Theor Chem Acc (1999) 101:426–434
DOI 10.1007/s00214990m185

Theoretical
Chemistry Accounts

Regular article

Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects

Jayashree Srinivasan, Megan W. Trevathan, Paul Beroza, David A. Case

Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

Received: 30 November 1998 / Accepted: 31 December 1998 / Published online: 16 March 1999

Abstract. The Poisson–Boltzmann (PB) continuum solvent model shows considerable promise in providing a description of electrostatic solvation effects in biomolecules, but it can be computationally expensive to obtain converged results for large systems. Here we examine the performance of a pairwise generalized Born approximation (GB) method on multiple conformations of a small peptide, three proteins (protein A, myoglobin, and rusticyanin) and four RNA and DNA duplexes and hairpins containing 20–24 nucleotides. Charge and dielectric radii models were adapted from the CHARMM and Amber force fields. Finite difference PB calculations were carried out with the Delphi and PEP programs, and for several examples the matrix of all pairwise interaction energies was determined. In general, this parameterization of the GB model does an excellent job of reproducing the PB solvation energies for small molecules and for groups near the surface of larger molecules. There is a systematic tendency for this GB model to overestimate the effects of solvent screening (compared to PB) for pairs of buried atoms, but individual errors tend to cancel, and a good overall account of conformational energetics is obtained. A simple extension to the GB model to account for salt effects (in the linearized Debye–Hückel approximation) is proposed that does a good job of reproducing the salt dependence of the PB calculations. In many cases, it should be possible to replace PB calculations with much simpler GB models, but care needs to be taken for systems with extensive burial of charges or dipoles.

Key words: Solvation – Electrostatics – Generalised Born theory – Salt effects – Continuum solvent

1 Introduction

There are many circumstances in molecular modeling studies where a simplified description of solvent effects has advantages over explicit modeling of each solvent molecule. One of the most popular models, especially for water, treats the solvent as a high dielectric continuum, interacting with charges that are embedded in solute molecules of lower dielectric. The solute charge distribution, and its response to the reaction field of the solvent dielectric, can be modeled either by quantum mechanics or by partial atomic charges in a molecular mechanics description. In spite of the severity of the approximation, this model often gives a good account of equilibrium solvation energetics, and is widely used to estimate p*K*s, redox potentials, and the electrostatic contributions to molecular solvation energies [1–5]. For molecules of arbitrary shape, the Poisson–Boltzmann (PB) equations that describe electrostatic interactions in a multiple-dielectric environment are typically solved by finite-difference or boundary-element numerical methods [1, 6–11]. These can be efficiently solved for small molecules, but become quite expensive for proteins or nucleic acids. For example, the DelphiII program, which is a popular program that computes a finite-difference solution, takes about 25 min on a 195 Mhz SGI processor to solve problems on a 145^3 grid, which would accommodate a myoglobin molecule with a grid spacing of 0.5 Å. Obtaining derivatives with respect to atomic positions adds to the time and complexity of the calculation [12]. Even though progress continues to be made in numerical solutions, there is a clear interest in exploring more efficient, if approximate, approaches to this problem.

One more approximate method that has received considerable attention is the generalized Born (GB) approach [13, 14]. In this model, the electrostatic contribution to the free energy of solvation is

$$\Delta G_{\mathrm{pol}} = -\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{i,j}\frac{q_i q_j}{f_{\mathrm{GB}}} \quad , \tag{1}$$

where q_i and q_j are partial charges, ε is the solvent dielectric constant, and f_{GB} is a function that interpolates between an "effective Born radius" α_i when the distance r_{ij} between atoms is short, and r_{ij} itself at large distances [14]. In the original model, values for α_i were determined by a numerical integration procedure, but it

Correspondence to: D.A. Case

has recently been shown that "pairwise" approximations, in which α_i is estimated from a sum over atom pairs, can be nearly as accurate and provide a simplified approach to energies and their derivatives [15–18]. Here we compare the performance of one of these models to finite-difference solutions to the PB equation, for a variety of systems ranging from peptides to proteins of about 150 amino acids, to oligonucleotides with about two dozen bases. The work here extends earlier similar comparisons in several ways:

1. We compare not only the total energies, but also each of the pairwise interaction energies in the two models.
2. We evaluate the performance of GB to reproduce PB results for the relative solvation energies of many conformers of the same molecule.
3. We propose and parameterize a simple extension of the GB model to include salt effects in the external medium.

The results show certain systematic differences between the GB and PB interaction energies for large systems, but suggest that for an important subset of biochemical problems the GB approach can serve as an efficient alternative to more elaborate calculations.

2 Methods

2.1 Systems studied

The principal aim of this study is to look at the ability of GB models to track changes in solvation energy as a function of conformational degrees of freedom. The systems studied here are listed in Table 1. For six of the molecules, multiple conformations were examined, extracted as snapshots from explicit solvent molecular dynamics simulations that have been previously described in the literature.

The AYPYD peptide conformations were extracted from molecular dynamics simulations that used the Amber/OPLS force field [19], and include extended conformers and more compact ones that form a type VI β turn. The range in solvation energies among these conformers is about 60 kcal/mol, with conformations in which the N and C termini are relatively close together being the least well solvated, and extended conformations having larger solvation free energies.

A second set of conformations comes from a study of the unfolding of protein A by Boczko and Brooks [20]. We chose a set of 100 snapshots from a CHARMM19/TIP3 molecular dynamics simulation that spanned a variety of conformations with radii of gyration ranging from 9 Å (for conformations near the native three-helix bundle) to 16 Å (for more extensively unfolded states). The range of solvation free energies is about 500 kcal/mol, again with the larger solvation energies generally corresponding to less compact configurations.

