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

Bernd Kuhn and Peter A. Kollman*

Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143-0446

Received June 5, 2000

We report calculations of free energies of binding, ΔG_{bind}, between a diverse set of nine ligands and avidin as well as between a peptide and streptavidin using the recently developed MM/PBSA approach. This method makes use of a molecular dynamics simulation of the ligand−protein complex to generate a thermally averaged ensemble of conformations of the molecules that are involved in the complex formation. Based on this set of structures, a free energy of binding is calculated using molecular mechanical and continuum solvent energies as well as including estimates of the nonpolar solvation free energy and solute entropy. We compare in our simulations different classes of ligands, involving biotin derivatives, the dye 2-(4′-hydroxyazobenzene)benzoic acid (HABA), and a cyclic hexapeptide, which cover a large range of binding free energies from −5 to −20 kcal/mol. Our calculations are able to reproduce experimental ΔG_{bind} values with a very good correlation coefficient of $r^2 = 0.92$. This agreement is considerably better than the results obtained with an alternate approach, the linear interaction energy approximation, for this system ($r^2 = 0.55$).

Introduction

Recent computational developments in structure-based drug design promise to accelerate the process of finding a therapeutic agent for a given medicinal target.[1] Despite those advances, it remains a major methodological challenge to give an accurate theoretical prediction of ligand−receptor binding affinities with as little computational effort as possible. Most often in industrial applications, docking and scoring methods are used to evaluate binding affinities.[2,3] While those computationally inexpensive approaches are suitable for ranking ligands of large databases with the purpose of identifying lead compounds, they often lack an accurate representation of important energy contributions, such as protein flexibility, solvation, and entropy. In the stages of lead refinement where more computer time can be spent on a particular compound, it might be desirable to include those energy terms at a physically more rigorous level.

One interesting approach in this respect has been Åqvist's linear interaction energy (LIE) approximation.[4] In this method, free energies of binding are obtained from averages of the interaction energies between the ligand and its surroundings, which are the receptor in the bound state and water in the unbound state, using a total of two molecular dynamics (MD) simulations. Electrostatic interaction energies are evaluated in a linear response context, and van der Waals energies are usually scaled by an empirical factor, which is fitted to best reproduce the observed free energies of binding. This approach was applied to a number of protein−ligand associations with reasonable accuracy.[4,5]

Recently, Srinivasan et al.[6] proposed a new method to predict free energies of complex macromolecules, which was termed MM/PBSA (Molecular Mechanics/Poisson−Boltzmann Surface Area). An MD simulation (typically in explicit solvent) is first carried out which yields a representative ensemble of structures. The average total free energy of the system, G, is then evaluated as:

$$G = G_{\text{PB}} + G_{\text{np}} + E_{\text{MM}} - TS_{\text{solute}} \quad (1)$$

where G_{PB} is the polar solvation energy, which is computed in continuum solvent, usually using a finite-difference Poisson−Boltzmann (PB) model, and G_{np} is the nonpolar solvation term, which can be derived from the solvent-accessible surface area (SA). E_{MM} denotes the sum of molecular mechanical (MM) energies of the molecule and can be further divided into contributions from electrostatic (E_{es}), van der Waals (E_{vdW}), and internal (E_{int}) energies:

$$E_{\text{MM}} = E_{\text{es}} + E_{\text{vdW}} + E_{\text{int}} \quad (2)$$

The last term in eq 1, TS_{solute}, is the solute entropy and is usually estimated by a combination of classical statistical formulas and normal-mode analysis. Since the MD simulation generates a thermally averaged ensemble of molecular conformations, one uses an arithmetic rather than Boltzmann-weighted average in the calculation of the free energy G. Using eqs 1 and 2, the binding free energy of a noncovalent association, ΔG_{bind}, can be computed as:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}}) \quad (3)$$

The ensemble of structures for the uncomplexed reactants are generated either by running separate MD simulations for them or by using the trajectory of the complex and simply removing the atoms of the protein

* To whom correspondence should be addressed. Phone: (415) 476-4637. Fax: (415) 476-0688. E-mail: pak@cgl.ucsf.edu.

