Reports

Free Energy Calculations by Computer Simulation

P. A. BASH,* U. C. SINGH,† R. LANGRIDGE, P. A. KOLLMAN

A fundamental problem in chemistry and biochemistry is understanding the role of solvation in determining molecular properties. Recent advances in statistical mechanical theory and molecular dynamics methodology can be used to solve this problem with the aid of supercomputers. By using these advances the free energies of solvation of all the chemical classes of amino acid side chains, four nucleic acid bases and other organic molecules can be calculated. The effect of a site-specific mutation on the stability of trypsin is predicted. The results are in good agreement with available experiments.

A "FREE ENERGY PERTURBATION" method for studies of the solvation properties of basic biological and other organic molecules and of the site-specific mutagenesis of proteins is described. The statistical mechanical foundations for this method have existed for many years, but they have only recently been used in computer simulations. A number of studies suggested that the method should be useful (*1–7*), but its applications have been limited. Thus we implemented this method in a molecular simulation program (*8, 9*), and have demonstrated its capabilities for problems in areas such as drug design (*10*). We have simulated the solvation properties of the amino acids and their side-chain analogues, and the results provide necessary reference points for studies of proteins. We can predict computationally determined free energies of solvation for four pyrimidines and purines and present a detailed analysis of the solvation properties of acetamide and two of its methyl derivatives and solvation nonadditivity in *p*-nitrophenol. We also apply the method to a site-specific mutant of trypsin.

The free energy difference between two states of a system, A and B, can be described by the statistical mechanical relation (*11, 12*)

$$\Delta G = -RT \ln \langle \exp\left(-\frac{H_{AB}}{RT}\right)\rangle_A \quad (1)$$

where H_{AB} is the difference in the Hamiltonian of states A and B, ΔG is the free energy difference between these states, R is the gas constant, T is the absolute temperature, and the symbol $\langle\rangle_A$ indicates that an ensemble average is to be taken over the reference state A. An equivalent representation describes each state with suitable parameters and a coupling constant λ:

$$\Delta G = \int_{\lambda=0}^{\lambda=1} \langle \frac{\partial H}{\partial \lambda} \rangle_\lambda d\lambda \quad (2)$$

where ΔG is the free energy change between states A ($\lambda = 0$) and B ($\lambda = 1$), H is the Hamiltonian for the system, and $\langle\rangle_\lambda$ indicates that an ensemble average is to be taken over the different states characterized by the coupling parameter λ.

The computational implementation of Eq. 1 is as follows. If the two states differ significantly, the states between A and B are described by the Hamiltonian

$$H_\lambda = \lambda H_A + (1 - \lambda)H_B \quad (3)$$

The free energy difference between states at λ' and λ is

$$\Delta G_\lambda = -RT \ln \langle \exp[(-H_{\lambda'} - H_\lambda)/RT]\rangle_\lambda \quad (4)$$

where $\lambda' - \lambda$ is a small quantity such that $H_{\lambda'} - H_\lambda$ is on the order of RT. The total change is obtained by taking the sum of these averages or "thermodynamic windows" (*4*)

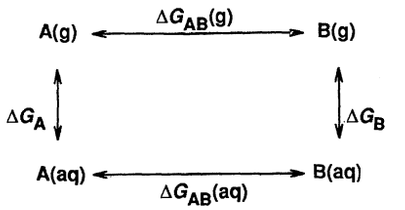
$$\Delta G = \sum_{\lambda=0}^{\lambda=1} \Delta G_\lambda \quad (5)$$

Molecular dynamics (MD) allows the free energy difference in Eq. 2 to be described by "slow growth" (*1*). If the Hamiltonian is changed a very small amount at each MD time step so that the system remains essentially in equilibrium, then the ensemble average may be approximated by a simple difference in the Hamiltonian to obtain

$$\Delta G = \sum_{\lambda=0}^{\lambda=1} H_{\lambda'} - H_\lambda \quad (6)$$

where H_λ is the Hamiltonian at one time and $H_{\lambda'}$ is the perturbed Hamiltonian with the same configuration as H_λ. In this case λ is a function of time.

In the thermodynamic cycle



where the molecules of states A and B are either in the gas phase [A(g) or B(g)] or in aqueous solution [A(aq) and B(aq)], the free energy of solvation of A and B, ΔG_A and ΔG_B, can be measured. Although the "mutation" of A into B in the gas phase $\Delta G_{AB}(g)$ or solution $\Delta G_{AB}(aq)$ cannot be determined experimentally, it can be calculated with computer simulations. Since the free energy is a state function, $\Delta G_{AB}(aq) - \Delta G_{AB}(g) = \Delta G_B - \Delta G_A$, so that the results from computer calculations can be related to experimental values.

