Biophysical Journal, Volume 97

Supporting Material

Common structural transitions in explicit solvent simulations of villin headpiece folding

Peter L. Freddolino and Klaus Schulten

Supplementary material for "Common structural transitions in explicit-solvent simulations of villin headpiece folding"

Peter L. Freddolino and Klaus Schulten

August 5, 2009

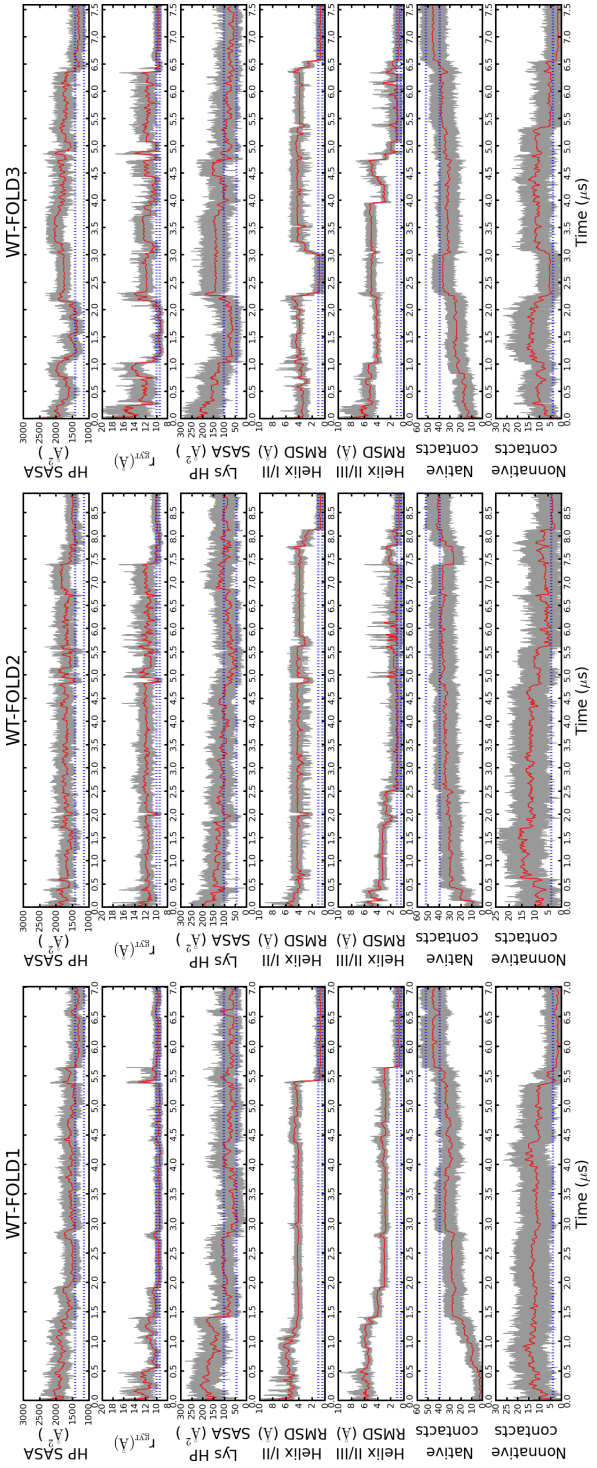

Figure 1: Statistics throughout the WT folding trajectories. Running averages over 30 ns are shown in red, and the range defined by the mean ± two standard deviations from simulation WT-NAT as blue bars. "HP SASA" refers to solvent accessible surface area (SASA) of hydrophobic groups. Lys HP SASA is the SASA of the aliphatic chains of residues 65 and 70. Native contacts are defined as favorable polar (heavy atom distance <3.5 Å) or hydrophobic (heavy atom distance <4.0 Å) sidechain-sidechain contacts formed in >25% of timesteps in simulation WT-NAT formed between residues separated by at least two residues in primary sequence (*n.b.* backbone hydrogen bonding interactions are included).

