Title: Examination of the Effect of N-terminal Diproline and Charged Side Chains on the Stabilization of Helical Conformation in Alanine-based Short Peptides: A Molecular Dynamics Study

Authors: Bhupesh Goyal, Ph.D.; Kinshuk Srivastava; Susheel Durani

To be cited as: 10.1002/slct.201601381

Link to VoR: https://doi.org/10.1002/slct.201601381

Examination of the Effect of N-terminal Diproline and Charged Side Chains on the Stabilization of Helical Conformation in Alanine-based Short Peptides: A Molecular Dynamics Study

Bhupesh Goyal,[*‡] *Kinshuk Raj Srivastava,*[§] *and Susheel Durani*[*] [a]

[a] *Dr. Bhupesh Goyal, Dr. Kinshuk R. Srivastava, Prof. Susheel Durani*

Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai–400076, India

E-mail: bhupesh@iitbombay.org; sdurani@iitb.ac.in

Present address

[‡] *Department of Chemistry, School of Basic and Applied Sciences, Sri Guru Granth Sahib World University, Fatehgarh Sahib–140406, Punjab, India*

[§] *Department of Physics and Astronomy, Michigan State University, East Lansing, USA*

Abstract

The effect of N-terminal diproline segment and charged side chains on the stabilization of helical conformation in alanine-based short peptides are examined using molecular dynamics (MD) simulations. The cationic peptides, Ac–Pro_1–Pro_2–Ala_3–Lys_4–Ala_5–Lys_6–Ala_7–Lys_8–Ala_9–NH_2 (**Ia**) and Ac–$^{D}Pro_1$–Pro_2–Ala_3–Lys_4–Ala_5–Lys_6–Ala_7–Lys_8–Ala_9–NH_2 (**IIa**) are examined for the role of lysine side chains on the inducement of helical conformation in alanine-based short peptides. To examine the influence of lysine and glutamic acid in the *i*, *i* + 4 arrangement on the stabilization of helical conformation, cationic peptides, **Ia** and **IIa**, are modified as ion-pair peptides, Ac–Pro_1–Pro_2–Glu_3–Glu_4–Ala_5–Ala_6–Lys_7–Lys_8–Ala_9–NH_2 (**Ib**) and Ac–$^{D}Pro_1$–Pro_2–Glu_3–Glu_4–Ala_5–Ala_6–Lys_7–Lys_8–Ala_9–NH_2 (**IIb**), respectively. MD simulations manifest enhanced occupancies in the α basin of φ, ψ space for ion-pair peptides as compare to cationic peptides. The radial distribution function (RDF) analysis highlight that large side chain substituents of lysine and glutamic acid assist in helix formation by blocking water molecules from solvating backbone CO and NH groups.

Introduction

The native state of proteins is stabilized by a complex interplay of hydrophobic effect, van der Waals forces, hydrogen bonding, electrostatic contribution of salt bridges, and helix-dipole interactions, however, relative contribution of these interactions remains unclear. The elucidation of quantitative contribution of these interactions remains a formidable challenge given the size of a typical protein and the complexity of its interactions.[1] To illuminate the underlying principles that govern stabilization of the native state of proteins, bottom-up approach of simple to incrementally complex models have been adopted. The alanine-based short peptides have been utilized to analyze the conformational preferences of the polypeptide chain and addressing the basis using computer simulations.[2] The empirical force fields have been applied for simulation of equilibria to address the thermodynamics with rigor.[3] During last years, alanine-based short peptides are emerged as the protein main chain models to elucidate the underlying fundamental forces that govern protein folding-unfolding equilibrium.[4] The models highlighted that unfolded proteins adopt appreciable order as semi-extended structures in correspondence of PPII conformation.[4h-4j,4l,4n,4o]

The α-helix is one of the most important structural domains in proteins and peptides that control numerous biological activities and functions.[5] The studies focused on the enhancement of overall helicity and stability of short helical peptides have contributed to the fundamental understanding of protein folding-unfolding equilibrium and have led to improvement in the biological and pharmaceutical activities.[6] The use of short synthetic peptides encompassing helical segments to modulate protein-protein interactions (PPIs) associated with human diseases represent great pharmacological interest. The majority of the PPIs involves α-helices, and has large interaction areas and shallow surfaces; thus, small molecule inhibitors are not effective for modulating PPIs.[7] During last years, constrained

peptides have been developed for modulating PPIs by enhancing the helicity of short peptides.[8] Thus, a better understanding of the α-helix structure and elucidation of factors that dictate its structure is of key importance.

The role of non-covalent interactions and their effects in folding and stability of α-helix and β-sheet peptides have been highlighted in literature.[9] The electrostatic interactions among main chain, side chains are well recognized for their role in promoting conformational folding in polypeptide structure.[10] The folding simulations of all-alanine peptides and a number of short alanine-based helical peptides with positively or negatively charged residues have highlighted the role of hydrophobic interaction and charged side chains in the folding of α-helical peptides.[9a,9e,9f,9i] In addition, computer simulations of these short helical alanine-based peptides and the peptides with salt bridge pairs have provided deeper insights into the role of charged side chains in the stability of helical peptides.[2e,9g] Meuzelaar *et al.* have investigated the effect of salt bridges between different types of charged amino-acid residue pairs on α-helix folding using a combination of ultraviolet circular dichroism, temperature-jump transient-infrared spectroscopy, and molecular dynamics simulations.[11] The authors highlighted that stabilizing salt bridges speed up α-helix formation by up to 50% and slow down the unfolding of the α-helix, whereas salt bridges with an unfavorable geometry have the opposite effect. Walker *et al.* highlighted the contribution of arginine-glutamate salt bridges to the helix stability of the peptide with the sequence AAARAAAAEAAEAAAARA.[12] The present study aim to exploit N-terminal diproline of homochiral and heterochiral structure, and charge-group effect over side chains to assess the role of specific structure modification planned in conformational equilibria of the model peptides.

The effect of N-terminal residue stereochemical mutation from L- to D-structure and chain-length have been addressed with molecular dynamics for the possibility of inducement of helical conformation in the oligoalanine models.[13] MD simulations reveal promotion of small fraction of helical conformation in oligoalanine models that involve specific effect of stereochemical modification in the N-terminal residue and chain-length. The role of stereochemical modification of amino acid residue from L- to D-structure in delineation of protein folding mechanism, [14] to increase the stability of proteins,[15] in the redesign an active and specific ion channel, [16] and in the design of novel folds is reported.[17] The present study address the alanine-based nonapeptide of poly-L structure for the stereochemical effect in N-terminal residue and charge-group effect over side chains that are capable of ordering the peptide as a helical fold. The model peptides are examined with molecular dynamics (MD) to assess the possible contribution of extended-β, semi-extended PPII and helical conformations in the equilibrium ensemble. The implications for understanding of the inducement of helical conformation in alanine-based short peptides are discussed. The results of the present study will aid in the design of novel peptides with helical structures. The role of helical structures in the rational design of biocompatible hydrogels and inhibition of disease-relevant intracellular or extracellular PPIs is a topic of current research.[18]

Results and discussion

The study is implemented with nine-residue length peptides (Table 1). The models are primarily alanine-based short peptides, and thus the sequences of an intrinsically helix favoring residue. The model alanine-based short peptide is substituted with internal charged side chains, Lys and Glu residues, to promote helical fold. The model peptides are equipped with diproline segment of homochiral and heterochiral structure for possible inducement of helical conformation. The diproline segments have been reported as potential nuclei for

initiating helical folding in peptides.[19] Kemp *et al.* have highlighted that covalently constrained diproline surrogate as the effective template for inducing helical conformations in short acyclic sequences.[20] Thus, a combination of N-terminal diproline of homochiral, heterochiral structure and charge-group effect over side chains is examined for possible inducement of helical folds. The cationic peptides, **Ia** and **IIa**, are designed to assess the effect of N-terminal diproline of homochiral and heterochiral structure, respectively, and the role of lysine side chains for the inducement of helical conformation. The ion-pair peptides, **Ib** and **IIb**, with N-terminal diproline of homochiral and heterochiral structure, respectively, are designed to assess the influence of large side chain substituents of lysine and glutamic acid residues in the *i*, *i* + 4 arrangement on the nucleation of helical conformation.

