VOLUME 29 NUMBER 10

ACCOUNTS of CHEMICAL RESEARCH®

OCTOBER 1996

Advances and Continuing Challenges in Achieving Realistic and Predictive Simulations of the Properties of Organic and Biological Molecules

PETER A. KOLLMAN

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

Received October 26, 1995

Introduction

Can the theoretical/computational chemist correctly predict the structures and free energies of molecular systems? It is an assumption of our research that classical molecular mechanics/dynamics is or can be made sufficiently accurate to accurately represent not only conformational (intramolecular) energetics but also nonbonded (both intramolecular and intermolecular) energetics.

Even though these classical mechanical energy functions are not perfect representations of the Born−Oppenheimer surface, progress has been made during the last decade in improving them.

If the force fields are proven accurate on rigid or degenerate systems such as liquids, where we can sample enough of the system to be confident of convergence, then they can be usefully applied to systems of greater complexity, where the main difficulty is sampling enough configurations of the system.[1−3]

Given the importance of an accurate molecular mechanical model, we first present the recent efforts to develop an improved one, dividing our discussion into (a) philosophy and approach to a molecular mechanical model, (b) testing the molecular mechanical model on simple systems, (c) nonadditive molecular mechanical models, and (d) the importance of long-range electrostatic effects in complex systems.

We then turn to applications of such models to molecules of organic and biochemical interest. Here we highlight a number of themes: (a) solvation in simple and complex systems, (b) molecular recognition, (c) more complex systems, octanol and DNA, and (d) chemical reactions in solutions and in enzymes. We then conclude with perspectives for the future. [*Accounts of Chemical Research* articles are intended to emphasize research in the author's lab; thus, we have been quite selective in our discussions. A recent exhaustive review of free energy calculations has attempted to note all the contributions to 1993.[4] However, we note that the many exciting studies being carried out in other labs attest to the vitality of computational chemistry and biochemistry.]

Molecular Mechanical Models

A. Philosophy and Approach. Equation 1 represents the simplest functional form of a force field for studying molecules, in which one can vary all degrees of freedom. The earliest force fields, which attempted to describe the structure and strain of small organic molecules, focused considerable attention on more elaborate functions of the first two terms, as well as cross terms. The modern versions of this are MM2/MM3[5,6] and CVFF,[7] which have been built with this "top down" philosophy.

Peter A. Kollman received his Ph.D. from Princeton University in 1970 and, after a year of postdoctoral work at Cambridge University, has been on the faculty of the Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, since September 1971. He is the father of two children, Sarah, 21, and Eli, 19, and the husband of Mercedes Acosta. In 1995, he received the ACS award for Computers in Chemistry, and in the period since his last *Accounts of Chemical Research* article (1990), he has improved his consecutive free throw record to 55, slightly more than his age.

(1) Dill, K. A.; Bromberg, S.; Yue, K. Z.; Fiebig, K. M.; Yee, D. P.; Thomas, P. D.; Chan, H. S. *Protein Sci.* **1995**, *4*, 561−602.
(2) Kolinski, A.; Skolnick, J. *Proteins* **1994**, *18*, 338−352.

S0001-4842(95)00067-7 CCC: $12.00

$$U(\mathbf{R}) = \sum_{\text{bonds}} K_r(r - r_{\text{eq}})^2 \quad \text{bond}$$
$$+ \sum_{\text{angles}} K_\theta(\theta - \theta_{\text{eq}})^2 \quad \text{angle}$$
$$+ \sum_{\text{dihedrals}} (V_n/2)(1 + \cos[n\phi - \gamma]) \quad \text{dihedral}$$
$$+ \sum_{i<j}^{\text{atoms}} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right) \quad \text{van der Waals}$$
$$+ \sum_{i<j}^{\text{atoms}} q_i q_j/\epsilon R_{ij} \quad \text{electrostatic} \qquad (1)$$

On the other hand, our approach, guided by our interest in proteins and nucleic acids, has been "bottom up".[8−10] Thus, we focused on the atomic charges q_i first. Building on work by Momany[11] and Cox and Williams,[12] we felt that the best, most general method to derive the atomic charges was to fit them to quantum mechanically calculated electrostatic potentials on appropriately chosen molecules or fragments. In our earlier attempt to do this, because of computational limitations in quantum mechanical calculations, we used a minimal basis set STO-3G to derive the q_i.[8,9] However, in our latest efforts,[10] a 6-31G* basis set was used. This basis set has the fortunate property in that it leads to charges (dipole moments) that are enhanced over accurate gas phase experimental values and, thus, implicitly builds in "polarization" effects characteristic of polar molecules in condensed phases. The fact that this basis set enhances the polarity just about the same amount as the water models TIP3P[13] and SPC[14] (where the charges are empirically adjusted to reproduce the water enthalpy of vaporization) is a fortunate fact and is key in leading to *balanced* solvent−solvent and solvent−solute interactions.

Although 6-31G* electrostatic potential charges were well suited for intermolecular interactions, a key stumbling block in their use in a general force field was that they often are statistically ill-determined[15] for buried charges in the molecule and, in that case, can lead to a poor representation of conformational energies. The key breakthrough to solve this problem was the RESP model, developed by Christopher Bayly.[16] By employing a hyperbolic restraint and multimolecule and multiconformational fitting (the latter independently noted as useful by Reynolds and Richards[17]), a general and powerful method to derive 6-31G*-based charges for any organic/biochemical model emerged.

van der Waals parameters are generally dominated by the inner closed shell of electrons, and thus are fortunately far more transferable than atomic charges. Therefore, generally only one set of van der Waals parameters (radius and well depth) per atom type need be employed (with the important exception of hydrogen).[18,19] Unfortunately, it is harder to derive van der Waals parameters than charges using *ab initio* quantum mechanics.[20,21] The emergence of a general model that is empirically calibrated to fit liquid structures and enthalpies, the OPLS model,[22] led us to use this approach in our force field.

Continuing with the bottom up development of our force field, we come to the torsion energy term, where the V_n and γ come from either experiment or quantum mechanical calculations on small-molecule models. Whereas MM2/MM3 often uses many terms in the Fourier series for rotation around a given bond type and attempts to reproduce the conformational energy for a collection of molecules, we have taken a minimalist approach.[10] For example, we have only a single V_3 torsional term around an X−C−C−Y bond except when X or Y is electronegative, where another term can be rationalized from electronic effects and can be derived directly using quantum mechanical calculations. This helps our model to be more easily generalized to new molecules, albeit in some cases probably at the cost of some accuracy. Exceptions to this minimalist approach are the ψ and ϕ of peptides and χ of nucleic acids, where more terms were added to ensure as accurate as possible a reproduction of the conformational energies around these key bonds.

