**Supporting Information for:**

**"Preserved Transmembrane Segment Topology, Structure, and Dynamics in Disparate Micellar Environments" in *The Journal of Physical Chemistry Letters***

*David N. Langelaan[1], Aditya Pandey[1], Muzaddid Sarker[1] and Jan K. Rainey[1,2]**

1. Department of Biochemistry & Molecular Biology, Dalhousie University, Halifax NS B3H 4R2, Canada
2. Department of Chemistry, Dalhousie University, Halifax NS B3H 4R2, Canada

* Address correspondence to JKR at jan.rainey@dal.ca

## Experimental Methods

**Materials.** SDS-$d_{25}$, DPC-$d_{38}$, HFIP-$d_2$, dithiothreitol (DTT)-$d_6$, deuterium oxide ($D_2O$; 99.8 atom % D) and $D_2O$ containing 1% (w/w) DSS were obtained from C/D/N Isotopes (Pointe-Claire, QC). LPPG was purchased from Avanti Polar lipids (Alabaster, AL). AR55 was produced as a C-terminally hexahistidine-tagged recombinant protein (comprising residues 1-55 of human AR, with a C-terminal $KKGH_6$ extension) in *Escherichia coli* BL-21(DE3), extracted following inclusion body isolation, and purified using reverse-phase high performance liquid chromatography as detailed in our previous work.[1] $MnCl_2$ and all other chemicals were obtained from Sigma-Aldrich (Oakville, ON).

**NMR sample preparation and data analysis.** All NMR solutions with micelles contained 20 mM $Na^+CD_3COO^-$, 1 mM DSS, 1 mM $NaN_3$, 10 mM DTT-$d_6$ and were adjusted to pH 5.00 ± 0.05 (SDS and LPPG) or 4.00 ± 0.05 (DPC) without accounting for deuterium isotope effects. Samples of uniformly $^{13}C$- and $^{15}N$-enriched AR55 with SDS (0.8 mM AR55, 96 mM SDS-$d_{25}$, 95% $H_2O$, 5% $D_2O$) and LPPG (0.56 mM AR55, 72 mM LPPG, 95% $H_2O$, 5% $D_2O$) and a

sample containing HFIP (0.9 mM AR55, 1 mM DSS, 5 mM DTT-$d_6$, 10% $D_2O$, 40% $H_2O$, 50% HFIP-$d_2$) were prepared. $^{1}H$-$^{15}N$ HSQC, $^{1}H$-$^{13}C$ HSQC, HNCO, HNcaCO, HNCA, HNcoCA, HNCACB, $^{15}N$-edited TOCSY, HcCH TOCSY, $^{15}N$-edited NOESY and $^{13}C$-edited NOESY experiments were collected for the NMR samples at 700 MHz (Bruker Avance III, NRC BMRF, Halifax, NS) or 800 MHz (Varian INOVA, NANUC, Edmonton, AB). All NOESY experiments for AR55 in SDS or LPPG were collected with a 125 ms mixing time, while a 150 ms mixing time was used for AR55 in 50% HFIP.

For each NMR sample condition, data were processed, analyzed, and structures were generated using XPLOR-NIH ver. 2.19[2-3] as detailed for AR55 with DPC micelles.[1] The final 40 member structural ensembles of AR55, validated using PROCHECK-NMR[4] and analyzed for regions of structural convergence in the context of conformational sampling as detailed previously,[5] have been deposited in the PDB[6] using SMSDep (accession # 2LOT, 2LOV and 2LOW for AR55 solubilized with SDS, LPPG and 50% HFIP respectively). The chemical shift, spectral peak, and restraint data were deposited into the BMRB[7] (accession #18224, 18226 and 18227 for AR55 solubilized with SDS, LPPG and 50% HFIP respectively).

**Paramagnetic relaxation enhancement (PRE).** Additional NMR samples of uniformly $^{15}N$-enriched AR55 were prepared as above except with 10% $D_2O$ and 90% $H_2O$ and either 77 mM DPC or 125 mM LPPG with 0.5 mM AR55 or 110 mM SDS with 1 mM AR55. Samples were titrated with $Mn^{2+}$ up to 1 mM, with $^{1}H$-$^{15}N$ HSQC spectra collected at each point. Data for AR55 solubilized in SDS micelles were collected at Agriculture and Agri-Food Canada (AAFC, Charlottetown, PE) on a 600 MHz Bruker Avance III instrument equipped with a 1.7 mm room temperature TXI probe. Data for AR55 solubilized in DPC and LPPG micelles were collected on a 700 MHz Bruker Avance III spectrometer equipped with a 1.7 mm TCI cryoprobe at the NRCI

BMRF facility. For all datasets PRE was measured as the ratio of peak heights in attenuated ($I_{att}$) and reference ($I_{ref}$) spectra. Error ($\sigma_{PRE}$) was calculated by evaluating the standard deviation of the PRE as:

$$\sigma_{PRE} = | PRE | ((\sigma_{I_{att}} / I_{att})^2 + (\sigma_{I_{ref}} / I_{ref})^2)^{1/2} \quad (1)$$

where $\sigma_{I_{att}}$ and $\sigma_{I_{ref}}$ represent the standard deviations of the spectral noise (as measured using NMRDraw[8]) in the $I_{att}$ and $I_{ref}$ spectra. Structural ensembles were visualized using PyMOL (Schrödinger LLC, Cambridge, MA).

**NMR relaxation data collection and analysis.** NMR samples of $^{13}C/^{15}N$ labelled AR55 (1 mM AR55, 20 mM $Na^+CD_3COO^-$, 1 mM DSS, 5 mM DTT, 1 mM $NaN_3$, 90% $H_2O$/10% $D_2O$ ) were prepared with either 102 mM DPC or 150 mM LPPG. In addition, the previously prepared $^{13}C/^{15}N$ sample of AR55 in SDS micelles was used. For all samples, $^1H$-$^{15}N$ HSQC spectra were collected to measure the $R_1$ and $R_2$ relaxation rates as well as the heteronuclear NOE at 16.4 T. For $R_1$ measurements, recovery delays of 50, 100, 250, 600, 1000 and 1500 ms and for $R_2$ measurements delays of 10, 30, 50, 70, 90, 130 and 150 ms were used. For AR55 solubilized in DPC micelles a $R_1$ measurement at 250 ms was not collected. For AR55 solubilized in SDS micelles the $R_1$ time points at 50 and 100 ms were removed due to an experimental artifact. The $^1H$-$^{15}N$ NOE was measured as the ratio of a saturated spectrum to a reference spectrum and error was propagated by measuring the spectral noise.[9]

Since AR55 is composed of two distinct regions (TM and N-terminal tail), calculation of a meaningful and accurate correlation time for the complex is difficult. For this reason we analyzed our relaxation data with reduced spectral density mapping,[10] which does not rely on an assumption of the correlation time. Specifically, $R_1$, $R_2$ and the heteronuclear NOE for a given backbone $^{15}N$ nucleus were used to infer the values of J(0), $J(\omega_N)$ and $J(0.87\omega_H)$ at that site.

Mathematica (Wolfram Research Inc., Champaign, IL) notebooks[9] were used to fit the $R_1$ and $R_2$ relaxation data and fit the relevant reduced spectral density mapping parameters. 100 Monte Carlo simulations were performed to determine the errors of the fitted parameters.