The "UUCG" simulation is of an RNA hairpin r(GGAC-UUCG-GUCC), where the loop has the UUCG sequence. The molecular dynamics simulations used the force field of Cornell et al. [21], starting from an NMR structure [22]. Although the conformations are all roughly similar, with Watson–Crick base-pairing in the stem and a slightly flexible loop, fluctuations in the positions of the charges (especially of the phosphates) leads to a range of solvation energies of about 100 kcal/mol.

The "B-DNA" set of conformations was taken from a molecular dynamics simulation of d(CCAACGTTGG)$_2$ in water, using the force field of Cornell et al. [23]. The "A-DNA" snapshots are for the same sequence, but are taken from a simulation in an 85% ethanol-water mixture; under these simulation conditions, the DNA remains reasonably near an A-form conformation [24]. As with the RNA simulations, there are substantial fluctuations in the solvation energies even among qualitatively similar conformations.

Table 1. Systems studied

Molecule	No. of atoms	No. of conformations	Charges	Radii	Notes
1 Aspartate	12	1	Amber95	Bondi	
2 AYPYD	63	10	OPLS	OPLS	a
3 Protein-A	459	100	Charmm19	Charmm19	b
4 UUCG tetraloop	382	100	Amber95	Bondi	c
5 Myoglobin	2543	1	Amber95	Bondi	d
6 Rusticyanin	2353	1	Amber95	Bondi	e
7 d(CCAACGTTGG)$_2$: B form	632	100	Amber95	Bondi	f
8 d(CCAACGTTGG)$_2$: A form	632	100	Amber95	Bondi	g
9 d(GCGCAATTGCGC)$_2$	758	100	Amber95	Bondi	h

[a] Only polar hydrogens and those on aromatic rings were represented. Radii correspond to the minima in the OPLS Lennard-Jones potential. Snapshots from the simulation described in Ref. [19]
[b] Only polar hydrogens are represented. Radii correspond to the minima in the OPLS Lennard-Jones potential. Snapshots at varying radii of gyration from the simulation described in Ref. [20]
[c] Snapshots from 0.06 to 1.05 ns of the RNA tetraloop simulation described in Ref. [22]
[d] Coordinates for carbonmonoxy myoglobin, from PDB entry 1MBC
[e] Coordinates for reduced rusticyanin from Ref. [52] PDB code 1CUR. Lone-pair charges on the S atoms of the methionine residues were collapsed onto the corresponding S atoms
[f] Snapshots from 0.4 to 1.4 ns of the B-form DNA simulation described in Ref. [23]
[g] Snapshots from 0.4 to 1.4 ns of the A-form DNA simulation described in Ref. [24]
[h] Snapshots from 30 to 130 ps of the B-form DNA simulation described in Ref. [25]

A B-form dodecamer, with the sequence $d(CGCGAATTCGCG)_2$ was analyzed from a solvated molecular dynamics simulation in water. This also used the force field of Cornell et al., and is described in detail elsewhere [25].

Results for single conformations of three additional systems are presented here as well. The aspartate ion illustrates the performance of the GB approach for a very small ion, and myoglobin and rusticyanin give an idea of the behavior to be expected for larger proteins, of about 150 amino acids, where burial of polar groups is more extensive than it is for protein A. For these systems, as well as for selected conformations of protein A and the nucleic acids systems and AYPYD, we have compared GB and PB estimates for all pairwise interactions among atoms. This is an expensive calculation for the larger systems (which is why we did it for only a few conformations) but allows a detailed analysis to be made of the strengths and weaknesses of the pairwise GB approach.

2.2 Finite-difference PB calculations

Most of the PB calculations were carried out using the DelphiII program, with a 0.25 Å grid spacing, and with charges and dielectric radii given in Table 1. The boundary conditions at the edge of the box were computed from Coulomb's law (or the Debye–Hückel expression when salt was present) using a solvent dielectric of 80. The interior dielectric in all cases was set to unity, and a minimum of 500 over-relaxation iterations was specified to ensure convergence. (Further details of the convergence behavior are presented elsewhere [26].) Salt effects were estimated at the linear PB level for ionic strengths of 0.1 and 0.5 M, and using an ion-exclusion radius of 2 Å.

Atomic pairwise interactions were computed using the PEP program [27]. This begins with a coarse grid (2.0 Å spacing) to cover the entire molecule, and proceeds through a series of focusing calculations on each atom in turn, with a final grid spacing of 0.015 Å. Although the focussing calculation uses a relatively small 33^3 grid, the requirement to carry out the calculation on each atom makes this procedure time consuming. However, we can use the results not only to compare total solvation energies with MEAD, Delphi or GB results, but can also make a comparison with GB for every pair of atoms in the system, providing a different view of the strengths and weaknesses of the GB approach.

2.3 GB calculations

There are several variants of the GB approach, which appear to provide comparable but not identical results. The particular approach tested here was introduced by Hawkins et al. [15, 16]. It calculates analytically the contribution of each atom j to α_i, the effective Born radius for atom i, and adds the contributions together. Since this procedure ignores overlaps among the atoms surrounding atom i, empirical correction factors are introduced to partially account for this behavior. Here we adopt the model that the correction factors depend only on the identity of atom j [15], and use scaling factors derived by Ponder and incorporated in the TINKER package (http: //dasher.wustl.edu/tinker): these values are listed in Table 2, and are only slight modifications of those originally proposed by Hawkins et al. [15]. As with some, but not all, GB implementations, the calculation of effective Born radii begins with radii reduced slightly from those used in the numerical PB calculations (in this case the reduction factor is 0.09 Å). Other GB parameterizations have been proposed [17, 18, 28, 29], and further developments may be expected. However, we believe that the broad conclusions discussed here, which are for systems ranging from a few atoms to those with 2500 atoms, and with solvation free energies from 0 to −6000 kcal/mol, will continue to be representative of results based on Eq. (1).

