# Analysis of integral expressions for effective Born radii

John Mongan
*Bioinformatics Program, Medical Scientist Training Program, Center for Theoretical Biological Physics, and Department of Chemistry and Biochemistry, UC San Diego, La Jolla, California 92093-0365, USA*

W. Andreas Svrcek-Seiler
*Institute for Theoretical Chemistry, University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria*

Alexey Onufriev[a]
*Department of Computer Science and Department of Physics, Virginia Tech, 2050 Torgersen Hall, Blacksburg, Virginia 24061, USA*

Generalized Born (GB) models provide a computationally efficient means of representing the electrostatic effects of solvent and are widely used, especially in molecular dynamics (MD). Accurate and facile computation of the effective Born radii is a key for the performance of GB models. Here, we examine a simple integral prescription, $R6$, based on the exact solution of the Poisson-Boltzmann (PB) equation for a perfect sphere. Numerical tests on 22 molecules representing a variety of structural classes show that $R6$ may be more accurate than the more complex integral-based approaches such as GBMV2. At the same time, $R6$ is computationally less demanding. Fundamental limitations of current integration-based methods for calculating effective radii, including $R6$, are explored and the deviations from the numerical PB results are correlated with specific topological and geometrical features of the molecular surface. A small systematic bias observed in the $R6$-based radii can be removed with a single, transferable constant offset; when the resulting effective radii are used in the "classical" (Still *et al.*'s) GB formula to compute the electrostatic solvation free energy, the average deviation from the PB reference is no greater than when the "perfect" (PB-based) effective radii are used. This deviation is also appreciably smaller than the uncertainty of the PB reference itself, as estimated by comparison to explicit solvent. © *2007 American Institute of Physics*. [DOI: 10.1063/1.2783847]

## I. INTRODUCTION AND THEORY

Accurate representation of the effects of aqueous solvent is essential for physically meaningful simulation of biomolecules. Explicit solvation, in which the solvent is represented by a collection of discrete water molecule models, provides the most physically rigorous means of simulating solvent, but its application is limited by high computational cost. Implicit solvent models replace discrete water models with a high dielectric solvent region and additional apolar solvation terms. While this involves a greater degree of approximation in the solvent model, it can dramatically reduce computational cost, as implicit solvent models reduce the degrees of freedom over which sampling must occur, while often providing more efficient computation of each configuration.

Solvation effects are described by $\Delta G_{solv}$, the free energy of transferring a given configuration of a molecule from vacuum to solvent. $\Delta G_{solv}$ is typically decomposed into polar and nonpolar components, $\Delta G_{solv} = \Delta G_{pol} + \Delta G_{nonpol}$, where $\Delta G_{nonpol}$ is the free energy of introducing the solute molecule into solvent with no electrostatic interaction between the solute and solvent, and $\Delta G_{pol}$ is the free energy change in the system resulting from turning on solute-solvent electrostatic interactions. This paper focuses on methods for calculating polar solvation energies.

Within the limitations of a continuum dielectric solvent representation, the Poisson-Boltzmann (PB) equation provides the most physically rigorous means of calculating $\Delta G_{pol}$ and has been widely used over the past decade.[1–7] While generally more efficient than explicit solvation, solution of the PB equation for nontrivial geometries remains fairly computationally intensive and commonly used approaches do not easily yield forces, although recent advances in the numerical PB methodology have improved the situation somewhat.[1,8–11] Generalized Born (GB) models have been developed as a computationally efficient approximation of numerical solutions of the PB equation,[6,12–24] and have had particularly successful application in dynamics.[25–35]

GB models evaluate polar solvation free energy as a sum of pairwise interaction terms between atomic charges. For a typical case of aqueous solvation of molecules with interior dielectric of 1, these interactions are closely approximated by

---
[a]Author to whom correspondence should be addressed. Electronic mail: alexey@cs.vt.edu

an analytical function introduced by Still *et al.*,[13] which interpolates between Born or Onsager limits at small distances and the Coulombic limit at long distances,

$$\Delta G_{pol} \approx \Delta G_{GB} = -\frac{1}{2}\sum_{i,j}\frac{q_i q_j}{\sqrt{r_{ij}^2 + R_i R_j \exp(-r_{ij}^2/4R_i R_j)}} \times \left(1 - \frac{1}{\epsilon_w}\right), \quad (1)$$

where $r_{ij}$ is the distance between atoms $i$ and $j$, $q_i$ and $q_j$ are partial charges, and $\epsilon_w \gg 1$ is the dielectric constant of the solvent. The key parameters in the GB function are the effective radii of the interacting atoms, $R_i$ and $R_j$, which represent each atom's degree of burial within the solute. More specifically, the effective radius of an atom is defined as the radius of a corresponding spherical ion having the same $\Delta G_{pol}$ as the solvation part of self-energy of this atom in the molecule. The latter is defined as the polar solvation free energy of the molecule with partial charges set to zero for all atoms except the atom of interest. The effective radius of an atom is larger than the intrinsic radius of its atom sphere because of the descreening effects of surrounding atoms, reducing the extent to which the atom charge is screened by solvent. The principal way to compute an $R_i$ is to solve for it in Eq. (1) with $i=j$, provided that the value of $\Delta G_{pol}^i$ has already been computed. The most accurate way of computing $\Delta G_{pol}$ within the implicit solvent framework is by solving the PB equation directly, which in practice requires numerical solution. This yields the so-called "perfect" effective radii: their use in Eq. (1) was shown to result in $\Delta G_{pol}$ being close to the corresponding PB values.[36] Although the perfect radii are impractical in most applications due to the computational demands of obtaining them, they have served as a natural point of reference for assessing the accuracy of GB, since current GB models are an approximation to the more fundamental formalism of the PB equation. In particular, they serve as a measure of error of other, fast and practical, but approximate methods for estimating effective radii.

In practice, $\Delta G_{pol}^i$ and the corresponding effective radius for each atom are generally calculated by first approximating the electrostatic energy density $(1/2)\mathbf{E}(\mathbf{r})\cdot\mathbf{D}(\mathbf{r})$ due to the atom of interest by some reasonably simple expression and then integrating over the appropriate volume;[5,14,15,22,33,37,38] equivalent formulations based on surface integrals have also been proposed.[20,24]

Historically, the first, but still widely used, approximation of that nature is the Coulomb field approximation (CFA), $\mathbf{E}(\mathbf{r})\cdot\mathbf{D}(\mathbf{r}) \propto q_i|\mathbf{r}|^{-4}$, which assumes that the electric field generated by the atomic point charge is unaffected by the nonhomogeneous dielectric environment created by the solute, so that the field has the form described by Coulomb's law. With this assumption, the CFA inverse effective radius ($\alpha_4$) is

$$\alpha_4 = R_i^{-1} = \frac{1}{4\pi}\int_{ext}\frac{1}{|\mathbf{r}-\mathbf{r}_i|^4}dV = \rho_i^{-1} - \frac{1}{4\pi}\int_{r>\rho}|\mathbf{r}|^{-4}dV, \quad (2)$$

where in the first expression the integral is taken over the region outside of the molecule. The second formula in the equation above is often used for computational convenience:[37] the origin is moved to the atom of interest, and the integration region is the interior of the molecule outside of the atom's van der Waals (VDW) radius $\rho_i$.

