JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Effects of Lysine Substitution on Stability of Polyalanine α Helix

Sergei Y. Ponomarev, Qina Sa, and George A. Kaminski*

Department of Chemistry and Biochemistry, Worcester Polytechnic Institute, Worcester, Massachusetts 01609, United States

ABSTRACT: We have studied the stability of polyalanine α-helices with lysine residues added at C- and N-termini in the gas phase and aqueous solution. Monte Carlo simulations with the fixed-charge OPLS-AA and our polarizable POSSIM force fields were carried out. The results of the simulations confirm previously observed phenomena of the helix being stable with the LYS residue on the C-terminus and losing its helical structure if the charged LYS residue is located at the N-terminus of the polypeptide in the gas phase. Both OPLS-AA and POSSIM force fields performed essentially similarly; thus the validity of both for reproducing and predicting structures of such polypeptides has been confirmed. We have also studied the effect of replacing the normal N- and C-termini with methyl capping (this approach is often used in computational studies). Our results have demonstrated that the structure and stability of the polypeptides do not depend significantly on such a substitution, although details of the resulting structure may change. The liquid-state simulations produced stable α-helixes regardless of the position of the protonated lysine residue. Overall, we have validated our polarizable POSSIM force field and the techniques used in the simulations, since the change of the helix structure as a function of the position of the LYS residue depends on a fine balance of energy contributions, and our methodology reproduced this balance well.

I. INTRODUCTION

Computer simulations of proteins and peptides have become an integral part of biomedically relevant research. Since most of such systems are too large to be treated with accurate quantum mechanical techniques, simulations with empirical force fields are the method of choice in many such studies, especially those involving protein molecules. While a detailed comparison of experimental and computational results is not always possible, some benchmarking is still available, especially in checking the secondary structure of the proteins against the available experimental data.

However, there is significant difficulty which often makes such comparisons problematic. There are many factors in play in the overall energy balance, which is responsible for a particular protein or peptide structure. Most notably, it may be difficult to separate the effect of the robustness of the solute force field in itself and the energy component which is responsible for the hydration energy of the system if the experiment and the simulation both take place in the liquid phase. This is why the high resolution gas-phase ion mobility experiments[1,2] offer a unique opportunity to decouple those components from each other and thus to validate them separately.

The utility of a direct comparison of gas-phase simulation and experimental results is also popular because the protein and polypeptide structures in the gas phase are often less stable. They show a greater degree of dependence on various conditions (such as discussed in more detail below), thus offering a great way to confirm whether the empirical force field in question can reproduce this dependence.

Homopolymers of simple amino acids, such as polyalanines, are one of the favorite targets for the comparison of computational and experimental data, as they have been studied by the both means.[3] Computer simulations demonstrate that such systems tend to form α-helices below a certain critical temperature. At the same time, there is experimental evidence of instability of such helices in the gas phase.[4] It has been suggested that this discrepancy could result from the fact that the experiments deal with charged systems, while simulations are normally done on electrostatically neutral polypeptides.[4] It has been demonstrated that there is a strong dependence of the polyalanine helix stability on the location of the charged group. For instance, a positively charged protonated lysine residue located at the N-terminus destabilizes the helical structure, while the same residue at the C-terminus makes it more stable.[4] This observed distinction offers a great test case for the validation of empirical force fields for proteins.

We have carried out simulations of tridecaalanine polypeptides (ala-13) in the gas phase with a single protonated lysine residue at either the C- or N-terminus. Moreover, it has been demonstrated previously that explicit treatment of electrostatic polarization in simulations of proteins with empirical force fields can be crucial in evaluating energetic properties, such as protein−ligand binding energies and calculations of protein pK_a shifts.[5−7] At the same time, the accuracy of reproducing protein structures is not necessarily so sensitive to the inclusion of the polarization term. Thus, in order to assess the importance of including explicit electrostatic polarization into the total Hamiltonian, both our polarizable POSSIM force field[8,9] and the fixed-charges OPLS-AA[10] were employed in the calculations presented in this article. Parameters for the LYS residue represent an addition to the functional groups previously available in the POSSIM, and thus they had to be fitted in the course of the reported studies.

The polypeptides were also simulated in aqueous solution with the fast polarized POSSIM force field.

Another issue which is important in protein force field development and applications is in simulating the termini of the proteins or peptides in hand. It is often convenient to use methyl $-CH_3$ caps instead of the actual termini in running computer simulations.[9−12] We present results of the α-helix simulations of

Received: June 13, 2012
Published: October 2, 2012

the polyalanine chains with both the actual termini and the methyl caps. This is done to assess the importance of reproducing the termini exactly for obtaining the correct data on the stability of the helixes.

The rest of the paper is organized as follows. Methods are presented in the next section. Section III is devoted to results and discussion. Finally, the conclusions are shown in Section IV.

II. METHODOLOGY

A. Structures of the Molecules. We considered the following four initial structures in our simulations. Figure 1 shows

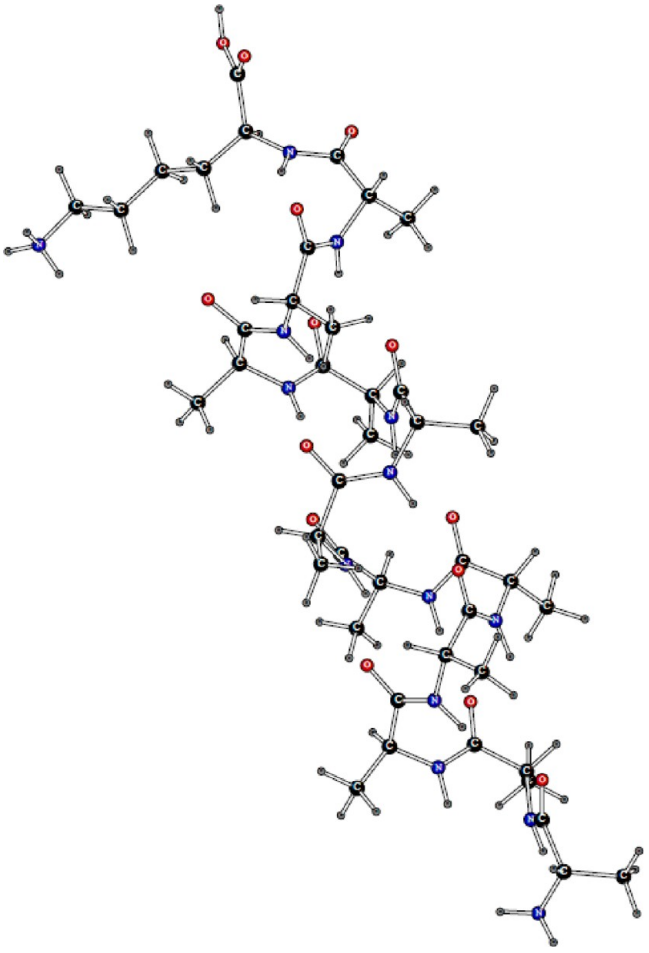

Figure 1. Ala13 structure with a lysine residue on the C-terminus.

the Ala13 structure with a lysine residue on the C-terminus (Ace-Ala$_{13}$-Lys^{+}). Presented in Figure 2 is the Ace-Lys^{+}-Ala$_{13}$ polypeptide with the protonated lysine on the N-terminus. Figures 3 and 4 show similar structures except that the normal termini are replaced with $-CH_3$ caps.

B. Force Fields. Both the fixed-charge OPLS-AA[10] and our polarizable POSSIM[8,9] force fields were employed. Their formalism can be described briefly as follows. The total energy E_{tot} is calculated by adding the electrostatic interactions $E_{electrostatic}$, the van-der-Waals energy E_{vdW}, harmonic bond stretching and angle bending $E_{stretch}$ and E_{bend}, and the torsional term $E_{torsion}$:

$$E_{tot} = E_{electrostatic} + E_{vdW} + E_{stretch} + E_{bend} + E_{torsion} \quad (1)$$

Electrostatic Energy. In the polarizable POSSIM force field, the electrostatic polarization energy as calculated with inducible

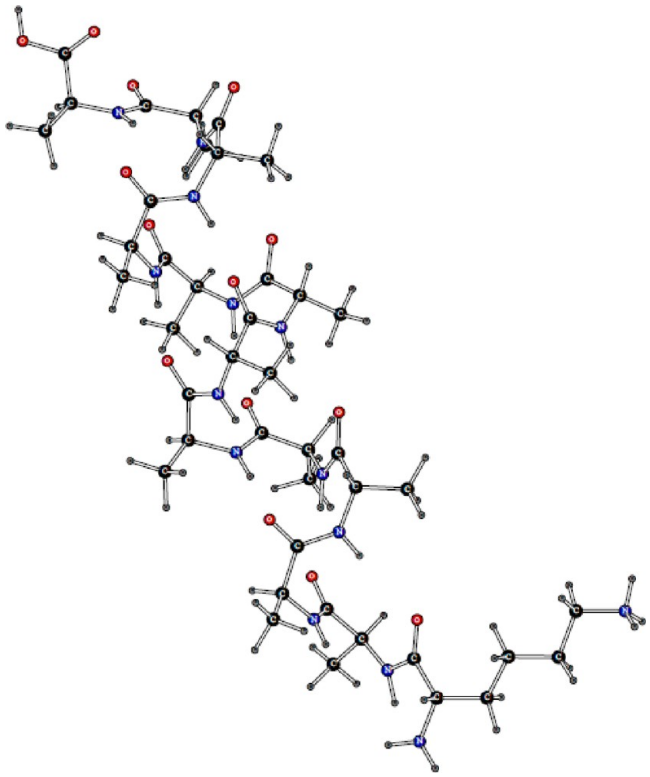

Figure 2. Ala13 structure with a lysine residue on the N-terminus.

point dipoles $\boldsymbol{\mu}$ is

$$E_{pol} = -\frac{1}{2}\sum_i \boldsymbol{\mu}_i \mathbf{E}_i^0 \quad (2)$$

where $\mathbf{E}^0$ is the electrostatic field in the absence of the induced dipoles.

$$\boldsymbol{\mu}_i = \alpha_i \mathbf{E}_i^0 + \alpha_i \sum_{j\neq i} \mathbf{T}_{ij}\boldsymbol{\mu}_j \quad (3)$$

Here, α represents scalar polarizabilities and $\mathbf{T}_{ij}$ is the dipole−dipole interaction tensor. The self-consistent eq 3 is usually solved iteratively. The first two iterations are

$$\boldsymbol{\mu}_i^{I} = \alpha_i \mathbf{E}_i^0 \quad (4a)$$

$$\boldsymbol{\mu}_i^{II} = \alpha_i \mathbf{E}_i^0 + \alpha_i \sum_{j\neq i} \mathbf{T}_{ij}\boldsymbol{\mu}_j^{I} = \alpha_i \mathbf{E}_i^0 + \alpha_i \sum_{j\neq i} \mathbf{T}_{ij}\alpha_j \mathbf{E}_j^0 \quad (4b)$$

Included in POSSIM is the second-order expression in eq 4b. It has been previously shown to yield a significant increase of the computational speed with no loss of accuracy.[8,13] The electrostatic energy also contains the pairwise-additive contribution from interactions of permanent charges:

$$E_{additive} = \sum_{i\neq j} \frac{q_i q_j}{R_{ij}} f_{ij} \quad (5)$$

The factor f_{ij} is equal to zero for 1,2 and 1,3 pairs (atoms which belong to the same valence bond or angle), to 0.5 for 1,4 interactions (atoms in the same dihedral angle), and to 1.0 otherwise.

The fixed OPLS force field contains only the additive term given in eq 5 for the electrostatic part of the Hamiltonian.