Wolfenden *et al.* (*13*) have proposed a "hydration potential" scale for the 20 amino acid side chains in water, where the experimental solvation energies for model compounds are used in place of the actual amino acids. We have performed a series of free energy calculations on amino acid side chains and small model compounds with the thermodynamic windowing procedure and slow growth methods.

Each simulation was done with about 400 TIP3P water molecules (*14*) at constant temperature (300 K) and pressure (1 atm) in a box with periodic boundary conditions (*15*) and an 8 Å nonbonded cutoff. SHAKE (*16*) was used to keep bond lengths fixed to allow a time step of 0.002 psec. The calculations were done in two stages with "electrostatic decoupling." First the charges were changed while the van der Waals parameters were kept at their initial values and then the van der Waals radii and well depths were slowly perturbed to their final values while the parameters for bond lengths and angles were left unchanged. For electrostatic changes the following protocols were used. Electrically neutral molecules were done with 20 thermodynamic windows with $\Delta\lambda = 0.05$; 500 steps of equilibration were followed by 500 steps of data collection, both of which were done in time steps of 0.002 psec. Molecules with net charge were done in 40 thermodynamic windows with $\Delta\lambda = 0.025$; 250 steps of equilibration were followed by 250 steps of data collection in time steps of 0.002 psec. Changes in van der Waals parameters were performed with slow growth rather than windowing because

Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143.

*Present address: Department of Chemistry, Harvard University, Cambridge, MA 02138.
†Present address: Department of Molecular Biology, Scripps Clinic and Research Foundation, La Jolla, CA 92037.

rather large energy fluctuations could occur with windowing even when $\Delta\lambda$ values were as small as 0.001. More reliable results were obtained when Eq. 6 was used with 30,000 steps at 0.001-psec time intervals and when the average of the results for both the forward (A → B) and backward (B → A) directions was used.

The free energies determined in these simulations were sensitive to the charge and van der Waals parameters used to describe each atom. We used the values of the van der Waals parameters described by Weiner *et al.* (*17*) [except for sp^2 C, see (*18*)] in an all-atom force field designed for simulations of proteins and nucleic acids. These van der Waals parameters can be used with comparable types of atoms in other molecules. However, the partial atomic charges were not so easily transferred but were important because for many molecules electrostatic energies dominated the calculated free energy changes. We used electrostatic potential fit charges from ab initio quantum mechanical calculations with a 6-31G* basis set for the amino acids (*19*). The charges were then scaled if necessary to fit the experimental electric dipole moment with the ratio of calculated to experimental dipole moments as the scaling factor.

Results for simulations that cover amino acid functional groups and methane are shown in Table 1. A number of points should be emphasized:

1) The work of Postma *et al.* (*1*) and Straatsma *et al.* (*20*) showed that the contributions to the free energy of solvation of nonpolar groups include the cavity term, which increases the free energy by incorporating the solute in water, and the dispersion term, which decreases the free energy of solvation. The simulation of the absolute free energy of solvation of methane produced a calculated value of 2.1 kcal/mol (Table 1), which compares well with the experimental value of 1.9 kcal/mol, and suggested that our van der Waals parameters were reasonable for carbon and hydrogen in size and dispersion and that adequate sampling was obtained in this case.

2) The simulated free energy changes were within about 10% of relevant experimental values. That is, changes less than 5 kcal/mol were within 0.5 kcal/mol of the experimental value, those between 5 and 10 kcal/mol were within 1.0 kcal/mol, and above 10.0 kcal/mol were within 1.5 kcal/mol.

3) Simulations on the analogous transformations of isobutane into methane and of 3-methylindole into methane were close to both the experimental values and to calculated values for the corresponding amino acid transformations of Leu into Ala and of Trp into Ala. These facts, together with the results for acetamide, acetic acid, and phenol, support Wolfenden's scale for amino acid solvation (*13*).

4) Free energies that were due to changes in van der Waals parameters were in general less accurate than calculated free energies due to electrostatics. This was related to the functional form of the van der Waals parameters and the amount of sampling performed. Nonetheless, these free energy changes could be calculated with much smaller errors than enthalpies (*1*).

5) The calculated free energies for charged amino acids are reasonable, and correctly ordered (*21*); however, the absolute values need to be corrected for reaction field effects to account for the 8 Å cutoff in the simulation. Such corrections are typically less than 0.5 kcal/mol for neutral systems (*22*).

