Solvation Model Based on Weighted Solvent Accessible Surface Area

Junmei Wang, Wei Wang, Shuanghong Huo, Matthew Lee, and Peter A. Kollman*

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

Received: January 22, 2001; In Final Form: March 5, 2001

We have developed a fast procedure to predict solvation free energies for both organic and biological molecules. This solvation model is based on weighted solvent accessible surface area (WSAS). Least-squares fittings have been applied to optimize the weights of SAS for different atom types in order to reproduce the experimental solvation free energies. Good agreement with experimental results has been obtained. For the 184-molecule set (model I), for which there are experimental solvation free energies in 1-octanol, we have achieved an average error of 0.36 kcal/mol, better than that of the SM5.42R universal solvation model[1] by Li et al. For the 245-molecule set (model II) that has experimental aqueous solvation free energies, our WSAS model achieves an average error of 0.48 kcal/mol, marginally larger than that of Li's model (0.46 kcal/mol). We have used a 401-molecule set, the largest training set (model IV) that we know of solvation model development, to derive the SAS weights in order to reproduce the experimental solvation free energies in water. For this model, we have achieved an average unsigned error of 0.54 kcal/mol and an RMS error of 0.79 kcal/mol.The advantage of this model lies in its simplicity and independence of charge models. We have successfully applied this model to predict the relative binding free energies for the five binding modes of HIV-1RT/efavirenz. The most favorable binding mode, which has an RMSD of 1.1 Å (for 54 C_α around the binding site) compared to the crystal structure, has a binding free energy at least 10 kcal/mol more negative than the other binding modes. Moreover, the solvation free energies with WSAS have a high correlation (the correlation coefficient is 0.92) to the solvation free energies calculated by the Poisson−Boltzmann/surface area (PBSA) model. As an efficient and fast approach, WASA is also attractive for protein modeling and protein folding studies. We have applied this model to predict the solvation free energies of the 36-mer villin headpiece subdomain in its native structure, a compact folding intermediate, and a random coil. The rank order of the solvation free energies and the free energies for the three kinds of conformational clusters are in reasonable agreement with those found by MM-PBSA, a widely used solvation free energy model.

Introduction

Solvation free energies in aqueous solution are important for structural and metabolic equilibria and kinetics.[2−5] For example, in enzyme catalysis, substrates are often dehydrated, and thus, the free energies of desolvation are important contributors to the free energies of activation. The electrostatic energy in a molecular mechanical model is usually coupled with the electrostatic part of the solvation free energy. Unfavorable solvation free energies of associating nonpolar groups may lead to hydrophobic stabilization of their mutual complexes, although the electrostatic energy may not be so favorable. On the other hand, a favorable electrostatic energy does not always mean a high affinity, owing to the high solvation free energy penalty. Thus, the solvation free energy plays an important role in both protein−ligand association and the protein folding.

How to treat the solvation effect efficiently and accurately is a long-term challenge in computational chemistry. At this point, there have been many different solvation models developed to address this problem, from the very simple atom and fragment-additive models[6,7] to methods that treating solvent molecules explicitly.[8−10] In 1975, Hine and Mookerjee developed a simple solvation model[6] based on the assumption that the solvation free energy is fragment additive. Therefore, in their model, the solvation free energy of a given molecule is the sum of the solvation free energy of all the fragments. Hansch and Leo also developed a similar solvation model[7] to predict the water/1-octanol partition coefficients of organic compounds. The additivity assumption of the above models work well for small organic molecules but has some problems when applied to medium and large molecules. In 1986, Eisenberg and McLachlan developed a simple solvation model[11] using the solvent accessible surface area to estimate the solvation free energies of proteins, and their model is still widely used today owning to its simplicity. In 1997, Hawkins, Cramer and Truhlar[12] developed an aqueous solvation model using geometry-dependent atomic surface tensions with implicit electrostatics. Unlike Eisenberg−McLachlan's model, this model applied a more complicated procedure to calculate the atomic surface tensions and there are three types of parameters needed to be parametrized (17 nonlinear parameters, 9 van der Waals radii, and 23 linear surface tension coefficients). In 1998, this model was extended to other solvents.[13] It is believed that the solvation free energy of a molecule is made up of two parts: the electrostatic contribution and the nonpolar contribution[2−5,14−15]. The latter contribution is usually modeled as proportional to the solvent accessible surface area. For most molecules, the electrostatic term dominates the total solvation free energy. All of the above solvation models are charge method independent and the electrostatic contribution of the solvation free energy is implicitly incorporated into the parameters. To more ac-

* To whom correspondence should be addressed. Tel: (415)-476−4673 (pak). Fax: (415)-476-0688. E-mails: Pak@cgl.ucsf.edu. Web site: http://www.amber.ucsf.edu/amber/.

10.1021/jp0102318 CCC: $20.00
Published on Web 05/05/2001

curately predict the solvation free energy, it is necessary to calculate the screening energy (the electrostatic part) using a more physically rigorous approach. The most popular models for screening energy calculations are the Generalized Born model (GB),[16−21] the Poisson−Boltzmann (PB) model,[14,15] and the conductor-like models (COSMO).[22−24] Although these models have a somewhat different theoretical basis, after appropriate parametrization, the electrostatic solvation effect can be well represented by each of these models and be applied in molecular docking, binding free energy calculations and protein folding studies.[25−34] One problem of those methods is that they are charge dependent, and the parameters developed with one charge method may not be transferable to others. Another problem is that these models are quite time-consuming compared to the fragment and surface area based models, especially for the PB and the conductor-like models.

In this work, we have developed a new solvation model based on the weighted solvent accessible surface area (WSAS) using a rather large molecular set (387 neutral molecules and 14 ions). The solvation free energy is evaluated using the following equation.

$$\Delta G_{\mathrm{WSAS}} = \sum_{i=0}^{m} \sum_{j=0}^{n_i} w_{\mathrm{i}} S_{\mathrm{j}} \qquad (1)$$

where m is the number of the atom types for a given model; w_{i} is the solvation free energy weight of atom type i; n_{i} are the number of atoms with atom type i in a molecule, and S_{j} is the solvent accessible surface area of atom j.

Although the basic idea is similar to the Eisenberg−McLachlan model,[11] the two models are quite different. The Eisenberg−McLachlan model has been mainly used to study the original conformational dependent solvation free energies of proteins. The parameters were derived using the water/octanol distribution coefficients of 20 amino acids. On the other hand, we have used the aqueous solvation free energies to develop our parameters not only for amino acid analogues but also for organic molecules. Therefore, our model has more potential applications. For example, it can be used to calculate the relative binding free energy of protein/ligand complexes, whereas the Eisenberg−Mclachlan model cannot be used in this fashion. Second, we have used a very large training set to derive the parameters for different atom types, whereas the Eisenberg−McLachlan model only applied a limited number of experimental data to fit the weight parameters for five basic atom types (C, N/O, O−, N+, S). The third difference is that we have used a probe radius of 0.6 Å to generate the solvent accessible surface (SAS), whereas Eisenberg−McLachlan model used a probe radius of 1.4 Å. Because, in both of the models, the contribution from the intrinsic charges is completely neglected, smaller probe radii and inclusion of more atom types allow the electrostatic solvent effect to be better represented.

In the following, we will discuss how we parametrized this solvation model. Two applications of this model, to binding free energy calculation and protein folding analysis, will also be presented.

Methodology

Solvent Accessible Surface (SAS) Area Calculations. The molecules were built up with the SYBYL software package[35] of Tripos Inc. and optimized using the MMFF94 force field with MMFF charges.[36] For each molecule, only the global minimum conformation was used in the SAS calculation and the subsequent parametrization. Simple conformational searches were performed for some molecules in cases where the global minimum conformation was not obvious. Molecular surface areas were calculated using the MSMS program[37] and the probe radius was set to 0.6 Å with density of 3.0 vertex/Å^2. For each atom, the atomic SAS in each surface component were added together. In the next step, we applied an automated atom type determination program[38] to judge the atom types for given molecules. The solvation free energies were then estimated by eq 1.

Least-Squares Fitting Procedure. Least-squares fittings were applied to derive the solvation free energy weights, w_{i} in eq 1. There are two different procedures to perform the fittings: one may derive all the parameters simultaneously or derive them step-by-step. For the latter, one first derives the parameters of a certain compound class, such as hydrocarbons, and the developed parameters are transferred to other compound classes without change during the subsequent fittings. We found that the first procedure is superior because it avoids the biases aroused from the order in which the compound classes are fitted.

With the first fitting procedure, we worked out the parameters for four different solvation models. In models I and II, all the molecules investigated are from Li et al.'s training sets for their SM5.42R models. However, in model I, all the molecules (184) have the experimental solvation free energies in 1-octanol, whereas the molecules (245) in model II have aqueous solvation free energies. We presented the two models for comparison purpose. Unlike the Generalized Born model and the Poisson−Boltzmann model, there is no dielectric constants introduced in our simple WSAS model; therefore, the parameters developed for one solvent cannot be transferred to other solvents. For other two-solvation models, apart from all the molecules in model II, we extended our molecular set by including another 142 neutral molecules and 14 charged ones that have aqueous solvation free energies. In model III, we divided our whole molecule set excluding charged molecules into training set (293 molecules) and test set (94 molecules). The purpose is to investigate how well do the parameters developed with the training set transfer to the test set. In model IV, we developed the parameters using all the molecules including the charged ones available in this work, which were then used for binding free energy calculations in the following two applications.