An additional feature in POSSIM which is not present in OPLS-AA is a cutoff parameter R_{cut}. If the overall distance R_{ij} is

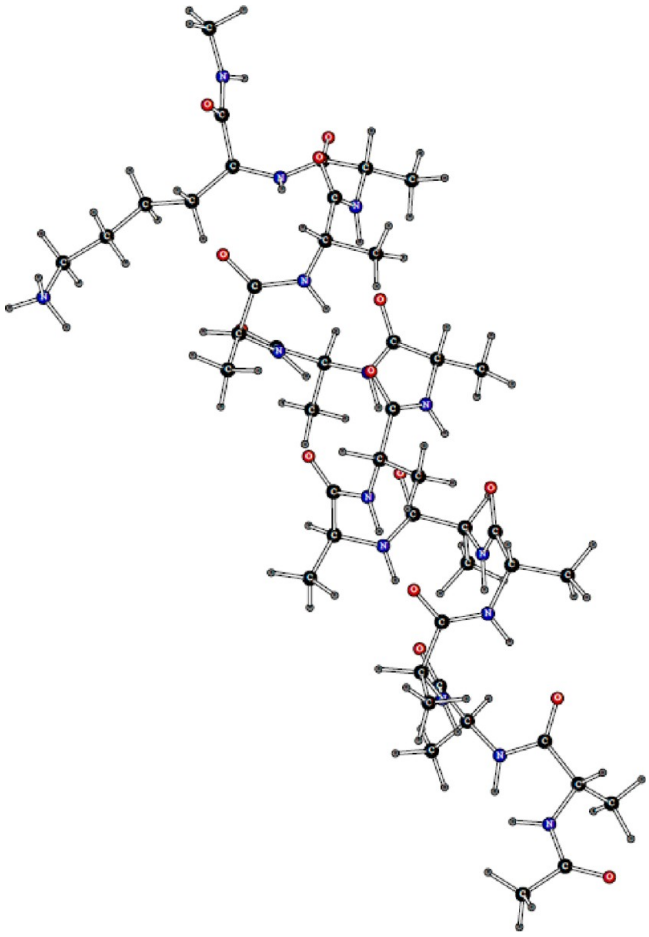

Figure 3. Ala13 structure with a lysine residue on the C-terminus with the termini themselves replaced by methyl caps.

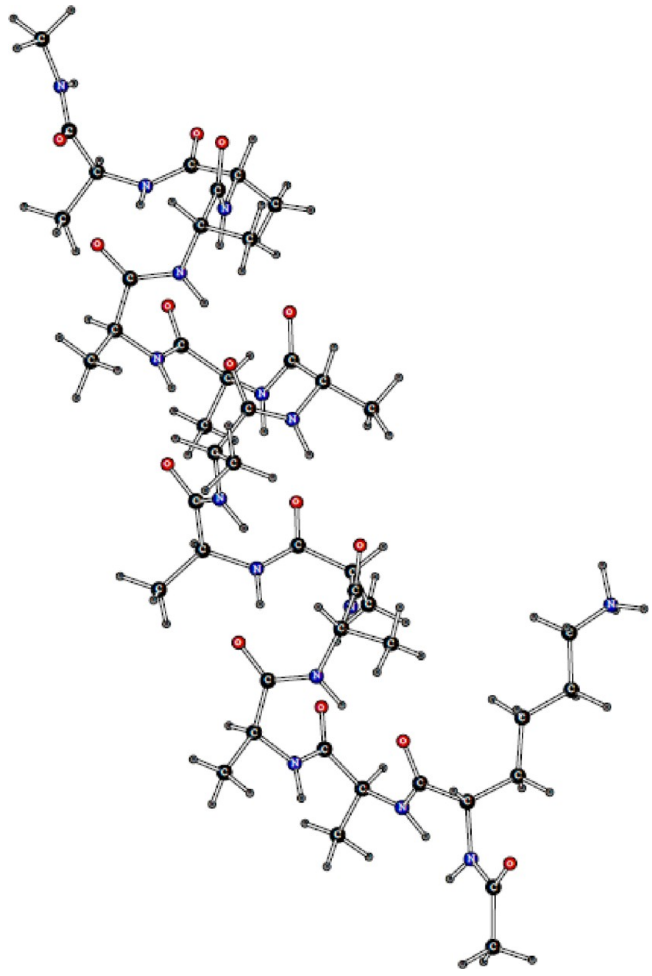

Figure 4. Ala13 structure with a lysine residue on the N-terminus with the termini themselves replaced by methyl caps.

smaller than the sum of these parameters $R_{\min}^{ij} = R_{\text{cut}}^{i} + R_{\text{cut}}^{j}$ for the atoms i and j, R_{ij} is replaced by a smoothing function

$$R_{ij}^{\text{eff}} = \left(1 - \left(\frac{R_{ij}}{R_{\min}^{ij}}\right)^2 + \left(\frac{R_{ij}}{R_{\min}^{ij}}\right)^3\right) \cdot R_{\min}^{ij} \tag{6}$$

While the approximation in eq 4b differs from the complete physical induced-dipole electrostatic model, we have demonstrated that it does not simply increase the computational speed but also does not compromise the accuracy of the simulations.[8,9,13] Moreover, the second-order approximation given by eq 4b makes the expression for the inducible dipoles into an analytical one, thus eliminating the possibility of a polarization catastrophe. This can also be useful in future developments, such as in creating continuum dielectric models, as convergence issues are known to be of importance for implicit solvation techniques.

The Rest of the Force Fields. We used the standard Lennard-Jones formalism for the van-der-Waals energy in both the OPLS-AA and POSSIM techniques:

$$E_{\text{vdW}} = \sum_{i \neq j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{R_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}}\right)^{6}\right] f_{ij} \tag{7}$$

Geometric combining rules are applied ($\varepsilon_{ij} = (\varepsilon_i \cdot \varepsilon_j)^{1/2}$, $\sigma_{ij} = (\sigma_i \cdot \sigma_j)^{1/2}$). Bond stretching and angle bending were calculated with the standard harmonic formalism, and the torsional term is obtained as

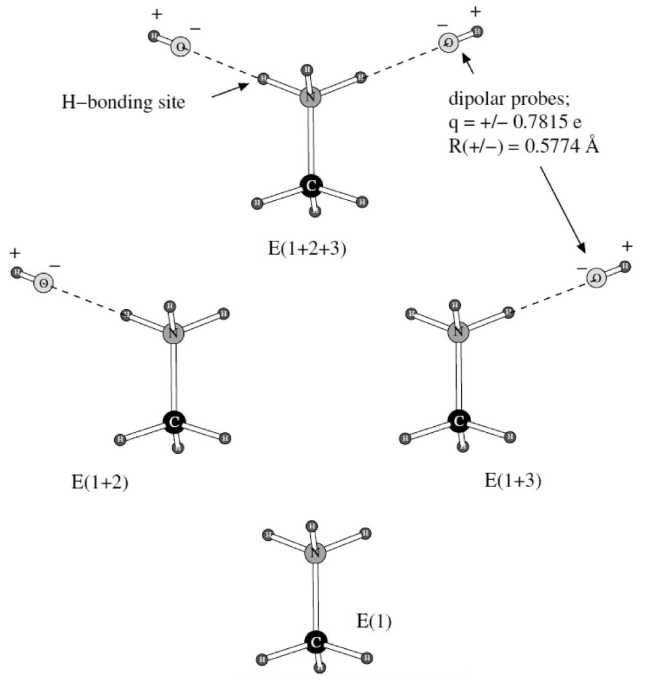

Figure 5. Calculating two- and three-body energies of the $CH_3NH_3^+$ molecule with dipolar probes.

dx.doi.org/10.1021/ct300492n | J. Chem. Theory Comput. 2012, 8, 4691−4706

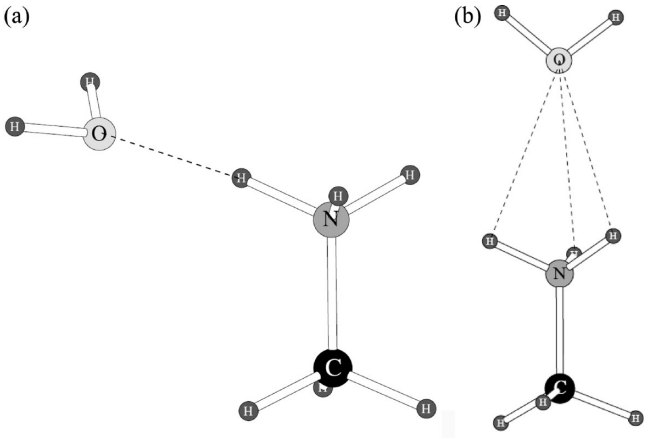

Figure 6. Gas-phase dimers of the $CH_3NH_3^+$ molecule with water. Configuration (a) was used for fitting of the parameters.

$$E_{\text{torsion}} = \sum_i \frac{V_1^i}{2}[1 + \cos(\phi_i)] + \frac{V_2^i}{2}[1 - \cos(2\phi_i)] + \frac{V_3^i}{2}[1 + \cos(3\phi_i)] \quad (8)$$

C. Additional Parameterization of the POSSIM Force Field. In addition to the previously produced parameters for the alanine part of the polypeptide, parameters for the protonated lysine residue and the protein N- and C-termini were needed, and they were obtained in the course of this project. For the lysine residue, the general scheme of this parameter fitting was the same as the one used in producing the previous version of the polarizable force field for proteins.[12] The first step was in generating parameters for the $CH_3NH_3^+$ molecule, which served as a prototype for the lysine side chain. We started with fitting the atomic polarizabilities to reproduce the three-body energies as was described, for example, in refs 8, 9, and 12. Briefly, we considered a configuration of the molecule with two electrostatic dipolar probes, composed of fixed charges with no Lennard-Jones or other parameters. The positions of the probes were chosen to correspond to hydrogen bonds. In the case of the $CH_3NH_3^+$ molecule, there is only one such pair of probes which has to be considered. It is shown in Figure 5.

Each dipolar probe consisted of two opposite charges of magnitude 0.78*e*, positioned 0.58 Å apart (for a dipole moment of 2.17D, which is similar to that of the nonpolarizable SPC/E water model[14]). The three body energies were calculated as follows:

$$E_{3\text{body}} = E(1 + 2 + 3) - E(1 + 2) - E(1 + 3) - E(2 + 3) + E(1) + E(2) + E(3) \quad (9)$$

The target quantum mechanical values of the energies were computed using density-functional theory (DFT) with the B3LYP method[15] and cc-pVTZ(-f) basis set. The Jaguar software suite[16] was employed. The resulting three-body energies were then used to fit atomic polarizablities α_i, which were assumed to be isotropic and are chosen to minimize the deviation of the three-body energy obtained with the POSSIM and the DFT calculations.

Next, atomic charges and Lennard-Jones parameters on the $-NH_3$ group were optimized to reproduce the quantum mechanical dimerization energies between the $CH_3NH_3^+$ molecule and water and the N···O distance. The quantum mechanical calculations were carried out following the general extrapolation protocol involving LMP2 data with the cc-pVTZ(-f) and cc-pVQZ basis sets, which has been described elsewhere.[17] The geometry of the complex is shown in Figure 6a. Finally, the H−C−N−H torsional parameters were fitted to reproduce the LMP2/cc-pVTZ(-f) quantum energy differences for angle values of 120° and 180°, as well as 150° and 180°. The parameters for the $-CH_3$ group were adopted from the general POSSIM aliphatic set[8] without any changes.