Finally, to ensure reasonable representation of bond and angle terms, we use empirical data (structures and vibrational frequencies). The use of this simple harmonic model precludes high accuracy, but in our opinion one would compromise the simplicity and generality of the model with more complex functional forms. Also, the use of a more complex form does not appear warranted, given the larger errors we are making in the electrostatic term (see below). On the other hand, in our opinion,[23,24] it is essential that bond angles be flexible for accurate reproduction of the above properties.

B. Testing the Model. A key test of our approach was the ability to accurately reproduce liquid struc-

(3) Srinivasan, R.; Rose, G. O. *Proteins − Struct., Funct., Genet.* **1995**, *22*, 81−99.
(4) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395−2417.
(5) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982.
(6) Allinger, N. L.; Yuh, Y. H.; Lii, J-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551, 8566, 8576.
(7) Maple, J. R.; Hwang, M. J.; Stockfisch, T. P.; Demur, U.; Waldman, M.; Ewig, C. S.; Hogle, A. T. *J. Comput. Chem.* **1994**, *15*, 162.
(8) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, Jr., S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.
(9) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230−252.
(10) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5196.
(11) Momamy, F. *J. Phys. Chem.* **1978**, *82*, 592.
(12) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1986**, *2*, 304.
(13) Jorgensen, W. L.; Chandresekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(14) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(15) Francl, M. M.; Casey, C.; Chirlian, L. E.; Gange, D. M. *J. Comput. Chem.* **1995**, *17*, 367−383.
(16) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(17) Reynolds, C. A.; Essex, J. W.; Richards, W. G. *J. Am. Chem. Soc.* **1002**, *114*, 9075−9079.
(18) Gough, C. A.; DeBolt, S. E.; Kollman, P. A. *J. Comput. Chem.* **1002**, *8*, 963−970.
(19) Veenstra, D. L.; Ferguson, D. M.; Kollman, P. A. *J. Comput. Chem.* **1992**, *8*, 971−978.
(20) Margenau, H.; Kestner, N. *The Theory of Intermolecular Forces*; Pergamon Press: Oxford, 1970.
(21) Prissette, J.; Kochanski, E. *J. Am. Chem. Soc.* **1978**, *100*, 6609−6613.
(22) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(23) Kollman, P. A.; Dill, K. A. *J. Biomol. Struct. Dyn.* **1991**, *8*, 1003−1107.
(24) Gibson, K. D.; Scheraga, H.A. *J. Biomol. Struct. Dyn.* **1991**, *8*, 1109−1111.

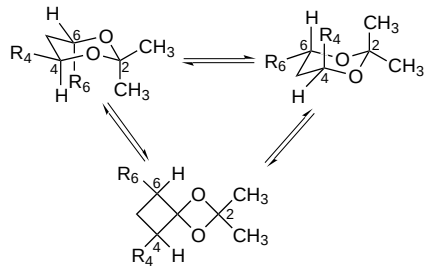


Figure 1. Chair and twist-boat conformations of 2,2-dimethyl-*trans*-4-*n*-hexyl-6-substituent-1,3-dioxanes.

tures and energies and free energies of solvation. The aqueous solvation free energies of a large number of molecules including substituted benzenes,[25] methanol,[26] hydrocarbons,[10] *N*-methylacetamide,[26] and dimethyl sulfide[10] as well as the liquid structure and energy of methanol and *N*-methylacetamide showed very good agreement with experiment. The point to emphasize is that little or no adjustment of parameters was done. Recently, Fox (unpublished results) has shown that our approach leads to a density and enthalpy of vaporization of liquid dimethyl sulfoxide (DMSO) within 2% of experiment, using RESP charges and van der Waals parameters taken without modification from the corresponding values in proteins. Liu *et al*.[28] had derived a united atom DMSO model by empirically adjusting the molecular mechanical parameters to exactly reproduce the experimental density and enthalpy of vaporization, but in the process, they had to make the equilibrium O−S bond length (R = 1.95 Å) significantly different from experiment (R = 1.80 Å).

A test of our electrostatic model was provided by Hobza *et al*.[29] Applying the highest level of *ab initio* theory practical, they calculated the 29 possible H-bonding base−base interaction energies. They then compared these with the energies determined by various force fields and semiempirical quantum mechanical models. Encouragingly, the Cornell *et al*.[10] model was, on balance, closer to the *ab initio* model than any others, even the OPLS[30] and CHARMM23[31] models. This was despite the fact that the Cornell *et al*. model simply fit the base charges with a RESP model, whereas OPLS and CHARMM23 adjusted them empirically to reproduce, among other things, hydrogen bond energies between bases or between base and water molecules.

The ability of our force field to model intramolecular (conformational) energies was provided by the studies of Rychnovsky *et al*.[32] They studied a well-defined conformational equilibrium between chair and twist-boat conformers of substituted 1,3-dioxanes (Figure 1). Even though high-level *ab initio* calculations reproduced the relative energies of these molecules well, MM2*/MM3* (Macromodel implementation of MM2 and MM3) and MM2/MM3 did not. However, our molecular mechanical model using RESP charges[19] had a correlation coefficient relative to these *ab initio* energies of $r^2 = 0.997$ with an average absolute error of 0.45 kcal/mol. In contrast, MM3 produced only an $r^2 = 0.749$ and an average error of 2.37 kcal/mol.[33] The important role of the electrostatic term in determining these energies (the r^2 of the relative electrostatic energies with the relative total energies was 0.99) explained the superior performance of the Cornell *et al*. approach compared to MM3. A qualitative insight into why the electrostatic and total energies were correlated suggested that, in addition to a steric effect favoring the twist-boat conformer, electron-withdrawing substituents favored the chair conformation because of electrostatic attraction with O_3. This led to the idea that a 6-CF_3 substituent would have a greater tendency than 6-CH_3 to be axial in the chair conformation, despite the smaller size of 6-CH_3. This idea was tested and supported in a joint experimental and theoretical study involving Rychnovsky's lab and ours.[34] Thus, even though this is one limited example, it provides encouragement that the approach described in ref 16 will be able to accurately represent intramolecular energies.

