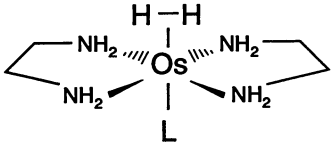
Use of a Dihydrogen Osmium Complex as a Versatile ^{1}H NMR Recognition Probe

Zai-Wei Li and Henry Taube*

A new recognition probe for biomolecules, $[en_2Os(\eta^2\text{-}H_2)]^{2+}$ (**1**; en, ethylenediamine), is reported. In aqueous solution, **1** binds readily to a variety of biomolecules, including nucleotides, RNA, amino acids, peptides, and phospholipids. In each case, binding leads to a characteristic proton nuclear magnetic resonance (^{1}H NMR) for the dihydrogen that appears in a spectral window in the range $\delta = 0$ to -20 parts per million, and as well to characteristic values of the coupling J_{HD} and of the relaxation time T_1. Small structural differences in molecules such as DGMP (2′-deoxyguanosine 5′-monophosphate) and IMP (inosine 5′-monophosphate) or Asp and Glu can readily be distinguished, such as when **1** binds to the N-7 position of the nucleobase of DGMP or IMP and when **1** binds to the carboxylate of Asp or Glu. Upon one-electron oxidation of the metal center, diamagnetic **1** is converted to a paramagnetic probe.

Proton NMR has become one of the most powerful tools for investigating biomolecules (*1*, *2*). Although higher field NMR instruments continue to be developed, the congestion of resonance signals in a small spectral range, usually between $\delta = 0$ and 10 ppm, complicates the interpretation of the observations.

The species $[trans\text{-}en_2Os(\eta\text{-}H_2)]^{2+}$, **1**, depicted below, has been described (*3*). Herein we report the results of experiments undertaken to explore its application as a recognition probe for biomolecules.



When a salt of **1** with PF_6^- or $CF_3SO_3^-$ as counterion is dissolved in water the variable ligand L is water. Water (or any other oxygen donor solvent) is rapidly replaced by other ligands in great variety. For each new composition, a characteristic value of δ for the dihydrogen unit is registered in the NMR spectrum. All of the values of δ are negative and are clearly separated from the shifts arising from protons of organic molecules. This property provides a convenient identification of the variable ligand L and also provides ready access to J_{H-D} and T_1 (H_2). The latter property can be the source of important information about the dynamics of proteins and other molecules (*4*).

The results obtained with DGMP (2′-deoxyguanosine 5′-monophosphate) illustrate several important facets of the properties of the probe. When DGMP is added to a solution of **1** (as the PF_6^- salt) in D_2O, each solute at 0.010 M, the original peak ($\delta = -13.45$ ppm) diminishes in intensity and a doublet grows in at $\delta = -13.83$ and -13.86 ppm (Fig. 1A). This phase of the reaction is complete in 10 min. We conclude that the probe is attached to the phosphate ($^3J_{PH} = 5.46$ Hz) and, because the reaction is not complete at equilibrium, that the affinity for the dinegative phosphate is not high. The equilibrium quotient, K_a (*6*), for the formation of the probe-RPO_4^{2-} complex is 3×10^2. Moreover, this attachment is labile, as appears to be the case for all σ oxygen-donor ligands. A peak at $\delta = -9.99$ ppm, already discernible after 10 min, continues to grow at the expense of the others, and after 24 hours the conversion to the most stable form is almost complete (Fig. 1B). Accompanying the growth of the peak at $\delta = -9.99$ ppm, H-8 and H-1′ shift from 7.96 to 8.65 ppm and from 6.09 to 6.32 ppm, respectively, indicating binding of the probe at N-7. Addition of a small amount of acid catalyzes exchange between η-H_2 and D_2O. For binding at N-7, J_{HD} is 16.54 Hz, a coupling constant in the range characteristic of η^2-H_2 (*7*). For this case only, the ^{1}H NMR spectra covering the usual range of values of δ are included and the NMR resonances for the NH_2 and CH_2 protons are identified. The former show structure corresponding to the three species shown in Fig. 1A.