It should be noted that the dipole moment of the $CH_3NH_3^+$ molecule is not aligned with any of the N−H bonds. Therefore, it could be possible to use a different configuration for parameter

Table 1. Atomic Parameters Used for the $CH_3NH_3^+$ Molecule

atom	q, electrons	σ, Å	ε, kcal/mol	α, Å^3	R_{cut}
N	−0.08	3.60	0.280	1.0000	0.80
H(N)	0.36	0.00	0.000		0.80
C	−0.18	3.50	0.066	1.9728	0.80
H(C)	0.06	2.50	0.030		0.80

Table 2. Results of Torsional Fitting for $CH_3NH_3^+$, $C_2H_5NH_3^+$ Molecules

	energy, kcal/mol		
system/angle	LMP2/cc-pVTZ(-f)	POSSIM	error, kcal/mol
	$CH_3NH_3^+$, H−C−N−H		
180°	0.000	0.000	0.000
150°	1.131	1.064	0.067
120°	2.343	2.282	0.061
average error			0.043
	$C_2H_5NH_3^+$, H−C−C−N		
180°	0.000	0.000	0.000
150°	1.540	1.542	0.002
120°	3.284	3.305	0.021
average error			0.008

Table 3. Values of Fitted Torsional Parameters, kcal/mol

dihedral angle	V_1	V_2	V_3
	$CH_3NH_3^+$, $C_2H_5NH_3^+$, lysine		
H−C−N−H	0.000	0.000	0.249
H−C−C−N	0.000	0.000	0.210
C−C−N−H	0.000	0.000	0.355
	lysine residue only		
N−C−C−C (χ_1)	−3.862	−0.355	5.035
C(O)−C−C−C (χ_1')	−6.000	−3.905	0.454
C−C−C−N (χ_4)	0.286	−2.595	−5.020
C−C−C−C (lys side-chain only)	0.980	−0.570	0.640
	neutral C−terminus		
C−C−C−O	2.840	0.0	2.090
N−C−C−O	4.343	−0.698	−3.634
	deprotonated C terminus		
N−C−C−O	5.000	2.980	0.000
C−C−C−O	0.000	0.370	1.000
	neutral N-terminus		
H−C−C−N	0.000	0.000	0.210
H−C−N−H	0.000	0.000	0.246
C−C−N−H	0.151	1.648	0.920
	protonated N terminus		
H−C−C−N	0.000	0.000	0.210
H−C−N−H	0.000	0.000	0.246
C−C−N−H	0.792	3.914	−0.435

dx.doi.org/10.1021/ct300492n | J. Chem. Theory Comput. 2012, 8, 4691−4706

Table 4. Quantum Mechanical (LMP2/cc-pVTZ(-f)) and POSSIM Conformational Energies (kcal/mol) and Angles for Lysine Dipeptide

	energy							
conf.	QM	POSSIM	ϕ	ψ	χ_1	χ_2	χ_3	χ_4
1	17.16	17.15	−98.50°	5.00°	−73.70°	172.20°	67.20°	174.80°
2	21.45	21.83	−147.40°	−53.90°	−172.80°	170.00°	67.20°	174.30°
3	16.70	17.41	−148.80°	158.20°	47.60°	73.10°	69.30°	70.00°
4	0.00	−0.74	−87.50°	65.40°	−162.10°	−162.20°	57.70°	53.70°
5	15.21	14.98	−88.60°	80.80°	−73.30°	163.50°	61.90°	174.30°
6	13.25	13.12	−155.10°	142.60°	20.90°	64.40°	66.90°	169.80°
av. error		0.50						

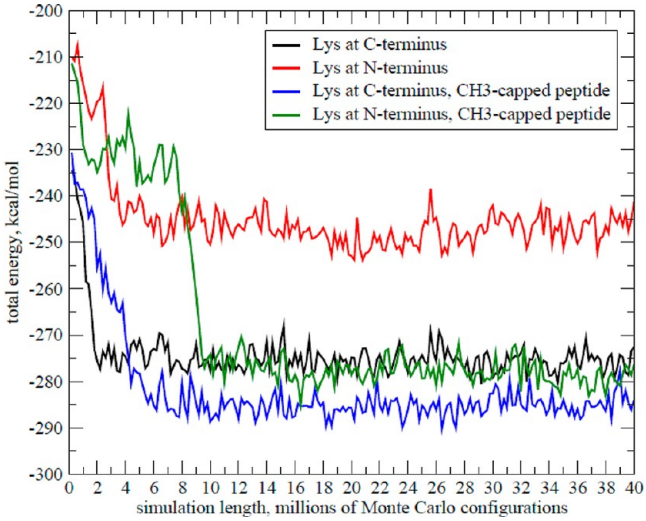

Figure 7. Energy of the polypeptides simulated with the OPLS-AA force field as a function of the number of Monte Carlo configurations.

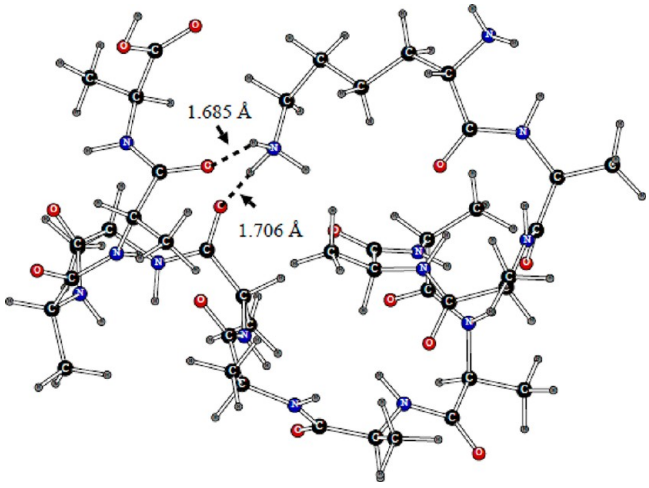

Figure 9. Snapshot of the OPLS-AA Lys-N structure after 40 × 10^6 Monte Carlo configurations.

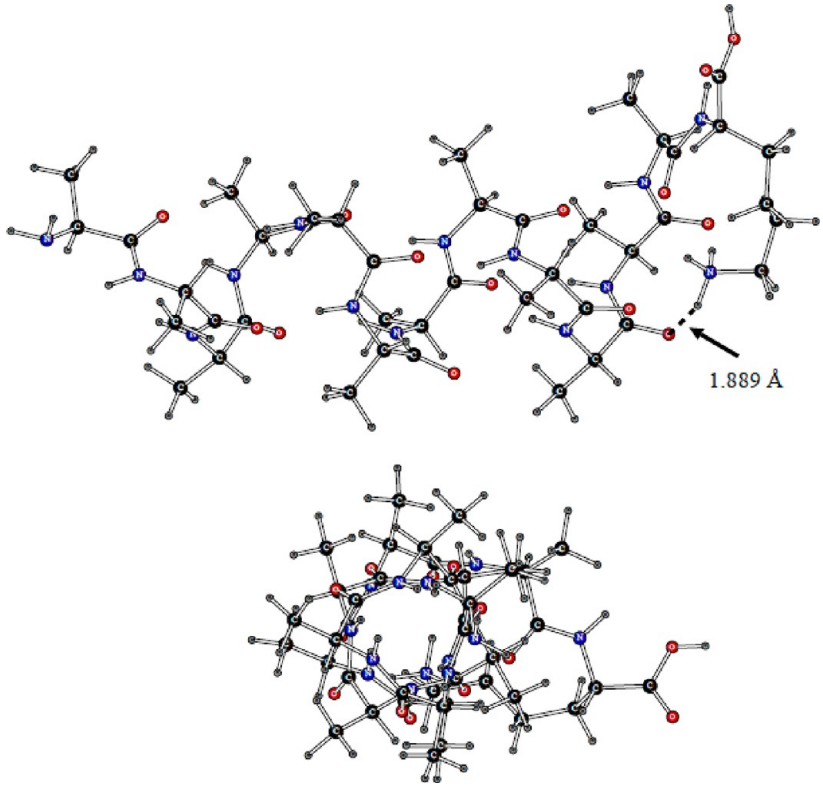

Figure 8. Snapshot of the OPLS-AA Lys-C structure after 40 × 10^6 Monte Carlo configurations.

dx.doi.org/10.1021/ct300492n | J. Chem. Theory Comput. 2012, 8, 4691−4706

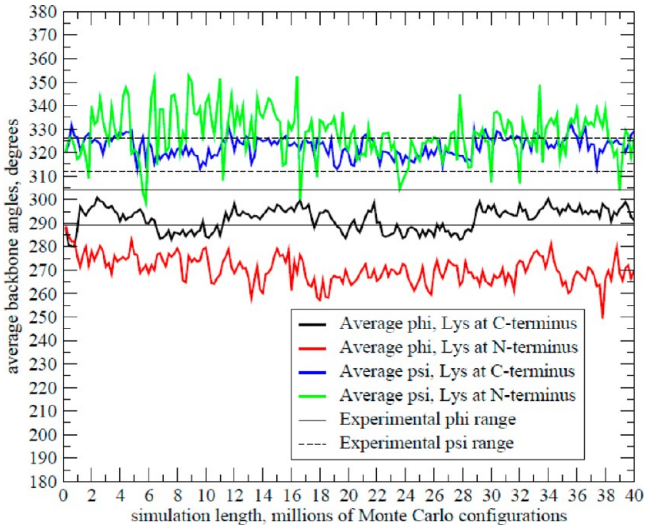

Figure 10. Averaged ϕ and ψ backbone angles in Lys-C and Lys-N systems as a function of the number of Monte Carlo configurations.

Table 5. Average ϕ and ψ Angles in the Peptide Backbone

force field	system	ϕ	ψ	$-(360° - \phi + 360° - \psi)$
OPLS	Lys at C terminus	291.8°	322.6°	105.6°
OPLS	Lys at N terminus	268.5°	325.2°	126.3°
OPLS	Lys at C terminus, CH_3 − capped peptide	291.9°	319.4°	108.7°
OPLS	Lys at N terminus, CH_3 − capped peptide	292.8°	327.5°	99.7°
POSSIM	Lys at C terminus	283.7	332.6	103.7°
POSSIM	Lys at N terminus	274.3	334.8	110.9°
POSSIM	Lys at C terminus, CH_3 − capped peptide	292.7°	325.8°	101.5°
POSSIM	Lys at N terminus, CH_3 − capped peptide	281.1°	317.2°	121.7°
POSSIM, in water	Lys at C terminus	277.8°	335.1°	107.1°
POSSIM, in water	Lys at N terminus	281.0°	326.9°	112.1°
POSSIM, in water	Lys at C terminus, CH_3 − capped peptide	289.6°	325.5°	104.9°
POSSIM, in water	Lys at N terminus, CH_3 − capped peptide	282.7°	336.5°	100.8°
experiment[a]		296°	319°	105°

[a]Reference 19

fitting in both the three-body and dimerization calculations. This configuration is shown in Figure 6b, and it has the water and the $CH_3NH_3^+$ dipole moments colinear. However, our calculations demonstrated that this configuration does not correspond to either the global or a local minimum in both POSSIM and quantum mechanical (LMP2/cc-pVTZ(-f)) simulations. Therefore, we employed the hydrogen-bonded structures shown in Figures 5 and 6a in our parameter development. If the geometry of the dimer is constrained to be the same as in Figure 6b (with the $CH_3NH_3^+$−water distance being variable), the extrapolated quantum mechanical binding energy is −15.51 kcal/mol and the N···O distance is 2.803 Å. The corresponding POSSIM values are −15.99 kcal/mol and 2.998 Å. Thus, the accuracy (achieved with no specific fitting) is within 0.5 kcal/mol and 0.2 Å. This is quite acceptable given that this configuration does not correspond to even a local minimum on the potential energy surface.

Furthermore, torsional parameters for the H−C−C−N and C−C−N−H dihedral angles in $C_2H_5NH_3^+$ were also fitted for

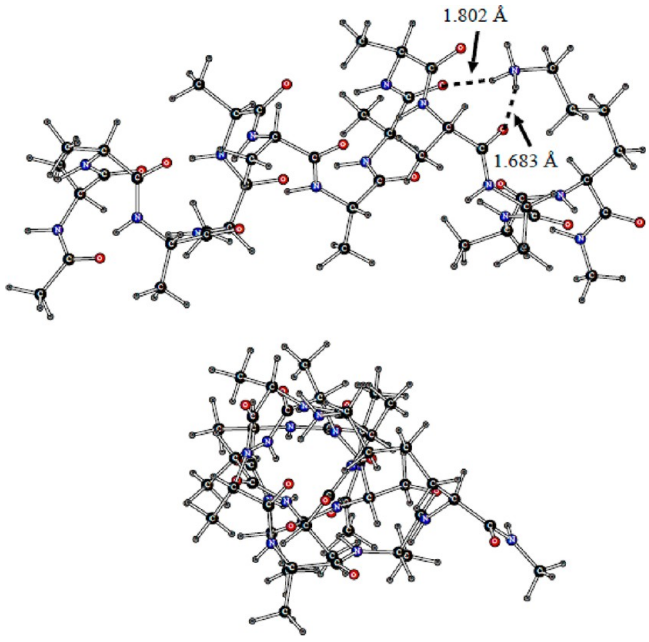

Figure 11. Snapshot of the OPLS-AA Lys-C structure after 40 × 10^6 Monte Carlo configurations. CH_3-capped polypeptide.

