Supplementary Information for

Influence of Branched Ligand Architectures on Nanoparticle Interactions with Lipid Bilayers

Carlos A. Huang-Zhu[1] and Reid C. Van Lehn[1,2,*]

[1]Department of Chemical and Biological Engineering, University of Wisconsin−Madison, Madison, Wisconsin 53706, United States

[2]Department of Chemistry, University of Wisconsin−Madison, Madison, Wisconsin 53706, United States

*to whom correspondence should be addressed: vanlehn@wisc.edu

Table S1. Number of water beads for simulated systems. Water beads are equal among replicas of the same chain architectures because the same initially solvated AuNP-bilayer system was used for all string method calculations. The number of lipids and counterions are the same across all systems.

Ligand R group	Chain Architecture	Number of water beads
11-C	Linear	30,285
	Double-Branched	30,216
15-C	Linear	29,770
	Triple-Branched	29,936
19-C	Linear	29,297
	Double-Branched	29,739
27-C	Double-Branched	29,148
	Triple-Branched	29,420

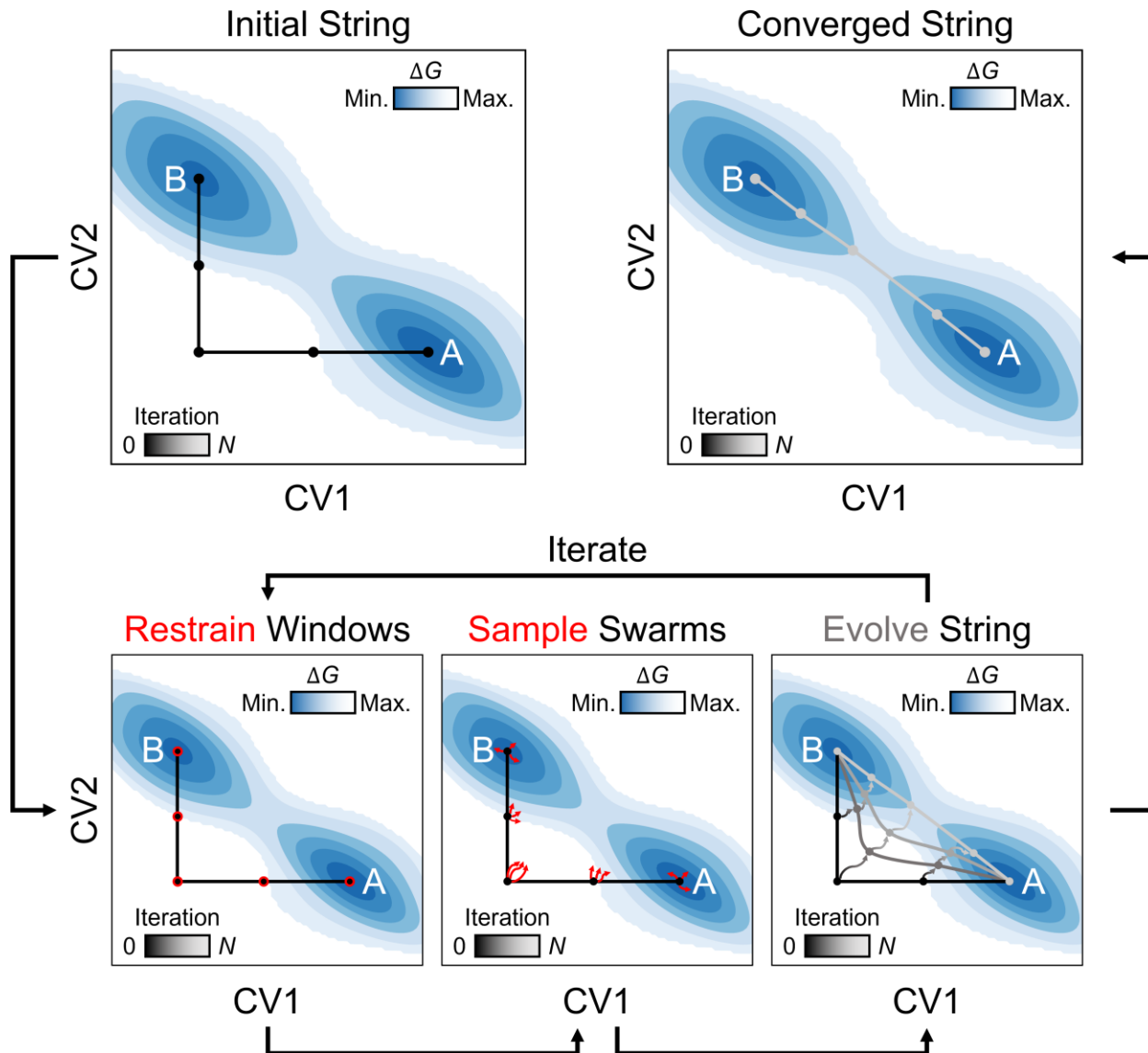Convergence of minimum free energy path (MFEP) calculations.

Figure S1. Schematic illustrating the string method with swarms of trajectories workflow for a simple 2D free energy landscape. An initial string composed of independent windows is generated connecting state A to B. At each iteration, the CV values from each window are restrained according to the string. From the restrained windows, a series of short unbiased trajectories are generated (*i.e.*, the swarms of trajectories, indicated as short red arrows). The drift of the CVs is averaged over all trajectories for each swarm and added to the old CVs to evolve the string, which is then reparameterized to ensure equal distances between consecutive windows. The process is repeated until the string is converged.
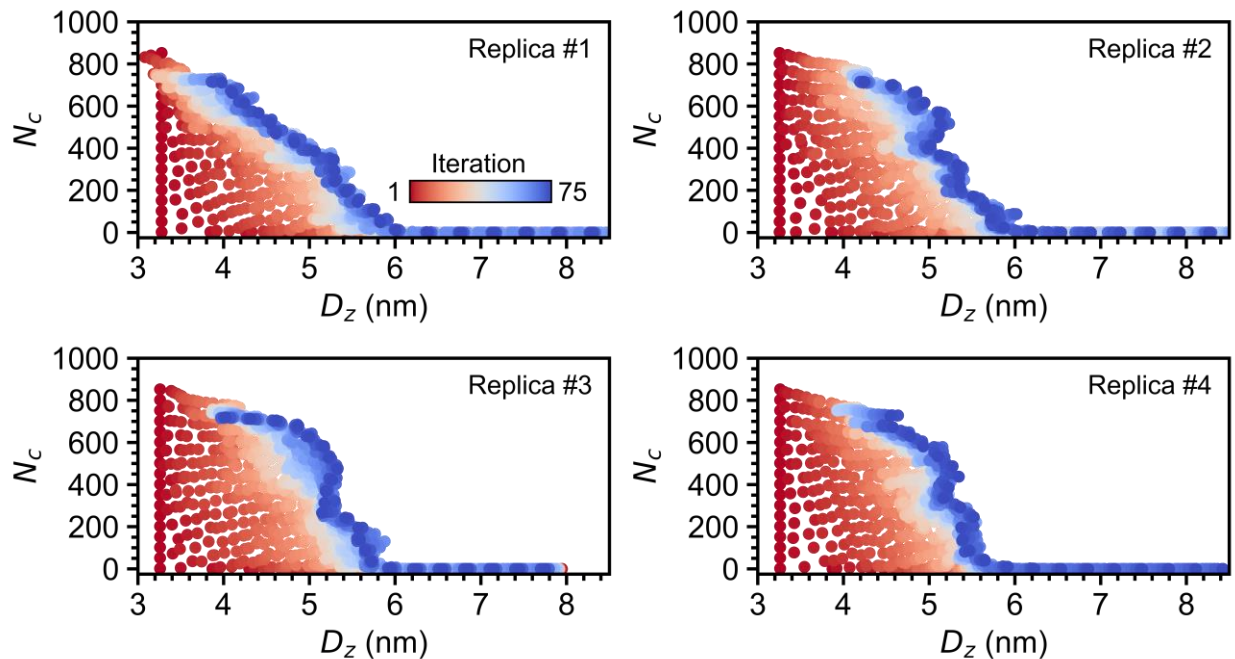
Figure S2. String evolution for all replicas of the linear 11-C ligand R group.
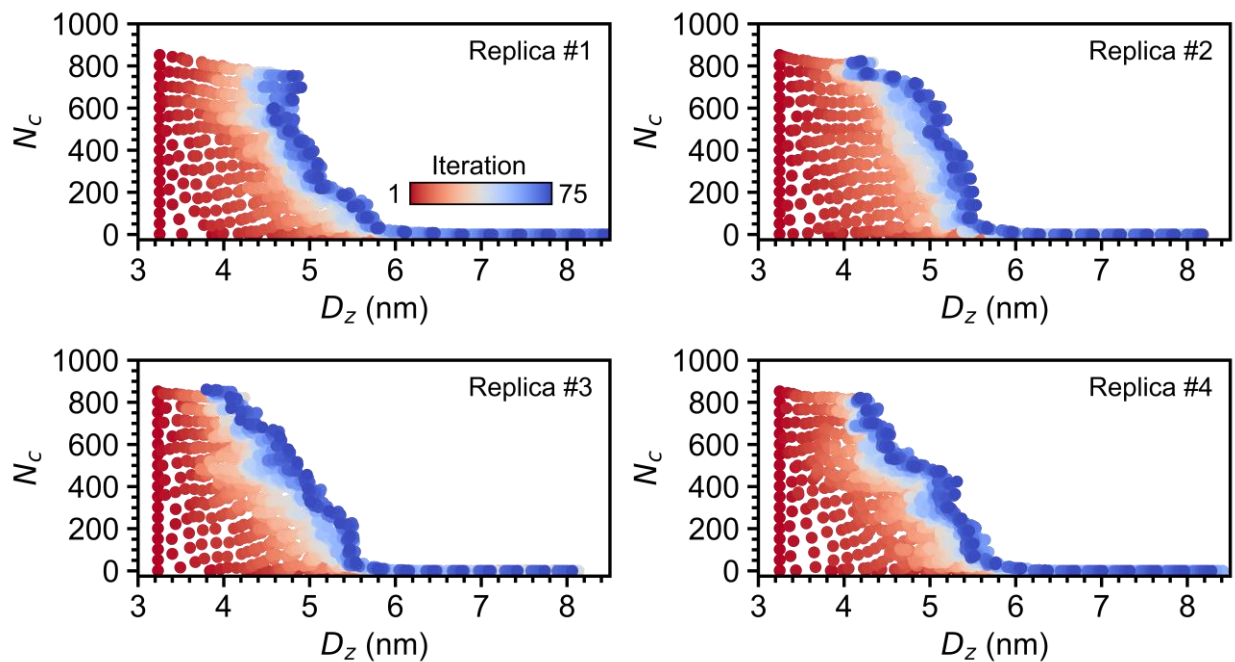

Figure S3. String evolution for all replicas of the double-branched 11-C ligand R group.
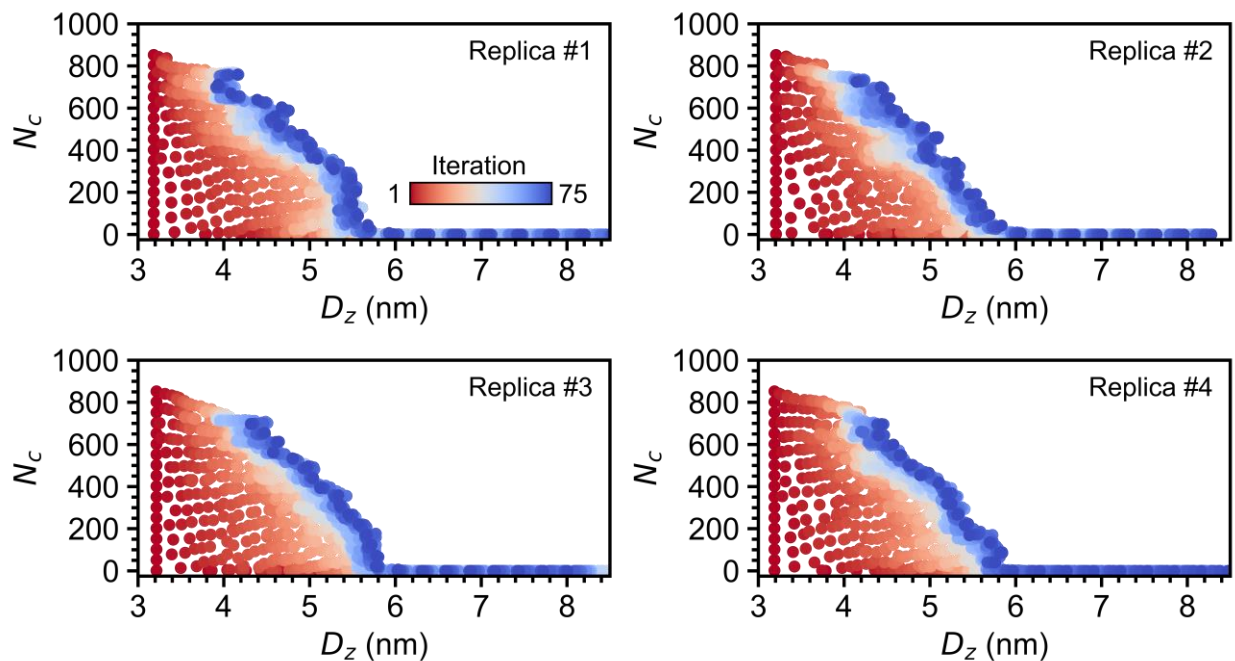
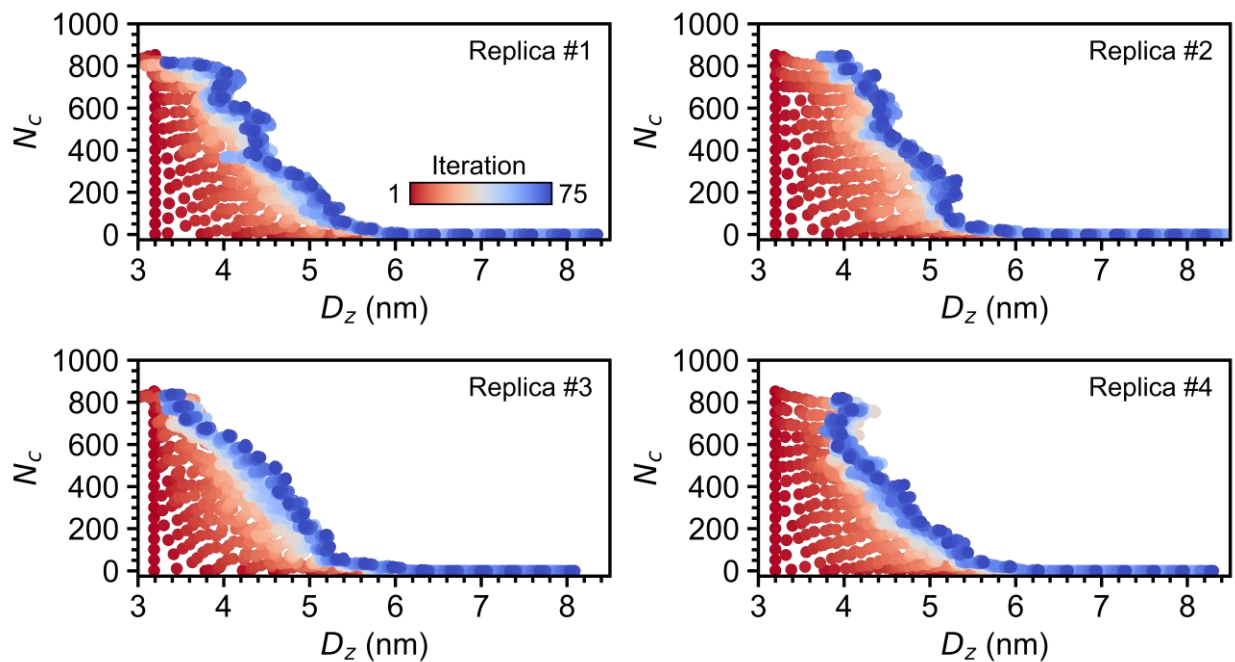Figure S4. String evolution for all replicas of the linear 15-C ligand R group.

Figure S5. String evolution for all replicas of the triple-branched 15-C ligand R group.
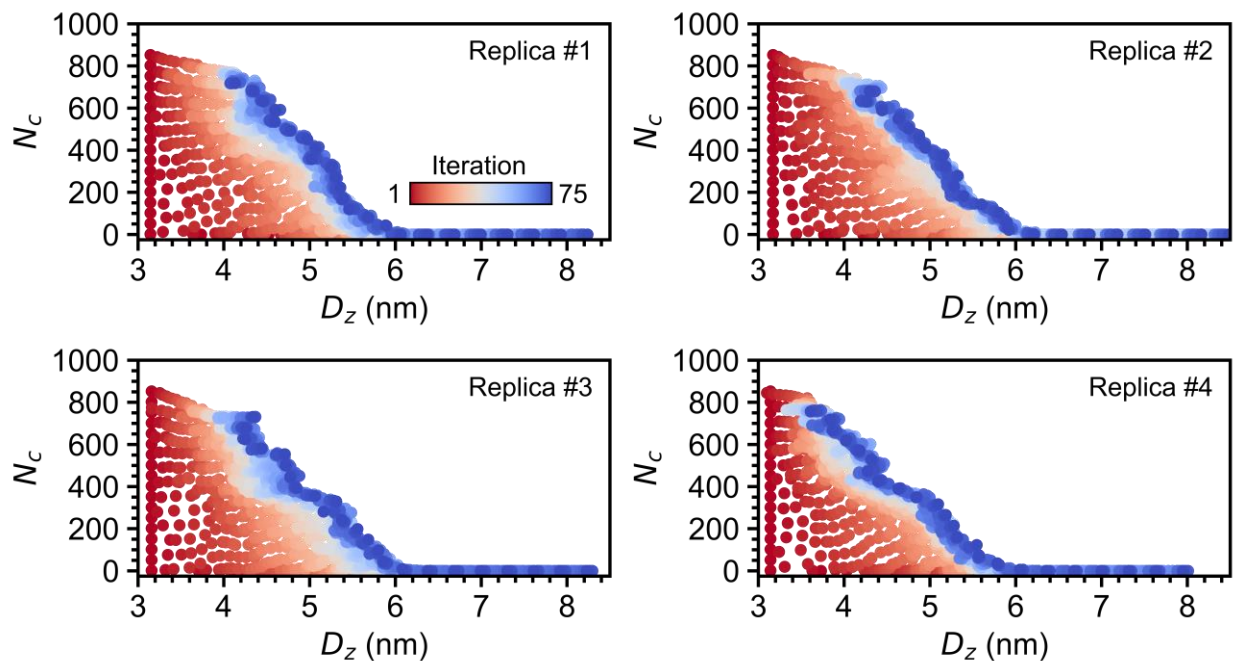
Figure S6. String evolution for all replicas of the linear 19-C ligand R group.
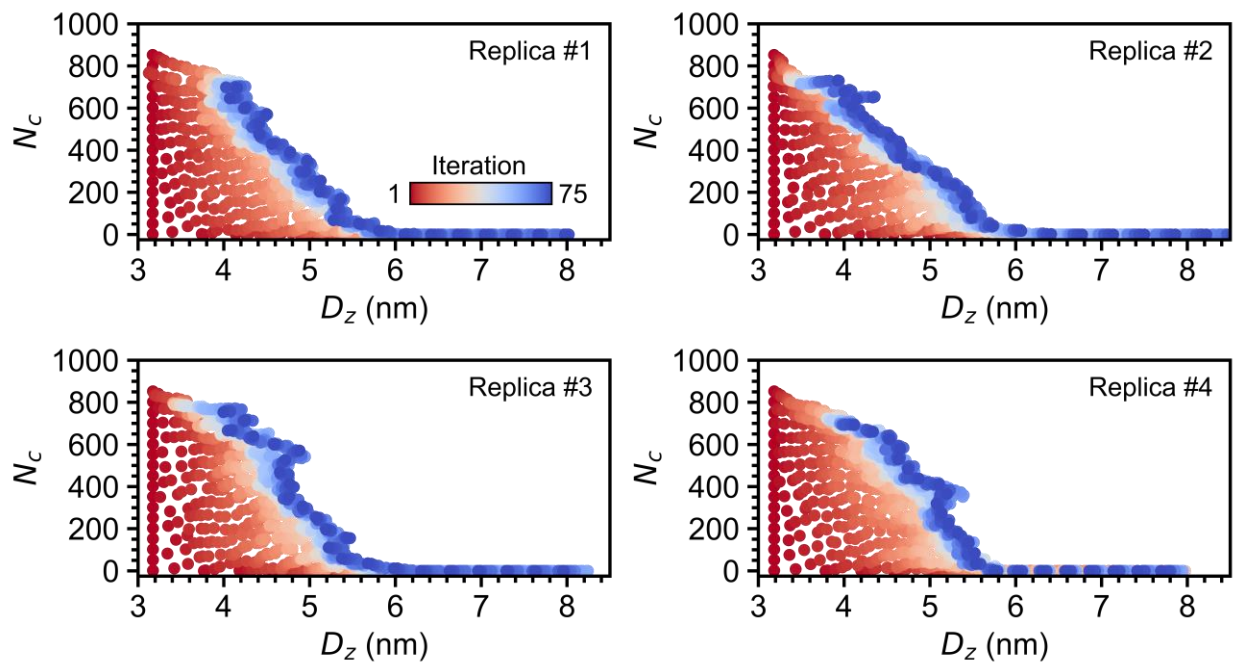

