JCTC Journal of Chemical Theory and Computation

A Nonradial Coarse-Grained Potential for Proteins Produces Naturally Stable Secondary Structure Elements

Davide Alemani,[†] Francesca Collu,[‡] Michele Cascella,*[,‡] and Matteo Dal Peraro*[,†]

Laboratory for Biomolecular Modeling, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland and Departement für Chemie und Biochemie, Universität Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

Received August 28, 2009

Abstract: We introduce a nonradial potential term for coarse-grained (CG) molecular simulations of proteins. This term mimics the backbone dipole−dipole interactions and accounts for the needed directionality to form stable folded secondary structure elements. We show that α-helical and β-sheet peptide chains are correctly described in dynamics without the need of introducing any a priori bias potentials or ad hoc parametrizations, which limit broader applicability of CG simulations for proteins. Moreover, our model is able to catch the formation of supersecondary structural motifs, like transitions from long single α-helices to helix−coil−helix or β-hairpin assemblies. This novel scheme requires the structural information of C_α beads only; it does not introduce any additional degrees of freedom to the system and has a general formulation, which allows it to be used in synergy with various CG protocols, leading to an improved description of the structural and dynamic properties of protein assemblies and networks.

1. Introduction

Molecular dynamics (MD) simulations have proven to be a powerful tool to investigate the structure and function of biomolecular systems. Among different approaches to modeling proteins, nucleic acids, and biological membranes, all-atom MD has been shown to provide a reasonable compromise between the accuracy of the force fields used to describe molecular interactions and the computational cost required to simulate relevant systems.[1] The continuous increase in computational power allows a routine application of such techniques to systems as large as 10^2 kDa at the multi-nanosecond time scale, when using high-performance-computing resources. Moreover, benefiting from such improved throughput, the scientific community has been recently very active in testing the reliability of current force fields[2,3] and in developing new models, protocols, and algorithms to increase both the accuracy and the performance of current MD codes. Nonetheless, within the atomistic framework, it remains computationally unaffordable to thoroughly sample size and time scales that are critical to most of the biological processes both in vitro and in vivo. In fact, fundamental events like protein folding, signal transduction, or DNA transcription, while all triggered by interactions at atomistic dimensionality, occur at very different time scales (from the femtosecond to the second and more) and span over different sizes (from few tens to millions of atoms). Such a scale range implies a dimensionality of the corresponding phase-space so large that its complete sampling by brute-force all-atom MD remains at the moment unfeasible.[4]

To extend the boundaries of time and size, one can abandon the atomistic representation for a coarser description of molecular systems. From the original framework of elastic networks used by Levitt and Warshel[5] and Go and Scheraga[6,7] to describe protein folding, more recent coarse-grained (CG) models make use of beads to represent groups of atoms that interact through effective potentials, which have a functional form similar to that of atomistic force fields. Successful CG models were introduced to describe polar/nonpolar interactions, such as in lipid self-assembly processes in water.[8−10] On the basis of these seminal works, new generations of CG models were developed to study aqueous surfactant solutions,[11,12] and membrane lipid assemblies.[13−16] Current

* Corresponding author e-mail: michele.cascella@iac.unibe.ch (M.C.); matteo.dalperaro@epfl.ch (M.D.P.).

[†] EPFL.

[‡] Universität Bern.

10.1021/ct900457z
Published on Web 12/22/2009

computational power allows using these schemes to access size scales on the order of 10^6 CG particles (i.e., ≈10 million atoms) and time scales approaching relevant experimental regimes (i.e., millisecond).[4] Recently, effective potentials based on similar atom-to-bead mapping have been extended to proteins[17−20] and nucleosomes.[21−25] These models have been used in the past years to study very large membrane−protein complexes, showing promising results for the study of structural rearrangements, which are functional to many biological processes.[26−36] Even coarser levels of approximating protein potentials have been proposed on the basis of a single bead representation of amino acids[37−39] or extended to describe well-defined structural intraprotein domains or even single proteins in multimolecular assemblies.[26,40]

Progressive efforts in linking CG models to the atomistic descriptions have also been produced by the computational community, with the scope of developing multiscale frameworks able to synergically exploit the complementary advantages of descriptions of molecular systems at different levels of resolution.[25,41−53]

Nonetheless, to date, many issues still afflict CG schemes, which limit in turn their general applicability to a vast class of relevant biological problems. The functional form of CG potentials is not univocally defined and strongly depends on the level of coarse-graining one chooses to adopt.[52,54] Moreover, accurate potential function parameters for CG MD simulations are still laking full transferability; therefore, they have to be recalibrated according to the system of interest.[55,56] The lack of both universality and transferability results in significant drawbacks, which reduce their general applicability. In particular, fully transferable CG models based on pair potentials can face difficulties in reproducing anisotropic properties, which are crucial for accounting for both stability in secondary structure elements and correct tertiary structure assembling. One possibility to circumvent such problems is to maintain a higher resolution (atomistic or quasi-atomistic) at least in the backbone region.[57−59] In the case where a one-bead per backbone unit mapping is chosen, it is usually necessary to introduce additional bias potentials, which are defined on a target conformation and do not naturally adapt to secondary structure modifications, which may occur during dynamics.[17,26,31]

In the past years, significant results have been achieved by Scheraga and co-workers, by defining coarse-grained potentials where dipole−dipole interactions for the backbone are explicitly taken into account.[60−63] Such a model has proven to be particularly effective for energy calculations in folding predictions. Recently, backbone dipole−dipole interactions within a similar scheme were also implemented in a MD scheme.[64]

In a recent previous paper, we have shown how the orientation of the backbone dipole can be directly reconstructed from the positions of C_α atoms only.[65] From that standpoint, we have here formulated a computational protocol able to account for the intrinsic nonradial nature of backbone−backbone interactions in a CG−molecular dynamics framework using a single-bead representation for backbone units. We have evaluated its performance in a series of test systems, finding that our designed potential term is able to naturally stabilize secondary structure motifs, such as α-helices and β-sheets, to describe basilar structure rearrangements and to reliably modulate structural transitions into supersecondary assemblies. The proposed potential energy term is computationally very efficient, as it is fully defined as a function of the backbone C_α coordinates; therefore, it does not introduce any additional degrees of freedom to the system. Together with analytical modifications in the bending and torsional potentials, also discussed in this work, this novel scheme represents a promising step toward the development of a CG force field, which can accurately describe the structural and dynamic properties of protein assemblies and networks.

