Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models

PETER A. KOLLMAN,*,† IRINA MASSOVA,†,× CAROLINA REYES,† BERND KUHN,† SHUANGHONG HUO,† LILLIAN CHONG,† MATTHEW LEE,† TAISUNG LEE,†,▽ YONG DUAN,†,¶ WEI WANG,† OREOLA DONINI,†,# PIOTR CIEPLAK,†,⊥ JAYSHAREE SRINIVASAN,†,∥ DAVID A. CASE,‡ AND THOMAS E. CHEATHAM, III§

Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, The Scripps Research Institute, LaJolla, California 92037, and Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112

Received February 22, 2000

ABSTRACT

A historical perspective on the application of molecular dynamics (MD) to biological macromolecules is presented. Recent developments combining state-of-the-art force fields with continuum solvation calculations have allowed us to reach the fourth era of MD applications in which one can often derive both accurate *structure* and accurate relative *free energies* from molecular dynamics trajectories. We illustrate such applications on nucleic acid duplexes, RNA hairpins, protein folding trajectories, and protein−ligand, protein−protein, and protein−nucleic acid interactions.

I. Introduction

Classical Monte Carlo and molecular dynamics (MD) simulations have been very useful in the study of liquids,[1] enabling a wide variety of properties to be calculated in good agreement with experiment. Until recently, such methods have been less useful in studies of complex biological molecules, because they could not accurately represent structures and free energies.

In fact, even though macromolecules have been studied with classical molecular dynamics since 1976, one could call the first era (1976−1985) the Dark Ages, because all one could do is to take a crystallographer's structure of a protein and ruin it, by moving it significantly away from the correct structure. A major reason for the difficulties was simply that the computational power was not adequate to include the environment around the macromolecule in any but the most primitive fashion. This is not to denigrate the efforts of those who did MD in this era, since their brave efforts[2,3] laid the groundwork for later eras. The work in this era also showed that proteins had both solid- and liquid-like properties, in contrast to the prevailing wisdom from X-ray structures of that era, that proteins were static structures.

The second era, the era of free energy calculations (1985−1994), was stimulated by papers by Berendsen,[4] McCammon,[5] and Jorgensen.[6] In the application of free energy perturbation approaches to biological molecules, one would typically solvate a 10−15-Å sphere around the active site with water molecules and "mutate" one ligand into another both when free and when bound to a protein. Calculating relative free energies of binding enabled contact with experiment in a useful and meaningful way.[7] These simulations typically kept the protein residues outside the active sphere fixed, to remain near the correct structure, and were limited in the size of the mutation whose free energy difference could be accurately calculated.

The third era, the era of structure calculations, was catalyzed by the computationally efficient particle mesh Ewald (PME) algorithm of Darden et al.[8] and the increased computational power available on parallel computers. This era was exemplified by the calculations of Cheatham, who showed that in a molecular dynamics simulation of a DNA double helix $d(CCAACGTTGG)_2$,[9] with full inclusion of water and counterions, the final structure after ∼500 ps was independent of whether the structure was begun in an A or B structure, even though these structures are ∼6 Å RMSD apart. The common average MD structure was in the B family, in agreement with experiment. The goal of such studies, simply stated, is to stay at the "correct" structure if started there and to get there from anywhere else. However, what if the barrier between the starting structure and the experimental structure is too large to surmount in about nanoseconds of simulation time? Can one see which structure is lower in free energy, and will the correct native structure be calculated as the lowest in free energy? This would be an excellent test of molecular mechanical and solvation models.

The ability to calculate structure and free energy, initiated by the papers by Srinivasan et al.,[10] Hermans et

Peter Kollman is a Professor of Pharmaceutical Chemistry in the School of Pharmacy, University of California, San Francisco. Carolina Reyes, Lillian Chong, Matthew Lee, and Wei Wang are graduate students and Irina Massova, Bernd Kuhn, Shuanghong Huo, Taisung Lee, Yong Duan, and Oreola Donini are postdocs in his laboratory, having received their Ph.D. degrees at Wayne State (Massova), ETH Lausanne (Kuhn), Boston University (Huo), Duke (Lee), University of Pittsburgh (Duan), and Queens University (Donini). Piotr Cieplak, who received his Ph.D. at Warsaw University, has been a Visiting Professor in Kollman's laboratory, and Thomas Cheatham is a former graduate student who did postdoctoral work at the NIH. David Case is a Professor of Chemistry and Molecular Biology at the Scripps Research Institute, and Jaysharee Srinivasan is a former postdoc in Case's laboratory, who received her Ph.D. from Wesleyan University. The initial development of the methodology described herein was a collaboration (ref 10) between the Case and Kollman groups.

* To whom correspondence and reprint requests should be addressed.
† University of California, San Francisco.
‡ The Scripps Research Institute.
§ University of Utah, Salt Lake City.
× Current address: Roche Biosciences, Palo Alto, CA.
▽ Current address: Molecular Simulations, Inc., San Diego, CA.
¶ Current address: Department of Chemistry, University of Delaware, Newark, DE.
# Current address: Kinetek Pharmaceuticals, Vancouver, BC, Canada.
⊥ Current address: Department of Chemistry, University of Warsaw, Poland.
∥ Current address: Combichem, Palo Alto, CA.

10.1021/ar000033j CCC: $19.00
Published on Web 10/04/2000

al.,[11] and Jayaram et al.[12] in 1998, heralds the beginning of a fourth era of macromolecular MD. This era combines the advances of the second and third eras, in that we can now calculate structure and free energy. In this Account we describe the new methodology, which we call MM−PBSA (molecular mechanics Poisson−Bolzmann surface area), and how it and variants have been applied to a wide variety of macromolecules and complexes of macromolecules with ligands or each other.

II. The MM−PBSA Approach

In this method to estimate the free energy of a complex system, one carries out a molecular dynamics simulation, typically in a periodic box with water and counterions, and correct representation of long-range electrostatic effects such as PME, saving a set of representative structures. Then one postprocesses these structures, removing any solvent and counterion molecules, and calculates the free energy, $\bar{G}$, according to the following equation:

$$\bar{G} = \bar{E}_{MM} + \bar{G}_{PBSA} - TS_{MM} \qquad (1)$$

where $\bar{G}$ is the calculated average free energy, and $\bar{E}_{MM}$ is the average molecular mechanical energy,

$$\bar{E}_{MM} = \bar{E}_{bond} + \bar{E}_{angle} + \bar{E}_{tors} + \bar{E}_{vdw} + \bar{E}_{elec} \qquad (2)$$

where these correspond to the bond, angle, torsion, van der Waals, and electrostatic terms in the molecular mechanical force field,[10] evaluated with no nonbonded cutoff. G_{PBSA} is the solvation free energy calculated with a numerical solution of the Poisson−Bolzmann equation and an estimate of the nonpolar free energy with a simple surface area term,[13] and $-TS_{MM}$ is the solute entropy, which can be estimated by quasi harmonic analysis[10] of the trajectory or, in selected cases, by using normal-mode analysis.[10] This final term is likely to be much smaller than the other two in many applications of estimating relative free energies. The free energy due to ionic strength effects can be added with a continuum approach, as described below for nucleic acids.[10]

