Effective Born Radii in the Generalized Born Approximation: The Importance of Being Perfect

ALEXEY ONUFRIEV, DAVID A. CASE, DONALD BASHFORD
Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, California 92037

Received 7 December 2001; accepted 19 April 2002

Abstract: Generalized Born (GB) models provide, for many applications, an accurate and computationally facile estimate of the electrostatic contribution to aqueous solvation. The GB models involve two main types of approximations relative to the Poisson equation (PE) theory on which they are based. First, the self-energy contributions of individual atoms are estimated and expressed as "effective Born radii." Next, the atom-pair contributions are estimated by an analytical function f^{GB} that depends upon the effective Born radii and interatomic distance of the atom pairs. Here, the relative impacts of these approximations are investigated by calculating "perfect" effective Born radii from PE theory, and enquiring as to how well the atom-pairwise energy terms from a GB model using these perfect radii in the standard f^{GB} function duplicate the equivalent terms from PE theory. In tests on several biological macromolecules, the use of these perfect radii greatly increases the accuracy of the atom-pair terms; that is, the standard form of f^{GB} performs quite well. The remaining small error has a systematic and a random component. The latter cannot be removed without significantly increasing the complexity of the GB model, but an alternative choice of f^{GB} can reduce the systematic part. A molecular dynamics simulation using a perfect-radii GB model compares favorably with simulations using conventional GB, even though the radii remain fixed in the former. These results quantify, for the GB field, the importance of getting the effective Born radii right; indeed, with perfect radii, the GB model gives a very good approximation to the underlying PE theory for a variety of biomacromolecular types and conformations.

 J Comput Chem 23: 1297–1304, 2002

Key words: generalized Born approximation; Poisson equation; effective Born radii; molecular dynamics; macromolecules

Introduction

An accurate description of electrostatic interactions in aqueous environments is essential for theoretical modeling of biomolecules. The semicontinuum electrostatics model makes a fundamental approximation by replacing discrete water molecules with a continuum medium with the dielectric properties of water. Within this approximation the electrostatic potential can be found—with a desired degree of accuracy—by numerical solution of the Poisson equation (PE). If salt is present, the Poisson-Boltzmann equation can be used, but the present focus is on the limit of zero salt concentration. Such models have been successfully applied for many years to calculate various macromolecular properties.[1–4] The standard PE approach to the calculation of the electrostatic component of solvation free energy ΔG_{el} involves consideration of charging processes in vacuum versus solvent, leading to the expression

$$\Delta G_{el} = \frac{1}{2}\sum_i q_i(\phi_{sol}(\mathbf{r}_i) - \phi_{vac}(\mathbf{r}_i)) \quad (1)$$

$$= \frac{1}{2}\sum_{ij} q_i q_j(\Phi_{sol}(\mathbf{r}_i, \mathbf{r}_j) - \Phi_{vac}(\mathbf{r}_i, \mathbf{r}_j)) \quad (2)$$

In the more conventional first form, ϕ_{sol} is the potential of the solute atomic partial charges q_i in an environment where the region inside the molecular surface (as determined by the atomic coordinates, radii, and probe sphere radius) has a dielectric constant of 1.0, and the exterior region has the dielectric constant of water ε_w; ϕ_{vac} is the potential of the same solute charges in an

Correspondence to: D. A. Case

Contract/grant sponsor: NIH; contract/grant number: GM57513

environment of uniform dielectric constant 1.0. (models with solute interior dielectric constants other than 1.0 are also commonly used). The second form, eq. (2), which displays the full dependence of ΔG_{el} on the solute charges, makes use of the Green functions of the Poisson problems, $\Phi(\mathbf{r}_i, \mathbf{r}_j)$. The Green function of a Poisson problem is defined as the solution of PE for the potential at $\mathbf{r}_i$ with the charge distribution of the original Poisson problem replaced by a single unit charge at $\mathbf{r}_j$. The Green function is related to the actual potential of a particular charge distribution by $\phi(\mathbf{r}) = \sum_j q_j \Phi(\mathbf{r}, \mathbf{r}_j)$. The Green function for the vacuum case Φ_{vac} is simply the Coulombic Green function $|\mathbf{r}_j - \mathbf{r}_i|^{-1}$, while Φ_{sol} is much more complex due to the dielectric boundary—it includes the Coulombic term, but also a reaction-field term.