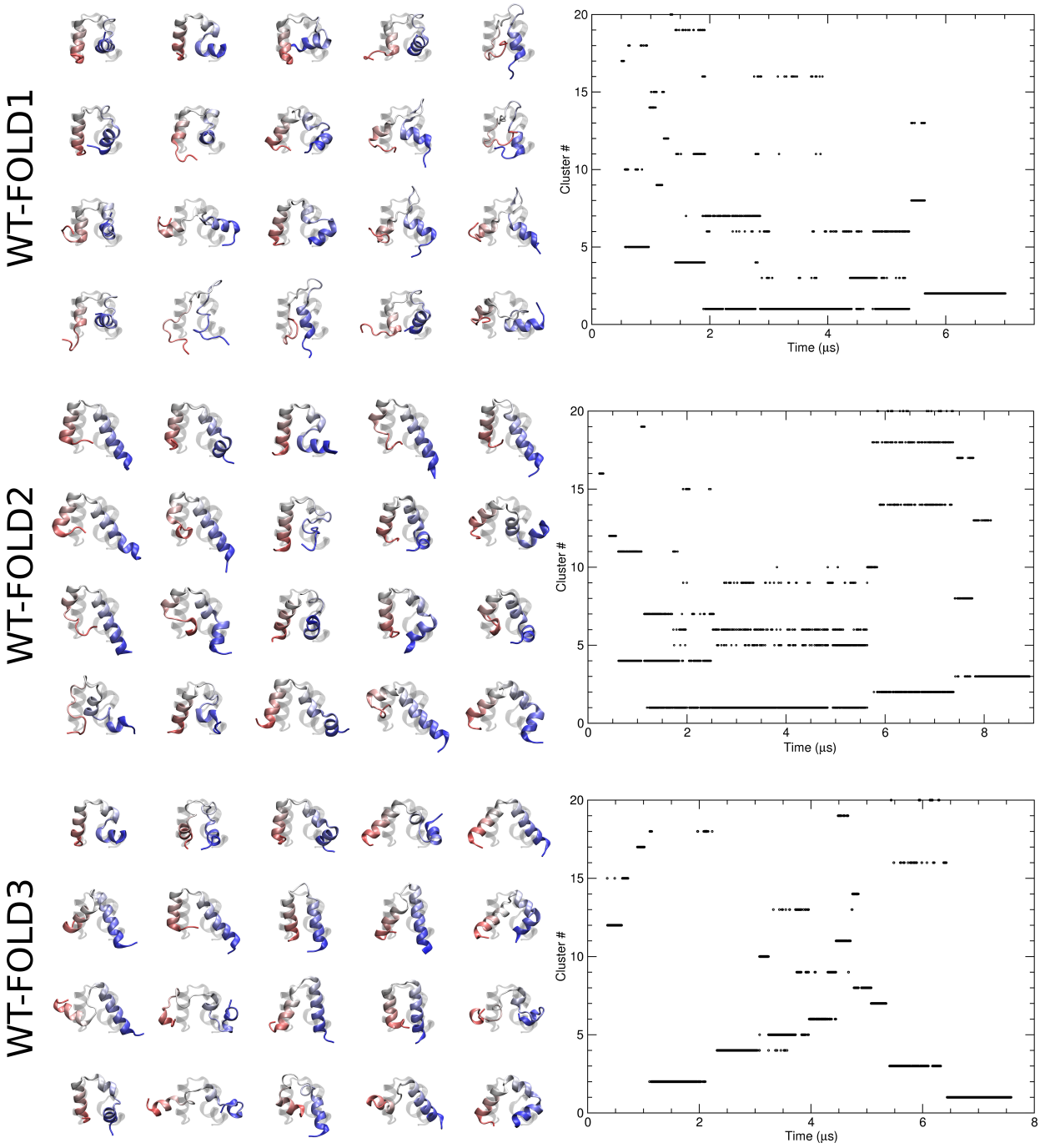

Figure 2: Clustering analysis for the WT simulations; the clustering procedure is described in the Materials and Methods section of the main text. For each simulation representative conformations for the twenty most occupied clusters are shown, in descending order of occupancy (running left to right, top to bottom). In addition, the cluster occupied at each point in time is plotted to the right of the conformations. Coloring of the cartoon diagrams for protein conformations runs blue to red from N terminus to C terminus. The crystal structure is shown in a transparent gray cartoon representation in all cases.

