JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Importance of the Ion-Pair Interactions in the OPEP Coarse-Grained Force Field: Parametrization and Validation

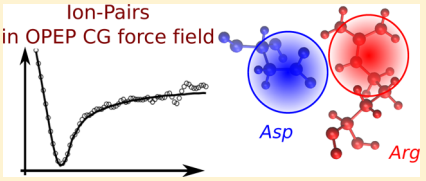
Fabio Sterpone,*[,†] Phuong H. Nguyen,[†] Maria Kalimeri,[†] and Philippe Derreumaux[†,‡]

[†]Laboratoire de Biochimie Théorique, IBPC, CNRS, UPR9080, Université Paris Diderot, Sorbonne Paris Cité, 13 rue Pierre et Marie Curie, 75005, Paris, France

[‡]Institut Universitaire de France, Bvd St Michel, 75005, Paris, France

Supporting Information

ABSTRACT: We have derived new effective interactions that improve the description of ion pairs in the Optimized Potential for Efficient protein structure Prediction (OPEP) coarse-grained force field without introducing explicit electrostatic terms. The iterative Boltzmann inversion method was used to extract these potentials from all-atom simulations by targeting the radial distribution function of the distance between the center of mass of the side chains. The new potentials have been tested on several systems that differ in structural properties, thermodynamic stabilities, and number of ion pairs. Our modeling, by refining the packing of the charged amino acids, impacts the stability of secondary structure motifs and the population of intermediate states during temperature folding/unfolding; it also improves the aggregation propensity of peptides. The new version of the OPEP force field has the potentiality to describe more realistically a large spectrum of situations where salt-bridges are key interactions.



1. INTRODUCTION

Salt-bridges are considered key interactions for protein stability.[1] They act as structural clamps that rigidify the protein matrix and link the interface of multidomain proteins[2] or protein complexes.[3,4] Their role in amyloid fibril formation is also well documented.[5,6]

From the thermodynamic point of view, their contributions to the stability of the folded state depend on the magnitude of the desolvation penalty.[7,8] The latter is alleviated when the salt-bridge is formed in a polar environment that compensates for the vanished exposure of the charged side chains to water. Therefore stabilizing salt-bridges are generally part of extended networks of electrostatic interactions.[9,10] The free energy contribution to stability has been estimated experimentally to be in the range of −0.5 to −5 kcal/mol depending on the system and on the location of the pairs.[11−14] In addition, the local rigidity caused by the ion pairing might tune the kinetic stability of the protein.[15]

The importance of salt-bridges has been specifically addressed in the context of thermophilic proteins[8,16,17] that are stable and functional at extreme temperatures, up to 100 °C.[18,19] In fact a common feature shared by these proteins with respect to their mesophilic homologues is the higher number of charged amino acids and salt-bridges.[20] Theoretical calculations pointed out that at the high working temperature of thermophiles the desolvation penalty reduces because of the smaller value of water's dielectric constant;[7] therefore making ion pairs more stable. According to other investigations, the stability gain associated to salt-bridges derives from an intrinsic overall higher polarity of the thermophilic proteins[21] or from the specific location of the ion pairs.[8,16]

In silico modeling of ion pairs and their stability is a challenging task. When charged amino acids are buried in the interior of proteins, it is not trivial to assess their protonation states and their consequent propensity to pair. In the framework of implicit solvent calculation the ion-pair energetics depends on the assumed dielectric response of the protein,[8] but the latter is shown to be not spatially uniform.[22,23] On the other hand in an explicit solvent representation, the local polarity of the solvent/protein interface where the majority of charged amino acids localizes, it is probably dependent on the model used for the water. Polarizable force fields should cure these weaknesses but at a computational cost.

The modeling is naturally more approximate in coarse-grained (CG) simulations of biomolecules where the leading philosophy is to accelerate the sampling by reducing the model accuracy to "essential" degrees of freedom or to targeted properties of the system. In Martini force field,[24] Unres,[25] Scorpion,[26] or Shea's model,[27] the electrostatics is explicitly introduced by localized effective charges, therefore salt-bridges form as the direct consequence of attractive Coulombic forces. In this context environmental effects depend on how the solvent polarity is treated. In the Scorpion force field[26] the polarity of the aqueous solution is described at a high level of theory by introducing a CG polarizable model for water. In the Unres force field the ion-pair interactions are described by specific analytical potentials.[28] These potentials are made of several physical terms, i.e., electrostatic, van der Waals, polarization, and cavity formation, whose constituent parameters are optimized to match the potential of mean force of side-

Received: April 30, 2013
Published: August 5, 2013

   dx.doi.org/10.1021/ct4003493 | J. Chem. Theory Comput. 2013, 9, 4574−4584

chain-like ions in water. For the same CG model an important effort has been undertaken toward the development of state-dependent interactions[25] to account for the temperature dependence of ion-pair stability.

In other CG models electrostatic interactions are hidden in effective pair potentials.[29−32] In the Optimized Potential for Efficient protein structure Prediction (OPEP) force field,[33,34] that is the object of the present study, the potentials acting on opposite charged amino acids are modeled with the Lennard-Jones-like function (LJ) used for other side-chain−side-chain interactions. As a consequence of this nonspecific treatment, ion pairs pack at a distance larger than that expected in real systems. For the four ion pairs, Arg/Asp, Arg/Glu, Lys/Asp, and Lys/Glu, we have compared the radial distribution functions of the distance between the side chains center of mass computed using all-atom simulations and CG simulations based on the OPEP force field (see Figure 1). Using the OPEP CG model, the side chains pack at larger distances, and for pairs involving the arginine amino acid, the doubly peaked structure of the distribution is lost.

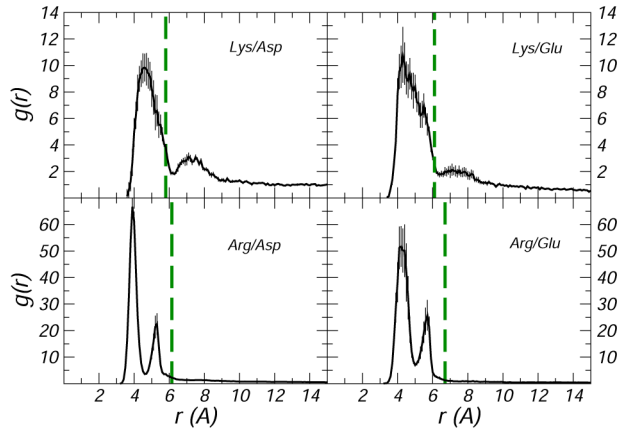


Figure 1. Radial distribution function $g(r)$ between the center of mass of the side chains computed for the four ion pairs considered in this work and obtained via all-atom simulations at ambient conditions (black lines). The vertical dashed green lines report the value of the relative distance at which the side-chain grains of the ionic amino acids of the same four pairs localize in the OPEP CG simulations; this value corresponds to the peak of the $g(r)$. Error bars were estimated by block analysis.

In this work we have derived new effective interactions that improve the description of ion pairs in the OPEP CG force field without introducing explicit electrostatic terms. The iterative Boltzmann inversion (IBI) method[35] was used to extract these potentials from all-atom simulations by targeting the radial distribution function of the distance between the center of mass of the side chains. The new potentials have been tested on several systems that differ in structural properties, thermodynamic stabilities, and numbers of ion pairs. The new model, that in future work will be referred to as OPEP v.5, by refining the packing of the charged amino acids affects the stability of secondary structure motifs and the population of intermediate states during temperature folding/unfolding. Moreover it reduces the overstabilization previously reported for OPEP v.3,[36] and in the context of protein aggregation, it eases interpeptide contacts.

The remainder of this manuscript is organized as follows: In Section 2, we describe the modeling of ion-pair potentials following a bottom-up approach, and in Section 3, we report the accurate testing for several study-cases; finally we conclude by highlighting the main finding of this work.

2. METHODS

In this section we provide details of the methodology used to model the salt-bridge propensity in the OPEP CG force field v.4 for the following four amino acid pairs: Arg/Asp, Arg/Glu, Lys/Asp, and Lys/Glu. The parametrization is reserved for the side-chain−side-chain interactions of oppositely charged amino acids. In the simulations performed to validate our modeling other ionic interactions, such as like-charged amino acids or charged C- or N-terminals with charged side chains, are modeled using the OPEP v.4 potentials as described in Supporting Information (SI).