10.1021/jm000241h CCC: \$19.00 © 2000 American Chemical Society
Published on Web 09/09/2000

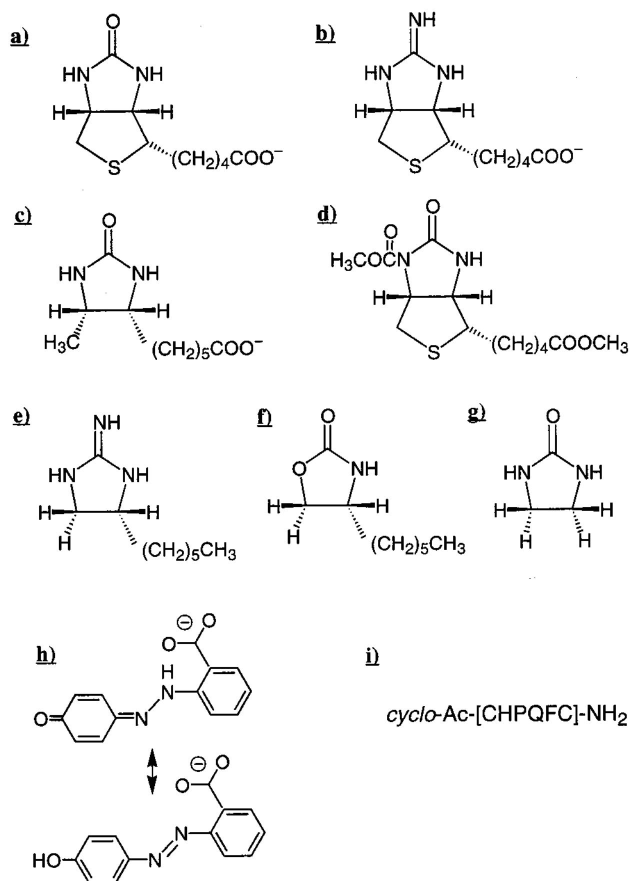

Figure 1. Chemical structures of the nine ligands investigated. Ligand names and abbreviations used throughout the text: (a) biotin (BTN1), (b) 2′-iminobiotin (BTN2), (c) desthiobiotin (BTN3), (d) 1′-*N*-methoxycarbonylbiotin methyl ester (BTN4), (e) D-4-*n*-hexyl-2-iminoimidazolidine (BTN5), (f) D-4-*n*-hexyloxazolidone (BTN6), (g) imidazolidone (BTN7), (h) 2-(4′-hydroxyazobenzene)benzoic acid (HABA) in its two tautomeric forms, (i) *cyclo*-Ac-[CHPQFC]-NH_2 (PEP).

or ligand, respectively. We applied the latter variant, and consequently, E_{int} of eq 2 cancels in the calculation of the free energy of binding. Apart from the calculation of binding affinities of noncovalent associations, MM/PBSA can be applied to a variety of other problems and has proven to be a powerful tool there, as reviewed recently.[7]

In this work, we use the MM/PBSA approach to examine the binding of a diverse set of nine ligands to avidin as well as of the disulfide-bridged cyclic hexapeptide *cyclo*-Ac-[CHPQFC]-NH_2 (PEP) to streptavidin. Those two proteins were chosen because of their good structural characterization by X-ray crystallography[8−11] and because of the very wide range of experimental binding free energies for a number of ligands,[12,13] which provide a good testing ground for our method. As can be seen from Figure 1, the ligands that we compare differ considerably in structure, size, and molecular charge and, moreover, cover a range of ΔG_{bind} from −5 to −20 kcal/mol. Our study is the first application of MM/PBSA to the binding of a diverse set of ligands to a protein.