2. Computational Methods

Parameterized potential functions in all-atom (AA) molecular dynamics usually take the form

$$V_{AA} = \sum_{\text{bonds}} k_r(r - r_0)^2 + \sum_{\text{angles}} k_\gamma(\gamma - \gamma_0)^2 + \sum_{\text{dihedrals}} [1 + \cos(n\varphi + \delta)] + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right] + \sum_{i<j}\frac{q_i q_j}{R_{ij}} \quad (1)$$

where the first three terms represent stretching, bending, and torsional potentials, and the last two terms are nonbonded van der Waals and Coulomb interactions. Our model potential for CG backbone beads is expressed in a similar form, but with some important modifications. In particular, we change the functional form of the angular potential, we introduce a correlation term for consecutive dihedral angles, and we replace point-charge Coulombic interactions with backbone dipole−dipole interactions:

$$V_{CG} = \sum_{\text{bonds}} k_r(r - r_0)^2 + \sum_{\text{angles}} k_\gamma P^4(\gamma - \gamma_0) + \sum_{\text{dihedrals}} k_\varphi[1 + \cos(n\varphi + \delta)] + \sum_{\text{dihedrals}-1} k_{\text{corr}}(\varphi_i - \varphi_{i+1})^2 + \sum_{k<l}\left(\frac{A_{kl}}{R_{kl}^{12}} - \frac{B_{kl}}{R_{kl}^{6}}\right) + \sum_{i<j} V[\boldsymbol{\mu}_i(C_{\alpha,i-1}, C_{\alpha,i}, C_{\alpha,i+1}), \boldsymbol{\mu}_j(C_{\alpha,j-1}, C_{\alpha,j}, C_{\alpha,j+1})] \quad (2)$$

In this expression, indices k and l run on all beads, while indices i and j run on all groups of three consecutive bonded $C_{\alpha,i-1}$, $C_{\alpha,i}$, and $C_{\alpha,i+1}$ beads (called C_i triplet hereafter, Figure 1).

2.1. Dipole−Dipole Interactions. Backbone atomistic dipole−dipole interactions are modeled as directional interactions between two C_i triplets. An electrostatic dipole $\boldsymbol{\mu}_i$ is associated with each C_i triplet. $\boldsymbol{\mu}_i$ mimics the electrostatic dipole associated with the peptide bond located between amino acids i and $i + 1$ (Figure 1). The geometrical orientation of these dipoles is constrained to that of the C_i triplet to which they belong. In this way, forces acting on dipoles can be transmitted to the relative C_α beads, without increasing the number of degrees of freedom to be propagated during dynamics.

The protocol works as follows (Figure 2):

• Determination of the spatial orientation of all dipoles $\{\boldsymbol{\mu}_i\}$ starting from the coordinates of all C_i triplets

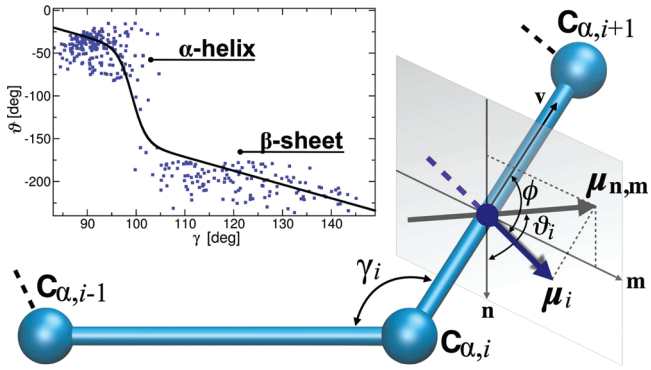

Figure 1. Backbone dipole reconstruction. The backbone dipole moment $\boldsymbol{\mu}_i$ as a function of the local reference system ($\boldsymbol{v}$, $\boldsymbol{n}$, $\boldsymbol{m}$) associated with the C_i triplet. ϕ is the angle between the $\boldsymbol{\mu}_i$ and the v axis. ϑ_i is the angle between the component of $\boldsymbol{\mu}_i$ on the plane $\boldsymbol{n}$, $\boldsymbol{m}$ ($\boldsymbol{\mu}_{n,m}$), and the $\boldsymbol{n}$ axis. The inset shows the statistical distribution of θ_i as a function of γ extracted from the PDB (square symbols) and its analytical fitting function (solid line). The local reference system is represented at the middle point of the $C_{\alpha,i}$–$C_{\alpha,i+1}$ bond for the sake of clarity.

• Computation of forces $\boldsymbol{F}_{ij}$ and torques $\boldsymbol{\tau}_{ij}$ acting on each dipole $\boldsymbol{\mu}_i$, due to dipole–dipole interactions
• Distribution of $\boldsymbol{F}_{ij}$ and $\boldsymbol{\tau}_{ij}$ to the corresponding C_i triplet
• Time propagation of C_i coordinates, following classical equations of motion (EoM)

Below, each step of this protocol is explained in detail.

Computation of the Backbone Dipole Moment. Given a triplet C_i, let γ_i be the angle formed by the C_α beads of the triplet and S_i be the local reference system of axes ($\boldsymbol{v}$, $\boldsymbol{n}$, $\boldsymbol{m}$), centered at $C_{\alpha,i}$, defined as follows: $\boldsymbol{v}$ is directed along the direction given by $C_{\alpha,i}$ and $C_{\alpha,i+1}$; $\boldsymbol{n}$ is orthogonal to the plane formed by the C_i triplet; $\boldsymbol{m}$ is orthogonal to $\boldsymbol{v}$ and $\boldsymbol{n}$ (Figure 1 and Figure 2, panel 2).

In a recent paper,[65] we showed that the components of the backbone dipole along such axes can be obtained with good approximation by fitting their experimental measurements from the PDB databank as functions of the angle γ. For the purposes of this work, we have reformulated such a fit by using the following representation in spherical coordinates:

$$\begin{aligned} \mu_{i,v} &= \mu_0 \cos\phi \\ \mu_{i,n} &= \mu_0 \sin\phi \cos\vartheta \\ \mu_{i,m} &= \mu_0 \sin\phi \sin\vartheta \end{aligned} \tag{3}$$

where μ_0 is the modulus of the vector $\boldsymbol{\mu}$ (3.6 D, modeled as two charges of ±0.33 au, located at a distance $\delta = 2.273$ Å, mimicking the atomistic intensity)[66] and ϕ and ϑ are defined as in Figure 1. In this way, by construction, the angle ϕ is a constant, with $\cos\phi = 0.177$. The angle ϑ_i remains as the only variable, which is fitted as a function of the angle γ_i:

$$\vartheta_i = -1.607\gamma_i + 0.094 + \frac{1.883}{\exp\left[\dfrac{(\gamma_i - \gamma_0)}{\sigma}\right] + 1} \tag{4}$$

where $\gamma_0 = 1.730$ rad and $\sigma = 0.025$ rad (see inset of Figure 1).