2.1. OPEP v.4. The OPEP CG model represents each amino acid by six centers of force: the side chain (Sc) is represented by a unique bead, while atomistic resolution is reserved for the backbone that includes N, H_n, C_α, C, O atoms. Exception is proline which is represented by all heavy atoms. The potential energy of the model is described in detail elsewhere[33,34] and is briefly summarized in SI. As many force fields, the Hamiltonian consists of short-range interactions for bond lengths and angles, improper torsions and rotations, and long-range nonbonded interactions. In OPEP v.3 the van der Waals nonbonded interactions were originally described by either 6−12 or repulsive terms depending on the pairs of amino acids[33] while in the version OPEP v.4 the Sc−Sc and C_α−C_α interactions are described by a new analytical form which limits the energy values at longer distances and optimizes the Sc−Sc interaction in α-helix.[34] Hydrogen-bond propensity of backbone atoms is also accounted for via two- and four-body potentials. Solvent effects are included implicitly by optimizing the torsion and nonbonded parameters. The previous versions of the model (v.3 and v.4) have been successfully applied to protein folding and structure prediction, amyloid formation, and protein stability. OPEP has been used in combination with a broad range of simulation techniques such as the molecular dynamics (MD),[37] Monte Carlo (MC),[38] enhanced sampling methods (i.e., replica exchange MD (REMD), metadynamics, simulated tempering),[36,39−42] and greedy algorithm.[43]

2.2. Iterative Boltzmann Inversion. In order to patch the OPEP force field we have derived the effective ion-pair potentials directly from atomistic simulations by applying the IBI method.[35] The four amino acid pairs have been simulated using classical MD at ambient conditions, T = 300 K and P = 1 atm, using the OPLS force field[44,45] for the amino acids and the TIP3P model for water, see Figure 1. In order to avoid extra ionic interactions and mimic the environment along a polypeptide chain the backbone terminals were kept in neutral form with NH_2 and COOH groups (see SI). In our calculations we have simulated the pairs of entire amino acids instead of side-chain-like ions[28] in order to include side-chain length effects and the hydrophobic contributions stemming from the methyl groups CH_xs. In order to improve the convergence of the procedure the radial distribution functions (RDF) between the center of mass of the amino acid side chains extracted from the atomistic simulations have been first smoothed with a high-order polynomial fit. Moreover, as Figure 1 shows, the RDFs involving the arginine amino acid are characterized by two peaks. We have treated separately these two states by decomposing the RDFs in two parts; we will discuss later the microscopic origin of these two states and the strategy adopted

dx.doi.org/10.1021/ct4003493 | J. Chem. Theory Comput. 2013, 9, 4574−4584

for mixing the associated potentials. We anticipate here that in order to decompose the RDF, the tail of the first peak had to be smoothed to the asymptotic value of 1 using an exponential decay.

In the IBI procedure[35,46] the effective potential $V^{\mathrm{eff}}(r)$ to be used in the CG simulations is derived iteratively by targeting the atomistic RDF. At the nth step of the iteration the relationships at work are

$$V^{(n+1)} = V^{(n)} + \lambda * \Delta V^{(n)}$$

$$\Delta V^{(n)} = k_{\mathrm{b}}T * \ln \frac{g^{(n)}}{g^{\mathrm{ref}}} = V_{\mathrm{PMF}}^{\mathrm{ref}} - V_{\mathrm{PMF}}^{(n)} \quad (1)$$

where the factor $\lambda \in (0,1]$ is a coefficient that controls the numerical stability of the iteration and the term $V^{(1)} = -k_{\mathrm{b}}T \ln(g^{\mathrm{ref}})$. In the formulas the dependence on the distance r is dropped to make notation light. The term g^{ref} denotes the RDF between the centers of force representing the effective particles (in our case the center of mass of the amino acid side chains) and is extracted from atomistic simulations. Our potential V^{eff} was derived by applying triangular smoothing on a regular grid of spacing d = 0.1 Å. For each iteration the CG simulation was run for 1.5 ns at T = 300 K using a time step of 1.5 fs. The amino acid pair was confined in a sphere of radius 35 Å with reflecting boundary conditions, and the temperature was kept constant by applying the Berendsen thermostat (τ_{B} = 0.1 ps). A previous test of the OPEP v.3 FF reported a very weak effect of the thermostat (Berendsen vs Langevin) on the average structures and heat capacities of the simulated proteins.[47] Bonds were kept rigid using the SHAKE procedure.[48] The effective forces $\mathbf{F}^{\mathrm{eff}} = -\nabla V^{\mathrm{eff}}$ were calculated using linear interpolation. Each IBI procedure consisted of ∼100−200 iterations even if the convergence could generally be attained with less effort (30−40 iterations), with $\delta g_n^2 = \int (g^{(n)} - g^{\mathrm{ref}})^2 \mathrm{d}r$ saturating at 10^{-2}. The resulting potentials were further polished by extrapolating the behavior at short distance via a LJ repulsion, r^{-12}, and modeling the tail at long distance either with an exponential function or a r^{-6} attractive part; the minima and desolvation peaks were also refined by polynomial smoothing.[46] The effect of this smoothing on two representative cases is shown in SI Figure 1. The smoothing cleans the potentials at long distance where the noise due to poor sampling is large. The V^{eff} potentials and the associated RDFs compared to those obtained via all-atom simulations are reported in Figures 2 and 3, respectively. All the modeling was produced with in house developed codes interfacing the OPEP MD engine.

2.3. Two State Model for Arg. We discuss now how the two-state distribution associated with the arginine amino acid was treated. These states correspond to two different hydrogen-bond (HB) patterns between the carboxylate CO_2^- and the guanidinium $(NHC(NH_2)_2)^+$ groups (see the amino acid structures represented in SI Figure 2). An effective intuitive order parameter that distinguishes them is the relative orientation of the side chains, here defined as the scalar product between the rigid unit vector that connects the backbone C_α to the center of mass of the side chain, $\cos(\theta) = \mathbf{n}_i \cdot \mathbf{n}_j$. The 2D joint probability distribution $p(r,\cos(\theta))$ (see SI Figure 3) shows that approximately one state is associated to a head-to-head configuration while another to a parallel configuration. Instead of working with a 2D potential, we have simplified the task by introducing a mixing rule:

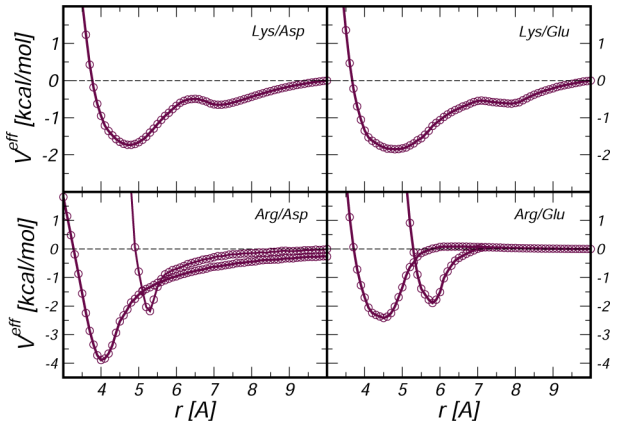


Figure 2. Effective potential V^{eff} for the four ion pairs Lys-Asp, Lys-Glu, Arg-Asp, and Arg-Glu. In the lower panels for each pair involving Arg, we distinguished the potentials for states 1 and 2.

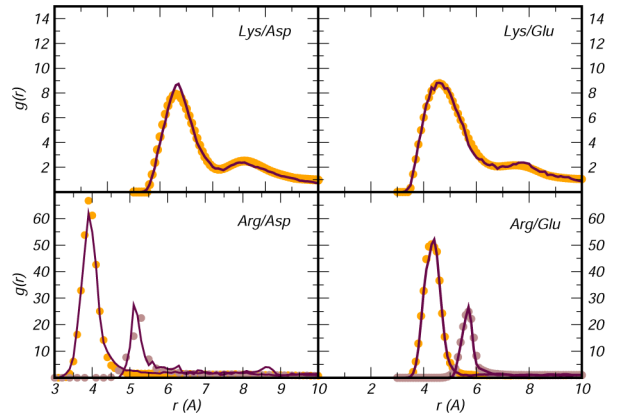


Figure 3. Radial distribution function $g(r)$ between the center of mass of the side chains computed for the four ion pairs considered in this work and obtained via all-atom simulations (dotted black lines) and the converged CG simulations during the IBI procedure (solid maroon lines). The $g(r)$ from all-atom simulations have been smoothed before applying the IBI procedure. The smoothed curves are represented in orange and brown circles (see second state in Arg/x curves).

$$V_{\mathrm{Arg}/x}^{\mathrm{eff}} = a_1^{\theta} * V_1^{\mathrm{eff}} + a_2^{\theta} * V_2^{\mathrm{eff}} \quad (2)$$

where $V_{1(2)}^{\mathrm{eff}}$ refers to the potential associated to state 1(2) for the interaction of Arg with the amino acid x = Asp, Glu, and the coefficients $a_{1(2)}^{\theta}$ are weights that depend on the value of $\cos(\theta)$. In order to determine the weights we have computed the population of states 1 and 2 as a function of $\cos(\theta)$ and used a convenient sigmoid function to fit them, see Figure 4. In the present implementation the weights $a_{1(2)}$ are parameters and do not contribute directly to the force via their angular dependence.