The case of the related molecules acetamide, *N*-methylacetamide, and *N,N*-dimethylacetamide (Fig. 1) is interesting because these molecules all have approximately equal free energies of solvation, yet the methyl-substituted compounds have one and two less hydrogen-bonding groups, respectively. Both experiments and our simulations suggest that the hydrogen-bond potential of the amide group is relatively unimportant in determining these free energies of solvation (*23*). There is reasonable agreement between the calculated and experimental free energies of solvation (Table 2). The relative electrostatic free energies parallel the relative experimental solvation free energies. That the total free energies do not correspond as well may be a reflection of the difficulties in calculating the van der Waals

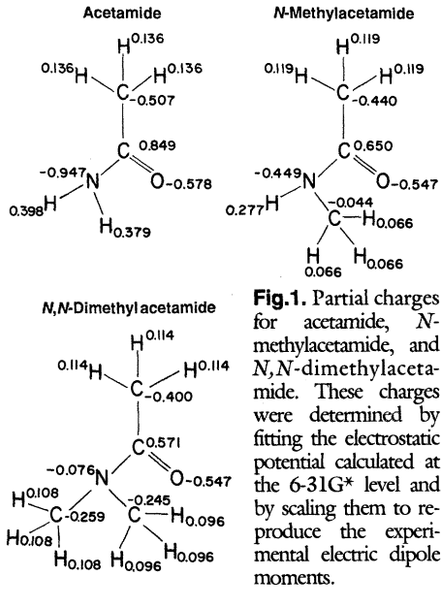


Fig. 1. Partial charges for acetamide, *N*-methylacetamide, and *N,N*-dimethylacetamide. These charges were determined by fitting the electrostatic potential calculated at the 6-31G* level and by scaling them to reproduce the experimental electric dipole moments.

Table 1. Potentials for amino acid side chains. Free energies and net free energies [$\Delta(\Delta G)$] are given for transformation of a reference state molecule into the indicated solute molecule. The electrostatic (Elec) free energy was determined with thermodynamic windowing. The van der Waals (VdW) energies were determined by two simulations, each with 30,000 steps at a Δt = 0.001 psec for the transformations of reference into solute and of solute into reference. The VdW free energies given are the average values and the standard deviation for these two simulations. Experimental free energy (Expt) are from Wolfenden's study of side-chain analogs of the amino acids (*13*). For those cases in which we have directly simulated the systems studied by Wolfenden (3-methyl indole, isobutane, phenol, acetic acid, and acetamide), experimental values are reported. In the cases where we stimulated only the amino acids directly. Expt shows the values for appropriate side-chain analogs. In each case, we took the difference in the values in column 3, table I, of (*13*) and multiplied these by $2.303RT$ = 1.363 kcal/mol. For the charged side chains, we did not decouple the energies, so only the net free energy is reported. There are larger absolute errors for the free energies for ionic groups and relevant experiments (*21*), but because of the uncertainty in the reaction-field corrections and solvation free energies in the neutral reference, we do not report experimental values for these.

Solute	Reference	Free energies (kcal/mol)			
		Elec	VdW	$\Delta(\Delta G)$	Expt
Arginine (+)	Arginine			−38.28 ± 0.71	
Arginine	Alanine	−12.51 ± 0.28	−1.89 ± 1.21	−14.40 ± 1.49	−12.86
Cysteine	Alanine	−2.02 ± 0.14	−1.02 ± 0.23	−3.04 ± 0.37	−3.19
Histidine(+)	Histidine			−35.49 ± 0.38	
Histidine	Alanine	−11.02 ± 0.51	−2.30 ± 0.71	−13.32 ± 0.20	−12.19
Leucine	Alanine	0.30 ± 0.05	0.35 ± 0.16	0.65 ± 0.20	
Lysine(+)	Lysine			−58.85 ± 0.60	
Lysine	Alanine	−7.03 ± 0.28	0.28 ± 0.18	−6.75 ± 0.10	−6.33
Phenylalanine	Alanine	−2.35 ± 0.11	−0.51 ± 0.34	−2.87 ± 0.45	−2.71
Serine	Alanine	−6.16 ± 0.21	−0.88 ± 0.15	−7.04 ± 0.35	−7.02
Tryptophan	Alanine	−4.10 ± 0.06	−3.19 ± 0.40	−7.29 ± 0.35	
Methane	Nothing	0.03 ± 0.02	2.14 ± 0.94	2.12 ± 0.90	1.95
3-Methylindole	Methane	−3.11 ± 0.21	−3.58 ± 0.73	−6.68 ± 0.93	−7.85
Isobutane	Methane	−0.11 ± 0.02	1.18 ± 0.33	1.08 ± 0.35	0.34
Phenol	Benzene	−3.80 ± 0.15	−0.76 ± 0.04	−4.57 ± 0.19	−5.36
Acetamide	Methane	−8.93 ± 0.16	−1.37 ± 0.37	−10.30 ± 0.21	−11.66
Acetate(−)	Acetamide			−71.86 ± 0.88	
Acetic acid	Methane	−6.80 ± 0.30	−2.30 ± 0.81	−9.03 ± 0.45	−8.64

