A Generalized-Born Solvation Model for Macromolecular Hybrid-Potential Calculations

Eric Pellegrini and Martin J. Field*

Laboratoire de Dynamique Moléculaire, Institut de Biologie Structurale, Jean-Pierre Ebel, 41 Rue Jules Horowitz, F-38027 Grenoble Cedex 01, France

Received: September 12, 2001; In Final Form: December 4, 2001

Generalized-born surface-area (GBSA) models have proved to be effective tools for estimating rapidly and with reasonable accuracy the solvation energies of molecular and macromolecular systems, and they have been employed in conjunction with both molecular mechanical (MM) and quantum mechanical (QM) potentials. In this article, we present our work to develop a GBSA model for calculations on macromolecules using hybrid potentials in which part of the system is treated with a semiempirical QM potential and the remaining atoms with a MM potential. Our efforts have centered principally on finding an approach for the calculation of the Born radii which is appropriate for MM and QM potentials and for small and large molecules, but inevitably, the competing requirements of these goals have meant a compromise in the design and parametrization of the model. We have, however, produced a scheme that we feel is suitable for macromolecular hybrid potential studies of processes, such as protein−ligand binding.

1. Introduction

The accurate description of solvation effects is a major challenge for theoretical and computational chemistry and biochemistry. The importance of such a goal is obvious, as many chemical phenomena of interest, and most of those in biology, occur in an at least a partially aqueous environment and solvation can have a powerful influence on the properties of molecular systems and the types of processes that they can undergo.

Arguably the most accurate way of simulating solvent effects is to handle the solute, which is the molecular system of interest, and the solvent at the same level of approximation. This is the basis behind explicit solvent models in which the solute is immersed in a bath of explicitly treated solvent molecules.[1] Although precise, calculations with explicit solvent models are expensive because proper solvation of the solute usually requires that the number of solvent atoms far exceeds those of the solute.

An alternative to the explicit solvent models are the implicit models in which an atomic-level representation of the solvent molecules is replaced by a much simpler description (for nice reviews, see refs 2 and 3). One of the most powerful implicit solvent models is based upon the Poisson−Boltzmann (PB) equation in which the solvent is represented by a continuum with dielectric constant and ionic strength of appropriate value and the solute occupies a cavity of the correct shape within the continuum. The electrostatic solvation energy within this model is obtained by solving the PB equation for the electrostatic potential throughout the system. PB-type models have given results of great utility, but unfortunately, algorithms for the solution of the PB equation and the accurate calculation of the derivatives of the electrostatic solvation energy are slow, which means that PB methods have found limited application in macromolecular geometry optimization and molecular dynamics calculations.

To speed up calculations, a number of approximate implicit-solvent methods have been proposed; one of the most successful of which has been the generalized-born surface-area (GBSA) approach. The use of the GB approximation for the estimation of solvation effects has a relatively long history in quantum chemistry, particularly for semiempirical quantum mechanical (QM) potentials (see, for example, refs 4 and 5 and references therein), but it appears to have been first applied to simulations with molecular mechanical (MM) potentials by Still and co-workers.[6] The appeal of the GBSA-type approach is that it can give results of a quality that compare favorably with those obtained by solving the PB equation but at a fraction of the cost.[7] It is also relatively easy to construct GBSA models that are differentiable and so can be used with simulation techniques that require derivatives.

GBSA methods are still the subject of active research. Apart from applications, recent work has concerned the introduction of analytic methods for the calculation of Born radii[8] and the parametrization of models for different force fields,[9] to account for salt effects[10] and improve the applicability of the models to macromolecular calculations.[11] Also notable is the extensive work of Cramer, Truhlar, and co-workers on a series of solvation models, related to the GBSA approach, that they have developed for use with semiempirical and ab initio QM potentials.[12]

One of the main interests in our laboratory has been the development of hybrid QM/MM potentials which we employ to study such processes as enzymatic reactions.[13] The concept underlying hybrid potentials is that they treat different portions of a system with potentials of differing accuracy and computational cost.[14,15] Thus, for example, a study of an enzyme reaction could treat the substrate and active site with a QM method and the rest of the system with a simpler MM potential. The aim of the work presented in this paper has been to come up with a GBSA model that is compatible with the range of hybrid QM/MM potentials that we have developed[16] and that is applicable to both small molecule and macromolecular systems. The outline of this paper is as follows. Section 2 describes our formulation of the GBSA model, section 3 presents

* Corresponding author. Tel: (33)-4-38-78-95-94. Fax: (33)-4-38-78-54-94. E-mail: pellegri@ibs.fr and mjfield@ibs.fr.

10.1021/jp0135050 CCC: $22.00
Published on Web 01/29/2002

the results of the parametrization of our models, and section 4 summarizes the paper.

2. Methods

In this section, we outline briefly the GBSA method and then go on to describe our treatment of the electrostatic part of the solvation energy.

2.1. GBSA Method. Following Still et al.,[6] the total solvation free energy in a GBSA model is written as the sum of three terms:

$$G_{\mathrm{solv}} = G_{\mathrm{cav}} + G_{\mathrm{vdW}} + G_{\mathrm{pol}} \tag{1}$$

where G_{cav} is the energy required to create a cavity for the solute in the solvent, G_{vdW} is the energy of the solute−solvent van der Waals interactions, and G_{pol} is the solute−solvent electrostatic polarization energy.

The first two terms are normally grouped together to form the SA part of the GBSA model, G_{SA}, and are expressed as

$$G_{\mathrm{SA}} = \sum_{i=1}^{N_{\mathrm{atoms}}} \sigma_i A_i \tag{2}$$

where A_i is the solvent accessible surface of atom i, σ_i is an empirical atomic solvation parameter, and N_{atoms} is the number of atoms in the system.

The polarization energy is approximated by a generalized-born equation of the form

$$G_{\mathrm{pol}} = -\frac{1}{2}\gamma\left(\frac{1}{\epsilon_{\mathrm{I}}} - \frac{1}{\epsilon_{\mathrm{O}}}\right)\sum_{i=1}^{N_{\mathrm{atoms}}}\sum_{j=1}^{N_{\mathrm{atoms}}}\frac{q_i q_j}{f_{ij}(r_{ij})} \tag{3}$$

where ϵ_{I} and ϵ_{O} are the solute and solvent dielectric constants, respectively, q_i is the fixed partial charge on atom i, f_{ij} is a function dependent upon the distance r_{ij} between the atoms i and j, and γ is a unit conversion factor. There is no unique form for the function f_{ij}, but Still et al. proposed the following:[6]

$$f_{ij}(r_{ij}) = \sqrt{r_{ij}^{\,2} + \alpha_{ij}^{\,2}\exp\left(-\frac{r_{ij}^{\,2}}{4\alpha_{ij}^{\,2}}\right)} \tag{4}$$

where $\alpha_{ij} = \sqrt{\alpha_i \alpha_j}$.

α_i is the Born radius for atom i, and it indicates how shielded the atom is from solvent. It can take values that range from the normal atomic radius, for atoms that are completely exposed to solvent, right up to the "radius" of the molecule of which the atom forms a part, for atoms that are buried at the center of the molecule. The efficient calculation of accurate Born radii is a crucial aspect of the GBSA method. In the original approach of Still et al.,[6] the computation was done numerically by determining the polarization energy, $G_{\mathrm{pol},i}$, for each atom in the system, assuming that that atom had a unit charge and that all the other atoms were neutral but still displaced the solvent. The radii were then determined from the formula

$$\alpha_i = -\frac{1}{2}\gamma\left(\frac{1}{\epsilon_{\mathrm{I}}} - \frac{1}{\epsilon_{\mathrm{O}}}\right)\frac{1}{G_{\mathrm{pol},i}} \tag{5}$$

The numerical evaluation of the radii (by a finite-difference method) is accurate but slow and does not readily provide the derivatives of the radii with respect to atomic positions which are needed for geometry optimizations and molecular dynamics simulations. This led to a number of workers, including Still et al.[8] and Hawkins et al.,[17] to seek approximate, analytic formulas to the Born radii that could be differentiated and rapidly evaluated.