2.4. The Energy Scale. In the OPEP force field the interactions between the center of forces are all scaled with weighting factors (w_x in SI), which are optimized to reproduce the energy spectra of folded, excited, and unfolded configurations of a large set of peptides[33] as well as the stability of selected proteins.[34] In the IBI procedure only isolated amino acid ion pairs have been considered, therefore the relative energy scale of the effective ion-pair potentials with respect to other side-chain interactions is a missing information. This has been tuned by a systematic testing on selected systems and discussed in Section 3. This scaling factor is indicated by f

dx.doi.org/10.1021/ct4003493 | J. Chem. Theory Comput. 2013, 9, 4574−4584

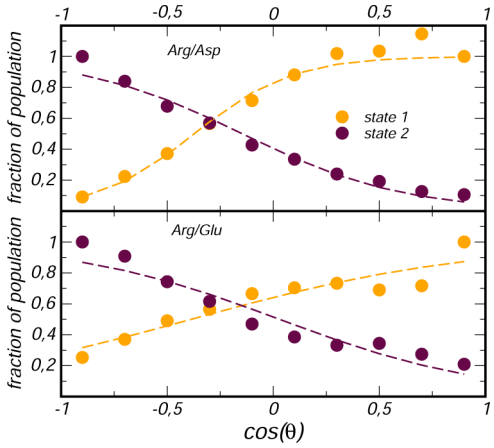


Figure 4. Fraction of population occupying the states 1 and 2 (yellow and maroon circles, respectively) as a function of the order parameter $\cos(\theta)$ and for the two ion pairs Arg/Asp and Arg/Glu (top and bottom panels, respectively). Data from all-atom simulations are represented in circles, the fitting function, $1/e^{\alpha(x-\beta)}+1$, is represented with dashed lines.

throughout the paper and is defined as the ratio between the energy at the minimum, ε^0, calculated for the ion pair Lys/Glu and the hydrophobic interaction Ile/Ile taken as reference value of the OPEP energy scale, $f = (\varepsilon^0_{\text{Lys/Glu}})/(\varepsilon^0_{\text{Ile/Ile}})$. The term $\varepsilon^0_{\text{Ile/Ile}}$, which includes its weighting factors, is set to 4.46 kcal/mol. In this work we used the version 4 of the OPEP force field,[34] and when the new potentials for ion pairs are included, the force field is referred to as OPEP/IP. In the limiting condition of $f = 0$ no potential acts between the amino acids of opposite charge.

2.5. The Dependence on Temperature. The effective potentials V^{eff}, as CG models in general, are thermodynamically state dependent, and their use in thermodynamic conditions different from the reference state could introduce artifacts.[49] However for many purposes CG models are routinely used to simulate systems in different temperature ranges and often combined with enhanced sampling methods like REMD that fully exploit temperature excitation. In doing so one should pay attention to some aspects. Herein we have first investigated the response of the reference RDFs to temperature increase, see SI Figure 4. After Boltzmann inversion we noticed minor temperature changes (~0.2 kcal/mol) on the pairs involving the Lys amino acid. On the contrary the temperature has a negligible effect for the pairs involving the Arg amino acid. It was previously reported[25] that Lys/Glu(Asp)-like interactions modeled by representative ions in water solution show a temperature dependence at the contact distance minimum, while the Arg/Asp(Glu) interactions are not affected by temperature changes. However, it is important to stress that for salt-bridges confined at biomolecular interfaces, like those in multidomain proteins, the increase of the temperature could reduce the desolvation cost of an amount as large as −5 kcal/mol,[50] therefore making ion pairs more stable in the high-temperature regime.

Moreover when REMD method is used, as in the current work, the standard acceptance rate must be carefully rederived to account for the dependence of the potential energy on the temperature. Given the set of replicas $X[.., q_m^{(i)}, ..., q_n^{(j)}, ...]$ the probability for exchanging the copies i and j and generating the new set $X'[.., q_m^{(j)}, ..., q_n^{(i)}, ...]$ can be derived from the detailed balance conditions.[51] For a potential energy composed by a temperature independent $(E(q))$ and dependent part $(V^{\text{eff}}(q,T))$ the swapping move is controlled by

$$\frac{w_{X\to X'}}{w_{X'\to X}} = e^{\Delta\beta_{mn}\Delta E+\Delta v^{\text{eff}}} \tag{3}$$

where the first contribution $\Delta\beta_{mn}\Delta E = (\beta_m - \beta_n)(E(q^{(i)}) - E(q^{(j)}))$ is that used in standard REMD, and the term Δv^{eff} is given by

Table 1. CG and All-Atom Simulations Performed in This Work Using Different Methodologies: Standard MD, REMD, and IBI[a]

system	run type	time [ns]	temperature [K]	force field
AA pairs	MD	200, 600	300, 360	OPLS
AA pairs	IBI	1.5 per step	300	OPEP v.4
RNase C-peptide RN24′[c]	REMD (24)	600 per replica	[300−475],[200−501][b]	OPEP v.4/OPEP/IP
RNase C-peptide RN24′[c]	REMD (48)	50 per replica	[280−410]	OPLS
RNase C-peptide WT	REMD (24)	200 per replica	[300−475]	OPEP v.4/OPEP/IP
β-hairpin	REMD (24)	400, 600 per replica	[300−475],[200−501][b]	OPEP v.4, OPEP/IP
β-hairpin	REMD (48)	70 per replica	[280−410]	OPLS/Amber99sb-ildn/Gromos
ccβ-p2 monomer	REMD (24)	200 per replica	[200−501]	OPEP v.4, OPEP/IP
ccβ-p2 dimer	REMD (24)	100, 200 per replica	[200−501]	OPEP v.4, OPEP/IP
2DA1	MD	60	290	OPEP/IP
1CLB	MD	30	290	OPEP/IP
1FCL	MD	30	290	OPEP/IP
1E0L	MD	60	290	OPEP/IP
1SHG	MD	30	290	OPEP/IP
1CM2	MD	60	300	OPEP v.4/OPEP/IP
1CM2	REMD (36)	300 per replica	[220−536]	OPEP v.4, OPEP/IP
1CM2	MD	100	300	Charmm22

[a]For REMD simulations the number of replicas is reported in parentheses. OPEP/IP indicates the OPEP force-field including the ion-pair interactions modeled in this work. [b]Range of temperature used for OPEP/IP and with a scaling factor $f = 1.0$ and 1.3. [c]Mutated system (K1A/E9L/M13A).

dx.doi.org/10.1021/ct4003493 | J. Chem. Theory Comput. 2013, 9, 4574−4584

$$\Delta v^{\text{eff}} = \beta_m(V_m^{\text{eff}}(q^{(j)}) - V_m^{\text{eff}}(q^{(i)})) + -\beta_n(V_n^{\text{eff}}(q^{(j)}) - V_n^{\text{eff}}(q^{(i)})) \quad (4)$$

In this work all the REMD simulations were based on the $V^{\text{eff}}_{T=300K}$. An accurate test of temperature-dependent potentials is reserved for a future investigation.

3. RESULTS

The derived potentials V^{eff} have been tested against several systems comprising three short monomeric peptides, a dimeric peptide system, and a set of six proteins. These study cases differ in structural properties and thermodynamic stabilities as well as the number of ion pairs. In some cases all-atom simulations of the study cases have been performed. In Table 1 the simulation details are summarized, while extra information concerning the protocols of the simulations are given in SI.

3.1. RNase C-peptide. The first system is the ribonuclease A C-peptide that corresponds to the first 13 residues stretch at the N-terminal of the protein. First, we have considered the system with three single point mutations (K1A/E9L/M13A) → (AETAAAKFLRNHA), this peptide is indicated below as RN24′. This peptide capped with succinyl and NH_2 groups (RN24),[52] the wild type[53] and other mutants,[54] have been well characterized experimentally by CD and NMR techniques as well as using computer simulations.[55] The RN24′ was already used as a test case to assess the performance of the OPEP FF.[36] The C peptide is a representative model of α-helix, and in the RN24′ mutant the number of possible ion-pair combinations is only two (E2/K7 and E2/R10). The system RN24′ was simulated in its zwitterionic form by REMD. We used 24 copies of the system, and the temperature range spanned by the replicas changed depending on the energy scaling factor f. Each individual replica was evolved using a time step of 1.5 fs for an effective OPEP total time of 600 ns. For a selected energy scale ($f = 1.3$), we have also simulated the peptide without mutations (WT) for 200 ns per replica.