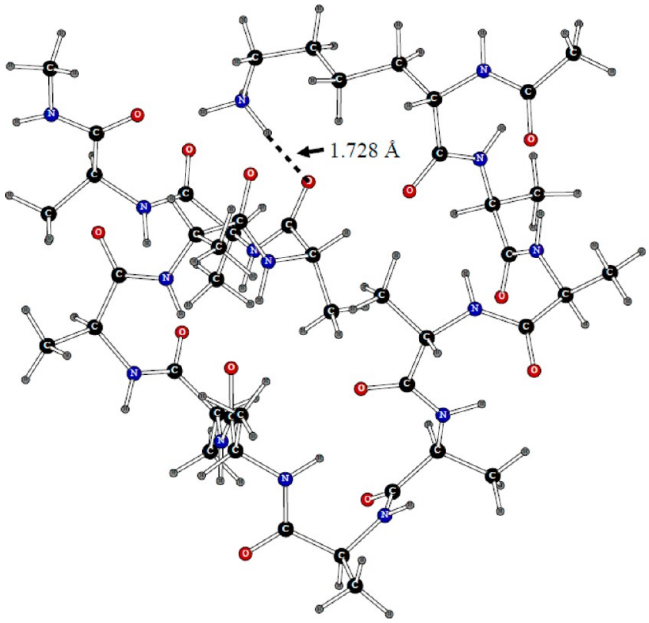

Figure 12. Snapshot of the OPLS-AA Lys-N structure after 40 × 10^6 Monte Carlo configurations. CH_3-capped polypeptide.

further use in the lysine residue. To achieve this, the H−C−C−N dihedral angle was fixed at 180°, 150°, and 120°, and comparison was once again made between the POSSIM and LMP2/cc-pVTZ(-f) relative conformational energies. The H−N−C−C angle was fixed at 180°. The torsional parameters were adjusted to minimize the difference.

Having obtained these new parts of the force field, we have transferred them to the protonated lysine residue. The only other required development was in fitting the appropriate side-chain torsional parameters (we followed our usual protocol of protein force field development[10,12] and took the backbone part of the lysine potential from the alanine parameter set[9]). The lysine side chain torsional parameters were fitted for the χ_1 and χ_1' angles

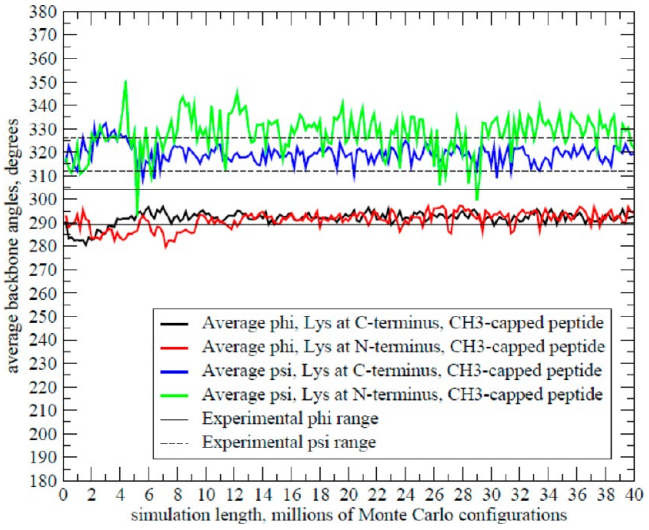

Figure 13. Averaged ϕ and ψ backbone angles in Lys-C and Lys-N systems as a function of the number of Monte Carlo configurations for the methyl-capped OPLS-AA systems.

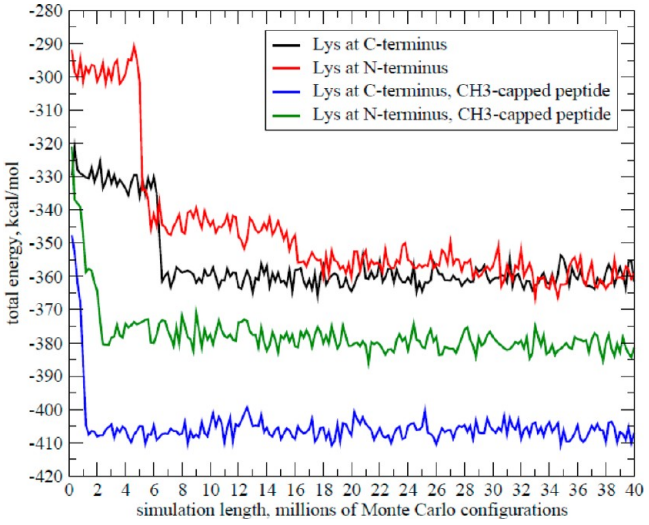

Figure 14. Energy of the polypeptides simulated with the polarizable POSSIM force field as a function of the number of Monte Carlo configurations.

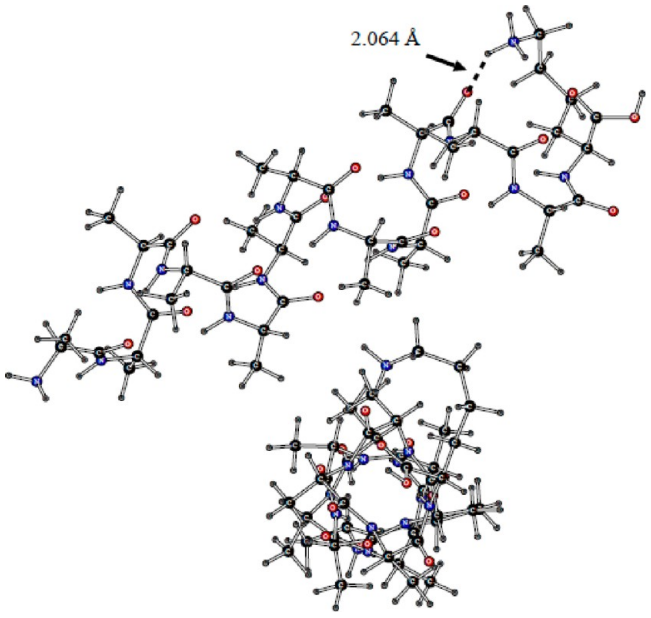

Figure 15. Snapshot of the polarizable POSSIM Lys-C structure after 40×10^6 Monte Carlo configurations.

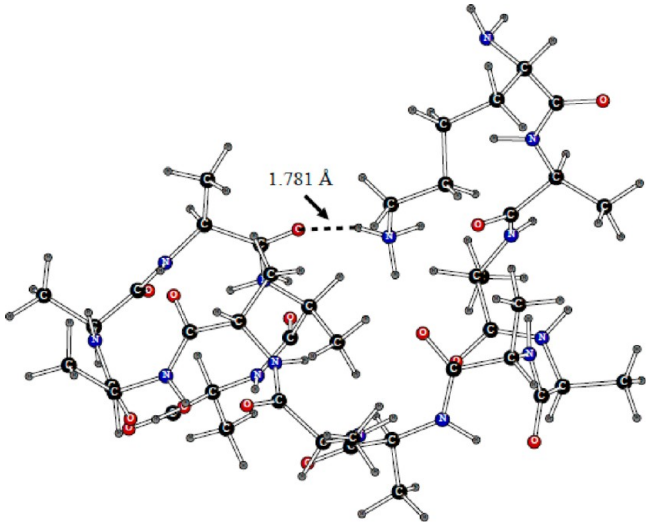

Figure 16. Snapshot of the polarizable POSSIM Lys-N structure after 40×10^6 Monte Carlo configurations.

(N−C−C−C and C(=O)−C−C−C, respectively), the χ_2 angle C−C−C−C, and the χ_4 angle C−C−C−N. The remaining H−C−C−C and H−C−C−H torsional Fourier coefficients were taken directly from our previously produced set for aliphatic hydrocarbons.[8]

We also had to produce POSSIM parameters for the deprotonated and protonated protein termini. The nonbonded parts of these sets were taken from the POSSIM CH_3COOH, CH_3COO^-, and CH_3NH_2 systems[18] and the $CH_3NH_3^+$ cation parameter set developed in the course of this work. The torsional component of the parameters was obtained by fitting the POSSIM against the LMP2/cc-pVTZ(-f) quantum mechanical energies for the rotamers having different angles for the terminal groups (−COOH, $-COO^-$, $-NH_2$, and $-NH_3^+$). The neutral versions of the termini were used in gas-phase calculations, and the ionized ones were employed in the simulations of solvated polypeptides.

Having completed this part of the project, we had all the necessary POSSIM parameters available for the actual polypeptide runs.

D. Monte Carlo Simulations of the Polypeptide. We carried out simulations of a tridecaalanine (ala-13) peptide with one additional lysine residue in the gas phase and in aqueous solution at 25 °C. The initial structure was set at the α-helix conformation, with $\phi = 296°$ and $\psi = 319°$, and the simulations proceeded with all the degrees of freedom completely unconstrained.

The simulations consisted of at least 40×10^6 Monte Carlo configurations to ensure convergence. A 7 Å dipole−dipole cutoff was used. The electrostatic interactions were quadratically feathered over the last 0.5 Å before the cutoff distance.

When solvated simulations were carried out with POSSIM, ca. 480 water molecules with the previously developed POSSIM water parameters[8] were employed.

dx.doi.org/10.1021/ct300492n | J. Chem. Theory Comput. 2012, 8, 4691−4706

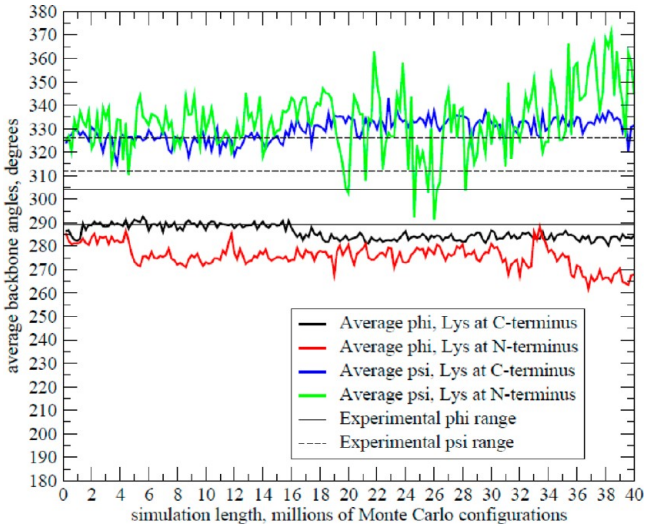

Figure 17. Averaged ϕ and ψ backbone angles in Lys-C and Lys-N systems as a function of the number of Monte Carlo configurations for the normally terminated polarizable POSSIM systems.

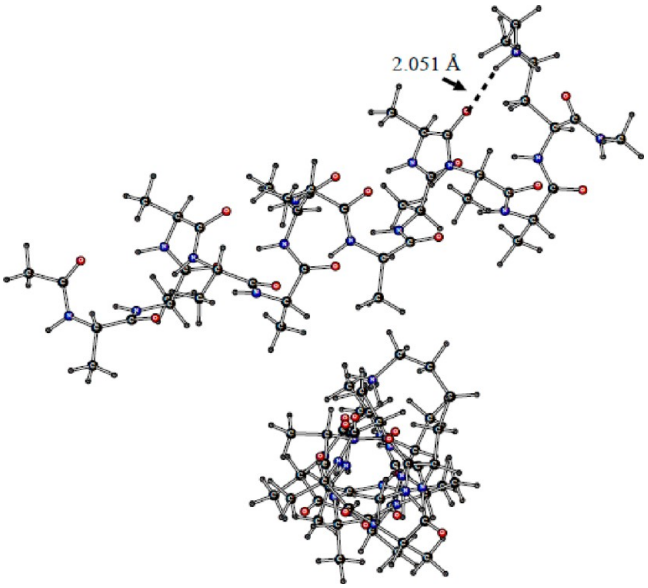

Figure 18. Snapshot of the POSSIM Lys-C structure after 40×10^6 Monte Carlo configurations. CH_3-capped polypeptide.

All the calculations which did not involve quantum mechanics (i.e., geometry optimizations and Monte Carlo runs) were performed with our previously introduced POSSIM software suite.[8] Whenever comparison with the fixed-charge OPLS-AA force field was done, the OPLS-AA results were also obtained with the POSSIM program.