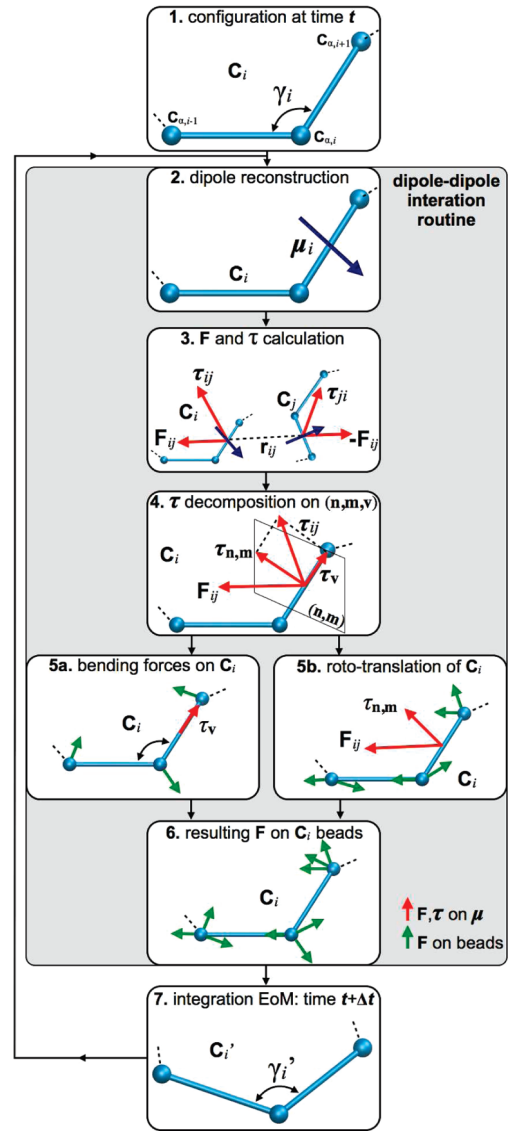

Figure 2. Dipole–dipole interaction protocol. The theoretical protocol of the new algorithm used to describe the dipole–dipole interactions.

Computation of Forces and Torques. The electrostatic dipole–dipole interaction potential reads

$$V(\boldsymbol{\mu}_i, \boldsymbol{\mu}_j) = \frac{\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j}{\mathbf{r}^3} - \frac{3}{\mathbf{r}^5}(\boldsymbol{\mu}_i \cdot \mathbf{r})(\boldsymbol{\mu}_j \cdot \mathbf{r}) \tag{5}$$

where $\mathbf{r}$ is the vector connecting the two dipoles. Forces $\boldsymbol{F}_{ij}$ and torques $\boldsymbol{\tau}_{ij}$ and $\boldsymbol{\tau}_{ji}$ acting between dipole couples are derived accordingly (Figure 2, panel 3):

$$\boldsymbol{F}_{ij} = \frac{3}{\mathbf{r}^5}(\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j)\mathbf{r} - \frac{15}{\mathbf{r}^7}(\boldsymbol{\mu}_i \cdot \mathbf{r})(\boldsymbol{\mu}_j \cdot \mathbf{r})\mathbf{r} + \frac{3}{\mathbf{r}^5}[(\boldsymbol{\mu}_j \cdot \mathbf{r})\boldsymbol{\mu}_i + (\boldsymbol{\mu}_i \cdot \mathbf{r})\boldsymbol{\mu}_j] \tag{6}$$

and

$$\begin{aligned} \boldsymbol{\tau}_{ij} &= -\frac{1}{\mathbf{r}^3}(\boldsymbol{\mu}_i \times \boldsymbol{\mu}_j) + \frac{3}{\mathbf{r}^5}(\boldsymbol{\mu}_j \cdot \mathbf{r})(\boldsymbol{\mu}_i \times \mathbf{r}) \\ \boldsymbol{\tau}_{ji} &= -\frac{1}{\mathbf{r}^3}(\boldsymbol{\mu}_j \times \boldsymbol{\mu}_i) + \frac{3}{\mathbf{r}^5}(\boldsymbol{\mu}_i \cdot \mathbf{r})(\boldsymbol{\mu}_j \times \mathbf{r}) \end{aligned} \tag{7}$$

Equations 6 and 7 are valid under the condition that the dipole length δ is much shorter than the distance between the same dipoles, that is, $\delta \ll |\mathbf{r}|$. At short distances, that is, when δ is on the same order of magnitude as $|\mathbf{r}|$, the dipole−dipole interactions are treated by considering the opposite charges of absolute value $q = 0.33$ au, placed at the edges of the dipole and at a fixed distance of δ, and computing point-charge interactions.

Dynamics of the C_i Triplets. The spatial orientation of the dipole $\boldsymbol{\mu}_i$ is fully determined by the geometry of the C_i triplet. Therefore, forces and torques acting on the dipole can be explicitly transformed into forces acting on the three beads forming C_i.

In fact, because any γ_i angle is associated with the respective ϑ_i angle via eq 4, the application of a torque on $\boldsymbol{\mu}_i$, which produces a variation of ϑ_i, is necessarily associated with a variation of γ_i (Figure 1 and Figure 2, panel 5a). On the contrary, the angle ϕ is constrained to a fixed values, therefore, torques that would produce a modification of that angle will instead result into a rigid rotation of the triplet (Figure 1 and Figure 2, panel 5b).

Specifically, the torque $\boldsymbol{\tau}_i$,

$$\boldsymbol{\tau}_i = \sum_j \boldsymbol{\tau}_{ij} \tag{8}$$

which is the total torque acting on $\boldsymbol{\mu}_i$ due to the interactions of $\boldsymbol{\mu}_i$ with all of the other dipoles $\boldsymbol{\mu}_j$, is expressed as the sum of two components:

$$\boldsymbol{\tau}_i = \boldsymbol{\tau}_{n,m} + \boldsymbol{\tau}_v \tag{9}$$

$\boldsymbol{\tau}_{n,m}$ is the projection of $\boldsymbol{\tau}_i$ on the plane $(\boldsymbol{n}, \boldsymbol{m})$, and $\boldsymbol{\tau}_v$ is the component along the $\boldsymbol{v}$ axis (Figure 2, panel 4). The $\boldsymbol{\tau}_v$ component, associated with a rotation around $\boldsymbol{v}$, affects ϑ only (Figure 2, panel 5a). Therefore, by inverting the biunivocal relationship between ϑ_i and γ_i (eq 4), such precession can be easily transformed into a bending force, $\boldsymbol{B}_i(\gamma_i)$ acting on the γ_i angle:

$$\boldsymbol{B}_i(\gamma_i) = \boldsymbol{\tau}_v \frac{\partial \vartheta_i}{\partial \gamma_i} \tag{10}$$

The $\boldsymbol{\tau}_{n,m}$ component is associated with a rotation of μ_i on a plane containing the $C_{\alpha,i}$, $C_{\alpha,i+1}$ direction, thus with a variation of the constrained angle ϕ. Therefore, this rotation of $\boldsymbol{\mu}_i$ has to be rigidly transferred to the triplet C_i (Figure 2, panel 5b).

In practice, interactions among dipoles result into rigid rototranslations of the triplets, produced by the combined action of $\boldsymbol{\tau}_{n,m}$, and forces $\boldsymbol{F}_{ij}$, and bending motions, under the action of $\boldsymbol{\tau}_v$ (Figure 2, panel 6).

Integration of the EoM. Our procedure rigorously transforms all forces acting on dipoles into forces acting on C_α beads (Figure 2); therefore, integration of the equations of motion is performed on C_α's only, without the introduction of additional degrees of freedom to the system or the need to impose any fictitious dipole dynamics. The generation of dipole geometries and the redistribution of forces from dipoles to beads constitute negligible additional computational costs to the simulations. In fact, they are constituted by simple operations, most of which (like the determination of internal reference systems and computation of the γ angles) are already routinely required for bending interactions and, thus, performed by standard MD operations. The only non-negligible cost is related to the dipole−dipole interaction part, which has to be added ex novo to the nonbonded interactions. Anyway, its cost is comparable to that of an additional nonbonded solute−solute interaction, which typically constitutes less than 1% of the total simulation cost per step, thus being acceptable for simulation-speedup criteria in CG simulations.

