PROTEINS: Structure, Function, and Genetics 45:144–158 (2001)

Protein Molecular Dynamics With the Generalized Born/ACE Solvent Model

Nicolas Calimet,[1] Michael Schaefer,[2*] and Thomas Simonson[1*]
[1]*Laboratoire de Biologie et Génomique Structurales (CNRS), Institut de Génétique et Biologie Moléculaire et Cellulaire, Strasbourg-Illkirch, France*
[2]*Laboratoire de Chimie Biophysique, Institut Le Bel, Université Louis Pasteur, Strasbourg, France*

ABSTRACT **Implicit solvent models are increasingly important for the study of proteins in aqueous solution. Here, the generalized Born (GB) solvent polarization model as implemented in the analytical ACE potential [Schaefer and Karplus (1996) J Phys Chem 100:1578] is used to perform molecular dynamics simulations of two small, homologous proteins: the immunoglobulin-binding domain of streptococcal protein G and the Ras binding domain of Raf. Several model parameterizations are compared through more than 60 ns of simulation. Results are compared with two simpler solvent models—an accessible surface area model and a distant-dependent dielectric model, with finite-difference Poisson calculations, with existing explicit solvent simulations, and with experimental data. The simpler models yield stable but distorted structures. The best GB/ACE implementation uses a set of atomic Voronoi volumes reported recently, obtained by averaging over a large database of crystallographic protein structures. A 20% reduction is applied to the volumes, compensating in an average sense for an excessive de-screening of individual charges inherent in the ACE self-energy and for an undersolvation of dipolar groups inherent in the GB screening function. This GB/ACE parameterization yields stable trajectories on the 0.5–1-ns time scale that deviate moderately (~1.5–2.5 Å) from the X-ray structure, reproduce approximately the surface distribution of charged, polar, and hydrophobic groups, and reproduce accurately backbone flexibility as measured by amide NMR-order parameters. Over longer time scales (1.5–3 ns), some of the protein G runs escape from the native energy basin and deviate strongly (3 Å) from the native structure. The conformations sampled during the transition out of the native energy basin are overstabilized by the GB/ACE solvation model, as compared with a numerical treatment of the full dielectric continuum model. Proteins 2001;45:144–158.**

Key words: solvation; protein structure; dielectric continuum; computer simulation

INTRODUCTION

Molecular dynamics simulations are a powerful tool for the study of biological molecules in solution. They are used routinely for conformational searching, refinement of structural models versus nuclear magnetic resonance (NMR) or crystallographic data, and generating thermodynamic ensembles to calculate free energies.[1,2] In many cases, a large proportion of the computing time is spent on the explicit treatment of aqueous solvent surrounding the solute of interest. To avoid this, considerable effort has gone into developing implicit solvent models for molecular dynamics.[3–5] Examples are surface area models[6,7] and Gaussian solvent-exclusion models.[8] Another promising class of models treats the solvent as a dielectric continuum.[5,9–11] The latter models have been used abundantly within the context of finite-difference Poisson–Boltzmann calculations, for both small organic molecules[9,12–15] and proteins.[5,10,11]

In addition to the applications mentioned above, molecular dynamics with a dielectric continuum solvent could replace current finite-difference Poisson–Boltzmann (FDPB) approaches for many problems, such as proton binding and redox properties. Indeed, current FDPB calculations are usually performed for only one or a few protein structures; averaging over a statistical ensemble is only implicit, through a mean field approximation.[16,17] It would be explicit with molecular dynamics (MD) and should yield a more accurate description of dielectric relaxation of the solute on an atomic scale.[18–20]

In addition to electrostatic solvation of polar and ionized atom groups, hydrophobic interactions yield an important contribution to the stability and dynamics of macromolecules. Thus, continuum electrostatic models have often been combined with surface area treatments of nonelectrostatic solute–solvent interactions.[13,14] This works well for linear and moderately branched saturated alkanes, but rather poorly for cyclic saturated alkanes,[14] and is of uncertain quality for large solutes such as proteins. More complex hydrophobic models were recently proposed[21,22] but are probably too expensive to be used with MD in the near future.

Grant sponsor: CNRS.

*Correspondence to: Thomas Simonson, Laboratoire de Biologie et Génomique Structurales (CNRS), Institut de Génétique et Biologie Moléculaire et Cellulaire, 1 rue Laurent Fries 67404 Strasbourg-Illkirch, France. E-mail: simonson@igbmc.u-strasbg.fr, or Michael Schaefer, Laboratoire de Chimie Biophysique, Institut Le Bel, 4 Rue Blaise Pascal, 67000 Strasbourg, France. E-mail: schaefer@piaf.u-strasbg.fr.

Received 26 February 2001; Accepted 11 May 2001

The molecular dynamics of large polypeptides with dielectric continuum solvent remains a challenging technique.[12,23–26] Continuum electrostatic forces are usually obtained from the Maxwell stress tensor,[16,27,28] which involves solving the Poisson–Boltzmann equation numerically. While this is faster than simulating explicit solvent, it is still costly and complex, and the first application to a protein was only reported very recently.[29] A simpler model was proposed 10 years ago by Still et al.,[30] known as the generalized Born (GB) model. It uses a simple screening function to renormalize charge–charge interactions in a way that depends on their proximity to solvent; it reproduces exactly the continuum solvent screening for a pair of very distant or very close charges (of nonzero sum). The "proximity to solvent" of a charge is measured by its "self-energy," i.e., its interaction energy with the solvent.[31] In principle, to calculate the self-energies, one must also solve the Poisson–Boltzmann equation numerically,[30,32] so there is little gain compared with a full dielectric continuum treatment. The key to efficient GB models has been the development of analytical approximations for the self-energies,[33–37] leading to fully analytical solvation models. Four analytical approximations have been used thus far.[34–37] The association of the GB screening function with a particular analytical approximation for the solvation radii is referred to as a GB model in what follows.

For charges highly exposed to solvent or deeply buried inside a solute, the GB screening function is expected to be reasonably accurate. For example, good agreement with FDPB calculations, MD with explicit solvent, and experimental data are obtained for the solvation of small organic molecules.[30,34–36,38–41] However, for moderately buried charges, the physics of the GB models may depart from the full dielectric continuum model. Therefore, careful parameterization and testing are needed to determine the usefulness of GB for large biomolecules. Current GB variants have given good results for solvation and conformational free energies of polypeptides,[42] nucleic acids,[43] and, to a lesser extent, proteins.[35,37,40,41,44,45] Recently, GB was used for the first time in molecular dynamics simulations of DNA and RNA[46,47] and of three proteins.[29,41,48] MD provides a stringent test of the model, as not only the average electrostatic potential, but the local electric field as well, must be reasonably accurate throughout the structure, with a correct balance between intra-solute and solute–solvent interactions. The protein simulations[29,41,48] used the solvation radii of Qiu et al.,[36] with numerical parameters adjusted against a library of fragment solvation energies. Conformations close to the crystal structure were sampled. However, additional studies are needed to determine the robustness of the method and its sensitivity to different parameterizations.

The purpose of this article is to test the GB approach further for protein molecular dynamics. We focus on a variant of the model based on an approximation for the solvation radii developed by Schaefer and Karplus,[35] called analytical continuum electrostatics (ACE). This variant was used to study the folding of small polypeptides[42,49] but has never been used for protein MD. It is referred to in this report as GB/ACE, to distinguish it from the other, substantially different GB variants.[34,36,37] We apply it to simulations of two small, homologous proteins (Fig. 1): domain B_1 of streptococcal protein G (protein G) and the Ras binding domain of Raf (Raf-RBD). We combine it with the hydrophobic term proposed by Schaefer et al.,[42] which is similar to a surface area treatment. Several parameterizations of the model are compared. We make use of two different sets of atomic volumes: the original set, referred to as V98[42] and a more recent set, referred to as V01.[50] The latter were developed by averaging atomic Voronoi volumes over a large set of proteins from the Protein Data Bank (PDB). The two volume sets are combined with one or two scaling factors for the volumes of selected atom types. Three possible values for the hydrophobic "surface tension" are compared. The model is compared with two simpler, commonly used models—a distant-dependent dielectric model (RDIE) and an atomic surface area model (ASP),[6,7] with finite-difference Poisson calculations (FDP), with explicit solvent simulations,[51] and with experiments.

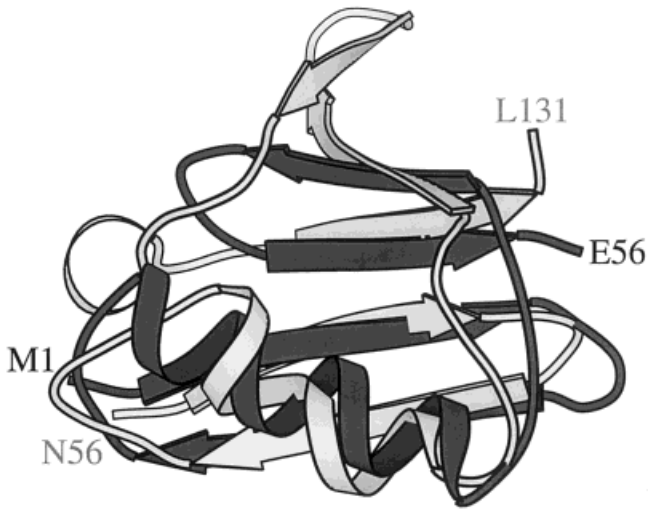

Fig. 1. Superposition of protein G (dark) and Raf-RBD (light). The chain termini are labeled by their residue type and number. Produced with Molscript.[72]

The following section describes the solvent models and computational methods. Simulations with the original GB/ACE parameterization are then presented under Results to determine the main sources of error in the solvation energy term, as well as the optimal hydrophobic surface tension. The parameterization is then modified by scaling selected atomic volumes. Finally, the recent set V01 of atomic volumes[50] is tested, along with different scaling factors. The best parameterizations reduce the atomic Voronoi volumes uniformly by 10–20%. This compensates in an average sense for an excessive de-screening of surface charges inherent in the ACE self-energy term, and for an undersolvation of dipolar groups inherent in the GB screening function (see Discussion). Stable trajectories are obtained for both proteins on the nanosecond time scale. The root-mean-square deviation (RMSD) from the crystal structure is ~2–3 Å with the earlier set of volumes, and 1.5–2.5 Å with the recent set when combined with a

suitable reduction factor. The ASP and RDIE models yield poorer structures despite similar RMSD.

