Molecular dynamics simulations of peptides and proteins with a continuum electrostatic model based on screened Coulomb potentials

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · APRIL 2003

Impact Factor: 2.63 · DOI: 10.1002/prot.10330 · Source: PubMed

4 AUTHORS, INCLUDING:

Harel Weinstein
Weill Cornell Medical College
365 PUBLICATIONS **13,170** CITATIONS

Molecular Dynamics Simulations of Peptides and Proteins With a Continuum Electrostatic Model Based on Screened Coulomb Potentials

Sergio A. Hassan, Ernest L. Mehler,* Daqun Zhang, and Harel Weinstein
Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York, NY, USA

ABSTRACT **A continuum electrostatics approach for molecular dynamics (MD) simulations of macromolecules is presented and analyzed for its performance on a peptide and a globular protein. The approach incorporates the screened Coulomb potential (SCP) continuum model of electrostatics, which was reported earlier. The model was validated in a broad set of tests some of which were based on Monte Carlo simulations that included single amino acids, peptides, and proteins. The implementation for large-scale MD simulations presented in this article is based on a pairwise potential that makes the electrostatic model suitable for fast analytical calculation of forces. To assess the suitability of the approach, a preliminary validation is conducted, which consists of (i) a 3-ns MD simulation of the immunoglobulin-binding domain of streptococcal protein G, a 56-residue globular protein and (ii) a 3-ns simulation of Dynorphin, a biological peptide of 17 amino acids. In both cases, the results are compared with those obtained from MD simulations using explicit water (EW) molecules in an all-atom representation. The initial structure of Dynorphin was assumed to be an α-helix between residues 1 and 9 as suggested from NMR measurements in micelles. The results obtained in the MD simulations show that the helical structure collapses early in the simulation, a behavior observed in the EW simulation and consistent with spectroscopic data that suggest that the peptide may adopt mainly an extended conformation in water. The dynamics of protein G calculated with the SCP implicit solvent model (SCP-ISM) reveals a stable structure that conserves all the elements of secondary structure throughout the entire simulation time. The average structures calculated from the trajectories with the implicit and explicit solvent models had a cRMSD of 1.1 Å, whereas each average structure had a cRMSD of about 0.8Å with respect to the X-ray structure. The main conformational differences of the average structures with respect to the crystal structure occur in the loop involving residues 8–14. Despite the overall similarity of the simulated dynamics with EW and SCP models, fluctuations of side-chains are larger when the implicit solvent is used, especially in solvent exposed side-chains. The MD simulation of Dynorphin was extended to 40 ns to study its behavior in an aqueous environment. This long simulation showed that the peptide has a tendency to form an α-helical structure in water, but the stabilization free energy is too weak, resulting in frequent interconversions between random and helical conformations during the simulation time. The results reported here suggest that the SCP implicit solvent model is adequate to describe electrostatic effects in MD simulation of both peptides and proteins using the same set of parameters. It is suggested that the present approach could form the basis for the development of a reliable and general continuum approach for use in molecular biology, and directions are outlined for attaining this long-term goal. Proteins 2003;51:109–125.** © 2003 Wiley-Liss, Inc.

Key words: MD simulation with an implicit solvent model; screened Coulomb potentials; simulation of protein G; simulation of dynorphin

INTRODUCTION

By describing the dynamic evolution of a system, molecular dynamics (MD) simulations are the appropriate tool for studying time-dependent properties of macromolecules and their interactions. Thus, thermodynamics properties can be evaluated from time averages of measurable quantities. However, two limitations of this computational approach stem from (i) the approximations in the force field used to define the system and (ii) the difficulties in representing the mostly polar-solvent environment in which biological macromolecules are usually immersed. Although far from complete, the development of force fields has made substantial progress, and high-quality force fields are available from various sources.[1–4] An accurate representation of the polar solvent is important because of its known role to regulate the dynamics and

Grant sponsor: National Institutes of Health; Grant numbers: P01 DA12923, DA 124080, K05 DA-00060, R01 DA 15170.

*Correspondence to: Ernest L. Mehler, Department of Physiology and Biophysics, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029. E-mail: mehler@inka.mssm.edu. URL: http://fulcrum.physbio.mssm.edu/~mehler/

Received 26 August 2002; Accepted 18 October 2002

stabilization of the macromolecules, thus playing a key role in their biological functions. Explicit representation of the solvent has been the approach of choice in computational simulations because of its reliability and conceptual simplicity. Although explicit solvent models are approximate, and improvements are still being reported,[5,6] they have been used successfully in many simulations that have provided interesting and important results, making possible the investigation of the effect of solvent on the structure, dynamics, and thermodynamic properties of a number of biological macromolecules (see, e.g., Refs. 7–11).

Unfortunately, an appropriate description of the effects of the solvent on the macromolecule requires a large number of solvent molecules to be included in the calculation, thus leading to a significant computational burden. For this reason, MD simulations of biological macromolecules in aqueous solution are usually limited to relatively short trajectories, precluding simulations in biologically relevant time regimes, which can often reach beyond microseconds to the millisecond range.[12–16] This computational problem has long been recognized and has led to the development of implicit representations of the solvent, in which its effects are included in the potential function.[17–26] In general, an implicit solvent model (ISM) must (i) capture the main physical properties of the solvent (including electrostatic and thermodynamic effects), (ii) be of general applicability and reliable in predicting macromolecular properties, (iii) have the potential for further refinements based on physical considerations, and (iv) be computationally fast. A fundamental component of this complete ISM is the description of the electrostatics (i.e., the model must properly account for the effect on the embedded macromolecule of the reorientation and polarization of the solvent molecules in the medium). These electrostatic effects are manifested by (i) the damping of the Coulomb interactions between the charges in the macromolecule and (ii) the interactions of the solute with the solvent, termed the self-energy contribution. Because of the difficulty in describing these effects, earlier attempts to develop an ISM attained only partial success, primarily because the models proposed were incomplete (see, e.g., Refs. 17 and 27–33).

The availability of a computationally fast and accurate ISM is critical for conducting simulations of sufficient length to answer biological questions in the molecular regime that allows mechanistic insight to be developed. An ISM is also necessary in studies using MC simulations where they are essential for proper sampling and convergence, and in other areas of computational biophysics where the use of explicit solvent makes the calculation too difficult or beyond the available computational power. These include prediction of three-dimensional (3D) structures of proteins,[34–39] protein-folding studies,[36,37,40–46] molecular recognition studies and ligand-docking techniques,[47–55] pKa calculations in proteins with many simultaneously ionizable sites,[56–59] and calculation of binding free energies.[60–62] ISM have also been used to incorporate the effects of the solvent in quantum mechanical calculations.[22,29,63,64] Another increasingly important area of application of ISM is the design of drugs where huge libraries of molecules have to be screened to select suitable lead compounds as possible candidates.[65]

In the last few years, several new methods have been proposed, some of which may be able to break the computational bottleneck mentioned above. First, are the various approaches for numerical solution of the Poisson-Boltzmann (PB) equation that have been used as ISM.[27,66–69] Although the PB equation approaches are probably still too slow, careful programming and massive parallelization can overcome this to some extent, but these computational improvements can equally well apply to other approaches that are inherently simpler and computationally faster. One such approach is the generalized Born (GB) method, which is based on the original idea formulated by Hoijtink.[70] Several variations of this approach have been proposed,[23,24,71,72] and applications of this method to MD simulations have been reported.[73,74] They appear to be substantially more efficient than solving the PB equations. An alternative approach that was recently advanced, called SCP-ISM, is based on screened Coulomb potentials (SCP).[75,76] This approach is formally derived from the classic microscopic theory of polar liquids, based on earlier developments by Lorentz, Debye, and Sack,[77–81] and leads to dielectric functions of sigmoidal form that incorporate reaction field effects.[82–84] The present study describes the extension of the SCP-ISM to MD simulations.

In the SCP-ISM approach, the system is described as a solute immersed in a continuum that permeates all of space and is completely characterized by a screening function $D(r)$.[75,76] This description differs from most methods that assume the transferability of the macroscopic formulation of electrostatics to microscopic systems. Instead, the SCP-ISM is derived from microscopic considerations and proper Boltzmann averaging to arrive at the continuum description.[76] This procedure eliminates the need to define an internal dielectric constant and a (discontinuous) boundary between the protein and the solvent, both of which are required in methods such as the GB approaches. This development is important, because it is well documented[85–90] that neither the internal dielectric constant nor the precise location of the boundary is well defined, so these quantities reduce essentially to arbitrary parameters. The use of a homogeneous internal dielectric constant for the protein interior remains an extended practice, despite the recognition that the protein interior is, in general, heterogeneous and anisotropic.[91–93] Recently, for example, fluorescence experiments were conducted to probe the electrostatic properties of the local environment around residues in the B1 domain of protein G, which showed that the protein interior is polar and heterogeneous.[93] Moreover, the definition of a boundary between the solute and the solvent is particularly problematic in MD simulations using the PB implicit model, where the so-called "dielectric pressure" must accurately be accounted for in the calculation of forces[94–96] to avoid unphysical results. The sensitivity of the results to the location of the interface is not only related to the particular methodology used for the calculation of elec-

trostatic energy but has a more fundamental root that is related to the slow rate of increase of the dielectric function with the distance, as shown from basic physical principles.[75,77–80,82–84,97] The bulk dielectric value of the solvent is reached only at distances > 5–7 Å,[75] so the boundary usually falls in a very critical region (vdW plus probe radius ~ 4 Å) where the dielectric screening is still increasing.

The SCP-ISM has been validated in a number of tests, all using the same functional form of the potential energy and the same parameterization. Relying on a single functional form and parameterization is essential for being able to apply the method to a variety of unknown systems. Indeed, the validation calculations conducted to assess the performance of the SCP-ISM involved different size scales, that is, individual amino acids, small peptides, and whole proteins, as well as comparisons with other theoretical approaches (e.g., finite difference PB and explicit water calculations) and with experimental results.[75,76,98–100] For example, for the alanine dipeptide, the energy surface provided by the SCP-ISM was shown to be close to the energy surface obtained with an explicit treatment of water molecules using the same force field.[76] Calculations of solvation energies of small peptides (up to nine residues long) with different net charges were shown to correlate well with the energies obtained from numerical solution of the PB equation.[75] The model was also successful in discriminating the native conformation of proteins from a large number of misfolded conformations (decoys) for medium (~60 amino acids) and large (~300 amino acids) proteins.[76] In addition, the SCP-ISM was used in the calculation of secondary structure of α-helix and β-hairpin forming peptides and yielded results in good agreement with the experimental data.[99] Moreover, the same model served in the calculations of several solvent-exposed loop-containing segments in the G-protein transducin using MC methods, and the results compared well to the crystal structures.[100] Modulation of the electrostatics by the local environment in the protein interior has been tested in the calculation of pKa of titratable residues.[56,57,92]

In most of the applications summarized above, the SCP-ISM was used in conjunction with MC simulations. In this article, the model is extended to MD simulations. As with the earlier MC implementation, the SCP-ISM MD algorithm has been programmed for use in combination with CHARMM[1] and the all-atom PAR22 force field.[101] In the next section, after a very brief overview of the SCP-ISM approach, a parameter-free method is proposed for evaluating the Born radii needed to calculate self-energy terms. This method is based on a contact model[102] for estimating the degree of solvent exposure and allows the electrostatic potential to be written entirely in terms of pair functions. The equations for calculating the forces with this force field are then given. A validation run of the model for long MD simulations is reported for two systems, a peptide of 17 amino acid residues and a small protein of 56 residues. The choice of these two systems was based on the need to evaluate the performance of the model at two different size scales because of the known requirement, in other approaches, to use different parameters to describe the energetics of peptides and proteins (e.g., Refs. 103 and 104). The results are compared with corresponding explicit solvent simulations and with experimental data.

MATERIALS AND METHODS

Molecular Dynamics With the SCP-ISM: Born Radii and the Calculation of Forces

A complete derivation the SCP-ISM was reported previously;[75,76] only a brief description is given here for completeness. The electrostatic component, E_T, of the energy of a macromolecule composed of N atoms with point charges q_i, in a polar solvent, is given in the SCP-ISM by

$$E_T = \frac{1}{2}\sum_{i\neq j}^{N}\frac{q_i q_j}{D_s(r_{ij})r_{ij}} + \frac{1}{2}\sum_{i=1}^{N}\frac{q_i^2}{R_{i,Bs}}\left[\frac{1}{D_s(R_{i,Bs})} - 1\right] \quad (1)$$

where the first term corresponds to the interaction energies and the second term to the self-energies. The function $D_s(r)$ is a nonlinear, distance-dependent screening function that accounts for all the screening mechanisms in the systems,[75,76] and $R_{i,Bs}$ is the effective Born radius of atom i in the solvated macromolecule.[75,76] The functional form of $D_s(r)$ used in the SCP-ISM is sigmoidal, as obtained from theoretical studies (e.g., Refs. 82 and 97) and supported by experiments.[105–111] Note that although Eq. 1 resembles superficially other continuum electrostatics approaches (e.g., GB model in any of its several variants), it also contains fundamental differences: (i) there are no internal and external dielectric constants defined in the system; (ii) it is based on screening functions instead of dielectric functions; (iii) there is no boundary between the solvent and the solute; (iv) Born radii appear only in the self-energy terms; and (v) the distance dependence of the dielectric properties appears both in the interaction and self-energy terms.

