JOURNAL OF PHYSICAL ORGANIC CHEMISTRY, VOL. 10, 563–576 (1997)

FREE ENERGIES OF SOLVATION IN CHLOROFORM AND WATER FROM A LINEAR RESPONSE APPROACH†

NORA A. McDONALD, HEATHER A. CARLSON and WILLIAM L. JORGENSEN*
Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, USA

Monte Carlo statistical mechanics simulations were used to compute absolute free energies of solvation in chloroform for 16 organic molecules. The intermolecular interactions were described by classical potential functions consisting of Coulomb and Lennard–Jones interactions. The partial charges for the solutes were derived from fitting to the electrostatic potential surfaces of *ab initio* 6–31G* wavefunctions. First, free energy perturbation (FEP) calculations yielded relative free energies of solvation. These were converted to absolute quantities through perturbations to the reference molecule, methane, which was annihilated. The average error in the FEP-computed free energies of solvation is 0·8 kcal mol^{-1}. Then, a linear response equation, which contains terms proportional to the Lennard–Jones (van der Waals) and Coulombic components of the solute–solvent energy and to the solvent-accessible surface area of the solute, was optimized and reproduced both the FEP-calculated and experimental free energies of solvation with average errors of *ca* 0·5 kcal mol^{-1}. In addition, an existing solute dataset for water, which had previously been fitted to the same equation, was expanded from 16 to 35 molecules. The fit of the Monte Carlo results for this set of molecules in TIP4P water to the experimental free energies of hydration yielded an average error of 0·8 kcal mol^{-1}. Combination of the predictions of free energies of solvation in water and chloroform yields partition coefficients, log *P*, with an average error of 0·3–0·4 log unit. © 1997 John Wiley & Sons, Ltd.

J. Phys. Org. Chem. **10**, 563–576 (1997) No. of Figures: 5 No. of Tables: 10 No. of References: 33
Keywords: free energies of solvation; chloroform; water; linear response
Received 17 October 1996; revised 24 January 1997; accepted 27 January 1997

INTRODUCTION

It was discovered in 1985 that differences in free energies of hydration for organic molecules could be computed with high precision in simulations using statistical perturation theory with explicit representation of solute and solvent molecules.[1] This opened up routine, quantitative theoretical studies of a wide range of fundamental problems in chemistry and biochemistry involving equilibria.[2] Importantly, two such perturbations for a pair of ligands unbound and bound to a receptor combine in a thermodynamic cycle to yield relative free energies of binding.[2,3] Absolute free energies of solvation are more difficult to compute since rather than interconverting two solutes, one solute must be perturbed to nothing (annihilated).[4] In practice, it is common to create a ladder of conversions to progressively smaller molecules ending with methane, which is annihilated.[4,5] A mutation between two molecules is further broken down into a series of substeps or windows, which increase in number as the dissimilarity of the two molecules increases. In spite of the computational demands, comparisons between computed and experimental free energies of hydration have now become an important test for the quality of force fields intended for use in molecular dynamics (MD) and Monte Carlo (MC) simulations of condensed-phase systems.[6] However, such free energy perturbation (FEP) calculations have practical limitations because mutations between systems which are very different structurally may require a prohibitive number of windows and/or it may be desired to treat a prohibitively large number of molecules for a structure–activity study. Consequently, approximate methods have been developed for more rapid calculation of free energies of solvation.[7–11]

Recently, Åqvist and co-workers[12] introduced a procedure based on linear response (LR) theory for estimating free energies of binding. In this model, the free energy of interaction of a solute with its environment is given by half the electrostatic (Coulombic) energy plus the van der Waals energy scaled by an empirical parameter, α. For binding a ligand to a protein, the differences in the interactions between the ligand in the unbound state and bound in the complex then provide an estimate of the free energy of binding, ΔG_b, via the equation

$$\Delta G_b = \tfrac{1}{2}\langle \Delta U_{elec} \rangle + \alpha \langle \Delta U_{vdw} \rangle \quad (1)$$

* *Correspondence to:* W. L. Jorgensen.
† Dedicated to Professor Frederick G. Bordwell.
Contract grant sponsor: National Institutes of Health.
Contract grant sponsor: National Science Foundation.

 CCC 0894–3230/97/070563–14 $17.50

A value of 0·162 for α was determined empirically by fitting to experimental data for a series of inhibitors.[12] The required energy components were obtained from molecular dynamics simulations for the inhibitors in water and for the protein–inhibitor complexes in water. Key advantages over FEP methods are (a) absolute free energies of binding are readily obtained and (b) only simulations at the endpoints of a mutation are requred, i.e. the simulations for the intermediate windows are eliminated, which allows much easier application to structurally diverse sets of molecules. In spite of the approximations in equation (1), the approach has yielded promising results for several applications.[12, 13]

Subsequently, Carlson and Jorgensen[14] tried the same approach to compute free energies of hydration. However, since both α and the coefficient of half for the electrostatic term in equation (1) are positive, and since both the electrostatic and van der Waals contributions to the solute-solvent energy are always negative, there was no provision for obtaining a positive free energy of hydration. Consequently, a cavitation term proportional to the solvent-accessible surface area (SASA) was added to the LR equation and, furthermore, the coefficient for the electrostatic term was allowed to vary. Tests of the resulting equation:

$$\Delta G_{\mathrm{hyd}} = \beta\langle U_{\mathrm{elec}}\rangle + \alpha\langle U_{\mathrm{vdw}}\rangle + \gamma\langle \mathrm{SASA}\rangle \qquad (2)$$

were performed on a set of 17 diverse molecules: methane, ethane, methyl chloride, methanethiol, dimethyl ether, acetonitrile, methylamine, methanol, acetone, methyl acetate, acetic acid, acetamide, benzene, aniline, nitrobenzene, chlorobenzene and pyridine. All were modeled with Lennard–Jones parameters from the OPLS all-atom force field and partial charges derived from fitting to electrostatic potential surfaces (EPS) from *ab initio* RHF/6–31G* wavefunctions.[15] The energy components and SASA values came from MC simulations in TIP4P water.[16] A fit of equation (2) to the experimental data, which cover a 12 kcal mol^{-1} range (1 kcal = 4·184 kJ), gave an rms error of 0·76 kcal mol^{-1} for the 17 solutes.

The present study extends this work by considering the applicability of equation (2) for free energies of solvation in an aprotic solvent, chloroform, and by expanding the number of molecules treated in water to 35. Combination of free energies of solvation in water and chloroform allows computation of the chloroform–water partition coefficient for the solute, log *P*. This provides a gauge of lipophilicity that can be valuable in determining biological activity.[17] For chloroform, the experimenal free energies of solvation were obtained by combining free energies of hydration and free energies of transfer from water to chloroform from the log *P* measurements. The experimental data are somewhat limited and a dataset of 16 molecules was used: methanol, phenol, methylamine, dimethylamine, trimethylamine, aniline, pyridine, acetaldehyde, acetone, acetic acid, acetamide, methyl acetate, acetonitrile, benzene, chlorobenzene and cyclohexane. For water, the expanded dataset incorporates ethylamine, dimethylamine, trimethylamine, dimethyl sulfide, methyl fluoride, methylene chloride, chloroform, tetrahydrofuran, ethanol, isopropyl alcohol, *tert*-butyl alcohol, phenol, acetaldehyde, ethylene, propylene, naphthalene, propane, butane and cyclohexane.