The ability to accurately calculate $\bar{G}$, the average free energy for a given macromolecular system in various different conformations or structures, adds a very important methodology to our computational arsenal. This has been possible before with free energy perturbation, but only for small systems and very limited conformational or topological changes. By using a continuum model, we are implicitly integrating out all the solvent coordinates and simplifying the problem. Also, by calculating the absolute free energy directly with eq 1 between the two "end points" instead of calculating the relative free energy along a mapping coordinate, we are avoiding computations on less interesting intermediate states. Nonetheless, it is clear that applications of eq 1 will have intrinsically much larger errors than free energy perturbation/thermodynamic integration calculations. What is surprising is that, despite these larger uncertainties, we can often calculate ΔG in respectable agreement with experiment.

III. Applications of MM−PBSA

A. Relative Free Energies of Macromolecules. 1. Duplex DNA and RNA. As noted, molecular dynamics simulations of d(CCAACGTTGG)$_2$ transition to B-DNA if started in the A form within ∼500 ps and stay in the B form if started there. On the other hand, the structure remains A if simulated in 85% ethanol,[14] so one has a stable A-DNA trajectory under such conditions. One can then calculate $\bar{G}$(B-DNA) and $\bar{G}$(A-DNA) in continuum water using eq 1 using "snapshots" from the trajectories in water and in 85% ethanol. As shown by Srinivasan et al.,[10] $\bar{G}$(B-DNA) − $\bar{G}$(A-DNA) ≈ −20 kcal/mol. Thus, the B form is more stable by a qualitatively reasonable ∼2 kcal/mol per base pair. This is particularly encouraging given that $\bar{E}_{MM}$(B-DNA) − $\bar{E}_{MM}$(A-DNA) is ∼ −300 kcal/mol, which is dominated by the electrostatic energies because the 18 negatively charged phosphate groups are farther apart in B-DNA than in A-DNA. However, $\bar{G}_{PBSA}$(B-DNA) − $\bar{G}_{PBSA}$-(A-DNA) ≈ +280 kcal/mol, since the phosphate groups of higher charge density in A-DNA are also more effectively solvated.

When one considers sequences that are more "A-philic", such as d(ACCCGCGGGT)$_2$[10] and d(C$_{10}$)·d(G$_{10}$),[15] the calculated ΔG (B-DNA vs A-DNA) becomes less negative, and if one considers the "A-phobic" sequence dA$_{10}$·dT$_{10}$, the calculated ΔG(B-DNA vs A-DNA) becomes more negative, consistent with experimental tendencies. The dominant reason for these tendencies appears to come from the $\Delta\bar{E}_{vdw}$ and $\Delta\bar{E}_{tors}$ components.

Molecular dynamics trajectories of phosphoramidate helices[10] (3′NH instead of the 3′O of normal phosphodiesters) go to the A-DNA form, independent of starting geometry. $\bar{G}$(B-DNA) − $\bar{G}$(A-DNA) is 12 kcal/mol, consistent with the A preference of such helices.[16] Interestingly, but not surprisingly, the calculations find this reversal of relative stabilities dominated by the torsional energies, since the NCCO anomeric effect is considerably smaller than that in OCCO.

When one carries out simulations on r(CCAACGTTGG)$_2$, both A and B trajectories are locally stable,[17] even though there is no evidence for the existence of B-RNA for any sequence. Consistent with experiment, $\Delta\bar{G}$(B-RNA) − $\bar{G}$(A-RNA) = +10 kcal/mol, and ionic strength effects increase this number.[10] Unfavorable van der Waals interactions of the 2′OH play a key role in destabilizing B-RNA, as suggested by simple model-building long ago.[18]

2. Hairpin RNA. Both an incorrect and a correct structure of a hairpin loop RNA, with a 4-base-pair stem and a UUCG loop, are locally stable for 5.5 ns. The incorrect form does convert to the correct form upon locally enhanced sampling in 250 ps[19] or if the ribosugars in the loop are converted to deoxyribosugars.[20] These structures differ mainly in the hydrogen bonding of the closing GU base pair in the loop. Encouragingly, the free energy of the correct form is calculated to be more stable than that of the incorrect form by a reasonable magnitude of 4 kcal/mol.[21]

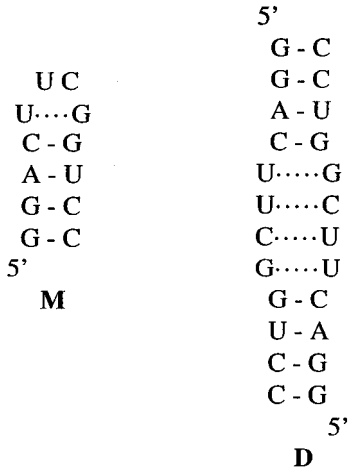

FIGURE 1. Schematic representation of monomer/dimer equilibrium in RNA hairpin loop, where M is observed in solution and D in crystal.

In the crystal, this hairpin loop dimerizes, with the complementary stems base-paired (Figure 1). This suggests that the free energies of two monomer (M) hairpin loops should be comparable to that of a dimer (D). The $\Delta\bar{E}_{\mathrm{elec}}$ favors the monomers by ~ -1850 kcal/mol, but $\Delta\bar{G}_{\mathrm{PBSA}}$ favors the dimer by a comparable amount. $\Delta\bar{G}(2\mathrm{M}-\mathrm{D}) = -18$ kcal/mol without considering salt. The ionic strength free energy at 0.1−1M salt favors D by ∼19−25 kcal/mol. Thus, MM−PBSA successfully supports the comparable free energy of 2M and D and, as is consistent with experiment, higher salt concentration and higher concentration of nucleic acid favor the form that is observed in the crystal, and lower salt and nucleic acid concentration favor the solution form.[21]

3. Proteins. Recently, Duan carried out a 1-μs folding trajectory on the small protein villin, starting with an unfolded form of the protein.[22] Although the protein did not reach the native structure (nor should it, since its estimated folding time is ∼10−100 μs), it came closest (∼4 Å Cα RMSD for the core residues, 9−32) to the native structure at 240−400 ns before leaving this structure and continuing significant fluctuations and changes in radius of gyration and RMSD from the native structure. A 100-ns "control" simulation starting from the NMR structure, and whose core remained within 1.5 Å of the experimental structure, was also carried out. The free energies of the structures along both the folding trajectory and the native trajectory were calculated.[23] The metastable folding intermediate found in the period 240−400 ns was the most stable structure during the folding trajectory, on average ∼10 kcal/mol more stable than structures during any other period of that trajectory. The free energies during the native structure were ∼30 kcal/mol more stable than structures sampled during the folding trajectory (i.e., 20 kcal/mol more stable than the metastable intermediate). Experimentally, the free energy of unfolding of villin at 298 K is ∼5 kcal/mol. How can one reconcile the apparent discrepancy between the experimental ΔG of 5 kcal/mol and the calculated ΔG between native structure and folding trajectory of 30 kcal/mol? If one considers the unfolded structure as an ensemble of many more different conformations than found in the native conformation, one can estimate this ΔG due to conformational entropy as $-RT$ ln x[36], where x is the ratio of the number of conformations/residue for the denatured state compared to the native state. (There are 36 residues in villin.) To reproduce the ΔG "discrepancy" of 25 kcal/mol, x should be near 3, a sensible relative conformational degeneracy ratio between denatured and native states.

