

The Native GCN4 Leucine-Zipper Domain Does Not Uniquely Specify a Dimeric Oligomerization State

Kaylyn M. Oshaben,[†] Reza Salari,[†] Darrell R. McCaslin,[‡] Lillian T. Chong,^{*,†} and W. Seth Horne^{*,†}

[†]*Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260*

[‡]*Department of Biochemistry, Biophysics Instrumentation Facility, University of Wisconsin – Madison, Madison, WI 53706*

SUPPORTING INFORMATION

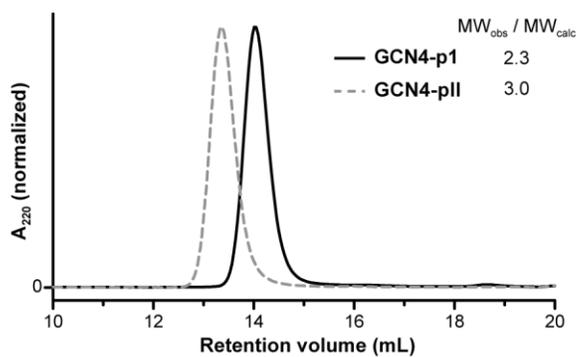


Figure S1. Gel-permeation chromatograms for GCN4-pI and GCN4-pII. Injections consisted of 100 μ M peptide and were eluted with 0.05 M phosphate, 0.15 M NaCl, pH 7. The ratio of molecular weight as determined by calibration of the column (MW_{obs}) to the molecular weight of monomeric peptide (MW_{calc}) is shown.