Computational Details

The starting point for our simulations was the X-ray structure of biotin complexed with tetrameric egg-white avidin at 2.7 Å resolution.[10] Since the four biotin binding sites are separated at least 22 Å and are independent from each other, we considered only one site in our calculations while keeping the other three as static. We derived atomic partial charges for each ligand by semiempirical PM3 geometry optimization and subsequent single-point Hartree−Fock (HF)/6-31G* calculation of the electrostatic potential, to which the charges were fitted using the RESP[14] procedure. Test calculations for biotin and 2-(4′-hydroxyazobenzene)benzoic acid (HABA) with different semiempirical methods (MNDO, AM1, PM3) showed that the PM3/HF combination reproduced best charges that are based on fully HF optimized geometries. The overall charge of the protein−ligand complex was neutralized by turning off the minimum number of outermost charged residues in each avidin monomer, i.e. residues separated at least 16 Å from the binding site. We then solvated the neutralized complex with a 20 Å sphere of TIP3P water molecules[15] and, using the *parm94* force field,[16] equilibrated the biotin−avidin complex for a total of 180 ps at $T = 300$ K. Equilibrated structures for the six biotin derivatives (BTN2−BTN7) were generated by mutating biotin in its equilibrated complex with avidin to the desired ligand and subsequently performing 90 ps of MD simulation. The structurally different ligands HABA and PEP were incorporated into avidin by superimposing active site residues of the X-ray complexes of biotin−avidin[10] with those of HABA−avidin[17] and PEP−streptavidin[18] (using also the biotin−streptavidin complex[8,9]), respectively. The resulting structures were first minimized and then equilibrated for 150 ps. Detailed conditions of our equilibration procedure are described elsewhere.[19]

The complex binding free energy was calculated from the difference in MM/PBSA free energies for the ligand−protein complex and uncomplexed reactants according to eqs 1−3. Ensembles of structures (50 snapshots) for the MM/PBSA calculation were obtained from 300-ps MD simulations of the solvated complex. The average molecular mechanical energy, E_{MM}, was calculated using AMBER with no cutoff for the evaluation of nonbonded interactions.[20] The PB calculation was done with the *DelPhi* program[21] using PARSE atomic radii[22] and Cornell et al. charges[16] with interior and exterior dielectric constants of 1 and 80, respectively. A grid spacing of 2/Å, extending 20% beyond the dimensions of the solute, was used. G_{np} was calculated from $G_{np} = \gamma SA + b$ ($\gamma = 0.00542$ kcal/(mol·Å^2), $b = 0.92$ kcal/mol)[22] using the surface area estimation of the program *MSMS*.[23] We estimated the change in solute entropy upon association, $-T\Delta S$, with the AMBER module *nmode*. In the first step of this calculation, an 8-Å sphere around the ligand was cut out from an MD snapshot for each ligand−protein complex. On the basis of the size-reduced snapshots of the complex, we generated structures of the uncomplexed reactants by removing the atoms of the protein and ligand, respectively. Each of those structures was minimized using a distance-dependent dielectric constant of $\epsilon = 4r$, to account for solvent screening, and its entropy was calculated using classical statistical formulas and normal-mode analysis. We found that individual MD snapshots adopted different conformations after minimization, resulting in differences in entropy of up to ±5 kcal/mol in unfavorable cases. Consequently, we averaged our entropy estimate over six snapshots and quoted the standard error of the mean as a measure of the variance. Test calculations showed that an 8-Å sphere around the ligand was large enough to yield converged mean changes in solute entropy.

Since the binding free energy for the cyclic peptide has been measured for the association with streptavidin but not with avidin, we additionally calculated ΔG_{bind} for the complex between PEP and streptavidin. In those calculations, we started directly from a 1.9-Å crystal structure of the peptide−streptavidin complex[18] and used an analogous computational procedure as described above.

All calculations have been performed on SGI Origin R10000/225 and DEC AlphaServer 4100/533 computers. Converted to CPU times of a single processor of the SGI Origin, the following approximate CPU hours are needed for the individual computational steps of each ligand−protein system: 450 ps of MD

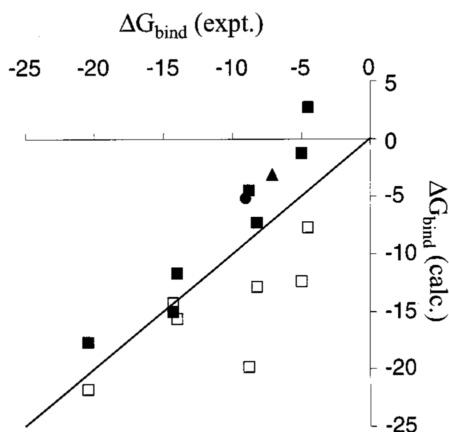

Figure 2. Correlation between calculated and experimental free energies of binding for the ligand–avidin complexes as well as the peptide–streptavidin complex. Black and white markers represent MM/PBSA and modified linear interaction energy[24] calculations, respectively. Different marker shapes distinguish between biotin analogues–avidin (squares), HABA–avidin (triangle), and cyclic peptide–streptavidin (circle). MM/PBSA free energies are listed in detail in Table 1.

equilibration and data collection (165 h), MM/PBSA free energy calculation for 50 snapshots (25 h), normal-mode analysis for six snapshots (10 h). This corresponds to a total CPU time (single processor) of 200 h for the calculation of ΔG_{bind} of each complex or 83 days for the 10 systems investigated in this study. Since MD programs are mostly well-parallelized and sets of snapshots of the MM/PBSA and normal-mode calculations can be easily distributed among different processors, the above-mentioned CPU time can be drastically reduced in multiprocessor computing environments.

