Structural and Kinetic Characterization of a β-hairpin Peptide in Solution: A Molecular Dynamics Study of a Synthetic Analog of Gramicidin S

Laura Zanetti Polzi[1], Andrea Amadei[2] and Alfredo Di Nola[1]

1. *Dipartimento di Chimica, Università di Roma "La Sapienza," Piazzale Aldo Moro 5, I-00185 Roma, Italy*
2. *Dipartimento di Scienze e Tecnologie Chimiche, Universitá di Roma "Tor Vergata," Via della Ricerca Scientifica 1, I-00133 Roma, Italy*

Although protein folding studies have considerably evolved during the past several years, the mechanism of folding is still uncertain. Protein folding is a problem of great interest, both to predict structures from sequences and to enhance the knowledge on protein misfolding related diseases. Peptide models have been demonstrated to be precious experimental systems in the study of the folding process, allowing the investigation of the basic structural elements formation mechanism in proteins: α–helices and β-structures.[1,2]

Studies of helical peptides are particularly successful,[3,4] while those of β-structures are more difficult due to the scarcity of suitable model systems. Recently, several β–sheet models have been designed, including peptide mimetics, protein fragments and artificial proteins.[5,6] These model peptides have been studied both with experimental and computational techniques,[7,8] mainly focusing on the β–hairpin formation, which seems to be important in several folding pathways.[9] The β-hairpin structural motif is the simplest protein motif involving two beta strands that look like a hairpin. The motif consists of two strands that are adjacent in primary sequence oriented in an antiparallel arrangement (where the N-terminus of one sheet is adjacent to the C-terminus of the next) and linked by a short loop of two to five amino acids (Figure 1). The first β–hairpin forming peptide whose folding kinetics was investigated, is the 16-residues C-terminal fragment of protein G. The dynamics of folding of this system has been studied using laser induced temperature jump (T-jump) and time-resolved fluorescence measurements.[10] Moreover, its folding kinetics was characterized by means of all-atoms molecular dynamics (MD) simulations, obtaining results in good agreement with experimental data.[11]

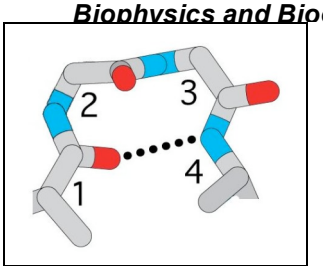

Fig. 1: Schematic representation of the β-hairpin structure. Black dots indicate hydrogen bonds.

Molecular dynamics simulations can, in principle, completely characterize the folding mechanism and the kinetics of the folding processes: indeed, all the information, about whichever process, is contained in the trajectories calculated using all-atoms MD. However, many long trajectories must be calculated in order to obtain sufficient statistical sampling to describe the folding kinetics and this implies a high computational cost. In addition, the reliability of a MD simulation depends on the model used to mimic the experimental condition (force field, pH, ionic strength etc.). Proteins that fold in short time are therefore the best candidates for MD folding studies. MD trajectories can be used to obtain the free energy surface that characterizes the folding process and subsequently, according to the energy landscape theory, to describe the folding kinetics and calculate the folding time constant.[11,12] In the energy landscape theory, the folding process is described as a diffusion under the effect of a free energy gradient toward the native structure.[13,14] The diffusion coefficient, as well as the free energy surface, can be evaluated from the MD trajectories thus allowing, by solving the diffusion equation, the complete characterization of the folding kinetics.

In this study, all-atoms MD simulations of a cyclic β–hairpin 6-meric peptide, analog of Gramicidin S (GS), were performed, in order to investigate the folding kinetics. GS is a cyclic peptide made of ten residues [cyclo(-Val-Orn-Leu-DPhe-Pro)$_2$] having antibiotic properties and an amphipatic nature. A series of synthetic cyclic peptides homologues of GS with different chain lengths (even number residues from 6 to 16), have been experimentally investigated both from structural and from dynamic point of view.[15,16,17]

