Assessing the influence of electrostatic schemes on molecular dynamics simulations of secondary structure forming peptides

2006 J. Phys.: Condens. Matter 18 S329

(http://iopscience.iop.org/0953-8984/18/14/S15)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **18** (2006) S329–S345 doi:10.1088/0953-8984/18/14/S15

Assessing the influence of electrostatic schemes on molecular dynamics simulations of secondary structure forming peptides

Luca Monticelli[1,2], **Carlos Simões**[3], **Laura Belvisi**[1,2] **and Giorgio Colombo**[1,3,4]

[1] Centre for Biomolecular Interdisciplinary Studies and Industrial Applications, University of Milano, I-20131 Milano, Italy
[2] Department of Organic and Industrial Chemistry, University of Milano, Via Venezian 21, 20133 Milano, Italy
[3] Istituto di Chimica del Riconoscimento Molecolare, CNR, Via Mario Bianco 9, 20131 Milano, Italy

E-mail: giorgio.colombo@icrm.cnr.it

Received 8 August 2005, in final form 15 September 2005
Published 24 March 2006
Online at stacks.iop.org/JPhysCM/18/S329

Abstract

Electrostatic interactions play a fundamental role in determining the structure and dynamics of biomolecules in solution. However the accurate representation of electrostatics in classical mechanics based simulation approaches such as molecular dynamics (MD) is a challenging task. Given the growing importance that MD simulation methods are taking on in the study of protein folding, protein stability and dynamics, and in structure prediction and design projects, it is important to evaluate the influence that different electrostatic schemes have on the results of MD simulations. In this paper we performed long timescale simulations (500 ns) of two peptides, beta3 and RN24 forming different secondary structures, using for each peptide four different electrostatic schemes (namely PME, reaction field correction, and cut-off schemes with and without neutralizing counterions) for a total of eight 500 ns long MD runs. The structural and conformational features of each peptide under the different conditions were evaluated in terms of the time dependence of the flexibility, secondary structure evolution, hydrogen-bonding patterns, and several other structural parameters. The degree of sampling for each simulation as a function of the electrostatic scheme was also critically evaluated. Our results suggest that, while in the case of the short peptide RN24 the performances of the four methods are comparable, PME and RF schemes perform better in maintaining the structure close to the native one for the β-sheet peptide beta3, in which long range contacts are mostly responsible for the definition of the native structure.

[4] Author to whom any correspondence should be addressed.

1. Introduction

Molecular dynamics (MD) simulations are widely used in the study of the structure and function of biopolymers. The results of simulations depend on a number of factors ranging from the quality of the force field, to the solvent model, to the extent of sampling, to the timescales accessed.

The treatment of electrostatic interactions, in particular, is a fundamental issue in all-atom molecular dynamics (MD) simulations of biological systems. The proper calculation of electrostatic forces has in fact a critical influence on the theoretical determination of biomolecular properties. Moreover, the computation of these forces represents the most demanding part of MD simulations in terms of computer power. As a consequence, several approximations have been introduced to allow the study of large inhomogeneous systems consisting of a solute (a protein or DNA molecule) and thousands of solvent molecules. The different methods range from simple neglect of charge–charge interactions beyond a certain cut-off distance, to the introduction of reaction fields (RF) [1], to the use of lattice–sum methods (particle mesh Ewald [2], particle–particle–particle mesh (P3M) [3]) that entail the calculation of all long range interactions within the system.

Considering the importance that MD simulations have been taking on in modern molecular biophysics in general, and in particular in the study of protein related problems such as folding, misfolding, and design, it is of interest to assess the influence of different electrostatic schemes on the calculation of the structural and dynamical properties of experimentally characterized systems. In this context, several authors have studied ideal polyalanine models, or small compact proteins for limited timescales [4–6].

In this work, we have decided to run several long timescale simulations (500 ns each) on two different peptide sequences, namely beta3 and RN24 (figure 1), representative of the most diffused secondary structure motifs. beta3, a 20-residue long peptide whose sequence is TWIQNGSTKWYQNGSTKIYT, was designed by de Alba and co-workers to fold into a three-stranded β-sheet and its structure in solution was probed by NMR analysis [7]. RN24 is a 13-residue long peptide, with sequence AETAAAKFLRAHA, able to populate bent and α-helical conformations in aqueous solution, as determined by NOE analysis [8]. The amount of α-helix is around 15% of the overall accessible conformations. In a previous paper on beta3, we reported on the importance of accessing long simulation times to accurately reproduce experimental NOE constraints and investigated the influence of PME or cut-off use on the stability of the native three-stranded β-sheet structure [9]. In this new study, we add to the previous analysis of beta3 the use of the RF method, and extend our investigations to RN24, in which conformational interconversion times are of the order of tens to hundreds of nanoseconds. In particular, we ran four different simulations of 500 ns each for both beta3 and RN24, using four different electrostatic approaches: the particle mesh Ewald (PME) [2], the reaction field (RF) [1], the simple cut-off truncation with and without neutralizing counterions. To our knowledge, this is the first systematic study of the effects of electrostatic approaches on experimentally characterized systems over biologically relevant timescales.

1.1. The electrostatic schemes

In this subsection, we will briefly review the electrostatic schemes used in the present study. Interactions within a particular biomolecular system are typically described through a potential function [10] such as $V(\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_N})$, where the $\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_N}$ vectors represent the positions of atoms in space:

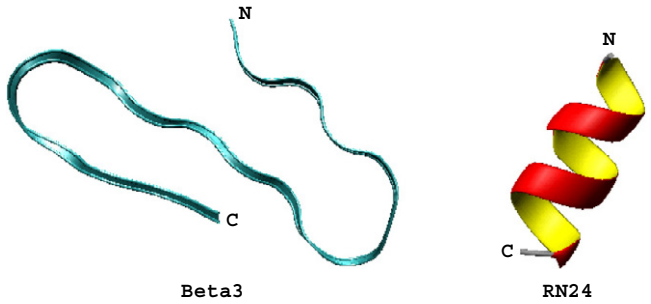

Figure 1. The native conformations of the two peptides analysed in this study beta3 and RN24. (This figure is in colour only in the electronic version)

$$V(\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_N}) = V_{\text{bond}} + V_{\text{non-bond}} = \sum_{\text{bonds}} \tfrac{1}{4} k_b (b^2 - b_0^2)^2 + \sum_{\text{angles}} \tfrac{1}{2} k_\theta (\cos(\theta) - \cos(\theta_0))^2 + \sum_{\text{impr.dihed.}} \tfrac{1}{2} k_\zeta (\zeta - \zeta_0)^2 + \sum_{\text{dihedrals}} K_\varphi |1 + \cos(n\varphi + \varphi_0)| + \sum_{(i,j)} \left[\frac{C_{12}(i,j)}{\mathbf{r}_{ij}^{12}} - \frac{C_6(i,j)}{\mathbf{r}_{ij}^{6}} \right] + \sum_{(i,j)} \frac{q_i q_j}{4\pi\varepsilon_0 \mathbf{r}_{ij}}. \quad (1)$$

The summation runs in principle over all pairs of atoms with sequence numbers i and j. The last term, in particular, represents the Coulomb description of interactions among particles bearing a (partial) charge. The problem here is that Coulombic interactions are long ranged, implying a high number of interaction evaluations, which in turn decrease dramatically the computational efficiency of MD simulations. To overcome this problem, the schemes quoted in the introduction have been introduced and will be briefly discussed here.