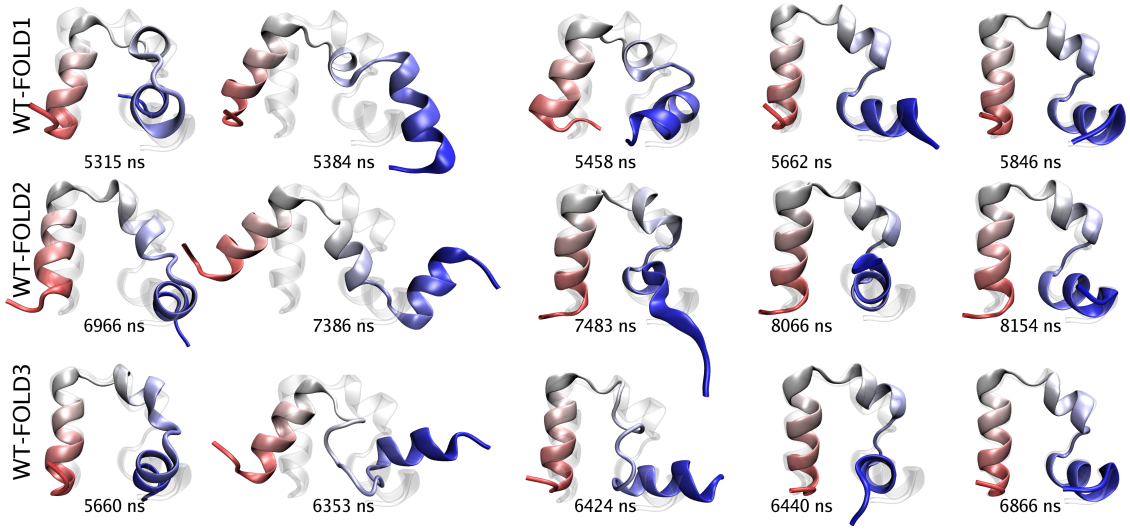

Figure 3: Representative structures from the transition between the flipped, open and post-flip conformation for each of the WT trajectories. Protein coloring runs blue to red from N terminus to C terminus; the crystal structure is shown as a transparent gray cartoon for comparison.