**Diffusion-ordered spectroscopy (DOSY) and hydrodynamics analysis**. The translational diffusion coefficient ($D_C$) of each AR55-micelle system was determined using $^1$H diffusion ordered spectroscopy (DOSY) (32 scans, sweep width 14 ppm, relaxation delay 1.5 s) employing pulsed field gradient (PFG) NMR[11] on a Bruker Avance 500 MHz (11.74 T) spectrometer equipped with a z-axis gradient and a BBFO SmartProbe. The DOSY pulse program incorporated stimulated echo and LED with bipolar gradient pulses and two spoil gradients[12] alongside presaturation during the relaxation delay for water suppression. The envelope of $^1$H signals was attenuated by linearly ramping the gradient strength up from 2% to 95% of the maximum (53.5 G/cm) in 16 steps. $D_C$ was calculated using Dynamics Center (Bruker) from the exponential fit of the observed signal decay as a function of gradient strength employing the Stejskal-Tanner formula:[13]

$$I = I(0)\exp[-D_C\times(2\pi\gamma g\delta)^2\times(\Delta-\delta/3)\times 1e4] \quad (2)$$

where I is the observed signal intensity, I(0) is the unattenuated signal intensity, $\gamma$ is the gyromagnetic ratio of $^1$H (4257.7 Hz/G), g is the gradient strength, $\delta$ is the gradient pulse length (optimized for a given sample between 6 and 8 ms), and $\Delta$ is the diffusion time (100 ms). The hydrodynamic diameter ($d_H$) of each system was calculated using the observed $D_C$ from the Stokes-Einstein equation[14]:

$$d_H = k_B T/3\pi\eta D_C \quad (3)$$

where $k_B$ is the Boltzmann constant, T is the absolute temperature, and $\eta$ is the solution viscosity (estimated based on the ratio of $H_2O$ and $D_2O$ for a given sample).

**Molecular dynamics simulations and trajectory analysis.** All simulations were performed using the GROMACS 4.6.1 software package[15-16] with graphical processing unit (GPU) card support employing the GROMOS 54A7 force field[17] and SPC water model.[18] DPC parameters were obtained from the automatic topology builder repository[19] based on charges and topology reported by Tieleman et al.[20] Simulations were carried out in rhombic dodecahedron boxes employing periodic boundary conditions with 54 DPC monomers, ca. 35000-36000 waters and 0.15 M NaCl as a balance of system electrostatic charge. Each of the first four ensemble members of AR55 in DPC (PDB entry 2LOU[1]) was employed as the starting polypeptide conformation for a series of four independent MD simulations.

Following randomized DPC placement and solvation, with steepest-decent energy minimization following each step, temperature and pressure equilibration were carried out with position restraints on the protein. Production runs at constant temperature, pressure, and number of particles were then carried out for ~100 ns following converged micelle formation (170-270 ns total), as judged by following the radius of gyration ($R_g$) of DPC phosphate headgroups. In production runs, the particle-mesh Ewald algorithm[21-22] was employed for long-range electrostatics and the Verlet scheme[23] for short-range electrostatics, with a 1.0 nm cut-off for short vs. long range. In each simulation, a 2 fs time step was employed for integration, the LINCS algorithm[24] was used to constrain protein and detergent bond lengths and the SETTLE algorithm[25] to constrain water bond lengths and angles. Temperature was maintained at 310 K with a velocity rescaling thermostat[26] with 0.1 ps coupling constant with detergent, protein, and solvent coupled separately. Pressure was held at 1.0 bar by the Parrinello-Rahman barostat,[27-28] with a time constant of 2 ps and compressibility of $4.5 \times 10^{-5}$ $bar^{-1}$.

Each trajectory was processed before analysis using the trjconv tool available with the GROMACS package. The molecules were first made whole and then the protein and DPC molecules were clustered. Periodic boundary jumps were removed using the first frame of the respective trajectory as a reference. The system was then centered around the protein and a compact unit cell representation was employed. Every 100$^{th}$ frame (i.e., at 50 ps intervals) of the trajectory thus obtained was written to obtain the final trajectory used for analysis. The radius of gyration (Rg) was calculated for the DPC molecules for the period of simulation using g_gyrate tool and was used to evaluate micelle formation. The solvent accessible surface area (SASA) was calculated for DPC+protein over the period of the simulation for all trajectories using the g_SAS function. Based on converged Rg and SASA, the last 100 ns from each trajectory was extracted for further analysis of the protein-micelle complex behavior. Deviation and order parameter (S, as defined by Hyberts et al.[29]) of the $\phi$ and $\psi$ dihedral angles for the 2000 frames for last 100 ns of the simulation for each model were calculated using an in-house tcl/Tk script. MD trajectories and frames were visualized using VMD.[30]

**Table S1.** The number of possible assignments and the numbers of assigned resonances for different isotopes and resonance types of AR55 in various membrane mimetics. Data for this table were generated using CCPNMR Analysis.[31] Data for DPC were published previously[1] and are provided strictly for comparative purposes.

| Category | Total possible | AR55_SDS assignments | | AR55_DPC assignments | | AR55_LPPG assignments | | AR55_HFIP assignments | |
|---|---|---|---|---|---|---|---|---|---|
| | | # | % | # | % | # | % | # | % |
| Carbon | 315 | 262 | 83 | 270 | 86 | 258 | 82 | 258 | 82 |
| Proton | 363 | 348 | 96 | 358 | 99 | 350 | 96 | 350 | 96 |
| Nitrogen | 88 | 70 | 80 | 70 | 80 | 70 | 80 | 70 | 80 |
| Amide | 126 | 124 | 98 | 124 | 98 | 124 | 98 | 124 | 98 |
| Backbone | 254 | 252 | 99 | 249 | 98 | 246 | 97 | 246 | 97 |
| Backbone non-H | 192 | 190 | 99 | 187 | 97 | 184 | 96 | 184 | 96 |
| Side chain H | 301 | 286 | 95 | 296 | 98 | 288 | 96 | 288 | 96 |
| Side chain non-H | 211 | 142 | 67 | 153 | 73 | 144 | 68 | 144 | 68 |

**Table S2.** Summary of the restraints employed for the final ensemble of 40 retained structures from 100 calculated structures. XPLOR-NIH energies, violation occurrences and structural statistics for AR55 solubilized in SDS, DPC, LPPG and 50% HFIP are shown. Data for DPC were published previously[1] and are provided strictly for comparative purposes.

| | **SDS** | **DPC** | **LPPG** | **50% HFIP** |
|---|---|---|---|---|
| **Rounds of structure calculation** | 9 | 11 | 18 | 15 |
| **Unique NOE restraints** | | | | |
| Total | 1556 | 1608 | 1654 | 2355 |
| Intraresidue | 740 | 788 | 756 | 776 |
| Sequential | 444 | 424 | 470 | 594 |
| Medium range ($\|i-j\| \leq 4$) | 320 | 322 | 330 | 728 |
| Long range ($\|i-j\| > 4$) | 2 | 0 | 0 | 6 |
| Ambiguous | 50 | 74 | 98 | 251 |
| **Residue Ramachandran plot statistics** | | | | |
| Core | 43.0% | 39.2% | 30.6% | 63.0% |
| Allowed | 45.5% | 47.3% | 55.6% | 31.4% |
| Generously allowed | 10.1% | 9.3% | 12.1% | 4.2% |
| Disallowed | 1.4% | 4.2% | 1.7% | 1.5% |
| **XPLOR-NIH energies (kcal/mol)**[a] | | | | |
| Total | 43.5 ± 1.4 | 39.8 ± 1.7 | 42.3 ± 2.5 | 44.6 ± 2.9 |
| NOE | 1.3 ± 0.6 | 1.1 ± 0.6 | 1.5 ± 0.9 | 2.0 ± 1.2 |
| **Violations** | | | | |
| NOE violations > 0.5 Å | 0 | 0 | 0 | 0 |
| NOE violations of 0.3-0.5 Å | 0 | 1 | 0 | 0 |
| NOE violations > 0.2-0.3 Å | 1 | 0 | 1 | 5 |

[a] Ranges are given by average deviations for XPLOR-NIH energies.