Cut-off methods. In these approaches, non-bonded electrostatic interactions are simply ignored beyond a certain cut-off. In order to calculate these intrinsically long ranged interactions with sufficient accuracy, a very long cut-off radius of at least 1.5 nm ought to be used. However, such a range would be very expensive computationally. That is why alternative cut-off schemes have been developed. One of these is the twin-range method [11], in which electrostatic interactions are evaluated at every integration step in a sphere of 0.8 nm around the particle (in the typical GROMOS force field approach), and every nth time step in a second larger spherical section comprising the area between the radii of 0.8 and 1.4 nm from the particle. This scheme, despite being extremely simplified, partly avoids the significant artefacts which straight truncation may introduce. This scheme has been used with a high degree of success in several applications, e.g. peptide folding and molecular recognition studies [12–14].

Reaction field. In the RF method each charge is surrounded by a cut-off sphere, within which the interactions with other particles are described explicitly [1]. The world outside the sphere is treated as a homogeneous dielectric medium with a certain permittivity and ionic strength, typical of the particular solvent and conditions that one wants to simulate. The Poisson–Boltzmann approach is then used to evaluate the forces from the continuum on the

particle. This long range force is then added to the short range one calculated explicitly. The electrostatic contribution can thus be written as

$$V = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_1}\left(\frac{1}{r_{ij}} - \frac{0.5C_{\mathrm{rf}}r_{ij}^2}{R_{\mathrm{rf}}^3} - \frac{1-0.5C_{\mathrm{rf}}}{R_{\mathrm{rf}}}\right) \tag{2}$$

where ϵ_1 equals 1. C_{rf} is the coefficient governing the size of the reaction field force:

$$C_{\mathrm{rf}} = \frac{(2\epsilon_1 - 2\epsilon_2)(1+\kappa R_{\mathrm{rf}}) - \epsilon_2(\kappa R_{\mathrm{rf}})^2}{(\epsilon_1 - 2\epsilon_2)(1+\kappa R_{\mathrm{rf}})\nu\epsilon_2(\kappa R_{\mathrm{rf}})^2} \tag{3}$$

where ϵ_2 is the relative dielectric permittivity of the electrostatic continuum and κ is an inverse Debye screening length outside the reaction field cut-off radius. If ϵ_2 equals 1 and κ equals 0, the reaction field terms are zero. If ϵ_2 is very large and κ equals 0, the cut-off spheres are surrounded by a highly conducting medium, which reduces the pair forces to 0 at r_{ij} equal R_{rf}.

This procedure results in a slightly higher computational expense, but improvements in simulation results compared to simple cut-off methods have been observed by several authors [6].

Particle mesh Ewald method. Finally, in recent years, the Ewald and other related mesh methods (PME, P3M etc) have been gaining more and more importance. They, and the particle mesh Ewald (PME) summation in particular, have been used in membrane and lipid simulations, in protein folding, and in nucleic acid simulations where the treatment of the highly charged molecules is of fundamental importance [15–18]. Ewald methods were initially developed for perfectly periodic crystalline systems, while at present they are being used for liquid and inhomogeneous systems. The enforcement of periodicity on non-crystalline systems has been shown to introduce artefacts if the simulation box is not large enough.

The total electrostatic energy of N particles and the periodic images are given by

$$V = \frac{1}{2}\sum_{\mathbf{n}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{4\pi\epsilon_0|\mathbf{r_{ij}}+\mathbf{n}|} \tag{4}$$

where $\mathbf{n} = (n_x, n_y, n_z)$ is the box index vector, and $i = j$ terms should be omitted when $(n_x, n_y, n_z) = (0, 0, 0)$. The simplest way to deal with this problem is to set a *cut-off* distance.

In Ewald summation, one starts from the exact treatment described by equation (4). The sum over n in equation (4) is conditionally convergent, but very slow. The idea is to convert the single slowly converging sum into two quickly converging terms and a constant term. The Ewald sum considers each point charge as surrounded by a Gaussian charge of opposite sign with the form

$$\rho_i(\mathbf{r}) = \frac{q_i\alpha^3}{\pi^{\frac{3}{2}}}\exp(-\alpha^2 r^2) \tag{5}$$

and in this way the summation over the real space becomes

$$V = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{|\mathbf{n}|=\mathbf{0}}^{\infty}\frac{q_i q_j}{4\pi\epsilon_0}\frac{\mathrm{erfc}(\alpha|\mathbf{r_{ij}}+\mathbf{n}|)}{|\mathbf{r_{ij}}+\mathbf{n}|} \tag{6}$$

where the erfc is the *complementary error function*:

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty}\exp(-t^2)\,\mathrm{d}t. \tag{7}$$

The trick is that equation (6) converges rapidly, with a rate that depends upon the width of the cancelling Gaussian distribution. A second charge distribution is then added to counteract the neutralizing distribution:

$$V = \frac{1}{2}\sum_{k\neq 0}\sum_{i=0}^{N}\sum_{j=1}^{N}\frac{1}{\pi L^3}\frac{q_i q_j}{4\pi\epsilon_0}\frac{4\pi^2}{k^2}\exp\left(-\frac{k^2}{4\alpha^2}\right)\cos(\mathbf{k}\cdot\mathbf{r_{ij}}). \tag{8}$$

This summation is performed in the reciprocal space. The complete Ewald sum requires an additional correction, known as the self-energy correction, which arises from a Gaussian acting on its own site, and is constant:

$$V = -\frac{\alpha}{\sqrt{\pi}}\sum_{k=1}^{N}\frac{q_k^2}{4\pi\epsilon_0}. \tag{9}$$

The final expression consists thus of the summation of these three terms (plus, if needed, a fourth that describes the correction for the surrounding medium of the simulation boxes).

This method is computationally quite expensive to implement; thus the particle mesh Ewald (PME) [2] method was developed to improve the performance of the reciprocal sum. Instead of directly summing wavevectors, the charges are assigned to a grid; this grid is then Fourier transformed and the reciprocal energy term is obtained by a single summation over the grid in k-space. The potential at the grid points is calculated by inverse transformation, and by using interpolation factors one get the forces on each atom. PME is in any case twice as expensive as cut-off schemes and 1.5 times more expensive than RF.

2. Materials and methods

The starting structure for the simulations of beta3 was one of the NMR-derived structures [7] (kindly provided by Professor Jimenez), with no violation of experimental restraints higher than 0.02 nm. The peptide was protonated to give a zwitterionic form (with N-terminal NH_3^+ and C-terminal COO^- groups), and the total charge was +2, due to the presence of two lysine residues. For RN24, the starting structure was a totally extended one and the peptide was capped with an acetyl group at the N terminal and an amide group at the C terminal [8].

