A modification of the generalized Born theory for improved estimates of solvation energies and *p*K shifts

B. Jayaram[a)]
Department of Chemistry, Indian Institute of Technology, Hauz Khas, New Delhi-110016, India

Y. Liu and D. L. Beveridge[b)]
Department of Chemistry, Wesleyan University, Middletown, Connecticut 06459

(Received 5 December 1997; accepted 24 March 1998)

We present herein an appraisal on the performance of the generalized Born (GB) model in estimating the solvation energies of small molecules and pK_a shifts of dicarboxylic acids. The quality of the solvation energy results obtained with the GB model was exceedingly good as already reported in the literature but the pK_a shift estimates fell short of expectations. Analysis of the problem on a simple prototype system revealed that with the GB model, the estimates of the two components, viz. the shielding and the self-energy terms, to be somewhat in error. These errors compensate each other in the calculation of solvation energies but affect the intramolecular interaction energies and hence *p*K shifts differently. We examine here the feasibility of introducing modifications to the GB model for a simultaneous evaluation of both solvation and intramolecular interaction energies. © *1998 American Institute of Physics.* [S0021-9606(98)50325-X]

I. INTRODUCTION

The generalized Born (GB) model[1–10] is a treatment of solvation electrostatics which provides accurate estimates of the energetics of solvation for a wide range of small molecules and molecular ions. These estimates as compared with corresponding calculations with the finite difference Poisson–Boltzmann theory, the free energy perturbation theory and experiment are generally within 5%, a quite impressive accomplishment considering the essential simplicity of the model. Recently, we examined the use of the GB model on the problem of pK_a shifts in dicarboxylic acids and found the results to be surprisingly at variance with the predictions of other theoretical methods for solvation. Analysis of the problem on a simple prototype system revealed that with the GB model, the estimates of the two components, the dielectric shielding of charges by the solvent and the self-polarization terms, to be somewhat in error. Furthermore the effects turn out to be mutually compensating in the calculation of solvation energies but affect differently the calculations of the effect of solvent on *p*K shifts. This problem is expected to be critical in the estimates of solvent mediated forces and thus in the use of the GB model in lieu of explicit solvent in molecular simulations. We propose and examine here some modifications to the GB model to remedy this problem.

II. BACKGROUND

The critical role that solvent plays in dictating the stability, flexibility, and interactions of molecules necessitates a good description of solvent effects in theoretical approaches to chemical and biomolecular problems. Explicit all atom representations of the molecular system together with its solvent environment treated at the Born–Oppenheimer level via Monte Carlo and molecular dynamics computer simulations provide one rigorous, well defined path to deal with solvent effects.[11–16] These involve extensive searches in the configuration space of the solute and solvent. The rapidity and ease with which the simulations can be performed continues to improve with the technological advances and increasingly accessible computational power but does the dimensionality of the problems of biological interest. This scenario has fostered the growth of dielectric continuum approaches in parallel. A clear goal of the continuum solvent methods has been to eliminate nonessential degrees of freedom or capture them in some approximate manner, and increase the aperture of investigation in the space and time domains of the biological system without compromising its structural, dynamic, and thermodynamic features. In simpler treatments, the solvent is traditionally assumed to influence the electrostatic interactions and an accurate description of the electrostatics is a prime concern in all theoretical endeavors. The nonelectrostatic terms in the framework of continuum solvent models, which account for the energy expense to form a cavity in the solvent to accommodate solute and the van der Waals interaction energy of the solute with solvent, are included by a surface tension/area or free energy density/volume type terms. Simple Coulomb's law $(q_i q_j / \epsilon r_{ij})$ used to describe the electrostatic interactions between any two charged sites q_i and q_j separated by a distance r_{ij}, in a solvent of dielectric constant ϵ tends to overdamp the interactions. Extension of the theories of Born, Onsager, and Kirkwood for solvation energies (Refs. 17–20, and references therein) requires that the overall molecular charge distribution be symmetric, spherical, elliptical or cylindrical. Computing the solvation energies and solvent mediated intramolecular interactions of molecules with arbitrary shapes is analytically intractable.

[a)]Currently at Wesleyan University on leave from IIT.
[b)]Author to whom correspondence should be addressed.

This triggered the development of a number of sophisticated numerical approaches such as the finite difference Poisson–Boltzmann (FDPB) method (Ref. 21, and references therein) for an accurate treatment of electrostatic interactions and solvent effects, in chemical and biomolecular systems.[21–40] The generalized Born (GB) model,[1–10] a computationally simpler alternative, combines Born expression for ion solvation with Coulomb's law and attempts to extend its applicability to molecular systems with arbitrary shapes.

