Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information for

Toward Accurately Modeling *N*-Methylated Cyclic Peptides

Diana P. Slough, Hongtao Yu, Sean M. McHugh and Yu-Shan Lin*

Department of Chemistry, Tufts University, Medford, Massachusetts, 02155, United States

***N*-methylated alanine parameters:** Because the RSFF1 force field was reparameterized based on the OPLS-AA/L force field, the parameters for *N*-methylated alanine were determined based on chemical similarity for RSFF1, following the OPLS philosophy (**Figure S1A, left**). The dihedral parameters involving the backbone *N*-methyl group of an *N*-methylated alanine are listed in **Table S1** for RSFF1; the three missing dihedral parameters were taken from OPLS/2005 and are marked with an * in **Table S1**.[1] For comparison, the dihedral parameters involving the backbone NH group of an alanine are listed in **Table S2**. The corresponding atom types for the dihedrals for alanine and *N*-methylated alanine in RSFF1 are given in **Figure S2**. On the other hand, the RSFF2 force field was reparameterized based on the Amber99sb force field. Following the philosophy of the Amber99 force fields, the partial charges for *N*-methylated alanine were determined using the multi-conformational RESP method[2] for RSFF2 simulations. Three different initial configurations of an *N*-methylated alanine dipeptide were generated with (ϕ, ψ) at (–160°, 160°), (–80°, 80°) and (80°, –60°) (the C5, C7eq and C7ax conformations) using Chimera,[3] with the *N*-methylated amide bond in the *trans* configuration. To better describe $\underline{\text{Ala}}_5$ in cyclo-(a$\underline{\text{A}}$AA$\underline{\text{A}}$A), which formed a *cis* peptide bond with Ala_4, an additional configuration with the *N*-methylated amide bond in the *cis* configuration and (ϕ, ψ) at (–132°, 49°), the dihedral angles of $\underline{\text{Ala}}_5$ in cyclo-(a$\underline{\text{A}}$AA$\underline{\text{A}}$A), was included as well. After energy minimization in Chimera, only two structures were obtained for the *trans* isomers with (ϕ, ψ) = (–137°, 62°) and (48°, 47°); the minimized (ϕ, ψ) for the *cis* isomer were (–136°, 56°) (**Figure S1B**). These three minimized structures were then optimized at the HF/6-31G* level using the Gaussian09 package.[4] The RESP charges were then fitted using the R.E.D. tools.[5] Atom types and charges for *N*-methylated alanine in the RSFF1 and RSFF2 force fields are shown in **Figure S1A**. Additionally, the dihedral parameters involving the backbone *N*-methyl group of an *N*-methylated alanine are listed in **Table S3** while those involving the backbone NH group of an alanine are listed in **Table S4** for RSFF2. The corresponding atom types for the dihedrals for alanine and *N*-methylated alanine in RSFF2 are given in **Figure S3.**

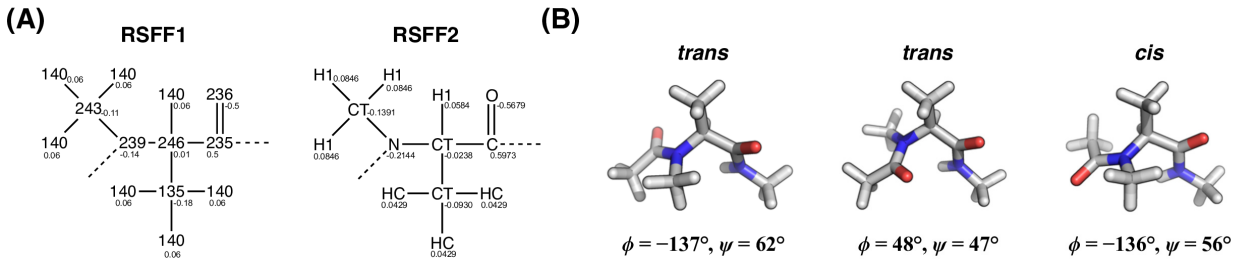


Figure S1. (**A**) Atom types and charges for an *N*-methylated alanine residue in RSFF1 and RSFF2 simulations. (**B**) Initial structures used for the G09 calculations to generate the RESP charges.

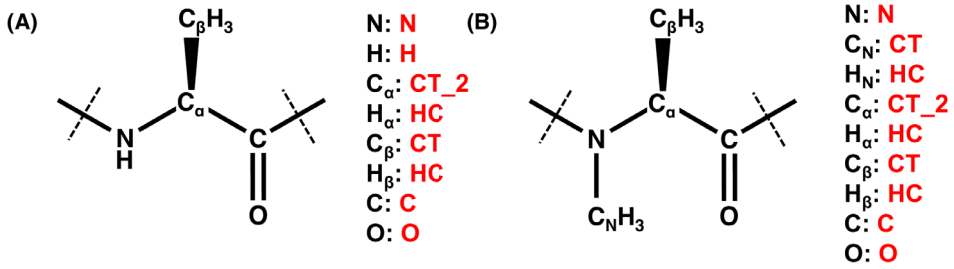


Figure S2. Atom names for the dihedrals involved in the backbone N group (NH or NCH_3) for (**A**) alanine and (**B**) *N*-methylated alanine. Atoms names in red correspond to **Tables S1 and S2**, respectively, for RSFF1.

Table S1. The dihedral parameters involving the backbone N and C of the methyl group of an *N*-methylated alanine in the RSFF1 simulations. Dihedrals with an * are taken from OPLS/2005.

Dihedral	Dihedral parameters in GROMACS, Ryckaert-Bellemans form (kJ/mol)					
	C_0	C_1	C_2	C_3	C_4	C_5
CT-N-C-CT_2	30.28798	–4.81160	–25.47638	0.0	0.0	0.0
CT-N-C-O	25.47638	0.0	–25.47638	0.0	0.0	0.0
CT-N-CT_2-CT*	9.71734	–7.44961	9.94118	–0.94558	–11.26333	0.0
CT-N-CT_2-C*	1.49369	0.0	–1.49369	0.0	0.0	0.0
CT-N-CT_2-HC*	0.0	0.0	0.0	0.0	0.0	0.0
C-N-CT-HC	–0.29079	–0.87237	0.0	1.16315	0.0	0.0
CT_2-N-CT-HC	0.0	0.0	0.0	0.0	0.0	0.0

Table S2. The dihedral parameters involving the backbone NH of an alanine in the RSFF1 simulations.

Dihedral	Dihedral parameters in GROMACS, Ryckaert-Bellemans form (kJ/mol)					
	C_0	C_1	C_2	C_3	C_4	C_5
H-N-C-CT_2	20.50160	0.0	–20.50160	0.0	0.0	0.0
H-N-C-O	20.50160	0.0	–20.50160	0.0	0.0	0.0
H-N-CT_2-CT	0.0	0.0	0.0	0.0	0.0	0.0
H-N-CT_2-C	0.0	0.0	0.0	0.0	0.0	0.0
H-N-CT_2-HC	0.0	0.0	0.0	0.0	0.0	0.0

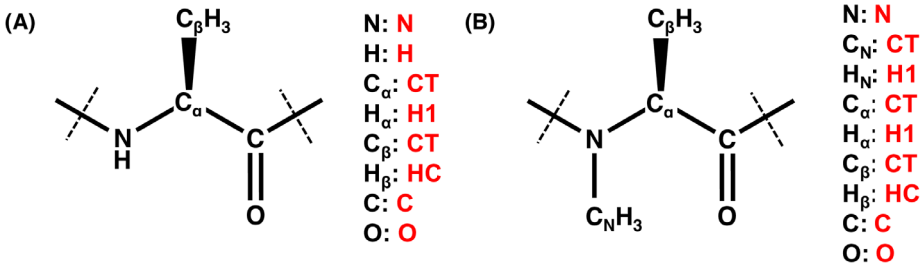


Figure S3. Atom names for the dihedrals involved in the backbone N group (NH or NCH_3) for (**A**) alanine and (**B**) *N*-methylated alanine. Atoms names in red correspond to **Tables S3 and S4**, respectively, for RSFF2.

Table S3. The dihedral parameters involving the backbone N and C of the methyl group of an *N*-methylated alanine in the RSFF2 simulations.

Dihedral atoms	Dihedral type	Dihedral parameters in GROMACS, periodic type		
		ϕ (°)	K (kJ/mol)	Multiplicity
CT-N-C-CT CT-N-C-O	X-N-C-X	180.0	10.460	2
C-N-CT-H1 CT-N-CT-H1 CT-N-CT-CT CT-N-CT-C	X-N-CT-X	0.0	0.000	0

Table S4. The dihedral parameters involving the backbone NH of an alanine in the RSFF2 simulations.