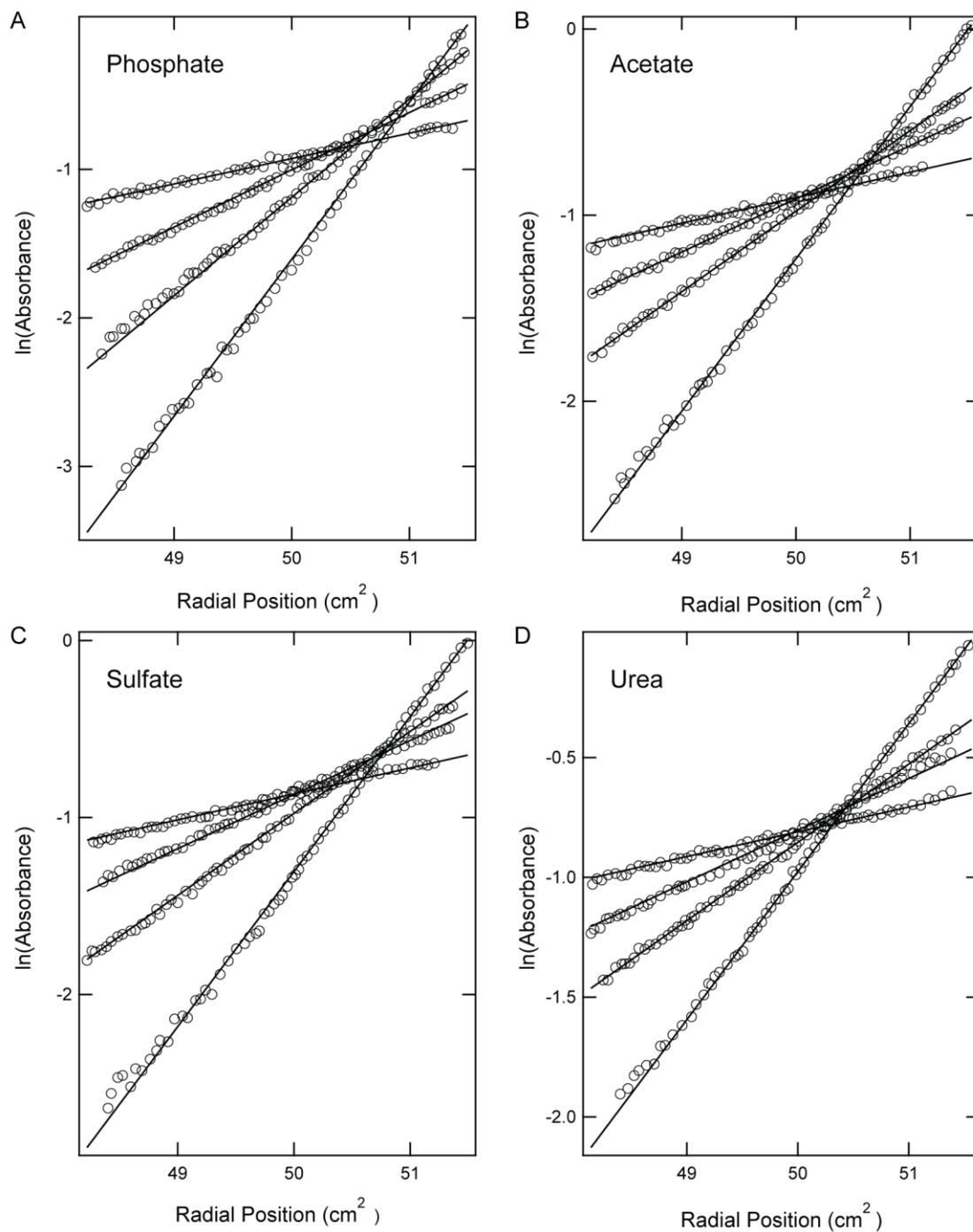


Figure S2. Analytical ultracentrifugation data for GCN4-p1 in four different buffers: (A) 0.05 M phosphate, 0.15 M NaCl, pH 7.0; (B) 0.1 M sodium acetate, 0.1 M sodium citrate tribasic, pH 5.3; (C) 0.2 M ammonium sulfate, 0.1 M MES, pH 6.6; (D) 10 mM phosphate, 6 M urea, pH 7.0. The points have been corrected for the non-sedimenting absorbance from the fitted model. The lines are the resulting best fit as a single macromolecular species from results shown in Table 2.

Sedimentation Equilibrium Data Analysis. Gradients were globally fit to a model which included a single macromolecular species and a small contribution from a non-sedimenting component (Figure S2). For panel A the analysis included speeds of 11000, 14000, 19300, 22000, 29000, 38000, 48000 at three initial loading concentrations (The data shown are for the loading concentration of 220 μM and speeds 19300, 29000, 38000 and 48000 rpm; data at other loading concentrations and speeds in this buffer are also linear). This experiment was designed to look for evidence of multiple macromolecular species, but none was found. Conditions in panels B, C and D were carried out to examine the effects of other buffer conditions similar to those employed in crystallization and spectroscopic studies on the association state. These were run at the same time with each buffer using a single loading concentration of peptide (between 245 and 267 μM). Data were collected at speeds of 18000, 26000, 32000 and 44000 rpm. At speeds greater than 50000 rpm, as yet unexplained spectral anomalies were observed at high radial positions; therefore we restricted the speeds for analysis to those 48000 and below; we note that even at higher speeds data at smaller radial positions were consistent with the finding from the global analysis at lower speeds.

In our global analysis the reduced molecular weight served as the fitting variable. The reduced molecular weight is $M(1 - \bar{v}\rho)$ where M is the weight average molecular weight, \bar{v} is the partial specific volume and ρ is the solvent density; this approach permits the impact of errors in the partial specific volume and density on M to be evaluated after the fitting process. Inclusion of a reduced molecular weight for a second species did not improve fits, and generally assuming the same partial specific volume yielded a second molecular weight not significantly different from the other component. In Table 2 the weight average molecular weights for each data set were computed from the reduced molecular weight using partial specific volumes and densities as indicated in the text.