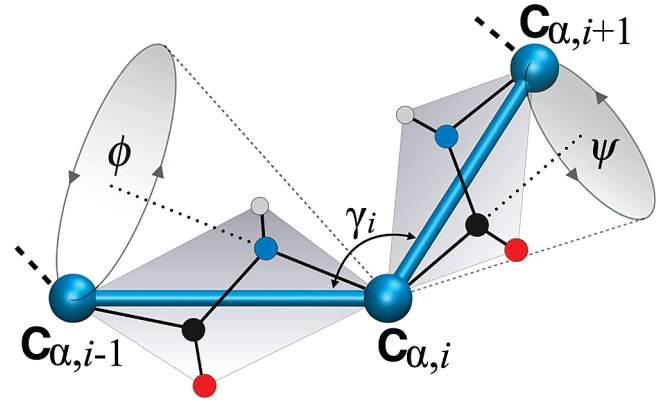

Figure 3. Relationship between the CG angle γ_i and the all-atom Ramachandran angles ϕ and Ψ. The all-atom structure is sketched in lines; the CG beads are represented in cyan balls-and-sticks. Rotation of any Ramachandran angle corresponds to a precession of the respective CG bead, resulting in a variation of the angle γ_i.

2.2. Bending Potential. Within the harmonic approximation for bending motion, only small oscillations of the angle γ formed by three consecutive backbone beads around a defined equilibrium value γ_0 are allowed. Instead, variations in such angles are related, in the respective all-atom representation, to structural variations of the Ramachandran torsional angles (Figure 3). Therefore, we decided to use a double-well potential to describe angular interactions. The angular potential function takes the form of a fourth degree polynomial function in the $(\gamma-\gamma_0)$ variable:

$$P^4(\gamma - \gamma_0) = \frac{1}{2}k_1(\gamma - \gamma_0)^2 - \frac{1}{3}k_2(\gamma - \gamma_0)^3 + \frac{1}{4}k_3(\gamma - \gamma_0)^4 \tag{11}$$

where γ_0 is 91°.

Such a kind of potential form for the CG bending term, originally proposed by Warshel and Levitt,[5] and recently successfully used in studies on CG modeling of HIV-1 protease,[22,37−39] allows the determination of two local minima, corresponding to the α-helical and β-strand basins. Thus, this allows for a more general description of CG interactions, without the need of specifying *biasing* secondary structure-dependent angular potentials. Rather, by modulating its shape, secondary structure propensities can be attributed to local portions of the amino acid (aa) sequence, thus making conformational transitions possible at finite energy costs. The two parameter sets chosen for simulations in this paper (see Table 1) are meant to reproduce the qualitative shape of a Boltzmann inversion of the statistical distributions coming

Table 1. Potential Energy Parameters[a]

structure	k_φ, n, δ	k_{corr}	γ_0	k1	k2	k3
α-helix	0.8, 3, 0.0	>0.4	1.588	100.0	295.3	218.0
β-sheet	0.8, 3, 0.0 and 2.6, 1, 0.0	0.6	1.588	70.0	308.5	302.2

[a] Different secondary structure propensities are modeled using the reported values. n = multiplicity of dihedral term; [γ_0], [δ] = rad; [k_φ], [k_{corr}] = kcal mol^{-1}; [k1] = kcal mol^{-1} rad^{-2}; [k2] = kcal mol^{-1} rad^{-3}; [k3] = kcal mol^{-1} rad^{-4}.

from triplets of amino acids with high α and high β propensities, respectively.

2.3. Dihedral Correlation Potential. The final modification of the molecular Hamiltonian consists of defining a correlation potential term for consecutive dihedrals:

$$V_{corr}(\varphi_i, \varphi_{i+1}) = k_{corr}(\varphi_i - \varphi_{i+1})^2 \quad (12)$$

This effective potential is meant to account for all of those steric and electrostatic interactions that would be present among those atoms that are represented by neighboring CG beads, and which result in a decreased rigidity of the chain. With such a potential, structures which are characterized by large variations in consecutive dihedral angles will pay for extra potential energy. In particular, the potential in eq 12 disfavors randomly collapsed structures, and it enhances the formation of ordered structures that show local conformational self-identity along the sequence. Rather, a minimal number of deformations (e.g., turns) can be allowed whenever they account for any favorable pairing of structural elements, like in β-hairpin or α-helical assemblies. Different values of k_{corr} can in principle be used for different portions of the chain, to represent local variations in its chemical-physical nature.

2.4. Choice of Simulation Parameters. Different parameter sets where chosen, as reported in Table 1, representing different physical natures of the peptide chain. In particular, we define increasing rigid polymers by progressively increasing the values of k_{corr}, and we modulate the α or β propensity by modulating the shape of the bending and torsional potentials (see Table 1). The van der Waals parameters were modeled as alanine beads of the MARTINI[17] potential (equilibrium distance, 5.28 Å; binding energy, 0.836 kcal mol^{-1}). Solvent−solute interactions were also taken from the MARTINI potential (equilibrium distance, 5.28 Å; binding energy, 0.741 and 1.195 kcal mol^{-1} for solute−solvent and solvent−solvent interactions, respectively). Dihedral potentials are in general expressed as a combination of multiple cosine terms. The dipole−dipole, bending, dihedral, and dihedral correlation potential terms account, on average, for ≈−0.29, +0.21, +0.13, and +0.02 kcal mol^{-1} per bead, respectively (data taken from simulation of the short helix, see below). Stretching distances were constrained to their equilibrium value (3.8 Å). In our simulations, interactions within consecutive dipoles were excluded. All interactions were screened by a relative global dielectric constant of 10.

2.5. Molecular Dynamics Simulations. Molecular dynamics simulations of different test systems were performed. In particular, we first tested the structural stability of 10-aa-long peptides in initial helical conformation and that of 25-aa-long peptides in initial antiparallel β-sheet conformation for different values of the potential parameters. This part of the work allowed us to define basins of stability for α-helices and β-sheet structural elements. Then, longer (25 to 30 aa) peptides in the helical conformation were studied, to investigate the capabilities of our model in predicting structural reorganizations of the polymer.

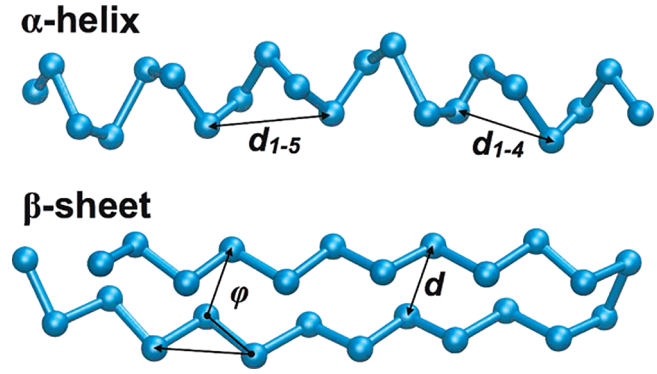

Figure 4. Structural parameters for α-helices and β-sheets. d_{1-5} and d_{1-4} are the distances between the helical pitches of five and four consecutive beads, respectively. d is the interstrand distance, and ϕ is the interstrand torsional angle.