METHODS

The GB/ACE Solvent Model

In continuum electrostatics, the solute is viewed as a set of (fractional) atomic charges in a cavity delimited by the solvent accessible surface, embedded in a high dielectric solvent medium.[9] The solvation free energy can be expressed as a sum of the interaction energies of each solute charge with solvent (their "self-energies," see below), and the solvent-screened interaction energies between all pairs of solute charges. In the GB approach, the interaction energy E_{ij}^{int} between charges q_i and q_j is approximated by the following expression[30]:

$$E_{ij}^{\text{int}} = \frac{q_i q_j}{r_{ij}} - (1 - 1/\epsilon_w)\frac{q_i q_j}{[r_{ij}^2 + b_i b_j \exp(-r_{ij}^2/4b_i b_j)]^{1/2}} \quad (1)$$

where r_{ij} is the distance between the charges, ϵ_w is the solvent dielectric constant, b_i is the effective solvation radius of charge i (respectively j):

$$b_i = -\frac{(1 - 1/\epsilon_w)q_i^2}{2\Delta E_i^{\text{self}}} \quad (2)$$

and ΔE_i^{self} is the self-energy of charge i (the difference between its free energy inside the solute cavity in solution, in the absence of the other solute charges, and its free energy in the gas phase). If the charge q_i were at the center of a sphere of radius b_i, its self-energy would be ΔE_i^{self}.[52] Thus, b_i can be thought of as the charge–solvent distance.

In addition to the screening function (eq. 1), the other main approximation of current GB models concerns the charge self-energies. Up to a small error,[53] the self-energy of a charge can be calculated by an expression involving an integral of the square electric field E^2 over the volume of the solute. All existing analytical approximations start by replacing E by the field E_C that would exist in the absence of any dielectric boundaries, i.e., the Coulombic field, as proposed by Schaefer and Frömmel.[33] It is then possible to construct an approximation for the function $E_C^2 \propto 1/r^4$ integrated over the solute volume. In the ACE approach, Schaefer and Karplus introduced gaussian atomic volumes, thus expressing the integral over the solute volume as a sum of integrals over the volumes of individual atoms[35,42]:

$$\Delta E_i^{\text{self}} = -\frac{(1 - 1/\epsilon_w)q_i^2}{2R_i} + \sum_{k \neq i} E_{ik}^{\text{self}} \quad (3)$$

where R_i is the radius of atom (charge) i (in general, R_i is the van der Waals radius[42]), and the integral expression E_{ik}^{self} is approximated by

$$E_{ik}^{\text{self}} = \frac{(1 - 1/\epsilon_w)q_i^2}{\omega_{ik}} \exp(-r_{ik}^2/\sigma_{ik}^2) + \frac{(1 - 1/\epsilon_w)q_i^2 V_k}{8\pi}\left(\frac{r_{ik}^3}{r_{ik}^4 + \mu_{ik}^4}\right)^4 \quad (4)$$

Here, ω_{ik}, σ_{ik}, and μ_{ik} are simple functions of the atomic volume V_k, the atomic van der Waals radii R_i, R_k, and of an adjustable "smoothing" parameter α that rescales the width of the atomic gaussian distributions.[42]

A hydrophobic term, W_h, proposed by Schaefer et al.,[42,49] was also added to the energy function to approximate the nonpolar contribution to the solvation free energy. It is proportional to the self-energy term (eq. 3) and has been shown to correlate closely with the total surface area S of the solute.[42,49] Therefore, it can be viewed approximately as a surface area term, $W_h \approx \sigma S$, where σ is an effective surface tension. The calculation of the hydrophobic energy term is very efficient because the self-energy is already available as part of the electrostatic energy. Three different values were assigned to σ; for the proteins considered in this work, they correspond to surface tensions of 0, 26, and 51 cal/mol/Å^2; 26 cal/mol/Å^2 corresponds to the average value for the water → cyclohexane transfer of small aliphatic hydrocarbons.[54,55]

In what follows, the atomic charges q_i are taken from the CHARMM param19 molecular mechanics force field[56]; the radii R_i are set equal to the atomic van der Waals radii (corresponding to the minimum in the van der Waals potential[57]). The remaining model parameters are the atomic volumes V_i and the smoothing parameter α. The volumes are taken from one of two libraries: the recent library, termed V01, was obtained by averaging Voronoi volumes over a large database of protein structures[50]; the earlier volume library V98 is as given in Ref. 42. The more recent volumes will be reported in the literature shortly[50] and are currently available from Schaefer et al.[50] on request. Overall scaling factors were applied to some or all of the atomic volumes in some simulations. In conjunction with the atom volume sets V98 and V01, the smoothing parameter α is set equal to 1.2 and 1.3, respectively.[42,50]

In a few of the simulations, an additional approximation was made: the atomic self-energies (eq. 3) were calculated once for the starting structure and then treated as constants. This simplifies the analytical form of the energy gradients and speeds up the calculations by about a factor of 2.

Atomic Surface Area Model ASP

In the atomic surface area (ASP) model, a solvation energy $W(r_1, r_2, \ldots, r_n)$ is added to the usual vacuum energy function:

$$W(r_1, r_2, \ldots, r_n) = \sum_i \sigma_i A_i(r_1, r_2, \ldots, r_n) \quad (5)$$

where r_i and A_i are the position and the solvent accessible surface area of atom i, and σ_i is an atomic "surface tension," which depends on the atom type. The surface tensions, taken from Wesson and Eisenberg,[6] result from fitting solvation free energies of small organic molecules,[58] and implicitly include both electrostatic and nonelectrostatic contributions.

Distance-Dependent Dielectric Model RDIE

In the distance-dependent dielectric (r-dependent dielectric) model (RDIE) electrostatic interactions between protein atoms are reduced by the screening function $\epsilon(r) = r$, leading to an interaction energy that scales as $1/r^2$. In

addition, the charges of ionized groups are reduced by a factor of four in some runs. This last variant has sometimes been used for NMR structure refinements.[59]

Preparation of the Structure, Force Field

The crystal structure of domain B_1 of streptococcal protein G was obtained from the Protein Data Bank (entry 1PGB). Raf-RBD was taken from the crystal structure of the Rap:RBD complex (PDB entry 1GUA). Hydrogens of polar groups were positioned based on stereochemical considerations[60]; aliphatic carbons and their associated hydrogens were treated as united atoms. Ionizable groups, which are all solvent exposed in protein G and Raf-RBD, were modeled in their most common charge states at pH 7. The total net charge of the proteins is then -4 and $+6$, respectively.

The CHARMM param19 force field was used,[56,61] along with the implicit solvent models described above. Electrostatic and van der Waals interactions were switched to zero between 13 and 16 Å separations. Covalent bond lengths involving hydrogens were held fixed with the SHAKE algorithm.[62] The protein dielectric constant was set to 1.

Simulation Protocol

The X-ray structure was subjected to 300 steps of minimization with an adopted basis Newton Raphson method, including the GB/ACE solvation energy term. Harmonic restraints with a 40-kcal/mol/Å^2 force constant were applied to nonhydrogen atoms to limit departures from the initial structure. The final RMSD from the crystal structure was 0.05 Å for all atoms. This starting structure was used for all the simulations reported below, regardless of the solvent model.

The minimized structure was used to start molecular dynamics at either 0 or 50 K; velocities were then rescaled every 0.6 ps to a target temperature, which was increased at each rescaling by 10 K. At 298 K, an equilibration phase was begun, using Langevin dynamics with a friction coefficient of 10 ps^{-1} for non-hydrogen atoms. Overall rotation and translation were removed every 20 ps. The harmonic restraints used during minimization were decreased gradually over a period of 250 ps (unless otherwise mentioned), following one of two possible schedules. In the first, the force constant was decreased in five equal steps to a final value of zero for side-chains and 1 kcal/mol/Å^2 for backbone atoms. In the second, the force constant was decreased in 10 steps by applying a multiplicative factor of 0.5 for backbone and 0.33 for side-chain atoms. The final restraint energy with the second schedule was only ~20 kcal/mol. Restraints were then removed, and production carried out for 1–3 ns, using Langevin dynamics with the same friction coefficient as during equilibration.

For the GB/ACE runs initiated at 0 K, this preparation phase was done once with each surface tension (i.e., 0, 26, and 51 cal/mol/Å^2); subsequent runs with a given surface tension were started from the corresponding equilibrated structure. To test the dependence of the results on the initial conditions, several trajectories were calculated for each σ, starting from the same structure but scaling the initial velocities by ±3%. The strong divergence of the trajectories of multi-body systems such as proteins as a consequence of infinitesimal changes in the initial conditions is then sufficient to randomize the structures in a few tens of picoseconds. For the runs initiated at 50 K, each run used a different random number seed for the starting velocities. For the ASP and RDIE runs, equilibration was performed once, and subsequent runs used scaled velocities. A few GB/ACE runs were also started at room temperature and used Verlet dynamics instead of Langevin dynamics, demonstrating that the starting temperature and temperature control mechanism do not bias the results significantly (data not shown). Calculations were done with the CHARMM program (versions 27a1 and 28a1).[56]

Finite-Difference Poisson Calculations

For comparison with GB/ACE, electrostatic solvation free energies were calculated by numerical solution of the Poisson equation, using the finite-difference program UHBD,[63] i.e., by a numerically accurate treatment of the continuum electrostatic solvent model. In the finite-difference approach, the spatial distribution of the dielectric constants, the atomic partial charges, and the electrostatic potential are described on a three-dimensional grid, and the discretized Poisson equation is solved iteratively.[64] The electrostatic solvation free energy is obtained as

$$\Delta E^{\text{solv}} = \frac{1}{2}\sum_i q_i(\phi_s(r_i) - \phi_g(r_i)) \qquad (6)$$

where the sum is over all atoms, q_i and r_i are the charge and position of atom i, and ϕ_g and ϕ_s are the potentials of the protein in the gas phase and in solution, respectively. To be consistent with the parameters of the GB/ACE potential, the protein and the solvent dielectric constants were set to 1 and 80, and the atomic partial charges were the same as in the MD simulations. The finite-difference calculations used the focusing technique[65]; in the first grid setup, the protein was surrounded by ≥20 Å of "water" (dielectric) on each side; in the second grid setup, it was surrounded by 5 Å of water. The grid spacing of the initial and final grids were 1.5 and 0.5 Å, respectively. The Shrake and Rupley dot surface[66] was calculated with 2,500 points per atom and used to define the solute–solvent boundary, along with the boundary smoothing method in UHBD.[67] To test the dependence of the results on the grid spacing, finite difference calculations of the solvation free energy were also performed for 200 structures of protein G using final grid spacings of 0.3 and 0.4 Å. Changing the grid spacing changed the solvation energies by less than 1% (~10 kcal/mol), so that there is no significant influence of grid spacing on the following analysis.