Table 1. The end-protected cationic peptides, **Ia** and **IIa**, and ion-pair peptides, **Ib** and **IIb** chosen for molecular dynamics. The peptides **Ia** and **Ib** have N-terminal homochiral diproline segment while peptides **IIa** and **IIb** have N-terminal heterochiral diproline segment for the inducement of helical conformation.

Model	Alanine-based short peptides
Ia	$Ac–Pro_1–Pro_2–Ala_3–Lys_4–Ala_5–Lys_6–Ala_7–Lys_8–Ala_9–NH_2$
Ib	$Ac–Pro_1–Pro_2–Glu_3–Glu_4–Ala_5–Ala_6–Lys_7–Lys_8–Ala_9–NH_2$
IIa	$Ac–{}^{D}Pro_1–Pro_2–Ala_3–Lys_4–Ala_5–Lys_6–Ala_7–Lys_8–Ala_9–NH_2$
IIb	$Ac–{}^{D}Pro_1–Pro_2–Glu_3–Glu_4–Ala_5–Ala_6–Lys_7–Lys_8–Ala_9–NH_2$

The molecular dynamics ensembles were prepared with GROMOS96 43a1 force field as described in the computational section. The GROMOS force field was adopted for the present study as it has been widely used for the conformational analysis of peptides.[21] Best *et al.* have reported large deviations with the experimental data in the computational studies of $(AAQAA)_3$ peptide using replica exchange simulations with CHARMM22/CMAP, AMBER99SB, and AMBER03 force fields.[3d] CHARMM22/CMAP and AMBER03 overstabilized the helix *i.e.* 95% and 87% helix at 300 K, respectively, whereas AMBER99SB understabilized the helix *i.e.* 2% at 300 K. Relative to experimental

measurements, the α-helical propensity of AMBER03 force field[22] is too high, while α-helical propensity is too low for AMBER99SB.[3g,23] The OPLS force field was regarded as the best force field in the description of the microstructures of organic molecules (*i.e.* liquid benzene).[24] Gerben *et al.* have compared the popular atomistic force fields, AMBER03, CHARMM22 + CMAP, GROMOS96, and OPLS-AA, to examine the force field that yield results in accordance with the experimental results using Aβ peptide.[25] The authors highlighted that AMBER03 and CHARMM22 + CMAP over-stabilize helical structure within Aβ by comparing secondary structure content, NMR shifts, and radius-of-gyration (R_g) to available experimental data. On the other hand, OPLS-AA and GROMOS96 yield helical and β-strand content, calculated NMR shifts, and radius-of-gyration that agree well with experimental data.

The model peptides were submitted to molecular dynamics for long enough duration to achieve equilibrium. For assessment of attainment of equilibrium, MD trajectories were evaluated for time dependent evolution in microstates of the polypeptide structure. This enumeration was approached with clustering algorithm of Daura *et al.*[26] The polypeptide structures populating MD trajectory were clustered in Cartesian space with root-mean-square deviation (RMSD) cutoff ≤ 0.15 nm over backbone atoms (N, Cβ, Cα, C), giving microstates diminishing in population, *viz.*, diminishing thermodynamic stability. The molecular dynamics ensembles were assessed in radius-of-gyration (R_g) over populated conformers. In addition to R_g, the ensembles were assessed in the occupancy of specific φ, ψ basins, and in the main-chain hydrogen bonds that are short-ranged (SR; i→i ± 2), medium-ranged (MR; i→i ± 3; i→i ± 4), and long-ranged (LR; i→i ± 5; i→i ± ≥ 6) according to sequence separation between donor and acceptor residue.

The time dependent evolution of MD trajectories in microstates of polypeptide structure is shown in Figure 1. The ion-pair peptides, **Ib** and **IIb**, are noted to achieve equilibrium early and saturate to a defined population in microstates. The cationic peptides, **Ia** and **IIa**, are noted to evolve more slowly and do not attain robust asymptote over the observation time of 250 ns. Assuming reasonable approximation of equilibria, simulations were terminated at the time points noted in Figure 1 and the ensembles were compared as macrostates and over the microstates. As noted in Table 2, the ion-pair peptides, **Ib** and **IIb**, populate in much smaller number of microstates, 280 and 219, respectively, compare to cationic peptides, **Ia** and **IIa**, 564 and 590, respectively. Correspondingly, the population of most-populated microstate increased from 14.8% and 14.5% for **Ia** and **IIa**, respectively, to 16.9% and 23.3% for **Ib** and **IIb**, respectively. Thus, conformational ensemble is less heterogenous for ion-pair peptides, **Ib** and **IIb**, as compare to cationic peptides, **Ia** and **IIa**. Thus, a combined effect of N-terminal diproline of homochiral, heterochiral structure and large side chain substituents of lysine and glutamic acid residues in the *i*, *i* + 4 arrangement reduce the conformational diversity in ion-pair peptides, **Ib** and **IIb**.

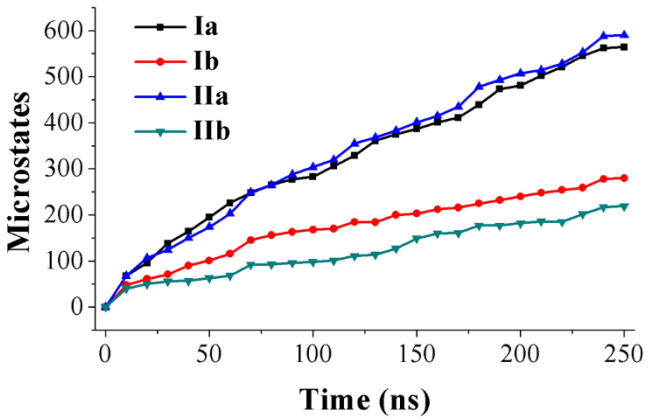


Figure 1. The evolution of microstates over end-protected cationic peptides, **Ia** and **IIa**, and ion-pair peptides, **Ib** and **IIb**, during molecular dynamics in water as explicit-solvent. Y-axis represent the number of microstates and X-axis represent molecular dynamics simulation time in ns.

Table 2. Population statistics, specific structural and conformational properties of macrostates of end-protected cationic peptides, **Ia** and **IIa**, and ion-pair peptides, **Ib** and **IIb**.

Model	No. of microstates	% Pop. in m_1[a]	R_g (nm)		φ, ψ distribution in M[b]			Average number of hydrogen bonds in M[b]			
			M[b]	m_1[a]	% α[c]	% β[c]	% PPII[c]	Avg. /Conf.[d]	SR[e]	MR[e]	LR[e]
Ia	564	14.8	0.56 ± 0.13	0.46 ± 0.02	**19.2**	25.9	44.3	1.1	0.28	**0.46**	0.33
Ib	280	16.9	0.52 ± 0.08	0.49 ± 0.02	**28.5**	23.4	39.2	1.2	0.17	**0.70**	0.32
IIa	590	14.5	0.52 ± 0.10	0.47 ± 0.02	**19.9**	25.4	45.1	1.2	0.27	**0.53**	0.44
IIb	219	23.3	0.49 ± 0.07	0.46 ± 0.02	**28.1**	22.6	40.7	1.3	0.19	**0.78**	0.36

[a] m_1: first microstate (most-populated); [b] M: Macrostate; [c] Basin definitions are, α: $^{L/D}\varphi$= −/+ 20 to −/+ 100, $^{L/D}\psi$ = −/+ 20 to −/+ 80; β: $^{L/D}\varphi$= −/+ 90 to −/+ 170, $^{L/D}\psi$ = +/− 80 to +/− 180; PPII: $^{L/D}\varphi$= −/+ 30 to −/+ 90, $^{L/D}\psi$ = +/− 80 to +/− 170; [d] The total number of hydrogen bonds during the entire simulation divided by the total number of conformations sampled during simulation is defined as average number of hydrogen bonds per conformation (Avg./Conf.); [e] Hydrogen bonds are short-ranged (SR; i→i ± 2), medium-ranged (MR; i→i ± 3, i→i ± 4) and long-ranged (LR; i→i ± 5, i→i ± ≥ 6) according to sequence separation between donor and acceptor residue.

