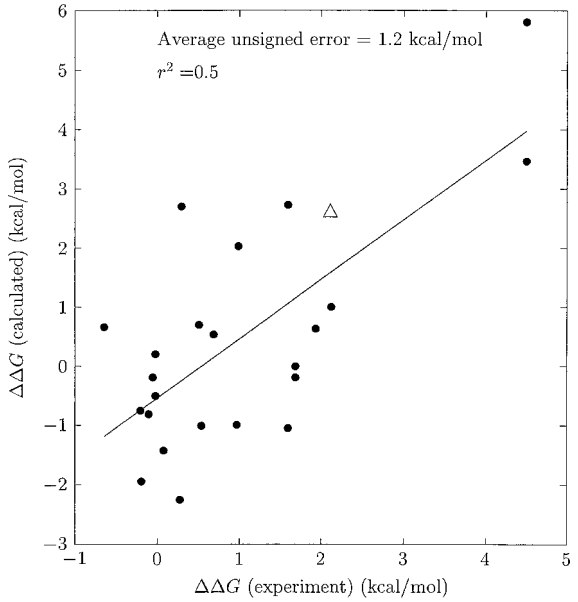
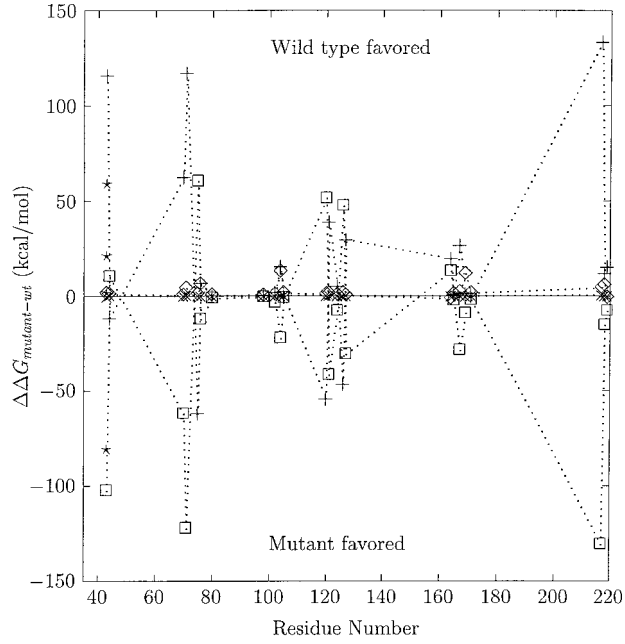


Figure 4. The experimental results vs. the calculated $\Delta\Delta G_{\text{binding}}$. • Represents the $\Delta\Delta G_{\text{binding}}$ for alanine mutations in both hGH and hGHbp, except R43A, D164A, R217A, and S219A in hGHbp and R64A and D171A in hGH. Δ represents $\Delta\Delta G_{\text{binding}}$ for R43A recalculated from the trajectory of R43A.

such that we chose conformations generated after 140 ps at 4 ps intervals for the free energy analyses.

In Figure 7, we present rms deviations between the average structure of the trajectory of mutant complex and the average structure of the trajectory of the wild-type complex for each residue within 10 Å of hGHbp W104, the center of the binding interface. Figure 8a–c illustrates the prominent deviations of the 79 total residues in this region, and Figure 7 shows that for around 77% of the residues at the interface of R43A, the rms deviations from the average wild type structure are less than 1.0 Å. Large movements of the side chains are quite localized, concentrated in the conformation and sequence vicinity of hGHbp S219, which is beneath hGHbp R43 in the wild-type structure. After the R43A mutation, hGHbp S219 shifts to make the hydrogen bond connecting to hGH D171 stronger. The occupancy of this hydrogen bond along the trajectory of the mutant complex is 99% (∼±10%), while only 76% (∼±10%) occupancy is found for the wild-type complex. Accompanying this shift, the aromatic ring of hGH F25 rotates to enhance vdW contacts with hGHbp S219. The salt bridge between hGH R64 and hGHbp E44 is located on the surface in both the wild-type and mutant complexes, but in the average mutant complex, these residues move apart, weakening the strength of the salt bridge. While the occupancy of the salt bridge over the trajectory of the wild-type complex is 79% (∼±10%), it is lowered to 61% (∼±10%) for the mutant complex. Compared with the average wild-type structure, the backbone of hGHbp Q216–G220 changes slightly in the average mutant complex. Neither hydrogen bonds nor vdW stacking between hGHbp Q216 and hGH is found in either the wild-type or mutant complexes. The movement of hGHbp R217 in the mutant complex breaks the vdW stacking with hGH