All simulations were carried out using the GROMACS package version 3.1 [19, 20], using the GROMOS96 43A1 force field [10]. All bond lengths were constrained to their equilibrium values, using the SETTLE algorithm [21] for water and the LINCS algorithm [22] for all other bonds. The aromatic rings in Tyr and Trp side chains were kept rigidly planar using dummy atom constructions, according to a published procedure [23], in order to remove degrees of freedom from the system. According to Berendsen and co-workers [23], removal of these degrees of freedom from the system allows the time step to be increased to 7 fs with negligible influence on the thermodynamical and dynamical properties of the system. In our simulations, the time step was 2 fs during solvent equilibration (initial 50 ps) and 5 fs in the following steps; the neighbour list for the calculation of non-bonded interactions was updated every ten time steps in the first case, and every four in the latter.

Each peptide was solvated with water in a dodecahedral box large enough to contain 1.2 nm of solvent around the peptide. The simple point charge (SPC) water model was used [24].

Four simulations were carried out for each peptide, using the same starting structures; the box size and simulation parameters were identical except for the treatment of electrostatic interactions. Each simulation is labelled according to the name of the peptide and the electrostatic scheme used. The two simulations using PME are thus labelled B3-PME and RN24-PME, the two using the reaction field are labelled B3-RF and RN24-RF, the two cut-off simulations without counterions (*vide infra*) are labelled B3-NOION and RN24-NOION, and finally the two simulations using the cut-off scheme with counterions are labelled B3-ION and RN24-ION.

Table 1. Summary of the simulations performed for beta3, differences in set-up, and structural features of the peptide conformations sampled in the four simulations. The root mean square deviation (RMSD) was determined using Cα atoms of the peptide's core (3–18 segment); the radius of gyration (R_g) was calculated for all peptide atoms.

System	Electrostatics	No of H_2O molecules	No of Cl^- ions	RMSD (nm)	R_g (nm)
B3-PME	PME (0.9)[a]	2679	2	0.259	0.752
B3-RF	RF[a]	2679	2	0.443	0.721
B3-NOION	Cut-off (0.8/1.4)[a]	2683	0	0.503	0.701
B3-ION	Cut-off (0.8/1.4)[a]	2681	2	0.565	0.766

[a] Cut-off distances in nanometres.

Table 2. Summary of the simulations performed for RN24, differences in set-up, and structural features of the peptide conformations sampled in the four simulations. The root mean square deviation (RMSD) was determined using Cα atoms of the peptide's core (2–12 segment); the radius of gyration (R_g) was calculated for all peptide atoms.

System	Electrostatics	No of H_2O molecules	No of Cl^- ions	RMSD (nm)	R_g (nm)
RN24-PME	PME (0.9)[a]	2894	1	0.387	0.677
RN24-RF	RF[a]	2894	1	0.437	0.676
RN24-NOION	Cut-off (0.8/1.4)[a]	2338	0	0.408	0.654
RN24-ION	Cut-off (0.8/1.4)[a]	2388	1	0.394	0.675

[a] Cut-off distances in nanometres.

In the PME, RF, and cut-off simulations with counterions, the appropriately charged counterions necessary to ensure electroneutrality of the system replaced two water molecules to yield an electrically neutral system. The water molecules to be replaced were chosen randomly, using the GENION algorithm (included in the GROMACS package).

The set-up of all simulations is summarized in tables 1 and 2.

In the simulations using the particle mesh Ewald (PME) method [2] the simulations were carried out as follows: the real space interactions were evaluated using a 0.9 nm cut-off, and the reciprocal space interactions were evaluated on a 0.12 nm grid with fourth-order spline interpolation. A twin-range cut-off of 0.8–1.4 nm was used for the calculation of Lennard-Jones interactions. The relative accuracy of direct/reciprocal space is controlled by the parameter ewald_rtol (using GROmacs methodology) which is set to 1×10^{-5}.

In the reaction field (RF) calculations [1], interactions beyond a cut-off of 1.4 nm were taken into account by using the reaction field correction described in equations (2) and (3), assuming an electrostatic continuum with the relative dielectric permittivity calculated for SPC water ($\epsilon_2 = 54.0$).

In the cut-off simulations, a twin-range cut-off of 0.8–1.4 nm was used for both Coulombic and Lennard-Jones interactions. The cut-off values are the same as those used for the parametrization of the GROMOS96 43A1 force field [10]. In all cases, periodic boundary conditions were used.

The peptide, the water, and the counterions were coupled separately to a temperature bath at 300 K with $\tau_T = 0.1$ ps using the Berendsen algorithm [25]. The pressure was kept at 1 bar using weak pressure coupling with $\tau_P = 1.0$ ps [25].

The systems were energy minimized with a steepest descent method for 5000 steps. In all the simulations the solvent was equilibrated in a 50 ps MD run with position restraints on the peptide. The force constant on the peptide atoms was 1000 kJ mol^{-1} nm^{-2}. The solvent equilibration run was followed by another 100 ps run without position restraints on the peptide, in which all atoms were given an initial velocity obtained from a Maxwellian distribution at

the desired initial temperature. Each production run, after equilibration, was 500 ns long, and peptide structures were sampled every 1 ps. To our knowledge, the present simulations are among the longest carried out on this kind of system using an all-atom representation of the solvent.

Conformational clustering analysis was performed on a subset of the conformations sampled in the trajectory, containing 25 000 structures (taken at 20 ps intervals), using the method described by Daura and co-workers [12] (as implemented in GROMACS). The positional root mean square deviation (RMSD) calculated on the backbone atoms of residues 3–18 for beta3 and 2–12 for RN24 was used as a similarity criterion, and a cluster radius of 0.1 nm was chosen.

Secondary structure assignments were based on the DSSP algorithm [26]. The graphical representations of the peptides were realized with the program MOLMOL [27].

All calculations were performed on clusters of six PCs (dual processor AMD Athlon XP 1800+). The speed of the calculation is about one CPU hour per nanosecond of simulation for PME, 0.7 h ns^{-1} of simulation for RF, and 0.5 h ns^{-1} for the cut-off simulations.

3. Results

Several parameters were analysed to characterize the behaviour of the peptides under different simulation conditions.

3.1. Structural parameters

We first concentrated on the parameters which are typically used in the analysis of MD simulations. The time evolution of the root mean square deviations (RMSD) of the C_α atoms from the starting structure is reported in figures 2(a)–(d) for beta3 and in figures 2(e)–(h) for RN24. In the case of beta3 simulations, B3-PME and B3-RF reach a structural equilibrium with RMSD values around 0.3 and 0.38 nm respectively. Interestingly, B3-RF takes about 150 ns to equilibrate, but after reaching the plateau, it shows a stability comparable to those of its PME counterpart. In the case of the two beta3 cut-off simulations, the RMSD values immediately diverge to around 0.6 nm and display higher fluctuations than in the previous case. The situation changes dramatically in the case of the simulations of the α-helical forming peptide RN24 (figures 2(e)–(h)). Despite being characterized by different profiles, all the graphs show comparable patterns of RMSD fluctuations, characteristic of a system which is undergoing multiple conformational interconversions. In particular, it is interesting to observe this trend for the RN24-PME and RN24-RF simulations, in contrast to the apparent higher conformational stability observed for B3-PME and B3-RF.

