Existence of specific "Folds" in Polyproline-II Ensembles of an "Unfolded" Alanine Peptide Detected by Molecular Dynamics

Vibin Ramakrishnan[#], Ranjit Ranbhor[#] and Susheel Durani*

Department of Chemistry* and School of Biosciences and Bioengineering[#],
IIT Bombay, Mumbai-400076, India.

Supporting Information

Materials and Methods

Molecular Dynamics simulation.

The simulations were performed on a 16 node cluster with Intel Pentium 4, 1.6 GHz processors in Linux 7.3 operating system. Molecular dynamics (MD) simulations were performed using GROMACS implemented GROMOS96 43A1 forcefield. The dynamics of alanine dipeptide (Ac-Ala-NHMe) and end protected octa-alanine (Ac-Ala_8-NHMe) were studied at 298, 323 and 373K under NVT conditions (canonical ensemble) with periodic boundary. Alanine dipeptide and octa-alanine in α or β conformation (ϕ = -57, ψ = -47 for α conformation; and ϕ = -120, ψ = +120 for β conformation), were energy minimized using steepest descent algorithm in 2000 steps. Solvent molecules (SPC water) were then added around the peptide to fill the cubic box of edge length 4.5 nm. The density of water was set to the values 0.998, 0.985 and 0.958 g/ml, respectively, as required for the temperatures 298, 323 and 373K.[1] The peptide was once again energy minimized in presence of solvent in 2000 steps, then position restrained and submitted to 200 ps MD for solvent relaxation. The production run for octa-alanine was fixed as 8 and 33 ns at 298K, 8 ns at 323K and 8 ns at 373K for α and β starting structures. For alanine dipeptide, the simulation length was set as 28 ns at 298K. The first 3 ns time period was exempted from each MD trajectory as an equilibration period. The non bonded list cut-off of 1.4 nm was used. The nonbonded interactions (VdW and electrostatics) were shifted to zero between 0.8 nm to 1.4 nm cutoff. An integration step of 2 fs was used.

Initial velocities were taken from a Maxwellian distribution at the chosen temperature. The temperature was maintained to the reference value by a weak coupling to an external temperature bath with a coupling relaxation time of 0.1 ps. Throughout the simulations, the bond lengths were constrained to ideal values with the shift procedure with a geometric accuracy of 10^{-4}.

Cluster analysis.

The clustering analysis was performed in GROMACS using g_cluster program according to the procedure of Daura et. al.[2] Briefly, the analysis were performed on PDBs extracted from MD trajectories at 5 ps intervals. The 40 ns MD trajectory of octa-alanine (Ac-Ala_8-NHMe) is populated with a total of 8000 members. The clustering was performed in Cartesian space on these 8000 peptide structures stored in PDB files. For each pair of structures, a least squares translational and rotational fit was performed using the backbone (N, C_β, C_α, C) atoms, and the atom positional root mean square deviation (RMSD) for this set of atoms was calculated. Using the criterion of similarity of two structures as an RMSD $\leq$ 0.15 nm for the backbone atoms, the number of neighbors of each clusters in the 6000 member pool was determined. The structure with the highest number of neighbors was then taken as the central member of the first cluster of structures. All the structures belonging to this cluster were thereafter removed from the pool. For each of the remaining structures the number of neighbors was again computed and the process was iterated until all structures were assigned to a cluster.

Free Energy calculation

Free energies derived from relative populations of individual clusters (conformational microstates at an RMSD cut-off of 0.15nm) were calculated according to Daura et. al by the following equation.[3]

$$\Delta G_{A\rightarrow B} = -k_B T \ln (P_B/P_A)$$

Where k_B is the Boltzmann constant, T the temperature, and P_A, P_B the relative probabilities of finding the system in states A and B. A and B in the system studied refers to the clusters, and P_A and P_B are taken as the number of structures in clusters A and B.

References

1. Kell, G. S *Journal of Chemical Engineering Data* **1975**, 20, 97-105.
2. Daura, X.; Gademann, K.; Jann, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem. Intl. Ed.* **1999**a, 38, 236-240.
3. Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins* **1999**, 34, 269-280.

α_R	ϕ = -20 to -100 Ψ = -20 to -80	**PPII**	ϕ = -30 to -90 Ψ = 80 to 170
β	ϕ = -90 to -150 Ψ = 80 to 180	**α_L**	ϕ = 40 to 80 Ψ = 10 to 100

Table S1. The definition of six primary ϕ, ψ basins in Ramachandran diagram adopted in this study.

Temp (K)	**Initial Conf**	**α_R %**	**β %**	**PPII %**	**α_L %**
298	**α**	8.4	24.0	25.6	7.8
	β	13.2	23.2	31.8	7.3
323	**α**	27.0	19.5	29.1	3.8
	β	12.6	16.9	27.6	20.1
373	**α**	17.1	24.2	27.6	7.3
	β	17.6	23.4	28.2	8.8

Table S2. Percentage population of specific ϕ, ψ basins, defined in **Table S1,** over last 5 ns of 8 ns MD trajectories of octa-alanine (Ac-Ala_8-NHMe) in water, initiated from the specified conformation (**α** or **β)** at the specified temperature.