The GBSA method as detailed by Still et al. was for use with MM potentials. Similar models have, however, been used in conjunction with QM potentials. Early work was done by Constanciel and Contreras[4] and by Kozaki et al.,[18] who combined GB models with semiempirical QM methods. More recent work has been done by Cramer, Truhlar, and co-workers, who have developed an extensive series of GB solvation models for use with both semiempirical and ab initio MO and density functional theory QM methods.[12]

The difference between applying a GBSA model to MM and QM potentials is that the charge or electron density of the quantum atoms changes in response to the surrounding dielectric medium. To see this explicitly, let us consider a solvated closed shell molecule being treated with a semiempirical QM method of the AM1, MNDO or PM3 type[19,20,21]. The energy, E_{QM}, of such a system is given by

$$E_{\mathrm{QM}} = \frac{1}{2}\sum_{\mu\nu} P_{\mu\nu}(H_{\mu\nu} + F_{\mu\nu}) + V_{\mathrm{nuc}} + G_{\mathrm{pol}} + G_{\mathrm{SA}} \tag{6}$$

where V_{nuc} is the repulsion energy between the nuclei of the quantum atoms and G_{pol} and G_{SA} are the GBSA energy terms of eqs 3 and 2, respectively. $P_{\mu\nu}$, $H_{\mu\nu}$, and $F_{\mu\nu}$ refer to the electronic density matrix, the one-electron matrix and the Fock matrix of the quantum atoms and the indices μ and ν to the basis functions used to expand the system's molecular orbitals.

The optimum orbitals and, hence, wave function and electron density are obtained by optimizing the energy expression (eq 6) with respect to the electronic variables, in this case the MO coefficients. In a vacuum, this leads to an eigenvalue equation of the form

$$\mathbf{F}\mathbf{c}_s = \epsilon_s \mathbf{c}_s \tag{7}$$

where $\mathbf{F}$ is the Fock matrix, ϵ_s is the energy of the orbital s and $\mathbf{c}_s$ is the vector of MO coefficients. In solution, the situation is a little more complicated, as the polarization energy, G_{pol}, is a function of the charges of the QM atoms and, hence, of the MO coefficients. As is well-known, there is no unique way of decomposing the electronic density of a quantum system into atomic contributions, and so a variety of schemes are employed. However, the simplest is a Mulliken analysis which, for the semiempirical methods used here, relates the charge, q_i, of the atom to the atom's nuclear charge Z_i and to its diagonal density matrix elements through

$$q_i = Z_i - \sum_{\mu\in_i} P_{\mu\mu} \tag{8}$$

With this definition of the atomic charges, optimization of the polarization energy, G_{pol}, with respect to the MO coefficients produces an equation similar to eq 7 except that the diagonal elements of the vacuum Fock matrix must be modified by adding terms of the form

$$F'_{\mu\mu} = \gamma\left(\frac{1}{\epsilon_I} - \frac{1}{\epsilon_O}\right)\sum_j \frac{q_j}{f_{ij}} \qquad \mu\in i \tag{9}$$

Solution of eq 7, with the extra term 9, proceeds in exactly the same way as in the vacuum case and is usually carried out using an iterative self-consistent field (SCF) algorithm.

2.2. GBSA Models and Hybrid Potentials. The implementation of a GBSA method for use with a hybrid QM/MM potential poses no problems of principle. The hybrid potentials that we treat in this paper divide the atoms of the system into two regions—one of which is treated with a semiempirical QM method and the other with a standard MM force field—but extension to more complicated hybrid potential schemes is straightforward.

As there is now a large literature concerning hybrid potentials, description of them here will be cursory, and so interested readers are referred to references[14,15] for further details. The potential energy, E, of a two-region hybrid potential can be written as the sum of three terms, one for the energy of the QM region, E_{QM}, one for the energy of the MM region, E_{MM}, and one for the energy of interaction between the two, $E_{QM/MM}$. Use of the GBSA model adds the two extra terms, G_{pol} and G_{SA}, where

$$E = E_{QM} + E_{MM} + E_{QM/MM} + G_{pol} + G_{SA} \qquad (10)$$

The GBSA terms necessitate little modification of a standard hybrid potential calculation (which is described, for example, in ref 16). The surface area term G_{SA} is independent of the nature of the atoms, whether QM or MM, and so can be evaluated independently. The polarization term is not so simple but, for the QM/MM case, can be rewritten as

$$G_{pol} = -\frac{1}{2}\gamma\left(\frac{1}{\epsilon_I} - \frac{1}{\epsilon_O}\right)\left[\sum_{i\in QM}\sum_{j\in QM}\frac{q_i q_j}{f_{ij}} + \sum_{i\in QM}\sum_{j\in MM}\frac{q_i q_j}{f_{ij}} + \sum_{i\in MM}\sum_{j\in MM}\frac{q_i q_j}{f_{ij}}\right] \qquad (11)$$

The polarization energy can be calculated easily from this expression, once the charges on the QM atoms have been determined, whereas the presence of interactions between the charges on the QM and MM atoms (the second term in brackets on the right-hand side of eq 11) means that the expression for the Fock matrix terms arising from the polarization energy, eq 9, must be modified by extending the sum to include all atoms, both QM and MM, in the system. The presence of bonds between atoms of the QM and MM regions does not introduce any complications, and so they can be handled in the normal way.

To finish this section, we note that for calculations on macromolecular systems with MM potentials, the solute dielectric constant is often taken to be different from one. Thus, for example, values ranging from 4 to 20 are common for calculations on proteins. For macromolecular calculations with hybrid potentials, it may, in certain cases, be advantageous to do the same thing, but as a high internal dielectric constant would not necessarily be appropriate for the QM region, it would mean that two solute regions with different dielectrics would be required. We do not explore this point here but leave it for future work.

2.3. Calculating the Born Radii. Once the basic GBSA model that we are going to use and its extension to hybrid potentials have been defined, it remains to be specified how the Born radii are to be calculated. In particular, we seek a method that is analytic, so that the expression can be differentiated, that is reasonably inexpensive computationally and that can be used, without change, for calculations with QM, MM, and QM/MM potentials and on both molecular and macromolecular systems.

We investigated a number of different routes and approaches, including that of Hawkins et al.,[17] but in the end, we obtained the best results by starting with the analytic expression introduced by Still and co-workers[8] for the polarization energies of the individual atoms, $G_{pol,i}$, that is used to define the Born radii (see eq 5). Their expression has the form

$$G_{pol,i} = \left(\frac{1}{\epsilon_I} - \frac{1}{\epsilon_O}\right)\left[\frac{-\gamma/2}{\rho_i + \phi + P_1} + \sum_j^{stretch}\frac{P_2 V_j}{r_{ij}^4} + \sum_j^{bend}\frac{P_3 V_j}{r_{ij}^4} + \sum_j^{non\text{-}bonding}\frac{P_4 V_j \mathcal{O}(r_{ij})}{r_{ij}^4}\right] \qquad (12)$$

where ρ_i is the van der Waals radius of atom i and ϕ, P_1, P_2, P_3, and P_4 are parameters. The three sums on the right-hand side of the equation run over the atoms that are bonded directly to atom i (the 1−2 interactions), those that are separated by two bonds (the 1−3 interactions) and those that are separated by three or more bonds (the 1−4 and higher interactions), respectively. $\mathcal{O}(r_{ij})$ is a function whose form need not concern us here, and V_j is the effective atomic volume of atom j which is given by

$$V_j = \frac{4\pi\rho_j^3}{3} - \sum_k \frac{\pi}{3} h_{jk}^2 (3\rho_k - h_{jk}) \qquad (13)$$

where the sum runs over all atoms k that are bonded to atom j and the function h_{jk} is defined as

$$h_{jk} = \rho_j\left(1 + \frac{\rho_k^2 - \rho_j^2 - r_{jk}^2}{2\rho_j r_{jk}}\right) \qquad (14)$$

Although we used eq 12 as our starting point, we extensively modified it. After much trial and error we arrived at our final expression which differs from the original one in three major ways:

(1) The expression 12 was developed for MM potentials, but it is not ideal for use with QM potentials, as the sums on the right-hand side of the equation require the definition of the bonds in the molecule, information which is unnecessary when performing QM calculations. Instead, therefore, we chose a form which replaced the sums over 1−2, 1−3, and 1−4 and higher interactions by sums which involved interactions defined by distance.