Theory of the generalized Born model. The total electrostatic energy G_{es} of a molecular system is expressed [Eq. (1), in units of kcal/mol] as a sum of the Coulomb interaction energy between each pair of charges in a solvent of dielectric constant ϵ, and Born's solvation (self) energy of each individual charge. $\alpha_i s$ are the Born radii. Each charge is assumed to be embedded in a low dielectric cavity (wherein the dielectric constant is set at unity) in a dielectric continuum solvent. This is rewritten as a sum of the Coulomb interaction energy in vacuum and polarization energy which quantifies the influence of solvent on the molecule. The GB model further proceeds to capture all the effects due to solvent G_{pol}, in one single term [Eq. (3)] with a clever choice of the effective distance parameter f_{GB}. The polarization energy in Eq. (3) is in fact a sum of solvent shielding terms and self-energy terms as shown in Eq. (5),

$$G_{es}=332\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{q_iq_j}{r_{ij}\epsilon}-166\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{n}\frac{q_i^2}{\alpha_i} \tag{1}$$

$$=332\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{q_iq_j}{r_{ij}}+G_{pol}, \tag{2}$$

$$G_{pol}=-166\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{n}\sum_{j=1}^{nb}\frac{q_iq_j}{f_{GB}}, \tag{3}$$

$$f_{GB}=(r_{ij}^2+\alpha_{ij}^2e^{-D})^{0.5};\quad \alpha_{ij}=(\alpha_i\alpha_j)^{0.5};$$

$$D=r_{ij}^2/(4\alpha_{ij}^2), \tag{4}$$

$$G_{pol}=\underbrace{-166\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{q_iq_j}{f_{GB}}}_{\text{shielding}}\underbrace{-166\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{n}\frac{q_i^2}{\alpha_i}}_{\text{self}}. \tag{5}$$

This set of equations for G_{pol} after a suitable calibration of the Born radii, namely $\alpha_i s$ in the above equations, enables an estimation of the solvation energy of a given molecule in the specified conformation, in a quantitative manner at very little computational expense. Equation (1) suffers from some logical inconsistency. The usage of the Coulomb term with ϵ (the first term on the right-hand side) implies the absence of dielectric boundaries in the system and is valid for point charges in a continuum solvent. The Born term however, (the second term which contains $(\{(1/\epsilon_{int})-(1/\epsilon_{ext})\}$, with $\epsilon_{int}=1$), vanishes if there is no dielectric boundary between the charges and the solvent. Equation (2) masks this inconsistency. This notwithstanding, the agreement between the GB results and the experimental solvation energies for a wide range of molecules is exceedingly good, by virtue of the choice of the functional form for f_{GB} and a calibration of the Born radii, the $\alpha_i s$.[1,2] Thus the GB model provides an attractive and expeditious approach to calculate the total electrostatic energy and solvation energy of a molecule, and promises to bring biomolecular problems from the domain of supercomputers/high-end workstations to PCs. An important question to be resolved before this goal is realized however is, whether an accurate treatment of solvation energy at the continuum solvent level automatically implies an accurate treatment of solvent effects which include solvent mediated interactions between any two charged sites in a molecule. To fully appreciate this issue we set up the following test case and some illustrative calculations on solvation energies of small molecules and pK_a shifts in dicarboxylic acids.

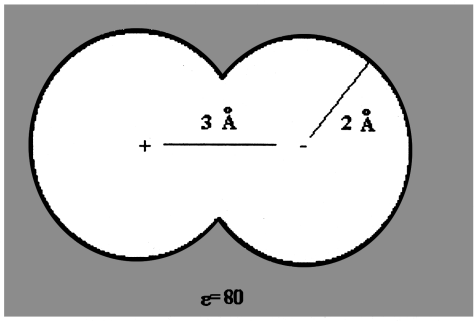

FIG. 1. A definition of the parameters for the test case (see text) of two interacting charges embedded in a continuum solvent.

III. CALCULATIONS AND RESULTS

A. Test case

Consider two overlapping spheres each of radius 2 Å, whose centers are located 3 Å apart and both are embedded in a solvent (Fig. 1) of dielectric constant 80. One sphere carries a unit positive charge and the other a unit negative charge placed at their respective centers. The electrostatics of this system is examined via three different continuum solvent methods viz. the FDPB method as detailed in Ref. 35, the Kirkwood theory adopting the procedure illustrated in Refs. 39 and 41–44 and the GB model described above. The radius of the Kirkwood sphere enclosing these two charged spheres was calibrated to yield the same solvation energy as with the FDPB method. For the GB model calculations, the Born radii are derived by using the dielectric descreening approach developed by Hawkins *et al.*,[2] so as to give the exact solvation energy as with the FDPB approach to facilitate a comparative analysis. The results are collected in Table I.

The solvation energy with the FDPB method (column 2, Table I) is calculated by taking the difference between the total energy of the system obtained with $\epsilon_{int}=2$; $\epsilon_{ext}=80$ set up and that with $\epsilon_{int}=2$; $\epsilon_{ext}=1$. The FDPB calculations with PARSE parameters give the best results[35] for solvation energies with this procedure. Note that the Coulomb contribution is computed with $\epsilon=1$ although the solute dielectric ϵ_{int} is set at 2. Usage of the solute dielectric $\epsilon=2$ for Cou-

TABLE I. Calculated electrostatic energies (in kJ/mol) for the two-charge test case partitioned into individual components.