C. Nonadditive and More Complex Models. What are the most important weaknesses in the above-described parametrizational approach and the use of eq 1? In our opinion, the main ones are two: the use of an effective two-body potential and the use of only atom-centered charges.

$$E_{\mathrm{pol}} = -(1/2)\sum_{i}^{\mathrm{atom}} \mu_i \cdot \mathbf{E}_i^{(0)} \qquad \text{polarization}$$

$$+ \sum_{\mathrm{ligands}}^{\mathrm{ion}\ i} \sum_{j,k} A_{ijk} e^{-\alpha_{ij}R_{ij}} e^{-\alpha_{ik}R_{ik}} e^{-\beta_{jk}R_{jk}} \qquad \text{three-body exchange} \tag{2}$$

In the last year, we have made substantial progress in laying the foundation for the development of a complete force field including explicit nonadditive effects (adding eq 2 to eq 1). Firstly, we have shown that such models, in contrast to additive models, lead to good agreement with experimental solvation free energies of representative organic ions $CH_3NH_3^+$ and $CH_3CO_2^-$ without any adjustment of van der Waals parameters.[35] Secondly, we have shown that such nonadditive terms are essential in accurately describing cation−π interactions.[36] Thirdly, we have shown that one can equally well describe liquid CH_3OH and NMA with additive models or a nonadditive model in which the charges are uniformly reduced (by 0.88).[26] Finally, the interaction free energy of Li^+ with the hexaanisole spherand is more accurately described by nonadditive than additive molecular mechanical models.[37] In addition, considering off-center charges in electrostatic potential fit models of atoms with "lone

(25) Kuyper, L.; Ashton, D.; Merz, K. M.; Kollman, P. A. *J. Phys. Chem*. **1991**, *95*, 6661−6666.
(26) Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem*. **1995**, *99*, 6208−6219.
(27) Sun, Y.; Spellmeyer, D.; Pearlman, D.; Kollman, P. A. *J. Am. Chem. Soc*. **1992**, *114*, 6798−6801.
(28) Liu, H. Y.; Mullerplathe, F.; van Gunsteren, W. F. *J. Am. Chem. Soc*. **1995**, *117*, 4313−4366.
(29) Hobza, P.; Hubalek, F.; Kabelac, M.; Spooner, J.; Mezzlik, P.; Vondrasek, J. *Chem. Phys. Lett*. **1996**, *257*, 31−35.
(30) Pranata, J.; Jorgensen, W. L. *Tetrahedron* **1991**, *47*, 2491−2501.
(31) Mackerell, A. D.; Wiorkiewieczkuczera, J.; Karplus, M. *J. Am. Chem. Soc*. **1995**, *117*, 11946−11975.
(32) Rychnovsky, S. D.; Yang, G.; Powers, J. P. *J. Org. Chem*. **1993**, *58*, 5251−5255.
(33) Howard, A. E.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem*. **1995**, *16*, 243−261.
(34) Cieplak, P.; Howard, A. E.; Powers, J.; Rychnovsky, S.; Kollman, P. A. *J. Org. Chem*. **1996**, *61*, 3662−3668.
(35) Meng, E. C.; Cieplak, P.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc*. **1994**, *116*, 12061−12062.
(36) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc*. **1995**, *117*, 4177−4178.
(37) Sun, Y.; Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem*. **1995**, *99*, 10081−10085.

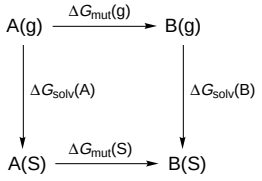
pairs" shows that they can often be important in leading to a very accurate description of H-bond directionality.[38]

D. Long-Range Electrostatic Effects. To accurately describe the energy and structure of complex systems, not only are the functional form and parameters of molecular models described by eqs 1 and 2 important, but also the manner in which the long-range electrostatic effects are represented. The standard approach is to use a nonbonded cutoff for both electrostatic and van der Waals interactions, which seems to be a reasonable method for proteins, but appears to be a poor method to describe highly charged molecules such as nucleic acids. Darden and coworkers have shown the impressive efficiency and accuracy of the particle mesh Ewald (PME) method in representing protein crystals[39] (0.3 Å rms deviation from the observed crystal structure for bovine pancreatic trypsin inhibitor (BPTI) in a 1 ns simulation with an increase in computer time of only ~50% over standard cutoff methods); in collaboration with Darden, Cheatham *et al.*[40] have shown that the PME method leads to very accurate simulations of proteins, DNA, and RNA in solution.

Applications

A. Solvation. Free energy is certainly one of the most important concepts in physical chemistry. The groundwork on calculating free energies was laid by Kirkwood[41] and Zwanzig,[42] and the first key "modern" developments and applications came from the work of Berendsen,[43] Jorgensen,[44] McCammon,[45] and Warshel.[46] The fundamentals of computational approaches to calculating free energies were reviewed by Beveridge and Mezei,[47] and we attempted to exhaustively review applications up to 1993.[4] The molecular dynamics codes AMBER,[48] CHARMM,[49] ENZYMIX,[50] and GROMOS,[51] among others, support free energy calculations, and the extensive capabilities for such simulations within AMBER are owed to the yeoman efforts of U. Chandra Singh (1985−1988) and David Pearlman (1988−present).

To calculate the relative solvation free energies of molecules A and B in solvent S, we can use a thermodynamic cycle such as



The relative solvation free energy of A and B determined experimentally is $\Delta\Delta G_{solv} = \Delta G_{solv}(B) - \Delta G_{solv}(A)$, and since the free energy is a state function, also $\Delta\Delta G = \Delta G_{mut}(S) - \Delta G_{mut}(g)$, which are the free energies determined by computational means by "mutating" the molecular mechanical model of A into B in solvent S and in the gas phase (g). Of course, if B consists of all "dummy" (noninteracting) atoms, this approach leads to the calculation of the absolute solvation free energies of A.

The calculation of these $\Delta\Delta G_{solv}$ values was a key element in validating the Cornell *et al.* force field.[10] Of course there are many other interesting structural properties involved in solvation, but these appear less sensitive to the molecular mechanical model than the energetic effects. Calculating $\Delta\Delta G_{solv}$ in agreement with experiment may be fortuitous, but a $\Delta\Delta G_{solv}$ in disagreement with experiment certainly indicates a flaw.