For many applications, PE methods present practical problems because of the computational cost of solving the PE and the difficulty of obtaining energy gradients with respect to atomic coordinates. Generalized Born (GB) models[1,5–8] offer an approximation to PE theory that is computationally more convenient, particularly in its pairwise analytic versions,[9,10] which are well suited to molecular mechanics calculations. Insofar as the GB models are fundamentally based on the same underlying approximation as PE theory, that is continuum electrostatics, the PE (PB) theory provides a natural reference point in the analysis of their accuracy. The GB model approximates eq. (1) or (2) by an analytical formula[5]:

$$\Delta G_{el} \approx \Delta G_{GB} = -\sum_i \frac{q_i^2}{2R_i}\left(1-\frac{1}{\varepsilon_w}\right) - \frac{1}{2}\sum_{ij,i\neq j}\frac{q_i q_j}{f^{GB}(r_{ij}, R_i, R_j)}\left(1-\frac{1}{\varepsilon_w}\right) \quad (3)$$

where the R_i are the *effective Born radii* of the atoms i and

$$f^{GB} = [r_{ij}^2 + R_i R_j \exp(-r_{ij}^2/4R_iR_j)]^{1/2} \quad (4)$$

Comparing the terms of eqs. (2) and (3) we find that each of the $1/R_i$ terms of eq. (3), which are just the traditional Born formulae[11] for the solvation free energy of isolated ions, corresponds to a self-energy term from eq. (2), that is

$$\frac{1}{R_i} \sim (\Phi_{sol}(\mathbf{r}_i, \mathbf{r}_i) - \Phi_{vac}(\mathbf{r}_i, \mathbf{r}_i)) \quad (5)$$

In other words, R_i should be that radius which, if inserted into the Born formula, would give the same electrostatic solvation free energy that PE theory would give for a hypothetical molecule having the same dielectric boundary as the original, but having only the single charge q_i at position $\mathbf{r}_i$.

Usually, the effective Born radii are estimated by an expression involving an integral of the energy density of a Coulomb field over the molecular volume, and are evaluated by numerical integration[5] or analytical approximations to such an integral.[9,10] Thus, the conventional GB model can be seen as containing two levels of approximation relative to the Poisson model on which it is based: the assumption of the functional form, eqs. (3) and (4), and the Coulomb and integration approximations used in determining the effective Born radii. The method was originally developed[5] for small compounds where it was found to reproduce solvation energies and individual charge-charge interactions quite well when compared to solutions of PE. However, its performance on larger molecules fell short of expectations, especially for molecules with larger interior regions. Modifications have since been proposed[12–15] that improved the model's performance in various ways, and similar work is in progress in a number of research groups.

Here we ask to what extent could the GB model be improved if "perfect" values for the effective Born radii could be used? Numerical solutions of the PE for the potentials due to individual charges, incorporating the full complex dielectric shape, are used to determine R_i values that make eq. (5) exact (to within the accuracy of the numerical Poisson solver). The resulting "perfect radii" are then used in eqs. (3) and (4), and we enquire as to how well this perfect-radii version of GB compares to a standard GB (as implemented in AMBER), in terms of its ability to reproduce the results of a full Poisson treatment of solvation [eq. (1)]. By comparison with GB models using one of the standard estimates of effective radii, we draw some conclusions about which approximations within the GB model cause the most serious problems and where efforts at improvement ought to be focused. We also consider the practicality of using computed perfect-radii in certain classes of applications.

Methods and Data

We have carried out tests on macromolecules representing different structural classes: native myoglobin (PDB ID 2MB5), B-DNA (10 base-pair duplex), and completely unfolded apomyoglobin. The unfolded apomyoglobin structure is generated by molecular dynamics simulations that model the conditions of acid-induced unfolding.[24] It has no tertiary structure and about 20% residual secondary structure, and is similar to the experimentally observed acid-unfolded state of myoglobin. The calculations reported in the Appendix also include bacteriorhodopsin (PDB ID 1QHJ). For each atom of each test structure, a Poisson problem is set up and solved having the dielectric boundary shape of the full molecule present, but keeping only the charge of that particular atom. The van der Waals (vdW) radii of Bondi[16] and a solvent-probe radius of 1.4 Å are used to define molecular surface,[17] which is taken as the dielectric boundary. The accumulation of these solutions gives the necessary Green-function information for use in eq. (2) for the full Poisson solvation energy, and in eq. (5) for the perfect effective Born radii. The computer program PEP, developed by P. Beroza[3] and available via the Internet,[25] is used for the set-up and solution of these Poisson problems. The finest grid spacing used in all calculations is 0.07 Å, decreasing from 4 Å in eight steps of focusing on the atom in question.

For the molecular dynamics test, the crystal structure of *E. coli* thioredoxin (PDB ID 2TRX) is used as a starting point. A GB model with specifiable fixed effective Born radii is incorporated into the AMBER 6.0 molecular dynamics package.[15,18] An all-atom force field[19] is employed. SHAKE is used to restrain hydrogen-heavy atom bond distances. The integration time-step is 2 fs, with an essentially infinite (300 Å) cut-off for long-range interactions. The average temperature of the system is maintained at 300

Table 1. Solvation Energy Components of Various Models.

Structure:	Myoglobin native			DNA			Myoglobin unfolded		
Model	Self [d]	Cross[e]	Total	Self	Cross	Total	Self	Cross	Total
PE[a]	−11843	8880	−2963	−8928	4253	−4675	−19053	15163	−3890
Perf. GB[b]	−11843	8866	−2976	−8928	4217	−4711	−19053	15108	−3945
Stand. GB[c]	−14949	11784	−3165	−9582	4998	−4586	−20671	16652	−4019

[a]Energy from eq. (1) or (2).
[b]Generalized born model [eqs. (3) and (4)] using "perfect" radii.
[c]GB model of ref. 10.
[d]Sum of the ii terms in eq. (2) or q_i^2/R_i terms in eq. (3).
[e]Sum of the ij terms in eq. (2) or (3). All energies in kcal/mol.