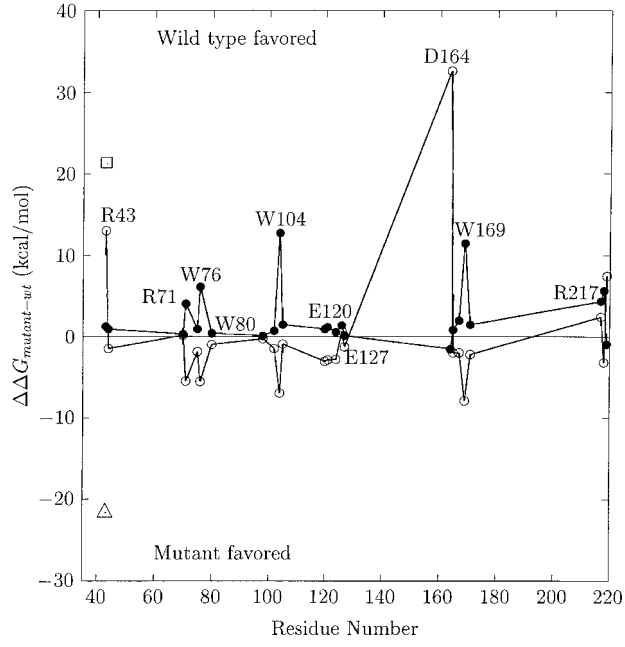


Figure 5. (a) Decomposition of $\Delta\Delta G_{\text{binding}}$ into $\Delta\Delta E_{\text{vdW}}$ (◊), $\Delta\Delta E_{\text{electrostatic}}$ (+), $\Delta\Delta G^{\text{polar}}_{\text{solvation}}$ (□), and $\Delta\Delta G^{\text{nonpolar}}_{\text{solvation}}$ (×). ∗ Represents the results recalculated from the trajectory of the mutant complex R43A. From top to bottom, the data points are $\Delta\Delta E_{\text{electrostatic}}$, $\Delta\Delta E_{\text{vdW}}$, $\Delta\Delta G^{\text{nonpolar}}_{\text{solvation}}$, and $\Delta\Delta G^{\text{polar}}_{\text{solvation}}$ for R43A. (b) Decomposition of $\Delta\Delta G_{\text{binding}}$ into $\Delta\Delta E_{\text{vdW}}$ and $\Delta\Delta E_{\text{electrostatic}} + \Delta\Delta G^{\text{polar}}_{\text{solvation}}$. • Represents $\Delta\Delta E_{\text{vdW}}$. ○ Represents $\Delta\Delta E_{\text{electrostatic}} + \Delta\Delta G^{\text{polar}}_{\text{solvation}}$. □ and Δ represent $\Delta\Delta E_{\text{vdW}}$ and $\Delta\Delta E_{\text{electrostatic}} + \Delta\Delta G^{\text{polar}}_{\text{solvation}}$ calculated from the trajectory of the mutant complex R43A.

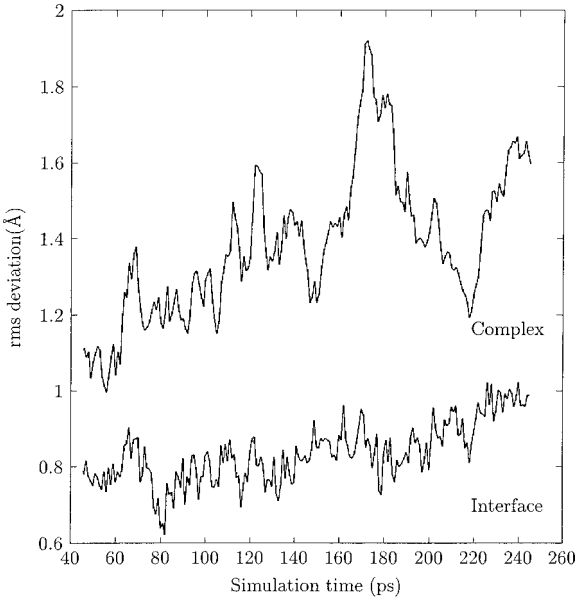


Figure 6. C_α rms deviation of the mutant complex R43A and of the interface as a function of simulation time. The data regions include 200 ps production dynamics, and the equilibration phase is not included.