RESULTS

Optimization of the GB/ACE Model Parameters

In this section, we describe simulations with the GB/ACE solvent potential, and modifications of the volume

TABLE I. Simulations Performed With Each Solvent Model*

Solvent model	Atomic volumes	Scaling factor[a]	Surface tension (cal/mol/Å^2)	Final RMSD[c] (Å)
		Protein G		
GB/ACE	V98	1.0	0	2.9
GB/ACE	V98	1.0	26	1.9, 2.1, 2.7, 3.2, 4.0
GB/ACE	V98	1.0	51	2.0, 2.8, 2.8, 3.1, 3.6
GB/ACE	V98	1.15, 0.85[b]	26	1.9, 2.8, 3.1
GB/ACE	V98	1.3–1.6, 0.5–0.7[b]	26	3.0, 3.1, 3.5, 3.7
GB/ACE	V98	0.8	26	2.4, 3.2
GB/ACE	V98	0.9	26	1.9, 2.4, 2.6, 2.9, 3.1
GB/ACE	V01	1.0	26	2.2, 3.8
GB/ACE	V01	0.9	26	2.4
GB/ACE	V01	0.8	26	1.4, 1.6, 2.0, 2.8
RDIE			0	1.1, 1.2, 2.0,[d] 2.2[d]
ASP			—	2.3–2.7
Explicit solvent[f]				1.2
		Raf-RBD		
GB/ACE	V98	1.0	0	1.9
GB/ACE	V98	1.0	27	1.7, 1.8, 3.6, 3.9
GB/ACE	V98	1.0	53	2.7
GB/ACE-CR[e]	V98	1.0	27, 53	2.9, 2.9, 3.2
GB/ACE	V98	0.9	27	2.2
GB/ACE	V01	0.9	27	3.4
GB/ACE	V01	0.8	27	2.1

*List of simulations performed with different solvent models. The root-mean-square deviation (RMSD) relative to the crystal structure is averaged over C_α values and over the last 50 ps of the simulation.
[a]Applied to atomic volumes in the GB/ACE models.
[b]Scaling factors applied only to the volumes of lysine NZ and of carboxylate oxygens. The large values (1.15–1.6) apply to NZ, the small values (0.5–0.85) apply to the oxygens.
[c]Each number refers to a separate run.
[d]These runs used the full charges for ionized groups. The other RDIE runs used charges reduced by one-fourth.
[e]Effective solvation radii treated as constants (see text).
[f]From Sheinerman and Brooks[51]; averaged over all backbone atoms.

parameter set that improve the accuracy of the potential for proteins. In the following sections, we describe in more detail the behavior of the model using its best parameterization and compare it with two other models—the RDIE model and the ASP model, as well as with finite-difference Poisson calculations (FDP), with explicit solvent simulations, and with experiment. To assess the quality of the models, it is important to use a variety of criteria. We report RMSD from the protein crystal structures, solvent accessible surface areas, and backbone amide NMR-order parameters. We also rely on visual inspection of the trajectories to detect significant local distortions that may not be evident from the RMSD.

The simulations performed with GB/ACE are listed in Table I. Three principal groups of parameterizations were used: (1) the V98 atomic volume set[42] combined with various parameters σ ("surface tensions") for the hydrophobic energy term; (2) the V98 atomic volume set combined with one global or two atom-type-dependent scaling factors and the best hydrophobic parameter; (3) the more recent V01 volume set[50] combined with a global scaling factor and the best hydrophobic parameter from the previous simulations. The results for protein G and Raf-RBD are similar; protein G is discussed first.

Protein G

The RMSD from the crystal structure, averaged over C_α values, are reported in Table I for all runs and plotted as a function of time in Figure 2 for selected runs. A 2-ns explicit solvent simulation with the same protein force field gave a backbone RMSD of 1.2 Å.[51] Most of the GB/ACE protein G runs with the unscaled V98 volumes are unstable on the nanosecond time scale, with rms deviations increasing by abrupt steps to 2.5–4 Å. They usually exhibit severe local distorsions of the protein core [Fig. 4(a)], where Glu27 in the middle of the α-helix folds back on itself and becomes completely buried in the center of the protein. Less severe, but related distortions are observed for most other carboxylate groups, which reorient to interact systematically with electrically neutral hydrogen bond donors such as the backbone amide groups. Positive lysine side-chains are not available for interactions because they are strongly oversolvated, pointing directly away from the protein bulk and into solvent [Fig. 4(a)]. Residues from the hydrophobic core, such as Trp43, shift outward and become solvent exposed. The stable GB/ACE runs with these parameters (3 out of 12 runs) exhibit similar features, although Glu27 remains in place longer (1.5–2.5 ns) before shifting into the protein core. During the stable segments of these trajectories, the total

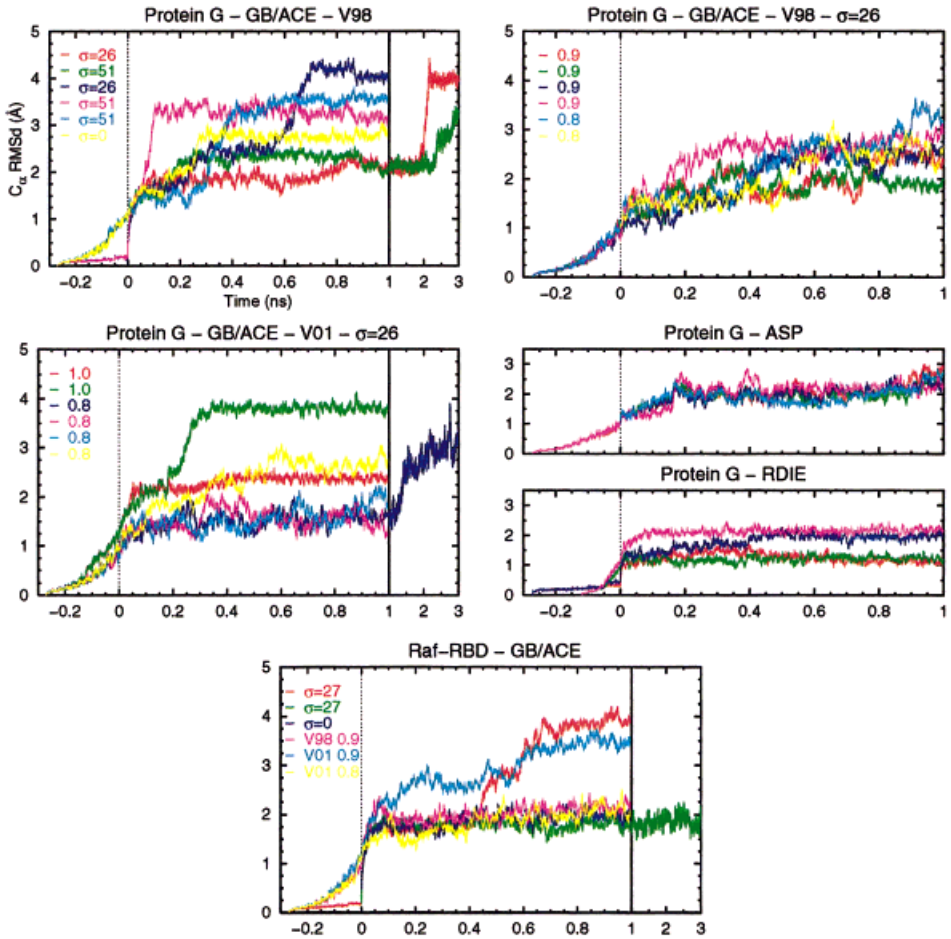

Fig. 2. Root-mean-square deviations (Å) from the crystal structure for C_α as a function of time. The bottommost panel shows Raf-RBD data; the other panels show protein G data. Panel headings indicate the solvent model used. Individual GB/ACE curves are labelled with the hydrophobic parameter σ and/or the volume scaling factor. The time origin (dashed vertical line) corresponds to the release of all harmonic restraints and the beginning of the production period.

solvent accessible surface of the protein is well reproduced (Fig. 3), although hydrophobic groups are overexposed and charged and neutral polar groups are underexposed, producing compensating errors. The longest stable segments and the best accessible surface areas are obtained with an effective hydrophobic surface tension of 26 cal/mol/Å^2. This value is used in all subsequent GB/ACE runs (i.e., the runs performed with scaled V98 volumes or the V01 volumes).

As a simple way to reduce the most obvious structural distorsions, the V98 atomic volumes were modified by applying one or two scaling factors. A similar strategy was used by other investigators[41,44] with the GB variants of Qiu et al.[36] and Hawkins et al.[34,38] Either a single overall factor is applied to all atomic volumes, or two distinct factors are applied to the carboxylate oxygens (to shrink them) and to the ammonium nitrogens (to enlarge them). The atomic smoothing parameter α (see discussion following eq. 4) was not varied. From the RMSD with respect to the crystal structure and from visual inspection of the structures, the best results were obtained with a single overall scaling factor of 0.9, reducing all the V98 atomic volumes by 10%. This reduction leads to increased protein–solvent interactions and decreased protein–protein interactions. The V98 volumes scaled by a factor 0.9 are referred to as the V98sc set. The resulting structures were much improved, with smaller RMSD of 2–3 Å and more realistic carboxylate behavior. In particular, Glu27 remained solvated at all times and several Asp side-chains transiently formed salt bridges to nearby lysines. There were no large abrupt increases in the RMSD. The total solvent accessible surface area is about 8% too high. The areas associated with polar and hydrophobic groups are within 5% of the experimental results, while charged groups are about 27% more exposed than in the crystal structure.