Cluster Number	Cluster Size (Number of Members in each cluster)	Percentage Population of each cluster	ΔG (kJ/mol) (with respect to cluster1)
1	1422	17.8	0
2	491	6.1	2.63
3	247	3.1	4.33
4	151	1.9	5.55
5	149	1.9	5.59
6	130	1.6	5.92
7	120	1.5	6.12
8	106	1.3	6.43
9	105	1.3	6.45
10	100	1.3	6.57
11	90	1.1	6.83
12	88	1.1	6.89
13	83	1.0	7.03
14	82	1.0	7.06

Table S3: Conformational clusters at 0.15 nm RMSD cut-off (conformational microstates) populating 298K equilibrium (40ns) ensemble of octa-alanine in water. There are a total of 374 clusters of which those with at least 80 members, out of the total 8000 structures contained in the 40 ns MD trajectory, are listed in the table along with their populations on % basis. The free energy difference (**Δ**G) was calculated from the relative populations of the clusters, as described in **Materials and Methods**.

	Cluster 1		Cluster 2		Cluster 3	
Res No	**ϕ**	**ψ**	**ϕ**	**ψ**	**ϕ**	**Ψ**
Ala_1	-96	154	-49	125	-49	168
Ala_2	-109	-47	-125	-70	-62	154
Ala_3	-118	101	-85	84	25	-77
Ala_4	35	66	48	138	-74	-25
Ala_5	67	7	59	-92	-75	-57
Ala_6	-121	142	-122	153	-59	-29
Ala_7	57	131	59	156	-127	142
Ala_8	-48	128	-56	-86	-117	-60

	Cluster 4		Cluster 5		Cluster 6	
Res No	**ϕ**	**ψ**	**ϕ**	**ψ**	**ϕ**	**Ψ**
Ala_1	-62	118	-116	136	-58	148
Ala_2	-47	136	-75	140	-107	171
Ala_3	-60	144	-152	118	-52	-51
Ala_4	-118	99	-61	-47	-70	113
Ala_5	-64	149	-134	133	-74	150
Ala_6	-61	133	85	-98	39	133
Ala_7	76	111	-70	147	64	171
Ala_8	-65	143	-96	137	-56	153

	Cluster 7		Cluster 8		Cluster 9	
Res No	**ϕ**	**ψ**	**ϕ**	**ψ**	**ϕ**	**Ψ**
Ala_1	-48	148	-135.0	142	-80	132
Ala_2	-74	151	-152	94	-92	127
Ala_3	-147	-93	-164	126	-130	118
Ala_4	-113	137	-55	-38	72	120
Ala_5	35	96	-72	-36	-53	-55
Ala_6	68	137	-70	-35	-63	154
Ala_7	-7	109	-82	130	-76	154
Ala_8	61	108	-76	-25	55	100

Table S4. Continued on next page

	Cluster 10		Cluster 11		Cluster 12	
Res No	**ϕ**	**ψ**	**ϕ**	**ψ**	**ϕ**	**Ψ**
Ala_1	-143	166	-136	105	-74	94
Ala_2	-52	136	-63	152	-107	99
Ala_3	46	85	-90	114	-63	-19
Ala_4	70	117	75	-36	-122	155
Ala_5	29	-174	-68	126	-47	139
Ala_6	-106	153	57	105	-13	76
Ala_7	-88	119	-47	147	-114	128
Ala_8	-64	134	-52	163	-78	107

	Cluster 13		Cluster 14	
Res No	**ϕ**	**ψ**	**ϕ**	**Ψ**
Ala_1	-96	88	-99	149
Ala_2	66	106	73	167
Ala_3	-68	156	-19	-76
Ala_4	-53	113	-64	-29
Ala_5	74	84	-90	-39
Ala_6	-65	122	-62	-39
Ala_7	-118	152	-127	147
Ala_8	-58	-49	-64	94

Table S4. The residue-by-residue ϕ, ψ values of the central members of the fourteen most populous conformational microstates (corresponding to the clusters 1 to 14 of **Table S3**) in 298K equilibrium (40ns) ensemble of octa-alanine (Ac-Ala_8-NHMe) in water.