The CFA is exact for a point charge at the center of a spherical solute, but it overestimates the effective radii for molecular geometries[22] as well as for spherical regions when the charge is off center;[39] in the latter case an exact calculation[21] shows that as the point charge approaches the surface, the value of $R_i$ can be off by a factor of 2. Some of the success of early GB models on small molecules may be attributed to fortuitous cancellation of errors in effective radius calculations between the overestimates of a CFA-based integrand and the underestimates of a VDW based region of integration.[33] The problems with the CFA were well known for quite some time, and effective approximations based on empirical corrections to the CFA (Refs. 22, 36, and 40) were proposed, albeit without apparent physical justification. These improvements typically take the form of a simple linear or a rational combination of correction terms such as

$$\alpha_N = \left(\frac{1}{4\pi}(N-3)\int_{ext}\frac{dV}{|\mathbf{r}-\mathbf{r}_i|^N}\right)^{1/(N-3)}, \quad (3)$$

where $N > 4$.

Lee *et al.* were the first to propose an accurate alternative to the CFA along these lines, an expression involving $\alpha_4$ and $\alpha_5$,[22] and later an even more accurate expression was put forth based on $\alpha_4$ and $\alpha_7$, respectively.[38] These works represented a major advance for the GB field. Later, massive fits to numerical PB calculations resulted in expressions involving even larger number of terms such as Eq. (3), see, e.g., Ref. 24 who used surface-integral analogs of Eq. (3).

Despite the greatly improved accuracy of such empirical expressions over the CFA in computing the effective Born radii, problems remained. From a practical standpoint, keeping multiple correction terms results in the need to compute multiple independent volume integrals which leads to decreased computational efficiency, especially compared to the CFA-based models.[41] Fundamentally, the boundaries of applicability are rarely well defined for an approach that lacks clear physical grounds. Likewise, it is not clear whether these types of corrections can be improved further. A physically well-grounded alternative would be desirable, but can it be computationally effective and accurate for complex biomolecular geometries? And fundamentally, how far should one try to go in improving the accuracy of computing the effective Born radii? After all, Still *et al.*'s GB model itself is an approximation, even if "perfect" radii are used.

Grycuk has recently shown[39] that, in the conductor limit, *a single* volume integral provides an *exact, rigorously derived* expression for the solvation part of the self-energy

$\Delta G_{\text{pol}}^i$ of a charge $q_i$ at any location inside a perfect sphere surrounded by infinite dielectric medium,

$$\lim_{\epsilon_w \to \infty} \Delta G_{\text{pol}}^i = -\frac{q_i^2}{2}\left(\frac{3}{4\pi}\int_{\text{ext}}\frac{dV}{|\mathbf{r} - \mathbf{r}_i|^6}\right)^{1/3}. \tag{4}$$

The formula immediately provides a unique prescription for computing the effective radii,[39] referred to in this paper as "$R6$,"

$$\alpha_6 = R_i^{-1} = \left(\frac{3}{4\pi}\int_{\text{ext}}\frac{dV}{|\mathbf{r} - \mathbf{r}_i|^6}\right)^{1/3}$$

$$= \left(\rho_i^{-3} - \frac{3}{4\pi}\int_{r > \rho}|\mathbf{r}|^{-6}dV\right)^{1/3}. \tag{5}$$

Since our main focus here is the accuracy of the effective Born radii, we use the conductor limit $\epsilon_w \to \infty$ throughout to avoid complicating the picture with issues arising from the limitations of the canonical GB model Eq. (1) in the finite $\epsilon_w$ regime.[42] A computationally effective analytical alternative to the GB model that addresses this specific problem was proposed,[43] and can be used in conjunction with any set of accurate effective Born radii.

For technical reasons such as computational efficiency, it may sometimes be of advantage to replace the volume integrals in Eq. (3) by the appropriate surface integrals.[20,21,24] Here, we list the appropriate formulae for completeness' sake. Using Gauss's theorem and the identity

$$\frac{1}{|\mathbf{r}|^N} = -\frac{1}{N-3}\nabla \cdot \frac{\mathbf{r}}{|\mathbf{r}|^N}, \tag{6}$$

one obtains, for $\alpha_6$,

$$\alpha_6 = R_i^{-1} = \left(-\frac{1}{4\pi}\oint_{\partial V}\frac{\mathbf{r} - \mathbf{r}_i}{|\mathbf{r} - \mathbf{r}_i|^6}\cdot d\mathbf{S}\right)^{1/3}, \tag{7}$$

where $\partial V$ denotes the molecular surface and $d\mathbf{S}$ denotes the vector-valued infinitesimal surface element. Given a triangulation of the molecular surface, this leads to

$$\alpha_6 \approx \left(-\frac{1}{4\pi}\sum_{\text{triangles } k}\frac{(\mathbf{c}_k - \mathbf{r}_i)\hat{\mathbf{n}}_k S_k}{|\mathbf{c}_k - \mathbf{r}_i|^6}\right)^{1/3}, \tag{8}$$

with $c_k$ the position of the centroid of triangle $k$, $S_k$ its area, and $\hat{\mathbf{n}}_k$ the unit vector orthogonal to triangle $k$ and pointing towards the inside of the solute.

The potential usefulness of the $R6$ prescription for practical calculations of effective radii depends critically on how sensitive the accuracy of the prescription is to inevitable deviations of realistic molecular surfaces from perfect spherical. Before we address this question in detail on a test set of biomolecules, we would like to have a rough estimate of what to expect. Exact solutions of the Poisson problem on simple nonspherical geometries can provide such an estimate; here, we consider an example shown in Fig. 1. The image charge solution yields an exact expression for $\Delta G_{\text{pol}}$ of the charge $q$,
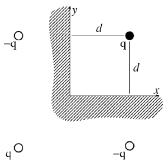


FIG. 1. Example of a nonspherical dielectric boundary for which the Poisson problem has an exact solution. A charge $q$ (black circle) is located at $(d, d, 0)^T$ on the bisector inside the 90° angle. The three outside quadrants $[(x < 0) \vee (y < 0)]$ are filled with a conductor. The reaction field potential inside the quadrant $[(x > 0) \wedge (y > 0)]$ is given exactly by the three image charges (empty circles).

$$\Delta G_{\text{pol}}(d) = \frac{q^2}{2d}\left(1 - \frac{\sqrt{2}}{4}\right). \tag{9}$$

The exact inverse radius can be obtained as

$$\alpha_{\text{corner}} = \frac{1}{d}\left(1 - \frac{\sqrt{2}}{4}\right) \simeq \frac{1}{d}\cdot 0.64645. \tag{10}$$
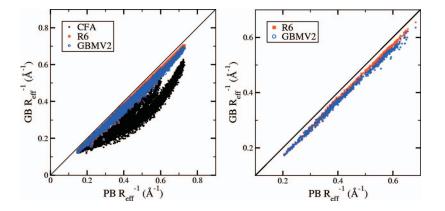
At the same time, the inverse radius via the $R6$ expression (5) is

$$\alpha_6 \simeq \frac{1}{d}\cdot 0.61753. \tag{11}$$

It is reassuring, and perhaps even a bit surprising, that for this "very" nonspherical geometry the $R6$ expression approximates the exact result within just a 4% error.