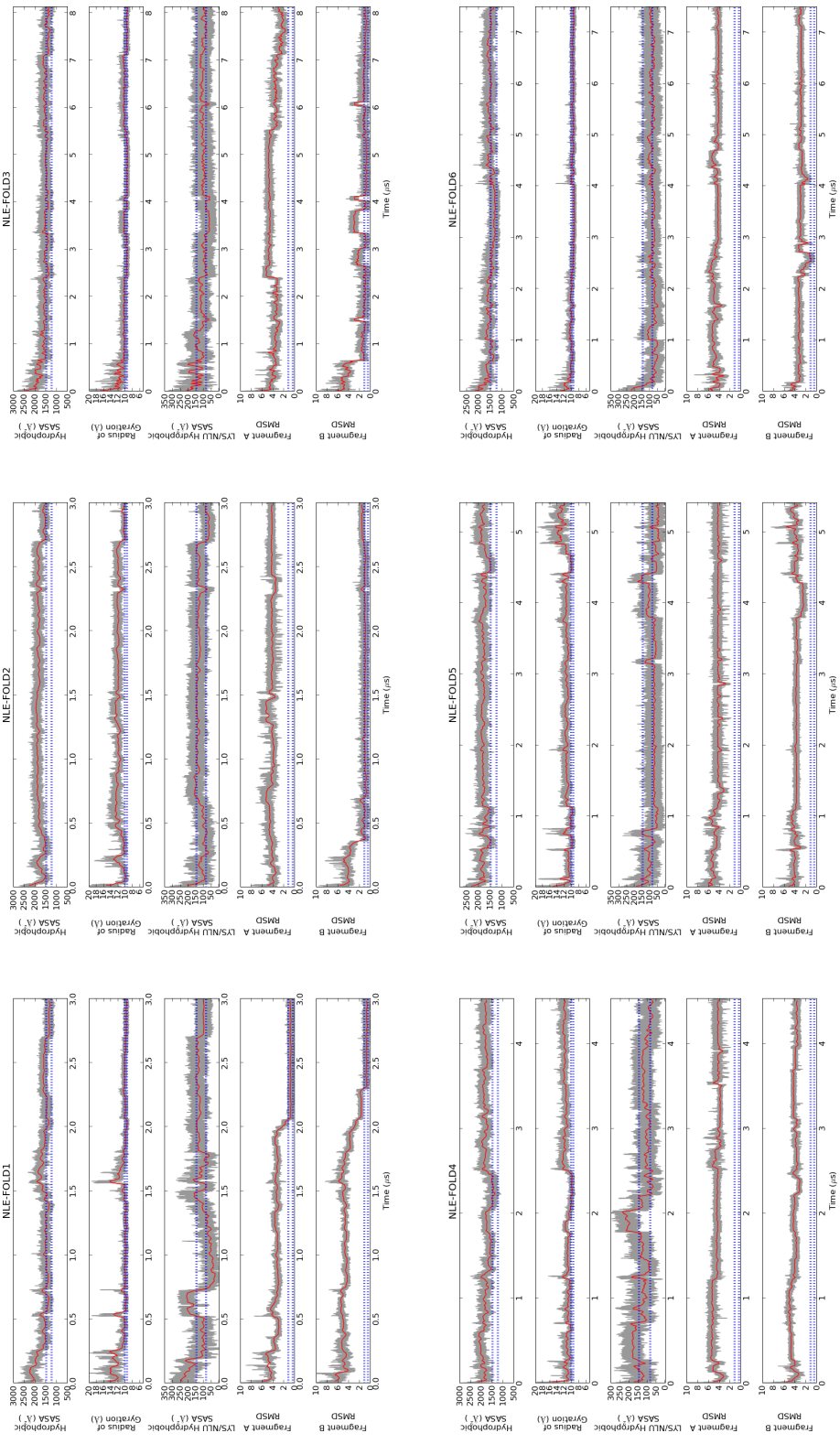

Figure 4: Statistics throughout NLE folding trajectories. Definitions are identical to those in SI Fig. 1; reference values are taken from simulation NLE-NAT.