Dihedral atoms	Dihedral type	Dihedral parameters in GROMACS, periodic type		
		ϕ (°)	*K* (kJ/mol)	Multiplicity
H-N-C-CT	X-N-C-X	180.0	10.460	2
H-N-C-O	H-N-C-O	180.0	10.460	2
		0.0	8.368	1
H-N-CT-CT H-N-CT-C H-N-CT-H1	X-N-CT-X	0.0	0.000	0

Normalized probability overlap analysis: To monitor convergence of each set of BE-META simulations, we calculated probability density overlap as a function of simulation time. The population density in 3D principal subspace, called the normalized integrated product (NIP), is given by:

$$\mathrm{NIP} = \frac{2\sum_i \rho_i \rho_{i,ref}}{\sum_i \rho_i^2 + \sum_i \rho_{i,ref}^2}$$

where ρ_i and $\rho_{i,ref}$ are the population density of grid point *i* and its reference value, respectively. The summation is over all the grid points in the 3D principal subspace. An NIP value of 0 shows no overlap while a value of 1 indicates perfect overlap. The population density of S2 during the last time portion was used as the reference for S1 and vice versa.[6]

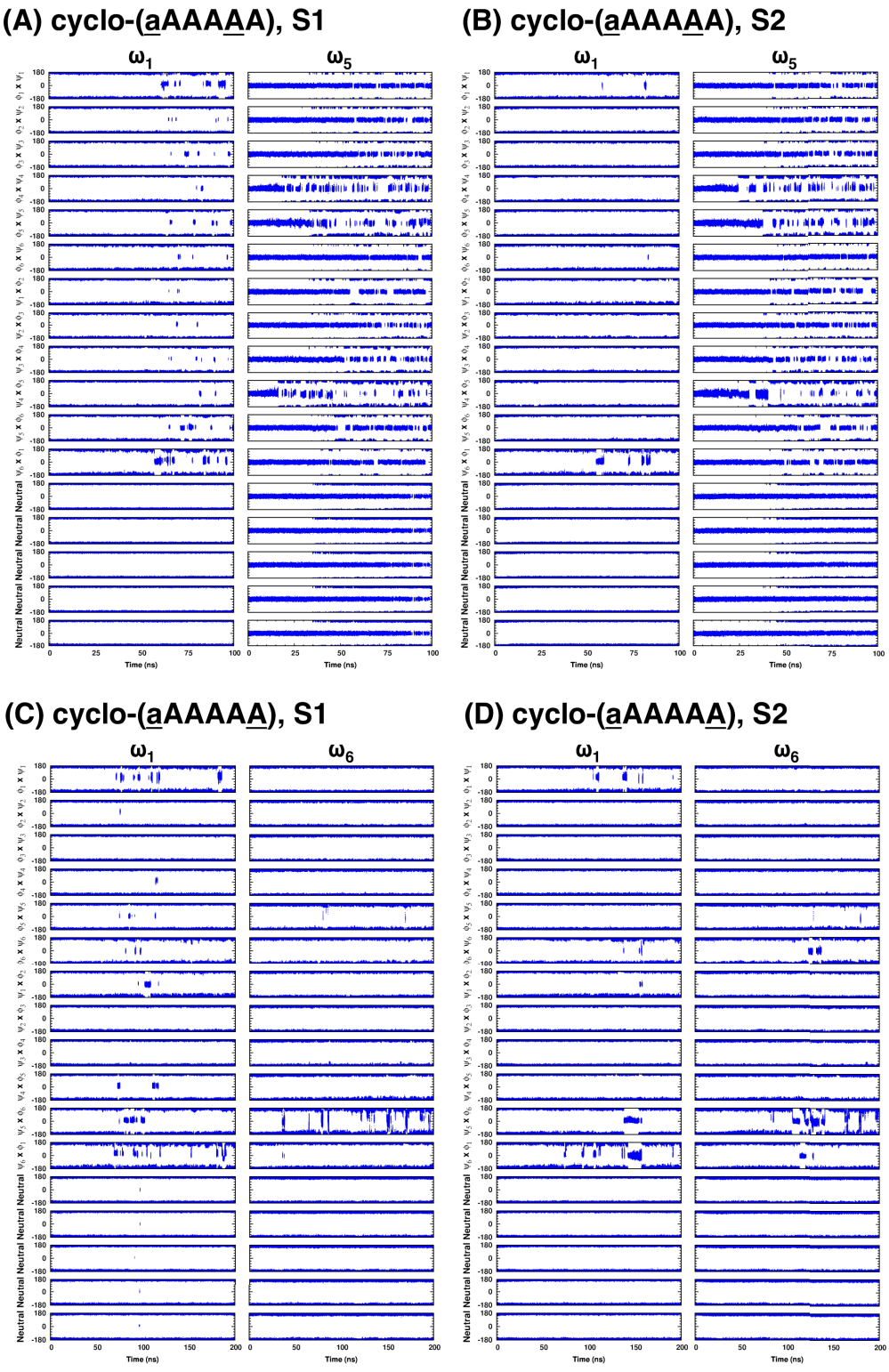


Figure S4. ω_1 and ω_5 trajectories of cyclo-(aAAAAA) for all replicas of (**A**) S1 and (**B**) S2 and ω_1 and ω_6 trajectories of cyclo-(aAAAAA) for all replicas of (**C**) S1 and (**D**) S2 with RSFF1. These simulations bias $\phi_i \times \psi_i$ and $\psi_i \times \phi_{i+1}$ with the correct isomers provided.