The very good to excellent agreement for both the absolute and relative aqueous solvation free energies of CH_4, C_2H_6, C_3H_8,[27] H_2O,[13] CH_3OH,[26] and $(CH_3)_2O$ (D. Venstra, unpublished results) supports the idea that model 1 contains the essential features of both the *hydrophobic effect* and *hydrogen bond stabilization*. It was thus a surprise when there were some simple molecules where the agreement with experiment was less satisfactory, i.e., the relative solvation free energies of methylated amines. The relative solvation free energy of replacing a hydrogen with a methyl group on an sp^3 carbon or oxygen is calculated in agreement with experiment (Table 1), and the *calculated* free energy of adding a methyl group to an sp^3 nitrogen is close to halfway between the values for the carbon and oxygen compounds.[52] However, the experimental value for the amines is considerably more negative, by ~1.1 kcal/mol per methyl group. Intuitively, it makes sense that replacing a hydrogen-bonding N−H with an N−CH_3 should make the molecule less soluble, both because of the hydrophobic effect and because the dipole moment of the molecule decreases. The H^+ affinity, Li^+ affinity, and H-bond acceptor abilities of successively methylated amines is a balance between electrostatics (which is more favorable for NH_3) and polarization/charge transfer (which is more favorable for successively methylated amines).[53] When we and others[54,55] repeat the free energy calculations using the nonadditive model which contains the polarization effect, the disagreement with experiment is lessened; our calculations find that the $\Delta\Delta G$ for $NH_3 \rightarrow CH_3NH_2$ and $CH_3NH_2 \rightarrow (CH_3)_2NH$ still deviates from experiment by ~1 kcal/mol, whereas the $\Delta\Delta G$ for $(CH_3)_2NH \rightarrow N(CH_3)_3$ is now quite consistent with experiment.

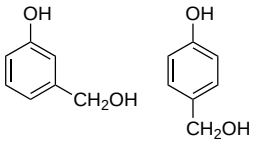
(38) Dixon, R. W.; Kollman, P. A. Going Beyond the Atom-Centered Model in Additive and Non-Additive Molecular Mechanics. *J. Comput. Chem.*, submitted for publication.
(39) York, D. M.; Wlodawer, A.; Petersen, L.; Darden, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8715.
(40) Cheatham, T. E., III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.
(41) Kirkwood, J. *J. Chem. Phys.* **1935**, *3*, 300.
(42) Zwanzig, R. *J. Chem. Phys.* **1954**, *22*, 1420.
(43) Postma, J. P. M.; Berendsen, H. J. C.; Houk, J. R. *Faraday Symp. Chem. Soc.* **1982**, *17*, 55.
(44) Jorgensen, W.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050.
(45) Tembe, B. L.; McCammon, J. A. *J. Comput. Chem.* **1984**, *8*, 281.
(46) Warshel, A. *J. Phys. Chem.* **1982**, *86*, 2218.
(47) Beveridge, D. L.; Mezei, M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431.
(48) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1−41.
(49) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; Slater, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(50) Described in ref 85 and available from Molecular Simulations, Inc.
(51) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *19*, 992. van Gunsteren, W. F.; Berendsen, H. J. C. BIOMOS B. V., Nijenborgh 4, 9747 AG, Groningen, The Netherlands.
(52) Morgantini, P.-Y.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 6057−6063.
(53) Kollman, P.; Rothenberg, S. *J. Am. Chem. Soc.* **1977**, *99*, 1333−1342.
(54) Meng, E. C.; Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 2367−2371.
(55) Ding, Y. B.; Bernardo, D. N.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1995**, *99*, 11575−11583.

Table 1. Relative Solvation Free Energies of X−H and X−CH_3 Mutations (kcal/mol)[a]

	molecules	$\Delta\Delta G_{calc}$ ($\Delta\Delta G_{exp}$)
X = sp^3 C	$CH_4 \rightarrow C_2H_6$	0.0 (0.2)
	$C_2H_6 \rightarrow C_3H_8$	0.2 (0.2)
X = sp^3 N	$NH_3 \rightarrow CH_3NH_2$	0.7 (−0.3) [0.4][b]
	$CH_3NH_2 \rightarrow (CH_3)_2NH$	1.6 (0.3) [1.3][b]
	$(CH_3)_2NH \rightarrow (CH_3)N$	2.3 (1.1) [1.2][b]
X = sp^3 O	$H_2O \rightarrow CH_3OH$	1.1 (1.2)
	$CH_3OH \rightarrow (CH_3)_2O$	3.3 (3.6)
X = sp^2 N	acetamide → NMA[b,c]	2.2 (−0.5)
	NMA → NNDMA[b,c]	1.0 (1.5)

[a] See ref 52. [b] With a polarizable model; see ref 54. [c] NMA = *N*-methylacetamide; NNDMA = *N*,*N*-dimethylacetamide.

Table 2. Relative ΔG_{solv} of *m*- (*m*-HMP) and *p*-(*p*-HMP) (Hydroxymethyl)phenol (kcal/mol)



m-HMP *p*-HMP

solvents	method	$\Delta\Delta G$[a]
gas → water	calculation	−0.4[b]
octanol → water	experiment	−0.4[d]
hexane → water	experiment	1.0[d]
cyclohexane → water	experiment	1.6[d]
toluene → water	experiment	3.2[c]

[a] A negative value means the para isomer is more favorably solvated in water. [b] Reference 58. [c] Reference 56. [d] Reference 59.

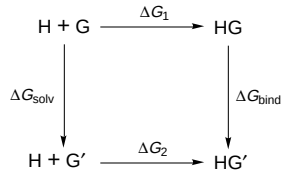
One can plausibly imagine a lack of inclusion of "charge transfer" [53] effects in molecular mechanical models of the N:···HOH interaction as being responsible for the larger than usual discrepancy with the experimental $\Delta\Delta G_{solv}$ in the amines. However, the discrepancy in the relative solvation free energy of acetamide and *N*-methylacetamide (NMA) (calculated, +2.2 kcal/mol; experimental, −0.4 kcal/mol) is not so easily explained (Table 1). Here, one replacing a considerably more acidic hydrogen with a methyl group, and in replacing the second one (NMA → *N*,*N*-dimethylacetamide) theory and experiment agree ($\Delta\Delta G_{solv}$ = 1−1.5 kcal/mol).