Comparison of the performance of the various integral expressions for calculating effective Born radii based on results in the literature is difficult because these results are typically drawn from complete GB models which include methods of performing integration coupled with a molecular surface definition (often implicit and often different from what one may think of as "consensus"). Variations in integration scheme and surface definition between models may have a larger effect on results than differences in the underlying integral expression, and there is no clear way to separate the effects of the integral expression from the other parts of the model when comparing results. Here, these complications are avoided through the following strategy. We use the classical Connolly molecular surface, which implies a sharp boundary between solute and solvent. The performance of integral expressions is compared directly through results produced by evaluating them using exactly the same high-resolution numerical quadrature over the molecular volume corresponding to the Connolly surface.

This paper is structured as follows. We begin by outlining a rigorously exact method of computing the effective radii for a spherical molecule. We then show that despite its simplicity, the method is still more accurate for typical biomolecular shapes than its empirical counterparts. Errors in the computed effective radii and their physical origins are discussed. By using the "perfect" (PB-based) effective radii as reference, we can separate the errors in the total GB sol-

FIG. 2. (Color) Comparison of approximate GB effective with the "perfect" (PB-based) radii for 19 small protein structures (left) and B-form DNA (right). GB effective radii were calculated using high precision numerical volume integration of three different expressions: the CFA, GBMV2, and the $R6$ expression shown in Eq. (5). For the DNA structure (right panel), only the two best performing expressions ($R6$ and GBMV2) are shown.

vation free energy into two components: one due to inaccuracies of the approximate effective radii and the other due to inherent errors of Still's GB formula itself. These are compared with the errors of the fundamental PB model itself, estimated from a comparison to the explicit solvent treatment as reference. An analysis of the relative sizes and likely origins of these errors allows us to make suggestions for the future development of GB models and methods for computing the effective Born radii.

## II. RESULTS AND DISCUSSION

### A. Comparison of integral expressions for the effective radii

Several integral expressions for calculating effective radii were applied to a variety of biomolecular geometries using the same very high-resolution quadrature of the integrals over molecular volumes (see Sec. III). These expressions are the Coulomb field approximation (CFA) described in Eq. (2), the "$R6$" expression shown in Eq. (5), and the GBMV2 parametrization[38] of the correction terms from Eq. (3). The GBMV2 prescription is a "correction" to the CFA in that it is based on a linear combination of $\alpha_4$ (CFA) and $\alpha_7$, a detailed analysis of this formula will be given in Sec. II C. Scatter plots comparing the approximate effective Born radii computed by these prescriptions to PB effective radii calculated using exactly the same molecular surfaces are presented in Figs. 2 and 3.

Note that the CFA effective radii presented here are uniformly larger than the PB effective radii, in agreement with previous theoretical work.[21,39] This result may appear to be in conflict with plots in some other publications, such as

Onufriev et al.,[33] which show underestimated CFA-based effective radii, especially as the effective radii become large. The underestimation in those plots is due to comparison of CFA effective radii calculated with a van der Waals-type surface to PB effective radii calculated with a Connolly surface; here, the same Connolly surface is used for all calculations, so there is no underestimation.

Both the $R6$ expression of Eq. (5) and the corrected CFA expression employed by GBMV2 produce radii that are in closer agreement with PB effective radii than the uncorrected CFA. It is worth mentioning that the GBMV2 is not the only expression that has been successful in improving upon the CFA by using a combination of $\alpha_N$ terms.[44] However, the effective radii computed by these models do not appear to offer a clear-cut accuracy advantage over GBMV2 on our test set (results not shown). We have therefore chosen GBMV2 to represent the class of expressions that use a combination of $\alpha_N$ to go beyond the CFA; other expressions of similar nature will not be considered here further. As seen in Table I, $R6$ outperforms GBMV2 by a small but statistically significant margin. The surprising similarity in results of these two expressions with rather different functional forms is explored below. Apart from the small improvement in accuracy, the $R6$ expression has other advantages over the GBMV2 expression. Perhaps most important is the fact that the $R6$ has a clear theoretical basis and a limiting case where it becomes exact, while GBMV2 is apparently empirical. In addition, the $R6$ should be more computationally efficient, as it involves only a single integral, rather than two. Furthermore, the $R6$ integral is likely to be somewhat easier to compute than the second GBMV2 integral, as it involves a lower-order integrand. Indeed, we encountered numerical stability
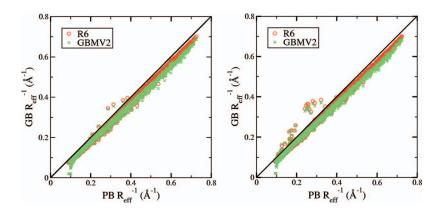


FIG. 3. (Color) Comparison of $R6$ and GBMV2 effective radii calculated for thioredoxin (left) and lysozyme (right). Substantial underestimates of some small radii by the GBMV2 expression (lower left of each plot) are due to numerical instability (see text). Inverse radii exhibiting these instability problems were changed to 0.02 Å$^{-1}$.

TABLE I. RMS errors of inverse GB effective radii relative to inverse PB-based (perfect) effective radii in $Å^{-1}$. The level of accuracy for each method is remarkably consistent across four disparate systems: a collection of small proteins, B-form DNA, and two larger proteins. For each system, F-tests show the improvements of GBMV2 over CFA and of $R6$ over GBMV2 to be significant with $p \ll 10^{-10}$. The F-test gives the probability that two samples are drawn from populations with the same standard deviation.

|                | CFA   | GBMV2 | $R6$  |
|----------------|-------|-------|-------|
| Small proteins | 0.135 | 0.037 | 0.029 |
| B-DNA          | 0.134 | 0.035 | 0.027 |
| Thioredoxin    | 0.134 | 0.036 | 0.025 |
| Lysozyme       | 0.134 | 0.035 | 0.025 |

problems when using the GBMV2 expression to calculate effective radii for deeply buried hydrogen atoms, where very small errors in the second integral lead to very large effective radii or arithmetic exceptions.