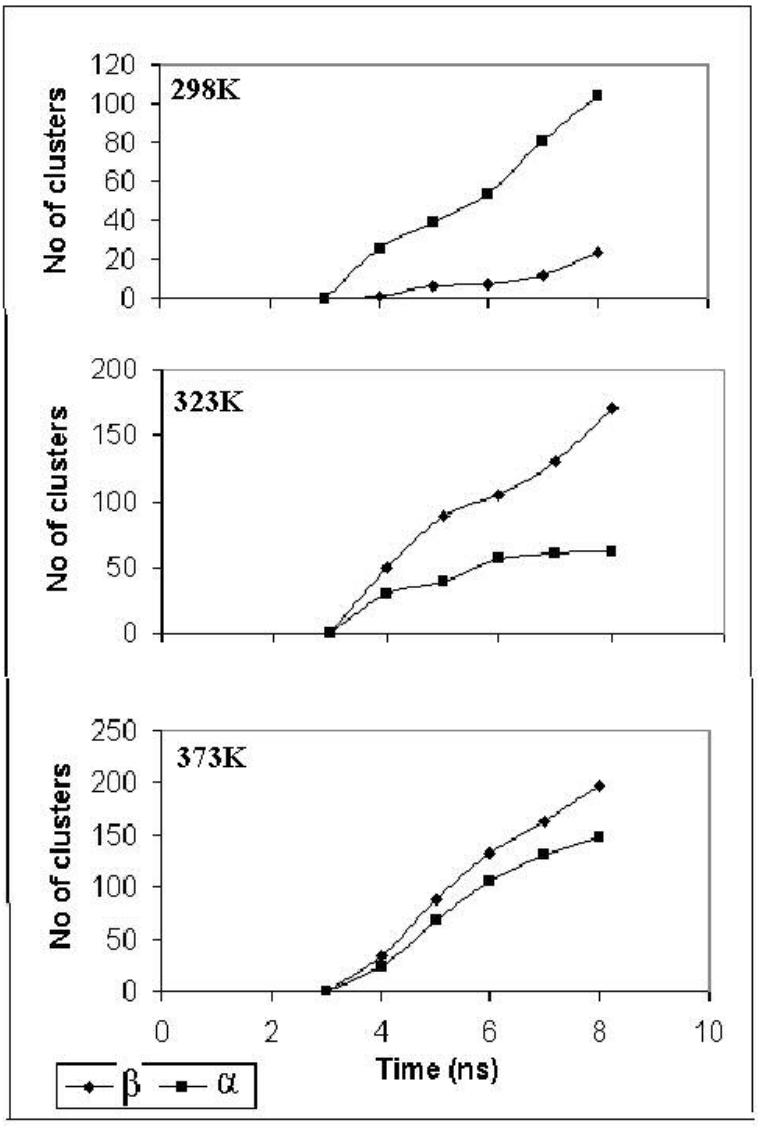


Figure S1: Progress curves for the accumulation of conformational microstates (clusters at 0.15 nm RMSD cut-off) during 3-8 ns of MD trajectories of octa-alanine in water, initiated in the specified starting conformation (**β** or **α**) at the indicated temperature.

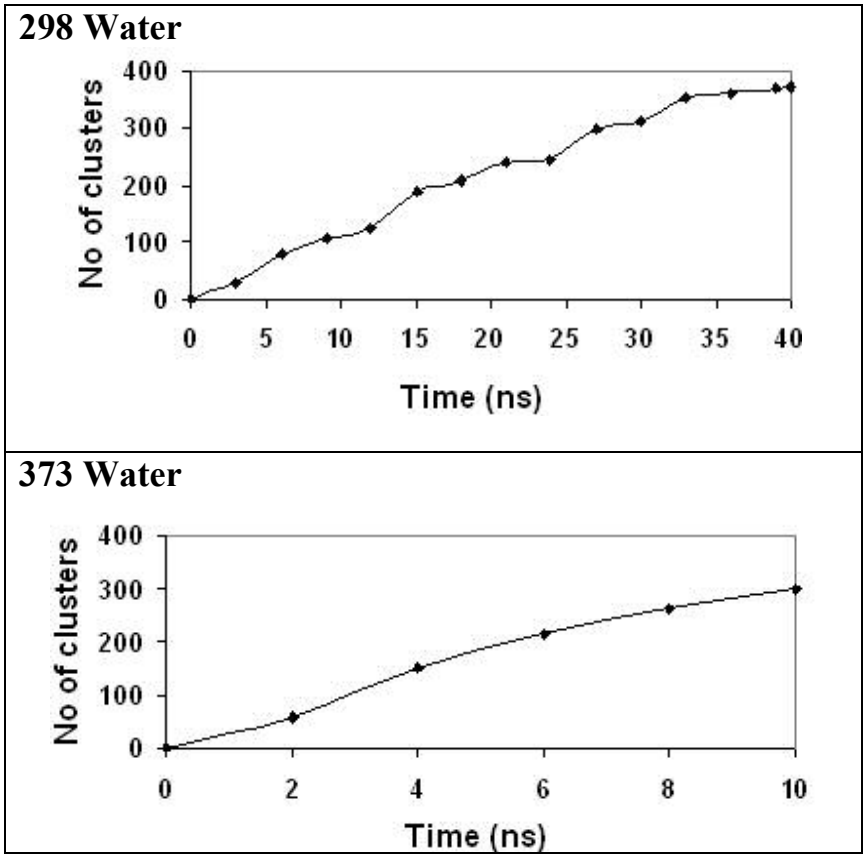
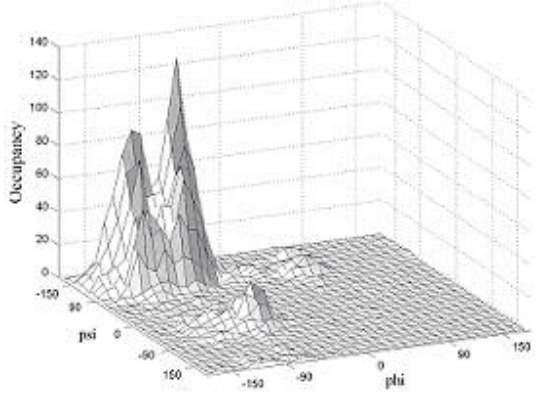
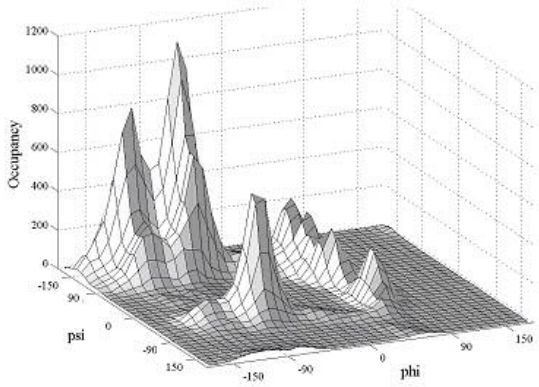