The qualitatively reasonable relative $\bar{G}$ for the villin folding and native states[23] has encouraged us to assess the usefulness of calculated $\bar{G}$ for various decoy structures of villin, to compare with native.[24] We found that the four models suggested by Baker's Rosetta,[25] among the most successful approaches in CASP3, were all ∼30 kcal/mol less stable than native. However, after 300 ps, one of these structures (number 18) transitions to a structure that is not only significantly closer to native (∼2.3 Å RMSD for Cα backbone vs 4 Å for the beginning structure of number 18), but whose $\bar{G}$ is also comparable to that of native. This further supports the usefulness of calculating $\bar{G}$ in evaluating the likelihood of existence of any "predicted" protein structure. Of course, if one does not know the native structure, all one can do is rank the predicted structures, but this application of $\bar{G}$ could be a powerful contributor in assessing predicted protein structures and should be an essential element used by both predictors and assessors in CASP4.

Sauer and co-workers[26] have studied a mutant of the arc repressor, in which two wild-type residues $N^{12}L^{13}$ are switched to the mutant $L^{12}N^{13}$. This switch causes a substantial conformational change, in which α helices turn into a β sheet. We have analyzed the MM−PBSA free energies of the two sequences in both structures.[27] In contrast to the above examples involving villin,[23,24] where the free energies are very similar for parm94[28] and parm96[29] force fields, parm96 successfully predicts that the native sequence is more stable in the native structure and the mutant sequence in the mutant structure, whereas parm94 makes the α helical structure more stable for both sequences. The two force fields differ only in their ψ,ϕ torsional potential, and parm96 was developed because it was clear that parm94 overstabilized α helices relative to β sheets. Thus, parm94 seems adequate for comparing structures with similar secondary structures, but parm96 seems to more accurately represent the relative stability of α and β secondary structures. It should be emphasized that one can run representative trajectories with one force field and then postprocess them with any number of others.

B. Binding Free Energies. 1. Protein−Ligand Interactions. One can estimate the ΔG for ligand association to proteins using eq 3:

$$\Delta G = \bar{G}_{\mathrm{complex}} - \bar{G}_{\mathrm{protein}} - \bar{G}_{\mathrm{ligand}} \qquad (3)$$

One can evaluate eq 3 by two methods: (a) run separate trajectories of complex, protein, and ligand or (b) evaluate all three terms in eq 3 using just the snapshots from a trajectory on the complex. Option (b) is 2−3 times more efficient but assumes that the snapshots of the protein and ligand taken from the complex trajectory are of

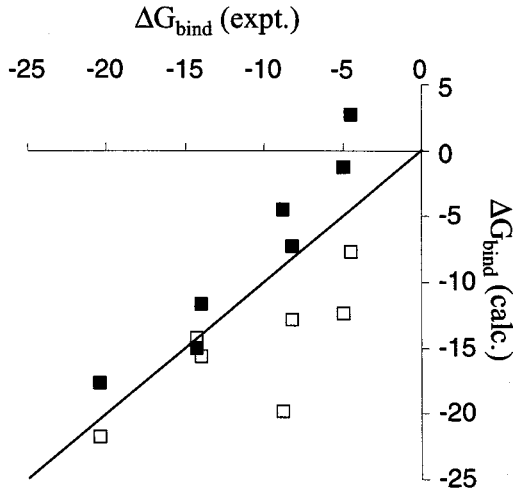

FIGURE 2. Correlation between calculated and experimental protein−ligand binding free energies for avidin and several biotin analogues. Black squares denote MM−PBSA calculated free energies,[30] and white squares refer to LIA calculations.[31] The solid line indicates perfect correlation ($r^2 = 1$).

comparable free energy to those that would emerge from separate trajectories of protein and ligand. Approach (b) seems a good approximation in some applications presented below.

We have applied eq 3 with approach (b) to the binding of biotin and analogues to avidin.[30] This system had been studied earlier with the linear interaction analysis (LIA) approach of Aqvist (eq 4).[31]

$$\Delta G = \alpha\langle\Delta E_{\text{elec}}\rangle + \beta\langle\Delta E_{\text{vdw}}\rangle \qquad (4)$$

In this approach, one carries out two molecular dynamics trajectories: one with the ligand free in solution and the other with the ligand bound to the binding site of the macromolecule. In each trajectory, one calculates the average electrostatic and van der Waals interaction energies of the ligand with other molecules. $\langle\Delta E_{\text{elec}}\rangle$ and $\langle\Delta E_{\text{vdw}}\rangle$ are the average differences between electrostatic interactions and van der Waals interactions in ligand bound and free trajectories, respectively. α and β are semiempirical parameters: $\alpha = 0.5$ from linear response theory, adjusted for the presence of OH groups, and β varies from 0.15 to 1.0, depending on the hydrophobicity of the ligand binding site.[32]

As noted, the LIA approach[31] and MM−PBSA[30] were applied to seven biotin analogues whose free energies for binding to avidin ranged from −4.5 to −20.8 kcal/mol.[30] If one does a regression analysis with the experimental and theoretical free energies, LIA fits experiment with an $r^2 = 0.55$ and an average absolute error ($\Delta\Delta G$) to the fitted line of 2.3 kcal/mol; MM−PBSA has an $r^2 = 0.92$ and an $\Delta\Delta G$ = 1.7 kcal/mol (Figure 2). We also extended the MM−PBSA calculation to two other structurally unrelated molecules known to bind to avidin, for which there was a crystal structure, 2-(4′-hydroxyazobenzene)benzoic acid (HABA), and a cyclic peptide, cyclo-Ac-[CHPQFC]-NH_2, as well as the binding of the peptide to streptavidin. The addition of these molecules did not cause deterioration of the correlation.

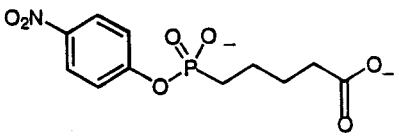

FIGURE 3. Dianionic hapten of 48G7.