The equilibria were analyzed on basis of radius-of-gyration (R_g) over polypeptide structure to highlight the state of "folding" or "unfolding". The results depicted in Figure 2 manifest "folding" to a bent conformation involving close spatial proximity between *N*- and *C*-termini in the model peptides.

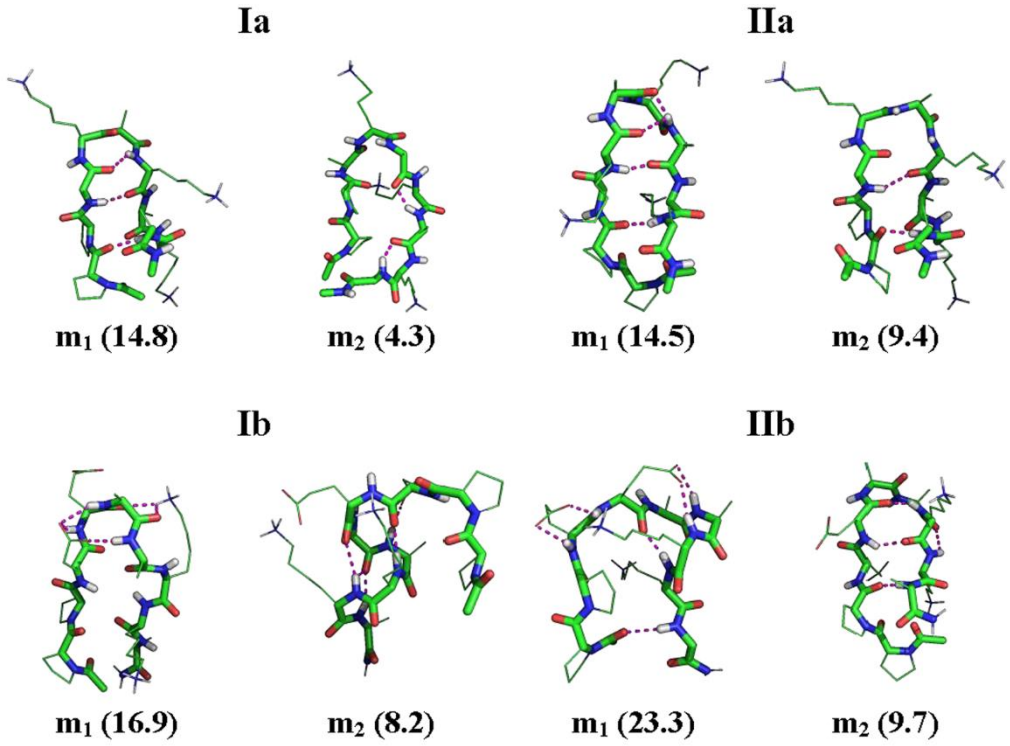


Figure 2. The central member of two most-populated microstates ($\mathbf{m_1}$ and $\mathbf{m_2}$) of end-protected cationic peptides, **Ia** and **IIa** (upper panel) and ion-pair peptides, **Ib** and **IIb** (lower panel) are shown in the stick representation. The hydrogen bonds among NH, C=O groups are shown in purple dashed lines. The percent population of each microstate is shown in parenthesis.

From the statistics of basin occupancy reported in Table 2, the model peptides have occupancy in α, β and PPII basins which highlight that model peptides adopt multiple conformations in water. These results are consistent with those previously reported by Dalgicdir *et al.* which highlight that two synthetic peptides, LKKLLKLLKKLLKL (LK) and EAALAEALAEALAE (EALA), adopt neither random coil nor fully formed α-helical

structure in water.[27] Using molecular dynamics simulations, the authors reported that the peptides adopt multiple conformations with short lifetimes in water. [27]

As noted from Table 2, the occupancy of α basin is much higher for ion-pair peptides as compare to cationic peptides. The occupancy in α basin is 28.5%, 28.1% for **Ib**, **IIb**, respectively, while the occupancy is only 19.2%, 19.9% for **Ia**, **IIa**, respectively. The φ, ψ spread and preferential basin occupancies of macrostate over end-protected cationic peptides, **Ia**, **IIa** and ion-pair peptides, **Ib**, **IIb** during molecular dynamics simulation is shown in Figure 3. As depicted in Figure 3, the occupancy of α basin is much higher for ion-pair peptides than cationic peptides. Correspondingly, the occupancy in PPII basin decrease from 44.3%, 45.1% for **Ia**, **IIa**, respectively, to 39.2%, 40.7% for **Ib**, **IIb**, respectively (Table 2).

We have not observed a significant change in the occupancies of α, β, and PPII-basin in the ϕ, ψ space by varying the N-terminal diproline segment from homochiral (^{L}Pro–^{L}Pro) to the heterochiral structure (^{D}Pro–^{L}Pro) which is consistent with the results from the previous study.[13b] The molecular dynamics simulations highlighted that two diastereomeric peptides, Ac-^{L}Pro-^{L}Pro-^{L}Ala-^{L}Ala-NHMe and Ac-^{D}Pro-^{L}Pro-^{L}Ala-^{L}Ala-NHMe, sample identical structures in the conformational ensemble.[13b] These results are consistent with the results reported by M. Oba *et al.*[28] The authors have investigated the solid-state conformation of diastereomeric -Pro-Pro-$(Aib)_4$ sequences. The authors have attached the two diastereomeric diproline (L-Pro-L-Pro and D-Pro-L-Pro) segments on the N-terminus of H-$(Aib)_4$-OMe segment. X-ray crystallographic analysis highlighted that the two diastereomeric hexapeptides, Cbz-L-Pro-L-Pro-$(Aib)_4$-OMe (**1**) and Cbz-D-Pro-L-Pro-$(Aib)_4$-OMe (**2**), formed identical structures with different N-terminal Pro residues.