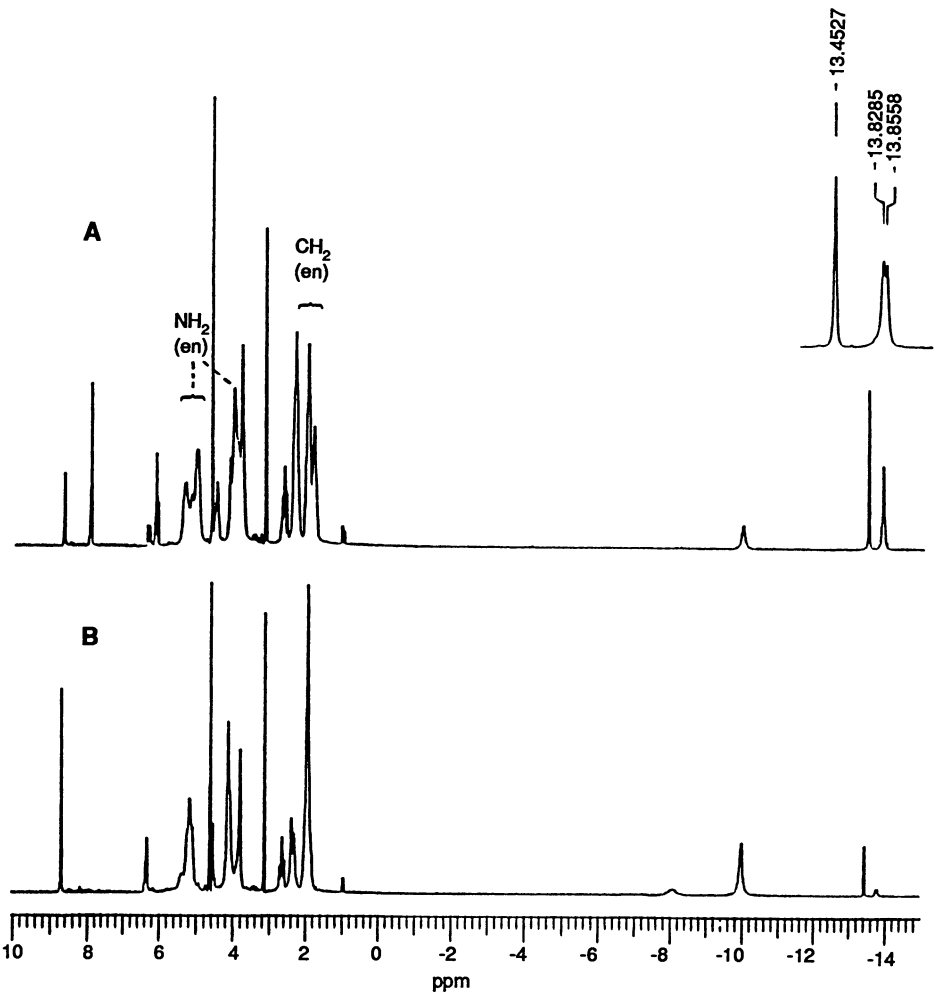


Fig. 1. Proton NMR spectra (200 MHz) of **1** with DGMP in D_2O (both 0.010 M, pD = 7). (**A**) After 10 min. (**B**) After 24 hours.

Department of Chemistry, Stanford University, Stanford, CA 94305.

*To whom correspondence should be addressed.

In many cases, partial isomerization to a cis form accompanies substitution. This secondary process does not interfere with many applications of the probe and may in fact enhance its power when the effect is further investigated. The small peak at $\delta = -7.99$ ppm (Fig. 1B) is assignable to the cis form. The equilibrium quotient for binding to the two forms is 2.9×10^3.

The ^{1}H NMR traces for several nucleotides are compared in Fig. 2, and data collected for them are summarized in Table 1. The observations with IMP (inosine 5′-monophosphate) are similar to those with DGMP. Binding to nucleobases is observed also for AMP (adenosine 5′-monophosphate), as well as for the simpler bases UMP (uridine 5′-monophosphate) and TMP (thymidine 5′-monophosphate), but in these cases the affinity is considerably less than it is for DGMP or IMP. The value of K_a for binding to cytidine 5′-monophosphate is <0.1.