Figure S2. Progress curves for accumulation of microstates (conformational clusters at 0.15 nm RMSD cutoff) in 298 K (40 ns) and 373 K (10 ns) MD trajectories of octa-alanine (Ac-Ala_8-NHMe) in water.

a) Alanine Dipeptide (298K)



b) Octa-alanine (298K)



c) Octa-alanine (373K)

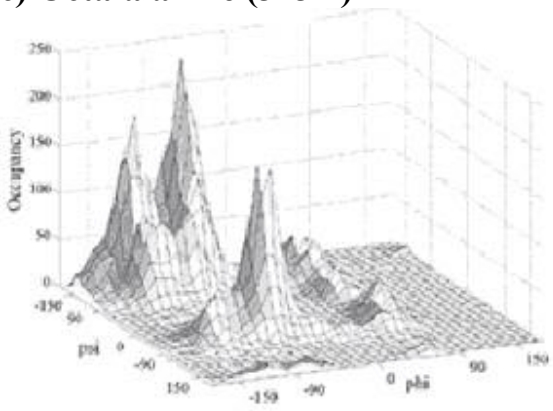


Figure S3. Occupancies of ϕ, ψ basins in equilibrium ensembles of alanine model peptides prepared by MD in water. **a)** Ac-Ala_1-NHMe at 298K (25ns MD), **b)** Ac-Ala_8-NHMe at 298K (40ns MD) and **c)** Ac-Ala_8-NHMe at 373K (10ns MD).

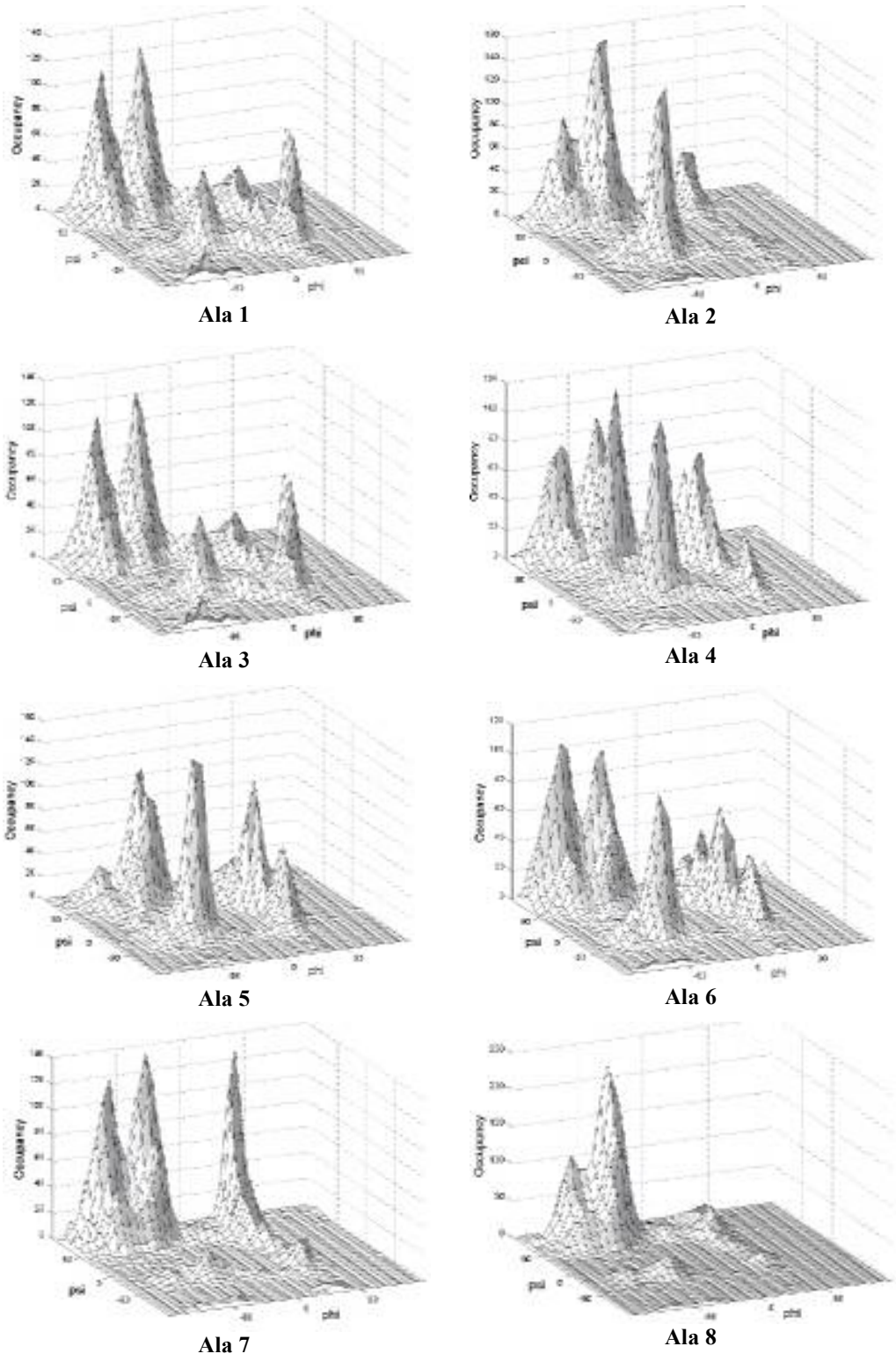


Figure S4. Residue-by-residue occupancies of ϕ, ψ basins in the 298 K (40ns MD) equilibrium ensemble of octa-alanine (Ac-Ala_8-NHMe) in water.