Energy (1)	FDPB (2)[a]	GB (3)[b]	Kirkwood (4)[c]	GB (5)[d]	Kirkwood (6)[e]	mGB (7)[f]
Total[g]	−683.8	−683.8	−683.8	−452.7	−452.7	−683.8
Coulombic	−462.7	−462.7	−462.7	−231.6	−231.6	−462.7
Solvation	−221.1	−221.1	−221.1	−221.1	−221.1	−221.1
Self	−647.9	−614.0	−576.4	−435.6	−415.5	−647.9
Shielding	426.8	392.9	355.3	214.4	194.4	426.8
Interaction	−35.9	−69.8	−107.1	−17.1	−37.2	−35.9

[a]van der Waals radius=2 Å; $\epsilon_{int}=2$; $\epsilon_{ext}=80$; probe radius 1.4 Å; resolution=4 grids/Å.
[b]Born radius=2.2319 Å; $\epsilon_{int}=1$.
[c]Radius of Kirkwood sphere=3.079 Å; $\epsilon_{int}=1$.
[d]Born radius=1.5528 Å; $\epsilon_{int}=2$.
[e]Radius of Kirkwood sphere=2.491 Å; $\epsilon_{int}=2$.
[f]Born radius=2.115 Å; $\epsilon_{int}=1$; $c=1.64$ in Eq. (6).
[g]Total=Coulombic+self+shielding=Coulombic+solvation=self+interaction.

lomb contribution as well leads to unphysical results. (The calculated interaction energy between a positive and negative charge is around +46.8 kcal. The binding energy can be positive for such a system but not the interaction energy.) No matter what the solute dielectric is, if the process involves the transfer of solute from vacuum to solvent, the Coulomb contribution is evaluated with $\epsilon=1$ in the FDPB method. In the GB and Kirkwood models however, the internal dielectric of the solute has to be used to estimate the Coulomb contribution as well. These results are shown in columns 3 and 4. Each method comes with its own prescription and the methodologies adopted above for obtaining the solvation energies conform to current practices. The total electrostatic energy (row 2), the Coulomb energy (computed with a dielectric constant of unity) (row 3) and the solvation energy (row 4) are nearly identical in all these three cases. We now examine whether the interaction energy between the two charges in all these cases is the same as well. To do this, one needs to compute the potential at one charge site (j) due to the other (i) taking account of the solvent effect and multiply this potential by the charge (q_j). Alternatively, the solvation energy may be partitioned into self-energy and shielding energy contributions and the latter added to the Coulomb energy to obtain the net interaction energy. These results are presented in rows 5–7 of Table I. As is noticeable from Table I, the algebraic sum of the self and shielding terms is the same in all the three cases, but the individual values are quite different. Because of these differences, the interaction energies (and hence forces) predicted by these three methods are different. To allay any apprehensions that these differences might be artefactual and originate in the choice of the solute dielectric, the GB and the Kirkwood model results with $\epsilon_{int}=2$ are also shown in Table I in columns 5 and 6. All the methodologies considered above give nearly the same solvation energies but no two methods agree on the interaction or total energies and only one of the estimates can be correct. The test case here merely illustrates the point that *a good solvation energy description or an accurate estimation of the total energy does not guarantee that the solvent mediated (intrasolute) interactions between charged sites are treated correctly*.

Results for the GB model in Table I (column 3) suggest that the shielding energies are underestimated. Shielding energies when combined with the self-energies give the solvation energies. Since the individual atomic self-energies can be adjusted by varying the Born radii, a good description of the solvation energy is always feasible, whatever the shielding energy. Shielding energy when combined with the vacuum Coulomb energy gives the interaction energy and this Coulomb contribution is fixed for a given charge distribution and geometry and admits no leverage. Hence a viable option in the development of empirical electrostatic models appears to be to fix the shielding energies first based on intramolecular interactions (as given by pK shifts for instance) and then address the self-energies by calibrating the radii to yield accurate solvation energies. We illustrate this strategy within the framework of the GB model in view of its advantages and simplicity. A general option is to derive the electrostatic contributions from experiment and employ them as a target. For the test case here, we require that the GB model reproduce results which are identical to the FDPB results in column 2 since the FDPB method is known to work well on both solvation energies and pK shifts. One possibility which would accomplish this task involves a modification of the f_{GB} term in Eq. (4) as follows:

$$f_{GB}=(r_{ij}^2+\alpha_{ij}^2 e^{-D})^{0.5};\quad D=r_{ij}^2/c\,\alpha_{ij}^2;\quad c\neq 4. \tag{6}$$

A value less than 4 for c improves the shielding energies but the self-energies have to be increased too commensurately for maintaining the solvation energy, which is easily done by decreasing the Born radius. Thus for a choice of $c=1.64$ and a slightly reduced Born radius, the FDPB results in column 2 are reproduced exactly by the GB model as shown in column 7 of Table I. The test case here is indicative of what is to be expected of the GB model on molecular systems and the direction to take for a reliable prediction of the solvation and intramolecular interaction energies simultaneously.

B. Solvation energies and *p*K shifts

We adopted a set of 25 molecules for testing the solvation energies for which experimental solvation energies as

well as the electrostatic contributions computed by the FDPB method (Table III of Ref. 35) were available.

The geometries of these molecules were generated using Insight II (Ref. 45) and minimum energy structures under vacuum conditions were obtained with the DISCOVER module using the CFF-91 force field. Ideally, one would have preferred to use solution conformation of each of these molecules which is beyond the minimization protocols with a continuum solvent approach. Arriving at a good continuum force field which can lead to the correct equilibrium structures is in fact one of the goals of the GB and the FDPB models. A suitable procedure would involve determining the average equilibrium structure(s) with molecular dynamics or Monte Carlo simulations using explicit solvent and employing these structures in turn for assessing/calibrating continuum solvent methods. This approach takes away the simplicity of the methodology. While bringing in rigor into structural assumptions, it introduces further force field/simulation protocol dependent variables into the problem. Since the goal of this study is a comparison of the electrostatic contribution and its components as estimated by the GB model with that of the FDPB method on a given structure, there appears to be no risk of losing the generality of the conclusions drawn.

