Structural, Thermodynamic, and Kinetic Properties of Gramicidin Analogue GS6 Studied by Molecular Dynamics Simulations and Statistical Mechanics

Laura Zanetti-Polzi,[1] Massimiliano Anselmi,[1] Maira D'Alessandro,[1,2]
Andrea Amadei,[2] Alfredo Di Nola[1]

[1] *Department of Chemistry, University of Rome "La Sapienza", Rome, Italy*

[2] *Department of Chemical Sciences and Technology, University of Rome "Tor Vergata", Rome, Italy*

ABSTRACT:

Gramicidin S (GS) analogues belong to an important class of cyclic peptides, characterized by an antiparallel double-stranded β-sheet structure with Type II′ β-turns. Such compounds can be used as model systems to understand the folding/unfolding process of β-hairpins and more in general of β-structures. In the present study, we specifically investigate the folding/unfolding behavior of the hexameric Gramicidin S analogue GS6 by using all-atoms molecular dynamics (MD) simulations at different temperatures, coupled to a statistical mechanical model based on the Quasi Gaussian Entropy theory. Such an approach permits to describe the structural, thermodynamic, and kinetic properties of the peptide and to quantitatively characterize its folding/unfolding transitions. © 2009 Wiley Periodicals, Inc. Biopolymers 91: 1154–1160, 2009.

Keywords: MD; Gramicidin; peptide folding

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

In honor of Professor Lelio Mazzarella.

INTRODUCTION

The study of peptides, and in particular of peptide folding/unfolding, has become one of the most innovative and challenging areas of biophysical–biochemical research which provided in the past several years a large body of relevant data based on both experimental and computational methods.[1–6]

Peptides proved to be excellent model systems to study protein folding/unfolding, allowing the characterization of the mechanism of formation of the protein basic structural elements.[7–9]

MD simulations have been widely used in the last decade to study the atomistic behavior of solvated peptides including structural/conformational transitions, clearly providing detailed information with temporal and spatial resolution unachievable by experimental methods.[10–13] Given the rapid increase in computing power, the higher efficiency of algorithms and the improving of the force fields over the past years, computer simulations can play an essential role in folding/unfolding studies, allowing nowadays the explicit simulation of fast peptide folding/unfolding events.[14–17] In previous articles,[18,19] we quantitatively characterized the folding/unfolding kinetics and thermodynamics of simple peptides by using MD simulations and advanced theoretical models based on statistical mechanics.

In this study, we use as simple model system the cyclic β-hairpin 6-meric peptide (GS6) analog of Gramicidin S (GS).

GS analogues containing 6, 8, 10, 12, 14, and 16 residues were largely studied with circular dichroism spectroscopy, showing, in the series composed by GS6, GS10, G14, a higher propensity to form ordered β-sheet with respect to the others.[20] The solution structures of GS6, GS10, GS14 have been solved by NMR spectroscopy,[20,21] revealing that they form stable antiparallel β-hairpin structures, bordered by

two Type II′ β-turns. The dynamics of β-turn formation and the folding/unfolding rates of the same peptides were investigated using equilibrium Fourier transform infrared spectroscopy and T-jump relaxation probed by time-resolved infrared spectroscopy.[22]

The aim of the present study is to quantitatively reproduce the experimentally observed properties to test the accuracy of the theoretical–computational approach employed, providing a deeper understanding of the thermodynamic and kinetic mechanism involved in the folding/unfolding process of GS6 and, more in general, of small peptides.

The article as organized as follows. In the theory section, we briefly outline the statistical mechanical model, based on the Quasi Gaussian Entropy (QGE) theory,[23–26] we utilize to describe the peptide folding/unfolding thermodynamics. Such a model, specifically designed to treat conformational transitions, has been introduced and described in details in a previous article.[19]

In the theory section, we also illustrate how to model in a simple way the folding/unfolding kinetics, discussing the effects of the possible presence of kinetically coupled degrees of freedom. Finally, in the results section, we show the thermodynamics and kinetics of GS6 folding/unfolding as provided by the theoretical–computational model employed, comparing our results with the available experimental data.