2.4 Salt effects

GB models have not traditionally considered salt effects, but the model can be extended to low salt concentrations at the Debye–Hückel level by the following arguments. The basic idea of the GB approach can be viewed as an interpolation formula between analytical solutions for a single sphere and for widely separated spheres. The form of f_{GB} in Eq. (1) can also closely reproduce results for a dipole in a sphere as well [14], and so the theory may plausibly work over a range of other geometries. For widely separated spheres, the solvation contribution in the linearized PB model becomes

$$\Delta G_{\mathrm{pol}} = -\left(1 - \frac{\mathrm{e}^{-\kappa r_{ij}}}{\varepsilon}\right)\frac{q_i q_j}{r_{ij}} \quad , \tag{2}$$

where κ is the Debye–Hückel screening parameter. The first term removes the gas-phase interaction energy, and the second term replaces it with a screened Coulomb potential. For a single spherical ion, the result is [30, 31]

$$\Delta G_{\mathrm{pol}} = -\frac{1}{2}\left(1 - \frac{1}{\varepsilon}\right)\frac{q_i^2}{b} - \frac{q_i^2 \kappa}{2\varepsilon(1+\kappa a)} \quad , \tag{3}$$

where b is radius of the sphere and a is the radial distance to which salt ions are excluded, and so $a - b$ is the ion exclusion radius. To a close extent, these two limits can be obtained by the simple substitution

$$\left(1 - \frac{1}{\varepsilon}\right) \rightarrow \left(1 - \frac{\mathrm{e}^{-\kappa f_{\mathrm{GB}}}}{\varepsilon}\right) \tag{4}$$

Table 2. Generalized Born scaling parameters

Atom	S_x
H	0.85
C	0.72
N	0.79
O	0.85
P	0.86
S	0.96
Fe	0.88

in Eq. (1). This reduces directly to Eq. (2) for large distances, and as r_{ij} goes to zero the salt-dependent terms become

$$-\frac{q_i^2\kappa}{2\varepsilon(1+\frac{1}{2}\kappa\alpha_i)} \tag{5}$$

through terms in κ^2. To terms linear in κ, Eqs. (5) and (3) agree, but the quadratic terms differ by the replacement of a with $\frac{1}{2}\alpha_i$. In practise, as we show later Eq. (4) gives salt effects that are somewhat larger than those predicted by finite-difference linearized PB calculations, but which are strongly correlated with them. One likely reason is that the GB model outlined here does not have the concept of an ion exclusion radius, and hence tends to overestimate salt effects compared to the PB model it is being compared to. We have found that acceptable results can be obtained by a simple scaling of κ by 0.73 in Eq. (4).

3 Results

3.1 Solvation energies

The overall comparison of PB and GB estimates of electrostatic solvation energies is given in Fig. 1. This illustrates that this GB model matches that from numerical PB solutions over a wide range of solvation energies, from near zero to −6000 kcal/mol. The optimal linear fit is very close to $y = x$, i.e., large and small systems are equally well predicted. Figure 2 is an expansion of Fig. 1 to look at 100 conformers of the B-DNA dodecamer. Here the range of solvation energies among various conformers is about 350 kcal/mol, and the root-mean-square (rms) error of GB predictions compared to PB is 3.8 kcal/mol, with an excellent correlation between the two data sets (correlation coefficient $r = 0.998$).

The two points significantly off the line in Fig. 1 correspond to the proteins rusticyanin and myoglobin; we believe that the poorer performance of GB for these systems is related to the fact that they have a larger interior, so that more atoms are shielded from solvent exposure than in nucleic acid duplexes. This point is discussed more fully below.

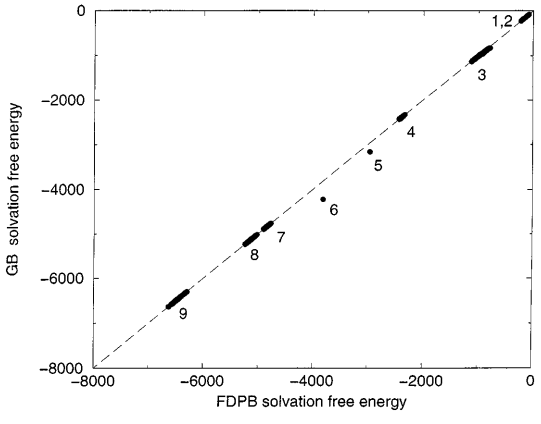

Fig. 1. Comparison of finite-difference Poisson–Boltzmann (*FDPB*) and generalized Born (*GB*) solvation free energies (in kcal/mol) at zero ionic strength; numbers refer to the nine sets of configurations listed in Table 1

3.2 Salt effects

The overall performance of our proposed method of including salt contributions in GB theory is shown for 0.1 M ionic strength in Fig. 3. The "salt contribution" is defined here as the difference between the predicted solvation energies at 0 and 0.1 M monovalent added salt. As with the total solvation energies, it is clear that there is little systematic bias in this GB method, compared to PB, and that the larger nucleic acid systems, where salt effects are highest, are treated nearly