III. RESULTS AND DISCUSSION

A. Parameterization of $CH_3NH_3^+$, $C_2H_5NH_3^+$, Lysine Residue, and C- and N-Termini. The first step was in fitting the polarizabilities for the $CH_3NH_3^+$ ion. The methyl group parameters were not changed from the previously created aliphatic set. The hydrogen atoms were not polarizable, and the final value of the nitrogen polarizability was equal to 1.00 Å^3. The error in the three-body energy was 0.27 kcal/mol. The two-body electrostatic energy of binding between the molecule and one of the probes shown in Figure 5 was reproduced within

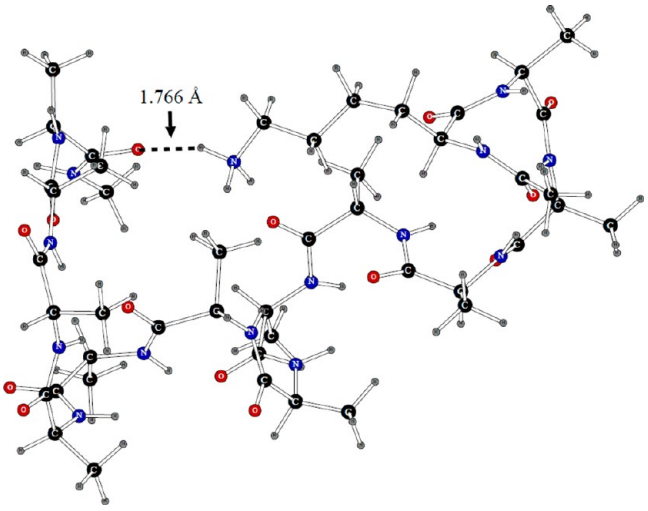

Figure 19. Snapshot of the POSSIM Lys-N structure after 40×10^6 Monte Carlo configurations. CH_3-capped polypeptide.

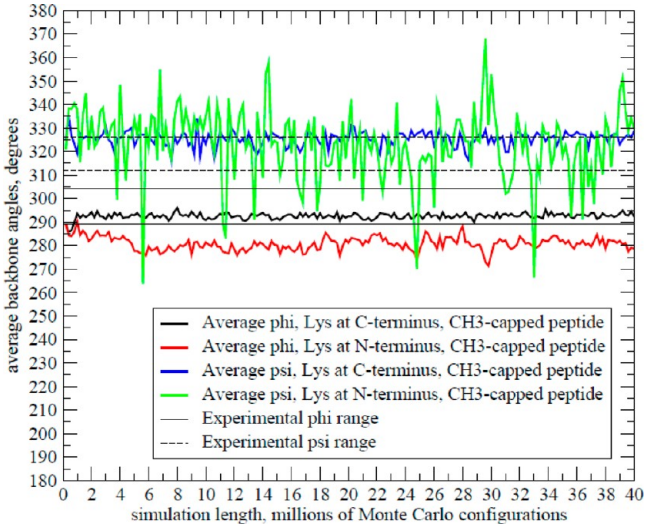

Figure 20. Averaged ϕ and ψ backbone angles in Lys-C and Lys-N systems as a function of the number of Monte Carlo configurations for the CH_3-terminated polarizable POSSIM systems.

1.60 kcal/mol from the quantum mechanical data. It is not unusual to have errors of ca. 2.0 kcal/mol in this parameter,[12] and so this error is calculated for reference only.

The electrostatic charges and Lennard-Jones parameters were adjusted to reproduce the quantum mechanical binding energy and N···O distance of the dimer shown in Figure 6. The calculated POSSIM values of the energy and the distance were −20.10 kcal/mol and 2.81 Å, respectively, while their quantum mechanical counterparts computed as described in the Methods section were −19.63 kcal/mol and 2.69 Å. Thus, the errors were in the acceptable range of ca. 0.5 kcal/mol and 0.12 Å for the energy and the distance. All the atomic parameters are listed in Table 1, where α stands for the atomic polarizability and R_{cut} designates the distance below which the effective R is scaled to avoid unphysical growth of the magnitude of the electrostatic energy, as described in ref 8.

The results of the torsional fitting for the $CH_3NH_3^+$, $C_2H_5NH_3^+$, molecules are shown in Table 2. The final fitted values of the torsional parameters for these systems are presented in Table 3. It can be seen that adjusting the V_3 torsional term within very reasonable limits leads to a high accuracy level.

dx.doi.org/10.1021/ct300492n | J. Chem. Theory Comput. 2012, 8, 4691−4706

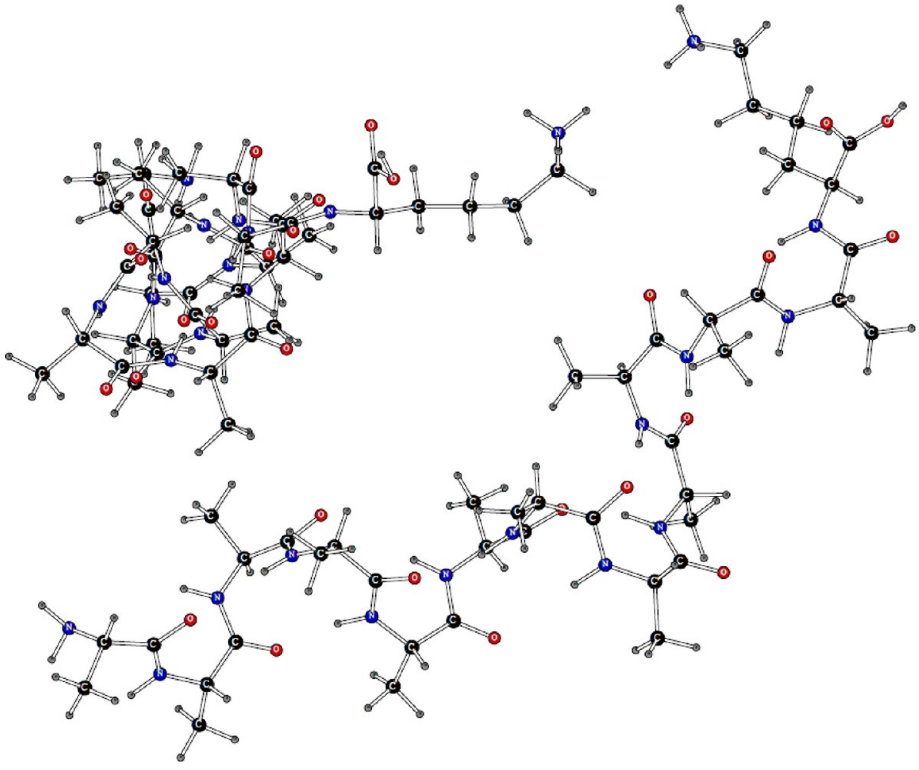

Figure 21. Snapshot of the POSSIM Lys-C structure in water after 40×10^6 Monte Carlo configurations. Normally terminated polypeptide. Water molecules are not shown.

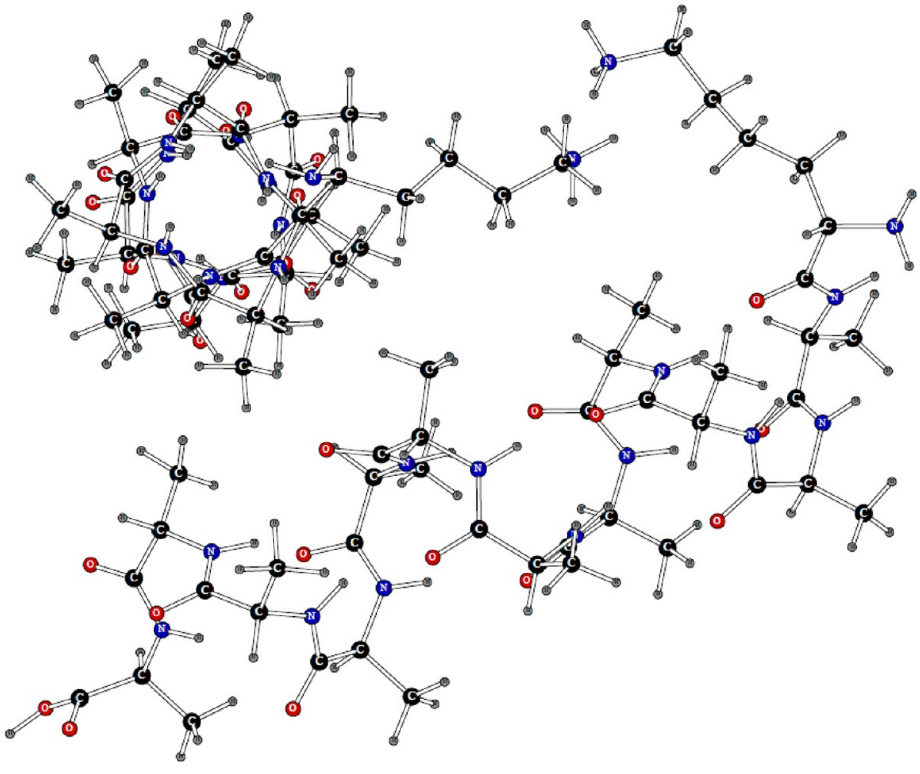

Figure 22. Snapshot of the POSSIM Lys-N structure in water after 40×10^6 Monte Carlo configurations. Normally terminated polypeptide. Water molecules are not shown.

The $CH_3NH_3^+$ molecule has an average error of 0.043 kcal/mol in calculating the H−C−N−H torsional energy, and the $C_2H_5NH_3^+$ one allows an even lower error of 0.008 kcal/mol for the H−C−C−N torsional energy.

The torsional fitting for the lysine side-chain parameters was carried out in the same way as described in refs 10 and 12. The resulting parameters are shown in the second part of Table 3. The fitting was done to minimize deviations from the quantum

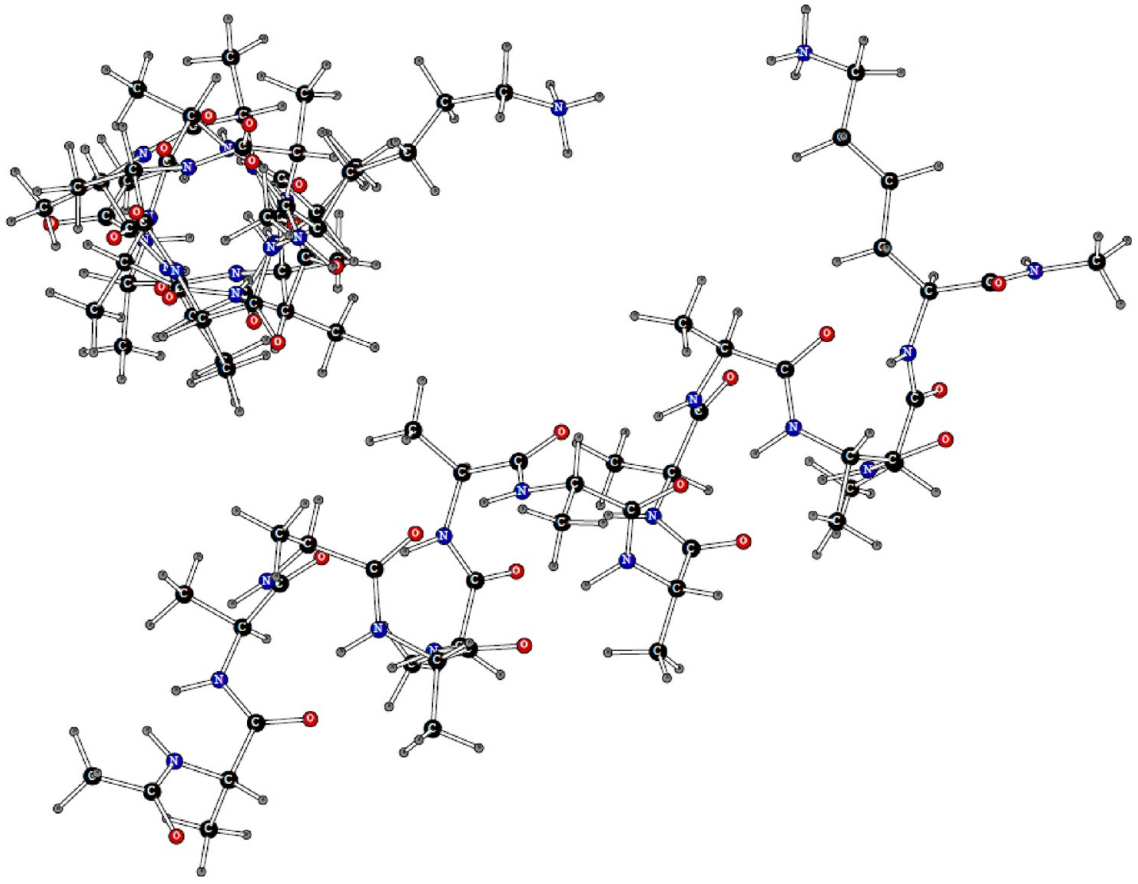

Figure 23. Snapshot of the POSSIM Lys-C structure in water after 40×10^6 Monte Carlo configurations. CH_3-capped polypeptide. Water molecules are not shown.