Born Radii of an Atom in the Macromolecular Environment

The effective Born radii, $R_{i,Bs}$, which appear in the self-energy terms of Eq. 1 are defined as linear combinations of the Born radii of the particle in bulk solvent, $R_{i,w}$, and in bulk protein interior, $R_{i,p}$, weighted by the degree of exposure of the atom to each medium,[75], that is,

$$R_{i,Bs} = R_{i,w}\xi_i + R_{i,p}(1 - \xi_i) \quad (2)$$

where ξ_i is the fraction of atom i exposed to the solvent. In the original implementation of the SCP-ISM for MC simulations,[75,76] this fraction was defined as $\xi_i = SASA_i/4\pi R_i^2$ where $SASA_i$ is the solvent-accessible surface area of the atom and $R_i = R_{i,vdW} + R_{pb}$, where $R_{i,vdW}$ is the van der Waals radius of the atom and R_{pb} is the probe radius.

The calculation of the "bulk" Born radii $R_{i,w}$ and $R_{i,p}$ is based on the physical idea of a cavity formed by the atom in the dielectric medium and its boundaries. A detailed account of the manner in which these quantities are calculated and a discussion of their physical origin is given in Ref. 75. Briefly, these quantities were defined in terms

of the covalent radii, $R_{i,COV}$, and an extension factor, that is, $R_{i,w} = R_{i,COV} + h_i$ and $R_{i,p} = R_{i,COV} + g_i$, where h_i and g_i are positive numbers that quantify the extension of the covalent cavity in each medium. The values of h_i and g_i depend, in general, on the volume and chemical nature of each atom (including the effect of charge). h_i is assigned the value 0.35 Å if the partial charge is negative and 0.85 Å if it is positive (different values arise because of the asymmetric location of the center of the permanent dipole moment in the water molecule)[112]; the value of g_i was estimated from an extensive search of packing in the interior of proteins and was roughly set to $g_i = h_i + 0.5$ Å. Refinements to these values can be made on the basis of the particular atom and the local environment defined by the protein interior, as described.[56]

For MD simulations, the calculation of forces involves the calculation of gradients of the Born radii of Eq. 2. Although relatively fast algorithms exist for approximate numerical calculation of derivatives of $SASA_i$,[113,114] they are still too slow for large-scale MD simulations of macromolecular systems. Therefore, an alternative approach is desirable for the calculation of the Born radii that allows for fast evaluation of forces. At the same time, the reasonableness of the *SASA*-based approach (cf. Eq. 2) for calculating Born radii has already been shown and deemed appropriate within the context of the SCP-ISM approximation. In particular, it was found that the solvation free energy and its components (i.e., the interaction and self-energy terms) were well correlated with FDPB results.[75] Consequently, it is advantageous for any new approach to estimate the Born radii within the SCP-ISM, based on the idea of solvent exposure (cf. Eq. 2), to preserve the quality of the self-energies given by the *SASA*-based approach.

Therefore, an analytical expression for the Born radii $R_{i,Bs}$ was developed that correlates with the expression given in Eq. 2. Such an expression for the Born radii is proposed here based on a contact model (CM) reported earlier.[102] In this case, the fraction exposed to the solvent ξ_i of Eq. 2 is replaced by the pairwise function

$$\xi_i = \left[A_i - B_i \sum_{j \neq i}^{N} \exp(-C_i r_{ij}) \right] / 4\pi R_i^2 \tag{3}$$

so that the Born radii $R_{i,Bs}$ of Eq. 2 become

$$R_{i,Bs} = \alpha_i + \beta_i \sum_{j \neq i}^{N} \exp(-C_i r_{ij}) \tag{4}$$

where $\alpha_i = R_{i,p} - (R_{i,p} - R_{i,w})\, A_i/4\pi R_i^2$ and $\beta_i = (R_{i,p} - R_{i,w})\, B_i/4\pi R_i^2$. The positive constants A_i, B_i, and C_i are determined by maximizing the correlation between $R_{i,Bs}$ calculated with Eq. 2 and with Eq. 4 in a set of representative proteins. Note that (i) B_i and C_i quantify the degree of packing of the atoms j around the atom i (i.e., they are related to the buried fraction inside the protein) and (ii) the exponential form of Eq. 3 reflects the fact that the contribution of the protein local environment to blocking access of water to an atom decreases rapidly beyond the nearest neighbor shell.

A technique was reported in a previous article for the incorporation of hydrogen-bonding (HB) interactions in the context of the SCP-ISM. In this approach, the system formed by the macromolecule immersed in the solvent is treated as a three-component dielectric medium composed of the solvent, the macroscopic protein medium, and the region surrounding the proton acceptors (PA). For polar hydrogen atoms (PH), the expression given by Eq. 2 for the Born radii is modified by introducing explicitly the environment created by the PA around a PH, that is,

$$R_{i,Bs} = R_{i,w}\xi_i + R_{i,A}\zeta_i + R_{i,p}(1 - \xi_i - \zeta_i) \tag{5}$$

where $R_{i,A}$ is the Born radius of the atom i in an ideal bulk PA medium, and ζ_i is the fraction of the atom buried inside the PA. Although the fraction ζ_i could be calculated numerically, in the implementation for MC calculations reported earlier this quantity was calculated by using an analytic approximation. The surface S_{ij} of a sphere of radius ρ_i buried inside a sphere of radius ρ_j, when the distance between their centers is $r_{ij} < \rho_i + \rho_j$, is given by $S_{ij} = \pi\rho_i[\rho_j^2 - (\rho_i - r_{ij})]/r_{ij}$, and then the fraction buried is given by $\zeta_{ij} = S_{ij}/4\pi\rho_{ij}^2$. In the general case of a PH immersed in more than one PA, Eq. 5 is approximated by

$$R_{i,Bs} \approx R_{i,w}\xi_i + \sum_{i=1}^{m} R_{ij}\zeta_{ij} + R_{i,p}\left(1 - \xi_i - \sum_{j=1}^{m} \zeta_{ij}\right) \tag{6}$$

where m is the number of PA that overlap the atom i (PH), and R_{ij} is the Born radius R_A corresponding to the particle i immersed in the environment created by the acceptor j. Note that with this approximation the term multiplying $R_{i,p}$ might become negative if the buried surface is shared by two or more PAs. In practice, when this is the case, the term is set to zero (because of the directionality of the HB geometry included in the model for MC calculations, this overlapping of buried fraction is infrequent). For the MD implementation conducted here, incorporating Eq. 3 into Eq. 6 yields

$$R_{i,Bs} = \alpha_i + \beta_i \sum_{j(\neq i)}^{N} \exp(-C_i r_{ij}) + \sum_{j=1}^{m} \gamma_{ij} \frac{[R_j^2 - (R_i - r_{ij})^2]}{4R_i r_{ij}} \tag{7}$$

where $\gamma_{ij} = R_{ij} - R_{i,p}$. The values of R_{ij} are adjusted to obtain the proper HB stabilization energies as described earlier.[98] In the implementation reported here, the parameters *A, B,* and *C* contained in Eqs. 4 and 7, which optimize the Born radii, were based on the five standard atom types, H, C, N, O, and S (see Table I) appearing in proteins. It should be noted that the optimal values of *A, B,* and *C* in proteins may not be optimal for other macromolecules (e.g., nucleic acids).

The most stringent test of the validity of the method is to base the optimization on a single protein and then test the transferability of the parameters to other proteins. Here the X-ray structure of the amino-terminal fragment of

TABLE I. Atom-Type Dependent Parameters That Determine the Fraction Exposed to the Solvent Defined in Eq. 3

	C	O	N	S	Non-PH	PH
A	31.0	40.0	60.0	60.0	17.0	1.20
B	6.60	7.10	6.90	6.50	9.60	3.68
C	0.80	0.74	0.62	0.62	1.10	1.82

A, B given in $Å^2$; C in $Å^{-1}$.

phage 434 repressor (PDB entry 1r69), determined at 2 Å resolution, was arbitrarily chosen for evaluating the parameters. The protein contains 63 residues and about 1000 atoms (including H). For each atom in the test protein, the Born radius and corresponding self-energy are calculated with the original SASA approach[75] and with the CM approach using Eq. 4 or Eq. 7. The results are presented as scatter plots for the Born radii [Fig. 1(a)] and the self-energies [Fig. 1(b)] of each atom of the protein used for optimization (correlation $r > 0.99$ in both cases). To show the transferability to other proteins, the values of A_i, B_i, and C_i optimized by using the single protein above were then used to calculate the Born radii and self-energies in 40 different proteins comprising mainly α, β, α + β, and α/β secondary structural motifs. This set of proteins was used earlier in the optimization of $R_{i,p}$ (cf. Eq. 2)[75] and consists of proteins ranging in size from 43 to 386 amino acids adding up to >55,000 atoms in total, including H atoms. All the structures were determined at high resolution (≥2 Å). Figure 2(a) displays the scatter plot of the Born radii, showing that the correlation between the two approaches is still good. Note that there is a small number of atoms (about 100 of 55,000) for which the difference with respect to the *SASA*-based approach is larger (~30–40%) but, even for these outliers (mostly heavy atoms), the self-energies are well correlated, as shown in Figure 2(b), indicating that the energetics are similar in the two approaches. The Born radii of most of the polar H atoms, including those forming H-bonds, show a correlation close to one, because these quantities are less sensitive to the values of the constants *A, B,* and *C* (PH atoms are usually buried inside the proton donor environment in the all-atom representation of the force field used here). In all these calculations, the crystal structures were completed by adding H atoms using the HBUILD option in CHARMM[1] and then minimized for 500 steps to eliminate bad contacts. Scatter plots of the self-energies in small peptides also showed a good correlation between the two approaches (data not shown).

Analytic Calculation of Forces in the SCP-ISM

With the analytical expressions for the Born radii $R_{i,Bs}$ given by Eqs. 4 and 7, the total electrostatic energy E_T of Eq. 1 is completely expressed as a pairwise potential for which gradients can be easily calculated. The force $\mathbf{F}_i$ on particle i is given by $\mathbf{F}_i = -\nabla_i V = -\nabla_i E_T - \nabla_i W$ where V is the total potential energy, E_T is the electrostatic component given by Eq. 1, and W is the rest of the force field; index i indicates that the derivatives are performed with respect to the coordinates $\mathbf{r}_i$ of particle i. Therefore, the electrostatic component of the force $\mathbf{F}_i$ is given by

$$\mathbf{F}_i^{\text{elec}} = -\nabla_i E_T = -\left(\frac{\partial E_T}{\partial x_i}, \frac{\partial E_T}{\partial y_i}, \frac{\partial E_T}{\partial z_i}\right) = -\sum_{j(\neq i)}^{N} \frac{1}{r_{ij}} \frac{\partial E_T}{\partial r_{ij}} \Delta \mathbf{r}_{ij} \tag{8}$$

where $\Delta \mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. The derivative of E_T in Eq. 8 can be divided into two parts involving the interaction and self-energy terms, that is,

$$\frac{\partial E_T}{\partial r_{ij}} = \frac{\partial E_T^{\text{int}}}{\partial r_{ij}} + \frac{\partial E_T^{\text{self}}}{\partial r_{ij}} \tag{9}$$

where the interaction component is given by (see Eq. 1)

$$\frac{\partial E_T^{\text{int}}}{\partial r_{ij}} = \frac{\partial}{\partial r_{ij}}\left[\frac{1}{2}\sum_{i\neq j}^{N} \frac{q_i q_j}{D_s(r_{ij}) r_{ij}}\right] = \frac{1}{2} q_i q_j \frac{\partial}{\partial r_{ij}} [D_s(r_{ij}) r_{ij}]^{-1}$$

$$= -\frac{1}{2} \frac{q_i q_j}{D_s(r_{ij}) r_{ij}} \left[\frac{1}{r_{ij}} + \frac{1}{D_s(r_{ij})} \frac{dD_s(r_{ij})}{dr_{ij}}\right] \tag{10}$$

The screening function in the SCP-ISM is given analytically by

$$D_s(r_{ij}) = \frac{1 + \varepsilon_s}{1 + k \exp(-\alpha_{ij} r_{ij})} - 1 \tag{11}$$

where ϵ_s is the dielectric constant of the bulk solvent, $k = (\epsilon_s\text{-}1)/2$ and α_{ij} is a parameter[75] that determines the rate of increase of the dielectric screening with the interparticle distance r_{ij}. The derivative of the screening function with respect to r_{ij} in Eq. 9 can be written in a compact form as

$$\frac{dD_s(r_{ij})}{dr_{ij}} = \alpha_{ij}(1 + \varepsilon_s)^{-1}[1 + D_s(r_{ij})][\varepsilon_s - D_s(r_{ij})] \tag{12}$$

Introducing Eq. 12 into Eq. 10 yields the magnitude of the force between particles i and j because of their mutual electrostatic interaction.