contribution. The loss of potential hydrogen-bonding groups in the methyl derivatives does not significantly affect the dipole moments or the solvation free energies. This implies that the overall distribution of charge is primarily responsible for the solvation properties of these molecules.

Solvation effects can be nonadditive. For example, *p*-nitrophenol is significantly more water soluble than one might predict from considering the relative solvation energies of phenol, nitrobenzene, and benzene. This has been noted by Wolfenden (*24*) as an example of solvation nonadditivity. We determined this nonadditivity as follows. First, we calculated the atom-centered partial charge of benzene, nitrobenzene, phenol, and *p*-nitrophenol with a 6-31G* basis set (*19*). Second, an "additive" partial charge model for *p*-nitrophenol was constructed with the charges of nitrobenzene, phenol, and benzene as described and shown in Fig. 2. This is compared with the real charge distribution (*19*) determined for *p*-nitrophenol also shown in Fig. 2. The free energy calculations were then performed by perturbing the *p*-nitrophenol molecule from the additive charge distribution into its "real" charge distribution. This process resulted in a free energy change of -0.9 ± 0.2 kcal/mol, which is in agreement with the experimental nonadditivity of -0.8 kcal/mol. The results of this simulation show quantitatively how the resonance interaction between the nitro and hydroxyl groups increases the polarity of nitrophenol, which leads to a stronger interaction with water.

Nucleic acid solvation is a difficult case experimentally because the base "side chains" of nucleic acids appear to be too water-soluble to measure their free energy of solvation directly (*25*). It may be possible to measure the free energy of solvation of these bases experimentally; thus we can make a prediction of their solvation free energies. We used an N^1-methyl model for pyrimidines and an N^9-methyl model for purines and the molecular mechanical model for nucleic acids (*17*). The calculated relative and absolute solvation energies of these pyrimidines and purines (Table 2) show that the solvation free energy of thymine should be measurable with current methods but that the values for the other bases will be difficult to determine directly. These values could also be used as reference values for simpler solvation theories (*26*).

Adenine and thymine have almost equal dipole moments, yet adenine has a much larger electrostatic solvation energy (Table 2). If partial charges are derived to fit electrostatic potentials, then the higher order electrostatic multipoles are usually well represented. Both the actual quantum mechanical calculations and the partial charge models showed that adenine had a substantially larger quadrupole moment than thymine, which can account for the larger electrostatic free energy found for adenine. Thus an accurate representation of at least the first two nonvanishing multipole moments, which had earlier been shown to play a role in determining the directionality of intermolecular interactions (*27*), has been shown to also be important in describing solvation effects.

The understanding of solvation and desolvation of nucleic acids is critical if the intermolecular interactions of nucleic acids with drugs or proteins is to be described. As an example, a comparison was made between the molecular mechanical interaction energy of deoxyguanosine 5′-phosphate (dGMP) and deoxythymidine 5′-phosphate (dTMP) with staphylococcal nuclease. We calculated that dGMP would interact more strongly with the protein by about 6 kcal/mol (*28*). However, it requires 11.8 kcal/mol more energy to desolvate guanine than thymine; thus the theoretical model that includes solvation effects correctly reproduces the experimental observation that dTMP is a better inhibitor of the enzyme than dGMP (*29*).

One can also use these free energy methods to calculate protein stability (*30*). We performed a free energy perturbation calculation in which a "buried" leucine (residue 33 in the chymotrypsin numbering scheme) side chain in trypsin is changed into alanine. This is compared to a mutation of Leu into Ala in solution as a reference to determine the relative stability of the native structure relative to a denatured structure, in which this group may be fully exposed.

The simulation was performed by allowing only those residues within 8 Å of the perturbed groups to move. An all-atom representation (*17*) was used for the protein together with the computational protocols we have described to perform the simulation. The various free energy changes (Table 2) suggest that this mutation should destabilize the enzyme by about 1 kcal/mol. Further studies to "design" more stable proteins should allow the critical assessment of these methods for predictions on the effects of site-directed mutants in proteins. By simulating the Asn 155 → Ala 155 mutation in both Michaelis complexes and tetrahedral intermediates, we have been able to predict the effect of mutations on binding as well as on the catalytic properties of subtilisin (*31*).