Comparing MM−PBSA and LIA, the former (with approach (b)) requires only one simulation/ligand and the latter requires two, and the former has no empirical parameters while the latter has one or two (depending on how α is used). The major difficulty with MM−PBSA lies in the need to carry out a classical statistical normal mode calculation of protein, ligand, and complex to determine the translational, rotational, and vibrational free energy change upon binding of the ligand. In the case of biotin−avidin above, the calculated $T\Delta S$ values were determined as the average over six snapshots. We expect that if the goal is merely to calculate the relative binding free energies of a series of similarly sized molecules binding to a common protein through eqs 1 and 3, this last term (TS_{MM}) might be neglected.

We have also[33] applied "computational fluorine scanning" to see if any of the CH_2 groups in the valeryl side chain of biotin could be replaced with fluorine and thus improve the binding free energy to avidin relative to that to biotin. We have identified one position that MM−PBSA and free energy calculations suggest would bind more tightly to avidin than does biotin.

Another application of eqs 1 and 3 with approach (b) is the study of a dianionic hapten (Figure 3) binding to a germline and affinity-matured antibody.[34] Experimentally, the absolute free energies of binding are −10.9 and −5.3 kcal/mol,[35] and the calculated values are −15.7 and −9.1 kcal/mol.

How does the affinity-matured antibody, which differs in nine amino acids from the germline one, achieve this improvement in binding? Interestingly, the average van der Waals interaction energies $\Delta\bar{E}_{\text{vdw}}$ (eqs 2 and 3) are almost identical for each one (∼35 kcal/mol). The change in free energy due to solute entropy, $T\Delta S_{\text{MM}}$, is 18−19 kcal/mol for both antibodies. The net electrostatic energy $\Delta\bar{E}_{\text{elec}} + \Delta\bar{G}_{\text{PB}}$ is ∼ +13 kcal/mol for germline and ∼ +5 kcal/mol for the mature antibody. Thus, affinity maturation occurs by making the electrostatic free energy less unfavorable in ligand binding.

We have gone further with this analysis by mutating the charges of Tyr 33H to zero in the topology files, reanalyzing the $\Delta\bar{E}_{\text{elec}}$ and $\Delta\bar{G}_{\text{PB}}$ using eqs 1 and 3, and calculating the (free) energy differences between normal and "mutated" Tyr 33H. This gives us the contribution of the partial charges of this residue to $\Delta\Delta\bar{G}$, $\Delta\Delta\bar{E}_{\text{elec}}$, and $\Delta\Delta\bar{G}_{\text{PB}}$ of binding. Orientations of Tyr 33H and the hapten differ in the germline and mature complexes: the O−H of Tyr 33H forms a hydrogen bond with the terminal CO_2^- of the hapten in the germline complex and with the PO_2^- of the hapten in the mature complex. Interestingly, the net sum of electrostatic interactions formed by the Tyr 33H with the hapten appears stronger by ∼1 kcal/mol ($\Delta\Delta\bar{E}_{\text{elec}}$ is 1 kcal/mol larger) in the germline complex, but

the desolvation penalty ($\Delta\Delta\bar{G}_{PB}$) is ~3.5 kcal/mol lower in the mature antibody, leading to the net result that Tyr 33H contributes ~2.5 kcal/mol to affinity maturation because it is more effectively buried in the mature complex and has to pay less of a desolvation price upon hapten binding. As argued more fully elsewhere,[34] minimizing the desolvation penalty that must be paid upon hapten binding is a less geometrically demanding strategy than optimizing specific van der Waals or hydrogen-bonded interactions upon residue mutation, and that is the apparent strategy of affinity maturation in this esterase antibody.

Computational alanine scanning (see below) can also be applied to the nine residues that differ in germline and mature antibodies, to make their sequences equivalent. When this is done, the ΔG_{bind} values of both antibodies are essentially unchanged, further supporting that affinity maturation in this case occurs not by improved direct interactions of mutated residues with the antibody, but by an indirect mechanism whereby the mutations change the geometry of hapten binding. This change is clear from the two crystal structures, and, as illustrated above with Tyr 33H, this can lead to improved binding by changing the electrostatic desolvation penalty.

Donini[36] has used a single trajectory of a ligand binding to a matrix metalloprotease to calculate the free energy of binding of five other analogue inhibitors. The relative binding free energies of the neutral (zwitterionic) inhibitors and charged inhibitors were correctly ranked within their series, but the neutral inhibitors were calculated to bind more strongly than the charged inhibitors, relative to experiment.

2. Nucleic Acid−Ligand Interactions. We[37] have applied MM−PBSA to study the interaction of actinomycin with $d(GAAGCTTC)_2$, and netropsin with $d(CGCGAATTCGCG)_2$, as well as acridine and the actinomycin chromophore with the former sequence. Qualitatively, the results are reasonable, with both actinomycin and netropsin leading to significantly more favorable association than the actinomycin chromophore. However, the absolute free energies of association appear to be significantly less favorable than those for the protein−ligand complexes, which may be due to (thus far) not considering ionic strength effects on association, or, in the case of actinomycin and its chromophore, an overestimate of the free energy for forming an intercalation site for binding.

3. Protein−Protein Interactions. The oncoprotein MDM2 and tumor suppressor p53 play important roles in regulation of the cell cycle. Extensive studies have been done on the interactions of truncated p53 and MDM2, where the N-terminal α helical portion of p53 that binds to MDM2 has been utilized.[38] This 12-residue sequence of p53 binds with micromolar affinity to MDM2. We have studied this truncated p53−MDM2 interaction using MM−PBSA[39] and eqs 1 and 3, with approaches (a) and (b).[35,38] The calculated absolute binding free energy was comparable to that found experimentally, but the most interesting use of MM−PBSA in this study was the implementation of computational alanine scanning, approach (b),[38] in which one mutates the complex trajectory by truncating the side chains of the peptide one at a time and recalculates $\Delta\Delta G_{bind}$ due to each alanine mutation.

Since all 20 amino acids have been substituted at all 12 residue positions in truncated p53, one can compare the experimental data at each position with the calculated $\Delta\Delta G_{bind}$ for alanine mutation, which is the difference between ΔG_{bind} for wild-type and alanine-mutated p53 at a given position. The four residues which have the largest $\Delta\Delta G_{bind}$ are the same as the four residues which are most sensitive to being substituted. Equally encouraging is the fact that the three fully charged residues, D, E, and K, which can be mutated to alanine without significant loss of binding, are calculated to have small $\Delta\Delta G_{bind}$, which requires an effective cancellation of $\Delta\Delta\bar{E}_{elec}$ and $\Delta\Delta G_{PB}$, each of which is ~50 kcal/mol in magnitude.

A separate trajectory was run on W23A, which was one of the critical residues that could not be substituted without substantial loss of binding. The $\Delta\Delta G_{bind}$ calculated by using only the native trajectory and mutating the topology to alanine in the trajectory snapshots was compared to the same quantity obtained when one calculated ΔG_{bind} using eqs 1 and 3 from each trajectory and then took the difference of the ΔG_{bind} of the trajectories (approach (a)). Both approaches led to a $\Delta\Delta G_{bind}$ of ~6 kcal/mol.