The PARSE charges and radii[35] have been shown to describe the solvation energies of these molecules quite well. This set reduces all the atoms in diverse molecular contexts into just six atom types (two for carbon, one each for oxygen, nitrogen, hydrogen, and sulfur) for the purposes of defining van der Waals radii. In contrast to the atomic descriptions in other force fields, this constitutes an extremely simplified approach to arrive at the energetics but suffices for the discussion here. Both the FDPB and GB models admit in principle, introduction of more atom types for better accuracies. The PARSE parameters[35] were assigned to each of these molecules and the electrostatic contribution to the solvation energies were recomputed with the FDPB method (column 2 of Table II) following a protocol identical to that described in Ref. 35. Solvation energies were also calculated with the GB model using the same set of charges and radii as in the FDPB calculations, with $c=4$ in Eq. (6) as proposed by Still and co-workers[1] together with the pairwise dielectric descreening approach of Cramer, Truhlar, and co-workers.[2–5]

The dielectric descreening approach is a method to calculate the effective Born radius α_i for each of the atoms in a specified molecular context. These $\alpha_i s$ are then employed in the solvation energy calculations [in Eqs. (3) and (4)]. The input parameters are the Cartesian coordinates of each of the atoms in the molecule, their van der Waals radii, and screening parameters for each atom type. In principle, the screening parameters can be as many as the number of atom types. Since the PARSE set adopted here has only six distinct van der Waals radii, we have decided to use just five screening parameters, one for each of the elements, namely, carbon, oxygen, nitrogen, hydrogen, and sulfur, in the set of molecules considered. As already noted above, a larger set of atom types and screening parameters is likely to ensure better accuracies.[2–5] A good set of screening parameters is not available *a priori* for the geometries, radii, and charges employed here. We extracted optimal screening parameters by a simulated annealing procedure via a minimization of the root mean square deviation between the FDPB solvation energies and those predicted by the GB model. The calculated solvation energies with these optimized screening parameters are shown in column 3 of Table II. We recover the result that the GB model works extremely well in predicting the electrostatic contribution to the solvation energies.

TABLE II. Calculated electrostatic contributions (in kJ/mol) to solvation energies.

System	FDPB	GB	m_1GB	m_2GB
1. Methanol	−30.76	−34.90	−31.02	−32.94
2. Ethanol	−29.76	−31.98	−27.13	−29.89
3. Ammonia	−24.62	−29.01	−23.70	−25.29
4. Methylamine	−28.76	−31.77	−28.30	−28.76
5. Ethylamine	−27.21	−25.83	−22.36	−23.20
6. Methylthiol	−13.42	−13.38	−13.41	−13.42
7. Acetone	−25.62	−32.19	−25.54	−28.17
8. 2-butanone	−24.29	−28.76	−20.82	−24.95
9. Acetaldehyde	−21.40	−25.21	−21.65	−21.70
10. Propionaldehyde	−20.57	−23.66	−19.31	−20.27
11. Acetic acid	−37.12	−41.63	−34.69	−38.16
12. Propionic acid	−36.83	−39.79	−33.27	−36.70
13. Acetamide	−50.91	−53.59	−39.54	−52.17
14. Propionamide	−48.99	−50.87	−35.90	−49.41
15. Benzene	−13.67	−13.25	−13.92	−13.45
16. Toluene	−13.33	−13.33	−13.38	−13.75
17. Pyridine	−30.35	−23.49	−30.35	−19.27
18. Phenol	−36.66	−36.07	−32.52	−34.86
19. Methyl imidazole	−48.82	−62.57	−73.90	−58.69
20. Ammonium	−410.31	−402.95	−424.02	−409.77
21. *N*-butyl ammonium	−332.64	−332.44	−344.26	−343.05
22. Acetate ion	−345.94	−345.81	−351.04	−345.94
23. Propionate ion	−341.63	−340.75	−341.59	−340.13
24. Methyl imidazolium	−278.85	−278.76	−278.85	−278.89
25. N-p-guanidinium	−284.28	−285.16	−270.11	−284.198
Mean error		1.63	2.55	0.42
Mean unsigned error (Relative to FDPB)		3.05	4.77	1.96

We then proceeded to evaluate the pK_a shifts (Table III) in dicarboxylic acids with both the FDPB and GB models using the same parameters as above,

$$\underset{\text{I}}{\overset{1}{\text{HOOC}}-\text{R}-\overset{2}{\text{COOH}}} \Leftrightarrow \underset{\text{II}}{\overset{1}{^{-}\text{OOC}}-\text{R}-\overset{2}{\text{COOH}}} \Leftrightarrow \underset{\text{III}}{\overset{1}{^{-}\text{OOC}}-\text{R}-\overset{2}{\text{COO}^{-}}}.$$

Interaction energies between the functional groups 1 and 2 in each of the above three species I, II, and III, denoted here as ΔE_{I}, ΔE_{II}, and ΔE_{III}, respectively, were computed and converted to pK_a shifts as follows:

$$\Delta pK_{a1} = (\Delta E_{\text{I}} - \Delta E_{\text{II}})/(2.303\ RT)$$

and

$$\Delta pK_{a2} = (\Delta E_{\text{III}} - \Delta E_{\text{II}})/(2.303\ RT).$$

To judge the success of the pK_a shift estimates with various theoretical models, we required that the results be better than the predictions based on Coulomb's law. As also noted pre-

TABLE III. Calculated $p\mathrm{K}_a$ shifts for dicarboxylic acids.