Prior computational work on computing free energies of solvation or log *P* for chloroform should be noted. Several groups have computed relative log *P* values using FEP methods with explicit representation of the solvent in MD or MC simulations.[18–20] Absolute log *P* values have been computed by Reynolds[21] for 30 solutes using Still's GB/SA models with a continuum representation of the solvents. Quantum mechanical treatments for solvation in chloroform have also been developed with continuum models of the solvent in self-consistent reaction field calculations.[22, 23]

COMPUTATIONAL DETAILS

Molecular structures and charges

The structures, parameters and charges for the original dataset have been reported previously.[5, 14] As before, standard geometries based on microwave structures[24] were used for the additional solutes, and are presented in Figure 1. The methyl groups have been symmetrized with C–H bond lengths of 1·09 Å and bond angles of 109·47°, and they have been oriented in the lowest energy conformation.

The computational protocol for generating the EPS charges was the same as that used for the original dataset;[5] briefly, single-point *ab initio* calculations were carried out with Gaussian 92 at the HF/6–31G* level and the EPS charges were generated from CHELPG calculations.[15] The charges are presented in Table 1 and the resultant calculated dipole moments are listed in Table 2 along with the experimental values.[25] The charges for any hydrogen atoms in CH_3 or NH_2 groups which could interconvert due to rotation have been averaged. This causes the differences between the computed dipole moments from the 6–31G* wavefunctions and from the partial charge models in Table 2. As always, the 6–31G* dipole moments are typically 20% greater than experimental gas-phase values, which is desirable for condensed-phase simulations.

Monte Carlo simulations

All solutes were represented by an all-atom force field of the common Coulomb plus Lennard–Jones format. The solute–solvent interaction energies are given by the equation

$$\Delta E_{ab} = \sum_{i}^{\text{on } a} \sum_{j}^{\text{on } b} \left\{ \left(\frac{q_i q_j}{r_{ij}} \right)^2 + 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \right\} \qquad (3)$$

where $\Delta E_{\mathbf{ab}}$ is the potential energy between molecules *a* and *b* which contain interaction sites *i* and *j* with partial charges *q* and Lennard–Jones parameters σ and ε. Geometric combining rules are used for σ and ε, and their values were

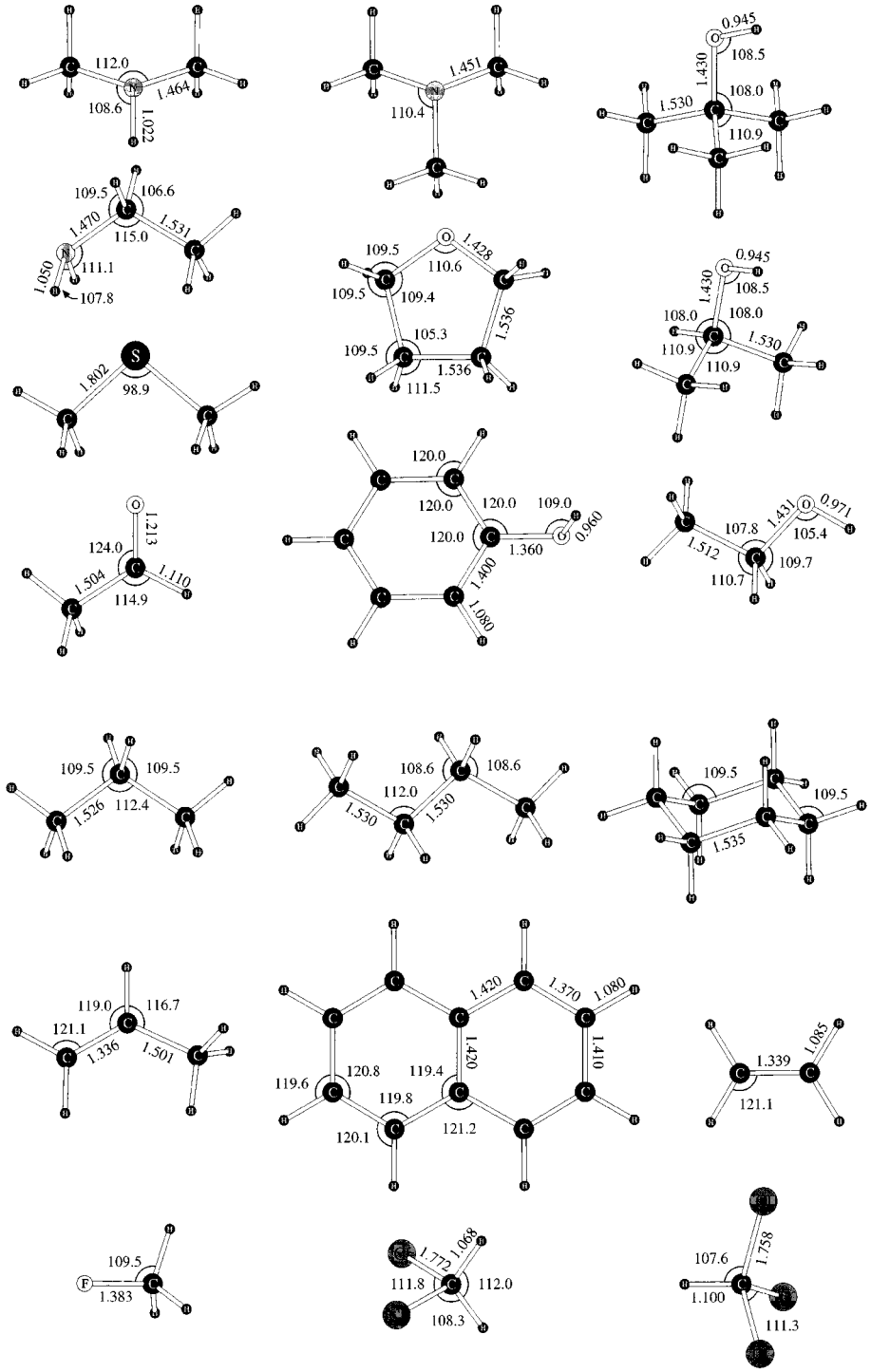

Figure 1. Standard geometries for new solutes used in the calculations. All distances are in Å and all angles are in degrees. Structures for the remaining solutes are in Ref. 5

Table 1. Computed HF/6–31G* EPS atomic charges and OPLS Lennard–Jones parameters[a]

Molecule	Atom	q (EPS)	σ (Å)	ε (kcal mol^{-1})
$CH_3CH_2NH_2$	CH_3	−0·259	3·55	0·066
	HC2	0·039	2·50	0·030
	CH_2	0·615	3·55	0·066
	HC1	−0·066	2·50	0·030
	N	−1·091	3·25	0·170
	H	0·375	0·0	0·0
$(CH_3)_2NH$	C	0·071	3·50	0·066
	HC	0·036	2·50	0·030
	N	−0·724	3·25	0·170
	H	0·366	0·0	0·0
$(CH_3)_3N$	C	−0·137	3·50	0·066
	HC	0·0846	2·50	0·030
	N	−0·3504	3·25	0·170
$(CH_3)_2S$	C	−0·078	3·50	0·066
	H	0·0715	2·50	0·030
	S	−0·273	3·55	0·250
CH_3CHO	CH_3	−0·233	3·50	0·066
	HC2	0·070	2·50	0·030
	C	0·606	3·55	0·076
	HC1	−0·042	2·42	0·030
	O	−0·541	2·96	0·210
THF	O	−0·598	3·00	0·170
	C1	0·435	3·50	0·066
	HC1	−0·055	2·50	0·030
	C2	0·004	3·50	0·066
	HC2	−0·015	2·50	0·030
CH_3CH_2OH	CH_3	−0·224	3·50	0·066
	HC2	0·060	2·50	0·030
	CH_2	0·475	3·50	0·066
	HC1	−0·063	2·50	0·030
	O	−0·703	3·12	0·170
	H	0·398	0·0	0·0
$(CH_3)_2CHOH$	CH_3	−0·437	3·55	0·066
	HC2	0·0955	2·50	0·030
	C	0·657	3·55	0·066
	HC1	−0·014	2·50	0·030
	O	−0·772	3·12	0·170
	H	0·430	0·0	0·0
$(CH_3)_3COH$	CH_3	−0·392	3·55	0·066
	HC	0·0762	2·50	0·030
	C	0·879	3·55	0·066
	O	−0·818	3·12	0·170
	H	0·4292	0·0	0·0
C_6H_5OH	O	−0·623	3·07	0·170
	H	0·390	0·0	0·0
	C1	0·500	3·55	0·070
	C2	−0·321	3·55	0·070
	H2	0·153	2·42	0·030
	C3	−0·039	3·55	0·070
	H3	0·112	2·42	0·030
	C4	−0·198	3·55	0·070
	H4	0·121	2·42	0·030
$CH_3CH_2CH_3$	CH_3	−0·260	3·50	0·066
	HC1	0·050	2·50	0·030
	CH_2	0·372	3·50	0·066
	HC2	−0·076	2·50	0·030
$CH_3CH_2CH_2CH_3$	CH_3	−0·257	3·55	0·066

Table 1. Continued.