Figure S7. String evolution for all replicas of the double-branched 19-C ligand R group.
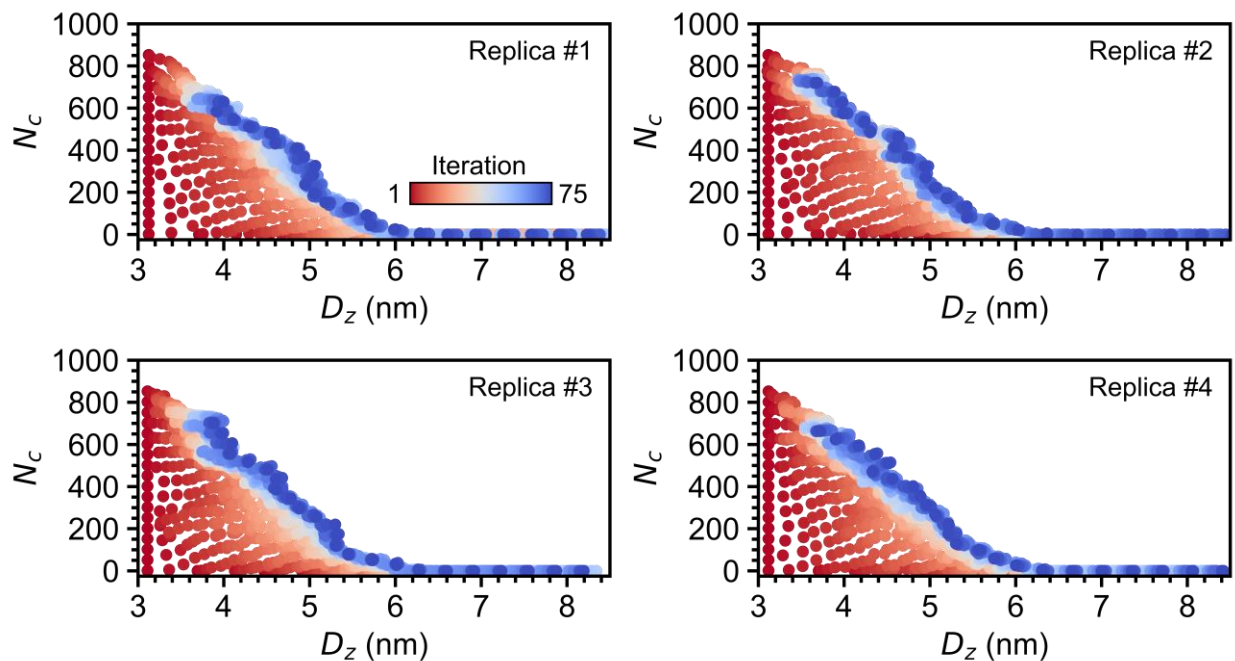
Figure S8. String evolution for all replicas of the double-branched 27-C ligand R group.
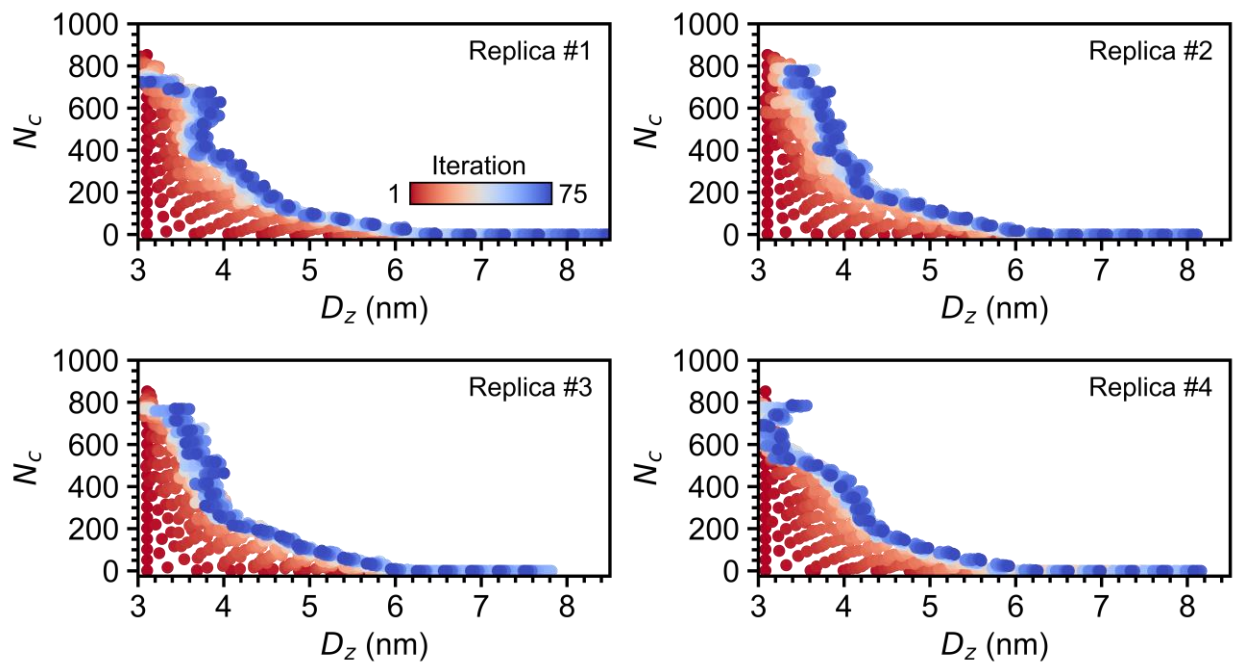

Figure S9. String evolution for all replicas of the triple-branched 27-C ligand R group.
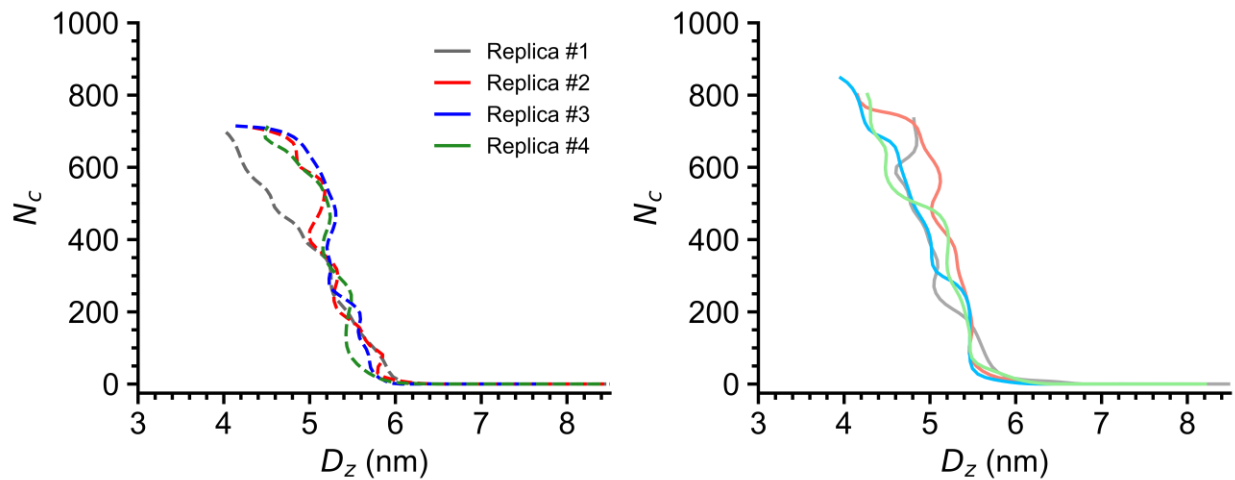
Figure S10. Converged minimum free energy paths for 11-C ligand pair. Left plot (dashed lines) correspond to the linear R group and right plot (solid lines) correspond to the double-branched R group.

Figure S11. Converged minimum free energy paths for 15-C ligand pair. Left plot (dashed lines) correspond to the linear R group and right plot (solid lines) correspond to the triple-branched R group.

Figure S12. Converged minimum free energy paths for 19-C ligand pair. Left plot (dashed lines) correspond to the linear R group and right plot (solid lines) correspond to the double-branched R group.
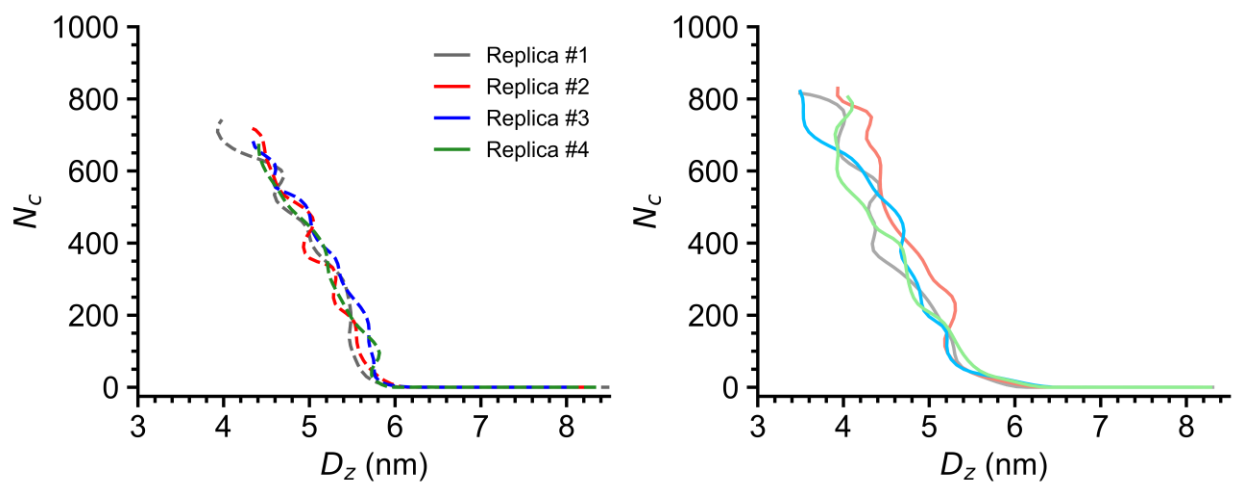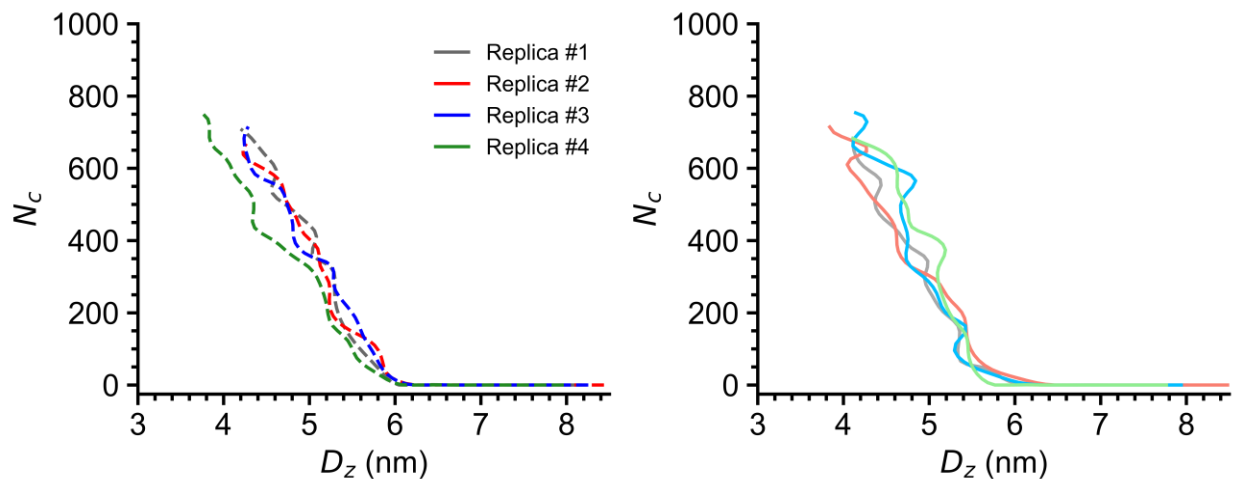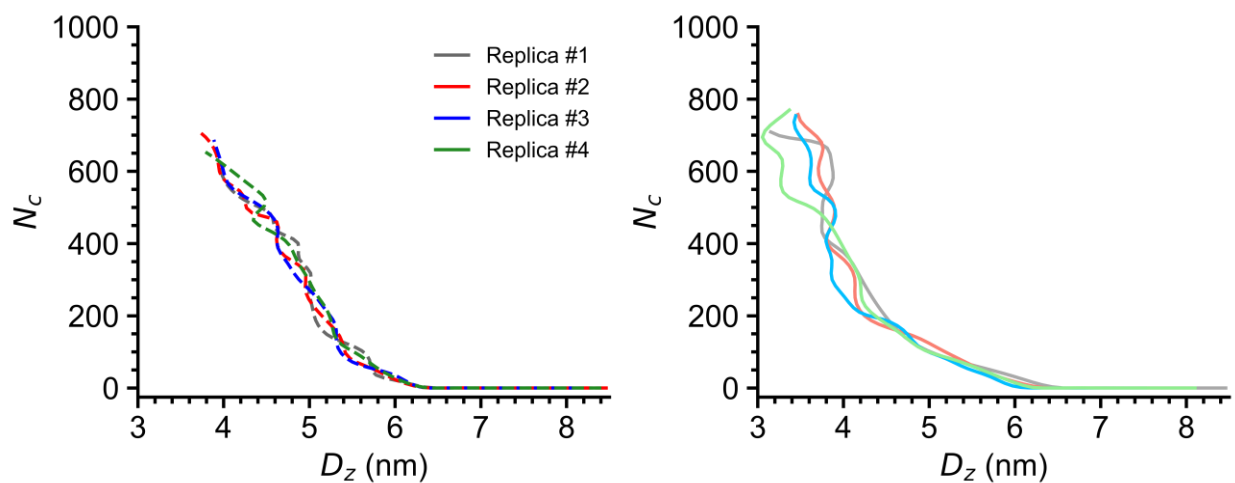

Figure S13. Converged minimum free energy paths for 27-C ligand pair. Left plot (dashed lines) correspond to the double-branched R group and right plot (solid lines) correspond to the triple-branched R group.

Figure S14. Converged strings used for the spline fit of all replicas of the linear 11-C R group. The collective variables computed from the umbrella sampling windows are plotted as dots with the black dot representing the window corresponding to the adsorption barrier.

Figure S15. Converged strings used for the spline fit of all replicas of the double-branched 11-C R group. The collective variables computed from the umbrella sampling windows are plotted as dots with the black dot representing the window corresponding to the adsorption barrier.
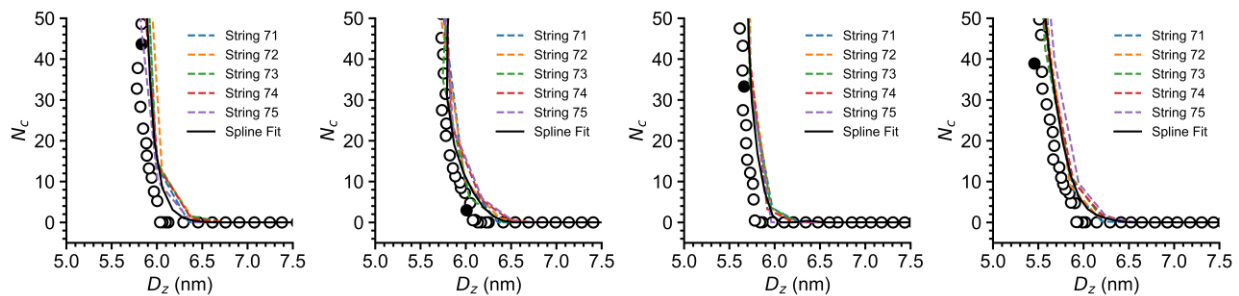
Figure S16. Converged strings used for the spline fit of all replicas of the linear 15-C R group. The collective variables computed from the umbrella sampling windows are plotted as dots with the black dot representing the window corresponding to the adsorption barrier.

Figure S17. Converged strings used for the spline fit of all replicas of the triple-branched 15-C R group. The collective variables computed from the umbrella sampling windows are plotted as dots with the black dot representing the window corresponding to the adsorption barrier.
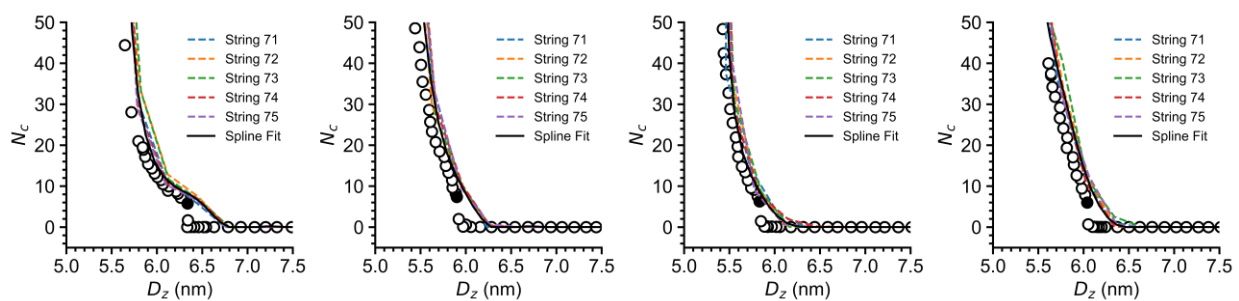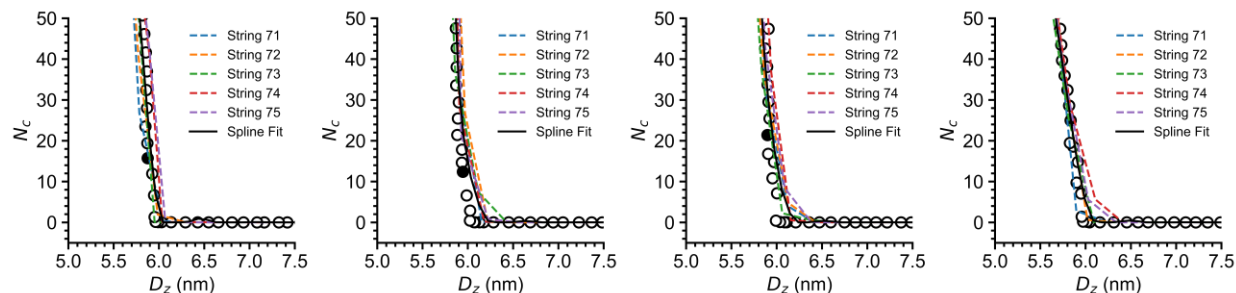
Figure S18. Converged strings used for the spline fit of all replicas of the linear 19-C R group. The collective variables computed from the umbrella sampling windows are plotted as dots with the black dot representing the window corresponding to the adsorption barrier.
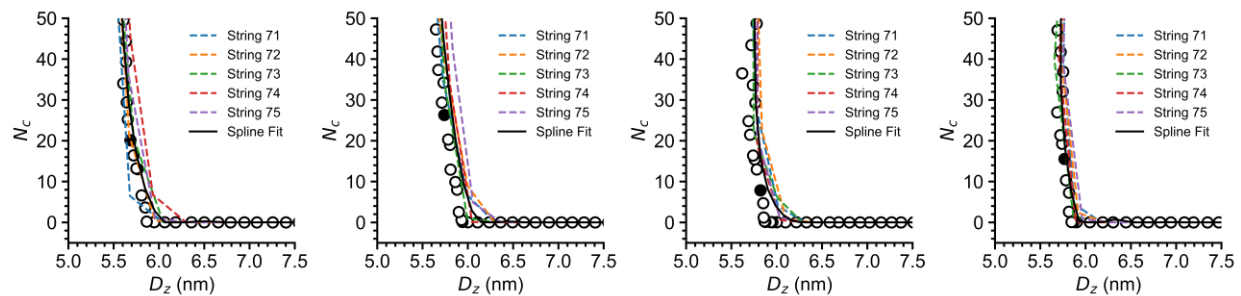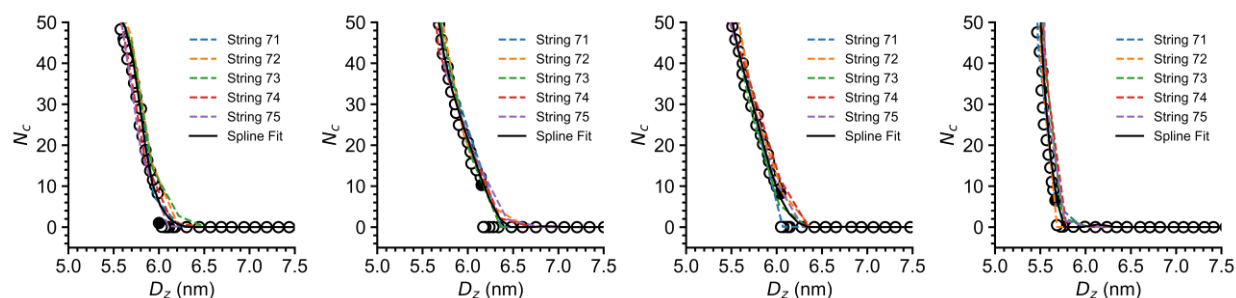

Figure S19. Converged strings used for the spline fit of all replicas of the double-branched 19-C R group. The collective variables computed from the umbrella sampling windows are plotted as dots with the black dot representing the window corresponding to the adsorption barrier.
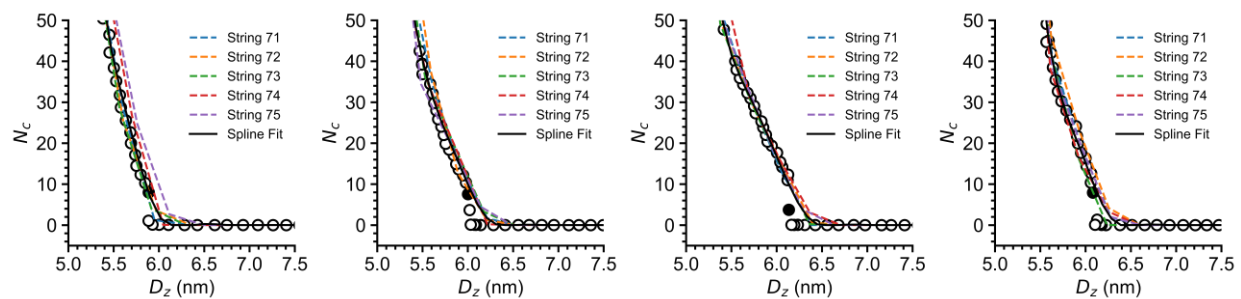
Figure S20. Converged strings used for the spline fit of all replicas of the double-branched 27-C R group. The collective variables computed from the umbrella sampling windows are plotted as dots with the black dot representing the window corresponding to the adsorption barrier.