Table 2. Solvation free energies and site-specific mutagenesis. All free energies and net free energies [$\Delta(\Delta G)$] are given for transformation of the reference state into solute or residue. The electrostatic free energy (Elec) was determined with 20 windows and 500 steps of equilibration and 500 steps of data collection and $\Delta t = 0.002$ psec. The van der Waals energies (VdW) were determined with 30,000 steps at a $\Delta t = 0.001$ picoseconds for each transformation of solute (residue) into reference and of reference into solute (residue). The values given are the average values for these two simulations with the standard deviation. The experimental free energy differences in kcal/mol are from (*23*). The electric dipole moments were calculated with the solute charge model (*19*). As shown in (*19*), this value is within 1 to 2% of the actual quantum mechanically calculated value. The quadrupole moments in the principal axis system were calculated with the solute charge model (*19*). For the adenine into guanine and thymine into cytosine transformations, the perturbations were done without decoupling the electrostatic and van der Waals contributions.

Solute (residue)	Reference	Free energies (kcal/mol)				μ (D)	Q_a, Q_b, Q_c (Buckingham)
		Elec	VdW	Δ(ΔG)	Expt		
			Pyrimidines and purines				
Adenine	Methane	−10.55 ± 0.40	−4.17 ± 0.35	−14.73 ± 0.76		2.3	15.2, −10.4, −4.8
Thymine	Methane	−6.57 ± 0.25	−3.04 ± 0.90	−9.61 ± 1.15		3.4	−10.3, 7.3, 2.9
Cytosine	Thymine			5.24 ± 0.33		5.5	13.3, −8.5, −4.8
Guanine	Adenine			6.95 ± 0.54		6.4	25.9, −19.7, −6.2
			Amides				
Acetamide	Methane	−8.93 ± 0.16	−1.37 ± 0.37	−10.30 ± 0.21	−11.77 ± 0.21		
N-Methylacetamide	Methane	−9.52 ± 0.28	−0.77 ± 0.48	−10.29 ± 0.76	−12.17 ± 0.20		
N,N-Dimethylacetamide	Methane	−8.35 ± 0.39	−0.82 ± 0.49	−9.16 ± 0.10	−10.64 ± 0.30		
			Site-specific mutagenesis				
Leucine (solution)	Alanine	0.30 ± 0.05	0.35 ± 0.16	0.65 ± 0.20			
Leucine (trypsin)	Alanine	−0.02 ± 0.01	−0.33 ± 0.32	−0.33 ± 0.34			

The studies of Wong and McCammon (*6*), Lybrand *et al.* (*5*), and Bash *et al.* (*10*), have shown that solvation free energies are critical in determining the strength of ligand binding. For example, we found (*10*) a significantly different relative solvation energy for anionic phosphonamidate as compared to phosphonate ester, in contrast to results on neutral phosphoramidates and esters (*32*). These calculations were instrumental in showing why phosphonamidates bind more strongly to thermolysin by approximately 4 kcal/mol than do the corresponding esters (*33*).

The statistical mechanical formulations of free energy differences described by Eqs. 1 and 2 are exact and any errors that result from a computational implementation of these equations are due to the accuracy in representing the Hamiltonian and sampling limitations. For the problems described above, we have shown that the first potential source of error is not a problem. The use of an empirical force field (*17*) that contains only exchange repulsion, dispersion, and electrostatic terms has led to simulated solvation free energies that are in good agreement with experiment. This suggests that aqueous solvation may have contributions from: (i) a cavity-exchange repulsion term in which the free energy is raised by incorporating a solute molecule in water; (ii) dispersion attraction between the solute and the surrounding water molecules; and (iii) electrostatic interactions between solute and solvent. The TIP3P water model (*14*) used here had an enhanced dipole moment ($\mu \approx 2.35$ D) compared to gas-phase water (1.85 D), whereas the solute molecules had partial charges scaled to reproduce relevant gas-phase dipole moments. The fact that no polarization of the solute was included in the model does not appear to be a problem in reproducing experimental solvation free energies of neutral molecules. In addition, another useful feature of the free energy perturbation method was that solvation energies could be calculated with a low statistical error. Thus the method can be used to refine the Hamiltonian representation (partial charges and other parameters) of the system (*34*).