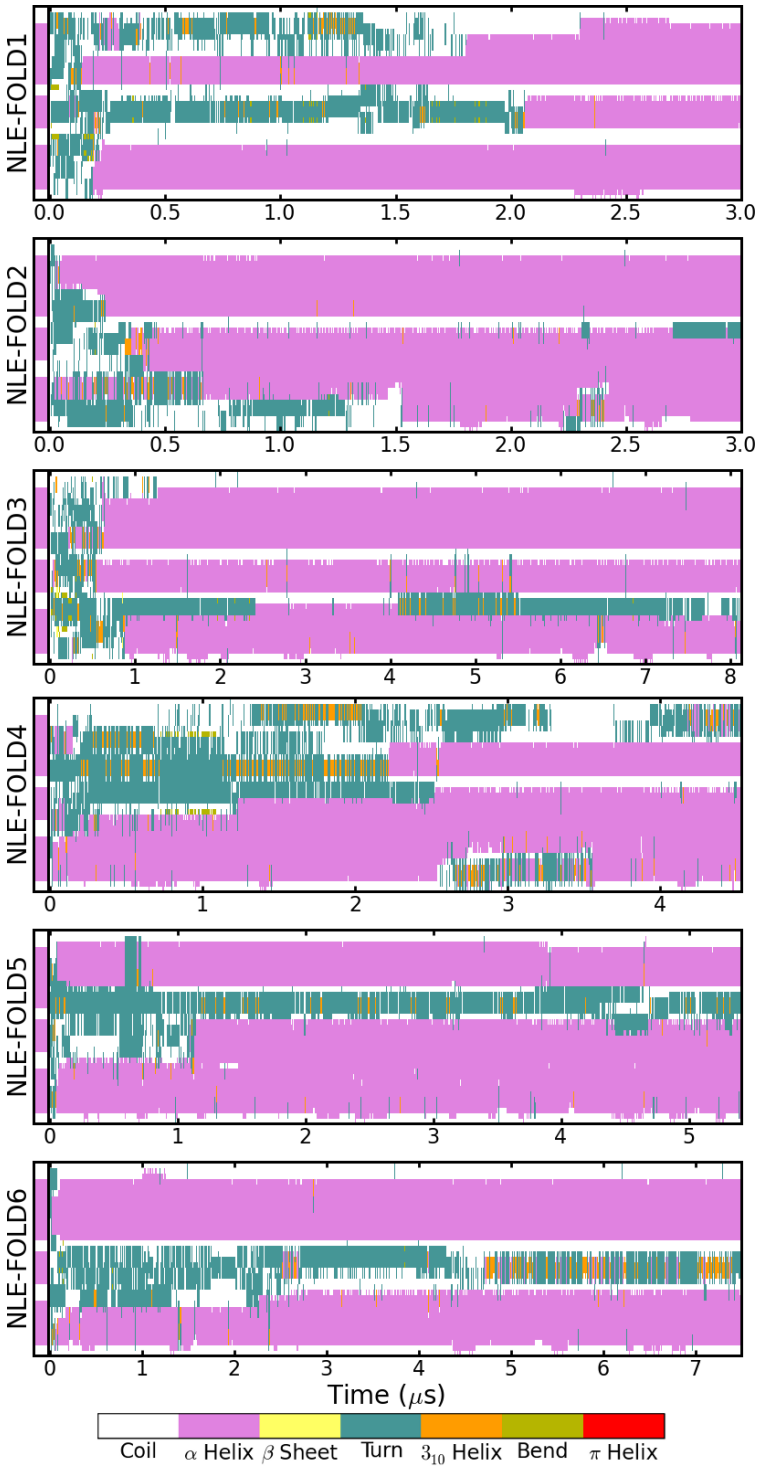

Figure 5: Secondary structure throughout the NLE folding trajectories; the values from the crystal structure are shown at the left side for reference.