Figure S21. Converged strings used for the spline fit of all replicas of the triple-branched 27-C R group. The collective variables computed from the umbrella sampling windows are plotted as dots with the black dot representing the window corresponding to the adsorption barrier.
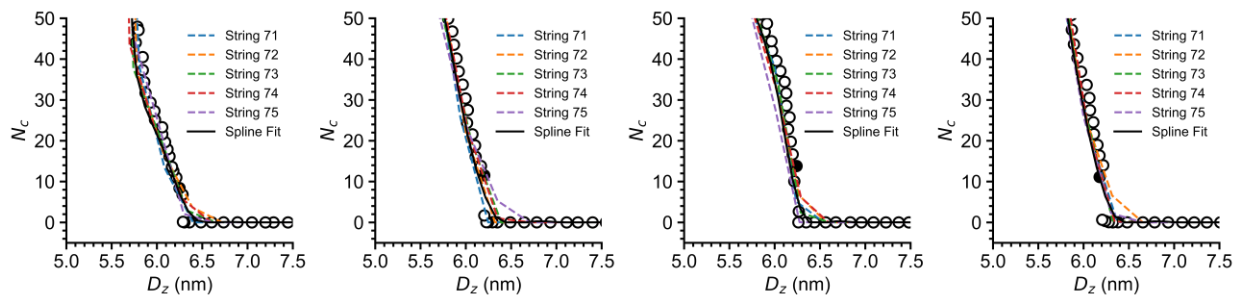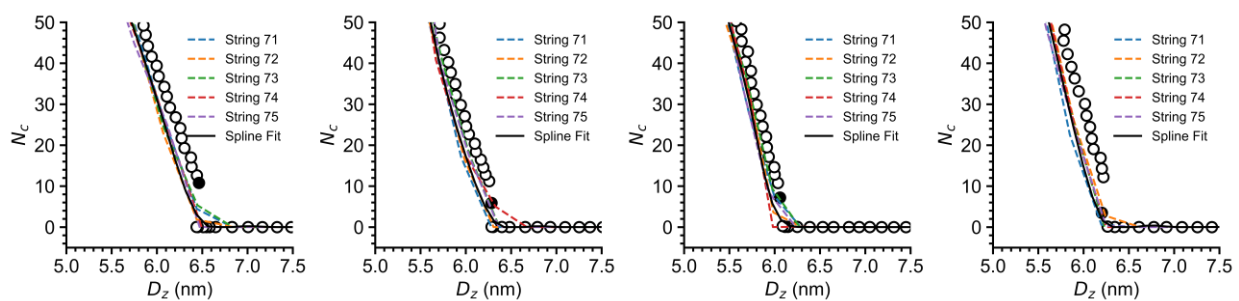
Convergence of potential of mean force (PMF) calculations.

Umbrella sampling simulations with 72 windows were performed to compute the potential of mean force along the minimum free energy path (MFEP) for each replica. For some windows the spring constant was individually tuned to achieve numerical stability (*i.e.*, regions where free energy barriers would be present). We biased the distance along the path, S, while restraining the normal distance away from the path, Z, using harmonic potentials. The values for S and Z, along with their respective spring constants are included in a separate CSV file for each AuNP and each replica.

Initial configurations for the umbrella sampling windows were taken from the five strings used to compute the MFEP. Each window was run for 50 ns. Convergence for each replica was assessed by computing the PMF using different timeframes from the umbrella trajectories by forward and backward sampling methods. In the forward sampling method, we sampled from 0 ns to t ns, where t was spaced out every 5 ns up to 50 ns. In the backward sampling method, we sampled from t ns to 50 ns, where t was spaced out every 5 ns starting from 0 ns. Configurations were saved every 1 ns, while collective variables were saved every 2 ps for the PMF calculation. PMFs were computed for each set of sampling times using the Weighted Histogram Analysis Method (WHAM)[1]. Umbrella sampling trajectories were deemed converged when the PMF barrier did not change with further sampling time (within a bound of ~1 $k_B T$). We also plotted the histograms to ensure sampling along the chosen reaction coordinate. For the linear ligand R groups, we used the last 45 ns (*i.e.*, from 5 ns to 50 ns) as the production trajectory, while for the branched ligand R groups we used the last 40 ns (*i.e.*, from 10 ns to 50 ns). Under-sampled regions at very large S observed in some replicas correspond to highly energetic and unfavorable states that suggest a subsequent transition event (*e.g.*, insertion into the bilayer) that is out of the scope of this study. Therefore, further tuning of the harmonic potentials for those umbrella windows was not performed, and those configurations are not considered for any analyses. The free energy barriers for adsorption and the adsorption free energies correspond to lower values of S and are thus not affected by these under-sampled regions. All plots shown below use an unscaled value of S; values reported in the main text were normalized such that $0 \leq S \leq 1$ for easier comparison between MFEPs and interpretability.

Figure S22. Potential of mean force (PMF) convergence for Replica #1 of the linear 11-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of *S* of interest.
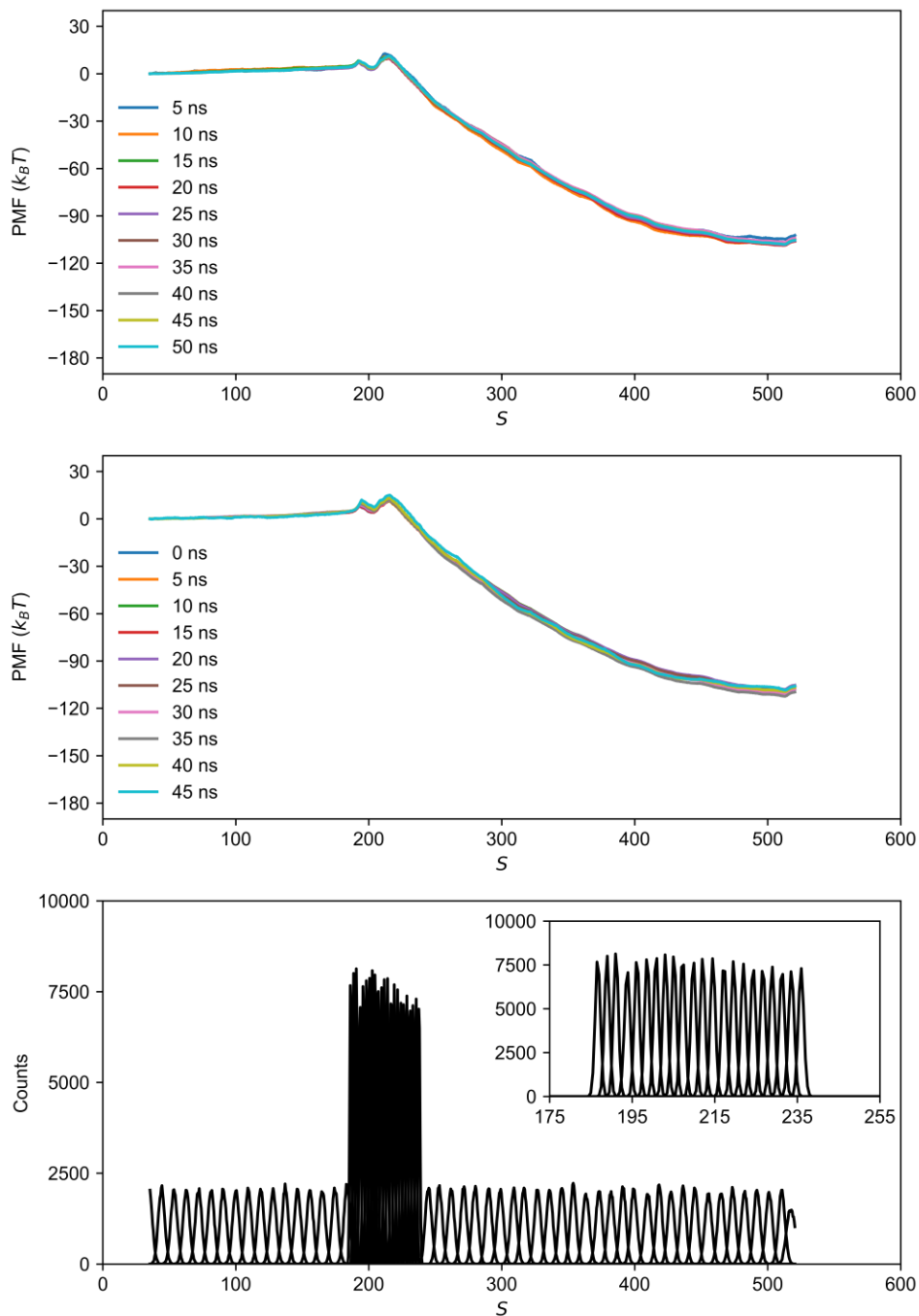
Figure S23. Potential of mean force (PMF) convergence for Replica #2 of the linear 11-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
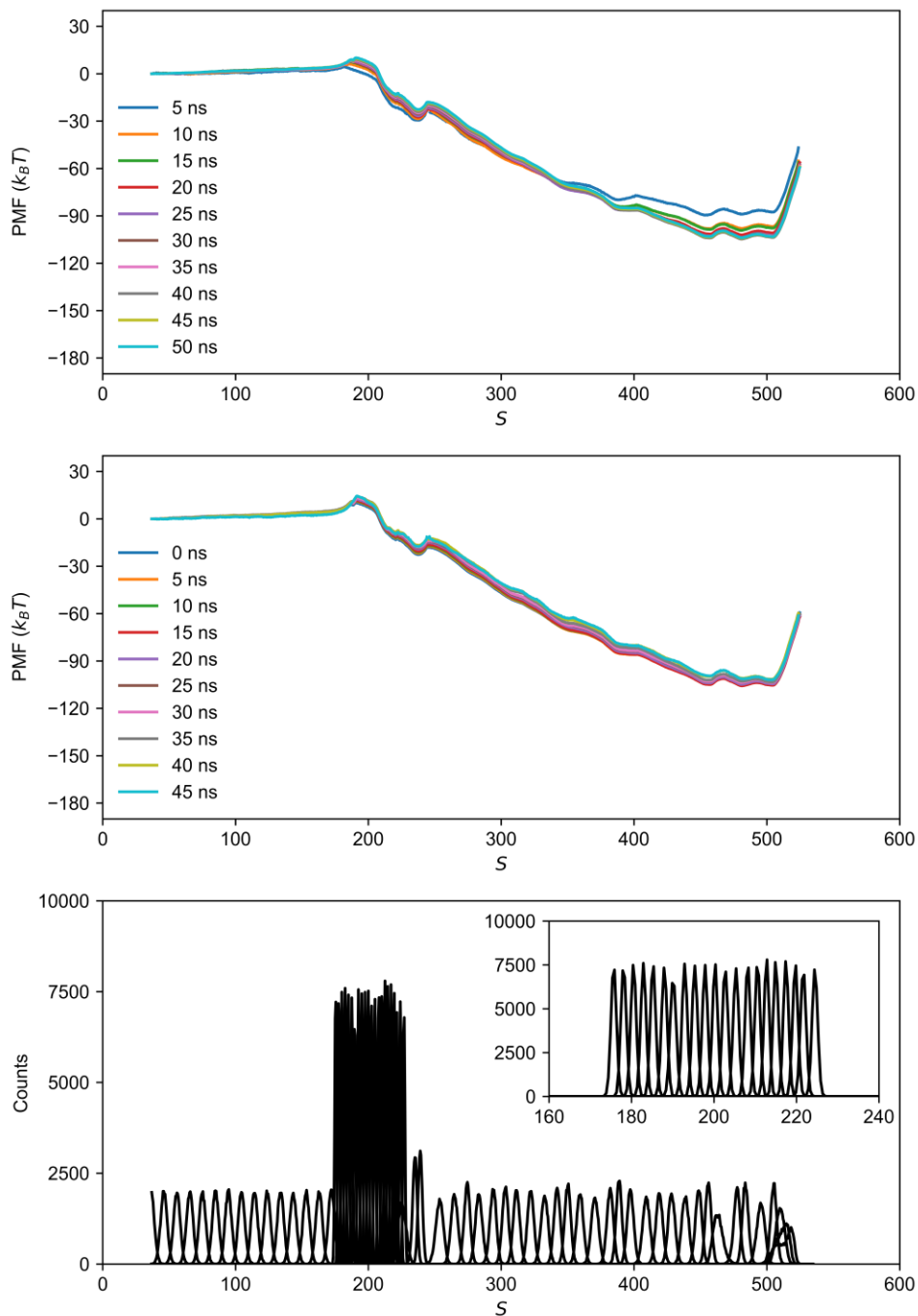
Figure S24. Potential of mean force (PMF) convergence for Replica #3 of the linear 11-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
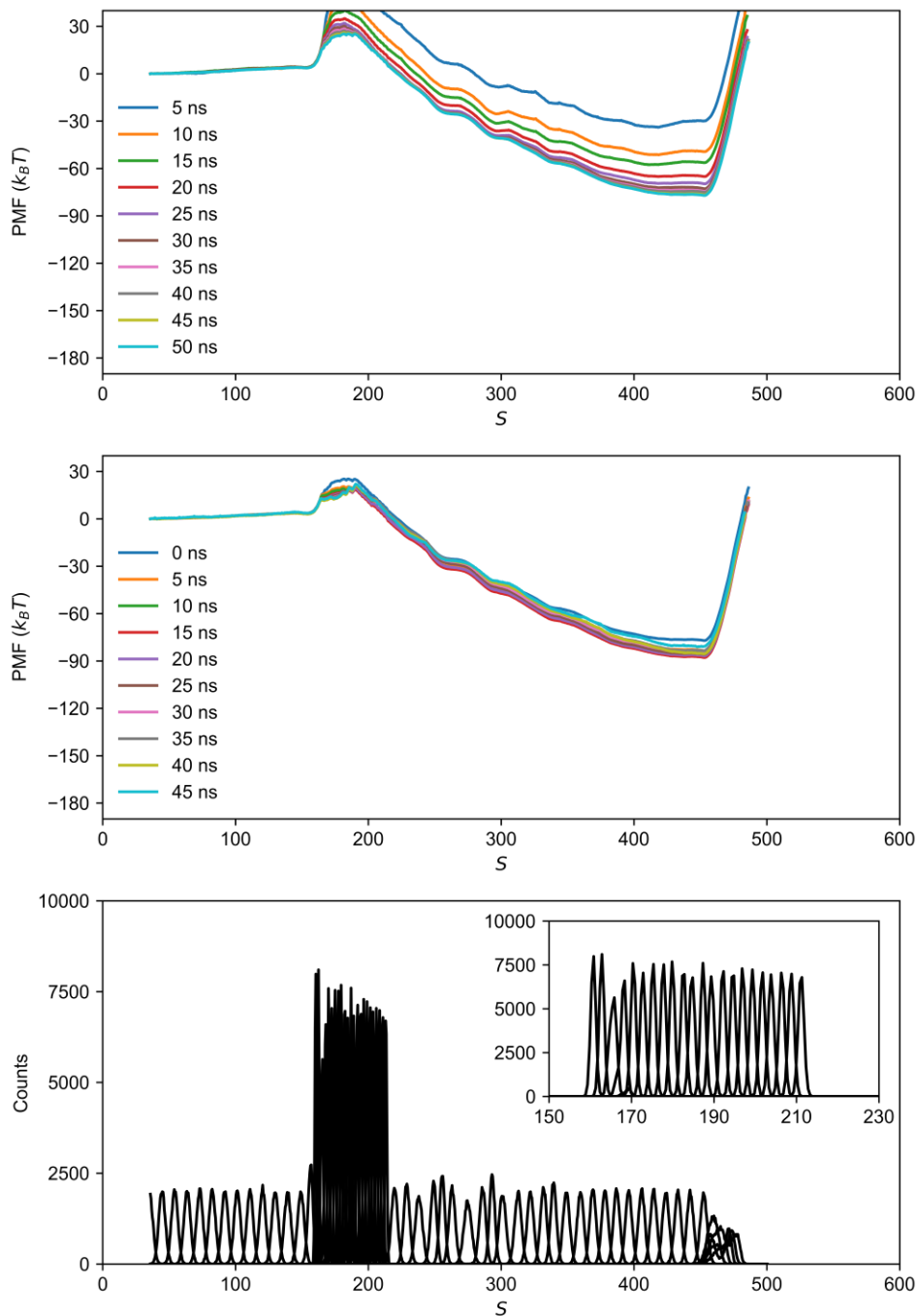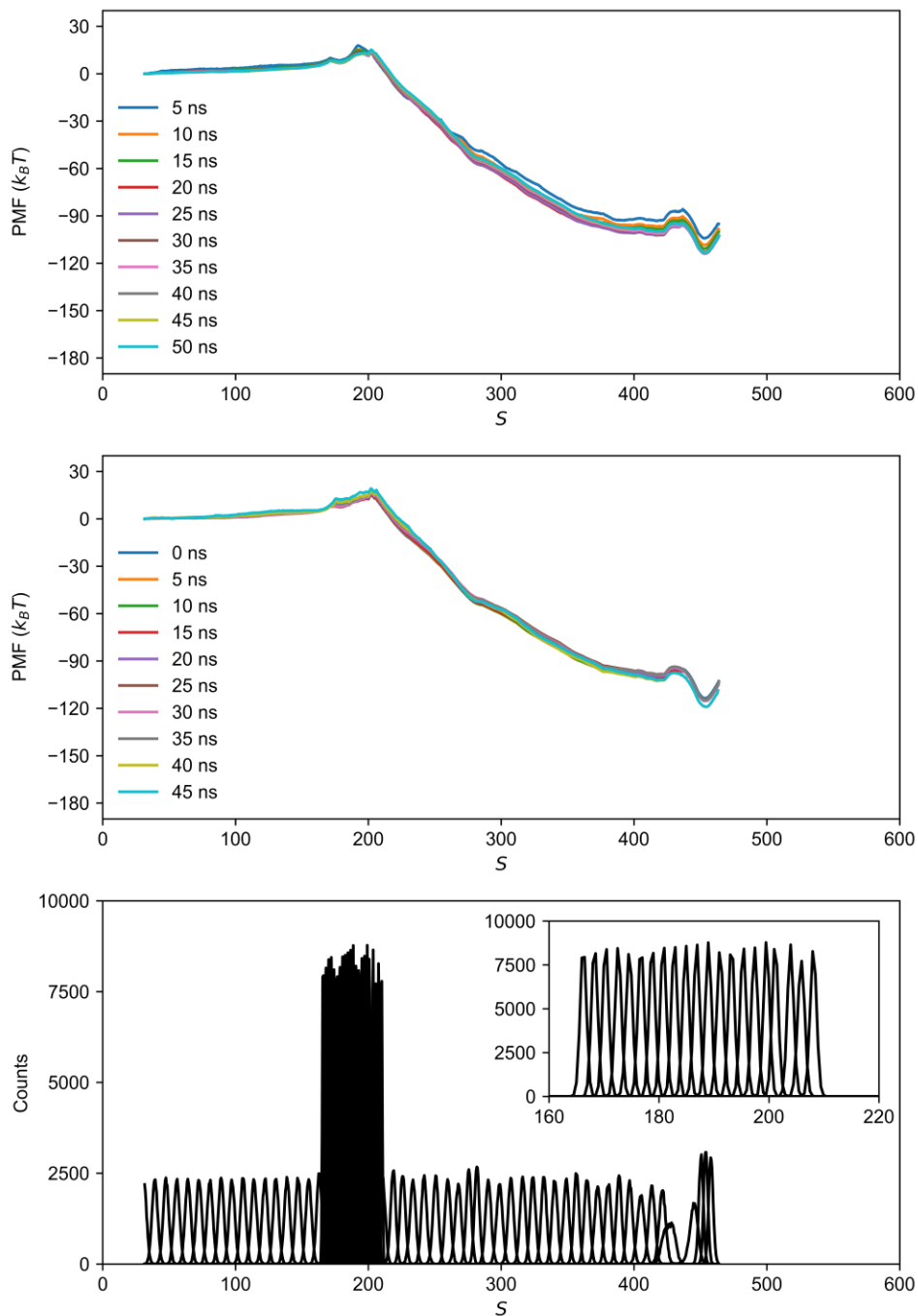
Figure S25. Potential of mean force (PMF) convergence for Replica #4 of the linear 11-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.

