Implementation and Testing of Stable, Fast Implicit Solvation in Molecular Dynamics using the Smooth-Permittivity Finite Difference Poisson–Boltzmann Method

NINAD V. PRABHU,[1] PEIJUAN ZHU,[2] KIM A. SHARP[1]
[1]*Johnson Research Foundation and Department of Biochemistry and Biophysics, University of Pennsylvania, 37th and Hamilton Walk, Philadelphia, Pennsylvania 19104*
[2]*Department of Pharmacology, University of Pennsylvania, Philadelphia, Pennsylvania 19104*

Received 17 March 2004; Accepted 10 August 2004
DOI 10.1002/jcc.20138
Published online in Wiley InterScience (www.interscience.wiley.com).

Abstract: A fast stable finite difference Poisson–Boltzmann (FDPB) model for implicit solvation in molecular dynamics simulations was developed using the smooth permittivity FDPB method implemented in the OpenEye ZAP libraries. This was interfaced with two widely used molecular dynamics packages, AMBER and CHARMM. Using the CHARMM-ZAP software combination, the implicit solvent model was tested on eight proteins differing in size, structure, and cofactors: calmodulin, horseradish peroxidase (with and without substrate analogue bound), lipid carrier protein, flavodoxin, ubiquitin, cytochrome *c*, and a *de novo* designed 3-helix bundle. The stability and accuracy of the implicit solvent simulations was assessed by examining root-mean-squared deviations from crystal structure. This measure was compared with that of a standard explicit water solvent model. In addition we compared experimental and calculated NMR order parameters to obtain a residue level assessment of the accuracy of MD-ZAP for simulating dynamic quantities. Overall, the agreement of the implicit solvent model with experiment was as good as that of explicit water simulations. The implicit solvent method was up to eight times faster than the explicit water simulations, and approximately four times slower than a vacuum simulation (i.e., with no solvent treatment).

 J Comput Chem 25: 2049–2064, 2004

Key words: finite-difference Poisson–Boltzmann; implicit solvent; molecular dynamics; continuum solvent

Introduction

Realistic molecular dynamics (MD) and Monte Carlo (MC) simulation of proteins in water requires an accurate and computationally tractable way to treat the solvent. Two treatments are currently used: explicit solvent, and implicit solvent. The explicit solvent treatment requires, as its name implies, a detailed atomic level description of the solvent, so it is computationally demanding. Simulations of 1–10 ns for medium to large proteins represent a major investment in computer time, despite improvements in computer memory and speed, while 100-ns to 1 μs simulations are still uncommon and are usually only attempted only for small systems.[1,2] Even as these barriers recede with improvements in computers, the computational biochemist will always want to do larger and longer simulations. Thus, there has been a continuous effort over more than a decade to develop implicit solvent models for molecular dynamics simulations. Implicit solvent models may be briefly defined as models that represent one of more of the physical properties of the solvent directly in the MD/MC simulation without requiring explicit water and/or averaging over water configurations.[3,4] The main motivation for implicit solvent models is to reduce the computational demand, which inevitably entails the sacrifice of the solvent's molecular detail. They are thus not appropriate for every application of MD/MC simulation. However, there are some macroscopic average solvent properties that are hard to simulate accurately with explicit solvent, such as dielectric effects. The ability to build these into an implicit solvent model may actually give them some advantage for certain kinds of simulations, which is another impetus for their development.

There are two major contributions of solvent to the energetics of a solvated biomolecule: electrostatic (including polar) interactions, and nonpolar (hydrophobic and van der Waals) interactions.

Correspondence to: K. A. Sharp; e-mail: sharp@mail.med.upenn.edu
Contract/grant sponsor: NSF; contract/grant number: MCB02-35440.

Early implicit solvent models for MD used the finite different Poisson–Boltzmann (FDPB) to represent the electrostatic contributions.[5–11] This was later combined with a surface area term to represent the nonpolar term.[12] While advances have been made in developing the FDPB model for dynamics,[13] and it is routinely used to postprocess snapshots from MD simulations,[14–16] it appears that several problems have limited its use within MD simulations. These include its still rather significant computational demand[13] despite various algorithmic improvements,[17–19] and the extraction of smoothly varying stable forces. The standard implementation of the FDPB method uses a two dielectric mapping (molecule and solvent) with a defined boundary surface between them. The boundary surface is typically either the molecular surface (MS) or the solvent accessible surface (SAS).[20,21] As the molecular conformation changes from step to step in MD the precise mapping of this boundary onto the grid varies in a rather abrupt way that can introduce significant noise into the forces, degrading the stability of the simulation. Ironically, the major computational demand of FDPB now also comes largely from generating the solvent/molecular boundary surface and mapping of the molecule onto the finite difference grid, because the actual solution of the PB equation has been speeded up so successfully.[17–19] However, if the grid resolution is decreased to speed up the finite difference algorithm, the numerical errors from boundary mapping are more acute.[13] A discussion of the accuracy and speed issue in the dielectric mapping when using FDPB in MD simulations, and one successful solution to the problem has recently been described by Lu and Luo.[22]

An alternative implicit solvent model is based on the generalized Born (GB) model,[23] with various modifications, improvements, and parameterizations for use with proteins and DNA.[24–32] For a review of GB models see Bashford and Case.[33] A related approach uses screened Coulomb potentials.[34] In these models a smoothly varying functional form (the generalized Born expression) describes a combination of charge–charge and charge–solvent interactions for each pair of atoms. This is more rapid to evaluate than the FDPB equations, and produces smoothly varying forces. The physical model underlying both the FDPB and GB models is that of a set of partial atomic charges contained in a low dielectric cavity surrounded by a high dielectric representing the solvent. The position of the charges is determined by the atomic positions, and the shape of the cavity is determined by the volume occupied by the molecule. However, in the GB method the shape of the solvent/molecule boundary is not used directly. Instead, the key parameter in GB model is the effective Born radius for each atom. For each atom that is the radius of the sphere that would give the same solvation energy calculated using the Born equation as would be obtained for that atom in its actual position with the real molecular shape. For an isolated atom its electrostatic solvation energy would be given by the Born equation with the assigned (typically van der Waals) radius of that atom. However, when this atom is part of a molecule the other atoms displace solvent or "descreen" the atom, reducing its solvating energy, or equivalently, increasing its effective Born radius. A variety of algorithms have been used to calculate the effective Born radii, the dominant strategy being to parameterize the method so that the GB electrostatic contribution to the solvation energy reproduces as closely as possible that from the FDPB method for a series of test molecules/conformations.[24–30] Improvements to the effective Born radius calculation algorithm have been made by increasingly more sophisticated "descreening" algorithms, and a detailed comparison of GB and FDPB methods (including the ZAP implementation of the FDPB method used here) focussing on solvation energies and computational speed has recently been made.[35] More sophisticated descreening methods tend to increase the computational time of GB, however, and reduce its relative speed advantage over the FDPB method. In the GB model, as in the FDPB model, the nonpolar contribution to solvation is usually treated with a solvent-accessible surface area term. In implementing either method in MD this surface area term adds a significant computation cost, because the derivative of the SAS with respect to each atomic coordinate must be calculated to obtain the contribution to the solvent forces. This has motivated improvements in the accuracy and speed of the nonpolar parts.[31]

Parameterization of GB against FDPB to match total solvation free energies has been a very fruitful approach for developing the GB models. In both models the total solvation energy is composed of two terms: the charge–charge and charge–solvent terms. It has, however, been pointed out that because the GB and FDPB models have different functional forms, concurrence of the total energy between the two models means that the two components may not, indeed probably must not, be the same.[36] The same argument applies to the components of the electrostatic forces used in MD. Interestingly, in one direct comparison of MD using either GB or FDPB implicit solvent models[13] it was found that GB yielded less accurate forces than energies compared to the FDPB model. It is not clear at this time whether this is a serious discrepancy between GB vs. FDPB parameterizations when doing dynamics, because there have not been enough simulations with a full FDPB treatment within MD to compare with. One of the goals of this work, by improving the implementation of FDPB in MD simulations, is to enable such direct comparisons to be made.

A potential solution to both the speed and stability problems in using the FDPB model with MD has been provided with the development and testing of a smooth permittivity methods for the FDPB calculations.[37,38] In brief, this method represents the atomic volumes by three-dimensional Gaussian density functions rather than van der Waals (VDW) radius spheres. This implementation of the FDPB model provides some significant advances. First, the mapping of the molecule onto the lattice is much faster. Second, it also avoids some artifacts incurred in the standard FPDB method, which generates the molecular surface from the atomic VDW spheres (as defined by the rolling sphere construct of Lee & Richards[20]) to define the dielectric boundary. Third, the dielectric varies smoothly from the interior to exterior over a subatomic distance. This results in less grid-mapping errors and smoother forces as the molecular changes conformation. Fourth, with the Gaussian representation the solvent accessible area and its derivative with respect to atomic coordinates are obtained at almost no extra cost, obviating an extra calculation for the nonpolar solvation term. Finally, the algorithm has been implemented in a fast, robust, user-friendly, and portable set of library routines known as ZAP, available from OpenEye Software (Santa Fe).

Given that the ZAP FDPB implementation is fast and has good algorithmic properties, and that the "gold standard" for accuracy of Generalized Born models is how well they reproduce the FDPB

Table 1. Proteins Used in This Study.

Name	PDB ID (resolution)	Number of atoms	Net charge[a]	Cofactors	Bound Ions	Bound ligand
Horseradish peroxidase	1ATJ (2.15)	4771	+1	Heme	2 Ca^{2+}	CO
Horseradish peroxidase	2ATJ (2.0)	4802	+1	Heme	2 Ca^{2+}	Benzyhydroxamic Acid, CO
Lipid binding protein (mouse)	1LIB (1.7)	2057	+2	None	None	None
Cytochrome *c* (horse heart)	1HRC (1.9)	1745	+6	Heme	None	None
De novo designed 3-helix bundle	1LQ7 (NMR)	1113	+2	None	None	None
Flavodoxin (Anabaena 7120)	1FLV (2.0)	2622	−19	Flavin	None	None
Calmodulin (*H. sapiens*)	1CDL (2.4)	2511	−10	None	4 Ca^{2+}	smMLCK peptide
Ubiquitin (*H. sapiens*)	1UBQ (1.8)	1231	0	None	None	None

[a]At pH 7, including bound ions, and assuming standard residue pKa values.

results for a series of static structures, we decided that a robust implementation of the smooth permittivity model implicit solvent model in MD simulations would be the best of both worlds: an acceptable computational cost combined with a full treatment of the PB model. In this article we describe implementation of the FDPB implicit solvent model in both AMBER and CHARMM using the ZAP libraries, a comparison of its computational cost vs. explicit solvent models, and tests of the realism of the model against experimental data on protein dynamics.