We further explored mutations of W23 that might improve binding by using PROFEC[40] to suggest where there was room to put additional methyl groups on the aromatic rings of W23. Both PROFEC and MM−PBSA simulations suggested that one might realize substantially improved binding if a methyl group were put at the $\zeta 2$ or $\eta 2$ positions of W23, but not at the $\epsilon 3$ position.

A second application of computational alanine scanning was to the "classic" case of human growth hormone−human growth hormone receptor (HGH−HGHr) complex.[41] Twelve residues were mutated, and in all cases but two (R43A and R217A), the calculated $\Delta\Delta G_{bind}$ values were in very good agreement with the experimental $\Delta\Delta G_{bind}$ (average absolute error of ~1 kcal/mol). For R43A and R217A, the $\Delta\Delta G_{bind}$ upon alanine mutation was substantially overestimated, with the calculated value 10−15 kcal/mol and the experimental value 1 kcal/mol.

This system is very large (involving ~70 000 atoms including water and counterions), so the simulation of even 400 ps takes considerable computer time. Thus, we carried out only 200 ps of simulation on the R43A system, since the R43A alanine scanning mutation had its $\Delta\Delta G_{bind}$ overestimated by a larger amount than the alanine scanning mutation R217A. During the R43A simulation, there was considerable conformational change around the region of mutation, since R43 was involved in a number of hydrogen bonds. When we recalculated the $\Delta\Delta G_{bind}$ for R43A based on using the ΔG (eq 3) separately calculated from the two trajectories, we found a value of ~1 kcal/mol, in excellent agreement with experiment, and in contrast to the much larger $\Delta\Delta G$ calculated with alanine scanning (Table 1). We expect this to be a general phenomenon—if the residue to be mutated is an ionic one

Table 1. Free Energy Components in the HGH−HGHr Complex[a]

mutation	$\Delta\Delta\bar{G}_{vdw}$	$\Delta\Delta\bar{G}_{PB} + \Delta\bar{E}_{elec}$	$\Delta\Delta\bar{G}_{SA}$	$\Delta\Delta\bar{G}_{TOT}$	$\Delta\Delta\bar{G}_{exp}$
W76A	6.1	−5.5	0.1	0.7	0.5
W80A	0.4	−1.0	0.1	−0.5	−0.02
W104A	12.7	−7.0	0.0	5.7	>4.5
W169A	11.5	−7.9	−0.1	3.5	>4.5
R43A	1.3	13.0	−0.3	14.0	2.12
R43A (separate)	21.2	−21.6	1.8	1.4	2.12

[a] The free energies are in kilocalories per mole and correspond to the $\Delta\Delta\bar{G}$ (binding) for the mutant Ala vs native proteins. See eqs 1 and 2 for the meaning of the various comoponents. We have separated the Poisson Bolzmann (G_{PB}) from the surface area (G_{SA}) contributions in this table, whereas they are combined in eq 1.

involved in a complex network of hydrogen bonds, the structural rearrangement of the environment upon mutation to a nonionic residue is likely to be substantial, and computational alanine scanning using only the native trajectory will then overestimate the $\Delta\Delta G$ upon mutation.

Wang[42] has recently studied HIV protease dimerization using MM−PBSA. He first was able to show that heterodimers of HIV protease had a stronger association than the native homodimer and the rank order of the different dimers agreed with experiment. He also developed a qualitative geometrical criterion to seek other mutations that could affect dimerization free energy and then used a rapid, minimization-based method to evaluate their MM−PBSA dimerization free energy. Several new mutants that might further stabilize heterodimer stability were identified.

4. Protein−RNA Interactions. Molecular dynamics simulations have been carried out on the spliceosomal protein U1A,[43] which binds to both an internal loop (IL) and a hairpin loop (HL) of comparable sequence. Crystallographic and NMR data for both complexes and the free protein and internal loop are available.[44−47] It is clear that, in contrast to the p53-MDM2 and HGH−HGHr complexes discussed above, both the protein and RNA undergo substantial conformational changes upon binding. Thus, one must consider such factors in calculating the free energy of RNA−protein association.[48] If there is a substantial conformational change, one should apply eq 3 using separate trajectories for the two monomers as well as the complex. Carrying out additional simulations on the separate monomers gives one additional information on the free energy of conformational change (ΔG_{conf}) that accompanies complex formation (eq 4):

$$\Delta G_{conf} = \bar{G}(\mathrm{A})(\text{complex trajectory}) - \bar{G}(\mathrm{A})(\text{monomer trajectory}) \quad (5)$$

where $\bar{G}$(A)(complex trajectory) is the average MM−PBSA free energy of molecule A taken from the trajectory of its complex with another molecule and $\bar{G}$(A)(monomer trajectory) is the average MM−PBSA free energy of molecule A taken from a trajectory of A alone.

The main difficulty in applying the above approach for macromolecule−macromolecule complexes is that experimental structure may not be available for the separate components (e.g., free A) and a simulation of about nanoseconds length may not lead to an accurate structure of the separate monomers. However, in the case of U1A complexes with both IL and HL, the following picture emerges (Figure 4): the ΔG_{conf} values for the U1A and HL or IL components are ∼10 kcal/mol, whereas the ΔG_{bind} obtained using monomers from the complex trajectory is ∼ −30 kcal/mol, leading to a net ΔG_{bind} of ∼ −10 kcal/mol, in the correct range as compared to experiment,