(2) Quite often, particularly for large molecules, the approximate formulas for $G_{pol,i}$, eq 12, and for the atomic volume, eq 13, give values that are of the wrong sign—positive for $G_{pol,i}$, implying a negative Born radius from eq 5, and negative for the atomic volume. This we cured by introducing functions that ensured that $G_{pol,i}$ was always negative and the atomic volume always remained positive.

(3) It is known (see, for example, ref 8 and the work of Onufriev et al.[11]) that the approximate analytic formulas for $G_{pol,i}$ often give values which are much too negative and, hence, radii that are much too small, for atoms that are buried inside a macromolecule. We were therefore forced to add terms which corrected for this.

We describe in more detail the changes that we made below, but the formulas that define our final model are as follows:

$$G_{\mathrm{pol},i} = \begin{cases} G_{\max} & G'_{\mathrm{pol},i} > 0 \\ \frac{1}{G_0^2}\left[\frac{2G_{\max}}{G_0} - 1\right](G'_{\mathrm{pol},i})^3 + \frac{1}{G_0}\left[2 - \frac{3G_{\max}}{G_0}\right] & G'_{\mathrm{pol},i} \in [G_0,0] \\ G'_{\mathrm{pol},i} & G'_{\mathrm{pol},i} < G_0 \end{cases}$$

$$(G'_{\mathrm{pol},i})^2 + G_{\max} \quad (15)$$

where

$$G'_{\mathrm{pol},i} = \left(\frac{1}{\epsilon_{\mathrm{I}}} - \frac{1}{\epsilon_{\mathrm{O}}}\right)\left[-\frac{\gamma}{2}\left(\frac{H_i}{\rho_i + P_1}\right) + \sum_{j<D_1}\frac{P_2 V_j}{r_{ij}^4} + \sum_{D_1<j<D_2}\frac{P_3(r_{ij})V_j}{r_{ij}^4} + \sum_{j>D_2}\frac{P_4 S_j V_j}{r_{ij}^4}\right] \quad (16)$$

In this equation, the sums in eq 12 over different classes of interactions have been replaced by sums over atoms, the atoms to include in which are determined by the values of two distance parameters D_1 and D_2. Likewise, the parameter P_3 in the original formulation has been replaced by a function $P_3(r_{ij})$ which interpolates smoothly between the constants P_2 and P_4 in the range $[D_1,D_2]$. It has the form

$$P_3(r_{ij}) = P_2 + \frac{(P_4 S_j - P_2)(D_1^2 - r_{ij}^2)^2(D_1^2 + 2r_{ij}^2 - 3D_2^2)}{(D_1^2 - D_2^2)^3} \quad (17)$$

We have modified the expression for the atomic volume such that

$$V_j = \begin{cases} V'_j & V'_j > V_0 \\ \frac{1}{V_0^2}\left[\frac{2V_{\min}}{V_0} - 1\right](V'_j)^3 + \frac{1}{V_0}\left[2 - \frac{3V_{\min}}{V_0}\right](V'_j)^2 + V_{\min} & V'_j \in [0,V_0] \\ V_{\min} & V'_j < 0 \end{cases} \quad (18)$$

with V'_j having a similar form to eq 13 except for the introduction of a switching function $\mathscr{W}'_{jk}$ which permits the sum to be over all atoms that are less than a distance cutoff D_s:

$$V'_j = \frac{4\pi\rho_j^3}{3} - \sum_{k\neq j}\frac{\pi}{3}h_{jk}^2(3\rho_k - h_{jk})\mathscr{W}'_{jk} \quad (19)$$

$\mathscr{W}'_{jk}$ has the form

$$\mathscr{W}'_{jk} = \begin{cases} 1 & r_{jk} \le D_1 \\ \frac{(D_s^2 - r_{jk}^2)(D_s^2 + 2r_{jk}^2 - 3D_1^2)}{(D_s^2 - D_1^2)^3} & r_{jk} \in [D_1,D_s] \\ 0 & r_{\mathrm{jk}} \ge D_{\mathrm{s}} \end{cases} \quad (20)$$

The introduction of the different cases in eqs 15, 18, and 20 ensures that the polarization energy of an atom remains negative and that its effective volume stays positive.

The last two functions, H_i and S_j, were introduced to improve Born radii when performing calculations on macromolecules. They are defined as

$$H_i = \begin{cases} 1 & A_i \ge A_H \\ 1 - [1 - H_0]\left[1 - \left(\frac{A_i}{A_H}\right)^2\right]^2 & A_i \ge A_H \end{cases} \quad (21)$$

and

$$S_{\mathrm{j}} = \begin{cases} 1 & A_j \ge A_S \\ 1 + S_{\max}\left[1 - \left(\frac{A_i}{A_H}\right)^2\right]^2 & A_j \ge A_S \end{cases} \quad (22)$$

In summary, eqs 15−22 complete the specification of our calculation of the Born radii. The parameters that need to be determined are D_1, D_2, D_s, $G_{\max}$, G_0, P_1, P_2, P_4, $V_{\min}$, V_0, A_H, A_S, H_0, and $S_{\max}$. It should be noted that all the parameters, except the last four (namely those that occur in eqs 21 and 22), are independent of the type of atom involved in the interactions. In contrast, it proved beneficial to consider different values for different atom types for the parameters A_H, A_S, H_0, and $S_{\max}$.

2.4. Technical Details. The GBSA model as described above was implemented for semiempirical QM, MM, and hybrid semiempirical QM/MM potentials in the molecular simulation program DYNAMO.[22,23] For the MM calculations, the OPLS-AA force field of Jorgensen et al.[24] was employed, whereas, for the QM potential, we always used the AM1 semiempirical method,[19] although the MNDO[20] and PM3 [21] methods gave results of similar quality. For the GBSA calculations, the atomic radii were taken to be half the values of the Lennard-Jones σ-parameters used in the OPLS-AA force field except for hydrogens where we used a value of 1.15 Å. The solvent accessible surface areas for the atoms and their derivatives were calculated using the algorithm of Wesson and Eisenberg[25] with a solvent probe radius of 1.4 Å. The cavity and van der Waals contributions to the solvation energy were obtained from eq 2 with the same value of σ_k of 0.0072 kcal mol^{-1}Å^{-2} for all atoms. Two different types of charge population analyses were tried in our QM and hybrid potential calculations—the Mulliken scheme of eq 8 and the class IV charge scheme developed by Cramer, Truhlar, and co-workers[26] for the AM1 potential.[19] The class IV scheme is more complex than the Mulliken approach but gives point charges that better reproduce molecular dipole moments. The solute and solvent dielectric constants were taken to have values of 1 and 80, respectively.