The systems were first minimized by 2000 steps of conjugate gradients; then they were slowly heated up from 0 to 300 K in 10 ns and simulated in the canonical ensemble on the nanosecond to microsecond time scales. We tested the use of different timesteps, starting from 1 fs and progressively increasing them. We found our simulations to be consistent among themselves for timesteps up to 20 fs. For larger timesteps, problems in conservation of the energy start to arise, leading eventually to inconsistent simulations.

Our potential function was implemented in the LAMMPS suite of programs[67−69] for MD, all simulations were run with the same program.

3. Results

Molecular dynamics runs were performed on model systems representing fundamental secondary and supersecondary structure elements. In this way, we tested the capability of our new potential function of predicting such motifs as locally stable structures. We have compared our results to those obtained with the MARTINI force field,[17] which is, at present, one of the most successful and reliable CG force fields proposed in the literature.

Short α-Helices. We simulated helical peptides composed of 10 units for 1.5 μs. Helices are stable in a regime of k_{corr} > 0.4 kcal mol^{-1} rad^{-2} and for bending potentials which define a strong helical propensity (Table 1). The helical structures remain stable along all of the simulation times, with average values for the 1−4 and 1−5 pitches in the helical turn of $d_{1-4} = 5.2 \pm 0.2$ Å and $d_{1-5} = 5.9 \pm 0.5$ Å, respectively (Figure 4). The structural data are in good agreement with the experimental measurement as obtained from PDB structures (Table 2).

Simulations with different dihedral potential intensities show the same behavior and do not lead to significant

Table 2. Structural Parameters for α-Helices and β-Sheets in CG MD[a]

	present results	PDB exptl. values
α-helices	$d_{1-4} = 5.2 \pm 0.2$ $d_{1-5} = 5.9 \pm 0.5$	$d_{1-4} = 5.1 \pm 0.1$ $d_{1-5} = 6.2 \pm 0.1$
β-sheets	$d_{is} = 5.1 \pm 0.1$ $\varphi_{is} = 91 \pm 8$	$d_{is} = 4.8 \pm 0.1$ $\varphi_{is} = 90 \pm 6$

[a] Values for 1−4 and 1−5 helical pitches and interstrand distances (d_{1-4}, d_{1-5}, d_{is}, respectively) are reported in Å; values for interstrand torsional angles are given in degrees.

differences in their results. The stability of the structure is not affected by variations of the dihedral energy term. The structures fluctuate with an average rmsd of 1.2 ± 0.2 Å from the initial structure, which corresponds to a mobility similar to that found using the MARTINI force field.[17]

Antiparallel *β*-Sheets. Molecular dynamics simulations of two antiparallel strands kept their average structural parameters close to optimal values. In particular, the interstrand distance and orientation (Figure 4) are $d = 5.1 \pm 0.1$ Å and $\phi = 91 \pm 8°$. These relevant structural parameters are in very good agreement with experimental values (Table 2). The β-hairpin structure remains within a rmsd as small as 1.6 ± 0.2 Å from the initial native geometry. In particular, the structure shows conserved eclipsed configurations between corresponding C_α's in the two strands and the correct overall average parallel orientation of their planes (Table 2). β-sheet geometries are stabilized by dihedral potential terms with an absolute minimum at $\varphi = 180°$, which corresponds to defining a stiff rigidity for the chain. Structural features characterizing the β-sheet pair cannot be easily described using standard CG models with radial nonbonded interactions. In fact, radial potentials do not discriminate among all those configurations that keep a similar number of contacts between the two strands (e.g., structures with wrongly oriented strand planes or with coiled strands). This results in a disordered ensemble of structures, different from the native one, which are found during CG MD runs with the MARTINI-ff and present a higher rmsd (3.0 ± 0.3 Å). The only way of overcoming this problem would be to explicitly define contact potentials between the two strands, which would, on the other hand, bias the structure. In our model, the presence of dipole interactions among strands automatically introduces preferential directionality, thus, producing the right geometry without the need of external biases.

Long α-Helices. Single long α-helices with nonspecific aa sequences are usually not stable and tend to undergo conformational transitions leading to supersecondary structure elements like, for example, helix−coil−helix or β-hairpin motifs.[70] We have tested the behavior of our backbone potential by simulating a 25-aa-long chain starting from an α-helical conformation. Attributing parameters previously found to produce stable helices, we find that the initial structure corresponds to a local minimum of the energy, but MD simulations at 300 K univocally lead to rapid conformational transition, to form helix−loop−helix structures (Figure 5). Such transition is induced by the attractive coupling between the finite electrostatic dipoles associated with the two antiparallel helical portions formed when the initial structure breaks. Our protocol, in fact, intrinsically

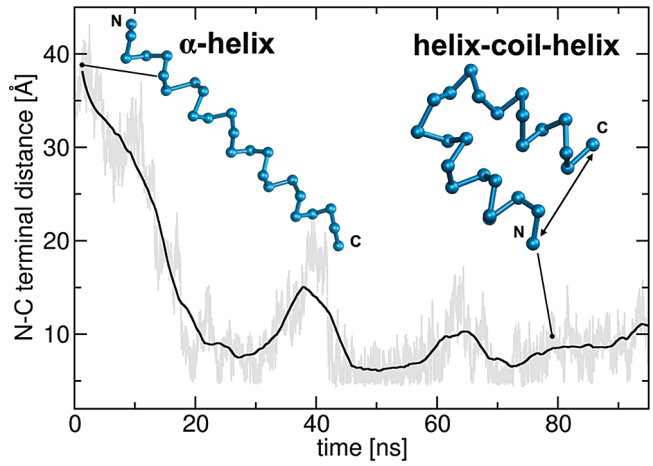

Figure 5. Helix−coil−helix formation. The N−C terminal distance as a function of the simulation time for a 25-aa-long chain. Averaged distances are represented with a solid black line. The insets show the conformational transition undergone by the initial helical structure into a supersecondary helix−coil−helix structure.

takes into account the interaction between macromolecular dipoles induced by backbone helical packing and therefore is able to correctly describe helical N-to-C polarity, which is a fundamental ingredient for the correct assembling of helical bundles. We note that in our simulations the terminal residues were not bringing any zwitterionic charge. Therefore, no instabilities due to attraction of the ends of the chain were present. The new structure is stable for values of k_{corr} of 0.4−0.5 kcal mol^{-1} rad^{-2}. For lower values of k_{corr}, instead, we find that, after formation of the helix−coil−helix motif, the two helices tend to break again, yielding collapsed structures that resemble a four-helix bundle (which cannot be fully formed due to the shortness of the peptide chain).

α-Helix−*β*-Hairpin Transition. Our potential was tested toward the possibility of describing α−β structural transitions. We used again the 25-aa-long α-helical peptide as a starting conformation and potential parameters, as in Table 1, which reproduce stable β-sheet structures. Within such potential parameters, the α-helix is only marginally stable and rapidly unfolds when heated up to 300 K. The peptide chain tends to stretch into an extended structure, which eventually bends to pair its two arms, leading to a β-hairpin structure (Figure 6). For smaller values of k_{corr}, the reorganization into a three-antiparallel-sheet structure is also observed.