We have first monitored the stability of the RN24′ peptide and identified the melting temperature by computing the specific heat curve as a function of temperature. The specific heat is calculated from the fluctuation of the potential energy $C_v = <\delta U^2>/k_bT^2$, and its temperature dependence is handled by the PTWham algorithm.[56] The curves obtained for different values of the scaling factor f are plotted in Figure 5. The shape of the curves is rather similar, all showing a peak at ∼390−400 K. The WT exhibits a quite similar behavior with $T_m \simeq 410-420$ K (data not shown). Experimentally the C-peptide was found to unfold above 303 K,[53] thus OPEP FFs overstabilize these helical systems.

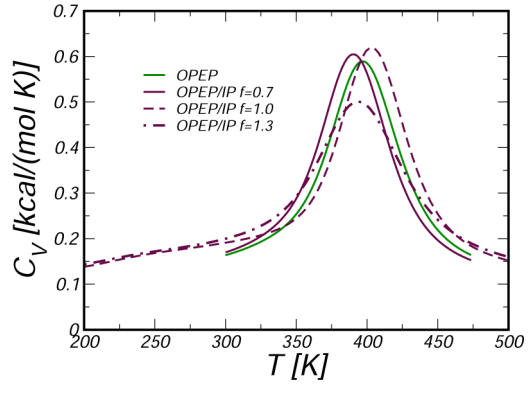


Figure 5. Specific heat C_v curves for the RNase C-peptide system.

In the folded state, the peptide adopts an α-helix which is stabilized by intramolecular HBs and salt-bridges. Experiments[52,54] and simulation[55] on the RN24 and other mutants have examined the role of the salt-bridge interaction Glu2-Arg10 on the helix stability. This ion pair becomes key for the stability since the K1A/E9L mutations remove a competitive stronger ion pair between Lys1 and Glu9. The Glu2-Arg10 ion pair is well represented in the OPEP/IP with the two side chains interacting at ∼4.5 Å, while in the original OPEP the equilibrium distance is much larger, ∼6 Å (see Figure 7). The close packing guided by the new potentials models well the experimental evidence of a direct HB between the two residues as it was suggested by NMR technique[52,54] and as it was observed in the structures of the RNase protein. Moreover the OPEP/IP stabilizes this packing with minor fluctuations, while in the original force field larger noninteracting distances, $d > 10$ Å, are also sampled (6%).

We have introduced an *ad hoc* parameter for counting the number of ion pairs formed during the simulation:

$$S_{IP} = \sum_{i,j} \frac{1 - (d_{ij}/d_0)^6}{1 - (d_{ij}/d_0)^{12}} \quad (5)$$

with $d_0 = 7$ Å and i,j running over the Asp, Glu and Arg, Lys amino acids, respectively.

As shown in Figure 6 we find a higher value of S_{IP} when the effective IP potentials are in use. The new potentials cause also a more extended/persistent and temperature resistant ion-pair network along the helix. In fact in the CG simulations a second pair is visible between Glu2-Lys7. REMD simulations have been carried out at the atomistic resolution for the same system confirming that a close arrangement of the side chain center of mass of these two residues is probable (see the bottom-left panel of Figure 7). However this interaction does not represent a true ion pair but mirrors, instead, the hydrophobic packing between the side chains. The HB between the ionic groups of the amino acids is formed rarely and occurs mainly in unfolded conformations. Upon unfolding and using the energy scaling factors $f = 1.0$ and 1.3, residual ion pairings are still observed at high T, e.g. the bound states Glu2-Lys7 and Glu2-Arg10 are

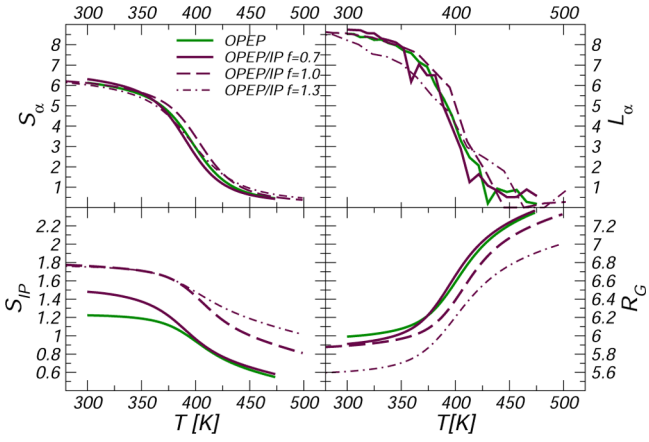


Figure 6. Average value of the HB number (S_α), gyration radius (R_G), number of ion pairs (S_{IP}) and α-helix length (L_α) formed along the RNase C-peptide as a function of temperature. Averages for all parameter but L_α are calculated using the PTwham algorithm.[56].

dx.doi.org/10.1021/ct4003493 | J. Chem. Theory Comput. 2013, 9, 4574−4584

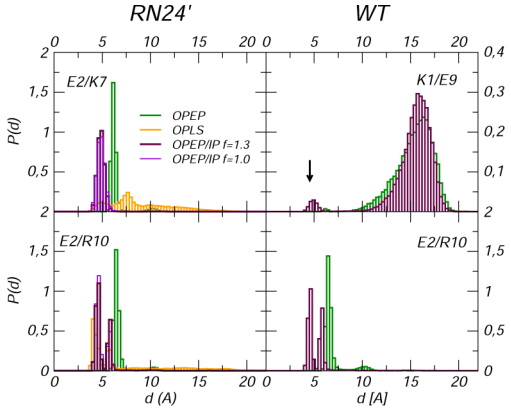


Figure 7. Peptide C. Left panels: Probability distribution of the distances between oppositely charged amino acids in the mutated C-peptide from OPEP v.4 (green) and OPEP/IP (maroon) and all-atom OPLS (orange) REMD simulations. In the right panels distribution of the distances between the pair K1/E9 and E2/R10 in the WT C-peptide from OPEP v.4 (green) and OPEP/IP (maroon). Data refer to T = 300K.

formed 70% and 50% of the simulation time at 420 K. Our finding suggests a stabilizing effect of these ion pairs on the intermediate structures sampled during the helix folding. For sake of comparison we have calculated the distribution of the ion-pair distances in the WT system and inquired the propensity of the amino acids Lys1 and Glu9 to form a salt-bridge. Brooks et al. have shown using atomistic simulations[55] that this pair competes with the Glu2-Arg10 salt bridge. Interestingly we note that even if with a low probability only the OPEP/IP potentials predict a bound state for the pair Lys1-Glu9, see the top-right panel in Figure 7.

Intramolecular backbone HBs (S_α), gyration radius (R_g) and α-helix length (l_α) of the mutated system take very similar values at ambient temperature, and their thermal disruption follows the same trend in both force fields, see Figure 6. (For counting the HBs we have defined the parameter S_α in the same spirit of eq 5 and monitoring the distance between the amide hydrogens and the carbonyl oxygens in the backbone (d_0 is set to 3 Å) and we consider only (i,i + 4) interactions with i indicating the residue index. Secondary structures were predicted using the STRIDE algorithm.[57]) It is worth noting that in our simulations at ambient conditions, the average helix lengths of the WT and RN24′ are comparable and encompass both ∼7−8 residues, ∼54−62% of the peptide length. In both peptides the helix starts at the level of the residue Ala4, and the last three residues of the N-terminus are not structured in agreement with the results from NMR[52] and previous calculations.[36,55]

Overall on this system because of the small content of ionic interactions, the refined potentials did not impact the global stability compared to the original force field; however they improved the relative packing of ionic amino acids.

3.2. β-hairpin. The second study-case is the GB1 β-hairpin (GEWTYDDATKTPTVTE) corresponding to the fragments 41−56 in the C-terminal of the GB1 immunoglobulin binding protein. For this system an increased number of ion pairs is expected to form in folded and unfolded states. Again we used REMD with 24 copies and tested several combinations of the scaling factor f. The peptide was simulated in zwitterionic form. Each individual replica was evolved for an effective OPEP time of 400−600 ns depending on the scaling factor and using a time step of 1.5 fs.