Results and Discussion

The results of our calculations are illustrated in Figure 2, in which the computed free energies of binding are displayed versus the experimental ΔG_{bind}. Using a regression analysis, we find a very good correlation ($r^2 = 0.92$) for the MM/PBSA values of the seven biotin analogues (black squares) with experiment, indicating an accurate reproduction of relative free energies of binding by our method. The average absolute error (AAE) to the regression line ($y = 1.23x + 5.4$) is 1.7 kcal/mol. This agreement is less so for LIE, which has been previously used to calculate the affinity of those ligands to avidin.[24] Using the original approach (electrostatic and van der Waals coefficients: $\alpha = 0.5$, $\beta = 0.16$) the correlation coefficient is small ($r^2 = 0.12$, values not shown in Figure 2). While this can be improved by fitting the nonpolar coefficient to experiment ($\beta \approx 1$) and including an entropic term into the LIE approach (white squares), there remains still less agreement with experiment ($r^2 = 0.55$, AAE = 2.3 kcal/mol, $y = 0.61x - 8.4$) compared to MM/PBSA. Note also that the slope of the MM/PBSA approach is much closer to experiment than found with LIE.

As can be seen from Figure 2, the good correlation between our method and experiment can be extended when we include the more diverse systems HABA–avidin (triangle) and PEP–streptavidin (circle) into our comparison ($r^2 = 0.92$). This is especially encouraging for the PEP–streptavidin complex (a) because sufficient conformational sampling seems most challenging for the flexible cyclic hexapeptide in the range of ligands and (b) because of the different protein streptavidin. In the comparison of the HABA–avidin complex, it should be noted that the calculated ΔG_{bind} depends on the position of the tautomeric equilibrium for HABA, shown in Figure 1h, in both the protein complex and aqueous solution. While spectroscopic data clearly indicate that the hydrazone form is more stable in the low-dielectric protein environment,[25] different opinions exist about this equilibrium in water, and latest spectroscopic results suggest a mixture of both forms at the slightly acidic pH at which the experimental ΔG_{bind} was determined.[26] Because of its predominance in the protein–ligand complex we have modeled HABA in the hydrazone form. We further addressed the tautomeric equilibrium of HABA in aqueous solution by additional MM/PBSA free energy calculations of both uncomplexed tautomers in water, based on individual MD simulations each, as well as by quantum mechanical polarized continuum model (PCM)[27] calculations. Both approaches suggest that in solution the azo form is thermodynamically more stable than the hydrazone tautomer by 4.8 ± 1.0 kcal/mol (MM/PBSA) and 4.0 kcal/mol (PCM). Assuming that this is the real situation in solution one would have to increase the computed ΔG_{bind} for HABA to +1.7 kcal/mol. Despite this uncertainty of 4–5 kcal/mol for HABA, the overall agreement between calculated and experimental values for the diverse set of ligands is very good.

Further insight into the forces involved in substrate binding can be obtained by analyzing the MM/PBSA free energy contributions, which are listed in Table 1 for the nine ligands. Comparing the van der Waals/nonpolar ($\Delta E_{vdW} + \Delta G_{np}$) with the electrostatic ($\Delta E_{es} + \Delta G_{PB}$) contributions, we find that the association between the nine ligands and avidin, and between the peptide and streptavidin, is mainly driven by more favorable nonpolar interactions in the complex than in solution. This has been proposed as a general scheme for noncovalent association.[28] However, as indicated by the energy components of BTN4 and PEP, this driving force can be considerably weakened when the polar groups do not find an adequate bonding pattern in the protein compared to water. The free energy penalty for this ($\Delta E_{es} + \Delta G_{PB}$) is least for BT1–BT3 and together with their large van der Waals contribution consequently leads to the highest binding affinity in the set of ligands. It is particularly encouraging for our method that the large van der Waals contribution to ΔG_{bind} for the peptide is compensated sensibly by the other energy contributions, resulting in only slight deviation from experiment. Although we cannot compare the computed binding free energy for PEP–avidin with experiment, its calculated value of $\Delta G_{bind} = -7.5$ kcal/mol, which is predicted to be more stable by 2.3 kcal/mol compared to the complex with streptavidin, is reasonable. For a given ligand, the difference in binding free energy between those two homologous proteins is only a few kilocalories/mole due to their very similar active sites. The calculated changes in solute entropy, $-T\Delta S$, are physically reasonable. The rigid imidazolidone (BTN7) with its single ring reveals the smallest change in solute entropy upon binding, and this value increases both by connecting the thiophan ring and attaching the flexible valeric acid side chain.