System	Experiment	Coulomb's Law	FDPB	GB	m_1GB	m_2GB
$\Delta p\mathrm{K}_{a2}$						
1. Oxalic acid	2.36	0.93	2.11	5.02	2.50	3.84
2. Malonic acid	2.26	0.82	1.78	3.27	1.79	2.41
3. Succinic acid	0.84	0.57	0.69	0.87	0.61	0.81
4. Glutaric acid	0.47	0.47	0.49	0.54	0.47	0.59
5. Adipic acid	0.38	0.39	0.34	0.40	0.40	0.46
6. Pimelic acid	0.34	0.34	0.31	0.35	0.34	0.38
7. Suberic acid	0.28	0.30	0.27	0.30	0.30	0.32
8. Azelaic acid	0.26	0.27	0.25	0.27	0.27	0.28
9. Methyl malonic acid	1.89	0.85	2.64	4.04	2.38	2.83
$\Delta p\mathrm{K}_{a1}$						
1. Oxalic acid	3.22	0.15	3.98	9.14	7.92	9.84
2. Malonic acid	1.75	0.07	2.74	2.76	1.98	2.36
3. Succinic acid	0.36	0.02	0.20	0.43	0.15	0.21
4. Glutaric acid	0.18	0.01	0.08	0.12	0.02	0.04
Mean error		0.72	0.10	0.99	0.35	0.75
Mean unsigned error		0.72	0.29	1.00	0.51	0.80
Mean error[a]		0.49	0.04	0.54	0.01	0.24
Mean unsigned error[a]		0.49	0.23	0.55	0.15	0.29

[a]Without the $\Delta p\mathrm{K}_{a1}$ of oxalic acid.

viously with formal charges and OPLS radii,[36,40] the FDPB method with PARSE partial atomic charges and radii meets the above success criterion reasonably well. The GB model however falls short of expectations. This is to be anticipated based on the test case discussed above. An inspection of the computed $p\mathrm{K}_a$ shifts reveals that the shielding energies are underestimated, particularly when the charge distributions are in proximity as in oxalic and malonic acid cases. It is also likely that the transferability assumed in assigning the partial charges to the dicarboxylic acids breaks down particularly for oxalic acid.

The performance of the GB model on both solvation energies and $p\mathrm{K}_a$ shifts on the systems considered here with PARSE parameters, is depicted in Fig. 2 in terms of the mean unsigned error computed as a function of the coefficient c in Eq. (6). Clearly a choice of $c=3.2$ is optimal for solvation energies and $c=1.7$ works best for $p\mathrm{K}_a$ shifts.

This dual behavior seen with the effective distance parameter f_{GB} as given by Eqs. (4) and (6) is not unique. During the course of this investigation we discovered a number of effective distance parameters which work well either on the solvation energies or on the $p\mathrm{K}_a$ shifts but not on both simultaneously. For instance, an entirely different functional form such as

$$f(r)=\left\{r_{ij}+\left[(r_{ij}+\alpha_{ij})\Big/\left(1+\frac{2dr_{ij}}{\alpha_{ij}}\right)\right]\right\} \tag{7}$$

with a choice of $d=0.63$, leads to reasonable values for solvation energies of all the molecules in Table II, with a mean unsigned error of 1.1 kcal without further parameterization but the $p\mathrm{K}_a$ shifts are off. For a choice of $d=2.0$, the $p\mathrm{K}_a$ shift predictions for all the systems in Table III, are exceedingly good with a mean unsigned error of 0.25 units, but the solvation energies are not at all satisfactory.

With regard to the function f_{GB} in Eq. (6), the modification proposed under the test case was attempted by reducing the value of c from 4 to 2 and by scaling down all the PARSE radii uniformly by a factor of 0.85. The corresponding solvation energies are given in the last column (m_1GB) of Table II. These can be further improved by fine tuning the size parameters (radii) instead of a flat reduction across the board as carried out here, to attain better accuracies. The $p\mathrm{K}_a$ shifts (last column in Table III) show a definite improvement. The optimized screening parameters are given in Table IV.