Figure S26. Potential of mean force (PMF) convergence for Replica #1 of the double-branched 11-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
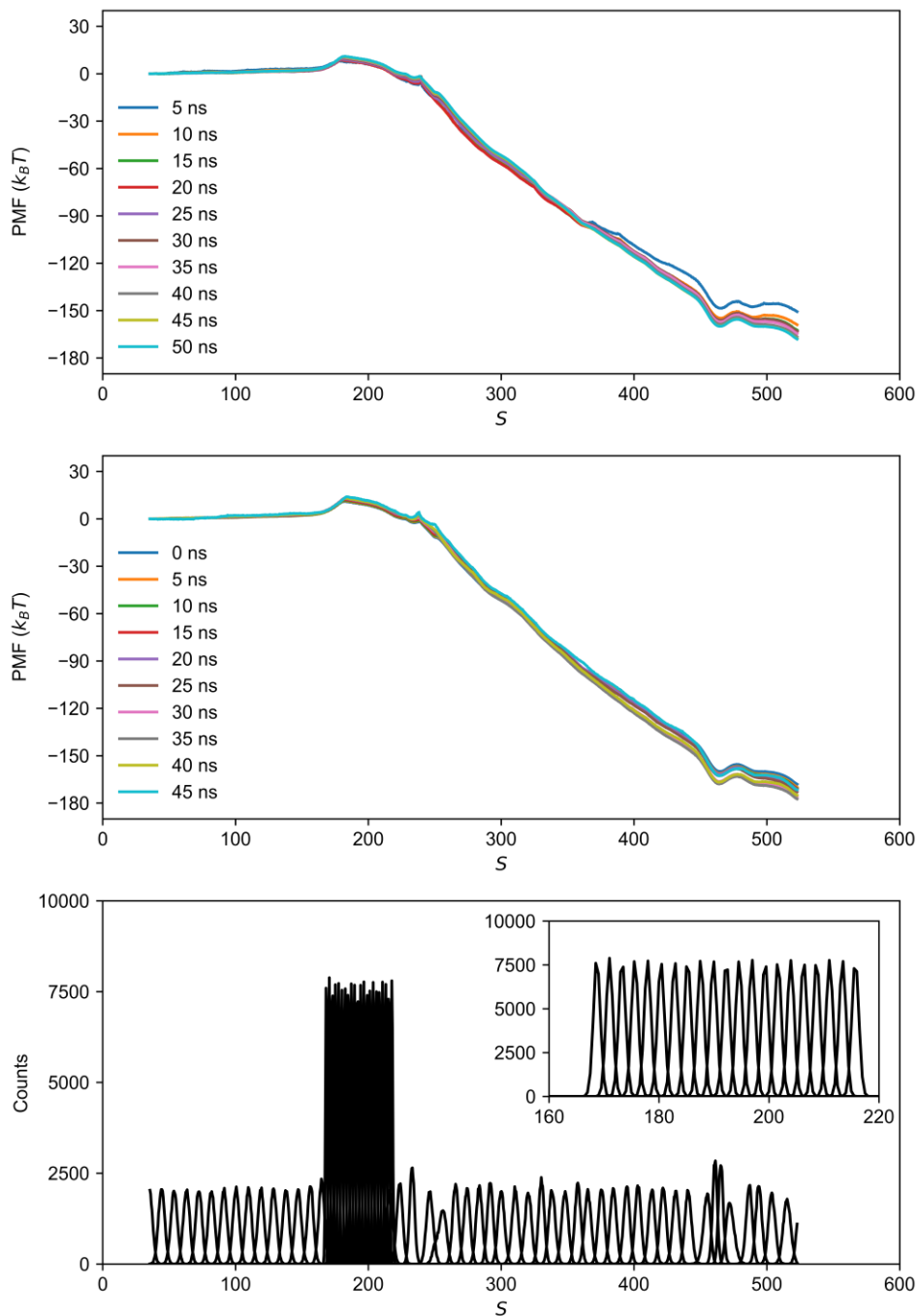
Figure S27. Potential of mean force (PMF) convergence for Replica #2 of the double-branched 11-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
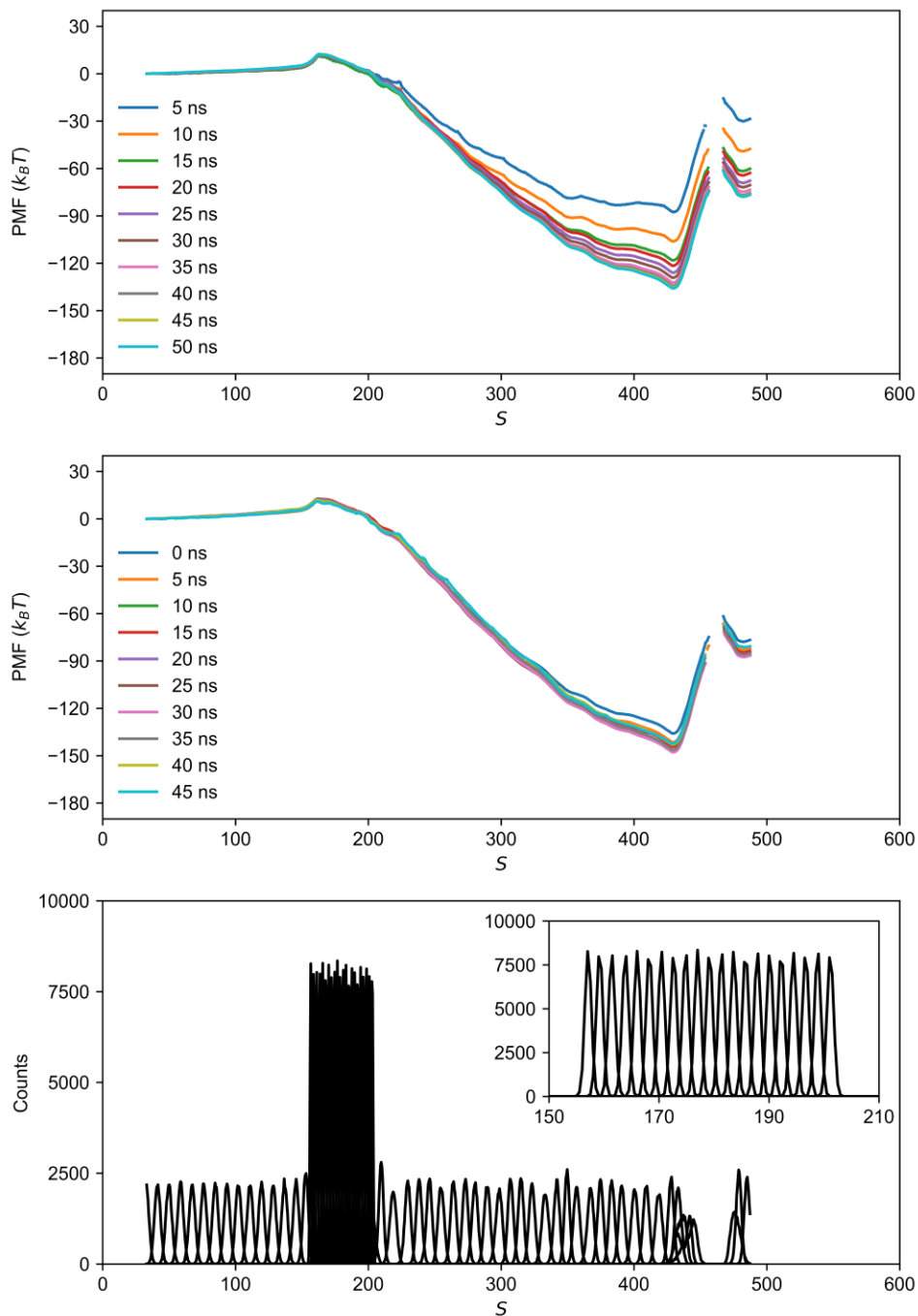
Figure S28. Potential of mean force (PMF) convergence for Replica #3 of the double-branched 11-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
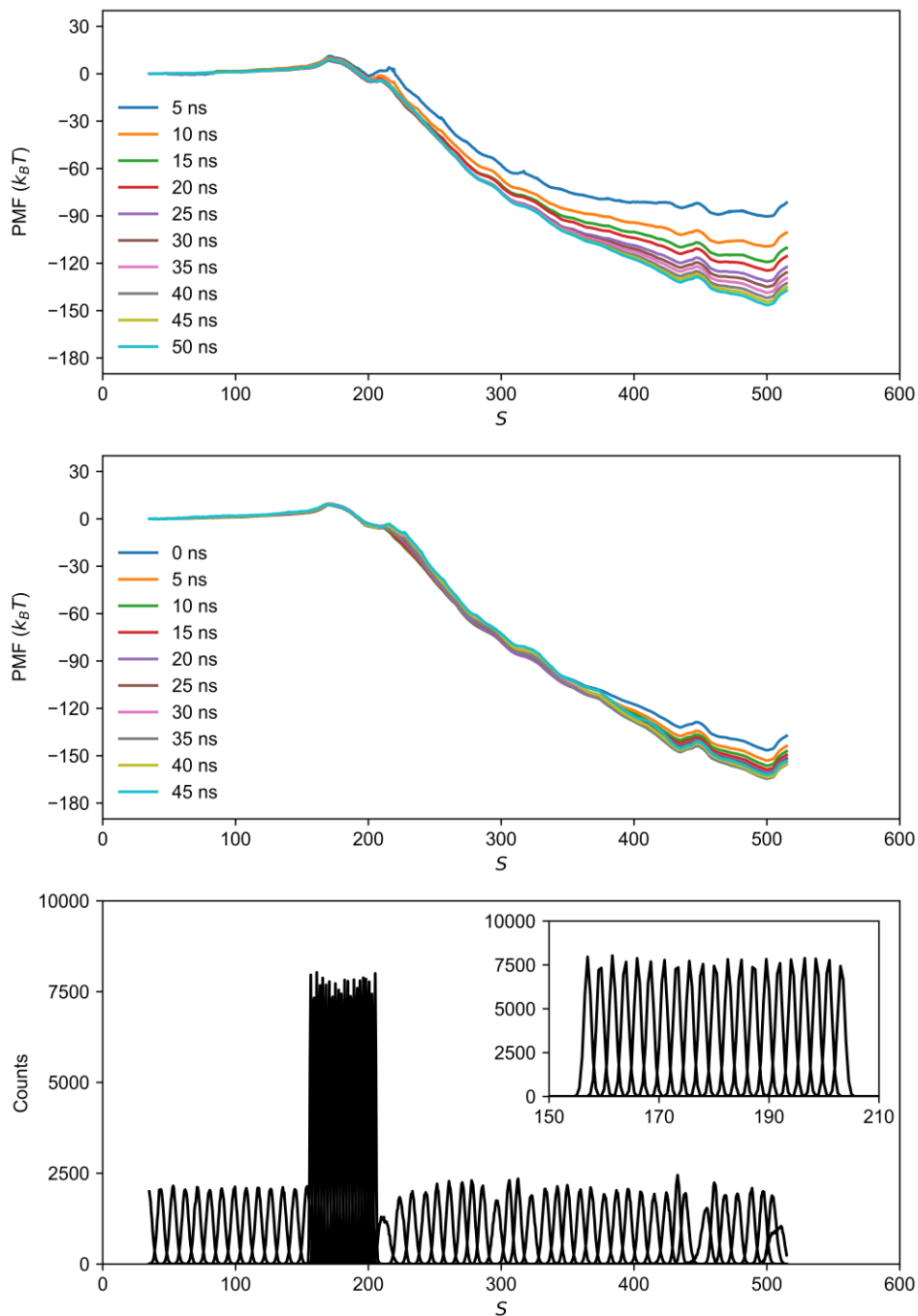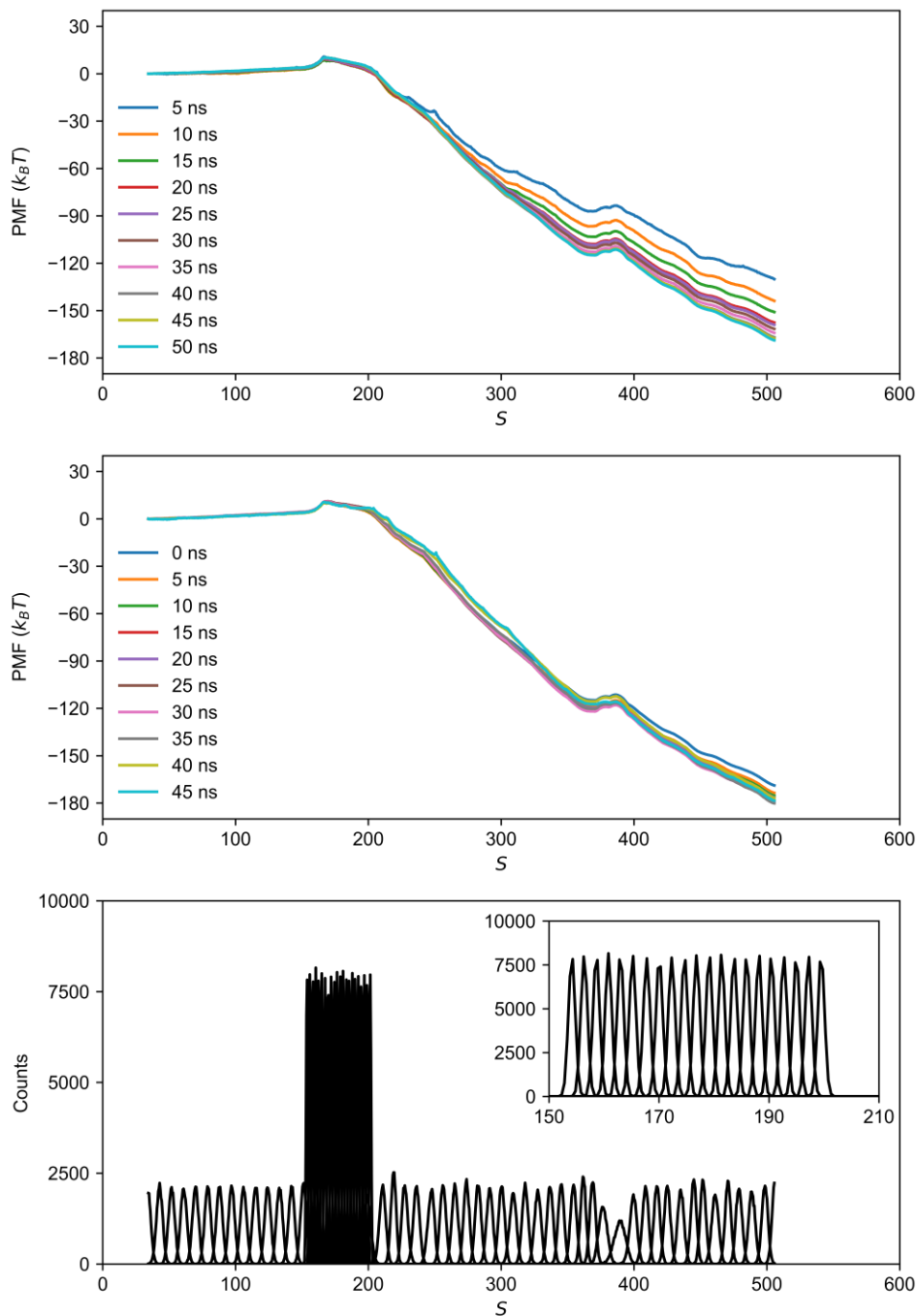
Figure S29. Potential of mean force (PMF) convergence for Replica #4 of the double-branched 11-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.