Table 1. Energy Contributions (see eqs 1, 2) to the Free Energy of Binding, ΔG_{bind}, Between Avidin and the Set of Nine Ligands as well as Between Streptavidin and the Peptide[a]

compd	ΔE_{es}	ΔE_{vdW}	ΔG_{PB}	ΔG_{np}	$-T\Delta S$	ΔG_{bind}	$\Delta\Delta G$
				Avidin			
BTN1	−154.4 (1.2)	−36.4 (0.3)	158.5 (0.9)	−3.5 (0.1)	18.1 (1.3)	−17.7 (1.2)	2.7[b]
BTN2	−163.4 (1.1)	−36.4 (0.5)	168.5 (1.1)	−3.5 (0.2)	19.8 (0.7)	−15.0 (0.9)	−0.7[b]
BTN3	−154.9 (1.2)	−33.5 (0.5)	161.9 (1.0)	−3.4 (0.1)	18.2 (1.0)	−11.7 (1.1)	2.3[b]
BTN4	−37.9 (0.7)	−50.2 (0.4)	71.8 (0.4)	−4.9 (0.2)	16.7 (1.9)	−4.5 (1.7)	4.3[b]
BTN5	−34.2 (0.6)	−32.9 (0.4)	49.5 (0.4)	−3.7 (0.2)	14.0 (1.6)	−7.3 (1.5)	0.9[b]
BTN6	−15.6 (0.6)	−34.3 (0.3)	39.8 (0.4)	−3.5 (0.2)	12.3 (1.9)	−1.3 (1.6)	3.7[b]
BTN7	−17.8 (0.5)	−15.7 (0.3)	28.3 (0.3)	−2.2 (0.1)	10.1 (0.8)	2.7 (0.8)	7.2[b]
HABA	−96.7 (1.0)	−32.2 (0.4)	114.1 (0.7)	−3.9 (0.1)	15.6 (0.9)	−3.1[c] (1.0)	4.0[d]
PEP	−43.3 (1.1)	−73.4 (0.6)	96.4 (1.1)	−7.5 (0.1)	20.3 (1.9)	−7.5 (1.9)	
				Streptavidin			
PEP	−53.9 (1.1)	−59.4 (0.4)	95.6 (1.1)	−6.7 (0.1)	19.2 (2.4)	−5.2 (2.2)	3.8[e]

[a] $\Delta\Delta G$ denotes the difference between calculated and experimental binding free energy. All energies are given in kcal/mol and are averaged over 50 snapshots, except for the change in solute entropy, $-T\Delta S$, for which an ensemble of six snapshots was used. The values in parentheses represent the standard error of the mean. Structures and names of the ligands are given in Figure 1. [b] Experimental free energies of binding in kcal/mol:[12] BTN1 (−20.4), BTN2 (−14.3), BTN3 (−14.0), BTN4 (−8.8), BTN5 (−8.2), BTN6 (−5.0), BTN7 (−4.5). [c] When using the azo tautomer for HABA in aqueous solution, ΔG_{bind} increases to +1.7 kcal/mol and $\Delta\Delta G$ = 8.8 kcal/mol. [d] Experimental free energy of binding:[13] −7.1 kcal/mol. [e] Experimental free energy of binding:[41] −9.0 kcal/mol.