K. The surface-area-based apolar solvation term with a surface tension parameter set to 0.005 kcal/mol Å^2 is also included in the potential energy function.

Results

For macromolecules representing different structural classes, native myoglobin, B-DNA, and completely unfolded myoglobin, the electrostatic contribution to the solvation free energy (ΔG_{el}) is calculated using the perfect-radius version of GB theory and compared to PE calculations. Comparisons are also made to a standard GB model, as implemented in AMBER 6.0.[15] By "standard" or "conventional," we mean an analytical GB model starting with the same vdW radii used in the PE calculations and using the analytical method of Hawkins et al.[10] for the calculation of effective Born radii, and eq. (4) for f^{GB}. This particular GB model has been used for quite some time and provides a convenient point of comparison with the perfect-radii GB model, although it is certainly not the only currently used GB model. The perfect-radii GB model gives a significantly better estimate of ΔG_{el} than the standard GB model (Table 1). By construction, the perfect-radii GB model gives the same self-term contributions as Poisson theory. What is striking is that the use of perfect radii also significantly improves the agreement of the sum of the cross terms ($i \neq j$) of the GB model with the equivalent part of the PE energy (Table 1). The perfect-radii GB model gives a sum of cross terms well within 1% deviation from the PE results for all three structures, while for the standard GB model, the corresponding deviation can be as high as 30% (native myoglobin).

Having found that the sum of the cross terms is improved by the use of perfect radii, we can enquire as to whether individual cross terms are similarly improved. A scatter plot of GB versus PE cross terms for native-state myoglobin shows that the use of perfect radii does indeed improve the GB cross terms significantly (Fig. 1). The RMSD from the PE cross terms for the perfect-radii and standard GB models are 0.07 and 0.29 kcal/mol, respectively, and unlike the standard model, the accuracy of the perfect-radii

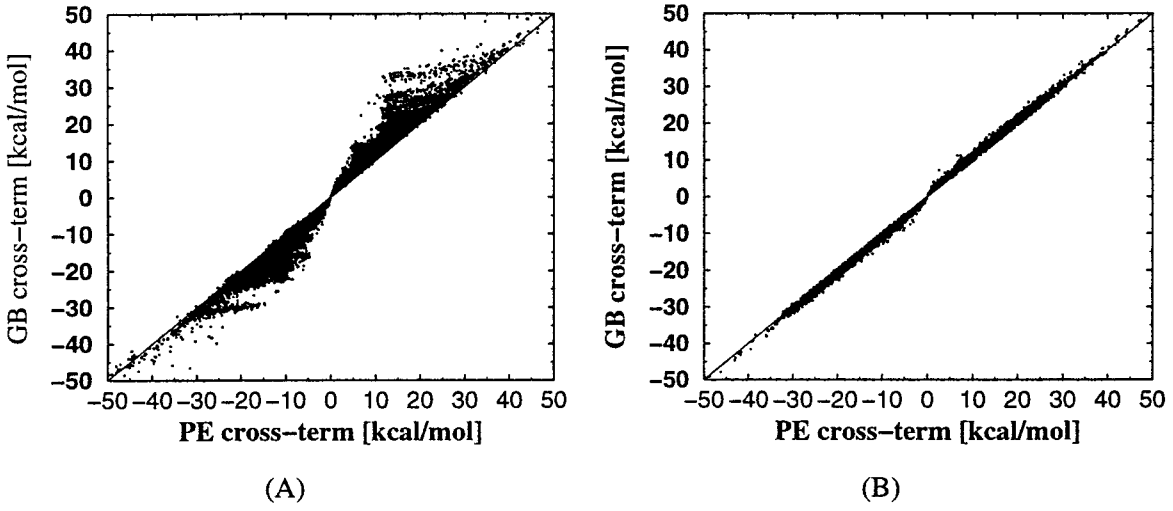

Figure 1. Comparison of individual cross terms in standard (A) or perfect-radii (B) GB models [ij terms of eq. (3)] to cross terms from PE theory [ij terms of eq. (2)] for the native state of myoglobin. The line $x = y$ represents a perfect match between the GB and PE theories.