Methods

Choice of Test Proteins

In developing any new molecular dynamics method that is aimed to be of general use, it is important to ensure that it works robustly; not just on common test proteins (e.g., small well-characterized cofactor-less proteins such as ubiquitin) but on the more challenging proteins that are bigger, contain cofactors, etc., that are often encountered in modeling applications. Eight proteins were chosen for this study based on these criteria (see Table I). They were a heterogeneous set in size, structure, and composition (see Fig. 1). Sizes ranged from the 67 residue Alpha-3 W *de novo* designed three-helix bundle (1LQ7) to the 306 residue horseradish peroxidase C (1ATJ and 2ATJ). The secondary structure elements also differed, with some proteins like Alpha-3 W and calmodulin (1CDL) having a large helical content, while others like the adipocyte lipid-binding protein (1LIB) and ubiquitin (1UBQ) having a beta-sheet as their dominant secondary structure. The proteins also in some cases had a well-defined hydrophobic core that was absent in others. They also varied in regard to presence and identity of cofactors, ions, and ligand. Three had a heme in the

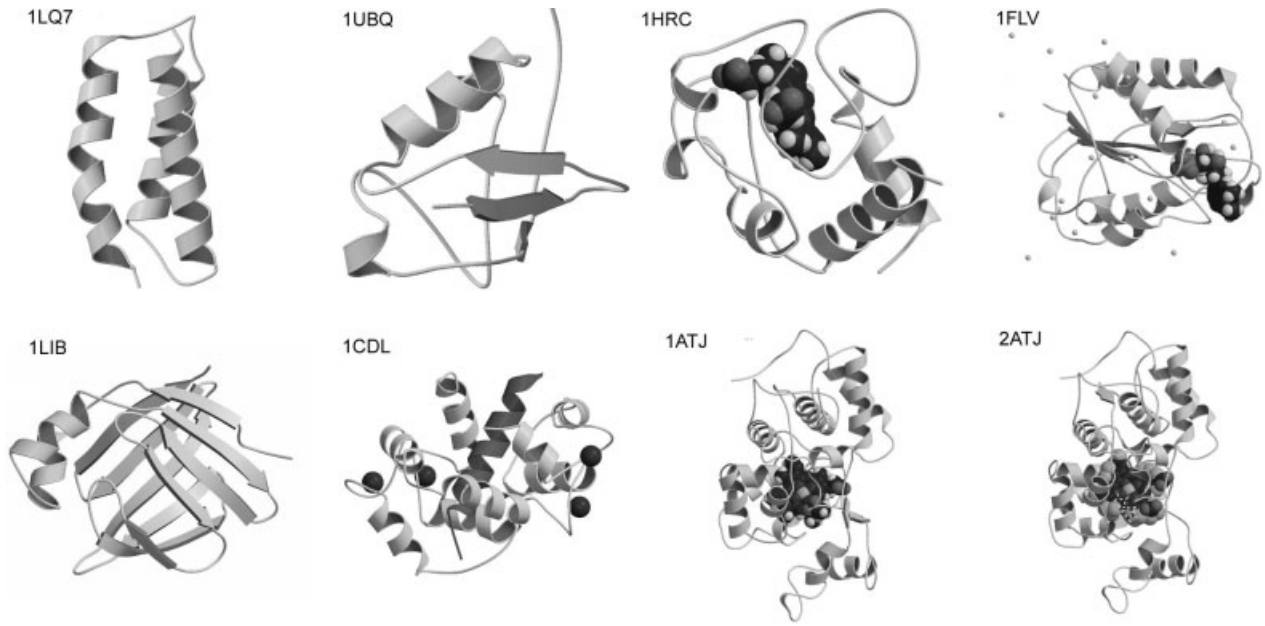

Figure 1. Ribbon diagrams of proteins used in this study. Cofactors, ligands, and ions are shown in CPK rendering.

interior, for example, cytochrome *c* (1HRC), and HRPC (1ATJ and 2ATJ), another had a flavin (1FLV). Two had ligands bound (2ATJ and 1CDL). Importantly with regard to testing any solvent model, the proteins vary widely in net charge (Table 1), with examples of high negative and positive net charge. With regard to the simulation of dynamics, for four of the proteins (1CDL, 1UBQ, 1LIB, and 1FLV) there is extensive experimental dynamics data in the form of close to complete amide (NH) and methyl NMR generalized order parameters (S^2) sets.[39–43]

Explicit Solvent Molecular Dynamics Simulations

In all cases the starting structures were taken from the Protein Database,[44] as follows. Horseradish peroxidase (HRPC) with the bound inhibitor Benzohydroxamic acid (BHA) used the PDB entry 2ATJ[45] and without BHA used entry 1ATJ.[46] Cytochrome *c* starting coordinates were obtained from entry 1HRC.[47] Minor modifications were made to the CHARMM force field to introduce connectivities from Met80, Cys14, and Cy17 to the heme (described in Prabhu et al.[48]). The NMR determined structure of the *de novo* three-helix bundle protein maquette was used for initial coordinates (structure A from PDB entry 1LQ7[49]). Coordinates of the calcium loaded Calmodulin (CaM) smooth muscle light-chain kinase peptide (smLCKp) complex were obtained from entry 1CDL of the PDB database.[50] Coordinates of ubiquitin were obtained from entry 1UBQ.[51] Coordinates of flavodoxin were obtained from entry 1FLV.[52] All structures were determined by X-ray crystallography except 1LQ7, where an NMR determined solution structure was used.

The simulations used the following CHARMM MD parameters and settings unless noted otherwise. The CHARMM force field and simulation software; the leapfrog Verlet integration scheme with a time step of 1 fs, cubic box periodic boundary conditions (PBC) with the minimum image convention; a constant temperature of 298 K was maintained using the Berendsen constant temperature algorithm with a coupling constant of 5 ps,[53] a pressure of 1 atm using a piston mass of 500; a nonbonded force cutoff at 13 Å using the force shift algorithm. The SHAKE algorithm[54] was used to fix the length of all bonds involving hydrogen atoms. Hydrogen atoms were added with CHARMM and structures minimized for 3000 steps using the ABNR method. TIP3P water was added to produce the required solvent box size. In some cases solvent ions were also added to produce a net neutral system. Structures were heated for 50 ps, and equilibrated for 50 ps. Structures were saved every 0.1 ps of the simulations for analysis. Small variations in protocol specific to the individual systems studied are the result of some of these simulations being part of other studies. These variations and other simulation parameters are given in Table 2, and described below.

For simulations using the particle mesh Ewald (PME) boundary condition,[55] unless otherwise noted a B-spline order of 4, an FFT grid a little over one point per Å and a real-space Gaussian-width kappa of 0.32 Å was used. The list of nonbonded interactions was truncated at 10 Å, using the by-cubes algorithm for computing the nonbonded interaction list. The nonbonded van der Waals and real-space electrostatic interactions were smoothly switched off at 7.5–9 Å. For 1LQ7 an orthorhombic solvent box of initial dimensions $70 \times 47 \times 47$ Å was used. The longest dimension of the box was allowed to vary in response to the constant pressure and temperature thermodynamic ensemble. The electrostatic interactions were computed with no truncation using particle mesh ewald (PME) with a B-spline order of 6 and an FFT grid around one point per Å, a real-space Gaussian-width kappa of 0.5 Å was used.

Table 2. Explicit Water Simulation Parameters and Conditions.

PDB ID	Number of waters	Box size Å[a]	Solvent ions	Boundary conditions
1ATJ	13,430	77	1 Cl^-	PME
2ATJ	13,464	77	1 Cl^-	PME
1LIB	7601	63	2 Cl^-	PBC
1HRC	5116	56	6 Cl^-	PME
1LQ7	4745	$70 \times 47 \times 47$	2 Cl^-	PME
1FLV	7400	63	19 Na^+	PME
1CDL	7465	63	0	PBC
1UBQ	4596	53	0	PBC

[a]Cubic box unless indicated otherwise.

Cofactor Force Field Parameters

Parameters for Heme and calcium were taken from the CHARMM22 parameter set. Parameters for flavin were obtained from AMBER,[56] at the Web site http://pharmacy.man.ac.uk/amber/index.html, and reformatted for CHARMM input. Parameters for Benzyhydroxamic acid were taken from Kaposi et al.[57]

Implementation of the ZAP FDPB Solvent Model

CHARMM version 27b2 with source code was obtained from Professor Martin Karplus.[58,59] Amber 6.0 with source was obtained from Professor Peter Kollman.[56] The ZAP C++ libraries for SGI and LINUX were obtained from OpenEye Scientific Software, Santa Fe (http://www.eyeseopen.com). To interface the ZAP libraries with the MD code interface subroutines (named zapchm or zapamb, respectively) were written in Fortran77, based on the example file written by J. Andrew Grant, which is distributed with the ZAP library. This routine passes the atomic coordinates, charges, radii, and dielectric constants from the MD program to the ZAP libraries, and creates the necessary "handles" in the ZAP library to perform the FDPB and accessible area derivative calculations and extract the forces. The routine then passes the forces back to the calling MD program. A copy of this routine is available from the authors upon request. To invoke this routine the MD programs were modified as follows. The existing generalized Born initialization routine was used to assign atomic radii and charges and the solvent dielectric constant, because these are common to the GB and FDPB methods. A call to the ZAP libraries was then inserted in the nonbonded energy/force subroutine, and after the call the FDPB and surface area solvation forces were added to the total MD forces. Two extra MD control parameters were used to specify the frequency of ZAP calculations and the hydrophobic surface free energy term.

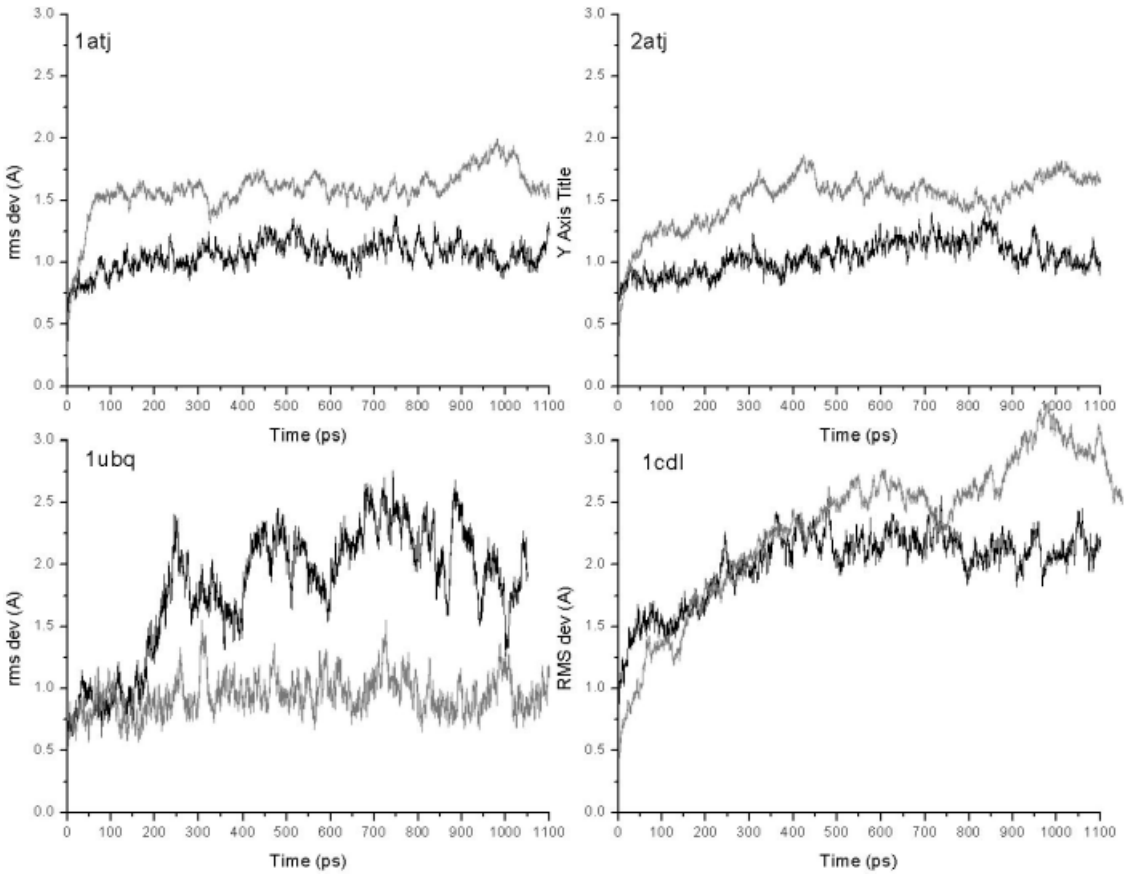

Figure 2. Root-mean-square deviation of backbone atoms as a function of simulation time for 1ATJ, 2ATJ, 1UBQ, and 1CDL. Bold trace: Explicit solvent. Light trace: ZAP implicit solvent model. The elapsed time includes the heating and equilibration phases.