Although the phosphate groups in most of nucleotides give almost identical shifts, those in AMP, ADP (adenosine 5′-diphosphate), and ATP (adenosine 5′-triphosphate) are readily distinguished (Table 1). For AMP a simple pattern is observed. At pD 8.0, the dominant form is the dinegative ion and K_a is measured as 1.9×10^2, whereas for Na_2HPO_4 it is 9×10^2. For ADP and ATP only the dominant peaks are recorded and minor peaks provide distinct evidence of multiple binding sites.

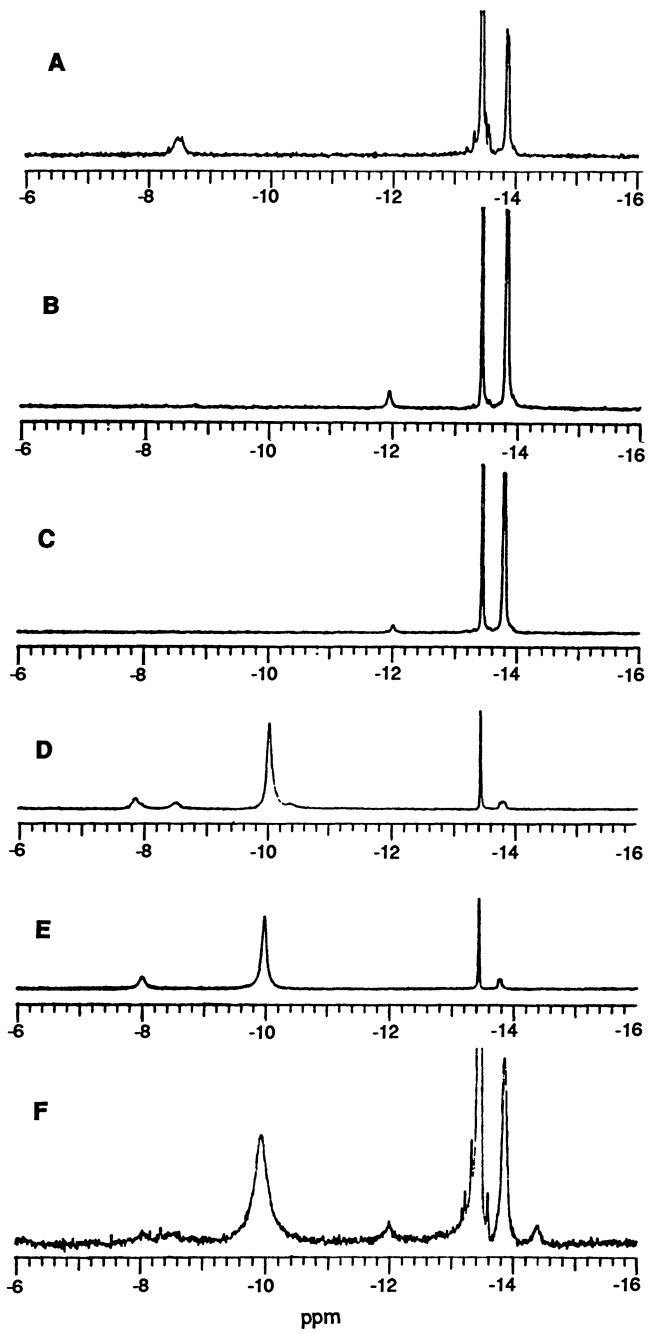


Fig. 2. Proton NMR spectra (200 MHz) of **1** (0.010 M) with nucleotides (all 0.010 M; pD = 7) and transfer RNA in D_2O 12 hours: (**A**) AMP, (**B**) UMP, (**C**) TMP, (**D**) IMP, (**E**) DGMP, and (**F**) tRNA (type X-SA, from Baker's yeast, 100 units), pD = 6.5.

In Fig. 2 are included the NMR signals of the indicator ligand also for a transfer RNA. A doublet at about −13.9 ppm indicates binding at dinegative phosphate for each of the nucleotides. The signals at about −12 ppm for TMP and UMP correspond to binding at the carbonyl of the nucleobases, and there is a corresponding signal for RNA. The signal for the latter at −10 ppm corresponds to N-7 binding by DGMP or IMP. There is also a peak at −8 ppm, which corresponds to the cis signals of the N-7 binding. The major signal for AMP is at −8.45 ppm, which is close to a −8.5-ppm peak of RNA. Because there is unpairing of bases in loops (8), further investigation of the resonances may provide useful information on the unpaired regions. For RNA, the peak at −14.38 ppm corresponds to binding to bridging mononegative phosphate. The strong peak at −13.92 ppm is close to that registered for dinegative phosphate, which may be an impurity in the RNA sample.