An example of a disagreement between free energy calculations and an *interpretation* of experiments involves the relative solvation free energies of ortho, meta, and para isomers of (hydroxymethyl)phenol. Given the potential intramolecular H-bonding of the ortho isomer, we concentrate on the relative ΔG_{solv} of the meta and para isomers (Table 2). Interestingly, the meta isomer has an experimental water/toluene partition coefficient that is much greater than those for the ortho and para isomers.[56] Ben-Naim has interpreted this difference in partitioning as evidence of a greater stability of the meta isomer in water[57] because the two oxygens are approximately the optimum distance to fit into the water structure. Our calculations[58] and experimental data for other solvents[59] suggest that the difference in partition coefficients of these isomers is in no small part due to the "noninteracting" less polar solvent, and this argues against the Ben-Naim interpretation.

(56) Haberfield, P.; Kwuls, J.; Hadded, M.; Rizzo, T. *J. Phys. Chem.* **1984**, *88*, 1913.
(57) Ben Naim, A. *Biopolymers* **1990**, *29*, 567.
(58) Sun, Y.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 6760−6763.
(59) Wilf, J.; Ben Naim, A. *J. Phys. Chem.* **1994**, *98*, 8594−8595.

B. Molecular Recognition. Being able to accurately calculate free energies of solvation suggests a reasonable balance in solute−solvent and solvent−solvent interactions. The next key challenge is to calculate $\Delta\Delta G_{bind}$ of guests G and G′ to a host H, all in aqueous (or other) solution.

A typical cycle for free energy calculations[45] where H is a host, G is a guest, and HG is the host−guest complex is given here:



$$\Delta\Delta G_{bind} = \Delta G_{bind} - \Delta G_{solv} = \Delta G_2 - \Delta G_1 \quad (3)$$

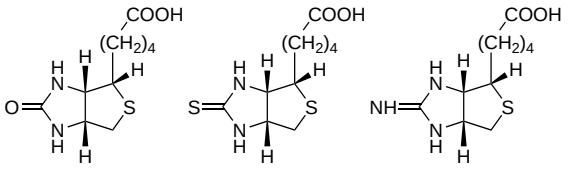
Now one requires a correct balance of solute (host)−solute (guest), solute (host or guest)−solvent, and solvent−solvent interactions to correctly calculate $\Delta\Delta G_{bind}$, although there clearly can be compensating errors in the calculation of ΔG_{solv} and ΔG_{bind}.

The relative binding free energies of alkali metal ions to the octaanisole spherand are fascinating because this system is, to our knowledge, the only example of two minima in the free energy as a function of ion size for the alkali metal cations, with the global minimum at the largest ion Cs^+ and a secondary minimum at Na^+. Calculations by Bayly[60] showed that, whereas the size of the binding site in the host is ideal for Cs^+, Na^+ and two water molecules can also "fill the site" as follows: Na^+ would be coordinated to four ether oxygens and two water oxygens, each of which can use a hydrogen to H-bond to two ether oxygens. Thus, all of the ether oxygens have an electrophilic "partner", and the Na^+ has a favorable coordination number of 6. This aesthetically satisfying rationalization of the secondary minimum led to calculations with solvent CH_3OH replacing water.[61] As expected, with this solvent, since each ether oxygen will not have an electrophilic partner if $Na^+(HOCH_3)_2$ is in the cavity, only the primary minimum at Cs^+ remains. We eagerly await experimental study of the complexation in anhydrous conditions with MeOH as the hydrophilic solvent.

Rebek and co-workers[62] have found that CH_4 and C_2H_4 are encapsulated in a "tennis ball" host, with a free energy of binding in the range of −2 to −3 kcal/mol, but there is no detectable association of CF_4 with this host. We have carried out free energy calculations[63] on the relative free energy of association of CH_4, $CHCl_3$, and CF_4. Although we quantitatively overestimate the $\Delta\Delta G$ for CH_4 vs $CHCl_3$ (calculated, 6.8 kcal/mol; experimental, 5.2 kcal/mol), we agree that $CHCl_3$ is very weakly bound. On the other hand the $\Delta\Delta G$ for CH_4 and CF_4 is only 0.9 kcal/mol, leaving a considerable puzzle of why encapsulation of CF_4 is not detected. Recent studies by Rebek[64] have found CF_4 encapsulated in the host, with binding affinities about $1-1^1/_2$ orders of magnitude less than we calculated. Despite the significant errors in the $\Delta\Delta G_{bind}$, the order of affinities $CH_4 > CF_4 > CHCl_3$ is as we predicted. The underestimate of the affinity of $CHCl_3$ is probably due to the hardness of our B_{ij}/R_{ij}^{12} repulsion, but it is not clear why we overestimate the affinity of CF_4 by so much.

Now let us turn to more complex systems, involving protein−ligand interactions. Here we focus on the biotin−streptavidin interaction, which is probably the strongest protein−organic ligand interaction on a per atom basis (with a K_a of $\sim 10^{13}$ corresponding to $-\Delta G_{bind} \simeq 18$ kcal/mol). This system is a beautiful example of important contributions to binding from hydrophobic/van der Waals and electrostatic interactions in which the "balance" between ΔG_{bind} and ΔG_{solv} is both delicate and very interesting.

However, before attempting to calculate the absolute binding free energy between biotin and streptavidin, we used free energy methods to calculate the relative binding free energy of biotin and thiobiotin (replacing the C=O with C=S, $\Delta\Delta G_{bind} \approx 4$ kcal/mol) and biotin and iminobiotin (replacing C=O with C=NH, $\Delta\Delta G_{bind}$



$\approx$ 6 kcal/mol).[65] Not only were our calculated $\Delta\Delta G$ values in agreement with experiment, but they were "textbook" examples of different ways the $\Delta\Delta G_{bind}$ can be changed by changes in either ΔG_{solv} or ΔG_{bind}. Thiobiotin binds more weakly than biotin because ΔG_{bind}, the interaction with the protein, is reduced much more by the O → S substitution (12 kcal/mol) than ΔG_{solv} (8 kcal/mol). This is not too surprising, given the three protein hydrogen bonds to the ureido oxygen. On the other hand, changing biotin to iminobiotin only raises ΔG_{bind} by ~1 kcal/mol, despite the presence of the hydrogen on the imino group which would lead one to expect more net loss in protein hydrogen bonding to the ligand. However, iminobiotin is ~5 kcal/mol more favorably hydrated ($\Delta G_{solv} = -5$ kcal/mol), leading to $\Delta\Delta G_{bind} \approx 6$ kcal/mol.