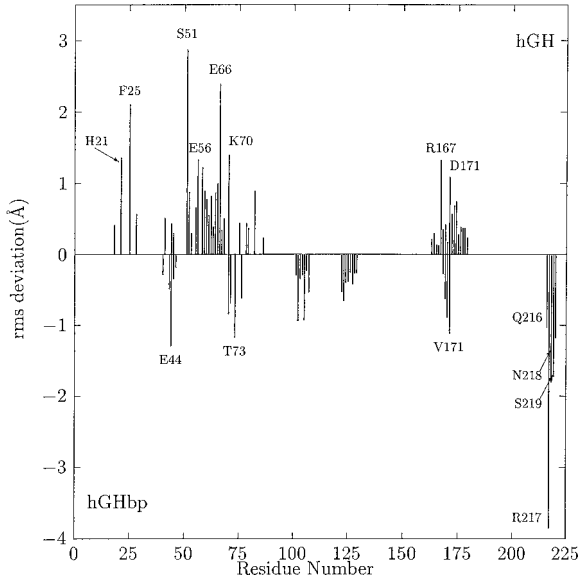
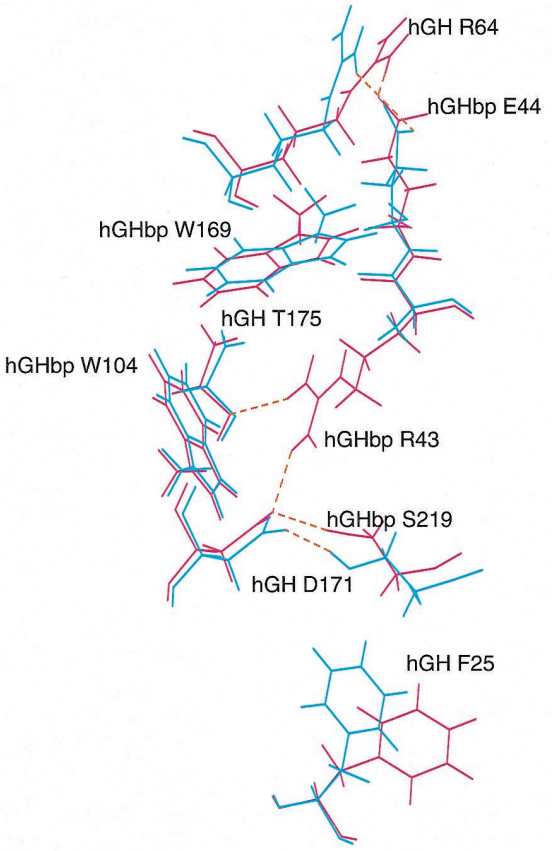


Figure 7. hGHbp W104 is chosen as the center of the binding interface. For all residues within 10 Å of it, the rms deviations between the average mutant structure of R43A and the average wild-type structure over the MD trajectories are analyzed.

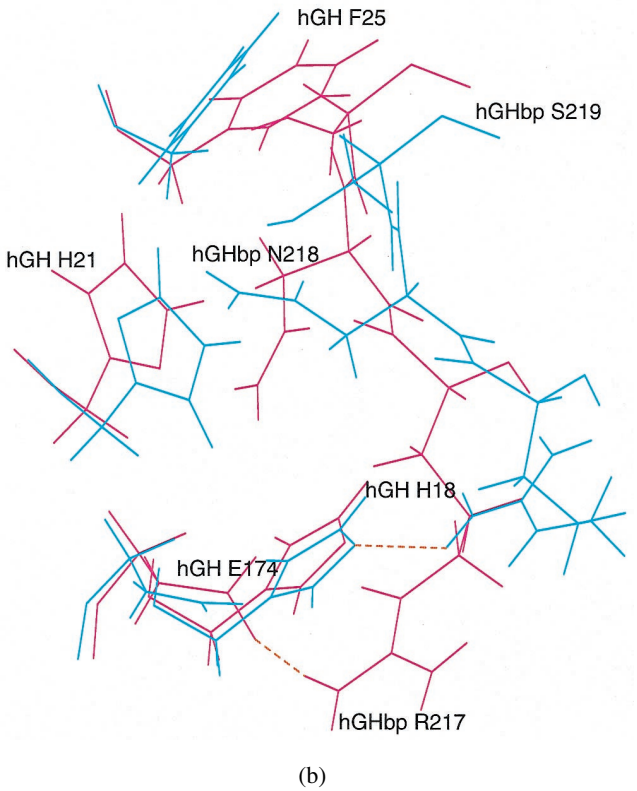
H18 and the salt bridge connecting hGH E174, but a new hydrogen bond concomitantly forms between hGHbp R217 (side chain HN) and hGH H18 (side chain N). hGHbp R217 also causes the nearby V171 in hGHbp to move with it. The vdW contacts between hGHbp N218 and hGH F25 and hGH H21 in the wild-type structure are not significantly changed after the mutation. hGH R167 is originally at the edge of the binding interface, but it moves inward towards the interface in the average mutant complex. Accordingly, this movement destroys the salt bridge with hGHbp E127 and forms a new one with hGHbp D126, which is beside hGHbp W104. Although hGH S51 and hGH E56 are not neighbors of hGHbp S219, their movements also contribute to the change of the binding affinity. Originally, hGH S51 does not make any direct contacts with the receptor, but after the mutation, it moves toward the interface to form a hydrogen bond between its carbonyl O and the $HN^{\eta 1}$ of hGHbp R71. The salt bridge between hGH E56 and hGHbp R71 near the surface of the wild-type complex becomes slightly stronger in the average mutant complex structure due to a simultaneous shift of the two side chains. The occupancy of the salt bridge over the trajectory of the mutant complex is 96% ($\sim\pm10\%$), compared to 89% ($\sim\pm10\%$) over the trajectory of the wild-type structure. The movements of hGH E66, hGH K70, and hGHbp T73 make only indirect contributions to the change of binding affinity, because none of them make direct contacts with their binding partners in either the mutant or wild-type complexes.