METHODS

MD Simulations Protocol

The initial structure of cyclo[(Lys-dTyr-Pro)$_2$] was created using Pymol version 0.99 (DeLano Scientific) on the basis of the experimental chemical structure.[20,21] Geometry optimization of that structure was done by using the empirical potential energy function of the GROMOS96 43a1 force field.[27] The peptide, in its starting conformation, was solvated with water and placed in a periodic cubic box large enough to contain the peptide and 0.5 nm of solvent on all sides. The two lysine side chains were protonated as to reproduce a pH of about 7: two negative counterions (Cl$^-$) were then added by replacing two water molecules to achieve a neutral condition. Molecular Dynamics simulations, in the NVT ensemble, with fixed bond lengths[28] were performed with the GROMACS software package[29] and with the GROMOS96 43a1 force field. Water was modeled by the simple point charge (SPC) model.[30] A nonbond pairlist cutoff of 9.0 Å was used, and the long-range electrostatic interactions were treated with the particle mesh Ewald method.[31] The isokinetic temperature coupling[32] was used to keep the temperature constant at the desired value. After various equilibration MD runs, six all-atom MD simulations in explicit water at six different temperatures and with different time lengths were carried out: 400 ns at 280 K, 300 ns at 310 K, 300 ns at 360 K, 180 ns at 400 K, 60 ns at 500 K, and 60 ns at 600 K. For the last two simulations, a time step of 1 fs was used, while for the others the time step was 2 fs.
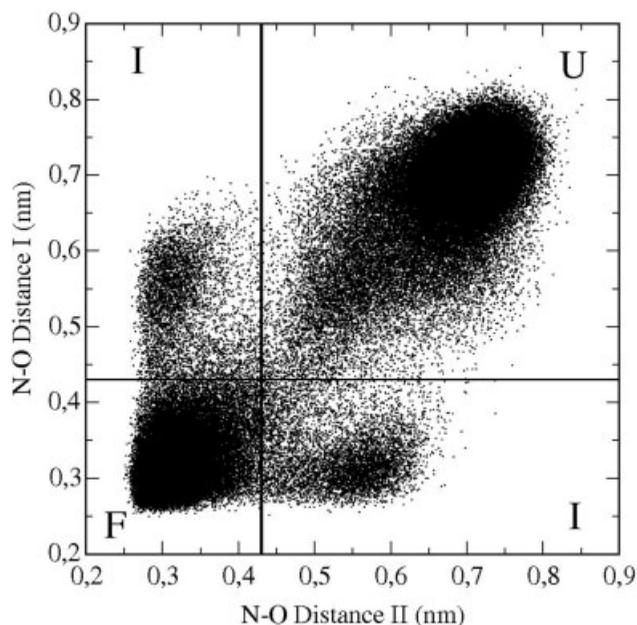
FIGURE 1 Projection of the trajectory at 310 K on the plane of the two N—O (H-bond) distances between Lys residues. Three conformational states can be defined: the folded state (F) in which both H-bonds are formed, the intermediate state (I) in which only one of the two H-bonds is formed, and the unfolded state (U) in which none of the H-bonds is formed.

THEORY

Thermodynamic Characterization

Given a system in thermodynamic equilibrium, the change in free energy on going from a reference state, ref, of the system to a generic state, i (e.g., from unfolded to folded), at constant temperature and constant volume can be evaluated as:

$$\Delta A_{\mathrm{ref}\rightarrow\mathrm{i}} = -RT \ln \frac{p_{\mathrm{i}}}{p_{\mathrm{ref}}} \tag{1}$$

where R is the ideal gas constant, T is the absolute temperature and p_{i} and p_{ref} are the equilibrium probabilities of finding the system in state i and state ref, respectively. We consider the conformational space and relative free energy as defined by the reaction coordinates given by the two distances between the oxygen and nitrogen atoms involved in the two hydrogen bonds (H-bonds). By projecting the MD trajectories onto the plane defined by these conformational coordinates (see Figure 1), we were able to identify three regions corresponding to three secondary structure states (conformational states): the folded state (F) in which both H-bonds are formed; the intermediate state (I) in which only one of the two H-bonds is formed; and the unfolded state