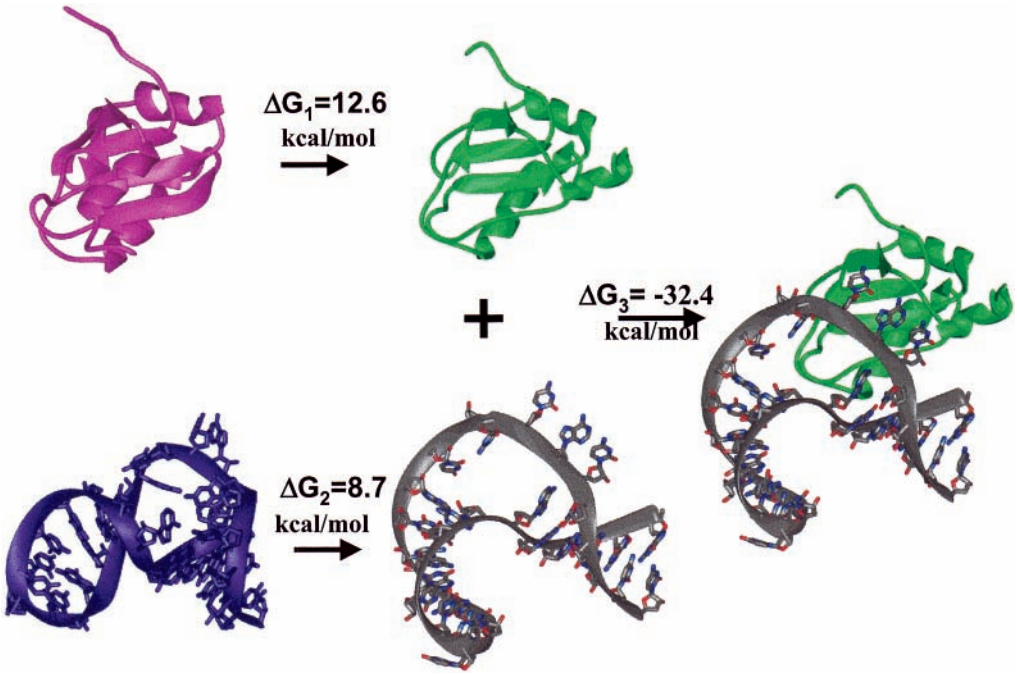

FIGURE 4. Conformational change upon binding of U1A protein and internal loop (IL) RNA. ΔG_1 and ΔG_2 are the MM−PBSA free energy differences between free and bound protein and RNA, respectively. ΔG_3 is the free energy of association of protein and RNA in their bound structures.

which gave −11 and −13 kcal/mol for the U1A-HL and U1A-IL complexes, respectively. Encouragingly, the absolute free energies of binding come out qualitatively correctly using, as in the examples above, normal-mode analyses to estimate $T\Delta S$. We also estimated the free energy cost of conformational change, ΔG_{conf}, for U1A-RNA binding to be ∼20 kcal/mol.

In the study of the U1A-HL and U1A-IL complexes, Reyes[49] set up a hierarchy of methods to estimate the $\Delta\Delta G_{bind}$ due to mutations in either the protein or the RNA. The first procedure is to start with the experimental structure and simply minimize the residue and its mutated form for a common number of steps and then use eq 3 on the single final structures to calculate ΔG for native and mutant structures, with the $\Delta\Delta G$ of mutation being the difference between these.

The second procedure is to run a molecular dynamics simulation on the native complex and postprocess the trajectory with both native and mutated topologies. This is simply a generalization of alanine scanning but can work, in principle, at least as long as the mutant topology involves not making the residue larger, such as Y → F, I → V, guanine → adenine, or thymine → cytosine.

The third procedure is to run separate trajectories on native and mutant structures, as we described for the R43A mutant of the HGHr−HGH complex. Nine mutations were carried out with the first procedure, with a calculated $\Delta\Delta G$ in reasonable agreement with experiment for most, with the clearest exception R52K, whose $\Delta\Delta G$ was calculated to be 9 kcal/mol, but the experimental $\Delta\Delta G$ was 0.5 kcal/mol. Two of the cases were repeated with the second procedure, V45A and Y13F, giving improved quantitative agreement with experiment. Finally, $\Delta\Delta G_{bind}$ for R52K was determined using separate trajectories. Although the error bar of such calculations is large, the calculated $\Delta\Delta G$ is now within ∼1 kcal/mol of experiment.

IV. Perspective and Conclusion

We have described variants of an approach, MM−PBSA, which we feel shows considerable promise in calculating a wide variety of free energies for complex molecular systems. The methods are much more efficient than traditional free energy methods but less accurate; they are, however, much more broadly applicable to systems which differ substantially in structure. Limited space precludes us from discussing in detail related applications using similar methods described in refs 11 and 12.[50−52] We also see this MM−PBSA methodology as nicely complementary to traditional free energy calculations.[53] In cases where both can be applied,[33] a consistent answer gives more confidence in a prediction.

MM−PBSA combines an explicit molecular mechanical model for the solute with a continuum method for the solvation free energy. This distinguishes it from methods which have had successes using just continuum models to calculate ΔG_{bind}.[54] The advantage of including molecular mechanical energies is substantial, allowing van der Waals effects and torsional energies to be explicitly included, and as shown for biotin−avidin, the van der Waals contribution to protein−ligand complexes is considerable.[55] As illustrated most clearly by the W104A and W16941 mutations in HGH−HGHr (Table 1), only by including all three of $\bar{E}_{vdw}$, $\bar{E}_{elec}$, and ΔG_{PB} can an accurate quantitative reproduction of $\Delta\Delta G$ due to mutation be calculated.

Of course, when one averages over many large numbers and a molecular dynamics trajectory, the calculated $\bar{G}$ range is often large. By averaging over "batch means", the standard deviation is considerably smaller than the range, but it is still considerably larger than "chemical" accuracy. Of course, in some mentioned applications above (e.g., B- vs A-DNA and RNA and villin native vs folding), such accuracy is not necessary to answer the question.

Besides the large statistical error and any inherent sampling and force field errors, the use of a continuum model to describe the solvation free energy might be expected to have limitations in situations where explicit water molecules form critical hydrogen-bonded interactions. We have found that, in contrast to the biotin−avidin interaction, where the absolute free energy is calculated in quite good agreement with experiment, in biotin−streptavidin the absolute ΔG_{bind} is calculated to be significantly too positive,[33] presumably because the $-CO_2^-$ of biotin remains partially solvated when bound to streptavidin. Considering explicit water molecules in the MM−PBSA analysis with free and bound ligand improves the agreement in absolute binding free energy.

Also, when binding involves divalent ions such as Mg^{+2} [56] and Zn^{+2},[36] use of the van der Waals radii from molecular mechanics can lead to a significantly too positive calculated ΔG_{bind}, presumably because continuum models overestimate the desolvation penalty of moving a divalent ion from a high to a low dielectric medium. Including some explicit waters as part of the environment of ions with absolute charge >1 may alleviate some of these problems.

Given that the rate-limiting step in MM−PBSA is carrying out the aqueous MD simulation, and waters are not even used in the free energy calculation, could one dispense with the waters in the simulation? There are encouraging aspects of using generalized Born (GB) methods to describe continuum solvation in MD of nucleic acids,[57,58] but it is not clear that simple surface area terms can describe hydrophobic effects adequately for protein simulations. The main necessity for MM−PBSA is to generate representative configurations of the system. This might be accomplished with an efficient simulation in which only the "relevant" residues experience dynamics or, if appropriate, full periodic boundary condition simulations with PME. This choice can depend on the question being addressed.

As presented above, we have shown how MM−PBSA can be used in site-specific mutagenesis, alanine scanning, and protein and ligand design. Much of the technology for carrying out these calculations is contained in AMBER6.[59] One can use a hierarchy of techniques[49]—rapid minimization to screen many possibilities, followed by longer simulations on the systems of particular interest.

In ligand design, it makes sense to first employ rapid but less accurate methods such as DOCK[60] to filter 10^5-10^6 or even more possibilities and then molecular dynamics techniques to perform more accurate calculations on the most interesting ligands. Thus, we feel that MM−PBSA should help to make molecular mechanics approaches (minimization and dynamics) more useful in the "end games" of protein folding, protein structure prediction, and ligand design.