**Table S3.** RMSD values of the various superpositions of AR55 illustrated in Figure 1 of the main manuscript. Data for DPC were published previously[1] and are provided strictly for comparative purposes.

| | Superposed Residues | RMSD (Å) |
|---|---|---|
| **AR55 in SDS micelles** | 13-18 | 1.16 |
| | 20-25 | 0.61 |
| | 31-43 | 0.71 |
| | 46-58 | 0.80 |
| **AR55 in DPC micelles** | 14-19 | 0.70 |
| | 20-25 | 0.66 |
| | 29-46 | 1.33 |
| | 47-57 | 0.99 |
| **AR55 in LPPG micelles** | 22-25 | 0.30 |
| | 27-45 | 1.16 |
| | 46-58 | 1.20 |
| **AR55 in 50% HFIP** | 13-27 | 0.71 |
| | 30-41 | 0.33 |
| | 42-59 | 0.56 |

**Table S4.** Translational diffusion coefficient ($D_C$) and hydrodynamic diameter ($d_H$) of pure detergent micelles and AR55-micelle complexes.

| | δ range (ppm) | SDS | SDS+AR55 | DPC | DPC+AR55 | LPPG | LPPG+AR55 |
|---|---|---|---|---|---|---|---|
| $D_C$ ($10^{-10}$ $m^2/s$) | 1.16-1.34[A] | 1.33± 0.02 | 1.14 ± 0.02 | 1.32 ± 0.02 | 1.09 ± 0.05 | 0.76±0.01 | 0.69±0.01 |
| | 6.40-8.93[B] | X | 0.89 ± 0.07 | X | 0.94±0.05 | X | 0.69±0.07 |
| $d_H$ (nm) | 1.16-1.34[A] | 4.88 ± 0.08 | 5.72 ± 0.08 | 4.89 ±0.06 | 5.89±0.26 | 8.52±0.14 | 9.45±0.16 |
| | 6.40-8.93[B] | X | 7.27 ± 0.53 | X | 6.82 ±0.34 | X | 9.42 ±0.91 |

[A] The peak P1 (from 1.16 to 1.34 ppm) corresponds to methyl/methylene protons, so represents detergent (pure micelle sample) or convoluted micelle without and with protein incorporated (micelle + protein sample).

[B] The peak P2 (from 6.40 to 8.93 ppm) corresponds to protein amide protons, so represents exclusive AR55 signal (bound to micelle complex).

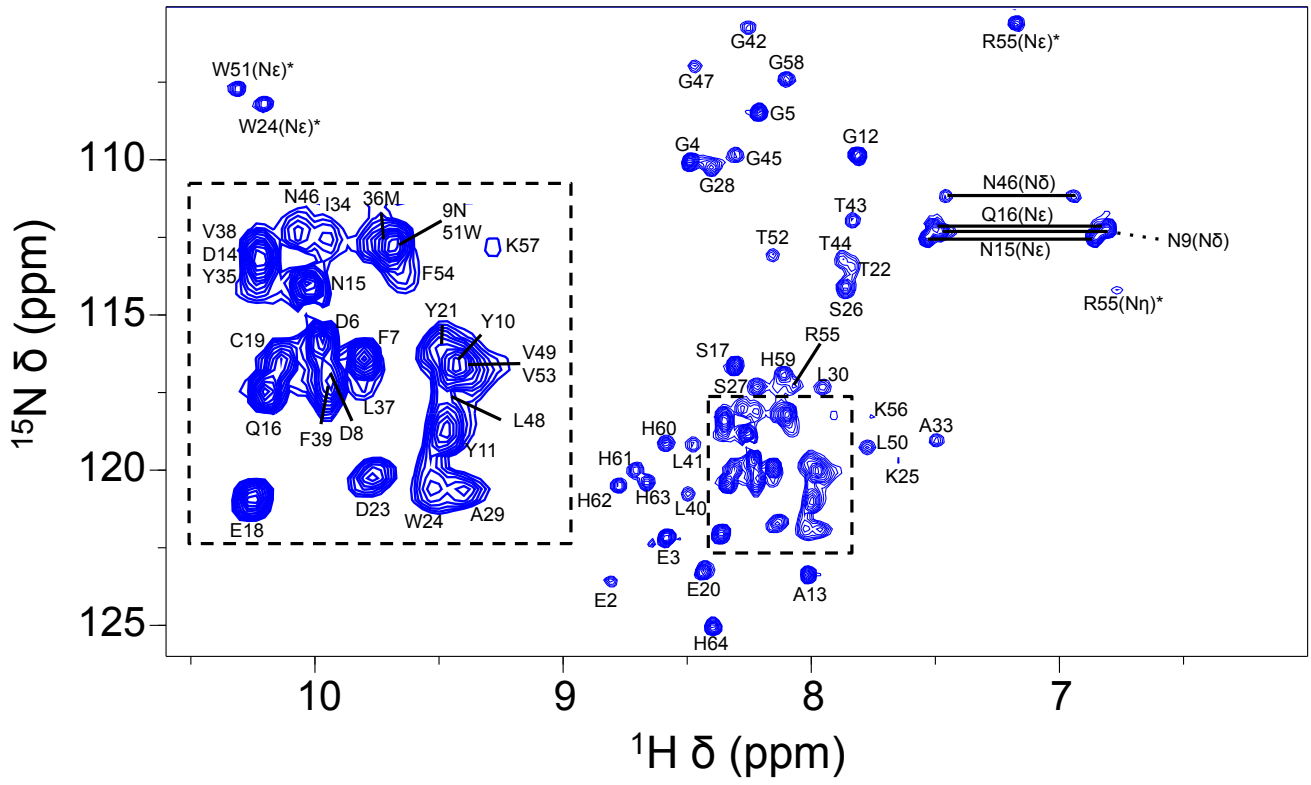


**Figure S1.** $^{1}$H-$^{15}$N HSQC for AR55 (acquired at 16.4 T and 37 °C) in LPPG micelles with assignments annotated. Note that sidechain cross-peaks denoted with an asterisk are aliased (un-aliased frequencies reported in the BMRB).

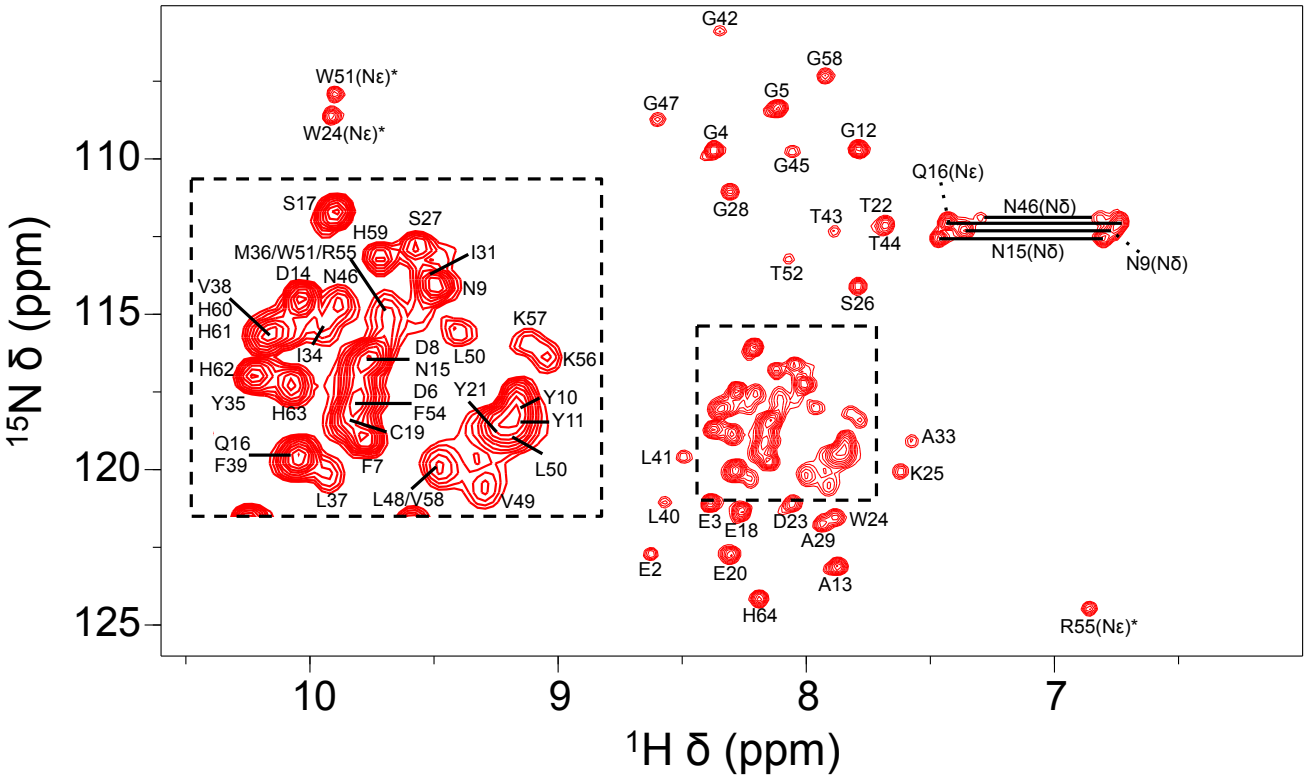


**Figure S2.** $^{1}$H-$^{15}$N HSQC for AR55 (acquired at 16.4 T and 37 °C) in SDS micelles with assignments annotated. Note that sidechain cross-peaks denoted with an asterisk are aliased (un-aliased frequencies reported in the BMRB).