In our study of the absolute $\Delta\Delta G$ for biotin−streptavidin, we used the double annihalation technique,[66] "disappearing" biotin in water (ΔG_{solv}) and in the protein (ΔG_{bind}). This corresponds, in the above thermodynamic cycle, to the case where G′ is all "dummy atoms", without interactions with either host or solvent. For convenience, one can choose a pathway where the partial charges on the atoms and then the van der Waals parameters disappear in separate calculations (eqs 4 and 5). This is a very ambitious

$$\Delta G_{bind} = \Delta G_{bind}^{elec} + \Delta G_{bind}^{vdw} \quad (4)$$

$$\Delta G_{solv} = \Delta G_{solv}^{elec} + \Delta G_{solv}^{vdw} \quad (5)$$

calculation, particularly the annihalation of such a relatively large molecule bound to streptavidin, and so one could only expect "semiquantitative" free energies. It turns out that ΔG_{bind}^{vdw} is very large due to the loss of dispersion attraction from the four tryptophans in the binding pocket; thus, the $\Delta\Delta G_{bind}$ is dominated by the van der Waals contribution.

This interpretation of why biotin is so tightly bound to (strept)avidin was different from that suggested by Weber *et al.*,[67] who focused attention on the hydrogen bonding at the ureido group. Our calculations indicate that the electrostatics at the ureido group are important, with biotin bound more tightly than thiobiotin or aminobiotin by 4−6 kcal/mol, but that the van der Waals contribution of ~15 kcal/mol dominates.

Our calculated $-\Delta\Delta G$ for biotin (relative to "nothing") was ~21 kcal/mol, in reasonable agreement with the experimental value, but with a large uncertainty. We thus used the same protocol on a completely different protein−ligand association, *N*-acetyltryptophanamide (NATA) interacting with α-chymotrypsin, where the observed free energy of association was −5 kcal/mol. In this case, our $-\Delta\Delta G$ was ~9 kcal/mol, in qualitative agreement, overestimating $-\Delta\Delta G$ by a comparable amount as in streptavidin−biotin.[68] Also in contrast to biotin−streptavidin the $\Delta\Delta G$ of NATA binding had a larger electrostatic than van der Waals contribution. This further suggested that the biotin−streptavidin complex derives its unusual affinity from the van der Waals/hydrophobic energy term.[68]

We are currently carrying out free energy calculations on the effect of mutations on key amino acid side chains on biotin binding to streptavidin, to complement ongoing experimental work in this area.[69] Another reason for doing such calculations is the ongoing debate in the literature[70,71] on the "usefulness" of separating calculated free energies into their "components", either in terms of energy components (van der Waals vs electrostatic) or in terms of energy contributions due to specific residues. There is general agreement that, whereas the calculation of the total free energy is, in principle, path independent, the calculation of the individual component free energies is not. However, we feel that a demonstration of a significantly greater effect on the biotin binding free energy upon mutating key tryptophan residues than upon mutating residues in the ureido hydrogen-bonding region such as Ser27 → Ala would support the "usefulness" of our previous separation of the free energies into van der Waals and electrostatic components.

Along these same lines, we have studied the Ile 96 → Ala mutation in barnase,[72] earlier studied by Prevost *et al.*[73] Both studies found comparable $\Delta\Delta G_{stab}$ (relative stability of native and mutant proteins) values of 3−5 kcal/mol, in good agreement with experiment. However, Prevost *et al.*[73] found a large contribution from intraresidue effects, whereas Sun *et al.*[72] found that essentially all of the free energy

(60) Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 697−703.
(61) Thomas, B. E.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 3449−3452.
(62) Branda, N.; Wyler, R.; Rebek, J., Jr. *Science* **1994**, *263*, 1267−1268.
(63) Fox, T.; Thomas, B. E.; McCarrick, M.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 10779−10883.
(64) Garcias, X.; Rebek, J. R. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1225−1228.
(65) Miyamoto, S.; Kollman, P. A. *Proteins* **1993**, *16*, 226−245.
(66) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. *J. Chem. Phys.* **1988**, *89*, 3742−3746.
(67) Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R. *Science* **1989**, *243*, 85.
(68) Miyamoto, S.; Kollman, P. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8402−8406.
(69) Dixon, R. Unpublished results.
(70) Mark, A.; van Gunsteren, W. *J. Mol. Biol.* **1994**, *240*, 167.
(71) Boresch, S.; Achonitis, S.; Karplus, M. *Proteins* **1994**, *20*, 25.
(72) Sun, Y-C.; Veenstra, D. L.; Kollman, P. A. *Protein Eng.* **1996**, *9*, 273−281.
(73) Prevost, M.; Wodak, S.; Tidor, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10880.

differences came from nonbonded interactions with residues distant in sequence but near in space in the folded state and with water in the unfolded state.

A similar approach can be used to study DNA−ligand interactions. Wemmer and co-workers[74] have shown that one can form noncovalent complexes in the minor groove of a DNA double helix with two distamycin-like molecules side by side. They have been able to characterize the binding geometries with 2D NMR and to estimate the relative binding free energies of a TTGAA·TTCAA binding site with ligands that contain solely pyrrole rings, compared to ligands in which one pyrrole is replaced with imidazole. We have carried out free energy calculations on these systems[75] in which the pyrrole ring is "mutated" to the imidazole, which corresponds to changing a CH to an N:. Since the two ligands are not in equivalent locations in the minor groove, one has four possible dimeric ligands that can bind in the groove: Pyr:Pyr, Pyr:Im, Im:Pyr, and Im:Im, where Pyr refers to a ligand that has only pyrrole rings and Im refers to one with a pyrrole replaced by imidazole. Im:Pyr differs from Pyr:Im in that the former has the imidazole ring closer to the G of the first strand and Pyr:Im has the imidazole closer to the C in the second strand. Compared to Pyr:Pyr, we calculate the $\Delta\Delta G_{bind}$ to be −0.7 kcal/mol for Im:Pyr (experiment, −1.6 kcal/mol), +2.8 for Pyr:Im (not observed), and +1.8 for Im:Im (experiment, +2.3 kcal/mol). This is another beautiful example of the competition between ΔG_{solv} and ΔG_{bind}. The relative solvation free energy of Pyr and Im is 2.8 kcal/mol, with Im more favorably solvated. When we mutate Pyr:Pyr to Im:Pyr, when found in the minor groove, we pay a solvation price of 2.8 kcal/mol, but the Im is in position to hydrogen bond with the G in the minor groove, and thus, $\Delta G_{bind} = -3.5$ kcal/mol. On the other hand, changing Pyr:Pyr to Pyr:Im costs the solvation price without much change in ΔG_{bind}, because the imidazole is now closer to the C, too far away to form a hydrogen bond. Mutating Pyr:Pyr to Im:Im costs $2 \times 2.8 = 5.6$ kcal/mol solvation, but one gains back the hydrogen bond between Im and G, leading to $\Delta\Delta G_{bind} \approx 2$ kcal/mol.