***β*-Helix.** Finally, we tested the stability of parallel β-sheets, by modeling an 81-aa-long peptide in a β-helical conformation. We used the core structure of the antifreeze protein from *Choristoneura fumiferana* (pdb code: 1M8N[71]) as a template model (Figure 7). Parallel β-sheet pairings are not characterized by optimal dipole alignment, and therefore they fold only in the presence of other stabilizing factors, for example, side-chain packing in supersecondary assemblies. In fact, the template protein presents a compact hydrophobic core in the middle of the helix, which stabilizes the whole structure. Therefore, we modeled our helix by filling the inner space with nonbonded hydrophobic beads with a VdW radius of 2.0 Å. We find that, apart from an initial sudden shrinking

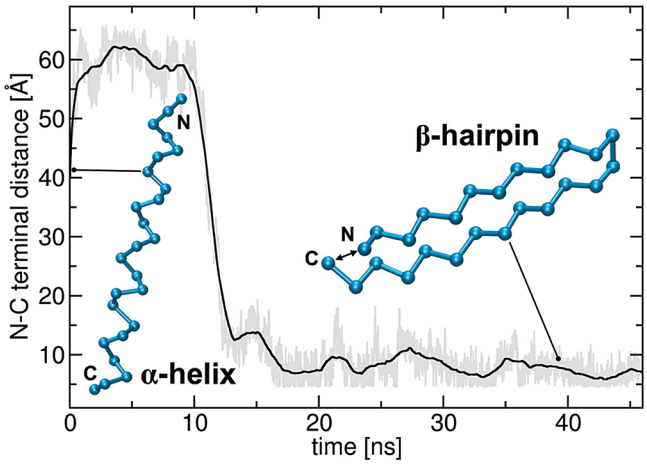

Figure 6. α-helix−β-hairpin transition. The N−C terminal distance as a function of the simulation time for a 25-aa-long chain. Averaged distances are represented with a solid black line. The insets show the conformational transition from the initial helical configuration to a β-hairpin structure.

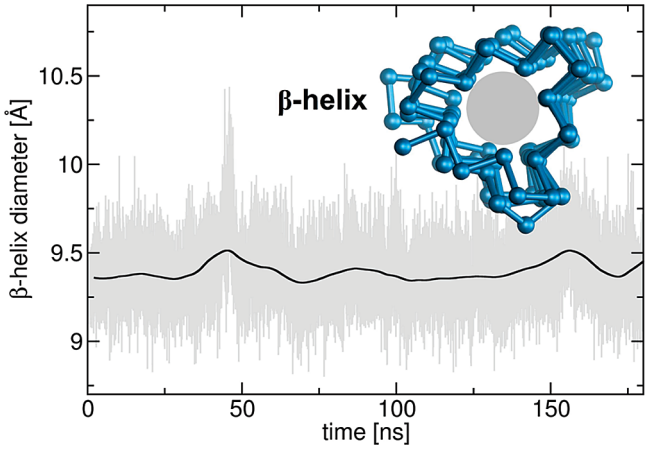

Figure 7. β-Helix structure. Diameter of a 81-aa-long peptide in a β-helical conformation as a function of the simulation time. The equilibrated structure is shown in the inset (PDB code: 1M8N).[71] The average diameter value remains stable during the simulation time.

of the structure, which has to be attributed to a nonoptimal space-filling by the hydrophobic-core mimicking beads, our model structure remains stable along the 160 ns of MD (Figure 7). Importantly, we see that the major deformations of the structure are associated with local distortions occurring at the vertices of the triangular assembly. The three β-sheet edges, on the contrary, keep all their contacts paired, as in the canonical structure of parallel β-sheets.

4. Discussion

Published data based on toy potentials and statistical mechanics modeling[72] showed that the minimal ingredients required to reproduce folded states of proteins are backbone flexibility and relative hydrophobicity of the chain, while H-bonding is required to ensure diversification, modulation, and ordered assembly of folded motifs. More generally, structuring of polymer chains can be associated with an induced coupling between consecutive Frenet terns, that is, of the relative orientations of the local reference systems associated with each bead.[73] While CG models based on molecular potentials (eq 1) can easily treat hydrophobicity (e.g., through relative solute−solute, solute−solvent, and solvent−solvent nonbonded interactions), both H-bonding and flexibility remain a more complex issue. In fact, H-bonds are intrinsically determined by the relative spatial orientation of objects (atoms) smaller than the resolution of the same CG models. The amino acid chain flexibility, at the all-atom level, is determined by the mixture of bonded and nonbonded energy terms which drive interactions of atoms in neighboring amino acids. This means that energy contributions for bonded interactions in CG should be able to represent such a mixture of both bonded and nonbonded interactions from the all-atom picture. In fact, in the CG picture, the pseudobonds which link the backbone beads are connecting objects that, in the all-atom picture, can be topologically distant. Therefore, simple harmonic approximations for bonded interactions, like those used in all-atom molecular Hamiltonians, can be too limiting for a correct description of the physics of the amino acid chain.

Our model potential for the protein backbone introduces an effective directional potential which mimics the presence of permanent electrostatic dipoles at peptide bond locations. Such directional interaction enables the reproduction of, at short-range, the local structure of contacts formed by the H-bonding network in all-atom systems, which is fundamental to defining secondary structure elements (Figure 4, Table 1). At the same time, it introduces long-range interactions which are able to drive the recognition and pairing of elements and, thus, formation of supersecondary structure assemblies, like in the case of the collapse of long helical structures into helix−coil−helix arrangements (Figure 5). Such assembly is favored by the presence of macromolecular dipoles associated to each helix, which pair in an energetically advantageous configuration. Such a feature is well represented by our model; in fact, within its representation, unspecific packing of backbone beads does not necessarily lead to an energy gain, and therefore it is not necessarily stabilized. This is confirmed by the results from our simulations, where we see the spontaneous assembly of ordered helix−turn−helix (Figure 5) and β-hairpin (Figure 6) motifs. In particular, our model predicts a very good relative orientation of two paired β-strands with respect to that predicted by potentials based on radial nonbonded interactions only (see Results section).