3. Results and Discussion

In this section, we discuss the results of calculations on a set of small molecules and on a set of proteins that we used to parametrize and test our GBSA model. Although the small molecule results are presented first and the macromolecular results second, several cycles of calculations on both sets of molecules were necessary before we obtained our optimum parameter set.

3.1. Small Molecules. The set of small molecules that we tested consisted of 73 molecules with a range of functional groups—alcohol, aliphatic, amine, aromatic, carbonyl, carboxylate—that cover many of those that are biologically important. For each of the molecules, the experimental solvation free energies are available.[8,9,27] The complete list of molecules is not given here but can be found in the tables contained in the Supporting Information.

To test the flexibility of our model, we performed five different types of parametrization which were as follows:
(i) *MM*. The calculations with this model were performed with a MM potential. The OPLS-AA force field was used to param-

TABLE 1: Values of the Parameters P_1, P_2, P_4, D_1, D_2 and D_s in the Various Optimized Models. The Model Names and Parameter Meanings Are Defined in the Text. P_1, D_1, D_2 and D_s Have Units of Å, P_2 and P_4 of kcal mol^{-1}Å, and $\mathscr{R}_1$ and $\mathscr{R}_2$ of kcal mol^{-1}.

model	P_1	P_2	P_4	D_1	D_2	D_s	$\mathscr{R}_1$	$\mathscr{R}_2$
MM	0.0002	1.9296	14.3298	1.5015	1.9038	3.9811	2.8	
QM(Mull.)	0.0001	1.6555	13.4871	1.5009	1.9211	3.9766	3.8	
QM(IV)	0.1555	1.0026	13.2668	1.5259	2.0322	3.9601	3.2	
Hybrid(Mull.)	0.0273	1.4588	14.1609	1.4756	2.1272	3.9624	2.9 (MM) 3.9 (QM)	3.4
Hybrid(IV)	0.0764	1.6396	15.3168	1.5424	2.0928	3.8850	3.0 (MM) 3.3 (QM)	3.2
MBest	0.0999	1.0053	9.5973	1.5533	2.7991	1.7477		

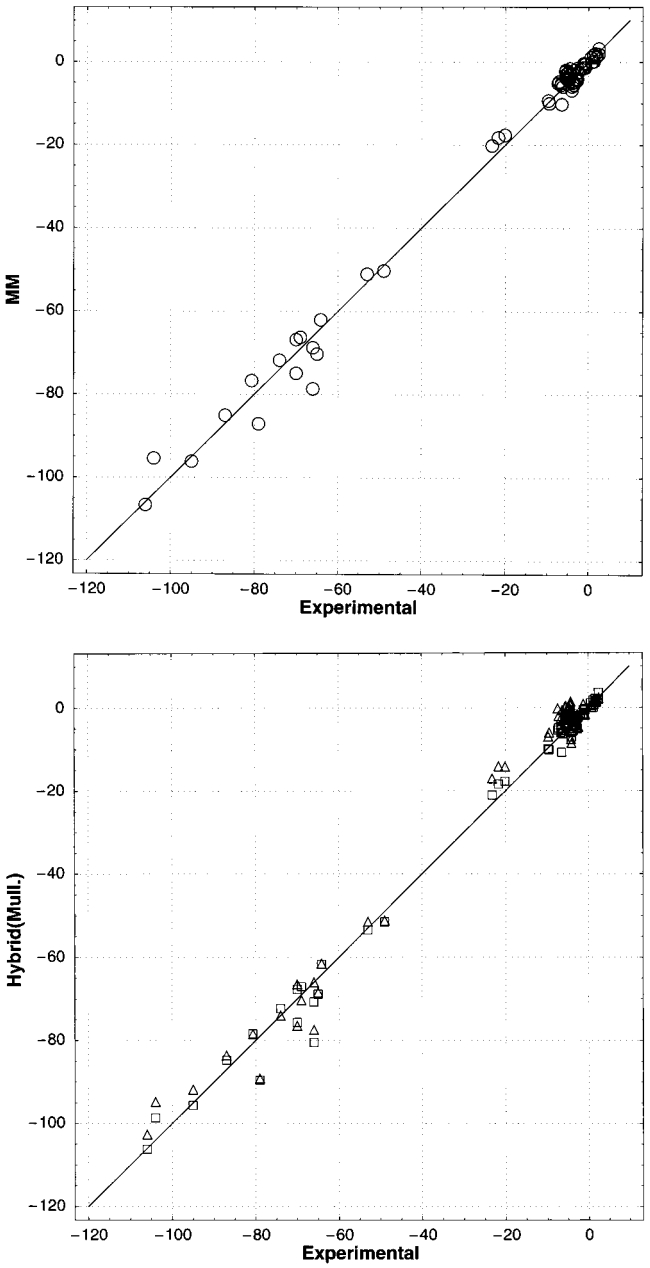

Figure 1. Plots of calculated vs experimental solvation energies for the small molecules in our test set: (a) the model MM; (b) the model Hybrid(Mull.), with MM values represented by triangles and QM values by squares. All energies are in kcal mol^{-1}.

etrize the molecules, but for those molecules where charge parameters were unavailable, we obtained point charges from the electrostatic potential-fitting procedure in the Jaguar ab initio program.[28] All calculations were done with the Hartree−Fock method and the 6-31G** basis set.

TABLE 2: Table Showing the Values of the $\mathscr{R}_1$ Function (in kcal mol^{-1}) Obtained When Different Sets of Reference Structures Are Used to Calculate the Small Molecule Solvation Free Energies[a]

parameter set/structures	vacuum	QM(Mull.)	QM(IV)
MM	2.8	2.8	2.8
QM(Mull.)	3.9	3.8	4.0
QM(IV)	3.3	3.3	3.3

[a] A full explanation of how the structures were prepared is given in the text. The diagonal elements represent the best fitted values for each of the models.

TABLE 3: PDB Codes and Corresponding Names of the Protein Structures Used for Parametrization of the GBSA Model

PDB code	protein name
1CRN	crambin
1STP	streptavidin
2IFB	intestinal fatty acid binding protein (I-FABP)
1MNC	neutrophil collagenase
4DFR	dihydrofolate reductase
1RBP	retinol binding protein
3PTB	β-trypsin
1DKX	substrate binding domain of DNAk
1ULB	purine nucleoside phosphorylase
1EED	endothiapepsin
1ABE	L-arabinose binding protein
2GBP	D-glucose binding protein
1THL	thermolysin
1CBX	carboxypeptidase A
1IVG	neuraminidase
1NSD	neuraminidase
2UAG	UDP *N*-acetylmuramoyl L-alanine D-glutamate ligase (MurD)
MurE	UDP *N*-acetylmuramoyl L-alanine D-glutamate L-lysine ligase
1YVE	acetohydroxy acid isomeroreductase

(ii) QM(Mull.) consisted of calculations using the AM1 QM potential with the charges on the quantum atoms obtained by a Mulliken population analysis.

(iii) QM(IV) was the same as model QM(Mull.) except that a class IV population analysis was done to obtain QM atom charges.

(iv) Hybrid(Mull.) consisted of a simultaneous parametrization of two copies of our set of 73 molecules. One set was treated with a MM potential and the other with a QM potential and a Mulliken population analysis.

(v) Hybrid(IV) was similar to model Hybrid (Mull.) but with a class IV population analysis.