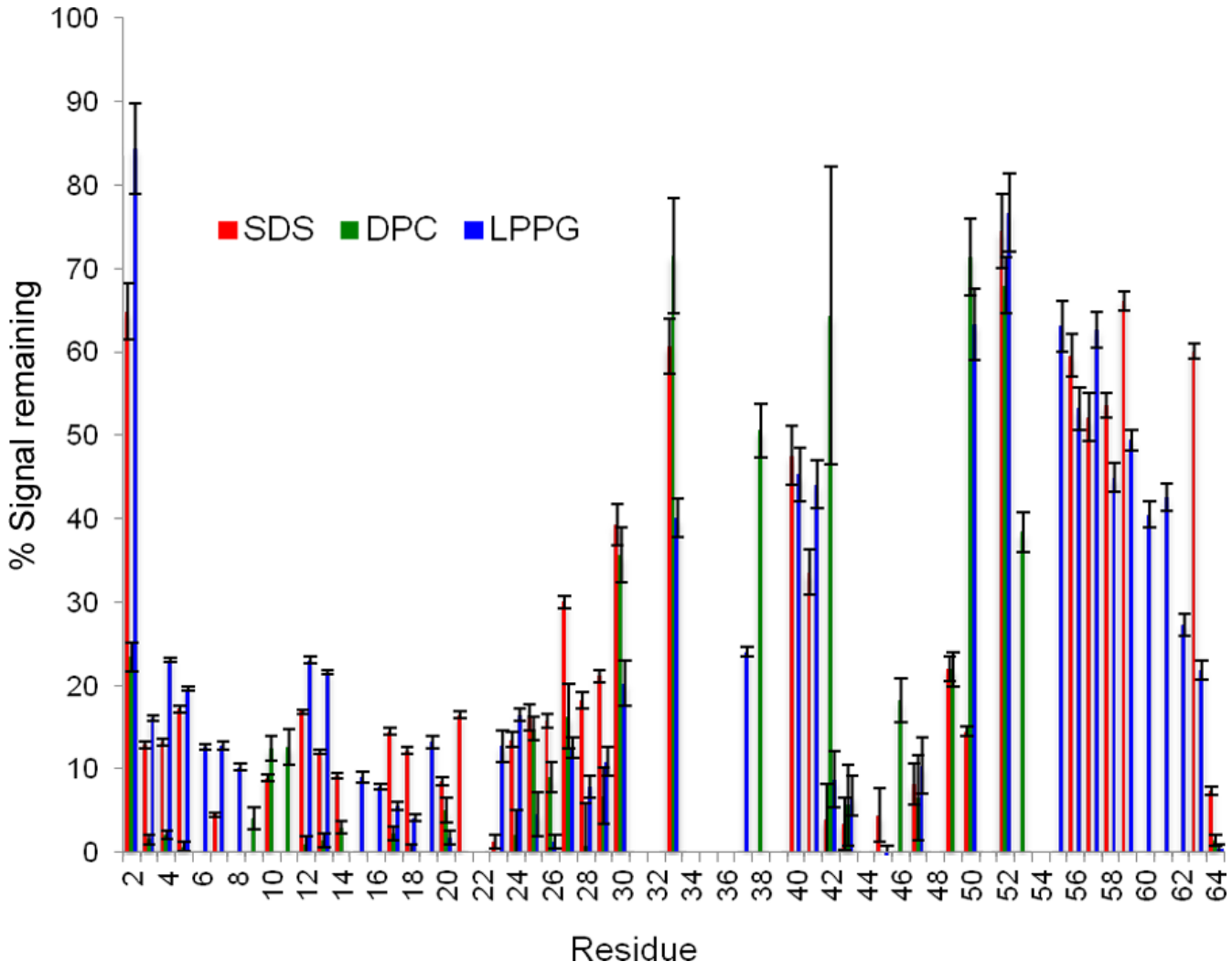


**Figure S3.** The percentage signal intensity remaining for selected signals after addition of $Mn^{2+}$ to the indicated NMR sample. Data is shown for AR55 solubilized in SDS with 150 μM $Mn^{2+}$, DPC with 1 mM $Mn^{2+}$ and LPPG with 1 mM $Mn^{2+}$.

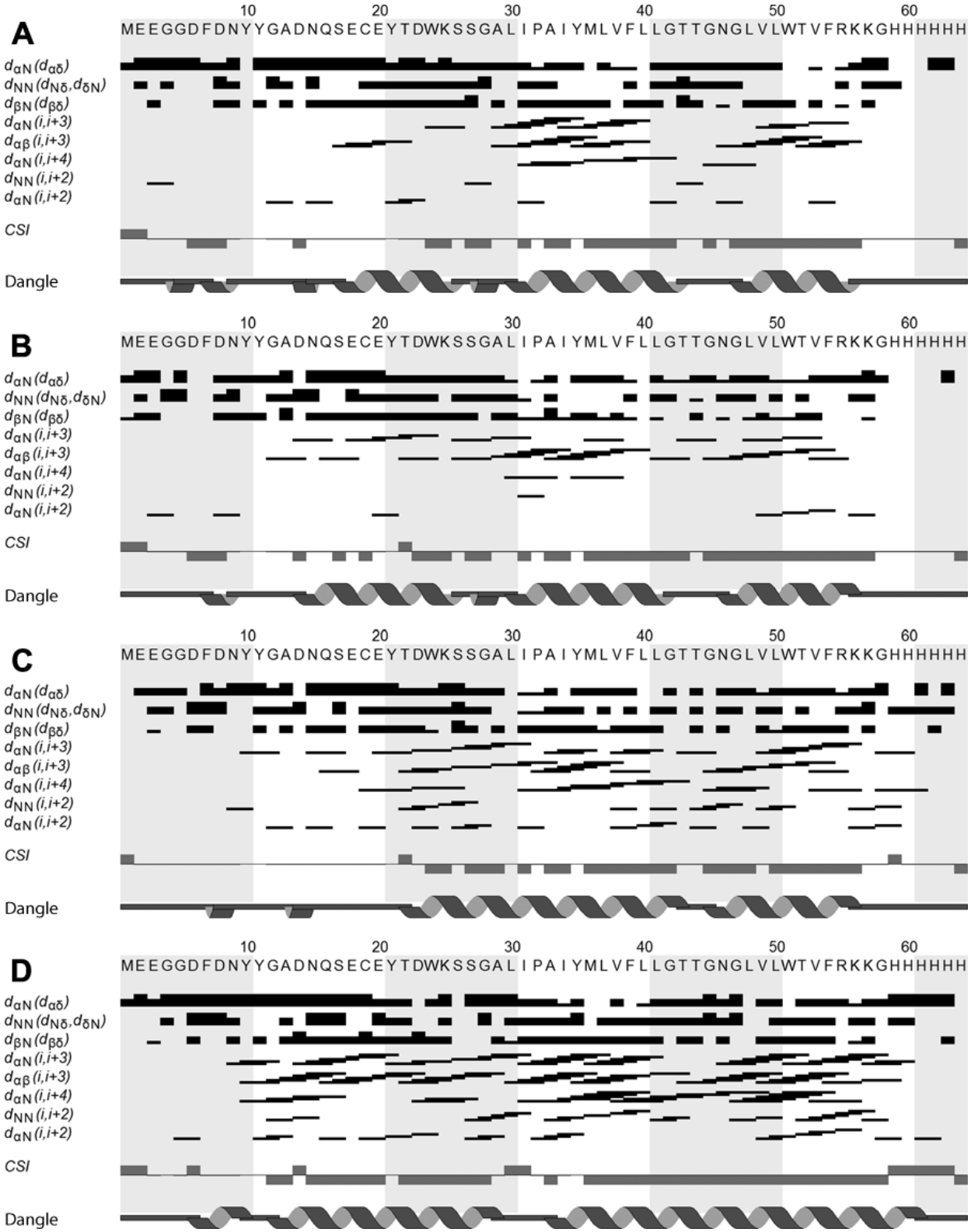


**Figure S4.** Through-space contacts observed for AR55 solubilized in (A) SDS micelles, (B) DPC micelles, (C) LPPG micelles and (D) 50% HFIP. Values of CSI[32] as well as secondary structure predictions using the DANGLE algorithm[33] are shown. This figure was generated using CCPNMR Analysis.[31]