The specific heat curves obtained with OPEP and OPEP/IP using f = 0.7 are very similar with a peak at the melting temperature $T_M \simeq 340$ K. Increasing the energy scaling factor to f = 1.0 and 1.3, the melting temperature downshifts to 280−290 K, values very close to that measured experimentally T_M^{exp} = 297 K,[58] see Figure 8. When folded, the stability of the β-sheet

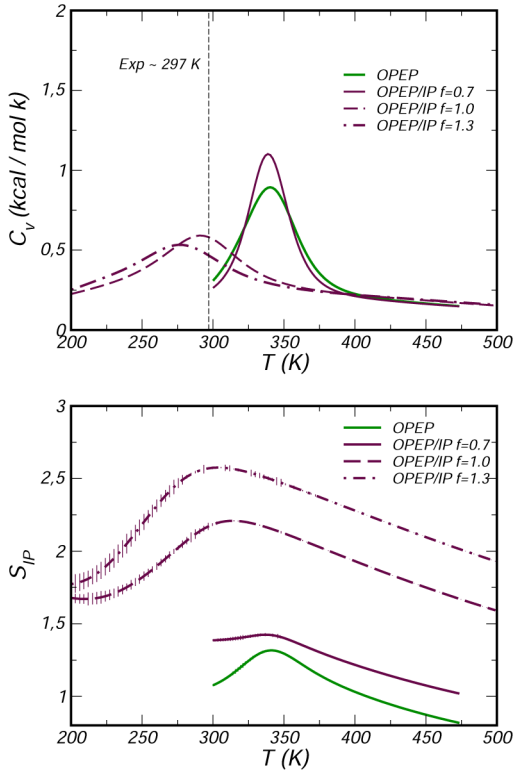


Figure 8. Top panel: Specific heat C_v curves for the β-hairpin system. Bottom panel: Average number of ion-pairs computed for β-hairpin as a function of the temperature and using the PTwham algorithm.[56]

is ensured by interbackbone HBs, hydrophobic packing of side chains, and a very stable salt-bridge formed at the turn region (Asp7-Lys10).[59] This pairing is very well described by OPEP/IP that enforces the CG beads to pack at the correct distance (see SI Figure 5). In the CG simulations we note that a conformational change of the turn region makes also favorable the interaction between Asp6 and Lys10. The frequency of this interaction and the packing distance distribution depend on the scaling factor f. We have checked whether this interaction is reproduced in REMD atomistic simulations, and this is not the case for the commonly used force fields, i.e., Amber99sb-ildn, OPLS, and Gromos. For these atomistic simulations we have reported in SI Figure 5 the distributions of the distances between the center of mass of opposite charged amino acids. Since Asp6-Lys10 packing is detected also using the original OPEP even if at a larger distance (∼ 6 Å, see SI Figure 5), we deduce that this liaison did not derive from the specific form of the nonbonded interactions. It is mainly caused by the fluctuations of the torsional angles in the backbone and/or by the reduced resolution of the side chains in the CG model. The former aspect has an impact on the accessible conformations of the turn region, and the latter could remove both steric clashes and screening effects. On the other hand, it

dx.doi.org/10.1021/ct4003493 | J. Chem. Theory Comput. 2013, 9, 4574−4584

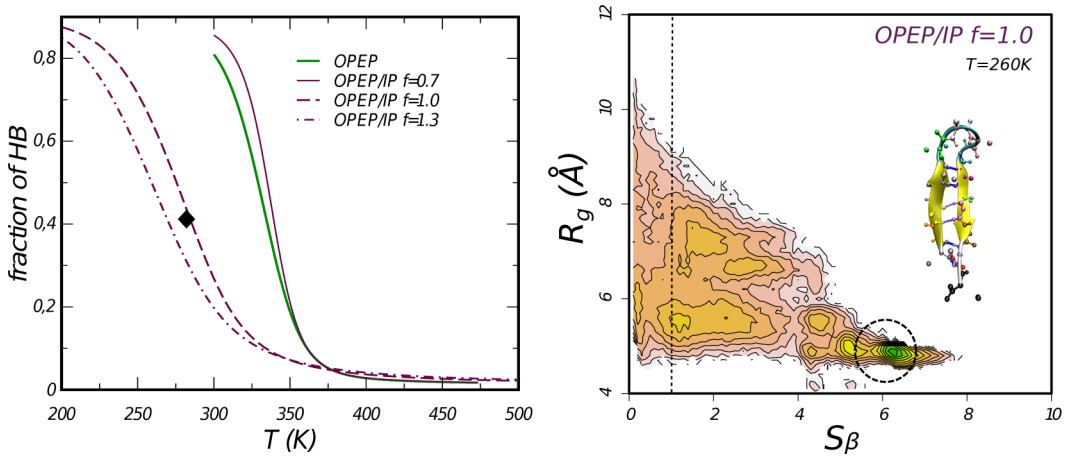


Figure 9. Left panel: Fraction of HBs vs T. The black diamond indicates the experimental value.[62] Right panel: 2D FES of the β-hairpin computed at T = 260 K for the OPEP/IP model and using a scaling factor f = 1.0. Equi-surface lines at 1.5 k_bT. A representative configuration of the folded state with root-mean-square displacement of ∼0.8 Å with respect to the native configuration extracted from the crystallographic structure pdbcode 1GB1.

is worth noting that the residue Asp6 forms a stable interaction with the residue Thr1, and this bonding has been observed in previous atomistic simulations[59] and pointed as an essential key for the stability of the folded state.

For all values of the scaling factor f, the number of ion pairs counted according to eq 5 is nonmonotonic with temperature as it is shown in the bottom panel of Figure 8, it first increases across the melting temperature because partially unfolded conformations allow long-range ion pairing between the terminal acidic residues Glu2 and Glu16 and the basic residue Lys10. These conformations are characterized by a transition of the turn region toward an α-helix. The overpairing smoothly disappears at higher temperatures. It is worth noting that in a recent study[60] some microstates characterized by partial α-helix content have been observed to act as gateways for the folding of the β strand. The presence of helical structures in the unfolded peptide is a long-standing issue since they were detected in computer simulations, see discussion in refs 60 and 61 but not in experiments.[58,62] The conflict might originate from some deficiencies of the atomistic force fields. In our study we observe a bell-like distribution of the helix content[57] as a function of temperature. When the OPEP/IP FF is used the distribution has a peak around the T_M, and at the peak the percentage to find an α-helix is 30 and 40% for f = 1.0 and 1.3, respectively.

The folded state of the β-hairpin is effectively described by counting the number of HBs formed across the backbone, or S_β. The order parameter S_β is defined using the same filtering as in the eq 5 but now considering the distance between the backbone amide hydrogens and carbonyl oxygens of residues i,j with $j > i + 4$; and by setting d_0 = 3 Å. S_β decreases with temperature according to a two-state model in agreement with experiment[58] and previous computations,[60,61,63] see left panel of Figure 9. At the temperature T = 282 K the OPEP/IP f = 1.0 predicts a percentage of HBs of 43%, in excellent agreement with the estimate from NMR at the same T (42%).[62]

A more global view of the conformational properties of the system is obtained by projecting the free energy landscape on a set of collective variables. For the β-hairpin the number of formed HB (S_β) was used in combination with the gyration radius of the hydrophobic core of the peptide (R_g).[36,61] The hydrophobic core is calculated including the heavy atoms and the side-chain beads of the residues Trp3, Tyr5, Phe12, and Val14. The free energy surfaces (FES) were computed using the PTwham algorithm.[56] In Figure 9 we report the landscape for OPEP/IP f = 1.0 at T = 260 K. The overall shape of the FESs obtained using OPEP and OPEP/IP looks similar below, close, and above the associated T_M (data not shown) and compares fairly well to what reported from atomistic simulations in ref.[36]

We have computed the free energy of unfolding, $\Delta F_{FU} = -k_bT \ln(P_F/P_U)$, by considering the probability distribution of the parameter S_β and defining a convenient two-state dividing surface:[36] the folded conformations correspond to values of S_β > 1 and the unfolded ones to values of S_β < l. In Figure 10 we report ΔF_{FU} as a function of temperature. We notice a shift in the melting temperature (here defined at ΔF_{FU} = 0) and a different linear dependency of ΔF_{FU} vs T upon increasing the factor f. The magnitude of the unfolding free energy compares fairly well with the estimates from atomistic simulations performed at T = 300 K, and based on the OPLS and Amber99sb FFs, see data in ref 36, these values are reported as dashed lines in the top panel of Figure 10. Finally, in the bottom panel of Figure 10 we compared the temperature variation of the folded population calculated in our simulations to that deduced experimentally from tryptophan fluorescence technique.[58] There is a good agreement below the melting temperature (293 K) when f = 1.0; on the contrary, the stronger energy scale f = 1.3 compromises the stability of the peptide and the β conformation becomes readily unfavorable.

For the β-hairpin the optimal scaling factor f = 1.0 allows to reproduce the experimental melting temperature as well as the number of HB and the content of β conformations in the folded state.

3.3. The ccβ-p2 Switch Peptide. The ccβ-p2 switch peptide is a repeat of alternate polar and hydrophobic residues, Ac-SREELEAKIRELELRIG-NH_2.[64] This peptide was designed to fold as a α-helix, but increasing temperature and in concentrated solution, it forms irreversibly amyloid fibril. At low temperature, coiled-coil structures stabilized by intra- and intermolecular salt-bridges are expected.