## B. Sources of error in integral calculation of effective radii

Understanding of the limitations of the $R6$ expression for calculating effective radii may be gained by examining outlier atoms: those that have the largest differences between $R6$ and PB effective radii. As seen in Fig. 4, all these outlier atoms are located near local concavities in the dielectric surface. Interestingly, the sign of the error is determined by whether the high dielectric region forming the surface concavity is continuous with the bulk solvent (surrounding high dielectric region): exterior surface invaginations lead to underestimates of the inverse effective radii, while completely enclosed or buried solvent pockets lead to overestimates of the inverse effective radii.

The overestimates of inverse effective radii near buried solvent pockets imply that the contributions of these pockets to the solvation free energy of nearby atoms are overestimated by the $R6$ expression, that is, these atoms are less solvated in reality than the $R6$ prescription predicts. This overestimate can be explained by analyzing the solvation effects that an enclosed pocket of solvent has on a nearby atom that lies in the low dielectric region of the solvent (such an atom is shown as red dot in Fig. 4). The favorable solvation free energy is produced by induction of complementary surface charge on the face of the solvent region nearest to the atom. However, the net induced charge on any contiguous dielectric region must be zero, so this complementary surface charge is balanced by an opposite-signed surface charge (i.e., the same sign as the charge on the atom) on the far face of the dielectric region. For an enclosed pocket of solvent, such as shown in Fig. 4, the far face is, in fact, fairly close to the atom in question, so the energetically unfavorable interaction between the surface charge there and the atom is significant in decreasing the magnitude of the solvation free energy provided by the pocket. This topology-specific cancellation effect is not accounted for by the $R6$ integration prescription, or by any combination of $r^{-N}$ integral expressions in Eq. (3) for that matter—the specific connectivity information needed to account for the above physical effect is simply absent
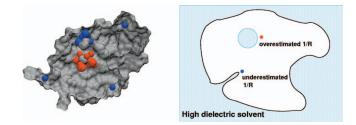


FIG. 4. (Color) Effects of solute's local geometry and topology on the accuracy of the $R6$ approximation. Left panel: Cut-away view of hen egg-white lysozyme illustrating positions of outlier atoms where the $R6$ prescription results in largest errors. The molecular surface is represented in gray, atoms where Eq. (4) overestimates inverse effective radii by more than 0.02 $Å^{-1}$ are red spheres, and atoms where inverse effective radii are underestimated by more than 0.05 $Å^{-1}$ are blue spheres. All the red spheres are clustered about a solvent pocket that is disconnected from the bulk solvent. The blue spheres are located near narrow "local" surface concavities and invaginations that communicate with bulk solvent; notably these are not found near the large-scale (nonlocal) cleft on the right side of the molecule. Right panel: A 2D schematic illustrating these trends.

from integrals over uniform space. As seen in Fig. 3, the number of atoms close enough to a solvent pocket to have their inverse effective radii significantly overestimated is quite small, so the impact on total solvation energy is likely to be fairly negligible unless these atoms carry significant charge. It is possible, though, that the local oversolvation effects of small solvent pockets could lead to physically unrealistic formation of such pockets near buried polar groups in molecular dynamics (MD) runs. These caveats, however, pertain to any integral-based method for calculating effective Born radii.

Nearly all atoms show some degree of underestimation of inverse effective radii, as seen in Fig. 2. The cause of this underestimation is identified by examining the atoms with the most dramatically underestimated inverse effective radii. These atoms are found near long and narrow concavities or invaginations of the exterior molecular surface.
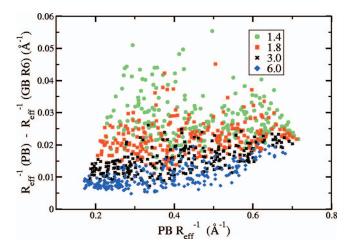


FIG. 5. (Color) Underestimation of the $R6$ inverse effective radii relative to "perfect" PB inverse effective radii diminish as the molecular surface is smoothed by increasing the probe radius from 1.4 to 6.0 Å (see Fig. 6). Results shown are for the beta hairpin structure, see Sec. III, with magnitude of underestimation on the vertical axis and PB effective radius on the horizontal axis. Mean underestimates and 95% confidence intervals for the four probe radii are $2.76 \pm 0.08 \times 10^{-2}$, $2.23 \pm 0.06 \times 10^{-2}$, $1.54 \pm 0.05 \times 10^{-2}$, and $1.02 \pm 0.04 \times 10^{-2}$ $Å^{-1}$.
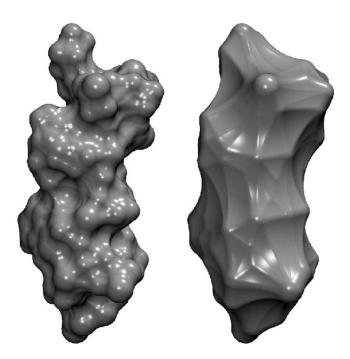
FIG. 6. Molecular surface of the beta hairpin showing the smoothing effects of increasing the solvent probe radius from 1.4 Å (left) to 6.0 Å (right). Molecular surfaces are calculated with MSMS (Ref. 52), renderings performed with VMD (Ref. 48).

The source of the concavity-induced underestimate of the $R6$ inverse radii and, therefore, solvation self-energies can be best understood by considering the conductor limit $\varepsilon_w \to \infty$. Suppose a partial charge in the low dielectric region is positioned near the tip of a conductor "inlet" (such an atom is shown as blue dot in Fig. 4). As is known from classical electrostatics, the induced opposite charge that builds up on the tip creates a very strong electric field around it. The effect of this field is that the nearby partial charge in the low dielectric region will be better solvated than it would be in the vicinity of a surface of moderate curvature. The $R6$ expression is derived for a spherical surface that has uniform radius of curvature and, thus, does not account for this effect of high local curvature. This geometry-specific underestimation of the solvation self-energy is inherent to any integral prescription for effective radii.

Since some degree of local concavity or invagination is found across any biomolecular surface, it is reasonable to expect that the small systematic underestimation of the inverse effective radii within the $R6$ prescription seen in Fig. 2 is a cumulative effect of underestimation of the solvation effect of dielectric in the concavities found across the entire surface.