Using NMR spectroscopy and Circular Dichroism, it has been demonstrated that, depending on the chain length, GS analogs alternate between structured and unstructured forms. In particular, peptides containing 2(2n+1) residues (where n=1,2,3...) show stable β–hairpin structures, while the others form disordered "random-coil" structures. Moreover, when the hairpin is present, it is bordered by two type II' β–turns. It has been in fact demonstrated that only type I' and II' β–turns have the correct geometry to align the residues at the end of the hairpin.[15] β–turn is a four-residue sequence (denoted *i* to *i+3*) that causes a reversal in direction of the peptide backbone. It has been shown[16] that the heterochirality of the turn is required for type II' turn

formation and therefore for hairpin formation. Moreover, the role of side-chain effect and N-substitution on type II' turn formation is crucial: in particular, turns in which D-tyrosine and proline occupy respectively the position i+1 and i+2 promote the β–hairpin formation.

GS analogs containing 6, 10 and 14 residues (the correct number of residues to form the β–structure), were experimentally studied in order to characterize the dynamics of hairpin formation.[17] Using Fourier transform infrared spectroscopy (FTIR) the absorption spectrum of the three peptides has been studied in the temperature range from 1°C to 85°C. The equilibrium changes in the FTIR spectra as a function of temperature were modeled in terms of two states, corresponding to the folded and unfolded structures. The relaxation kinetics of the folded/unfolded transition following a laser induced T-jump were studied, obtaining, by means of single exponential fit, the relaxation rate of each peptide. Folding and unfolding rates were extracted from the observed relaxation rates using the equilibrium constant determined from the static FTIR data. In particular, for the 6-meric peptide, that is for the peptide chosen for our MD simulations, the evaluated unfolding rate is $1.4\ 10^7 s^{-1}$.

In the present study we have performed MD simulations to characterize the kinetics and dynamics of the folding process. The 6-meric peptide was chosen among the three GS analogs because of its shortness and the consequent reduced computational power required. Unfortunately, the structural data obtained by NMR experiments are not deposited in any data bank: the starting atomic coordinates in the simulation were thus reconstructed starting from the primary sequence of the peptide [cyclo(Lys-dTyr-Pro)$_2$]. The 6 amino acids were arranged as an extended linear structure. This extended structure was then made cyclic and equilibrated before the productive MD simulation.

Two all-atoms MD simulations of the 6-meric analog of GS were carried out: the first one at the melting temperature of the peptide, that is at T=310K, with a simulation length of 300ns; the second at T=400K with a simulation length of 250ns. The simulations were performed with the GROMACS software package,[18] using the GROMOS96 forcefield[19] and a time step of 2 fs for numerical integration. The peptide was solvated with water, that was modeled by the simple point charge (SPC) model,[20] and placed in a cubic box. The simulations were performed in the canonical (NVT) ensemble with temperature-coupling obtained by the isogaussian method,[21] In both simulations two negative counterions were added to obtain a neutral condition.

The analysis of the time evolution of secondary structure, performed with the DSSP program (Figure1), showed the formation of β–turns but it didn't give any evidence of the folded/unfolded transition. As shown in Figure 2, the structure is substantially stable during the simulated time, both at 310K and at 400K. The formation of the first turn is clear in the central residues of the sequence, that is Pro 3 and Lys 4. The second turn is not observable in the figure: the corresponding residues, Pro 6 and Lys 1, are reported as unstructured. Actually, this is a sham caused by the inability of the DSSP program to treat with cyclic structures. It has been demonstrated, by

changing the numbering of the residues, that also Proline 6 and Lysine 1 constitute a turn.

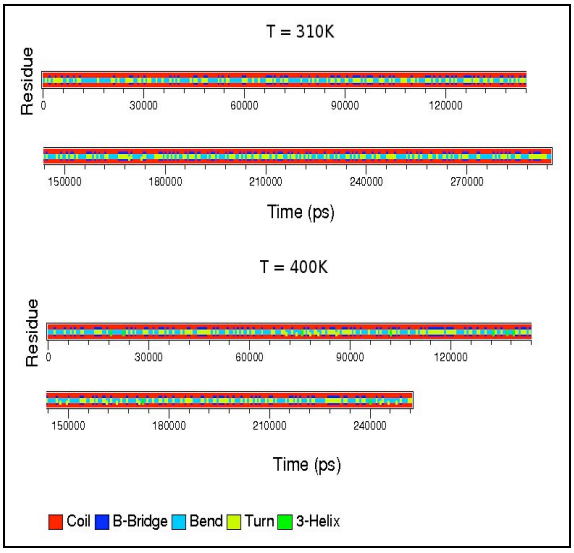

Fig. 2: Time evolution of the secondary structure.Y-axis denotes residue number and x-axis the simulation time. Secondary structures are color-coded as shown at the bottom of the figure.