An alternative to the reduction of the radii proposed above is to modify the effective distance function itself. At

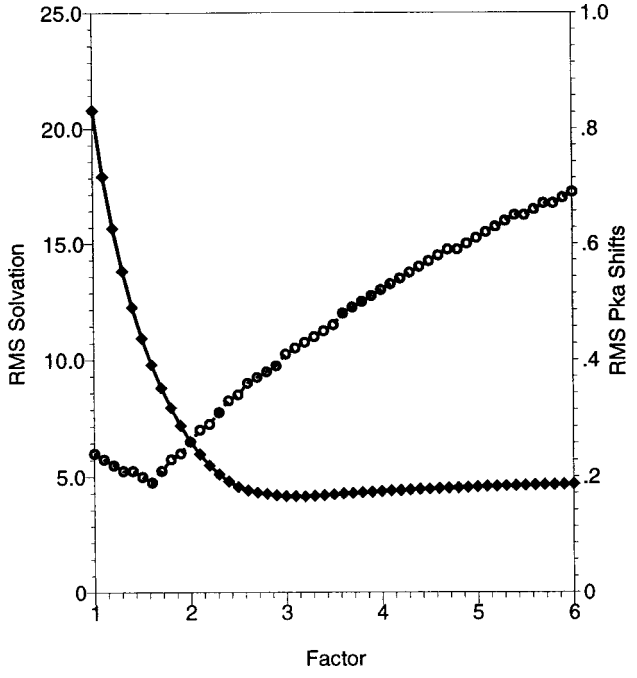

FIG. 2. Mean unsigned error in solvation energies (solid line) and $p\mathrm{K}_a$ shifts (dashed line) shown as a function of the factor c in Eq. (6).

TABLE IV. Screening parameters.

	GB	m_1GB	m_2GB
H	0.7649	0.9470	0.7672
C	0.7526	0.9019	0.7021
O	0.7506	1.0336	0.7731
N	0.8494	0.9137	0.8522
S	1.6367	1.6541	1.0202

the risk of looking cumbersome, the following expression:

$$f_{m2\mathrm{GB}}=f_{m\mathrm{GB}}\{(\epsilon\gamma-\gamma)/(\epsilon\gamma-1)\},$$

where $f_{m\mathrm{GB}}=f_{\mathrm{GB}}$ with $c=2$ in [Eq. (4)], employed for calculating shielding interactions, yields results which show a modest improvement over the GB model (last columns in Tables II and III). The predicted solvation energies with this function come closest to the FDPB values with the data set considered here. γ in the above equation is a sigmoidal function given as

$$\gamma=[1-((\epsilon-4)/2)(\beta^2+2\beta+2)e^{-\beta}];$$

$$\beta=(0.4rij+\alpha_{ij}).$$

This type of expression for γ has been utilized successfully elsewhere[46–52] in computing intra and intermolecular interactions in biomolecules.

The results presented above with the test case and on some molecular systems bring out the merits and limitations of the generalized Born set of equations and point to a direction for improvement for a simultaneous and accurate evaluation of both solvation and interaction energies. The seemingly arbitrary effective distance parameter in the generalized Born theory has its origins in the solutions to Poisson equation as described in the Appendix. Obviously, formulating an effective distance function that can at once account for solvent effects in evaluating solvation, intra and intermolecular energies in diverse molecular contexts though not impossible is not an easy task. It is hoped that this note will stimulate further research into effective distance parameters better than those given in Eqs. (6)–(8) for a simultaneous evaluation of all related thermodynamic quantities via a suitable conformational averaging.

IV. SUMMARY

The performance of the generalized Born model was examined on both solvation and interaction energies with a prototypical system and with some small molecules. The solvation energy estimates, in general, are in good accord with other theoretical predictions but the calculated pK_a shifts fall short of expectations. Since the shielding energy when combined with the self-energy gives the solvation energy but when added to the Coulomb energy gives the interaction energy, *solvation and interaction cannot be dealt with separately* in developing an empirical model for treating electrostatic interactions. Some modifications are proposed and examined which involve altering both the shielding and self-energies for improved estimates of solvation and interaction energies. Overall, the GB model and its modification proposed here appear to give a satisfactory account of the energetics of solvation of small molecules. The dynamical behavior of these models however, is yet to be ascertained.

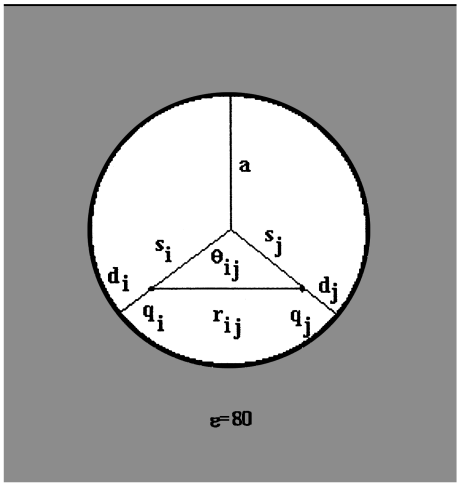

FIG. 3. A definition of the parameters in the Kirkwood theory for two interacting charges embedded in a continuum solvent.

ACKNOWLEDGMENTS

Funding from the NIH Grant No. GM 37909 from the National Institutes of General Medical Sciences (to D.B.) and from the Council of Scientific and Industrial Research, India (to B.J.), is gratefully acknowledged. B.J. wishes to express his gratitude to Professor Beveridge for providing financial support and facilities to carry out research at Wesleyan.

APPENDIX

The relation between the generalized Born theory, Eqs. (3) and (4) in particular, and the solutions to Poisson equation for an arbitrary charged distribution with an overall spherical symmetry embedded in a continuum solvent (Kirkwood theory) is examined and the nature of the functional form of the effective distance parameter f_{GB} [Eq. (4)] is discussed here.