In an application of free energy calculations to a more challenging "sampling" problem, Sun calculated the free energy of association of K^+ with 18-crown-6 and pentaglyme (its open chain analog) in methanol and water.[76] Only in methanol is the association of K^+ to pentaglyme strong enough to be observed; in this solvent, the experimental free energy of association of K^+ to 18-crown-6 is −8.4 kcal/mol, and to pentaglyme, −3.0 kcal/mol. By evaluating the free energy of annihilating K^+ in solution, when constrained to bind to the D_{3d} conformation of 18-crown-6 and when constrained to bind to the corresponding "D_{3d}-like" conformation of pentaglyme, Sun found a $\Delta\Delta G$ due to a difference in interaction of K^+ with 18-crown-6 and pentaglyme of 1.9 kcal/mol and a difference between the two hosts in constraint free energy of 0.8 kcal/mol while annihilating the K^+, both more favorable for 18-crown-6 than for pentaglyme. By using a more approximate method to estimate conformational energies of pentaglyme, Sun showed that the free energy cost of forcing pentaglyme from its conformational ensemble to the D_{3d}-like conformation was 3.0 kcal/mol, whereas for 18-crown-6, this D_{3d} conformation appears to be the global minimum in solution. Thus, the sum of the above calculated free energies, 5.7 kcal/mol, for the "macrocyclic effect", is in excellent agreement with the above-noted observed value of 5.4 kcal/mol.

The key role of not only conformational degeneracy but also the relative energies of nearby low-energy conformations has also been noted by Lightstone and Bruice,[77] who found an excellent correlation of the reactivity of molecules capable of intramolecular anhydride formation with the relative conformational free energies of conformations which have the atoms suitably placed to react.

C. More Complex Systems: Octanol and DNA. In the past few years, one has become able to carry out simulations on more complex systems. Given the importance of water-saturated 1-octanol as a model for the hydrophobic part of biological systems, and the computational challenge of describing the configurations of complex liquids, Steve DeBolt initiated a thorough study of the structure, dynamics, and solvation of 1-octanol and its water-saturated solution.[78]

First, the OPLS model was slightly adjusted to reproduce, within molecular dynamics, the conformational distribution and liquid structure and enthalpy of methanol, hexane, and octane. Calculations of the percent of non-H-bonded OH groups in liquid octanol at 40 and 75 °C were in excellent agreement with experiment. Octanol liquid and its water-saturated solution were characterized as fluctuating regions that were rich in either OH or CH_n groups ∼8 Å in diameter. Calculations of the dielectric relaxation function led to three dielectric relaxation times in good relative agreement with observations, allowing a new interpretation of the slowest relaxation time as the time for breakup and reformation of the regions rich in OH groups. The calculation of the dielectric constant of octanol at 40 and 75 °C was in reasonable agreement with the experimental values. Finally, the relative octanol/water partition coefficients of phenol and benzene were calculated in excellent agreement with experiment. All the above simulations required nanoseconds of computing to achieve converged properties, but the free energy calculation in water-saturated octanol required particularly long simulations for convergence and were enabled by DeBolt's development of an approach to free energy calculations using coarse grained parallel computing on a multiprocessor NCUBE computer.[79]

As noted above, Cheatham *et al.* have reported stable trajectories of DNA and RNA with full representation of water and counterions, using the Cornell *et al.* force field and PME electrostatics.[40] More recently, Cheatham[80] has been able to carry this much further with his studies of DNA double helices. He has carried out four unrestrained ∼1 ns length MD trajectories in aqueous solution on the DNA duplex

(74) Gerenstanger, B. H.; Jacobsen, J. P.; Maksich, M.; Dervan, P. B.; Wemmer, D. E. *Biochemistry* **1994**, *33*, 3055.
(75) Singh, S. B.; Ajay; Wemmer, D. E.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7673−7677.
(76) Sun, Y.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 3599−3604.
(77) Lightstone, F. C.; Bruice, T. C. *J. Am. Chem. Soc.* **1994**, *116*, 71789.
(78) DeBolt, S. E.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5316−5340.
(79) DeBolt, S. E.; Pearlman, D. A.; Kollman, P. A. *J. Comput. Chem.* **1994**, *15*, 351−373.
(80) Cheatham, T. E.; Kollman, P. A. *J. Mol. Biol.* **1996**, *259*, 434−444.

d(CCAACGTTGG)$_2$, two beginning in a canonical A-DNA structure and two beginning in a canonical B-DNA structure. As judged by root mean squared coordinate deviations, average structures computed from all four of the trajectories converge to within ~0.8−1.6 Å (all atoms) of each other, which is 1.3−1.7 Å (all atoms of the central 6 residues from each strand) and 3.1−3.6 Å (all atoms) away from the B-DNA-like X-ray structure reported for this sequence. This is apparently the first example of independent nanosecond molecular dynamics trajectories with full representation of DNA charges, solvent, and long-range electrostatics that demonstrates both internal consistency (two different starting structures and four different trajectories lead to a consistent average structure) and considerable agreement with the X-ray crystal structure of this sequence and NMR data on duplex DNA in aqueous solution. This internal consistency of structure for a given sequence suggests that one can now begin to realistically examine sequence dependent structural effects in DNA duplexes using molecular dynamics.

D. Chemical Reactions in Solution and in Enzymes. Up to now, we have focused on the study of models based on eqs 1 and 2 to study noncovalent interactions and conformational changes. One would like to also study chemical reactions in condensed phases, which requires electronic structure theory for the atoms where bonding is changing and a simpler representation of the remainder of the system (surrounding the solvent or protein). The pioneering study in this area vis-à-vis enzyme catalysis was that of Warshel and Levitt.[81]

There is exciting research activity in these areas,[82−92] but space limitations preclude their inclusion.