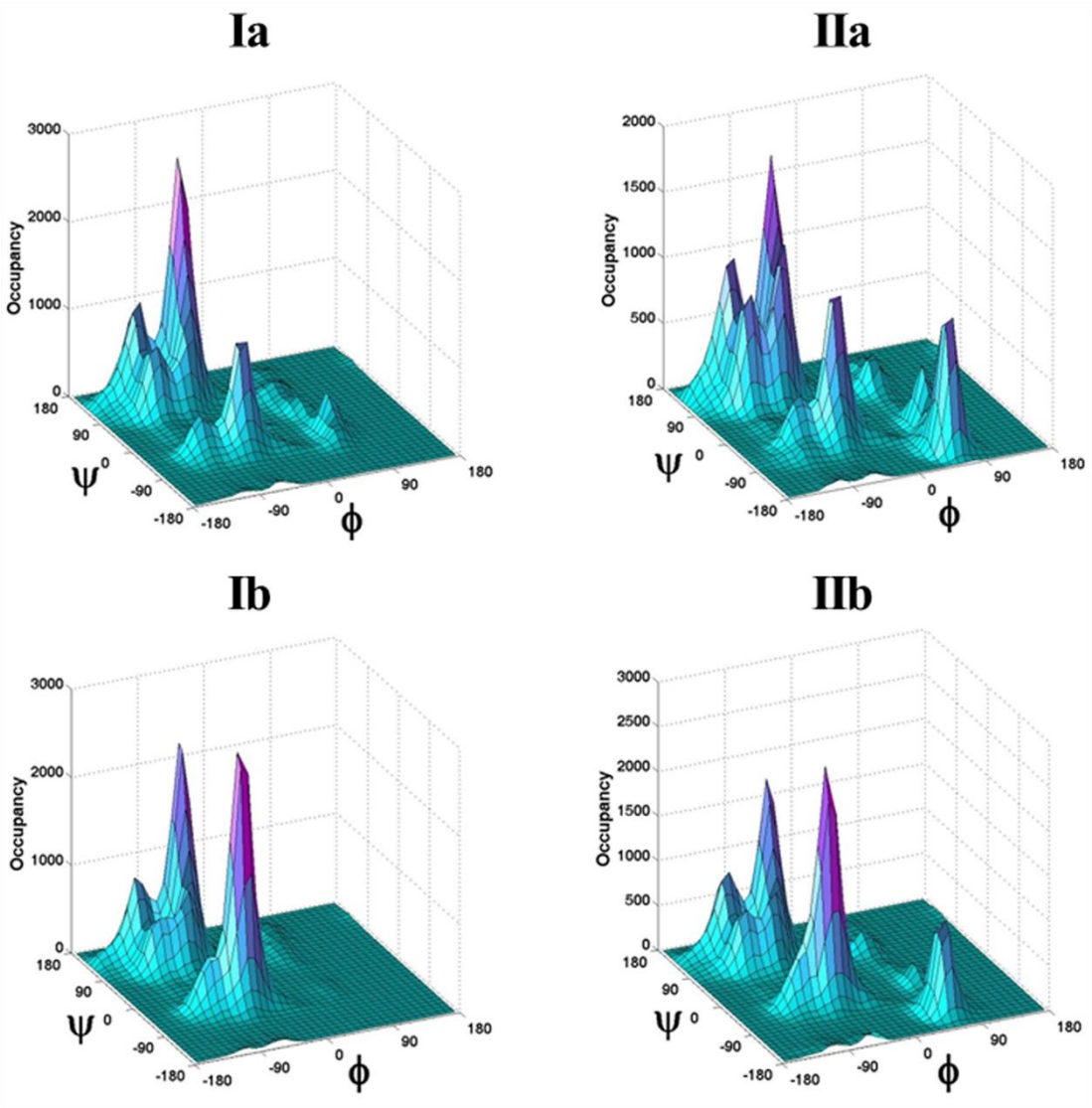


Figure 3. The φ, ψ spread and preferential basin occupancies of macrostates over end-protected cationic peptides, **Ia** and **IIa** (upper panel) and ion-pair peptides, **Ib** and **IIb** (lower panel) during molecular dynamics simulation.

Correlated with higher occupancy of α conformation for ion-pair peptides, a higher average number of MR hydrogen bonds are observed for ion-pair peptides as compare to cationic peptides. The average number of MR hydrogen bonds increased from 0.46, 0.53 for **Ia**, **IIa**, repectively, to 0.70, 0.78 for **Ib**, **IIb**, respectively. The total number of hydrogen bonds during the entire simulation divided by the total number of conformations sampled during simulation have been evaluated for each ensemble and we defined it as average number of

hydrogen bonds per conformation (Avg./Conf.). As listed in Table 2, the average number of hydrogen bonds per conformation are higher for ion-pair peptides as compare to cationic peptides which indicate more folded conformations were sampled for ion-pair peptides. Thus, a combined effect of N-terminal diproline of homochiral, heterochiral structure and large side chain substituents of lysine and glutamic acid residues in the *i*, *i* + 4 arrangement in ion-pair peptides, **Ib** and **IIb**, promote higher occupancy of the macrostate in the α basin.

To analyze the influence of N-terminal diproline of homochiral and heterochiral structure, cationic lysine side chains, and large side chain substituents of lysine and glutamic acid residues in the *i*, *i* + 4 arrangement, we resolve the canonical ensembles to the microstates. The stick representation of two most-populated microstates, $\mathbf{m_1}$ and $\mathbf{m_2}$, with percent population in parenthesis is shown in Figure 2. As depicted in Figure 2, the microstates of peptides **Ia**, **Ib**, **IIa**, **IIb** adopt a U-shape conformation in water, apparently maximized in intrachain interaction involving short, medium and long ranged hydrogen bonds of peptides, viz. SR, MR, and LR hydrogen bonds. Thus, according to molecular dynamics, the model peptides exist in the bent-shaped conformational folds in water apparently maximized in intrachain interaction. The ion-pair peptide **Ib** with large side chain substituents of lysine and glutamic acid residues in the *i*, *i* + 4 arrangement has the partially folded helical structure in water as observed in the second microstate (Figure 2).

Solvation shell analysis

To elucidate the effect of solvent on the conformation of model peptides, we calculated the radial distribution function (RDF) of solvent atoms against atoms of the model peptides. The calculated RDF of oxygen atom of solvent water around N, O, Nζ (lysine), Cβ (alanine), and Oε (glutamic acid) atoms of the model peptides are shown in Figure 4 and 5. The absence of first RDF maxima of oxygen atom of water against peptide-NH in the cationic as well as ion-

pair peptides suggest the involvement of peptide-NH groups in intrapeptide hydrogen bonding (Figure 4 and 5). Thus, the model peptides adopt U-shaped conformation in water which is also reflected in stick representation of the conformational folds of two most-populated microstates shown in Figure 2.

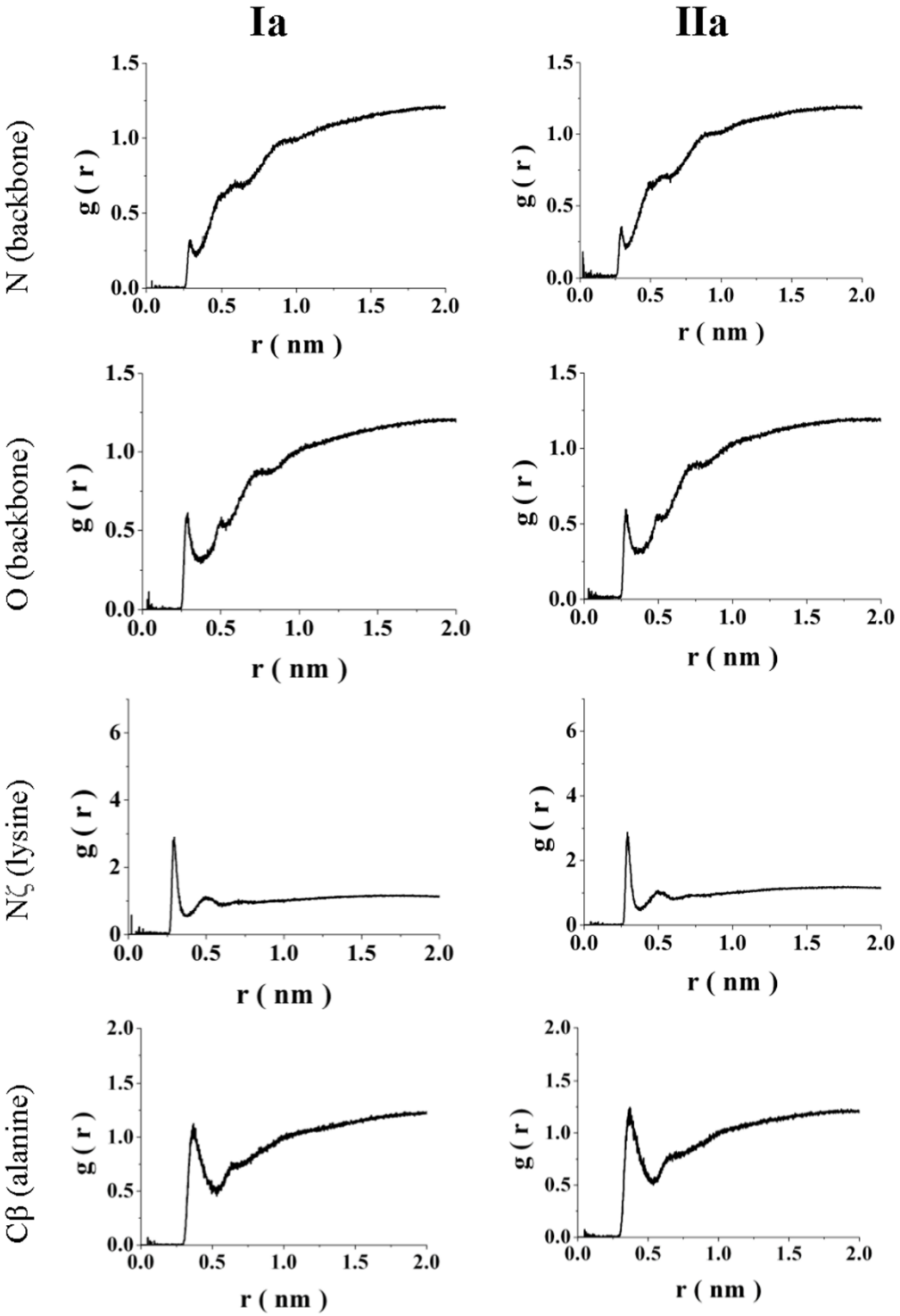


Figure 4. Radial distribution of oxygen atom of solvent water against N, O, Nζ (lysine), and Cβ (alanine) atoms of the most-populated microstate of cationic peptides, **Ia** (left panel) and **IIa** (right panel).