Running MD with the ZAP Implicit Solvent Model

The implementation of ZAP in both CHARMM and AMBER was checked for correctness by evaluating forces for several configurations of a small molecule N-methyl acetamide, and comparing with the results from a FPDB solvent force calculation using a stand alone program supplied with the ZAP libraries. We selected CHARMM for detailed comparison with explicit water simulations and experiment because several existing explicit water simulations were already available to us.

The general protocol follows that of a standard explicit water MD simulation with the omission of explicit solvent and periodic boundary conditions. The molecule is built, assigned force field parameters, and minimized in the MD package to remove internal strain before switching on the implicit solvent model. The ZAP implicit solvent was switched on for a heating period of 50 ps, an equilibration period of 50 ps, and then sequential batches of 100 ps sampling simulations. Simulations were performed at 298 K using the same MD parameters as for the explicit water simulations where appropriate, except as follows. There was no pressure control needed, and no periodic boundary conditions. A constant temperature of 298 K was maintained using velocity reassignment. In addition to the Langevin simulation, the temperature is controlled through the fluctuation/dissipation effect. Because in the FDPB treatment there is effectively no distance cut off on the solvent forces, the electrostatic nonbonded interaction for the molecule was used with no cutoff and the extended electrostatics option, with a constant molecular dielectric of 1. For the ZAP part of the simulation, the solvent dielectric was $\varepsilon_o = 80$ with zero ionic strength and a surface free energy term of $\gamma = 5$ cal/Å^2. To be consistent with the MD force-field treatment, the ZAP interior dielectric was $\varepsilon_i = 1$. We note, however, that because the ZAP implementation of the FDPB method uses 3D Gaussian density functions to represent atomic volume, the dielectric assigned to the interior of the molecule is determined by the set of overlapping Gaussian density profiles, and so while still low compared to the solvent, it can be somewhat higher than $\varepsilon_1 = 1$ in the interstices between atoms. We have found, as in previous work[38] that this gives solvation energies quite close to those obtained from the DelPhi two-dielectric molecular surface boundary implementation of FDPB calculations, which is usually done with an interior dielectric of 2 (to allow for electronic polarization). Atomic radii and charges for ZAP were taken directly from CHARMM, i.e., were identical to those used in the CHARMM implementation of the GB solvent model.[25] The ZAP forces are applied at every time

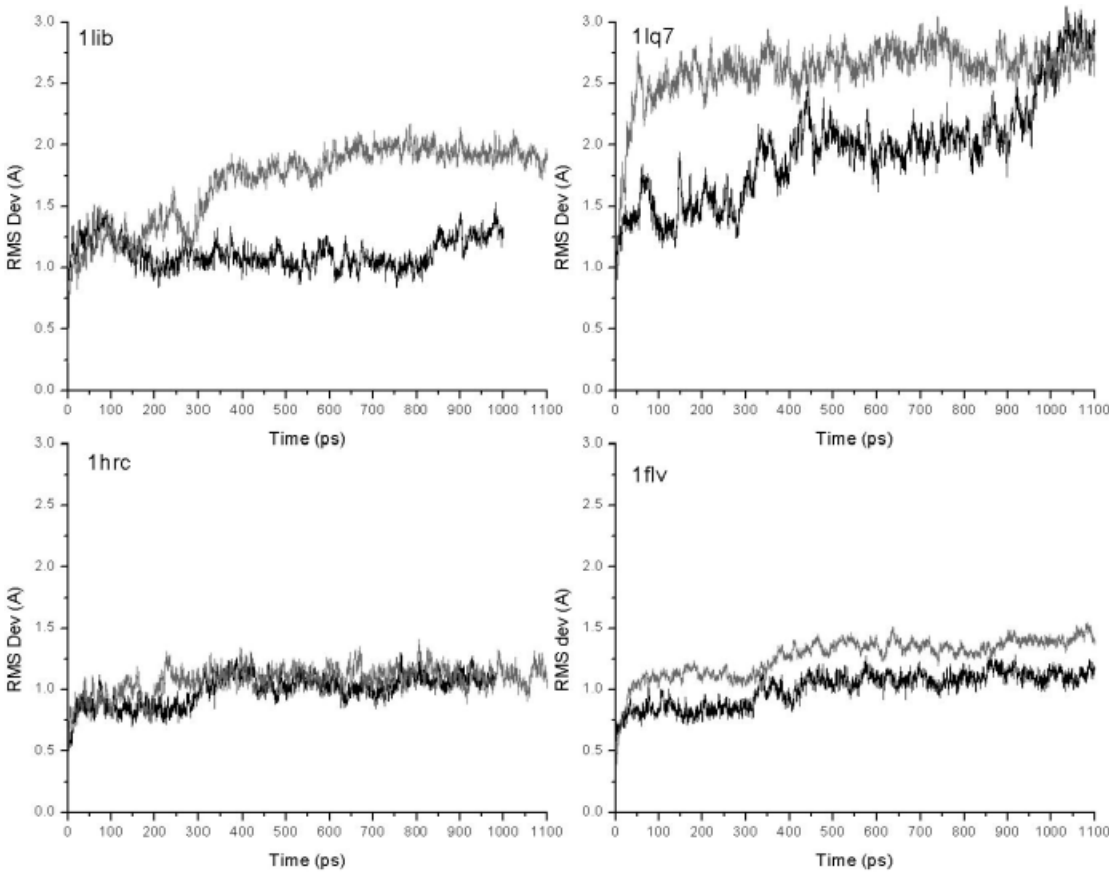

Figure 3. Root-mean-square deviation of backbone atoms as a function of simulation time for 1LIB, 1LQ7, 1HRC, and 1FLV. Bold trace: explicit solvent. Light trace: ZAP-implicit solvent model. The elapsed time includes the heating and equilibration phases.

step but for computational efficiency they are updated with a new FDPB/area calculation once every 10 MD steps because the protein/solvent boundary changes shape rather little over a time of 10–20 fs. To establish whether this update frequency is sufficient, some simulations were repeated with a ZAP update every 5-MD steps, producing simulations with virtually identical stability, as judged by the root-mean-square (rms) deviation from crystal structure over time. For the 1FLV simulation, the CHARMM-ZAP simulations were also run with the accessible solvent area Langevin or stochastics dynamics model to mimic the effect of explicit solvent fluctuation and friction.[60] The friction coefficient beta for each atom was proportional to its fractional surface accessibility. Surface accessibility of the atoms was calculated using the program solvent accessible surface area program SURFCV.[61] Completely buried atoms had a coefficient of 0, i.e., no Langevin term, while a fully exposed atom had the CHARMM recommended coefficient of 70 ps^{-1}. Results presented here for 1FLV are for the stochastic dynamics CHARMM-ZAP simulation.

Analysis of Trajectories

To analyze the stability of the trajectories, the realism of the simulation, and to compare explicit and implicit solvent simulations, the following analyses were performed: time evolution of the radius of gyration and rms deviation of coordinates with respect to the X-ray/NMR structure, to determine stability. Time evolution of native H-bonds and native tertiary contacts (defined as a donor H and acceptor O/N within 2.2 Å and as a C_β–C_β pair within 6 Å in the original X-ray/NMR structure, respectively). In the ZAP/MD simulations the time evolution of the surface area and solvent electrostatic energy terms was also used to assess stability. For charged sidechains, the rms fluctuation in position compared to their mean position was computed. The square of the amide NH and methyl generalized NMR order parameter, S^2 was computed for each NH/methyl containing residue using the method of Chatfield et al.[62] as described previously,[63] and compared to NMR data where available.

Results and Discussion

Root-Mean-Squared Deviations

The root-mean-square deviation (RMSD) of the coordinates from the original PDB structures were computed for each protein for

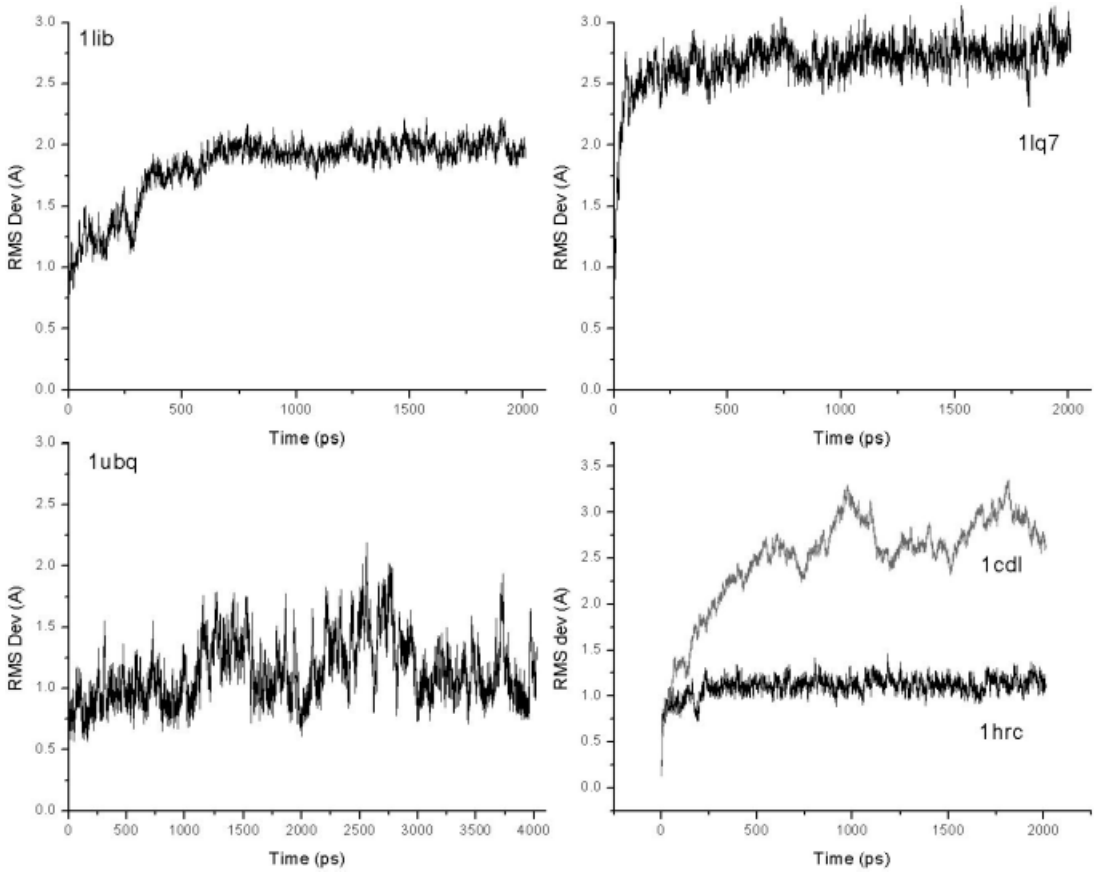

Figure 4. Root-mean-square deviation of backbone atoms as a function of extended simulation time for selected proteins using the implicit solvent model.