These preliminary observations are confirmed by the analysis of the residue based root mean square fluctuation (RMSF) calculations, which are generally used to investigate conformational flexibility. The simulations of beta3 using PME and RF indicate similar low flexibility values for all residues, while the two simulations employing the plain cut-off method with or without the presence of counterions display globally higher conformational flexibilities (see figure 3(a)). In contrast, the conformational flexibility of the RN24 sequence does not display such a well defined dependence on the electrostatic calculation method used (see figure 3(b)).

The time evolution of the secondary structure for the two peptides is reported in figures 4 and 5. B3-PME and B3-RF show that the NMR-determined three-stranded β-sheet structure is the dominant one during MD (see figures 4(a) and (b)). It is worth noting that in B3-RF the β-hairpin structure covering residues 10 to 18 is lost for more than 130 ns and then

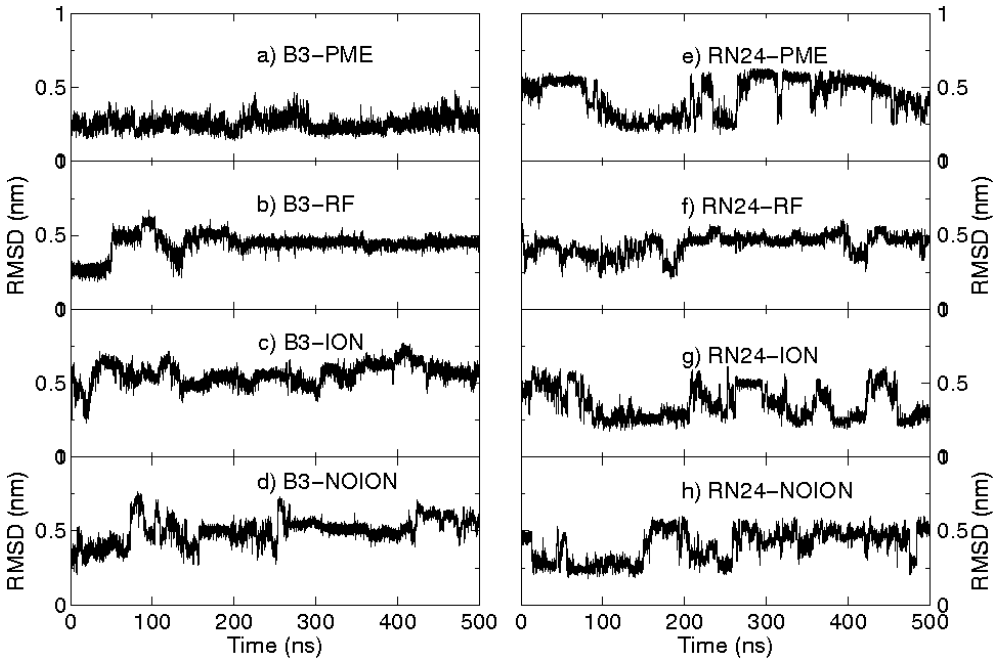

Figure 2. The time evolution of the positional root mean square deviation for beta3 ((a)–(d)) and RN24 ((e)–(h)).

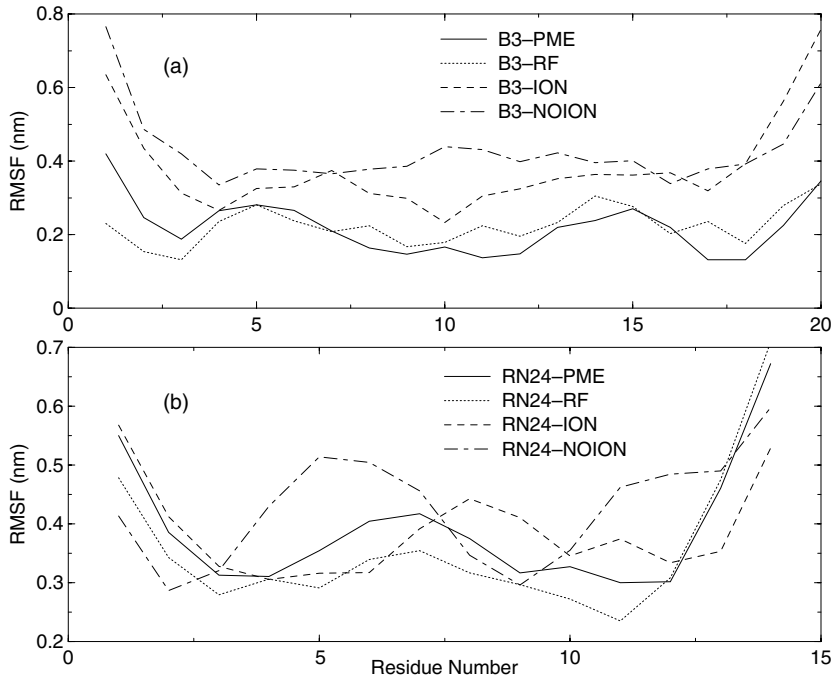

Figure 3. Root mean square positional fluctuations calculated over the 500 ns of each trajectory for beta3 (a) and RN24 (b).

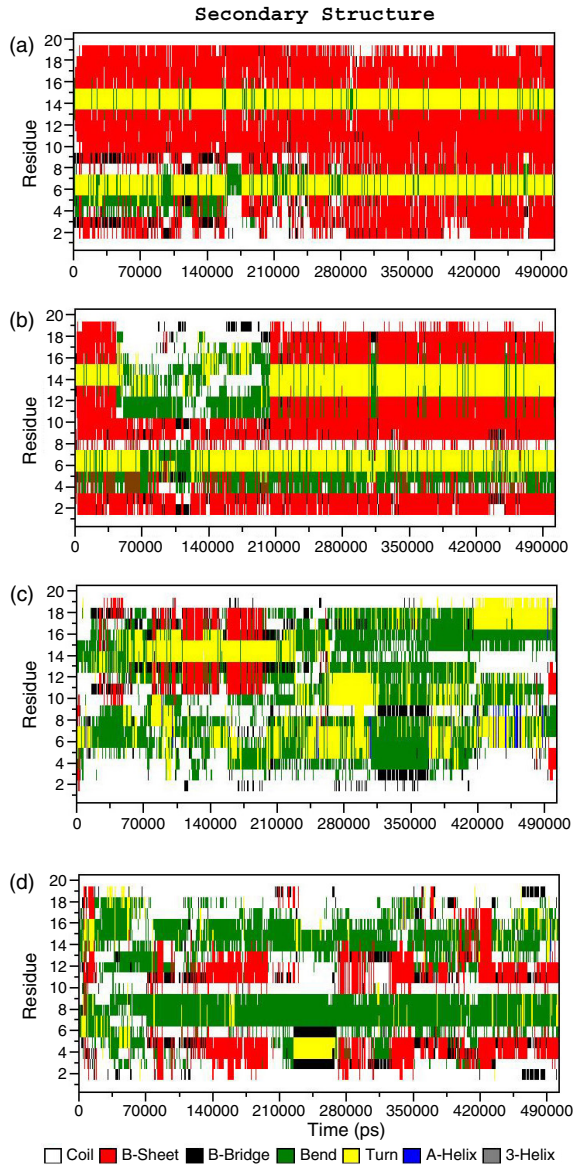

Figure 4. Time evolution of the secondary structure calculated according to the DSSP algorithm for beta3: (a) B3-PME; (b) B3-RF; (c) B3-NOION; (d) B3-ION.