Molecule	Atom	q (EPS)	σ (Å)	ε (kcal mol^{-1})
	HC1	0·051	2·50	0·030
	CH_2	0·184	3·55	0·066
	HC2	−0·040	2·50	0·030
C_6H_{12}	C	0·001	3·50	0·066
	H(ax)	0·013	2·50	0·030
	H(eq)	−0·014	2·50	0·030
$C_{10}H_8$	C9 (fusion)	0·194	3·55	0·070
	C1	−0·264	3·55	0·070
	H1	0·146	2·42	0·030
	C2	−0·084	3·55	0·070
	H2	0·105	2·42	0·030
C_2H_4	C	−0·292	3·55	0·076
	H	0·146	2·42	0·030
CH_3CHCH_2	CH_3	−0·027	3·50	0·066
	HC3	0·022	2·50	0·030
	CH	0·015	3·55	0·076
	HC2	0·080	2·42	0·030
	CH_2	−0·498	3·55	0·076
	HC1	0·182	2·42	0·030
FCH_3	F	−0·281	2·94	0·061
	C	0·161	3·50	0·066
	H	0·040	2·50	0·030
Cl_2CH_2	Cl	−0·140	3·40	0·300
	C	−0·016	3·50	0·066
	H	0·148	2·50	0·030
Cl_3CH	Cl	−0·051	3·47	0·300
	C	−0·063	3·50	0·066
	H	0·216	2·50	0·030
$C_6H_5NH_2$	C1	0·557	3·55	0·070
	C2	−0·370	3·55	0·070
	H2	0·161	2·42	0·030
	C3	−0·004	3·55	0·070
	H3	0·098	2·42	0·030
	C4	−0·230	3·55	0·070
	H4	0·118	2·42	0·030
	N	−0·963	3·25	0·170
	H(N)	0·374	0·00	0·000
C_6H_5Cl	C1	0·104	3·55	0·070
	C2	−0·087	3·55	0·070
	H2	0·114	2·42	0·030
	C3	−0·122	3·55	0·070
	H3	0·120	2·42	0·030
	C4	−0·100	3·55	0·070
	H4	0·112	2·42	0·030
	Cl	−0·166	3·47	0·300
$C_6H_5NO_2$	C1	0·043	3·55	0·070
	C2	−0·110	3·55	0·070
	H2	0·142	2·42	0·030
	C3	−0·108	3·55	0·070
	H3	0·119	2·42	0·030
	C4	−0·067	3·55	0·070
	H4	0·114	2·42	0·030
	N	0·716	3·25	0·170
	O	−0·446	2·96	0·210

[a] Charges and Lennard–Jones parameters for other molecules in this study have been reported in Ref. 5.

Table 2. Computed and experimental dipole moments

Molecule	BOSS with EPS partial charges	6–31G*	Experiment[25]
Ethylamine	1·41	1·42	1·22
Dimethylamine	1·59	1·25	1·03
Trimethylamine	1·20	0·88	0·61
Dimethyl sulfide	2·02	1·78	1·50
Acetaldehyde	3·17	3·10	2·69
Tetrahydrofuran	2·05	2·08	1·63
Ethanol	2·02	1·82	1·69
Isopropyl alcohol	1·89	1·81	1·66
tert-Butyl alcohol	2·26	1·68	1·66
Phenol	1·82	1·82	1·45
Propane	0·13	0·02	0·08
Butane	0·00	0·00	0·00
Cyclohexane	0·00	0·00	0·00
Naphthalene	0·00	0·00	0·00
Ethylene	0·00	0·00	0·00
Propylene	0·30	0·28	0·37
Methyl fluoride	2·08	2·08	1·85
Methylene chloride	2·19	2·19	1·60
Chloroform	1·53	1·53	1·01

taken from the OPLS all-atom force field.[6a, 26] For the molecules which were simulated in both water and chloroform, the geometries, Lennard–Jones parameters and charges were the same for both simulations. The internal geometries of the molecules were fixed with the exception of the C–C–O–H and the C–C–N–H torsions for phenol and aniline.[26a, b] Sampling of methyl rotations was not included. The potential function parameters for the solvents came from the TIP4P model of water[16] and the OPLS 4-site model of chloroform.[18b]

The simulations in water and chloroform were performed for one solute molecule in a periodic cube of 260 solvent molecules, roughly 20 and 33 Å on a side, respectively. The computational details are the same as in the previous studies including use of Metropolis sampling in the NPT ensemble at 25 °C and 1 atm.[5, 14] The FEP calculations were performed using statistical perturbation theory with double-wide sampling, as described before.[5] A solute A was gradually mutated to a solute B in a series of four or five separate simulations with a coupling parameter, λ, that ranged from 0 (A) to 1 (B). For the annihilation of methane in chloroform, 10 simulations with $\Delta\lambda = 0{\cdot}05$ were used. MC simulations were also run at the end-points, A and B, in order to obtain the energy components for the correlation equation (2). The BOSS program was used for all MC simulations.[27]

Each system underwent 10^6 configurations of equilibration and 4×10^6 configurations of averaging. Solute moves were attempted every 60 configurations and volume changes were tried every 1625 configurations. The ranges for translations and rotations of the solute and solvent molecules were adjusted to give approximately 40% acceptance rates for new configurations. In TIP4P water, the solvent–solvent interactions were truncated at O–O distances of 8·5 Å and solute–solvent interactions were cut off at 10 Å based approximately on the distance between the solute's center-of-mass and the water oxygen. In chloroform, both solvent–solvent and solute–solvent interactions were truncated at 12 Å. In all cases the potential functions were quadratically feathered to zero over the last 0·5 Å.

The solvent-accessible surface areas for the solutes were determined with the SAVOL2 program.[28] All radii were taken from Rashin *et al.*,[29] with the exception of chlorine, which was not available and was estimated at 2·0 Å. Probe radii of 1.4 Å for water and 3·2 Å for chloroform were used and are consistent with estimates based on van der Waals radii. The solvent-accesible surface areas along with the energy components were then fitted to equation (2) with a simplex algorithm-based program.