explicit and implicit solvent simulations as a function of time. A least-squares fitting to align the alpha-carbon atoms from the trajectory with those of the reference coordinate set was followed by calculation of the RMS deviation of either backbone atoms or of the whole molecule. Figures 2 and 3 show the RMSD behavior for backbone atoms. The common feature of all the simulations is an initial steep rise due to thermalization of the protein, followed by a plateau. In almost all cases a definite plateau is seen, indicating that a stable simulation is achieved. The position of the plateau (i.e., the mean extent of deviation from experimental structure) varies from case to case, but is mostly between 1.5 to 2.9 Å, which is typical of stable MD simulations of globular proteins in solvent. Extending the implicit solvent simulations to 2 ns for four of the proteins (1LIB, 1LQ7, 1HRC, and 1CDL), and for one protein to 4 ns (1UBQ) showed that the proteins maintained their stability (Fig. 4).

The one exception to the clear RMSD plateau behavior is seen for the explicit water simulations of 1LQ7, in which the RMSD appears to be still climbing towards 3 Å at the end of the simulation. Interestingly, the implicit solvent simulation for this protein also has a rather large RMSD, but a clear plateau value is evident, from early in the simulation, about 100 ps onwards. The average RMSD over the last 500 ps of the simulation is summarized in Table 3 for backbone and all atoms. Overall the ZAP simulations show a somewhat higher mean RMSD, in most cases not significant, and in all cases within the acceptable range for stable simulations. The exception is for ubiquitin, where the ZAP simulation has a significantly smaller RMSD than the explicit water simulation.

Table 3. Root-Mean-Square Deviations from Crystal Structure.

	Explicit TIP3P water		ZAP Implicit solvent	
Protein	Backbone	All atoms	Backbone	All atoms
1ATJ	1.1	1.5	1.4	1.8
2ATJ	1.1	1.4	1.4	2.2
1LIB	1.1	1.6	1.9	2.4
1HRC	1.0	1.6	1.1	1.7
1LQ7	2.1	2.6	2.7	3.4
1FLV	1.1	1.6	1.3	3.0
1CDL	2.1	2.7	3.1	3.4
1UBQ	2.1	2.6	1.0	1.4

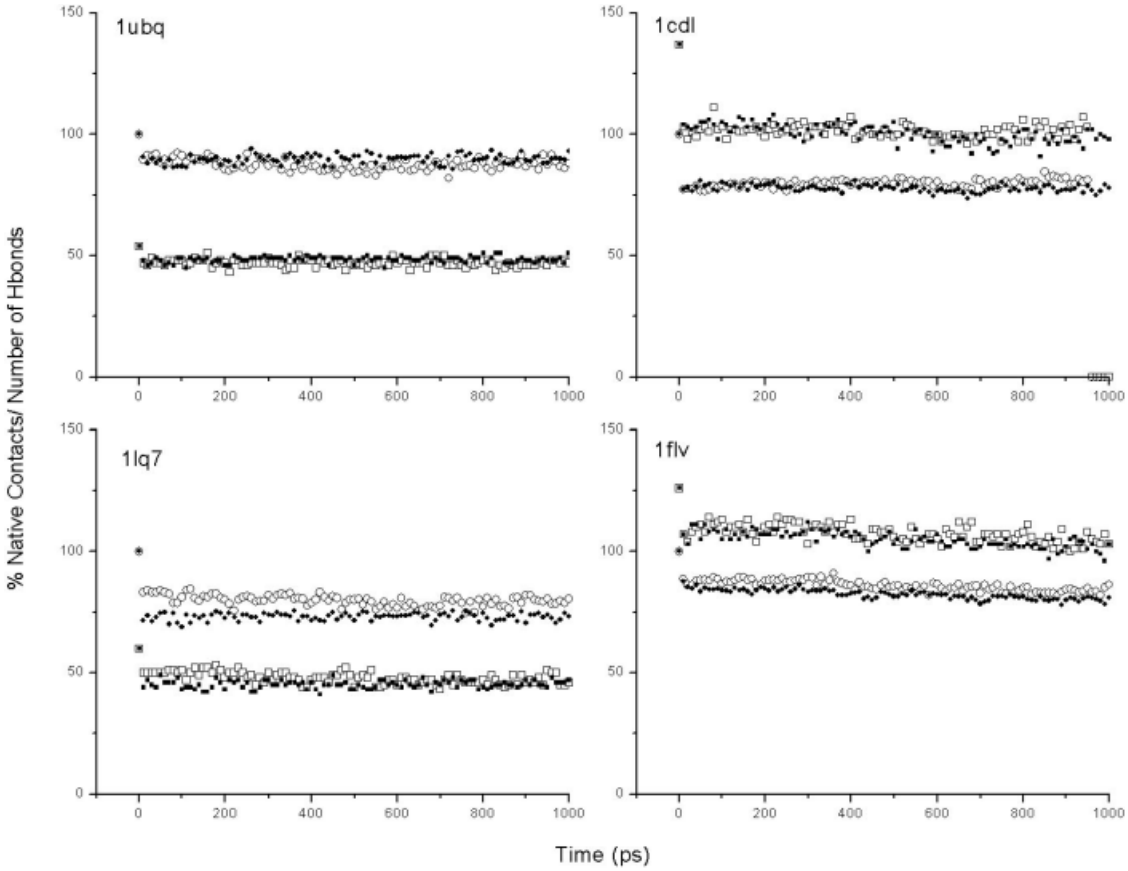

Figure 5. Native tertiary contacts (circles) and H-bonds (squares, defined in text) as a function of simulation time for 1UBQ, 1CDL, 1LQ7, and 1FLV. Open symbols: explicit water model. Filled symbols: ZAP implicit solvent model. The elapsed time includes the heating and equilibration phases.

Figure 5 shows native H-bond and native tertiary contact plots for four of the proteins as a function of time. The two solvent treatments give almost identical results: H-bonds and native contacts drop to about 80% of their initial value very quickly (within a few picoseconds) and then fluctuate slightly around their plateau value for the rest of the simulation. Results for 1ATJ, 2ATJ, 1HRC, and 1LIB (not shown) were very similar. Overall, the implicit solvent model gave marginally less native contacts and H-bonds than the explicit water model, again excepting 1UBQ, where the explicit water model had lower values. These results attest to the stability of the simulations, even in the case of 1LQ7 with explicit water, which has a rather large RMSD. In interpreting RMSD behavior it should be noted that some significant deviation from X-ray structures is expected due to thermal motion and the absence of crystal-contact constraints. Thus, providing the simulation is stable and the RMSD within usual limits (typically <2.5 Å for backbone) a lower RMSD cannot be simply be interpreted as reflecting a more accurate simulation, and the values in Table 3 indicate that the two solvent treatments produce stable simulations of very similar quality. The two cases of 1UBQ and 1LQ7 are somewhat different, in that the RMSD profiles are quite different in the two solvent models, and these two cases were analyzed in further detail.

Graphical examination and analysis of snapshots of 1UBQ from the last 500 ps of the explicit and implicit solvent simulations provided an explanation for the much larger deviation of the former from the X-ray structure (Fig. 6). The figure shows a representative snapshot from the latter part of the explicit and implicit water simulations overlaid on the X-ray structure. For both simulations the deviations from the X-ray structure are fairly small over the most of the protein and are typical of those expected from thermal motion. The exception is in the C-terminal end. In the crystal structure there is a salt bridge between D39 and R74. This salt bridge is maintained in the ZAP simulation. In the explicit water simulation, however, an alternative salt bridge, D39–R72 is formed. The C-terminal with its now unattached R74 shows a large deviation from crystal structure, hence giving a much larger overall RMSD than the implicit solvent simulation. Salt bridge stabilities are known to depend on a subtle balance between charge–charge and charge solvent energy terms[64,65] and the effect seen in the C-terminus of ubiquitin. We speculate that the ZAP model does a better job in this case than the explicit water simulation because the correct bulk water dielectric behavior is already built into the model, and does not have to be simulated—a challenging task for any explicit water model—although a single case of better agreement should not be over interpreted.

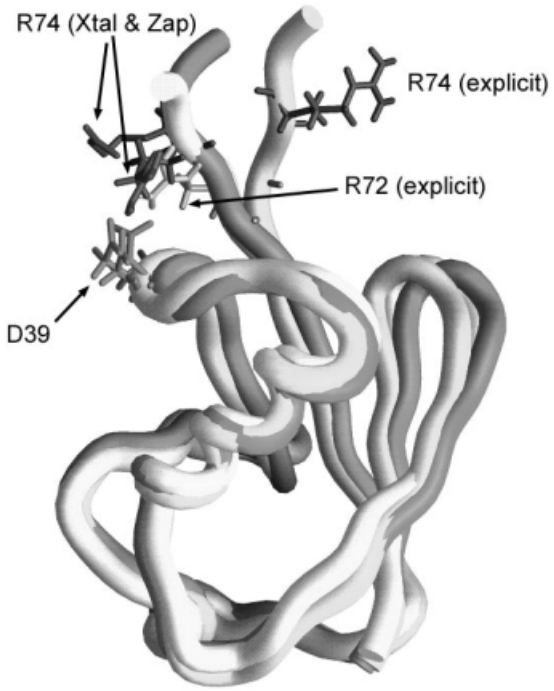

Figure 6. Backbone worm traces of 1UBQ from X-ray structure (shaded) and simulations with ZAP implicit solvent and explicit solvent (indicated on figure).

Graphical analysis of snapshots of 1LQ7 during the explicit and implicit simulations revealed that the three helix bundle structure was quite dynamic during the simulations compared to the other proteins. This is probably a consequence of 1LQ7 being a *de novo* design protein, in which the core is primarily stabilized by hydrophobic interactions, which allow for a malleable set of tertiary interactions. This contrasts with more highly evolved native proteins with more structurally specific packing and H-bond interactions producing a more defined core structure. In the case of the explicit solvent simulation of 1LQ7, these structural fluctuations resulted in a gradual sliding of the middle helix in its N-terminal direction with respect to the other two helices (Fig. 7). By the end of the simulation this helix had moved almost one turn, resulting in a large and gradually increasing RMSD over the course of the simulation. In contrast, the sliding was much smaller in the ZAP simulation, resulting in a more stable simulation with a smaller RMSD. Again, this might be a consequence of having the correct bulk water behavior (in this case the hydrophobic term) built into the implicit solvent model. Again, a single case should not be over interpreted as indicating the general superiority of the implicit solvent model for treating the hydrophobic term.