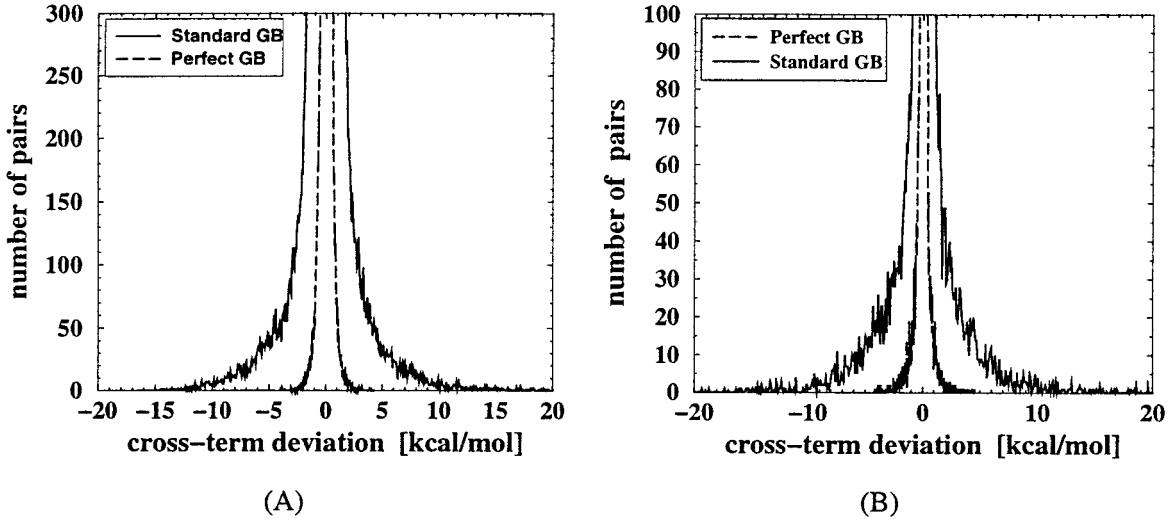

Figure 2. The distribution of deviations of individual cross terms of GB models from the PE cross terms for myoglobin (A) and B-DNA (B). The energy interval from −40 to +40 kcal/mol is divided into 500 bins of 0.16 kcal/mol each. The number of atom pairs whose cross-term deviation between GB and PE falls within a given bin is plotted on the ordinate.

model is fairly uniform across the entire range of energies. Plots of the distributions of cross-term errors for the two models for native myoglobin and B-DNA (Fig. 2) show that the relative widths of the distributions do not appear to be system specific. We conclude that the approximation embodied in the choice of the functional forms, eqs. (3) and (4), is, by itself, a very reasonable one, and getting the effective radii and hence the self energy right allows one to estimate the pairwise charge-charge interactions quite accurately.

Given perfect radii, the origins of any remaining error in the cross terms can be explored by comparing $f^{\mathrm{GB}}(\mathbf{r}_{ij}, R_i, R_j)$ to a corresponding f value calculated by matching all the terms of PE and GB theory. Setting equal each of the ij terms of eqs. (2) and (3), with the new f replacing f^{GB} in the latter, one obtains an expression for the new, "perfect" f values:

$$f_{ij}^{\mathrm{PE}} = -\frac{1}{2}\left(1 - \frac{1}{\varepsilon_w}\right)(\Phi_{\mathrm{sol}}(\mathbf{r}_i, \mathbf{r}_j) - \Phi_{\mathrm{vac}}(\mathbf{r}_i, \mathbf{r}_j))^{-1} \qquad (6)$$

The performance of the standard f^{GB} expression, eq. (4), is measured against these "perfect" values in Figure 3, where the error is displayed as two-dimensional surfaces by taking advantage of the fact that the standard expression can be written as a function of two variables, the interatomic distance r_{ij} and the harmonic average of the effective radii, $\langle R\rangle = (R_i R_j)^{1/2}$. The relative systematic errors of the standard f^{GB} are small: at most 15% at $\langle R\rangle \sim 6$ Å, $r_{ij} \sim 10$ Å for myoglobin, and only a few percent for the DNA, while the random component of the error is larger, especially in the important region of small r_{ij}. Note that significantly smaller systematic deviation in the DNA case does not correspond to a noticeably smaller random error (see also the distributions in Fig. 2). An effort to reduce the systematic component of the error is reported in the Appendix.

The perfect-radii GB model is used in a 6 ns molecular dynamics simulation of thioredoxin in its native state at 300 K. The potential energy function incorporates a GB solvation term [eq. (3)] for which perfect-radii are supplied from a preliminary set of Poisson equation solutions. These radii are kept fixed throughout the simulation. The back-bone RMSD from the X-ray structure remains reasonably small, staying below 1.5 Å in the first 2 ns of simulation and ranging from 1.6 to 1.8 Å during the final 4 ns (Fig. 4). This contrasts favorably with recent experience with the use of more conventional GB models in molecular dynamics simulations of proteins, in which deviations from the native structure were larger, even though the effective Born radii were updated at every step of MD.[18,23] However, the RMSDs obtained in our simulations using the perfect-radii GB model are not as small as in an equivalent high-level explicit-solvent simulation (in a separate simulation of thioredoxin[18] that used explicit water and particle-mesh Ewald, the RMSD was 1.2 Å at $t = 6$ ns). The slight (~7% relative to the starting structure) decrease in the over-all protein volume that we observe in our perfect-radii simulation may be due to the underestimation of charge-solvation forces in the fixed-radii procedure: charged groups may shift towards the protein interior, as the self-energy of a single charge would not increase if it moved towards the interior regions.