We first focus on the monomeric state. The original OPEP predicts a quite high melting temperature, T_M = 360 K (see Figure 11). Introducing the IP effective interactions, the stability of the peptide strongly depends on the scaling factor f. The melting temperature shifts to much lower values for low

dx.doi.org/10.1021/ct4003493 | J. Chem. Theory Comput. 2013, 9, 4574−4584

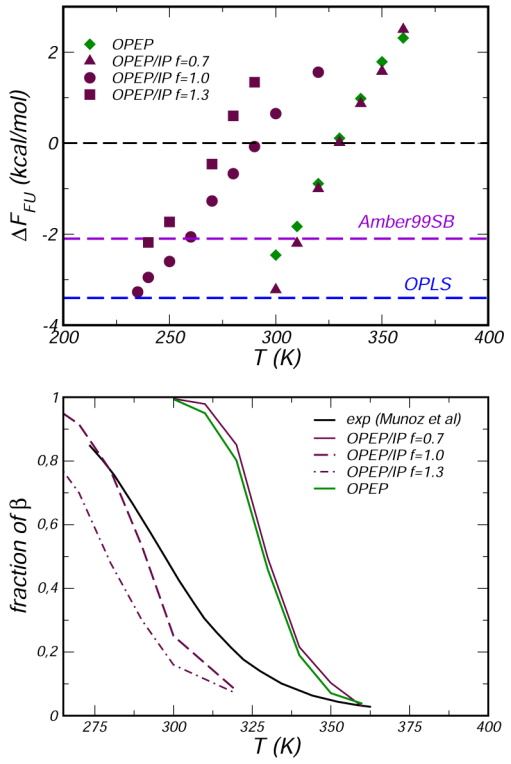


Figure 10. Top panel: Unfolding free energy as a function of the temperature for the β-hairpin system. The probability distribution of the number of HBs ($P(S_\beta)$) is used to compute the population of the folded ($S_\beta > 1$) and unfolded ($S_\beta < 1$) states. Colored dashed lines indicate the value of ΔF_{FU} reported in ref 36 and estimated using atomistic simulations at T = 300 K by combined parallel tempering and metadynamics. Bottom panel: fraction of β population estimated using the S_β parameter and compared to experimental data from Munoz et al.[58]

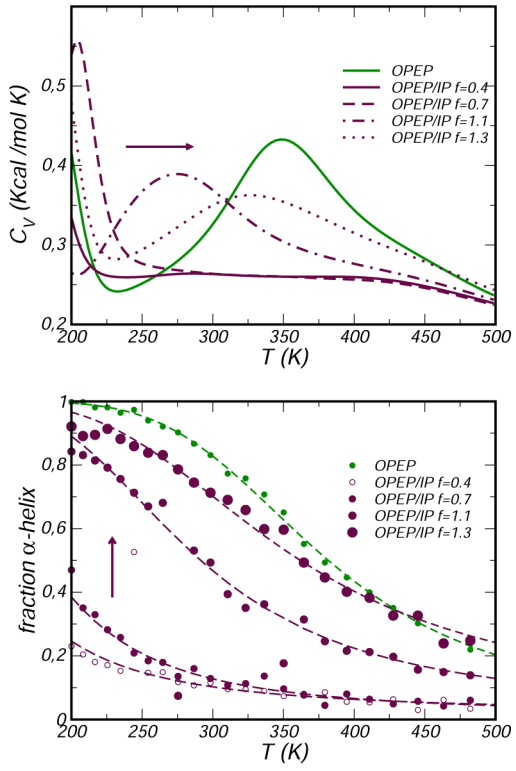


Figure 11. Top panel: Specific Heat C_V curves for the ccβ-p2 monomer system. Bottom panel: Population of the pure α-helix state computed for the monomer ccβ-p2 as a function of temperature and for several energy scales f. The population decay is fitted by a sigmoid function deriving from a two-state model.

f. For higher scaling factors f = 1.1 and 1.3, the T_M increases to 270 and 330 K, respectively. We show below that these values stabilize the α-helix in the folded state in agreement with experiments.[64] The population of the α structure (p_α) with respect to the total ensemble of sampled conformations is reported as a function of temperature in the lower panel of Figure 11 (the secondary structure prediction was based on the STRIDE algorithm).[57] Using the original OPEP the peptide fully occupies the α-helix state at low temperature ($p_\alpha \sim 100\%$ at T = 200 K), and increasing the temperature it unfolds transforming into a random coil. The population $p(\alpha)$ decreases as a function of temperature according to $p(\alpha) = 1/1 + e^{\beta\Delta F_{FU}}$) as predicted by a two-state model. This feature is reproduced by OPEP/IP only with a high value of the scaling factor. In particular the values f = 1.1 and 1.3 produce $p_\alpha \sim 80\%$ and ~90% at T = 200 K; on the contrary for smaller values of f the peptide samples solely the unfolded conformations. We have estimated the unfolding enthalpy (ΔH) and entropy (ΔS) by fitting the temperature decay of $p(\alpha)$ and deduced that the achieved stability associated to high scaling factors f is purely enthalpic with a negligible entropic contribution. When formed, the α-helix covers ~65−70% (OPEP/IP f = 1.1−1.3) and 70% (OPEP) of the peptide's total length in very good agreement to what was reported in the original work by Kammerer and Steinmetz[64] by applying the AGADIR algorithm[65] (70%). It is worth noting that the original OPEP does not make favorable the β structure, while introducing the effective IP potentials, this motif becomes accessible even if not dominant. For example with OPEP/IP f = 1.1, the probability to occupy a pure β state at low temperature is $p_\beta \sim 10\%$ and the motif extends to ~45% of the peptide.

We turn our attention to the dimeric state. The two peptides are expected to form a coiled-coil aggregate at low T. Using the original OPEP force field we observe that in the low-temperature regime (T = 200−270 K) the pure α-helix structure is favored ($p_\alpha \simeq 70\%$), but pure β structures can be also accessed ($p_\beta \simeq 25\%$). The effective potential in OPEP/IP reverses the α/β populations when f = 1.1, with the pure β structure being slightly more favorable (45%) than pure α (25%). Further increase of the scaling factor, f = 1.3, makes the α-helix dominant in the folded state, $p_\alpha \sim 60\%$, hence favoring the coiled-coil pairing. Moreover the refined ion-pair potentials increase the propensity to aggregate as shown in the right top panel of Figure 12. This is a long-range effect controlled by the tail of the new potentials that substantially differ from the previous modeling that was specifically constructed to be short-range in nature.[34] In this regard, a comparison of the two models is provided in SI Figure 1. Several ion-pair patterns link the two peptides (see the frequency matrix in Figure 12) with a variety of relative orientations and not an unique parallel orientation as predicted by the designers.[64] It is worth mentioning here a previous work by Missimer and co-workers[66] on the ccβ-p peptide. Using all-atom simulations of the aggregated trimeric state of the peptide in water solution,[66] they investigated the role of interhelical ion pairs on the stability of the coiled-coil, and it was noticed that at high temperature the

dx.doi.org/10.1021/ct4003493 | J. Chem. Theory Comput. 2013, 9, 4574−4584

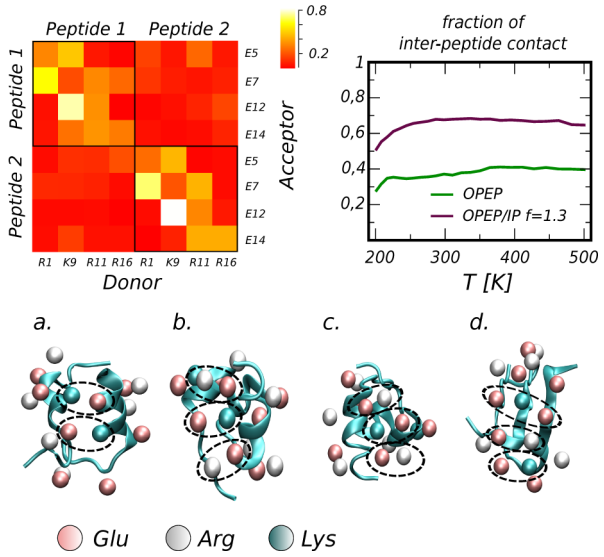


Figure 12. Top panels, left: Frequency matrix of intra- and interpeptide ion pairs and right: fraction of states in which the peptides are in contact vs temperature. Lower panels: Representative aggregation states extracted from the simulation using OPEP/IP f = 1.1 (a–c) are representatives of coil–coiled states while (d) is a β-sheet aggregate. We have highlighted several connectivity patterns and the underlying secondary structure of the two peptides in (a) the formed ion pairs are $E^{(1)}7/K^{(2)}9$ and $K^{(1)}9/E^{(2)}5$; in (b) and (c) $E^{(1)}12/R^{(2)}16$,$K^{(1)}9/E^{(2)}14$ and $E^{(1)}5/R^{(2)}11$; in (d) $E^{(2)}7/K^{(1)}9/E^{(1)}14$ and $R^{(2)}1/E^{(1)}7/R^{(1)}16$.

network of possible interhelical ion pairs can accommodate stochastically the disorder induced by high temperature.