P.A.K. would like to acknowledge research support from the NIH (GM-29072 and CA-25644).

References

(1) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.
(2) McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of Folded Proteins. *Nature* **1977**, *267*, 585−590.
(3) Levitt, M. Computer Simulation of DNA Double-helix Dynamics. *Cold Spring Harbor Symp. Quant. Biol.* **1983**, *1*, 251−262.
(4) Postma, J. P. M.; Berendsen, H. J. C.; Hook, J. R. Thermodynamics of Cavity Formation in Water: A Molecular Dynamics Study. *Faraday Symp. Chem. Soc.* **1982**, *17*, 55−67.
(5) Tembe, B. L.; McCammon, J. A. Ligand−Receptor Interactions. *Comput. Chem.* **1984**, *8*, 281−283.
(6) Jorgensen, W. L.; Ravimohan, C. Monte Carlo Simulation of Differences in Free Energies of Hydration. *J. Chem. Phys.* **1985**, *83*, 3050−3054.
(7) Kollman, P. A. Free Energy Calculations−Applications To Chemical And Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.
(8) Darden; T.; York, D.; Pedersen, L. Particle Mesh Ewald−An N.LOG(N). Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(9) Cheatham, T. E.; Kollman, P. A. Observation of the A-DNA to B-DNA Transition During Unrestrained Molecular Dynamics in Aqueous Solution. *J. Mol. Biol.* **1996**, *259*, 434−444.
(10) 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.
(11) Vorobjev, Y. N.; Almagro, J. C.; Hermans, J. Discrimination between Native and Intentionally Misfolded Conformations of Proteins: ES/IS, a New Method for Calculating Conformational Free Energy that Uses Both Dynamics Simulations with an Explicit Solvent and an Implicit solvent continuum model. *Proteins: Struct., Funct. Genet.* **1998**, *32*, 399−413.
(12) Jayaram, B.; Sprous, D.; Young, M. A.; Beveridge, D. L. Free Energy Analysis of the Conformational Preferences of A and B Forms of DNA in Solution. *J. Am. Chem. Soc.* **1998**, *120*, 10629−10633.
(13) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Continuum Models. *J. Phys. Chem.* **1998**, *98*, 1978−1983.
(14) Cheatham, T. E.; Crowley, M. F.; Fox, T.; Kollman, P. A. A Molecular Level Picture of the Stabilization of A-DNA in Mixed Ethanol−water Solutions. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 9626−9630.
(15) Cheatham, T. E.; Srinivasan, J.; Case, D. A.; Kollman, P. A. Molecular Dynamics and Continuum Solvent Studies of the Stability of poly dG poly dC and poly dA−poly dT DNA Duplexes in Solution. *J. Biomol. Struct. Des.* **1998**, *16*, 265−280.
(16) Cieplak, P.; Cheatham, T. E.; Kollman, P. A. Molecular Dynamics Simulations Find That Phosphoramidate Modified DNA Duplexes Undergo a B to A Transition and Normal DNA Duplexes an A to B Transition. *J. Am. Chem. Soc.* **1997**, *119*, 6722−6730.
(17) Cheatham, T. E.; Kollman, P. A. Molecular Dynamics Simulations Highlight the Structural Differences among DNA:RNA, RNA:RNA, and DNA:RNA Hybrid Duplexes. *J. Am. Chem. Soc.* **1997,** 119, 4805−4825.
(18) Arnott, S.; Hutchinson, F.; Spencer, M.; Wilkins, M. H. F.; Fuller, W.; Langridge, R. X-ray Diffraction Studies of Double Helical Ribonucleic Acid. *Nature* **1966**, *211*, 227−232.
(19) Simmerling, C.; Miller, J.; Kollman, P. A. Combined Locally Enhanced Sampling and Particle Mesh Ewald as a Strategy to Locate the Experimental Structure of a Nonhelical Nucleic Acid. *J. Am. Chem. Soc.* **1998**, *120*, 7149−7155.
(20) Miller, J.; Kollman, P. A. Observation of an A-DNA to B-DNA Transition in a Nonhelical Nucleic Acid Hairpin Molecule Using Molecular Dynamics. *Biophys. J.* **1997**, *73*, 2702−2710.
(21) Srinivasan, J.; Miller, J.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stabilities of RNA Hairpin Loops and Helices. *J. Biomol. Struct. Dyn.* **1998**, *16*, 671−675, 677−682.
(22) Duan, Y.; Kollman, P. A. Pathways to a Protein Folding Intermediate Observed in a 1 Microsecond Simulation in Aqueous Solution. *Science* **1998**, *282*, 740−744.
(23) Lee, M.; Duan, Y.; Kollman, P. A. On the Use of MM−PBSA in Estimating the Free Energies of Proteins: Application to Native, Intermediate and Unfolded Villin Headpiece. *Proteins* **2000**, *39*, 309−316.
(24) Lee, M.; Baker, D.; Kollman, P. A. Getting 1.4 Å C_α RMSD Structure Predictions on Two Small Proteins, HP-36 and S15, with Molecular Mechanics. *J. Am. Chem. Soc.*, submitted.
(25) Simons, K. T.; Bonneau, R.; Ruczinski, I.; Baker, D. Ab initio Protein Structure Prediction of CASPIII Targets Using ROSETTA. *Proteins: Struct., Funct. Genet.* **1999**, *53*, 171−176.
(26) Cordes, M. H. J.; Walsh, N. P.; Mcknight, C. J.; Sauer, R. T. Evolution of a Protein Fold In Vitro. *Science* **1999**, *284*, 315−327.
(27) Lee, T.; Kollman, P. A. Manuscript in preparation.
(28) Cornell, W.; Cieplak, P.; Bayley, C. I.; Gould, I.; 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 and Nucleic Acids. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(29) Kollman, P. A.; Dixon, R.; Cornell, W.; Fox, T.; Chipot, C.; Pohorille, A. The Development/Application of a Minimalist Molecular Mechanics Force Field Using a Combination of Ab Initio Calculations and Experimental Data, in Computer Simulations of Biomolecular Systems: *Computer Simulations of Biomolecular Systems*, ESCOM, 1997.
(30) Kuhn, B.; Kollman, P. A. 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 (MM/PBSA). *J. Med. Chem.*, in press.
(31) Wang, J.; Dixon, R.; Kollman, P. A. Ranking Ligand Binding Affinities with Avidin: A Molecular Dynamics Based Interaction Energy Study. *Proteins* **1998**, *34*, 69−81.
(32) Wang, W.; Wang, J.; Kollman, P. A. What Determines the Van der Waals Coefficient in the LIE (Linear Interaction Energy) Method to Estimate Binding Free Energies in Molecular Dynamics Simulations? *Proteins* **1999**, *34*, 395−402.
(33) 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.
(34) Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A. Molecular Dynamics and Free-Energy Calculations Applied to Affinity Maturation in Anti-body 48G7. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14330−14335.
(35) Patten, P. A.; Gray, N. S.; Yang, P. L.; Marks, C. B.; Wedemeyer, G. J.; Boniface, J. J.; Steven, R. C.; Schultz, P. G. The Immunological Evolution of Catalysis. *Science* **1996**, *271*, 1086−1091.
(36) Donini, O.; Kollman, P. A. Calculation and Prediction of Binding Free Energies for Matrix Metalloproteinases. *J. Med. Chem.*, in press.
(37) Cieplak, P.; Kollman, P. A. Manuscript in preparation.
(38) Bottger, A.; Bottger, V.; Garcia-Echeverria, C.; Chere, P.; Hochkeppel, H.-K.; sampson, W.; Ang, K.; Howard, S. F.; Pricksley, S. M.; Lane, D. P. Molecular Characterization of the HDM2-P53 Interaction. *J. Mol. Biol.* **1997**, *269*, 744−756.
(39) Massova, I.; Kollman, P. A. Computational Alanine Scanning to Probe Protein−Protein Interactions: A Novel Approach to Evaluate Binding Free Energies. *J. Am. Chem. Soc.* **1999**, *121*, 8133−8143.
(40) Radmer, R. J.; Kollman, P. A. The Application of Three Approximate Free Energy Calculations Methods to Structure Based Ligand Design: Trypsin and its Complex with Inhibitors. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 215−227.
(41) Huo, S.; Massova, I.; Kollman, P. A. Computational Alanine Scanning of the 1:1 Human Growth Hormone−receptor Complex. *Proteins: Struct., Funct. Genet.*, submitted for publication.
(42) Wang, W.; Kollman, P. A. Free Energy Calculations of Dimer Stability of the HIV Protease Using Molecular Dynamics and Continuum Solvent Model. *J. Mol. Biol.*, in press.
(43) Reyes, C.; Kollman, P. A. Molecular Dynamics Studies of UIA-RNA Complexes. *RNA* **1999**, *5*, 235−244.
(44) Avis, J. M.; Allain, F. H. T.; Howe, P. W. A.; Varani, G.; Nagai, K.; Neuhaus, D. Solution Structure of the N-terminal RNP Domain of U1A Protein−The Role of C-terminal Residues in Structure Stability and RNA Binding. *J. Mol. Biol.* **1996**, *257*, 398−341.