The more usual way of incorporating GB models into MD is to use an analytic estimate of the radii at each step, and to include gradient terms due to the coordinate dependence of the radii in the forces. The use of fixed effective Born radii here implies that the "perfect" radii do not change very much during the simulation. This is checked by comparing the perfect radii computed from the native structure (and used in the simulation) with radii computed from the final, $t = 6$ ns structure. The comparison shows that final values cluster around the initial values, and that the bulk of the

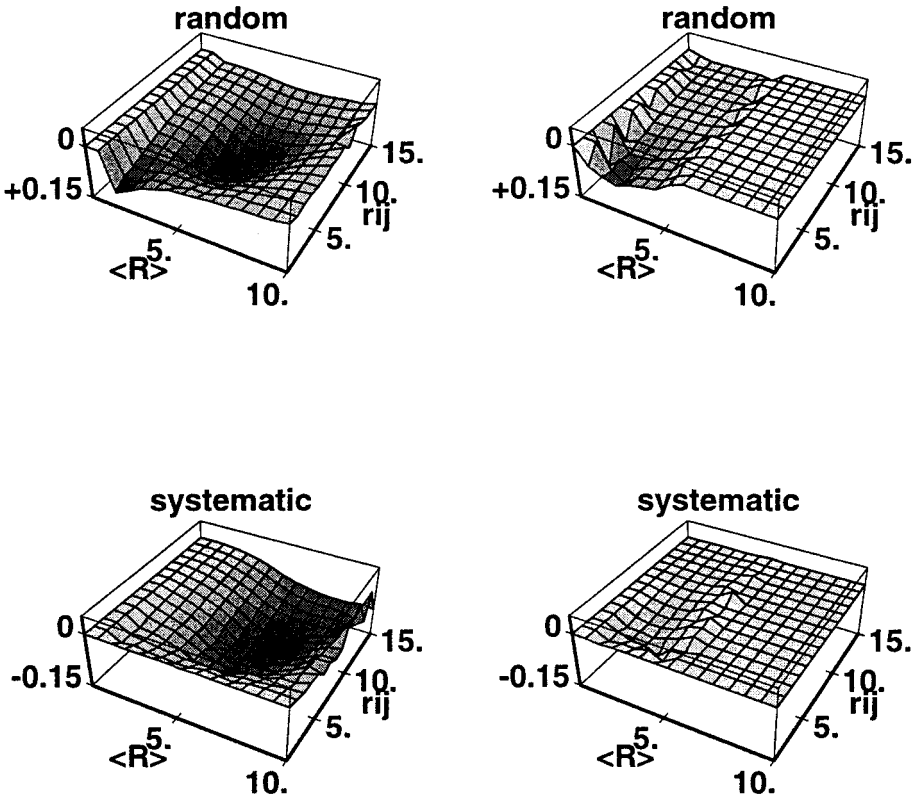

Figure 3. Deviations between "perfect" values of the effective interatomic interaction distance, as calculated by eq. (6), and the usual functional form, eq. (4). Left panel: native myoglobin, right panel: B-DNA. Vertical axis: relative systematic error $\langle (f^{\mathrm{GB}}(ij) - f_{ij}^{\mathrm{PE}})/f_{ij}^{\mathrm{PE}} \rangle$ and standard relative deviations from the mean $[\langle (((f_{ij}^{\mathrm{PE}} - \langle f_{ij}^{\mathrm{PE}} \rangle)/f_{ij}^{\mathrm{PE}})^2 \rangle]^{1/2}$ (random error) are calculated as averages over sets of atom pairs $\{ij\}$ belonging to square bins of 1×1 Å^2 in the $\langle R \rangle$, r_{ij} plane ($\langle R \rangle \equiv (R_i R_j)^{1/2}$). The random error plot has the positive axis downward to ease visual comparison with the systematic error. Distances are in Å.

inverse radii are within 0.1 Å^{-1} of their initial values (Fig. 5a). It is also instructive to see how different the self- and pair-interaction energies (self and cross terms) are in the final structure computed using the final set of perfect radii from the same energies computed with the initial set of perfect radii (Fig. 5b). Note that for most of the atoms, the differences in the self-energy are within 1 kcal/mol, and the differences are smaller for the cross terms.

Discussion

For macromolecules, the origin of much of the error of the standard GB model comes from the fact that the approach tends to underestimate the effective radii for buried atoms,[13] mostly because the standard integration procedure treats the small vacuum-filled crevices between the vdW spheres of protein atoms as being filled with water, even for structures with little interior, such as unfolded apomyoglobin (Fig. 6). Smaller R_i values yield self-energy terms that are too negative compared to the correct PE values [see eq. (3)]. The trend is the same for cross terms involving like charges, as smaller R_i's produce smaller values of f^{GB}, via eq. (4), making the factor, $-q_i q_j / f^{\mathrm{GB}}$, too negative. By the same token, however, the remaining interaction between the opposite charges is too positive, and the net effect, as seen in Table 1, is a cancellation of errors in the total solvation energy ΔG_{GB}. Obviously, this is fortuitous. It is also somewhat fragile as it depends on the predominance of opposite-charge over like-charge interactions, which may not hold over a wide variety of macromolecules and conformations.