In short, due to side-chain reorganizations after the mutation, the complex loses not only its original hydrogen bond network between the hormone and R43 of hGHbp, but also other intermolecular electrostatic interactions and vdW contacts; however, it also gains some contacts. To further investigate the change in the hydrogen bond pattern, we analyzed all the hydrogen bonds between the hormone and its receptor from the trajectories of the wild-type structure and of the mutant complex as shown in Figure 9. All the hydrogen bonds shown in Figure 9 are those with occupancy of greater than 60%. The total number of the hydrogen bonds is the same in both the wild-type and mutant complexes; however, only 10 out of 16 hydrogen bonds involve the same hydrogen bond donor and acceptor.

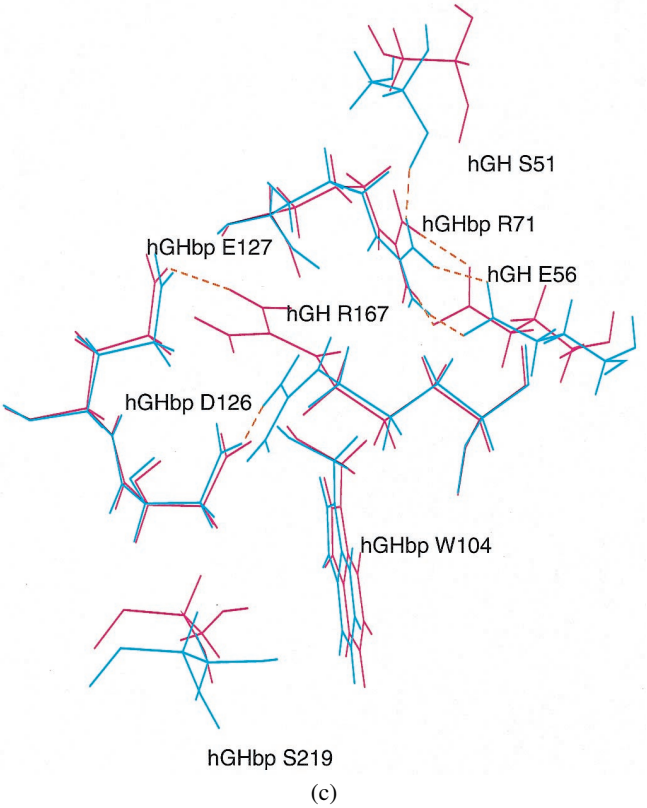
We next recalculated $\Delta G_{\text{binding}}$ from the trajectory of the mutant complex (in Table 3). Compared to the wild-type complex, the intermolecular vdW and electrostatic interactions between hormone and its mutant receptor become weaker, but the electrostatic desolvation penalty is also smaller. The smaller nonpolar desolvation free energy of the mutant complex implies that less binding interface is buried in the mutant complex than that in the wild type. The recalculated $\Delta\Delta G_{\text{binding}}$, denoted as a triangle in Figure 4, is in good agreement with the experimental result. We also decompose the recalculated $\Delta\Delta G_{\text{binding}}$ into $\Delta\Delta E_{\text{vdw}}$ and $\Delta\Delta E_{\text{electrostatic}}$ combined with $\Delta\Delta G^{\text{polar}}_{\text{solvation}}$ (Fig. 5). Unlike the previous result based on the first order approximation, $\Delta E_{\text{electrostatic}}$ plus $\Delta\Delta G^{\text{polar}}_{\text{solvation}}$ is now negative, suggesting that from an electrostatic point of view it is unfavorable to bury a salt bridge inside the interface. To decouple the contribution of vdW and electrostatics, Wells and coworkers mutated hGHbp R43 to leucine and found that the binding free energy of R43L was only 0.52 kcal/mol higher than that of the wild-type complex.[15] Because Leu has nearly the same steric shape as the alkyl portion of Arg, but does not have a charged head group, it was reasoned that hGHbp R43 contributed to binding affinity mainly by vdW contacts rather than electrostatic interactions.[15] Our calculations suggest that the situation is more complicated. The intermolecular vdW contribution to the $\Delta\Delta G_{\text{binding}}$ of hGHbp R43A is a large favorable value, and the intermolecular electrostatic interactions plus the electrostatic desol-