(45) Allain, F. H. T.; Gubser, C. C.; Howe, P. W. A.; Nagai, K.; Neuhaus, D.; Varani, G. Specificity of Ribonucleoprotein Interaction Determined by RNA Folding During Complex Formation. *Nature* **1996**, *380*, 646−650.
(46) Grubser, C. C.; Varani, G. Structure of Polyadenylation Regulatory Element of the Human U1A Pre-MRNA 3′ Untranslated Region and Interaction with the U1A Protein. *Biochemistry* **1996**, *35*, 2253−2267.
(47) Dubridge, C.; Ito, H.; Evans, P. R.; Teo, C. H.; Nagai, K. Crystal Structure at 1.92 Angstrom Resolution of the RNA-binding Domain of the U1A Sliceosomal Protein Complexed with an RNA Hairpin. *Nature* **1994**, *372*, 432−438.
(48) Reyes, C.; Kollman, P. A. Structure and Thermodynamics of RNA-protein Binding: Using Molecular Dynamics and Free Energy Analysis to Calculate Both the Free Energies of Binding and Conformational Change. *J. Mol. Biol.* **2000**, *297*, 1145−1158.
(49) Reyes, C.; Kollman, P. A. Investigating the Binding Specificity of U1A-RNA by Computational Mutagenesis. *J. Mol. Biol.* **2000**, *295*, *1*−*6*.
(50) Beveridge, D. L.; McConnell, K. J. Nucleic Acids: Theory and Computer Simulation, Y2K. *Curr. Opin. Struct. Biol.* **2000**, *10*, 182−196.
(51) Jayaram, B.; McConnell, K. J.; Dixit, S. B.; Beveridge, D. L. Free Energy Analysis of Protein−DNA Binding: The EcoRI Endonuclease−DNA Complex. *J. Comput. Phys.* **1999**, *151*, *333*−*357*.
(52) Vorobjev, C. Y. N.; Hermans, J. ES/IS Estimation of Conformational Free Energies by Combining Dynamics Simulations with Explicit Solvent with an Implicit Continuum Model. *Biophys. Chem*. **1999**, *78*, 195−205.
(53) Lee, T.; Kollman, P. A. Theoretical Studies Suggest a New Antifolate as a More Potent Inhibitor of Thymidylate Synthetase. *J. Am. Chem. Soc.* **2000**, *122*, 4385−4393.
(54) Froloff, N.; Windemuth, A.; Honig, B. On the Calculation of Binding Free Energies Using Continuum Methods: Application to MHC Class I Protein−peptide Interactions. *Protein Sci.* **1997**, *6*, 1293−1301. Hunenberger, P. H.; Helms, V.; Narayana, N.; Taylor, S. S.; and others. Determinants of ligand binding to cAMP-dependent protein−kinase. *Biochemistry* **1999**, *38*, 2358−2366. Shen, J.; Wendoloski, J. Electrostatic Binding Energy Calculation Using the Finite Difference Solution to the Linearized Poisson−Boltzmann Equation—Assessment of its Accuracy. *J. Comput. Chem.* **1996**, *17*, 350−357.
(55) 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. Miyamoto, S.; Kollman, P. A. Absolute and Relative Binding Free Energy Calculations of the Interaction of Biotin and its Analogs with Streptavidin Using Molecular Dynamics Free Energy Perturbation Approaches. *Proteins: Struct., Funct. Genet.* **1993**, *16*, 226−245.
(56) Minehardt, T.; Pate, E.; Cooke, R.; Kollman, P. A. A Molecular Dynamics Study of the Energetic, Mechanistic and Structural Implications of a Closed Phosphate Tube in NCD. Submitted to *Biophys. J.*
(57) Williams, D. J.; Hall, K. B. Unrestrained stochastic dynamics simulations of the UUCG tetraloop using an implicit solvation model. *Biophys. J.* **1999**, *76*, 3192−3205.
(58) Tsui, V.; Case, D. A. Molecular Dynamics Simulations of Nucleic Acids Using a Generalized Born Solvation Model. *J. Am. Chem. Soc.* **2000**, *122*, 2489−2498.
(59) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E.; 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. K.; Kollman, P. A. Amber 6; University of California, San Francisco, 2000.
(60) Zou, X. Q.; Sun, Y. X.; Kuntz, I. D. Inclusion of solvation in ligand binding free energy calculations using the generalized-born model. *J. Am. Chem. Soc.* **1999**, *121*, 8033−8043.

AR000033J