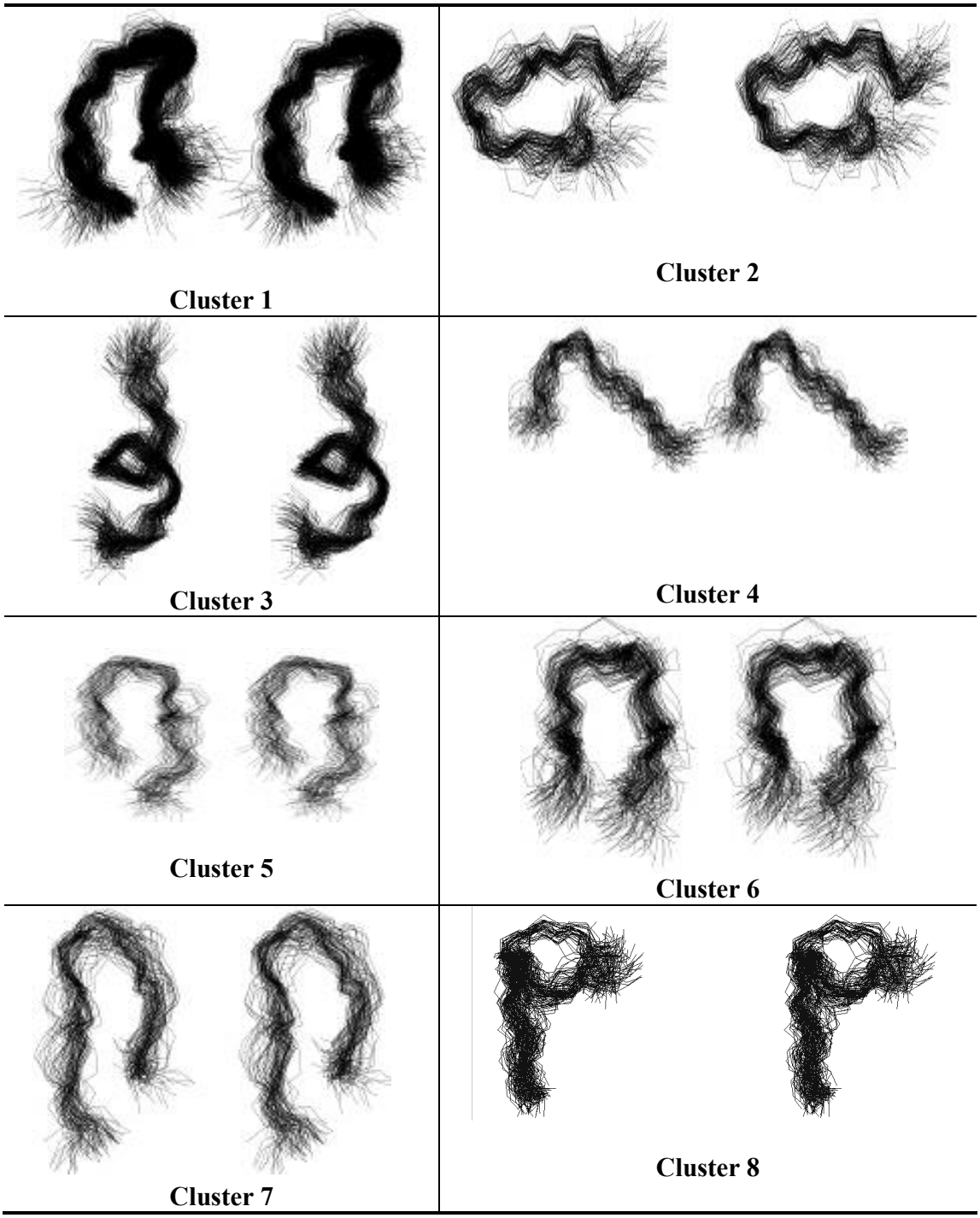


Figure S5. Continued on next page

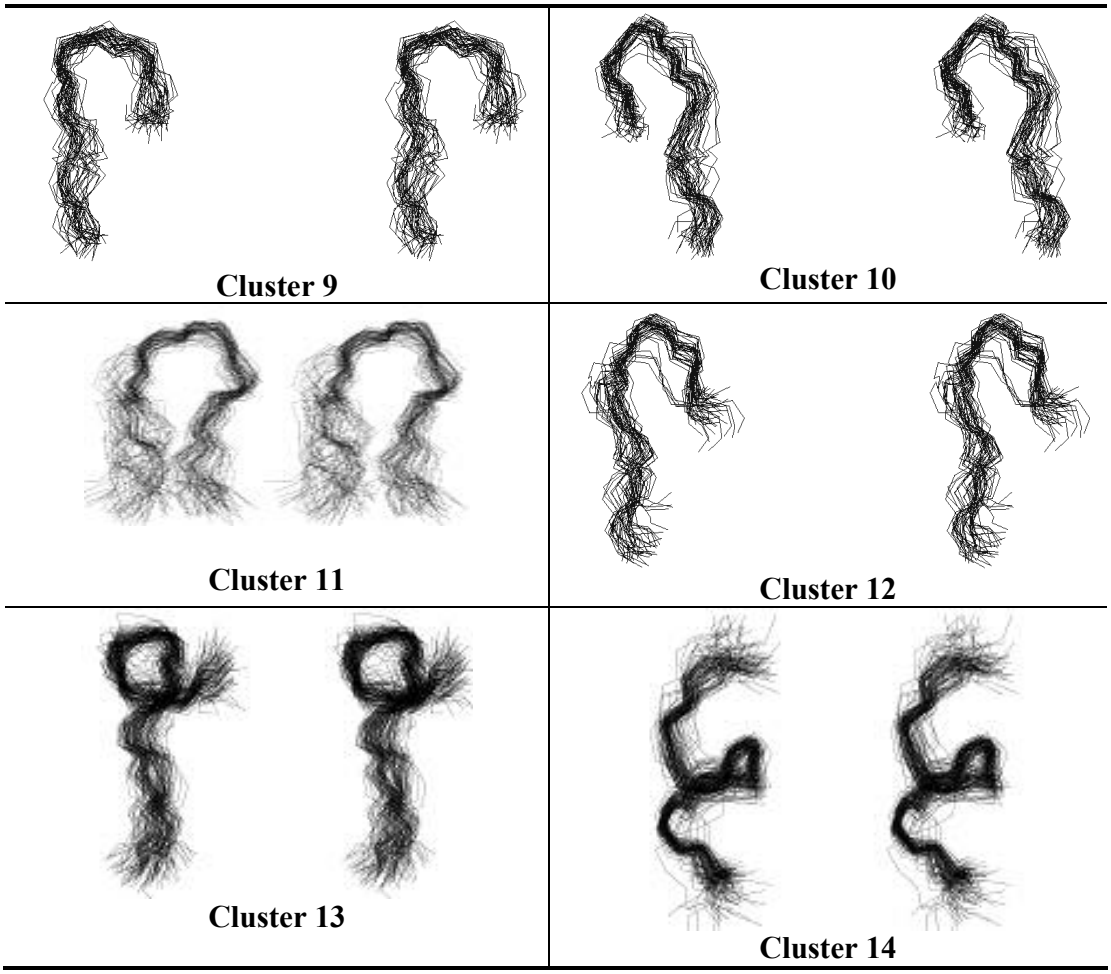


Figure S5. Stereo images of the fourteen most populous conformational microstates (clusters 1-14 of **Table S3)** each comprising more than 1% of the 298K equilibrium (40ns) ensemble of octa-alanine (Ac-Ala_8-NHMe) in water. The clusters add up to overall 43% of the 298K equilibrium (40 ns) ensemble. The figure was prepared with MOLMOL program. Randomly chosen one-third of the members are shown in the case of the clusters 1, 2 and 3 and half of the members are shown in the case of the clusters 4 to 8. All the members contained in the clusters 9 – 14 are shown in the figure.