The two residues of tyrosine assume a β–bridge conformation stabilizing the hairpin formation with two hydrogen bonds, the first between the C=O group of Tyr 2 and the N-H group of Tyr 5, the second between the C=O group of Tyr 5 and the N-H group of Tyr 2. The stability of these two hydrogen bonds and the consequent stability of the secondary structure during the simulated time was confirmed by analyzing the distance between the backbone atoms O and N of the tyrosine residues (Figure 3).

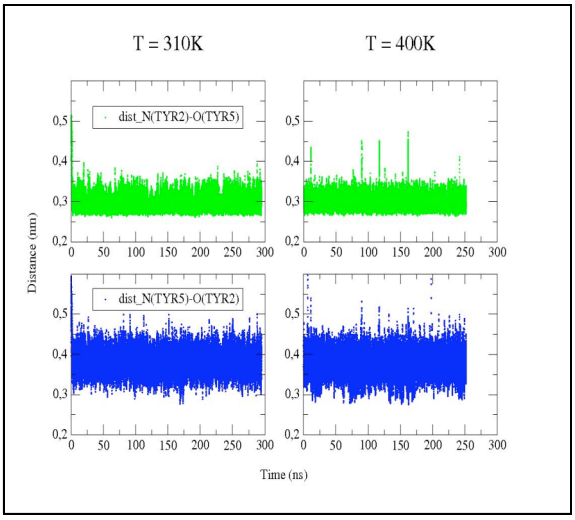

Fig. 3: Time evolution of the distance between the groups C-O and N-H of the two tyrosine residues at both simulated temperatures.

The outcome structure is then a β–hairpin bordered by two β–turns in which Proline and Lysine occupy respectively the position *i+1* and *i+2* and the residues of both tyrosine occupy the positions *i* and *i+3* in each turn (Figure 4). It has to be noted that the β–hairpin is different from the one predicted by the NMR experimental data in which tyrosine occupies the *i+1* position. Moreover, as previously reported, no secondary structure transition is evident from the simulated data.

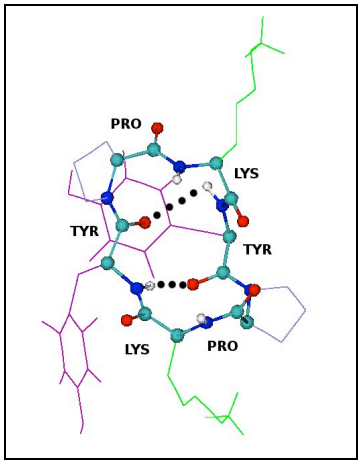

Fig. 4: Snapshot of the structure at T=310K. The hydrogen bonds between the two tyrosines are shown with dotted lines.

Although our structure does not correspond to the NMR one and our simulation does not show any structural transition, we have calculated the infrared spectrum due to the carboxilic groups. As the IR spectrum largely depends on the electric field on the C=O groups, exerted by all the atoms of the system, we have calculated the projection of the electric field on the C=O groups of the proline and lysine residues. The experimental IR spectra were interpreted as caused by a folding/unfolding transition as the C=O vibrational frequencies depend on the electric field that acts on them.
The evaluation of the projection of the electric field, which has been carried out with a purpose-built program for such computations, shows that the electric field acting on the C=O groups is substantially stable during the simulated time, both at 310K and at 400K (Figure 5A). The histogram of the amplitudes of the electric field, which is proportional to the frequency spectrum, shows that at both temperatures the absorption frequency is centered on a unique value, as no transitions are present (Figure 5B).