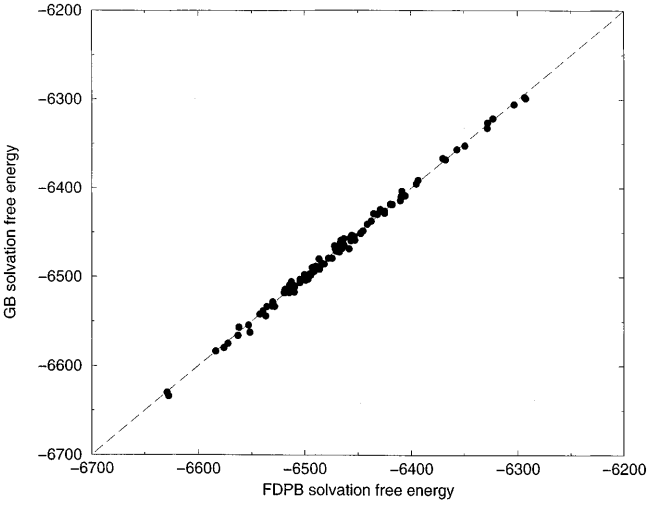

Fig. 2. Expansion of Fig. 1 for the B-DNA dodecamer (conformation set 9)

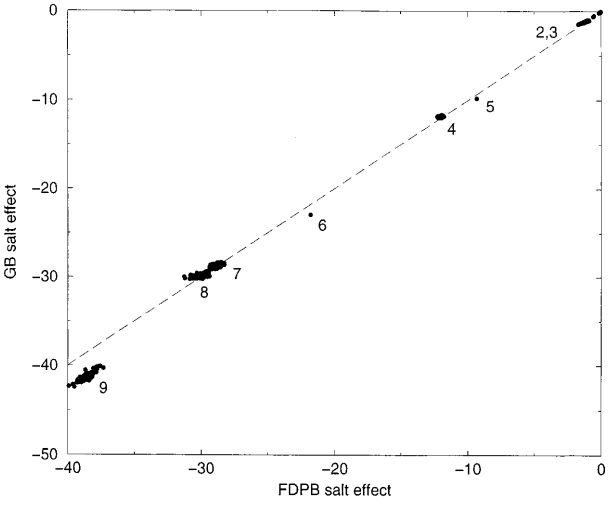

Fig. 3. Same as Fig. 1, but for the salt contribution to ΔG(solv) at 0.1 M ionic strength. Numbers refer to configurations listed in Table 1

as well as are smaller systems, although the correlation is certainly not perfect. The results for each of the 100 B-DNA conformers are given in Fig. 4. Here the range of salt effects is fairly small, varying by only about 3 kcal/mol over the various conformers. The correlation between GB and PB estimates of the salt effect is reasonably good (although poorer than for total solvation energies), with a correlation coefficient of 0.82 and an rms difference between the two estimates of 2.7 kcal/mol. Overall, it is clear that the GB estimates are capturing most of the behavior of PB for these examples.

The dependence of the salt contribution to solvation on the concentration of added monovalent salt is illustrated for one B-DNA conformer in Fig. 5. There is a rough accord between GB and PB estimates all the way to 1 M added salt, with the saturation of the salt effect occuring at about the same salt concentration in each theory. It should be noted, of course, that the linear PB model itself is expected to be valid only at low ionic strengths, and the use of either theory needs to be treated with caution. For many purposes though, the GB and PB theories are close enough to be effectively interchangeable, and so GB theory allows a preliminary investigation of salt effects at very low cost. For example, we have shown elsewhere [32] that the predicted effects of salt on an RNA hairpin-duplex equilibrium is nearly the same at 0.1 M added salt with linearized PB, nonlinear PB, or the present GB theory.

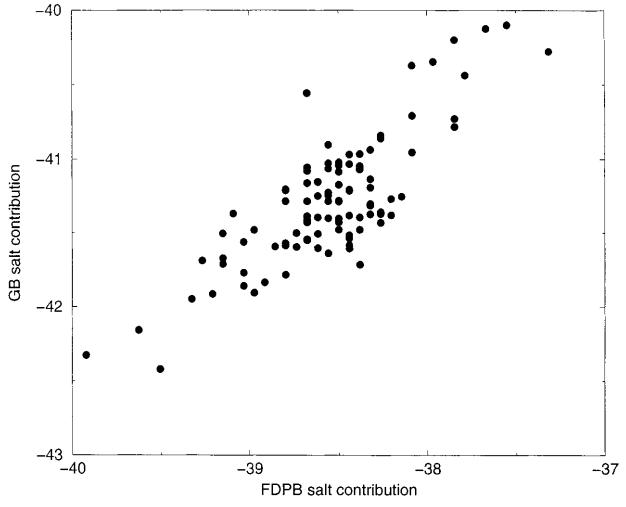

Fig. 4. Expansion of Fig. 3 for the B-DNA dodecamer (conformation set 9)

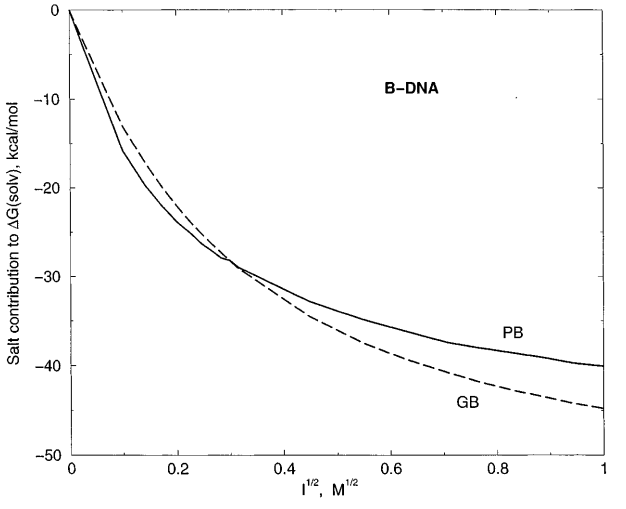

Fig. 5. Salt contribution for a B-DNA conformer as a function of the square root of the concentration of added monovalent salt