Summarizing the RMSD time course results, typical plateau values of 1–1.9 Å were seen in both solvent models for 1ATJ, 2ATJ, 1LIB, 1HRC, and 1FLV, indicating that both simulation protocols maintain an overall structure that is very close to the one determined by X-ray crystallography. Rather larger RMSDs are seen for 1UBQ for the explicit water case and 1LQ7 for the reasons given above. In addition, 1CDL has a rather higher plateau RMSD than the other proteins in both solvent models. This is due to movement of the two globular end domains with respect to each other due to flexing of the long helix linking them. This is a well-documented dynamic feature of calmodulin[66] related to its binding and allosteric action, which is reproduced in both solvent models, again indicating that the simulations are realistic. The fact that stable simulations can be obtained on a wide variety of proteins with one implicit solvent protocol and one uniform set of parameters reinforces conclusions from earlier work by Lu and Luo[22] about the crucial importance of using a smooth description of the solute–solvent dielectric boundary. This, in turn, produces smoothly varying electrostatic solvent forces that are essential for a robust and accurate FDPB implicit solvent model.

Intramolecular and Solvation Energy Balance

Figure 8 shows the time course of the CHARMM intramolecular electrostatic energy, and the ZAP solvation energy for 1CDL during the implicit solvent simulation. At the beginning of the simulation the solvation energy rapidly becomes more favorable as the implicit solvent model drives the protein charged and polar groups into better solvated conformations. There is a concomitant weakening, or screening of the intramolecular electrostatic interaction. These two effects are on the order of several hundred kcal/mol, but largely cancel each other. Within a few picoseconds a stable balance is achieved between the two effects, which persists throughout the simulation. Fluctuations in the molecule–solvent term tend to be smaller than in the intramolecular term because the long range nature of the solvent reaction field term makes it less sensitive to angstrom level fluctuations in protein atom positions. The general behavior of a rapid and mutually compensating strengthening of solvation electrostatics and weakening of intramolecular electrostatics, each on the order of several hundred kcal/mol, followed by a stable balance was seen in the other seven proteins (results not shown).

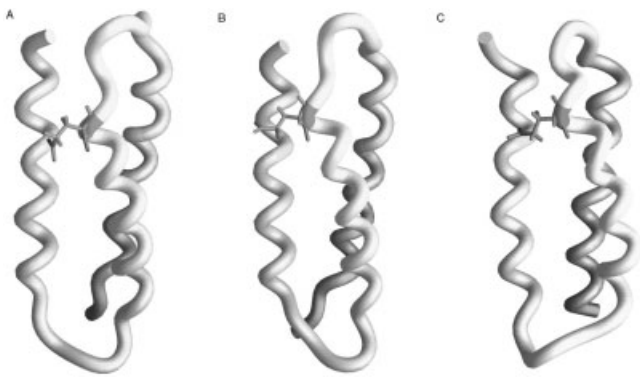

Figure 7. Backbone worm traces of 1LQ7 from X-ray structure (left) and simulations with explicit solvent (middle) and ZAP implicit solvent (right). The side chain of residue 29 on helix 2 is shown to indicate the position of helix 2 with respect to the other two helices.

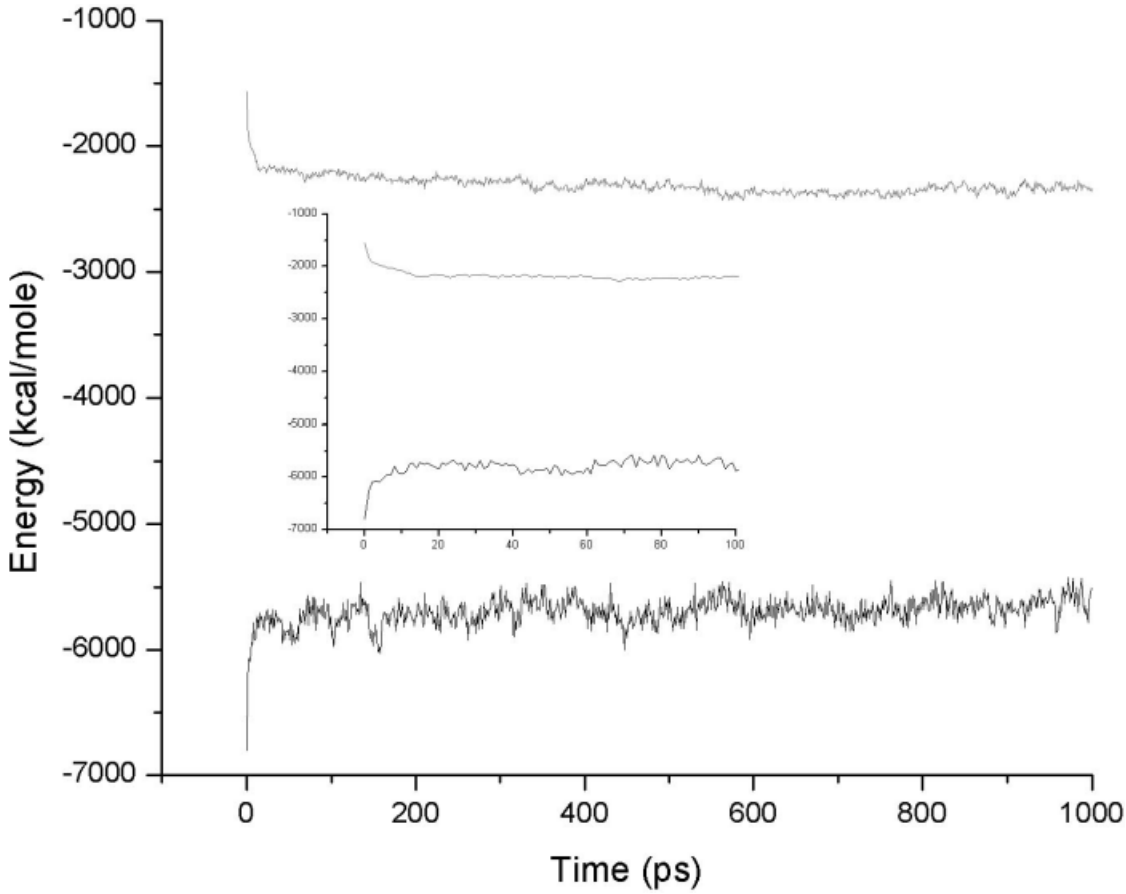

Figure 8. Time course of intramolecular electrostatic energy (lower trace) and ZAP solvation energy (upper trace) for the 1CDL simulation. The elapsed time includes the heating and equilibration phases. Inset plot shows expanded time scale of the initial 100 ps of the simulation.

Side Chain Fluctuations

Because in our initial implementation of the ZAP model there is no solvent friction, one might expect that surface side chains, especially charged ones, might exhibit more mobility compared to explicit water simulations. To examine this, the RMS fluctuation in position of the charged side chains, with respect to their average position was computed for three proteins, 1LQ7, 1HRC, and 1FLV (Table 4). Surprisingly, the mobility of these side chains is little different in the explicit and implicit solvent models. The mobility is slightly less with the implicit model in two of the cases, and higher in the third, 1FLV. For the 1FLV simulation, solvent friction was added to solvent exposed atoms through CHARMM's Langevin or stochastic dynamics option. The results are little different from a no-friction CHARMM-ZAP simulation, comparing for instance with 1HRC in Table 4, which exhibits a very similar overall RMSD from initial structure to 1FLV. For 1LQ7, there is more charged side-chain mobility, but this appears equally in both solvent models, and just reflects the more dynamic overall structure of the *de novo* designed three-helix bundle. Overall, the mobility of side-chain mobility seems to be determined primarily by their protein environment as differences in side-chain mobility are larger between proteins than between solvent models.

NMR Order Parameters

The square of the generalized NMR order parameter, S^2, provides information on the extent of angular motion of amide and methyl groups in a protein on the picosecond time scale.[67] It is straightforward to calculate NMR order parameters from an MD trajectory of sufficient length,[62] so comparison of measured and calculated parameters is useful as a site resolved probe of the realism of an implicit solvent dynamics simulation.[24] Nearly complete (on a

Table 4. Root-Mean-Square Deviation and Fluctuation in Charged Side Chains (Å).

	Explicit		Implicit	
Protein	Deviation[a]	Fluctuation[b]	Deviation[a]	Fluctuation[b]
1LQ7	2.7	0.17	3.1	0.11
Cytochrome *c*	2.1	0.12	2.1	0.09
1FLV	2.0	0.06	1.9	0.08

[a]Root-mean-squared deviation from X-ray/NMR structure.
[b]Root-mean-squared fluctuation around average simulation position.

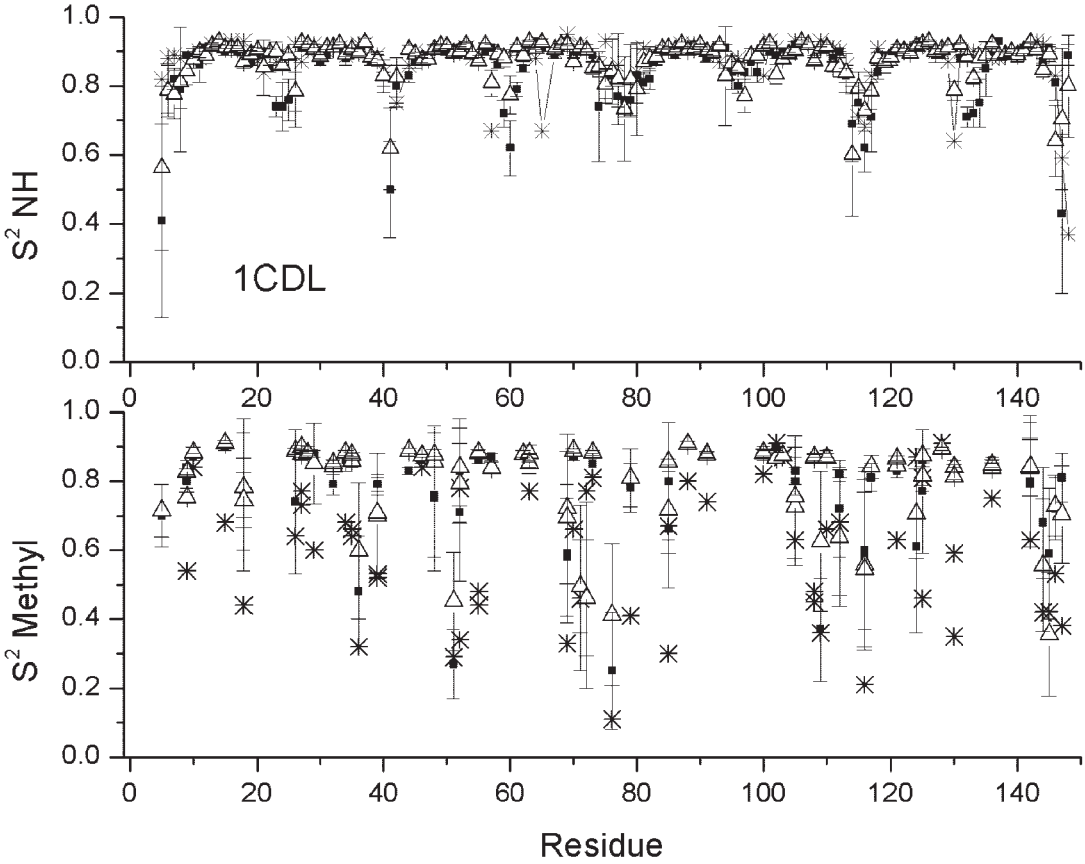

Figure 9. Square of the generalized order parameter S^2 vs. sequence for calmodulin–smMLCKp complex. (*) NMR data.[39,40] (■) Explicit solvent. (△) Zap solvent. Upper plot: amide NH groups; lower plot: methyl groups.