With the more recent V01 atomic volumes and no scaling, results are poor [Fig. 4(c)]. However results are quite good if an overall 0.8 scaling factor is applied, reducing the volumes by 20% (V01sc). One run is unstable, but in the other three runs, stability is good, backbone rms deviations are small (1.5–2 Å) and roughly constant over time, and the general aspect of the structures is reasonable [Fig. 4(d)]. Solvent accessible surface areas (total, polar, hydrophobic, and ionic) are in fair agreement with the

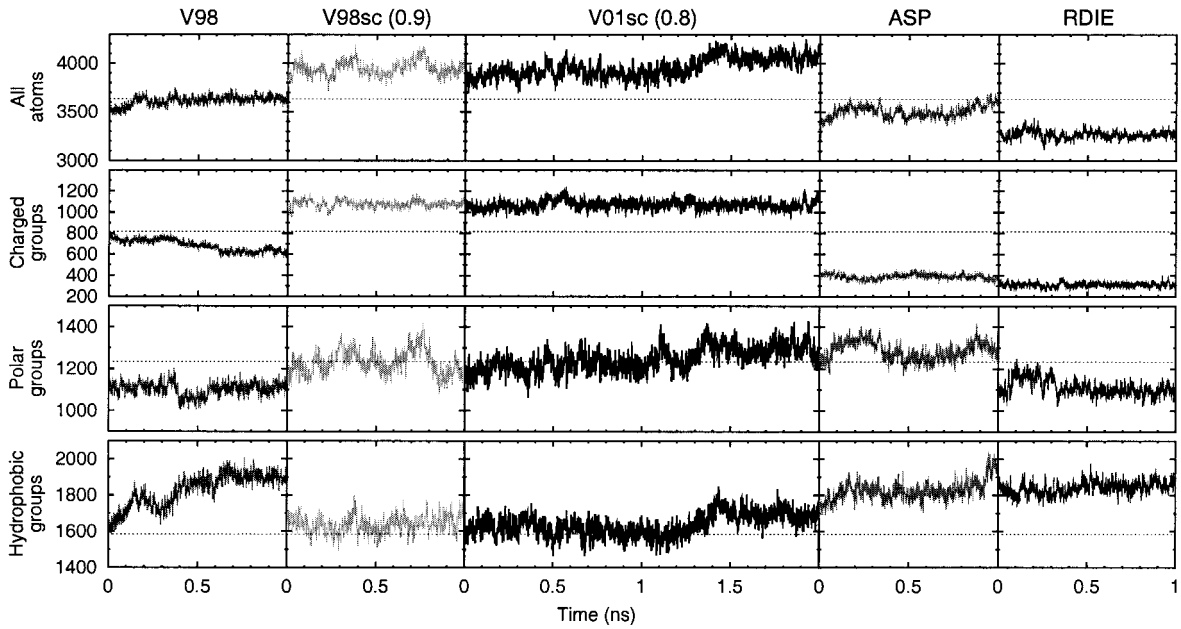

Fig. 3. Protein G solvent accessible surface area (Å^2) as a function of time during selected simulations, performed with different GB/ACE parameterizations, with RDIE, and with ASP. Different panels show sums over all atoms, polar atoms, charged side chain groups, and hydrophobic atoms, as indicated. Dotted horizontal lines correspond to the crystal structure. The V01sc run deviates sharply from the crystal structure after 1.3 ns (see text).

crystal structure (Fig. 3), similar to the behavior obtained with the V98sc volumes.

Extending one of the V01sc runs beyond 1 ns, the structure deteriorated sharply after 1.3 ns, similar to some of the V98sc runs above. The time dependence of selected energy terms is shown for this run in Figure 5. The sum of the potential and solvation energies is very stable, fluctuating around its initial value with no visible drift. Note that the total potential energy includes contributions from bonded terms, in addition to the van der Waals and Coulomb terms. The structural instability coincides with a decrease in the self-component of the solvation energy and a compensating increase in the van der Waals energy, so that the overall energy is essentially constant. The screening and Coulombic energy components do not shift significantly. At the same time, the total accessible surface area increases (Fig. 3) as a result of increases in its hydrophobic and polar components, while the accessible area of charged groups remains constant. This behavior is consistent with a loosening of packing and a slight expansion of the structure.

The accuracy of the GB/ACE solvation energy is analyzed further for this run and several others below, by comparison with finite-difference Poisson calculations.

RBD

For Raf-RBD, six GB/ACE runs were performed with the unscaled V98 atomic volumes. Three were stable on the nanosecond time scale, with final RMSD of 1.7–1.9 Å from the crystal structure (Fig. 2). One of these used a hydrophobic "surface tension" of 0 cal/mol/Å^2; two used 27 cal/mol/Å^2. Extending one of the latter at up to 3 ns, it remained stable, with a roughly constant RMSD. Thus, the behavior of this GB/ACE parameterization for this homologous, but slightly larger protein appears to be slightly better than with protein G. Note that the position equivalent to the protein G Glu27 is occupied by Leu78, and there are no glutamates in the α-helix. Three additional runs treated the effective atomic solvation radii as constants (see Methods), an approximation that speeds up the method by about a factor of 2. These runs behaved significantly worse (Table I).

Figure 2 also shows a run with the V98sc volumes, and two runs combining the V01 volumes with two possible scaling factors: 0.8 and 0.9. The V98sc run and the V01 run with the optimal scaling factor (0.8) are very well behaved. The V01 run with a 0.9 scaling factor is unstable, shifting rapidly away from the starting structure.

Comparison With the RDIE and ASP Solvent Models

Surprisingly, the backbone RMSD with the ASP and RDIE models are of the same magnitude as the stable GB/ACE runs, or even smaller. With RDIE, two of the four runs actually have RMSD of $<$1.5 Å over the whole production period (1 ns). Analysis of the structures shows that despite these small deviations, there are significant structural distorsions; in particular, the structures are too compact and rigid, and reproduce poorly the distribution of charged groups on the protein surface (see below). Similar observations have been reported for other proteins.[68] The five ASP runs all have rms deviations of ~2 Å during most of the production. Similar rms deviations were obtained by Fraternali and van Gunsteren[7] for three other small proteins during 500 ps simulations with a similar, surface area-based solvation model.

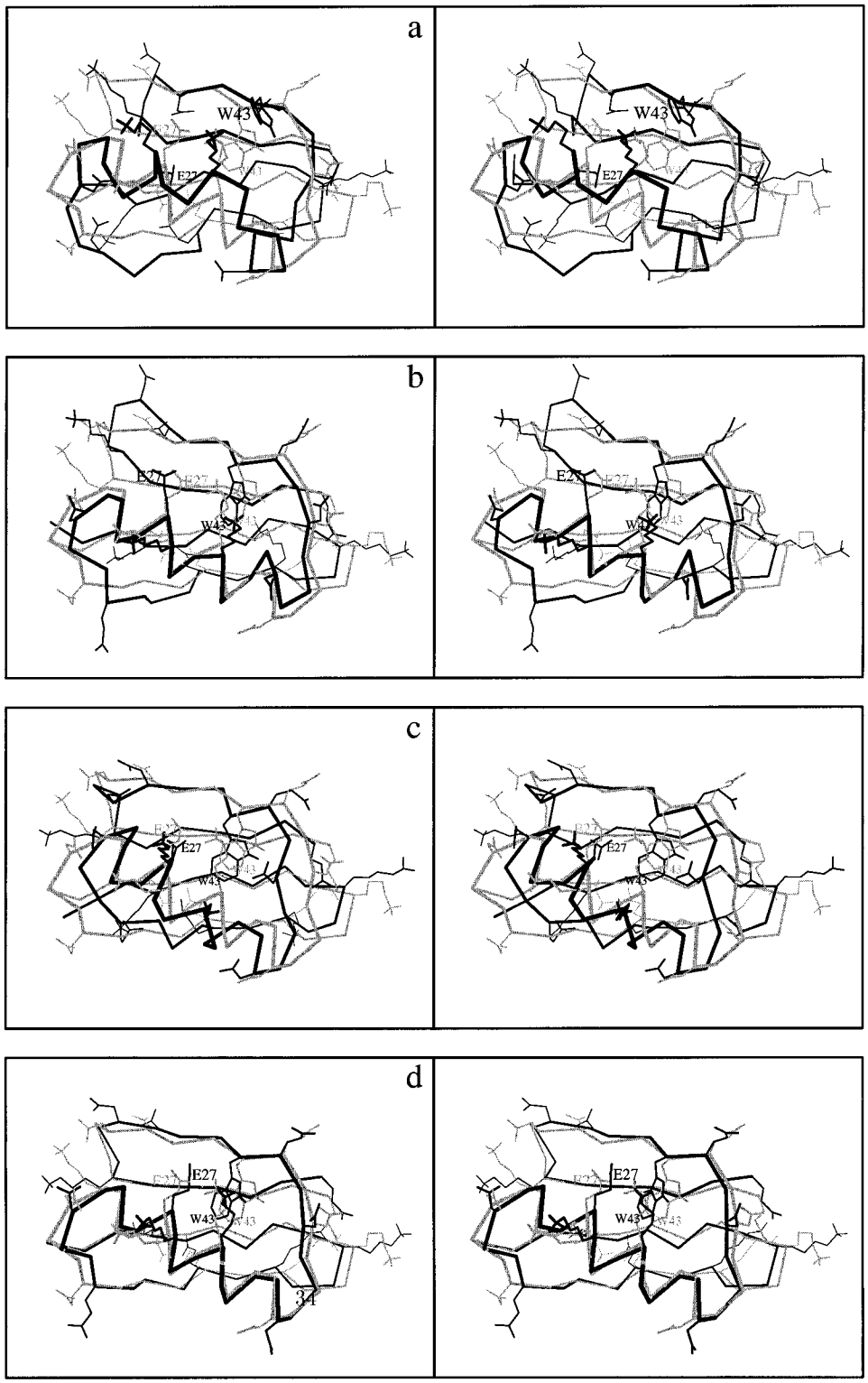

Fig. 4. Typical snapshots after 1 ns of production with GB/ACE using (**a**) the V98 volumes; (**b**) the V98sc volumes; (**c**) the V01 volumes; (**d**) the V01sc volumes. Crystal structure in gray. Stereo views showing the backbone and selected side-chains. With the unscaled V98 and V01 volumes, Glu27 becomes buried. Produced with Molscript.[72]

The solvent accessible surface area is reported as a function of time for ASP and RDIE simulations in Figure 3. The total surface area, as well as the areas corresponding to polar, charged, and hydrophobic groups, are shown. The ASP and especially the RDIE models give much poorer descriptions of the accessible surface area than GB/ACE.