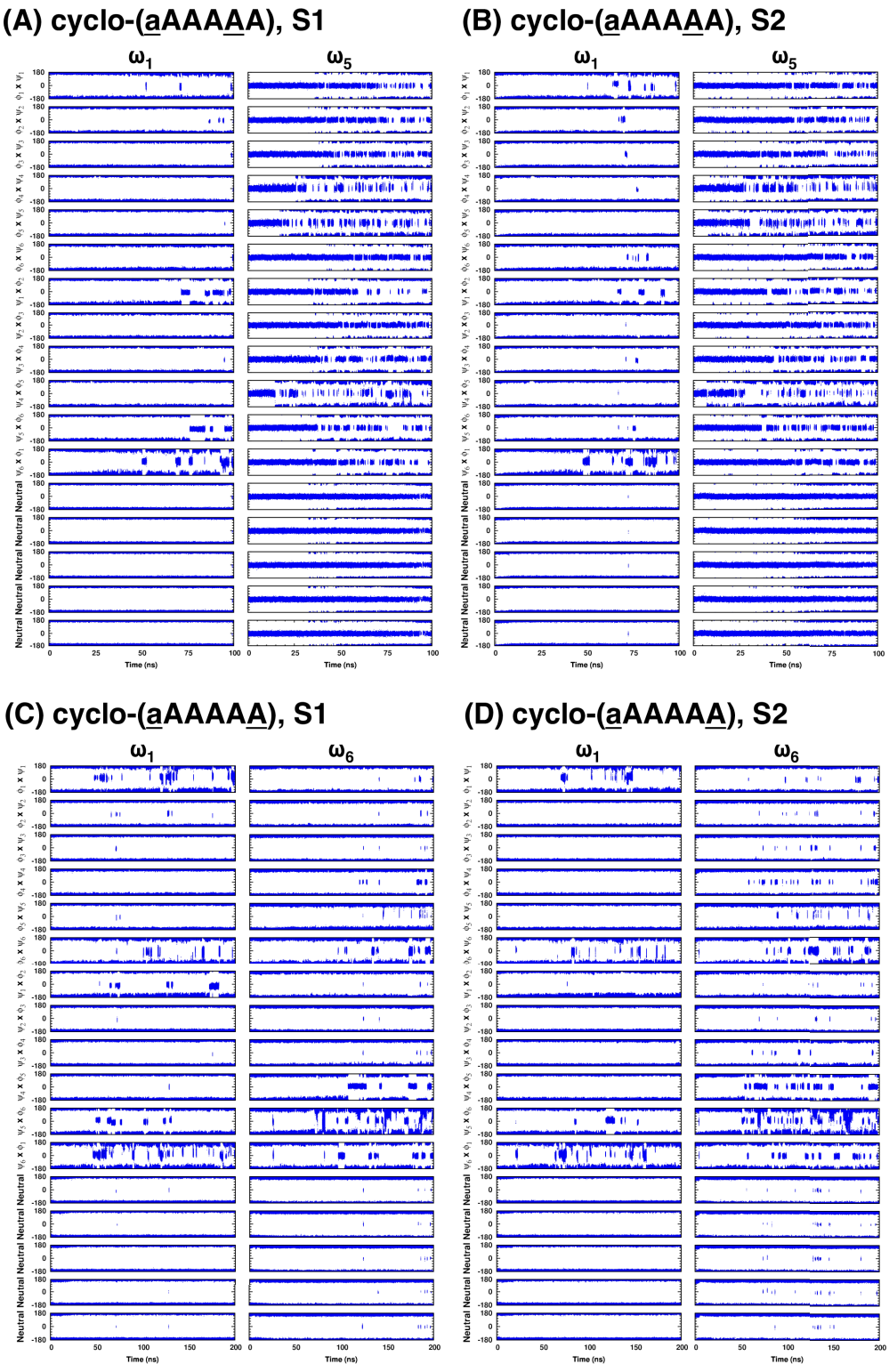


Figure S5. ω_1 and ω_5 trajectories of cyclo-(aAAAAA) for all replicas of (**A**) S1 and (**B**) S2 and ω_1 and ω_6 trajectories of cyclo-(aAAAAA) for all replicas of (**C**) S1 and (**D**) S2 with RSFF2. These simulations bias $\phi_i \times \psi_i$ and $\psi_i \times \phi_{i+1}$ with the correct isomers provided.

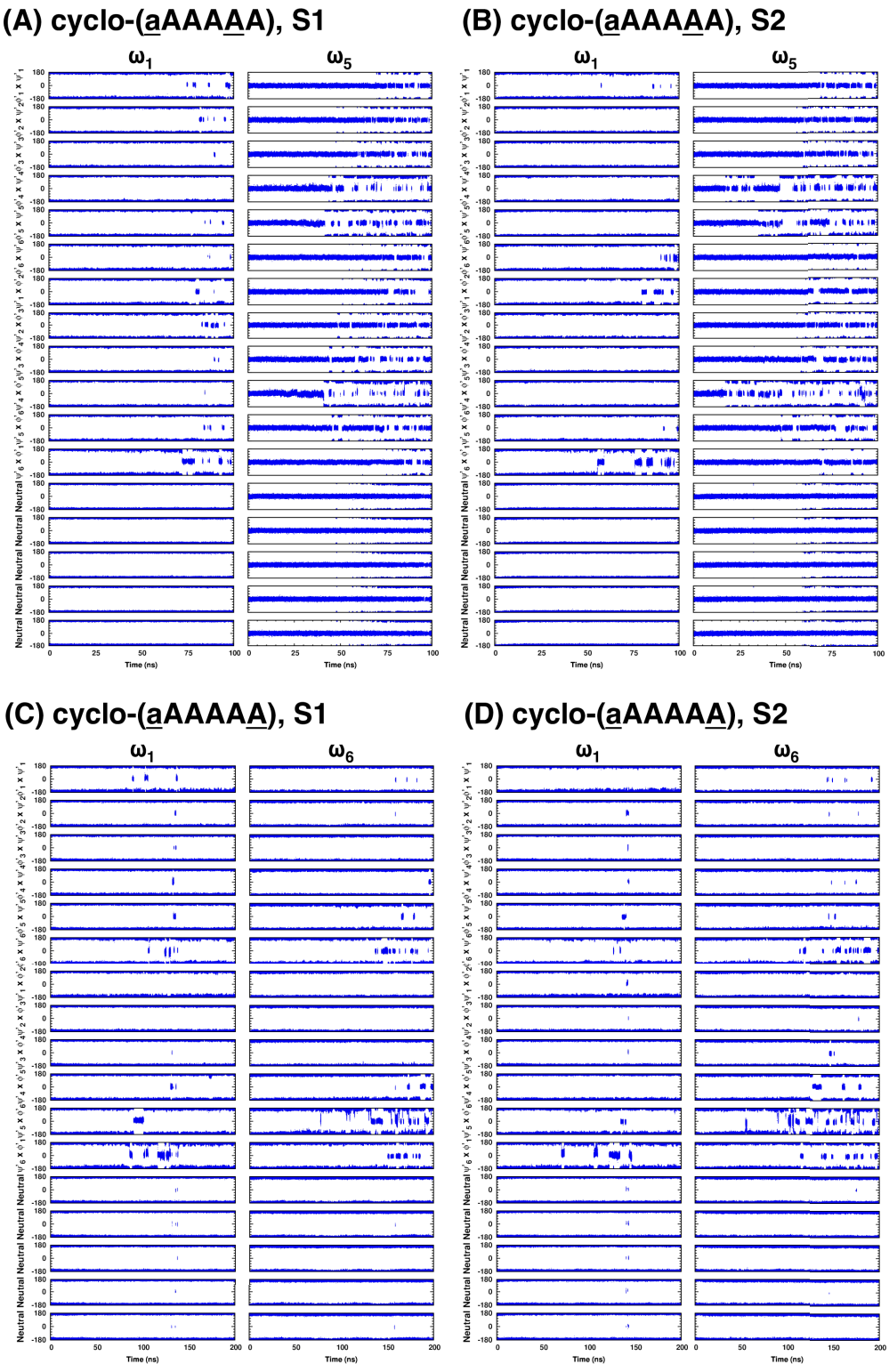


Figure S6. ω_1 and ω_5 trajectories of cyclo-(aAAAAA) for all replicas of (**A**) S1 and (**B**) S2 and ω_1 and ω_6 trajectories of cyclo-(aAAAAA) for all replicas of (**C**) S1 and (**D**) S2 with RSFF1. These simulations bias $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided.

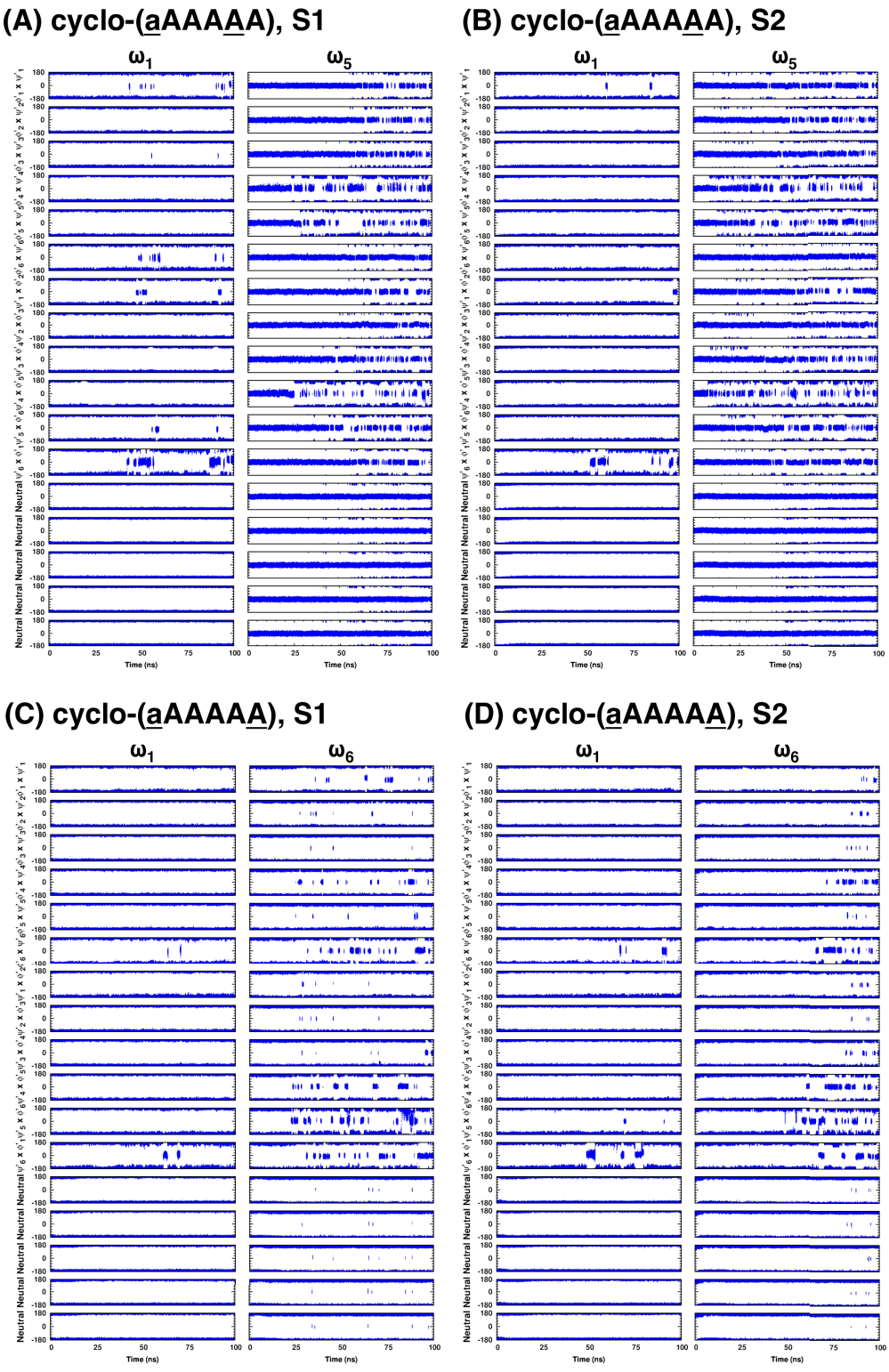


Figure S7. ω_1 and ω_5 trajectories of cyclo-(aAAAAA) for all replicas of (**A**) S1 and (**B**) S2 and ω_1 and ω_6 trajectories of cyclo-(aAAAAA) for all replicas of (**C**) S1 and (**D**) S2 with RSFF2. These simulations bias $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided.