Data on the binding of the probe to the carboxyl function of a number of amino acids are summarized in Table 2. Small structural differences, such as between Asp and Glu, can be distinguished by carboxyl binding of **1**.

When **1** reacts with imidazole in D_2O, two ^{1}H NMR peaks, −9.24 and −8.35 ppm, appear (Fig. 3A), which can be assigned to the trans and cis adducts, respectively. These features appear also with the tripeptide pGlu-His-Gly amide (Fig. 3C) and also with His (Fig. 3B), but in this case an additional peak is observed that may reflect the two possible modes of binding to the imidazole ring. Why this extra feature does not appear for the tripeptide is not clear. The patterns displayed in Fig. 3, A to

Table 1. Observed ^{1}H NMR data for binding of **1** at the nucleobase position and at the phosphate of nucleotides (Nt). Conditions: 20°C, 200 MHz, D_2O solvent, and all concentrations 0.010 M.

Nt	δ(ppm)	$^3J_{PH}$ (Hz)	$T_1(H_2)$ (ms)	K_a
	Nucleobase binding (pD 7)			
DGMP	−9.98*		15	2900
	−7.99†			
IMP	−10.02*		13	6900
	−7.87†			
AMP	−8.45		37	12
TMP	−12.01		35	7
UMP	−11.94		20	17
	Phosphate binding (pD 8)			
AMP	−13.83	5.30	84	170
ADP	−13.97	5.30	87	1200
ATP	−14.05	5.62	81	3800

*Trans. †Cis.

C, were obtained 24 hours after mixing, but in no case was equilibrium reached. The signal corresponding to the cis form continues to grow at the expense of the others. From data obtained after 10 days, a lower limit of 1.0×10^2 for the value of K_a for binding of imidazole is calculated. Equilibrium is reached even more slowly in the other two cases. After 15 days, lower limits of 1.4×10^4 and 8.1×10^2 for the values of K_a for binding of His and the tripeptide are obtained. We observe the values of T_1 (200 MHz; 20°C) to be sensitive to the change in molecular size of the species containing the binding site: For imidazole, histidine, and the peptide they are 75, 19, and 15 ms for the trans form and 68, 54, and 25 ms for the cis form, in each case, respectively. In water, binding to amide group is not observed, at least at the concentrations we used.

The two sulfur-containing amino acids Cys and Met show the effects of binding to sulfur with δ values of −8.97 and −7.86 ppm and T_1 values of 87 and 29 ms, each respectively (400 MHz). The K_a for sulfur binding for Met is 6.2×10 and for Cys is 15.