Relative Binding Free Energy Calculations with WSAS. In another work,[39] we combined molecular docking and molecular dynamics simulations to model the complex of HIV-1RT/efavirenz, for which the crystal structure has been solved by Dupont Pharmaceuticals, but not published. We performed 500 pico-second MD simulations for the five binding modes suggested by the flexible docking and estimated the binding free energies for each binding mode using the MM-PBSA approach. In this work, instead of using PBSA, a more time-consuming solvation free energy approach, we applied WSAS to estimate the relative binding free energies for the five binding mode with the same snapshots as used in the MM-PBSA analysis.

Energy-Analysis on the MD Trajectories of Villin Headpiece. In 1998, Duan and Kollman[40] carried out a 1-μs molecular dynamics folding simulation for villin headpiece with explicit water. They found a very stable folding intermediate (from 240−400 ns), which is closer to the native structure than any other structure found along the trajectory. Subsequently, Lee et al.[34] applied MM-PBSA to analyze both a 1-μs trajectory and a 100 ns native trajectory. They found that the free energy of the folding intermediate was less favorable than that of the native one, but much more favorable than the unfolded random coil

TABLE 1: Definitions of Atom Types Used in Solvation Free Energy Calculation, 1−47 Are the Basic Atom Types and 48−53 Are Special Atom Types for Ions

no.	atom type symbol	explanation
1	CT3	sp^3 carbon attached by three hydrogens
2	CT2	sp^3 carbon attached by two hydrogens
3	CT1	sp^3 carbon attached by one hydrogens
4	CT	sp^3 carbon connected with four heavy atoms
5	CA	sp^2 aromatic carbon
6	CD2	sp^2 carbon attached by two hydrogens
7	CD1	sp^2 carbon attached by one hydrogens
8	CD	sp^2 carbon without hydrogens attached
9	CS	sp^1 carbon
10	CO	carbon in the carbonyl function group
11	CN	sp^2 carbon in the amide function group
12	F	any fluorine
13	Cl	any chlorine
14	Br	any bromine
15	I	any iodine
16	HCT3	hydrogen connected to sp^3 carbon that has three electron withdrawn group
17	HCT2	hydrogen connected to sp^3 carbon that has two electron withdrawn group
18	HCT1	hydrogen connected to sp^3 carbon that has one electron withdrawn group
19	HCT	hydrogen connected to sp^3 carbon that has no electron withdrawn group
20	HCD2	hydrogen connected to sp^2 carbon that has two electron withdrawn group
21	HCD1	hydrogen connected to sp^2 carbon that has one electron withdrawn group
22	HCD	hydrogen connected to sp^2 carbon that has no electron withdrawn group
23	HCA	hydrogen connected to aromatic carbon
24	HCS	hydrogen connected to sp carbon
25	HO	hydrogen connected to oxygen
26	HN	hydrogen connected to nitrogen
27	HS	hydrogen connected to sulfur
28	OH	hydroxyl oxygen in alcohol
29	OC	ether oxygen
30	O	oxygen in carbonyl function group
31	OS	sp^3 oxygen in ester
32	OOH	hydroxyl oxygen in acid
33	SH	sulfur in thiols
34	SS	sulfur in disulfide and thioether
35	SP	sp2 S connected to phosphate
36	NS	sp nitrogen
37	ND2	sp2 nitrogen in the heteroring with two connected atoms
38	ND3	sp2 nitrogen in the heteroring with three connected atoms
39	NO	nitrogen connected to oxygen
40	NH2	amide nitrogen with two attached hydrogens
41	NH1	amide nitrogen with one attached hydrogen
42	N	amide nitrogen without attached hydrogens
43	NTH2	amine nitrogen with two attached hydrogens
44	NTH	amine nitrogen with one attached hydrogen
45	NT	amine nitrogen without attached hydrogen
46	P3	phosphate with three connected atoms
47	P4	phosphate with four connected atoms
48	O-	acetate ion and propionate ion
49	OA-	*p*-cresole ion and $C_6H_5O^-$
50	S-	C_6H_5S- and CH_3S-
51	HN+	hydrogen connected to charged nitrogen
52	NT+	sp3 nitrogen cation
53	ND+	sp2 nitrogen cation, such as methyl-imidazolium
54	N+	N−p-guanidinium

(500 − 1000 ns). For the solvation free energy, the sequence order for the three conformation cluster is the opposite of the free energy, i.e., $\Delta G^{\text{solvation}}{}_{\text{native}} > \Delta G^{\text{solvation}}{}_{\text{intermediate}} > \Delta G^{\text{solvation}}{}_{\text{random coil}}$. In this work, we performed energy analysis for the two MD trajectories with WSAS substituting for PBSA.

Results and Discussion

Atom Type Definition. The solvation free energy is made up of two parts: the electrostatic and the nonpolar (cavity and van der Waals) contributions. The electrostatic part, which is charge-method dependent, can be evaluated using several different models, such as finite difference Poisson−Boltzmann (FDPB), dielectric screening Poisson−Boltzmann (DSPB),[14−15] the Generalized Born (GB),[16−21] and the conductor-like screening models.[22−24] The nonpolar part can be simply estimated using the total solvation accessible surface area (SAS), according to the formula $\Delta G = \gamma A + b$, where g and b are constants, and A is SAS. For WSAS, the two contributions are considered simultaneously, and the electrostatic part is taken into account implicitly using different atom types. In principle, atoms that are the same atom type should bear similar charges. With this in mind, we selected the atom types in the following way. We began with only a few very basic atom types and introduced specific ones only if they could reduce the error significantly. After introducing new atom types, it was necessary to check whether one was overfitting the data or not. Table 1 lists the entire set of atom types in our WASA models.

Probe Radius. In other solvation models (PBSA, GBSA,

TABLE 2: Solvation Free Energies of Organic Compounds[a]

no.	compound name	model(I)	model(II)	model(III)	model(IV)	expt (a)	expt (b)
1	Methane	0.52	2.01	2.01	2.01	1.98	0.51
2	Ethane	−1.18	0.86	0.81	0.83	1.83	−0.64
3	Propane	−1.76	0.93	0.84[b]	0.86	1.96	−1.26
4	Cyclopropane	−1.21	0.01	0.25	−0.12	0.75	−1.60
5	2-methyl propane	−2.22	0.88	0.69[b]	0.77	2.32	−1.45
6	2,2-dimethyl propane	−2.31	1.37	1.20[b]	1.30	2.50	−1.74
7	*n*-butane	−2.26	1.06	0.92	0.96	2.08	−1.86
8	2,2-dimethylbutane			1.26[b]	1.36	2.59	
9	Cyclopentane	−2.71	0.56	0.46	0.36	1.20	−2.65
10	*n*-pentane	−2.75	1.19	1.00	1.05	2.33	−2.45
11	2-methylpentane		1.22	0.98	1.07	2.52	
12	3-methyl pentane			1.07	1.16	2.51	
13	2,4-dimethyl pentane		1.28	0.98[b]	1.11	2.88	
14	2,2,4-trimethyl pentane		1.60	1.32	1.46	2.85	
15	Methylcyclopentane			0.42	0.40	1.60	
16	*n*-hexane	−3.25	1.31	1.08	1.15	2.49	−3.01
17	Cyclohexane	−3.29	0.76	0.55[b]	0.54	1.23	−3.46
18	Methylcyclohexane	−3.62	0.90	0.61	0.66	1.71	−3.21
19	Cis-1,2-dimethylcyclohexane		1.02	0.71	0.78	1.58	
20	*n*-heptane	−3.74	1.44	1.16[b]	1.24	2.62	−3.74
21	*n*-octane	−4.24	1.57	1.24[b]	1.33	2.89	−4.18
22	Ethylene	−0.52	0.68	0.93	0.81	1.27	−0.27
23	Propylene	−1.58	0.92	0.93	0.87	1.27	−1.14
24	2-methyl propene	−1.56	1.16	1.09[b]	1.10	1.16	−2.03
25	1-butene	−2.02	1.42	1.24	1.14	1.38	−1.89
26	2-methyl-2-butene			1.21[b]	1.24	1.31	
27	3-methyl-1-butene			0.98	0.95	1.83	
28	1-pentene			1.33	1.24	1.66	
29	Trans-2-pentene		1.47	1.14	1.12	1.34	
30	4-methyl-1-pentene			1.16	1.14	1.91	
31	Cyclopentene		−0.36	−0.21	−0.24	0.56	
32	1-hexene	−3.01	1.69	1.41	1.34	1.66	−2.94
33	Cyclohexene			0.76	0.69	0.37	
34	Trans-2-heptene			1.32	1.31	1.66	
35	1-methylcyclohexene			0.90	0.91	0.67	
36	1-octene			1.57	1.52	2.17	
37	1,3-butadiene	−2.04	0.53	0.75	0.66	0.61	−2.10
38	2-methyl-1,3-butadiene			0.87	0.86	0.68	
39	2,3-dimethyl-1,3-butadiene			0.95	1.02	0.40	
40	1,4-pentadiene			1.72[b]	1.48	0.94	
41	1,5-hexadiene			1.73	1.52	1.01	
42	Acetylene	−0.53	−0.03	0.02	0.03	−0.01	−0.51
43	Propyne	−1.70	−0.35	−0.11	−0.05	−0.48	−1.59
44	1-butyne			−0.03	0.02	−0.15	
45	1-pentyne	−2.78	−0.09	0.05	0.12	−0.16	−2.79
46	1-hexyne	−3.27	0.04	0.13[b]	0.21	0.01	−3.43
47	1-heptyne			0.21	0.30	0.60	
48	1-octyne			0.29	0.40	0.71	
49	1-nonyne			0.38[b]	0.49	1.05	
50	Butenyne		−0.21	0.04	0.02	0.04	
51	Benzene	−4.38	−1.83	−1.44	−1.56	−0.89	−3.72
52	Toluene	−4.80	−1.11	−1.11	−1.15	−0.76	−4.55
53	1,2,4-trimethylbenzene			−0.57	−0.47	−0.86	
54	Ethylbenzene	−4.89	−0.95	−0.84	−0.91	−0.61	−5.08
55	*m*-xylene	−5.24	−0.40	−0.79	−0.76	−0.80	−5.25
56	*o*-xylene	−5.22	−0.52	−0.87[b]	−0.87	−0.90	−5.07
57	*p*-xylene	−5.22	−0.39	−0.77	−0.74	−0.80	−5.19
58	Propylbenzene			−0.76	−0.81	−0.53	
59	2-propylbenzene			−0.85[b]	−0.86	−0.30	
60	Butylbenzene			−0.49	−0.51	−0.40	
61	2-butylbenzene			−0.58	−0.64	−0.45	
62	*tert*-butylbenzene			−0.50	−0.47	−0.44	
63	*tert*-amylbenzene			−0.24	−0.23	−0.18	
64	Naphthalene	−7.12	−2.41	−2.52[b]	−2.58	−2.41	−6.97
65	Anthracene	−9.92	−2.99	−3.62	−3.64	−4.23	−10.47
66	Phenanthrene			−3.73	−3.72	−4.06	
67	Acenaphthene			−2.39[b]	−2.42	−3.40	
68	*p*-chlorotoluene		−0.94	−1.23[b]	−1.22	−1.92	