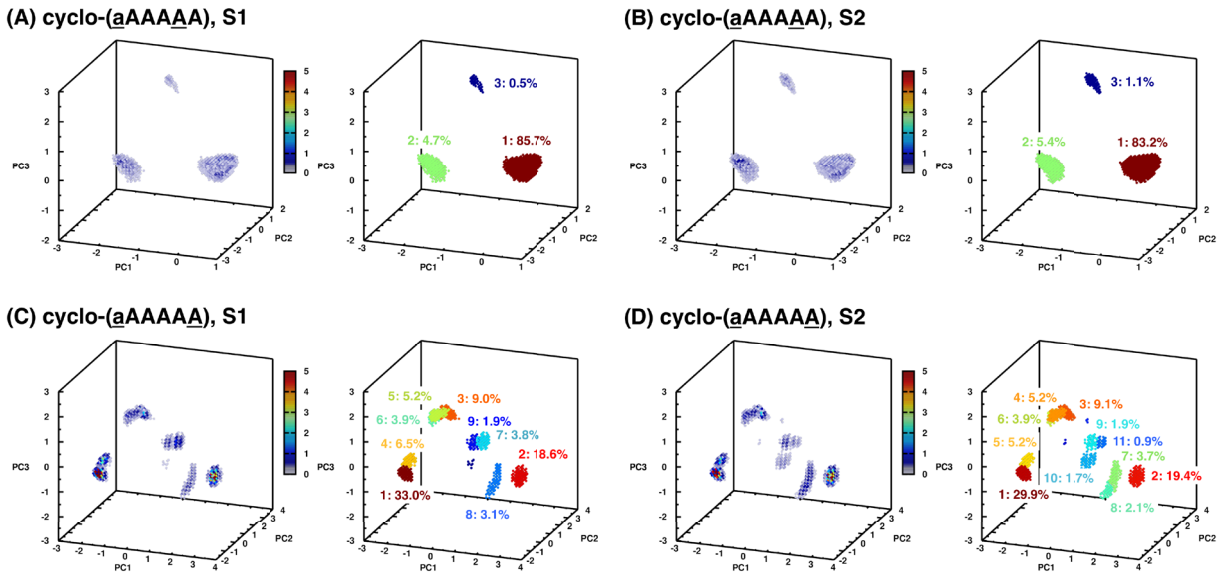


Figure S8. dPCA with the three largest principal components (left) and cluster analysis (right) for the correct isomer of cyclo-(aAAAAA), (**A**) S1 and (**B**) S2 and cyclo-(aAAAAA), (**C**) S1 and (**D**) S2 with RSFF1. Clusters are numbered based on populations, and clusters with populations > 0.5% are labeled. Results are from the last 50 ns of systems biasing $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided.

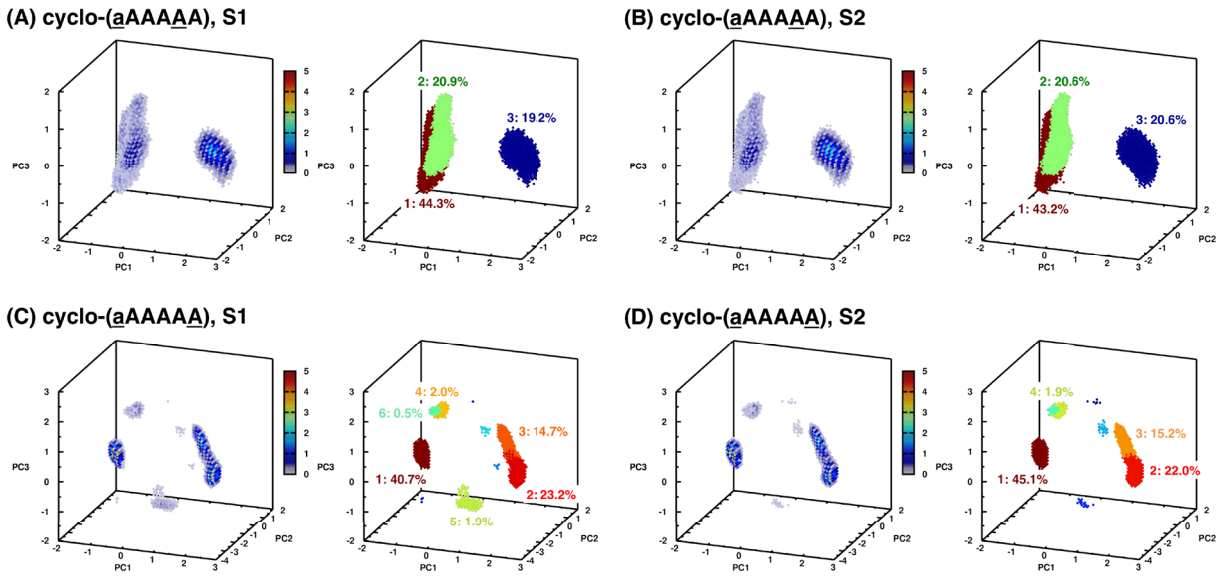


Figure S9. dPCA with the three largest principal components (left) and cluster analysis (right) for the correct isomer of cyclo-(aAAAAA), (**A**) S1 and (**B**) S2 and cyclo-(aAAAAA), (**C**) S1 and (**D**) S2 with RSFF2. Clusters are numbered based on populations, and clusters with populations > 0.5% are labeled. Results are from the last 50 ns of systems biasing $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided.

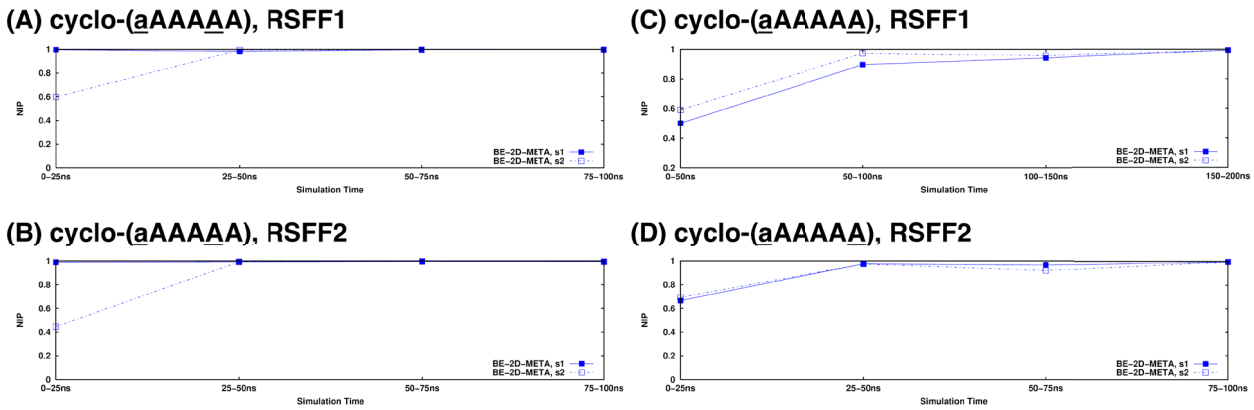


Figure S10. NIP of the neutral replicas of (**A**) cyclo-(aAAAAA), RSFF1; (**B**) cyclo-(aAAAAA), RSFF2; (**C**) cyclo-(aAAAAA), RSFF1 and (**D**) cyclo-(aAAAAA), RSFF2. Results are from the systems biasing $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided.