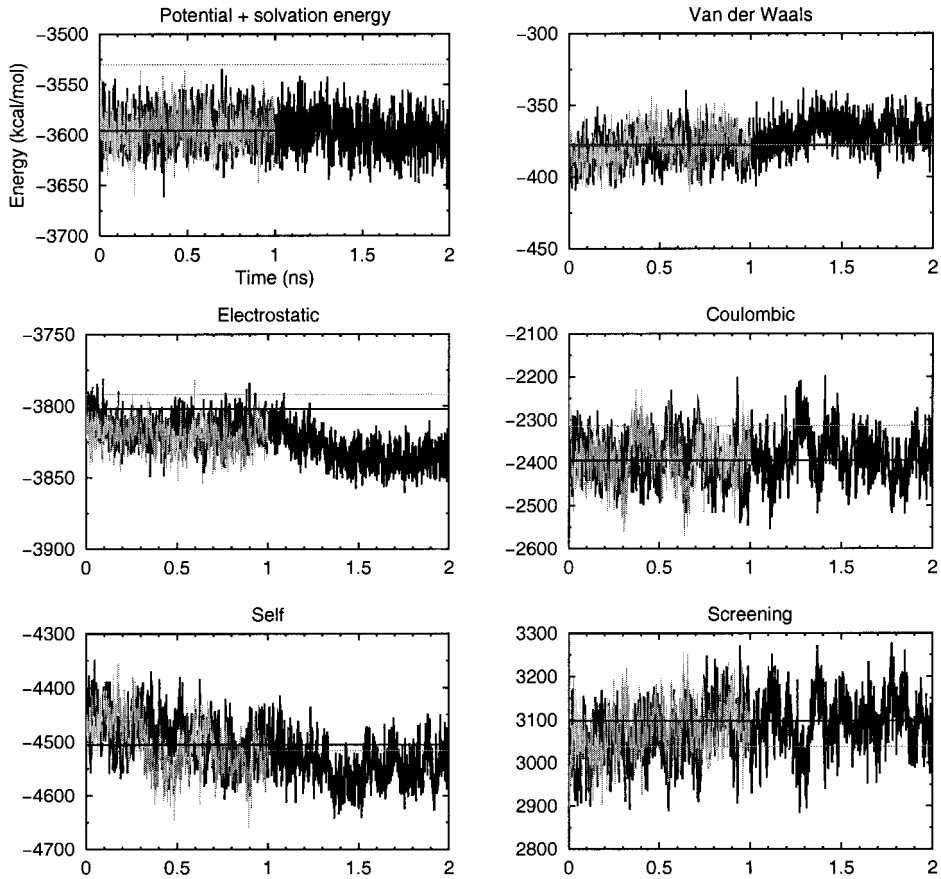

Fig. 5. Selected energy terms vs. time during a representative protein G run with GB/ACE and the V01sc atomic volumes. This run becomes unstable after 1.3 ns. Shown are the total potential energy plus solvation energy, van der Waals energy, total electrostatic energy, Coulomb energy, GB/ACE self-energy, and GB/ACE screening energy. Total electrostatic energy is the sum of the latter three terms. Horizontal lines correspond to the structures at the start of the production period (not the crystal structure). The total potential energy also includes "bonded" energy terms (corresponding to distorsion of bonds, angles, and dihedrals).

Charged residues are underexposed by as much as 50% with both models. The RDIE structures are much too compact; this suggests that although they have not shifted far in terms of RMS C_α deviation, they are too rigid. They are apparently trapped close to the starting structure by unrealistic energy barriers associated with overly tight steric packing or with the lack of solvation of polar and ionic groups.[68] With ASP, the amount of hydrophobic surface accessible to solvent increases dramatically after about 800–900 ps; at the same time, the rms deviation begins to increase, indicating a possible instability with this model over long time periods (>1 ns).

The relative efficiencies of the models in terms of computer time were also compared. GB/ACE is 8.6 times slower than RDIE; ASP is 10.1 times slower than RDIE. It should be noted that all three models scale in the same way with system size. Indeed, it has been pointed out[35] that even though GB/ACE involves three-body forces, the GB/ACE force calculation can be performed in a way that scales with N^2, where N is the number of atoms in the protein. Timing data for an explicit solvent simulation of another small protein with the same force field were also available. Solvating the protein with a spherical droplet of explicit TIP3P solvent of 28-Å radius, for a total of 10,000 atoms, and treating the region outside the droplet as a dielectric continuum, leads to an MD model 15 times slower than GB/ACE.

Protein G: N—H-Order Parameters

Backbone amide order parameters[51,69] characterize the angular correlations of the N—H amide bond associated with internal vibrations of the protein. If $\mathbf{u}(t)$ is a unit vector along the bond at time t, the order parameter is the limit as $\tau \to \infty$ of the correlation function $\langle P_2(\mathbf{u}(t) \cdot \mathbf{u}(t + \tau))\rangle$, where the brackets represent an average over the time t (approximated here by an average over the MD simulation) and $P_2(x) = \frac{1}{2}(3x^2 - 1)$ is the second Legendre polynomial. The order parameters were estimated from several of the GB/ACE simulations, using production periods of 800–1,200 ps. With these, the convergence of the correlation functions whose long-time plateau defines the order parameters is reasonable for most residues (not shown).

The best order parameters are obtained with the scaled V01sc volumes. The results for two runs are compared with experiment and an explicit solvent simulation in

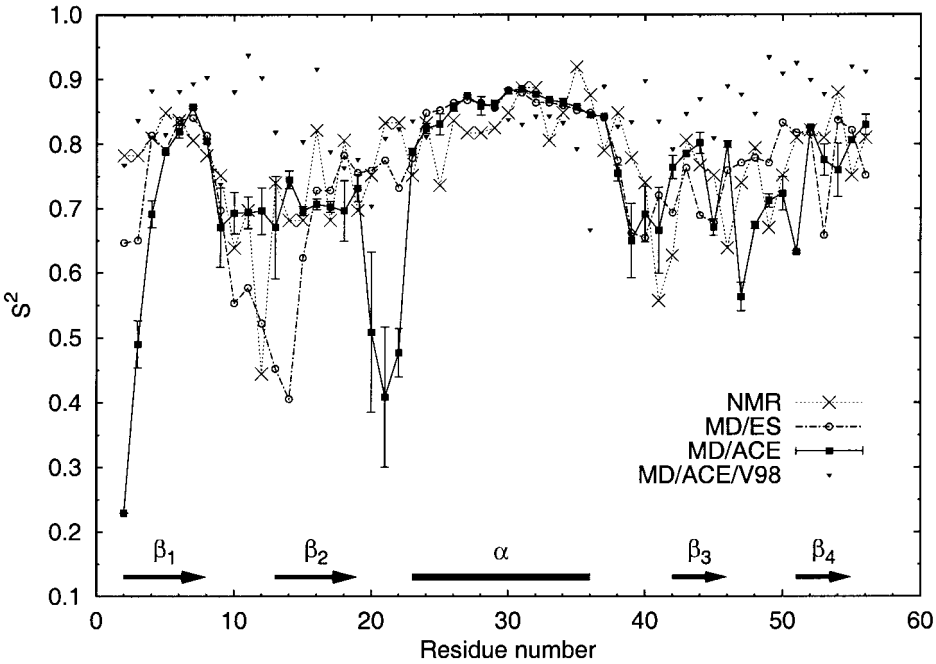

Fig. 6. Protein G backbone amide order parameters S^2 from GB/ACE with the V01sc atomic volumes (MD/ACE) and with the V98 volumes (MD/ACE/V98), from an explicit solvent simulation by Sheinerman and Brooks[51] (MD/ES) and from a nuclear magnetic resonance (NMR) experiment.[73] The GB/ACE V01sc results are averaged over two runs; error bars correspond to the differences between the two runs.

Figure 6. The agreement between the two runs is reasonable, except for a few residues in flexible loops, for which the ~1-ns sampling period is still too small to obtain converged results. Overall, the agreement with experiment is remarkably good, comparable to the explicit solvent simulation.

With the unscaled V98 atomic volumes, results are much poorer (Fig. 6): in the α-helix, the calculated order parameters are in the correct range, but in the loop and strand regions they are much too high. The results with RDIE are even worse (data not shown). This shows that the poorer models and GB/ACE parameterizations are unable to achieve the good agreement found with the best GB/ACE parameterization.

Sources of Instability: Comparing GB/ACE With Finite-Difference Poisson

Several of the protein G simulations deviate strongly from the native, experimental structure, either very early in the run or during the course of production. These instabilities presumably arise from inaccuracies in the implicit solvent model, since MD simulations with explicit solvent and the same protein force field behave much better.[18,51] One possible source of error in the GB/ACE model may be its departure from the "exact" treatment of the dielectric continuum solvent model (which can be obtained by the comparatively slow numerical solution of the Poisson equation). Alternatively, the continuum solvent model itself may be insufficiently accurate, compared with explicit solvent. To determine how well the GB/ACE model reproduces the "exact" dielectric continuum model, the GB/ACE solvation energy was compared with the solvation energy from numerical, finite-difference solution of the Poisson equation (see Methods) for structures sampled during several trajectories. The GB/ACE solvation energy is obtained as the sum of the self and screening energy contributions (Fig. 5). This analysis was applied to six protein G runs: a 3-ns V01sc run that becomes unstable after 1.3 ns, another stable and one unstable run with the V01sc volumes, two runs with the unscaled V01 volumes, and a V98sc run that becomes unstable after 500 ps. The data are summarized in Figures 7 and 8, illustrating several important points.

First, although the V01 volumes give a fair correlation with the FDP energies, they agree poorly in terms of absolute magnitude [Fig. 7(a)]. Thus, with these volumes, GB/ACE overstabilizes the non-native structures sampled in the two V01 trajectories by >600 kcal/mol.