The improved performances of our model can be related to the intrinsic anisotropic character of the electrostatic potential of proteins. In fact, the nonbonded potential terms for CG beads have to effectively describe, among the others, the electrostatic interactions between the different groups of atoms they represent. As the electrostatic potential of a group of atoms, expressed as a function of a single center, is formally defined by its multipolar expansion, it is reasonable to think that a fully transferable radial potential would be particularly suited in those cases where CG beads map groups of atoms with a non-negligible total charge Q_{tot}, or in the assumption that characteristic times for multipole reorganiza-

tion are much faster than the typical times of CG dynamics. On the contrary, the protein backbone is characterized by the presence of strong permanent dipoles localized at peptide bonds between C_α atoms, an orientation that is not free but directly correlated to the local conformational geometry (Figure 3). As the conformation of the polymer is a piece of information explicitly present at the CG-resolution, directional nonbonded contributions due to the presence of permanent dipoles remain a relevant quantity to describe the physics of proteins.[74]

More sophisticated modeling for bonded terms than the harmonic approximation also turns out to be useful for the stabilization and modulation of secondary structure elements. In particular, we make use of a double-well potential for bending, representing the two principal stable basins in the Ramachandran plot. Moreover, we introduce a correlation potential energy term for consecutive dihedrals. These modifications ensure a direct control on the flexibility properties of the polymer. Specifically, such potential disfavors random orientations of the chain, which would eventually lead to collapsed structures, promoting structures characterized by conformational self-identity along the chain, like in extended (e.g., β-strands) or coiled structures (e.g., helical motifs). For parameters corresponding to highly flexible polymers (k_{corr} <0.4 kcal mol^{-1} rad^{-2}), random collapsed structures tend to appear. On the contrary, for high values of k_{corr}, combined with strong dihedral potentials of multiplicity $n = 1$ and $k_\varphi > 2.8$ kcal mol^{-1}, all initial structures tend to form single, long, stretched motifs. Our findings on the correlation between flexibility and structure stability are in very good qualitative agreement with the phase diagram reported in ref 72.

Our potential energy term for nonbonded interactions does not need to be defined according to the starting secondary structure; rather, it allows for structural reorganizations or deformations which can lead to energetically more favorable assemblies. This was seen in our simulations of long α-helical structures, where, starting from the same conformation, for two peptides with different defined secondary structure propensities, we observed the spontaneous formation of helix−coil−helix or β-hairpin motifs, respectively. This is also confirmed by the structural deformations seen in the β-helix simulations, where the strain of the structure due to nonoptimal modeling of the hydrophobic core remains localized at the ends of the β-sheets, keeping ordered elements in their place. We remark on the fact that we did not define different specific potentials for the beads in the sheets and in the turns; rather, the system itself finds it energetically convenient to pay for some strain in the turn regions, while keeping the strands paired, as expected. The possibility of explicitly defining bead-specific bending, dihedral, and correlation terms makes possible the future development of new CG force fields based on the proposed functional form, able to incorporate both local flexibility and secondary-structure propensity into the backbone potential.

The presented algorithm allows for a full description of dipole locations and orientations as functions of the coordinates of the C_α beads, and, thus, representation of their interactions with forces acting on the same C_α beads. Therefore, it does not introduce further degrees of freedom to the system, limiting the increase of computational time to the explicit evaluation of the dipole−dipole interactions. The possibility of reconstructing structural features from the C_α coordinates in proteins was already shown in the case of side-chain localization[75] and effectively used to improve the performance of CG network models.[76,77] The dipolar interactions, within our model, have to be considered as an effective potential which represents, through an explicit potential energy term, the energy gain in the formation of ordered contacts in a folded protein backbone. Its functional form is such that it allows it to be easily incorporated into any CG representation of proteins which make use of C_α beads. In this respect, the intensity of the dipoles, as well as that of the interactions in general, can be in principle rescaled or re-equilibrated to optimally match pre-existing CG models or the multiscaling criteria used to develop the original CG potential.

The use of a dipole that has a definition consistent with the corresponding permanent electrostatic dipole of a protein backbone has specific advantages. In fact, we had already shown that, within a very similar approach,[65] electrostatic potentials of proteins can be reproduced by CG structural information, in particular by accounting for the contribution only of backbone and side-chain permanent dipoles. Therefore, such multipolar definition of CG electrostatics can be helpful in describing more reliably protein−protein assemblies or protein−ligand recognition in multiscale frameworks.

5. Concluding Remarks

In conclusion, our protocol is able to introduce backbone dipole interactions in CG MD simulations of proteins, which, in turn, allows for an unbiased representation of stable secondary structure elements, as well as prediction of their dynamical arrangement into supersecondary structure assemblies. The proposed directional potential has a general form and can in principle be coupled to existing CG protocols (single or multibead), which retain structural information of the C_α trace of proteins. Together with modifications of the bonded energy terms, with respect to simplest harmonic approximations (in particular, of bending and dihedrals), our scheme constitutes a promising step toward the development of a more universal and transferable CG force field for proteins, not plagued with knowledge-based biases on the secondary structure. In particular, the directionality of the backbone structure is directly connected to the bending angle. Thus, the secondary structure propensity of amino acids, which is chemically encoded in the side-chain, can be elegantly controlled by using the backbone bending potential as an order parameter. The present results anticipate the development of a new CG force field able to take into account the intrinsic anisotropy of protein structures, leading to an improved description of the structural and dynamic properties of protein assemblies and networks.

Acknowledgment. The presented work was supported by the Swiss National Science Foundation (Grants No.

PP002_118930, 200021_122120). The authors wish to thank A. Giorgetti and E. Spiga for useful discussions.

References

(1) Karplus, M.; McCammon, J. *Nat. Struct. Biol.* **2002**, *9*, 646–652.

(2) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–725.

(3) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T.; Laughton, C.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817–3829.

(4) Klein, M. L.; Shinoda, W. *Science* **2008**, *321*, 798–800.

(5) Levitt, M.; Warshel, A. *Nature* **1975**, *253*, 694–698.

(6) Go, N.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 535–542.

(7) Tanaka, S.; Sheraga, H. A. *Macromolecules* **1976**, *9*, 945–950.

(8) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J, Phys. Chem. B* **2001**, *105*, 4464–4470.

(9) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J, Phys. Chem. B* **2001**, *105*, 9785–9792.

(10) Saiz, L.; Klein, M. L. *Acc. Chem. Res.* **2002**, *35*, 482–489.

(11) Shinoda, W.; DeVane, R.; Klein, M. L. *Soft Matter* **2008**, *4*, 2454–2462.

(12) Shinoda, W.; Devane, R.; Klein, M. L. *Mol. Simulat.* **2007**, *33*, 27–36.

(13) Marrink, S. J.; de Vries, A.; Mark, A. *J. Phys. Chem. B* **2004**, *108*, 750–760.

(14) Brannigan, G.; Lin, L. C. L.; Brown, F. L. H. *Eur. Biophys. J.* **2006**, *35*, 104–124.

(15) Shi, Q.; Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2006**, *110*, 15045–15048.

(16) Lu, L.; Voth, G. A. *J. Phys. Chem. B* **2009**, *113*, 1501–1510.

(17) Monticelli, L.; Kandasamy, S.; Periole, X.; Larson, R.; Tieleman, D.; Marrink, S. J. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.

(18) Arkhipov, A.; Freddolino, P.; Imada, K.; Namba, K.; Schulten, K. *Biophys. J.* **2006**, *91*, 4589–4597.

(19) DeVane, R.; Shinoda, W.; Moore, P. B.; Klein, M. L. *J. Chem. Theory Comput.* **2009**, *5*, 2115–2124.

(20) Basdevant, N.; Borgis, D.; Ha-Duong, T. *J. Phys. Chem. B* **2007**, *111*, 9390–9399.

(21) Voltz, K.; Trylska, J.; Tozzini, V.; Kurkal-Siebert, V.; Langowski, J.; Smith, J. *J. Comput. Chem.* **2008**, *29*, 1429–1439.

(22) Trylska, J.; Tozzini, V.; Chang, C.; McCammon, J. *Biophys. J.* **2007**, *92*, 4179–4187.