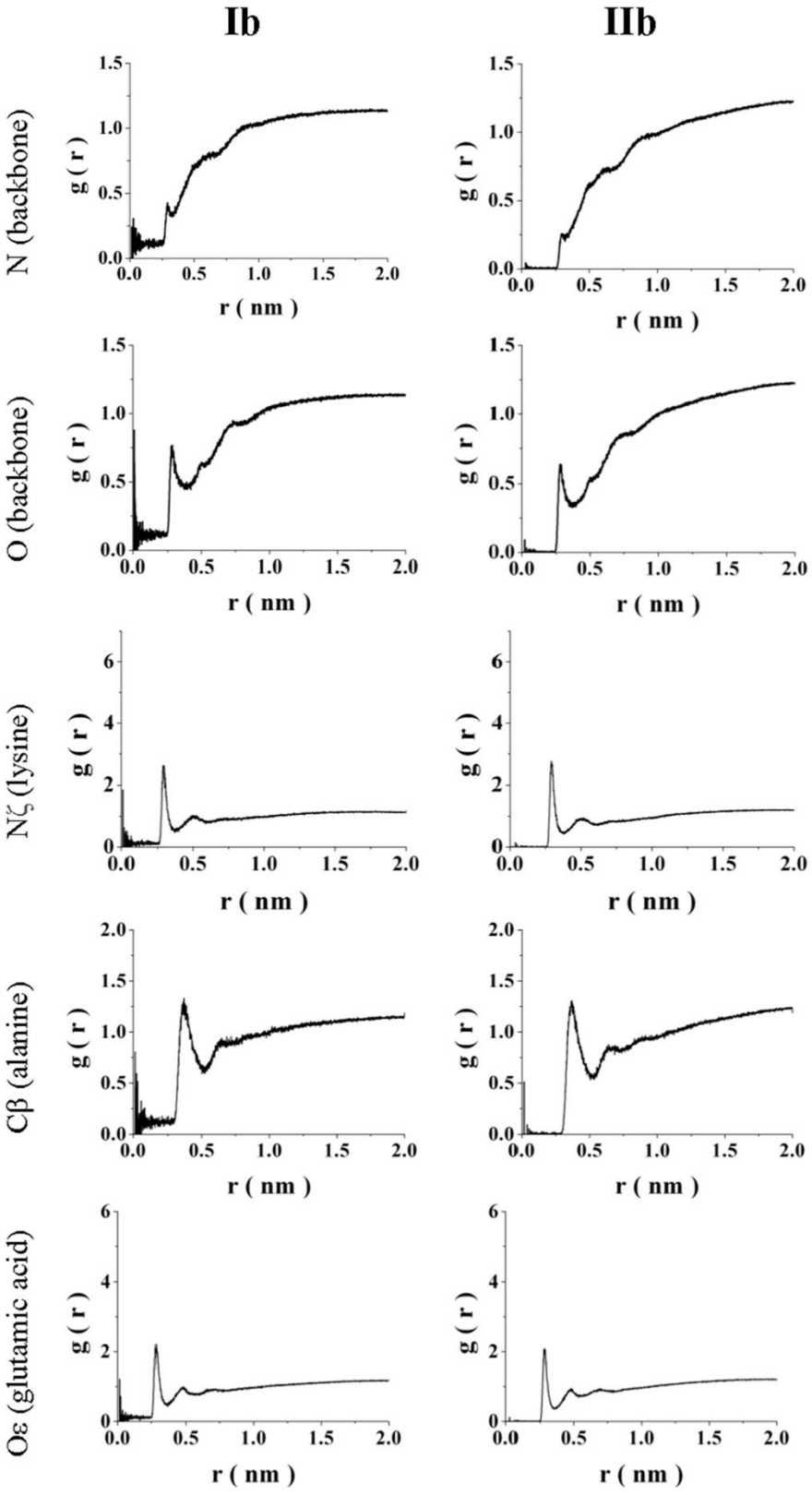


Figure 5. Radial distribution of oxygen atom of solvent water against N, O, Nζ (lysine), Cβ (alanine), and Oε (glutamic acid) atoms of the most-populated microstate of ion-pair peptides, **Ib** (left panel) and **IIb** (right panel).

The Cβ atom of alanine has first RDF maxima at 0.36 nm against oxygen of water which suggest methyl group of alanine may be solvated by combination of C-H…O hydrogen bonds and van der Waals (vdW) interactions. As depicted in Figure 4 and 5, Nζ atom of lysine has first RDF maxima at 0.30 nm against oxygen of water in cationic as well as ion-pair peptides that suggest solvation of the lysine side chain through hydrogen bonding with water. The Oε atom of glutamic acid has first RDF maxima at 0.30 nm against oxygen of water in ion-pair peptides that suggest solvation through hydrogen bonding with water (Figure 5). The RDF analysis highlight that charged side chains of the lysine and glutamic acid residues are solvated by water molecules. This imply that the large side chain substituents of lysine and glutamic acid residues assist in helix formation by blocking water molecules from solvating backbone CO and NH groups as observed in the formation of partially folded helical structure in the ion-pair peptide, **Ib**, which is in agreement with the earlier reports.[2d,2e,29,30]

It is reported in literature that short (n = 10) sequences of Ala peptides, Ala_n, do not form helices in water.[31] The NMR data for short polyalanine peptides, ($Ala_{3–7}$), highlighted that the peptides exist as polyproline II (PPII) helix-like structures with very little population in the α-helical conformation.[32] In the present study, MD simulations highlight inducement of helical conformation in a short alanine-based peptides by employing N-terminal diproline segment and lysine, glutamic acid residues in the *i*, *i* + 4 arrangement which is significant for a peptide of nine-residue length. The results of the present study will expand our understanding of peptide and protein folding, and will aid in the design of novel peptides with α-helical structures which can modulate PPIs.

Conclusions

In the present study, we have examined the effect of N-terminal diproline segment and charged side chains in the alanine-based short peptides as an approach to scrutinize the specific role of interactions within main chain and between side chains on the inducement of helical conformation. MD simulations reveal enhanced occupancies in the α basin of φ, ψ space for ion-pair peptides, **Ib** and **IIb**, as compare to cationic peptides, **Ia** and **IIa**, while the occupancies are nearly identical whether the N-terminal diproline segment is homochiral or heterochiral in structure. Thus, MD simulations highlight sampling of the α-conformation in the model peptides largely depend on the large side chain substituents of lysine and glutamic acid residues in the *i*, *i* + 4 arrangement rather than stereochemical structure of the N-terminal diproline segment. The RDF analysis highlight that charged side chains of lysine and glutamic acid residues are solvated by water molecules which imply that large side chain substituents of lysine and glutamic acid residues assist in helix formation, although partially folded, by blocking water molecules from solvating backbone CO and NH groups. MD simulations highlight the influence of N-terminal diproline segment as well as large side chain substituents of lysine and glutamic acid residues in the *i*, *i* + 4 arrangement on the inducement of helical conformation in alanine-based short peptides. The present study will enhance our understanding on nucleation of helical conformation in short peptides and hence aid in the design of novel peptides with helical structures.