Evidence supporting this idea is provided in Fig. 5, which shows a steady and significant decrease in the magnitude of underestimation as the molecular surface is made progressively smoother (by increasing the probe radius) and local concavities diminish. It may be suggested that this increase in accuracy is attributable to the molecular surface becoming more spherical and thus closer to the case in which Eq. (5) is exactly correct. However, the renderings in Fig. 6 show that even for the largest probe radius, the molecular surface of the beta hairpin is decidedly nonspherical, sug-

gesting that the improved results are, in fact, due to elimination of local concavities. Interestingly, Fig. 5 shows that the largest magnitude improvement is seen for the most deeply buried atoms (those with the smallest inverse effective radius, on the left of the plot). This might be because increasing the probe radius rapidly eliminates the "deep inlets" near buried atoms, thus removing the main reason for the underestimation of the inverse effective radii within the $R6$ integral prescription.

While the geometrically and topologically based sources of error discussed above are likely present in the CFA, the inherent error of the CFA itself is large enough in comparison to these more subtle errors that the latter cannot be appreciated within the framework of the CFA.

## C. The accuracy of some earlier integral-based prescriptions

As we have just seen, the $R6$ prescription, although being exact only for a sphere, provides an estimate of the effective Born radii for realistic molecular geometries that is more accurate than prescriptions that are not exact for a sphere. The accuracy of some of them, e.g., GBMV2, comes close to that of $R6$, while the other ones examined here are less accurate. As we will see below, this somewhat puzzling accuracy of empirical corrections based on combinations of several $\alpha_{N \neq 6}$ can be easily explained by examining how well these prescriptions approach the exact $\alpha_6$ expression on a sphere. To this end, consider a perfectly spherical cavity of radius $D$ embedded inside an infinite conductor region; a point charge $q$ located at distance $x$ from the cavity's center. Denoting $p = x/D$, we obtain for $\alpha_N$ of Eq. (3),

$$\alpha_4(p) = \frac{1}{2D}\left( \frac{1}{1-p^2} + \frac{1}{2p} \log \frac{1+p}{1-p} \right), \tag{12}$$

$$\alpha_5(p) = \frac{1}{D(1-p^2)}\left( 1 - \frac{p^2}{6} \right)^{1/2}, \tag{13}$$

$$\alpha_6(p) = \frac{1}{D(1-p^2)}, \tag{14}$$

$$\alpha_7(p) = \frac{1}{D(1-p^2)}\left( 1 + \frac{2}{3}p^2 - \frac{1}{6}p^4 \right)^{1/4}, \tag{15}$$

or, in general (for $N > 4$),

$$\begin{aligned} \alpha_N(p) = {}& \frac{1}{D(1-p^2)} \times (2(N-2)(N-4)p)^{-1/(N-3)} \\ & \times ((1+p)^{N-3}((N-3)-p) \\ & - (1-p)^{N-3}((N-3)+p))^{1/(N-3)}. \end{aligned} \tag{16}$$

Figure 7 shows why various combinations of $\alpha_N$'s can be found that lead to more accurate effective Born radii than those obtained from the CFA$= \alpha_4$: For $N > 6$, $\alpha_N$ is overestimated (the effective Born radii are underestimated), while for $N < 6$, $\alpha_N$ is underestimated relative to the exact value $\alpha_6$. Apparently, appropriate combinations of "overshoots" and "undershoots" can lead to large cancellations of errors and an overall reasonable agreement with the exact result. For ex-
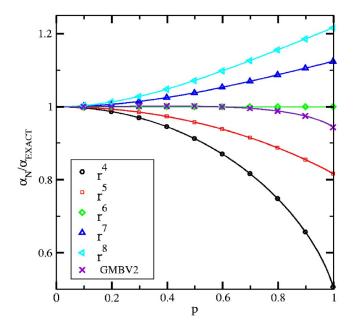
FIG. 7. (Color online) Relative deviations from the exact values of inverse effective radii $\alpha_N$, Eq. (16), with $N=4$, 5, 6, 7, and 8, computed for a perfect sphere. The GBMV2 expression is based on a combination of $\alpha_4$ and $\alpha_7$, Eq. (17). Distance from the sphere's center (dimensionless) is shown along the $x$ axis.

ample, one of the formulae proposed by Lee *et al.*,[38] termed as GBMV2,

$$\left(1 - \frac{\sqrt{2}}{2}\right)\alpha_4(p) + \frac{\sqrt{2}}{2}\alpha_7(p), \qquad (17)$$

does approximate the exact value $\alpha_6(p)$ very closely (see Fig. 7). This closeness illustrates the good performance of prescriptions that interpolate the effective Born radii via a combination of $\alpha_N$ terms.

## D. Further improvements and limitations

Given that the magnitude of the average underestimate of the $R6$ inverse effective radii is fairly constant across the wide variety of biomolecular shapes tested here, it seems reasonable to expect that the accuracy of the approximation may be further improved by adding a constant offset value to each radius to account for the average effect of surface invaginations. The strategy leads to the following prescription:

$$R_{R6*}^{-1} = R_{R6}^{-1} + 0.028 \text{ Å}^{-1}, \qquad (18)$$

where the constant is chosen to minimize the rms deviation between the "corrected" inverse radii $R_{R6*}^{-1}$ and the perfect inverse radii for the training set of 19 protein structures used in Fig. 2. It should be emphasized that this particular offset value is only appropriate for the typical solvent probe radius of 1.4 used here to compute the "uncorrected" $R6$ effective radii. This statement is likely to hold true for many other effective radii prescriptions that use empirical offsets to improve the accuracy of the radii: the offset values are optimal only for the specific way the molecular surface was computed in each reference model. The accuracy of the GB model in Eq. (1) with the "corrected" $R6^*$ effective radii from Eq. (18) can be tested by comparing the computed

TABLE II. A comparison of relative errors in $\Delta G_{pol}$ computed by various methods. First column: the error of the GB($R6^*$) model relative to the PB reference is computed as arithmetic mean of $|(\Delta G_{pol}(GB) - \Delta G_{pol}(PB))/\Delta G_{pol}(PB)|$ over all 22 molecular structures considered here. Second column: the error of the GB(perfect) model with perfect (PB-based) radii computed in the same manner. Third column: the relative error of the PB relative to explicit solvent treatment is estimated as arithmetic mean of $|(\Delta G_{pol}(PB) - \Delta G_{pol}(FEP))/\Delta G_{pol}(FEP)|$, $\Delta G_{pol}(PB)$ and $\Delta G_{pol}(FEP)$ values for a set of small proteins were taken from Table 5 ("Bondi radii" column) of Ref. 45. Fourth column: the variability of the PB procedures is estimated as the average of $|(\Delta G_{pol}(APBS) - \Delta G_{pol}(MEAD))/\Delta G_{pol}(MEAD)|$. The parameter values, such as finite-difference grid spacing, used here to perform computations with MEAD and APBS PB solvers, are given in the Sec. III. In all cases, the statistical margin of error ($\pm$) is computed as standard deviation.

| method reference | GB($R6^*$) PB | GB(perfect) PB | PB explicit solvent | PB PB |
|---|---|---|---|---|
| relative error, % | 1.13±0.97 | 1.28±0.99 | 3.73±3.18 | 0.50±0.17 |

$\Delta G_{pol}$ with the PB reference, Table II. The corresponding relative error in $\Delta G_{pol}$ of the $R6^*$ prescription is 1.13%, averaged across the 22 structures represented in Table I. When the 19 small proteins used to fit the offset in Eq. (18) are excluded, the relative error decreases to 0.33%, suggesting that this offset is likely applicable to a wide range of molecular shapes and sizes. To put the size of the GB($R6^*$) errors in perspective, one can compare them to a recent estimate[45] of the relative error of the PB standard itself, 3.73%, relative to the explicit solvent treatment. For completeness, in the fourth column of Table II we list the relative difference between $\Delta G_{pol}$ computed by two widely used PB solvers at high grid resolution.