recovered around 210 ns (see figure 4(b)). In B3-ION and B3-NOION, ordered secondary structure is lost and not recovered for the rest of the simulation. RN24, in contrast, can fold to the helical structure from the starting extended one in a time span ranging between 50 and 100 ns independently of the electrostatic calculation scheme employed (figure 5). Moreover the peptide samples a wide range of conformations, ranging from short β-hairpins to more disordered bent conformations. Experimental NMR results on RN24 have shown long range NOE peaks that are not typical of helices, suggesting multiple conformations in aqueous solution that include helical and bent structures [8]. Our simulation results agree with

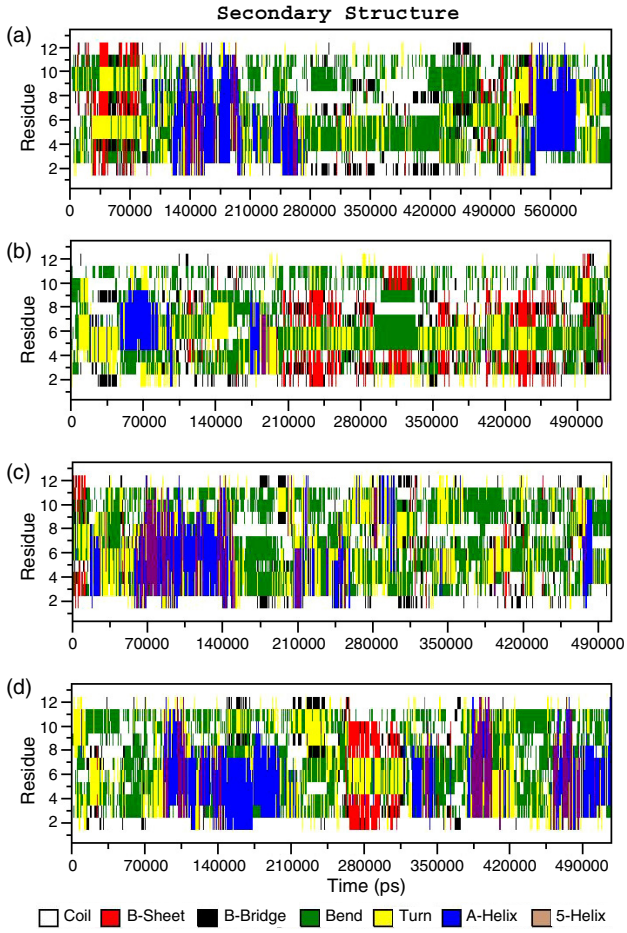

Figure 5. Time evolution of the secondary structure calculated according to the DSSP algorithm for RN24: (a) RN24-PME; (b) RN24-RF; (c) RN24-NOION; (d) RN24-ION.

these observations, showing a helical content around 10% in all simulation conditions for the RN24 peptide. Averaging over the entire trajectory, we found no violation over 0.05 nm of NOE-derived interproton distances [31].

To investigate further the helical propensities of each of the MD schemes employed, the time evolution of the number of hydrogen bonds between residues n to $n + 4$ was calculated (figure 6). Simulation RN24-RF yields a general lower estimate of the amount of helical conformations than the other three methods, which perform very similarly, RN24-PME being the one in which the amount of α-helical conformation is closer to the experimental determinations of about 20%. The analysis of the number of native interstrand H bonds in the β-sheet peptide (figure 7) shows once more the similarity in structural results obtained applying the PME and RF approaches in the simulation of systems in which long range (non-local) interactions are important for the determination of the folded conformation.

Tables 3 and 4 show the averaged values of the solvent-accessible surface area for the two peptides in the different simulations. What is clearly apparent is that PME and RF tend to give a lower solvent accessibility (more compact structures) in the case of beta3, while no sizable

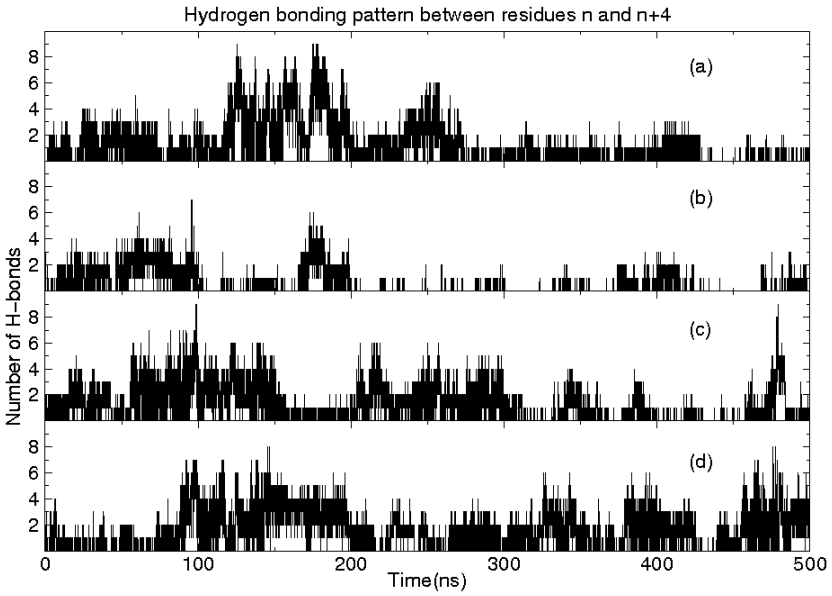

Figure 6. Time evolution of the n to $n + 4$ hydrogen bonds in the simulations of RN24: (a) RN24-PME; (b) RN24-RF; (c) RN24-NOION; (d) RN24-ION.

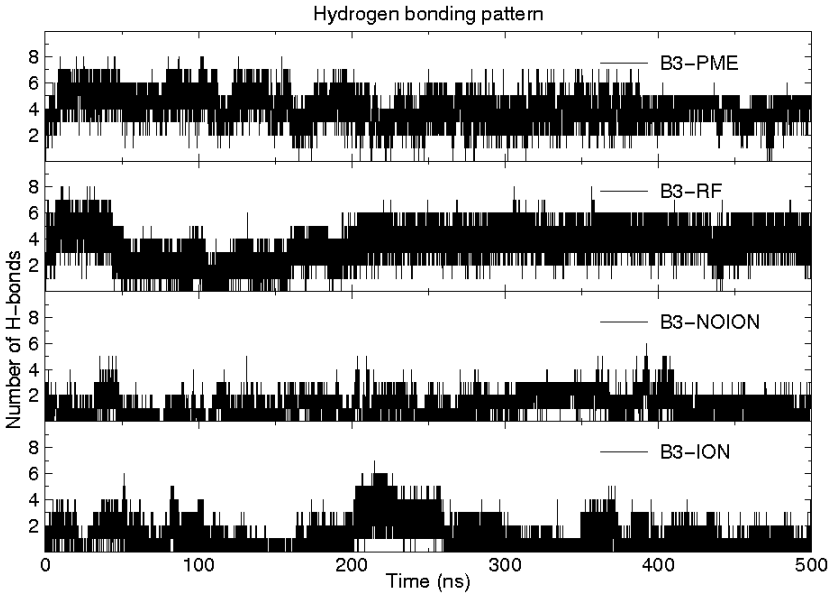

Figure 7. Time evolution of the native hydrogen bonds in the simulations of beta3.

differences are observed for RN24. Finally, the time evolution of the percentage of native contacts was considered. While no significant difference among all methods is evident from this analysis for RN24, in the case of beta3 PME can preserve more than 90% of the native contacts in time, RF more than 85%, and the two cut-offs cause this value to drop to less than 50%.