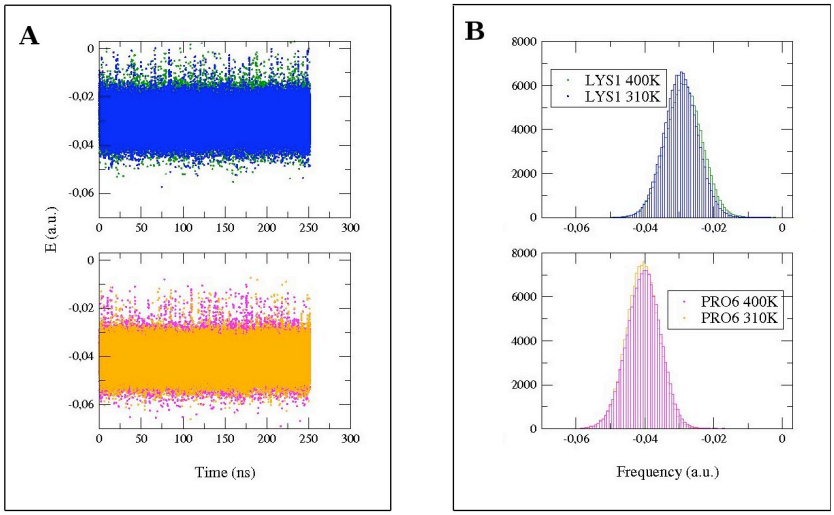

Fig. 5: A: time evolution of the projection of the electric field acting on the C=O group of Lys 1 at 310K (green line) and 400 K (blue line) and on the same group of Pro 6 at 310K (orange line) and at 400K (pink line). The electric field is measured in atomic units.

Fig. 5: B: Histogram of the amplitudes of the electric field acting on the C=O group af Lys 1 at 310K (green line) and 400 K (blue line) and on the same group of Pro 6 at 310K (orange line) and at 400K (pink line)

However, it has to be pointed out that the difference spectrum obtained by subtracting the spectrum at 310K from the spectrum at 400K has, for both residues, the same trend of the experimental temperature-dependent difference spectra (Figure 6).

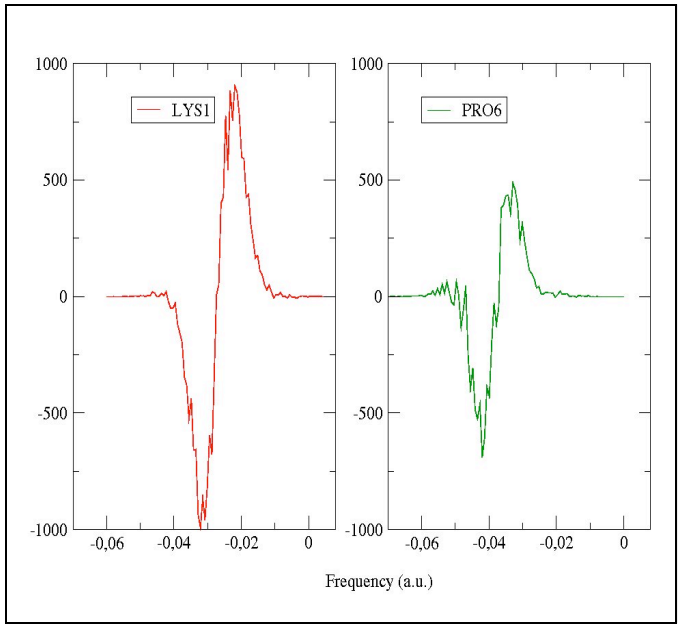

Fig. 6: Difference spectra obtained by subtracting the spectrum at 310K from the spectrum at 400K of the residues Lysine 1 (red line) and Proline 6 (green line)

As no transitions were observed in the MD simulation, according to our analysis, the difference spectrum behavior has to be ascribed to a different arrangement of the water molecules. Our data can be ascribed to a variation of the probability density of the state populated by the solvent molecules. This variation should proceed continuously with temperature raising.

However, we can not exclude that different conformations, such as the one obtained by NMR measurements, in which the tyrosine occupies a different position in the turn, or unfolded structures, give rise at the same spectral behavior.

The difference between the NMR structure and the simulated one requires a deepening of these preliminary data: the aim of our following works will be to investigate what causes the difference in the β-turns between the simulated and the NMR data. It is also important to understand if the difference between the NMR structure and the simulated one could be imputable to the model used to reproduce in silico the experimental environment.