residue basis) experimental data for NH and methyl S^2 was available for four of the proteins in this study: 1CDL, 1UBQ, 1FLV, and 1LIB. Figure 9 shows the comparison of NH and Methyl group S^2 values between experimental and the two solvent models for the calmodulin-smMLCKp complex. Error estimates on the calculated S^2 values were obtained from the standard deviation of S^2 calculated from successive 200-ps time frames of a 1-ns simulation. Errors are representative of the other three proteins, and so are only shown on one figure for clarity. The experimental data for NH groups show the characteristic pattern of high values (0.85–0.91) within secondary structure elements, with lower values in the loop regions and the last one or two residues of secondary structure elements. Methyl data show no simple correlation with secondary structure, and represent a complex balance of tertiary interaction and packing effects that produce quite heterogeneous dynamic behavior in the protein core.[63] Visual comparison of NMR data with the calculated amide S^2 values show good agreement with experiment for both explicit and implicit solvent models: the overall profile of highs and lows is reproduced, the absolute values in the high S^2/secondary structure regions are closely matched. The systematic pattern of lower values in the loop regions is also reproduced, although the absolute values are not matched as closely as within secondary structure elements. The larger error bars in the loop regions indicate that the simulations have significant uncertainty here, one source being a poorer sampling of more infrequent, but larger scale motions that depress the S^2 values. The story for comparison of NH S^2 values of 1UBQ (Fig. 10) is similar, the profile of highs and lows being reproduced perhaps even more closely. A notable deviation between the two solvent models occurs at the C-terminal end of 1UBQ (where, unfortunately, NMR values cannot be obtained). Here, the ZAP model gives consistently higher S^2 values than the explicit water model. This can again be explained in terms of the maintenance of the D39–R74 salt bridge in the implicit solvent simulation, vs. the formation of the D39–R72 salt bridge in the explicit water simulation (see above). In the latter case, formation of a salt bridge with an arginine further from the C-terminus allows a longer section of C-terminus to move, producing lower order parameters for these residues.

For the methyl values, there is no clear pattern in the experimental data as a function of sequence in either 1CDL or 1UBQ, and a larger range of values, which makes it harder to compare the calculated and experimental profiles. Both simulated sets tend to have lower values where the experiment does, and higher ones where the experiment does, but clearly the agreement is not as good as for NH groups. Visual inspection also suggests that the

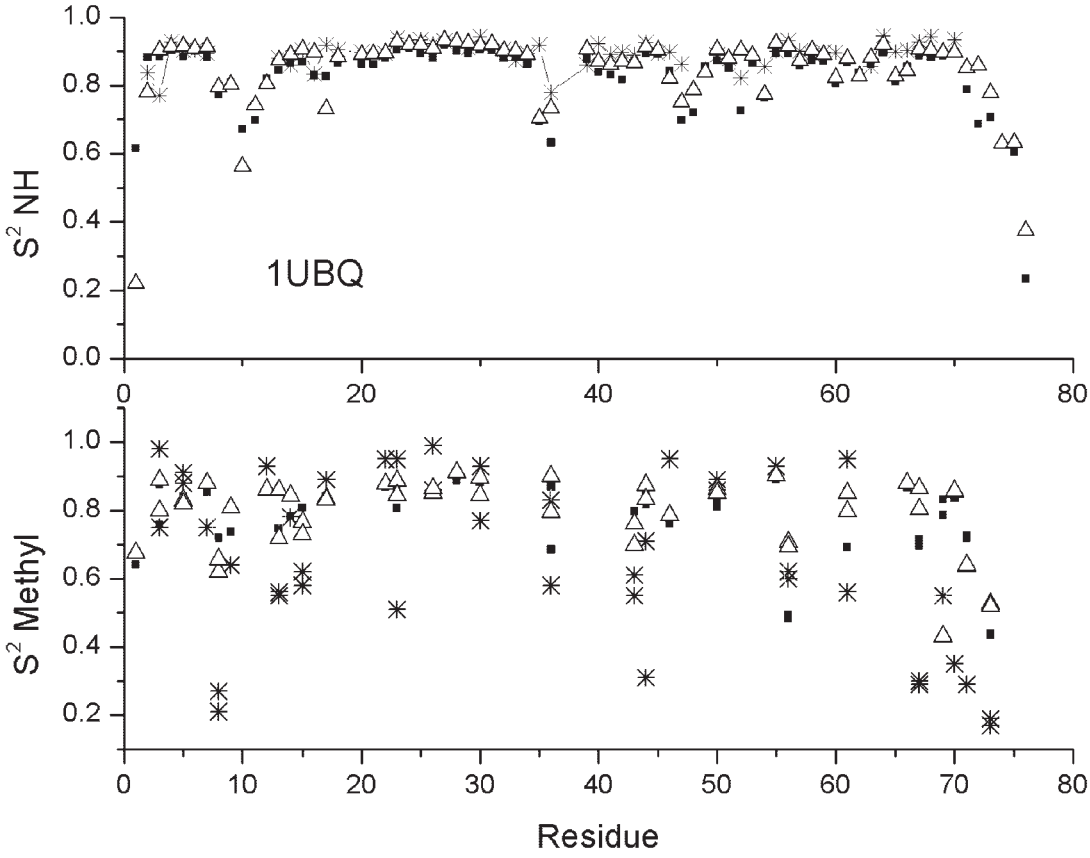

Figure 10. Square of the generalized order parameter S^2 vs. sequence for ubiquitin. (*) NMR data.[41] (■) Explicit solvent. (△) Zap solvent. Upper plot: amide NH groups; lower plot: methyl groups.

methyl S^2 profile of 1UBQ is fit better by the simulations than is 1CDLs. However, it is difficult to quantify this from straight comparison of S^2 vs. sequence plots. It is also difficult to assess from comparison of raw S^2 values the relative fidelity of the two solvent models. To address this, a more quantitative comparison was made as follows. Each experimental or simulated set of residue S^2 values was transformed into a set of "*Z*-values" (in the sense used in statistics) by

$$Z_i^{\text{set}} = (S_i^{2\,\text{set}} - \langle S^{2\,\text{set}}\rangle)/\sigma_{S^{2\,\text{set}}} \quad (1)$$

where the subscript i refers to the ith residue, $\langle S^{2\,\text{set}}\rangle$ and $\sigma_{S^{2\,\text{set}}}$ are the mean and standard deviation of the set of order parameters taken over all the residues of the protein, and set = NMR, explicit or ZAP. The residual "*Z*-value" for the explicit and ZAP data sets is

$$\Delta Z_i^{\text{explicit/ZAP}} = Z_i^{\text{explicit/ZAP}} - \Delta Z_i^{\text{NMR}} \quad (2)$$

Transformation into "*Z*-value" removes the effect of systematic over- or underestimate, and of different dynamic ranges, so that the residual *Z*-value enables one to better compare the overall sequence profiles of high/low order parameters. Put another way, if the residual *Z*-values of a simulation data set were all zero it would mean that it reproduced the experimental profile of order parameters exactly to within a single additive and single multiplicative constant. It thus enables one to focus on the principle use of the order parameters in this context: as a measure of the relative extent of motion of different parts of a protein, and to compare the two solvent models for relative accuracy. Figures 11 and 12 show the residual *Z*-value profiles for NH and methyl S^2 data, respectively, for all four proteins for which close to full NMR data was available. The good agreement inferred for the NH data sets for 1CDL and 1UBQ is confirmed by the residual *Z*-profiles, which generally lie well within ± 1 Z (one "standard deviation") over the whole sequence. The poorer agreement with experiment inferred for the methyl sets is also borne out in the residual *Z*-value profiles, as well as the better agreement for 1UBQ methyls than 1CDL methyls. The residual *Z*-value profiles provide a robust way to quantify agreement between two sets of residue S^2 values, and to identify regions of the protein where agreement is better or worse. From the ΔZ profiles one can conclude that generally the NH S^2 values are better reproduced than the methyl S^2 values (the exception is for 1FLV), and that the simulations reproduce the relative mobility profiles of 1CDL and 1UBQ best, followed by 1LIB, and last by 1FLV. 1FLV is protein with an unusually rigid backbone as judged by measured ND S^2 values.[42] This may arise from the rigidity of the flavin cofactor, but the exact reason is not known. However,

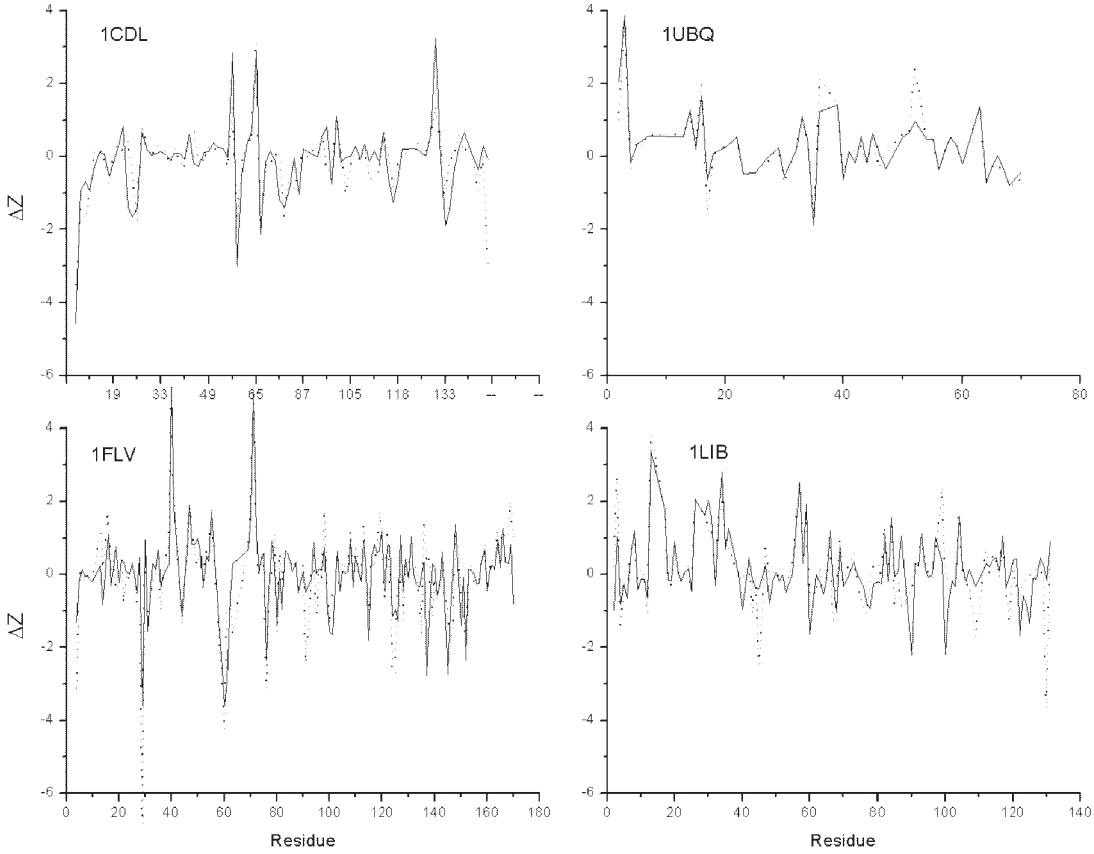

Figure 11. Residual order parameter *Z*-values (ΔZ) vs. sequence for amide groups using experimental NMR data[39–43] for reference. Solid line: explicit solvent; dashed line: implicit ZAP solvent.

this rigidity is not reproduced in either the explicit of implicit solvent simulation, perhaps because the flavin force field parameters, especially those governing out of plane motions, are not precise enough. Nevertheless, the most important conclusion to be drawn from the residual *Z*-profiles is that there is no clear distinction in accuracy between the explicit and implicit solvent simulations. In some protein or protein regions one treatment does better; in other regions the other does better.