In contrast, for runs with the scaled V01sc volumes, agreement between GB/ACE and FDP is very good [Fig. 7(b)]. The three runs analyzed sample both native-like structures and structures with rms deviations of up to 3 Å from native. Agreement is good even for the latter structures. With the V98sc volumes, agreement with FDP is slightly poorer (data not shown): the FDP energies are lower on average by ~50 kcal/mol, but the correlation between the FDP and GB/ACE energies is still very high. With both the V98sc and the V01sc volume sets, the slope of the FDP energies versus the GB/ACE energies is <1. The GB/ACE energies thus have a larger spread, indicating a greater sensitivity to conformation than FDP.

For the highly distorted structures sampled in the V01 runs, we recalculated the energies using the scaled V01sc volumes. This yields a large improvement, reducing the

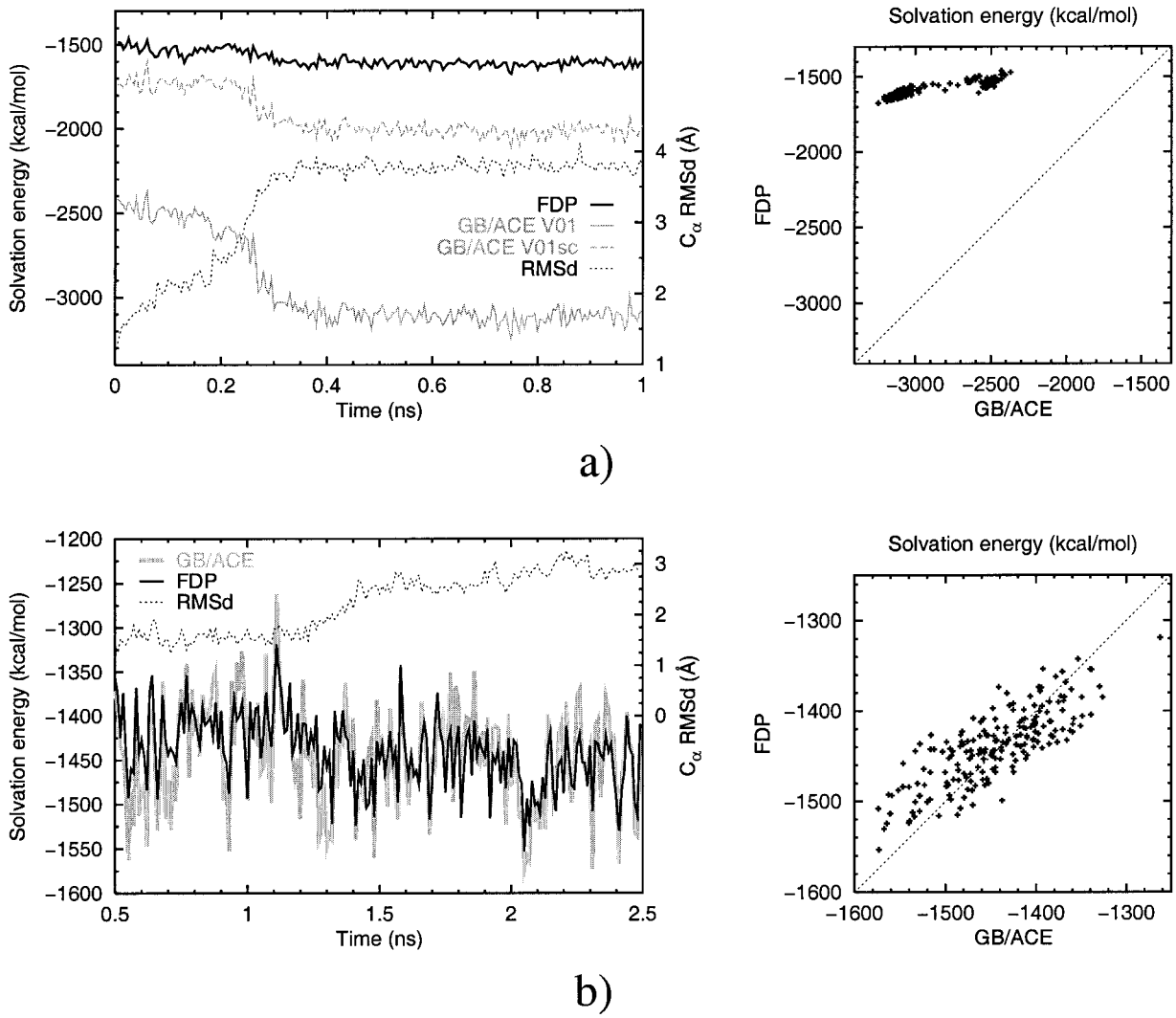

Fig. 7. Solvation energies (kcal/mol) for structures sampled during two protein G simulations using GB/ACE: comparison between GB/ACE and finite-difference Poisson calculations (FDP). **a:** Run with the unscaled V01 atomic volumes. **b:** Run with the scaled V01sc atomic volumes. Left: Solvation energies calculated with GB/ACE and with FDP. Root-mean-square deviations (RMSD) from the crystal structure (Å) are also shown (dots). **a:** GB/ACE energies calculated with the scaled V01sc volumes are also reported (dashed line; the trajectory was produced with the unscaled volumes). Right: Correlation between the GB/ACE and FDP solvation energies.

deviation from FDP by one-half; nevertheless, a ~250-kcal/mol discrepancy remains [Fig. 7(a)]. Thus, these distorted structures (generated with the unscaled V01 volumes) are apparently too far from the compact native geometry to obey the approximations of the GB/ACE model, even with optimal GB/ACE parameters. This contrasts with the moderately distorted structures sampled in some of the V01sc runs (discussed above), which gave energies very close to those of FDP.

Although the scaled V01sc volumes lead to excellent agreement with the FDP energies, structural instabilities do occur with these volumes in some runs. To understand this further, Figure 8 plots, for two V01sc runs, the backbone RMSD and the difference between the GB/ACE and FDP solvation energies. There is no obvious correlation between the two. For example, some native-like structures near the beginning of both runs are overstabilized by GB/ACE, while some non-native structures near the end of one of the runs are overstabilized by FDP. We do observe interesting correlations in some runs, where sudden departures from the native conformation occur simultaneously with large GB/ACE error levels. For example, during the 3 ns V01sc run [Fig. 8(a)], GB/ACE strongly overstabilizes the structures that occur between $t = 1.3$ and 1.5 ns; precisely the period during which the structure departs from the native structure (its RMSD increasing from 1.5 Å to 2.5 Å). Another example occurs around $t = 0.3$ ns in the V01sc run shown in Figure 8(b): the rms deviation increases abruptly from 1.2 Å to 2.0 Å; simultaneously, the GB/ACE energies deviate sharply from the "exact" FDP energies. Although this effect is not seen systematically in all runs, it does suggest that differences

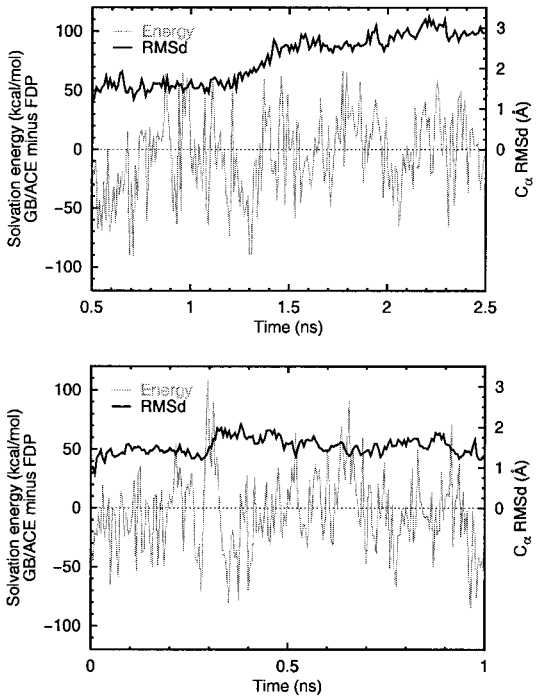

Fig. 8. GB/ACE and finite-difference Poisson solvation energies are compared for structures sampled during two protein G simulations that used GB/ACE. The difference between the "exact" finite-difference Poisson calculation and the GB/ACE energy is shown. Root-mean-square deviations (RMSD) from the starting structure (Å) are also shown. **a:** Same run as in Fig. 7b. **b:** Another run, using the V01sc volumes.

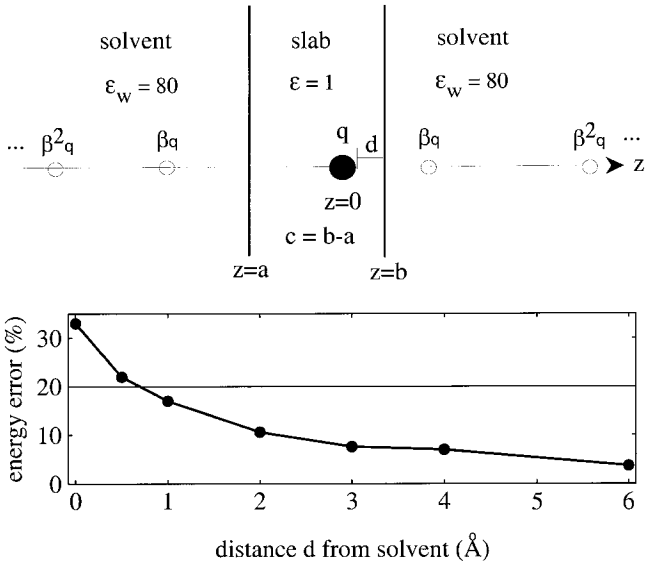

Fig. 9. Comparing the GB/ACE self-energy to an "exact" Poisson calculation for a spherical charge q in a low dielectric slab. The GB/ACE relative error is shown as a function of charge distance d from solvent. The charge radius is $R = 2$ Å. The left-hand charge–solvent distance a is held constant at 15 Å; only b (and $d = b - R$) are varied. There are infinite series of image charges on each side of the slab; only the first two in each series are shown. Each successive image is obtained by applying to the previous image a mirror reflection with respect to one or the other dielectric boundary plane; each reflection multiples the image charge by $\beta = (1 - \epsilon_w)/(1 + \epsilon_w)$, where ϵ_w is the solvent dielectric constant.[70]

in the GB/ACE and FDP energy surfaces account for at least part of the instability observed in some runs. Simulations with the FDP solvation potential[29] are needed to examine this in more detail.