Figure S30. Potential of mean force (PMF) convergence for Replica #1 of the linear 15-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
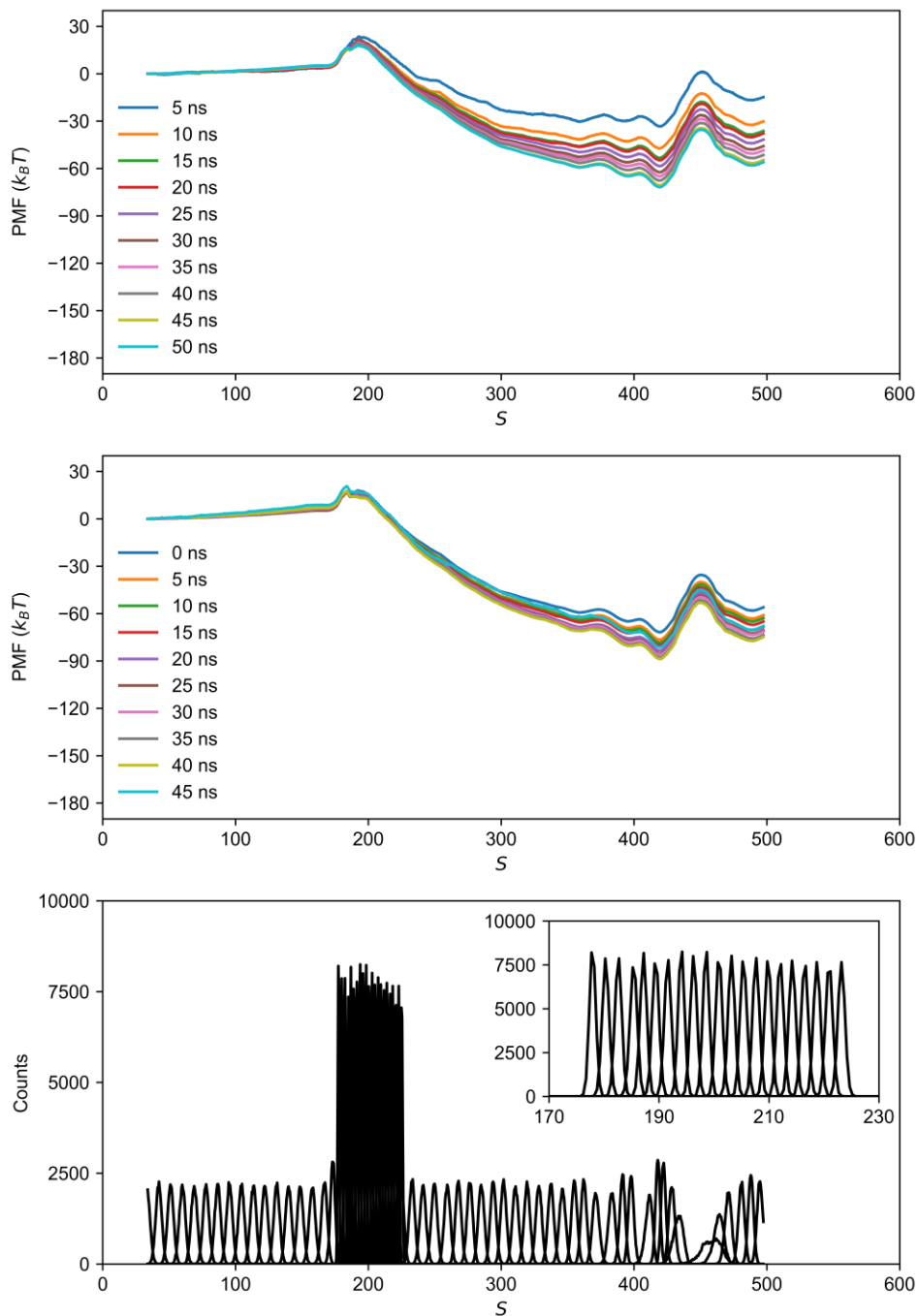
Figure S31. Potential of mean force (PMF) convergence for Replica #2 of the linear 15-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
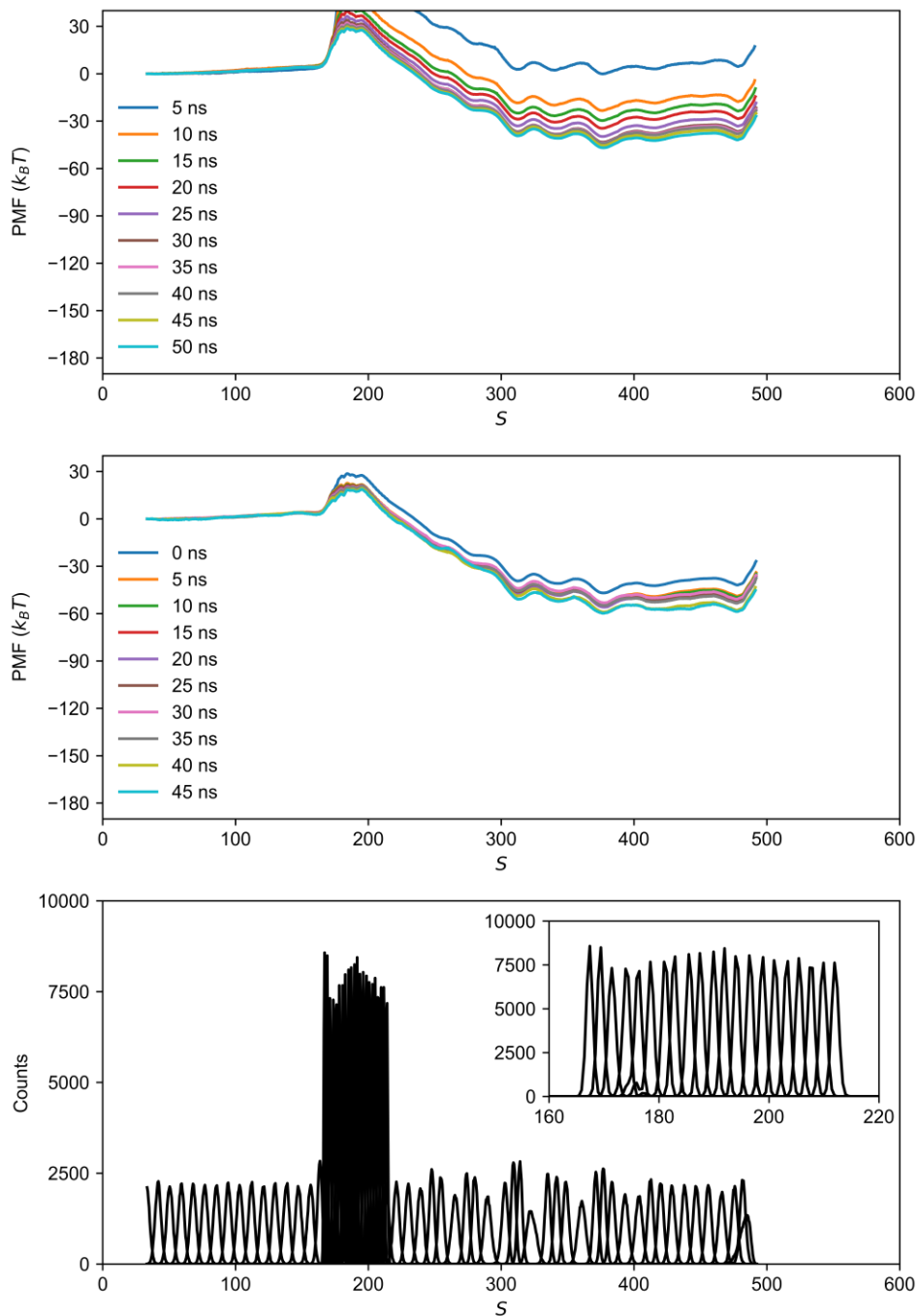
Figure S32. Potential of mean force (PMF) convergence for Replica #3 of the linear 15-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
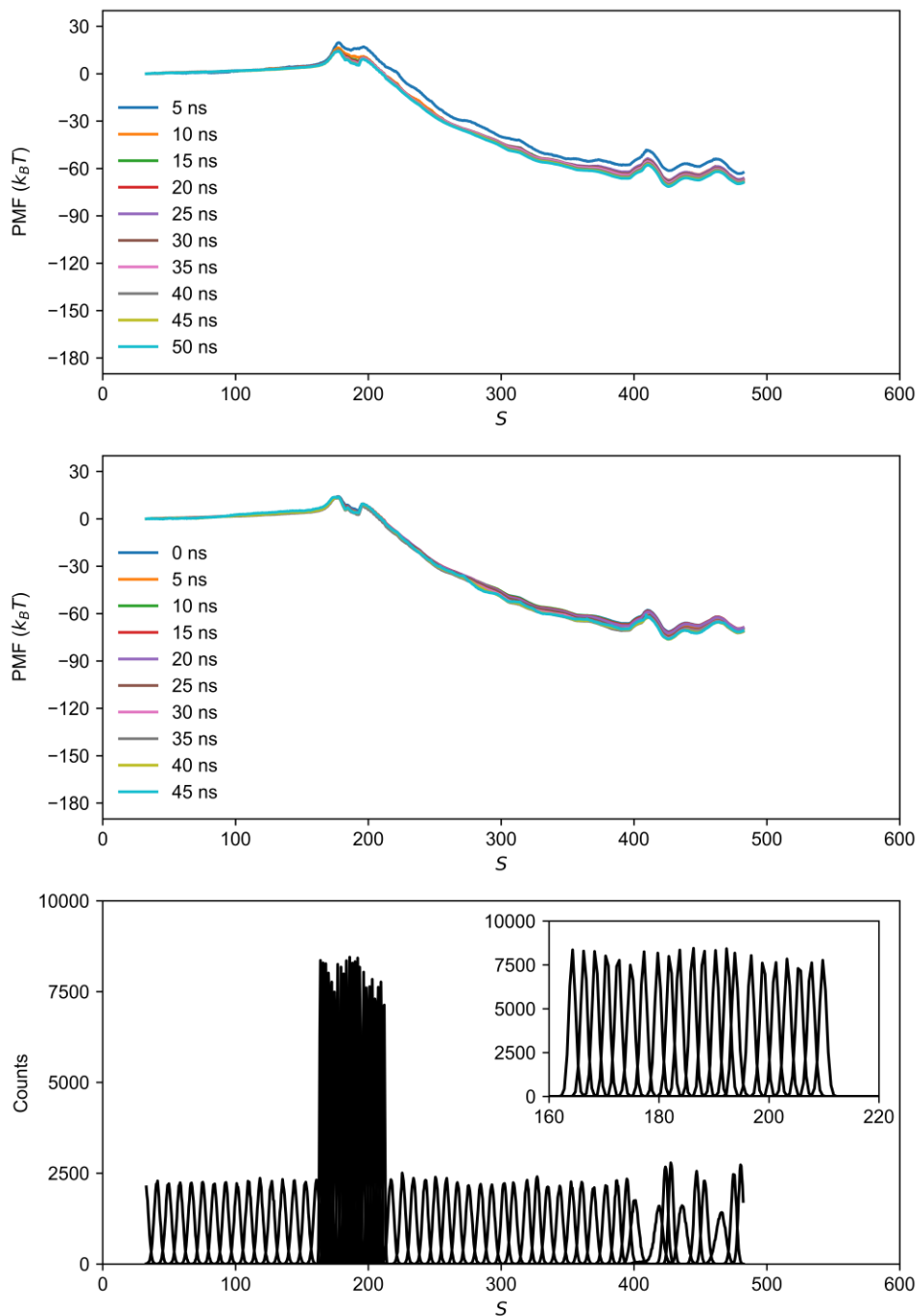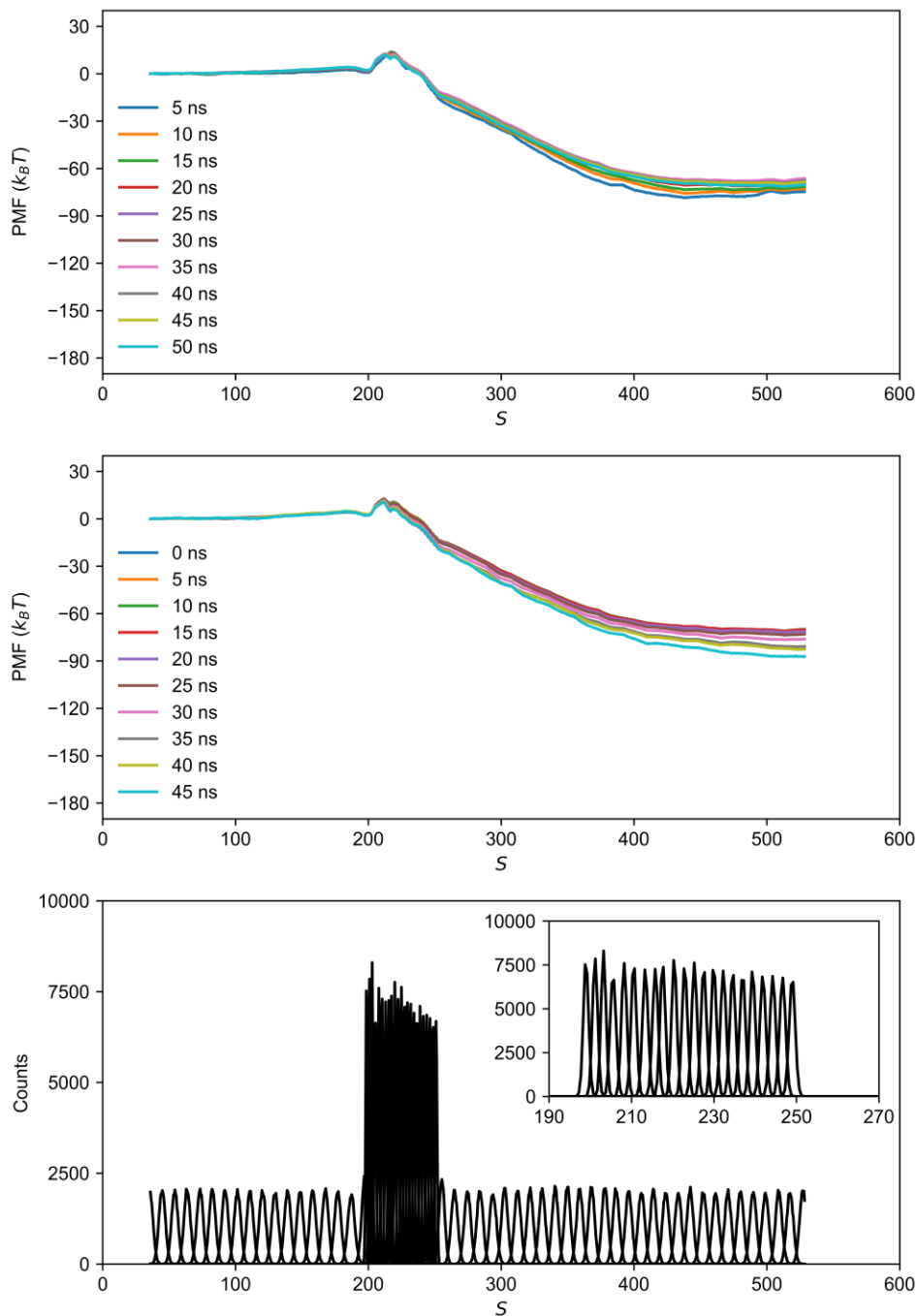
Figure S33. Potential of mean force (PMF) convergence for Replica #4 of the linear 15-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.

Figure S34. Potential of mean force (PMF) convergence for Replica #1 of the triple-branched 15-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of *S* of interest.

Figure S35. Potential of mean force (PMF) convergence for Replica #2 of the triple-branched 15-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
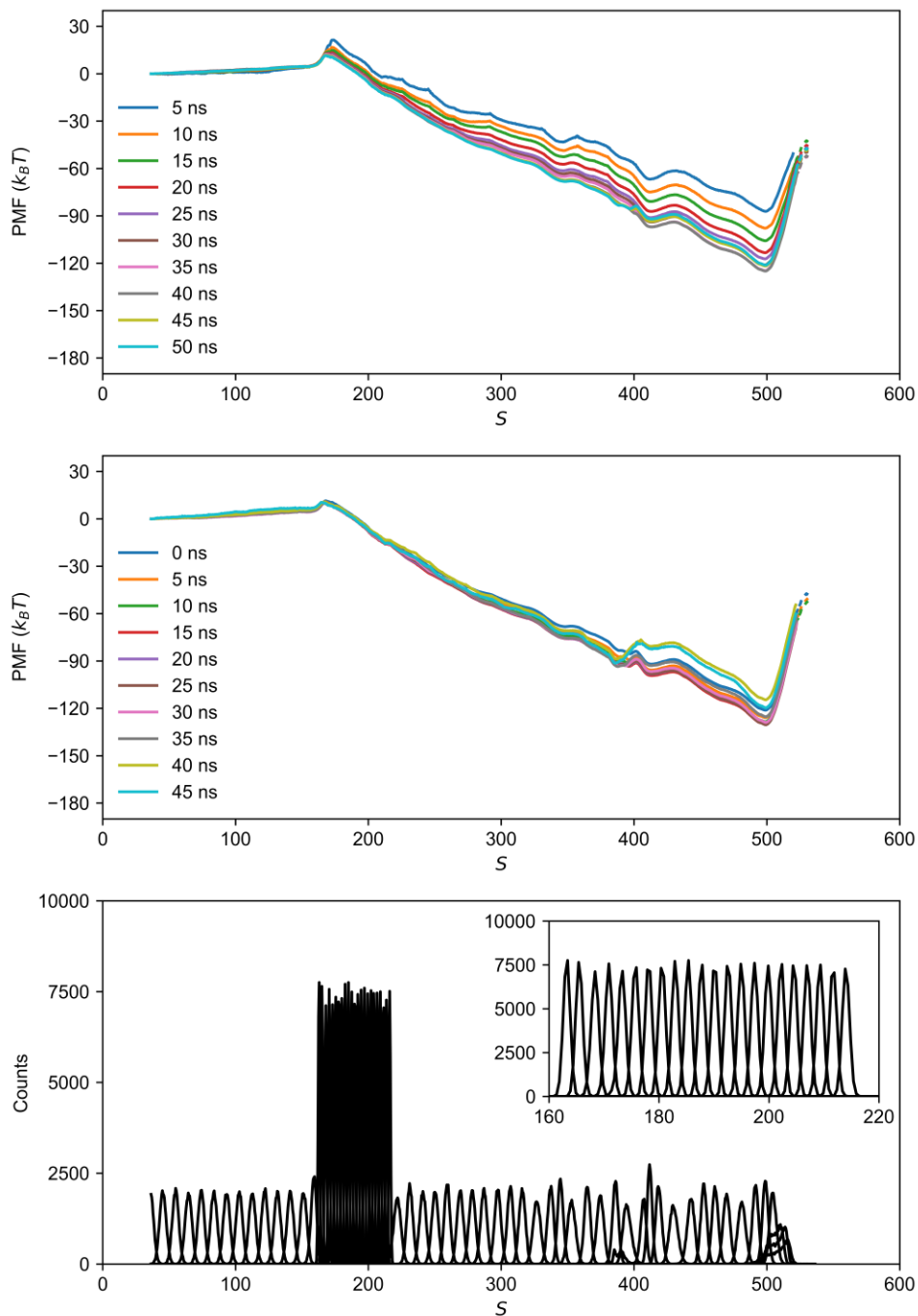
Figure S36. Potential of mean force (PMF) convergence for Replica #3 of the triple-branched 15-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
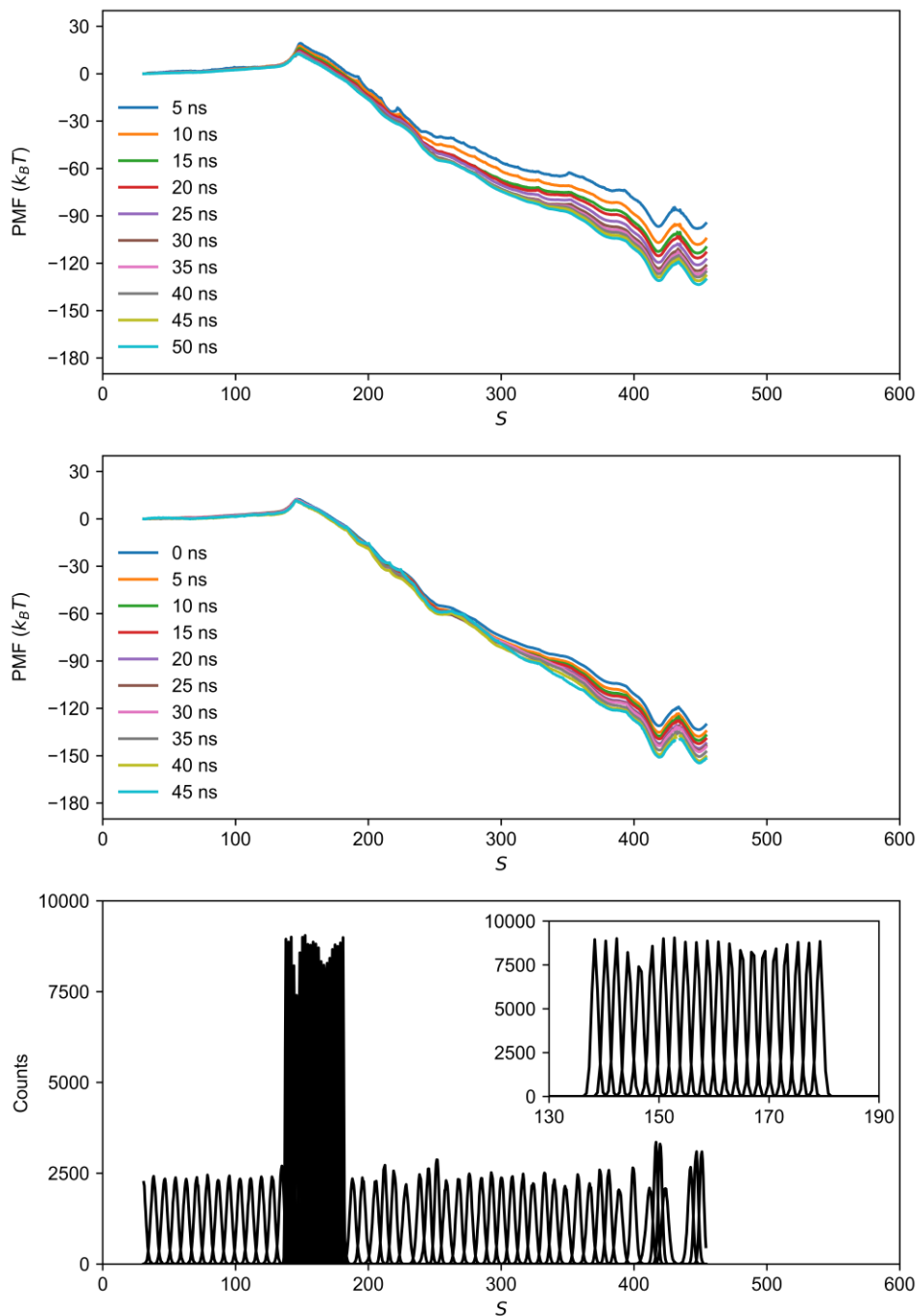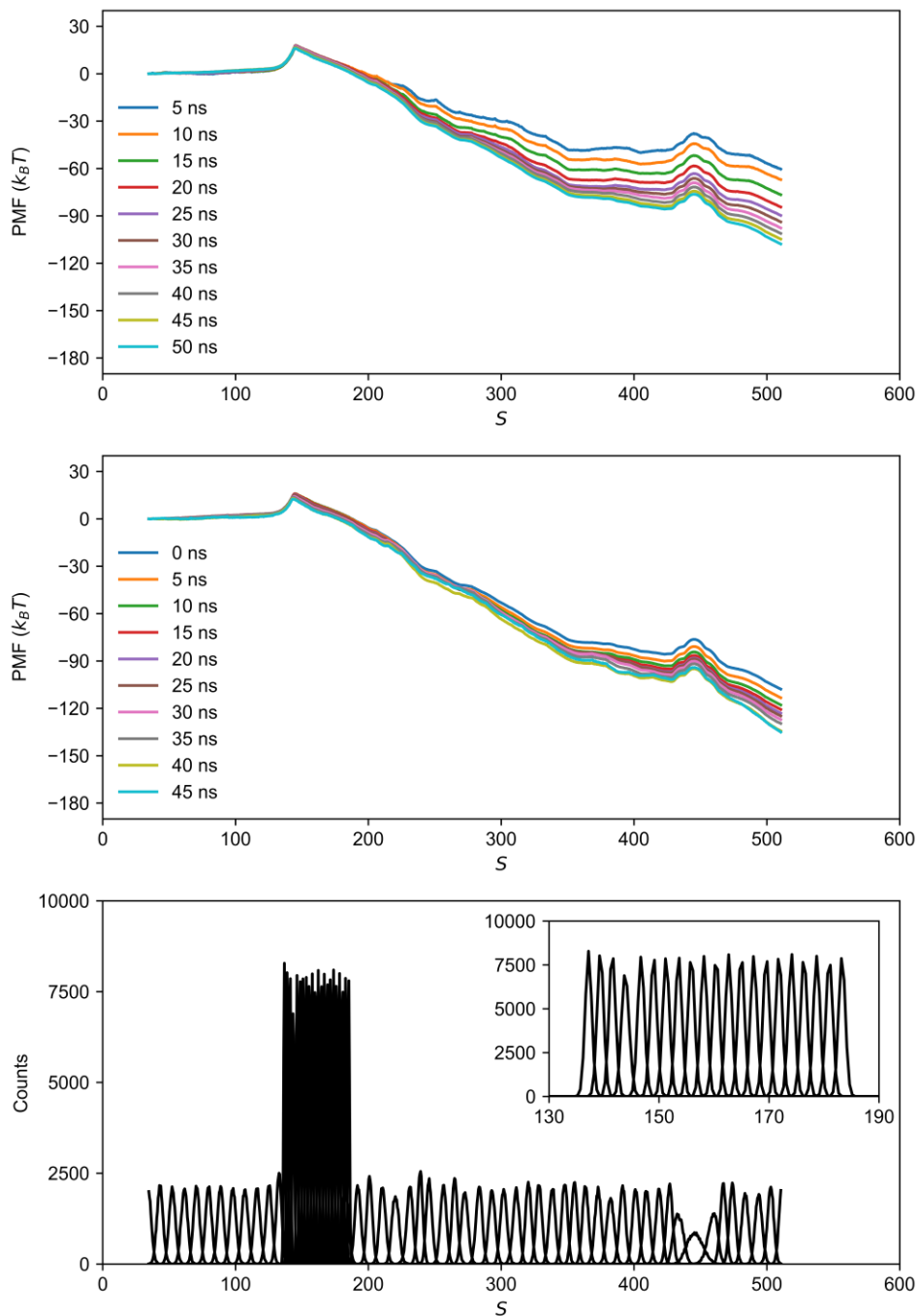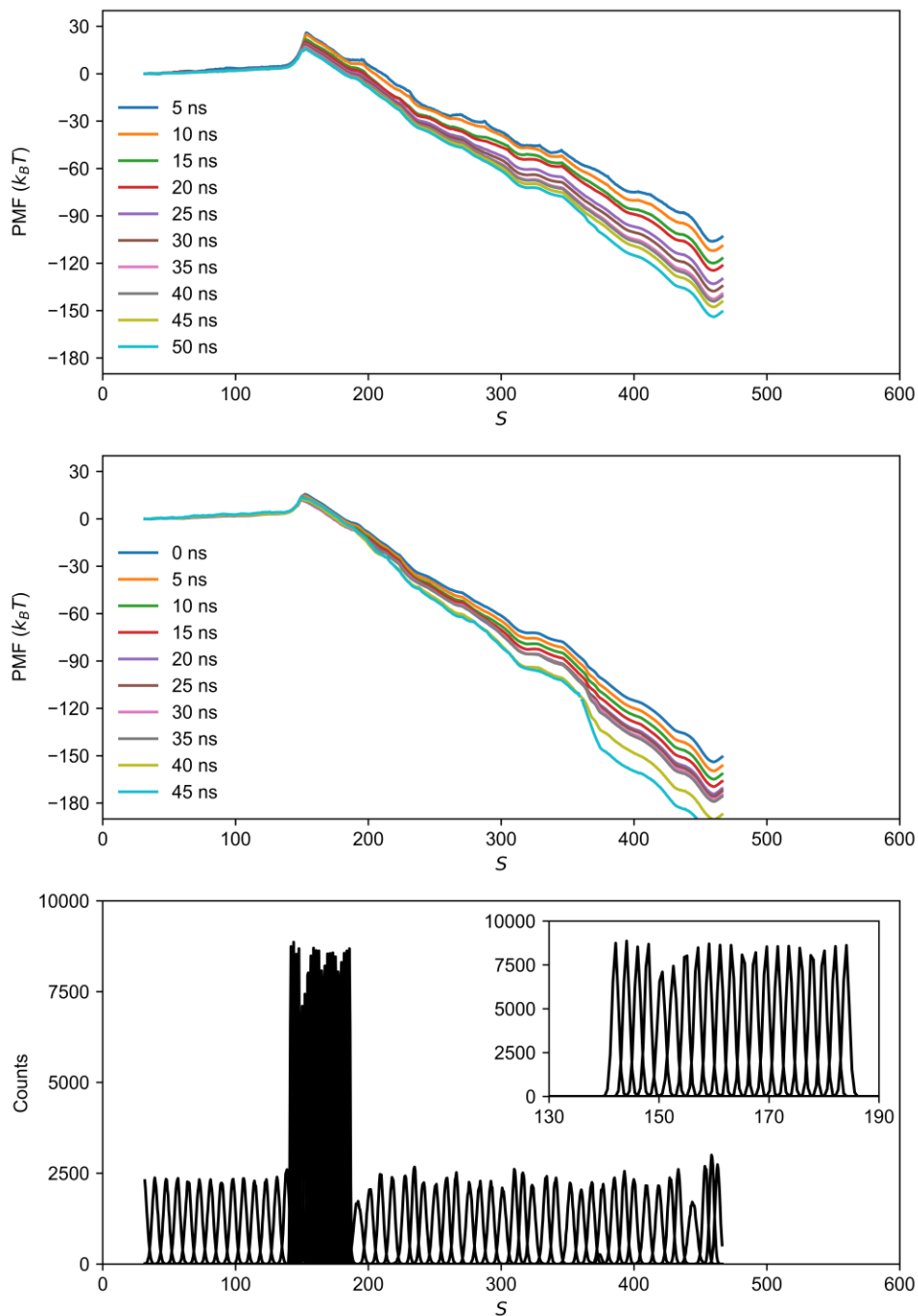
Figure S37. Potential of mean force (PMF) convergence for Replica #4 of the triple-branched 15-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.