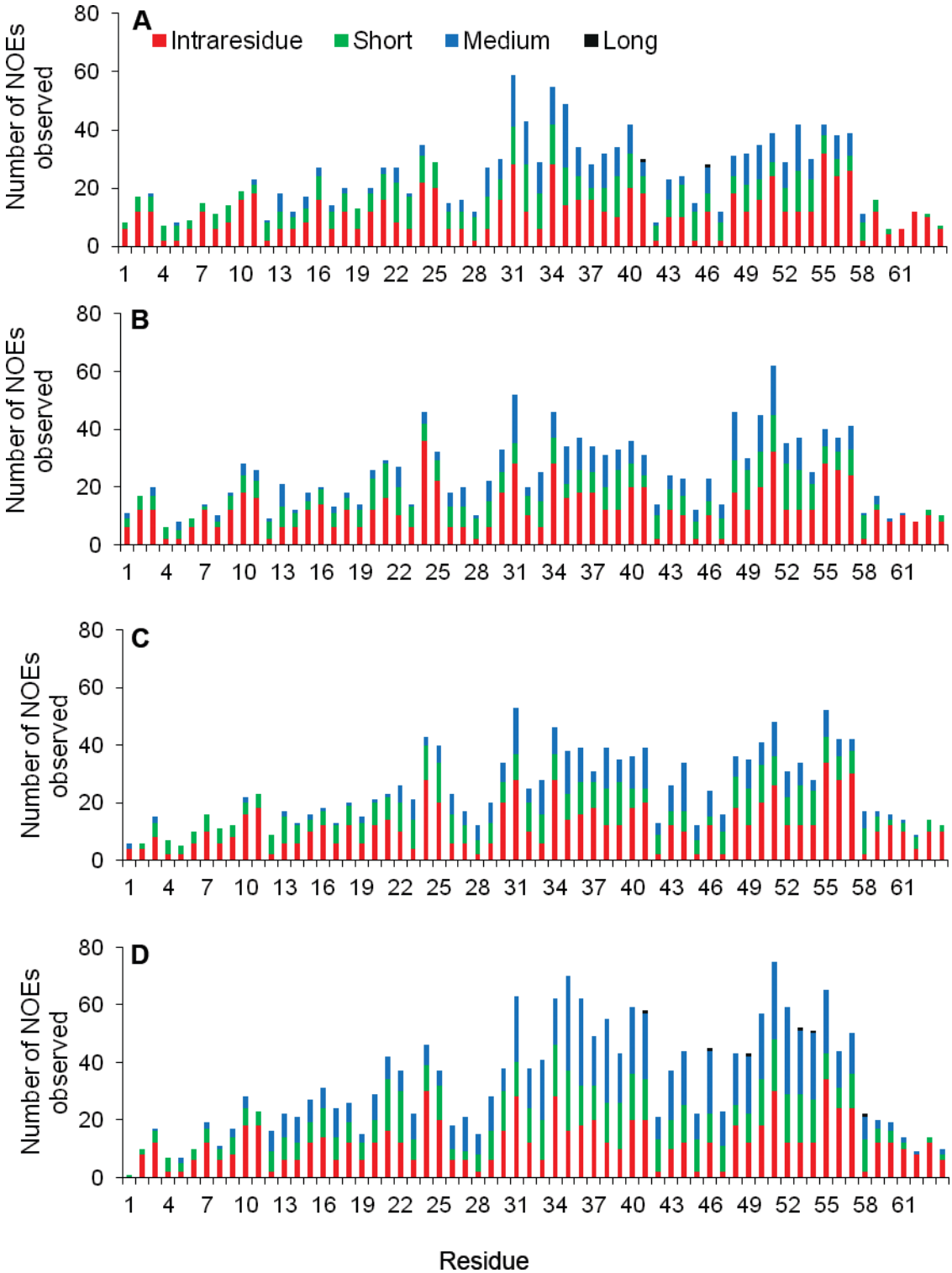


**Figure S5.** A breakdown of per-residue NOE contacts used to calculate the final ensembles of AR55 structures. Intraresidue, short (sequential), medium ($|i - j| \leq 4$) and long range NOE interactions shown for AR55 in (A) SDS, (B) DPC (previously published;[1] shown for comparative purposes), (C) LPPG and (D) 50% HFIP. Note that these plots do not include the ambiguous NOE contacts listed in Table S2.

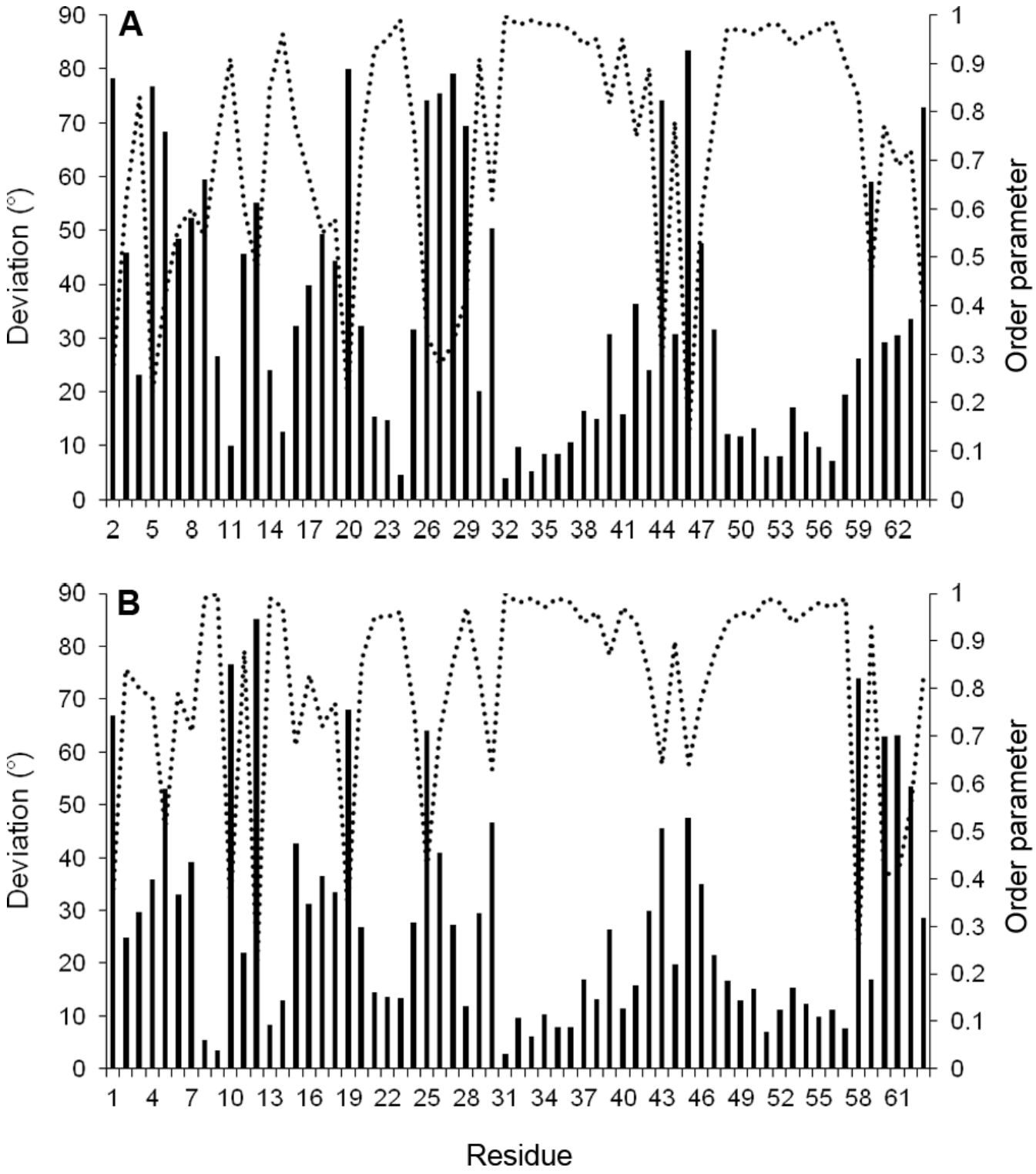


**Figure S6.** An overview of the convergence of ϕ (A) and φ (B) dihedral angles of the 40 members of the AR55 structural ensemble in SDS micelles. Both the deviation (bars) and order parameter (line) are indicated.