In each case, the parametrization was performed by minimizing the root-mean-square (RMS) deviation between the calculated and experimental solvation free energies with respect to the parameters in our GBSA model. For the MM, QM(Mull.),

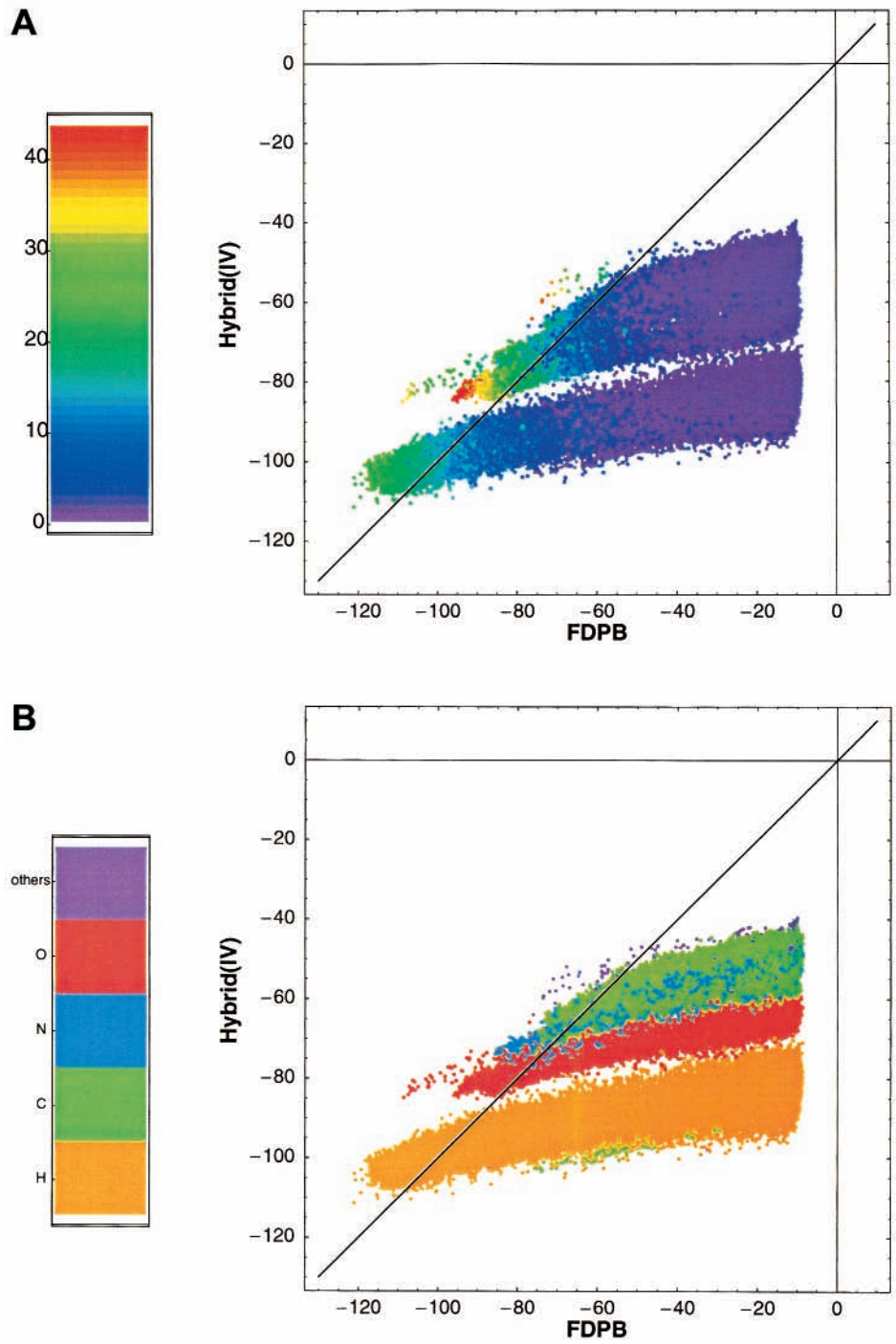

Figure 2. Plots of atomic polarization energies calculated with the Hybrid(IV) model vs reference values calculated using the UHBD program: (a) the points for each atom are colored according to their solvent accessible area (the color scale is in Å^2); (b) the points are colored according to atom type. All energies are in kcal mol^{-1}.

and QM(IV) models, the fitting function was

$$\mathscr{R}_1 = \sqrt{\frac{\sum_{m=1}^{N_{\mathrm{mol}}} (\Delta G_{\mathrm{solv},m}^{\mathrm{MM\ or\ QM}} - \Delta G_{\mathrm{solv},m}^{\mathrm{exp}})^2}{N_{\mathrm{mol}}}} \qquad (23)$$

whereas for the Hybrid models, it was

$$\mathscr{R}_2 = \sqrt{\frac{\sum_{m=1}^{N_{\mathrm{mol}}} (\Delta G_{\mathrm{solv},m}^{\mathrm{MM}} - \Delta G_{\mathrm{solv},m}^{\mathrm{exp}})^2 + \sum_{m=1}^{N_{\mathrm{mol}}} (\Delta G_{\mathrm{solv},m}^{\mathrm{QM}} - \Delta G_{\mathrm{solv},m}^{\mathrm{exp}})^2}{2N_{\mathrm{mol}}}} \qquad (24)$$

In these equations, N_{mol} is the number of molecules in the dataset (in our case 73) and $\Delta G_{\mathrm{solv},m}^{\mathrm{exp}}$, $\Delta G_{\mathrm{solv},m}^{\mathrm{MM}}$, and $\Delta G_{\mathrm{solv},m}^{\mathrm{QM}}$ are the experimental, the calculated MM, and the calculated QM free energies of solvation for molecule m, respectively.

The solvation free energies of the molecules were calculated as the differences between their energies in solution (i.e., with the GBSA model) and in a vacuum. In principle, the solution structure of each molecule needs to be reoptimized whenever the GBSA parameters change, but as this adds considerably to the cost of the parametrization, we used reference solution and vacuum structures for most of our calculations. This meant that only a single energy calculation needed to be done per molecule