The small molecular weight of the peptide makes it impossible even at maximal attainable speeds to fully deplete all of the peptide mass from the solution column; so the presence of non-sedimenting absorbances is not directly addressable by experiment. Such an absorbance can arise from imperfections in the cell windows, small molecule contaminants (TFA) or even slight mismatches between the sample prepared simply by dilution and the fresh buffer used as the reference. Data obtained at 60000 rpm near the solution meniscus were higher absorbance than expected even for a system composed entirely of monomer; moreover spectra recorded near the meniscus were distinct from those of the peptide recorded initially. Therefore a non-sedimenting component was included in the fitting process. In all cases the fitted value was less than the value observed at 60000 rpm near the meniscus but greater than zero and varied among the samples.

Figure S2, shows a plot of the logarithm of measured absorbance (minus the fitted non-sedimenting contribution) versus the squared distance for the center of rotation. In this presentation gradients composed of a single species are transformed into a series of straight lines whose slopes depend on speed and are also proportional to the weight average molecular weight. The lines shown are based on the results of the global fitting of data to the single species model. Clearly, with available data the single species model is adequate under all conditions and over the range of concentrations accessible in the centrifugation experiment. It should be noted that the data shown in panel A is only a subset of that collected, all of which are described by the single species in Table 2.

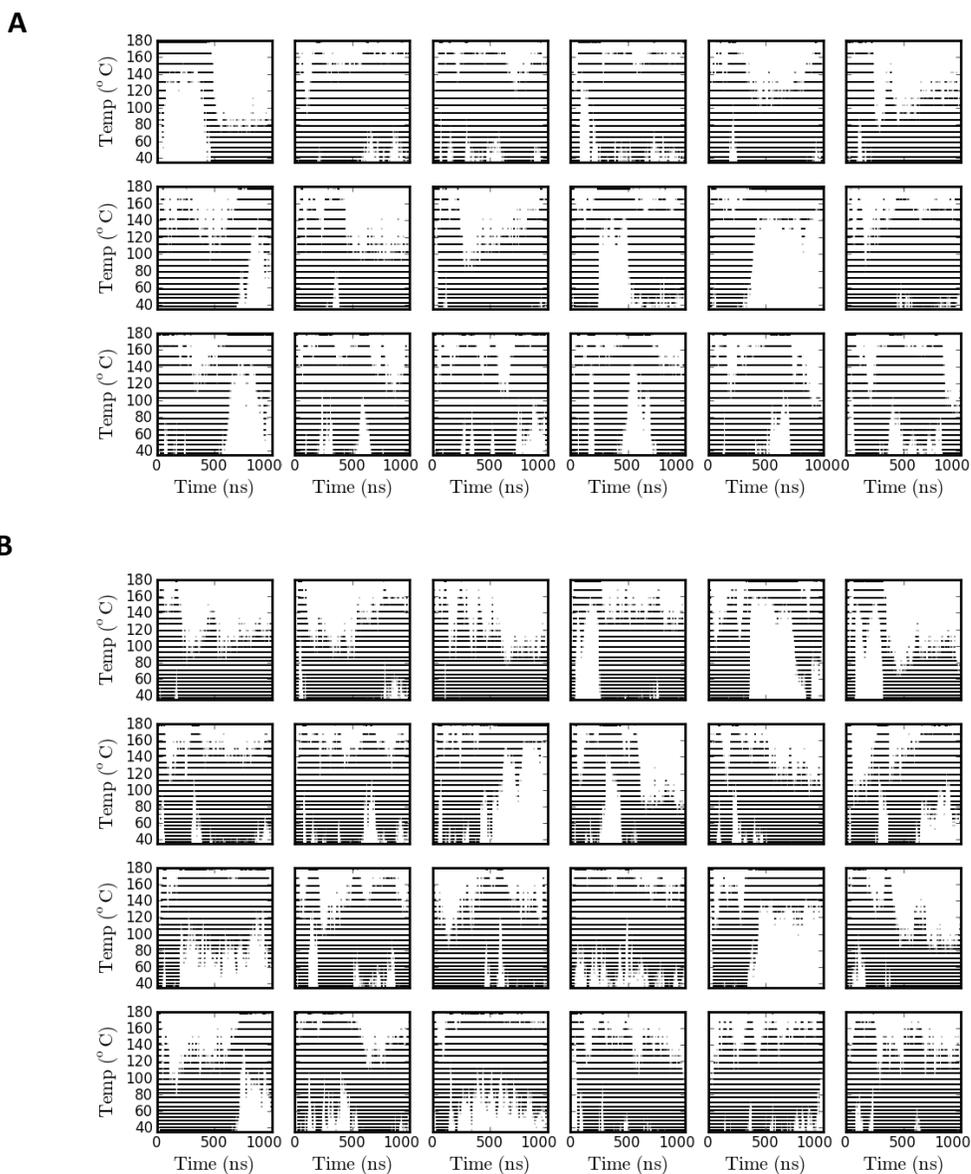


Figure S3. Convergence of parallel tempering MD simulations based on all replicas visiting most of the temperatures for the (A) dimer and (B) trimer. Each subplot represents one replica's visits of the different temperatures between 37 and 177 °C.

Table S1. Convergence of parallel tempering MD simulations based on the average number of round trips and the end-to-end transition times (τ_{end}) for the dimer and trimer.

	Average number of round trips per replica ^a	τ_{end} (ns)
Dimer	10.2 ± 4.0	98.0 ± 38.4
Trimer	5.6 ± 2.1	178.6 ± 70.0

^a Uncertainties in the number of round trips represent one standard deviation.

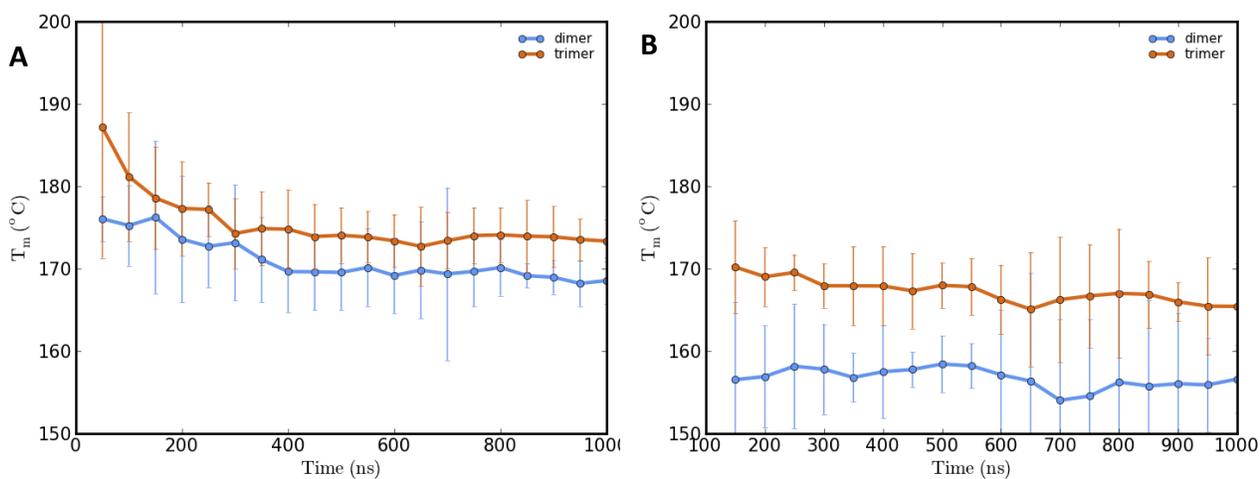


Figure S4. Convergence of melting temperatures of the dimer and trimer as estimated from parallel tempering MD simulations. A) Estimated T_m values as a function of simulation length. The melting temperatures change significantly during the first 100 ns. B) Estimated T_m values as a function of simulation length starting from 100 ns. The T_m values appear converged over various simulation lengths of 100-1000 ns, indicating at least 100 ns was required for convergence. Melting temperatures were estimated based on helical content (conformations were considered folded if the number of helical residues were more than 20 for the dimer or 30 for the trimer). Error bars represent one standard deviation in the estimated melting temperatures of five consecutive blocks in the corresponding time segment.