A single overall measure of agreement between measured and simulated S^2 profiles can be obtained from the unsigned average of ΔZ over all the residues:

$$\langle|\Delta Z^{\text{explicit/ZAP}}|\rangle = \frac{1}{N}\sum|\Delta Z_i^{\text{explicit/ZAP}}| \tag{3}$$

These values are collected in Table 5, and the numbers summarize in a concise form the relative accuracy of simulating NH vs. methyl motions, and explicit vs. implicit models.

The order parameter S^2 data in Figures 9–12 and the mean side chain fluctuation data (Table 4) describe the relative mobility of different groups in terms of their extent of motion, but do not describe the time scale of these motions. One measure of motional time scales can be obtained from the time constant τ for the angular motions described by the order parameter S^2. The time correlation function for angular motion of a unit vector $\boldsymbol{\mu}$ describing the N—H or methyl bond direction is[67]

$$C(\Delta t) = \langle P_2(\boldsymbol{\mu}(t)\cdot\boldsymbol{\mu}(t+\Delta t))\rangle_t \tag{4}$$

where the second Legendre polynomial is $P_2(x) = (3x^2 - 1)/2$, and $\langle\ \rangle_t$ indicates an average over all time origins. The long time limit of C (with the effect of molecular rotation removed) gives S^2. If the decay of C from its value of 1 at $\Delta t = 0$ to its limiting value of S^2 is described by a single time constant τ, i.e., $C(\Delta t)$ can be modeled by monoexponential decay kinetics as[67]

$$C(\Delta t) = S^2 + (1 - S^2)e^{-\Delta t/\tau} \tag{5}$$

then τ can be obtained from both NMR data and from simulations by calculating $C(\Delta t)$ and fitting to eq. (5). Methionine residues of 1CDL were chosen for comparison of experimental and simulated τ values for a number of reasons. First, methionine is the residue with a methyl group furthest from the backbone. It thus provide the "purest" reporter of side-chain dynamics. Second, calmodulin is unusually rich in methionine residues, having nine of them, and they show a wide range of order parameter values, indicating heterogeneous dynamics. Third, although precise determination of τ is difficult, good experimental data for τ of these 1CDL methi-

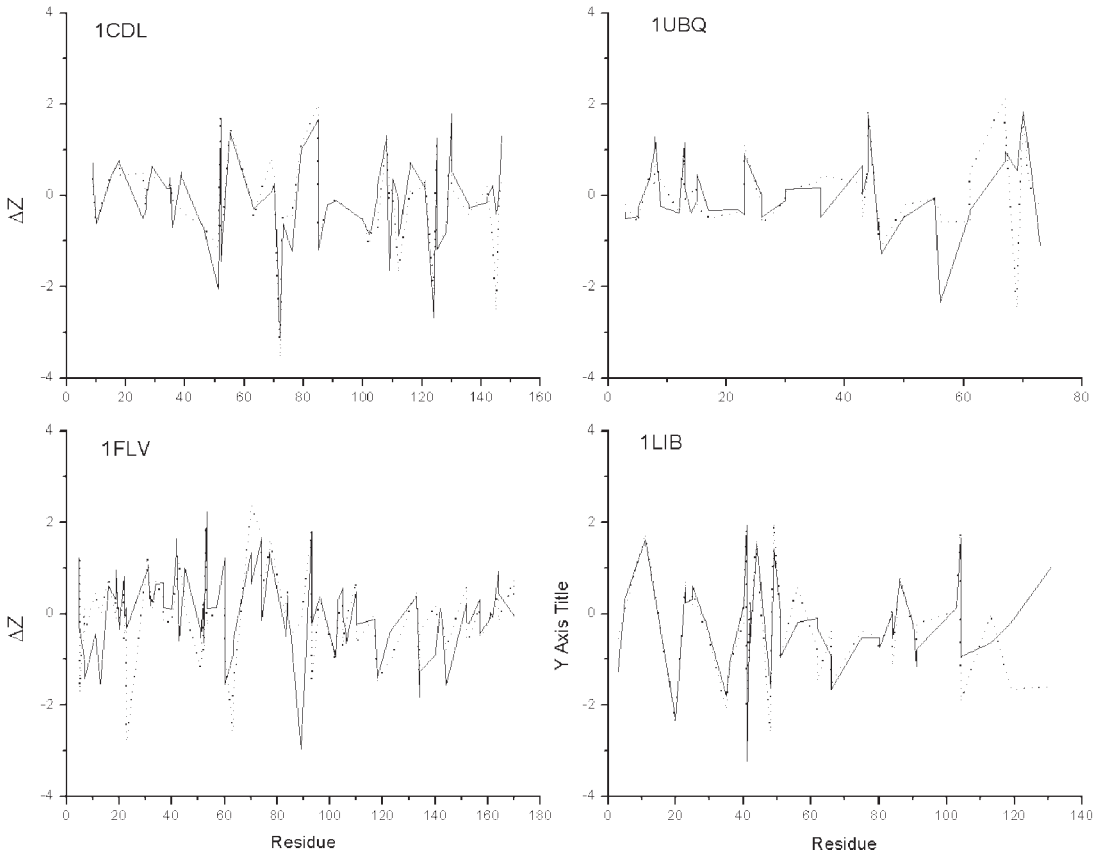

Figure 12. Residual order parameter *Z*-values (ΔZ) vs. sequence for methyl groups using experimental NMR data[39–43] for reference. Solid line: explicit solvent; dashed line: implicit ZAP solvent.

onine residues has been obtained.[40,68] 1CDL methionine τ values were calculated from the explicit and implicit solvent simulations and compared to experiment (Fig. 13). The magnitude and overall pattern of high and low values is surprisingly well reproduced in both simulations with the exception of M124 in the ZAP simulation. Comparing the overall rates of motion using the average methionine τ values, these are 14 ± 4 ps for the NMR values, 14 ± 3 ps for the explicit solvent simulations, and 10 ± 6 ps for the implicit solvent simulations. Overall angular motion time scales are statistically indistinguishable between the two solvent models, and between simulation and experiment. Arguably, solvent would

Table 5. Mean Unsigned Residual *Z*-Values for Calculated NMR Order Parameters.

	Explicit TIP3P water		ZAP Implicit solvent	
Protein	$\langle\lvert\Delta Z\rvert\rangle$NH	$\langle\lvert\Delta Z\rvert\rangle$Methyl	$\langle\lvert\Delta Z\rvert\rangle$NH	$\langle\lvert\Delta Z\rvert\rangle$Methyl
1LIB	0.68	0.82	0.70	0.90
1FLV	0.71	0.67	0.84	0.69
1CDL	0.69	0.73	0.62	0.73
1UBQ	0.60	0.65	0.66	0.59

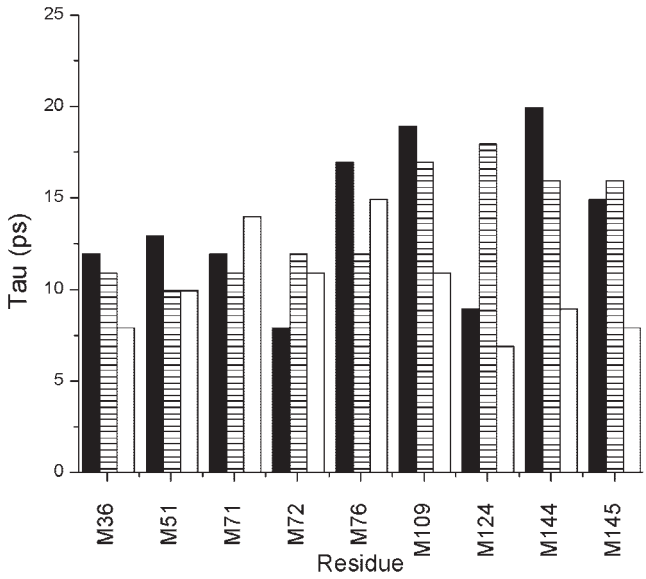

Figure 13. Relaxation time for the generalized order parameter (tau) for calmodulin–smMLCKp complex methionine methyl groups from NMR data[39] (solid bars), explicit water simulations (shaded bars), and implicit solvent (unfilled bars).

Table 6. Comparative Timings for Different Solvent Treatments.[a]

Molecule	Protein atoms (explicit solvent atoms)	No solvent	Zap[b]	Explicit[b]
Ubiquitin	1231 (15,019)	56	306 [0.25] (5.5)	1170 [0.95] (21)
Calmodulin	2155 (24,906)	124	639 [0.30] (5.1)	4750 [2.2] (38)
HRPC	4771 (45,063)	320	1591 [0.33] (5.0)	8640 [1.8] (27)

[a]In seconds, for 1 ps @ 1 fs/step on a 1.6 Ghz AMD processor with 1 Gb Ram.
[b]Figures in [ ] are timings per protein atom. Figures in ( ) are times relative to "no solvent" simulations.

have most effect on protein motion time scales for surface residues, rather than methionines, which are mostly buried. We thus identified all the methyl containing residues with significant (>50 Å^2) exposed side-chain surface area (residues T5, I9, T44, A46, A57, L69, T79, A103, L116, T117, I130, V142, and T146) and compared the mean τ values for explicit and implicit solvent simulations. The values were 16 ± 9 and 12 ± 9 ps, again statistically indistinguishable. Overall, the τ value results indicate that both solvent models have realistic angular motion time scales and that the time scale is independent of the solvent treatment. The latter implies that the time scale of these kinds of motion is governed principally by the temperature and the overall shape of the energy surface, including the rotamer torsion barriers, rather than by discrete solvent friction/fluctuation effects.