In many applications of PE theory to macromolecules, such as the calculation of pK_a values, one is interested not in the total solvation energy but in the values of individual terms that go into it, such as the differences in self-energy terms between an ionizable group in protein versus a model compound, or the energy of interaction between the charged forms of two ionizable groups.[20] Similarly, a theory of solvation to be used in molecular mechanics must do well not only for the total solvation energy of a native structure, but also for energy differences pertaining to particular atomic motions that pick out certain energy terms over others. Thus, the utility of GB as a replacement for other higher-level theories of solvation in these and many other contexts depends not only in its performance in predicting the total ΔG_{el} but also on the accuracy of the individual terms.

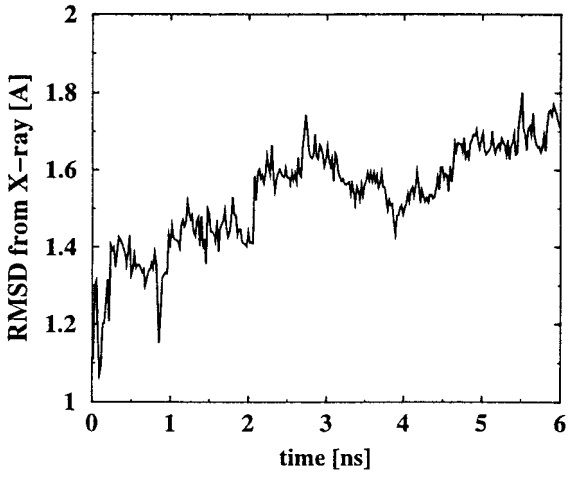

Figure 4. Back-bone RMSD from X-ray structure during a molecular dynamics simulation of thioredoxin using the perfect-radii GB model.

The present results suggest that the most important way to advance GB models is to get the effective Born radii right. If these can be chosen in a way that gives correct self-energy terms, the cross terms almost take care of themselves. The mostly random character of the remaining error in the cross terms suggests that the remaining problems have not so much to do with the particular choice of the functional form, eq. (4), as with the assumption that there can be a truly universal expression for the cross terms, simply as functions of the distance and self energies (through the R) of the interacting atoms. The results of our own efforts to improve the f^{GB} function (see Appendix) underscore this point.

The calculation of effective Born radii in the standard GB model involves two approximations. First, it is assumed that the energy density of the electrostatic field of the molecule in solvent can be approximated as the energy density of a Coulomb field, that is, the contribution of the reaction field is neglected. Second, approximations are introduced for the integration of this energy density over the molecular region, typically in order to obtain an expression involving only a sum over atom pairs.[9,10] We and others[12–14,21,22] have made efforts to improve the integration approximation, but it should be emphasized that the underlying Coulomb field approximation already introduces significant error, even if the integration can be carried out perfectly, as considerations of some simple analytical cases have shown.[8,9,21] Curiously, the Coulomb approximation tends to overestimate the effective Born radii, especially for the surface atoms, thereby partially canceling the integration error, which tends to underestimate them.[13] The net effect is, however, an underestimation of the effective Born radii, especially for the buried atoms, which are expected to have large R_i (Fig. 6).

The main motivation for the perfect-radii GB model was to provide a standard of comparison, and a test of which approximations of the standard GB theory are the main sources of error. The need to perform numerical solutions of the PE for every individual atomic charge in the molecule makes this perfect-radii variant of GB impractical for many applications. Nevertheless, it may have practical applications in cases where many energy evaluations are required, and effective radii are not expected to change much. The present MD simulations were a test case in this spirit. The results show that the deviations of the inverse radii from the beginning to the end of the simulation were not insignificant. On the other hand, the simulation was quite long, and the structure stayed fairly close to the native structure. This contrasts favorably with recent expe-

