Supporting Information:

Unfolding Pathways of Hen Egg White Lysozyme in Ethanol

Alice R. Walker[a,‡], Nikhil Baddam[b], G. Andrés Cisneros[a,*]

[a]Department of Chemistry, University of North Texas, Denton, TX 76201, USA
[b]Department of Chemistry, Wayne State University, Detroit, MI 48202, USA

molecular dynamics, lysozyme, protein folding

Archives of the theoretical initial starting structures, as well as parameters for each system, are available online as additional supplementary information.

Table S1. Protonation populations for constant pH simulations of lysozyme in water at pH of 4.

Residue Number	State 0	State 1	State 2	State 3	State 4
Residue: GLU 7	**0.951353** (0)	0.019701 (1)	0.003840 (1)	0.020586 (1)	0.004520 (1)
Residue: ASP 18	**0.443645** (0)	**0.310606** (1)	0.000146 (1)	**0.245434** (1)	0.000168 (1)
Residue: GLU 35	0.024694 (0)	**0.435374** (1)	0.038023 (1)	**0.465220** (1)	0.036688 (1)
Residue: ASP 48	**0.981755** (0)	0.010435 (1)	0.000436 (1)	0.007142 (1)	0.000232 (1)
Residue: ASP 52	**0.951048** (0)	0.006752 (1)	0.000203 (1)	0.041740 (1)	0.000257 (1)
Residue: ASP 66	**0.984633** (0)	0.014564 (1)	0.000244 (1)	0.000559 (1)	0.000000 (1)
Residue: ASP 87	**0.989800** (0)	0.003767 (1)	0.000237 (1)	0.005928 (1)	0.000268 (1)
Residue: ASP 101	**0.293935** (0)	0.**152567** (1)	**0.140748** (1)	**0.357563** (1)	0.055186 (1)
Residue: ASP 119	**0.969930** (0)	0.008761 (1)	0.002594 (1)	0.014965 (1)	0.003750 (1)