(a)



(b)



(c)

Figure 8. (continued)

Figure 8. Comparison of the side-chain conformations of the average wild-type complex (magenta) and the average R43A (cyan) complex over the trajectories. The dashed lines denote hydrogen bonds.

vation term is almost equal to the intermolecular vdW contribution but negative, such that they offset each other. As a result, the primary net contribution to the $\Delta\Delta G_{\text{binding}}$ of hGHbp R43A is the change in nonpolar solvation free energy, about 1.9 kcal/mol, corresponding to ~350 Å^2 less buried surface area at the interface in the mutant complex than that of the wild type.

Assumptions about Conformational Changes

Experimental evidence suggests that alanine mutations in the hGH:hGHbp complex have only a small disturbance in the protein fold and its structural characteristics.[13] For instance, the mutant proteins can still be expressed; most of the mutant receptors bind to the antireceptor mAb, and the CD spectra of hGHbp W104A are similar to those of the wild type. Therefore, the assumption that the global conformation does not change after alanine mutation is quite reasonable. As to the assumption of the absence of local conformational changes, in our single trajectory alanine scanning approach, we assume that the local side-chain rearrangements after alanine mutation do not significantly affect the binding affinity. This is, for many cases, a good approximation. For example, excellent agreement with the experimental results was obtained for alanine

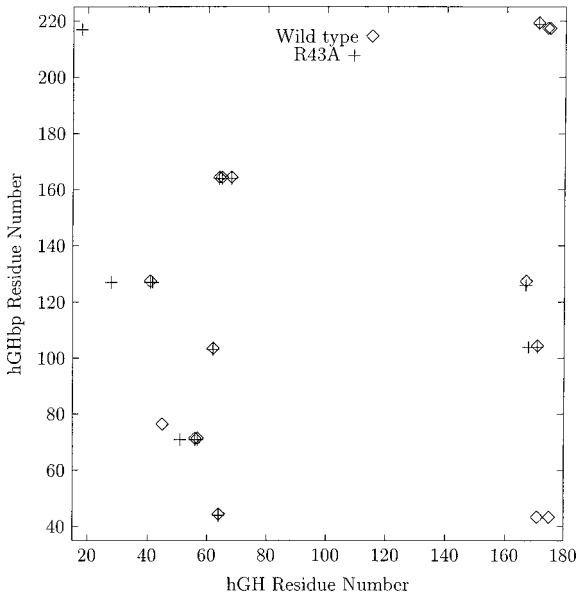


Figure 9. The hydrogen bond pattern at the hormone–receptor interface for the wild-type and R43A complexes.

scanning of p53-Mdm2, a small peptide bound to a protein.[10] Because the interface of this system is small, in the absence of global conformational changes upon binding, the reorganization of the local side chains to compensate the effect of alanine mutations seems insignificant. For the large protein–protein interface we have investigated here, the calculated $\Delta\Delta G_{\text{binding}}$ for all tryptophans are in excellent agreement with experimental data, and except the control, hGHbp W80A, we found that the intermolecular vdW interactions dominated the binding for all Trps at the interface. This implies that if a large hydrophobic side-chain contacts its binding partner via vdW interactions, it is difficult to adjust the stacking pattern of nearby side chains to compensate for the loss of these intermolecular vdW interactions after alanine mutation. The situation is different for charged/polar residues buried at the interface, such as R43A, D164A, R217A, and S219A in hGHbp, for which the assumption of no local side chain rearrangements fails.