GB/ACE Self-Energies: Effect of the Coulomb Field Approximation

One of the main sources of systematic error in GB/ACE is the use of the Coulomb field instead of the full, solvent-screened field in calculating the charge self-energies.[35] To estimate this error numerically, we consider a single charge q in a low dielectric planar slab, surrounded by high dielectric solvent (Fig. 9). This represents a highly simplified model of a charge near the surface of a protein; it was chosen because it lends itself readily to analytical and numerical treatment. For example, the solvent reaction field is equivalent[70] to the field of two infinite series of image charges (Fig. 9). Similar tests have been reported for spherical geometries[53] and for individual charges in actual proteins.[44]

We consider geometries typical of protein surface groups: 2–3-Å charge radius, 0–4-Å charge–solvent distance, and 14–24-Å "protein" diameter. With these, we find that the Coulomb field approximation overestimates the desolvation (i.e., the loss in solvation free energy) by about 10–30%. This is exactly the magnitude of the correction introduced above by scaling the atomic Voronoi volumes. Therefore, the volume scaling largely corrects on average for the Coulomb field approximation. The error decreases rapidly with increasing charge–solvent distance and is rather insensitive to protein thickness (parameter c in Fig. 9; data not shown). The results with spherical "proteins" were similar.[53] Because the charge–solvent distance can be approximated by the atomic solvation radius b_i (eq. 2; Methods), and since b_i is available in the ACE energy calculation, it should be possible to devise an efficient scaling procedure that corrects the error still more accurately, taking into account the degree of burial of individual charges. This possibility will be explored elsewhere.

CONCLUDING DISCUSSION

The efficiency of the generalized Born model and its good behavior for solvation and conformational energies suggest that it is currently the method of choice for modelling implicit solvent in molecular dynamics simulations. Compared with finite-difference Poisson models, for example, it provides a large speed enhancement[29] with only a moderate loss of accuracy. It is therefore important to determine its behavior in detail. So far, the GB screening function has been combined with four different approximations for the atomic solvation radii b_i (eq. 2).[34–37] Before this work, only one[36] had been used in MD simulations of proteins.[29,41,48]

The GB/ACE approximation was initially parameterized for small polypeptides,[35,42,49] leading to the V98 volume set. More recently, the V01 volume set was developed[50] by averaging Voronoi volumes over a large data base of structures from the PDB. Limited reparameterization was done here by varying the amplitude of the hydrophobic term and by scaling the atomic volumes, usually by

applying a single multiplying factor of 0.8 (with V01) or 0.9 (with V98) to all atom types.

With the unscaled V98 volumes, the best results are obtained with an effective surface tension of ~25 cal/mol/Å^2. With this value, two out of five protein G runs and two out of four Raf-RBD runs gave structures stable on the nanosecond time scale and in fair agreement with experiment; the other runs were unstable (Table I). The protein G instability was accompanied by undersolvation of certain charged groups, particularly carboxylate groups. Thus, the Glu27 side-chain had a strong tendency to become buried in the core of the protein. Protein backbone flexibility as measured by the N—H-order parameters was underestimated.

With the downscaled V98sc volumes, the model behaves much better. Carboxylate groups remain solvent-accessible and most runs are stable, with smaller deviations from the crystal structure. This variant of the model already performs significantly better than the two simpler models, ASP and RDIE, despite the low RMSD from the crystal structure obtained with the latter.

With the unscaled V01 atomic volumes, results are poor: the trajectories deviate rapidly and strongly from the experimental structure, and the calculated solvation energies deviate significantly from the "exact" finite-difference Poisson results, stabilizing non-native structures by up to 600 kcal/mol relative to the native structure, compared to an 80-kcal/mol stabilization with FDP.

In contrast, when the V01 volumes are reduced uniformly by 20%, the model behaves quite well. RMSD from the crystal structure are usually low (1.5–2.5 Å, compared with 1.2 Å with explicit solvent); the surface distribution of polar and hydrophobic groups is accurately reproduced, and N—H-order parameters are in good agreement with experiments. GB/ACE solvation energies agree closely with FDP results, showing no particular tendency to understabilize native, or overstabilize non-native structures, compared with FDP. This implementation of the model appears comparable to the best GB implementations developed in other laboratories.[29,41,44]

Even with the best GB/ACE parameterizations, while all the Raf-RBD runs were stable, two of the protein G runs showed stability problems, deviating by almost 3 Å from the X-ray structure after 500 ps and 1,500 ps, respectively. Thus, further improvement of the model is needed before very long simulations can be done with complete confidence. The main approximations in the GB model reside in the heuristic form of the screening function, eq. 1, and in the use of the Coulombic field[33] in calculating the self-energies (see Methods). Reducing the atomic volumes partly corrects for both approximations. Thus, the Coulomb field approximation was shown to overestimate charge desolvation by 10–30% for groups close to the surface[35,53] (Fig. 9). The surface region includes all the charged and most of the polar groups of the protein, so that its contribution is expected to dominate the error in the GB/ACE solvation energy. By reducing all atomic volumes, the contribution of this region is corrected for in an average sense, even though the local errors in the field are not corrected. More sophisticated corrections are possible and will be explored in future work. A recent variant of the GB model includes image charge effects from the solute–solvent boundary[71]; this idea could possibly be used to introduce a correction in the functional form of the self-energy integral.

Another systematic error in the GB/ACE potential arises from the GB screening function, with the same sign. Indeed, the GB screening function is designed to give the correct solvation energy for groups of like charges in the limit of very small separations. However, for small dipolar groups, the solvation energy is systematically underestimated by about 20%. This is seen by expanding the interaction energy of two opposite charges, q and $-q$, in the limit of small separations $r \ll b$, where b is the solvation radius of the individual charges. The leading term in the GB solvation energy is $-\frac{3}{8}q^2r^2/b^3$, instead of $-\frac{1}{2}q^2r^2/b^3$ with the "exact" dielectric continuum model (ignoring small terms of order $1/\epsilon_w^2$ in both cases, where ϵ_w is the solvent dielectric constant). This error could in principle be avoided by changing the functional form of the screening function, such that the leading terms for both a dipolar group and for a pair of nearby like charges are in accord with theory.

Another element of the model that could be improved is the hydrophobic energy term. The treatment used in this study is essentially a surface area model,[42] with a hydrophobic energy proportional to the total protein surface area. This is certainly simplistic; more sophisticated models will need to be implemented in the future.[21,22] However, the simulation results were found to be much less sensitive to the magnitude of this term than to the details of the electrostatic terms.

The simulations indicate that with the best GB/ACE parameterization, the position and local structure of the native energy basin are well described. However, the energy barriers surrounding this basin appear to be underestimated, as shown by the extended V01sc run that escapes from it after 1.3 ns. It should be noted that the time scale for motions with continuum solvent models (including GB) has been found to be faster than with explicit solvent,[24,46,47] presumably because of weaker mechanical and dielectric friction; e.g. the dielectric response of the solvent is treated as instantaneous. This may facilitate escape from the native energy basin. The non-native energy basin attained by the "unstable" 3-ns GB/ACE run has an enthalpy similar to that of the native basin. In addition, the solvation energies calculated with a full dielectric continuum model for the structures sampled during this run are in good agreement with the GB/ACE solvation energies. There is no systematic destabilization of the native structure by GB/ACE, compared with FDP. In some simulations, we observe correlations between the magnitude of the GB/ACE error (relative to FDP) and sudden departures from the native conformation. For example, during the 3-ns V01sc run, GB/ACE overstabilizes the structures that occur between $t = 1.3$ and 1.5 ns, where the protein departs from its native structure. This finding suggests that differences in the GB/ACE and FDP energy surfaces account for at least part of the instability

observed in this run. Inaccuracies in the dielectric continuum model itself may contribute as well.

In summary, with atomic Voronoi volumes uniformly reduced by 20%, the GB/ACE model provides a very good approximation to a dielectric continuum solvent model, as well as to an explicit solvent model. Molecular dynamics simulations of two small proteins were stable on the nanosecond timescale and in remarkably good agreement with both explicit solvent simulations and experiment. The quality of this GB/ACE implementation should be compatible with many applications, such as conformational searching, NMR, and X-ray structure refinement, and with the analysis of electrostatic interactions in proteins.

ACKNOWLEDGMENTS

The authors thank Martin Karplus, David Case, Charles Brooks, and Brian Dominy for helpful discussions. This work was supported by CNRS grants [Financement Jeune Equipe (to T.S.) and Physique et Chimie du Vivant program (to T.S.)]. Some of the simulations were carried out at the supercomputer center CINES of the French Ministry of Education, supercomputer allocation c20010722229 (to T.S.).