3.3 Pairwise interaction energies

For small molecules, it has been established that GB theories can give solvation free energies that closely match those from a numerical PB approach [17, 29, 33]. It is impressive, however, that this near equivalence extends to the individual terms in Eq. (1) as well. The GB and PB results for the aspartate ion, which, as an ion, has a significant total solvation energy, but for which all atoms are also near the surface, are compared in Fig. 6. The excellent correlation shows that the particular parameters used here for GB indeed mimic well the interactions of a numerical PB solution. It is important that GB implementations work well in situations like this, since many charged side chains in proteins have high solvent accessibility.

Figure 7 shows pairwise interactions for one configuration of protein A, and Fig. 8 shows similar results for one of the b12 conformers. Here there is a more complex behavior in which some of the individual terms are significantly different in the PB and GB models. In Fig. 7 individual terms are divided into off-diagonal terms, reflecting charge-charge interactions between atoms, and diagonal terms, that represent the energetic consequences of moving individual charges from a vacuum environment to the final protein/solvent environment. In each case the great majority of the contributions

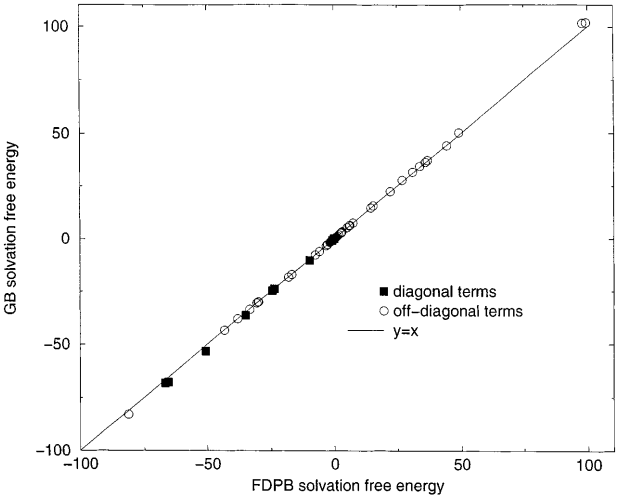

Fig. 6. Comparison of FDPB and GB estimates of solvation energies (in kcal/mol) for all pairs of charges in the aspartate ion. The *squares* represent the diagonal terms, the *circles* represent the off-diagonal terms, and the *line* represents the equation $y = x$

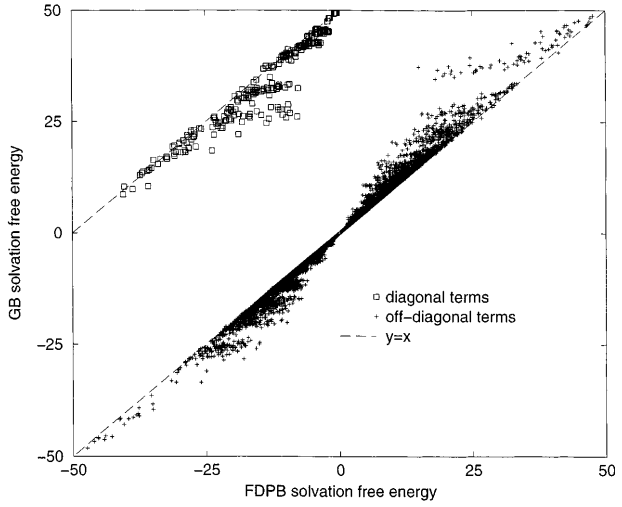

Fig. 7. As for Fig. 6, for one of the protein A configurations. Diagonal terms (□) have been shifted up by 50 kcal/mol for better visibility. The off-diagonal terms are represented by *crosses*. In each case, the *dashed line* ($y = x$) indicates equality between the FDPB and GB models

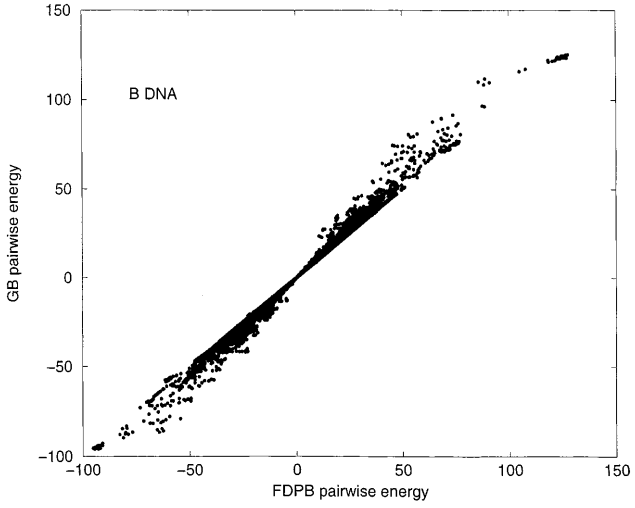

Fig. 8. As for Fig. 6, for one of the B-DNA configurations from set 7 in Table 1

(consisting of $N = 459$ diagonal terms and $N(N-1)/2 = 105,048$ off-diagonal terms) are about the same for both models, with a small percentage of the terms showing significant differences. In Fig. 7, for example, of the 105,048 off-diagonal terms, 276 (or 0.3%) show absolute differences between PB and GB of more than 5 kcal/mol, and 2560 (or 2.4%) have an absolute difference greater than 1 kcal/mol. Figures 7 and 8 visually emphasize the larger differences, since the many points on the $y = x$ line lie on top of each other.