RESULTS

For perturbations performed in water, the computed results can be directly compared with experimental free energies of hydration for transfer from the gas phase to aqueous solution with 1M standard states. Typically, though, the experimental data for organic solvents such as chloroform is available in the form of partition coefficients, which are directly related to the free energy of transfer, ΔG_t, between two solvents, in this case water and chloroform by the equation

$$\Delta G_t(\mathrm{A}) = -2{\cdot}3RT \log P_\mathrm{A} \qquad (4)$$

The absolute free energy of solvation in chloroform, $\Delta G_{solv}(\mathrm{A})$, is then just the sum of the free energies of hydration, ΔG_{hyd}, and of transfer [equation (5)].

$$\Delta G_{solv}(\mathrm{A}) = \Delta G_{hyd}(\mathrm{A}) + \Delta G_t(\mathrm{A}) \qquad (5)$$

The validity of using 6–31G* EPS changes in combination with the OPLS all-atom Lennard–Jones parameters for obtaining free energies of hydration via FEP calculations in TIP4P water was addressed previously; the average error for the original set of 13 organic molecules was 1·1 kcal mol^{-1}.[5] The utility of this approach for chloroform was tested here. In all, 19 perturbations were performed to connect the 16 molelcules in the dataset to ethane and methane, which was annihilated. The resultant FEP-computed free energies of solvation are compared with the experimental values[21, 30] from equation (5) in Table 3. The reported uncertainties ($\pm1\sigma$) for the computed values were obtained from the fluctuations in separate averages over batches of 2×10^5 configurations. It is not uncommon to have variations of *ca* 0·2 kcal mol^{-1} in experimental free energies of hydration and of *ca* 0·2 log units in partition coefficients,[20, 30] so the overall uncertainties in the experimental free energies of solvation in chloroform are 0·3–0·5 kcal mol^{-1}.

The average unsigned error for the calculated values of ΔG_{solv} is 0·83 kcal mol^{-1}; the largest individual error is for methylamine, which is not well enough solvated by

Table 3. Free energies of solvation (kcal mol^{-1}) in chloroform at 25 °C from experiment and from Monte Carlo/FEP calculations with EPS charges

Molecule	ΔG_{exp} [a]	ΔG_{FEP}	Error
Methanol	−3·32	−1·62±0·1	1·70
Phenol	−7·10	−6·62±0·2	0·48
Methylamine	−3·44	−1·38±0·1	2·06
Dimethylamine	−3·70	−2·65±0·2	1·05
Trimethylamine	−3·98	−3·77±0·2	0·21
Aniline	−6·70	−6·54±0·2	0·16
Pyridine	−6·58	−6·07±0·2	0·51
Acetaldehyde	−3·65	−2·92±0·1	0·73
Acetone	−4·96	−4·58±0·1	0·38
Acetic acid	−4·63	−4·14±0·1	0·49
Acetamide	−6·98	−5·59±0·1	1·39
Methyl acetate	−4·88	−5·17±0·2	−0·29
Acetonitrile	−4·49	−3·77±0·1	0·72
Benzene	−4·61	−5·07±0·2	−0·46
Chlorobenzene	−5·81	−6·51±0·2	−0·70
Cyclohexane	−4·48	−6·47±0·2	−1·99

[a] From Refs 21 and 30 and equation (5). ΔG values are in kcal mol^{-1} in all tables.

2·1 kcal mol^{-1}, while cyclohexane is too well solvated by 2·0 kcal mol^{-1}. The results are also displayed in Figure 2.

The computed solute–solvent energy components and the SASA for each solute in chloroform are presented in Table 4. The results from the fits of these data with equation (2) to the experimental and FEP free energies of solvation in chloroform are presented in Tables 5 and 6, respectively, and the correlations are shown graphically in Figures 3 and 4. In order to obtain a measure of how accurately the parametrized LR method could predict ΔG_{solv}, a cross-validation technique was employed. Specifically, four solutes were removed from the dataset, the fits to the FEP and the experimental free energies were performed again and the resulting parameters were used to predict the ΔG_{solv} for the solutes which had been omitted. The results for the fits of the 12 remaining molecules to equation (2) are given in Tables 5 and 6 in the columns labeled 'CV;' the predicted values are in the columns labeled '‡.'

The solute–solvent energy components and the SASA from the MC simulations for the 19 new solutes in TIP4P water are given in Table 7. These results were then combined with those from the previous study[14] to yield data for a total of 35 solutes; the energy components and SASAs were fitted to the experimental free energies of hydration using equation (2). The outcome from the fit to the entire dataset is summarized in Table 8 and Figure 5. The parameters, α, β and γ and the rms deviations are given in Table 9, including those from several fits performed on subsets of the complete 35-solute dataset. Finally, the linear response predictions were combined to yield predicted

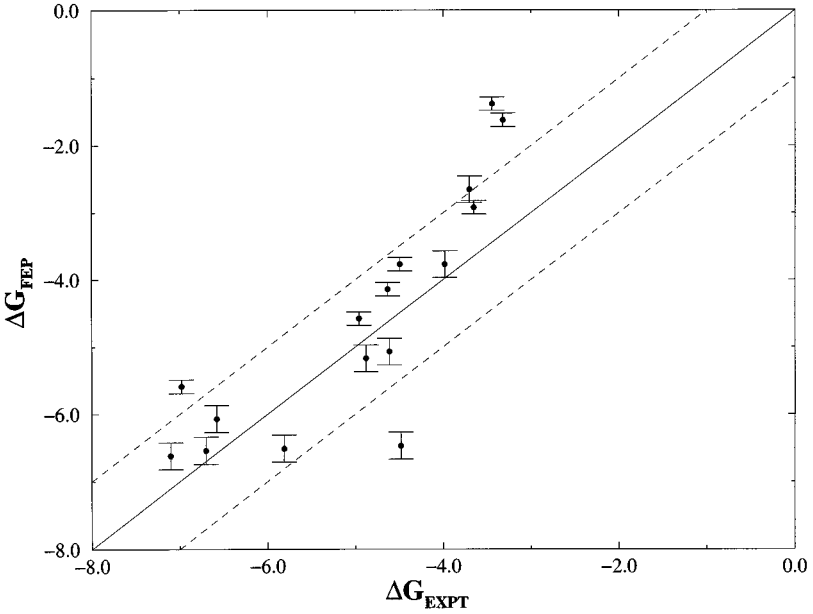

Figure 2. Correlation between the FEP results with EPS charges in chloroform and the experimental free energies of solvation. The solid line has unit slope and the dashed lines are 1 kcal mol^{-1} on either side

Table 4. Average Coulombic and Lennard–Jones solute–solvent energies and solvent-accessible surface areas from the Monte Carlo simulations with EPS charges in OPLS chloroform

Molecule	$\langle U_{elec} \rangle$	$\langle U_{vdw} \rangle$	SASA (Å^2)
Methanol	−2·24±0·14	−6·16±0·06	358·6
Phenol	−1·63±0·09	−15·53±0·09	484·3
Methylamine	−1·69±0·15	−6·29±0·08	387·8
Dimethylamine	−1·40±0·12	−9·56±0·08	429·3
Trimethylamine	−0·78±0·07	−12·77±0·08	467·5
Aniline	−1·87±0·13	−15·75±0·08	506·7
Pyridine	−1·59±0·10	−14·09±0·09	465·8
Acetaldehyde	−3·04±0·13	−8·63±0·07	287·4
Acetone	−3·17±0·31	−11·90±0·15	356·7
Acetic acid	−2·71±0·21	−10·71±0·11	411·9
Acetamide	−5·49±0·19	−11·04±0·06	338·3
Methyl acetate	−2·13±0·12	−13·88±0·08	462·4
Acetonitrile	−3·69±0·22	−9·20±0·08	553·6
Benzene	−0·60±0·14	−13·87±0·10	464·0
Chlorobenzene	−0·87±0·12	−16·87±0·14	517·1
Cyclohexane	−0·04±0·01	−16·43±0·08	502·4

chloroform–water partition coefficients, log *P*, for the 16 solutes. The results are compared with the experimental data in Table 10.