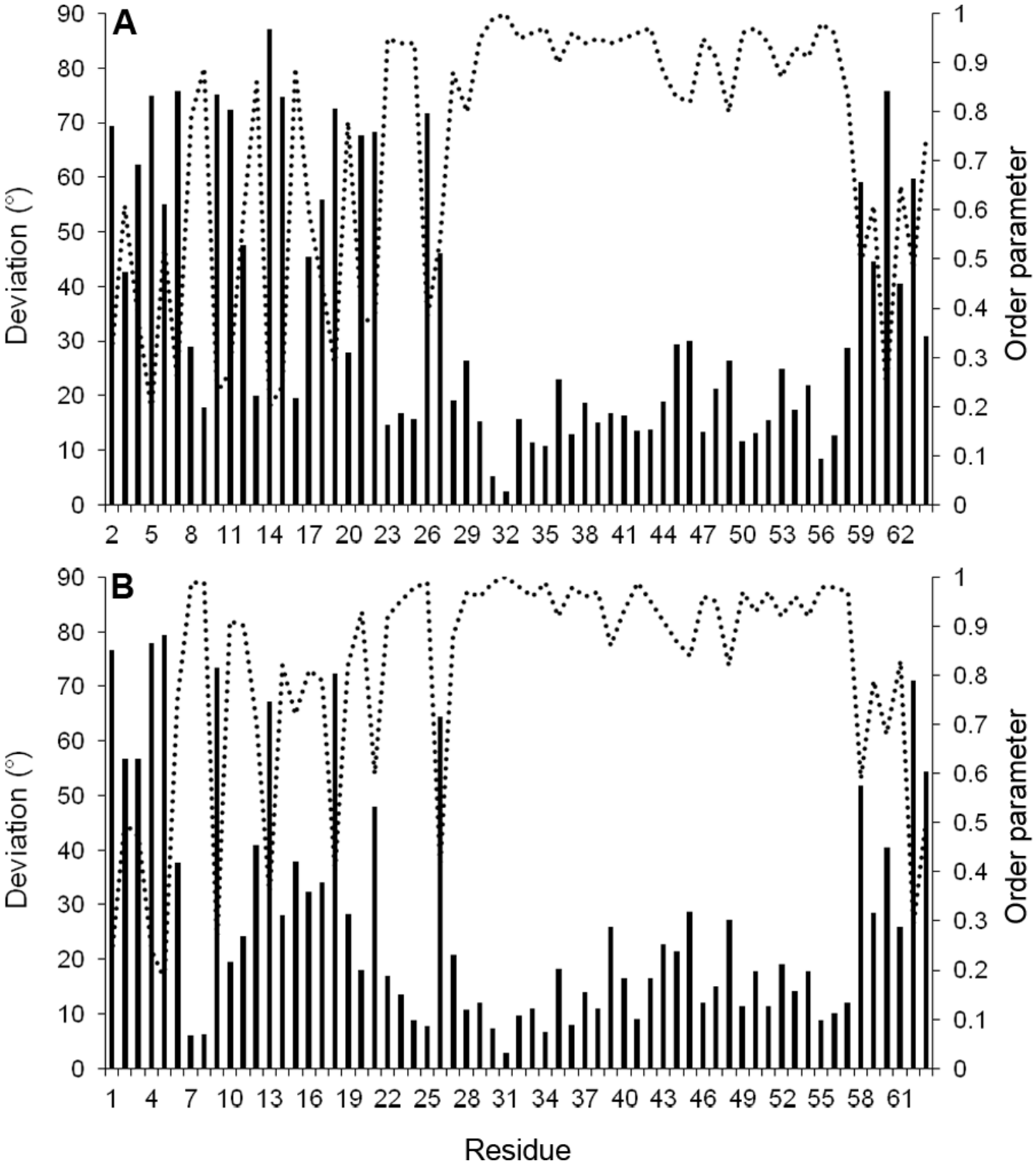


**Figure S7.** An overview of the convergence of ϕ (A) and φ (B) dihedral angles of the 40 members of the AR55 structural ensemble in LPPG micelles. Both the deviation (bars) and order parameter (line) are indicated.

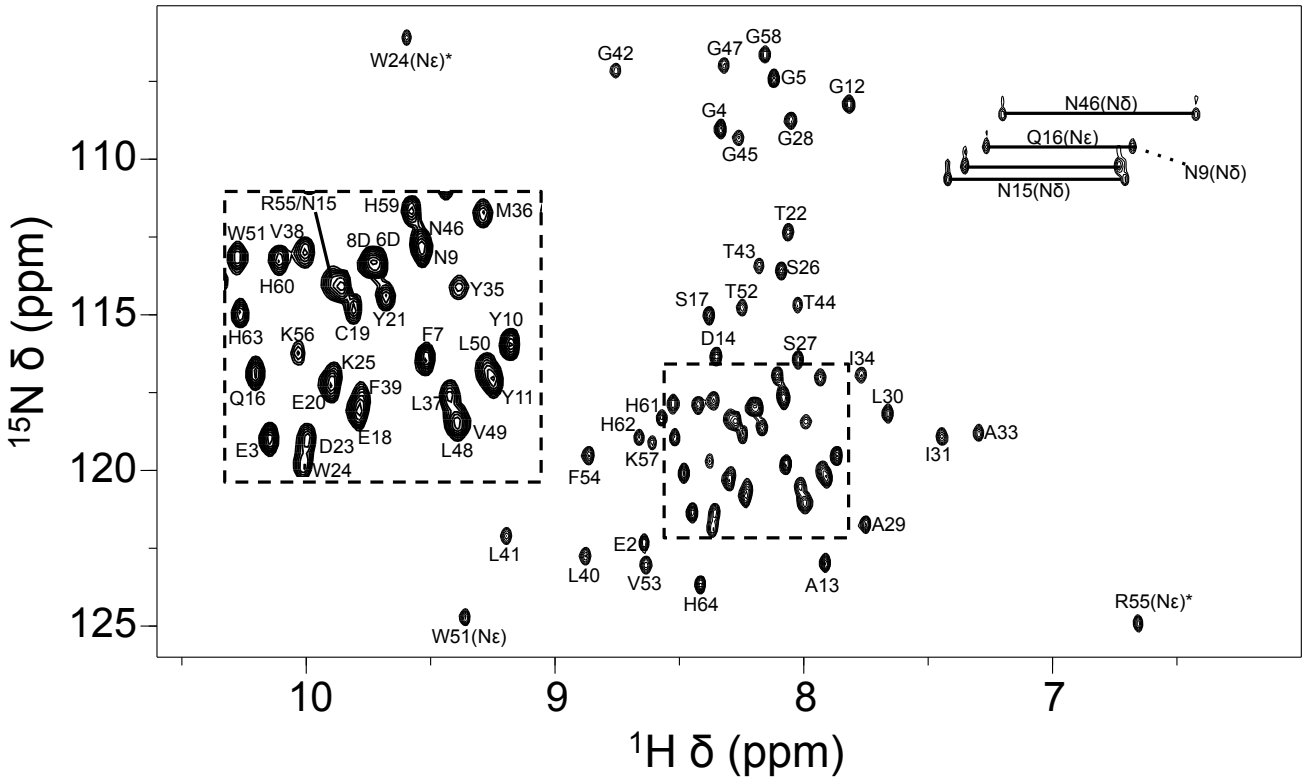


**Figure S8.** $^{1}H$-$^{15}N$ HSQC for AR55 (acquired at 16.4 T and 37 °C) in 50% HFIP/$H_2O$ with assignments annotated. Note that sidechain cross-peaks denoted with an asterisk are aliased (un-aliased frequencies reported in the BMRB).