Acknowledgments

The authors acknowledge Department of Science and Technology, Ministry of Science and Technology, Government of India (Sanction No: 09DST028) for financial support and IIT Bombay, Mumbai for computing facility "Corona". Bhupesh Goyal gratefully acknowledge Science and Engineering Research Board (SERB), Department of Science and Technology,

Ministry of Science and Technology, Government of India for the award of SERB Start-Up Research Grant (Young Scientists) (Sanction No: SB/FT/CS-013/2014).

Keywords: Alanine-based Short Peptides; Helical Conformation; Molecular Dynamics; N-terminal Diproline; Protein Folding

References

[1] a) S. Gianni, P. Jemth, *Biophys. Chem.* **2016**, *212*, 17–21; b) Y. Cote, G. G. Maisuradze, P. Delarue, H. A. Scheraga, P. Senet, *J. Phys. Chem. Lett.* **2015**, *6*, 1082–1086; c) P. G. Wolynes, *Biochimie* **2015**, *119*, 218–230; d) S. Sacquin-Mora, *J. R. Soc. Interface* **2015**, *12*, 20150876; e) S. W. Englander, L. Mayne, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15873–15880; f) K. A. Dill, J. L. MacCallum, *Science* **2012**, *338*, 1042–1046; g) K. A. Dill, S. B. Ozkan, M. S. Shell, T. R. Weikl, *Annu. Rev. Biophys.* **2008**, *37*, 289–316; h) V. Daggett, A. Fersht, *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 497–502.

[2] a) M. Compiani, E. Capriotti, *Biochemistry* **2013**, *52*, 8601–8624; b) G. S. Jas, K. Kuczera, *Mol. Simul.* **2012**, *38*, 682–684; c) W-Z. Wang, T. Lin, Y-C. Sun, *J. Phys. Chem. B* **2007**, *111*, 3508–3514; d) T. Ghosh, S. Garde, A. E. García, *Biophys. J.* **2003**, *85*, 3187–3193; e) A. E. Garcia, K. Y. Sanbonmatsu, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2782–2787.

[3] a) P. Faccioli, S. a Beccara, *Biophys. Chem.* **2016**, *208*, 62–67; b) R. Raucci, G. Colonna, G. Castello, S. Costantini, *Int. J. Pept. Res. Ther.* **2013**, *19*, 117–123; c) B. R. Brooks, et al. *J. Comput. Chem.* **2009**, *30*, 1545–1614; d) R. B. Best, G. Hummer, *J. Phy. Chem. B* **2009**, *113*, 9004–9015; e) O. Guvench, A. D. MacKerell Jr., *Methods Mol. Biol.* **2008**, *443*, 63–88; f) W. F. van Gunsteren, J. Dolenc, A.E. Mark, *Curr. Opin. Struct. Biol.* **2008**, *18*, 149–153; g) R. B. Best, N. Buchete, G. Hummer, *Biophys. J.* **2008**, *95*, L07–L09; h) Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, *J. Comput. Chem.* **2003**, *24*, 1999–2012.

[4] a) C. Calero-Rubio, B. Paik, X. Jia, K. L. Kiick, C. J. Roberts, *Biophys. Chem.* **2016**, *217*, 8–19; b) C. Liu, J. W. Ponder, G. R. Marshall, *Proteins: Struct., Funct., Bioinf.* **2014**, *82*, 3043–3061; c) J. C. Schulz, M. S. Miettinen, R. R. Netz, *J. Phys. Chem. B* **2015**, *119*, 4565–4574; d) M. Zhang, D. Wei, M. Tang, C. Shi, H.-L. Cui, C. Du, *Mol. Simul.* **2016**, *42*, 398–404; e) M. Rossi, V. Blum, P. Kupser, G. von Helden, F. Bierau, K. Pagel, G. Meijer, M. Scheffler, *J. Phys. Chem. Lett.* **2010**, *1*, 3465–3470; f) J. Makowska, A. Liwo, W. Zmudzinska, A. Lewandowska, L. Chmurzynski, H. A. Scheraga, *Biopolymers* **2011**, *97*, 240–249; g) D. Wang, B. Jaun, W. F. van Gunsteren, *ChemBioChem* **2009**, *10*, 2032–2041; h) J. Makowska, S. Rodziewicz-Motowildo, K. Baginska, J. A. Vila, A. Liwo, L. Chmurzynski, H.A. Scheraga, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1744–1749; i) V. Ramakrishnan, R. Ranbhor, S. Durani, *J. Am. Chem. Soc.* **2004**, *126*, 16332–16333; j) J. E. Shea, C. L. Brooks III, *Annu. Rev. Phys. Chem.* **2001**, *52*, 499–535; k) C. L. Brooks, III, *Acc. Chem. Res.*

2002, *35*, 447–454; l) I. H. McColl, E. W. Blanch, L. Hecht, N. R. Kallenbach, L. D. Barron, *J. Am. Chem. Soc.* **2004**, *126*, 5076–5077; m) A. Chakrabarty, R. L. Baldwin, *Adv. Protein Chem.* **1995**, *46*, 141–176; n) Z. Shi, C. A. Olson, G. D. Rose, R. L. Baldwin, N. R. Kallenbach, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9190–9195; o) Z. Shi, R. W. Woody, N. R. Kallenbach, *Adv. Protein Chem.* **2002**, *62*, 163–240.

[5] a) A. C. Engler, H. I. Lee, P. T. Hammond, *Angew. Chem.* **2009**, *121*, 9498–9502; *Angew. Chem. Int. Ed.* **2009**, *48*, 9334–9338; b) J. Y. Rao, Z. F. Luo, Z. S. Ge, H. Liu, S. Y. Liu, *Biomacromolecules* **2007**, *8*, 3871–3878; c) J. S. Guo, Y. B. Huang, X. B. Jing, X. S. Chen, *Polymer* **2009**, *50*, 2847–2855; d) D. E. Robertson, R. S. Farid, C. C. Moser, J. L. Urbauer, S. E. Mulholland, R. Pidikiti, J. D. Lear, A. J. Wand, W. F. Degrado, P. L. Dutton, *Nature* **1994**, *368*, 425–431; e) R. B. Hill, D. P. Raleigh, A. Lombardi, W. F. Degrado, *Acc. Chem. Res.* **2000**, *33*, 745–754; f) J. W. Bryson, S. F. Betz, H. S. Lu, D. J. Suich, H. X. X. Zhou, K. T. Oneil, W. F. Degrado, *Science* **1995**, *270*, 935–941; g) Z. B. Li, T. J. Deming, *Soft Matter* **2010**, *6*, 2546–2551; h) A. P. Nowak, V. Breedveld, L. Pakstis, B. Ozbas, D. J. Pine, D. Pochan, T. J. Deming, *Nature* **2002**, *417*, 424–428; i) M. S. Shim, Y. J. Kwon, *Biomaterials* **2010**, *31*, 3404–3413; j) H. Schlaad, B. Smarsly, M. Losik, *Macromolecules* **2004**, *37*, 2210–2214; k) H. A. Klok, S. Lecommandoux, *Solid-state structure, organization and properties of peptide–Synthetic hybrid block copolymers. In Peptide Hybrid Polymers*, Springer-Verlag, Berlin, **2006**, *202*, pp. 75–111; l) N. Hadjichristidis, H. Iatrou, M. Pitsikalis, G. Sakellariou, *Chem. Rev.* **2009**, *109*, 5528–5578.

[6] a) S. Lee, H. Yoshitomi, M. Morikawa, S. Ando, H. Takiguchi, T. Inoue, G. Sugihara, *Biopolymers* **1995**, *36*, 391–398; b) V. Munoz, L. Serrano, *Nat. Struct. Biol.* **1994**, *1*, 399–409; c) M. T. Ma, H. N. Hoang, C. C. G. Scully, T. G. Appleton, D. P. Fairlie, *J. Am. Chem. Soc.* **2009**, *131*, 4505–4512; d) D. K. Graff, B. PastranaRios, S. Y. Venyaminov, F. G. Prendergast, *J. Am. Chem. Soc.* **1997**, *119*, 11282–11294; e) Y. X. Chen, C. T. Mant, S. W. Farmer, R. E. W. Hancock, M. L. Vasil, R. S. Hodges, *J. Biol. Chem.* **2005**, *280*, 12316–12329; f) C. E. Schafmeister, J. Po, G. L. Verdine, *J. Am. Chem. Soc.* **2000**, *122*, 5891–5892; g) O. M. Stephens, S. Kim, B. D. Welch, M. E. Hodsdon, M. S. Kay, A. Schepartz, *J. Am. Chem. Soc.* **2005**, *127*, 13126–13127.