In additional simulations, we have further investigated different possible binding modes for the ligand HABA (see Figure 1h, hydrazone form). Crystallographic data of HABA complexed with avidin[17] and streptavidin[9] were in both cases interpreted as its CO_2^- group being buried in the protein, thereby replacing the ureido group (−HNCONH−) functionality of the naturally bound biotin and its C=O group partially solvated and pointing out of the protein. This binding mode seems counterintuitive because of the large desolvation penalty one has to pay for burying the negatively charged carboxylate in the protein and the similarity between the C=O group of HABA and the ureido group of biotin. Because of the not entirely unambiguous interpretation of the X-ray electron density, we docked HABA in a rotated conformation in which the C=O group of HABA overlaps with the ureido group of biotin, performed an MD equilibration, and calculated the free energy of binding using our above-mentioned protocol. The calculated ΔG_{bind} for this conformation is 11 ± 1 kcal/mol more positive than the X-ray determined one, which clearly confirms that HABA, despite the large free energy price for desolvating a negatively charged moiety, binds with the CO_2^- group in the interior of the protein. This suggests that MM/PBSA can also be used to rank qualitatively the relative free energy of different binding modes, in addition to its use in ranking the relative ΔG_{bind} of different ligands.

Conclusion

The ability of the MM/PBSA approach to quantitatively predict the relative binding affinities of avidin to eight ligands of diverse chemical nature, as well as of streptavidin to a hexapeptide, is a promising step toward a more accurate theoretical description of ligand−receptor complexes. It is particularly encouraging that our method not only performs better than previous LIE calculations on this system but also does not require empirical adjustment, which seems to be the case in the latter approach, where the nonpolar free energy contribution has to be evaluated on a protein-by-protein basis.[29] The fact that we can also reproduce $\Delta\Delta G_{bind}$ for the same ligand binding to two different proteins[30] makes our MM/PBSA approach a useful tool for the general evaluation of protein−ligand interactions. Clearly, our approach is most useful in applications in which good crystallographic data for protein−ligand complexes are available. Cases with conformational heterogeneity, such as when the binding mode of a novel ligand is not known, are more complex.

Probably the largest uncertainty of this method is the calculation of the change in solute entropy upon complexation. It has been estimated that the error in calculating entropies using a normal-mode analysis may exceed 0.5 kcal/mol for flexible molecules.[31] Moreover, different snapshots of the MD trajectory adopt different conformations after energy minimization, resulting in a variance of $-T\Delta S$ of ±5 kcal/mol in unfavorable cases. We have tried to overcome this problem by averaging over six structural snapshots, each, but alternate approaches such as covariance−matrix methods[32,33] might yield a better estimate of the change in solute entropy. In the assessment of the calculated binding free energies of this application, it should also be noted that changes in internal energy upon complex formation are neglected because of the use of structures from a single MD trajectory for all reaction partners. Given the high binding affinity of biotin and supported by calculations of Boström et al. on the biotin−streptavidin system,[34] it is unlikely that biotin and the closely related biotin derivatives have high strain energy. However, in general applications of the MM/PBSA method, the assumption of zero strain energy might be not fulfilled for some ligands. In those cases, the use of individual MD simulations for each molecule involved in the complex formation might be warranted.

Although MM/PBSA is a considerably faster technique compared to conventional free energy perturbation methods,[35] several possibilities exist to further reduce the computational effort per ligand, thereby enabling it to rank a larger set of ligands. First, there is the possibility to replace the PB continuum calculation, which is one of the more time-consuming steps in our approach, with a Generalized Born (GB) solvent model.[36] Preliminary calculations suggest that GB solvation energies correlate reasonably well with the PB values.[37] Additionally performing the MD simulation, which is used for generating a representative set of

structures, with the GB solvation model would make the explicit water simulation unnecessary. Since this is the slowest step in the process the computational cost could be further decreased. This option is available in the new version of AMBER 6.[38] For comparison, LIE calculations require two separate calculations of the complex and ligand in explicit solvent. Second, instead of calculating individual trajectories for each ligand, one might replace this by a simulation in which all ligands simultaneously interact with the protein. Here, the protein would move in the mean field of the ligands while each ligand feels the full force of the protein. A similar approach[39] proved successful in previous studies of binding free energies,[39,40] and we are currently investigating the two above-mentioned alternatives in our lab.

Acknowledgment. We thank W. Wang for helpful discussions. We gratefully acknowledge computer time at the NSF supercomputer centers and graphics support from the UCSF computer graphics lab, T. Ferrin, P.I. (RR-1081). P.A.K. is pleased to acknowledge research support through the NIH (GM-29072). B.K. is grateful for support from the German Academic Exchange Service (DAAD) through a research scholarship.