The free energy is a state function; its value must be independent of the path between states. However, not all pathways are identical, because there will be a different set of energy barriers that must be traversed in each case that requires different amounts of time to complete the transformation. For a finite computer simulation, the practical problem is to find a pathway that can be adequately sampled with the computational resources available. For these reasons electrostatic decoupling was necessary; when both van der Waals parameters and charges were changed simultaneously, artificially large free energy changes could result because electrostatic energies would be overestimated.

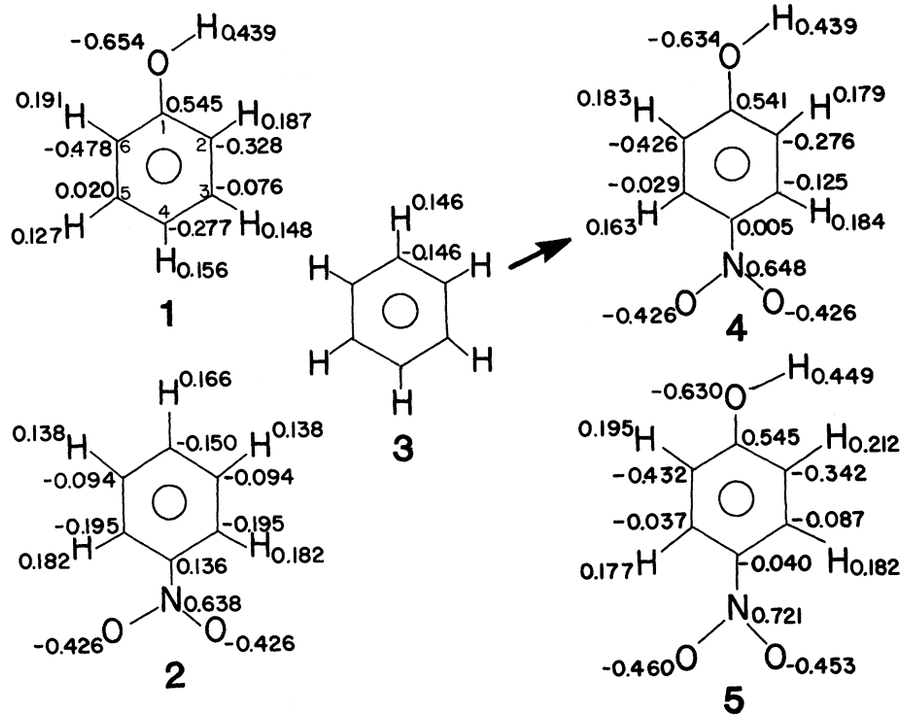


Fig. 2. Partial charges determined with the methods described in (*19*) with the use of a 6-31G* basis set for phenol (**1**), nitrobenzene (**2**), benzene (**3**), and *p*-nitrophenol (**4**). The partial charges for the additive model of *p*-nitrophenol (**5**) were determined as follows: An initial set of partial charges were taken from phenol for C-1, C-2, and C-6 and their substituents and from nitrobenzene for C-3, C-4, and C-5 and their substituents. These charges were corrected for polarization of the remote carbons by comparing, for example, the charges found at C-3, C-4, and C-5 and at the corresponding hydrogen atoms in phenol with the charges of −0.146 for carbon and 0.146 for hydrogen in benzene. For example, C-3 in phenol has a charge of −0.076 as compared to the value of −0.146 in benzene. Thus its net charge is −0.195 (from nitrobenzene) + (−0.076) − (−0.146) (correction for polarization in phenol) = −0.125. A similar analysis is done for all the ring carbon atoms and their substituents to produce the model (**4**).

Similar considerations were relevant for the evaluation of van der Waals energy changes. All such perturbations were performed with the slow growth method. For van der Waals changes, small configurational fluctuations during the simulation could result in relatively large fluctuations in the perturbed energy that could more easily cause artifacts in the statistical sampling in the window method. However, one difficulty with the slow growth method was that it was not clear how many time steps would be necessary for any particular change. To determine a reasonable value, three simulations were performed to evaluate the solvation free energy of neon with periods of 20, 30, and 40 psec in each direction. The simulations were done both by growing and disappearing neon to check for possible hysteresis. All three simulations produced values of about 2.3 ± 0.5 kcal/mol, which compared well with the experimental value of 2.4 kcal/mol, although the calculated values for growing neon were about 1 kcal/mol larger than for its disappearance. We chose the intermediate value of 30 psec (30,000 steps at 0.001-psec time steps) for all van der Waals calculations. The use of 30,000 steps for all our van der Waals simulations could be questioned, since more sampling should be required for 3-methylindole than neon. However, our interest was to test these methods on as many cases as possible and not to concentrate on only one or two examples. One would expect less hysteresis if the total number of steps was increased for larger changes (*20*).