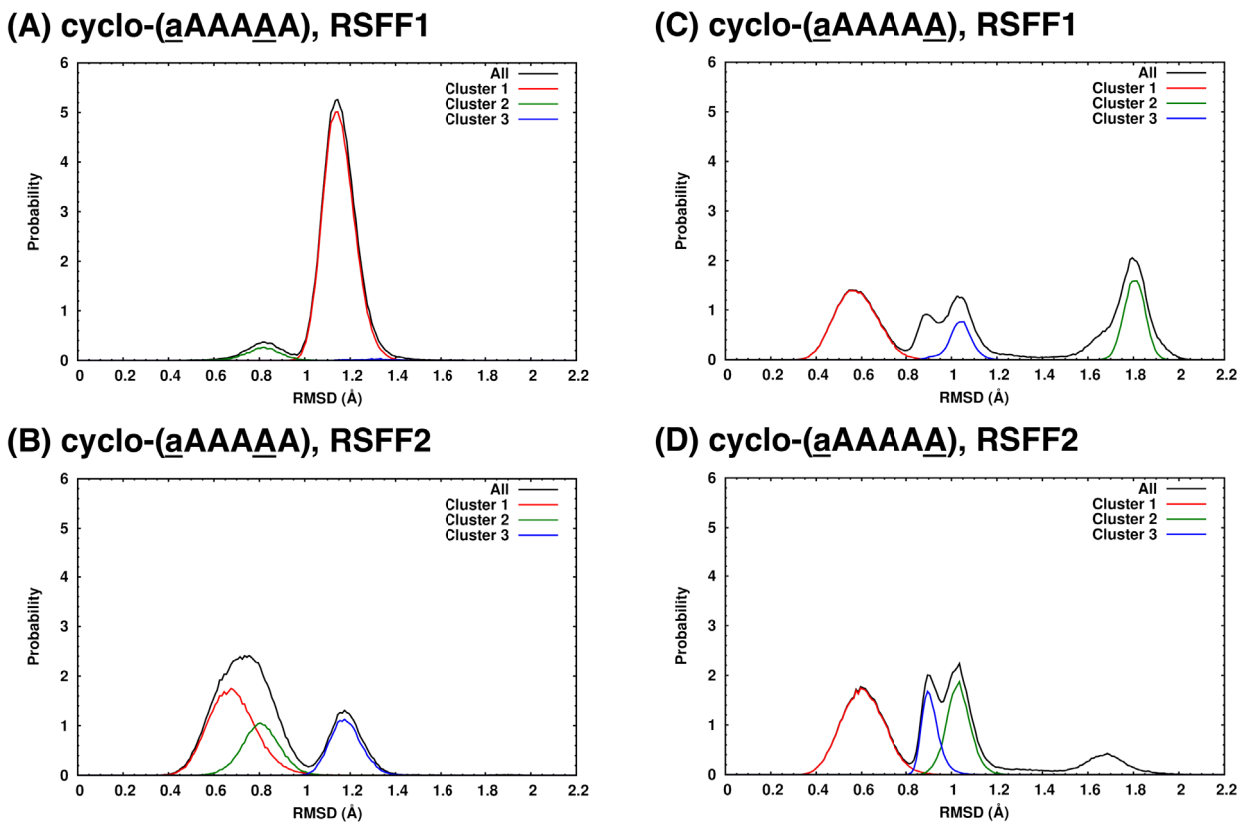


Figure S11. RMSD distributions of the correct isomer of (**A**) cyclo-(aAAAAA), RSFF1; (**B**) cyclo-(aAAAAA), RSFF2; (**C**) cyclo-(aAAAAA), RSFF1 and (**D**) cyclo-(aAAAAA), RSFF2. Note in (**A**), cluster 3 has a population of 0.5% and is located at ~1.3 Å. Results are from the last 50 ns of systems biasing $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided.

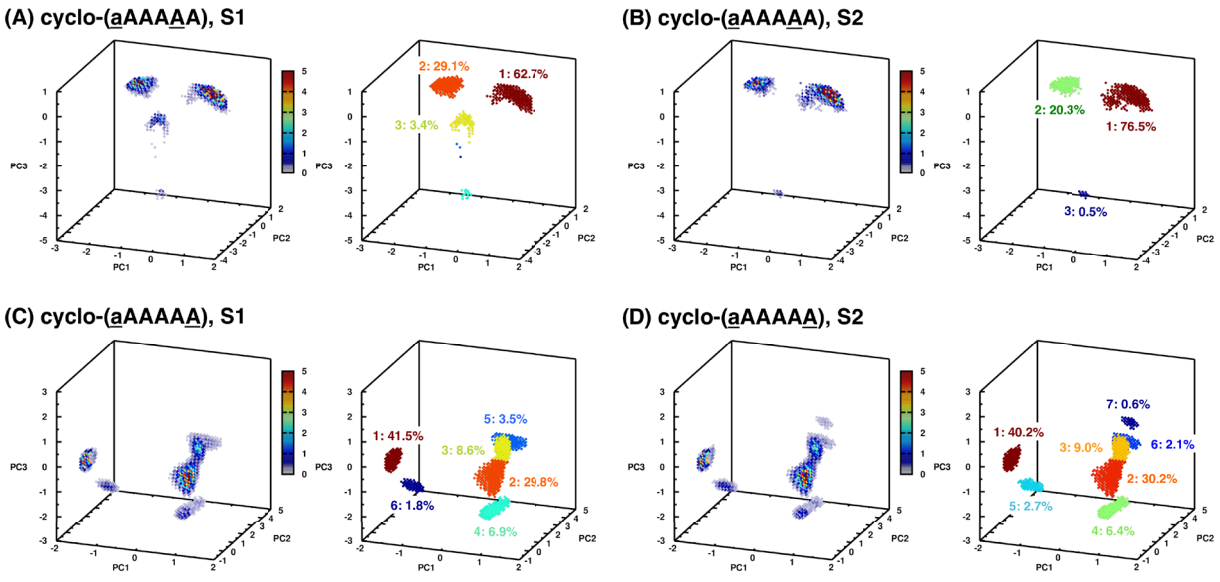


Figure S12. dPCA with the three largest principal components (left) and cluster analysis (right) for the correct isomer of cyclo-(aAAAAA), (**A**) S1 and (**B**) S2 and cyclo-(aAAAAA), (**C**) S1 and (**D**) S2 with RSFF1 in DMSO. Clusters are numbered based on populations. Results are from the last 50 ns of systems biasing $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided.

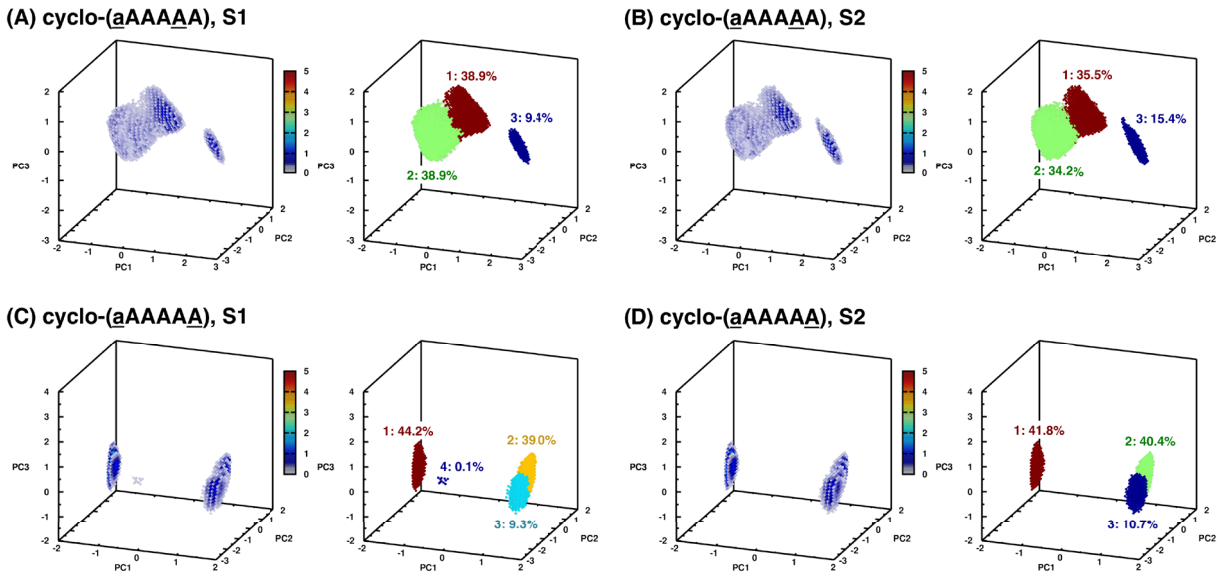


Figure S13. dPCA with the three largest principal components (left) and cluster analysis (right) for the correct isomer of cyclo-(aAAAAA), (**A**) S1 and (**B**) S2 and cyclo-(aAAAAA), (**C**) S1 and (**D**) S2 with RSFF2 in DMSO. Clusters are numbered based on populations. Results are from the last 50 ns of systems biasing $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided.