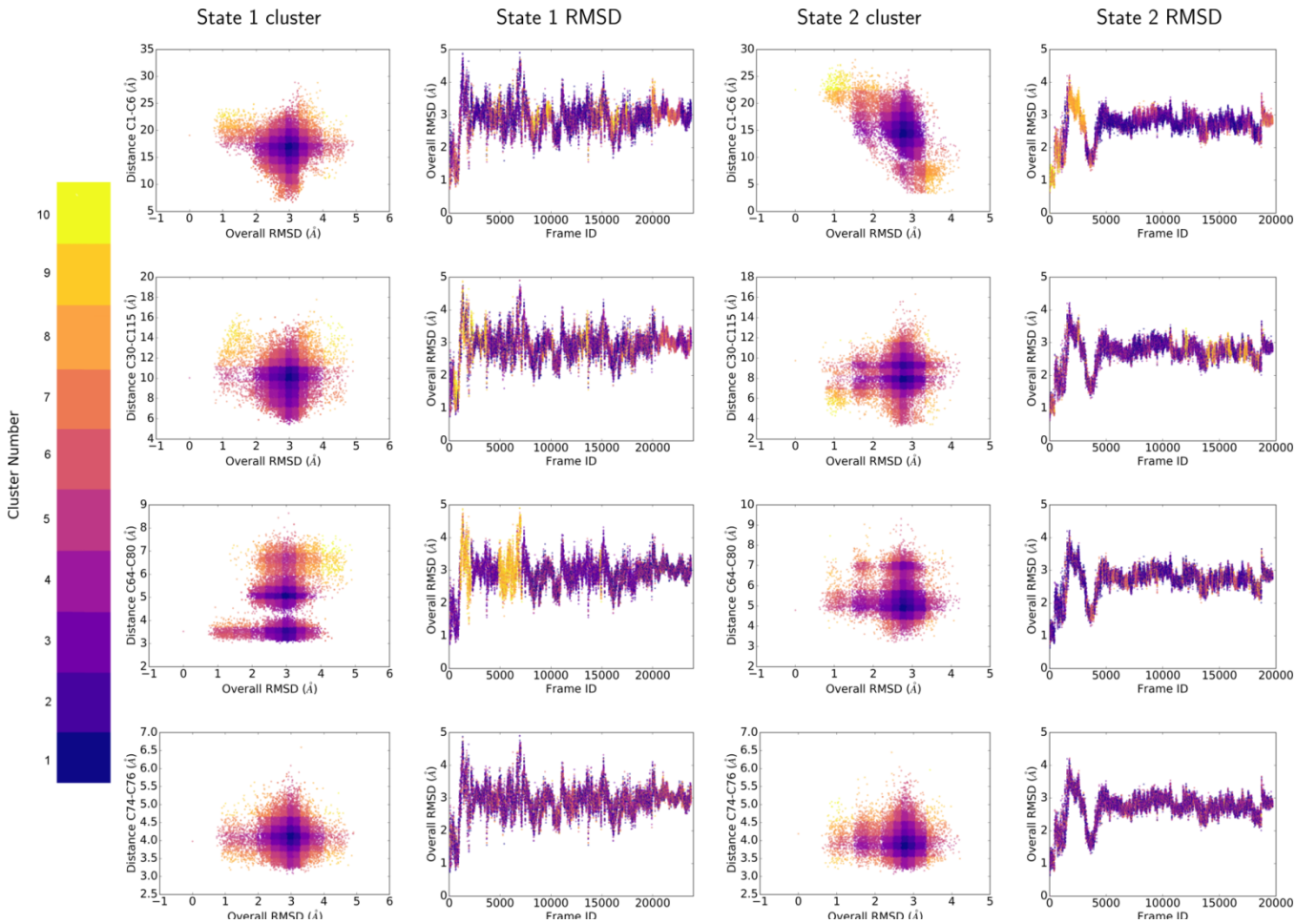
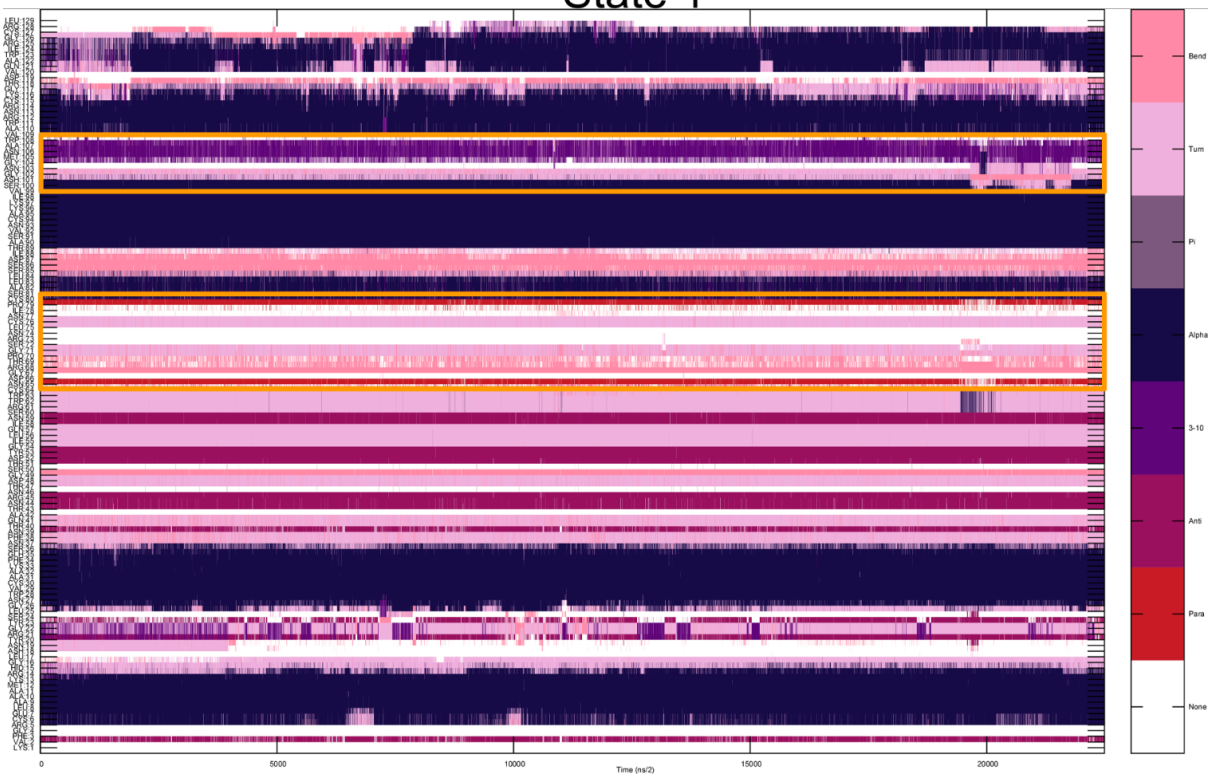


Figure S1. Representations of cluster analysis performed by sulfur-sulfur atomic distance for reduced disulfide bonds against overall backbone RMSD for the long timescale simulations of State 1 and State 2, ordered by average distance from largest to smallest. The first and third columns plot each respective set of distance clusters against overall RMSD, and are colored by their cluster percentage, with the most frequently occurring cluster as blue, medium as magenta-pink and lowest as yellow. The second and fourth columns show the overall RMSD over time colored by distance cluster according to the same color scheme.