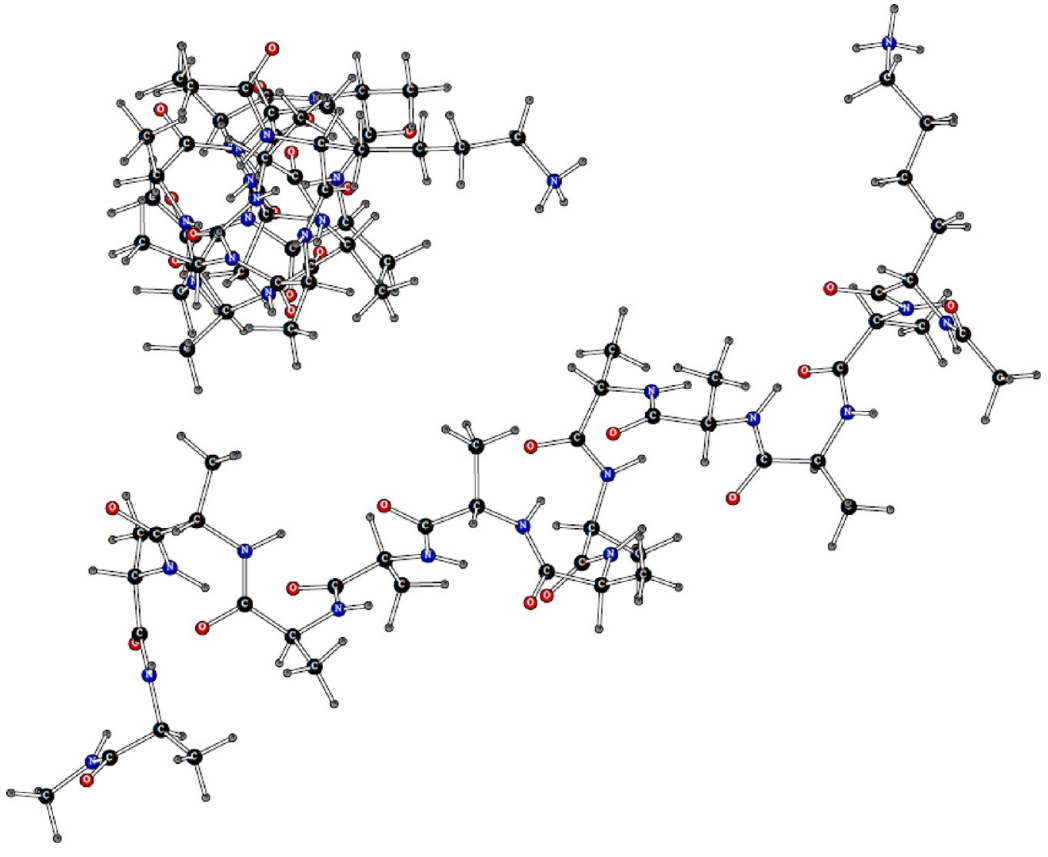

Figure 24. Snapshot of the POSSIM Lys-N structure in water after 40×10^6 Monte Carlo configurations. CH_3-capped polypeptide. Water molecules are not shown.

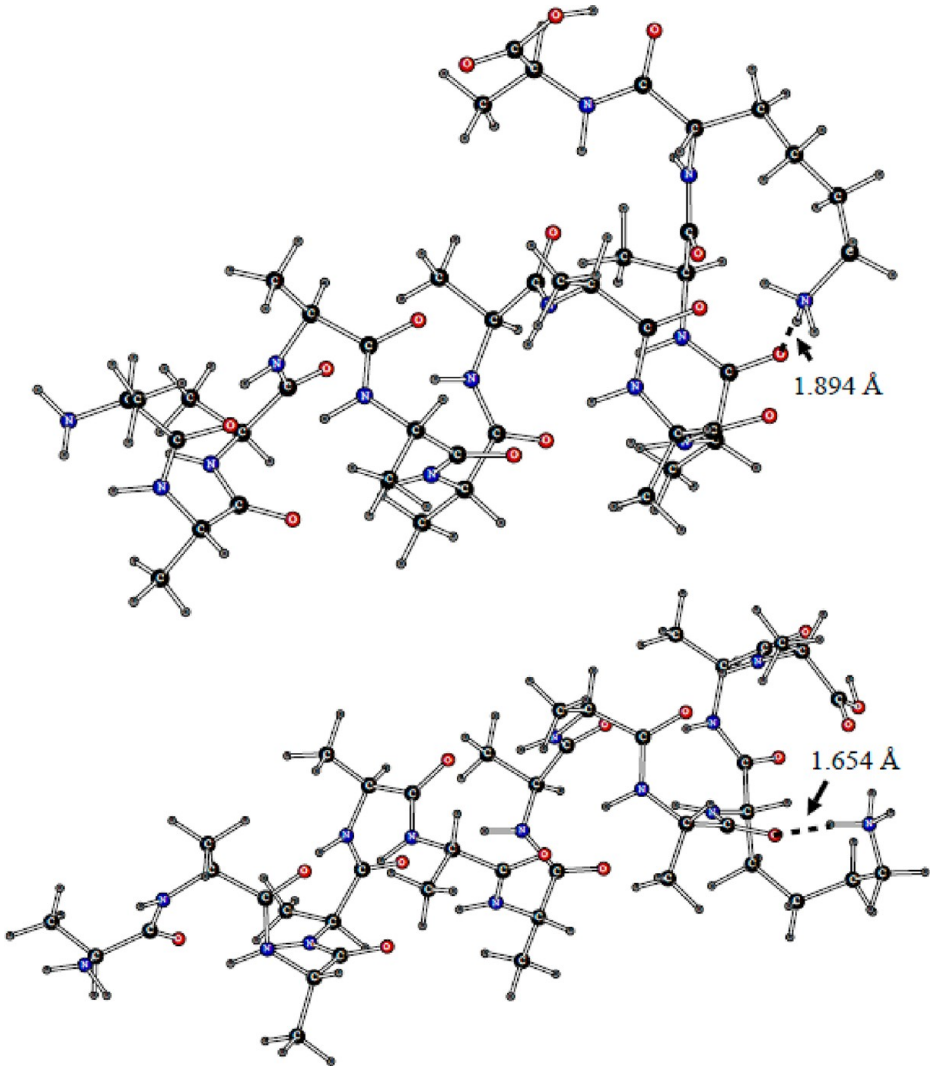

Figure 25. Snapshot of the OPLS Lys-C-1 (top) and Lys-C-2 (bottom) structures in the gas phase after 40×10^6 Monte Carlo configurations. Normally terminated polypeptide.

mechanical conformational energies for the lysine dipeptide. Following our standard methodology of torsional fitting for charged protein residues,[10,12] we kept the key dihedral angles fixed; thus their deviation from the quantum mechanical results is zero. The average error in the conformational energy is only 0.50 kcal/mol, which compares favorably with the result for the previous version of the polarizable force field (0.59 kcal/mol) and the fixed-charge OPLS-AA/L force field result of 0.88 kcal/mol.[12] The results are shown in Table 4.

The parametrization of the peptide and protein termini with the $-COOH$, $-COO^-$, $-NH_2$, and $-NH_3^+$ groups resulted in average rotamer energy errors of 0.331, 0.181 , 0.319 and 0.290 kcal/mol, respectively. Such errors in torsional energies are acceptable in the general framework of our force field development for proteins and peptides. The values of the fitted torsional parameters for these terminal groups are given in Table 3.

Once all the above parameters have been finalized, we proceeded to the simulations of the alanine13 α-helices with one additional lysine residue.

B. Simulations of the Polypeptides. *Using the OPLS-AA Force Field with Normally Terminated Polypeptides.* The first step in simulating the actual tridecaalanine polypeptides with an additional protonated lysine residue was in Monte Carlo runs for these peptides modeled with the OPLS-AA force field. The initial structures of these systems with the Lys residue at the C- and N-termini were the same as shown in Figures 1 and 2, respectively. In this case, the termini had the normal $-COOH$ and $-NH_2$ groups (the C-terminus was protonated and the N-terminus deprotonated as compared to the usual protein forms in aqueous solutions since the ionic forms of the termini are favored by hydration but are less stable in the gas phase). The resulting total energy of the systems as a function of the simulation length is shown in Figure 7. The black line corresponds to the lysine and the C-terminus and the red one, to the lysine residue at the N terminus. We will term them Lys-C and Lys-N, respectively.

It should be noted that the absolute values of the energies are not important since the chemical structure of the species is different, and force fields contain terms which may lead to different "baseline" energy values. The important part is in the behavior of the energies with progressing Monte Carlo simulations. One can observe the following pattern. The Lys-C polypeptide achieves equilibrium rather quickly after ca. 2.2×10^6 configurations, and the total energy stays constant within a few kilocalories per mole. At the same time, the Lys-N peptide takes another million configurations to equilibrate (and then

dx.doi.org/10.1021/ct300492n | J. Chem. Theory Comput. 2012, 8, 4691−4706

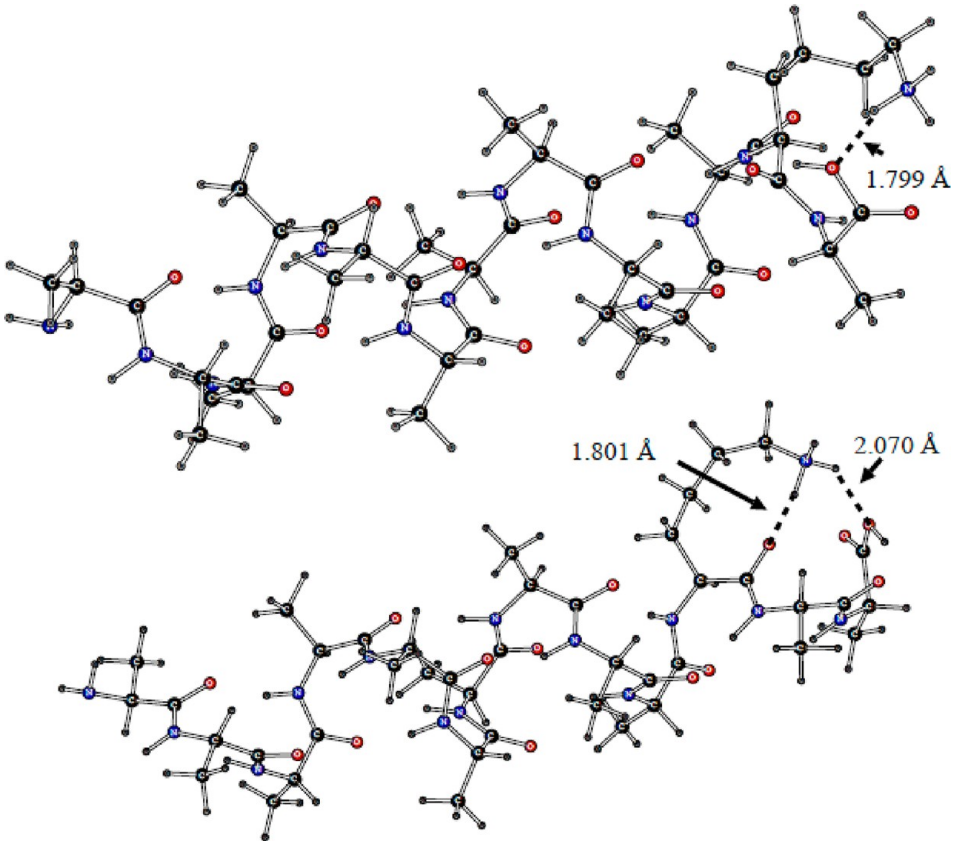

Figure 26. Snapshot of the POSSIM Lys-C-1 (top) and Lys-C-2 (bottom) structures in the gas phase after 40×10^6 Monte Carlo configurations. Normally terminated polypeptide.

there seems to be another slight drop in the energy after roughly 17×10^6 configurations from the start of the simulations).

It is interesting to take a look at the final equilibrated structures. They are given in Figures 8 and 9 for the Lys-C and Lys-N systems, respectively.