Figure S38. Potential of mean force (PMF) convergence for Replica #1 of the linear 19-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of *S* of interest.

Figure S39. Potential of mean force (PMF) convergence for Replica #2 of the linear 19-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
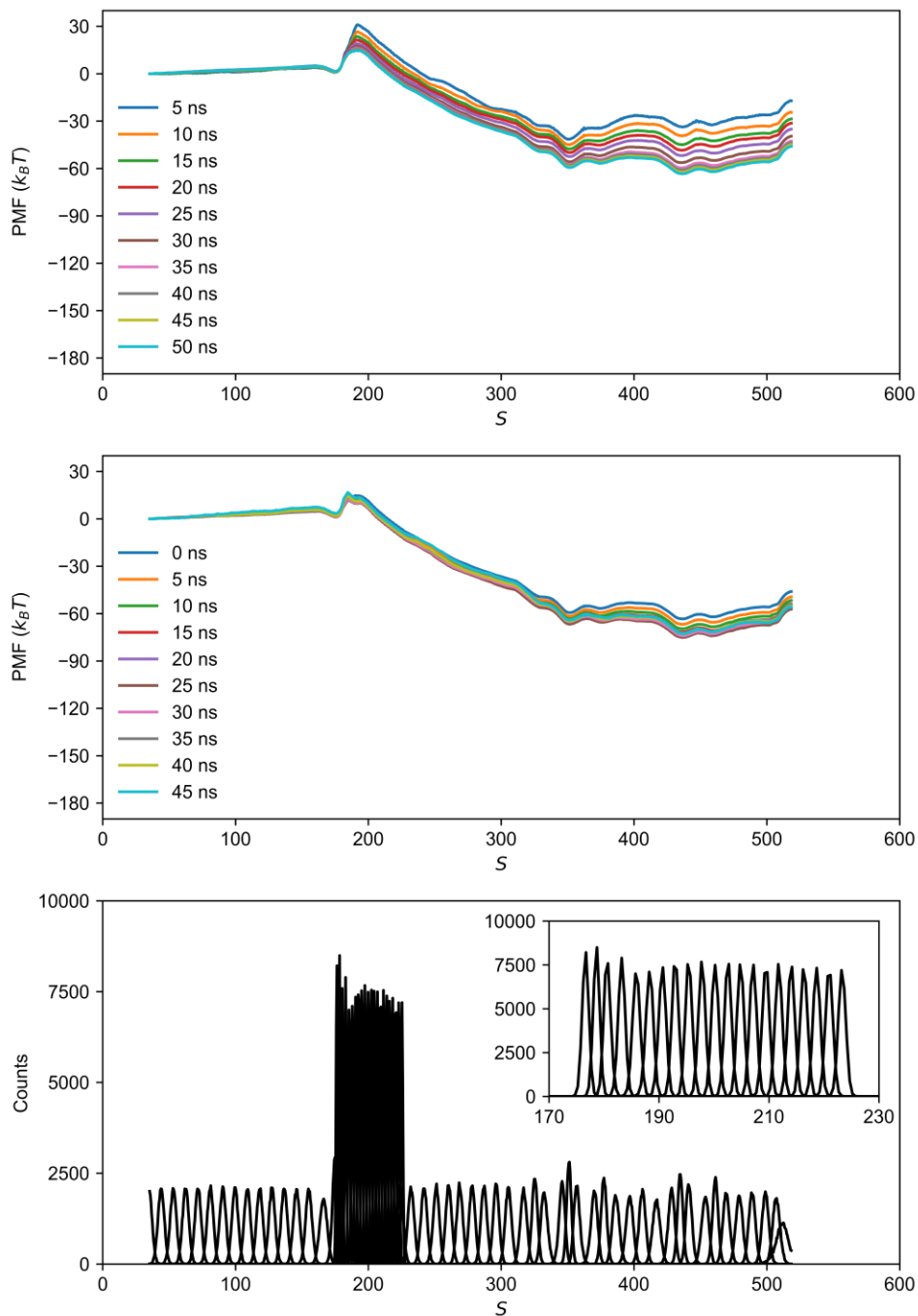
Figure S40. Potential of mean force (PMF) convergence for Replica #3 of the linear 19-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
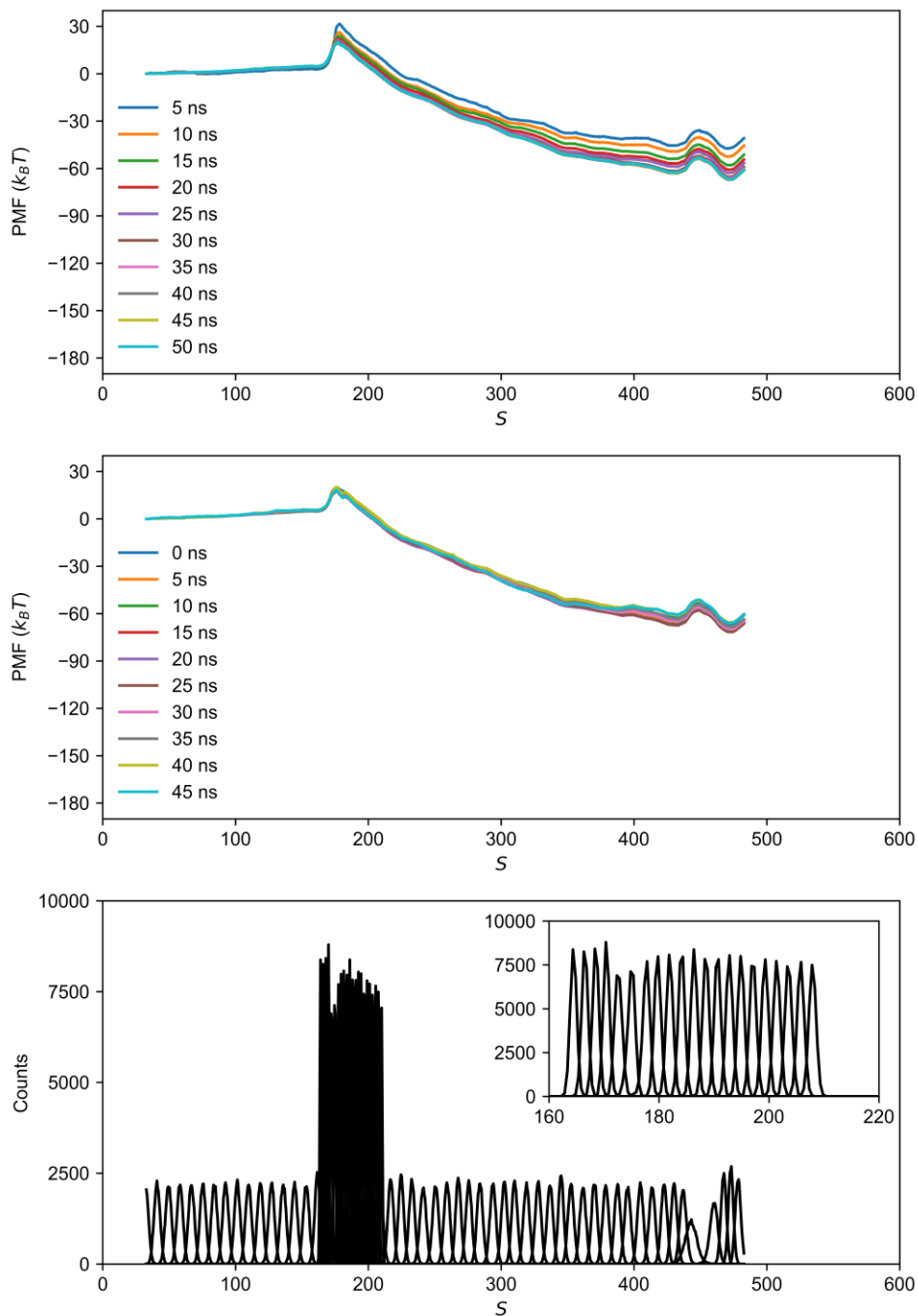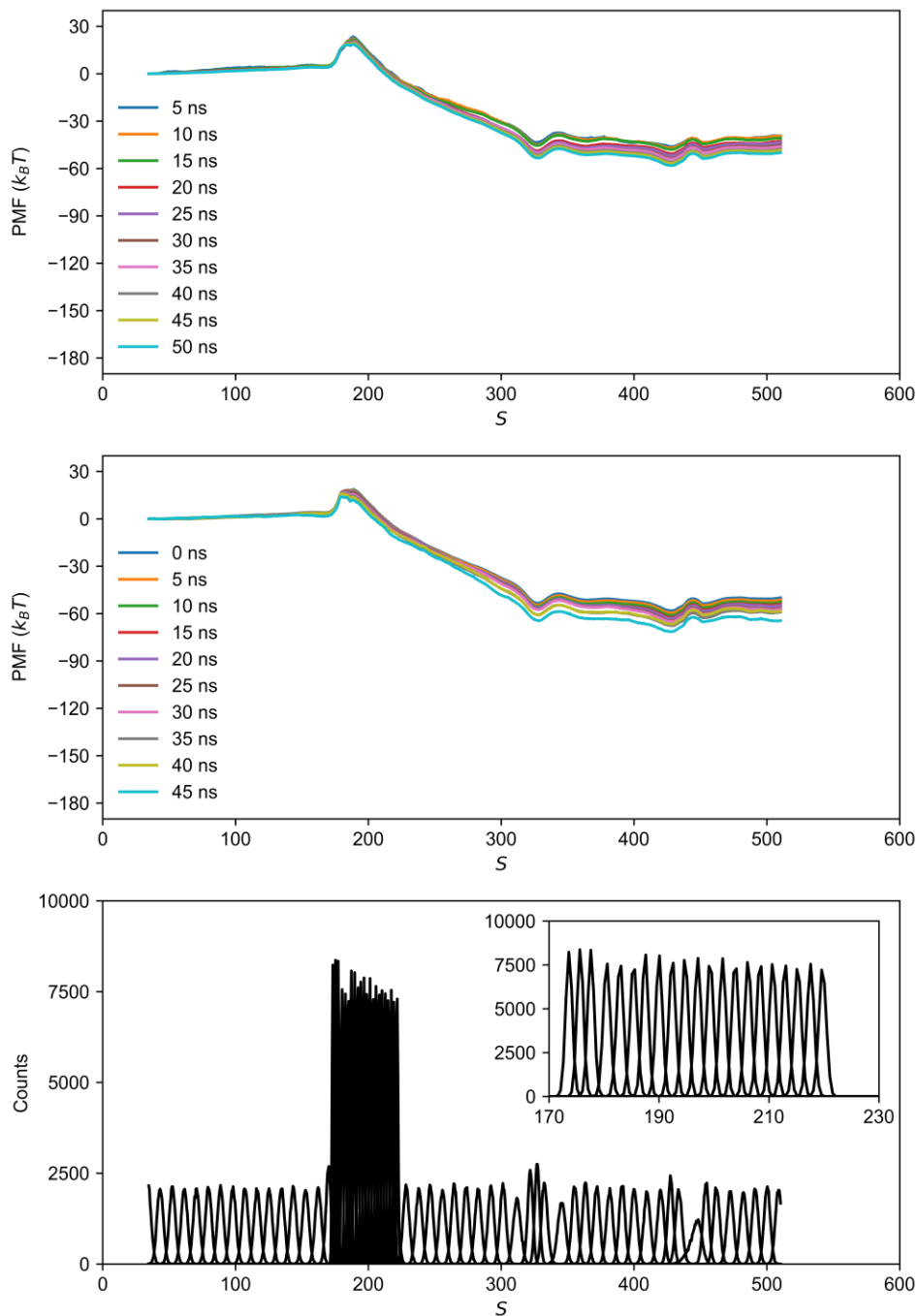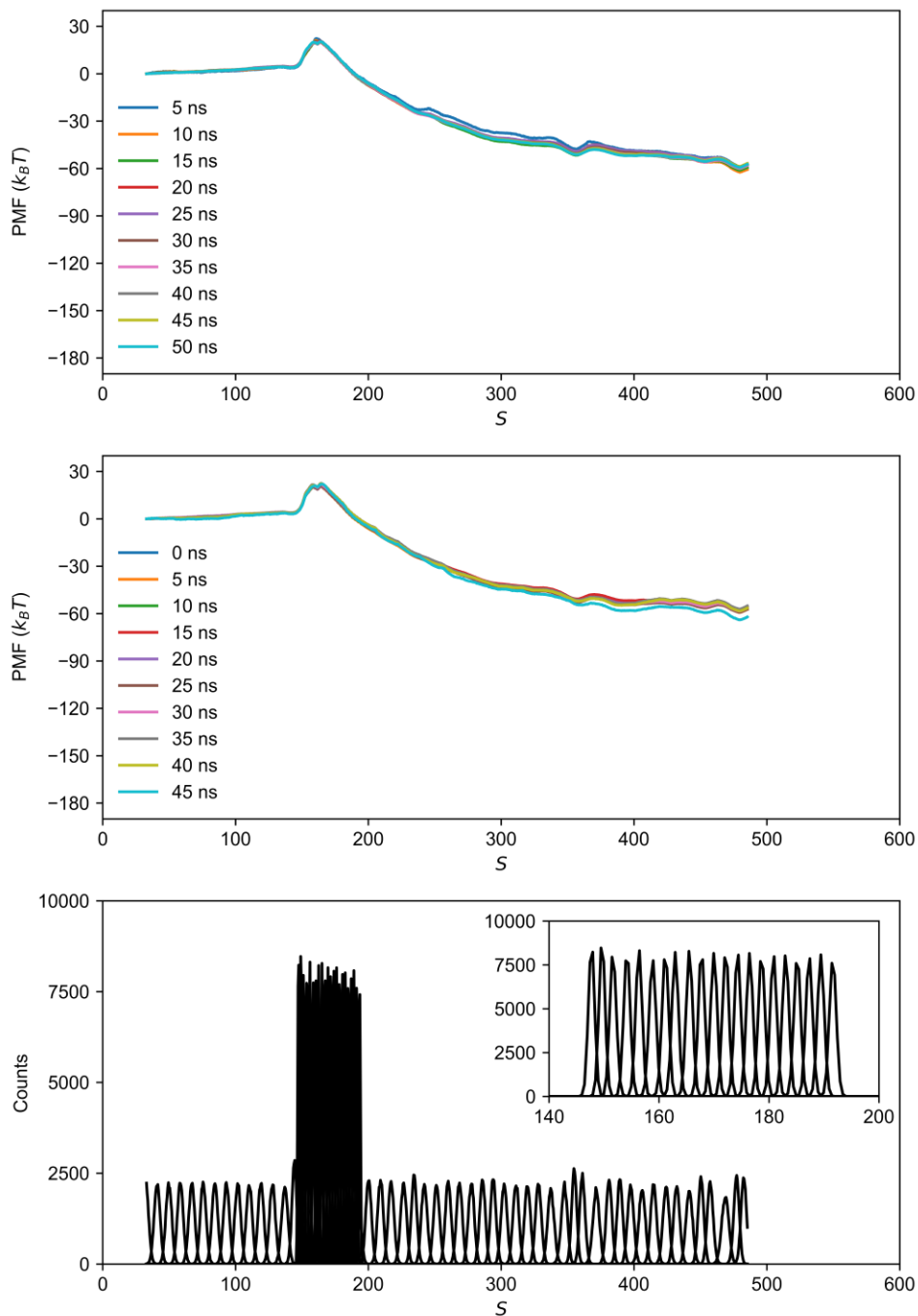
Figure S41. Potential of mean force (PMF) convergence for Replica #4 of the linear 19-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.

Figure S42. Potential of mean force (PMF) convergence for Replica #1 of the double-branched 19-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.