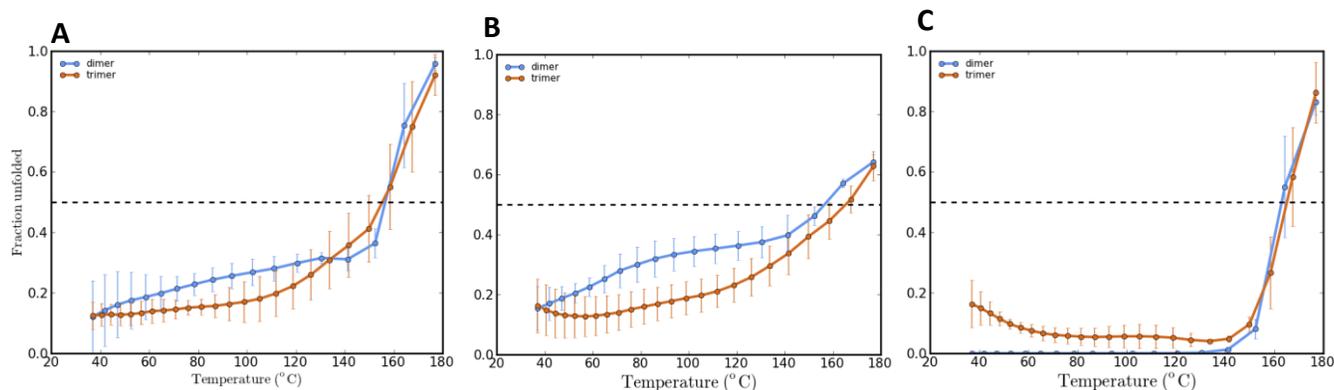


Figure S5. Fraction unfolded as a function of temperature based on parallel tempering MD simulations of the dimer and trimer forms, as monitored using three different order parameters. (A) C_{α} rmsd from the crystal structure. Conformations were considered unfolded if the C_{α} rmsd was more than one standard deviation higher than the average value at 37 °C. The average C_{α} rmsd at 37 °C relative to the crystal structure was 12.0 ± 1.0 Å for the dimer and 13.4 ± 2.4 Å for the trimer. (B) Fraction helicity. Conformations were considered unfolded if the number of helical residues was more than one standard deviation below the average at 37 °C. The average number of helical residues at 37 °C was 25.4 ± 4.9 and 38.6 ± 8.2 for the dimer and trimer, respectively. (C) Chain dissociation. Conformations were considered unfolded if at least one monomer was dissociated beyond 4.5 Å. Uncertainties represent one standard deviation in the average of the corresponding quantity calculated for five consecutive blocks of 180 ns in the simulations.

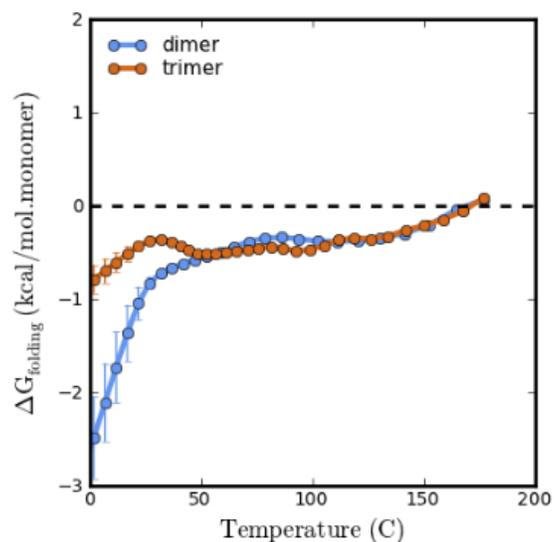


Figure S6. Folding free energies of the dimer and trimer computed from parallel tempering MD simulations at various temperatures, based on snapshots collected every 5 ns (total of 180 snapshots at each temperature). The folded state was defined based on helical content (see Methods). Uncertainties are provided for all data points and computed as described in Methods.