DISCUSSION

Since the range of observed free energies of solvation in chloroform for the 16 solutes is only 3·8 kcal mol^{-1}, the average error from the FEP results of 0·8 kcal mol^{-1} yields considerable scatter in Figure 2. There is a pattern in the errors. Most of the predicted ΔG_{solv} values are too positive and these are for polar solutes that can form hydrogen bonds. The five predicted values that are too favorable are for the less polar solutes including cyclohexane, benzene and chlorobenzene. This could result from the EPS charge distributions not being polar enough, although this seems unlikely in view of the dipole moment results in Table 2 and the prior results for free energies of hydration with the same source for charges.[5] The problem more likely stems from the united-atom CH group in the four-site chloroform model, which may not model C–H···X hydrogen bonds well. Another possibility is 'wet' chloroform in the

Table 5. Results from fitting equation (2) to the experimental free energies of solvation in chloroform[a]

		ΔG_{LR}			Error		
Molecule	ΔG_{expt}	All	CV	‡	All	CV	‡
Methanol	−3·32	−3·13	−3·13		0·19	0·19	
Phenol	−7·10	−5·88	−5·83		1·22	1·27	
Methylamine	−3·44	−2·85	−2·81		0·59	0·63	
Dimethylamine‡	−3·70	−3·77	—	−3·72	−0·07		−0·02
Trimethylamine	−3·98	−4·49	−4·39		−0·51	−0·41	
Aniline	−6·70	−6·09	−6·04		0·61	0·66	
Pyridine‡	−6·58	−5·39	—	−5·33	1·19		1·25
Acetaldehyde	−3·65	−4·44	−4·52		−0·79	−0·87	
Acetone‡	−4·96	−5·58	—	−5·62	−0·62		−0·66
Acetic acid‡	−4·63	−4·90	—	−4·92	−0·27		−0·29
Acetamide	−6·98	−6·58	−6·78		0·40	0·20	
Methyl acetate	−4·88	−5·62	−5·60		−0·74	−0·72	
Acetonitrile	−4·49	−4·85	−4·86		−0·36	−0·37	
Benzene	−4·61	−4·76	−4·66		−0·15	−0·05	
Chlorobenzene	−5·81	−5·90	−5·80		−0·09	0·01	
Cyclohexane	−4·48	−5·30	−5·15		−0·82	−0·67	

Set	α	β	γ	Rms error	Average error
All	0·338	0·553	0·0005	0·64	0·54
CV	0·340	0·607	0·0009	0·62	0·50

[a] Experimental data from Refs 21 and 30 and equation (5). ‡ Denotes the solutes which were removed for the fit to the data for the remaining 12 solutes.

Table 6. Results from fitting equation (2) to the FEP-calculated free energies of solvation in chloroform[a]

		ΔG_{LR}			Error		
Molecule	ΔG_{FEP}	All	CV	‡	All	CV	‡
Methanol	−1·62±0·1	−1·92	−1·98		−0·30	−0·36	
Phenol	−6·62±0·2	−6·12	−6·13		0·50	0·49	
Methylamine	−1·38±0·1	−1·73	−1·78		−0·35	−0·40	
Dimethylamine‡	−2·65±0·2	−3·20	—	−3·13	−0·55		−0·48
Trimethylamine	−3·77±0·2	−4·58	−4·57		−0·81	−0·80	
Aniline	−6·54±0·2	−6·20	−6·21		0·34	0·33	
Pyridine‡	−6·07±0·2	−5·45	—	−5·33	0·62		0·74
Acetaldehyde	−2·92±0·1	−3·74	−3·81		−0·82	−0·89	
Acetone‡	−4·58±0·1	−5·15	—	−5·05	−0·57		−0·47
Acetic acid‡	−4·14±0·1	−4·16	—	−4·08	−0·02		0·06
Acetamide	−5·59±0·1	−5·26	−5·39		0·33	0·20	
Methyl acetate	−5·17±0·2	−5·46	−5·50		−0·29	−0·33	
Acetonitrile	−3·77±0·1	−2·84	−2·95		0·93	0·82	
Benzene	−5·07±0·2	−5·14	−5·12		−0·07	−0·05	
Chlorobenzene	−6·51±0·2	−6·52	−6·50		−0·01	0·01	
Cyclohexane	−6·47±0·2	−6·19	−6·15		0·28	0·32	

Set	α	β	γ	Rms error	Average error
All	0·529	0·202	0·005	0·50	0·42
CV	0·525	0·231	0·005	0·50	0·42

[a] Experimental data from Refs 21 and 30 and equation (5). ‡ Denotes the solutes which were removed for the fit to the data for the remaining 12 solutes.

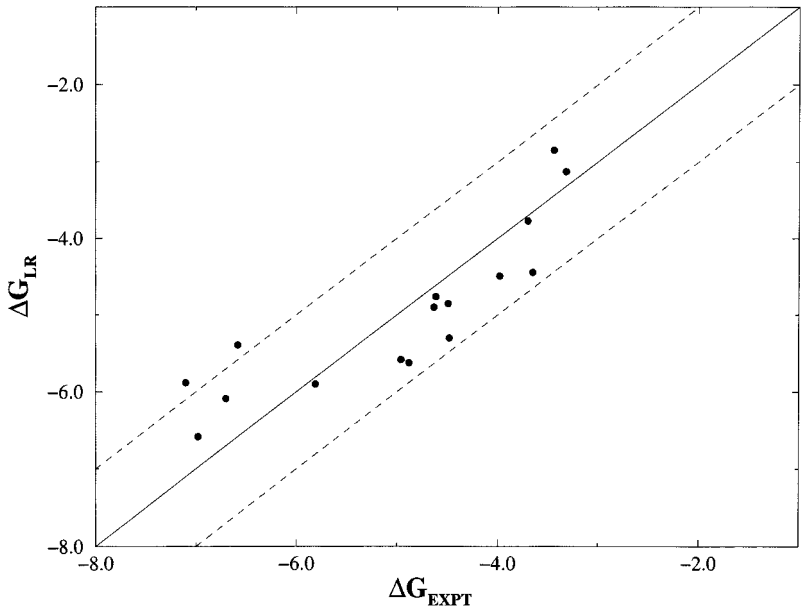

Figure 3. Correlation between the predictions from the linear response equation (2) and the experimental free energies of solvation in chloroform

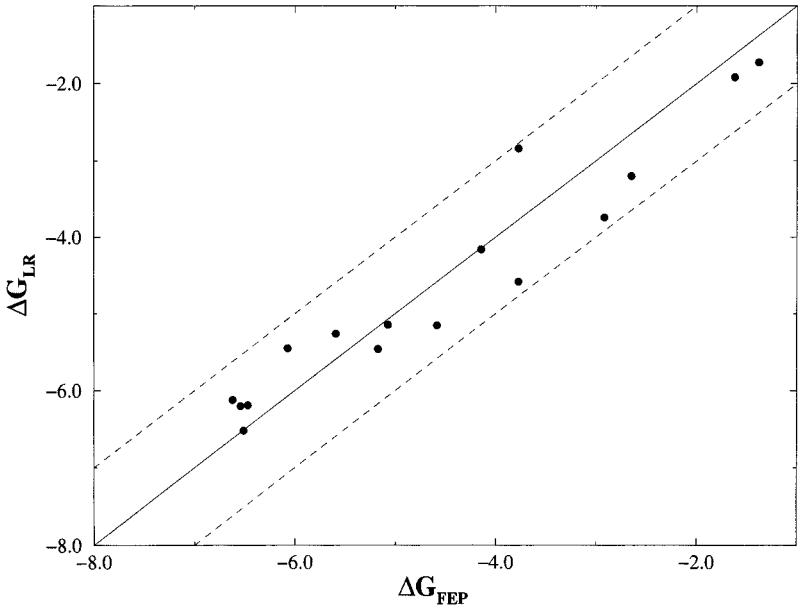

Figure 4. Correlation between the predictions from the linear response equation (2) and the FEP-calculated free energies of solvation in chloroform