TABLE 2: (Continued)

no.	compound name	model(I)	model(II)	model(III)	model(IV)	expt (a)	expt (b)
69	Fluoromethane		0.22	0.11	−0.08	−0.22	
70	1,1-difluoroethane	−1.62	0.10	0.03	0.03	−0.11	−1.13
71	Trifluoromethane			0.43	0.28	0.80	
72	Tetrafluoromethane	0.16	2.66	2.44	2.45	3.16	1.50
73	Hexafluoroethane		3.50	3.21	3.22	3.94	
74	Octafluoropropane		4.12	3.78[b]	3.80	4.28	
75	Fluorobenzene	−4.40	−0.86	−0.92	−0.97	−0.78	−3.87
76	2-chloro-1,1,1-trifluoroethane			1.07	0.97	0.05	
77	Chlorofluoromethane			−0.55[b]	−0.60	−0.77	
78	Chlorodifluoromethane			0.01	−0.13	0.11	
79	Chlorotrifluoromethane	−0.84	1.82	1.73	1.72	2.52	−1.97
80	Dichlorodifluoromethane	−1.79	1.03	1.05	1.03	1.69	−1.25
81	Fluorotrichloromethane	−2.71	0.27	0.40	0.38	0.82	−2.63
82	1,1,2-trichloro-1,2,2-trifluoroethane	−2.42	1.31	1.34	1.32	1.77	−2.54
83	1,1,2,2-tetrachlorodifluoroethane			0.74[b]	0.72	0.82	
84	1,1,2-trichlorotrifluoroethane			1.33	1.31	1.77	
85	Chloropentafluoroethane			2.56	2.56	2.86	
86	1,1-dichorotetrafluoroethane			1.94	1.93	2.50	
87	1,2-dichlorotetrafluoroethane			1.93[b]	1.92	2.31	
88	1-bromo-1-chloro-2,2,2-trifluoroethane	−2.84	0.38	0.45[b]	0.53	−0.13	−3.27
89	Bromotrifluoromethane	−1.47	1.35	1.23[b]	1.31	1.79	−0.75
90	1-bromo-1,2,2,2-tetrafluoroethane		1.02	0.98	1.07	0.52	
91	Chloromethane			−0.67	−0.85	−0.56	
92	Dichloromethane	−3.20	−1.76	−1.25[b]	−1.30	−1.36	−3.07
93	Trichloromethane	−4.08	−1.29	−0.97	−1.10	−1.07	−3.81
94	Tetrachloromethane			−0.22	−0.25	0.10	
95	Chloroethane	−2.45	−0.24	−0.27	−0.39	−0.63	−2.58
96	1,1-dichloroethane			−0.95	−0.95	−0.85	
97	E−1,2-dichloroethane			−1.22[b]	−1.46	−1.73	
98	1,1,1-trichloroethane	−3.31	0.03	0.14	0.15	−0.25	−3.69
99	1,1,2-trichloroethane	−4.38	−1.84	−1.64	−1.76	−1.95	−4.53
100	1,1,1,2-tetrachloroethane		−0.86	−0.73	−0.84	−1.15	
101	1,1,2,2-tetrachloroethane			−1.62	−1.61	−2.36	
102	Pentachloroethane			−0.88	−0.89	−1.36	
103	Hexachloroethane		−0.60	−0.30[b]	−0.34	−1.40	
104	1-chloropropane	−2.94	−0.05	−0.13	−0.23	−0.35	−3.06
105	2-chloropropane	−2.82	−0.02	−0.15[b]	−0.16	−0.24	−2.84
106	1,2-dichloropropane			−0.89	−1.03	−1.25	
107	1,3-dichloropropane			−1.06	−1.29	−1.90	
108	1-chlorobutane			−0.04	−0.13	−0.14	
109	2-chlorobutane		0.29	0.14	0.15	0.07	
110	1,1-dichlorobutane			−0.49	−0.46	−0.70	
111	1-chloropentane		0.21	0.04	−0.03	−0.07	
112	2-chloropentane		0.41	0.22	0.24	0.07	
113	3-chloropentane			0.42	0.43	0.07	
114	Chloroethylene		0.00	0.02	−0.04	0.49	
115	Cis-1,2-dichloroethylene	−3.59	−0.52	−0.75	−0.78	−1.17	−3.71
116	Trans-1,2-dichloroethylene	−3.60	−0.46	−0.69	−0.72	−0.76	−3.61
117	Trichloroethylene	−3.94	−0.05	−0.33[b]	−0.34	−0.44	−3.75
118	Tetrachloroethylene	−4.29	0.42	0.04	0.05	0.05	−4.24
119	3-chloropropane		0.30	0.19[b]	−0.06	−0.57	
120	Chlorobenzene	−5.28	−1.65	−1.56	−1.63	−1.01	−5.00
121	Chlorotoluene		−0.99	−1.28	−1.27	−1.92	
122	*o*-chlorotoluene		−0.93	−1.23	−1.22	−1.15	
123	1,2-dichlorobenzene	−6.12	−1.50	−1.70	−1.71	−1.36	−6.01
124	1,3-dichlorobenzene			−1.69[b]	−1.70	−0.98	
125	1,4-dichlorobenzene	−6.18	−1.47	−1.69	−1.70	−1.01	−5.67
126	2,2?-dichlorobiphenyl	−8.36	−2.61	−2.41	−2.52	−2.73	−9.41
127	2,3-dichlorobiphenyl	−8.59	−2.63	−2.51[b]	−2.61	−2.45	−9.23
128	2,2,3?-trichlorobiphenyl	−9.16	−2.46	−2.52	−2.59	−1.99	−9.12
129	Bromotrichloromethane		−0.81	−0.61[b]	−0.56	−0.93	
130	1-chloro-2-bromoethane			−1.60[b]	−1.74	−1.95	
131	Bromomethane	−2.68	−1.09	−1.11	−1.19	−0.82	−2.43
132	Dibromomethane	−4.34	−2.41	−2.00[b]	−1.86	−2.11	−4.18
133	Tribromomethane	−5.31	−1.76	−1.51	−1.36	−1.98	−5.62
134	Bromoethane	−3.02	−0.62	−0.68	−0.71	−0.70	−2.90
135	1,2-dibromoethane			−1.98	−2.01	−2.10	
136	1-bromopropane	−3.52	−0.43	−0.53	−0.53	−0.56	−3.42
137	2-bromopropane	−3.28	−0.30	−0.44	−0.36	−0.48	−3.40
138	1,2-dibromopropane			−1.47	−1.42	−1.94	
139	1,3-dibromopropane			−1.83	−1.87	−1.96	

TABLE 2: (Continued)