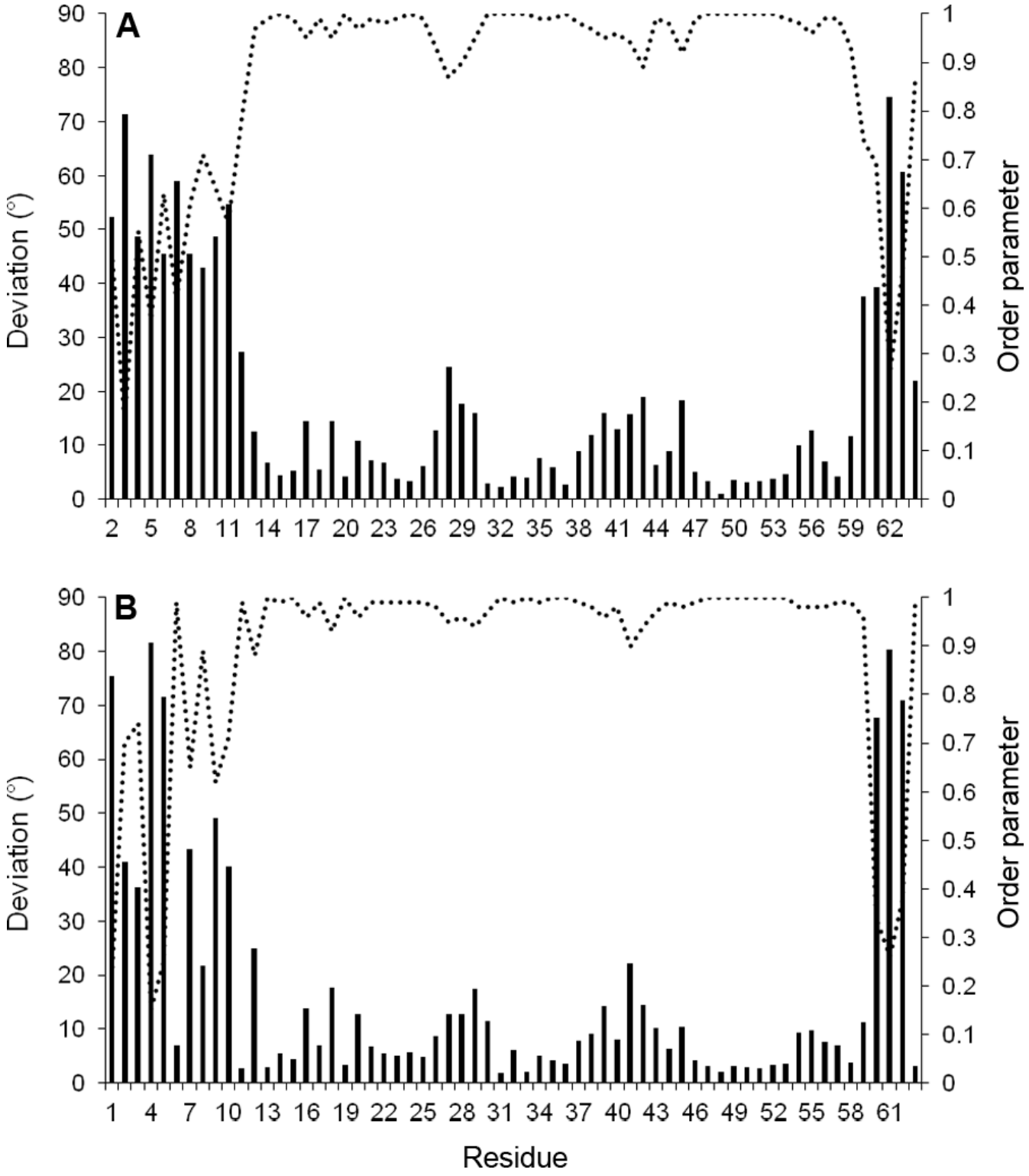


**Figure S9.** An overview of the convergence of ϕ (A) and φ (B) dihedral angles of the 40 members of the AR55 structural ensemble in 50% HFIP. Both the deviation (bars) and order parameter (line) are indicated.

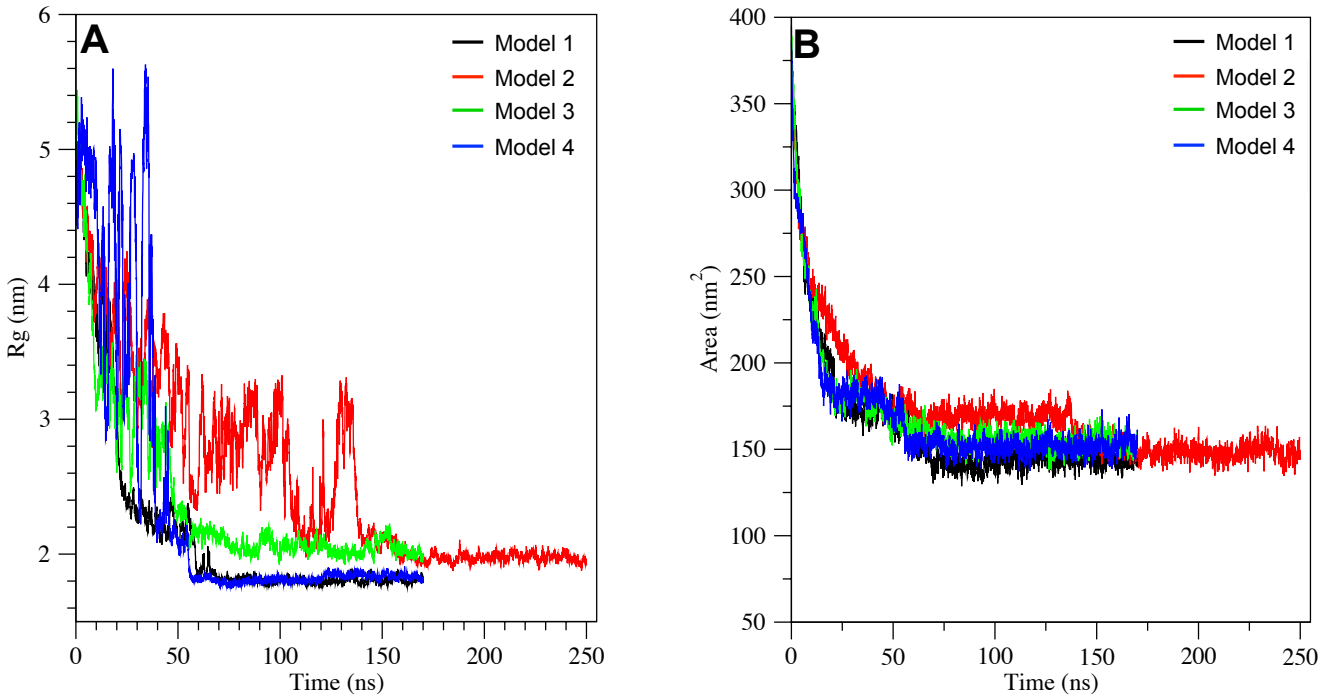


**Figure S10.** (A) Radius of gyration ($R_g$) for DPC molecules and (B) solvent accessible surface area for protein and DPC molecules tracked as a function of simulation time for each of the four replicate MD simulations performed using the indicated model from the DPC structural ensemble (PDB entry 2LOU) as a starting structure.