The ability to simulate free energies rather than enthalpies in the complex systems studied here is an exciting advance. Not only are there much smaller statistical errors inherent in calculating free energies than enthalpies, but free energy is the quantity of most relevance in physical chemical experiments.

REFERENCES AND NOTES

1. J. P. M. Postma, H. J. C. Berendsen, J. R. Haak, *Faraday Symp. Chem. Soc.* **17**, 55 (1982).
2. A. Warshel, *J. Phys. Chem.* **86**, 2218 (1982); _____ and S. Russell, *Q. Rev. Biophys.* **17**, 283 (1984).

3. B. L. Tembe and J. A. McCammon, *Comput. Chem.* **8**, 281 (1984).
4. W. L. Jorgensen and C. Ravimohan, *J. Chem. Phys.* **83**, 3050 (1986).
5. T. Lybrand, J. A. McCammon, G. Wipff, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 833 (1986).
6. C. F. Wong and J. A. McCammon, *J. Am. Chem. Soc.* **108**, 3830 (1986).
7. M. Mezei and D. L. Beveridge, *Free Energy Simulations* (New York Academy of Sciences, New York, in press).
8. U. C. Singh, F. Brown, P. A. Bash, P. Kollman, *J. Am. Chem. Soc.*, in press.
9. U. C. Singh, P. K. Weiner, J. W. Caldwell, P. A. Kollman, AMBER (UCSF), version 3.0, Department of Pharmaceutical Chemistry, University of California, San Francisco, 1986.
10. P. A. Bash, U. C. Singh, F. K. Brown, R. Langridge, P. A. Kollman, *Science* **235**, 574 (1987).
11. R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954).
12. J. P. Valleau and G. M. Torrie, in *Modern Theoretical Chemistry*, B. Berne, Ed. (Plenum, New York, 1977), vol. 5, p. 137.
13. R. Wolfenden, L. Andersson, P. M. Cullis, C. C. B. Southgate, *Biochemistry* **20**, 849 (1981).
14. W. L. Jorgensen, J. Chandrasekhar, J. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
15. H. J. C. Berendsen *et al.*, *ibid.* **81**, 3684 (1984)
16. W. F. van Gunsteren and H. J. C. Berendsen, *Mol. Phys.* **34**, 1311 (1977).
17. S. J. Weiner, P. A. Kollman, D. T. Nguyen, D. A. Case, *J. Comput. Chem.* **7**, 230 (1986).
18. A van der Waals parameter for sp^2 carbons of $R^* = 1.90$ Å, $\epsilon = 0.09$ kcal/mol [which is closer to the value suggested by D. E. Williams and T. L. Starr, *Comput. Chem.* **1**, 173 (1977)], is used instead of the $R^* = 1.85$ Å, $\epsilon = 0.12$ kcal/mol of (*17*).
19. U. C. Singh and P. A. Kollman, *J. Comput. Chem.* **5**, 129 (1984).
20. T. P. Straatsma, H. J. C. Berendsen, J. P. M. Postma, *J. Chem. Phys.* **85**, 6720 (1986).
21. R. Pearson, *J. Am. Chem. Soc.* **108**, 6109 (1986).
22. D. L. Beveridge and G. Schnuelle, *J. Phys. Chem.* **79**, 2562 (1975).
23. R. Wolfenden, *Biochemistry* **17**, 201 (1978).
24. ———, *Science* **222**, 1087 (1983).
25. P. M. Cullis and R. Wolfenden, *Biochemistry* **20**, 3024 (1981).
26. D. Eisenberg and A. McLachlan, *Nature (London)* **319**, 199 (1986).
27. P. Kollman, *Acc. Chem. Res.* **10**, 465 (1977).
28. J. Caldwell and P. Kollman, unpublished results.
29. P. C. S. Fuchs and C. Anfinsen, *J. Biol. Chem.* **242**, 3063 (1967).
30. C. F. Wong and J. A. McCammon, in *Structure, Dynamics and Function of Biomolecules*, A. Ehrenberg and R. Rigler, Eds. (Springer-Verlag, Berlin, 1986).
31. S. Rao, P. A. Bash, U. C. Singh, P. Kollman, unpublished results.
32. R. Wolfenden and R. Williams, *J. Am. Chem. Soc.* **105**, 1028 (1983).
33. P. A. Bartlett and C. K. Marlowe, *Science* **235**, 569 (1987).
34. W. F. van Gunsteren, P. Kollman, H. J. C. Berendsen, unpublished results.
35. All calculations were performed on the Cray-XMP/48 at the San Diego Supercomputer Center (SDSC) under grants awarded to P.A.B. and R.L. (SDSC grant) and P.A.K. (NSF-DMB-8419883). Grants NIH-CA-25644 and NIH-GM-29072 to P.A.K. and use of the UCSF Computer Graphics Laboratory (NIH RR-1081 to R.L.) are acknowledged. We thank H. Berendsen and W. van Gunsteren for useful discussions.