The second term of Eq. 9 is given by

$$\frac{\partial E_T^{\text{self}}}{\partial r_{ij}} = \frac{\partial}{\partial r_{ij}}\left\{\frac{1}{2}\sum_{i=1}^{N} \frac{q_i^2}{R_{i,Bs}}\left[\frac{1}{D_s(R_{i,Bs})} - 1\right]\right\}$$

$$= \frac{1}{2} q_i^2 \frac{\partial}{\partial r_{ij}}\left\{\frac{1}{R_{i,Bs}}\left[\frac{1}{D_s(R_{i,Bs})} - 1\right]\right\} + (\text{same with } i \rightarrow j) \tag{13}$$

where the Born radii are given by Eq. 7 if the atom is a H-bonded proton or by Eq. 4 otherwise. Carrying out the differentiation, Eq. 13 becomes

$$\frac{\partial E_T^{\text{self}}}{\partial r_{ij}} = \frac{1}{2} \frac{q_i^2}{R_{i,Bs}^2}\left[1 - \frac{1}{D_s(R_{i,Bs})} - \frac{R_{i,Bs}}{D_s^2(R_{i,Bs})} \frac{dD_s(R_{i,Bs})}{dR_{i,Bs}}\right] \frac{\partial R_{i,Bs}}{\partial r_{ij}}$$

$$+ (\text{same with } i \rightarrow j) \tag{14}$$

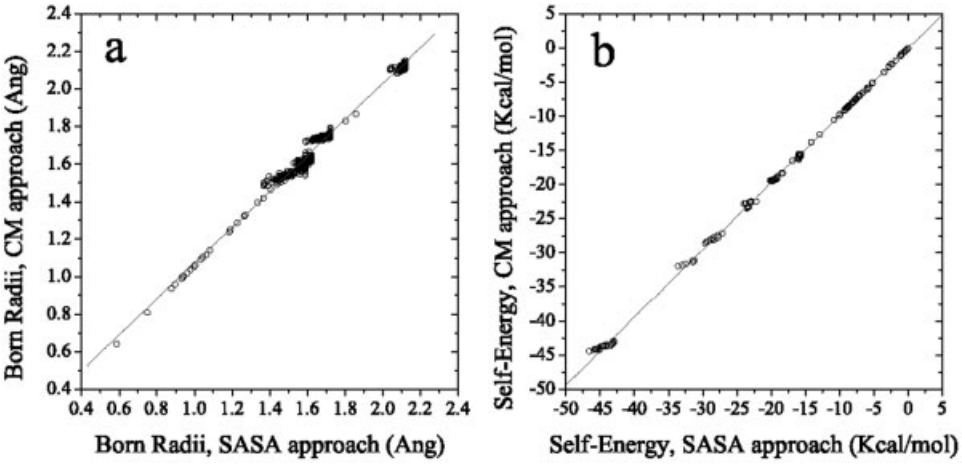

Fig. 1. Comparison of Born radii calculated with the solvent accessible surface area (SASA) approach[75] and the exponential contact model (CM) (Eqs. 3, 4, and 7): **a:** Scatter plot of the optimized Born radii calculated with the CM and with the SASA approach (correlation $r = 0.99$); **b:** Scatter plot of the self-energies obtained from the optimized Born radii plotted in (a) ($r = 1.00$). The protein used for the optimization of the Born radii (PDB entry 1r69) contains >1000 atoms, including hydrogens.

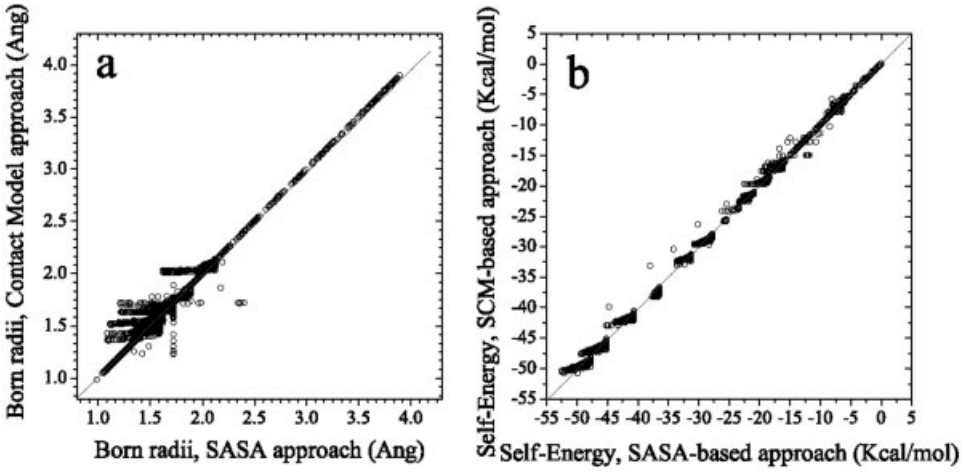

Fig. 2. Born radii and self-energies of atoms in 40 proteins (>55,000 atoms including H) obtained from the PDB, comprising the main structural motifs α, β, α/β, α + β. Protein 1r69 does not belong to the set. **a:** Scatter plot of the Born radii calculated with the parameters optimized on the single protein of Figure 1 (correlation $r = 0.99$); **b:** Self-energies of the atoms calculated with the Born radii plotted in (a) ($r = 1.00$).

where the derivative of D_s with respect to the Born radius is given by Eq. 12 but with r_{ij} replaced by $R_{i,Bs}$. The partial derivatives of $R_{i,Bs}$ in Eq. 14 are given by

$$\frac{\partial R_{i,Bs}}{\partial r_{ij}} = -\beta_i C_i \exp(-C_i r_{ij}) \tag{15}$$

if the Born radius of the atom is calculated from Eq. 4, or

$$\frac{\partial R_{i,Bs}}{\partial r_{ij}} = -\beta_i C_i \exp(-C_i r_{ij}) + \frac{\gamma_{ij}}{4R_i}\left(\frac{R_i^2 - R_j^2}{r_{ij}^2} - 1\right) \tag{16}$$

if it is given by Eq. 7. Note that at the point of contact between two spheres (i.e., $r_{ij} = R_i + R_j$), the last derivative is discontinuous, which creates an additional force when a polar hydrogen approaches an acceptor. Therefore, it is convenient to introduce a shift function $f(r_{ij})$ that multiplies the last terms of Eq. 7 and is defined by $f(r_{ij}) = [1 - r_{ij}^2/(R_i + R_j)]^2 2$ for $r_{ij} < R_i + R_j$ and $f(r_{ij}) = 0$ otherwise. Therefore, Eq. 16 is modified to

$$\frac{\partial R_{i,Bs}}{\partial r_{ij}} = -\beta_i C_i \exp(-C_i r_{ij}) + \frac{\gamma_{ij}}{4R_i}\left(\frac{R_i^2 - R_j^2}{r_{ij}^2} - 1\right) f(r_{ij}) + \gamma_{ij}\frac{[R_j^2 - (R_i - r_{ij})^2]}{4R_i r_{ij}}\frac{df(r_{ij})}{dr_{ij}} \tag{17}$$

where both f and f' are null at $r_{ij} \geq R_i + R_j$. Introducing Eq. 17 into Eq. 14 yields the magnitude of the force between particles i and j, due to their mutual influence in their self-energies.

MD Simulation of Peptides and Proteins With the SCP-ISM

The replacement of explicit solvent molecules by an implicit representation affects properties that are related

to steric and kinetic effects of the solvent, such as friction and pressure. The lack of directionality imposed by hydrogen-bonding interactions between the solvent and exposed side-chains is also expected to affect the dynamical behavior of the solute. Therefore, results obtained by using an ISM designed to represent *only* electrostatic effects (including entropic effects of the solvent[26,39]) must be analyzed with caution, and it is important to recognize which properties are expected to be reproduced with an ISM, and which are not. It is broadly assumed that a general ISM should describe reasonably well the electrostatic properties of both proteins and peptides, without requiring ad hoc modifications of the energy function or adjustments of the parameters. Improvements of the description should be based solely on physical considerations as, for example, incorporating the different dielectric properties inside a protein that are governed by the heterogeneous, local architectures,[56,57,92] or introducing a more sophisticated treatment of hydrophobic effects.[22,115–123]

To test the current implementation of the SCP-ISM for molecular dynamics simulation and to assess its potential to become a general and reliable approach for biological application, long simulations of the peptide Dynorphin (Dyn) and the globular protein, Protein G (ProtG), were conducted, and the results compared with simulations using an explicit solvent model and with available experimental data.

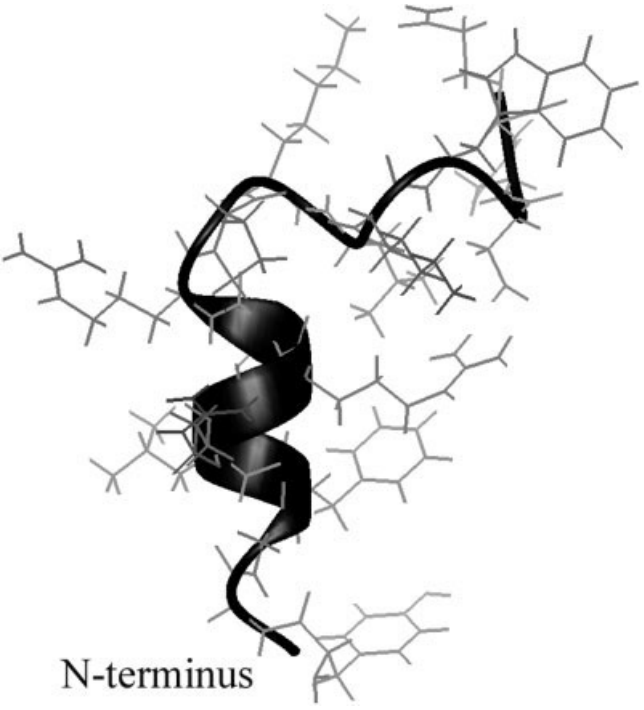

Fig. 3. Initial structure of Dynorphin used in the MD simulation with the implicit solvent model. This structure is based on data obtained from NMR studies in micelles and used earlier as the initial structure in a long MD simulation with an explicit solvent model (see Ref. 128).

MD Simulation of Dynorphin in Aqueous Solution

Dynorphin (H-YGGFLRRIRPKLKWDNQ-OH) is a 17-amino acids peptide hormone that binds with varying degrees of specificity to subtypes of opioid receptors.[124–129] Experimental and theoretical studies of the peptide in aqueous solvent have been reported, providing evidence that the peptide adopts an α-helical conformation between residues Gly3 and Arg9.[128,129] Starting from the NMR structure, a recent multinanosecond MD simulation showed that the peptide forms a stable α-helical conformation in the membrane environment, but in an aqueous solvent the helical structure is disrupted within the first nanosecond of simulation. These results are consistent with earlier data from FTIR and NMR spectroscopy on a shorter (biologically active) peptide obtained from naturally occurring Dyn, which suggested that the peptide adopts mainly an extended conformation in water.[130] Because of the detailed MD simulation reported recently[129] and the experimental results pertaining to both the structure and the dynamics of this peptide, it has been used here as a first test of the SCP-ISM MD implementation reported in the previous section. A 3-ns simulation was conducted following the same protocol reported recently in the 3-ns MD simulation using explicit waters.[129] The initial structure of the peptide (see Fig. 3) was taken from Ref. 129, where it was built from the internal parameters obtained from the NMR study[131]: the segment 1–9 was taken to be in an α-helical structure, whereas the rest of the peptide was in a random conformation. The peptide was simulated by using the SCP-ISM incorporated into CHARMM,[1] with the all-atom PAR22 force field.[101] The simulation was conducted in the microcanonical (NVE) ensemble at T = 300 K using the SHAKE algorithm and a cutoff of 13 Å for the nonbonded interactions. During the initial 50 ps of the equilibration phase, the C_α atoms of the peptide were restrained with a harmonic force that was gradually removed during 1.4 ns of simulation by decreasing the force constant in four steps as reported earlier.[129] Subsequently, a production run of 3 ns was conducted, and coordinates were saved every 10 ps for analysis. Throughout the simulation, the temperature remained close to 300 K, despite the conformational changes, with fluctuations on the order of 11 K.