In both force fields and for all tested scaling factors f we do not see dominant β structures at high temperature (even if rarely sampled, see Figure 12) as experimentally observed; this is because in our dimeric system, there is no reservoir of peptides for nucleating the growth of fibrils.

In conclusion, for the switch peptide, when the scaling factor is correctly tuned (f = 1.1 and 1.3), the OPEP/IP potentials provide a stability of the helical conformation close to the original OPEP in agreement with the experiments and also improve the propensity to aggregate and make accessible the β conformers as expected by the designers.[64]

3.4. Protein Stability. We turn our attention to larger systems. We have selected five proteins with size between 37 and 75 residues. These proteins (PDB codes: 2DA1, 1CLB, 1FCL, 1E0L, 1SHG) already served as test cases to assess the protein stability with the OPEP v.4 FF, see ref 34. We have performed MD simulations with the new OPEP/IP potentials and various scaling factors f. The systems were first equilibrated with harmonic restrains on C_α atoms, then the trajectory was evolved for 30 ns for each system and value of f. The C_α RMSD calculated on the protein rigid cores are reported in Table 2. For the lowest value f = 0.9 all proteins but one (1SHG) exhibit a remarkable stability on the explored time-scale; at a higher value f = 1.0 two systems (1FCL and 1SHG) experience a larger deviation with respect to the experimental crystal structure, a signature of a minor structural distortion. By visual inspection we observed that for these proteins, the long-range nature of the new potentials favors some local pairings that are not present in the X-ray configuration and that distort the relative orientation of secondary structures. The highest scaling factor f = 1.3 enhances even more this behavior. For the well-behaving proteins 2DA1 and 1E0L we have verified the stability on a longer time scale (f = 0.9 and t_{sim} = 60 ns) and found comparable C_α RMSD, see data in parentheses in Table 2, respectively. Finally, for 2DA1 we have monitored the effect of the thermostat on the structural stability by performing an extra simulation of 10 ns where temperature was controlled by the Langevin thermostat: On the same time scale the difference between the average C_α RMSD is as small as 0.1 Å, (2.9 ± 0.7 vs 2.8 ± 0.3 Å).

The final test is performed on a protein of larger size (85 residues): the histidine-containing protein (HPr) from *E. coli*. Its fold owns both α-helices and β-sheets. The number of charged amino acids is 19, of which 8 are basic and 11 acidic. Experimentally the protein is stable up to the melting temperature T_M = 336 K and has optimal stability at 290 K.[67] The protein was simulated with the single point mutation H15A (PDB code 1CM2).

We have first investigated the stability of the secondary structure of the protein as compared to what was observed in atomistic simulations at the same thermodynamic conditions. The CG trajectories were evolved at T = 300 K for ∼60 ns using a time step of 1.5 fs and several values of the scaling factor f. The original OPEP force field preserves quite well the overall structure of the protein as mirrored by the RMSD with respect to the crystallographic structure that fluctuates steadily around the value of 4 Å. However a major deficiency is manifested when looking at the stability of the secondary structure. The short α-helix Ser46-Thr52 unfolds shortly after the equilibration phase. This instability is cured by OPEP/IP with a scaling factor f = 1.3. However on longer time scale, as also observed for 1FCL and 1SHG proteins, a too strong ion paring causes instability in another part of the protein, the helix-1 (Thr16-Glu25).

The response to temperature was examined via REMD with 24 replicas spanning the range 220–530 K. Each replica has been simulated for 300 ns. The melting temperature derived from the OPEP simulations T_M = 400 K is rather high compared to the experimental value of 336 K, see Figure 13.

Table 2. C_α RMSD

system	size	rigid core	aa^-/aa^+	v.5 f = 0.9	v.5 f = 1.0	v.5 f = 1.3	v.4[a]
2DA1	60	16–63	8/13	2.9 ± 0.5(3.2)	3.4 ± 0.6	3.8 ± 0.5	3.2 ± 0.2
1CLB	75	3–14,20–39,47–53,60–73	17/10	3.3 ± 0.3	3.2 ± 0.3	5.0 ± 0.3	2.9 ± 0.2
1FCL	56	3–55	10/6	3.7 ± 0.9	4.3 ± 0.8	8.6 ± 0.4	2.8 ± 0.2
1E0L	37	7–33	6/5	3.1 ± 0.3 (3.1)	3.1 ± 0.3	5.4 ± 0.2	2.1 ± 0.5
1SHG	57	8–11,29–61	9/11	4.3 ± 0.5	5.0 ± 0.8	8.2 ± 0.7	2.4 ± 0.3
1CM2	85	2–9, 15–28, 31–52, 58–84	11/8	5.3 ± 0.6(5.9)	5.0 ± 0.5(4.9)	4.3 ± 0.4(4.6)	4.2 ± 0.4(4.3)

[a]The data for 2DA1, 1CLB, 1FCL, 1E0L, and 1SHG are taken from ref 34. For sake of comparison with ref 34 these systems were simulated at 290 K. For the HPr system we report two RMSD values, one calculated after 30 ns of simulation and a second one in parentheses after 60 ns.

dx.doi.org/10.1021/ct4003493 | J. Chem. Theory Comput. 2013, 9, 4574–4584

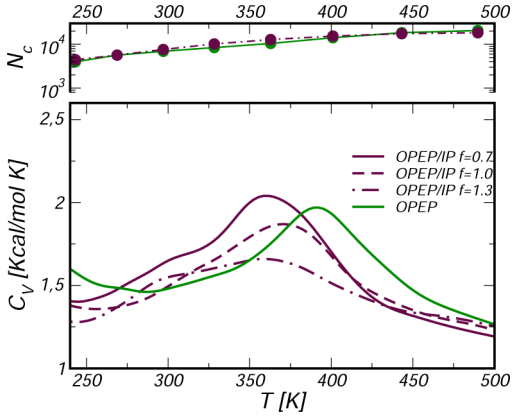


Figure 13. Specific heat curve for the HPr protein. In the inset graph the S_β parameter is reported. Top panel: number of conformational states visited by the protein as a function of the temperature.

However, as already observed for β-hairpin, when the IP potentials are switched on T_M shifts down approaching a more realistic value, $T_M \simeq 360$ K. It is worth noting that even classical all-atom force fields generally overestimate the melting temperature of proteins with a shift of ∼30−40 K.[68] The relative occupation of the folded state as a function of selected temperatures is reported in SI Figure 6.

In the temperature window 300−400 K the number of substates visited by the protein in OPEP/IP simulation (f = 1.3) is ∼20% larger than with OPEP. This has been quantified by computing the number of clusters formed at a selected temperature with C_α RMSD as a metric and a cutoff of 5 Å, see top panel in Figure 13. This finding suggests, as previously noted for the β-hairpin, that a refined description of ion pairing impacts not only the folded conformations but also the possible intermediates toward the unfolded state. It is however important to stress that in CG simulations the overpairing and the preserved local structures in partially unfolded conformers could be in principle favored by the lack of an explicit representation of the solvent and hence of its competitive solvation forces.

4. CONCLUSION

In this work we have introduced ad hoc potentials for describing ion-pair interactions in the OPEP force field[33,34] without including electrostatics explicitly. The potentials have been derived according to a bottom-up procedure using the IBI method that targets the radial distribution function between the side chain centers of mass of ionic amino acid pairs. The targeted RDFs are extracted from atomistic simulations with explicit solvent at ambient conditions. The force field including these effective interactions has been tested against a large set of study cases. An optimal range of values for the energy scale of the ion-pair interactions with respect to the other force field terms has been determined, f = 0.9−1.0. These values ensure the optimal stability of small size proteins on the time scale of tenths of nanoseconds; moreover for the β-hairpin and HPr proteins they allow to calculate melting temperatures which are closer to the experimental estimates. In some cases, higher values of f performed better as we observed for the ccβ-p2 switch peptide. However, we believe to be a general safe protocol, for any new system, to perform preliminary tests and assess the best scaling factor.

We observed that the new potentials, by refining the relative packing of ion pairs, have an impact on the stability of the folded conformations and on the population of intermediate states. They have also important effect on peptide aggregation. Moreover they reduce the overstabilization previously reported for the OPEP force field.[36] In short, the new version of OPEP, from now will be referred to as version 5, has the potentiality to describe more realistically a large spectrum of situations where salt-bridges are key interactions. In future work we will explicitly deal with the temperature dependence of these interactions and how simplified solvent effects can be introduced at the level of a hydrodynamics description.[69]

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

In the Supporting Information we provide the description of the OPEP coarse-grained force-field and the protocols of the atomistic simulations carried out to support the modelling and validation of the ion-pair potentials. We also provide data to complement the discussion of the results. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: fabio.sterpone@ibpc.fr

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

F.S. and M.K.: The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) grant agreement no.258748. We are grateful to CINES HPC for allocated computational time in Jade (DARI 2012 c2012086818). We acknowledge the financial support for infrastructures from ANR-11-LABX-0011-01.