no.	compound name	model(I)	model(II)	model(III)	model(IV)	expt (a)	expt (b)
140	1-bromo-2-methylpropane			−0.49[b]	−0.45	−0.03	
141	1-bromobutane	−4.01	−0.30	−0.45	−0.43	−0.41	−4.16
142	1-bromo-isobutane		−0.34	−0.49	−0.46	−0.03	
143	1-bromo-3-methylbutane			−0.41	−0.34	0.20	
144	1-bromopentane	−4.50	−0.17	−0.37	−0.34	−0.08	−4.68
145	3-bromopropene	−3.21	0.05	−0.13	−0.30	−0.86	−3.30
146	Bromobenzene	−5.69	−2.01	−1.91	−1.89	−1.46	−5.46
147	1,4-dibromobenzne	−7.01	−2.20	−2.38[b]	−2.23	−2.30	−7.47
148	*p*-bromotoluene	−6.11	−1.30	−1.56	−1.49	−1.39	−6.36
149	1-bromo-2-ethylbenzene			−1.34[b]	−1.30	−1.19	
150	*o*-bromocumene			−1.17	−1.08	−0.85	
151	Iodomethane	−3.41	−1.11	−1.18	−1.27	−0.89	−3.07
152	Diiodomethane	−5.52	−2.41	−2.07	−1.95	−2.49	−5.63
153	Iodoethane	−3.68	−0.62	−0.72[b]	−0.75	−0.72	−3.45
154	1-iodopropane		−0.42	−0.57	−0.57	−0.58	
155	2-iodopropane	−3.88	−0.27	−0.44[b]	−0.37	−0.46	−4.40
156	1-iodobutane		−0.29	−0.48	−0.48	−0.26	
157	1-iodopentane		−0.17	−0.40	−0.38	−0.12	
158	Iodobenzene	−6.29	−2.00	−1.94[b]	−1.93	−1.73	−6.18
159	Methanol	−4.64	−5.83	−5.95	−6.08	−5.07	−3.87
160	Ethanol	−4.44	−4.78	−5.15	−5.19	−4.90	−4.36
161	2,2,2-trifluoroethanol			−3.54	−3.53	−4.30	
162	Ethylene glycol	−8.23	−10.41	−10.62	−10.82	−9.30	−7.44
163	1-propanol	−4.93	−4.58	−5.01[b]	−5.02	−4.85	−5.02
164	2-propanol	−5.14	−4.70	−5.00	−4.96	−4.75	−4.62
165	1,1,1-trifluoro-2-propanol	−5.29	−3.75	−3.65	−3.70	−4.16	−5.12
166	2,2,3,3-tetrafluoropropanol			−4.33	−4.47	−4.90	
167	2,2,3,3,3-pentafluoropropanol			−2.96[b]	−3.10	−4.15	
168	Hexafluoro-2-propanol	−4.43	−2.21	−2.18	−2.22	−3.76	−5.76
169	2-methyl-1-propanol			−4.44[b]	−4.35	−4.51	
170	1-butanol	−5.42	−4.44	−4.92	−4.92	−4.72	−5.71
171	2-butanol			−4.32	−4.26	−4.61	
172	*tert*-butyl alcohol	−5.02	−3.21	−3.42	−3.31	−4.51	−4.78
173	2-methyl-1-butanol			−4.21	−4.12	−4.42	
174	3-methyl butanol			−6.49[b]	−6.42	−4.42	
175	2-methyl-2-butanol			−2.76	−2.60	−4.43	
176	2,3-dimethylbutanol			−4.10	−4.05	−3.91	
177	1-pentanol	−6.84	−4.94	−5.03	−5.10	−4.49	−6.40
178	2-pentanol			−4.21	−4.13	−4.39	
179	3-pentanol			−4.20	−4.19	−4.35	
180	2-methyl-1-pentanol			−4.01	−3.89	−3.93	
181	2-methyl-2-pentanol			−3.11	−3.04	−3.93	
182	2-methyl-3-pentanol			−3.82	−3.73	−3.89	
183	4-methyl-2-pentanol			−3.63[b]	−3.45	−3.74	
184	Cyclopentanol			−5.20[b]	−5.26	−5.49	
185	1-hexanol	−7.33	−4.81	−4.96	−5.01	−4.36	−7.06
186	3-hexanol			−3.48	−3.40	−3.68	
187	Cyclohexanol			−4.94	−4.91	−4.95	
188	4-heptanol			−3.38[b]	−3.28	−4.01	
189	Cycloheptanol			−4.80	−4.77	−5.49	
190	1-heptanol	−7.83	−4.69	−4.88	−4.92	−4.25	−7.75
191	1-octanol	−8.34	−4.57	−4.80	−4.83	−4.10	−8.13
192	1-decanol	−9.38					−9.88
193	Allyl alcohol	−4.76	−4.49	−4.85	−5.00	−5.03	−5.27
194	Cyclopentanol		−4.67	−5.20	−5.26	−5.49	
195	Phenol	−8.41	−6.54	−6.58	−6.57	−6.53	−8.69
196	4-bromophenol	−9.68	−6.69	−7.03[b]	−6.89	−7.10	−10.59
197	4-nitrophenol			−9.13	−8.83	−10.60	
198	4-*tert*-butylphenol			−5.64	−5.49	−5.92	
199	2-cresol	−7.63	−4.79	−5.54	−5.41	−5.86	−8.49
200	3-cresol	−8.86	−5.82	−6.25	−6.18	−5.49	−8.20
201	4-cresol	−8.82	−5.81	−6.24[b]	−6.17	−6.12	−8.84
202	2,2,2-trifluoroethanol	−5.26	−4.16	−3.95	−4.10	−4.31	−4.81
203	*p*-bromophenol		−6.69	−7.03	−6.89	−7.13	
204	2-methoxyethanol	−6.44	−7.79	−7.72	−7.85	−6.77	−5.83
205	Dimethoxymethane			−4.36[b]	−4.30	−2.93	
206	Methyl propyl ether	−3.58	−1.89	−1.59[b]	−1.65	−1.66	−3.63
207	Methyl isopropyl ether	−3.32	−1.76	−1.53	−1.52	−2.00	−4.64
208	Methyl *tert*-butyl ether	−3.19	−0.67	−0.46[b]	−0.36	−2.21	−3.49
209	Diethyl ether	−3.45	−1.40	−1.16	−1.21	−1.75	−2.89

TABLE 2: (Continued)

no.	compound name	model(I)	model(II)	model(III)	model(IV)	expt (a)	expt (b)
210	Ethyl propyl ether			−1.02	−1.04	−1.81	
211	Dipropyl ether			−0.88[b]	−0.88	−1.16	
212	Diisopropyl ether		−0.15	−0.12[b]	0.04	−0.53	
213	di-*n*-butyl ether			−0.71	−0.68	−0.83	
214	Tetrahydrofuran	−3.72	−2.78	−2.29	−2.44	−3.12	−3.93
215	2-methyltetrahydrofuran			−2.20[b]	−2.21	−3.30	
216	Anisole	−6.20	−3.41	−3.15	−3.16	−2.45	−5.47
217	Ethyl phenyl ether	−6.54	−2.69	−2.54	−2.50	−4.28	−5.65
218	1,1-diethoxyethane			−3.00	−2.91	−3.27	
219	1,2-dimethoxyethane	−4.96	−5.01	−4.25	−4.42	−4.84	−4.55
220	1,2-diethoxyethane			−3.11	−3.20	−3.53	
221	1,3-dioxolane			−5.07	−5.01	−4.09	
222	1,4-dioxane	−5.06	−5.70	−4.72	−4.87	−5.05	−4.89
223	2,2,2-trifluoroethyl vinyl ether		−0.03	0.26	0.33	−0.12	
224	1-chloro-2,2,2-trifluoroethyl difluoromethyl ether		0.15	0.44[b]	0.48	0.11	
225	1,1-dichloro-2,2-difluoroethylmethyl ether	−3.66					−4.02
226	Formaldehyde	−3.68					−3.23
227	Acetaldehyde		−3.48	−3.65	−3.74	−3.50	
228	Propanal	−4.04	−3.15	−3.38[b]	−3.52	−3.44	−4.13
229	Butanal	−4.56	−3.00	−3.29	−3.41	−3.18	−4.62
230	Pentanal		−2.88	−3.21	−3.32	−3.03	
231	Hexanal			−3.13	−3.23	−2.81	
232	Heptanal			−3.05[b]	−3.13	−2.67	
233	Octanal		−2.50	−2.97	−3.04	−2.29	
234	Nonanal			−2.88	−2.94	−2.07	
235	Trans-2-butenal			−3.94	−3.97	−4.23	
236	Trans-2-hexenal			−3.55[b]	−3.60	−3.68	
237	Trans-2-octenal			−3.39	−3.42	−3.44	
238	Trans,*trans*-2,4-hexadienal			−4.26	−4.24	−4.64	
239	Benzaldehyde	−7.13	−5.41	−5.48	−5.54	−4.02	−6.13
240	m-hydroxybenzaldehyde	−11.09	−10.07	−10.58	−10.53	−9.51	−11.39
241	p-hydroxybenzaldehyde	−11.10	−10.08	−10.59	−10.54	−10.47	−12.36
242	Acetone	−3.56	−4.04	−3.45	−3.50	−3.80	−3.15
243	2-butanone	−3.96	−3.57	−3.11[b]	−3.24	−3.71	−3.78
244	3-methyl-2-butanone			−2.88	−2.56	−3.24	
245	3,3-dimethylbutanone	−4.18	−2.32	−2.07	−2.14	−2.89	−4.53
246	2-pentanone	−4.39	−3.29	−2.86[b]	−2.87	−3.52	−4.35
247	3-pentanone	−4.24	−2.87	−2.63	−2.96	−3.41	−4.36
248	4-methyl-2-pentanone			−2.49	−2.50	−3.06	
249	2,4-dimethyl-3-pentanone			−1.34[b]	−1.40	−2.74	
250	Cyclopentanone	−4.72	−4.42	−3.89	−4.08	−4.68	−5.01
251	2-hexanone	−4.88	−3.16	−2.77	−2.77	−3.41	−5.02
252	2-heptanone	−5.38	−3.03	−2.69	−2.68	−3.04	−5.65
253	4-heptanone		−2.54	−2.27[b]	−2.23	−2.93	
254	2-octanone	−5.87	−2.90	−2.61	−2.58	−2.88	−6.38
255	2-nonanone			−2.53	−2.45	−2.48	
256	5-nonanone		−2.28	−2.10	−2.03	−2.67	
257	2-undecanone			−2.37	−2.30	−2.15	
258	Acetophenone	−7.15	−5.51	−5.06	−5.11	−4.58	−6.74
259	Acetic acid	−6.82	−6.91	−6.90	−6.58	−6.70	−6.35
260	Propionic acid	−6.62	−6.00	−6.29[b]	−6.04	−6.46	−6.86
261	Butyric acid	−7.17	−5.84	−6.19	−5.93	−6.35	−7.58
262	Pentanoic acid	−8.11	−5.98	−6.20	−5.83	−6.16	−8.22
263	Hexanoic acid	−8.61	−5.85	−6.12	−5.74	−6.21	−8.82
264	4-amino-3,5,6-trichloropyridine-2-carboxylic acid	−12.85	−13.32	−14.78[b]	−13.83	−11.96	−12.37
265	Methyl formate	−2.82	−2.81	−2.42	−2.59	−2.78	−2.82
266	Ethyl formate		−2.62	−2.19	−2.29	−2.65	
267	Propyl formate			−2.06	−2.14	−2.48	
268	Methyl acetate	−3.62	−3.26	−3.18[b]	−3.27	−3.31	−3.54
269	Isopropyl formate			−3.09	−3.08	−2.02	
270	Isobutyl formate			−3.21	−3.21	−2.22	
271	Isoamyl formate			−1.44	−1.45	−2.13	
272	Ethyl acetate	−4.13	−3.06	−2.93[b]	−2.96	−3.08	−4.06
273	Propyl acetate	−4.64	−2.89	−2.80	−2.81	−2.85	−4.55
274	Isopropyl acetate			−2.56	−2.52	−2.65	
275	Butyl acetate	−5.14	−2.76	−2.72	−2.72	−2.55	−4.96
276	Isobutyl acetate			−2.30	−2.26	−2.36	
277	Amyl acetate			−2.64[b]	−2.62	−2.45	
278	Isoamyl acetate			−2.21	−2.15	−2.21	