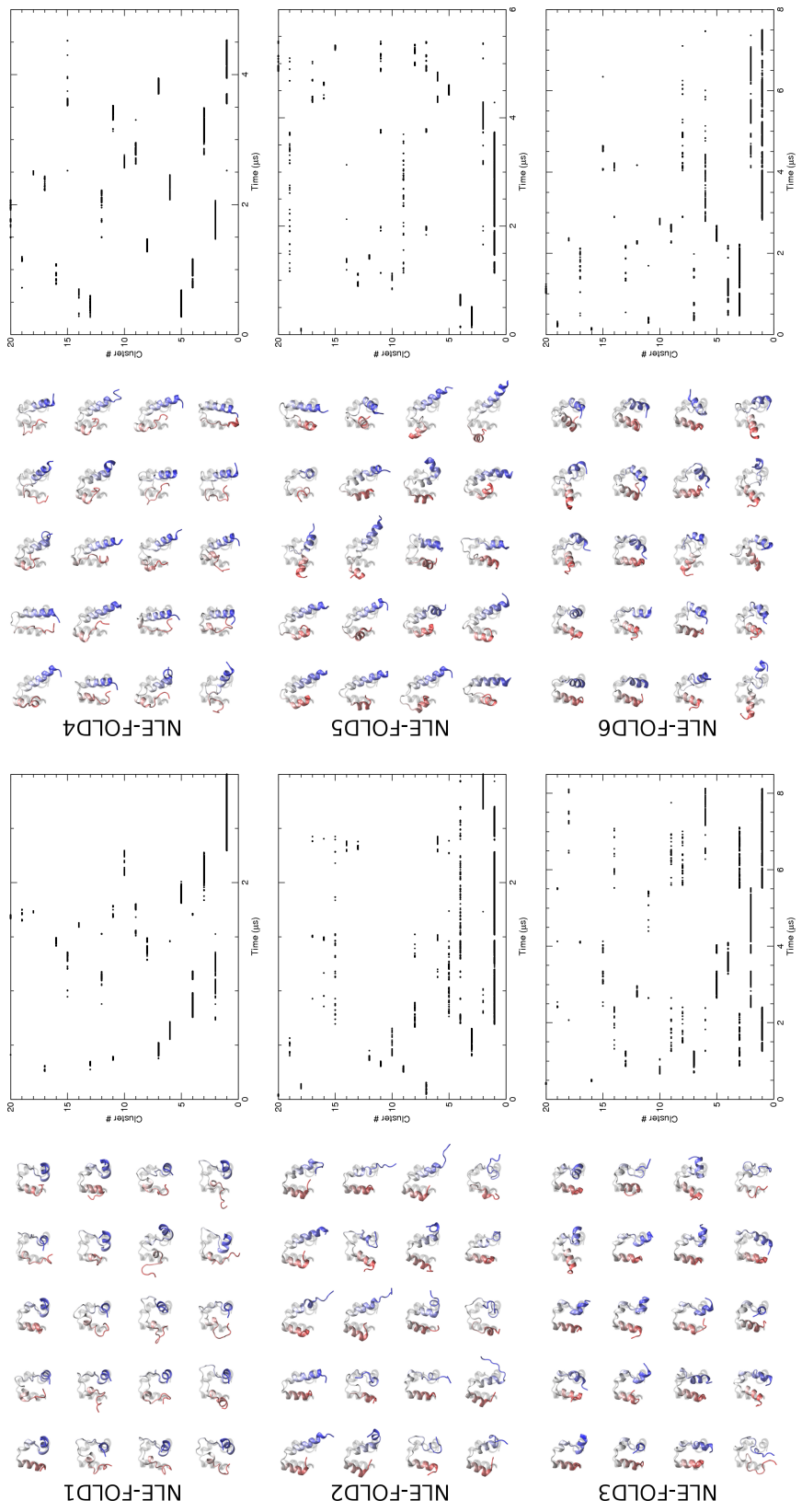

Figure 6: Clustering analysis for the NLE simulations. The clustering procedure is described in the Materials and Methods section of the main text. For each simulation representative conformations for the twenty most occupied clusters are shown, in descending order of occupancy (running left to right, top to bottom). In addition, the cluster occupied at each point in time is plotted to the right of the conformations. Coloring of the cartoon diagrams for protein conformations runs blue to red from N terminus to C terminus; the crystal structure is shown in each case as a transparent gray cartoon.

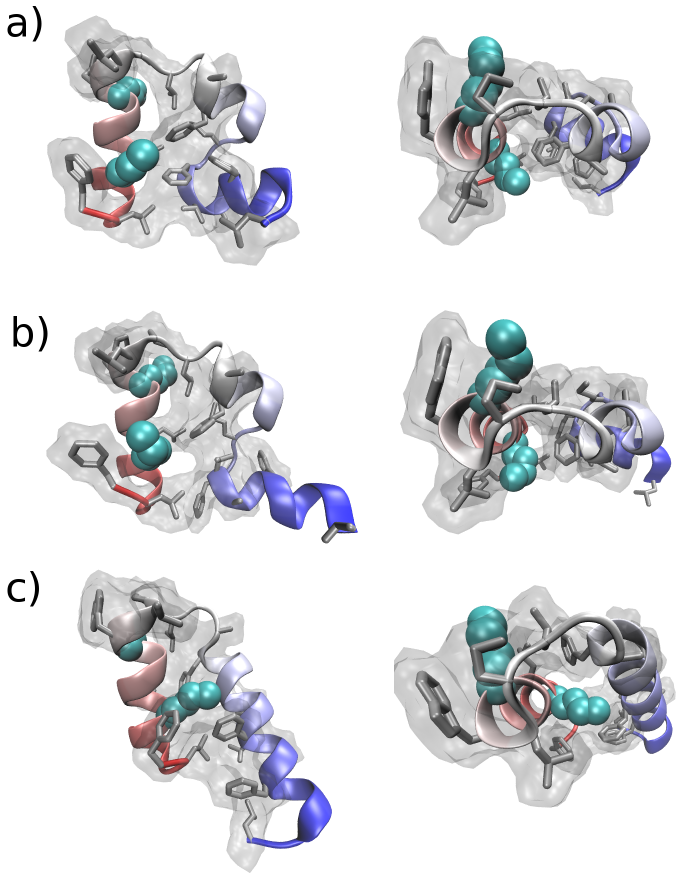

Figure 7: Representative conformations of highly occupied clusters from the NLE simulations. In all cases the protein backbone is colored blue to red from N terminus to C terminus; the norleucine residues are shown in cyan, other hydrophobic side chains in gray, and a molecular surface representations of the hydrophobic side chains as a transparent gray surface. a) Cluster 1 from simulation NLE-FOLD1 (native state). b) Cluster 1 from simulation NLE-FOLD3 (near-native state). c) Cluster 2 from simulation NLE-FOLD2 (non-native state with joined helix I/II).