Table 7. Average Coulombic and Lennard–Jones solute–solvent energies and solvent-accessible surface areas from the Monte Carlo simulations with EPS charges in TIP4P water[a]

Molecule	$\langle U_{elec} \rangle$	$\langle U_{vdw} \rangle$	SASA (Å^2)
Ethylamne	−22·20±0·50	−2·70±0·12	206·4
Dimethylamine	−12·38±0·26	−4·73±0·09	204·5
Trimethylamine	−7·06±0·21	−7·95±0·08	227·7
Dimethyl sulfide	−4·73±0·24	−6·26±0·08	212·9
Acetaldehyde	−15·27±0·30	−4·52±0·11	176·5
Tetrahydrofuran	−11·74±0·39	−7·57±0·11	212·6
Ethanol	−20·56±0·41	−2·87±0·13	188·4
Isopropyl alcohol	−20·53±0·52	−4·98±0·17	216·8
tert-Butyl alcohol	−23·79±0·49	−5·74±0·18	241·3
Phenol	−18·34±0·46	−8·80±0·15	242·0
Propane	−0·61±0·04	−6·88±0·04	207·1
Butane	−0·40±0·04	−8·76±0·07	236·5
Cyclohexane	−0·08±0·02	−11·55±0·07	254·8
Naphthalene	−6·70±0·35	−13·32±0·12	292·0
Ethylene	−1·57±0·11	−4·42±0·05	160·1
Propylene	−2·59±0·25	−6·31±0·10	195·6
Methyl fluoride	−6·29±0·20	−2·95±0·07	153·9
Methylene chloride	−4·37±0·13	−7·03±0·07	217·4
Chloroform	−2·14±0·12	−8·77±0·09	247·1

[a] Results for the 19 new solutes; data for the 16 original solutes in Ref. 14.

experimental shake-flask measurements. Hydrogen bonding solutes could associate with some of the water in chloroform, which would make log *P* (chloroform–water) greater and the measured free energies of solvation overly negative. However, comparisons of log *P* from shake-flask and direct Henry's law measurements in the few available cases suggest that errors from this source are covered by the 0·3–0·5 kcal mol^{-1} uncertainty noted above.[31]

An independent check was also made on the precision of the FEP calculations in chloroform. One cycle of mutations was evaluated: methane → ethane → methanol → methylamine → acetonitrile → acetone → methane. The ΔG for this closed path should, of course, be exactly zero; the hysteresis obtained was 0·25 kcal mol^{-1}, which is consistent with the uncertainties of 0·1–0·2 kcal mol^{-1} for the computed ΔG_{solv} values in Table 3.

The parameters, rms deviations and average error from the fit of equation (2) for the full set of 16 molecules to the experimental ΔG_{solv} values are given at the bottom of Table 5. Results are also listed for the fit with the smaller 12-molecule dataset. The rms and average errors for the entire set are 0·64 and 0·54 kcal mol^{-1}. This shows improvement over the comparison of the FEP results and experiment, as reflected in Figures 2 and 3. The remaining error is not much greater than the estimated uncertainty in the experimental data. The largest individual errors in the predictions are 1·2 kcal mol^{-1} for phenol and pyridine. The rms and average errors for the fit to the LR equation change little with the removal of four solutes from the dataset. The calculated ΔG_{solv} values for the 12 remaining solutes are

also affected only slightly; the largest change is 0·2 kcal-mol^{-1}, but most differences are only 0·1 kcal mol^{-1}. The ΔG_{solv} values predicted for the four removed solutes with the parameters from the 12-solute fit are remarkably close to the original calculated values from the full fit: none changed by more than 0·06 kcal mol^{-1}. The results were very similar when a different set of four solutes were omitted. Thus, the α, β and γ parameters for chloroform appear to be well converged and should be appropriate for use with new solutes.

Not surprisingly, and as in the previous LR study,[14] the fit of equation (2) to the FEP results is even better; the pertinent parameters, rms deviations and errors are given at the bottom of Table 6. The closer accord is also apparent in comparing Figures 3 and 4. Again, arbitrary removal of four molecules from the dataset has little effect on the parameters, rms and average errors and on the predicted free energies of solvation for the omitted molecules compared with the calculated values from the full fit. It is interesting that although the LR method works well at predicting either the experimental or the FEP data, the parameters from the fits to the two different datasets are different (Tables 5 and 6). The coefficient α is critical in chloroform since the intermolecular energies are dominated by the van der Waals (Lennard–Jones) component, as shown in Table 4. Thus, the increase in α from 0·338 for the fit to the experimental data to 0·529 for the fit to the FEP results is offset largely by the increase in γ from 0·0005 to 0·0050. The latter value of γ seems large since it is near the low-end of the range of values typical for water in which cavity formation is expected to be more difficult owing to the higher surface tension than for chloroform (72 vs 27 dyn cm^{-1}).[9,14]

Table 8. Experimental and computed [equation (2)] free energies of hydration at 25 °C

Molecule	ΔG_{expt}	ΔG_{LR}	Error
Trimethylamine	−3·2	−1·43	1·77
Dimethylamine	−4·3	−2·61	1·69
Phenol	−6·6	−5·01	1·59
Methylamine	−4·6	−3·30	1·30
Chloroform	−1·1	0·14	1·24
Methanol	−5·1	−4·01	1·09
Aniline	−4·9	−3·96	0·94
Dimethyl sulfide	−1·4	−0·48	0·92
Methylene chloride	−1·4	−0·49	0·91
Pyridine	−4·7	−3·91	0·79
Naphthalene	−2·4	−1·85	0·55
Tetrahydrofuran	−3·5	−2·99	0·51
Methanethiol	−1·2	−0·72	0·48
Methyl chloride	−0·6	−0·32	0·28
Chlorobenzene	−1·1	−0·93	0·17
Ethanol	−5·0	−4·93	0·07
Acetonitrile	−3·9	−3·83	0·07
Dimethyl ether	−1·9	−1·85	0·05
Benzene	−0·8	−1·07	−0·27
Isopropyl alcohol	−4·8	−5·08	−0·28
Acetaldehyde	−3·5	−3·79	−0·29
Nitrobenzene	−4·1	−4·39	−0·29
Acetamide	−9·7	−10·00	−0·30
Acetone	−3·8	−4·45	−0·65
Methyl fluoride	−0·2	−0·87	−0·67
Ethylamine	−4·5	−5·19	−0·69
Acetic acid	−6·7	−7·40	−0·70
Ethylene	1·3	0·34	−0·96
Cyclohexane	1·2	0·22	−0·98
Ethane	2·0	0·83	−1·17
Propylene	1·3	−0·02	−1·32
Propane	2·0	0·60	−1·40
tert-Butyl alcohol	−4·5	−6·00	−1·50
Butane	2·1	0·57	−1·53
Methyl acetate	−3·3	−5·17	−1·87

α=0·238; β=0·314; γ=0·012; Rms error=0·99; average error=0·84

[a] Ref. 30.

Turning to the results in water, the rms error for the fit of equation (2) to the experimental free energies of hydration has increased from 0·76 kcal/mol^{-1} for the original 16 molecules to 0·99 kcal mol^{-1} for the full dataset of 35 molecules in Table 8. The average unsigned error for the 35 molecules is 0·84 kcal mol^{-1}. In view of the simplicity of equation (2), the origin of the partial charges, the variety of solutes, and the 12 kcal mol^{-1} range for the experimental values, the quality of the fit is remarkably good (see Figure 5). The optimized parameters have also changed from 0·111, 0·287 and 0·006 to 0·238, 0·314 and 0·012 forα, β and γ respectively. The relatively small change for β is consistent with the general dominance of the electrostatic (Coulombic) term for the intermolecular energetics in water (Table 7).