For the off-diagonal terms, the absolute values for the GB solvation energies tend to be larger than those for the PB model; this means that the screening of charge-charge interactions is greater in the GB model. If we express this in terms of an effective dielectric constant ε_{ij}, defined by

$$\Delta G_{ij} = \frac{-q_i q_j}{r_{ij}} \left(1 - \frac{1}{\varepsilon_{ij}}\right) \tag{6}$$

then the effective dielectric between pairs of charges tends to be greater for GB theory than for the PB model. This is illustrated in Fig. 9, which shows the distribution of effective dielectrics as a function of distance between the charges for both the GB and PB models. The GB results look roughly like those of a sigmoidal distance-dependent dielectric function [34–36], whereas the PB results have more variable behavior. For most pairs closer than 10 Å, this effective pairwise dielectric constant is greater for the GB model than for the PB one. Since this over-screening of interactions leads to opposite errors in solvation energies (relative to PB) for like-charge interactions compared to opposite-charge interactions, and since there are roughly equal numbers of both types of interactions, the difference in total solvation energies between GB and PB models is reasonably small in spite of these individual errors. For example, the mean difference between PB and GB solvation energies for 10 conformers of protein A is about 10 kcal/mol, even though the differences in the individual contributions range from −16 to +22 kcal/mol. The over-screening in the GB model, though, could have significant effects on results for pK or redox potentials, which are sensitive to the mutual interactions of individual charges; this distinction between performance for pK calculations and for overall solvation energies has been analyzed recently by Jayaram et al. [37].

For the diagonal terms in the solvation energy expression, shown by the squares in Fig. 7, the GB solvation energies are often close to those of the PB theory, but tend to be more negative for certain buried atoms. This is illustrated in Fig. 10, which plots the difference of GB and PB diagonal contributions to the solvation energy against the effective Born radius α_i for the peptide oxygen atoms in protein A. These atoms all have the same partial charge and initial radius (1.4 Å), and so the variation in effective radius is a measure of the extent of burial from the solvent. For effective radii less than about 1.8 Å, the GB and PB theories are very close to each other (consistent with the results of Fig. 6), with increasingly large errors for greater amounts of protection from the solvent. This difference is consistent with that seen in the off-diagonal terms: the environment around a charge in the GB model has a higher effective dielectric in the GB than in PB model, leading to more negative solvation energies for individual charges.

It is natural to ask whether the systematic differences between the two models are unique to the particular implementation of the GB model used here, and whether they might not be minimized by a suitable reparameterization. A number of groups have looked at the development of modified GB theories that could minimize discrepancies (over some training set) of the GB and PB results [15–17, 33, 37, 38]. It is not easy to describe at the present time the relative strengths and weaknesses of different parameterizations as they might show up in macromolecular calculations, although the general performance of various versions of Eq. (1) should have

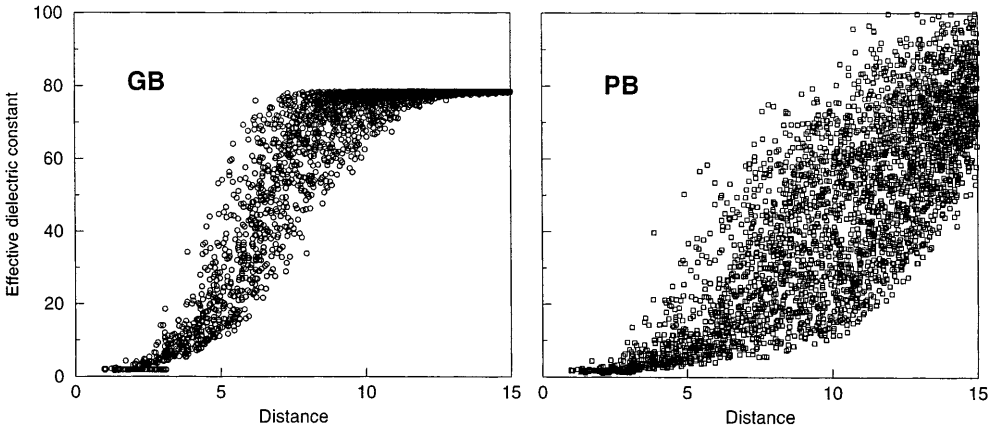

Fig. 9. Effective screening constant (defined by Eq. 6) for pairs of atoms in protein A, showing results for both the PB and the GB models

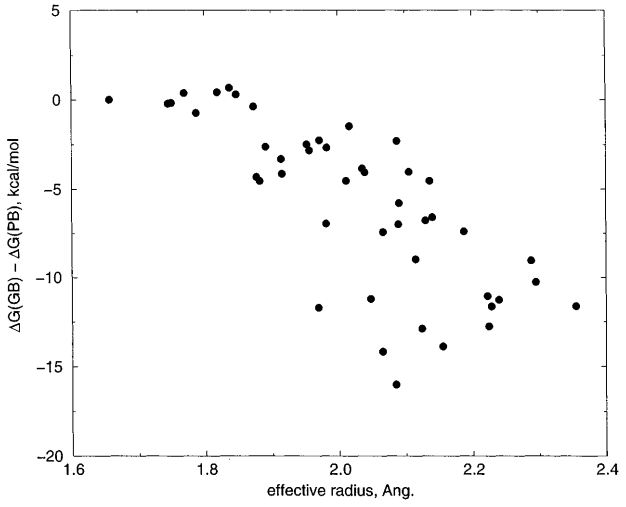

Fig. 10. Diagonal terms from Fig. 7, plotted against the effective Born radius α_i, for the peptide oxygen atoms in protein A