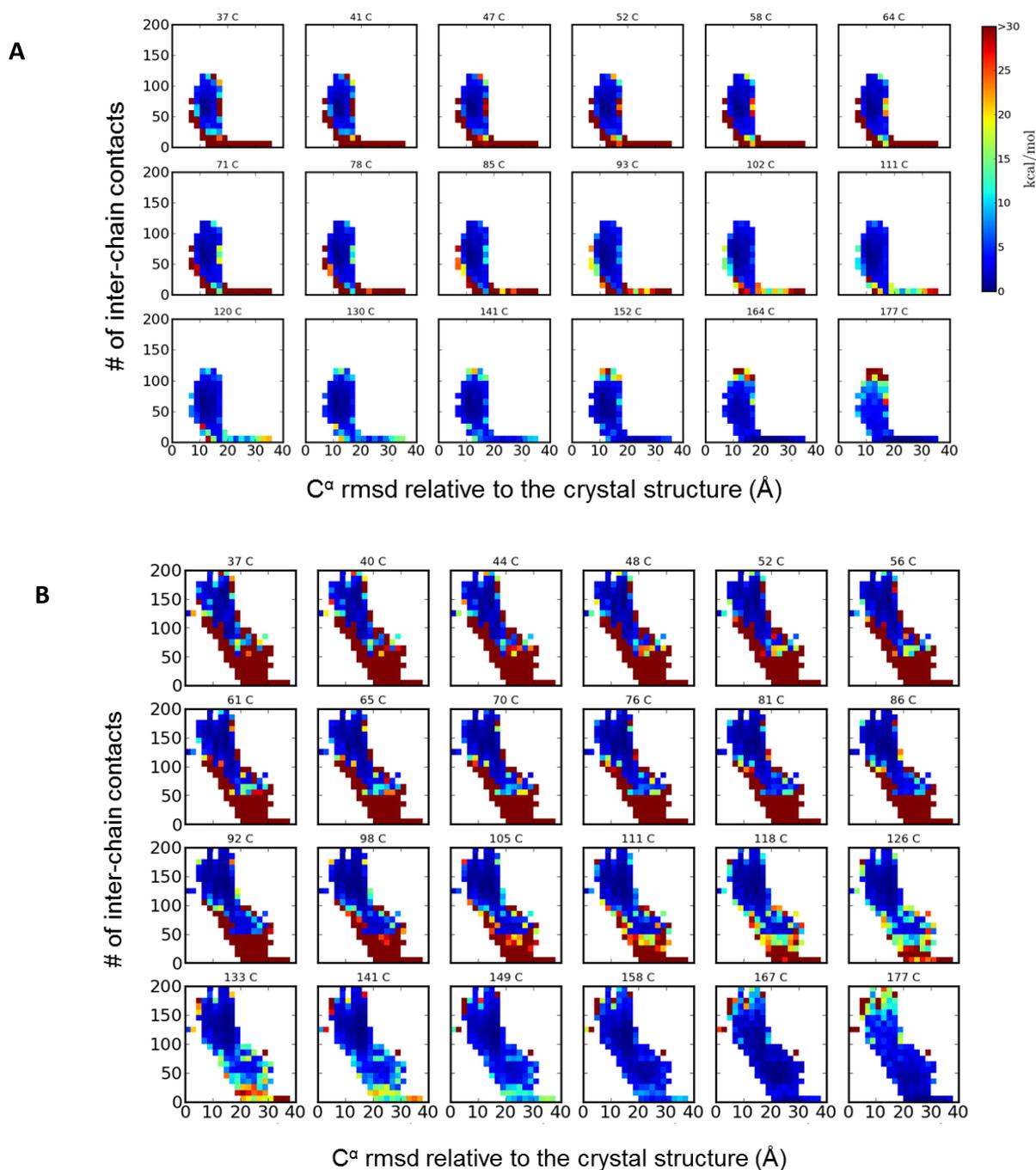


Figure S7. Comparison of the 2-D free energy profiles of the dimer (A) and trimer (B) as a function of C_{α} rmsd and the total number of inter-chain residue contacts at each replica temperature of the parallel tempering MD simulations. The color of each bin represents the corresponding free energy from blue (most favorable) to red (least favorable). The PMF plots were constructed using snapshots collected every 5 ns with bins of 2 Å width and 10 contacts height. Two residues were considered in contact if the distance between any two heavy atoms from the two residues was less than 5.5 Å.

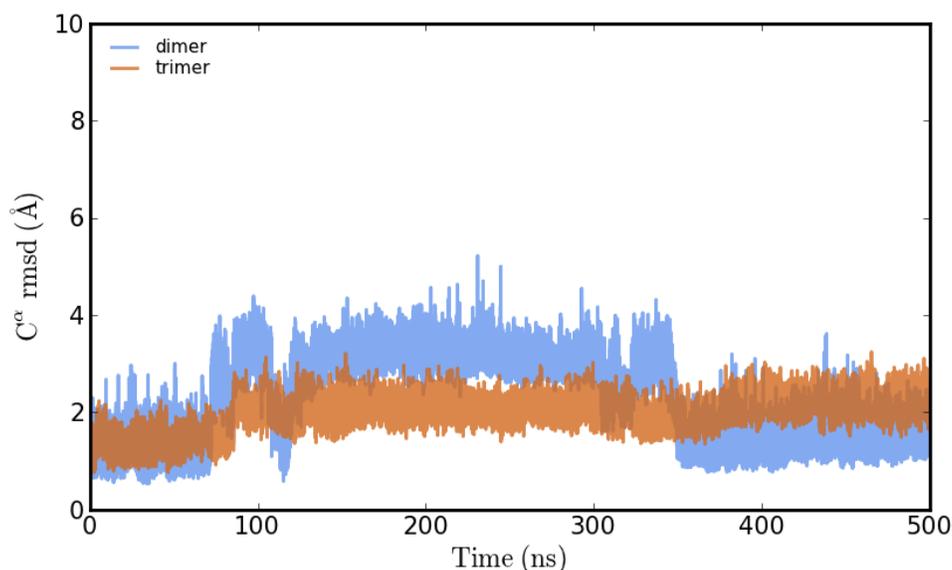


Figure S8. Reliability of the simulation model as gauged by low C_{α} rms deviations from the crystal structure of the dimer and trimer in standard MD simulations at 20 °C. The average C_{α} rmsd values are 2.3 ± 0.9 Å for the dimer and 2.0 ± 0.4 Å for the trimer. Uncertainties represent one standard deviation.

Table S2. Summary of a -position Asn residues in parallel, homotrimeric coiled coils in the PDB.^a

PDB ^b	Resolution (Å) ^c	Core Asn Residue #	Core Pocket Near Asn ^d	
			Volume (Å ²)	Content
1EBO	3.00	85	49	chloride
1WP7	2.20	155	50	water
1WT6	1.60	37	100	water
		51	34	water
2FYZ	2.20	133	50	water
		148	120	nothing
2IEQ	1.75	14	206	water
2POH	2.10	63	67	water
2W6B	2.80	608	35	nothing

^a Identified by a search of the CC+ database (ref. 76 from main text). ^b PDB accession code. ^c Refinement resolution of the crystal structure. ^d Volume of the core pocket in the vicinity of the a -position Asn and the contents of the pocket, if any, resolved in the crystal structure.