(U) in which none of the two H-bonds is formed. For every region, the corresponding MD frames were counted providing the equilibrium probabilities and hence, according to Eq. (1), the free energy changes $\Delta A_{\mathrm{ref} \rightarrow \mathrm{i}}$. Note that the variation of the Helmholtz free energy, due to the conformational transition of the peptide in the simulation box, exactly corresponds to the peptide chemical potential change (i.e., $\Delta A_{\mathrm{ref} \rightarrow \mathrm{i}} = \Delta \mu_{\mathrm{ref} \rightarrow \mathrm{i}}$). Such peptide chemical potential variations and the mean potential energies of the whole simulation box, taken for each conformational state for a set of temperatures (see Methods), were utilized to obtain, by a fitting procedure, a detailed model of the peptide partial molar thermodynamics according to the QGE theory as described in details in a recent article.[19] Such a QGE model provides the temperature dependence of the peptide chemical potential change and related partial molar properties according to[19]:

$$\Delta \mu = \Delta u_0' - \Delta c_{v0}' T_0 \Lambda(T) + p'(T) \Delta v - RT \ln \gamma \quad (2)$$

where $\Delta u_0'$ and $\Delta c_{v0}'$ are the variations of the partial molar excess internal energy and isochoric heat capacity with respect to the reference state at the reference temperature T_0 (in our case $T_0 = 310$ K), Δv is the partial molar volume change with respect to the reference state and $R \ln \gamma$ corresponds to a partial molar entropic term due to hard body effects. Moreover, p' is the pure solvent excess pressure and

$$\Lambda(T) = \frac{1}{\delta_0} + \frac{T}{T_0 \delta_0^2} \ln \left(1 - \frac{\delta_0 T_0}{T(1 - \delta_0) + \delta_0 T_0} \right) \quad (3)$$

with δ_0 a temperature independent dimensionless constant obtained by the pure solvent simulations.

Note that $\Delta u_0'$, $\Delta c_{v0}'$, Δv, and $R \ln \gamma$ are the parameters, corresponding to the physical properties fully defining the QGE model, which are obtained via the fitting procedure.

Kinetic Characterization

To describe in a simple way the kinetics of the folding–unfolding process, we may consider a single conformational degree of freedom q as defined by the bisector of the plane identified by the two hydrogen bonds coordinates in Figure 1 by using the free energy profile along q, as obtained via

$$\Delta A(q) = -RT \ln \frac{\rho(q)}{\rho(q_{\mathrm{ref}})} \quad (4)$$

with $\rho(q)$ the equilibrium probability density, we may readily obtain the complete kinetics of the folding/unfolding transitions by solving a Fokker-Plank type equation[33]:
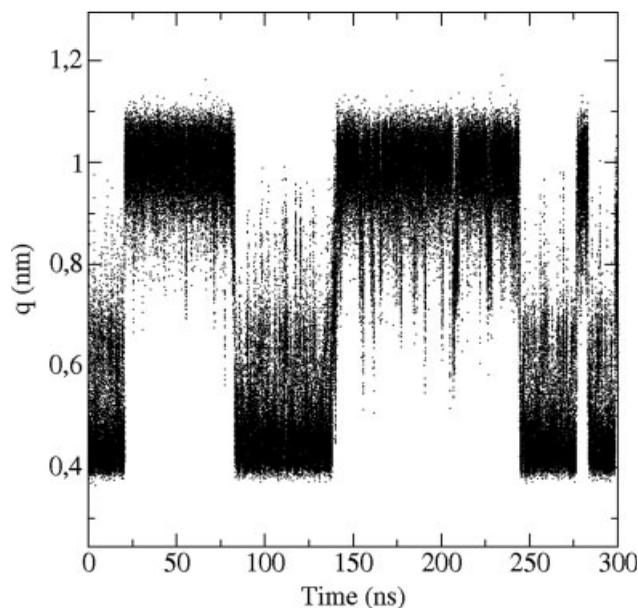
FIGURE 2 Time evolution of the reaction coordinate q during the MD run at 310 K.