The loss of secondary structure observed in the explicit water simulation of Dyn[129] is reproduced in the SCP-ISM run (i.e., the α-helical structure of the peptide collapses early in the simulation). To illustrate the character of the overall conformational changes that occur during the 3-ns trajectory, Figure 4 shows two snapshots, at 100 ps and 3 ns, superimposed on the initial NMR structure. Early in the simulation the helical segment from residues 3 to 9 opens slightly, with the 1–4 H-bonding pattern characteristic of the α-helix first transformed into a 1–5 H-bond pattern. Subsequently, the helical motif is completely disrupted and remains in a random conformation until the end of the simulation. In the explicit water (EW) simulation, the same process occurs much more slowly than in the ISM simulation. For the latter, the opening of the α-helix begins even before the constraints of the C_α atoms are fully released. This difference in time regimes has also been observed in studies using other continuum models[74] and is believed to be related to the lack of friction between

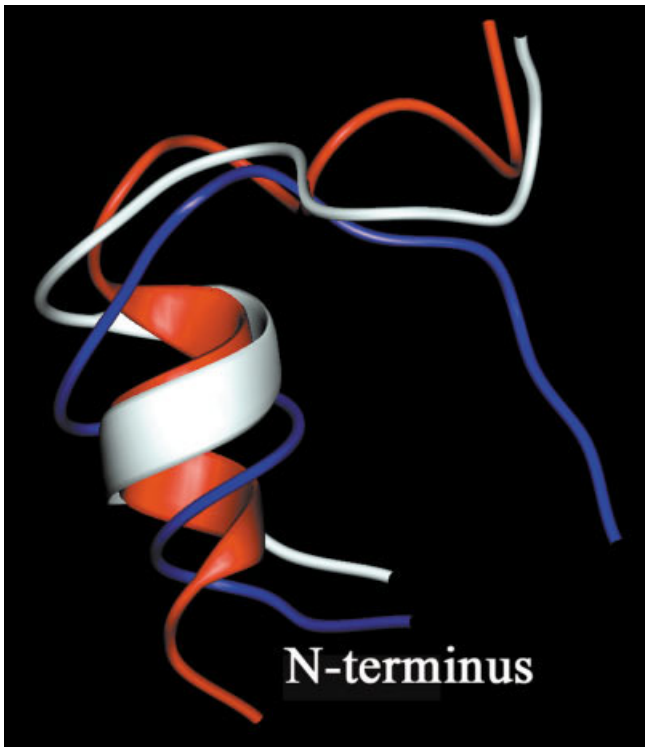

Fig. 4. Two snapshots at $t = 100$ ps (white) and at $t = 3$ ns (blue) of the Dyn peptide in the simulation using the implicit solvent; the initial NMR-based structure is shown in red. Early in the simulation, the 1–4 H-bond pattern characteristic of the α-helix (red) is transformed into a 1–5 H-bond pattern (white). Later on in the simulation, the helical structure collapses and the peptide remains open in a nonhelical conformation (blue) until the end of the simulation.

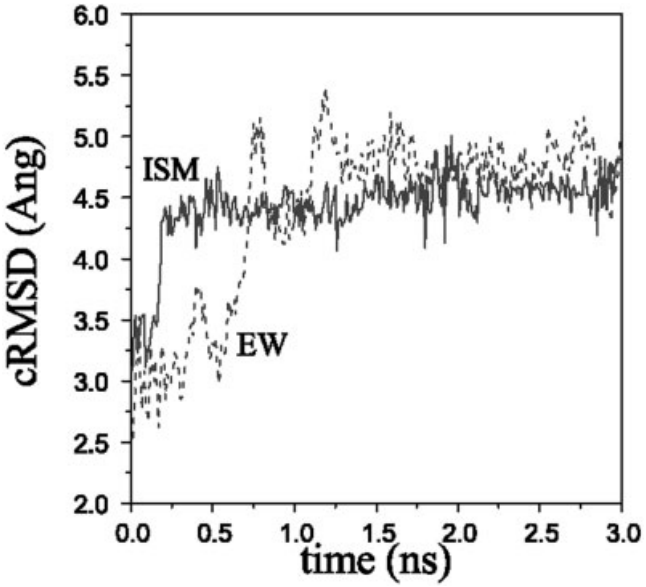

Fig. 5. C_α root-medium-square deviation (cRMSD) for Dynorphin as a function of time for the two simulations. The conformational changes occur faster in the implicit solvent simulation, where a stable structure is reached at $t \sim 0.3$ ns; in the explicit solvent simulation, the stable conformation of the peptide is reached at $t \sim 1$ ns. The two trajectories exhibit similar fluctuations after $t = 1.5$ ns.

solute and solvent and to the instantaneous response of the implicit solvent.

The time evolution of the EW and ISM RMSD of C_α atoms (cRMSD) with respect to the initial structure is displayed in Figure 5. The trajectories clearly show the much quicker loss of secondary structure in the ISM simulation. However, after the helix is disrupted in the explicit simulation (around 1 ns) and both structures have equilibrated (around 1.5 ns), the cRMSD and their fluctuations are quite similar. Fluctuations of the side-chains are larger in the ISM than in the EW simulation, which presumably has the same origin (discussed above) as the differences in time regimes shown in Figure 5.

Average, energy minimized, structures were calculated from the last 1.2 ns of the trajectory from each simulation and they are shown in Figure 6 along with the initial NMR structure. It is apparent that despite the differences in behavior during the early part of the trajectory (see Fig. 5), the average structures are quite similar. The segment comprising residues 3–15 displays a similar folding having a cRMSD of 2.2Å, with the segment 3–9 adopting an open conformation with no characteristic helical H-bonding pattern in both simulations. The main qualitative difference between the two simulations is observed in the N-terminus. In the SCP-ISM simulation, the side-chain of Y1 forms an internal H-bond that forces the terminus to

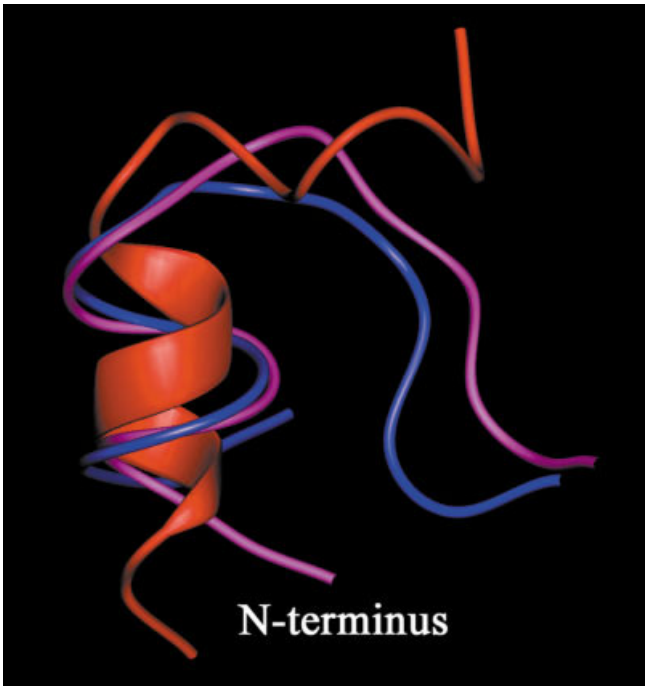

Fig. 6. Superposition of the average structures obtained from the last 1.2 ns of simulations obtained by using the implicit (magenta) and explicit (blue) solvents; the initial NMR-based structure is shown in red. The behavior of the peptide in the SCP-ISM simulation is qualitatively similar to the behavior observed in the explicit solvent simulation. The most important structural difference occurs in the N-terminus. In the EW, it appears fully solvated, forming an H-bond to water molecules, whereas in the ISM, it bends inward, forming an intramolecular H-bond.

bend inward. In the EW simulation, this terminus is completely exposed to the solvent, and the side-chain of Y1 is H-bonded to water molecules.

Long MD Simulation of Dynorphin in Water

Because the simulation with implicit solvent reproduced the main qualitative features of the peptide in explicit water, it is of interest to gain further insight into the behavior of this biologically active peptide by continuing the simulation to 40 ns. The results were analyzed, focusing mainly on the structural changes along the trajectory. For the entire 40-ns trajectory, the average temperature remained stable at 300 K with fluctuations of the order of 11 K. The segment 1–9 of the peptide remained unstructured for about 12 ns and then commenced a series of conformational changes consisting of frequent transitions between random coil (qualitatively similar to the structure shown in Fig. 6) and α-helical conformations. The transient helical structures in this part of the simulation were always observed within the segment comprising residues 3–9, whereas the segment 10–17 did not adopt a defined secondary structure and remained in a random conformation during the entire simulation. The N-terminus also undergoes conformational changes in which Y1, which remained H-bonded for most of the first 3 ns, frequently changes its orientation, being solvent exposed or H-bonded to different residues.

To quantify the conformational changes in the structure, especially the transitions between random coil and α-helix, the number of contiguous residues with backbone torsional angles in the interval $(\phi,\psi) = (-57° \pm 30°, -47° \pm 30°)$ is plotted as a function of time in Figure 7. It is assumed that three or more contiguous residues (3.6 residues on average) whose backbone angles lie in this interval define an α-helical turn. Figure 7 shows that, except for a fluctuation lasting <10 ps around 2.2 ns, where the peptide transiently adopts an α-helical motif, in the period up to 12 ns the peptide remains open with very few transitions to a short helical structure (three residues with the backbone torsional angles in the right intervals). After 12 ns, however, frequent, short-lived transitions are observed from an open, random coil conformation to an α-helical structure.

To illustrate the kind of structural changes that occur in this part of the trajectory, three snapshots of the peptide at t = 21.2, 22.1, and 22.4 ns are shown along with the initial structure in Figure 8. Note that the peptide seems to lose all structural similarity to a helix (at 22.1 ns) and presumably loses all memory of the previous helical conformation, but the helical motif is rapidly reformed (in this case, within 300 ps). The different helical structures that appeared during the simulation did not necessarily involve the same residues within the segment 3–9 but shifted their positions within this segment. These results indicate a propensity for the peptide to adopt an α-helical conformation in water, but apparently the small stabilization free energy establishes an equilibrium between the helical and random coil conformations that favors the latter. Apparently, starting from the NMR structure and using the PAR22 force field, it requires about 12 ns of simulation before this equilibrium is established. Moreover, for the reasons pointed out above, it can be expected that this type of equilibrium would take even longer to be reached in an explicit solvent simulation.

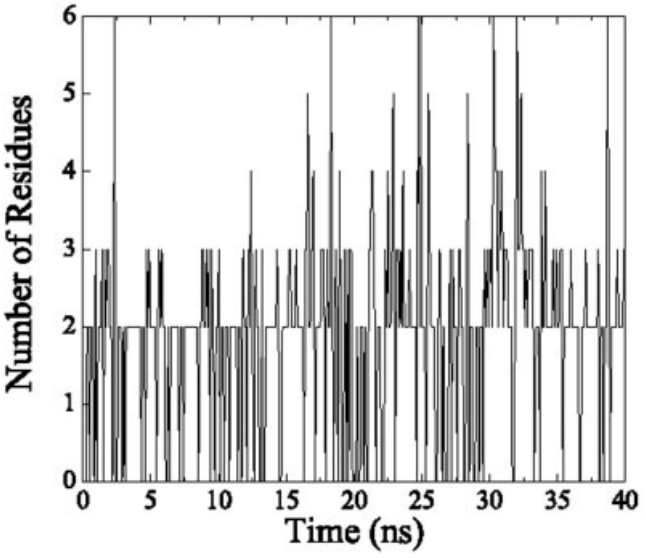

Fig. 7. Maximum number of contiguous residues whose backbone dihedral angles are in the range corresponding to an α-helix, that is $(\phi,\psi) = (-57° \pm 30°, -47° \pm 30°)$. An average of 3.6 contiguous residues in this range defines an α-helix turn. Note that after $t \sim 15$ ns, the peptide enters a cycle of fast interconversions between random-coil and α-helix conformations. At any point of the trajectory, a helical motif is found only in the segment comprising residues Y1–P10, whereas the segment between residues K11 and Q17 remains random during the entire simulation time.

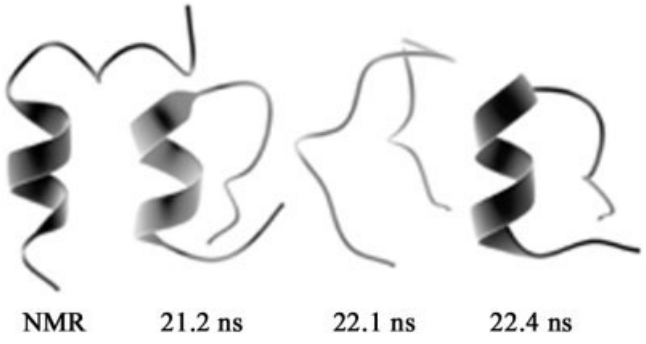

Fig. 8. Three snapshots of the peptide at different times along the simulation, illustrating the kind of conformational changes occurring in the dynamics. The initial NMR-based structure is also shown. In this 1.2-ns sequence, the peptide passes from an α-helical structure to a random conformation and returns rapidly to another α-helical structure.