TABLE 2: (Continued)

no.	compound name	model(I)	model(II)	model(III)	model(IV)	expt (a)	expt (b)
279	Hexyl acetate			−2.56	−2.53	−2.26	
280	Methyl propionae	−3.98	−2.68	−2.71	−2.78	−2.97	−4.06
281	Ethyl propionate			−2.46	−2.48	−2.80	
282	Propyl propionate		−2.38	−2.34	−2.33	−2.54	
283	Isopropyl propionate			−2.10	−2.04	−2.22	
284	Amyl propionate			−2.17	−2.14	−1.99	
285	Methyl butyrate	−4.46	−2.52	−2.59[b]	−2.65	−2.84	−4.59
286	Ethyl butyrate			−2.35	−2.35	−2.50	
287	Propyl butyrate			−2.22	−2.20	−2.28	
288	Methyl pentanoate	−4.96		−2.51	−2.55	−2.54	−5.13
289	Ethyl pentanoate			−2.26	−2.25	−2.52	
290	Methyl hexanonate		−2.26	−2.43	−2.46	−2.48	
291	Ethyl heptanoate			−2.10[b]	−2.06	−2.30	
292	Methyl octanoate		−2.01	−2.27	−2.27	−2.05	
293	Methyl benzoate	−7.29	−5.36	−5.19	−5.32	−4.28	−7.26
294	Methylamine	−3.80	−4.74	−5.20	−5.08	−4.60	−3.78
295	Ethylamine	−4.53	−4.45	−4.69	−4.56	−4.61	−4.09
296	Propylamine	−5.03	−4.27	−4.55[b]	−4.40	−4.50	−4.77
297	Butylamine	−5.52	−4.14	−4.46	−4.30	−4.38	−5.35
298	Pentylamine		−4.02	−4.38	−4.21	−4.09	
299	Hexylamine			−4.30	−4.12	−4.04	
300	Dimethylamine		−4.65	−4.79[b]	−4.78	−4.28	
301	Diethylamine	−4.23	−2.93	−3.02	2.98	−4.06	−4.75
302	Dipropylamine	−5.16	−2.52	−2.70	−2.61	−3.65	−6.02
303	Dibutylamine			−2.53	−2.42	−3.31	
304	Trimethylamine	−3.60	−3.89	−3.91	−4.16	−3.23	−3.60
305	Triethylamine			−1.64[b]	−1.74	−3.03	
306	Azetidine		−4.68	−4.98	−4.98	−5.56	
307	Piperazine	−6.77	−9.18	−9.34	−9.20	−7.40	−5.80
308	N,n?-dimethylpiperazine		−7.03	−7.08[b]	−7.43	−7.58	
309	*N*-methylpiperazine		−8.11	−8.22	−8.32	−7.77	
310	Aniline	−7.46	−5.71	−5.89	−5.70	−5.49	−6.71
311	1,1-dimethyl-3-phenyl urea		−11.72	−10.76	−10.41	−11.87	
312	N?N-dimethyl aniline			−3.73	−3.94	−2.90	
313	Ethylenediamine			−8.62	−8.36	−9.75	
314	HydrazineHydrazine	−5.62	−8.64	−9.51	−8.92	−9.30	−6.48
315	2-methoxyethanamine		−7.23	−7.13[b]	−7.08	−6.55	
316	Morpholine	−5.92	−7.45	−7.05[b]	−7.04	−7.17	−5.99
317	*N*-methylmorpholine		−6.36	−5.90	−6.15	−6.34	
318	n-methylpyrrolidine			−3.54	−3.81	−3.97	
319	n-methylpiperidine		−3.36	−3.07	−3.27	−3.89	
320	Pyrrolidine		−4.43	−4.58	−4.58	−5.47	
321	Piperidine	−5.32	−4.33	−4.38	−4.34	−5.10	−6.27
322	Pyridine	−5.62	−4.92	−3.61	−4.02	−4.69	−5.34
323	2-methylpyridine	−5.28	−4.43	−4.05	−4.11	−4.62	−6.14
324	3-methylpyridine	−6.05	−4.06	−3.51[b]	−3.79	−4.77	−6.40
325	4-methylpyridine	−6.13	−4.15	−3.27	−3.60	−4.92	−6.60
326	2-ethylpyridine	−5.56	−3.67	−3.57	−3.64	−4.32	−6.40
327	3-ethylpyridine			−3.38[b]	−3.63	−4.60	
328	4-ethylpyridine			−2.99	−3.34	−4.72	
329	2,3-dimethylpyridine			−3.70	−3.71	−4.81	
330	2,4-dimethylpyridine		−3.64	−3.72[b]	−3.69	−4.85	
331	2,5-dimethylpyridine		−3.74	−3.89	−3.82	−4.70	
332	2,6-dimethylpyridine		−3.99	−4.51	−4.22	−4.60	
333	3,4-dimethylpyridine			−3.19	−3.40	−5.21	
334	3,5-dimethylpyridine			−3.26[b]	−3.42	−4.84	
335	3,4-dimethylpyridine			−3.19[b]	−3.93	−5.21	
336	3,5-dimethylpyridine			−3.26	−3.41	−4.84	
337	2-methylpyrazine	−6.67	−7.26	−6.81[b]	−7.06	−5.51	−5.87
338	2-ethylpyrazine			−6.52	−6.74	−5.45	
339	2-isobutylpyrazine			−6.68	−6.68	−5.05	
340	2-ethyl-3-methoxypyrazine	−6.93		−6.67	−6.88	−4.39	−6.85
341	2-isobutyl-3-methoxypyrazine			−5.00	−5.31	−3.68	
342	9-methyladenine		−13.60	−13.94	−13.74	−13.60	
343	1-methylthymine		−10.63	−11.43	−11.84	−10.40	
344	Methylindole			−5.58	−5.63	−5.91	
345	Methylimidazole			−9.61	−9.40	−10.25	
346	N-propyl guanidine			−11.05	−11.18	−10.92	
347	Acetonitrile	−3.14		−3.19	−3.31	−3.89	−3.15
348	Propionitrile	−3.72	−3.23	−3.11[b]	−3.24	−3.85	−3.66

TABLE 2: (Continued)