Cluster analysis was performed on the long timescale simulations for State 1 and State 2 in order to group similar structures for investigation by both sulfur-sulfur reduced disulfide bond distance and by root mean square deviation (RMSD) of the backbone carbons from their original positions at the start of the simulation. Interestingly, despite the relatively similar overall RMSDs and overall disulfide bond distances for both State 1 and State 2, the cluster analysis reveals some striking differences between their structures as they unfold, as can be seen in Figure S3. The cluster analysis with respect to RMSD for the C6-C127 distance in State 1 indicates that a variety of states are sampled over time, and shows that the C6-C127 distance is constantly fluctuating between the clusters. State 2, on the other hand, has a diverging cluster relative to the RMSD and shows specific distances in stages throughout the simulation, indicating that the system populates specific regions while the protein unfolds. This difference between State 1 and State 2 is consistent with changes in hydrogen bonding and secondary structure stability (see below). The C30-C115 distance clustering shows that State 2 has two distinguishable centroids between RMSD and distance whereas State 1

again shows a smaller number of clusters with respect to RMSD, though they both display fluctuation behavior. For the C64-C80 distance this behavior is flipped, with State 1 showing a more highly diverging character from RMSD and a clear distance-based state progression over time, and State 2 showing a more centralized cluster and average-fluctuation over time.

State 1



State 2

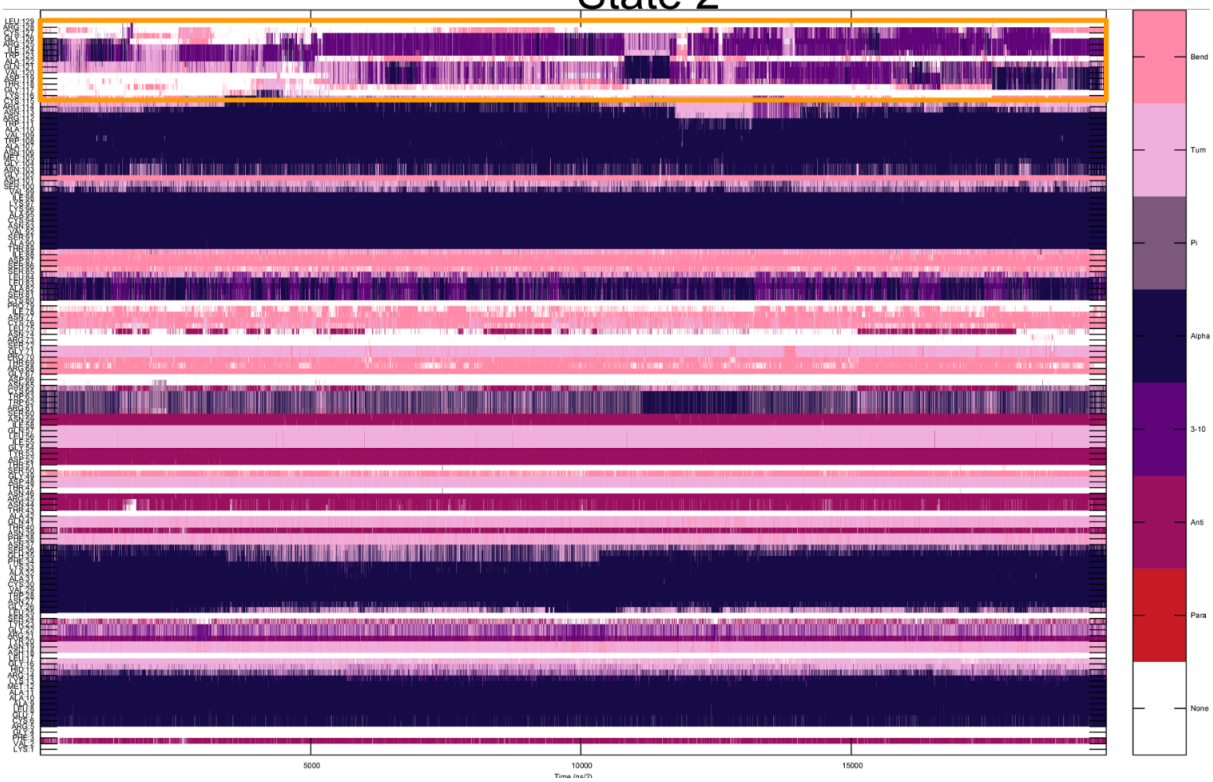


Figure S2. Transition in character of secondary structure per residue over time (ns*5) for State 1 (top) and State 2 (bottom), with salmon representing bend character, pink representing turn character, gray representing pi helix character, navy blue representing alpha helix character, purple representing 3-10 helix character, mauve representing antiparallel beta sheet, red representing parallel beta sheet, and white representing unstructured loops. Areas of strong differences between the states are highlighted in orange.

Figure S4 shows the per-residue secondary structure nature over time for each trajectory—in both cases, though as previously discussed the extent of the change differs in location between the states, there is a clear lowering of alpha helix character and increase in beta sheet character over time for both protonation states. State 1 shows more overall change to beta sheet structures over time, while State 2 shows more overall destabilization of the alpha helices. It is also clear from Figure 7 that the beta sheet section of the lysozyme is the only consistently stable secondary structure between both State 1 and State 2.