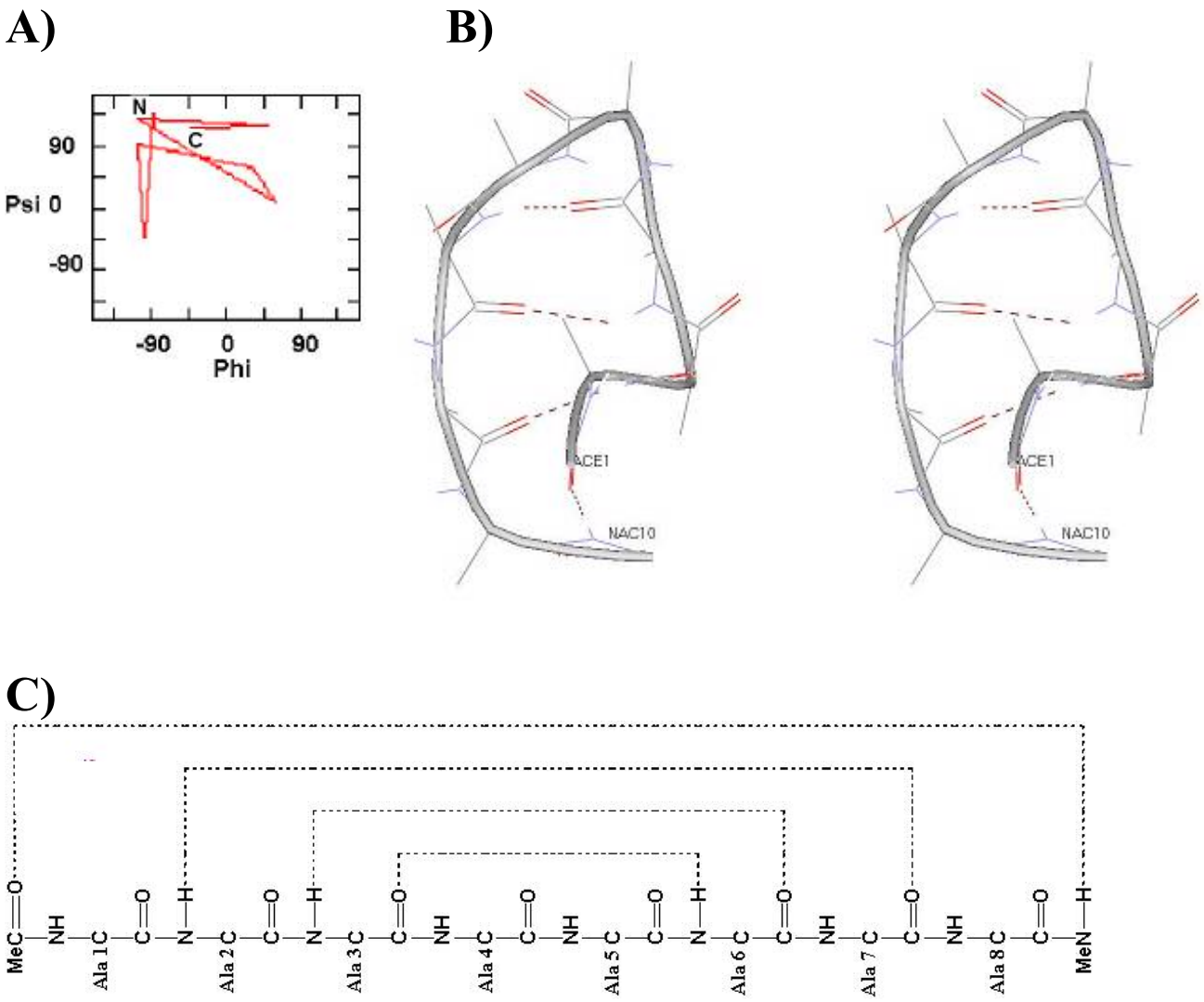


Figure S6 A. Ramachandran coordinates of the central member of the Cluster 1 comprising 18 % of the 298K equilibrium (40 ns) ensemble of octa-alanine in water (A). Stereo view of the central member showing its hydrogen bond registry in the dotted lines (B). The peptide describes a 2:2 β hairpin of standard right handed twist with a type I' turn centered on Ala_4 and Ala_5. The hydrogen bond between terminal NHMe and CH_3CO groups, out of register with the canonical hairpin-ladder, is responsible for the wrap around of the carboxy end, and therefore for the unusual placement of Ala_2 in α_R and Ala_7 in α_L basins.