With the systematic offset between $R6$ and perfect radii eliminated by Eq. (18), the average total error in $\Delta G_{pol}$ computed by the GB approximation based on the $R6$ becomes, within the statistical margin of uncertainty, equal to the error in $\Delta G_{pol}$ based directly on the perfect radii. Also, the errors of $\Delta G_{pol}$ computed via the GB($R6^*$) model appear to be significantly smaller than the corresponding errors in PB relative to the explicit solvent treatment. These observations suggest that, at least as far as deviations in the total molecular $\Delta G_{pol}$ are concerned, there is no pressing need for further improvement of the $R6^*$ effective radii prescription. Nevertheless, there may be specific cases in which further improvement is desirable. For example, in MD simulations, it is not just the overall $\Delta G_{pol}$ that matters, but the components of the free energy for each atom that determine the forces acting on them. As discussed previously, significant deviations between the GB($R6$) and the PB effective Born radii may exist for a small subset of atoms near specific regions of solvent space, e.g., internal cavities. However, improving the accuracy of the effective radii beyond $R6^*$ is, by itself, unlikely to fix the problem. This is because the "canonical" GB Eq. (1) has the same underlying physical basis as the $R6$ approximation for the effective radii, the Kirkwood spherical model,[39] so it may be necessary to go beyond the assumptions of a spherical solute to reduce these errors. Furthermore, as some of the solvent regions where the canonical GB model breaks down represent small water clusters or even single water molecules more or less separated from the bulk solvent, they

are likely cases where the fundamental PB approximation of solvent as a uniform dielectric breaks down. Therefore, improving agreement with PB effective radii for these atoms may well be an overfitting to the errors of PB rather than a true improvement in physical realism. Finally, it should be kept in mind that in practical GB models, the accuracy of the $R6^*$ prescription may be significantly reduced by approximations made by fast routines employed to compute the integral in Eq. (5) or its equivalents.

The problems to be overcome in any practical implementation of the $R6$ are nontrivial. For example, the excellent accuracy of the complete GBMV2 effective radii prescription draws on both an integral expression for the effective radii that is nearly as accurate as $R6$ and an innovative and effective numerical method for defining the dielectric boundary that closely approximates the molecular surface. Unfortunately, the computational efficiency of this method is considerably lower compared to routines used by many "fast," but less accurate GB models.[41] An alternative approach, embodied by the GBn (Ref. 46) model we have recently developed, involves analytic approximations of the integral over atoms and the solvent-excluded regions between atoms and is far more computationally efficient (though perhaps less accurate) than that employed by GBMV2. While the current GBn model was developed with the CFA approximation in mind, we believe that the approach it takes for analytical integration is also promising in combination with the much more accurate $R6$. The use of a surface-integral alternative of Eq. (5) may also help. A naive implementation of Eq. (8) based on a triangulation of the molecular surface provided by the program SURF (Ref. 47) [available with VMD (Ref. 48)] resulted in good agreement between effective radii calculated by integrating either over the molecular interior or surface. One might expect that integration over the molecular surface will show better scaling behavior for large systems, but a detailed comparison of performance and accuracy is nontrivial and beyond the scope of the present work.

## III. METHODS

The test structures used here are randomly selected from a larger set of representative protein structures from Feig *et al.*,[41] the selection criterion being that the compounds are small enough to allow for high-resolution grid computations. Their PDB IDs are 1az6, 1bh4, 1bku, 1brv, 1byy, 1cmr, 1dfs, 1dmc, 1eds, 1fct, 1fmh, 1fwo, 1g26, 1ha9, 1hzn, 1paa, 1qfd, 1qk7, 1scy, 2trx, and 2lzt. We have chosen chain "A" or "model 1" where appropriate. The assignment of partial charges, protonation states, etc. are described in Ref. 41. In addition, we used a canonical B-DNA 10 base pair structure from Ref. 27 and $\beta$ hairpin PDB ID 2gb1. The Bondi radius set was used for all molecules. The random selection has resulted in a fairly representative sampling of various structural classes and charge states in the test set. The total charge of the structures varies from $-18$ (B-DNA) to $+9$ (lysozyme), with most of the structures (17) falling in the range from $-4$ to $+4$. The structural composition of the proteins is as follows: 7 mostly $\alpha$ helical, 4 mostly $\beta$ sheet, 5 roughly equal

mix of $\alpha / \beta$, and 5 mostly disordered. The size of most of these proteins is about 30 amino acids, although two of them are larger: 2trx and 2lzt have 108 and 129 residues, respectively.

"Perfect" PB equation-based radii were calculated using PB treatment as implemented in APBS 0.4.0.[11] A separate calculation was performed for each atom of each molecule. In each calculation, the partial charge of the atom of interest was set to 1, while partial charges of all other atoms were set to zero. A 129-point cubic grid centered on the atom of interest was used to discretize the problem. Multiple Debye-Huckel boundary conditions were used for the initial grid, which was sufficiently large that no portion of the molecule was closer than 4 Å to the edge of the grid. Each focusing step halved the grid spacing, while maintaining the same number of grid points. Focusing step boundary conditions were derived from the potential calculated on the immediately preceding grid. Focusing continued until the grid spacing reached 0.1 Å. Except where otherwise indicated, all calculations used a nonsmoothed molecular surface definition with a probe radius of 1.4 Å and a surface probe point density of 50. A four-level finite-difference multigrid solver was employed in conjunction with the linearized Poisson-Boltzmann equation (which reduces to the Poisson equation since ion concentrations were zero). Charge was discretized using cubic B-splines. All solvated calculations used a dielectric constant of $\epsilon_w = 1000$ to mimic the conductor limit $\epsilon_w \rightarrow \infty$, and therefore, avoid masking the geometry-specific deficiencies of the standard GB model by its inaccuracies arising from finite $\epsilon_w$.[42] The dielectric constant of the solute region was set to 1; a parallel set of reference calculations was performed with a spatially uniform dielectric constant of 1 to determine the gas-phase charge discretization reference energy. The self-energy of each atom was calculated by subtracting the reference energy from the solvated energy from the most focused grid. Radii were calculated from self-energies using the Born equation. MEAD 2.2.5* with double precision and otherwise default parameter settings is used as reference PB solver in Table II. The dielectrics are as described above. Six focusing steps are used with the coarsest cubic grid having 81 points in each direction and 3.2 Å grid spacing, and the finest grid of 315 points in each direction and 0.1666 Å spacing.[49]