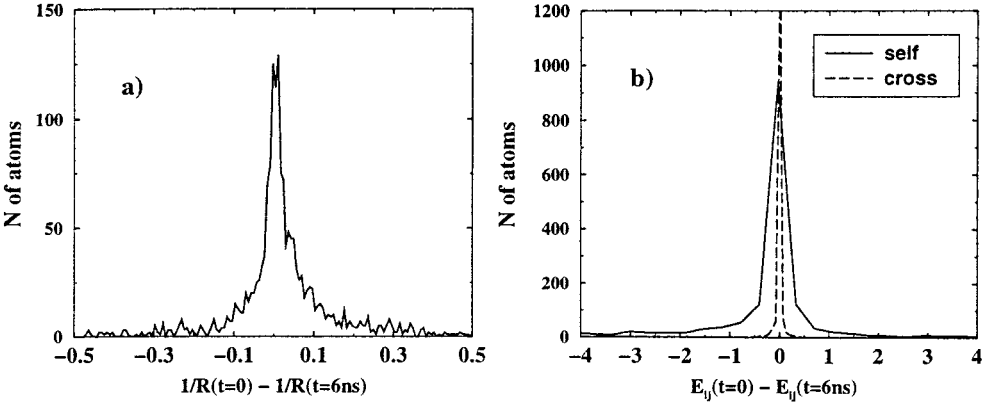

Figure 5. (a) The distribution of the difference between the inverse of the effective radii $1/R_i$ at the beginning and the end of the 6 ns MD simulation of thioredoxin. All 1654 atoms are considered. For comparison, the average value of $1/R_i$ at $t = 0$ is 0.34 Å^{-1}. (b) The distribution of the differences in self- and pair-interaction (cross-terms) energies in the final structure is computed in two different ways: either using the initial ($t = 0$) set of perfect radii, or the final set ($t = 6$ ns). For the self-term all 1654 atoms are considered, and a set of randomly selected 1654 atom pairs is chosen in the computation of the cross-terms. Note that for most of the atoms, the differences in the self-energy are within 1 kcal/mol, and the differences are smaller for the cross-terms.

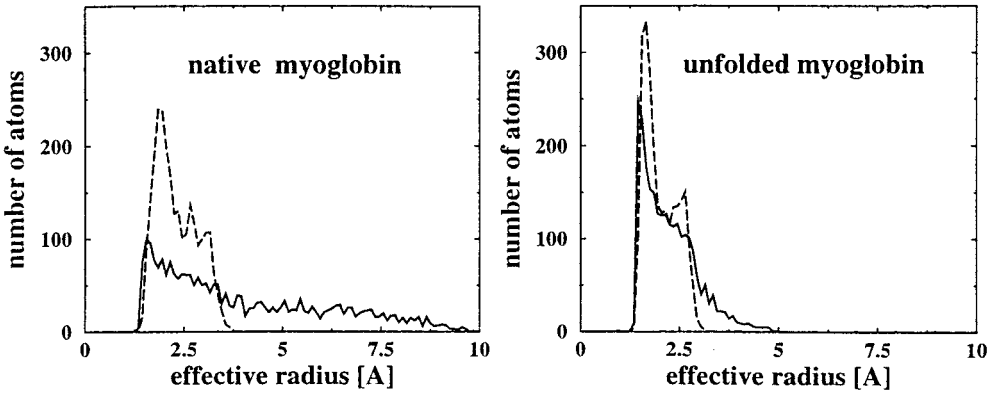

Figure 6. The distribution of effective radii for the native and acid unfolded myoglobin, showing considerable underestimation of R_i's by the standard GB model (dashed line) compared to the PE (solid line) one. The range of R_i from 0 to 15 Å is divided into 150 bins, and the number of atoms whose effective radii fall within a bin is plotted on the ordinate.

rience in the use of more conventional GB models in molecular dynamics simulations of proteins.

Acknowledgments

We thank C. L. Brooks III for helpful discussions.

Appendix

Here we attempt to find modifications to the standard f^{GB} function, eq. (4), so that when used with perfect radii, its systematic error (Fig. 3) is as small as possible, and the calculated solvation free energy contributions are as close to PE theory as possible. One clue as to how one might improve formula eq. (4) comes from an observation that, unless salt is present, one does not expect any exponentially decaying components in the cross terms, because the electrostatic potential ultimately comes from a solution of the Poisson, rather than the Poisson-Boltzmann, equation. With this in mind, we have explored replacing $\exp(-r_{ij}^2/4R_iR_j)$ in eq. (4) with more slowly varying functions of r_{ij}, having polynomial or even logarithmic behavior. For generality we also consider introducing a longer-range term, $\beta r_{ij}(R_iR_j)^{1/2}$. The following set of functions is examined:

$$f^{GB} = [r_{ij}^2 + \beta r_{ij}(R_iR_j)^{1/2} + R_iR_jS_k]^{1/2} \tag{7}$$

where S_k can be one of

$$S_\infty = \exp(-r_{ij}^2/4R_iR_j) \tag{8}$$

$$S_1 = \frac{1}{(1 + r_{ij}^2/4R_iR_j)} \tag{9}$$

$$S_{1/2} = \frac{1}{\sqrt{(1 + r_{ij}^2/2R_iR_j)}} \tag{10}$$

$$S_0 = \frac{1}{(1 + \ln(1 + r_{ij}^2/4R_iR_j))} \tag{11}$$

We test the performance of the functions above by computing the total solvation energy for a set of structures. In agreement with our physical intuition, the slowly varying functions of r_{ij}, such as eqs. (10) and (11), give the closest agreement with the PE calculations (Table 2, where the $\beta = 0$, $k = \infty$ row corresponds to the standard function). As for the long-range term, we find that although $\beta \approx 0.1$ and $k = 1$ [eq. (9)] result in a considerably smaller systematic deviation between GB and PE for native myoglobin, at least according to a plot similar to Figure 3 (not shown), the calculated solvation energies are worse than those computed using the standard function (Table 2). This shows the inherent difficulties in finding an optimal f^{GB} for all structures, which are emphasized further by Figure 7—clearly demonstrating that f_{ij}^{PE} is not even a single-valued continuous function of R_i, R_j, and r_{ij}. The above does not imply, however, that a more sophisticated form of f^{GB} would not work better in principle. Among the set of functions we have discussed above, eqs. (10) and (11) appear to be the optimum

Table 2. Comparison of the Total Electrostatic Solvation Energy between the PE and the Perfect GB Models with Different f^{GB} Functions.