$$\frac{\partial \rho(q, t)}{\partial t} = \frac{D}{KT} \left[\rho \left(\frac{d^2 \Delta A}{dq^2} \right) + \left(\frac{\partial \rho(q, t)}{\partial q} \right) \left(\frac{d \Delta A}{dq} \right) \right]$$
$$+ D \left(\frac{\partial^2 \rho(q, t)}{\partial q^2} \right) \quad (5)$$

where $\rho(q,t)$ is the time-dependent probability density and D is the diffusion coefficient along q (in the present case at 310 K, $D = 1.66 \times 10^{-4}$ nm^2/ps). It must be remarked that such a model is sufficiently accurate only within the approximation of an instantaneous relaxation of all the other degrees of freedom during the diffusion along q. When some degrees of freedom other than q relax at a comparable rate of the diffusion along the chosen reaction coordinate, a proper kinetic model must include such degrees of freedom and their coupling. This can be accomplished either explicitly including all the relevant reaction coordinates in the Fokker-Plank type equation or implicitly considering the effects of the other degrees of freedom by monitoring the fluctuation time-behavior of q as provided by MD simulations, hence still describing the kinetics via a single reaction coordinate.

In the present case, the trajectory of the reaction coordinate q (see Figure 2) reveals the presence of two distinct fluctuation regimes corresponding to two free energy basins (see Figure 3) and characterized by different fluctuation distributions: one peaked at about 0.4 nm (corresponding to the folded state basin) and the other peaked at about 1.0 nm (corresponding to the unfolded state basin). It must be noted that the two identified distributions significantly overlap in
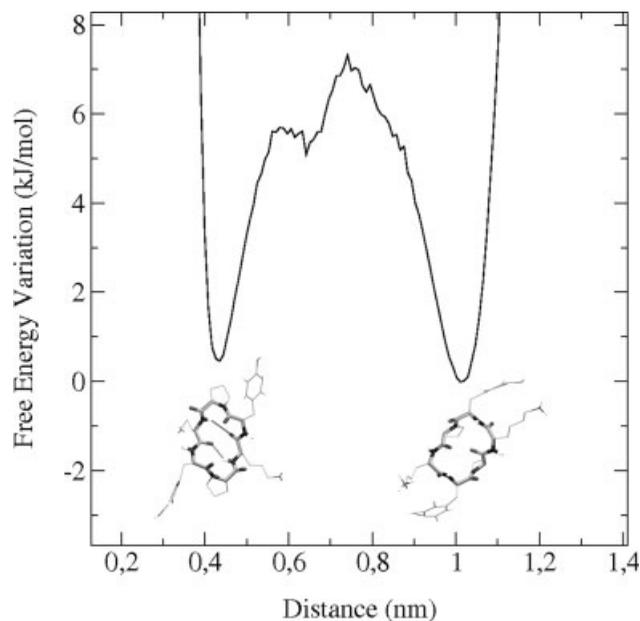
FIGURE 3 Free energy variation (at 310 K) along the reaction coordinate q, defined by the bisector of the plane given by the two N—O (H-bond) distances between Lys residues. Two local minima can be observed corresponding to the folded (left) and the unfolded (right) states. Two representative snapshots of the folded and unfolded states extracted from the MD run at 310 K are also reported below the corresponding minimum. Hydrogen bonds are represented with dotted lines.

the 0.7–0.9 nm range, hence indicating that GS6 folding/ unfolding kinetics is characterized by the relaxation of a set of coupled degrees of freedom corresponding to the inter-conversion kinetics of the two q-fluctuation regimes. In this article, we define the transition from one fluctuation regime to the other as occurring when the reaction coordinate reaches the value corresponding to the peak of the other regime (e.g., starting from the folded state, when q reaches 1.0). In this way, we may easily subdivide the q trajectory into subparts corresponding to a single fluctuation regime and hence evaluate the mean lifetime for the folded to unfolded and reverse transition. In practice, the use of such a criterion for the MD simulation data of GS6 at 310 K, provides 4 and 3 blocks of the q trajectory for the folded and unfolded state, respectively (the lower and upper fluctuation-blocks in Figure 2).