Table 3. Average solvent-accessible surface areas for hydrophobic (non-polar) and hydrophilic (polar) peptide regions in each of the four beta3 simulations.

System	Hydrophobic (NP) SASA (nm^2)	Hydrophilic (P) SASA (nm^2)	Total (T) SASA (nm^2)	NP/T ratio (%)	P/T ratio (%)
Native B3	11.08	9.22	20.29	55	45
B3-PME	10.81	6.96	17.78	61	39
B3-RF	10.47	7.19	17.65	59	41
B3-NOION	11.56	7.16	18.72	62	38
B3-ION	11.92	6.92	18.84	63	37

Table 4. Average solvent-accessible surface areas for hydrophobic (non-polar) and hydrophilic (polar) peptide regions in each of the four RN24 simulations.

System	Hydrophobic (NP) SASA (nm^2)	Hydrophilic (P) SASA (nm^2)	Total (T) SASA (nm^2)	NP/T ratio (%)	P/T ratio (%)
Native RN24	7.92	5.57	13.49	59	41
RN24-PME	8.71	5.14	13.85	63	37
RN24-RF	8.55	5.28	13.84	62	38
RN24-NOION	8.63	5.06	13.68	63	37
RN24-ION	8.74	5.12	13.86	63	37

3.2. Sampling properties

The conformational clustering algorithm developed by Daura and co-workers [12] was applied to characterize the amount of conformational space sampled in the simulations. To find clusters of structures in a trajectory the RMSD of the C_α atom positions between all pairs of structures is determined. For each structure the number of other structures for which the RMSD was 0.1 nm (on all residues except the N and C termini) or less was calculated. The structure with the highest number of neighbours was taken as the centre of the cluster and formed together with all its neighbours the first cluster. The structures of this cluster were then removed from the pool of structures. The process is repeated until the whole pool of structures is empty. In this way a series of non-overlapping clusters can be obtained. The number of clusters found as a function of simulation times is reported in figure 8. Both in B3-PME and B3-RF, the number of clusters sampled as a function of time equilibrates in the last 200 ns. In the cut-off simulations the number of clusters keeps increasing with time. Once again, it is worth noting the similarity in the results using either PME or RF for the simulations of the β-sheet system. In contrast, in the simulations of the RN24 peptide the number of clusters sampled as a function of time keeps increasing linearly as a function of time (figure 9). The absolute number of clusters sampled in RN24-PME is however strikingly lower than in the other cases, showing that the conformational dynamics of the peptide may be damped by the artificial periodicity induced by the use of lattice sums, as already noticed by other authors [4, 5]. This is particularly true for peptides with the ability to sample non-compact, extended structures. In this case, one should use very large simulation cells in order to make the perturbations due to periodicity negligible. This choice, however, would in turn have a huge impact on the computational expense required to reach timescales comparable to those of experiments. This observation is valid also if one considers the conformational transitions among different clusters as a function of time (figure 10). In the case of RN24-PME simulation, this transition number is still much lower than in all the other cases.

In contrast, in beta3 simulations, the total number of clusters visited and the transitional dynamics are very similar for B3-PME and B3-RF (figure 11). Big differences are observed

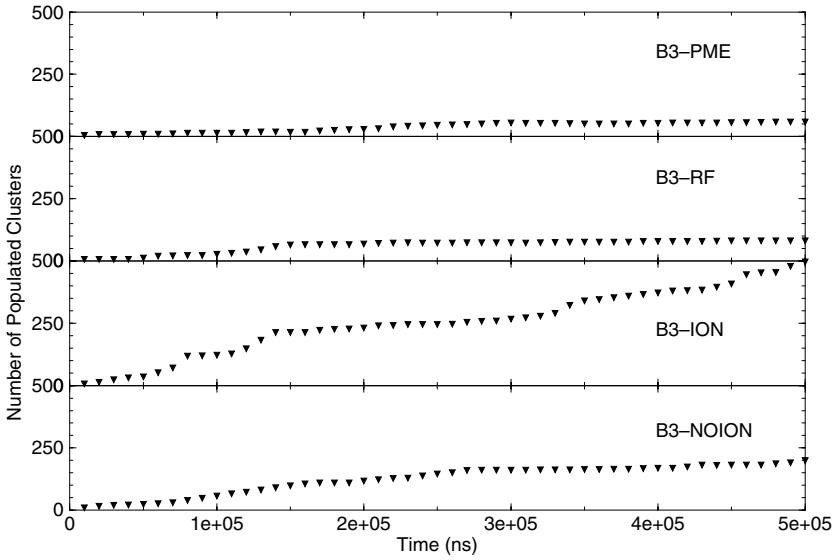

Figure 8. Time dependence of the number of clusters sampled in the course of time in simulations of beta3.

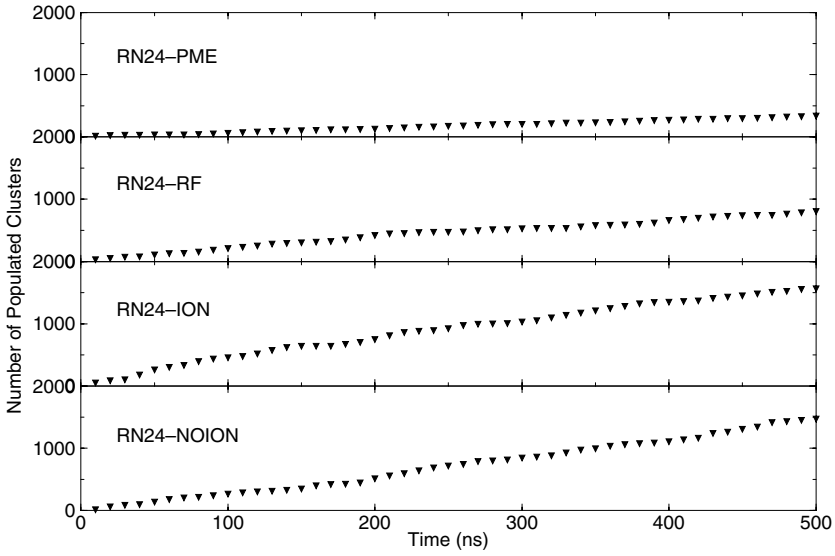

Figure 9. Time dependence of the number of clusters sampled in the course of time in simulations of RN24.

for the cut-off simulations: a much higher number of conformational clusters is observed, together with a very high number of cluster–cluster transitions. These results, combined with the previous structural observations, suggest that in B3-NOION and B3-ION, a high degree of instability is introduced at long timescales due to the neglect of important long range interactions. Interestingly, in B3-PME and B3-RF, the most populated clusters, and the clusters to which most of the transitions are directed, are the first three clusters containing native-like conformations of the peptide.