[7] a) H. Yin, A. D. Hamilton, *Angew. Chem.* **2005**, *117*, 4200–4235; *Angew. Chem. Int. Ed.* **2005**, *44*, 4130–4163; b) V. Azzarito, K. Long, N. S. Murphy, A. J. Wilson, *Nat. Chem.* **2013**, *5*, 161–173; c) L.-G. Milroy, T. N. Grossmann, S. Hennig, L. Brunsveld, C. Ottmann, *Chem. Rev.* **2014**, *114*, 4695–4748.

[8] a) D. Y. Jackson, D. S. King, J. Chmielewski, S. Singh, P. G. Schultz, *J. Am. Chem. Soc.* **1991**, *113*, 9391–9392; b) J. R. Kumita, O. S. Smart, G. A. Woolley, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3803–3808; c) A. M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris, A. F. Spatola, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 11273– 11278; d) N. E. Shepherd, H. N. Hoang, G. Abbenante, D. P. Fairlie, *J. Am. Chem. Soc.* **2005**, *127*, 2974–2983; e) S. Cantel, A. L. C. Isaad, M. Scrima, J. J. Levy, R. D. DiMarchi, P. Rovero, J. A. Halperin, A. M. D'Ursi, A. M. Papini, M. Chorev, *J. Org. Chem.* **2008**, *73*, 5663–5674; f) A. Muppidi, K. Doi, S. Edwardraja, E. J. Drake, A. M. Gulick, H. G. Wang, Q. Lin, *J. Am. Chem. Soc.* **2012**, *134*, 14734–14737; g) Y. H. Lau, P. de Andrade, S.-T. Quah, M. Rossmann, L. Laraia, N. Skold, T. J. Sum, P. J. E. Rowling, T. L. Joseph, C. Verma, M. Hyvonen, L. S. Itzhaki, A. R. Venkitaraman, C. J. Brown, D. P. Lane, D. R. Spring, *Chem. Sci.* **2014**, *5*, 1804–1809; h) C. M. Haney, W. S. Horne, *J. Pept. Sci.*

2014, *20*, 108–114; i) Y. Zou, A. M. Spokoyny, C. Zhang, M. D. Simon, H. Yu, Y.-S. Lin, B. L. Pentelute, *Org. Biomol. Chem.* **2014**, *12*, 566–573; j) G. J. Hilinski, Y.-W. Kim, J. Hong, P. S. Kutchukian, C. M. Crenshaw, S. S. Berkovitch, A. Chang, S. Ham, G. L. Verdine, *J. Am. Chem. Soc.* **2014**, *136*, 12314–12322; k) D. Mazzier, C. Peggion, C. Toniolo, A. Moretto, *Biopolymers* **2014**, *102*, 115–123; l) T. E. Speltz, S. W. Fanning, C. G. Mayne, C. Fowler, E. Tajkhorshid, G. L. Greene, T. W. Moore, *Angew. Chem.* **2016**, *128*, 4324–4327; *Angew. Chem. Int. Ed.* **2016**, 55, 4252–4255.

[9] a) S. Chowdhury, W. Zhang, C. Wu, G. Xiong, Y. Duan, *Biopolymers* **2003**, *68*, 63–75; b) R. Geney, M. Layten, R. Gomperts, V. Hornak, C. Simmerling, *J. Chem. Theory Comput.* **2006**, *2*, 115–127; c) J. A. Vila, D. R. Ripoll, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 14812–14816; d) A. Liwo, M. Khalili, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2362–2367; e) V. S. Pande, I. Baker, J. Chapman, S. P. Elmer, S. Khaliq, S. M. Larson, Y. M. Rhee, M. R. Shirts, C. D. Snow, E. J. Sorin, B. Zagrovic, *Biopolymers* **2003**, *68*, 91–109; f) P. Ferrara, J. Apostolakis, A. Caflisch, *Proteins: Struct., Funct., Genet.* **2002**, *39*, 252–260; g) H. Nymeyer, A. E. Garcia, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13934–13939; h) H. Wang, S.-S. Sung, *J. Am. Chem. Soc.* **2000**, *122*, 1999–2009; i) C. D. Snow, E. J. Sorin, Y. M. Rhee, V. S. Pande, *Ann. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 43–69; j) X. Cheng, G. Cui, V. Hornak, C. Simmerling, *J. Phys. Chem. B* **2005**, *109*, 8220–8230; k) E. J. Sorin, V. S. Pande, *J. Comput. Chem.* **2005**, *26*, 682–690; l) J. E. Shea, J. N. Onuchic, C. L. Brooks, III. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16064–16068; m) W. Zhang, C. Wu, Y. Duan, *J. Chem. Phys.* **2005**, *123*, 154105; n) S. Chowdhury, H. Lei, Y. Duan, *J. Phys. Chem. B* **2005**, *109*, 9073–9081; o) A. Cavalli, P. Ferrara, A. Caflisch, *Proteins: Struct., Funct., Bioinf.* **2002**, *47*, 305–314; p) C. C. Wei, M. H. Ho, W. H. Wang, Y. C. Sun, *J. Phys. Chem. B* **2005**, *109*, 19980–19986.

[10] a) E. G. Baker, G. J. Bartlett, M. P. Crump, R. B. Sessions, N. Linden, C. F. Faul, D. N. Woolfson, *Nat. Chem. Biol.* **2015**, *11*, 221–228; b) Y. Gao, Y. Li, L. Mou, B. Lin, J. Z. H. Zhang, Y. Mei, *Sci. Rep.* **2015**, *5*, 10359–10364; c) J. Huang, A. D. Mackerell Jr., *Biophys. J.* **2014**, *107*, 991–997; d) J. Huang, P. E. M. Lopes, B. Roux, A. D. MacKerell Jr., *J. Phys. Chem. Lett.* **2014**, *5*, 3144−3150; e) A. Vishnyakov, D. S. Talaga, A. V. Neimark, *J. Phys. Chem. Lett.* **2012**, *3*, 3081−3087; f) Z. Wu, Q. Cui, A. Yethiraj, *J. Phys. Chem. Lett.* **2011**, *2*, 1794–1798; g) R. F. Sommese, S. Sivaramakrishnan, R. L. Baldwin, J. A. Spudich, *Protein Sci.* **2010**, *19*, 2001–2005; h) S. Sivaramakrishnan, J. Sung, M. Ali, S. Doniach, H. Flyvbjerg, J. A. Spudich, *Biophys. J.* **2009**, *97*, 2993–2999; i) S. Sivaramakrishnan, B. J. Spink, A. Y. L. Sim, S. Doniach, J. A. Spudich, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13356–13361; j) D. Süveges, Z. Gáspári, G. Tóth, L. Nyitray, *Proteins: Struct., Funct., Bioinf.* **2009**, *74*, 905–916.

[11] H. Meuzelaar, M. Tros, A. Huerta-Viga, C. N. van Dijk, J. Vreede, S. Woutersen, *J. Phys. Chem. Lett.* **2014**, *5*, 900−904.

[12] K. D. Walker, T. P. Causgrove, *J. Mol. Model.* **2009**, *15*, 1213–1219.

[13] a) B. Goyal, A. Kumar, K. R. Srivastava, S. Durani, *J. Biomol. Struct. Dyn.* **2016**, *6*, 1–13; b) B. Goyal, A. Kumar, K. R. Srivastava, S. Durani, *RSC Adv.* **2016**, *6*, 74162–74176.

[14] R. M. Culik, S. Annavarapu, V. Nanda, F. Gai, *Chem. Phys.* **2013**, *422*, 131–134.