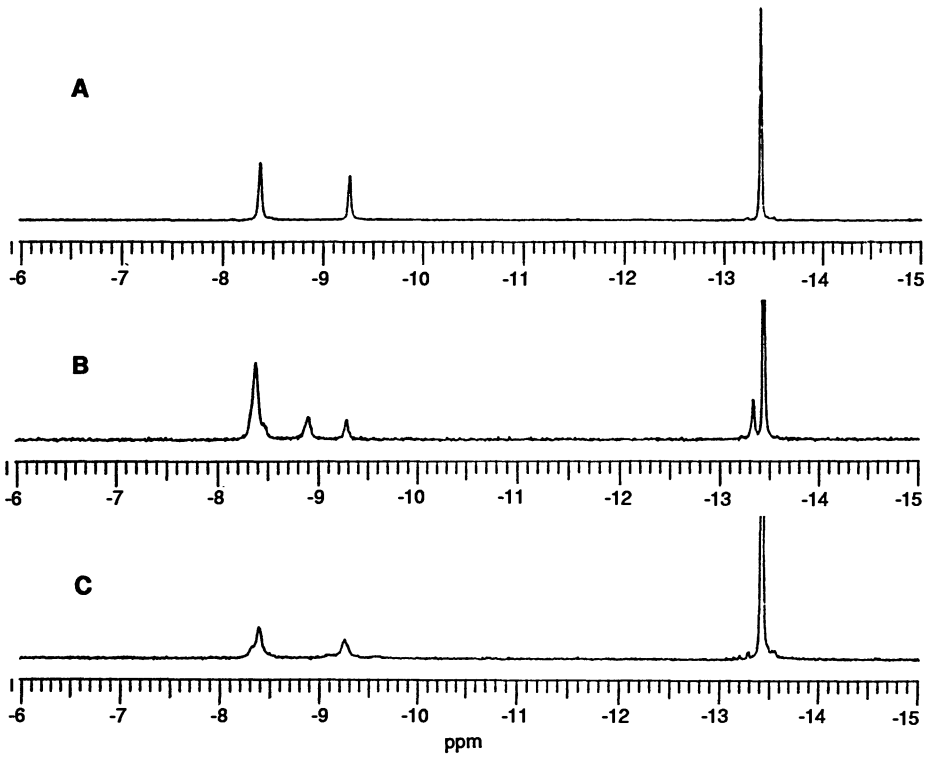


Fig. 3. Proton NMR spectra (200 Mz) of **1** with (**A**) imidazole, (**B**) His, and (**C**) pGlu-His-Gly amide in D_2O (all 0.010 M, pD = 7, 24 hours after mixing).

Table 2. Proton NMR data for binding of **1** with carboxylate in amino acids. Conditions: 400 MHz, D_2O solvent, all concentrations 0.010 M, and pD = 6.5.

Parameter	Gly	Ser	Asn	Cys	Met	His	Asp		Glu	
δ(ppm)	−13.28	−13.27	−13.25	−13.28	−13.16	−13.27	−13.24	−13.33	−13.13	−13.28
$T_1(H_2)$(ms)	124	104	112	98	78	96	96	80	116	98
K_a	8.5	19	19	7.3	12	<7	169	157	42	27

Table 3. Proton relaxation time data of **1** (0.01 M) with DGMP (0.04 M) in D_2O. T_{1o} and T_1 are the relaxation times in the absence and presence of Os(III), respectively, and T_d is the ^1H relaxation time for the probe bound to N-7 of DGMP, before oxidation ($1/T_1p = 1/T_1 - 1/T_{1o}$) (*12*).

Time	H-8	H-1′	H-2′	H-3′	H-4′	H-5′
T_{1o} (ms)	712	1257	327	817	1048	314
T_d (ms)	445	1043	228		872	235
T_1 (ms)	287	443	232		516	255
$1/T_1p$ (s^{-1})	2.08	1.46	1.25		0.99	0.74

Characteristic resonances are also observed for phosphate binding in phospholipids. A doublet (−13.90 ppm; $^3J_{PH}$ = 6.88 Hz) characteristic of phosphate binding was obtained for a sample of phosphatidyl-L-serine sodium salt dissolved in CD_3OD (4.0×10^{-3} M), the carboxylate binding peak for this molecule appearing at −12.97 ppm. Although the phosphate is mononegative, the K_a value is 5×10^2; the enhanced stability is attributable to the change in solvent, D_2O to CD_3OD. Two others, sphingomyelin and L-*d*-phosphatidylethanolamine dilauroyl, both give phosphate binding peaks at δ = −13.90 ppm, but with distinguishable $^3J_{PH}$ values of 6.72 and 6.84 Hz, respectively.

On oxidation to Os(III), the metal center becomes paramagnetic and advantage can be taken of this change (*9–11*). In this process H_2 is lost, and commonly, when Os(II) ammines are oxidized to Os(III), the bonds to σ donor ligands, of the type featured in most of the cases we have described, become more robust. In most cases, again for σ donors, the oxidation occurs without relocation of the metal center.

In an experiment in which **1** (0.010 M) and GMP (0.040 M) were mixed, after 5 hours all of the probe was bound to N-7 of the nucleobase. The NMR tube was opened to the atmosphere for 24 hours, and T_1 values for the protons of the nucleotide were then measured.

The values of $(1/T_1 - 1/T_{1o})$ (Table 3) are in the order expected for a dipole-dipole interaction (*12*). The measurements of T_1 for the diamagnetic form of the probe located at the same position are included.