Examination of the results in Table 9 reveals that fitting various subsets of the whole produces improved correlations. The molecules are ordered in Table 8 according to the error, which equals $\Delta G_{eq.(2)} - \Delta G_{expt}$. The worst errors are for amines, which are predicted to be too hydrophobic, and for hydrocarbons, whch are computed to be too hydrophilic. Removal of the amines improves the rms error to 0·90 kcal-mol^{-1} with little alteration of the parameters. Amines have also proven problematic in other computations of free energies of hydration by FEP or SCRF methods.[32] Specific polarization effects have been proposed to be important in this case,[32] although it may be that the atom-centered charge model is inadequate here. The present results for ethylamine are surprising since they show the opposite error than for the other amines. The MC simulations were repeated in this case, but the results changed little. Another problem molecule is methyl acetate. Although its dipole moment with the EPS charges, 1·83 D, is in reasonable accord with the experimental value, 1·72 D, both the FEP results[14] and the linear response prediction (Table 8) find it too hydrophilic by 1·9–2·0 kcal mol^{-1}. The results are sensitive to the details of the EPS charge distributions, which may yield some anomalies.

Greater improvement to an rms error of 0·82 kcal mol^{-1} is obtained by removing the hydrocarbons. In this case, all three parameters are reduced in magnitude; in particular, γ is smaller by a factor of three (Table 9). This reflects the

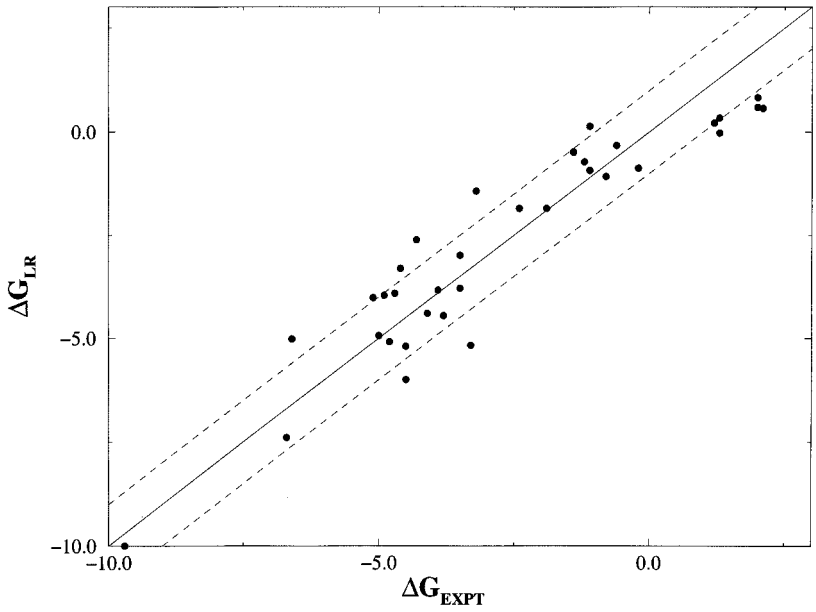

Figure 5. Correlation between the predictions from the linear response equation (2) and experimental free energies of hydration for the full dataset of 35 molecules

lessened importance of the cavitation term for the remaining molecules, which have stronger electrostatic interactions with the aqueous medum (Table 7 here and Table 1 in Ref. 14). Separate fits for the solutes that do or do not form hydrogen bonds with water also provided lower rms errors near 0·8 kcal mol^{-1}. For the hydrogen-bonding molecules, the fit is dominated by the electrostatic term with negligible contributions from the van der Waals and cavitation terms. A more balanced fit is obtained for the molecules which do not form hydrogen bonds.

Fits were also made to simplified versions of equation (2). As listed in Table 9, a two-term expression with just β and γ optimized increases the rms error from 0·99 to 1·07 kcal mol^{-1} for the 35 solutes in water. The two-term alternative with the SASA term left out yields an rms error of 1·10 kcal mol^{-1}. Furthermore, α becomes negative in order to provide a means for obtaining positive values for ΔG_{hyd} since both solute–solvent energy components are always negative (Table 7). These patterns were found in the prior study and led to the inclusion of the SASA term.[14] If the $\gamma\langle$SASA$\rangle$ term is replaced by a constant δ, the rms is the same as for equation (2) with α=0·122, β=0·303 and δ=1·642. The success in this case is probably attributable to the relatively similar size for the 35 solutes, which is reflected in the limited range of SASA values in Table 7. Upon inclusion of much larger solutes, it is expected that the benefit of the SASA term should emerge. In chloroform, the cavitation term is not expected to be as significant, as witnessed by the small γ values in Table 5. In this case, a fit to the experimental free energies of solvation with the SASA term omitted only increases the rms error in Table 5 from 0·639 to 0·641 kcal mol^{-1} with α=0·321 and β=0·540.

Given the correlation equations for predicting free

Table 9. Results of fits to equation (2) for subsets of the 35 solutes in water

Data set (No. of molecules)	Rms error	α	β	γ
Whole set (35)	0·99	0·238	0·314	0·012
Whole set (35)	1·07	(0·000)[a]	0·276	0·002
Whole set (35)	1·10	−0·032	0·260	(0·000)[a]
Whole − amines (30)	0·90	0·296	0·322	0·015
Whole − hydrocarbons (27)	0·82	0·142	0·253	0·004
No H-bonds to water (13)	0·78	0·402	0·542	0·022
With H-bonds to water (22)	0·85	0·070	0·233	0·000

[a] Fixed.

Table 10. Predicted and experimental chloroform–water partition coefficients

Molecule	$\Delta G_{LR}^{H_2O}$	$\Delta G_{LR}^{CHCl_3}$	$\Delta\Delta G_{calc.}$	$\Delta\Delta G_{expt}$	Error	Log $P_{calc.}$	Log P_{expt}	Error
Methanol	−4·01	−3·13	−0·88	−1·78	0·90	−0·65	−1·31	0·66
Phenol	−5·01	−5·88	0·87	0·50	0·37	0·64	0·37	0·27
Methylamine	−3·30	−2·85	−0·45	−1·16	0·71	−0·33	−0·85	0·52
Dimethylamine	−2·61	−3·77	1·16	−0·60	1·76	0·85	−0·44	1·29
Trimethylamine	−1·43	−4·49	3·06	0·78	2·28	2·25	0·57	1·68
Aniline	−3·96	−6·09	2·13	1·80	0·33	1·56	1·32	0·24
Pyridine	−3·91	−5·39	1·48	1·88	−0·40	1·09	1·38	−0·29
Acetaldehyde	−3·79	−4·44	0·65	0·15	0·40	0·48	0·11	0·37
Acetone	−4·45	−5·58	1·13	0·98	0·15	0·83	0·72	0·11
Acetic acid	−7·40	−4·90	−2·50	−2·07	−0·43	−1·84	−1·52	−0·32
Acetamide	−10·00	−6·58	−3·42	−2·72	−0·70	−2·51	−2·00	−0·51
Methyl acetate	−5·17	−5·62	0·45	1·58	−1·13	0·33	1·16	−0·83
Acetonitrile	−3·83	−4·85	1·02	0·59	0·43	0·75	0·43	0·32
Benzene	−1·07	−4·76	3·69	3·81	−0·12	2·71	2·80	−0·09
Chlorobenzene	−0·93	−5·90	4·97	4·71	0·26	3·65	3·46	0·19
Cyclohexane	0·22	−5·30	5·52	5·68	−0·16	4·05	4·17	−0·12

energies of solvation in chloroform and water, log *P* values can then be predicted from equations (4) and (5). The results in Table 10 for the 16 compounds that were modeled both in chloroform and water are gratifying. The average error for the log *P* predictions is 0·49 log unit. The largest error by far is for trimethylamine and the second largest error is for dimethylamine. It is likely that the experimental log *P* values are erroneously low in these cases owing to protonation of the amines in the water layer, as noted previously for diethylamine.[21] With these two datapoints left out, the average error for the predicted log *P* values in Table 10 is 0·35 log unit. This shows an improvement over the average error of 0.61 log unit with diethylamine excluded in the application of Still's GB/SA models by Reynolds.[21] In view of the uncertainties in the experimental log *P* values and the 6·2-unit range of values in Table 10, the present approach has notable predictive value. There is far more experimental data available for octanol–water partition coefficients and numerous structure-based approaches have been devised for their prediction.[17,33] The best, highly parametrized regression equations yield average errors of *ca* 0·5 log unit for datasets with *ca* 1000 compounds.