References

(1) Marrone, T. J.; Briggs, J. M.; McCammon, J. A. Structure-based Drug Design: Computational Advances. *Annu. Rev. Pharmacol. Toxicol.* **1997**, *37*, 71−90 and references therein.
(2) Kuntz, I. D.; Meng, E. C.; Shoichet, B. K. Structure-Based Molecular Design. *Acc. Chem. Res.* **1994**, *27*, 117−123.
(3) Lybrand, T. P. Ligand Protein Docking and Rational Drug Design. *Curr. Opin. Struct. Biol.* **1995**, *5*, 224−228.
(4) Aqvist, J.; Medina, C.; Samuelsson, J. E. New Method For Predicting Binding Affinity in Computer-Aided Drug Design. *Protein Eng.* **1994**, *7*, 385−391.
(5) Hulten, J.; Bonham, N. M.; Nillroth, U.; Hansson, T.; Zuccarello, G.; Bouzide, A.; Aqvist, J.; Classon, B.; Danielson, U. H.; Karlen, A.; Kvarnstrom, I.; Samuelsson, B.; Hallberg, A. Cyclic HIV-1 Protease Inhibitors Derived From Mannitol: Synthesis, Inhibitory Potencies, and Computational Predictions of Binding Affinities. *J. Med. Chem.* **1997**, *40*, 885−897.
(6) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate − DNA Helices. *J. Am. Chem. Soc.* **1998**, *120*, 9401−9409.
(7) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.-S.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D.; Cheatham, T. E., III. *Acc. Chem. Res.* **2000**, in press.
(8) Weber, P. C.; Ohlendorf, D. H.; Wendolowski, J. J.; Salemme, F. R. Structural Origins of High-Affinity Biotin Binding to Streptavidin. *Science* **1989**, *243*, 85−88.
(9) Weber, P. C.; Wendoloski, J. J.; Pantoliano, M. W.; Salemme, F. R. Crystallographic and Thermodynamic Comparison of Natural and Synthetic Ligands Bound to Streptavidin. *J. Am. Chem. Soc.* **1992**, *114*, 3197−3200.
(10) Pugliese, L.; Coda, A.; Malcovati, M.; Bolognesi, M. 3-Dimensional Structure of the Tetragonal Crystal Form of Egg-White Avidin in Its Functional Complex With Biotin At 2.7 Å Resolution. *J. Mol. Biol.* **1993**, *231*, 698−710.
(11) Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L. 3-Dimensional Structures of Avidin and the Avidin Biotin Complex. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5076−5080.
(12) Green, N. M. Avidin. *Adv. Protein Chem.* **1975**, *29*, 85−133.
(13) Green, N. M. Avidin and Streptavidin. *Methods Enzymol.* **1990**, *184*, 51−67.
(14) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints For Deriving Atomic Charges − the Resp Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(15) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(16) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field For the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(17) Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L. The Structure of the Complex Between Avidin and the Dye, 2-(4′-Hydroxyazobenzene) Benzoic Acid (Haba). *FEBS Lett.* **1993**, *328*, 165−168.
(18) Katz, B. A. Binding to Protein Targets of Peptidic Leads Discovered By Phage Display − Crystal Structures of Streptavidin-Bound Linear and Cyclic Peptide Ligands Containing the Hpq Sequence. *Biochemistry* **1995**, *34*, 15421−15429.
(19) Kuhn, B.; Kollman, P. A. A ligand that is predicted to bind better to avidin than biotin: Insights from computational fluorine scanning. *J. Am. Chem. Soc.* **2000**, *122*, 3909−3916.
(20) A typical range of E_{MM} for the complexes between avidin and the biotin derivatives was −(2800−3000) kcal/mol with a standard deviation of approximately 40 kcal/mol for 50 snapshots.
(21) Honig, B.; Nicholls, A. Classical Electrostatics in Biology and Chemistry. *Science* **1995**, *268*, 1144−1149.
(22) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(23) Sanner, M. F.; Olson, A. J.; Spehner, J. C. Reduced Surface − an Efficient Way to Compute Molecular Surfaces. *Biopolymers* **1996**, *38*, 305−320.
(24) Wang, J.; Dixon, R.; Kollman, P. A. Ranking Ligand Binding Affinities With Avidin: A Molecular Dynamics-based Interaction Energy Study. *Protein-Struct. Funct. Genet.* **1999**, *34*, 69−81.