■ REFERENCES

(1) Kumar, S.; Nussinov, R. *J. Mol. Biol.* **1999**, *293*, 1241−1255.
(2) Zhang, X.; Meining, W.; Fischer, M.; Bacher, A.; Ladenstein, R. *J. Mol. Biol.* **2001**, *306*, 1099−1114.
(3) Sheinerman, F. B.; Honig, B. *J. Mol. Biol.* **2002**, *318*, 161−177.
(4) Salari, R.; Chong, L. T. *J. Phys. Chem. Lett.* **2010**, *1*, 2844−2848.
(5) López de la Paz, M.; Goldie, K.; Zurdo, J.; Lacroix, E.; Dobson, C. M.; Hoenger, A.; Serrano, L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16052−16057.
(6) Speare, J. O.; Rush, T. S.; Bloom, M. E.; Caughey, B. *J. Biol. Chem.* **2003**, *278*, 12522−12529.
(7) Elcock, A. H. *J. Mol. Biol.* **1998**, *284*, 489−502.
(8) Xiao, L.; Honig, B. *J. Mol. Biol.* **1999**, *289*, 1435−44.
(9) A.G., G.; Gribenko, A. V.; Makhatadze, G. I. *Protein Sci.* **2008**, *17*, 1285−1290.
(10) Donald, J. E.; Kulp, D. W.; DeGrado, W. F. *Proteins* **2011**, *79*, 898−915.
(11) Tissot, A. C.; Vuilleumier, S.; Fersht, A. R. *Biochemistry* **1996**, *35*, 6786−6794.
(12) Strop, P.; Mayo, S. L. *Biochemistry* **2000**, *39*, 1251−1255.
(13) Luisi, D. L.; Snow, C. D.; Lin, J.-J.; Hendsch, Z. S.; Tidor, B.; Raleigh, D. P. *Biochemistry* **2003**, *42*, 7050−7060.
(14) Makhatadze, G. I.; Loladze, V. V.; Ermolenko, D. N.; Chen, X.; Thomas, S. T. *J. Mol. Biol.* **2003**, *327*, 1135−1148.
(15) Gruia, A. D.; Fischer, S.; Smith, J. C. *Chem. Phys. Lett.* **2004**, *385*, 337−340.
(16) Kumar, S.; Ma, B.; Tsai, C. J.; Nussinov, R. *Proteins* **2000**, *38*, 368−83.

(17) Thomas, A. S.; Elcock, A. H. *J. Am. Chem. Soc.* **2004**, *126*, 2208−14.
(18) Vieille, C.; Zeikus, G. J. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1−43.
(19) Sterpone, F.; Melchionna, S. *Chem. Soc. Rev.* **2012**, *41*, 1665−1676.
(20) Karshikoff, A.; Ladenstein, R. *Trends Biochem. Sci.* **2001**, *26*, 550−6.
(21) Dominy, B. N.; Minoux, H.; Brooks, C. L., III *Proteins* **2004**, *57*, 128−41.
(22) Simonson, T.; Perahia, D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1082−1086.
(23) Guest, W. C.; Cashmanb, N. R.; Plotkin, S. S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6286−6295.
(24) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. *J. Chem. Theory Comput.* **2013**, *9*, 687−697.
(25) Sobolewski, E.; Ołdziej, S.; Wiśniewska, M.; Liwo, A.; Makowski, M. *J. Phys. Chem. B* **2012**, *116*, 6844−6853.
(26) Basdevant, N.; Borgis, D.; Ha-Duong, T. *J. Chem. Theory Comput.* **2013**, *9*, 803−813.
(27) Bellesia, G.; Shea, J.-E. *J. Chem. Phys.* **2009**, *130*, 145103.
(28) Makowski, M.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. B* **2011**, *115*, 6130−6137.
(29) M. R. Betancourt, D. T. *Protein Sci.* **1999**, *8*.
(30) Ding, F.; Buldyrev, S. V.; Dokholyan, N. V. *Biophys. J.* **2005**, *88*, 147−155.
(31) Májek, P.; Elber, R. *Proteins* **2009**, *76*, 822−836.
(32) Cheon, M.; Chang, I.; Hall, C. K. *Proteins* **2010**, *78*, 2950−2960.
(33) Maupetit, J.; Tuffery, P.; Derreumaux, P. *Proteins* **2007**, *69*, 394−408.
(34) Chebaro, Y.; Pasquali, S.; Derreumaux, P. *J. Phys. Chem. B* **2012**, *116*, 8741−8752.
(35) Pütz, D. R. M.; Müller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624−1636.
(36) Barducci, A.; Bonomi, M.; Derreumaux, P. *J. Chem. Theory Comput.* **2011**, *7*, 1928−1934.
(37) Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2007**, *126*, 025101.
(38) Derreumaux, P. *Phys. Rev. Lett.* **2000**, *85*, 206−209.
(39) Wei, G.; Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2003**, *119*, 6403.
(40) Nasica-Labouze, J.; Meli, M.; Derreumaux, P.; Colombo, G.; Mousseau, N. *PLoS Comput. Biol.* **2011**, *7*, e1002051.
(41) Chebaro, Y.; Jiang, P.; Zang, T.; Mu, Y.; Nguyen, P. H.; Mousseau, N.; Derreumaux, P. *J. Phys. Chem. B* **2012**, *116*, 8412−8422.
(42) Nguyen, P. H.; Okamoto, Y.; Derreumaux, P. *J. Chem. Phys.* **2013**, *138*.
(43) Thévenet, P.; Shen, Y.; Guyon, J. M. F.; Derreumaux, P.; Tufféry, P. *Nucleic Acids Res.* **2012**, *40*, W288−93.
(44) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(45) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(46) Rühle, V.; Junghans, C.; Lukyanov, A.; K. Kremer, K.; Andrienko, D. *J. Chem. Theory Comput.* **2009**, *5*, 3211−3223.
(47) Spill, Y. G.; Pasquali, S.; Derreumaux, P. *J. Chem. Theory Comput.* **2011**, *7*, 1502−1510.
(48) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(49) D'Adamo, G.; Pelissetto, A.; Pierleoni, C. *J. Chem. Phys.* **2013**, *138*, 234107.
(50) Salari, R.; Chong, L. T. *J. Phys. Chem. B* **2012**, *116*, 2561−2567.
(51) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96−123.
(52) Osterhout, J. J.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Dyson, H. J.; Wright, P. E. *Biochemistry* **1989**, *28*, 7059−7064.
(53) Bierzynski, A.; Kim, P. S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2470−2474.
(54) Fairman, R.; Shoemaker, K. R.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Biophys Chem.* **1990**, *37*, 107−119.
(55) Khandogin, J.; Chen, J.; Brooks, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18546−18550.
(56) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. *J. Chem. Theory Comput.* **2007**, *3*, 26−41.
(57) Frishman, D.; Argos, P. *Proteins* **1995**, *23*, 566−579.
(58) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196−199.
(59) Tsai, J.; Levitt, M. *Biophys. Chem.* **2002**, *101*, 187−201.
(60) De Sancho, D.; Mittal, J.; Best, R. B. *J. Chem. Theory Comput.* **2013**, *9*, 1743−1753.
(61) Zhou, R.; Berne, B. J.; Germain, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931−14936.
(62) Blanco, F. J.; Rivas, G.; Serrano, L. *Nat. Struct. Mol. Biol.* **1994**, *1*, 584−590.
(63) Nguyen, P. H.; Stock, G.; Mittag, E.; Hu, C.-K.; Li, M. S. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 795−808.
(64) Kammerer, R. A.; Steimetz, M. O. *J. Struct. Biol.* **2006**, *155*, 146−153.
(65) Lacroix, E.; Viguera, A.; Serrano, L. *J. Mol. Biol.* **1998**, *284*.
(66) Missimer, J. H.; Steinmetz, M. O.; Baron, R.; Winkler, F. K.; Kammerer, R. A.; Daura, X.; van Gunsteren, W. F. *Protein Sci.* **2007**, *16*, 1349−59.
(67) Nicholson, E. M.; Scholtz, J. M. *Biochemistry* **1996**, *35*, 11369−11378.
(68) Lindorff-Larsen, K.; Piana, S.; Dror, R.; Shaw, D. *Science* **2011**, *334*, 517−520.
(69) Fyta, M.; Kaxiras, E.; Melchionna, S.; Succi, S. *Comput. Sci. Eng.* **2008**, *10*, 10−19.