no.	compound name	model(I)	model(II)	model(III)	model(IV)	expt (a)	expt (b)
349	Butyronitrile	−4.22	−3.10	−3.03	−3.14	−3.64	−4.25
350	Nitroethane	−3.83	−3.58	−3.77	−3.94	−3.71	−3.93
351	Benzonitrile	−6.75	−5.28	−5.04	−5.24	−4.10	−6.09
352	2,6-dichlorobenzonitrile	−8.52	−4.94	−5.28[b]	−5.39	−5.22	−9.18
353	3,5-dibromo-4-hydroxybenzonitrile		−9.82	−10.09	−10.00	−9.00	
354	1-nitropropane	−4.33	−3.38	−3.62	−3.77	−3.34	−4.44
355	2-nitropropane	−4.57	−3.83	−4.03[b]	−4.23	−3.14	−4.23
356	1-nitrobutane	−4.82	−3.26	−3.54	−3.68	−3.08	−5.11
357	Nitrobenzene	−6.13	−3.60	−4.02	−3.84	−4.12	−6.63
358	2-nitrotoluene	−7.08	−4.18	−4.61	−4.67	−3.59	−6.80
359	3-nitrotoluene			−3.70	−3.44	−3.45	
360	N'N-dimethyl formamide			−5.47	−5.57	−4.90	
361	*N*-methylformamide			−9.37	−9.39	−10.00	
362	Acetamide		−9.72	−9.12	−9.56	−9.72	
363	E−*N*-methylacetamide		−9.39	−9.39	−9.29	−10.00	
364	Z−*N*-methylacetamide		−10.53	−10.10	−10.02	−10.00	
365	Propionamide			−8.59[b]	−8.92	−9.42	
366	Methanethiol		−1.61	−1.60	−1.66	−1.24	
367	Ethanethiol		−1.13	−1.10	−1.07	−1.30	
368	1-propanethiol	−3.52	−0.90	−1.00[b]	−1.01	−1.05	−3.52
369	Thiophenol	−5.99	−2.47	−2.42	−2.41	−2.55	−5.99
370	Thioanisole	−6.41	−2.66	−2.87	−2.69	−2.73	−6.47
371	Dimethyl sulfide	−3.10	−1.82	−1.98	−1.90	−1.54	−4.24
372	Diethyl sulfide	−3.76	−0.72	−0.94	−0.79	−1.43	−4.09
373	Methyl ethyl sulfide			−1.93	−1.73	−1.49	
374	Dipropyl sulfide		−0.33	−0.64	−0.44	−1.27	−3.89
375	Bis(2-chloroethyl) sulfide		−2.66	−2.84[b]	−2.92	−3.92	
376	Dimethyl disulfide		−2.47	−2.63[b]	−2.19	−1.83	
377	Diethyl disulfide		−1.37	−1.59	−1.09	−1.63	
378	Trimethyl phosphate	−6.77	−8.34	−8.34	−8.15	−8.70	−7.81
379	Triethyl phosphate	−7.73	−5.97	−6.07[b]	−5.99	−7.80	−8.88
380	Tripropyl phosphate	−9.23	−5.39	−5.64	−5.50	−6.10	−8.65
381	2,2-dichloroethenyl dimethyl phosphate	−8.69	−7.75	−7.67	−7.55	−6.61	−8.59
382	Dimethyl-4-bromo-2,5-dichlorophenyl thiophosphate	−12.24	−5.51	−5.34[b]	−5.49	−5.70	−12.30
383	Dimethyl 2,4,5-trichlorophenyl thiophosphate	−11.80	−5.21	−5.06	−5.22	−5.06	−11.69
384	Dimethyl 4-nitrophenyl thiophosphate	−11.39	−7.44	−7.42[b]	−7.40	−7.62	−11.70
385	Dimethyl 5-(4-chloro) bicyclo[3.2.0] heptyl phosphate		−7.46	−7.17	−7.33	−7.28	
386	Diethyl 2,4-dichlorophenyl thiophosphate	−11.13	−5.02	−4.82	−5.03	−3.86	−10.87
387	Dimethyl 3-methyl-4-thiomethoxyphenyl thiophosphate	−11.75	−6.20	−5.98	−6.14	−6.92	−12.55
388	Ethyl 4-cyanophenyl phenylthiophosphonate	−11.05	−5.07	−5.08	−5.07	−5.10	−11.06
389	Diethyl 4-nitrophenyl thiophosphonate	−12.12	−6.07	−6.26	−6.17	−6.27	−11.31
390	*o*-ethyl o?−4-bromo-2-chlorophenyl s-propyl phosphorothioate	−12.67	−6.14	−7.04[b]	−6.61	−4.09	−10.49
391	Ammonium				−81.53	−81.53	
392	CH_3NH_3+				−75.15	−73.99	
393	$(CH_3)_2NH_2$+				−67.00	−67.00	
394	$(CH_3)_3NH^+$				−63.00	−63.00	
395	$C_6H_5NH_3$+				−70.67	−72.00	
396	N-butylammonium				−69.32	−69.24	
397	Methyl imidazolium				−64.13	−64.13	
398	N−p-guanidinium				−66.07	−66.07	
399	Methylthiol ion				−77.33	−76.79	
400	$C_6H_5S^-$				−70.38	−70.99	
401	Acetate ion				−82.18	−80.65	
402	Propionate ion				−77.50	−79.12	
403	$C_6H_5O^-$				−75.71	−76.00	
404	*p*-cresole ion				−75.30	−75.01	

[a] Expt (a) lists the solvation free energies in water (Refs 1, 6, 12, and 46) and (b) lists the solvation free energies in 1-octanol (Ref 1). [b] Compounds in test set for model 3.

etc.), a probe radius of 1.4 Å is usually applied for SAS calculation. However, in this work, a probe radius of 0.6 was used. The reason is that with a smaller probe radius, a more precise description of the surface area can be achieved compared to that with a larger probe radius. For large molecules, the interior atoms, although, they have no accessible surface area, can still make a contribution to the electrostatic solvation free energy. However, in WSAS, those contributions are totally neglected. This deficiency of the simple models can be partially reduced by using a smaller probe radius. Moreover, a probe radius of 0.6 is still in the reliable radius range for the MSMS program. In the following, we also gave an example of how the probe radius affecting the performance of the WSAS models.

Solvation Models. Table 2 lists all the solvation free energies

TABLE 3: Performance of WSAS (Mode 4) by Solute Function Class[a]

	model I			model IV		
solute class	no. of solute	unsigned mean error	RMS error	no. of solute	unsigned mean error	RMS error
alkanes	14	0.32	0.39	21	1.15	1.20
alkenes	7	0.21	0.27	21	0.38	0.46
alkynes	3	0.09	0.11	8	0.28	0.32
aromatic hydrocarbons	8	0.25	0.33	18	0.33	0.42
fluorides	9	0.60	0.72	22	0.46	0.51
chlorides	17	0.25	0.36	40	0.32	0.40
bromides	12	0.20	0.23	20	0.26	0.33
iodinates	5	0.26	0.30	8	0.22	0.27
alcohols	22	0.45	0.56	45	0.56	0.76
ethers	10	0.50	0.62	20	0.66	0.73
aldehydes	6	0.53	0.69	15	0.45	0.60
ketones	10	0.27	0.31	17	0.53	0.61
acids	6	0.32	0.35	6	0.60	0.73
esters	9	0.09	0.11	29	0.28	0.40
amines	12	0.49	0.61	28	0.83	1.53
amides	0			6	0.44	0.51
nitriles	5	0.30	0.33	7	0.60	0.69
nitro compounds	6	0.25	0.38	6	0.58	0.70
compounds with N in heterorings	7	0.53	0.60	25	1.10	1.23
compounds with S	5	0.31	0.53	12	0.40	0.50
compounds with P	12	0.62	0.86	13	0.70	1.01
ions				14	0.53	0.80
total	184	0.36	0.49	401	0.54	0.79

[a] All the errors are in kcal/mol.

for four WSAS models. In model I, 184 molecules for which there are experimental solvation free energies in 1-octanol were studied. The WSAS model achieves an unsigned average error of 0.36 kcal/mol better than those of the SM5.42R universal solvation models (0.56 and 0.58 for SM5.42R/AM1 and SM5.42R/PM3, respectively). In model II, a 245-molecule set with aqueous solvation free energies, used as the training set by Li et al.[1] for their SM5.42R models, were investigated for comparison. Finally, our model and the universal solvation models achieve very similar performances: WSAS has a marginally larger unsigned average error than that of SM5.42R (0.48 vs 0.46 kcal/mol (SM5.42R/AM1 and SM5.42R/PM3)), but a slightly smaller RMS error (0.64 vs 0.66 (SM5.42R/AM1) and 0.68 (SM5.42R/PM3) kcal/mol). To investigate how the probe radius affects the fitting results, we also developed another model with the same molecular set of model II, for which the solvent accessible surface areas were calculated with a probe radius of 1.4 Å, a commonly-used probe radius for water molecule. In this model, both the unsigned average error and the RMS error are significantly larger than those in model II, which are 0.67 and 0.89 kcal/mol, respectively. Compared to the mode II, the solvation free energies of some branched molecules are significantly worse because of the reason stated above.

Models III and IV, the largest molecule sets in aqueous solvation free energy studies that we know, were applied to perform parametrizations. The molecule sets of model III (387) and IV (401) are same except that model IV includes 14 ions. For model III, the molecule set was further divided into a training set (293) and a test set (94). The WSAS model gives unsigned average errors of 0.50 and 0.66 kcal/mol for the training set and test set, respectively. In total, the average error for the 387 molecules is 0.538 kcal/mol, marginally larger than that of model IV (0.536 kcal/mol) without considering the ions. For the 14 ions, we achieved an average error of 0.53 kcal/mol. That means the fitting procedure is reliable and no overfitting has occurred.

For model IV, 47 independent variables entered the fitting procedure, and this number is a somewhat larger than that in the SM5.42 models (10 radii and 28 surface-tension coefficients). It is reasonable because our simple models need more variables to represent the electrostatic solvation energies. Compared to the SM5.0R, in which the electrostatic effect was implicit treated, the two models achieved similar performances and the mean unsigned errors are 0.536 (this work) vs 0.57 (SM5.0R) kcal/mol.

Table 3 lists the unsigned average errors for models I and IV by compound classes. The result of model IV is further summarized by the scatter plot in Figure 1. The correlation coefficient is 0.97 and the slope and intercept are 0.94 and −0.15, respectively. Table 4 lists the parameters for all of the four models. In the following applications of WSAS, we always used the parameters in model IV.

Relative Binding Free Energy Calculations on HIV-1RT/efavirenz. Table 5 lists the results of energy analyses with MM-WSAS and MM-PBSA. Binding mode I has the most favorable binding free energy (with or without intramolecular energy correction) in both of the models, which is about 10 and 7 kcal/mol more favorable than the second best mode, respectively. Because of the neglect of the contribution of the intrinsic

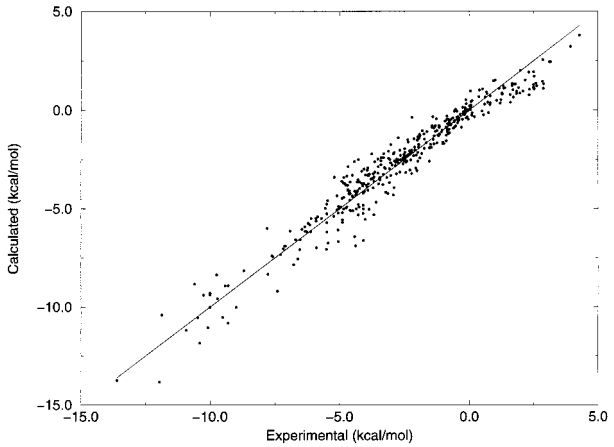

Figure 1. Calculated and experimental solvation free energies for the 401-molecule set (model 4)