(25) Thomas, E. W.; Merlin, J. C. Resonance Raman Spectroscopic Studies of 2-(4′-hydroxyphenylazo)-benzoic acid and Some Substituted Analogues-II. Binding to Avidin and Bovine Serum Albumin. *Spectrochim. Acta* **1979**, *35A*, 1251−1255.
(26) Ni, F.; Cotton, T. M. Surface-Enhanced Resonance Raman Study of Avidin-Dye Interactions: a Model for Chromophore-Containing Proteins. *J. Raman Spectrosc.* **1988**, *19*, 429−438.
(27) PCM (Miertus, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of ab initio Molecular Potentials for the Prevision of Solvent Effects. *J. Chem. Phys.* **1981**, *55*, 117) combines a Hartree−Fock self-consistent reaction field calculation of the electric polarization of the solvent with separate estimates of the cavity and dispersion−repulsion contributions to the free energy. PCM energies were obtained by minimization of the molecules in the gas phase and subsequent single-point calculations in the reaction field of water, both at the HF/6-31+G* level. Spheres with united atom Hartree−Fock (UAHF) radii (Barone, V.; Cossi, M.; Tomasi, J. A New Definition of Cavities for the Computation of Solvation Free Energies by the Polarizable Continuum Model. *J. Chem. Phys.* **1997**, *107*, 3210−3221) were used to build the cavity, and the polarization charge was normalized by means of additional effective charge, distributed according to the solute electronic density.
(28) Miyamoto, S.; Kollman, P. A. What Determines the Strength of Noncovalent Association of Ligands to Proteins in Aqueous Solution. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8402−8406.
(29) Wang, W.; Wang, J.; Kollman, P. A. What Determines the van der Waals Coefficient β in the LIE (Linear Interaction Energy) Method to Estimate Binding Free Energies Using Molecular Dynamics Simulations? *Protein-Struct. Funct. Genet.* **1999**, *34*, 395−402.
(30) Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A. Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14330−14335.
(31) Senderowitz, H.; Guarnieri, F.; Still, W. C. A Smart Monte Carlo Technique For Free Energy Simulations of Multiconformational Molecules, Direct Calculations of the Conformational Populations of Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 8211−8219.
(32) Karplus, M.; Kushik, J. N. Methods for Estimating the Configurational Entropy of Macromolecules. *Macromolecules* **1981**, *14*, 325−332.
(33) Schlitter, J. Estimation of Absolute and Relative Entropies of Macromolecules Using the Covariance Matrix. *Chem. Phys. Lett.* **1993**, *215*, 617−621.
(34) Bostrom, J.; Norrby, P. O.; Liljefors, T. Conformational energy penalties of protein-bound ligands. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 383−396.
(35) Beveridge, D. L.; DiCapua, F. M. Free Energy Via Molecular Simulation: Applications to Chemical and Biomolecular Systems. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431−492.
(36) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical Treatment of Solvation For Molecular Mechanics and Dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.

(37) Wang, W.; Kollman, P. A. Unpublished results.
(38) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C.; Darden, T.; Merz, K. M.; Stanton, R. V.; Chen, A.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. *Amber 6;* San Francisco, CA, 2000.
(39) Pitera, J.; Kollman, P. Designing an Optimum Guest for a Host Using Multimolecule Free Energy Calculations: Predicting the Best Ligand for Rebek's "Tennis Ball". *J. Am. Chem. Soc.* **1998**, *120*, 7557−7567.
(40) Eriksson, M. A. L.; Pitera, J.; Kollman, P. A. Prediction of the Binding Free Energies of New TIBO-like HIV-1 Reverse Transcriptase Inhibitors Using a Combination of PROFEC, PB/SA, CMC/MD, and Free Energy Calculations. *J. Med. Chem.* **1999**, *42*, 868−881.
(41) Giebel, L. B.; Cass, R. T.; Milligan, D. L.; Young, D. C.; Arze, R.; Johnson, C. R. Screening of Cyclic Peptide Phage Libraries Identifies Ligands That Bind Streptavidin With High Affinities. *Biochemistry* **1995**, *34*, 15430−15435.

JM000241H