The mean time interval of the folded/unfolded blocks (i.e., the average of the corresponding time intervals), gives hence the estimate of the unfolding/folding mean lifetime.

RESULTS AND DISCUSSION

As mentioned in the theory section, we used the plane defined by the two (H-bonds) N—O distances between Lys

residues as conformational space to describe the relevant conformational state of GS6. Using 0.43 nm as cut-off distance for the formation of each hydrogen bond (0.43 nm corresponds to the probability minimum of the N—O distance distribution, see Figure 4), the chosen conformational space may be divided into three regions representing different conformational states: the folded state (F) in which both H-bonds are formed; the intermediate state (I) in which only one of the two H-bonds is formed; and the unfolded state (U) in which none of the two H-bonds is formed.

It is worth to note that the density peaks of the folded and unfolded states in Figure 1, corresponding to the free energy minima and relative structures along the bisector of the plane in Figure 1 (see Figure 3), clearly indicate that the β-hairpin structure is relatively stable in GS6 in agreement with experimental data.[20,21] The Type II′ β-turns are stabilized by the intraturn hydrogen bond, involving the Lys residues. Moreover, our MD simulations provide for the folded state the correct (i.e., experimentally observed) arrangement of the residues in each β-turn with D-Tyr and Pro residues in positions i+1 and i+2, respectively, and their side chains relative arrangement in line with the so-called equatorial-axial rule. In the unfolded state, the β-hairpin structure is completely disrupted with the main chain of the peptide distorted and proline and tyrosine residues no more in the proper arrangement of the β-turn.
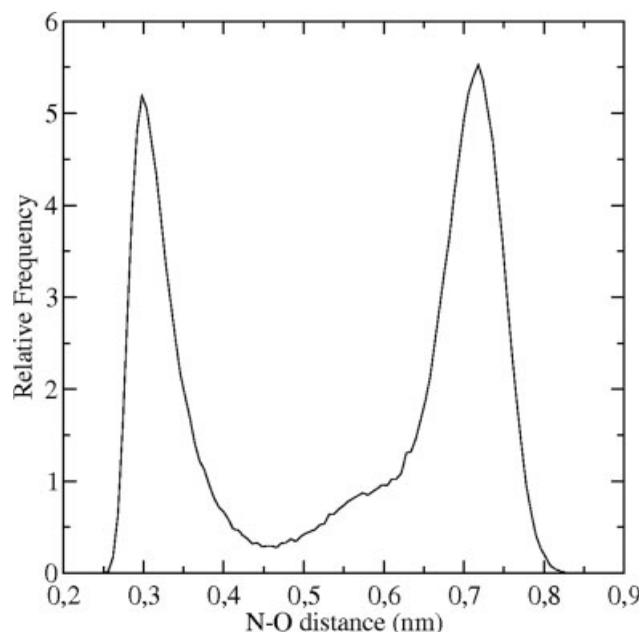
FIGURE 4 Distribution function (at 310 K) of the distance between the nitrogen and the oxygen atoms involved in the hydrogen bonds. The left peak corresponds to the folded structure while the right peak corresponds to the unfolded one.
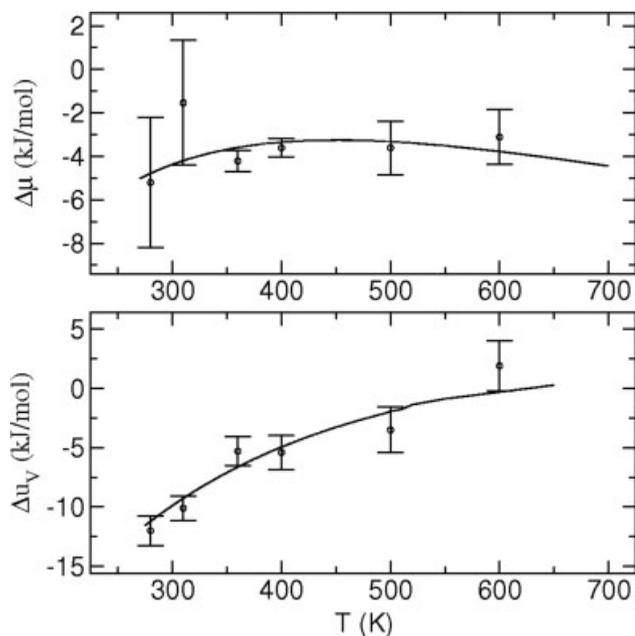
FIGURE 5 Plot of the unfolding chemical potential change (upper panel) and isochoric internal energy change (lower panel) as a function of temperature along the isochore. In the figure, the circles represent the values obtained by the direct use of MD data with their relative error bar, while the solid line represents the corresponding QGE model prediction.