TABLE 4: Coefficients of Different Atom Types in Four Solvation Models

no.	atom type	model 1	model 2	mode 3	model 4
1	CT3	0.0403158	0.0066970	0.0194019	0.0117970
2	CT2	0.0267465	−0.0216949	−0.0024938	−0.0183314
3	CT1	−0.1569365	−0.4100890	−0.4218311	−0.4449764
4	CT	0.0852502	0.0648943	0.0758596	0.0697875
5	CA	−0.1234892	−0.0181325	−0.0539652	−0.0480703
6	CD2	−0.0018089	−0.2702024	−0.1448279	−0.1236206
7	CD1	−0.0606230	−0.2417435	−0.1435261	−0.1226437
8	CD	−0.0502707	0.0683193	0.0216036	0.0246029
9	CS	−0.0119038	−0.0099579	−0.0066959	−0.0061025
10	CO	−0.0445866	−0.2228267	−0.1498968	−0.1178312
11	CN	−0.0887039	−0.4286707	0.0425598	0.0208413
12	F	0.0009828	0.0210116	0.0193639	0.0194227
13	Cl	−0.0213188	−0.0026045	−0.0013341	−0.0015111
14	Br	−0.0265121	−0.0083257	−0.0081012	−0.0068335
15	I	−0.0322793	−0.0070827	−0.0076022	−0.0066410
16	HCT3	−0.0679251	−0.0170203	−0.0011903	−0.0061394
17	HCT2	−0.0478811	−0.0492458	−0.0388482	−0.0365374
18	HCT1	−0.0351047	−0.0170883	−0.0198628	−0.0215288
19	HCT	−0.0247406	0.0080586	0.0038341	0.0062124
20	HCD2	0.0469531	0.1949421	0.1715518	0.1459805
21	HCD1	−0.0053710	0.1955127	0.1000105	0.0820232
22	HCD	−0.0073817	0.2082411	0.1211654	0.1036663
23	HCA	0.0523980	−0.0045301	0.0281887	0.0220654
24	HCS	0.0057524	0.0153421	0.0109157	0.0100682
25	HO	0.2705296	0.1066424	−0.0267598	−0.0029297
26	HN	−0.3419044	−0.5065868	−0.2288100	−0.2363502
27	HS	0.9477842	−0.0128937	0.0082210	0.0377756
28	OH	−0.4914185	−0.4078431	−0.2800215	−0.3006240
29	OC	−0.0816853	−0.1324043	−0.101900	−0.0816844
30	O	−0.0872033	−0.1199546	−0.1205119	−0.1357120
31	OS	0.0186137	0.1805694	0.1106798	0.1176048
32	OOH	−0.5085145	−0.2287141	−0.1347761	−0.1421025
33	SH	−0.4070334	−0.0196503	−0.0278438	−0.0372476
34	SS	−0.0351209	−0.0198278	−0.0243691	−0.0121603
35	SP	−0.0801803	−0.0463802	−0.0362589	−0.0477491
36	NS	−0.0727196	−0.1125512	−0.1097706	−0.1143369
37	ND2		−2.5894959	−0.5286290	−0.3270664
38	ND3	−0.0433070	0.1462724	−0.4619591	−0.4072273
39	NO	0.4712564	0.6334536	0.6014819	0.7646976
40	NH2		1.6028609	0.5258473	0.5234891
41	NH1		0.5537324	−0.4282425	−0.3474580
42	N		−1.1272568	−0.6351730	−0.4713683
43	NTH2	0.6526083	1.0130259	0.2320687	0.2720645
44	NTH	0.3987392	0.4780587	−0.0175972	0.0325122
45	NT	−0.1326350	−0.3665904	−0.3779258	−0.3476829
46	P3	1.4048622	1.7795549	1.3700100	1.5160293
47	P4	−0.0307978	−0.0549194	−0.3496134	−0.2509340
48	O-				−1.2154901
49	OA-				−2.1887474
50	S-				−1.1141940
51	HN+				−1.3254553
52	NT+				6.2614711
53	ND+				−5.7053672
54	N+				2.6350040

charges, the absolute solvation free energy with WSAS is smaller than that of PBSA. However, the solvation free energy with the two models has a high correlation (Figure 2) and the correlation coefficient is 0.92, which means the contribution of the intrinsic charges to the solvation free energy are comparable for the five binding modes and this deficiency of WSAS model may be acceptable if one is only interested in the relative solvation free energy. For the solvation free energy, both of WSAS and PBSA have very similar root-mean-square deviations (RMSD), which are about 2.0 kcal/mol; however, for the binding free energy, MM-PBSA has a somewhat smaller RMSD error than MM-WSAS (about 3.0 kcal/mol) due to the compensation of the electrostatic energies of the MM and the PBSA. We believe WSAS can be applied efficiently in evaluating the binding free energies for multiple ligands in a database.

TABLE 5: Summary of Energies from MD Simulations on HIV-1RT/Efavirenz[a]

binding mode	mode I	mode II	mode III	mode IV	mode V
E_{intra}	131.6	133.7	133.4	132.7	133.2
ΔE_{MM}	−54.9	−46.0	−50.7	−46.2	−46.1
ΔG_{WSAS}	18.9	25.6	22.8	20.0	16.5
ΔG_{PBSA}	28.1	33.2	28.8	25.6	23.0
$\Delta\Delta G_{MM-WSAS}$[b]	−36.0	−20.4	−27.9	−26.2	−29.6
$\Delta\Delta G_{MM-PBSA}$	−26.8	−12.8	−21.9	−20.7	−23.4
$-T\Delta S$	13.6	13.7	13.7	13.4	15.2
$\Delta\Delta G_a$[c]	−22.4	−6.7	−14.2	−12.8	−14.4
$\Delta\Delta G_b$	−13.2	0.9	−8.2	−7.3	−7.9
$\Delta\Delta G_a$'	−22.4	−4.6	−12.4	−11.7	−12.8
$\Delta\Delta G_b$'	−13.2	3.0	−6.4	−6.2	−6.3

[a] All the energies are in kcal/mol. [b] $\Delta G_{MM-WSAS} = \Delta E_{MM} - \Delta G_{WSAS}$; $\Delta G_{MM-PBSA} = \Delta E_{MM} - \Delta G_{PBSA}$. [c] $\Delta\Delta G_a = \Delta G_{MM-WSAS} - T\Delta S$; $\Delta\Delta G_b = \Delta G_{MM-PBSA} - T\Delta S$; $\Delta\Delta G_a$' and $\Delta\Delta G_b$' are corrected binding free energies by adding the relative intramolecular energy of the ligand to $\Delta\Delta G_a$ and $\Delta\Delta G_b$, respectively.

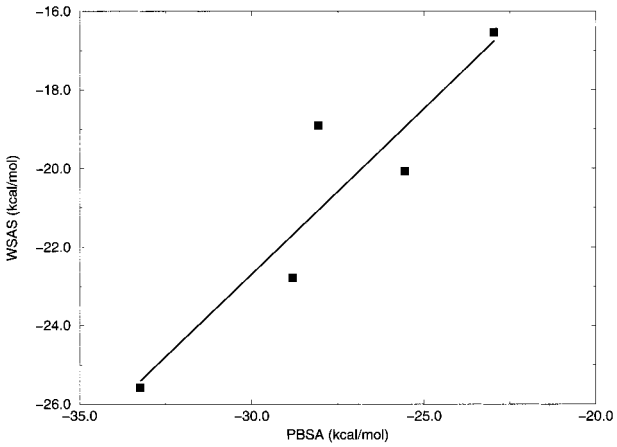

Figure 2. Relationship of the solvation free energies by WSAS and PBSA for the five binding mode of HIV-1RT/ efavirenz. The correlation coefficient, slope, and intercept are 0.92, 0.84, and 2.62, respectively.

We also carried out the solvation free energies of the five binding modes with the Eisenberg−McLachlan's model. In their model, the solvation free energy weighs (in kcal/mol/Å^2) of five basic atom types (C (0.016), N/O (−0.006), O− (−0.024), N+ (0.050), and S(0.021)) were derived with the solvation free energies of 20 amino acids, which were estimated using the distribution coefficients of octanol/water. Because there are no parameters for fluorine and chlorine in their model, we simply borrowed the parameters from model IV, which are 0.0194 kcal/mol/Å^2 for fluorine and −0.0015 kcal/mol/Å^2 for chlorine. Considering that the hydrogen is treated implicitly in Eisenberg−McLachlan's model, we deleted all of the hydrogens and applied the united atom radii in the MSMS program for SAS calculations. The following are the solvation free energies for the five binding modes in sequence: 8.0, 7.3, 7.5, 7.4, and 7.2 kcal/mol. It is clear that the solvation energies with this model cannot be well discriminated and the rank order is not satisfactory either. The reason lies in the fact that their model is too simple, and the parameters fitted with the distribution coefficients in other than the gas/water system may not be proper for this kind of calculations.