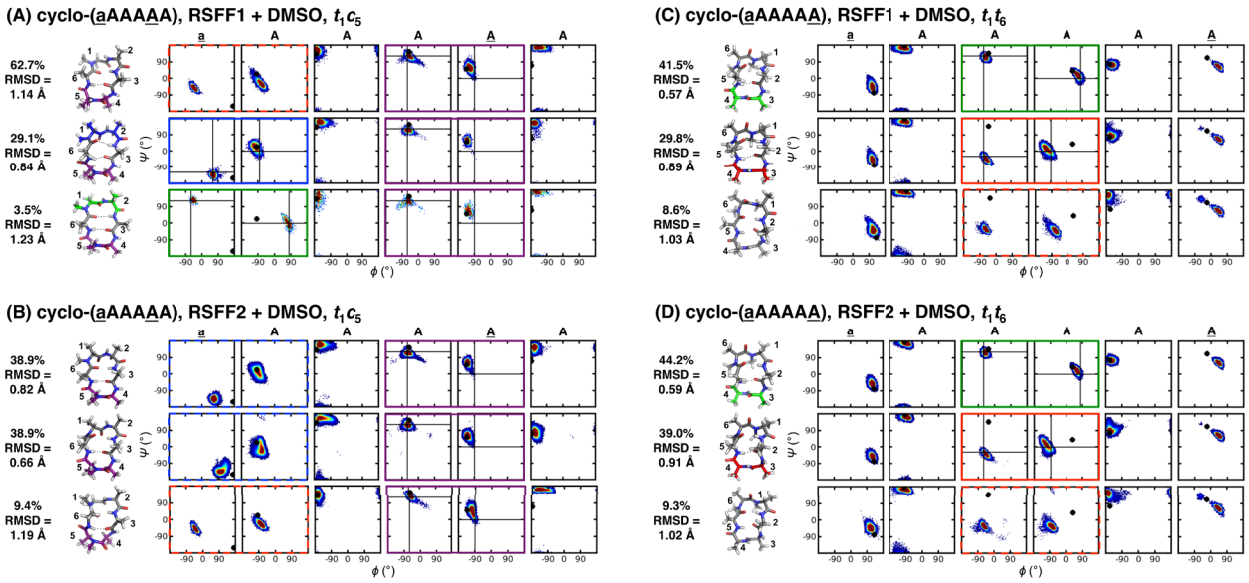


Figure S14. Structural ensembles of the CPs when the experimental isomers were given (*cis* peptide bond between residues 4 and 5 for cyclo-(aAAAAA) and all *trans* peptide bonds for cyclo-(aAAAAA)) and ω angles were not biased. Populations and Ramachandran plots of the top three clusters from cluster analysis with the correct isomer of cyclo-(aAAAAA), using (**A**) RSFF1 and (**B**) RSFF2; and cyclo-(aAAAAA) using (**C**) RSFF1 and (**D**) RSFF2. Representative structures can be found in **Figure 3**. Results are from the last 50 ns of systems biasing $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ in DMSO. ϕ and ψ of the NMR structures are shown as black dots. Type I, II, II' and VI turns were classified if they formed hydrogen bonds in > 50% of cluster frames and are colored in red, green, blue and purple boxes, respectively.

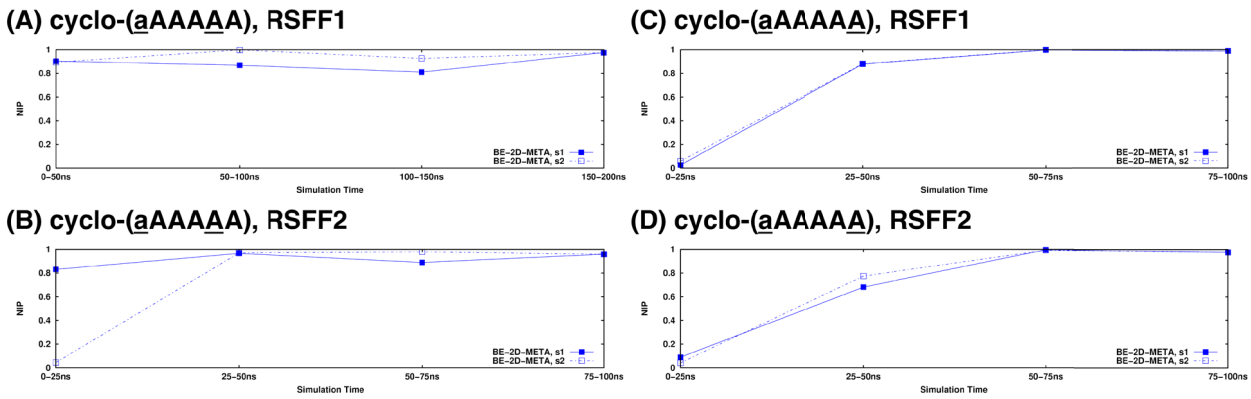


Figure S15. NIP of the neutral replicas of (**A**) cyclo-(aAAAAA), RSFF1 (**B**) cyclo-(aAAAAA), RSFF2, (**C**) cyclo-(aAAAAA), RSFF1 and (**D**) cyclo-(aAAAAA), RSFF2 in DMSO. Results are from the systems biasing $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided.

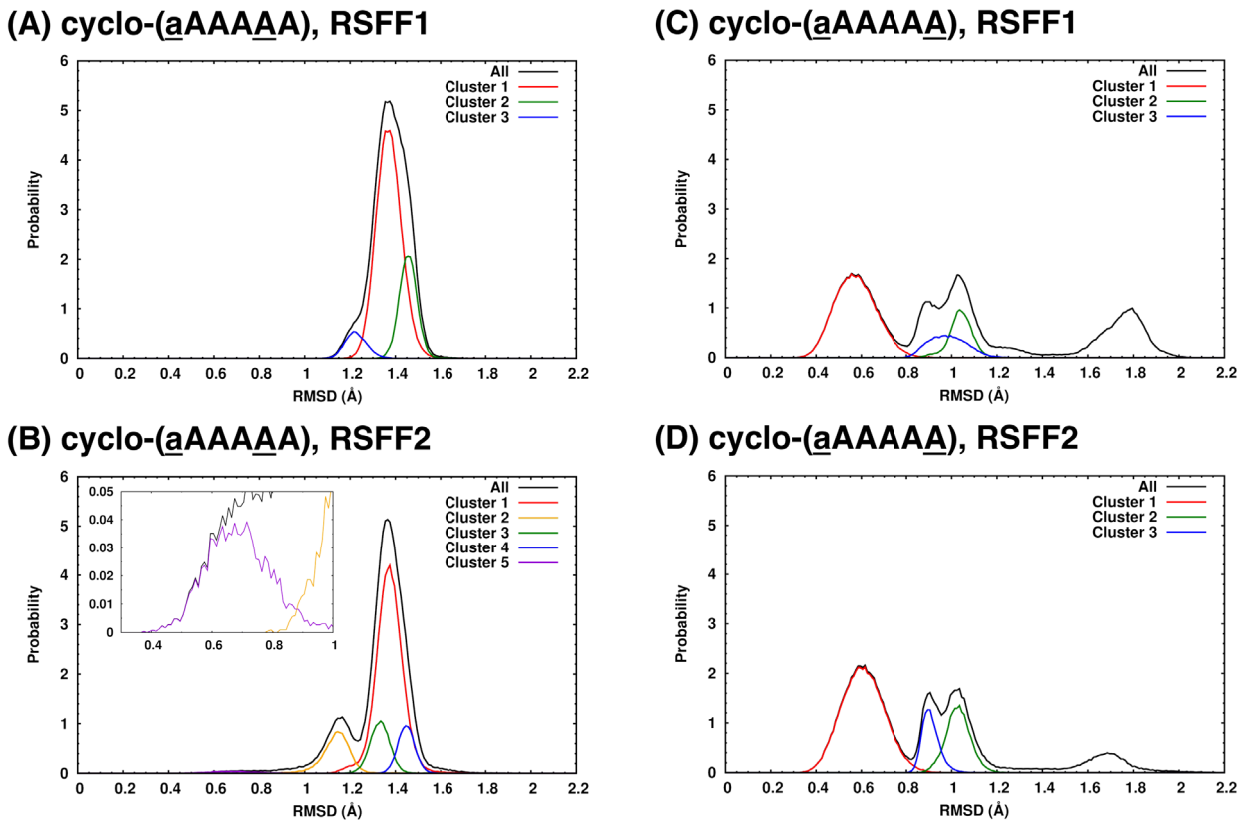
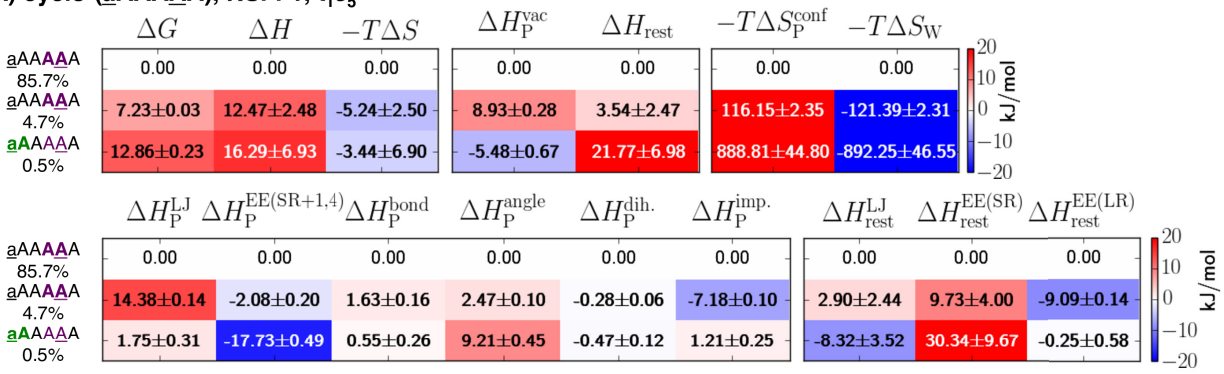