The integrals in the effective radius expressions described in Eqs. (2), (3), (5), and (17) were evaluated numerically. Integration was performed in Cartesian coordinates with rectilinear volume elements. Integrands were evaluated at the center of each volume element; a volume element was included if its center was outside the atom of interest and inside the molecular (Connolly) surface, as defined by APBS 0.4.0 using the same probe radius and point density as in the PB calculations. Because the values of these integrands change most quickly near the origin, the volume element is expanded as the integration moves away from the origin. Volume elements measure 0.015 Å on a side within 3 Å of the origin, 0.05 Å between 3 and 15.5 Å, 0.2 Å between 15.5 and 75.5 Å, and 0.5 Å beyond this point. This is by no means an optimally efficient integration scheme but it seems to be sufficient for the purposes required here.

Convergence of the numerical calculations was assessed by computing PB and integral-derived effective radii as described above and at twice this resolution. Since this increases the problem size by approximately a factor of 8, it was necessary to use a small test system: acetyl-phenylalanine-$N$-methyl-amide was selected. The test system had a collapsed conformation such that the largest effective radius was approximately 2.6 Å. All inverse effective radii calculated by the PB method decreased with the increased resolution: the largest absolute change was 0.004 Å$^{-1}$ and the largest relative change was 0.5%. All but one of the inverse effective radii calculated using the $R6$ integral increased or remained constant with the increased resolution: the largest absolute change was 0.001 Å$^{-1}$; the largest relative change was 0.3%. The small magnitude of change due to increased resolution relative to the average discrepancy observed between PB and integral-derived effective radii (see Table I) indicates that for the purposes required here, the values calculated using the processes described previously can be considered well converged.

## IV. CONCLUSIONS

In this study, the accuracies of several integral expressions for calculating effective Born radii have been assessed by comparing the radii they yield with well-converged "perfect" (PB-based) radii. Direct examination of these expressions using exactly the same molecular surface definitions in all cases and high precision numerical integration (as opposed to approximate fast routines often used in practice to estimate the integrals involved) leads to fewer uncertainties and clearer basic conclusions. We find that the empirical integral expression introduced by Lee *et al.* in the GBMV2 model and the $R6$ formula based on integrating $r^{-6}$ over solute volume or surface are both dramatic improvements over the still widely used CFA approximation. The $R6$ expression based on a single volume or surface integral may be preferred to the GBMV2 expression, or similar expressions involving combinations of multiple integral terms $r^{-N}$, $N \neq 6$. The main reason is lower computational complexity and stronger theoretical foundation of the $R6$, combined with its somewhat better accuracy. The clear physical basis of the $R6$ formula is especially appealing for future development of GB-type models. The accuracy of some approximate expressions that involve combinations of $r^{-N}$, $N \neq 6$ terms can be rationalized by their closeness to the exact formula for perfectly spherical geometry—the $R6$.

On the realistic biomolecular shapes we have analyzed, the $R6$-based radii contain two types of errors: a fairly systematic bias for nearly all atoms and a "random" component for a small subset of them. The use of progressively higher resolution grids shows that the systematic bias is not an artifact of inadequate accuracy of grid-based integration used here. Likewise, this bias does not appear to depend upon the overall molecular shape. Rather, the bias is a function of the fine-grain coarseness of molecular surface: the larger the probe radius used to compute the surface, the smoother the surface and the smaller the bias. For a given probe, the systematic bias of $R6$ approximation can be significantly re- duced by adding a single constant offset to the inverse effective radii. The offset value is transferable between structures. The resulting approximation that we termed as "$R6^*$" (with the offset tabulated for the commonly used probe value of 1.4) produces effective Born radii that are indistinguishable from the "perfect" ones in the following sense: the corresponding solvation free energies computed via the standard GB formula are well within the statistical error margin of each other. Notably, the remaining error margin of the GB($R6^*$) solvation free energies relative to the numerical PB is three times lower than the recently reported[50] error of the reference numerical PB methodology itself relative to the explicit solvent treatment.

In addition to the largely removable systematic error, some $R6$ effective radii show what appears to be appreciable nonsystematic error. While the latter type of error affects sufficiently few atoms that its impact on the solvation energy of the entire molecule is negligible, understanding the origins of this type of error is important for two reasons. First, these deviations may have a significant local effect, e.g., in MD simulations or pK calculations. Second, they point to the limitations of the GB model itself, understanding of their origins may eventually lead to improvements of the model. While we may not have covered every possible scenario in this work, one trend is definite: most significant differences between the $R6$ radii and the "perfect" PB-based ones occur in regions of solute whose local geometry and/or topology is significantly different from convex simply connected spherical. That conclusion should come at no surprise since the $R6$ approximation is itself derived on a perfect sphere—the geometry for which $R6$ yields exact effective radii.

Importantly, the above conclusions about the geometrical and topological sources of error in the $R6$ do not depend on the specific power (6) of the $r^{-6}$ integrand, and is likely to hold true for any other model based on volume/surface integration, providing that the model goes sufficiently beyond the CFA, and the integration is done accurately enough to reveal these effects.

On the basis of the above conclusions, two observations can be made relevant to the future development of the GB models. First, the $R6^*$ expression seems to represent a sufficient solution to the problem of calculating effective radii: further attempts to increase their agreement with PB results would be unlikely to succeed in improving the accuracy of the GB model itself in its canonical version due to Still *et al.* Instead, it appears more prudent to focus on developing analytical models that go beyond the current GB in capturing the effects of essentially nonspherical molecular topology/ connectivity within the implicit solvent PB framework. Still, in view of the relatively large inherent errors of the PB model itself relative to the more fundamental explicit solvent treatment, one has to be careful to distinguish overfitting to PB from real improvement. Perhaps, more immediate challenges in GB model development involve inventing fast and accurate methods to perform the integration in the $R6^*$ expression (or equivalent) over a physically realistic dielectric boundary: the use of simplified dielectric boundaries, such as those based on VDW surface, may offer computational advantages,[51] but at the same time may not be physically

realistic.[45] In fact, since the $R6$ no longer has the fortuitous cancellation of errors that "helped" VDW-based CFA, one has to be especially careful with choosing the right integration domain. The problems to be overcome are nontrivial and are present in every aspect of the problem, including the need for high numerical stability of the algorithms required to be used in MD simulations.