Moreover, it could be interesting to simulate the kinetics of the larger peptides, i.e. the 10-meric and the 14-meric analogs of GS, in order to

characterize their structural dynamics.
A complete characterization of the above mentioned GS analogs could aim to investigate in silico the effect of mutations of the peptide's structure on its dynamics and consequently to help the design of new antibiotics.

References

1. Callender, R.H.; Dyer, R.B.; Gilmanshin, R. and Woodruff W.H. 1998 *Annu. Rev. Phys. Chem*. **49**: 173-202
2. Eaton, W.A.; Munoz, V.; Thompson, P.A.; Henry, E.R. and Hofrichter J. 1998 *Acc. Chem. Res.* **31**: 745-753
3. Werner, J.H.; Dyer, R.B.; Fesinmeyer, R.M. and Andersen, N.H. 2002 *J. Phys. Chem. B*. **106**: 487-494
4. Williams, S.; Causgrove, T.P.; Gilmanshin, R.; Fang, K.S.; Callender, R.H., Woodruff, W.H. and Dyer, R.B. 1996 *Biochemistry*. **35**: 691-6973
5. Jäger M.; Nguyen H.; Crane J.C.; Kelly J.W. and Gruebele M. 2001 Journ. Molec. Biol. **311**(2): 373-393
6. Griffiths-Jones, SR; Maynard, AJ and Searle, MS. 1999 *Journ. Molec. Biol.* **292**(5): 1051-1069
7. Daidone, I.; Amadei, A. and Di Nola, A. 2005 *Proteins: Struct. Funct. Bioinf.* **59**:510-518
8. Gnanakaran, S.; Nymeyer, H.; Portman, J.; Sanbonmatsu, K. and Garcia, A. 2003 *Curr. Opin. Struct. Biol.* **13**: 168-174
9. Olsen, K.A.; Fesinmeyer, R.M.; Stewart, J.M. and Andersen N. H. 2005 *Proc. Nat. Acad. Sci.* **102**(43): 15483-15487
10. Munoz, V.; Thompson, P.; hofrichter, J. and Eaton, W. 1997 *Nature*. **390**: 196-199
11. Daidone, I.; D'Abramo, M.; Di Nola, A. and Amadei, A. 2005 *J. Am. Chem. Soc.* **127**: 14825-14832
12. Daidone, I.; Amadei, A.; Roccatano, D. and Di Nola, A. 2003 Biophys. Journ. **85**: 2865-2871
13. Plotkin, S. and Onuchic, J. 2002 *Q. Rev. Biophys.* **35**: 111-167
14. Plotkin, S. and Onuchic, J. 2002 *Q. Rev. Biophys.* **35**: 205-286
15. Gibbs, A.C.; Kondejewski, L.H.; Gronwald, W.; Nip, A.M.; Hodges, R.S., Sykes, B.D. and Wishart, D.S. 1998 *Nat. Struct. Biol.* **5**(4): 284-288
16. Gibbs, A.C.; Bjorndahl, T.C.; Hodges, R.S. and Wishart, D.S. 2002 *J. Am. Chem. Soc.* **124**(7): 1203-1213
17. Maness, S.J.; Franzen, S.; Gibbs, A.C.; Causgrove, T.P. And Dyer R.B. 2003 *Biophys. J.* **84**: 3874-3882
18. Van der Spoel, D.; Van Drunen, R. and Berendsen, H.J.C. 1994 GROningen Machine for Chemical Simulations, Department of Biophysical Chemistry, BIOSON Research Institute, Nijenborgh 4 NL-9717 AG Groningen
19. Van Gunsteren, W.F.; Billeter, S.R.; Eising, A.A.; Hunenberger, P.H.; Kruger, P.; Mark, A.E.; Scott, W.R.P. and Tironi, I.G. 1996 Zurich: Hochshulverlag AG an der ETH Zurich
20. Berendsen, H.J.C.; Grigera, J.R. and Straatsma, T.P. 1987 *J. Phys.Chem.* **91**:

6269-6271
21. Brown, D. and Clarke, J.H.R. 1993 *J. Chem. Phys*. **98**: 10089-10092