Figure S16. RMSD distributions of (**A**) cyclo-(aAAAAA), RSFF1; (**B**) cyclo-(aAAAAA), RSFF2; (**C**) cyclo-(aAAAAA), RSFF1 and (**D**) cyclo-(aAAAAA), RSFF2. Results are from the last 50 ns of systems biasing $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with additional biases on the ω angles involving the *N*-methylated nitrogen atoms.

Table S5. Thermodynamics decomposition of cyclo-(aAAAAA) using (A) RSFF1 and (B) RSFF2 when the experimental isomers were given and ω angles were not biased. Results shown are from the S1 simulations.

(A) cyclo-(aAAAAA), RSFF1, t_1c_5



(B) cyclo-(aAAAAA), RSFF2, t_1c_5

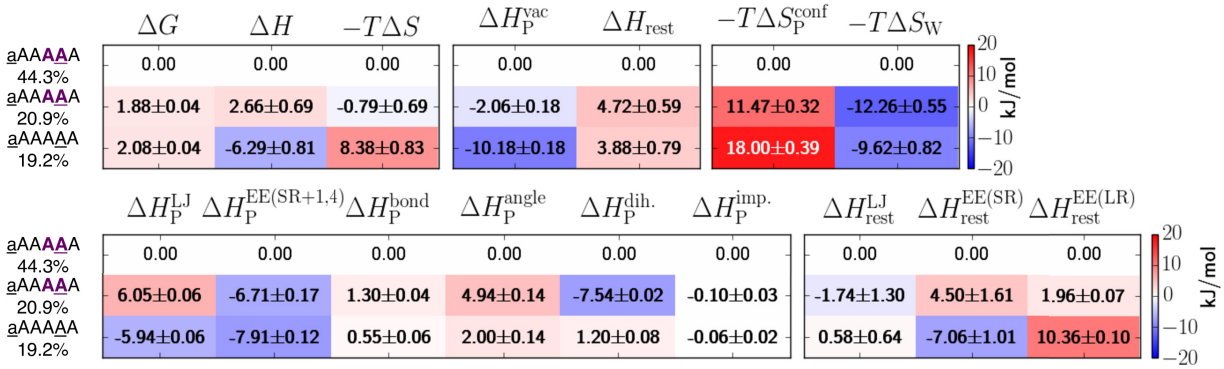
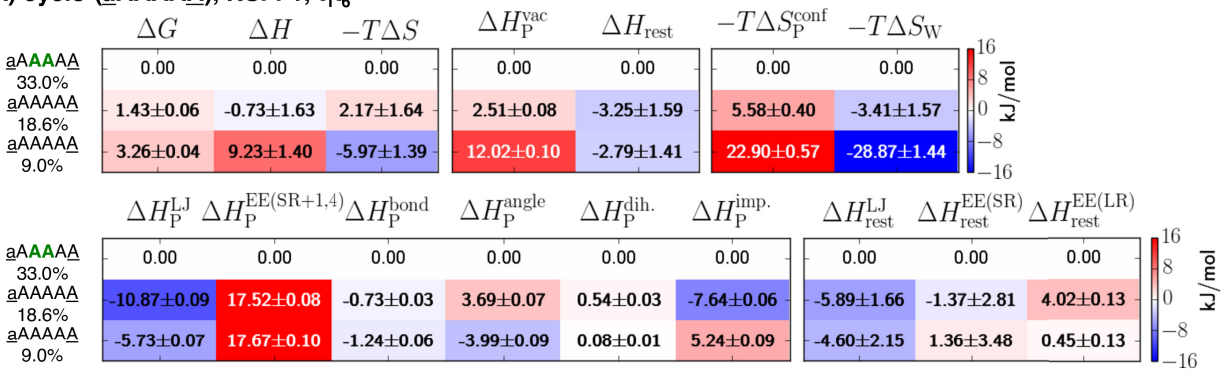


Table S6. Thermodynamics decomposition of cyclo-(aAAAAA) using (A) RSFF1 and (B) RSFF2 when the experimental isomers were given and ω angles were not biased. Results shown are from the S1 simulations.

(A) cyclo-(aAAAAA), RSFF1, t_1t_6



(B) cyclo-(aAAAAA), RSFF2, t_1t_6

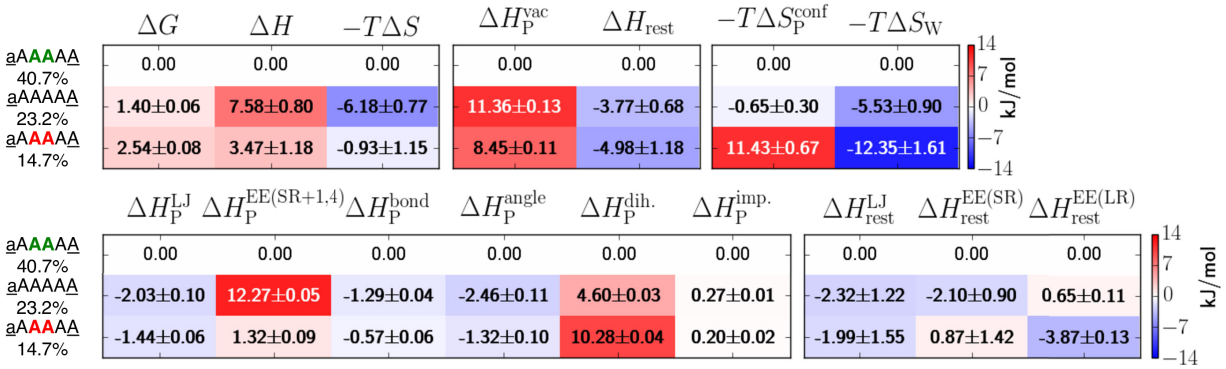


Table S7. Average number of violations of the NMR distance restraints[7] for the top three clusters of cyclo-(aAAAAA) and cyclo-(aAAAAA) using RSFF1 and RSFF2. These simulations bias $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided and correspond to the clusters shown in **Figure 3**.