Simulation of Protein G in Aqueous Solution

The B1 immunoglobulin-binding domain from streptococcal protein G (ProtG) is a 56-amino acid polypeptide. The three-dimensional structure of this protein has been determined by X-ray crystallography[132] and NMR spectroscopy.[133] The structure contains a packed hydrophobic core between a four-stranded β-sheet and a four-turn α-helix

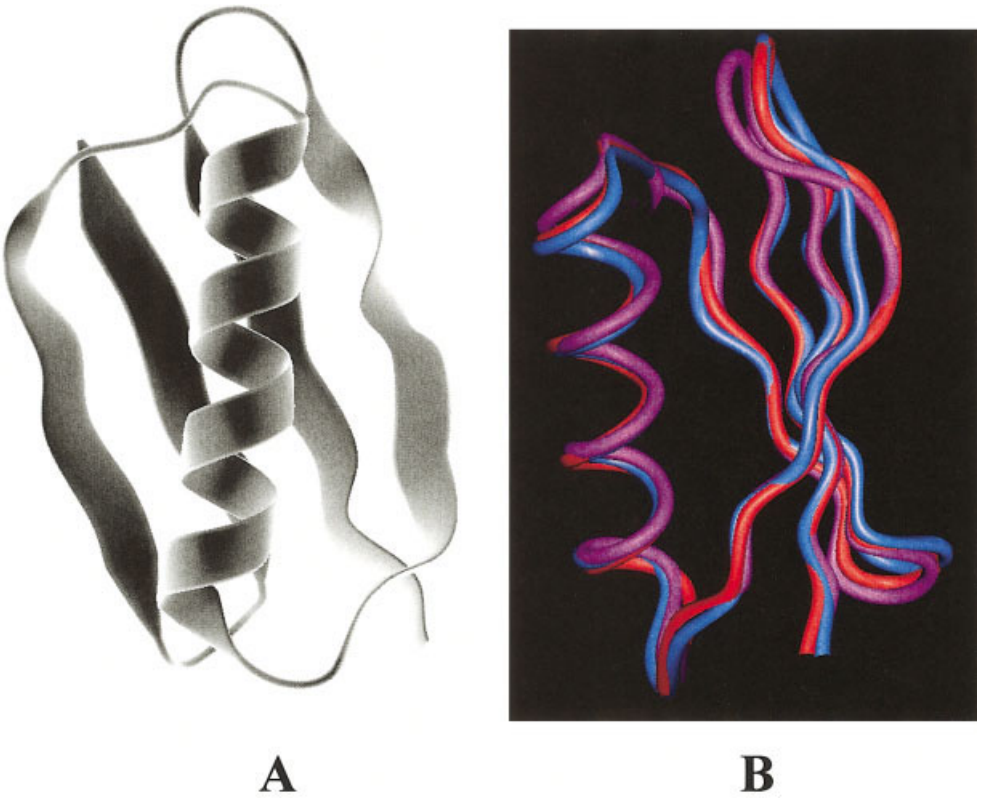

Fig. 9. **A:** Ribbon representation of the B1 domain of protein G used in the simulations (PDB entry 1pgb). The 56-amino acid protein is composed of an α-helix that faces a four-strand β-sheet, forming a tight hydrophobic core. Protein G does not contain proline residues and is not stabilized by disulfide bridges; **B:** Superposition of the average structures on the crystal structure (red: crystal structure; blue: ISM; magenta: EW).

[see Fig. 9(A)]. The structures in the crystal and in solution differ primarily in the loop involving residues 46–51, and other structural variations are observed in the helix.[132]

Because ProtG has no disulfide bridges, no proline residues and most of the residues participate in defined secondary structure, the B1 domain of protein G has been the focus of many experimental and theoretical studies.[73,134–144] The large amount of available data on this protein makes it a suitable system for the initial study intended here. Explicit water MD simulations have been reported starting from the NMR or X-ray structures and using the united atom representation.[143] In this section, we describe the results of two 3 ns simulations, one with EW and the other using the SCP-ISM. The simulations start from the crystallographic structure (PDB entry 1pgb) and the all-atom representation. The study and analysis reported here extends that on Dyn, reported above, to provide an initial assessment of the suitability of the approach to simulate the dynamics of small proteins as well as peptides.

To construct the starting structure for both implicit and explicit water simulations, all the crystallographic water molecules were removed from the crystal structure (1pgb). For the EW simulation, the water drop model with constraining shell was used as described previously.[145] A sphere of TIP3 water molecules with a radius of 32 Å was built around the protein. Solvent molecules that overlapped protein atoms were removed, leaving a total of ~5800 water molecules. The thickness of the water shell around the protein was at least 15 Å. The whole system (water and protein) was minimized, heated, and equilibrated at T = 298 K for 100 ps.

Both the implicit and explicit solvent 3-ns MD simulations were performed in the microcanonical ensemble. Although the simulations were conducted at constant energy, the average temperatures remained close to 298 K throughout the entire production phase of both simulations. The fluctuations were about 7 K in the SCP-ISM simulation and about 2 K in the EW simulation. The SHAKE algorithm and a cutoff of 13 Å for the nonbonded interactions were used. The ionization states of the titratable amino acid residues corresponded to those at pH 7, assuming unshifted pKa values.

Figure 10 shows the temporal evolution of the cRMSD for both simulations, with respect to the minimized starting structure for each solvent model. The average structures, used for the subsequent analysis, were calculated from the last 2.2 and 1.2 ns of the ISM and EW trajectories, respectively, eliminating rigid body displacements and minimized to eliminate bad contacts. Figure 10 shows that the trajectory of the SCP-ISM follows closely the trajectory of the explicit solvent simulation. The SCP-ISM trajectory

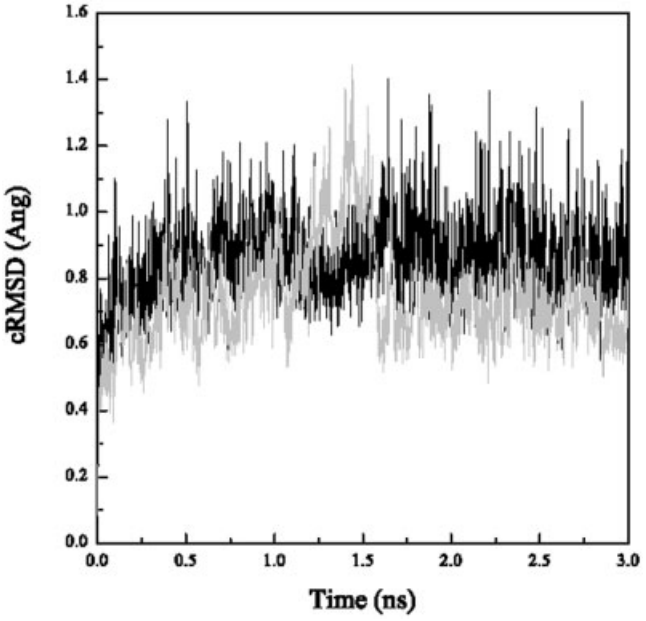

Fig. 10. cRMSD of protein G from the implicit (black) and explicit (gray) solvent model MD simulations, with respect to the corresponding minimized starting structures obtained from each model.

appears to be stable after about 0.6 ns, whereas the EW trajectory is stabilized after 1.6 ns of simulation. This trend of the ISM simulation to stabilize in a shorter time was also seen in the Dyn simulation described above. Figure 9(B) gives the superposition of the average structures from both simulations on the crystal structure, showing that both the ISM and EW simulations remain close to the crystal structure preserving all the elements of secondary structure throughout the entire simulation time. The cRMSD of the average structures with respect to the crystal structure and between each other are shown in Table II, indicating the overall similarity. The all-atom RMSD of the average structure from the ISM with respect to the crystal structure is larger than the corresponding RMSD of the EW average structure, which suggests that side-chains undergo larger conformational changes in the implicit solvent than in the explicit solvent. The most important conformational changes of the average structures compared with the crystal occur in the loop involving residues 8–13 in both cases, and to a lesser extent in the loop segment 46–50 [cf. Fig. 9(B)]. The cRMSD of the EW and ISM average structures with respect to the minimized NMR structure (PDB entry 2gb1) were found to be 1.15 Å and 1.48 Å, respectively. These values are larger than the corresponding cRMSD with respect to the crystal structure (Table II), in agreement with earlier calculations reported by other authors.[143]

Comparison of the results reported here with recent simulations on protein G[73,103,146] using the generalized Born (GB) approach[21] indicate reasonable agreement between the trajectories, although the trajectory shown in Figure 10 seems to be somewhat more stable, at least for the 3-ns run conducted so far. In addition, the overall Cα RMSD for both EW and ISM simulations obtained here are somewhat smaller than reported in the earlier studies.[73,103,146] In comparing the results reported here with these GB studies, it is important to note that the parameterization of the latter solvent model was conducted, in part at least, by using protein structural information and/or dynamic properties. Determination of the parameters for the SCP-ISM follows a different approach: As described in detail,[75] the parameters of the SCP-ISM were determined on the basis of solvation energies of amino acid analogs, and the only information used from proteins was that $D_s(r)$ was restrained to stay within a range that reflects the dielectric behavior of protein interiors (in contrast to pure solvent) as found from pKa calculations.[56] Subsequently, the same set of parameters have been used in all applications of the SCP-ISM reported to date. Therefore, the agreement (or discrepancy) found between the ISM and EW simulations reported here is due to the formulation of the model and not to any special protein structural or dynamic characteristics effectively incorporated into its parameters.

It is instructive to analyze the RMSD obtained from various groups of residues as given in Table II. Comparison of the RMSD of EW/X-ray and ISM/X-ray for the different groups indicates that the general trend of the side-chain conformational differences obtained in the SCP-ISM is not uniformly followed by all of the residues. Thus, for Asp, Glu, Ile, and Val, the structural differences are similar in both solvent models. It is also noted that the structural variability of the hydrophobic residues is smaller than for the polar or charged residues, for both simulation systems. This is also evident from the last two columns of Table II, where the RMSD of solvent-exposed and buried residues have been calculated. The RMSD of the buried residues in both solvent models are much smaller and closer together than the RMSD of the solvent-exposed residues. This result shows that, at least for ProtG and the 3-ns length of the simulation, the structural integrity of the interior of the protein is conserved in the ISM, despite the increased variability of the exposed residues.

Figure 11(A) displays the root-mean-square displacements, $\langle \Delta r^2 \rangle$, of the C_α atoms from both simulations, as a measure of the fluctuations in the elements of secondary structure, and Figure 11(B) shows the $\langle \Delta r^2 \rangle$ for the C_γ atoms, as representative of the side-chain movements. The approximate positions of each element of secondary structure are shown at the bottom of each panel. The $\langle \Delta r^2 \rangle$ obtained from the B-factors of the X-ray structure are also shown in Figure 11. Figure 11(A) indicates that the main fluctuations occur in the loop segments. Explicit and implicit simulations yield the same trend in the displacements, with the helical segment being somewhat more flexible in the SCP-ISM simulation. Note also that the qualitative behavior of the $\langle \Delta r^2 \rangle$ in both solvent models mimics qualitatively the temperature factors of the crystal structure.

Although the overall trend of the side-chain fluctuations is also similar in both simulations [Fig.11(B)], there are some important differences for a number of atoms. The largest differences occur in side-chains that are mostly

TABLE II. RMSD of Various Subsets of Atoms in the Crystal and Simulated Structures of ProtG

	All[a]	Cα[b]	Lys[c]	Asp Glu[c]	Asn Gln[c]	Leu[c]	Ile[c]	Val[c]	Thr[c]	Phe Tyr Trp[c]	Exp fr[d] >0.3	Exp f[e] <0.3
ew/pgb	1.65	0.81	1.51	1.94	1.75	1.32	0.56	1.32	0.87	0.82	1.93	1.27
ism/pgb	2.16	0.81	2.40	1.89	2.24	1.54	0.42	1.31	1.25	1.13	2.63	1.50
ew/ism	2.15	1.12	2.16	2.49	1.95	1.70	0.87	0.77	1.46	1.19	2.58	1.57

[a]RMSD of all heavy atoms.
[b]RMSD of Cα atoms.
[c]RMSD of all heavy atoms in residues of subset listed in column header. Note that the RMSD of the subsets from column 4 on are calculated for all heavy atoms based on the Cα chain superposition.
[d]All heavy atom RMSD of side-chains with average fractional solvent exposure >0.3.
[e]As d, but for all side-chains with fractional solvent exposure <0.3.