FIGURE 6 Plot of the excess isochoric heat capacity change due to the peptide insertion into the solvent. In the figure, the circles represent the values obtained by using the mean square potential energy fluctuation of the simulation box with the relative error bars, while the solid line represents the corresponding QGE model prediction.

In Figure 5, we report the temperature dependence of the unfolding chemical potential change and the isochoric internal energy change, $\Delta u_v = (\partial \beta \Delta \mu / \partial \beta)_v$, as provided by the QGE model. In the figure, we also report the corresponding values as obtained by the conformational probabilities and mean potential energies given by the MD simulations. To test severely the accuracy of the model prediction, we compare in Figure 6 the isochoric excess heat capacity change due to the peptide insertion into the solvent as obtained by the QGE model with the corresponding values as provided by MD data (the latter were evaluated by using the mean square fluctuations of the simulation box potential energy). Note that these MD-based values were not used within the parameterization procedure of the model. From these figures it is clear that the accuracy (within the noise) of the QGE model in describing and predicting the peptide thermodynamics as provided by the MD simulations.

Interestingly, the unfolding chemical potential change shows a negative broad maximum centred at ∼450 K indicating that the unfolded state is thermodynamically more stable in the whole temperature range considered with an increased stability for lower and higher temperatures. When considering the energetic and entropic contributions to the unfolding chemical potential change, as provided by the QGE model (see Figure 7), it clearly emerges that the unfolded state is characterized by a lower entropy and energy than the folded
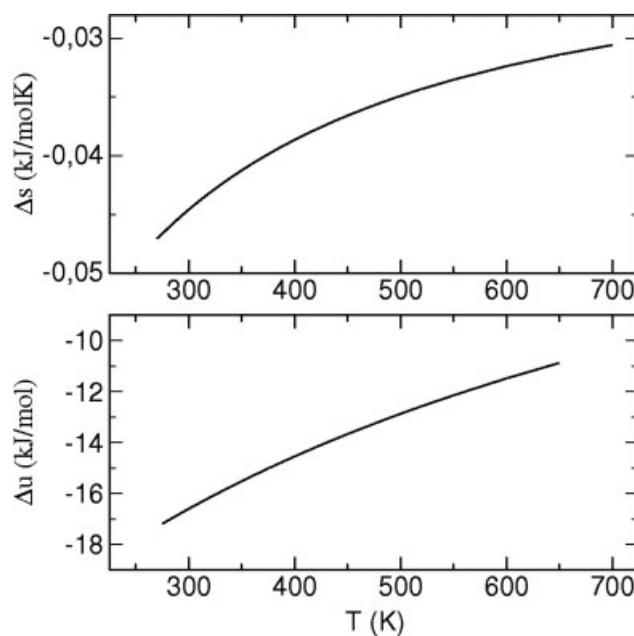
FIGURE 7 Plot of the unfolding partial molar entropy change (upper panel) and internal energy change (lower panel) of the peptide as a function of temperature along the isochore as provided by the QGE model.

state, hence indicating a reduction of the peptide partial molar entropy coupled to an energetic optimization due to the unfolding process. Such a counterintuitive thermodynamic behavior, observed and described in details in a recent article for a completely different peptide,[19] is driven by the charged and/or polar groups solvent exposure occurring upon unfolding and causing a relevant solvent rearrangement in the peptide first solvation shells leading to the peptide partial molar volume reduction (electrostriction).