The Kirkwood expressions [Fig. 3 and Eqs. (3)–(9) and (45)–(50) of Ref. 39] for the solvent polarization energy can be expressed in the form of Eq. (3) as

$$G_{\mathrm{pol}}=-166\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{q_iq_j}{f_{\mathrm{KW}}},$$

$$f_{\mathrm{KW}}=a\left[1-\frac{2s_is_j\cos\theta_{ij}}{a^2}+\left(\frac{s_is_j}{a^2}\right)^2\right]^{(1/2)}.$$

In arriving at the above equation the following simplifications have been used. Internal dielectric of the solute $\epsilon_{\mathrm{int}}=1$ and solvent dielectric $\epsilon\gg1$. Noting that

$$r_{ij}^2=s_i^2+s_j^2-2s_is_j\cos\theta_{ij}\quad\text{and}\quad a=s_i+d_i=s_j+d_j$$

and eliminating $\cos\theta_{ij}$ and a from the above equation for f_{KW} gives

$$f_{\mathrm{KW}}=\left[r_{ij}^2+\frac{(d_i^2+2s_id_i)}{(s_i+d_i)}\cdot\frac{(d_j^2+2s_jd_j)}{(s_j+d_j)}\right]^{(1/2)}.$$

d_i and d_j in the above equations are distances from the locations of the charges i and j to the exterior of the molecule, respectively. Similarly s_i and s_j are distances from the locations of the charges i and j to the center of the molecule, respectively. If the atoms i and j are exposed to the solvent d_i and d_j are identifiable with α_i and α_j, the respective Born radii. If the atoms i and j form a contact pair in solvent, the s_i and s_j are similarly identifiable with α_i and α_j. In a general case, the d and s terms may be written as

$$d_i=\xi_{1i}\alpha_i;\quad s_i=\xi_{2i}\,\alpha_i;\quad \text{and}\quad d_j=\xi_{1j}\alpha_j;$$

$$\text{and}\quad s_j=\xi_{2j}\alpha_j.$$

The ξ coefficients are functions of the distance between the charges i and j as well as their location in the molecule. The effective distance function with these conversions can be written as

$$f_{\mathrm{KW}}=\left[r_{ij}^2+\alpha_i\alpha_j\,\frac{(\xi_{1i}^2+2\xi_{1i}\xi_{2i})}{(\xi_{1i}+\xi_{2i})}\,\frac{(\xi_{1j}^2+2\xi_{1j}\xi_{2j})}{(\xi_{1j}+\xi_{2j})}\right]^{(1/2)}$$

or

$$f_{\mathrm{KW}}=[r_{ij}^2+\alpha_i\alpha_j\xi(r_{ij},d_i,d_j,s_i,s_j)]^{0.5}.$$

For a single ion i, $r_{ij}=0$; $\xi_{2i}=\xi_{2j}=0$; and $\xi_{1i}=\xi_{1j}=1$ giving $f_{\mathrm{KW}}=\alpha_i$ which leads to the Born expression. For two point charges i and j, $\alpha_i=\alpha_j=0$ which converts f_{KW} to r_{ij} giving the Coulomb's law. The generalized Born expression for the effective distance function

$$f_{\mathrm{GB}}=(r_{ij}^2+\alpha_i\alpha_je^{-D})^{0.5};\quad D=r_{ij}^2/(4\alpha_{ij}^2)$$

in fact looks very similar. The function ξ above of the Kirkwood theory is replaced by an exponential in the generalized Born theory. The exponential is designed to always be less than unity to give reliable results on solvation energies of small molecules. The ξ however, is less than unity only for distances less than half the sum of the Born radii. Thus in relation to the Kirkwood theory, the effective distances with the generalized Born theory are smaller leading to relatively stronger shielding contributions as is noticeable in columns 3 and 4 of Table I.

It is of interest to note that the solutions to Poisson equation for two charges separated by the solvent (the intermolecular problem),[20] when cast in the form of the generalized Born effective distance function (f_{inter}), take a slightly different form,

$$f_{\mathrm{inter}}=r_{ij}\left(\frac{\epsilon-1}{\epsilon-\delta_{ij}}\right),$$

$$\delta_{ij}=\left[\frac{\frac{1}{2}\left\{\left(\frac{r_{ij}}{\alpha_i}\right)^2\left(\frac{r_{ij}}{\alpha_j}-1\right)^2+\left(\frac{r_{ij}}{\alpha_j}\right)^2\left(\frac{r_{ij}}{\alpha_i}-1\right)^2\right\}}{\left\{\left(\frac{r_{ij}}{\alpha_j}-1\right)^2\left(\frac{r_{ij}}{\alpha_i}-1\right)^2-1\right\}}\right].$$

Thus arriving at a single effective distance function fit for all occasions is a daunting task.