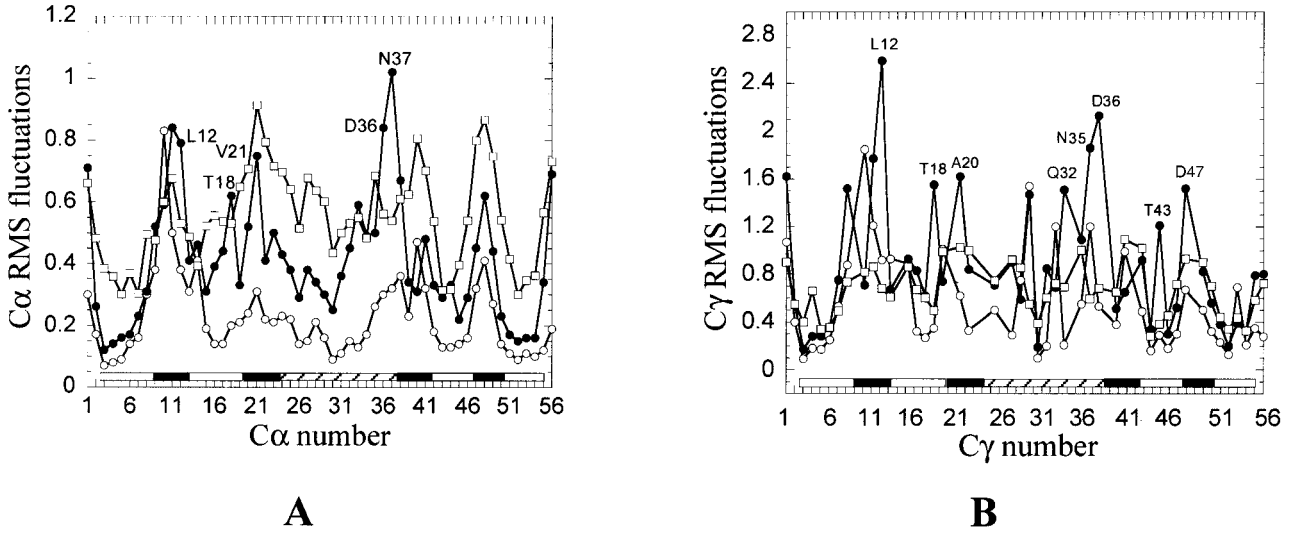

Fig. 11. **A:** cRMS fluctuations of protein G obtained in the simulations with implicit (●) and explicit (○) solvent models. The fluctuations assigned to the crystal structure (□) were obtained from the B-factors (see text). The approximate positions of the elements of secondary structure are displayed schematically at the bottom (white: β-strands; black: loops; crossed lines: α-helix). In both simulations and in the crystal, the larger fluctuations are observed in the loop segments. Although the overall trend of the fluctuations are similar, a few atoms (identified with residue name and number), mostly in the loop regions, show larger fluctuations in the SCP-ISM than in the EW simulation; **B:** Similar as in (a) but for C_γ atoms. As observed in the C_α atoms, the overall trend of the fluctuations in both the explicit and implicit solvent simulation is similar, except for the few atoms indicated. The fluctuations of C_γ atoms reflect the movements of side-chains. Most of the side-chains with large fluctuations are fully solvent exposed and/or belong to loop segments.

solvent exposed (indicated with residue name and number in the figure). This finding suggests that the side-chains of the protein undergo larger movements, as was found from the simulation of Dynorphin. It is likely that the lack of steric interactions with explicit solvent molecules and the intrinsic instantaneous response of the implicit solvent may be involved in the larger fluctuations of exposed side-chains. This finding is further confirmed by the results in the last two columns of Table II and by noting that fluctuations of longer, solvent-exposed side-chains (e.g., the CO_2^- group in 36-Asp) are much larger in the SCP-ISM than in the explicit solvent dynamics [Fig 11(B)]. The fluctuations of side-chains in the crystal are in general smaller than obtained in both simulations, probably because of the crystal packing and lower temperature of the crystal.

To further analyze local differences and similarities between the explicit and implicit average structures of protein G, it is instructive to compare the solvent exposed fractions of the side-chains. These are exhibited in Figure 12 for both average structures and for the crystal structure. The calculation was conducted with CHARMM,[1] and the plotted values represent the average obtained from the atomic *SASA* of the side-chains. Inspection of Figure12 shows that the side-chain exposed fractions (SCEF) of both the EW and ISM average structures closely follow the overall SCEF of the crystal structure. In particular, all deeply buried hydrophobic residues remain buried in both simulations as already indicated in Table II, with two exceptions, Leu12 and Thr25, where the SCEFs of the structures in both simulations increase by about 70%.

Among the polar and titratable residues, there are a few where one of the simulated structures shows somewhat anomalous behavior in the sense that its SCEF is substantially different from the crystal structure and the other simulated structure values. These residues are listed in Table III. It is of interest to note that this behavior is observed in both the EW and ISM simulations. Note, however, that the trend is systematically opposed: the SCEF of residues with anomalous behavior in the ISM are smaller than in the crystal structure, whereas they are larger for the corresponding residues in the EW structure.

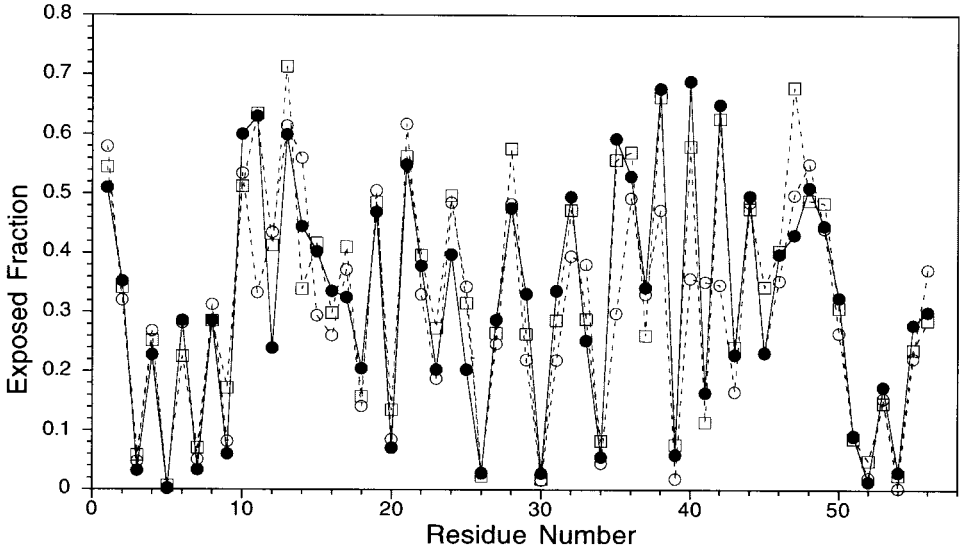

Fig. 12. Solvent-exposed fractions of the side-chains of the average structures calculated from the EW (□) and ISM (○). Note that with a few exceptions (see text and Table III), the side-chain-exposed fractions obtained from both simulations closely follow the corresponding values calculated from the X-ray structure (●).

TABLE III. Residues With Anomalous Exposed Surface Areas[a]

Res	solv mdl	xtl exp f	mdl exp f	Comment
THR11	ism	0.63	0.34	H-bonded to O of ASN37
LYS13	ew	0.60	0.71	conformational differences
THR16	ism	0.34	0.20	conformational differences
LYS28	ew	0.48	0.58	ism & xtl bent toward helix, in opposite directions
ASN35	ism	0.59	0.34	H-bonded to ASP40
ASP40	ism	0.69	0.35	H-bonded to ASN35, both residues turned inward
GLU42	ism	0.65	0.34	H-bonded to Thr55
TYR45	ew	0.23	0.34	ism and xtl in H-bond network with Asp47
ASP47	ew	0.43	0.68	ism and xtl in H-bond network to Tyr45 & Lys50

[a]The solvent exposed fraction (exp f) of a residue in a given solvent model is considered anomalous if it differs by >0.1 from the exp f of the residues in the other two structures (xtl = crystal; mdl = model).

The average structure obtained from the ISM simulation shows that residues Thr11, Asn35, Asp40, and Glu42 have bent inward, forming H-bonds with other residues, as indicated in Table III. In all these cases, the corresponding residues of the EW simulation and the crystal structure remain exposed to the solvent and do not show any tendency to interact with other side-chains. Although it is tempting to ascribe these results to insufficient solvation by the ISM, the situation is complicated by the behavior of Tyr45 and Asp47. Here residue Asp47 in the ISM and crystal structures participate in a complex (and fairly similar) H-bonding network with Tyr45 and Lys50. In contrast, in EW Asp47 is pointed outward, into the solvent and does not interact with any other side-chains. These results show the rather complex interplay between the various electrostatic and nonelectrostatic forces that govern the behavior of specific residues in the protein, as elaborated in the Discussion below.

DISCUSSION AND FUTURE DIRECTIONS

In this article, the implementation of the SCP-ISM in MD simulations was presented, complementing the earlier applications of the approach using Monte Carlo techniques. As discussed above, all these implementations used the same parameterization of the model that was obtained independently of any particular application.[75] The only modification of the potential function compared to the previously reported version of the model[75,76,98] is the implementation reported here of an analytical function for calculating the Born radii based on a contact model. This contact model was shown here to provide a procedure for calculating the Born radii based on the degree of exposure to solvent, and burial in the protein that is equivalent to the *SASA* approach (i.e., it conserves the quality of the energetics of the macromolecule as already validated in a number of problems). The optimization of the contact model presented here uses the minimal possible amount of structural information (i.e., one protein) and, therefore, its performance relies primarily on the physical basis of its formulation. Introduction of this contact model into the force field yields a pairwise potential function that makes possible the efficient evaluation of the forces. This fast analytical model reduces the time required to calculate the energy by a factor of two com-

pared to *SASA*-based calculations and makes possible the practical implementation of the SCP-ISM in MD simulations.

Although at the present stage of development, code optimization was not of primary concern, and priority was given to the quality of the energetics; the SCP-ISM also yielded a fast algorithm, scaling MD simulations by a factor of three relative to simulations in vacuum. It is to be emphasized that the SCP-ISM is a general approach, where no ad hoc assumptions were required to make the model agree with experimental or other theoretical findings. Another important, albeit practical rather than fundamental feature of the model, is the possibility of adjusting the strength of the short-range hydrogen-bonding interactions independently of the long-range Coulomb interactions.[98] In effect, HB stabilization energies can be adjusted to fit the values derived from *ab initio* calculations or from experimental data, without affecting the damping of the other electrostatic interactions in the system.

The MD simulations of the Dyn peptide and the ProtG protein were conducted to evaluate the performance of the SCP-ISM at two size scales. The results compared well with simulations using explicit solvent and experimental data. Thus, in the 3-ns MD simulation of Dynorphin using the SCP-ISM, the qualitative behavior of the peptide in EW was found to be reproduced: the α-helical motif of the peptide is gradually disrupted ending up in a turn with no characteristic HB interactions. Subsequently, after the simulation reached around 12 ns, the system began to fluctuate rapidly between partial helical and random coil conformations, suggesting a propensity for the peptide to adopt helical secondary structure, even in water, but with a weak stabilization energy, which results in the peptide acquiring predominantly a random conformation.

The simulation of protein G with the SCP-ISM resulted in a stable structure that shifted little from the X-ray structure. All the elements of secondary structure and also the tertiary structure of the protein were preserved throughout the entire simulation time. The stabilization of secondary structural elements, in particular the ends of the α-helix and β-strands, is a significant result, especially because the helical conformations of the peptide were not strongly stabilized in water. Although the overall behavior of the protein in both simulations (ISM and EW) was similar, the fluctuations generally were larger in the implicit solvent. Nevertheless, the general trend of the fluctuations was the same in both simulations, as the largest fluctuations were observed in loop segments. In both solvent models, the difference in fluctuations was larger in the side-chains, especially when fully solvated, as evidenced by the fluctuations of C_γ atoms.

The finding that the MD simulations with implicit solvent models reproduce the overall behavior of the system found in explicit solvent simulations is gratifying. However, the similarities and differences of the results obtained from these two models need to be evaluated with caution keeping in mind that in the implicit model the steric effects and the dynamics of the solvent are neglected. Implicit models lack both van der Waals interactions and explicit hydrogen bonding directionality between the solute and the solvent, which may contribute in maintaining polar and charged surface residues solvent exposed. In addition, the lack of friction and of random collisions of the water molecules should affect the dynamics of the surface residues. However, these effects could be mimicked by using, e.g., Langevin dynamics to partially account for the absent degrees of freedom of the water molecules around the peptide or protein. It is essential, therefore, that continuum models do not correct the large fluctuations or the degree of exposure to the solvent by reparameterization of the electrostatic component of the force field. Such a procedure may introduce artifacts into the results, because frictional effects (as well as others mentioned above) are nonelectrostatic in nature.

Hydrophobic effects can affect the dynamics and stabilization of macromolecules and are not accounted for in this initial MD algorithm of the SCP-ISM, although the electrostatic solvation of polar, ionized groups and hydrophobic side-chains are taken into account through the self-energy terms. To account for hydrophobic interactions, continuum models have usually been combined with simple surface area treatments of nonelectrostatic solute-solvent interactions. Although this proportionality can be derived under reasonable physical assumptions, the uncertain quality of this simple treatment of hydrophobic interactions suggests that more sophisticated models are needed, and more promising approaches have been recently reported,[22,115–123,147] which should be tested in the context of a general ISM in the near future.