Summarizing Remarks

Above, we have tried to give some examples of "advances", suggesting that significant improvements have been made in force field methodology and models, leading to improved agreement with experiment in simulations of structures and free energies. But what has been learned and of what use is it? The many examples of agreement with experiment validate that a model based on eq 1 (with PME and the addition of eq 2 (nonadditive effects) as necessary) contains much of the essential physics necessary to understand noncovalent effects in biochemistry. Hydrogen bonding and ionic species are well described by a model with accurate electrostatics, and nonpolar (hydrophobic) effects are well described using accurate van der Waals terms for hydrophobic solutes[27] and standard water models.[13,14]

How can these accurate simulations be used to give insight? For example, we have gained physical insight by examining the balance between ΔG_{solv} and ΔG_{bind} in leading to $\Delta\Delta G_{bind}$. The biotin−streptavidin study[65,68] suggested that van der Waals/dispersion forces are the dominant contribution to $\Delta\Delta G_{bind}$[65] (in contrast to the initial interpretation[67]), and our interpretation is consistent with subsequent site-specific mutagenesis studies on the dramatic effect on $\Delta\Delta G_{bind}$ of changing the binding site tryptophans to alanine.[93] The key effect of solvent water[60,61] in the biphasic free energy example in the octaspherand is another example of a beautiful insight from the simulation. Another example is the study of CF_3-substituted dioxanes.[33,34]

The ideal results of such insights are predictions. One has been made in the case of the spherands, where the free energy curve is predicted to become monophasic in methanol.[61] We have also predicted correctly that the phosphinic acid (−CH_2−) inhibitor of thermolysin binds to thermolysin nearly as tightly as the phosphoramidate analog (−NH−).[94,95] The prediction of CF_4 binding to tennis ball molecules has turned out to be correct.[63,64]

If one is generally successful in accurately simulating the properties of many systems, the exceptions to this become more interesting. The methyl effect in amine and amide solvation is such an example. While the inclusion of polarization[54,55] improves the accuracy of the free energy for methyl substitution in amine solvation, particularly $(CH_3)_2NH \rightarrow (CH_3)N$, the ΔG_{solv} for $NH_3 \rightarrow CH_3NH_2$ and $CH_3NH_2 \rightarrow (CH_3)_2NH$ still differs by significantly more than found for the analogous carbon and oxygen analogs. Also still to be understood is why the meta isomer of (hydroxymethyl)phenol has a larger toluene/water partition coefficient than the para isomer.

Resolving the discrepancy in those free energy calculations where there is disagreement between calculation and experiment can be viewed as one of the "continuing challenges", but the largest challenge facing us is the local minimum, or sampling, problem.

That challenge is to determine the relative conformational free energies of complex molecules in solution because we seek a Boltzmann distribution of all low-energy local minima, not just the global minimum. It is clear that a standard application of molecular dynamics and Monte Carlo methods is very inefficient at traversing the space between local minima. Sendorowitz and Still[96] have suggested interesting ways to improve the sampling, by using a combination of a continuum solvation model, stochastic dynamics, and Monte Carlo approaches.

One of the most promising approach for visiting low-energy structures using molecular dynamics with an explicit solvent is the locally enhanced sampling (LES) approach of Elber and co-workers.[97] In this approach, multiple copies of fragments of the molecule topology

(81) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.
(82) Blake, J. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 7430.
(83) Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1992**, 114, 10966.
(84) Gao, J. L.; Xiu, X. F. *Science* **1992**, *258*, 631.
(85) Cramer, C. J.; Truhlar, D. G. In *Solvent Effects and Chemical Reactivity*; Tapia, O., Bertran, J., Eds.; Kluwer: Dordecht, The Netherlands, in press.
(86) Warshel, A. In *Computer Modeling of Chemical Reactions in Enzymes and in Solution*; Field, M. J., Ed.; Wiley: New York, **1991**.
(87) Field, M.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
(88) Hartsough, D. S.; Merz, K. M. *J. Phys. Chem.* **1995**, *99*, 11266.
(89) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30*, 5826.
(90) Alagona, G.; Ghio, C.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 9855−9862.
(91) Daggett, V.; Schroder, S.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 8926−8935.
(92) Rastelli, G.; Thomas, B. E.; Kollman, P. A.; Santi, D. V. *J. Am. Chem. Soc.* **1995**, *117*, 7213−7227.
(93) Chilkoti, A.; Tan, P. H.; Stayton, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1854−1758.
(94) Merz, K. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 5649.
(95) Morgan, B. P.; Scholtz, J. M.; Ballinger, M. D.; Zephin, I. D.; Bartlett, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 297−307.
(96) Senderowitz, H. F.; Still, W. C. *J. Am. Chem. Soc.* **1995**, *117*, 8211.
(97) Roitberg, A.; Elber, R. *J. Chem. Phys.* **1991**, *95*, 9277.

are heated, and this is done in the context of explicit solvent representations. This approach has the additional advantage that the effective "barriers" to rotation of the fragments are reduced to $1/n$ (n = number of copies), the value of a single copy. This LES method has given exciting results for peptides,[98] and we are currently applying it to the very small protein CMTI and the $\alpha \rightarrow \beta$ equilibrium in hexosaccharides.[99]

I acknowledge the efforts of my collaborators, students, and postdocs, who have been the source of the research in which I had the pleasure to participate. Many are mentioned in the joint references, but I will mention explicitly those whose work is featured in this Account. Recent efforts in software within AMBER 4.1 are the work of David Pearlman, David Case, Jim Caldwell, Bill Ross, and Tom Cheatham, with important contributions by Steve DeBolt (now at Scripps), Tom Darden (NIEHS), and J. Vincent and K. Merz (Penn State). For the studies on force field development and long-range electrostatics Wendy Cornell, Piotr Cieplak, Chris Bayly, Allison Howard, Jim Caldwell, Tom Cheatham, Jennifer Miller, Thomas Fox, and Tom Darden have contributed a great deal. In the area of conformational sampling, the work of Yax Sun, Carlos Simmerling, and Steve DeBolt is featured prominently. Free energy calculations involved studies by Yax Sun, Shuichi Miyamoto, Chris Bayly, Bert Thomas, Thomas Fox, Piotr Cieplak, Suresh Singh, and Pierre-Yves Morgantini. QM/MM calculations were done by Kitty Ghio, Giuliano Alagona, Valerie Daggett, Stefan Schroeder, Giulio Rastelli, and Bert Thomas. To all of the above and others noted in the references, I give my thanks. I also gratefully acknowledge research support from the NIH (Grants CA-25644 and GM-29072) and NSF (Grant CHE-94-17458) and useful comments on this Account by Jim Caldwell, Rob Stanton, Bill Ross, and Thomas Fox.

AR9500675

(98) Simmerling, C. L.; Elber, R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3190.
(99) Simmerling, C. L. Manuscript in preparation.