CONCLUSION

A correlation approach using equation (2) has been explored here for computing free energies of solvation in water and chloroform. The approach retains the advantages of explicit-solvent simulations including the ability to readily change temperature, pressure, and the solvent model, and the incorporation of specific short-range interactions such as hydrogen bonding. It is also much more efficient computationally than free-energy perturbation calculations and can be much more easily applied to solutes with widely varying structures. Although some empiricism has been introduced, the predictions from the correlation equation also show improved accuracy over FEP results. The approach requires partial charges and Lennard–Jones parameters for the molecules in order to describe the intermolecular interactions. The Lennard–Jones parameters are quite standardized, while *ab initio* 6–31G* EPS charges have been utilized here. Alternative charge models warrant consideration. For the dataset of 16 molecules in chloroform, the correlation approach yielded an average error of 0·54 kcal mol^{-1} for predicted free energies of solvation. For the dataset of 35 molecules in water the average error is 0·84 kcal mol^{-1} for free energies of hydration. Combination of the correlations yields predictions for chloroform–water partition coefficients, log *P*, with average errors of *ca* 0·4 log unit. The quantitative utility of the present methodology is competitive with the best available alternative approaches, which are generally more empirical, contain more parameters and are less readily adapted to new solvents and conditions.

ACKNOWLEDGMENTS

Gratitude is expressed to the National Institutes of Health and National Science Foundation for support.

REFERENCES

1. W. L. Jorgensen and C. Ravimohan, *J. Chem. Phys.* **83**, 3050 (1985).
2. For a review, see P. A. Kollman, *Chem. Rev.* **93**, 2395 (1993).
3. T. P. Lybrand, J. A. McCammon and G. Wipff, *Proc. Natl. Acad. Sci. USA* **83**, 833 (1986).
4. For a review, see W. L. Jorgensen and J. Tirado-Rives, *Perspect. Drug Discovery Des.* **3**, 123 (1995).
5. H. A. Carlson, T. B. Nguyen M. Orozco and W. L. Jorgensen, *J. Comput. Chem.* **14**, 1240 (1993).
6. (a) G. Kaminski, E. M. Duffy, T. Matsui and W. L. Jorgensen, *J. Phys. Chem.* **98**, 13077 (1994); (b) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am.*

Chem. Soc. **117**, 5179 (1995).
7. Y. K. Kang, G. Némethy and H. A. Scheraga, *J. Phys. Chem.* **91**, 4109 (1987).
8. A. A. Rashin and K. Namboodiri, *J. Phys. Chem.* **91**, 6003 (1987).
9. W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.* **112**, 6127 (1990).
10. A. Warshel and Z.-T. Chu, *ACS Symp. Ser.* **568**, 71 (1994).
11. D. Eisenberg and A. D. McLachlan, *Nature (London)* **319**, 199 (1986).
12. (a) J. Åqvist, C. Medina and J.-E. Sammuelsson, *Protein Eng.* **7**, 385 (1994); (b) J. Åqvist and S. M. Mowbray, *J. Biol. Chem.* **270**, 9978 (1995); (c) T. Hansson and J. Åqvist, *Protein Eng.* **8**, 1137 (1995).
13. M. D. Paulsen and R. L. Ornstein, *Protein Eng.* **9**, 567 (1996).
14. H. A. Carlson and W. L. Jorgensen, *J. Phys. Chem.* **99**, 10667 (1995).
15. M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Robb, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart and J. A. Pople, *Gaussian 92, Revision A.* Gaussian, Pittsburgh, PA (1992).
16. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
17. C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology.* Wiley, New York (1979).
18. (a) W. L. Jorgensen, *Acc. Chem. Res.* **22**, 184 (1989); (b) W. L. Jorgensen, J. M. Briggs and M. L. Contreras, *J. Phys. Chem.* **94**, 1683 (1990).
19. J. W. Essex, C. A. Reynolds and W. G. Richards, *J. Chem. Soc., Chem. Commun.* 1152 (1989).
20. W. J. Dunn, III, P. I. Nagy and E. R. Collantes, *J. Am. Chem. Soc.***113**, 7898 (1991).
21. C. H. Reynolds, *J. Chem. Inf. Comput. Sci.* **35**, 738 (1995).
22. F. J. Luque, Y. Zhang, C. Alemán, M. Bachs, J. Gao and M. Orozco, *J. Phys. Chem.* **100**, 4269 (1996); M. Orozco, C. Colomines and F. J. Luque, *Chem. Phys.* **209**, 19 (1996).
23. D. J. Giesen, C. C. Chambers, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem.* **101**, 2061 (1997).
24. M. D. Harmony, V. W. Laurie, R. L. Kuczkowski, R. H. Schwendeman and D. A. Ramsay, *J. Phys. Chem. Ref. Data* **8**, 619 (1979).
25. R. D. Nelson, D. R. Lide and A. A. Maryott, *Natl. Stand. Ref. Data Ser., Natl. Bur. Stand.* **10**, 1 (1967).
26. (a) W. L. Jorgensen and T. B. Nguyen, *J. Comput. Chem.* **14**, 195 (1993); (b) W. L. Jorgensen, E. R. Laird, T. B. Nguyen and J. Tirado-Rives, *J. Comput. Chem.* **14**, 206 (1993); (c) W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.* **118**, 11225 (1996).
27. W. L. Jorgensen, *BOSS, v. 3.6.* Yale University, New Haven, CT (1995).
28. J. M. Skell and R. S. Pearlman, *SAVOL2.* University of Texas, Austin, TX (1988).
29. (a) A. A. Rashin, *Biopolymers* **23**, 1605 (1984); (b) A. A. Rashin, M. Iofin and B. Honig, *Biochemistry* **25**, 3619 (1986).
30. C. J. Cramer and D. G. Truhlar, *J. Comput.-Aided Mol. Des.* **6**, 629 (1993); A. Leo, *Pomona College Medicinal Chemistry Data Base.* Pomona College, Claremont, CA (1988).
31. W. J. Dunn, III, unpublished data.
32. Y. Ding, D. N. Bernardo, K. Krogh-Jespersen and R. M. Levy, *J. Phys. Chem.* **99**, 11575 (1995); E. C. Meng, J. W. Caldwell and P. A. Kollman, *J. Phys. Chem.* **100**, 2367 (1996); B. Marten, K. Kim, C. Cortis, R. A. Friesner, R. B. Murphy, M. N. Rignalda, D. Sitkoff and B. Honig, *J. Phys. Chem.* **100**, 11775 (1996).
33. G. Klopman, J.-Y. Li, S. Wang and M. Dimayuga, *J. Chem. Inf. Comput. Sci.* **34**, 752 (1994).