The analysis of protein G in terms of RMSD and *SASA* showed that with the SCP-ISM, the interior of the protein is largely insensitive to the approximations inherent in implicit solvent models, at least within the timescale of the simulation reported here. Although apolar side-chains buried in the protein did not become solvent exposed, a few polar and charged residues did decrease their solvent exposure in the ISM simulation. However, several polar or charged residues in the EW simulation also exhibited anomalous solvent exposure, albeit always as an increase in the degree of *SASA*. Therefore, in the ISM, the behavior of these side-chains cannot be attributed simply to insufficient solvation, because other neglected, nonelectrostatic effects may be involved. Notably, however, these results do show the potential pitfalls in parameterizing implicit solvent models on the basis of structure (or other properties) of proteins, because in conducting the fitting it is possible to mask characteristics arising from nonelectrostatic effects that are neglected in implicit models.

The correct description of solvation effects is crucial in most of the important applications arising in molecular biophysics, including calculations of binding free energies, molecular recognition, docking problems, and macromolecular interactions. Given the need for conducting simulations in realistic biological timescales, it is clear that much further work must be devoted to constructing a complete implicit solvent, in which the continuum electrostatic model is likely to remain a key component.

ACKNOWLEDGMENTS

Computational support was provided by the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputing Center and by the Advanced Scientific Computing Laboratory at the Frederick Cancer Research Facility of the National Cancer Institute (Laboratory of Mathematical Biology). The authors also acknowledge access to the computer facilities at the Institute of Computational Biomedicine (ICB) of the Mount Sinai Medical Center.

REFERENCES

1. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. J Comput Chem 1983;4: 187–217.
2. Weiner S, Kollman P, Case D, Singh UC, Ghio C, Alagona G, Profeta SJ, Weiner P. A new force field for molecular mechanical simulation of nucleic acids and proteins. J Am Chem Soc 1984; 106765–784.
3. van Gunsteren WF. GROMOS. Groningen molecular simulation computer program package. The Netherlands: University of Groningen; 1987.
4. Jorgensen WL, Tirado-Rives J. The OPLS potential functions for proteins. Energy minimization for crystals of cyclic peptides and crambin. J Am Chem Soc1988;110:1657–1666.
5. Halgren TA, Damm W. Polarizable force fields. Curr Opin Struct Biol 2001;11: 236–242.
6. Wang W, Donini O, Reyes CM, Kollman PA. BIOMOLECULAR SIMULATIONS: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. Annu Rev Biophys Biomol Struct 2001;30:211–243.
7. Cheatham TE, Kollman P. Molecular dynamics simulation of nucleic acids. Ann Rev Phys Chem 2000;51:435–471.
8. Karplus K. Molecular dynamics simulations of biomolecules. Acc Chem Res 2002;35:321–323.
9. Shen TY, Tai KH, Henchman RH, McCammon JA. Molecular dynamics of acetylcholinesterase. Acc Chem Res 2002;35:332–340.
10. Warshel A. Molecular dynamics simulations of biological reactions. Acc Chem Res 2002;35:385–393.
11. Karplus K, Petsko GA. Molecular dynamics simulations in biology. Nature 1990;347:631–639.
12. Duan Y, Kollman PA. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. Science 1998;282;740–744.
13. Feher VA, Cavanagh J. Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F. Nature 1999;400:289–293.
14. Volkman BF, Lipson D, Wemmer DE, Kern D. Two-state allosteric behavior in a single-domain signaling protein. Science 2001; 291:2429–2433.
15. Gao G, Semenchenko V, Arumugan S, van Doren SR. Tissue inhibitor of metalloproteinases-1 undergoes microsecond to millisecond motions at sites of matrix metalloproteinase-induced fit. J Mol Biol 2000;301:537–552.
16. Gottschalk M, Dencher NA, Halle B. Microsecond exchange of internal water molecules in bacteriorhodopsin. J Mol Biol 2001; 311:605–621.
17. Simonson T. Macromolecular electrostatics: continuum models and their growing pains. Curr Opin Struct Biol 2001;11:243–252.
18. Sharp KA, Honig B. Electrostatic interactions in macromolecules: theory and applications. Annu Rev Biophys Biophys Chem 1990;19:301–332.
19. Honig B, Sharp K, Yang A-S. Macroscopic models of aqueous solutions: biological and chemical applications. J Phys Chem 1993;97:1101–1109.
20. Tazaki K, Doi J. Effective atomic charge model for solvent and ion screening effect in iterative calculation and application to Monte Carlo simulation. J Phys Chem 1996;100:14520–14525.
21. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
22. Cramer CJ, Truhlar DG. An SCF solvation model for the hydrophobic effect and absolute free energies of aqueous solvation. Science 1992;256:213–217.
23. Qiu D, Shenkin PS, Hollinger P, Still WC. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. J Phys Chem B 1997;101:3005–3014.
24. Schaefer M, Karplus M. A comprehensive analytical treatment of continuum electrostatics. J Phys Chem 1996;100:1578–1599.
25. Resat H, Marrone TJ, McCammon JA. Enzyne-inhibitor association thermodynamics: explicit and continuum solvent studies. Biophys J 1997;72:522–532.
26. Lazaridis T, Karplus M. Effective energy functions for protein structure prediction. Curr Opin Struct Biol 2000;10:139–145.
27. Orozco M, Luque FJ. Theoretical methods for the description of the solvent effect in biomolecular systems. Chem Rev 2000;100: 4187–4226.
28. Arnold GE, Ornstein RL. An evaluation of implicit and explicit solvent model systems for the molecular dynamics simulation of bacteriophage T4 lysozyme. Proteins 1994;18:19–33.
29. Cramer CJ, Truhlar DG. Implicit solvation models: equilibria, structure, spectra, and dynamics. Chem Rev 1999;1999:2161–2200.
30. Geunot J, Kollman PA. Molecular dynamics studies of a DNA-binding protein. 2. An evaluation of implicit and explicit solvent models for the molecular dynamics simulation of the escherichia coli trp repressor. Protein Sci 1992;1:1185–1205.
31. Roux B, Simonson T. Implicit solvent models. Biophys Chem 1999;78:1–20.
32. Shen MY, Freed KF. Long time dynamics of met-enkephalin: comparison of explicit and implicit solvent models. Biophys J 2002;82:1791–1808.
33. Zhang LY, Gallicchio E, Friesner RA, Levy RM. Solvent models for protein-ligand binding: comparison of implicit solvent poisson and surface generalized Born models with explicit solvent simulations. J Comp Chem 2000;22:591–607.
34. Hardin C, Pogorelov TV, Luthey-Schulten Z. Ab initio protein structure prediction. Curr Opin Struct Biol 2002;12:176–181.
35. Liu YX, Beveridge SL. Exploratory studies of ab initio protein structure prediction: multiple copy simulated annealing AMBER energy functions, and a generalized Born/solvent accessibility solvation model. Proteins 2002;46:128–146.
36. Bonneau R, Baker D. Ab initio protein structure prediction: progress and prospects. Annu Rev Biophys Biomol Struct 2001;30: 173–189.
37. Gibbs N, Clarke AR, Sessions RB. Ab initio protein structure prediction using physicochemical potentials and a simplified off-lattice model. Proteins 2001;43:186–202.
38. Petrella RJ, Karplus K. A limiting-case study of protein structure prediction: energy-based searches of reduced conformational space. J Phys Chem B 2000;104:11370–11378.
39. Lazaridis T, Karplus M. Effective energy function for proteins in solution. Proteins 1999;35:133–152.
40. Godzik A, Kolinski A, Skolnick J. Lattice representation of globular proteins: how good are they? J Comp Chem 1993;14: 1194–1202.
41. Sikorski A, Kolinski A, Skolnick J. Computer simulations of de novo designed helical proteins. Biophys J 1998;75:92–105.
42. Pedersen JT, Moult J. Protein folding simulations with genetic algorithms and a detailed molecular description. J Mol Biol 1997;269:240–259.
43. Unger R, Moult J. Local interactions dominate folding in a simple protein model. J Mol Biol 1996;259:988–994.
44. Monge A, Lathrop EJP, Gunn JR, Shenkin PS, Friesner RA. Computer modeling of protein-folding conformational and energetic analysis of reduced and detailed protein models. J Mol Biol 1995;247:995–1012.
45. Kolinski A, Betancourt MR, Kihara D, Rotkiewicz P, Sknolnick J. Generalized comparative modeling (GENECOMP): a combination of sequence comparison, threading, and lattice modeling for protein structure prediction and refinement. Proteins 2001;44: 133–149.
46. Chiu TL, Goldstein RA. How to generate improved potentials for protein tertiary structure prediction: a lattice model study. Proteins 2000;41:157–163.
47. Mandell JG, Roberts VA, Pique ME, Kotlovyi V, Mitchell JC, Nelson E, Tsigelny I, Ten Eyck LF. Protein docking using continuum electrostatics and geometric fit. Protein Eng 2001;14: 105–113.