20 October 1986; accepted 4 March 1987

On the Prevalence of Room-Temperature Protein Phosphorescence

JANE M. VANDERKOOI, DOROTHY B. CALHOUN, S. WALTER ENGLANDER

A large number of proteins were tested for the property of intrinsic phosphorescence in deoxygenated aqueous solution at room temperature. The majority of proteins exhibit phosphorescence under normal solution conditions. Phosphorescence lifetimes from 0.5 millisecond to 2 seconds were observed in three-fourths of the proteins tested. The lifetime appears to correlate with relative isolation of the tryptophan indole side chain from solvent. With few exceptions, proteins in general can be expected to display a phosphorescence lifetime greater than 30 microseconds. This widespread characteristic of proteins has been largely overlooked because long-lived phosphorescence is highly sensitive to quenching by low levels of dissolved oxygen in solution. Protein phosphorescence offers a new time domain and a far wider dynamic range than has been used before for photoluminescence experimentation.

EXPERIMENTAL APPROACHES BASED on the measurement of intrinsic protein fluorescence have proven immensely useful for a broad range of problems in macromolecular science. However, tryptophan fluorescence, in which photon emission occurs on the nanosecond time scale, is only one of the possible consequences of photoexcitation. Intersystem crossing to the triplet state is another, and this can result in long-lived phosphorescence emission. The ability to measure protein phosphorescence in liquid solutions would greatly expand the potential for photoluminescence studies of protein interactions and dynamics on a longer and often more pertinent time scale. However, this possibility has received little attention from the biochemists, since phosphorescence is thought to be largely a property of the solid state. The results reported here demonstrate that this is not the case.

Debye and Edwards (*1*) reported in 1952 that phosphorescence can be observed from proteins at cryogenic temperatures. Under these conditions phosphorescence lifetimes for a variety of proteins attain a fairly standard value of 5.5 seconds. Twenty years later, Saviotti and Galley (*2*) found that two proteins, alcohol dehydrogenase from horse liver and alkaline phosphatase from *Escherichia coli*, phosphoresce in aqueous solution at room temperature with lifetimes on the order of 1 second. Kai and Imakubo (*3*) found three additional proteins with lifetimes between 1 and 20 msec.

In studies of protein photoluminescence we showed that oxygen quenches the phosphorescence of even the extraordinarily well-protected tryptophans of alcohol dehydrogenase and alkaline phosphatase with a rate constant of $10^9 M^{-1}$ sec^{-1} (*4*). Thus when oxygen is present at normal solution concentration (250 μM in air-exposed water), phosphorescence emission with an intrinsic lifetime as long as 5 seconds will be reduced in intensity by a factor of 10^6 and exhibit a lifetime of only 4 μsec. This explains why the phosphorescence of proteins in solution has escaped detection. The original discovery of long-lived protein phosphorescence by Saviotti and Galley (*2*) was connected with a fortunate accident in this regard.

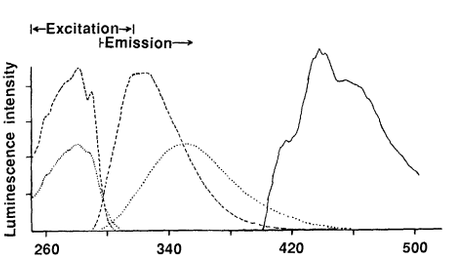


Fig. 1. Excitation and emission spectra of cod parvalbumin. The samples contained 2 mg of cod parvalbumin per milliliter in 0.01*M* sodium phosphate and 0.1*M* sodium chloride at *p*H 7.0. For phosphorescence measurements the deoxygenating system described in Table 1 was included. The ordinate for each spectrum is separately scaled for ease of visualization. Fluorescence spectra were taken with protein-bound calcium (- - -) and after addition of 3 m*M* EDTA to remove calcium (. . .). A detection wavelength of 330 nm was used with the excitation spectrum and an excitation wavelength of 280 nm was used with the emission spectrum. The phosphorescence emission spectrum (———) was taken with an excitation wavelength of 280 nm.

Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104.