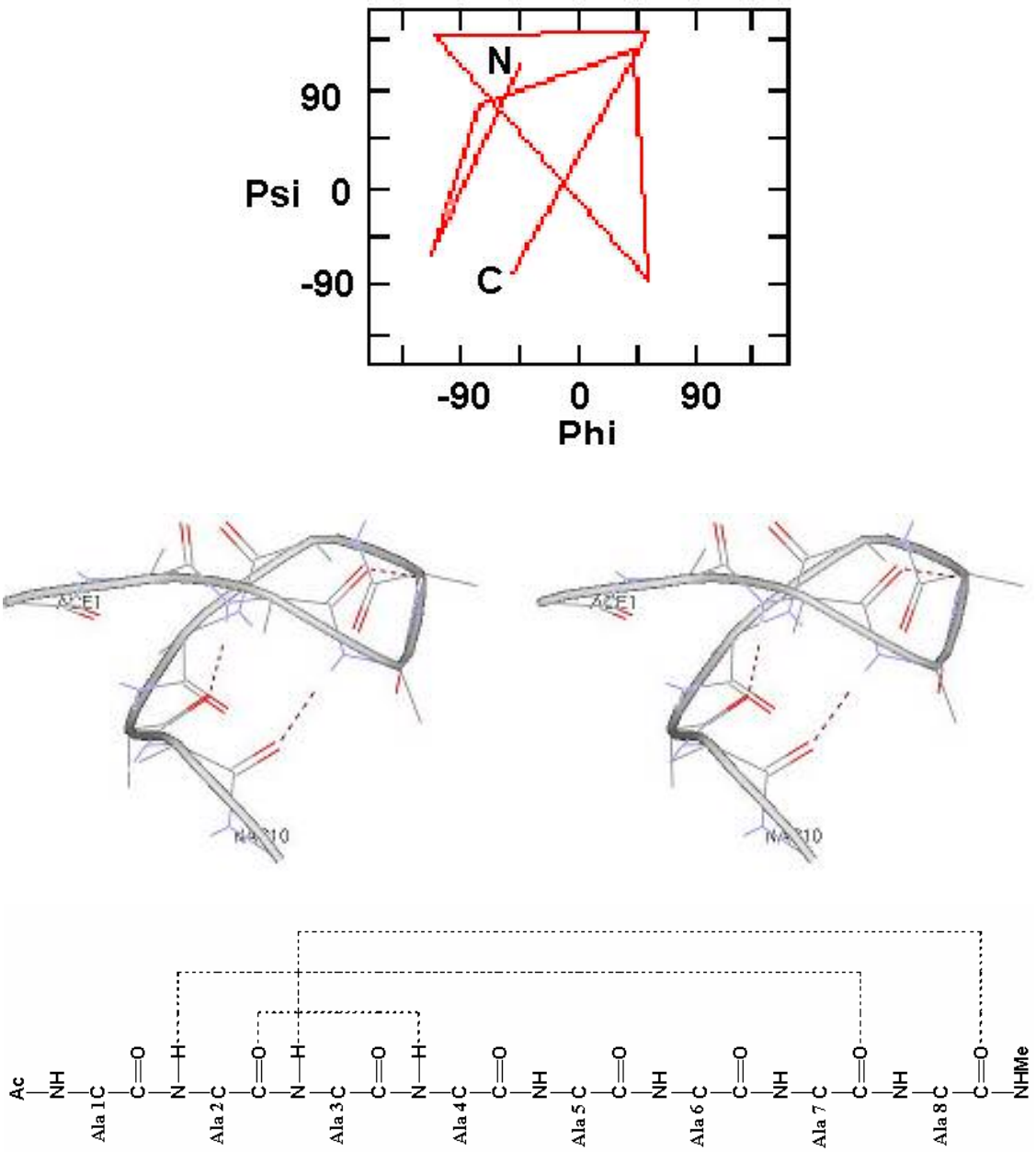


Figure S6 B. Ramachandran coordinates, stereo images and hydrogen bond registry of the central member of Cluster 2.