Energy Analysis on Villin Headpiece. Table 6 lists the energy analyses results for two MD trajectories of villin headpiece using both MM-WSAS and MM-PBSA. The RMSD of each energy term is indicated in the followed parenthesis. Overall, for the three conformational clusters (the native, the

TABLE 6: Summary of Energies from MD Simulations on Villin Headpiece

ensemble	native (1−100 ns)	folding-intermediate (240−400 ns)	unfolded coil (500−1000 ns)
E_{MM}	−254.5 (35.7)	−224.6 (38.0)	−128.2 (61.4)
ΔG_{WSAS}	−732.9 (24.9)	−761.5 (24.5)	−781.6 (38.9)
ΔG_{PBSA}	−803.2 (31.9)	−806.8 (31.6)	−891.2 (54.5)
TS	455.9 (3.8)	450.1 (2.0)	455.3 (5.6)
$\Delta G_{MM-WSAS}$	−986.6 (33.2)	−986.8 (31.4)	−909.6 (47.1)
$\Delta G_{MM-PBSA}$	−1057.7 (15.9)	−1031.5 (15.7)	−1019.2 (17.2)
ΔG_1[b]	−1442.5	−1436.9	−1364.9
ΔG_2	−1513.6	−1481.6	−1474.5

[a] All the energies are in kcal/mol. [b] $\Delta G_1 = \Delta G_{MM-WSAS} - TS$; $\Delta G_2 = \Delta G_{MM-PBSA} - TS$.

folding intermediate, and the unfolded random coil), the two solvation models give similar trends both for the solvation free energy and the free energy when the entropy contribution was also included. Quantitatively, the WSAS model makes the folding intermediate almost as stable as the native, whereas PBSA finds a large free energy gap between these two structures. Similar to the binding free energy calculation, the free energy with WSAS has a somewhat larger fluctuation, although for the solvation free energy, WSAS is actually more stable than PBSA (the RMSD of WSAS is about 60% of that of PBSA). The reason lies in the fact that the solvent accessible surface area is not as sensitive to the conformations along the MD trajectories as are the screening energies. Usually, the RMS deviation of screening energy is about 1% of the total screening energy in the MM-PBSA analysis, whereas the RMS of SAS is smaller than that of 0.5%. WSAS is nonetheless an attractive solvation model, especially for very large systems with many structures to be evaluated. For example, in protein modeling, both the threading[41−43] and the ab initio[44−45] prediction approaches need to evaluate hundreds, even thousands, of conformations. With WSAS, one can filter some unfavorable conformations (improper folding structures, too many buried-charge conformations) first. Then, one focuses on the more promising conformations using more physically rigorous approaches.

Other Applications of WSAS. Apart from the two applications discussed above, WSAS could be also very useful in other areas. For small- and medium-sized molecules, few atoms are completely buried, so this model can give very accurate solvation free energies. Therefore, the solvation free energies with this model could be applied as the parameters in QSAR studies. In QSAR, Hansch−Leo's solvation free energy models,[7] which intend to reproduce the distribution coefficient of the molecules between 1-octanol/water, are widely used. The WSAS model based on gas/water partitioning, could be an alternative model for such studies.

Finally, this model can be used to calculate the atomic hydrophobic dipolar moments μ[11] described by eq 2, where n is the total number of atoms in a group, r_i is a vector form the origin to the atom i; S_i is the solvent accessible surface area of atom i; w_i is the solvation free energy weight of atom i; and the brackets indicate the mean value of all the atoms in the group. Eisenberg−McLachlan[11] noted that the hydrophobic dipolar moments of neighboring segments of secondary structure tend to oppose each other in correctly folded proteins, but not in incorrectly folded ones, and that the hydrophobic moments can be used to classify helices (on the basis of amino acid sequences) and to detect periodicities. With the hydrophobic moment, one can detect incorrectly folded structures.

In the future work, we will apply this model to calculate the relative binding free energy for a set of protein complexes and incorporate this model into a docking program

$$\mu = \sum_{i=1}^{n} w_i S_i r_i - \langle w_i S_i \rangle \sum_{i=1}^{n} r_i \qquad (2)$$

Conclusions

In this work, we have developed a fast and efficient solvation model based on atomic solvent accessible surface area. For model I, where the training set has 184 molecules with experimental solvation free energies in 1-octanol, WSAS achieves an unsigned average error of 0.36 kcal/mol, better than those of SM5.42R universal solvation model. For model II, both the WSAS and the universal solvation models achieve similar performances in reproducing the experimental aqueous solvation free energies for 245 molecules and the unsigned average error is about 0.46 kcal/mol with WSAS. In model III and IV, we have studied the largest molecule set for aqueous solvation free energy studies that we know. For model III, the whole molecule set has been separated into training set (293 molecules) and test set (94 molecules). Although the training set has a little lower unsigned average error (0.50 kcal/mol) compared to that of test set (0.66 kcal/mol), the overall unsigned average error (0.538 kcal/mol) is very close that of the Model IV (0.536 kcal/mol), in which the parameters were derived using all the molecules.

We have applied WSAS in predicting the relative binding free energies for the five binding modes of HIV-1RT/efavirenz. With this model, the correct binding mode has a binding free energy at least 10 kcal/mol favorable than other binding modes. Moreover, the solvation free energies calculated with WSAS have a high correlation to those calculated by PBSA. Considering WSAS is a much faster solvation model than PBSA, it could be useful to implement WSAS in docking programs to take into account the solvation effect on ligand binding.

We have also applied WSAS in perform the energy analysis for the native and 1-μs MD trajectories of villin headpiece. For the three conformational clusters (the native, the folding intermediate, the unfolded random coil), we achieved the same solvation free energy and free energy trends as PBSA. Although this model has some larger inherent error in the calculation of absolute free energies compared to MM-PBSA, considering the simplicity and the very small computer time cost, WSAS could still be an attracting solvation model, especially for very large systems and with many structures to be evaluated.

Acknowledgment. P.A.K. is grateful to acknowledge research support from the NIH (GM-29072 and GM-56609 (E. Arnold, P. I.)).

References and Notes

(1) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J. *Theor. Chem. Acc.* **1999**, *103*, 9−63.
(2) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144−1149.
(3) Gilson, M. K.; Honig, B. *Proteins* **1998**, *4*, 7−18.
(4) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2200.
(5) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213-217.
(6) Hine, J.; Mookerjee, P. K. *J. Org. Chem.* **1975**, *40*, 292−298.
(7) Hansch, C.; Leo, A. *Exploring QSAR*; American Chemical Society: Washington, DC. **1995**.
(8) Kollman, P. A. *Chem. Rev*, **1993**, *93*, 2395−2417.
(9) Miller, J. L.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 8587−8594.
(10) Meng, E. C.; Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem. B* **1996**, *100*, 2367−2371.
(11) Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199−203.
(12) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, *101*, 7147−7157.

(13) Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *J. Org. Chem.* **1998**, 63, 4305−4313.
(14) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(15) Luo, R.; Moult, J.; Gilson, K. *J. Phys. Chem. B* **1997**, *101*, 11 226−11 236.
(16) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19 824−10 839.
(17) Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 629−666.
(18) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122−129.
(19) Jayaram, B.; Sprous, D.; Beveridge, D. L. *J. Phys. Chem. B* **1998**, *102*, 9571−9576.
(20) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(21) Beveridge, D. L.; Dicapua, F. M. *Annu. Rev. Biophys. Chem.* **1989**, *18*, 431−492.
(22) Dolney, D. M.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2000**, 340−366.
(23) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 799−805.
(24) Barone, V.; Cossi, M.; Tomasi, J. *J. Comput. Chem.* **1998**, *19*, 404−417.
(25) Zou, X.; Sun, Y.; Kuntz, I. D. *J. Am. Chem. Soc.* **1999**, *121*, 8033−8043.
(26) Rapp, C. S.; Friesner, R. A. Proteins: *Struct. Funct., & Gene.* **1999**, *35*, 1−11.
(27) Shen, J.; Quiocho, F. A. *J. Comput. Chem.* **16**, 445−448.
(28) Eriksson, M. A.; Pitera, J.; Kollman, P. A. *J. Med. Chem.* **1999**, *42*, 868−881.
(29) Srinivasan, J.; Miller, J.; Kollman, P. A.; Case, D. A. *J. Biol. Struct. Dyn.* **1998**, 16, 671−682.
(30) Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A. *Proc. Nat. Acad. Soc.* **1999**, *96*, 14 330−14 335.
(31) Massova, I.; Kollman, P. A. *Perspect. Drug Discov. Des.* **2000**, *18*, 113−135.
(32) Massova, I.; Kollman, P. A. *J. Am. Chem. Soc.* **1999**, *121*, 8133−8143.
(33) Reyes, C. M.; Kollman, P. A. *J. Mol. Biol.* **2000**, *295*, 1−6.
(34) Lee, M. R.; Duan, Y.; Kollman, *Proteins: Struct. Funct., & Gene.* **2000**, *39*, 309−316.
(35) *Sybyl Manual*; Tripos Inc.: 1699 South Hanley Rd., St. Louis, Mo 63144-2913, 1998.
(36) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490−641.
(37) Sanner, M. F.; Olson, A. J.; Spehner, J. *Biopolymers* **1996**, *38*, 305−320.
(38) Wang, J.; P.; Kollman, P. *J. Comput. Chem.*, in press.
(39) Wang, J.; Wang, W.; Morin, P.; Kollman, P. *J. Am. Chem. Soc.*, in press.
(40) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740−744.
(41) Marino, B. C.; Martin, M.; M.; Mizuguchi, K.; Siddle, K.; Blundell, T. L. *Biochem. Soc. Trans.* **1999**, *27*, 715−726.
(42) Srinivasan, N.; Antonelli, M.; Jacob, G.; Korn, I.; Romero, F.; Jedlicki, A.; Dhanaraj, V.; Sayed, M. F. R.; Blundell, T. L.; Allende, C. C.; Allende, J. E. *Protein Eng.* **1999**, *12*, 119−127.
(43) Sowdhamini, R.; Burke, D. F.; Deane, C.; Huang, J. F.; Mizuguchi, K.; Nagarajaram, H. A.; Overington, J. P.; Srinivasan, N.; Steward, R. E.; Blundell, T. L. *Acta Crystallogr. Sec. D-Bio. Cryst.* **1998**, *54*, 1168−1177.
(44) Kim, T. S.; Kooperberg, C.; Huang, E.; Baker, D. *J. Mol. Biol.* **1997**, *268*, 209−225.
(45) Yu, X.; Huang, E. S.; Levitt, M.; Samudrala, R. *J. Mol. Biol.* **2000**, *300*, 171−185.
(46) Pearson, R. G. *J. Am. Chem. Soc.* **1986**, *108*, 6109−6114.