It should be emphasized that these are not averaged geometries but snapshots, and thus the hydrogen bond lengths should be viewed as representative only and not as average lengths. But the trend is clear. The Lys-C polypeptide stays in the form of an α-helix. It is apparently stabilized by the protonated Lys residue making a hydrogen bond to the backbone oxygen atom of one of the neighboring alanine residues. At the same time, the Lys-N structure collapses into an essentially globular and much less ordered shape, with the lysine residue hydrogen bonded to oxygens at the other side of the molecule, which is close to the C-terminus. These conclusions as such are consistent with the previous findings[4] and offer an additional demonstration of the fact that a polyalanine α-helix can be stable or unstable in the gas phase, depending on the location of the lysine substitution.

In order to gain additional insight into the difference between the Lys-C and Lys-N results, let us turn to a more quantitative measure of the helicity and order of the system. Let us consider the backbone ϕ and ψ angles. Given in Figure 10 are the averaged values of these angles for the two systems we have been discussing (Lys-C and Lys-N).

The thin solid lines and the thin dashed lines show the experimental ranges of the ϕ and ψ angles, respectively (296° and 319°, with a 7° uncertainty[19]). These values, averaged over the last 20×10^6 Monte Carlo configurations, are also shown in Table 5. The last column in the table is essentially the angle by which the helix turns, per one residue. The Lys-C and Lys-N molecules simulated with the OPLS-AA force field correspond to the data in the first two rows of the table.

The general trend from the graphs and the tabulated data is clear. The Lys-C molecule demonstrates the backbone angles which are basically falling within the experimentally observed limits for ϕ and ψ. The overall change of −105.6° per residue is only ca. 0.6° away from the average experimental value of 105°. The helical order is obviously preserved. At the same time, the alanine polypeptide with the lysine residue at the N-terminus denaturates. And not only the average backbone angles have a greater tendency to be out of the appropriate experimental limits, but the averaged angles (especially the ψ angle) demonstrate a greater range of values, which is natural given that the disordered structure is likely to be less rigid than an ordered helix. It also looks like it is the ϕ angle which is, on average, responsible for most of the denaturation. The average value of ϕ is down 23.3° away from its counterpart in Lys-C, while the ψ one grows by only 2.6° (as shown in Table 5). Overall, the average angle per residue is now 126.3°, about 20° more in magnitude than for the Lys-C system. It appears that the denaturation in this case proceeds via the helix winding more tightly.

Using the OPLS-AA Force Field with CH_3-Capped Polypeptides. It is not uncommon in computational studies to simulate peptides and proteins by replacing the proper C- and N-termini with methyl caps. While this simplifies the structure, the question remains as to whether this procedure significantly perturbs the physical picture of the system involved or if it can be used safely and still lead to the same or similar conclusions as simulations with the actual C- and N-terminal groups.

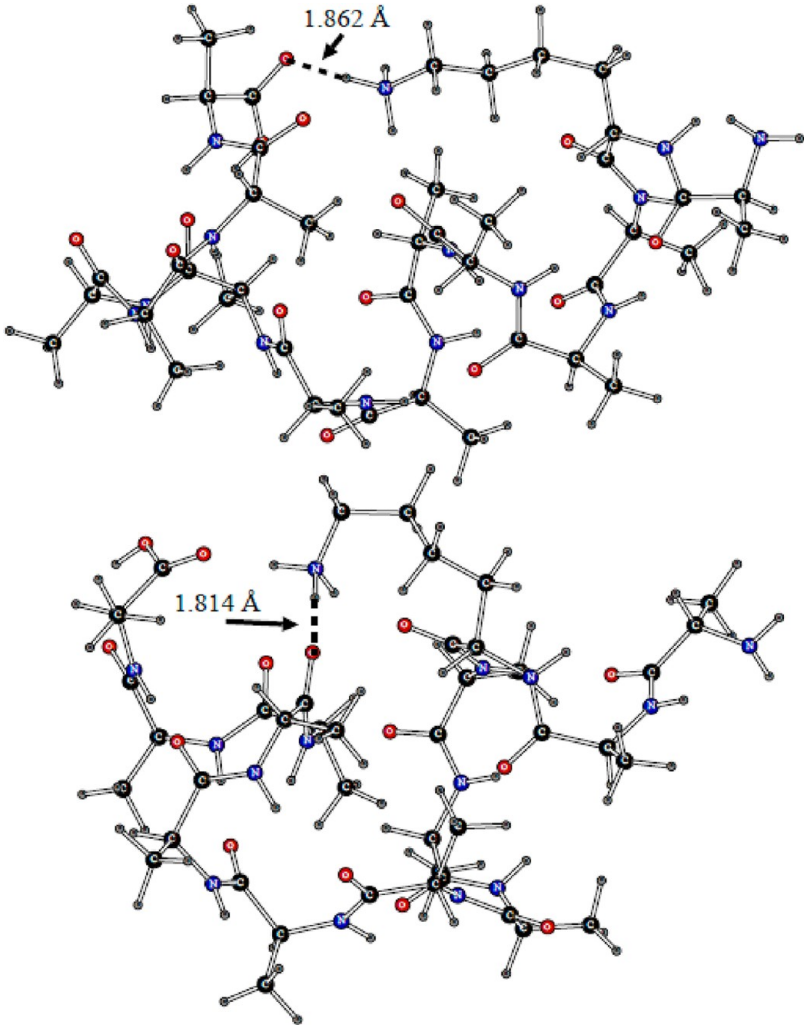

Figure 27. Snapshot of the OPLS Lys-N+1 (top) and Lys-N+2 (bottom) structures in the gas phase after 40 × 10^6 Monte Carlo configurations. Normally terminated polypeptide.

In order to test the validity of the methyl-capping approach, we have simulated modified Lys-C and Lys-N molecules with the termini replaced by CH_3 caps. These structures are shown in Figures 3 and 4, and they were modeled with the OPLS-AA force field.

The total energies of these systems as a function of the Monte Carlo simulation length are shown in Figure 7 along with those for the helices with the true termini. It can be seen that the equilibration behavior is generally the same as for the normally terminated peptide, although it takes longer (almost 10 × 10^6 Monte Carlo configurations) for the Lys-N structure to reach the equilibrium. Snapshots of the Lys-C and Lys-N structures with methyl caps are given in Figures 11 and 12, respectively.

Once again, the lengths of the shown hydrogen bonds are measured on the snapshots and are not averaged, and thus their values should be viewed as representative only. The general behavior of the systems is similar to that of the normally capped polypeptides. The system with the protonated lysine residue near the normal C-terminus site demonstrates stability of the helical structure. The Lys tends to form hydrogen bonds to backbone oxygen atoms of the neighboring alanine residues. In the case of the lysine being located near the place where the N-terminus would be, the structure collapses, and the Lys forms hydrogen bonds to the alanine residues on the opposite part of the system. Therefore, the qualitative behavior of these peptides does not seem to depend on whether the normal termini or methyl caps are placed at the ends of the structures.

Let us now study the more quantitative measure offered by the backbone ϕ and ψ dihedral angles. The graphs of these angles as a function of the simulation length for the CH_3-capped peptides are given in Figure 13.

In this case, average values of the angle ϕ for Lys-C and Lys-N are closer to each other than for the normally terminated peptides. The computed average values of angle ϕ are mostly within the experimental range, although they do show a certain tendency to be too low. The values of ψ are both more flexible and are largely higher than the experimental values. The average values are given in Table 5. The average change per residue for the methyl-capped Lys-N is −99.7°; thus the denaturation in this case seems to proceed via unwinding of the helix.

While the behavior of the averaged ϕ angles for normally and methyl-capped Lys-N helixes simulated with the OPLS force field is somewhat different, there is still a significant similarity between these two kinds of systems. First of all, in both cases, the Lys-C versions stay helical, and the Lys-N ones denaturate. Second, even the angles for the stable Lys-C versions are similar regardless of the capping. It would probably be reasonable to conclude that using methyl caps can lead to some deviations of

dx.doi.org/10.1021/ct300492n | J. Chem. Theory Comput. 2012, 8, 4691−4706

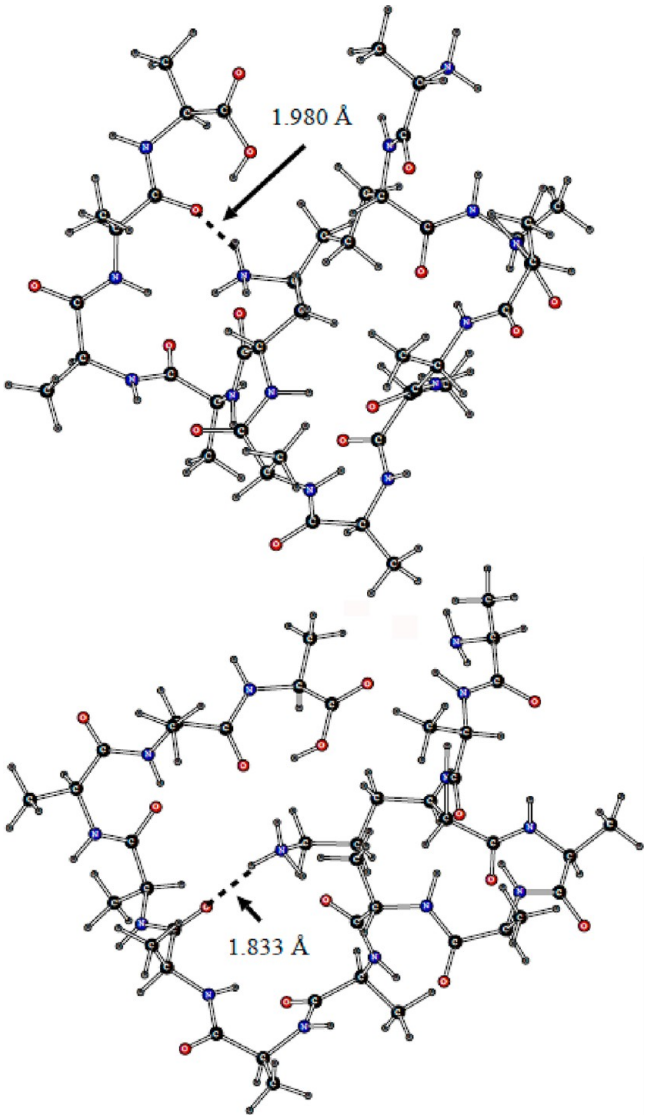

Figure 28. Snapshot of the POSSIM Lys-N+1 (top) and Lys-N+2 (bottom) structures in the gas phase after 40 × 10^6 Monte Carlo configurations. Normally terminated polypeptide.

the polypeptide behavior but should not lead to major qualitative differences.

Using the POSSIM Force Field with Normally Terminated Polypeptides. The values of the total system energies as a function of the simulation length for the polarizable Lys-C and Lys-N structures with normal termini are given in Figure 14.

The trends in these graphs are similar to those observed for the OPLS-AA Lys-C and Lys-N structures (as shown in Figure 7). The Lys-C energy reaches its equilibrium faster, but all the energies converge by ca. 17 × 10^6 configurations. Snapshots of the structures after 40 × 10^6 configurations are shown in Figures 15 and 16.

The physical picture seems to be similar to that emerging from studying the OPLS-AA versions on the polypeptides. The Lys-C system stays helical and ordered, with the protonated lysine group forming a hydrogen bond to one of the neighboring alanine residues. In case of the Lys-N polypeptide, the helix collapses, and the lysine is hydrogen bonded to residues at the other end of the system.

The averaged values of the backbone ϕ and ψ dihedral angles, as a function of the simulation length, as produced with the fast polarizable POSSIM formalism, are shown in Figure 17. It can be seen that the Lys-C system simulated with POSSIM tends to have ϕ lower and ψ higher than the average experimental numbers, but the total per residue given in Table 5 is −103.7°, which is close to the experimental number. At the same time, the Lys-N system shows greater deviations and significantly noisier averaged values of ψ. There is a difference with the OPLS-AA picture in the values of POSSIM Lys-N ψ ending up noticeably higher than those of Lys-C. The overall average turn per residue obtained in the POSSIM Lys-N is −110.9°, which is higher than the experimental value and the Lys-C result obtained with POSSIM, the difference being qualitatively the same although smaller in magnitude than in the OPLS-AA case.

Using the POSSIM Force Field with CH_3-Capped Polypeptides. In order to assess the effect of replacing the normal termini with methyl caps on the results of the POSSIM simulations, we have run our simulations with the CH_3-capped polypeptide. The final structures of the Lys-C and Lys-N systems are shown in Figures 18 and 19, respectively.