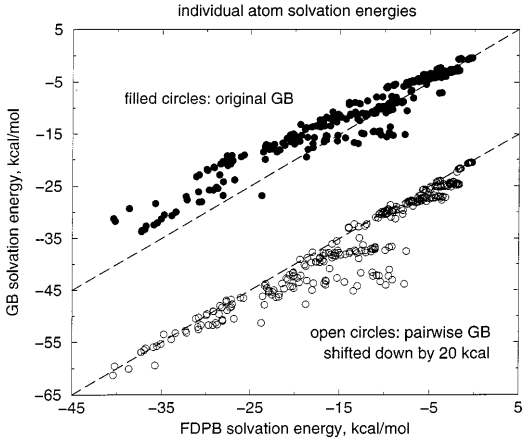

Fig. 11. Comparison of the diagonal contributions to the solvation energy for the protein A conformation shown in Fig. 7. *Filled circles* give results form the "original" GB model, as implemented in the Macromodel program; *open circles* (shifted downward by 20 kcal/mol for visibility) give results for the present pairwise GB model

some common characteristics. As a first step toward cross comparisons, the original and pairwise GB results for the diagonal contributions to solvation energy for protein A are compared in Fig. 11. (Here the "original" GB results are those of the Macromodel program.) The original GB model has differences of both signs with respect to the PB model for individual atom solvation energies; the pairwise GB model has differences of approximately equal magnitude, but which are more consistent in always being negative, and which more clearly distinguish surface atoms (for which GB and PB models give nearly identical results) from buried atoms (where the GB solvation energy is more negative than the PB counterpart).

4 Discussion

In this work, we have broadly adopted the perspective that one measure of the usefulness of GB models of solvation lies in their ability to reproduce results from PB calculations that also use a continuum model and adopt the same charges and Born radii. This is a straightforward way to evaluate the generalized Born model, but is certainly not the only course that could be taken. It should be emphasized that the PB model itself is only at best an approximate view of true solvation energetics, and that agreement between PB and GB theories is only of interest to the extent that either reflects the true energetics of solvation [39]. Several groups have recently shown that PB or GB estimates of differences in solvation energies between conformers (which is the property examined here) are effective in rationalizing conformational preferences of macromolecules in solution: these studies have included comparisons of "A" and "B" form nucleic acid double helices [26, 40, 41], loop conformations in proteins and in RNA [32, 42], extended and folded forms of peptides [29, 42–50], and folded versus misfolded forms of proteins [51]. This, combined with the generally excellent correlations with experiment found for small molecules [16, 17, 33], suggests that the GB approach can be an important tool for many types of investigations, especially since it is fast enough to be used as (part of) an effective solvation model for Monte Carlo or molecular dynamics simulations.

It also seems clear, however, that GB models, especially of the "pairwise" variety considered here, can only match the PB model for certain ranges of the charge and atomic radii parameters, which in some sense are near to the values used to calibrate the "overlap" or "correction" factors in Table 2 (or equivalent parameters in other formulations). Many plausible physical models fall outside this range, and would not be readily amenable to the GB approach considered here. For example, the use of a single large sphere to represent an entire protein (as in Tanford–Kirkwood theory) is straightforward to study with numerical PB methods, but would require an entirely different type of GB theory than the one considered here. Models that treat membrane bilayers as "slabs" of low dielectric would face similar problems. Warshel and Papazyan [39] have recently emphasized these limitations of GB theory, pointing out the restricted range of problems to which it should be applied. The results found here support this view, but suggest that the range of applicability extends to at least some interesting "macromolecular" problems. Figure 9 illustrates that for charge-charge interactions, GB theory looks roughly, but not exactly, like a model that uses a distance-dependent dielectric constant. It should be emphasized, however, that GB theory (unlike simple distance-dependent dielectric models) also includes "diagonal" contributions that penalize the burial of individual atomic partial charges, and hence has a more complete, if necessarily approximate, description of solvation energies.

In our view, a successful GB model ought to preserve the remarkable accuracy that is currently found for small molecules, where all of the atoms are near the surface, as illustrated in Fig. 6. This will help to ensure that the behavior of exposed side chains in proteins, which often tend to dominate electrostatic interactions, are described in a faithful way. In particular, it does not seem to us wise to effect some compromise which improves overall performance only at the expense of describing surface atoms in a worse fashion. This is in some sense what happens with the original, numerical, formulation of GB theory, as illustrated in Fig. 11.

Within these broad constraints, however, it does appear that GB models can be remarkably effective in providing approximations to solvation energies computed from numerical PB codes, and that a relatively straightforward extension that deals with salt effects shares this general accuracy and breadth. There are, however, significant differences that remain, especially for more strongly buried charges, and the results outlined here do not provide a general solution to this problem. It is certainly possible that alternative formulations of the GB model could substantially reduce this remaining level of disagreement, and there are a variety of recent efforts along such lines.