A striking feature of the results described is that the probe leads to complexes that are formed readily but which are sufficiently nonlabile so that characteristic, rather than average, values of δ are registered. Moreover, the usefully high affinity of the probe for a wide range of ligands, both saturated and unsaturated, is without a parallel among dipositive ions. The affinity pattern is quite different from that shown by $[Ru(NH_3)_5H_2O]^{2+}$, especially in the capacity for backbonding (*13*). The propensity of $Os(NH_3)_5^{2+}$ for backbonding is very much greater than that of $[Ru(NH_3)_5H_2O]^{2+}$, and thus we can be certain that K_a for the association of CH_3CN with $Os(NH_3)_5^{2+}$ in water would much exceed the value determined for $Ru(NH_3)_5^{2+}$, namely $\sim 1 \times 10^6$ (*14*). An upper limit to the corresponding value for $[Os(NH_3)_4(\eta^2\text{-}H_2)]$ is 5×10^3, which is the value measured in acetone. The diminished affinity of the probe for π acids when a molecule of ammonia is replaced by H_2 is expected when H_2 itself acts as a π acid: both are competing for π *d* electrons (*15*). On the other hand, K_a for

association of the probe with Cl^- is ~2.0 × 10^2 (*16*), whereas for $Ru(NH_3)_5^{2+}$ it is only 3 (*17*). The comparison illustrates the electron withdrawing power of η^2-H_2 acting as a π acid.

Dihydrogen-bound η^2-H_2 as a ligand may differ significantly from other ligands in that it may undergo substantial structural changes when the coligands are altered. The values of J_{HD} we have observed cover the range from >20 Hz to a value of <2, suggesting a change in H–D distance, and thus a change in the capacity of HD or H_2 to act as a π acid. If the charge on H_2 remains constant, this tendency is expected to increase as the H–H distance increases, but in remaining attached to the metal, and in the limit of being converted to a dihydride, the antibonding orbitals are filled, so that the value of the distance optimum for backbonding would lie between that of free H_2 and that found in dihydrides.

REFERENCES AND NOTES

1. O. Jardetsky and G. C. K. Roberts, Eds., *NMR in Molecular Biology* (Academic Press, New York, 1981).
2. K. Wüthrich, Ed., *NMR of Proteins and Nucleic Acids*, (Wiley, New York, 1986).
3. Z.-W. Li and H. Taube, *J. Am. Chem. Soc.* **113**, 8946 (1991).
4. See (*1*), chap. 12.
5. All spectra were recorded at 20°C.
6. The equilibrium quotient is defined as [Pr-L]/([Pr-H_2O] [L]), where [L] is the concentration of the free ligand and where [Pr-L]/[Pr-H_2O] is determined from the relative abundance of the two forms of the probe (Pr) as obtained from the ^1H NMR spectrum.
7. G. J. Kubas, *Acc. Chem. Res.* **21**, 120 (1988); R. H. Crabtree, *ibid.* **23**, 95 (1990).
8. L. Stryer, *Biochemistry* (Freeman, New York, ed. 2, 1981).
9. G. N. LaMar, W. D. Horrocks, Jr., R. H. Holm, *NMR of Paramagnetic Molecules* (Academic Press, New York, 1973).
10. A. S. Mildvan, J. Granot, G. M. Smith, M. N. Liebman, *Adv. Inorg. Biochem.* **2**, 211 (1980).
11. M. Pyle and J. K. Barton, *Prog. Inorg. Chem.* **38**, 413 (1990).
12. A. S. Mildvan and R. K. Gupta, *Methods Enzymol.* **49G**, 322 (1978).
13. H. Taube, in *Survey of Progress in Chemistry*, A. F. Scott, Ed. (Academic Press, New York, 1973), vol. 6, 1–46.
14. J. F. Wishart, H. Taube, K. J. Breslauer, S. S. Isied, *Inorg. Chem.* **23**, 2997 (1984).
15. S. S. Isied and H. Taube, *ibid.* **13**, 1545 (1974).
16. Z.-W. Li and H. Taube, unpublished results.
17. H. S. Lim, D. J. Barclay, F. C. Anson, *Inorg. Chem.* 11, 1460 (1972).