Another challenging situation is that there are probably some "false negatives" for the alanine mutations of the residues that have small direct contacts with their binding partners, but which cause conformational changes after being mutated to alanine, for example, hGHbp I165 and hGH I58. The burial percentage for hGHbp I165 upon binding is 47.3%, and it makes vdW contacts with hGH C189 while being loosely packed with C182 and R178 in hGH. But hGHbp I165 makes a large contribution to the binding (2.13 kcal/mol according to the experimental data[15]). Our calculated $\Delta\Delta G_{\text{binding}}$ underestimates this by ~1 kcal/mol.

Table 3. The Components of the Binding Free Energy of the R43A Complex Calculated from the R43A Trajectory.

Contribution	Complex	hGHbp	hGH
$E_{\text{electrostatic}}$	−12333.7 (19.6)	−5710.8 (12.3)	−6437.4 (11.2)
E_{vdW}	−1619.6 (6.4)	−768.9 (4.3)	−704.7 (4.6)
E_{internal}	5671.2 (9.9)	3017.7 (7.9)	2653.4 (6.7)
$G^{\text{polar}}_{\text{solvation}}$	−6499.5 (21.0)	−3457.0 (10.6)	−3341.8 (12.3)
$G^{\text{nonpolar}}_{\text{solvation}}$	107.5 (0.2)	64.3 (0.1)	60.4 (0.1)
G_{subtotal}	−14674.2 (6.6)	−6854.7 (5.9)	−7770.1 (7.0)
	Binding [Complex—(hGHbp + hGH)]		
$\Delta E_{\text{electrostatic}}$	−185.5 (5.8)		
ΔE_{vdW}	−146.0 (1.4)		
$\Delta G^{\text{polar}}_{\text{solvation}}$	299.3 (5.7)		
$\Delta G^{\text{nonpolar}}_{\text{solvation}}$	−17.2 (0.1)		
$\Delta G_{\text{subtotal}}$	−49.4 (1.6)		

The data presented in this table are mean values with the standard deviation of the mean in brackets. The detailed explanation of the symbols is in the text.

Comparison of Single Trajectory Mutation Protocol and Minimization Mutation Protocol

A less expensive method to screen site-directed mutagenesis has been proposed by Reyes and Kollman,[26] and uses simple energy minimization rather than full MD simulation. They found good agreement with experimental data when mutations involved residues of similar size; none were W → A or R → A. To test whether this minimization protocol is suitable for alanine mutation of large side chains, we studied alanine mutations of W76, W80, W104, and W169 on the receptor side. The alanine mutation was performed on the crystal structure, and then a series of minimizations were carried out on both the mutant and wild-type complexes.[26] As seen in Table 4, the protocol of alanine mutation based on minimization is not able to discriminate the "hot spots" of binding free energy (hGHbp W104 and hGHbp W169) from the non-"hot spot" (hGHbp W76) and the control (hGHbp W80). hGHbp W169 even makes unfavorable contributions to the total binding free energy. To further understand why the alanine mutation protocol based on minimization failed, we list the components of the binding free energy calculated from the minimized crystal structure in Table 5. The intermolecular vdW interaction is more favorable in the minimized structure than that in the structures sampled by the MD simulation at 300 K. The intermolecular electrostatic energy of the minimized structure, however, is much more unfavorable than that obtained from the representative structures sampled by the MD simulation. These results demonstrate that the bad intermolecular vdW contacts are easily released upon minimization, while the intermolecular electrostatic interactions seem to require a longer time to relax (Table 5). In addition, the variations in the electrostatic energy during the MD simulations may be meaningful, but are not captured in the single-point minimization protocol.

Conclusions

The advantage of the single trajectory alanine scanning strategy is that by postprocessing only one MD trajectory of the wild-type complex, we can reproduce the experimental $\Delta\Delta G_{\text{binding}}$ with the

Table 4. Comparison of $\Delta\Delta G_{\text{binding}}$ of Alanine Scanning for the Protocols of Trajectory Mutation and Minimization.