To characterize in a simple way the kinetics of the folding/unfolding transitions, we have considered a single reaction coordinate q as defined by the bisector of the plane utilized to define the conformational space (see Figure 1). The use of a Fokker-Plank type equation to obtain the kinetic rate constants is appropriate only within the approximation that all the other degrees of freedom may be considered as instantaneously relaxed along the reaction coordinate transition. As evidenced in the theory section the reaction coordinate fluctuation as provided by the MD simulations (see Figure 2) are incompatible with such an assumption, as clearly shown by the significant overlapping of the distributions for the folded and unfolded q-fluctuation regimes.

According to the method briefly outlined in the theory section, we define the transition from one fluctuation regime to the other (within our approximation, the folding/unfolding transition) as occurring when the reaction coordinate reaches the value corresponding to the peak of the other regime. At 310 K, we obtain 31 ns for the unfolding mean lifetime and 59 ns for the folding mean lifetime (10 ns and 27 ns are the corresponding standard errors), in good agreement with the experimentally measured values at 324 K (71 ns and 100 ns for the unfolding and folding mean lifetime, respectively).[22] Interestingly, the use of the Fokker-Plank type equation for the single reaction coordinate considered, provides a significantly faster kinetics (1.2 ns and 1.4 ns for the unfolding and folding lifetimes) indeed confirming its inaccuracy due to the presence of kinetically coupled degrees of freedom involved in the folding/unfolding transitions.

The nature of the hidden degrees of freedom coupled to the reaction coordinate chosen is elusive, probably involving the peptide side chains and solvent molecules, and beyond the scope of the present article. However, given the interest on the folding/unfolding kinetic mechanism, the characterization of the coupled degrees of freedom involved in the folding/unfolding transitions will be addressed in a future article.

CONCLUSIONS

In the present study, the folding/unfolding process for the Gramicidin analogue GS6 has been extensively investigated by the use of MD simulations coupled to statistical mechanical models to characterize the thermodynamics and kinetics of the process. Such an approach allowed to quantitatively obtain the peptide folding/unfolding partial molar thermodynamics and the corresponding kinetic rate constants. Results show that GS6 unfolded state is characterized by a lower entropy and energy than the folded one, as a consequence of the peptide partial molar volume decrease. Such a thermodynamic behavior driven by the charged and/or polar groups solvent exposure (electrostriction), is in line with our previous data on a completely different small peptide.[19] It is worth to note that the emerging electrostriction-driven folding/unfolding thermodynamics as evidenced by the small peptides we studied, is likely to be a specific feature of solvated small peptides where the effects of intramolecular interactions and configurational freedom are relatively small compared to the peptide-solvent thermodynamic coupling. However, the present data confirm that the solvent exposure of charged and/or polar chemical groups typically results in electrostriction effects, hence possibly implying that electrostriction may play a significant role also in larger peptides and even in proteins.

Characterization of the folding/unfolding kinetics provided clear indications that a set of coupled degrees of freedom is involved in the relaxation process, hence implying that a simple one-dimensional Fokker-Plank type equation cannot be used to model accurately the kinetics. The analysis of the reaction coordinate trajectory in terms of fluctuation regimes (i.e., fluctuation distributions) allowed a simple identification of folding and unfolding transitions, providing estimates of the corresponding mean lifetimes which match rather well the experimentally determined values.

REFERENCES

1. Callender, R. H.; Dyer, R. B.; Gilmanshin, R.; Woodruff, W. H. Annu Rev Phys Chem 1998, 49, 173–202.
2. Sporlein, S.; Carstens, H.; Satzger, H.; Renner, C.; Behrendt, R.; Moroder, L.; Tavan, P.; Zinth, W.; Wachtveitl, J. Proc Natl Acad Sci USA 2002, 99, 7998–8002.
3. Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. Nature 1997, 390, 196–199.
4. Ferrara, P.; Caflisch, A. Proc Natl Acad Sci USA 2000, 97, 10780–10785.
5. Pande, V. S.; Rokhsar, D. S. Proc Natl Acad Sci USA 1999, 96, 9062–9067.