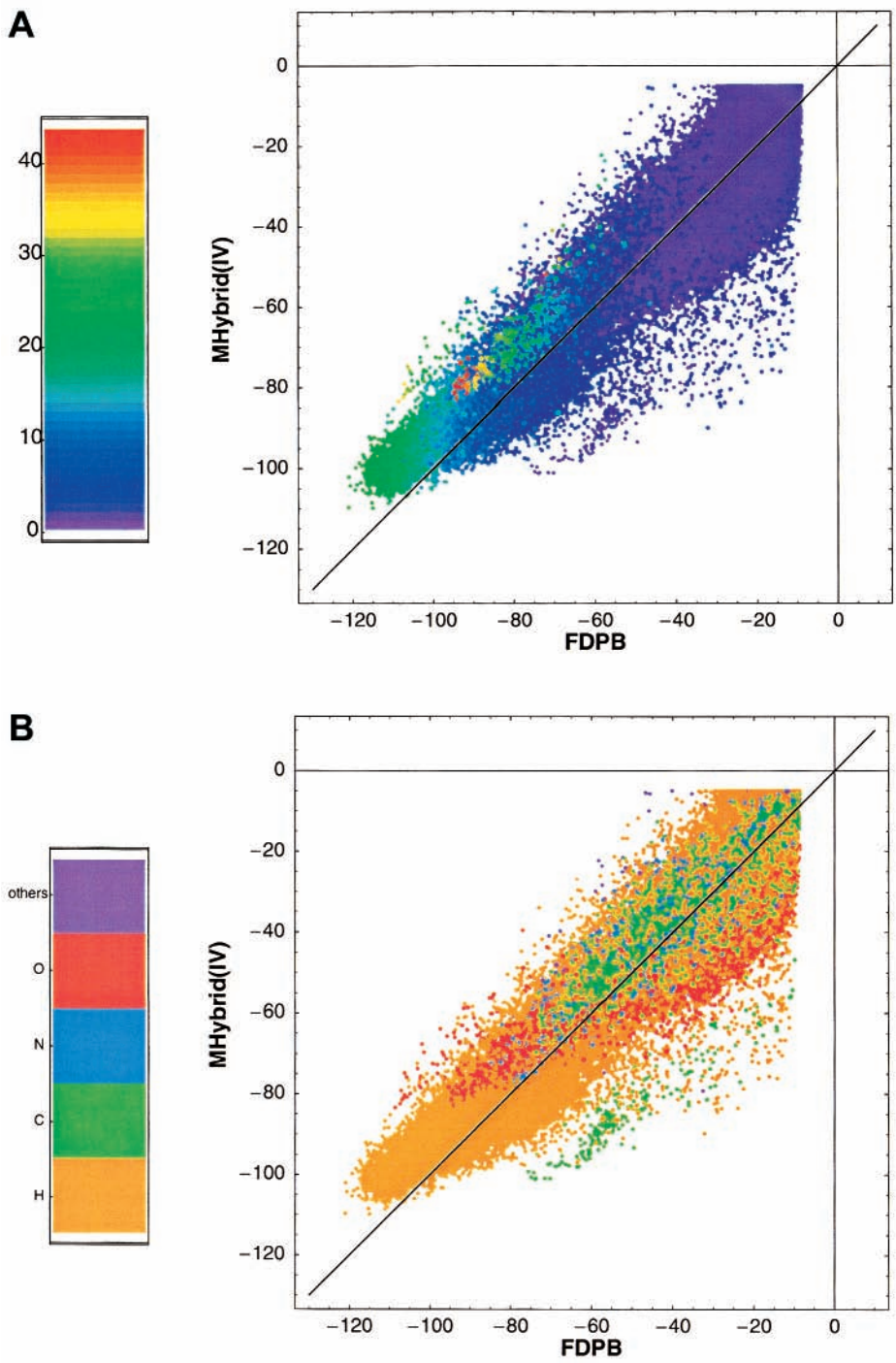

Figure 3. Plots of atomic polarization energies calculated with the MHybrid(IV) model vs reference values calculated using the UHBD program: (a) the points for each atom are colored according to their solvent accessible area (the color scale is in Å^2); (b) the points are colored according to atom type. All energies are in kcal mol^{-1}.

to obtain their solvation free energies (as the vacuum energies were constant and could be calculated once and stored). We did, however, test the validity of this assumption and found only small changes when different structures were used. Some results to illustrate this will be presented below.

The parameters that we found necessary to optimize for small molecules were D_1, D_2, D_s, P_1, P_2, and P_4. The other parameters, which enter into eqs 15, 18, 21, and 22, were found to be significant only for macromolecules and so will be discussed in the next section. The optimizations were carried out by minimizing the functions $\mathscr{R}_1$ and $\mathscr{R}_2$ with respect to the six free parameters using a genetic algorithm developed in our laboratory for various optimization tasks. The parameters resulting from the optimizations and the corresponding values of the fitting functions are shown in Table 1, whereas Figure 1 shows schematically the results of some of the fits. Tables containing the complete results for each molecule are provided in the Supporting Information.

Table 1 shows that the MM parametrization gives the lowest value of the $\mathscr{R}_1$ function, and hence the best fit, for the solvation energies. There is no reason a priori to suppose that the AM1 QM potential would give worse results than the MM potential—although it may be so—but it is clear that we performed a much more in-depth optimization for the MM models than for the QM and Hybrid models as calculations with the MM potentials were faster than the QM calculations by at least 1 order of

magnitude. This enabled us to include larger numbers of chromosomes and propagate them for more generations when running the genetic algorithm. A more detailed analysis of the solvation energies for the individual molecules shows that the MM and both QM models reproduce well the energies for hydrocarbon species but that the QM models are less accurate, in general, for groups containing heteroatom species—especially amino and hydroxyl groups for which the calculated solvation energies are systematically underestimated. As for the QM models, the QM(IV) parametrization gives results that are significantly better than QM(Mull.). This is reasonable given the more accurate charge distributions generated by a class IV than by a Mulliken population analysis. The results for the Hybrid models mirror those for the two QM models, with the Hybrid(IV) parametrization giving a better fit than Hybrid-(Mull.). As the hybrid models use a single set of parameters to fit both MM and QM solvation energies, they will not provide fits that are as good as those that result when the MM and QM solvation energies are fit individually. This is indeed the case as can be seen by comparing the values of the $\mathscr{R}_1$ functions for the different models in Table 1. Finally, we note that the values of the parameters obtained for the different models are quite similar with the parameters P_1 and P_2 showing proportionately the most variation.

To finish this subsection, we present results to show how altering the reference structures employed for the evaluation of the solvation free energies in the optimization procedure changes the value of the $\mathscr{R}_1$ function that is obtained. Thus, we recalculated the $\mathscr{R}_1$ function for the MM and QM models using three different sets of structures—a set obtained by minimizing the molecules with the AM1 potential in a vacuum and two other sets obtained by minimizing in solvent with the QM(Mull.) and QM(IV) GBSA parametrizations. The results are listed in Table 2, from which it can be seen that the differences for the MM and QM(IV) models are minor and for the QM(Mull.) model are of, at most, 0.2 kcal mol^{-1}.

3.2. Macromolecules. Our macromolecular tests were performed on a set of 19 proteins, which are listed in Table 3. The structures of 18 of these proteins were taken from the protein data bank (PDB,)[29] whereas the coordinates of the 19th were provided by O. Dideberg[30] (see Table 3). The same protocol to prepare the structures was followed for all the proteins. This involved determining the positions of the hydrogens using the INSIGHTII visualization program[31] and then performing a short conjugate-gradient energy minimization with respect to the positions of all atoms to remove any bad contacts using the CHARMM modeling program.[32]

In contrast to the small molecule tests, we did not fit to the solvation free energies of the proteins, as these are unavailable experimentally and, in any case, do not provide sufficient data for a fit, but took atomic polarization energies, $G_{\mathrm{pol},i}$, as data. The reference values we obtained by performing finite-difference Poisson−Boltzmann calculations with the UHBD program[33] and using the procedure defined by Still et al..[6] Thus, for each atom, we calculated the polarization energy assuming that the atom in question had a unit charge and all the others were neutral but displaced the dielectric. All calculations were done with zero ionic strength, and the frontier between the protein and solvent was defined by the solvent accessible surface with a probe sphere radius of 1.4 Å. Resolution of the finite-difference equations was done with grid-focusing, employing a coarse grid of size 41^3 and a spacing of 2.5 Å centered at the protein's center and a fine grid of size 21^3 and a spacing of 0.3 Å centered on the atom in question.

The fits were done in a way similar to the small molecule case, by minimizing the RMS deviation between the atomic polarization energies calculated with our GB model and the UHBD reference values. The optimization was performed, as before, using a genetic algorithm and the RMS function was

$$\mathscr{R}_M = \sqrt{\frac{\sum_{i=1}^{N_{\mathrm{atoms}}}(G_{\mathrm{pol},i} - G_{\mathrm{pol},i}^{\mathrm{UHBD}})^2}{N_{\mathrm{atoms}}}} \qquad (25)$$

Before discussing the results of our optimizations, we present in Figure 2 plots of the atomic polarization energies calculated with our small molecule Hybrid(IV) parametrization versus the UHBD reference values for the whole protein set (∼80000 atoms). For these calculations, we took the values of the H and S functions of eqs 21 and 22 to be one for all atoms. It can be seen that there is good agreement for atoms with very negative polarization energies (i.e., those that are exposed to solvent and so have small Born radii), but for buried atoms with low solvent accessible areas, the GB model severely overestimates the magnitude of the polarization energy. A closer examination shows that different atom types display deviations of differing magnitude. Thus, hydrogens deviate significantly more than oxygens which, in turn, show larger deviations than either carbons or nitrogens. This behavior is typical no matter which small molecule model is used.