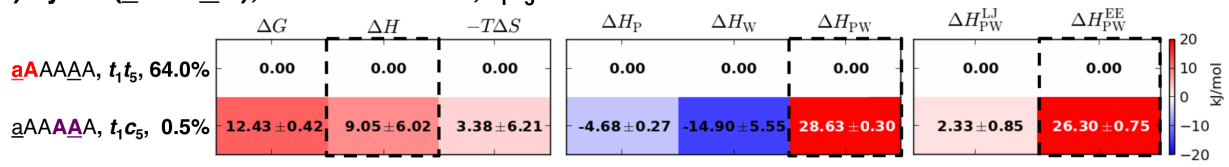
CP		Total Violations	Violations > 0.1 Å	Violations > 0.3 Å	Violations > 1.0 Å
cyclo-(aAAAAA), RSFF1	Cluster 1	11.04	8.82	5.64	1.53
	Cluster 2	7.54	5.21	2.66	0.32
	Cluster 3	10.17	7.37	4.21	1.74
cyclo-(aAAAAA), RSFF2	Cluster 1	6.58	4.65	2.71	0.56
	Cluster 2	7.27	4.67	2.26	0.17
	Cluster 3	10.60	8.72	5.85	1.39
cyclo-(aAAAAA), RSFF1	Cluster 1	4.51	2.17	0.90	0.05
	Cluster 2	17.43	15.57	11.13	5.21
	Cluster 3	7.96	6.06	3.35	0.54
cyclo-(aAAAAA), RSFF2	Cluster 1	4.09	2.19	0.85	0.06
	Cluster 2	7.56	5.53	3.17	0.75
	Cluster 3	5.14	3.57	2.55	0.66

Table S8. MaxSub scores and TM scores for the top three clusters of cyclo-(aAAAAA) and cyclo-(aAAAAA) using RSFF1 and RSFF2. These simulations bias $\phi'_i \times \psi'_i$ and $\psi'_i \times \phi'_{i+1}$ with the correct isomers provided and correspond to the clusters shown in **Figure 3**.

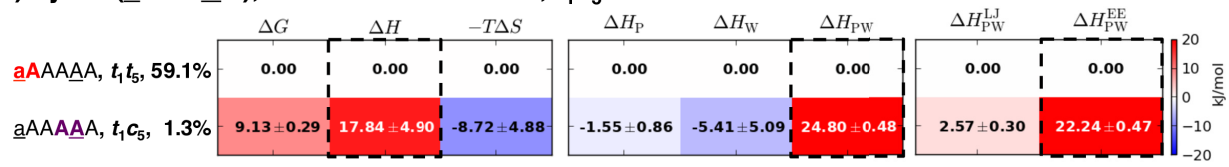
		MaxSubScore (d_0 = 3.5Å)			TM Score with d_0 = 1.0 Å		
		Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
cyclo-(aAAAAA)	RSFF1	0.92 ± 0.01	0.97 ± 0.01	0.95 ± 0.01	0.68 ± 0.03	0.76 ± 0.05	0.71 ± 0.03
	RSFF2	0.97 ± 0.01	0.97 ± 0.01	0.91 ± 0.01	0.77 ± 0.04	0.76 ± 0.05	0.69 ± 0.02
cyclo-(aAAAAA)	RSFF1	0.97 ± 0.01	0.83 ± 0.01	0.92 ± 0.01	0.80 ± 0.05	0.41 ± 0.02	0.62 ± 0.05
	RSFF2	0.97 ± 0.01	0.93 ± 0.01	0.97 ± 0.01	0.78 ± 0.05	0.62 ± 0.05	0.76 ± 0.05

Table S9. Thermodynamics decomposition of cyclo-(aAAAAA) using RSFF2 with different water models. Results are shown from the S1 simulations when ω angles involving *N*-methylated nitrogen atoms were also biased.

(A) cyclo-(aAAAAA), RSFF2 + SPC/E, t_1c_5



(B) cyclo-(aAAAAA), RSFF2 + TIP4P/2005, t_1c_5



(C) cyclo-(aAAAAA), RSFF2 + TIP4P/EW, t_1c_5

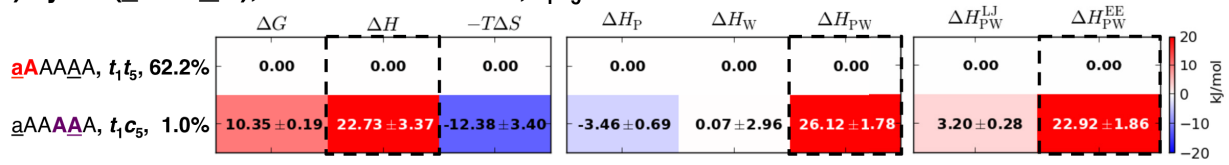
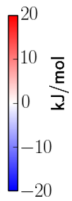


Table S10. Thermodynamics decomposition of the most populated t_1t_5 and t_1c_5 structures of cyclo-(aAAAAA) using RSFF2. Results shown are from the S1 simulations when the ω angles involving the *N*-methylated nitrogen atoms were also biased.

	ΔG	ΔH	$-T\Delta S$	$\Delta H_{\mathrm{P}}^{\mathrm{vac}}$	ΔH_{rest}	$-T\Delta S_{\mathrm{P}}^{\mathrm{conf}}$	$-T\Delta S_{\mathrm{W}}$
aAAAAA, t_1t_5, 57.1%	0.00	0.00	0.00	0.00	0.00	0.00	0.00
aAAAAA, t_1t_5, 11.4%	4.00±0.05	8.15±1.39	-4.15±1.38	-0.92±0.17	9.07±1.49	26.20±0.44	-30.34±1.50
aAAAAA, t_1t_5, 11.3%	4.04±0.04	-2.03±1.72	6.07±1.74	17.61±0.20	-19.64±1.63	29.52±0.69	-23.45±1.80
aAAAAA, t_1t_5, 9.0%	4.61±0.10	1.72±1.67	2.88±1.72	-12.85±0.15	14.57±1.64	44.75±1.76	-41.86±1.90
aAAAAA, t_1c_5, 1.0%	9.63±0.10	21.70±3.39	-12.06±3.35	42.42±0.61	-20.72±3.60	418.05±10.14	-430.11±12.35

	$\Delta H_{\mathrm{P}}^{\mathrm{LJ}}$	$\Delta H_{\mathrm{P}}^{\mathrm{EE(SR+1,4)}}$	$\Delta H_{\mathrm{P}}^{\mathrm{bond}}$	$\Delta H_{\mathrm{P}}^{\mathrm{angle}}$	$\Delta H_{\mathrm{P}}^{\mathrm{dih.}}$	$\Delta H_{\mathrm{P}}^{\mathrm{imp.}}$	$\Delta H_{\mathrm{rest}}^{\mathrm{LJ}}$	$\Delta H_{\mathrm{rest}}^{\mathrm{EE(SR)}}$	$\Delta H_{\mathrm{rest}}^{\mathrm{EE(LR)}}$
aAAAAA, t_1t_5, 57.1%	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
aAAAAA, t_1t_5, 11.4%	-1.83±0.08	-6.52±0.07	0.37±0.11	0.12±0.18	6.75±0.10	0.19±0.04	4.77±1.37	11.20±2.76	-6.90±0.09
aAAAAA, t_1t_5, 11.3%	-4.69±0.09	17.00±0.07	-0.27±0.14	0.18±0.10	5.44±0.07	-0.06±0.02	2.35±1.71	-21.64±2.65	-0.35±0.07
aAAAAA, t_1t_5, 9.0%	-0.73±0.10	-15.97±0.12	-0.06±0.06	3.23±0.13	0.71±0.04	-0.03±0.01	2.26±0.94	7.58±2.04	4.72±0.10
aAAAAA, t_1c_5, 1.0%	4.11±0.28	31.00±0.28	-0.84±0.39	-0.17±0.84	8.00±0.15	0.31±0.10	-3.59±2.35	14.87±4.26	-31.99±0.42



REFERENCES

1. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L., Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B.* **2001**, 105, 6474-6487.
2. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A., A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, 97, 10269-10280.
3. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera–A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, 25, 1605-1612.
4. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A., et al. *Gaussian 09*, Gaussian, Inc.: Wallingford, CT, USA, 2009.
5. Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P., The R.E.D. Tools: Advances in RESP and ESP Charge Derivation and Force Field Library Building. *Phys. Chem. Chem. Phys.* **2010**, 12, 7821-7839.
6. Damas, J. M.; Filipe, L. C. S.; Campos, S. R. R.; Lousa, D.; Victor, B. L.; Baptista, A. M.; Soares, C. M., Predicting the Thermodynamics and Kinetics of Helix Formation in a Cyclic Peptide Model. *J. Chem. Theory Comput.* **2013**, 9, 5148-5157.
7. Beck, J. G.; Chatterjee, J.; Laufer, B.; Kiran, M. U.; Frank, A. O.; Neubauer, S.; Ovadia, O.; Greenberg, S.; Gilon, C.; Hoffman, A., et al., Intestinal Permeability of Cyclic Peptides: Common Key Backbone Motifs Identified. *J. Am. Chem. Soc.* **2012**, 134, 12125-12133.