The qualitative physical picture remains the same as in all the other POSSIM and OPLS simulations. The structure with the lysine residue at the C-terminus retains its helicity, and the N-lys one denaturates and adopts a disordered geometry shown in Figure 19.

Evolution of the averaged backbone angles ϕ and ψ as a function of the simulation length is shown in Figure 20.

The general behavior of the backbone angles for the methyl-capped peptides simulated with the polarizable POSSIM force field is similar to that of the normally terminated ones. The values of the ϕ angle are lower for the CH_3-capped Lys-N systems as compared to Lys-C, while the values of ψ are much noisier than their counterparts for the Lys-C. The average turn of the helix per residue is −101.5° for Lys-C and −121.7° for Lys-N; thus the deviations from the average experimental value of −105° are somewhat bigger than for the peptide with the regular termini.

While the CH_3 capping of the systems does not alter the qualitative physical picture, it does appear that the methyl capping introduces more stability to the geometry of these polypeptides simulated with the POSSIM force field. The same conclusion is true for the OPLS case, as can be seen after comparing the graphs presented in Figures 10 and 13.

Simulating the Normally and CH_3-Capped Polypeptides in Solution with the POSSIM Force Field. Although this study is devoted first and foremost to investigating the behavior of the alanine polypeptide with one lysine residue in the gas-phase, and the helix-disordered transition has been reported for such gas-phase systems, we have also carried out an examination of normally terminated and CH_3-capped Lys-C and Lys-N systems in aqueous solution using the POSSIM force field (and with the previously developed polarizable POSSIM water model[8]). The termini (in the normally terminated cases) were present in their ionized zwitterionic form ($-COO^-$ and $-NH_3^+$).

Final snapshots of the resulting peptide structures after 40 × 10^6 Monte Carlo configurations are presented in Figures 21−24. Water molecules are not shown in these figures for the sake of clarity.

It can be seen that the physical picture in these solvated simulations is qualitatively different from that in the gas phase. The α-helix is preserved in all four cases. Both normally and CH_3-terminated polypetides act in the same way. Moreover, positioning of the protonated lysine residue at the C- or N-terminus makes no difference.

dx.doi.org/10.1021/ct300492n | J. Chem. Theory Comput. 2012, 8, 4691−4706

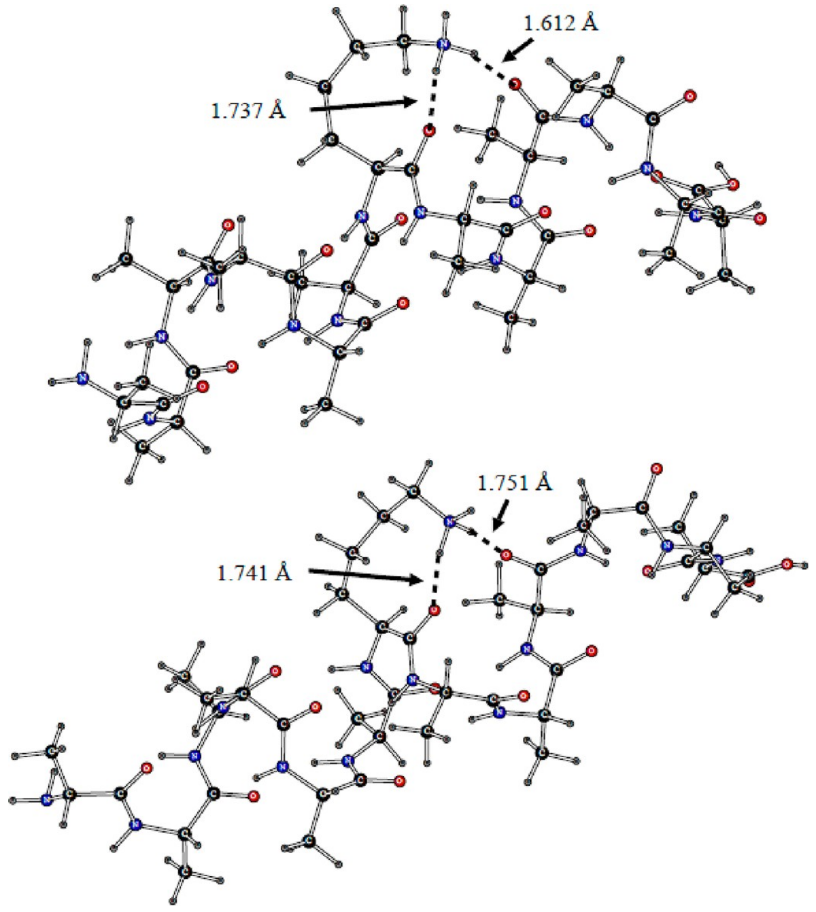

Figure 29. Snapshot of the OPLS (top) and POSSIM (bottom) Lys-Mid structures in the gas phase after 40 × 10^6 Monte Carlo configurations. Normally terminated polypeptide.

This result makes perfect sense given that all four structures have the lysine residue extended toward the bulk of the solvent. This ionized group is solvated in water and does not have to form a hydrogen bond to the backbone of the peptide. Therefore, the presence of the lysine residue and its position makes no difference and does not destroy the helicity of the structure.

Investigating the Effect of Placing the Lysine Residue in Nonterminal Positions of the α-Helix. We have also applied the OPLS-AA and polarizable POSSIM force fields to studying the effect of positioning the protonated lysine residue at places other than the C- and N-termini. We termed these additionally investigated systems Lys-C-1, Lys-C-2 (for the peptides with the lysine residue located one and two residues away from the C-terminus), Lys-N+1, Lys-N+2 (for the charged residue being one and two residues away from the N-terminus), and Lys-Mid (for the LYS residues being at the middle of the polypeptide). All these simulations were carried out in the gas phase with the normally terminated systems (−COOH and $-NH_2$).

Final snapshots of the OPLS Lys-C-1 and Lys-C-2 systems after 40 million Monte Carlo configurations are given in Figure 25. The results are very similar to those for the OPLS Lys-C, as shown in Figure 8. The $-NH_3^+$ group is hydrogen-bonded to carbonyl oxygen atoms near the C-terminus, and the overall structure of the system remains helical.

Resulting snapshots of the Lys-C-1 and Lys-C-2 structures simulated with the polarizable POSSIM force field are shown in Figure 26. The helicity is also preserved here, just like in the case of the Lys-C POSSIM geometry in Figure 15. At the same time, while the overall α-helical motif is the same, there is a difference in the observed hydrogen bonding of the Lys residue. It is bonded to the −COOH group of the C-terminus (as well as to a carbonyl oxygen of the polypeptide backbone in Lys-C-2). Therefore, while the main physical picture is the same as in the terminal Lys structure, there is a difference in details. This is a case when these details obviously do depend on the presence of the normal termini, as opposed to the methyl caps.

The OPLS-AA Lys-N+1 and Lys-N+2 systems do denaturate, just like in the Lys-N case in Figures 9 and 12, as can be seen from the structure in Figure 27. The $-NH_3^+$ group of the protonated lysine residue is forming hydrogen bonds with oxygen atoms in the backbone in the C-terminus region.

Likewise, the POSSIM force field yields a similar result with the α-helix destroyed and the Lys residue binding to backbone oxygens, as seen in Figure 28. It appears that there are no significant differences between these results and those obtained for the Lys-N systems simulated with POSSIM.

Finally, structures of the OPLS and POSSIM polypeptides with the protonated lysine residue located in the middle of the chain are presented in Figure 29. Both structures are similar. The lysine residue is bending toward the C-terminus and is forming a hydrogen bond with one of the backbone oxygen atoms.

The overall physical picture emerging from these simulations is as follows. The polypeptides with the protonated lysine residue on or near the C-terminus, as well as at the middle of the chain, remain helical in the course of the simulations, regardless of whether the OPLS or POSSIM force field is employed. The structures with

dx.doi.org/10.1021/ct300492n | J. Chem. Theory Comput. 2012, 8, 4691−4706

the lysine residue at or near the N-terminus lose their helical structure and become disordered. Depending on whether the normal termini or methyl caps are used, details of the structure may change, but the overall picture is likely to stay the same.

IV. CONCLUSIONS

We have simulated polyalanine α-helix with a protonated lysine residue added at or near the termini, as well as in the middle of the molecule. Fixed-charge OPLS-AA and fast second-order polarizable POSSIM force fields were employed. Our findings are in agreement with literature results. We have shown that such a structure with the lysine residue at or near the C-terminus forms a stable α-helix in the gas phase, while a similar system with the charged residue at or near the N-side has its secondary structure denaturating into a globular form.

We have also shown that capping the end residues of the polypeptide with methyl groups may impact details of the final structure but does not appear to change the above general physical conclusion. Thus, employing this practice in computational protein studies can probably be considered acceptable.

Simulations of the polypeptides in aqueous solution with the polarizable POSSIM force field predict that these systems retain their helical structure. This result is not unexpected given that the protonated lysine residue adopts an extended conformation and gets solvated in this case; thus it does not have any significant hydrogen-bonded interactions with the backbone and does not affect the helical structure. Capping the system with methyl groups does not appear to change this result.

We have shown in our previous work that both OPLS-AA and POSSIM predict that a pure polyalanine α-helix would retain its structure in the gas phase.[9] At the same time, a relatively small change in the backbone torsional parameters in the POSSIM simulations leads to denaturation of the helix. Although direct comparison with experimental data cannot be carried out for this system, this does show that the gas-phase polyalanine helicity depends on a very fine balance of energy effects.

We also conclude that our fast polarizable POSSIM technique and parameters have passed another test, and so far there has been no reason to doubt that they are adequate for use in simulating larger protein and protein−ligand molecular systems.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gkaminski@wpi.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The project described was supported by Grant Number R01GM074624 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

REFERENCES

(1) Dugourd, P.; Hudgins, R. R.; Clemmer, D. E.; Jarrold, M. F. *Rev. Sci. Instrum.* **1997**, *68*, 1122−1129.

(2) Hudgins, R. R.; Woenckhaus, J.; Jarrold, M. F. *Int. J. Mass Spectrom. Ion Process.* **1997**, *165*, 497−507.

(3) See, for example: (a) Peng, Y.; Hansmann, U. H. E.; Alves, N. A. *J. Chem. Phys.* **2003**, *118*, 2374−2380. (b) Kohtani, M.; Jarrold, M. F. *J. Am. Chem. Soc.* **2004**, *126*, 8454−8458. (c) Kohtani, M.; Schneider, J. E.; Jones, T. C.; Jarrold, M. F. *J. Am. Chem. Soc.* **2004**, *126*, 16981−16987.

(4) Wei, Y.; Nadler, W.; Hansmann, U. H. E. *J. Chem. Phys.* **2007**, *126*, 204307 and references therein.

(5) MacDermaid, C. M.; Kaminski, G. A. *J. Phys. Chem. B* **2007**, *111*, 9036−9044.

(6) Click, T. H.; Kaminski, G. A. *J. Chem. Theory Comput.* **2009**, *5*, 2935−2943.

(7) Click, T. H.; Ponomarev, S. Y.; Kaminski, G. A. *J. Comput. Chem.* **2012**, *33*, 1142−1151.

(8) Kaminski, G. A.; Ponomarev, S. Y.; Liu, A. B. *J. Chem. Theory Comput.* **2009**, *5*, 2935−2943.

(9) Ponomarev, S. Y.; Kaminski, G. A. *J. Chem. Theory Comput.* **2011**, *7*, 1415−1427.

(10) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.

(11) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908−5920.

(12) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X. X.; Murphy, R. B.; Zhou, R. H.; Halgren, T. A. *J. Comput. Chem.* **2002**, *23*, 1515−1531.

(13) Kaminski, G. A.; Zhou, R.; Friesner, R. A. *J. Comput. Chem.* **2003**, *24*, 267−276.

(14) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(15) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(16) *Jaguar*, v4.2; Schrödinger, Inc.: Portland, OR, 2000.

(17) Kaminski, G. A.; Maple, J. R.; Murphy, R. B.; Braden, D.; Friesner, R. A. *J. Chem. Theory Comput.* **2005**, *1*, 248−254.

(18) Li, X.; Ponomarev, S. Y.; Sa, Q.; Sigalovsky, D. L.; Kaminski, G. A. In preparation for publication.

(19) Kurt D. Berndt, Protein Secondary Structure. http://www.cryst.bbk.ac.uk/PPS2/course/section8/ss-960531_6.html (accessed on December 10, 2010).