[1] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson, J. Am. Chem. Soc. **112**, 6127 (1990).
[2] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, Chem. Phys. Lett. **246**, 122 (1995).
[3] D. Liotard, G. D. Hawkins, G. C. Lynch, C. J. Cramer, and D. G. Truhlar, J. Comput. Chem. **16**, 422 (1995).
[4] C. C. Chambers, G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. **100**, 16385 (1996).
[5] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. **100**, 19824 (1996).
[6] R. Constanciel and R. Contreras, Theor. Chim. Acta **65**, 1 (1984).
[7] R. Conteras and A. Aizman, Int. J. Quantum Chem. **27**, 293 (1985).
[8] T. Kozaki, K. Morihashi, and O. Kikuchi, J. Mol. Struct.: THEOCHEM **168**, 265 (1988).
[9] T. Kozaki, K. Morihashi, and O. Kikuchi, J. Am. Chem. Soc. **111**, 1547 (1989).
[10] P. J. Goodford, J. Med. Chem. **28**, 849 (1985).
[11] D. L. Beveridge and F. M. DiCapua, Annu. Rev. Biophys. Biophys. Chem. **18**, 431 (1989).
[12] W. L. Jorgensen, Acc. Chem. Res. **22**, 184 (1989).
[13] W. F. van Gunsteren and H. J. C. Berendsen, Acta Crystallogr., Sect. B: Struct. Sci. **29**, 992 (1990).
[14] J. A. McCammon, Curr. Opin. Struct. Biol. **1**, 196 (1991).
[15] P. Kollman, Chem. Rev. **93**, 2395 (1993).
[16] W. F. van Gunsteren, F. J. Luque, D. Timms, and A. E. Torda, Annu. Rev. Biophys. Biomol. Struct. **23**, 847 (1994).
[17] D. L. Beveridge and G. W. Schnuelle, J. Phys. Chem. **23**, 2562 (1975).
[18] B. Jayaram and D. L. Beveridge, Biopolymers **27**, 617 (1988).
[19] B. Jayaram and D. L. Beveridge, J. Phys. Chem. **94**, 4666 (1990).
[20] B. Jayaram, J. Phys. Chem. **98**, 5773 (1994).
[21] B. Honig and A. Nicholls, Science **268**, 1144 (1995).
[22] J. B. Matthew, Annu. Rev. Biophys. Biophys. Chem. **14**, 387 (1985).
[23] N. K. Rogers, Prog. Biophys. Mol. Biol. **48**, 37 (1986).
[24] S. Harvey, Proteins: Struct., Funct., Genet. **5**, 78 (1989).
[25] M. E. Davis and J. A. McCammon, Chem. Rev. **90**, 509 (1990).
[26] K. Sharp and B. Honig, Annu. Rev. Biophys. Biophys. Chem. **19**, 301 (1990).
[27] A. Warshell and J. Aqvist, Annu. Rev. Biophys. Biophys. Chem. **20**, 267 (1991).
[28] N. M. Allewell and H. Oberoi, Methods Enzymol. **202**, 3 (1991).
[29] C. J. Cramer and D. G. Truhlar, Science **256**, 213 (1992).
[30] D. Bashford, Curr. Opin. Struct. Biol. **2**, 40 (1992).
[31] J. Tomasi and M. Persico, Chem. Rev. **94**, 2027 (1994).
[32] M. Holst, R. E. Kozack, F. Saied, and S. Subramanian, Proteins: Struct., Funct., Genet. **18**, 231 (1994).
[33] T. Simonson and A. T. Brunger, J. Phys. Chem. **98**, 4683 (1994).
[34] A. B. Schmidt and R. Fine, Mol. Simul. **13**, 347 (1994).
[35] D. Sitkoff, K. A. Sharp, and B. Honig, J. Phys. Chem. **98**, 1978 (1994).
[36] E. Rajasekaran, B. Jayaram, and B. Honig, J. Am. Chem. Soc. **116**, 8238 (1994).
[37] M. K. Gilson, Curr. Opin. Struct. Biol. **5**, 216 (1995).
[38] G. N.-Szabo and G. G. Ferenczy, Chem. Rev. **95**, 829 (1995).
[39] H. Nakamura, Q. Rev. Biophys. **29**, 1 (1996).
[40] S. B. Dixit, R. Bhasin, E. Rajasekaran, and B. Jayaram, J. Chem. Soc., Faraday Trans. **93**, 1105 (1997).
[41] H. L. Friedman, Mol. Phys. **20**, 1533 (1975).
[42] J. G. Kirkwood and F. H. Westheimer, J. Chem. Phys. **6**, 506 (1938).
[43] F. H. Westheimer and M. W. Shookhoff, J. Am. Chem. Soc. **61**, 55 (1938).
[44] C. Tanford, J. Am. Chem. Soc. **79**, 5348 (1957).
[45] Insight II, Biosym Technologies (MSI), San Diego.
[46] N. Arora and B. Jayaram, J. Comput. Chem. **18**, 1245 (1997).
[47] B. Jayaram, A. Das, and N. Aneja, J. Mol. Struct.: THEOCHEM **361**, 249 (1996).
[48] J. Ramstein and R. Lavery, Proc. Natl. Acad. Sci. USA **85**, 7231 (1988).
[49] B. E. Hingerty, R. H. Richie, T. L. Ferrel, and J. E. Turner, Biopolymers **24**, 427 (1985).
[50] B. Jayaram, S. Swaminathan, D. L. Beveridge, K. Sharp, and B. Honig, Macromolecules **23**, 3156 (1990).
[51] M. O. Fenley, G. S. Manning, and W. K. Olson, Biopolymers **30**, 1191 (1990).
[52] H. A. Gabb, R. Lavery, and C. Prevost, J. Comput. Chem. **16**, 667 (1995).