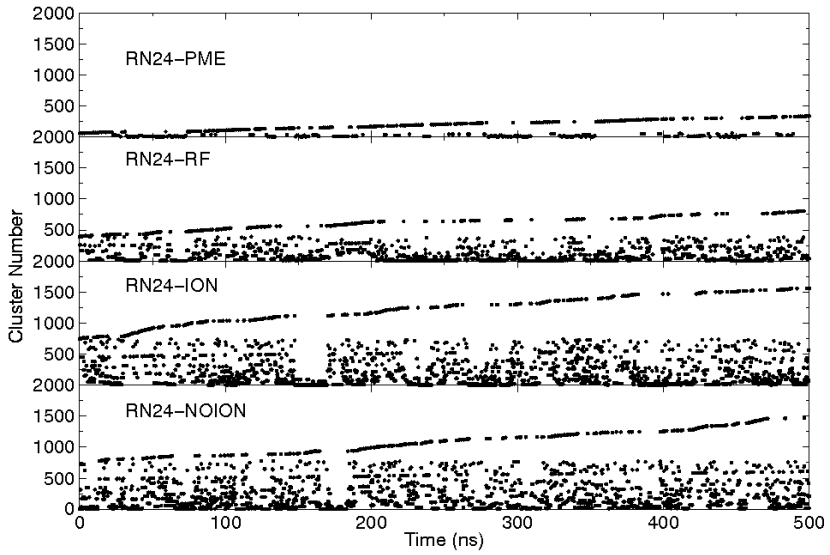

Figure 10. Transitions among clusters in the simulations of RN24.

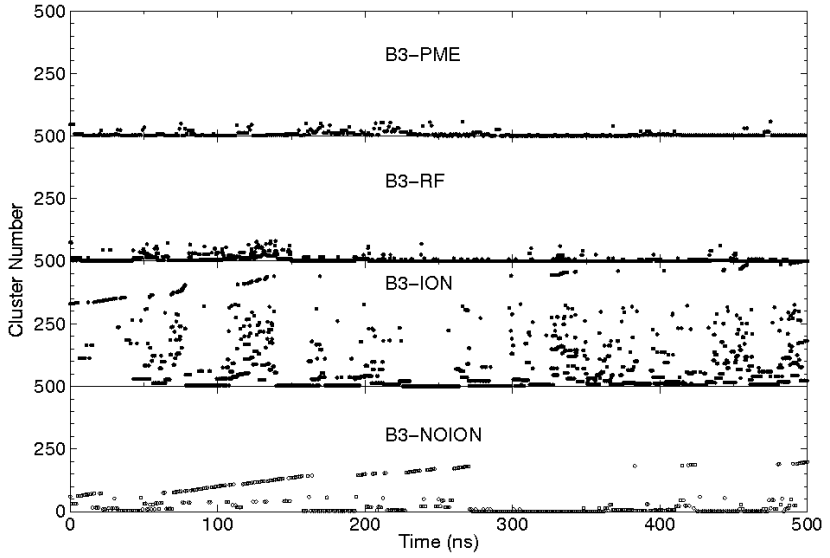

Figure 11. Transitions among clusters in the simulations of beta3.

4. Discussion

The analysis of the effect of different protocols on MD simulation results for peptides representing major secondary structure motifs is of critical importance, given the increasingly important role that simulation methods take in present day physical biology. The theoretical foundations of the methods that we discussed in this paper are different. PME is based on the exact calculation of all long range electrostatic interactions through the use of infinite replicas of the simulation box; the reaction field method is based on the truncation of Coulombic interactions after a certain cut-off distance and on the treatment of the medium outside the

cut-off as a continuum with a dielectric constant typical of the solvent under examination. Finally, pure cut-off methods rely on simple truncation of electrostatic force calculations after the predefined cut-off radius. As a consequence of these different representations, the structural and dynamical properties of the resulting MD trajectories might be different. It is important to know whether a particular scheme is better suited than others for the description of peptide systems derived from experimental NMR characterization in solution.

In this paper, we have shown that for beta3 in general PME and RF methods give comparable results in terms of structural and dynamical parameters, with RF allowing one to save almost a half of the computer time required for PME calculations. In both of these cases, the experimental structure is well reproduced for the 500 ns of the simulations. It is interesting to note that in B3-RF, a significant unfolding–refolding event during the trajectory traversal can be observed. Once refolded to the NMR three-stranded β-sheet structure, the peptide is stable for the remaining 300 ns of the B3-RF simulation, as should be expected given the long characteristic folding times (of the order of the microseconds) for β-sheet peptides.

Very different results are obtained for the β-sheet forming peptide beta3 using cut-off methods. In the latter case in particular, the starting native structure is lost after a few tens of nanoseconds and never recovered during the rest of the simulation. In contrast, the different electrostatic methods examined do not yield sensitive differences if the RN24 peptide is considered: structural fluctuations and secondary structure evolution all give in fact comparable results. The main difference is linked to the number of structural clusters visited during the simulations, which is dramatically lower for RN24-PME compared to the others. The latter simulation is also the one yielding the higher percentage of α-helical structure. Previous calculations have shown that artificial periodicity in PME calculations may overstabilize α-helices relative to other conformations and slow down the dynamics of conformational interconversion.

In all cases, multiple folding–unfolding events can be observed for RN24, consistent with the much shorter timescales (of the order of tens to hundreds of nanoseconds) required to fold to α-helical conformations. Moreover, recent NMR measurements of RN24 in aqueous solution have detected NOE peaks that are not typical of helices, suggesting the presence of bent and compact conformations, such as the ones we have observed in our simulations. The calculation of NOE distance violations by averaging over each of the RN24 500 ns trajectories found no violation over 0.05 nm of NOE-derived interproton distances [31]. This shows that in the case of a short sequence in which mainly local and short range interactions are responsible for the formation of a particular structure, such as RN24, the structural dynamics of the system turns out to be rather insensitive to the particular electrostatic treatment for long range interactions. Actually, the distances between the C_αs of residues n to $n+3$, n to $n+4$, and n to $n+5$ in an α-helical stretch vary between 4.8 and 8.1 nm, well within the radius of the sphere used to define the long range interaction calculation even in the pure cut-off methods.

The presence of long range interactions among residues far into the sequence is, in contrast, a typical signature of β-sheet forming peptides [7, 28]. That is why it should be expected that simulation protocols avoiding the crude neglect of a whole set of long range interactions could yield better results in terms of structural, dynamical, and benchmarking results against experiments. The results on beta3 actually confirm this hypothesis. Using simple cut-off methods introduces huge instabilities in both the secondary and tertiary structure, leading to a complete loss of β structure after a few tens of nanoseconds. PME and RF, in contrast, lead to stable trajectories with a very low number of NOE distance violations [9].