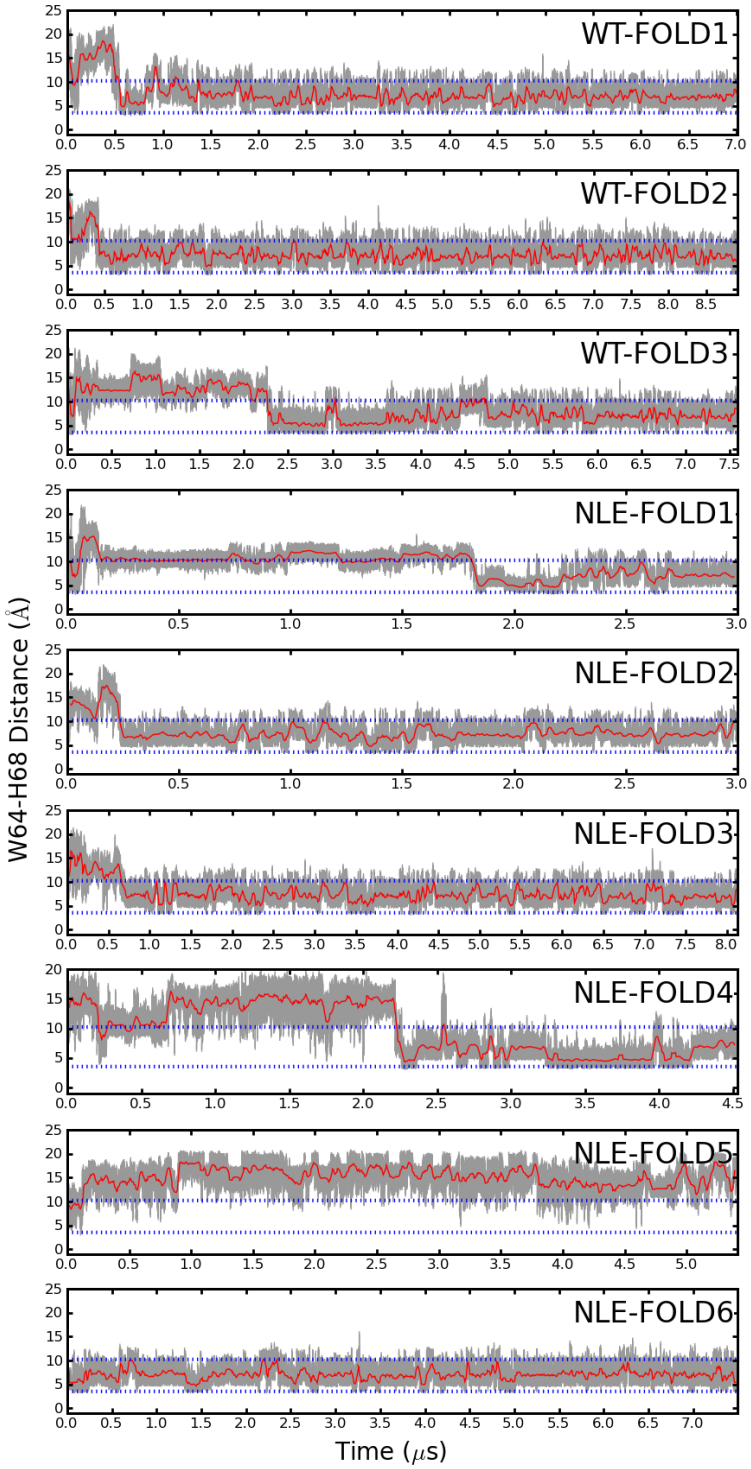

Figure 8: Distance between the center of mass of the W64 and H68 ring systems throughout the folding simulations. Average values ± two standard deviations from simulation WT-NAT are shown as blue dotted lines, and a running average over 30 ns is shown in red.