Acknowledgements. This work was supported by NIH grant RR 12255, and by an NSF predoctoral fellowship to MWT. We thank Doree Sitkoff for assistance with Delphi, Peter Kollman, Tom Cheatham, Jennifer Miller, and Charles Brooks for providing snapshots from their molecular dynamics simulations, and Brian Dominy, Charles Brooks, Don Bashford, and David Beveridge for helpful discussions and for providing results in advance of publication.

References

1. Tomasi J, Persico M (1994) Chem Rev 94: 2027–2094
2. Honig B, Sharp K, Yang AS (1993) J Phys Chem 97: 1101–1109
3. Davis ME, McCammon JA (1993) Chem Rev 93: 509–521
4. Beroza P, Case DA (1998) Methods Enzymol 295: 170–189
5. Chambers CD, Hawkins GD, Cramer CJ, Truhlar DG (1996) J Phys Chem 100: 16398–17385
6. Nicholls A, Honig B (1991) J Comput Chem 12: 435–445
7. Holst M, Saied E (1993) J Comput Chem 14: 105–113
8. Zauher RJ, Varnek A (1996) J Comput Chem 17: 864–877
9. Cortis CM, Friesner RA (1997) J Comput Chem 18: 1591–1608
10. Vorobjev YN, Scheraga HA (1997) J Comput Chem 18: 569–583
11. Bruccoleri RE, Novotny J, Davis ME, Sharp KA (1997) J Compu Chem 18: 268–276
12. Gilson MK, Davis ME, Luty BA, McCammon JA (1993) J Phys Chem 97: 3591–3600
13. Constanciel R, Contreras R (1984) Theor Chim Acta 65: 1
14. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) J Am Chem Soc 112: 6127–6129
15. Hawkins GD, Cramer CJ, Truhlar DG (1995) Chem Phys Lett 246: 122–129
16. Hawkins GD, Cramer CJ, Truhlar DG (1996) J Phys Chem 100: 19824–19839
17. Qiu D, Shenkin PS, Hollinger FP, Still WC (1997) J Phys Chem A 101: 3005–3014
18. Edinger SR, Cortis C, Shenkin PS, Friesner RA (1997) J Phys Chem B 101: 1190–1197
19. Demchuk E, Bashford D, Case DA (1997) Folding Des 2: 35–46
20. Boczko EM, Brooks CL III (1995) Science 269: 393–396
21. Cornell WD, Cieplak P, Bayly CJ, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DG, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117: 5179–5197
22. Miller JL, Kollman PA (1997) J Mol Biol 270: 436–450
23. Cheatham TE III, Kollman PA (1996) J Mol Biol 259: 434–444
24. Cheatham TE III, Crowley MF, Kollman PA (1997) Proc Natl Acad Sci USA 94: 9626
25. Cieplak P, Cheatham TE III, Kollman PA (1997) J Am Chem Soc 119: 6722–6730
26. Srinivasan J, Cheatham TE III, Kollman PA, Case DA (1998) J Am Chem Soc 120: 9401–9409
27. Beroza P, Fredkin DR (1996) J Comput Chem 17: 1229–1244
28. Scarsi M, Apostolakis J, Caflisch A (1997) J Phys Chem A 101: 8098–8106
29. Scarsi M, Apostolakis J, Caflisch A (1998) J Phys Chem B 102: 3637–3641
30. Kirkwood JG (1934) J Chem Phys 2: 351–361
31. Tanford G, Kirkwood JG (1957) J Am Chem Soc 79: 5333–5339
32. Srinivasan J, Miller J, Kollman PA, Case DA (1998) J Biomol Struct Dyn (in press)
33. Murphy RB, Pollard WT, Friesner RA (1997) J Chem Phys 106: 5073–5084
34. Hingerty BE, Ritchie RH, Ferell TL, Turner JE (1985) Biopolymers 24: 427
35. Lavery R, Sklenar H, Zakrzewska K, Pullman B (1986) J Biomol Struct Dyn 3: 989
36. Mehler EL (1996) J Phys Chem 100: 16006–16018
37. Jayaram B, Liu Y, Beveridge DJ (1998) J Phys Chem 109: 1465–1471
38. Dominy BN, Brooks CL III (1999) J Phys Chem (in press)
39. Warshel A, Papazyan A (1998) Curr Opin Struct Biol 8: 211–217
40. Jayaram B, Sprous D, Young MA, Beveridge D (1998) J Am Chem Soc 120: 10629–10633
41. Cheatham TE III, Srinivasan J, Case DA, Kollman PA (1998) J Biomol Struct Dyn 16: 265–280
42. Smith KC, Honig B (1994) Proteins 18: 119–132

43. Bashford D, Case DA, Choi C, Gippert GP (1997) J Am Chem Soc 119: 4964–4971
44. Demchuk E, Bashford D, Gippert G, Case DA (1997) J Mol Biol 270: 305–317
45. Yang AS, Honig B (1995) J Mol Biol 252: 351–365
46. Yang AS, Hong B (1995) J Mol Biol 252: 366–376
47. Yang, AS, Hitz B, Honig B (1996) J Mol Biol 259: 873–882
48. Marrone TJ, Gilson MK, McCammon JA (1996) J Phys Chem 100: 1439–1441
49. Smart JL, Marrone TL, McCammon JA (1997) J Comput Chem 18: 1750–1759
50. Luo R, Moult J, Gilson MK (1997) J Phys Chem B 101: 11216–11236
51. Vorobjev YN, Almagro JC, Hermans J (1998) Proteins 32: 399–413
52. Botuyan MV, Toy-Palmer A, Chung J, Blake RC II, Beroza P, Case DA, Dyson HJ (1996) J Mol Biol 263: 752–767