	$\Delta\Delta E_{\text{vdW}}$		$\Delta\Delta G^{\text{polar}}_{\text{solvation}} + \Delta\Delta E_{\text{ele}}$		$\Delta\Delta G^{\text{nonpolar}}_{\text{solvation}}$		$\Delta\Delta G_{\text{total}}$		$\Delta\Delta G_{\text{total}}$
Mutation	Traj.	Min.	Traj.	Min.	Traj.	Min.	Traj.	Min.	Exp.
W76A	6.14	7.46	−5.51	−4.13	0.07	0.15	0.70	3.48	0.51
W80A	0.41	−0.44	−0.96	4.63	0.05	0.01	−0.50	4.20	−0.02
W104A	12.73	15.88	−6.96	−11.53	0.04	0.01	5.81	4.36	>4.5
W169A	11.48	11.68	−7.86	−11.76	−0.15	−0.12	3.47	−0.20	>4.5

average unsigned error ∼1 kcal/mol for the alanine mutations of hydrophobic residues and polar/charged residues without buried salt bridges. This is particularly useful when one considers a very large protein system such as hGH:hGHbp, which, when explicit water is included, involves ∼75,000 atoms. It is also encouraging that the simulation, carried out over nearly half a nanosecond, was quite stable (average rms deviation of C_α ∼ 1.4 Å), which is probably due in no small part to our use of the accurate long-range electrostatics afforded by PME.[5] For the alanine mutation of large side chains, the sampling afforded by such an MD simulation is essential to obtain the accurate $\Delta\Delta G_{\text{binding}}$, because simple minimization is less able to accurately depict the $\Delta\Delta G_{\text{binding}}$ in those cases. However, when one wishes to analyze the effect of mutating a buried salt bridge, it is of importance to include structural relaxation. We have done an additional MD simulation on the hGHbp R43A mutant of hGH:hGHbp to suggest how the protein interface undergoes significant relaxation upon mutation and to demonstrate that this leads to a much better description of $\Delta\Delta G_{\text{binding}}$. If one did not know the experimental data, how would one know if an additional simulation was warranted? Based on our experience here and previously,[10] we expect that alanine scanning that both involves a buried charged residue and has a calculated $\Delta\Delta G_{\text{binding}} > 5$ kcal/mol warrants a separate trajectory if an accurate estimate of $\Delta\Delta G_{\text{binding}}$ is important.

Table 5. The Components of the Binding Free Energy of the Wild-Type Complex Calculated from the Minimized Crystal Structure.

Contribution	Complex	hGHbp	hGH
$E_{\text{electrostatic}}$	−11821.8	−5585.7	−6129.1
E_{vdW}	−2541.5	−1202.7	−1146.5
E_{internal}	2458.0	1334.8	1123.2
$G^{\text{polar}}_{\text{solvation}}$	−6062.6	−3236.1	−3157.6
$G^{\text{nonpolar}}_{\text{solvation}}$	103.7	62.9	58.1
G_{subtotal}	−17863.4	−8626.8	−9251.9
	Binding [Complex—(hGHbp + hGH)]		
$\Delta E_{\text{electrostatic}}$	−107.0		
ΔE_{vdW}	−192.3		
$\Delta G^{\text{polar}}_{\text{solvation}}$	331.1		
$\Delta G^{\text{nonpolar}}_{\text{solvation}}$	−17.4		
$\Delta G_{\text{subtotal}}$	14.4		

It is clear that the results of our calculations support the idea of considerable plasticity at the hGH:hGHbp interface,[14] in that significant structural changes occur upon limited mutation. As noted above, for the p53 helix interacting with Mdm2, single-trajectory alanine scanning calculations were in very good agreement with qualitative experiments on the $\Delta\Delta G$,[10] even for charged residues that were mainly solvent exposed. Thus, for p53–Mdm2, the interface is quite well defined and rigid. On the other hand, for hGH:hGHbp, with the exception of the four tryptophans, the agreement with experiment for the single-trajectory alanine scanning $\Delta\Delta G$ is qualitative at best. This is particularly evident for residues interacting with other charged residues (R43, D164, R217, and S219 on the receptor side and R64 and D171 on the hormone side, Table 2) where the single trajectory alanine scanning $\Delta\Delta G$s are significantly overestimated. Separate trajectories for the wild type and R43A show that one can simulate the observed free energy much better in this way, because it allows the structural relaxation that accompanies hGHbp R43A mutation. For the other polar and hydrophobic residues (excluding the tryptophans), single-trajectory alanine scanning leads to an average error of ∼1.3 kcal/mol. There are, however, a number of sign differences between the calculated and experimental data, suggesting that subtle conformational readjustments are likely to be important in accurate calculation of $\Delta\Delta G$ in such cases. Perhaps the single-trajectory alanine scannning method works best for the tryptophans because they tend to be in the most preorganized part of the binding site, such that a W → A mutation, which causes relatively little structural relaxation. Further calculations on various separate trajectory mutations in the hGH:hGHbp complex may shed farther light into the nature of structural relaxation in a more general way than our limited study of R43A, where encouraging agreement with experiment was obtained.