Relative Timings

Comparative timings for explicit solvent and ZAP implicit solvent models are presented in Table 6 for three proteins in the set, including the smallest, Ubiquitin, the largest, horseradish peroxidase. Timings are given for a 1-ps simulation at 1-fs/step on a 1.6-GHz AMD processor with 1 Gb RAM. For comparison, timings with no solvent treatment (i.e., evaluating only the intramolecular forces) are given. Exact timings will depend on several factors such as the number of explicit waters, nonbonded cutoff, boundary conditions for the explicit solvent, grid resolution, and update frequency for the ZAP solvent, and other algorithmic parameters, which may be varied to some degree. Thus, relative timings are somewhat approximate. For example, explicit water timings may be improved by use of octahedral rather than rectangular/cubic boxes. Both explicit and ZAP solvent treatments involve a significant computational cost. Nevertheless, the results show that one can expect a factor of 4 to 8 speedup using the ZAP solvent treatment instead of explicit water for a typical protein simulation. From the timings per protein atom it can be seen that the ZAP times increase only slightly more than linearly, and so should scale well for larger simulations. In contrast to the experience of Lu and Luo,[22] we find that the cost of the Gaussian distribution type of smooth permittivity model is quite acceptable in MD simulations, perhaps because we are using the highly optimized OpenEye ZAP libraries.

Conclusions

The smooth permittivity implementation of the FDPB algorithm has made possible a stable, fast implementation of the FDPB/solvent-accessible surface area model for implicit solvation within MD. It is shown here that the method is robust, and can produce good quality simulations. We emphasize that initial testing of the method has been on a rather diverse and challenging set of proteins, and used the same implicit solvent protocol and parameters for all proteins. The validity of the FDPB implicit solvent model is evaluated by comparison with experiment and with the current standard for simulations—the explicit solvent model. There are two main criterion for comparison; the first involves the structure of the simulated biomolecule, and the second the dynamics of the structure. The time histories of the RMSDs from the ZAP simulation show, first, that the simulations are stable over a nanosecond time scale. Second, the RMSDs from the implicit solvent model compare well with those from explicit solvent. In many cases the average RMSD at the end of the simulation for both the ZAP and explicit water simulations were around 1.5. In cases where the differences exceeded 2 Å, it was mainly due to fluctuations at the termini where the protein is not expected to have a rigidity defined structure (1UBQ), or conformational fluctuations that have been linked to function (1CDL). As might be expected if the two solvent models are of comparable accuracy, in some situations the explicit model did better; in others, the implicit model did better. In two cases where overall solvation energetics seems to be of particular importance, the implicit model appeared to do slightly better, viz in the C-terminal salt bridging of 1UBQ, and in maintaining the 1LQ7 hydrophobic core. Cases where the explicit model did better may depend more on molecular detail of solvent structure and H-bonding. The overall conclusion from the comparison of explicit and implicit solvent models over a very diverse set of proteins, and using a number of measures of simulation stability and realism is that they are (apart from the computer time used) almost indistinguishable in broad terms. Indeed, given any of the comparative measures used here stripped of identifying information, it would be difficult to say whether they came from an explicit or implicit simulation.

Future directions for the smooth permittivity Finite Difference Poisson–Boltzmann model for implicit solvation in molecular dynamics include a detailed comparison with GB models. Implementation of fast, stable FDPB implementation of the type described here now makes this feasible. In addition, more extensive testing on a wider range of proteins and nucleic acids is needed, and is also now more feasible.

Acknowledgments

We thank Professor A. Joshua Wand for providing the most recent NMR order parameter data for 1CDL, 1FLV, and 1UBQ and for

many helpful discussions. We thank Laura Marian for carrying out the calculations of angular correlation time constants.

References

1. Duan, Y.; Kollman, P. A. Science 1998, 282, 740.
2. Mayor, U.; Guydosh, N. R.; Johnson, C. M.; Grossmann, J. G.; Sato, S.; Jas, G. S.; Freund, S. M.; Alonso, D. O.; Daggett, V.; Fersht, A. R. Nature 2003, 421, 863.
3. Sharp, K. In Encyclopedia of Computational Chemistry; Schleyer, P., Ed.; John Wiley and Sons: Chichester, England, 1998.
4. Sharp, K. A. In Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications; van Gunsteren, W. F.; Weiner, P. K.; Wilkinson, A. J., Eds.; Escom: Leiden, 1993.
5. Davis, M. E.; McCammon, J. A. J Comput Chem 1990, 11, 401.
6. Sharp, K. A. J Comp Chem 1991, 12, 454.
7. Zauhar, R. J Comp Chem 1991, 12, 575.
8. Niedermeier, C.; Schulten, K. Mol Simulat 1992, 8, 361.
9. Gilson, M.; Davis, M.; Luty, B.; McCammon, J. J Phys Chem 1993, 97, 3591.
10. Gilson, M.; Davis, M.; Luty, B.; Madura, J.; McCammon, J. Biophys J 1993, 64, A 354.
11. Madura, J.; Gilson, M. Biophys J 1993, 64, A 220.
12. Sitkoff, D.; Sharp, K.; Honig, B. J Phys Chem 1994, 98, 1978.
13. David, L.; Luo, R.; Gilson, M. J Comp Chem 2000, 21, 295.
14. Schiffer, C.; Caldwell, J.; Kollman, P.; Stroud, R. Mol Simulat 1993, 10, 121.
15. Miyamoto, S.; Kollman, P. Proteins Struct Funct Genet 1993, 16, 226.
16. Tsui, V.; Case, D. A. J Phys Chem B 2001, 105, 11314.
17. Nicholls, A.; Honig, B. J Comp Chem 1991, 12, 435.
18. Holst, M.; Kozack, R. E.; Saied, F.; Subramaniam, S. Proteins 1994, 18, 231.
19. Luo, R.; David, L.; Gilson, M. K. J Comput Chem 2002, 23, 1244.
20. Lee, B.; Richards, F. M. J Mol Biol 1971, 55, 379.
21. Connolly, M. L. Science 1983, 221, 709.
22. Lu, Q.; Luo, R. J Chem Phys 2003, 119, 11035.
23. Still, C.; Tempczyk, A.; Hawley, R.; Hendrickson, T. J Am Chem Soc 1990, 112, 6127.
24. Calimet, N.; Schaeffer, M.; Simonson, T. Proteins 2001, 45, 144.
25. Dominy, B.; Brooks, C. J Phys Chem B 1999, 103, 3765.
26. Jayaram, B.; Sprous, D.; Beveridge, D. L. J Phys Chem B 1998, 102, 9571.
27. Lee, M. S.; Salsbury, F. R., Jr.; Brooks, C. L., III. J Chem Phys 2002, 116, 10606.
28. Onufriev, A.; Bashford, D.; Case, D. A. J Phys Chem B 2000, 104, 3712.
29. Qui, D.; Shenkin, P.; Hollinger, F.; Still, W. J Phys Chem 1997, 101, 3005.
30. Tsui, V.; Case, D. JACS 2000, 122, 2489.
31. Gallicchio, E.; Levy, R. M. J Comp Chem 2004, 25, 479.
32. Feig, M.; Im, W.; Brooks, C. L. J Chem Phys 2004, 120, 903.
33. Bashford, D.; Case, D. Annu Rev Phys Chem 2000, 51, 129.
34. Hassan, S.; Mehler, E.; Zhang, D.; Weinstein, H. Proteins 2003, 51, 109.
35. Feig, M.; Onufriev, A.; Lee, M.; Im, W.; Case, D.; Brooks, C. J Comp Chem 2003, 25, 265.
36. Jayaram, B.; Liu, Y.; Beveridge, D. L. J Chem Phys 1998, 109, 1466.
37. Im, W.; Beglov, D.; Roux, B. Comput Phys Comm 1998, 111, 59.
38. Grant, J. A.; Pickup, B. T.; Nicholls, A. J Comput Chem 2001, 22, 608.
39. Lee, A. L.; Wand, A. J. Nature 2001, 411, 501.
40. Lee, A. L.; Kinnear, S. A.; Wand, A. J. Nat Struct Biol 2000, 7, 72.
41. Lee, A. L.; Flynn, P.; Wand, A. J. JACS 1999, 121, 2891.
42. Liu, W.; Flynn, P.; Fuentes, E.; Kranz, J.; McCormick, M.; Wand, A. J. Biochemistry 2001, 40, 14744.
43. Constantine, K.; Friedrichs, M.; Wittekind, M.; Jamil, H.; Chu, C.; Parker, R. A.; Goldfarb, V.; Mueller, L.; Farmer, B. Biochemistry 1998, 37, 7965.
44. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; P. E. B. Nucleic Acid Res 2000, 28, 235.
45. Henriksen, A.; Schuller, D.; Meno, K.; Welinder, K.; Smith, A.; Gajhede, M. Biochemistry 1998, 37, 805.
46. Gajhede, M.; Schuller, D.; Henriksen, A.; Smith, A.; Poulos, T. Nat Struct Biol 1997, 4, 1032.
47. Bushnell, G.; Louie, G.; Brayer, G. J Mol Biol 1990, 214, 585.
48. Prabhu, N. V.; Dalosto, S. D.; Sharp, K. A.; Wright, W. W.; Vanderkooi, J. M. J Phys Chem 2002, 106, 5561.
49. Dai, Q.-H.; Tommos, C.; Fuentes, E. J.; Blomberg, M. R.; Dutton, P. L.; Wand, A. J. J Am Chem Soc 2002, 124, 10952.
50. Meador, W.; Means, A.; Quiocho, F. Science 1992, 257, 1251.
51. Vijay–Kumar, S.; Bugg, C. E.; Cook, W. J. J Mol Biol 1987, 194, 531.
52. Rao, S. T.; Schaffie, F.; Yu, C.; Satyshur, K. A.; Stockman, B. J.; Markley, J. L.; Sundarlingam, M. Protein Sci 1992, 1, 1413.
53. Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. J Chem Phys 1984, 81, 3684.
54. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical J Comp Chem 1977, 23, 327.
55. Essmann, U. L.; Perera, M. L.; Berkowitz, T.; Darden, T.; Lee, H.; Pedersen, L. J Chem Phys 1995, 103, 8577.
56. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179.
57. Kaposi, A. D.; Prabhu, N. V; Dalosto, S. D.; Sharp, K. A.; Wright, W. W.; Stavrov, S. S.; Vanderkooi, J. M. Biophys Chem 2003, 106, 1.
58. MacKerell, A. D.; Brooks, B.; Brooks, C. L.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. In The Encyclopedia of Computational Chemistry; Schleyer, P. v. R., Ed.; John Wiley & Sons: Chichester, 1998, p. 271.
59. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J Comp Chem 1983, 4, 187.
60. Pastor, R.; Karplus, M. J Phys Chem 1988, 92, 2636.
61. Sridharan, S.; Nicholls, A.; Sharp, K. A. J Comp Chem 1995, 16, 1038.
62. Chatfield, D.; Szabo, A.; Brooks, B. R. JACS 1998, 120, 5301.
63. Prabhu, N.; Lee, A.; Wand, J. W.; Sharp, K. A. Biochemistry 2002, 42, 562.
64. Soman, K.; Yang, A.-S.; Honig, B.; Fletterick, R. Biochemistry 1989, 28, 9918.
65. Hendsch, Z. S.; Tidor, B. Protein Sci 1994, 3, 211.
66. Kranz, J. K.; Flynn, P. F.; Fuentes, E. J.; Wand, A. J. Biochemistry 2002, 41, 2599.
67. Lipari, G. L.; Szabo, A. JACS 1982, 104, 4546.
68. Lee, A. L.; Sharp, K. A.; Kranz, J. K.; Song, X.; Wand, A. J. Biochemistry 2002, 41, 13814.