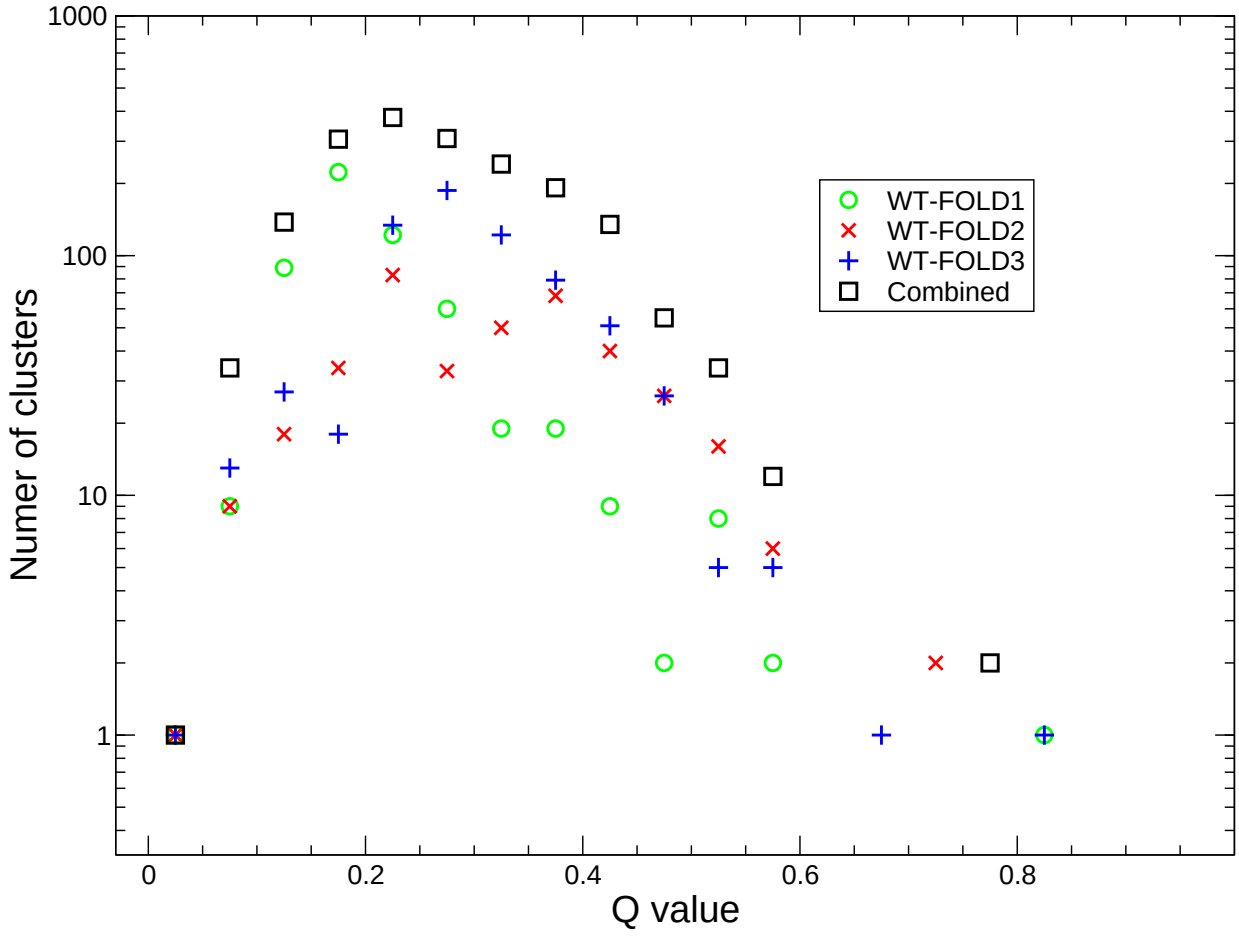

Figure 9: Histograms of the number of clusters found for each conformation in a given range values of Q_{res} to the native state (Q_{res} values are calculated for the central structure of each cluster); bins are spaced 0.05 units apart. Bins with zero occupancy are omitted. "Combined" refers to a separate clustering run on the concatenation of the WT-FOLD1, WT-FOLD2, and WT-FOLD3 trajectories, not a simple sum of the other data points.