Figure S43. Potential of mean force (PMF) convergence for Replica #2 of the double-branched 19-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
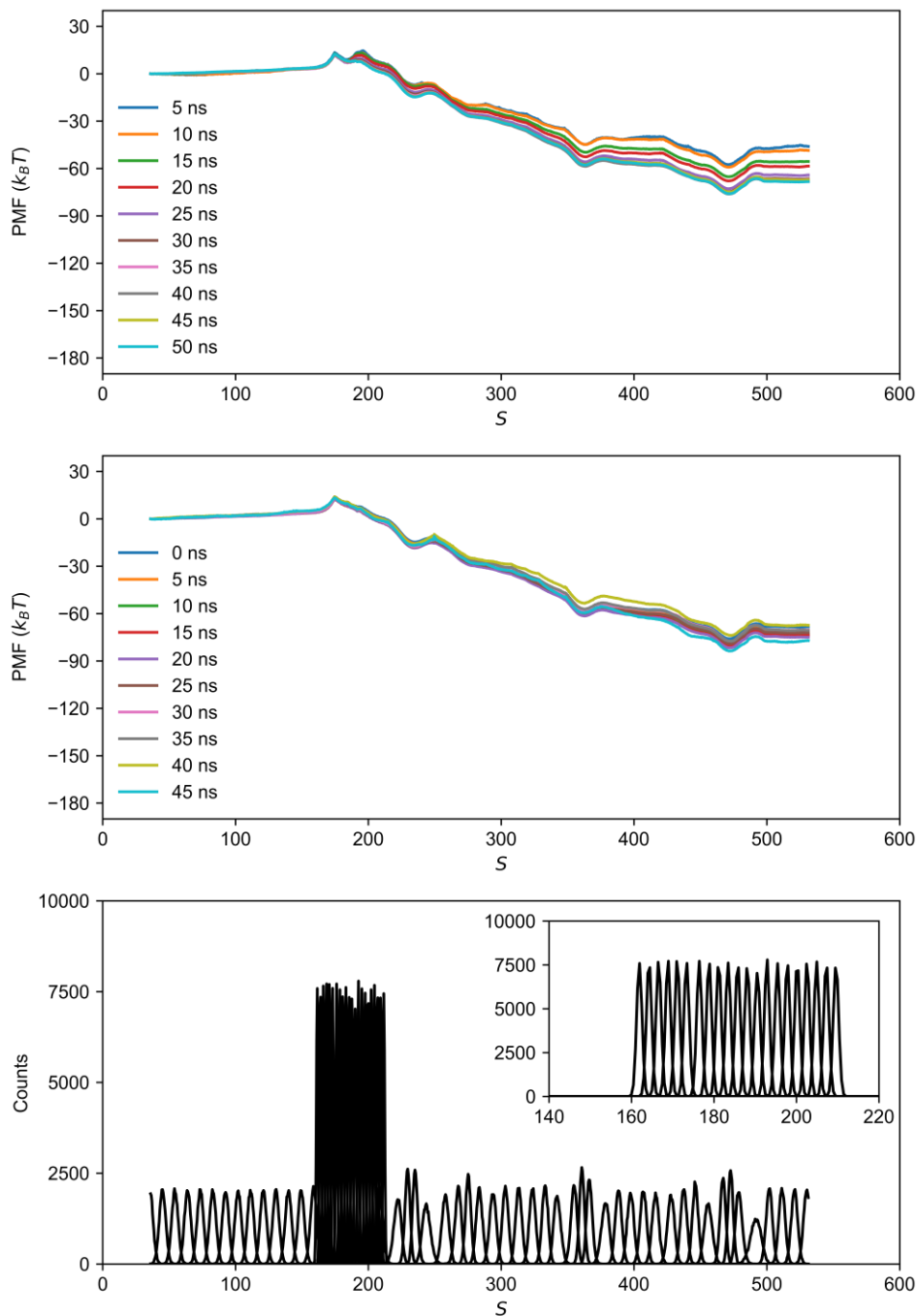
Figure S44. Potential of mean force (PMF) convergence for Replica #3 of the double-branched 19-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
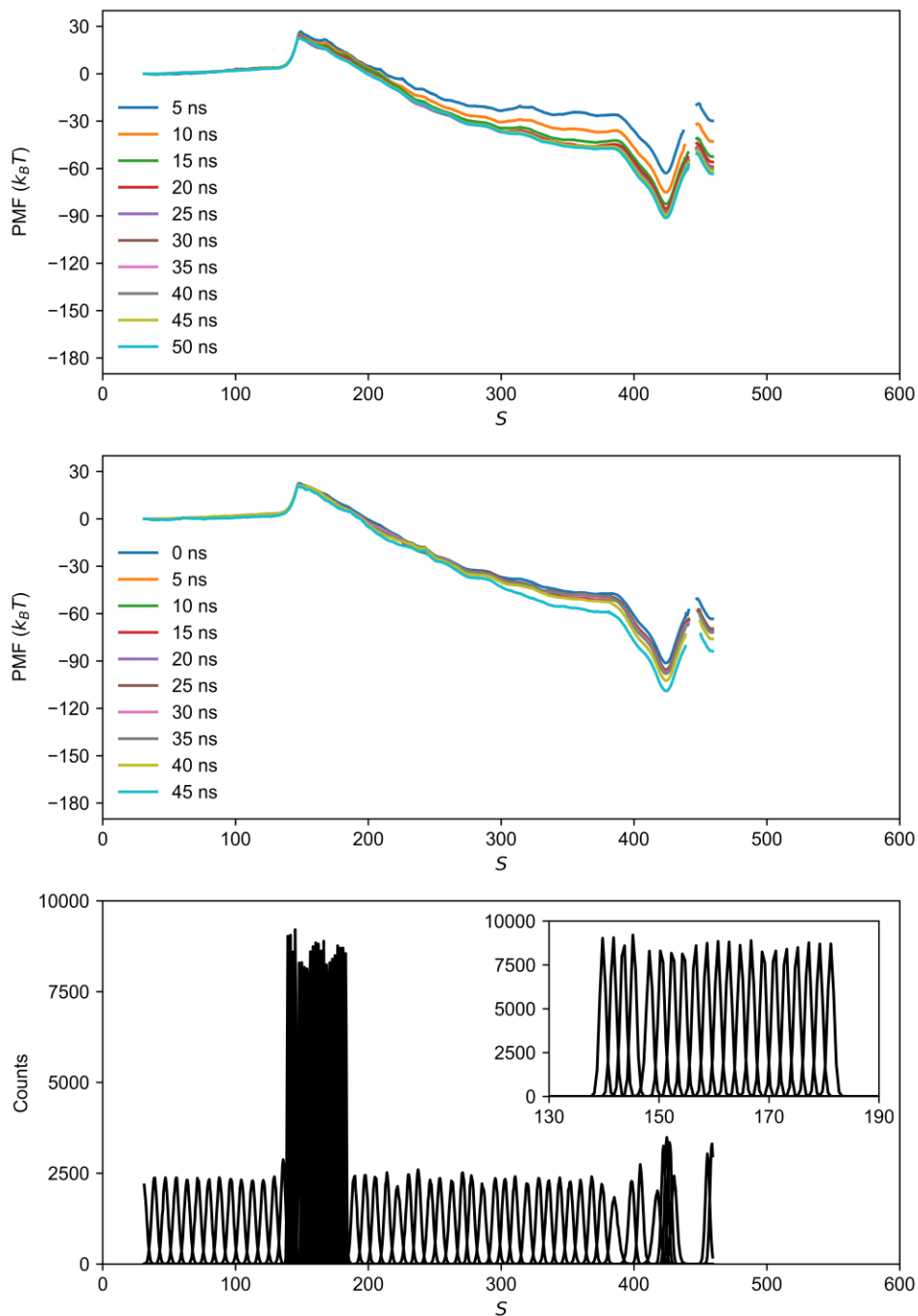
Figure S45. Potential of mean force (PMF) convergence for Replica #4 of the double-branched 19-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.

Figure S46. Potential of mean force (PMF) convergence for Replica #1 of the double-branched 27-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
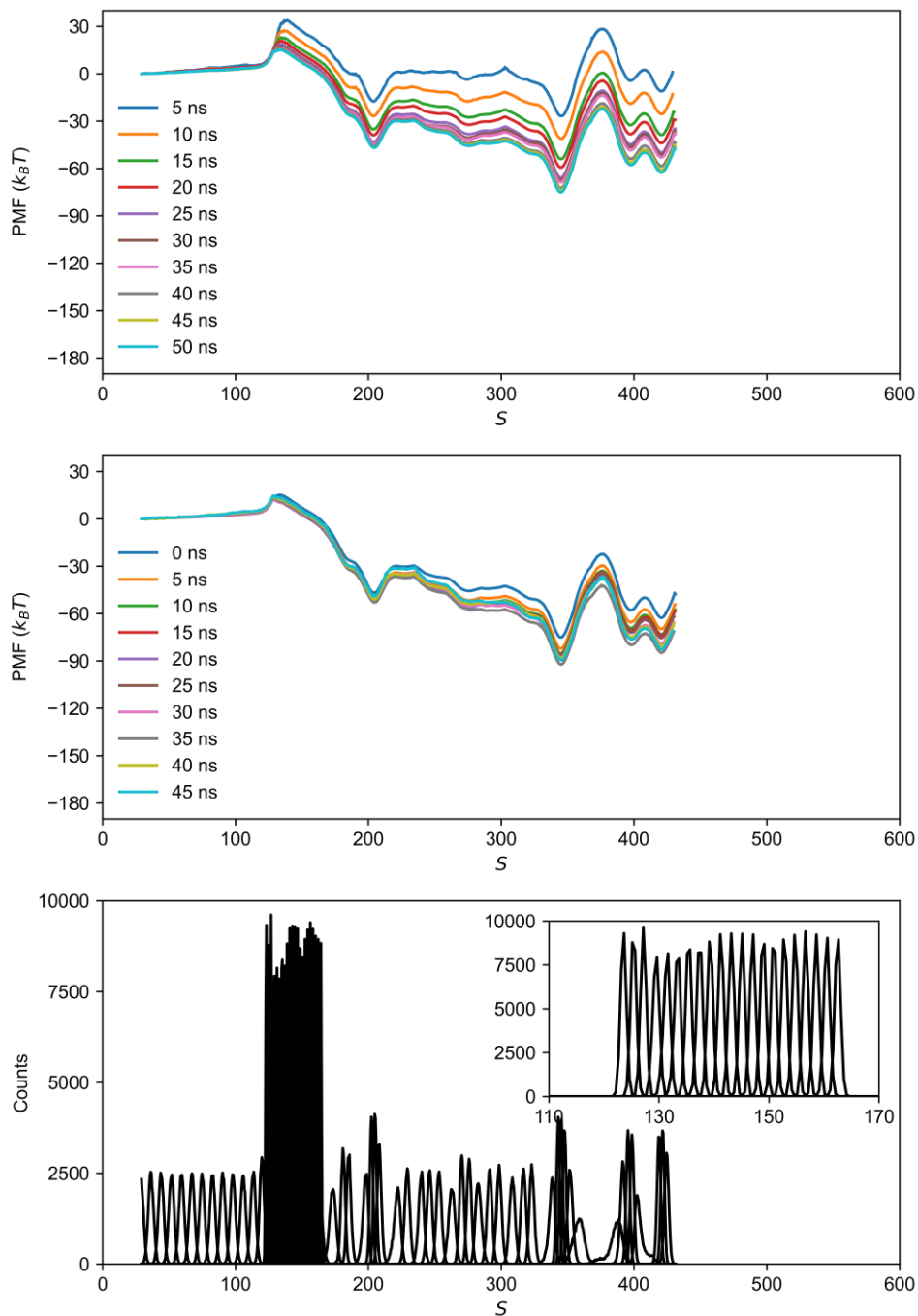
Figure S47. Potential of mean force (PMF) convergence for Replica #2 of the double-branched 27-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
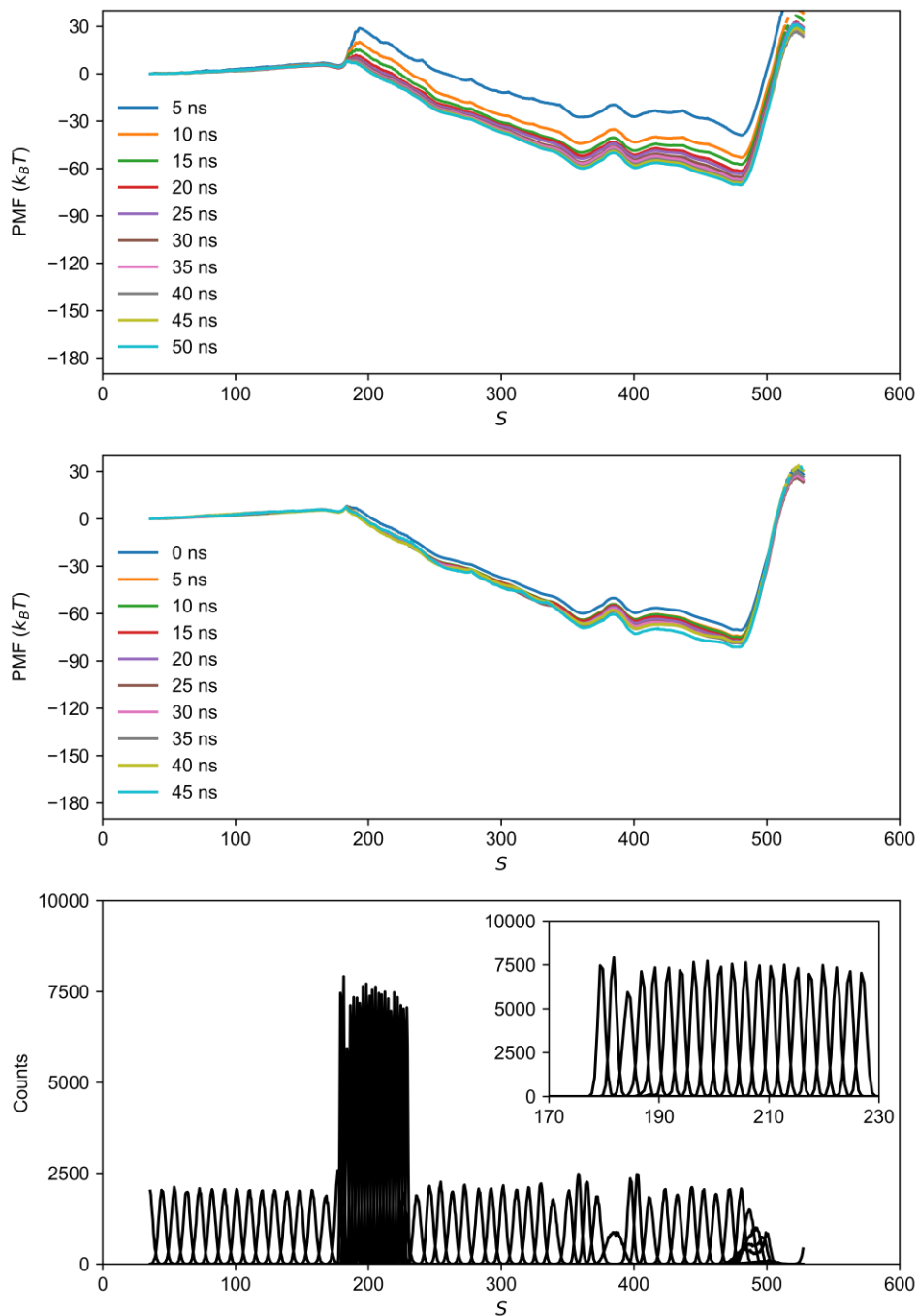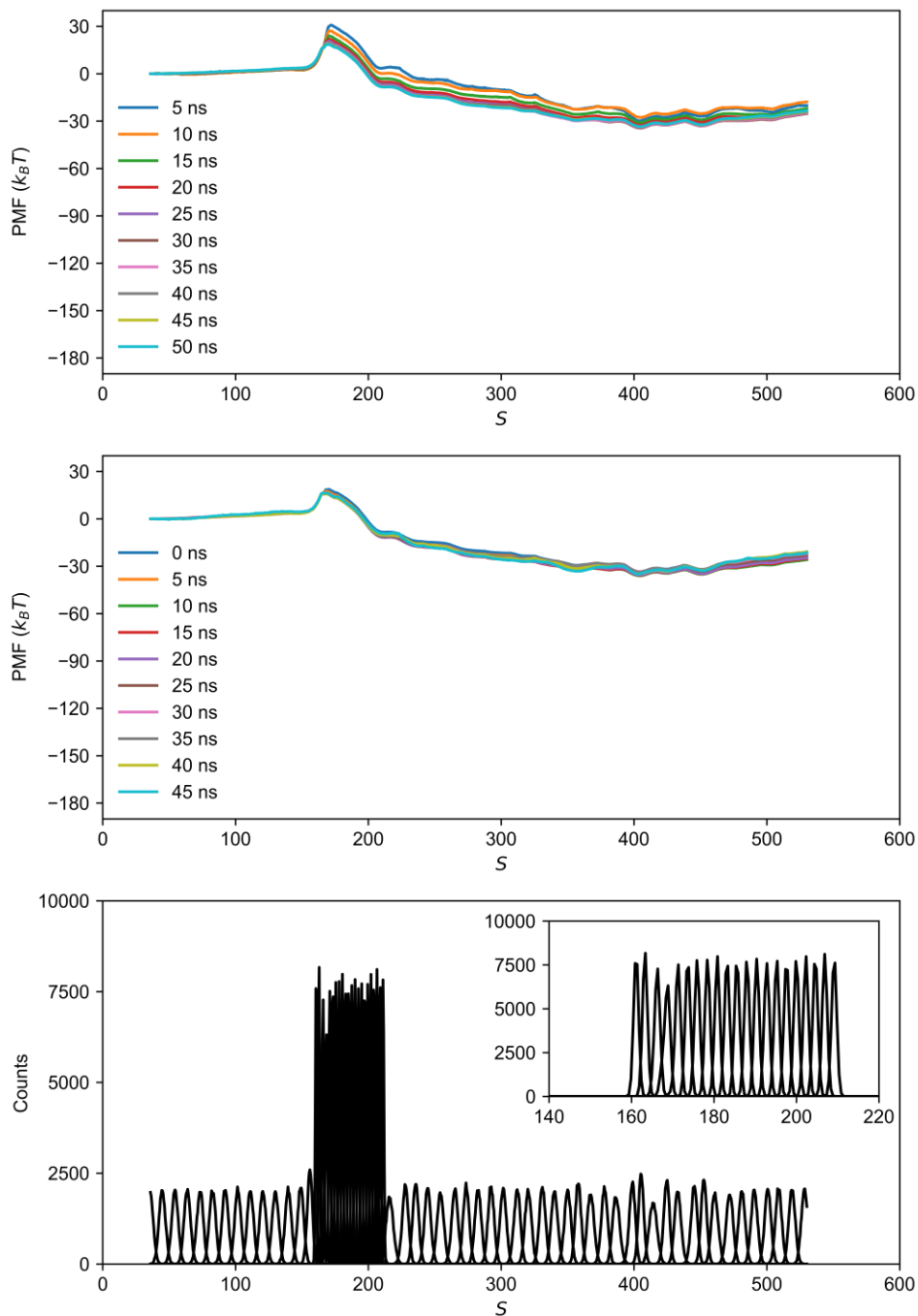
Figure S48. Potential of mean force (PMF) convergence for Replica #3 of the double-branched 27-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
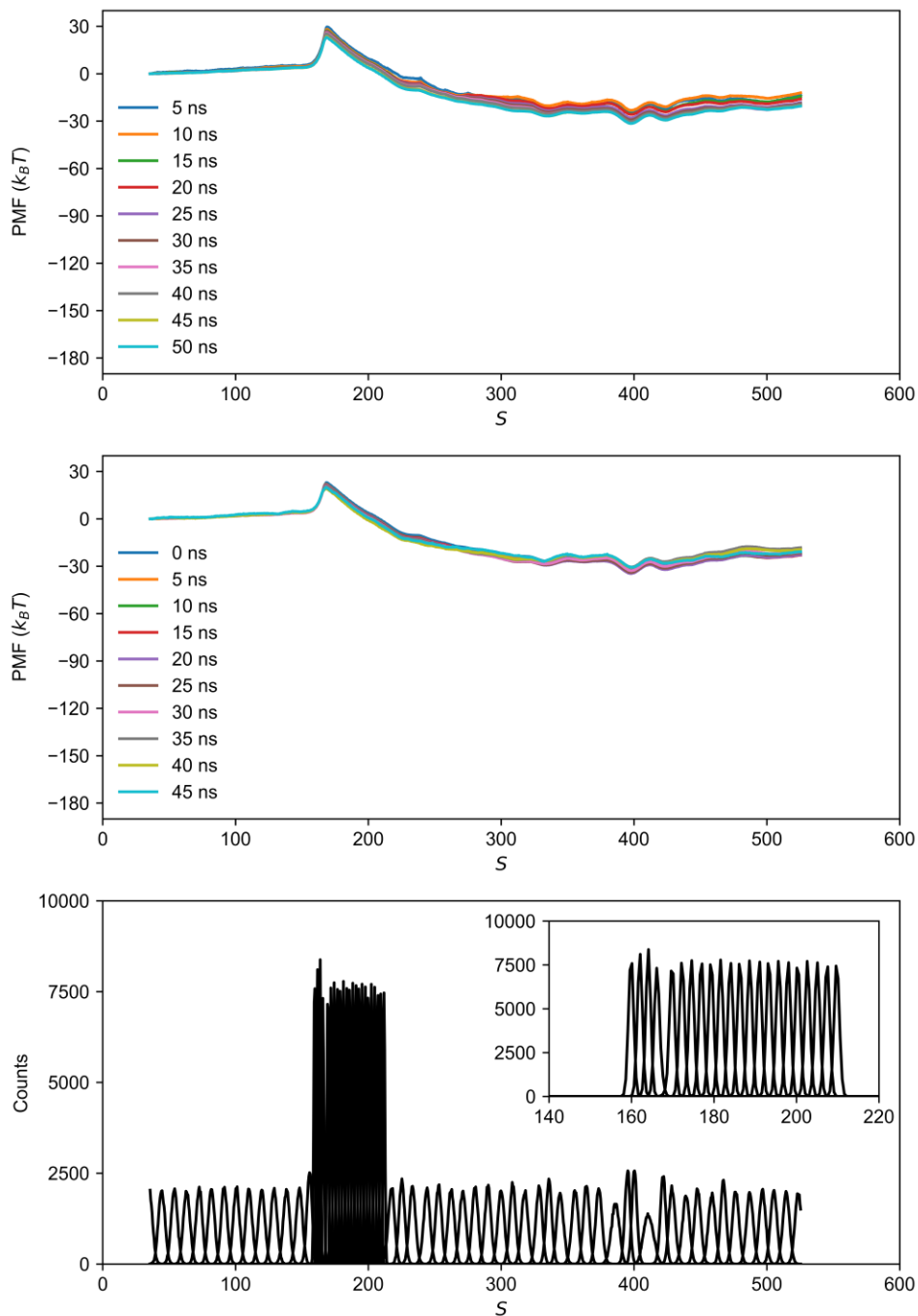
Figure S49. Potential of mean force (PMF) convergence for Replica #4 of the double-branched 27-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.