REFERENCES

1. Brooks C, Karplus M, Pettitt M. Proteins: A theoretical perspective of dynamics, structure and thermodynamics. Adv Chem Phys 1987;71:1–259.
2. Becker O, Mackerell A, Jr, Roux B, Watanabe M, editors. Computational biochemistry and biophysics. New York: Marcel Dekker; 2001.
3. Roux B, Simonson T, editors. Implicit solvent models for biomolecular simulations. Biophys Chem (Special Issue) 1999;78:1–218.
4. Orozco M, Luque F. Theoretical methods for the description of the solvent effect in biomolecular systems. Chem Rev 2000;100:4187–4226.
5. Simonson T. Macromolecular electrostatics: Continuum models and their growing pains. Curr Opin Struct Biol 2001;11:243–252.
6. Wesson L, Eisenberg D. Atomic solvation parameters applied to molecular dynamics of proteins in solution. Protein Sci 1992;1:227–235.
7. Fraternali F, van Gunsteren W. An efficient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution. J Mol Biol 1996;256:939–948.
8. Lazaridis T, Karplus M. Effective energy function for proteins in solution. Proteins 1999;35:133–152.
9. Kirkwood J, Westheimer F. The electrostatic influence of substituents on the dissociation constant of organic acids. J Chem Phys 1938;6:506–512.
10. Sharp K, Honig B. Electrostatic interactions in macromolecules: Theory and applications. Annu Rev Biophys Biophys Chem 1991; 19:301–332.
11. Schaefer M, Vlijmen HV, Karplus M. Electrostatic contributions to molecular free energies in solution. Adv Protein Chem 1998;51: 1–57.
12. Sharp K. Incorporating solvent and ion screening into molecular dynamics using the finite-difference Poisson–Boltzmann model. J Comp Chem 1991;12:454–468.
13. Sitkoff D, Sharp K, Honig B. Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 1994;98:1978–1988.
14. Simonson T, Brünger AT. Solvation free energies estimated from macroscopic continuum theory: An accuracy assessment. J Phys Chem 1994;98:4683–4694.
15. Nina M, Beglov D, Roux B. Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. J Phys Chem B 1997;101:5239–5248.
16. Landau L, Lifschitz E. Electrodynamics of continuous media. New York: Pergamon; 1980.
17. Fröhlich H. Theory of dielectrics. Oxford: Clarendon Press; 1949.
18. Simonson T, Perahia D. Microscopic dielectric properties of cytochrome c from molecular dynamics simulations in aqueous solution. J Am Chem Soc 1995;117:7987–8000.
19. Sham Y, Chu Z, Warshel A. Consistent calculations of pK_a's of ionizable residues in proteins: Semi-microscopic and microscopic approaches. J Phys Chem B 1997;101:4458–4472.
20. Simonson T, Archontis G, Karplus M. A Poisson–Boltzmann study of charge insertion in an enzyme active site: The effect of dielectric relaxation. J Phys Chem B 1999;103:6142–6156.
21. Hummer G. Hydrophobic force field as a molecular alternative to surface area models. J Am Chem Soc 1999;121:6299–6305.
22. Lum K, Chandler D, Weeks J. Hydrophobicity at small and large length scales. J Phys Chem B 2000;103:4570–4577.
23. Gilson M, McCammon J, Madura J. Molecular dynamics simulation with a continuum electrostatic model of the solvent. J Comp Chem 1995;16:1081–1095.
24. Potter M, Kirchhoff P, Carlson H, McCammon J. Molecular dynamics of cryptophane and its complexes with tetramethylammonium and neopentane using a continuum solvent model. J Comp Chem 1999;20:956–970.
25. Lu B, Wang C, Chen W, Wan S, Shi Y. A stochastic dynamics simulation study associated with hydration force and friction memory effect. J Phys Chem B 2000;104:6877–6883.
26. Marchi M, Borgis D, Levy N, Ballone P. A dielectric continuum molecular dynamics method. J Chem Phys 2001;114:4327–4385.
27. Gilson M, Davis M, Luty B, McCammon J. Computation of electrostatic forces on solvated molecules using the Poisson–Boltzmann equation. J Phys Chem 1993;97:3591–3600.
28. Im W, Beglov D, Roux B. Continuum solvation model: Computation of electrostatic forces from numerical solutions to the Poisson–Boltzmann equation. Comp Phys Commun 1998;109:1–17.
29. David L, Luo R, Gilson M. Comparison of generalized Born and Poisson models: Energetics and dynamics of HIV protease. J Comp Chem 2000;21:295–309.
30. Still WC, Tempczyk A, Hawley R, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
31. Bashford D, Case D. Generalized Born models of macromolecular solvation effects. Annu Rev Phys Chem 2000;51:129–152.
32. Scarsi M, Apostolakis J, Caflisch A. Continuum electrostatic energies of macromolecules in aqueous solutions. J Phys Chem A 1997;101:8098–8106.
33. Schaefer M, Froemmel C. A precise analytical method for calculating the electrostatic energy of macromolecules in aqueous solution. J Mol Biol 1990;216:1045–1066.
34. Hawkins G, Cramer C, Truhlar D. Pairwise descreening of solute charges from a dielectric medium. Chem Phys Lett 1995;246:122–129.
35. Schaefer M, Karplus M. A comprehensive analytical treatment of continuum electrostatics. J Phys Chem 1996;100:1578–1599.
36. Qiu D, Shenkin P, Hollinger F, Still W. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. J Phys Chem A 1997;101:3005–3014.
37. Ghosh A, Rapp C, Friesner RA. Generalized Born model based on a surface area formulation. J Phys Chem B 1998;102:10983–10990.
38. Hawkins G, Cramer C, Truhlar D. Parameterized model for aqueous free energies of solvation using geometry-dependent atomic surface tensions with implicit electrostatics. J Phys Chem B 1997;101:7147–7157.
39. Scarsi M, Apostolakis J, Caflisch A. Comparison of a GB solvation model with explicit solvent simulations: Potentials of mean force and conformational preferences of alanine dipeptide and 1,2-dichloroethane. J Phys Chem B 1997;102:3637–3641.
40. Wagner F, Simonson T. Implicit solvent models: Combining an analytical formulation of continuum electrostatics with simple models of the hydrophobic effect. J Comp Chem 1999;20:322–335.
41. Dominy B, Brooks C III. Development of a generalized Born model parameterization for proteins and nucleic acids. J Phys Chem B 1999;103:3765–3773.
42. Schaefer M, Bartels C, Karplus M. Solution conformations and thermodynamics of structured peptides: Molecular dynamics simulation with an implicit solvation model. J Mol Biol 1998;284:835–847.
43. Srinivasan J, Cheatham T, Cieplak P, Kollman P, Case D. Continuum solvent studies of the stability of DNA, RNA, and

phosphoramidate–DNA helices. J Am Chem Soc 1998;120:9401–9409.
44. Onufriev A, Case D. Modification of the generalized Born model suitable for macromolecules. J Phys Chem B 1999;104:3712–3720.
45. Majeux N, Scarsi M, Apostolakis J, Ehrhardt C, Caflisch A. Exhaustive docking of molecular fragments with electrostatic solvation. Proteins 1999;37:88–105.
46. Williams D, Hall K. Unrestrained simulations of the UUCG tetraloop using an implicit solvation model. Biophys J 1999;76: 3192–3205.
47. Tsui V, Case D. Molecular dynamics simulations of nucleic acids with a generalized Born model. J Am Chem Soc 2000;122:2489–2498.
48. Bursulaya B, Brooks C III. Comparative study of the folding free energy landscape of a three-stranded β-sheet protein with explicit and implicit solvent models. J Phys Chem B 2001;104:12378–12383.
49. Schaefer M, Bartels C, Karplus M. Solution conformations of structured peptides: Continuum electrostatics vs. distance-dependent dielectric functions. Theor Chem Acc 1999;101:194–204.
50. Schaefer M, Bartels C, Leclerc F, Karplus M. Effective atom volumes for implicit solvent models: Comparison between Voronoi volumes and minimum fluctuation volumes. J Comp Chem 2001; (in press).
51. Sheinerman FB, Brooks CL. A molecular dynamics simulation study of segment B1 of protein G. Proteins 1997;29:193–202.
52. Born M. Volumen und hydrationswärme der ionen. Z Phys 1920;1:45–48.
53. Schaefer M. Monte Carlo–Dynamik und Elektrostatik von Proteinen. PhD thesis, Heidelberg University; 1989.
54. Chothia C. Hydrophobic bonding and accessible surface area in proteins. Nature 1974;248:338–339.
55. Abraham MH. Free energies, enthalpies, and entropies of solution of gaseous nonpolar nonelectrolytes in water and nonaqueous solvents. The hydrophobic effect. J Am Chem Soc 1982;104:2085–2094.
56. Brooks B, Bruccoleri R, Olafson B, States D, Swaminathan S, Karplus M. Charmm: A program for macromolecular energy, minimization, and molecular dynamics calculations. J Comp Chem 1983;4:187–217.
57. Roux B, Yu H, Karplus M. Molecular basis for the Born model of solvation. J Phys Chem 1990;94:4683–4688.
58. Wolfenden R, Andersson L, Cullis P, Southgate C. Affinities of amino acid side chains for solvent water. Biochemistry 1981;20: 849–855.
59. Fry DC, Madison VS, Greeley DN, Felix AM, Heimer EP, Frohman L, Campbell RM. Solution structures of cyclic and dicyclic analogues of growth hormone releasing factor as determined by two-dimensional NMR and CD spectroscopies and constrained molecular dynamics. Biopolymers 1992;32:649–666.
60. Brünger AT, Karplus M. Polar hydrogen positions in proteins: An empirical energy placement and neutron diffraction comparison. Proteins 1988;4:148–156.
61. Reiher WE. Theoretical studies of hydrogen bonding. PhD thesis, Harvard University; 1985.
62. Ryckaert J, Ciccotti G, Berendsen H. Numerical integration of the cartesian equations of motion for a system with constraints: Molecular dynamics of n-alkanes. J Comp Phys 1977;23:327–341.
63. Madura J, Briggs J, Wade R, Davis M, Luty B, Ilin A, Antosiewicz J, Gilson M, Baheri B, Scott L, McCammon J. Electrostatics and diffusion of molecules in solution: Simulations with the University of Houston Brownian Dynamics program. Comp Phys Commun 1995;91:57–95.
64. Davis ME, McCammon JA. Electrostatics in biomolecular structure and dynamics. Chem Rev 1990;90:509–521.
65. Gilson M, Sharp K, Honig B. Calculating electrostatic interactions in bio-molecules: Method and error assessment. J Comp Chem 1988;9:327–335.
66. Shrake A, Rupley JA. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. J Mol Biol 1973;79:351–371.
67. Davis M, McCammon J. Dielectric boundary smoothing in finite-difference solutions of the Poisson equation: An approach to improve accuracy and convergence. J Comp Chem 1991;12:909.
68. Steinbach P, Brooks B. New spherical-cutoff methods for long-range forces in macromolecular simulation. J Comp Chem 1994;15: 667–683.
69. Chandrasekhar I, Clore G, Szabo A, Gronenborn A, Brooks B. A 500 ps molecular dynamics simulation study of interleukin-1 beta in water: Correlation with nuclear magnetic resonance spectroscopy and crystallography. J Mol Biol 1992;226:239–250.
70. Smythe W. Static and dynamic electricity. New York: McGraw-Hill; 1968.
71. Havranek J, Harbury P. Tanford–Kirkwood electrostatics for protein modeling. Proc Natl Acad Sci USA 1999;96:11145–11150.
72. Kraulis P. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr 1991;24: 946–950.
73. Barshi JJ, Grasberger B, Gronenborn A, Clore G. Investigation of the backbone dynamics of the IgG-binding domain of streptococcal protein G by heteronuclear two-dimensional 1H–15N nuclear magnetic resonance spectroscopy. Protein Sci 1994;3:15–21.