(23) Trovato, F.; Tozzini, V. *J. Phys. Chem. B* **2008**, *112*, 13197–13200.

(24) Villa, E.; Balaeff, A.; Schulten, K. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6783–6788.

(25) Ayton, G. S.; Voth, G. A. *Curr. Opin. Struct. Biol.* **2009**, *19*, 138–144.

(26) Arkhipov, A.; Yin, Y.; Schulten, K. *Biophys. J.* **2008**, *95*, 2806–2821.

(27) Treptow, W.; Marrink, S. J.; Tarek, M. *J. Phys. Chem. B* **2008**, *112*, 3277–3282.

(28) Yelimov, S.; van der Giessen, E.; Onck, P.; Marrink, S. J. *Biophys. J.* **2008**, *94*, 2994–3002.

(29) Marrink, S. J.; Mark, A. *J. Am. Chem. Soc.* **2003**, *125*, 11144–11145.

(30) Risselada, H.; Marrink, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17367–17372.

(31) Bond, P.; Wee, C.; Sansom, M. S. P. *Biochemistry* **2008**, *47*, 11321–11331.

(32) Carpenter, T.; Bond, P.; Khalid, S.; Sansom, M. S. P. *Biophys. J.* **2008**, *95*, 3790–3801.

(33) Marrink, S. J.; de Vries, A. H.; Tieleman, D. P. *BBA Biomembr.* **2009**, *1788*, 149–168.

(34) Balali-Mood, K.; Bond, P. J.; Sansom, M. S. P. *Biochemistry* **2009**, *48*, 2135–2145.

(35) Klingelhoefer, J. W.; Carpenter, T.; Sansom, M. S. P. *Biophys. J.* **2009**, *96*, 3519–3528.

(36) Sherwood, P.; Brooks, B. R.; Sansom, M. S. P. *Curr. Opin. Struct. Biol.* **2008**, *18*, 630–640.

(37) Tozzini, V.; McCammon, J. *Chem. Phys. Lett.* **2005**, *413*, 123–128.

(38) Tozzini, V.; Trylska, J.; Chang, C.; McCammon, J. *J. Struct. Biol.* **2007**, *157*, 606–615.

(39) Tozzini, V.; McCammon, J. *Protein Sci.* **2004**, *13*, 194–194.

(40) Arkhipov, A.; Freddolino, P.; Schulten, K. *Structure* **2006**, *14*, 1767–1777.

(41) Delle Site, L.; Abrams, C. F.; Alavi, A.; Kremer, K. *Phys. Rev. Lett.* **2002**, *89*, 156103.

(42) Villa, E.; Balaeff, A.; Mahadevan, L.; Schulten, K. *Multiscale Modell. Simulat.* **2004**, *2*, 527–553.

(43) Neri, M.; Anselmi, C.; Cascella, M.; Maritan, A.; Carloni, P. *Phys. Rev. Lett.* **2005**, *95*, 218102.

(44) Lyman, E.; Ytreberg, F. M.; Zuckerman, D. M. *Phys. Rev. Lett.* **2006**, *96*, 028105.

(45) Lyman, E.; Zuckerman, D. M. *J. Chem. Theory. Comput.* **2006**, *2*, 656–666.

(46) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 2469–2473.

(47) Izvekov, S.; Voth, G. A. *J. Chem. Theory Comput.* **2006**, *2*, 637–648.

(48) Ayton, G. S.; Voth, G. A. *J. Struct. Biol.* **2007**, *157*, 570–578.

(49) Ensing, B.; Nielsen, S. O.; Moore, P. B.; Klein, M. L.; Parrinello, M. *J. Chem. Theory Comput.* **2007**, *3*, 1100–1105.

(50) Heyden, A.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 217–221.

(51) Zacharias, M. *J. Chem. Theory Comput.* **2008**, *4*, 477–487.

(52) Ayton, G.; Noid, W.; Voth, G. A. *Curr. Opin. Struct. Biol.* **2007**, *17*, 192–198.

(53) Cascella, M.; Dal Peraro, M. *CHIMIA* **2009**, *63*, 14–18.

(54) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–150.

(55) Noid, W.; Chu, J.; Ayton, G.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. *J. Chem. Phys.* **2008**, *128*, 244114.

(56) Noid, W.; Liu, P.; Wang, Y.; Chu, J.; Ayton, G.; Izvekov, S.; Andersen, H.; Voth, G. A. *J. Chem. Phys.* **2008**, *128*, 244115.

(57) Bereau, T.; Deserno, M. *J. Chem. Phys.* **2009**, *130*, 235106.

(58) Maupetit, J.; Tuffery, P.; Derremaux, P. *Proteins* **2007**, *69*, 394–408.

(59) Gopal, S. M.; Mukherjee, S.; Cheng, Y.-M.; Feig, M. *Proteins* **2009**, DOI: 10.1002/prot.22645.

(60) Liwo, A.; Oldziej, S.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 849–873.

(61) Liwo, A.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Oldziej, S.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 874–887.

(62) Liwo, A.; Oldziej, S.; Czaplewski, C.; Kozlowska, U.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 9421–9438.

(63) Liwo, A.; Khalili, M.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 2362–2367.

(64) Májek, P.; Elber, R. *Proteins* **2009**, *76*, 822–836.

(65) Cascella, M.; Neri, M. A.; Carloni, P.; Dal Peraro, M. *J. Chem. Theory Comput.* **2008**, *4*, 1378–1385.

(66) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberg, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Hochschulverlag AG an der ETH Zurich: Zurich, Switzerland, 1996.

(67) LAMMPS Molecular Dynamics Simulator. http://lammps.sandia.gov/ (accessed November 25th, 2009).

(68) Plimpton, S. J. *J. Comp. Phys.* **1995**, *117*, 1–19.

(69) Plimpton, S. J.; Pollock, R.; Stevens, M. Particle-Mesh Ewald and rRESPA for Parallel Molecular Dynamics Simulations. In *Proceedings of the Eighth SIAM Conference on Parallel Processing for Scientific Computing* Minneapolis, 1997; SIAM, ISBN 0-89871-395-1.

(70) Forood, B.; Perezpaya, E.; Houghten, R.; Blondelle, S. *Biochem. Biophys. Res. Commun.* **1995**, *211*, 7–13.

(71) Leinala, E.; Davies, P.; Doucet, D.; Tyshenko, M.; Walker, V.; Jia, Z. *J. Biol. Chem.* **2002**, *277*, 33349–33352.

(72) Hoang, T. X.; Trovato, A.; Seno, F.; Banavar, J. R.; Maritan, A. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 7960–7964.

(73) Banavar, J. R.; Cieplak, M.; Hoang, T. X.; Maritan, A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6900–6903.

(74) Warshel, A.; Russell, S. *Q. Rev. Biophys.* **1984**, *17*, 283–422.

(75) Park, B.; Levitt, M. *Proteins* **1996**, *258*, 367–392.

(76) Micheletti, C.; Carloni, P.; Maritan, A. *Proteins* **2004**, *55*, 635–645.

(77) Neri, M.; Cascella, M.; Micheletti, C. *J. Phys., Condens. Matter* **2005**, *17*, S1581–S1593.

CT900457Z