[15] a) B. Anil, B. B. Song, Y. F. Tang, D. P. Raleigh, *J. Am. Chem. Soc.* **2004**, *126*, 13194–13195; b) D. V. Williams, B. Barua, N. H. Andersen, *Org. Biomol. Chem.* **2008**, *6*, 4287–4289; c) A. Rodriguez-Granillo, S. Annavarapu, L. Zhang, R. L. Koder, V. Nanda, *J. Am. Chem. Soc.* **2011**, *133*, 18750–18759.

[16] F. I. Valiyaveetil, M. Sekedat, R. MacKinnon, T. W. Muir, *Proc. Nat. Acad. Sci. USA* **2004**, *101*, 17045–17049.

[17] a) M. D. Struthers, R. P. Cheng, B. Imperiali, *Science* **1996**, *271*, 342–345; b) S. Durani, *Acc. Chem. Res.* **2008**, *41*, 1301–1308; c) K. Patel, B. Goyal, A. Kumar, N. Kishore, S. Durani, *J. Phys. Chem. B* **2010**, *114*, 16887–16893; d) B. Goyal, K. Patel, K. R. Srivastava, S. Durani, *RSC Adv.* **2015**, *5*, 105400–105408; e) B. Goyal, K. R. Srivastava, K. Patel, S. Durani, *ChemistrySelect* **2016**, *1*, 2050–2057.

[18] a) Y. Demizu, T. Misawa, T. Nagakubo, Y. Kanda, K. Okuhira, Y. Sekino, M. Naito, M. Kurihara, *Bioorg. Med. Chem.* **2015**, *23*, 4132–4138; b) T. T. Dinh, D. H. Kim, H. X. Luong, B. J. Lee, Y. W. Kim, *Bioorg. Med. Chem. Lett.* **2015**, *25*, 4016–4019; c) K. Estieu-Gionnet, G. Guichard, *Expert Opin. Drug Discov.* **2011**, *6*, 937–963; d) Y. Huang, Q. Feng, Q. Yan, X. Hao, Y. Chen, *Mini Rev. Med. Chem.* **2015**, *15*, 73–81; e) J. S. Khara, F. K. Lim, Y. Wang, X. Y. Ke, Z. X. Voo, Y. Y. Yang, R. Lakshminarayanan, P. L. Ee, *Acta Biomater.* **2015**, *28*, 99–108; f) N. Mehrban, B. Zhu, F. Tamagnini, F. I. Young, A. Wasmuth, K. L. Hudson, A. R. Thomson, M. A. Birchall, A. D. Randall, B. Song, D. N. Woolfson, *ACS Biomater. Sci. Eng.* **2015**, *1*, 431–439; g) S. E. Miller, P. F. Thomson, P. S. Arora, *Curr. Protoc. Chem. Biol.* **2014**, *6*, 101–116; h) M. Pelay-Gimeno, A. Glas, O. Koch, T. N. Grossmann, *Angew. Chem.* **2015**, *127*, 9022–9054; *Angew. Chem. Int. Ed.* **2015**, *54*, 8896–8927.

[19] Y. V. Venkatachalapati, P. Balaram, *Nature* **1979**, *281*, 83–84.

[20] a) D. S. Kemp, J. G. Boyd, C. C. Muendel, *Nature* **1991**, *352*, 451–454; b) D. S. Kemp, T. P. Curran, J. G. Boyd, T. J. Allen, *J. Org. Chem.* **1991**, *56*, 6683–6697; c) D. S. Kemp, T. P. Curran, W. M. Davis, J. G. Boyd, C. Muendel, *J. Org. Chem.* **1991**, *56*, 6672–6682; d) G. E. Job, B. Heitmann, R. J. Kennedy, S. M. Walker, D. S. Kemp, *Angew. Chem.* **2004**, *116*, 5767–5769; *Angew. Chem. Int. Ed.* **2004**, *43*, 5649–5651; e) B. Heitmann, G. E. Job, R. J. Kennedy, S. M. Walker, D. S. Kemp, *J. Am. Chem. Soc.* **2005**, *127*, 1690–1704.

[21] a) A. Kumar, S. Srivastava, S. Tripathi, S. K. Singh, S. Srikrishna, A. Sharma, *J. Biomol. Struct. Dyn.* **2016**, *34*, 1252–1263; b) L. Baweja, K. Balamurugan, V. Subramanian, A. Dhawan, *J. Mol. Graph. Model.* **2015**, *61*, 175–185; c) V. Minicozzi, R. Chiaraluce, V. Consalvi, C. Giordano, C. Narcisi, P. Punzi, G. C. Rossi, S. Morante, *J. Biol. Chem.* **2014**, *289*, 11242–11252; d) I. Autiero, E. Langella, M. Saviano, *Mol. BioSys.* **2013**, *9*, 2835–2841; e) S. T. Ngo, M. S. Li, *J. Phys. Chem. B* **2012**, *116*, 10165−10175; f) M. H. Viet, S. T. Ngo, N. S. Lam, M. S. Li, *J. Phys. Chem. B* **2011**, *115*, 7433–7446; g) C. Yang, J. Li, Y. Li, X. Zhu, *J. Mol. Struct. THEOCHEM* **2009**, *895*, 1–8; h) Y. Sun, Z. Qian, G. Wei, *Phys. Chem. Chem. Phys*. **2016**, *18*, 12582–12591.

[22] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, *J. Comput. Chem.* **2003**, *24*, 1999–2012.

[23] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.

[24] C.-F. Fu, S. X. Tian, *J. Chem. Theory Comput.* **2011**, *7*, 2240–2252.

[25] S. R. Gerben, J. A. Lemkul, A. M. Brown, D. R. Bevan, *J. Biomol. Struct. Dyn.* **2014**, *32*, 1817–1832.

[26] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *Angew. Chem.* **1999**, *111*, 249–253; *Angew. Chem. Int. Ed.* **1999**, *38*, 236–240.

[27] C. Dalgicdir, C. Globisch, C. Peter, M. Sayar, *PLoS Comput. Biol.* **2015**, *11*, e1004328.

[28] M. Oba, Y. Demizu, N. Yamagata, Y. Sato, M. Doi, M. Tanaka, H. Suemune, H. Okuda, M. Kurihara, *Tetrahedron* **2010**, *66*, 2293–2296.

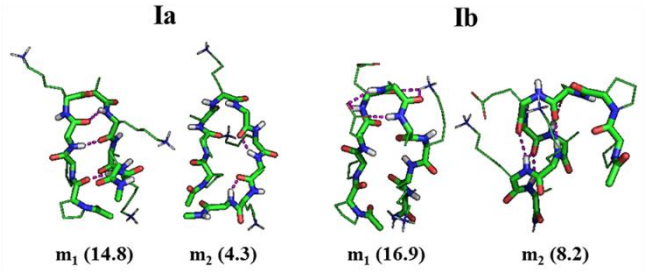
[29] J. A. Vila, D. R. Ripoll, H. A. Scheraga, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13075–13079.

[30] J. Dzubiella, *J. Am. Chem. Soc.* **2008**, *130*, 14000–14007.

[31] a) R. T. Ingwall, H. A. Scheraga, N. Lotan, A. Berger, E. Katchalski, *Biopolymers* **1968**, *6*, 331–368; b) K. E. B. Platzer, R. H. A. V. S. Anathanarayanan, H. A. Scheraga, *Macromolecules* **1972**, *5*, 177–187.

[32] J. Graf, P. H. Nguyen, G. Stock, H. Schwalbe, *J. Am. Chem. Soc.* **2007**, *129*, 1179−1189.

Table of contents graphic



N-terminal diproline of homochiral structure, ^{L}Pro–^{L}Pro, and large side chain substituents of lysine and glutamic acid residues in the *i*, *i* + 4 arrangement stabilize the helical conformation in alanine-based nonapeptide, Ac–Pro_1–Pro_2–Glu_3–Glu_4–Ala_5–Ala_6–Lys_7–Lys_8–Ala_9–NH_2 (**Ib**), as observed in the second most-populated microstate, $\mathbf{m_2}$, during molecular dynamics in explicit-water.