It was due to results of the type shown in Figure 2 that we introduced the H and S functions, with the parameters A_H, A_S, H_0, and $S_{\max}$, and the parameters $G_{\max}$, G_0, $V_{\min}$, and V_0. In an initial parametrization of these functions, we took the small molecule Hybrid(Mull.) and Hybrid(IV) models and then optimized only the parameters of the H and S functions. The parameters A_H and H_0 were optimized only for hydrogens and the H function was not fitted for other atom types. By contrast, the parameters A_S and $S_{\max}$ were applied to all atoms although they were assumed to have the same value, irrespective of atom type. The parameters $G_{\max}$, G_0, $V_{\min}$, and V_0 were not optimized but were assigned the values −5 kcal mol^{-1}, −20 kcal mol^{-1}, $V_{\mathrm{full}}/7$ Å^3, and $V_{\mathrm{full}}/5$ Å^3, respectively, where $V_{\mathrm{full}} \equiv 4\pi\rho_i^3/3$, i.e., the uncorrected volume for atom i with radius ρ_i. These values were not optimized so as to reduce the complexity of the optimization procedure but were chosen after inspection of the various cases in which positive atom polarization energies or negative atom volumes were found. The optimizations were performed for streptavidin, as this protein is of a reasonable size but not so large that calculations become too expensive.

The results of these parametrizations, termed models MHybrid(Mull.) and MHybrid(IV), are shown in Figure 3, Table 4, and Table 5, respectively. Compared to Figure 2, Figure 3 shows that agreement between the GB and UHBD polarization energies has been greatly improved for all atoms, no matter what their type or their position within the protein. This supports our reasoning behind the introduction of the correction factors, i.e., using the S function to correct the radii for all buried atoms and then the H function for the correction of the hydrogen radii which deviate the most. Application of these two models, MHybrid(Mull.) and MHybrid(IV), to the other proteins in our test set showed that the parameters were transposable to larger proteins without a noticeable degradation in the quality of the results (see, for example, the $\mathscr{R}_M$ values for these models in Table 4).

Although our parametrizations in this section were done to reproduce the macromolecular atomic polarization energies, we are ultimately interested in the ability of the Born radii determined from these energies to reproduce macromolecular

TABLE 4: Values of the $\mathscr{R}_M$ Fitting Function (in kcal mol^{-1}) Obtained for the Test Set of 19 Proteins with Various Parameter Models[a]

protein	N_{atoms}	Hybrid(Mull.)	Hybrid(IV)	MHybrid(Mull.)	MHybrid(IV)	MBest
1CRN	642	33.4	27.1	8.4	8.5	6.5
1STP	1744	39.3	32.7	8.7	8.4	6.5
2IFB	2113	43.6	36.7	9.6	9.6	7.8
1MNC	2367	43.8	37.0	9.4	9.4	7.4
4DFR	2486	42.7	35.8	9.1	8.9	6.8
1RBP	2754	44.1	37.3	10.2	9.9	7.8
3PTB	3221	46.7	39.6	9.2	9.1	6.7
1DKX	3342	40.3	33.4	9.1	8.9	6.7
1ULB	4504	47.3	40.1	9.4	9.2	7.4
1EED	4669	47.7	40.7	9.7	9.4	7.1
1ABE	4672	47.2	40.1	9.2	9.1	6.8
2GBP	4698	47.6	40.4	9.4	9.1	6.9
1THL	4705	47.6	40.6	9.9	9.7	7.7
1CBX	4790	49.2	41.9	9.4	9.4	7.4
1IVG	5882	48.6	41.4	9.9	9.9	7.5
1NSD	5964	48.4	41.2	9.4	9.3	6.9
2UAG	6535	47.0	39.9	10.1	9.7	7.7
MurE	7418	48.2	41.1	10.0	9.7	7.7
1YVE	7833	49.1	41.9	10.3	10.1	8.2

[a] The models are defined in the text.

TABLE 5: Values of the Parameters A_H (Å^2), A_S (Å^2), H_0, and S_{max} in the Various Optimized Models[a]

model	parameter	H	C	N	O	X
MHybrid(Mull.)	A_S	1.0222	1.0222	1.0222	1.0222	1.0222
	A_H	2.7000	unused	unused	unused	unused
	S_{max}	1.49081	1.49081	1.49081	1.49081	1.49081
	H_0	0.8009	1.0	1.0	1.0	1.0
MHybrid(IV)	A_S	0.0139	0.0139	0.0139	0.0139	0.0139
	A_H	2.0875	unused	unused	unused	unused
	S_{max}	1.31515	1.31515	1.31515	1.31515	1.31515
	H_0	0.8124	1.0	1.0	1.0	1.0
MBest	A_S	0.5667	0.2625	0.0653	0.0542	0.5333
	A_H	3.9486	1.4167	0.0653	1.8500	3.5972
	S_{max}	1.0984	1.1967	0.9002	1.3014	0.6235
	H_0	0.2690	0.0431	0.0387	0.1552	0.0174

[a] The model names and parameter meanings are defined in the text. Atom type X refers to atoms other than H, C, N, and O.

solvation energies. To test this, we determined the polarization energies for the protein test set using the radii obtained from the UHBD calculations and those from our Mhybrid(Mull.) model. These results are plotted in Figure 4 along with the energy components G_1 and G_2 which were obtained by decomposing the solvation energy into one-body and two-body terms, i.e.

$$G_{pol} = G_1 + G_2 = -\frac{1}{2}\gamma\left(\frac{1}{\epsilon_I} - \frac{1}{\epsilon_O}\right)\left[\sum_i \frac{q_i^2}{f_{ii}} + \sum_{i\neq j}\sum_j \frac{q_i q_j}{f_{ij}}\right] \quad (26)$$

From Figure 4, it is apparent that the radii obtained from the model Mhybrid(Mull.) give solvation energies that are in poor agreement with those obtained with the UHBD radii. The differences are due to the two-body terms, G_2, as the one-body contributions to the energy, G_1, match reasonably well.

The failure of the MHybrid models caused us to examine systematically the reasons why, despite the good agreement for the Born radii, the solvation energies were not well produced. A number of problems became clear. First, the parameters G_{max} and G_0 were shown to be inadequate as the value of G_{max} was too large, resulting in the radii of some atoms being too small, and the value of G_0 was too small, meaning that the polarization energies of too many atoms were being modified from their natural values (see eq 15). Second, it became clear that while the value of the parameter D_s was fine for small molecules, it was too large for macromolecules, as the atoms are more densely packed and so too many atoms were being included in the calculation of the effective atomic volumes. As a result of these observations, we decided to reoptimize all the parameters in our model including those optimized with the small molecule set of data. We also decided to relax our restrictions on the A_H, A_S, H_0, and S_{max} parameters and optimize all of them separately for each atom type. The parameters that we obtained as a result of the optimization form the set that we call MBest and are listed in Tables 1 and 4. The parameters G_{max}, G_0, V_{min}, and V_0 were found to be -12 kcal mol^{-1}, -13 kcal mol^{-1}, $V_{full}/6$ Å^3, and $V_{full}/5$ Å^3, respectively. Some of the results with the new model are given in Figure 3 and in Table 4, where it can be seen that the MBest model is now able to reproduce the UHBD polarization energies and that the values of the $\mathscr{R}_M$ function are much better for all the proteins tested than those given by the other macromolecular models.