We have used fully solvated trajectories with periodic boundary conditions and PME to generate snapshots for MM-PBSA, even though the waters are discarded in the free energy analysis. We did this to determine how quantitatively one could calculate $\Delta\Delta G_{\text{binding}}$ for this "classic" alanine scanning application of hGH:hGHbp. Of course, it would have been more efficient to include only waters near the interface and keep the protein further away rigid. Even more efficient would have been to use a GB model to create the trajectory.[28] At this point, we do not know how much the calculated $\Delta\Delta G_{\text{binding}}$ would have been affected using such methods to generate the trajectory; that is an issue for subsequent studies.

The alanine scanning approach presented here can also be extended to other "scanning," for example, fluorine scanning of C—H groups.[27] The strategy of combining MD simulations with a continuum solvation model to analyze the binding free energies thus provides a number of exciting methodologies to evaluate the contributions of individual groups to overall binding free energies. These methods are encouragingly robust, usually demonstrating a satisfactory balance between intermolecular vdW and net electrostatic interactions and between intermolecular electrostatics, calculated with molecular mechanics, and solvation electrostatics, calculated by solving the Poisson–Boltzmann equation.

Acknowledgments

Molecular graphics images were produced using the MidasPlus package from the Computer Graphics Laboratory at University of California, San Francisco (supported by NIH P41RR-01081). We thank the National Center for Supercomputing Applications for providing computational resources. S. Huo is grateful to Y. Duan, W. Wang, B. Krueger, and W. Arnold for helpful discussions.

References

1. Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. J Am Chem Soc 1998, 120, 9401.
2. Gilson, M. K.; Honig, B. Proteins 1988, 4, 7.
3. Honig, B.; Nicholls, A. Science 1995, 268, 1144.
4. Sanner, M. F.; Olson, A. J.; Spehner, J. C. Biopolymers 1996, 38, 305.
5. Darden, T.; York, D.; Pedersen, L. J Chem Phys 1993, 98, 10089.
6. Brooks, B. R.; Janežic, D.; Karplus, M. J Comput Chem 1995, 16, 1522.
7. Jayaram, B.; Young, M. A.; Beveridge, D. L. J Am Chem Soc 1998, 120, 10629.
8. Vorobjev, Y.; Almagro, J. C.; Hermans, J. Proteins 1998, 32, 399.
9. Cheatham, T. E., III; Srinivasan, J.; Case, D. A.; Kollman, P. A. J Biomol Struct Dyn 1998, 16, 265.
10. Massova, I.; Kollman, P. A. J Am Chem Soc 1999, 36, 8133.
11. Böttger, A.; Böttger, V.; Garcia-Echeverria, C.; Chéne, P.; Hochkeppel, H.; Sampson, W.; Ang, K.; Howard, S. F.; Picksley, S. M.; Lane, D. P. J Mol Biol 1997, 269, 744.
12. Massova, I.; Kollman, P. A. Perspect Drug Disc Design 2000, 18, 1.
13. Bass, S. H.; Mulkerrin, M.; Wells, J. A. Proc Natl Acad Sci USA 1991, 88, 4498.
14. Atwell, S.; Ultsch, M.; De Vos, A. M.; Wells, J. A. Science 1997, 278, 1125.
15. Clackson, T.; Ultsch, M. H.; Wells, J. A.; de Vos, A. M. J Mol Biol 1998, 277, 111.
16. Cunningham, B. C.; Wells, J. A. Science 1989, 244, 1081.
17. Cunningham, B. C.; Wells, J. A. J Mol Biol 1993, 234, 554.
18. Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. A. Comput Phys Commun 1995, 91, 1.
19. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, J. K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179.
20. Jorgensen, W. L. J Chem Phys 1982, 77, 4156.
21. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. J Comput Phys 1977, 23, 327.
22. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. J Chem Phys 1984, 81, 3684.
23. Sitkoff, D.; Sharp, K. A.; Honig, B. J Phys Chem 1994, 98, 1978.
24. Wang, W.; Kollman, P. A. J Mol Biol 2000, 303, 567.
25. Reyes, C. M.; Kollman, P. A. J Mol Biol 2000, 297, 1145.
26. Reyes, C. M.; Kollman, P. A. J Mol Biol 2000, 295, 1.
27. Kuhn, B.; Kollman, P. A. J Am Chem Soc 2000, 122, 3909.
28. Williams, D. J.; Hall, K. B. Biophys J 1999, 76, 3192.