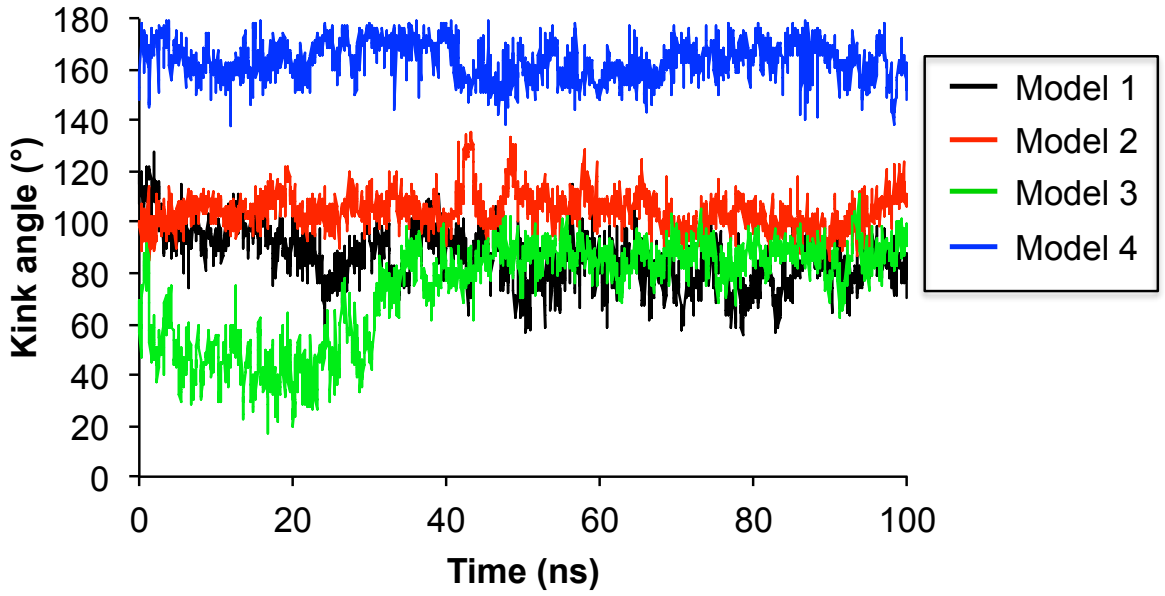


**Figure S11.** AR55 TM1 kink angle, calculated using MC-HELAN,[34] as a function of simulation time tracked as a function of simulation time for each of the four replicate MD simulations performed using the indicated model from the DPC structural ensemble (PDB entry 2LOU) as a starting structure. Simulation times correspond to the final 100 ns of each simulation (i.e., following convergence (Figure S10)). Average and standard deviations for each simulation are - model 1: 86° ± 11°; model 2: 104° ± 7°; model 3: 74° ± 20°; model 4: 164° ± 7°. For comparison, the initial kink angle for each NMR-derived ensemble member from PDB entry 2LOU was – model 1: 43°; model 2: 96°; model 3: 7°; model 4: 60°.

**REFERENCES**


1. Langelaan, D. N.; Reddy, T.; Banks, A. W.; Dellaire, G.; Dupre, D. J.; Rainey, J. K., Structural Features of the Apelin Receptor N-Terminal Tail and First Transmembrane Segment Implicated in Ligand Binding and Receptor Trafficking. *Biochim Biophys Acta* **2013,** *1828*, 1471-1483.

2. Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M., The Xplor-NIH NMR Molecular Structure Determination Package. *J Magn Reson* **2003,** *160*, 65-73.

3. Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M., The Xplor-NIH NMR Molecular Structure Determination Package. *Prog NMR Spectroscopy* **2006,** *48*, 47-62.

4. Laskowski, R. A.; Rullmannn, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M., AQUA and PROCHECK-NMR: Programs for Checking the Quality of Protein Structures Solved by NMR. *J Biomol NMR* **1996,** *8*, 477-486.

5. Rainey, J. K.; Fliegel, L.; Sykes, B. D., Strategies for Dealing with Conformational Sampling in Structural Calculations of Flexible or Kinked Transmembrane Peptides. *Biochem Cell Biol* **2006,** *84*, 918-929.

6. Berman, H.; Henrick, K.; Nakamura, H., Announcing the Worldwide Protein Data Bank. *Nat Struct Biol* **2003,** *10*, 980.

7. Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z. et al., Biomagresbank. *Nucleic Acids Res* **2008,** *36*, D402-408.

8. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J Biomol NMR* **1995,** *6*, 277-293.

9. Spyracopoulos, L., A Suite of Mathematica Notebooks for the Analysis of Protein Main Chain $^{15}$N NMR Relaxation Data. *J Biomol NMR* **2006,** *36*, 215-224.

10. Farrow, N. A.; Zhang, O.; Szabo, A.; Torchia, D. A.; Kay, L. E., Spectral Density Function Mapping Using $^{15}$N Relaxation Data Exclusively. *J Biomol NMR* **1995,** *6*, 153-162.

11. Morris, K. F.; Johnson, C. S., Diffusion-Ordered 2-Dimensional Nuclear Magnetic Resonance Spectroscopy. *J Am Chem Soc* **1992,** *114*, 3139-3141.

12. Wu, D. H.; Chen, A. D.; Johnson, C. S., An Improved Diffusion-Ordered Spectroscopy Experiment Incorporating Bipolar-Gradient Pulses. *J Magn Reson A* **1995,** *115*, 260-264.

13. Stejskal, E. O.; Tanner, J. E., Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J Chem Phys* **1965,** *42*, 288-292.

14. Cantor, C. R.; Schimmel, P. R., *Techniques for the Study of Biological Structure and Function*. W. H. Freeman: San Francisco, 1980.

15. Berendsen, H. J. C.; Vanderspoel, D.; Vandrunen, R., GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput Phys Commun* **1995,** *91*, 43-56.

16. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* **2008,** *4*, 435-447.

17. Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F., Definition and Testing of the GROMOS Force-Field Versions 54A7 and 54B7. *Eur Biophys J* **2011,** *40*, 843-856.

18. Hermans, J.; Berendsen, H. J. C.; Vangunsteren, W. F.; Postma, J. P. M., A Consistent Empirical Potential for Water-Protein Interactions. *Biopolymers* **1984,** *23*, 1513-1518.

19. Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E., An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J Chem Theory Comput* **2011,** *7*, 4026-4037.

20. Tieleman, D. P.; van der Spoel, D.; Berendsen, H. J. C., Molecular Dynamics Simulations of Dodecylphosphocholine Micelles at Three Different Aggregate Sizes: Micellar Structure and Chain Relaxation. *J Phys Chem B* **2000,** *104*, 6380-6388.

21. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* **1993,** *98*, 10089-10092.

22. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J Chem Phys* **1995,** *103*, 8577-8593.

23. Pall, S.; Hess, B., A Flexible Algorithm for Calculating Pair Interactions on SIMD Architectures. *Comput Phys Commun* **2013,** *184*, 2641-2650.

24. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A Linear Constraint Solver for Molecular Simulations. *J Comput Chem* **1997,** *18*, 1463-1472.

25. Miyamoto, S.; Kollman, P. A., Settle: An Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *J Comput Chem* **1992,** *13*, 952-962.

26. Bussi, G.; Donadio, D.; Parrinello, M., Canonical Sampling through Velocity Rescaling. *J Chem Phys* **2007,** *126*.

27. Nose, S.; Klein, M. L., Constant Pressure Molecular Dynamics for Molecular Systems. *Mol Phys* **1983,** *50*, 1055-1076.

28. Parrinello, M.; Rahman, A., Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J Appl Phys* **1981,** *52*, 7182-7190.

29. Hyberts, S. G.; Goldberg, M. S.; Havel, T. F.; Wagner, G., The Solution Structure of Eglin C Based on Measurements of Many Noes and Coupling Constants and Its Comparison with X-Ray Structures. *Protein Sci* **1992,** *1*, 736-751.

30. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual Molecular Dynamics. *J Mol Graph* **1996,** *14*, 33-38, 27-38.

31. Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D., The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins* **2005,** *59*, 687-696.

32. Wishart, D. S.; Sykes, B. D., The $^{13}C$ Chemical-Shift Index: A Simple Method for the Identification of Protein Secondary Structure Using $^{13}C$ Chemical-Shift Data. *J Biomol NMR* **1994,** *4*, 171-180.

33. Cheung, M. S.; Maguire, M. L.; Stevens, T. J.; Broadhurst, R. W., Dangle: A Bayesian Inferential Method for Predicting Protein Backbone Dihedral Angles and Secondary Structure. *J Magn Reson* **2010,** *202*, 223-233.

34. Langelaan, D. N.; Wieczorek, M.; Blouin, C.; Rainey, J. K., Improved Helix and Kink Characterization in Membrane Proteins Allows Evaluation of Kink Sequence Predictors. *J Chem Inf Model* **2010,** *50*, 2213-2220.