Having reoptimized the parameter set for the macromolecular case, we wanted to verify its performance on the small molecule test cases by recalculating the $\mathscr{R}_1$ and $\mathscr{R}_2$ functions using the MBest parameter set. Just for interest, we also did the calculations with the MBest parameter set but in which the macromolecular parameters were excluded (i.e., using only those listed in Table 1). The results are listed in Table 6. Not unexpectedly, the results are less good than those given by parameter sets specifically optimized for the small molecule test set. The effect is also more pronounced for the QM results than for the MM results. However, although less good for small molecules, the MBest parameter set has the great advantage that it can be employed with reasonable accuracy for both macromolecular and small molecule calculations using either MM or QM potentials, the latter with either a Mulliken or, preferably, a class IV-type population analysis. In any case, the availability of the other parameter sets means that one of these can be used if higher precision is needed in specific circumstances.

Conclusions

We have constructed and parametrized several GBSA implicit solvation models for use with MM, QM and hybrid QM/MM potentials and for calculations on molecular and macromolecular systems. All our models are continuous and differentiable, and so they can be employed straightforwardly for geometry

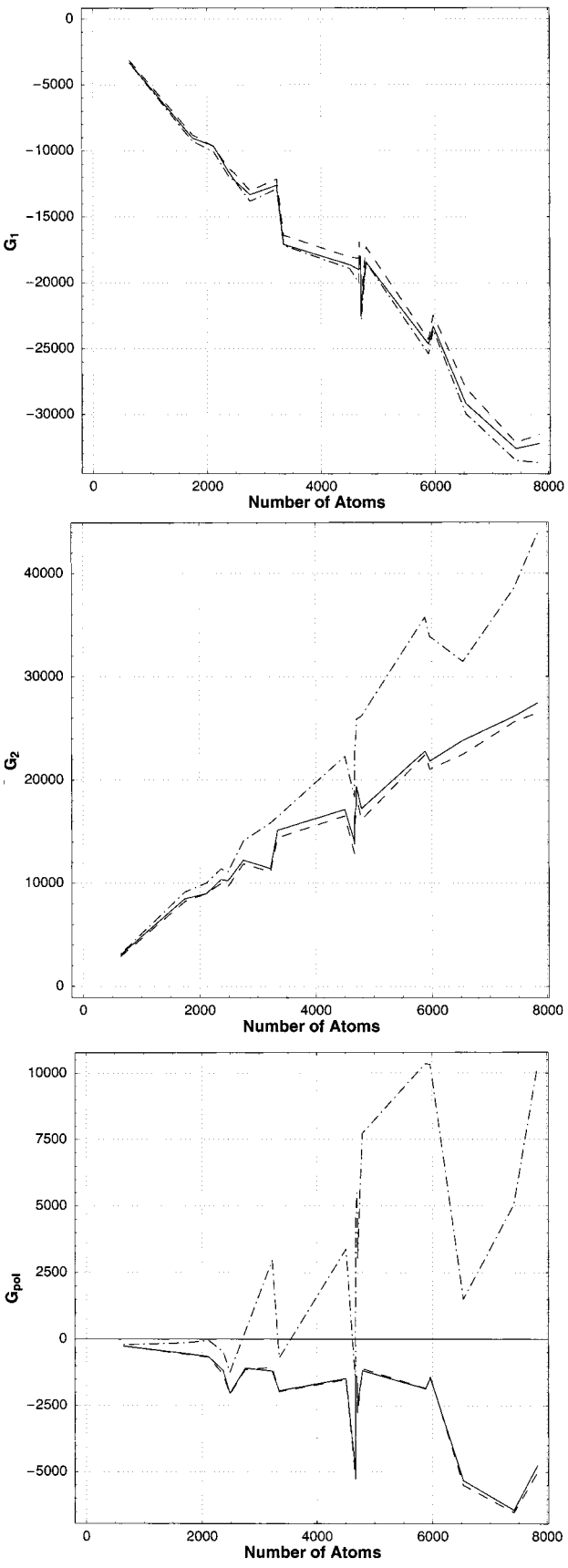

Figure 4. Plots of protein polarization energies versus the number of atoms in the protein: (a) G_1; (b) G_2; (c) G_{pol}. All energies are in kcal mol^{-1}. The polarization energies were calculated using Born radii obtained with different methods: reference UHBD values, full line; MHybrid(Mull.) radii, dotted−dashed line; Mbest radii, dashed line.

TABLE 6: Values of the $\mathcal{R}_1$ and $\mathcal{R}_2$ Functions (in kcal mol^{-1}) Calculated with the MBest Parameter Set (with and without Macromolecular Parameters) and with MM, QM(Mulliken), and QM(Class IV) Potentials

	$\mathcal{R}_1$			$\mathcal{R}_2$	
method	MM	Mulliken	class IV	Mulliken	class IV
MBest (no macromol. parameters)	3.6	4.8	4.6	4.3	4.1
MBest (full)	3.6	4.6	4.4	4.1	4.0

optimization and molecular dynamics simulations. The models differ in accuracy, depending upon how they were parametrized, but our most general model gives a good description of macromolecular atomic Born radii and solvation energies and yet also reproduces reasonably the MM and QM solvation energies of our test set of small molecules.

Our original aim in developing hybrid potential GBSA models was to have a method capable of quickly estimating the solvation energies in simulations of processes, such as enzyme reactions and protein−ligand binding, in which the substrate and part of the protein were treated quantum mechanically and the remainder of the system with a MM potential. We have already tested our GBSA models in some applications of this type, with encouraging results, and will report on these calculations and possible enhancements to the method in due course.

Acknowledgment. The authors would like to thank Otto Dideberg for providing the MurD and MurE structures and the Institut de Biologie Structurale−Jean-Pierre Ebel, the Commissariat à l'Energie Atomique and the Centre National de la Recherche Scientifique for support of this work.

Supporting Information Available: Tables showing the optimizations and experimental solvation free energies for the molecules. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford: 1987.
(2) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.
(3) Simonson, T. *Curr. Opin. Struct. Biol.* **2001**, *11*, 243.
(4) Constanciel, R.; Contreras, R. *Theor. Chim. Acta* **1984**, *65*, 1.
(5) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305.
(6) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
(7) Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. *J. Phys. Chem. B* **1997**, *101*, 1190.
(8) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005.
(9) Jayaram, B.; Sprous, D.; Beveridge, D. L. *J. Phys. Chem. B* **1998**, *102*, 9571.
(10) Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A. *Theor. Chem. Acc.* **1999**, *101*, 426.
(11) Onufriev, A.; Bashford, D.; Case, D. A. *J. Phys. Chem. B* **2000**, *104*, 3712.
(12) See, for example, Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2160 and references therein.
(13) Field, M. J. *J. Comput. Chem.*, in press.
(14) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York: 1995; Vol. 7., p 119.
(15) Amara, P.; Field M. J. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley and Sons: Chichester: 1998; Vol. 1, p 431.
(16) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
(17) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824.
(18) Kozaki, T.; Morihashi K.; Kikuchi, O. *J. Am. Chem. Soc.* **1989**, *111*, 1547.

(19) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(20) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.
(21) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209 and 221.
(22) Field, M. J. *A Practical Introduction to the Simulation of Molecular Systems*; Cambridge University Press: Cambridge: 1999.
(23) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088.
(24) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
(25) Wesson, L.; Eisenberg, D. *Protein Science* **1992**, *1*, 227.
(26) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Aid. Mol. Des.* **1995**, *9*, 87.
(27) Cramer, C. J.; Truhlar, D. G. *J. Comput. Aid. Mol. Des.* **1992**, *6*, 629.
(28) *Jaguar 3.5*; Schrödinger Inc.: Portland, OR, 1998.
(29) Protein Data Bank, Brookhaven National Laboratory, Upton, NY, 11973.
(30) Dideberg, O. Private communication.
(31) *Insight II*; Molecular Simulations Inc.: San Diego, CA, 1993.
(32) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States D. J.; Swaminanathan S.; Karplus M. *J. Comput. Chem.* **1983**, *4*, 187.
(33) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comput. Phys. Comm.* **1991**, *62*, 187.