From the results of this study, the RF method provides a good compromise between stable trajectories, good agreement with experimental structure determination, and efficient sampling in peptide simulations. The main caveat which should be considered at this stage regards its

applicability to inhomogeneous systems: the cut-off spheres around each particle may actually not be homogeneous dielectric, that can be characterized by just one single constant. However, in spite of this drawback, RF is the only method able to dramatically improve the simulation results compared to cut-off schemes where long range effects are important, and to require much less computational effort than PME in systems where short ranged interactions are the determinant for structure formation. A second caveat is linked with the possible use of different water models such as TIP3P [29] and TIP4P [30]. They may also be thought to have an impact on the dynamics of the peptides in the context of different electrostatic schemes, mainly due to the different representation of the charge distribution on the atoms of the solvent. In this study SPC was chosen, since the parameters of the GROMOS96 force field were optimized to work with this particular model.

Empirically speaking, RF seems to represent the method of choice for all-atom MD simulations of non-trivial peptide systems over long timescales, with a computational expense that can be considered reasonable with present day PC based computer systems.

References

[1] Tironi I G, Sperb R, Smith P E and van Gunsteren W F 1995 A generalized reaction field method for molecular dynamics simulations *J. Chem. Phys.* **102** 5451–9

[2] Darden T, York D and Pedersen L 1993 Particle mesh Ewald; an $n\log(n)$ method for Ewald sums in large systems *J. Chem. Phys.* **98** 10089–92

[3] Hockney R W and Eastwood J W 1998 *Computer Simulation Using Particles* (Bristol: Institute of Physics Publishing)

[4] Weber W, Hunenberger P H and McCammon J A 2000 Molecular dynamics simulations of a polyalanine octapeptide under Ewald boundary conditions; influence of artificial periodicity on peptide conformation *J. Phys. Chem.* B **104** 3668–75

[5] Kastenholz M A and Hunenberger P H 2004 Influence of artificial periodicity and ionic strength in molecular dynamics simulations of charged biomolecules employing lattice sum methods *J. Phys. Chem.* B **108** 774–88

[6] Gargallo R, Hunenberger P H, Aviles F X and Oliva B 2003 Molecular dynamics simulations of highly charged proteins: comparison of the particle–particle–particle mesh and reaction field methods for the calculation of electrostatic interactions *Protein Sci.* **12** 2161–72

[7] de Alba E, Santoro J, Rico M and Jimenez A A 1999 De novo design of a monomeric three-stranded antiparallel β-sheet *Protein Sci.* **8** 854–65

[8] Osterhout J J, Baldwin R L, York E J, Stewart J M, Dyson H J and Wright P E 1989 1 h nmr studies of the solution conformations of an analogue of the c-peptide of ribonuclease a *Biochemistry* **28** 7059–64

[9] Monticelli L and Colombo G 2004 The influence of simulation conditions in molecular dynamics investigations of model β-sheet peptides *Theor. Chem. Acc.* **112** 145–57

[10] van Gunsteren W F, Daura X and Mark A E 1998 GROMOS force field *Encyclopedia Comput. Chem.* **2** 1211–6

[11] Berendsen H J C 1985 Treatment of long range forces in molecular dynamics *Molecular Dynamics and Protein Structure* ed J Hermans (Western Spring, IL: Polycrystal Book Service) pp 18–22

[12] Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren W F and Mark A E 1999 Peptide folding: when simulation meets experiment *Angew. Chem. Int. Edn* **38** 236–40

[13] Bernardi A, Galgano M, Belvisi L and Colombo G 2001 Simulation of carbohydrate protein interactions: computer aided design of a second generation gm1 mimic *J. Comput. Aided Mol. Des.* **15** 117–28

[14] Colombo G, De Mori G M S and Roccatano D 2003 Interplay between hydrophobic cluster and loop propensity in β-hairpin formation: a mechanistic study *Protein Sci.* **12** 538–50

[15] Cheatham T E III 2004 Simulation and modeling of nucleic acid structure, dynamics and interactions *Curr. Opin. Struct. Biol.* **14** 360–7

[16] Duan Y and Kollman P A 1998 Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution *Science* **282** 740–9

[17] De Mori G M S, Micheletti C and Colombo G 2004 All-atom folding simulations of the villin headpiece from stochastically selected coarse-grained structures *J. Phys. Chem.* B **108** 12267–70

[18] De Mori G M S, Colombo G and Micheletti C 2005 Study of the villin headpiece folding dynamics by combining coarse-grained Monte Carlo evolution and all-atom molecular dynamics *Proteins Struct. Funct. Bioinformat.* **58** 459–71

[19] van der Spoel D, van Drunen R and Berendsen H J C 1994 *GRoningen MAchine for Chemical Simulations* Department of Biophysical Chemistry, BIOSON Research Institute, Nijenborgh 4 NL-9717 AG Groningen, 1994. e-mail to gromacs@chem.rug.nl

[20] Lindhal E, Hess B and van der Spoel D 2001 Gromacs3.0, a package for molecular simulation and trajectory analysis *J. Mol. Mod.* **7** 306–17

[21] Miyamoto S and Kollman P A 1992 Settle: an analytical version of the shake and rattle algorithms for rigid water models *J. Comput. Chem.* **13** 952–62

[22] Hess B, Bekker H, Fraaije J G E M and Berendsen H J C 1997 A linear constraint solver for molecular simulations *J. Comput. Chem.* **18** 1463–72

[23] Feenstra K A, Hess B and Berendsen H J C 1999 Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems *J. Comput. Chem.* **20** 786–98

[24] Berendsen H J C, Grigera J R and Straatsma T P 1987 The missing term in effective pair potentials *J. Phys. Chem.* **91** 6269–71

[25] Berendsen H J C, Postma J P M, van Gunsteren W F, Di Nola A and Haak J R 1984 Molecular dynamics with coupling to an external bath *J. Chem. Phys.* **81** 3684

[26] Kabsch W and Sander C 1983 Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features *Biopolymers* **22** 2576–637

[27] Koradi R, Billeter M and Wuthrich K 1996 Molmol: a program for display and analysis of macromolecular structures *J. Mol. Graphics* **14** 51–5

[28] Colombo G, Roccatano D and Mark A E 2002 Folding and stability of the three-stranded β-sheet peptide betanova: insights from molecular dynamics simulations *Proteins Struct., Funct. Genet.* **46** 380–92

[29] Jorgensen W L, Chandrasekhar J, Madura J and Klein M 1983 Comparison of simple potential functions for simulating liquid water *J. Chem. Phys.* **79** 926–35

[30] Jorgensen W L and Madura J 1985 Temperature and size dependence for Monte Carlo simulations of tip4p water *Mol. Phys.* **56** 1381–92

[31] Monticelli L, Tielman D P and Colombo G 2005 Mechanism of helix nucleation and propagation: microscopic view from microsecond time scale MD simulations *J. Phys. Chem.* B at press