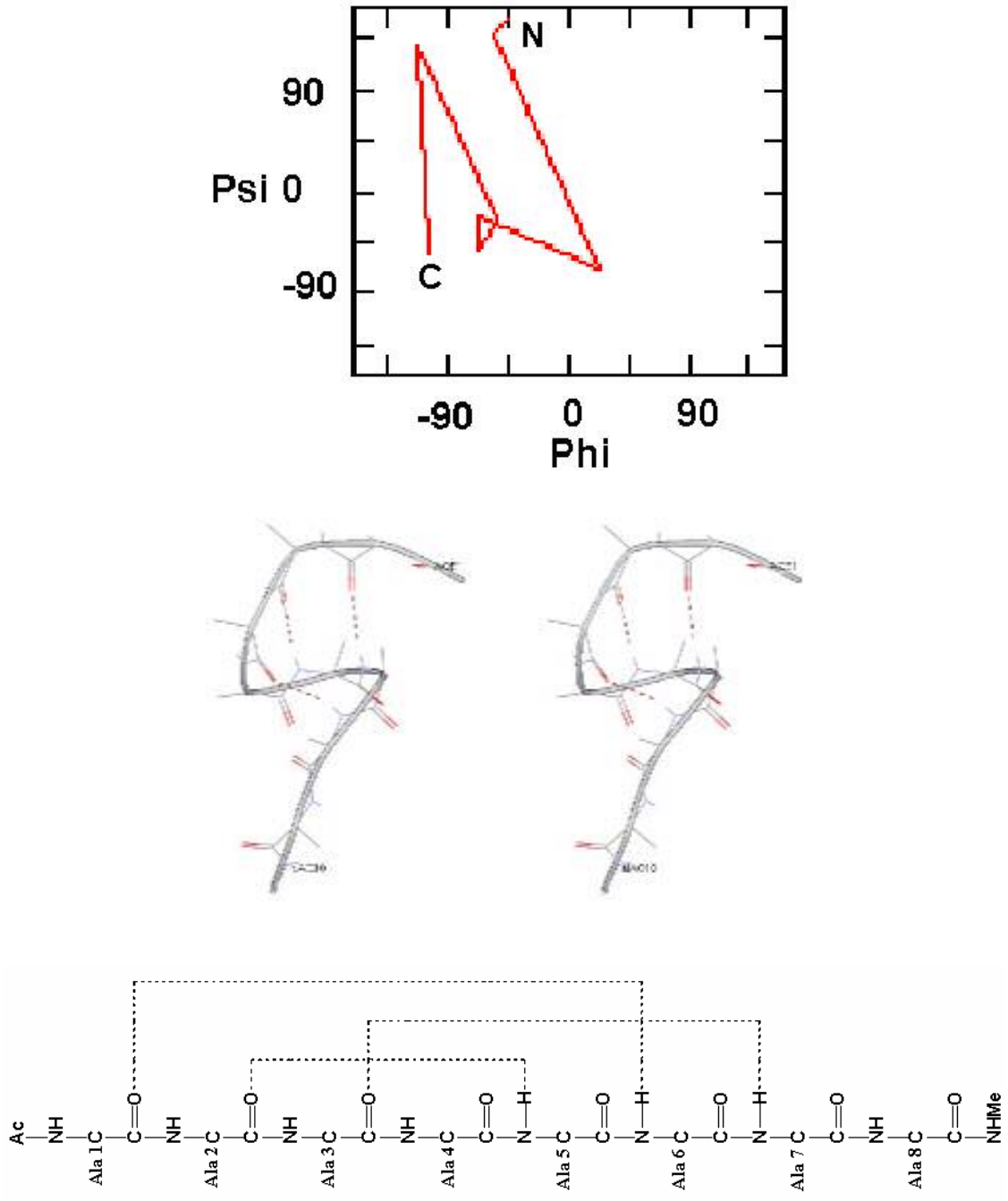


Figure S6 C. Ramachandran coordinates, stereo images and hydrogen bond registry of Cluster 3, a consecutive type II' turn-α-helical turn.