18. Support of this work by National Institutes of Health grant GM13638-24 is gratefully acknowledged.

30 October 1991; accepted 4 February 1992

An SCF Solvation Model for the Hydrophobic Effect and Absolute Free Energies of Aqueous Solvation

Christopher J. Cramer and Donald G. Truhlar

A model for absolute free energies of solvation of organic, small inorganic, and biological molecules in aqueous solution is described. This model has the following features: (i) the solute charge distribution is described by distributed monopoles, and solute screening of dielectric polarization is treated with no restrictions on solute shape; (ii) the energetic effects of cavity formation, dispersion interactions, and solute-induced restructuring of water are included by a semiempirical cavity surface tension; and (iii) both of these effects are included in the solute Hamiltonian operator for self-consistent field (SCF) calculations to allow solvent-induced electronic and geometric distortion of the solute. The model is parameterized for solutes composed of H, C, N, O, F, P, S, Cl, Br, and I against experimental data for 150 neutral solutes and 28 ions, with mean absolute errors of 0.7 and 2.6 kilocalories per mole, respectively.

Solubilities of molecules are directly related to the free energy of solvation, and reaction rates in solution are directly related to the differential free energy of solvation of transition states and reactants (*1*). Solvation free energies in aqueous solution are biochemically important for structural and metabolic equilibria and kinetics. For example, in enzyme catalysis a substrate is generally dehydrated, and thus the free energy of desolvation is an important contributor to the free energy of activation (*2*, *3*). Unfavorable solvation energies of associating nonpolar groups may lead to hydrophobic stabilization of their mutual complexes, and this may play an important role in enzyme-substrate or protein-inhibitor binding. Desolvation and hydrophobic stabilization are also important in DNA-binding interactions (*4*). The hydrophobic effect has long been implicated in protein folding, but it is hard to correlate the solvent affinities of residues with solution data for small molecules because nonterminal protein residues in enzymes are less polar than the zwitterionic monomers (*5*). In addition, protein stability may depend on a difficult-to-assess balance of hydrophobic and other noncovalent interactions (*6*), such as electrostatics and exchange repulsion. A theoretical approach can be valuable in allowing the quantitative estimation of the energetics of such solvation effects that are not accessible to direct experimental measurement and also in allowing a self-consistent treatment of hydrophobic effects with other noncovalent and even with covalent interactions.

Semiempirical molecular orbital theory (*7*) provides an increasingly powerful approach to modeling electronic energies of gas-phase molecules and solutes and their interactions with other molecules, including water (*8*–*10*). However, bulk solvation energies result from the cooperative effect of large numbers of solute molecules. For example, the effects of aqueous solvation on the activation energies of nucleophilic substitution reactions at methyl chlorides converged with respect to increasing the number of water molecules to within only 1 to 2 kcal mol^{-1} when this number was increased from 54 to 66 (*11*). Molecular dynamics and Monte Carlo simulation studies of aqueous solvation and the hydrophobic effect routinely involve 250 or more solvent molecules (*12*), and proper solvation of biomolecules may require many more than 1000 waters (*13*). One way to make the problem more manageable for small charged and polar solutes is to use models that treat the solvent as a continuum dielectric with an electric polarization field. Such effects are typically treated in terms of the solvent dielectric constant, and this approach has a long history for solvent polarization by monatomic ions (*14*), dipoles (*15*), and systems composed of distributed monopoles (*16*–*20*). Such effects can be incorporated into molecular orbital theory by including local field terms in the Hamiltonian, and the promise of local-field SCF models for biological chemistry has long been recognized (*21*). However, the biological applications of such models are still rather limited, perhaps because such methods are not well suited to treating the hydrophobic effect.

The hydrophobic effect is more subtle than the solvent polarization effects that dominate free energies of solvation for small charged and polar solutes, and it involves additional physical effects not accounted for in dielectric constants. It can, however, be treated by continuum models that assume a proportionality (*2*, *19*, *22*, *23*) between hydrophobicity (or all or part of the free energy of solvation) and the surface area of the solute, or—more physically—the sol-

C. J. Cramer, U.S. Army Chemical Research Development and Engineering Center, Research Directorate, Physics Division, Chemometric and Biometric Modeling Branch, Aberdeen Proving Ground, MD 21010.
D. G. Truhlar, Department of Chemistry, Supercomputer Institute, and Army High Performance Computing Research Center, University of Minnesota, Minneapolis, MN 55455.