## ACKNOWLEDGMENTS

[1] M. K. Gilson, M. E. Davis, B. A. Luty, and J. Andrew McCammon, J. Phys. Chem. **97**, 3591 (1993).
[2] B. H. Honig and A. Nicholls, Science **268**, 1144 (1995).
[3] J. D. Madura, J. M. Briggs, R. C. Wade, M. E. Davis, B. A. Luty, A. Ilin, J. M. Antosiewicz, M. K. Gilson, B. Bagheri, L. R. Scott, and J. Andrew McCammon, Comput. Phys. Commun. **91**, 57 (1995).
[4] P. Beroza and D. A. Case, Methods Enzymol. **295**, 170 (1998).
[5] M. Scarsi, J. Apostolakis, and A. Caflisch, J. Phys. Chem. A **101**, 8098 (1997).
[6] C. J. Cramer and D. G. Truhlar, Chem. Rev. (Washington, D.C.) **99**, 2161 (1999).
[7] N. A. Baker, Curr. Opin. Struct. Biol. **15**, 137 (2005).
[8] W. Im, D. Beglov, and B. Roux, Comput. Phys. Commun. **111**, 59 (1998).
[9] R. Luo, L. David, and M. K. Gilson, J. Comput. Chem. **23**, 1244 (2002).
[10] Q. Lu and R. Luo, J. Chem. Phys. **119**, 11035 (2003).
[11] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. Andrew McCammon, Proc. Natl. Acad. Sci. U.S.A. **98**, 10037 (2001).
[12] M. Feig and C. L. Brooks, Curr. Opin. Struct. Biol. **14**, 217 (2004).
[13] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson, J. Am. Chem. Soc. **112**, 6127 (1990).
[14] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, Chem. Phys. Lett. **246**, 122 (1995).
[15] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. **100**, 19824 (1996).
[16] M. Schaefer and M. Karplus, J. Phys. Chem. **100**, 1578 (1996).
[17] D. Qiu, P. S. Shenkin, F. P. Hollinger, and W. C. Still, J. Phys. Chem. **101**, 3005 (1997).
[18] S. R. Edinger, C. Cortis, P. S. Shenkin, and R. A. Friesner, J. Phys. Chem. B **101**, 1190 (1997).
[19] B. Jayaram, Y. Liu, and D. L. Beveridge, J. Chem. Phys. **109**, 1465 (1998).
[20] A. Ghosh, C. S. Rapp, and R. A. Friesner, J. Phys. Chem. B **102**, 10983 (1998).
[21] D. Bashford and D. A. Case, Annu. Rev. Phys. Chem. **51**, 129 (2000).
[22] M. S. Lee, F. R. Salsbury, and C. L. Brooks, J. Chem. Phys. **116**, 10606 (2002).
[23] A. K. Felts, Y. Harano, E. Gallicchio, and R. M. Levy, Proteins **56**, 310 (2004).
[24] A. N. Romanov, S. N. Jabin, Y. B. Martynov, A. V. Sulimov, F. V. Grigoriev, and V. B. Sulimov, J. Phys. Chem. A **108**, 9323 (2004).
[25] B. N. Dominy and C. L. Brooks, J. Phys. Chem. B **103**, 3765 (1999).
[26] L. David, R. Luo, and M. K. Gilson, J. Comput. Chem. **21**, 295 (2000).
[27] V. Tsui and D. A. Case, J. Am. Chem. Soc. **122**, 2489 (2000).
[28] N. Calimet, M. Schaefer, and T. Simonson, Proteins **45**, 144 (2001).
[29] V. Z. Spassov, L. Yan, and S. Szalma, J. Phys. Chem. B **106**, 8726 (2002).
[30] C. Simmerling, B. Strockbine, and A. E. Roitberg, J. Am. Chem. Soc. **124**, 11258 (2002).
[31] T. Wang and R. C. Wade, Proteins **50**, 158 (2003).
[32] H. Nymeyer and A. E. Garcia, Proc. Natl. Acad. Sci. U.S.A. **100**, 13934 (2003).
[33] A. Onufriev, D. Bashford, D. A. Case, Proteins **55**, 383 (2004).
[34] E. Gallicchio and R. M. Levy, J. Comput. Chem. **25**, 479 (2004).
[35] M. C. Lee and Y. Duan, Proteins **55**, 620 (2004).
[36] A. Onufriev, D. A. Case, and D. Bashford, J. Comput. Chem. **23**, 1297 (2002).
[37] A. Onufriev, D. Bashford, and D. A. Case, J. Phys. Chem. B **104**, 3712 (2000).
[38] M. S. Lee, M. Feig, F. R. Salsbury, and C. L. Brooks, J. Comput. Chem. **24**, 1348 (2003).
[39] T. Grycuk, J. Chem. Phys. **119**, 4817 (2003).
[40] W. Im, M. S. Lee, and C. L. Brooks, J. Comput. Chem. **24**, 1691 (2003).
[41] M. Feig, A. Onufriev, M. S. Lee, W. Im, D. A. Case, and C. L. Brooks, J. Comput. Chem. **25**, 265 (2004).
[42] G. Sigalov, P. Scheffel, and A. Onufriev, J. Chem. Phys. **122**, 094511 (2005).
[43] G. Sigalov, A. Fenley, and A. Onufriev, J. Chem. Phys. **124**, 124902 (2006).
[44] M. Feig, W. Im, and C. L. Brooks, J. Chem. Phys. **120**, 903 (2004).
[45] J. M. J. Swanson, J. Mongan, and J. Andrew McCammon, J. Phys. Chem. B **109**, 14769 (2005).
[46] J. Mongan, C. Simmerling, J. A. McCammon, D. A. Case, and A. Onufriev, J. Chem. Theory Comput. **3**, 156 (2007).
[47] A. Varshney, F. P. Brooks, and W. V. Wright, IEEE Comput. Graphics Appl. **14**, 19 (1994).
[48] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).
[49] D. Bashford, *Scientific Computing in Object-Oriented Parallel Environments*, edited by Y. Ishikawa, R. R. Oldehoeft, V. W. Reynders, and M. Tholburn, Lecture Notes in Computer Science, Vol. 1343 (Springer, Berlin, 1997), p. 233.
[50] J. M. J. Swanson, S. A. Adcock, and J. Andrew McCammon, J. Chem. Theory Comput. **1**, 484 (2005).
[51] H. Tjong and H. X. Zhou, J. Phys. Chem. B **111**, 3055 (2007).
[52] M. Sanner, A. Olsen, and J.-C. Spehner, *Proceedings of the 11th ACM Symposium on Computational Geometry* (ACM, New York, 1995), pp. C6–C7.