48. Jackson RM, Gabb HA, Sternberg MJE. Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. J Mol Biol 1998;276:265–285.
49. Jackson RM, Sternberg MJE. A continuum model for protein-protein interactions. Application to the docking problem. J Mol Biol 1995;250:258–275.
50. Majeux N, Scarsi M, Apostolakis J, Ehrhardt C, Caflisch A. Exhaustive docking of molecular fragments with electrostatic solvation. Proteins 1999;37:88–105.
51. Muegge I. A knowledge-based scoring function for protein-ligand interactions: probing the reference state. Perspect Drug Discov 2000;20:99–114.
52. Shoichet BK, Leach AR, Kuntz ID. Ligand solvation in molecular docking. Proteins 1999;34:4–16.
53. Arora N, Bashford D. Solvation energy density occlusion approximation for evaluation of desolvation penalties in biomolecular interactions. Proteins 2001;43:12–27.
54. Camacho CJ, Weng ZP, Vajda S, DeLisi C. Free energy landscapes of encounter complexes in protein-protein association. Biophys J 1999;76:1166–1178.
55. Perez C, Ortiz AR. Evaluation of docking functions for protein-ligand docking. J Med Chem 2001;44:3768–3785.
56. Mehler EL, Guarnieri F. A self-consistent, microenvironment modulated screened coulomb potential approximation to calculate pH dependent electrostatic effects in proteins. Biophys J 1999;77:3–22.
57. Mehler EL. A self-consistent, free energy based approximation to calculate pH dependent electrostatic effects in proteins. J Phys Chem 1996;100:16006–16018.
58. Bashford D, Karplus M. pK_a's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. Biochemistry 1990;29:10219–10225.
59. Lim C, Bashford D, Karplus M. Absolute pK_a calculation with continuum dielectric methods. J Phys Chem 1991;95:5610–5620.
60. MacKerell AD, Sommer MS, Karplus M. pH-Dependence of binding reactions from free energy simulations and macroscopic continuum electrostatic calculations. Application to 2′GMP/3′GMP binding to ribonuclease T-1 and implication for catalysis. J Mol Biol 1995;247:774–807.
61. Tsui V, Case DA. Calculations of the absolute free energies of binding between RNA and metal ions using molecular dynamics simulations and continuum electrostatics. J Phys Chem B 2001;105:11314–11325.
62. Simonson T, Archontis G, Karplus M. Continuum treatment of long-range interactions in free energy calculations. Application to protein-ligand binding. J Phys Chem 1997;101:8349–8362.
63. Gogonea V, Suarez D, van der Vaart A, Merz KM Jr. New developments in applying quantum mechanics to proteins. Curr Opin Struct Biol 2001;11:217–223.
64. Cortis CM, Langlois JM, Beachy MD, et al. Quantum mechanical geometry optimization in solution using a finite element continuum electrostatics method. J Chem Phys 1996;105:5472–5484.
65. Lybrand TP. Ligand-Protein Docking and Rational Drug Design. Curr Opin Struct Biol 1995;5:224–228.
66. Davis ME, Madura JD, Luty BA, McCammon JA. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston brownian dynamics program. Comp Phys Comm 1991;62:187–197.
67. Madura JD, Briggs JM, Wade RC, Davis ME, Luty BA, Ilin A, Antosiewics J, Gilson MK, Bagheri B, Scott LR, McCammon JA. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston brownian dynamics program. Comput Phys Commun 1995;91:57–95.
68. Rocchia W, Alexov EG, Honig B. Extending the applicability of the nonlinear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions. J Phys Chem B 2001;105:6507–6514.
69. Bashford D, editor. An object-oriented programming suite for electrostatic effects in biological molecules. Berlin: Springer; 1997.
70. Hoijtink GJ, de Boer E, van der Meer PH, Weijland WP. Reduction potentials of various aromatic hydrocarbons and their univalent anions. Rec Trav Chim 1956;75:487–503.
71. Hawkins GD, Cramer CJ, Truhlar DG. Pairwise descreening of solute charges from a dielectric medium. Chem Phys Lett 1995;246:122–129
72. Ghosh A, Rapp C, Friesner RA. Generalized Born model based on a surface area formulation. J Phys Chem B 1998;102:10983–10990.
73. Calimet N, Schaefer M, Simonson T. Protein molecular dynamics with the generalized Born/ACE solvent model. Proteins 2001;45:144–158.
74. Tsui V, Case DA. Molecular dynamics simulations of nucleic acids with a generalized Born solvation model. J Am Chem Soc 2000;122:2489–2498.
75. Hassan SA, Guarnieri F, Mehler EL. A general treatment of solvent effects based on screened coulomb potentials. J Phys Chem B 2000;104:6478–6489.
76. Hassan SA, Mehler EL. A critical analysis of continuum electrostatics: the screened coulomb potential-implicit solvent model and the study of the alanine dipeptide and discrimination of misfolded structures of proteins. Proteins 2002;47:45–61.
77. Debye P. Polar molecules. New York: Dover; 1929.
78. Lorentz HA. Theory of electrons. New York: Dover; 1952.
79. Sack VH. The dielectric constant of electrolytes. Phys Z 1926;27:206–208.
80. Sack VH. The dielectric constants of solutions of electrolytes at small concentrations. Phys Z 1927;28:199–210.
81. Mehler EL. The Lorentz-Debye-Sack theory and dielectric screening of electrostatic effects in proteins and nucleic acids. In: Murray JS, Sen K, editors. Molecular electrostatic potential: concepts and applications. Amsterdam: Elsevier Science; 1996; p 371–405.
82. Ehrenson S. Continuum radial dielectric functions for ion and dipole solution systems. J Comp Chem 1989;10:77–93.
83. Onsager L. Electric moments of molecules in liquids. J Am Chem Soc 1936;58:1486–1493.
84. Böttcher CJF. The dielectric constant of dipole liquids. Physica 1938;V:635–639.
85. Scarsi M, Apostolakis J, Caflisch A. Continuum electrostatic energies of macromolecules in aqueous solutions. J Phys Chem B 1997;101:8098–8106.
86. Antosiewicz J, McCammon JA, Gilson MK. Prediction of pH-dependent properties of proteins. J Mol Biol 1994;238:415–436.
87. Pollock EL, Alder BJ, Pratt LR. Relation between the local field at large distances from a charge or dipole and the dielectric constant. Proc Natl Acad Sci USA 1980;77:49–51.
88. Sham YY, Chu ZT, Warshel A. Consistent calculations of pKa's of ionizable residues in proteins: semi-microscopic and microscopic approaches. J Phys Chem B 1997;101:4458–4472.
89. Sham YY, Muegge I, Warshel A. The effect of protein relaxation on charge-charge interactions and dielectric constants of proteins. Biophys J 1998;74:1744–1753.
90. Warshel A, Papazyan A. Electrostatic effects in macromolecules: fundamental concepts and practical modeling. Curr Opin Struct Biol 1998;8:211–217.
91. Warshel A, Russell ST, Churg AK. Macroscopic models for studies of electrostatic interactions in proteins: limits and applicability. Proc Natl Acad Sci USA 1984;81:4785–4789.
92. Mehler EL, Fuxreiter M, Simon I, Garcia-Moreno BE. The role of hydrophobic microenvironment in modulating pKa shifts in proteins. Proteins 2002;48:283–292.
93. Cohen BE, McAnaney TB, Park ES, Jan YN, Boxer SG, Jan LY. Probing protein electrostatics with a synthetic fluorescent amino acid. Science 2002;296:1700–1703.
94. Davis ME, McCammon JA. Calculating electrostatic forces from grid-calculated potentials. J Comp Chem 1990;11:401.
95. Sharp K. Incorporating solvent and ionic screening into molecular dynamics using the finite-difference Poisson-Boltzman method. J Comp Chem 1991;12:454–468.
96. Gilson MK, Davis ME, Luty BA, McCammon JA. Computation of electrostatic forces on solvated molecules using the Poisson-Boltzmann equation. J Phys Chem 1993;97:3591–3600.
97. Bucher M, Porter TL. Analysis of the Born model for hydration of ions. J Phys Chem 1986;90:3406–3411.
98. Hassan SA, Guarnieri F, Mehler EL. Characterization of hydrogen bonding in a continuum solvent model. J Phys Chem B 2000;104:6490–6498.
99. Hassan SA, Mehler EL. A general screened coulomb potential based implicit solvent model: calculation of secondary structure of small peptides. Int J Quant Chem 2001;83:193–202.
100. Hassan SA, Mehler EL, Weinstein H. Structure calculations of protein segments connecting domains with defined secondary structure: a simulated annealing Monte Carlo combined with biased scaled collective variables technique. In: Schlick T, editor.

Lecture notes series in computational science. New York: Springer Verlag, Ag; 2002. p 197–231.
101. MacKerell AD, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher I, Roux WEB, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–3616.
102. Colonna Cesarei F, SanderC. Excluded volume approximation for protein-solvent interactions. The solvent contact model. Biophys J 1990;57:1103–1107.
103. Dominy BN, Brooks ICL. Development of a generalized Born model parametrization for proteins and nucleic acids. J Phys Chem B 1999;103:3765–3773.
104. Onufriev A, Bashford D, Case DA. Modification of the generalized Born model suitable for macromolecules. J Phys Chem B 2000;104:3712–3720.
105. Webb TJ. The free energy of hydration of ions and the electrostriction of the solvent. J Am Chem Soc 1926;48:2589–2603.
106. Debye P, Pauling L. The inter-ionic attraction theory of ionized solutes. IV. The influence of variation of dielectric constant on the limiting law for small concentrations. J Am Chem Soc 1925;47:2129–2134.
107. Schwarzenbach G. Der Einfluss einer Ionenladung auf die Acidität einer Säure. Z Physik Chem A 1936;176:133–153.
108. Harvey SC, Hoekstra P. Dielectric relaxation spectra of water adsorbed on lysozyme. J Phys Chem 1972;76:2987–2994.
109. Pennock BD, Schwan HP. Further observations on the electrical properties of hemoglobin-bound water. J Phys Chem 1969;73: 2600–2610.
110. Takashima S, Schwan HP. Dielectric dispersion of crystalline powders of amino acids, peptides and proteins. J Phys Chem 1965;69:4176–4182.
111. Hasted JB, Ritson DM, Collie CH. Dielectric properties of aqueous ionic solutions. Parts I and II. J Chem Phys 1948;16:1–21.
112. Latimer WM, Rodebush WH. Polarity and ionization from the standpoint of the Lewis theory of valence. J Am Chem Soc 1920;42:1419–1433.
113. Sridharan S, Nicholls A, Sharp K. A rapid method for calculating derivatives of solvent accessible surface areas of molecules. J Comp Chem 1995;16:1038–1044.
114. Weiser J, Shenkin PS, Still WC. Approximate solvent-accessible surface areas from tetrahedrally directed neighbor densities. Biopolymers 1999;50:373–380.
115. Hummer G,Garde S. Cavity expulsion and weak dewetting of hydrophobic solutes in water. Phys Rev Lett 1998;80:4193–4196.
116. Hummer G, Garde S, Garcia AE, Paulaitis ME, Pratt EA. Hydrophobic effects on a molecular scale. J Phys Chem B 1998;102:10469–10482.
117. Lum K, Chandler D, Weeks JD. Hydrophobicity at small and large length scales. J Phys Chem B 1999;103:4570–4577.
118. Hummer G, Garde S, Garcia AE, Pohorille A, Pratt EA. An information theory model of hydrophobic interactions. Proc Natl Acad Sci USA 1996;93:8951–8955.
119. Huang DM, Chandler D. The hydrophobic effect and the influence of solute-solvent attractions. J Phys Chem B 2002;106:2047–2053.
120. Hummer G, Garde S, Garcia AE, Pratt EA. New perspectives on hydrophobic effects. Chem Phys 2000;258:349–370.
121. Chandler D. Two faces of water. Nature 2002;417:491–491.
122. Garde S, Hummer G, Garcia AE, Pratt EA, Paulaitis ME. Hydrophobic hydration: inhomogeneous water structure near nonpolar molecular solutes. Phys Rev E 1996;53:R4310–R4313.
123. Simonson T, Brunger AT. Solvation free energies estimated from macroscopic continuum theory: an accuracy assessment. J Phys Chem 1994;98:4683–4694.
124. Chavkin C, Goldstein A. Specific receptor for the opioid peptide dynorphin: structure-activity relationships. Proc Natl Acad Sci USA 1981;78:6543–6547.
125. Claye LH, Unterwald EM, Ho A, Kreek MJ. Both dynorphin A(1-17) and [des-tyr(1)]dynorphin A(2-17) inhibit adenylyl-cyclase activity in rat caudate-putamen. J Pharmacol Exp Ther 1996;277:359–365.
126. Erne D, Sargent DF, Schwyzer R. Preferred conformation, orientation, and accumulation of dynorphin A-(1-13)-tridecapeptide on the surface of neutral lipid membranes. Biochemistry 1985;24:4261–4263.
127. Goldstein A. Biology and chemistry of the dynorphin peptides. In: Meienhofer J, Udenfriend S, editors. The peptides. New York: Academic Press; 1984. p 95–145.
128. Sankararamakrishnan R, Weinstein H. Molecular dynamics simulations predict a tilted orientation for the helical region of dynorphin A(1-17) in dimyristoylphosphatidylcholine bilayers. Biophys J 2000;79:2331–2344.
129. Sankararamakrishnan R, Weinstein H. Positioning and stabilization of dynorphin peptides in membrane bilayers: the mechanistic role of aromatic and basic residues revealed from comparative MD simulations. J Phys Chem B 2002;106:209–218.
130. Renugopalakrishnan V, Rapaka RS, Huang S-G, Moore S, Hutson TB. Dynorphin A(1-13) peptide NH groups are solvent exposed: FT-IR and 500MHz 1H NMR spectroscopic evidence. Biochem Biophys Res Commun 1988;151:1220–1225.
131. Tessmer MA, Kallick DA. NMR and structural model of dynorphin A(1-17) bound to dodecylphosphocholine micelles. Biochemistry 1997;36:1971–1981.
132. Gallagher T, Alexander P, Bryan P, Gilliland GL. Two crystal structures of the BI immunoglobulin-binding domain of streptococcal protein G and comparison with NMR. Biochemistry 1994; 33:4721–4729.
133. Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM. A novel, highly stable fold of the immunoglobulin binding domian of streptococcal protein G. Science 1991;253:657–661.
134. Alexander P, Fahnestock S, Lee T, Orman J, Bryan P. Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: why small proteins tend to have high denaturation temperatures. Biochemistry 1992;31:3597–3603.
135. Alexander P, Orban J, Bryan P. Kinetic analysis of folding and unfolding the 56 amino acid IgG-binding domain of streptococcal protein G. Biochemistry 1992;31:7243–7248.
136. Frank MK, Clore GM, Gronenborn AM. Structural and dynamic characterization of the urea denatured state of the immunoglobulin binding domain of streptococcal protein G by multidimensional heteronuclear NMR spectroscopy. Protein Sci 1995;4:2605–2615.
137. Kuszewski J, Clore GM, Gronenborn AM. Fast folding of a prototypic polypeptide: the immunoglobulin binding domain of streptococcal protein G. Protein Sci 1994;3:1945–1952.
138. Blanco FJ, Jimenez MA, Pineda A, Rico M, Santoro J, Nieto JL. NMR solution structure of the isolated fragment protein G B1 domain: evidence of trifluorethanol induced native-like β hairpin formation. Biochemistry 1994;33:6004–6014.
139. Blanco FJ, Serrano L. Folding of protein G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements. Eur J Biochem 1995;230:634–649.
140. Smith C, Regan L. Guidelines for protein design: the energies of the β sheet side chain interactions. Science 1995;270:980–982.
141. Barshi JJ Jr, Grasberger B, Gronenborn AM, Clore GM. Investigation of the backbone dynamics of the IgG-binding domian of streptococcal protein G by heteronuclear two-dimensional 1H-15N nuclear magnetic resonance spectroscopy. Protein Sci 1994; 3:15–21.
142. Orban J, Alexander P, Bryan P. Hydrogen-deuterium exchange in the free and immunoglobulin G-bound protein G B-domain. Biochemistry 1994;33:5702–5710.
143. Sheinerman FB, Brooks CL III. A molecular dynamics simulation study of segment B1 of protein G. Proteins 1997;29:193–202.
144. Ceruso MA, Amadei A, Di Nola A. Mechanics and dynamics of B1 domain of protein G: role of packing and surface hydrophobic residues. Protein Sci 1999;8:147–160.
145. Sankararamakrishnan R, Konvicka K, Mehler EL, Weinstein H. Solvation in simulated annealing and high temperature molecular dynamics of proteins: a restrained water droplet model. Int J Quant Chem 2000;77:174–186.
146. Zhu JA, Shi YY, Liu HY. Parametrization of a generalized Born/solvent-accessible surface area model and applications to the simulation of protein dynamics. J Phys Chem B 2002;106: 4844–4853.
147. 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.