	Absolute error (relative to PE) (kcal/mol)			
Function [eqs. (7–11)]	Mb (Native)	DNA	apoMb (Unfolded)	Bacteriorhodopsin
$\beta = 0$, $k = \infty$	−13	−36	−55	−22
$\beta = 0$, $k = 1$	−3	−13	−29	−10
$\beta = 0$, $k = \frac{1}{2}$	+5	+3	−11	−1
$\beta = 0$, $k = 0$	+5	+6	−9	0
$\beta = 0.1$, $k = 1$	−131	−127	−289	−209

The first row corresponds to the standard f^{GB} of eq. (4).

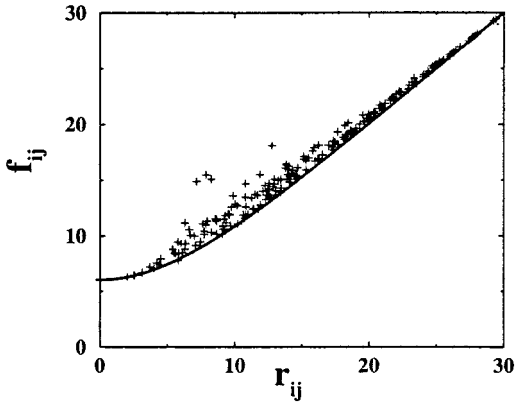

Figure 7. Comparison of the "ideal" f_{ij}^{PE} (discrete symbols) with the standard f^{GB} function (solid line). Native myoglobin coordinate set is used. Each discrete symbol of the plot corresponds to a pair of atoms i and j such that 6 Å $< R_i <$ 6.1 Å and 6 Å $< R_j <$ 6.1 Å. The solid line represents f^{GB} (R_i = 6.05 Å, R_j = 6.05 Å, r_{ij}). Both the atom-atom distance r_{ij} and f_{ij} are in Å. Note that f_{ij}^{PE} is not a single-valued continuous function of R_i, R_j, and r_{ij}, but rather a distribution of points for each r_{ij}, with a small systematic shift towards higher values compared to the standard f^{GB}.

ones because they give best solvation energies for all structures considered here. We emphasize however, that these new f^{GB} functions give closer agreement with the PE model when the "perfect" radii are used, and may not be optimal with less-than-perfect generalized Born radii. Furthermore, the difference in the systematic part of the error in the two molecules shown in Figure 3 suggests that different functional forms may be the most optimal for different macromolecular types.

References

1. Cramer, C. J.; Truhlar, D. G. Chem Rev 1999, 99, 2161.
2. Honig, B.; Nicholls, A. Science 1995, 268, 1144.
3. Beroza, P.; Case, D. A. Methods Enzymol 1998, 295, 170.
4. Madura, J. D.; Davis, M. E.; Gilson, M. K.; Wade, R. C.; Luty, B. A.; McCammon, J. A. Rev Comp Chem 1994, 5, 229.
5. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. J Am Chem Soc 1990, 112, 6127.
6. Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. J Phys Chem B 1997, 101, 1190.
7. Jayaram, B.; Liu, Y.; Beveridge, D. J. J Chem Phys 1998, 109, 1465.
8. Bashford, D.; Case, D. Annu Rev Phys Chem 2000, 51, 129.
9. Schaefer, M.; Froemmel, C. J Mol Biol 1990, 216, 1045.
10. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. J Phys Chem 1996, 100, 19824.
11. Born, M. Z Phys 1920, 1, 45.
12. Ghosh, A.; Rapp, C. S.; Friesner, R. A. J Phys Chem B 1998, 102, 10983.
13. Onufriev, A.; Bashford, D.; Case, D. J Phys Chem B 2000, 104, 3712.
14. Dominy, B. N.; Brooks III, C. L. J Phys Chem B 1999, 103, 3765.
15. Tsui, V.; Case, D. J Am Chem Soc 2000, 122, 2489.
16. Bondi, A. J Phys Chem 1964, 68, 441.
17. Connolly, M. L. Science 1983, 221, 709.
18. Tsui, V.; Case, D. Biopolymers 2001, to appear.
19. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz Jr, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179.
20. Bashford, D.; Karplus, M. Biochemistry 1990, 29, 10219.
21. Schaefer, M.; Karplus, M. J Phys Chem 1996, 100, 1578.
22. Lee, M. S.; Salsbury, F. R. J.; Brooks, C. L. I. J Chem Phys 2002, 116, 10606.
23. Caliment, N.; Schaefer, M.; Simonson, T. Proteins 2001, 45, 144.
24. Onufriev, A.; Case, D. A.; Bashford, D. Submitted for publication.
25. ftp://ftp.scripps.edu/case/beroza/pep