6. Ferrara, P.; Apostolakis, J.; Caflisch, A. J Phys Chem B 2000, 104, 5000–5010.

7. Galzitskaja, O. V.; Higo, J.; Kuroda, M.; Nakamura, H. Chem Phys Lett 2000, 326, 421–429.

8. Eaton, W. A.; Munoz, V.; Thompson, P. A.; Henry, E. R.; Hofrichter, J. Acc Chem Res 1998, 31, 745–753.

9. Werner, J. H.; Dyer, R. B.; Fesinmeyer, R. M.; Andersen, N. H. J Phys Chem B 2002, 106, 487–494.

10. Daidone, I.; Amadei, A.; Roccatano, D.; Di Nola, A. Biophys J 2003, 85, 2865–2871.

11. Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. J Mol Biol 1998, 280, 925–932.

12. Zagrovic, B.; Sorin, E. J.; Pande, V. J Mol Biol 2001, 313, 151–169.

13. Zhang, Z.; Shi, Y.; Liu, H. Biophys J 2003, 84, 3583–3593.

14. Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. J Am Chem Soc 2002, 124, 12952–12953.

15. Snow, C. D.; Zagrovic, B.; Pande, V. S. J Am Chem Soc 2002, 124, 14548–14549.

16. Zagrovic, B.; Pande, V. J Comput Chem 2003, 24, 1432–1436.

17. Kubelka, J.; Hofrichter, J.; Eaton, W. A. Curr Opin Struct Biol 2004, 14, 76–88.

18. Daidone, I.; D'Abramo, M.; Di Nola, A.; Amadei, A. J Am Chem Soc 2005, 127, 14825–14832.

19. Noe, F.; Daidone, I.; Smith, J. C.; Di Nola, A.; Amadei, A. J Phys Chem B 2008, 112, 11155–11163.

20. Gibbs, A. C.; Kondejewski, L. H.; Gronwald, W.; Nip, A. M.; Hodges, R. S.; Sykes, B. D.; Wishart, D. S. Nat Struct Biol 1998, 5, 284–288.

21. Gibbs, A. C.; Bjorndahl, T. C.; Hodges, R. S.; Wishart, D. S. J Am Chem Soc 2002, 124, 1203–1213.

22. Maness, S. J.; Franzen, S.; Gibbs, A. C.; Causgrove, T. P.; Dyer, R. B. Biophys J 2003, 84, 3874–3882.

23. Amadei, A.; Apol, M. E. F.; Di Nola, A.; Berendsen, H. J. C. J Chem Phys 1996, 104, 1560–1574.

24. Amadei, A.; Apol, M. E. F.; Berendsen, H. J. C. J Chem Phys 1997, 106, 1893–1912.

25. Amadei, A.; Apol, M. E. F.; Berendsen, H. J. C. J Chem Phys 1998, 109, 3004–3016.

26. Apol, M. E. F.; Amadei, A.; Berendsen, H. J. C. J Chem Phys 1998, 109, 3017–3027.

27. van Gunsteren, W. F.; Billeter, S.; Eising, A.; Hunenberger, P.; Kruger, P.; Mark, A. E.; Scott, W.; Tironi, I. Biomolecular Simulations: The GROMOS96 Manual and User Guide; BIOMOS bv Zurich: Groningen, 1996.

28. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. J Comput Chem 1997, 18, 1463–1472.

29. Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. Comput Phys Commun 1995, 91, 43–56.

30. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In Intermolecular Forces; Pullman, B., Ed.; D. Reidel Publishing Company: Dordrecht, The Netherlands, 1981; pp 331–342.

31. Darden, T.; York, D.; Pedersen, L. J Chem Phys 1993, 98, 10089–10092.

32. Brown, D.; Clarke, J. H. R. Mol Phys 1984, 51, 1243–1252.

33. Amadei, A.; D'Abramo, M.; Daidone, I.; D'Alessandro, M.; Di Nola, A.; Aschi, M. Theor Chem Acc 2007, 117, 637–647.