Figure S50. Potential of mean force (PMF) convergence for Replica #1 of the triple-branched 27-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
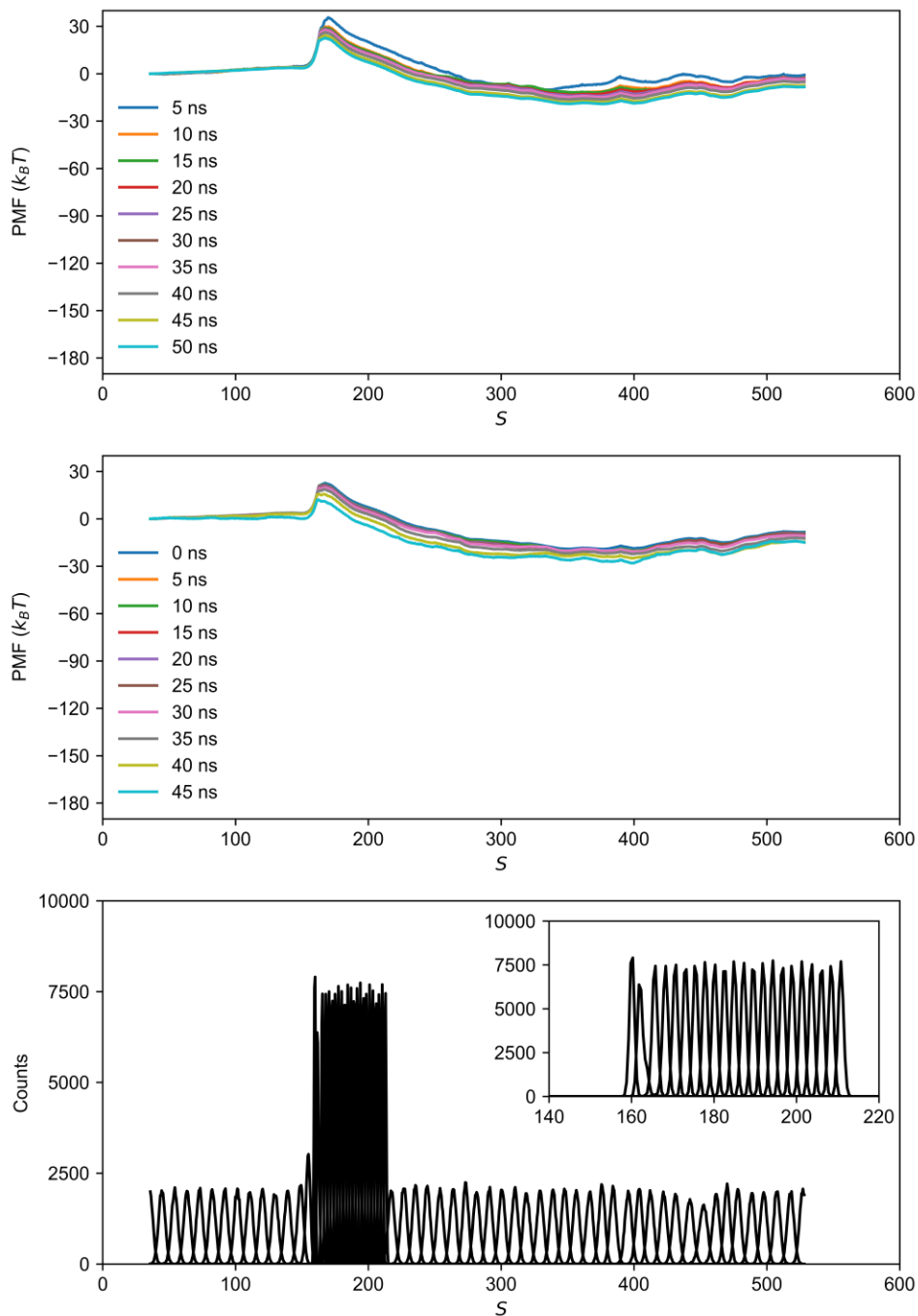
Figure S51. Potential of mean force (PMF) convergence for Replica #2 of the triple-branched 27-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
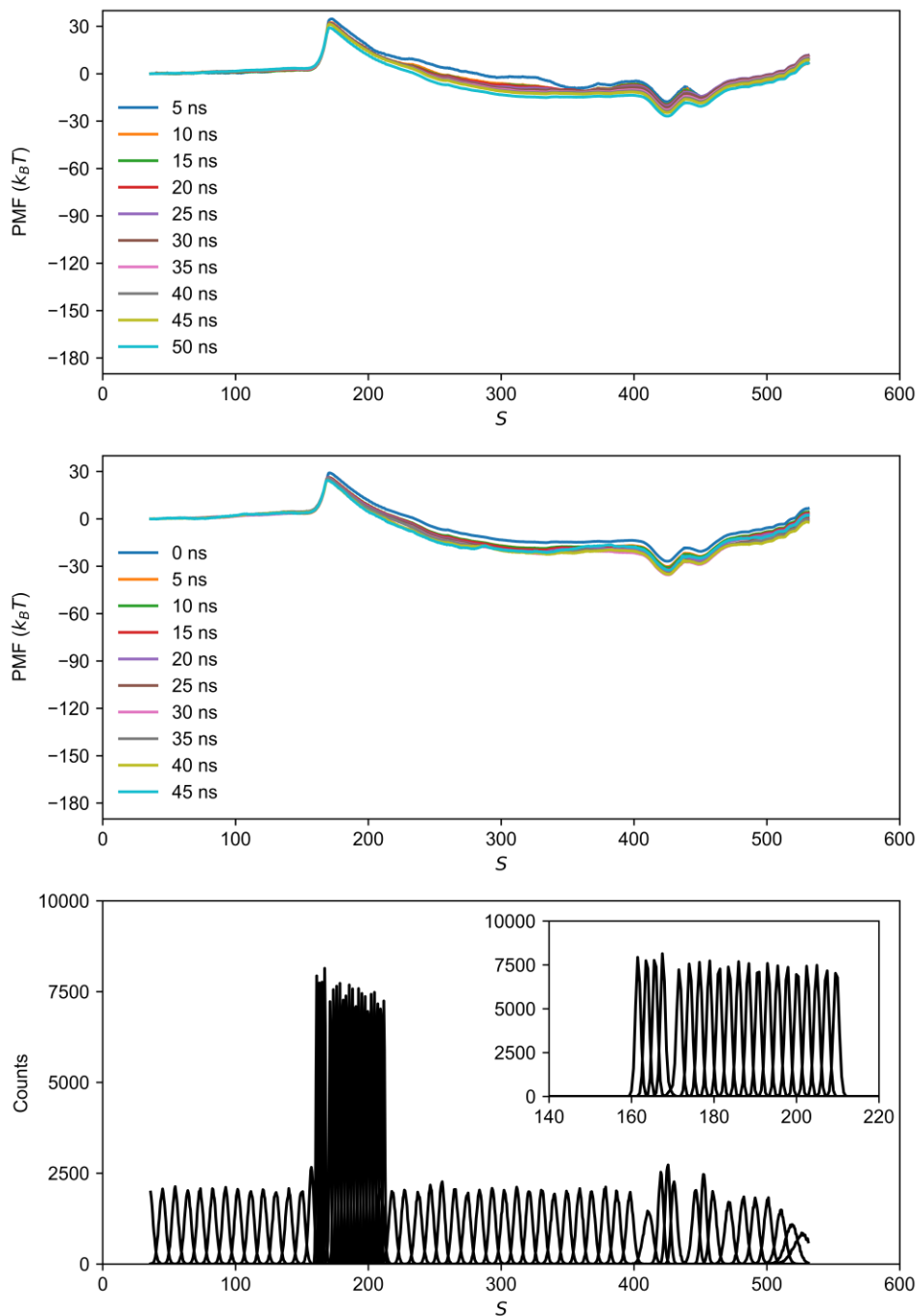
Figure S52. Potential of mean force (PMF) convergence for Replica #3 of the triple-branched 27-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.
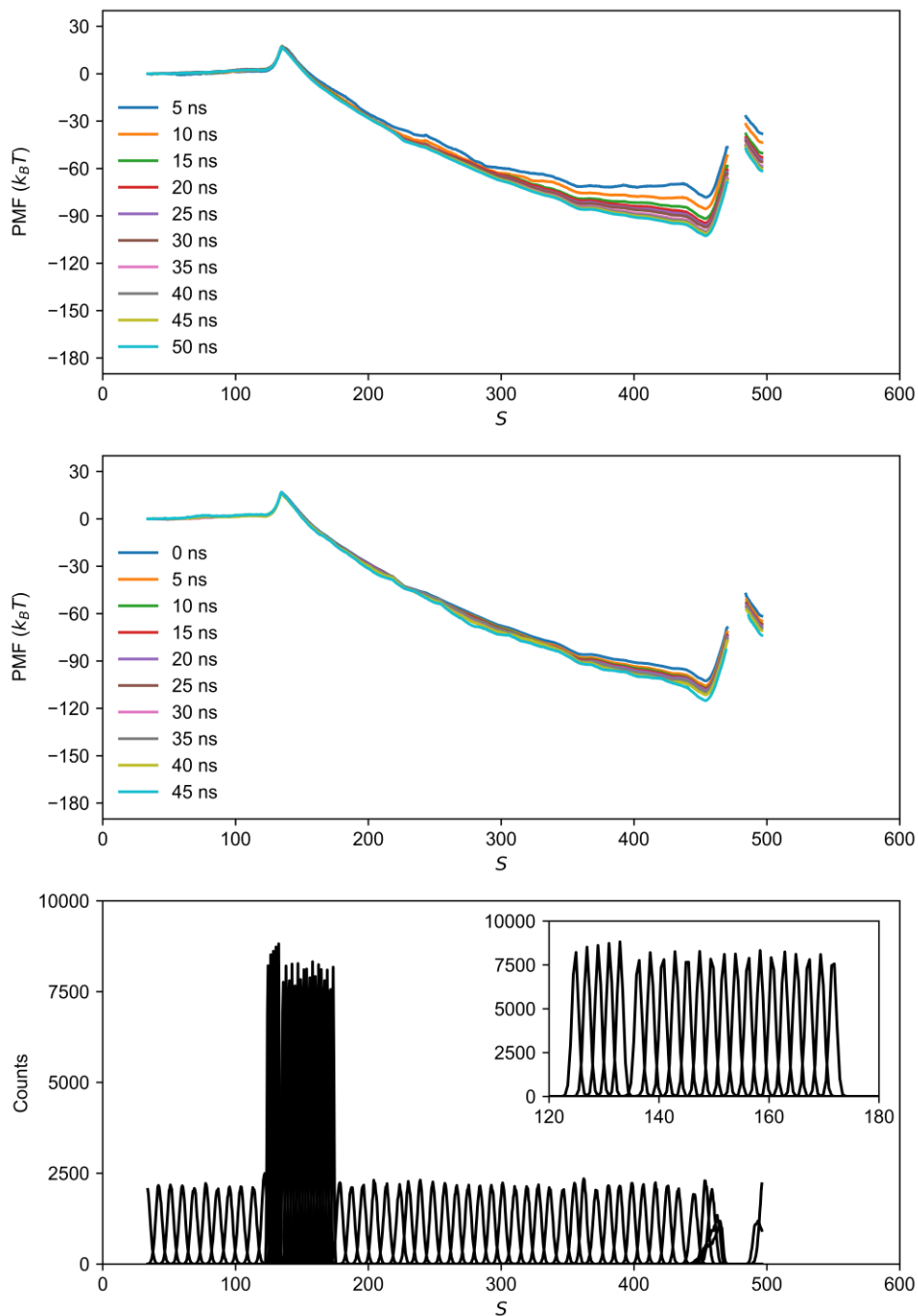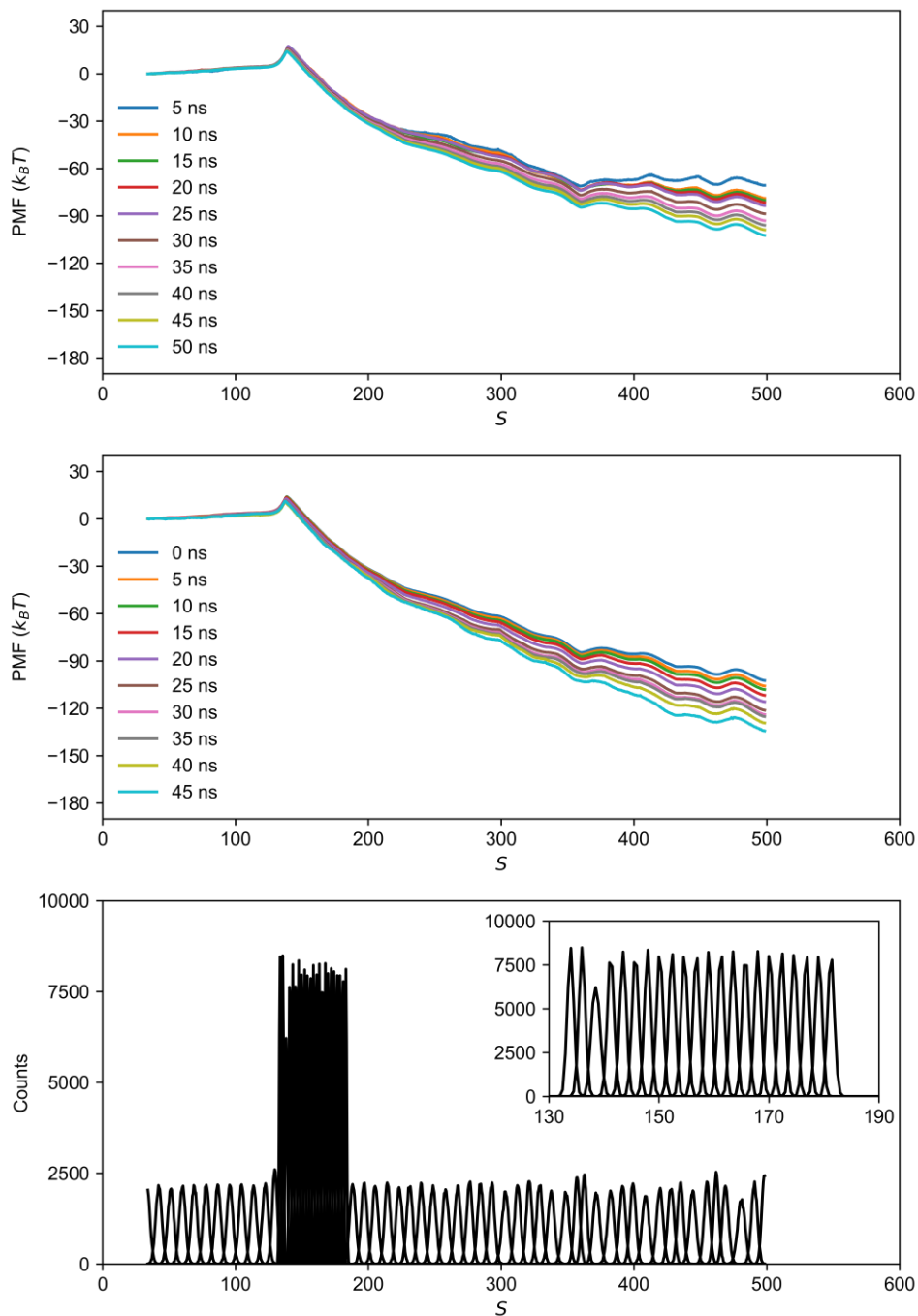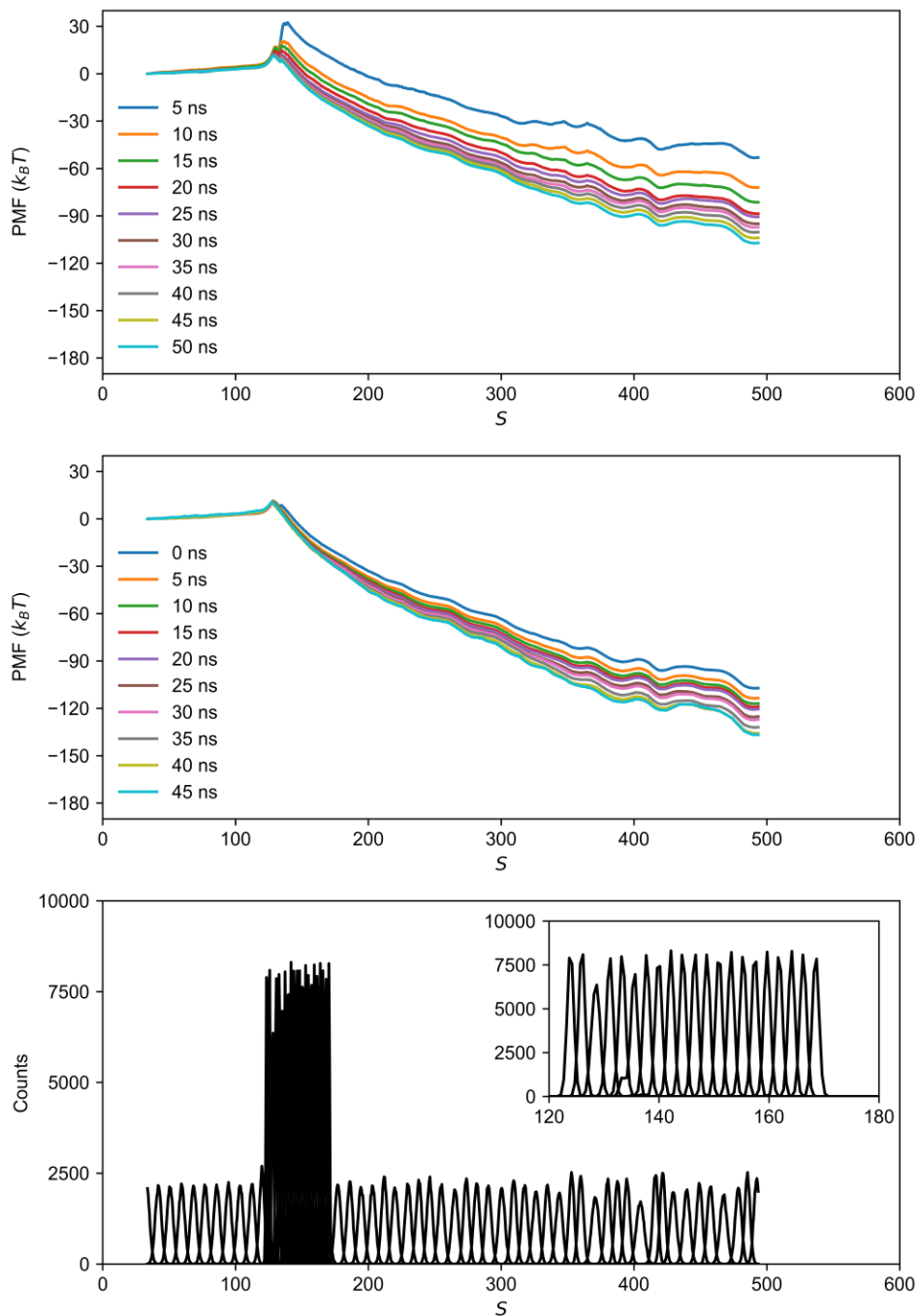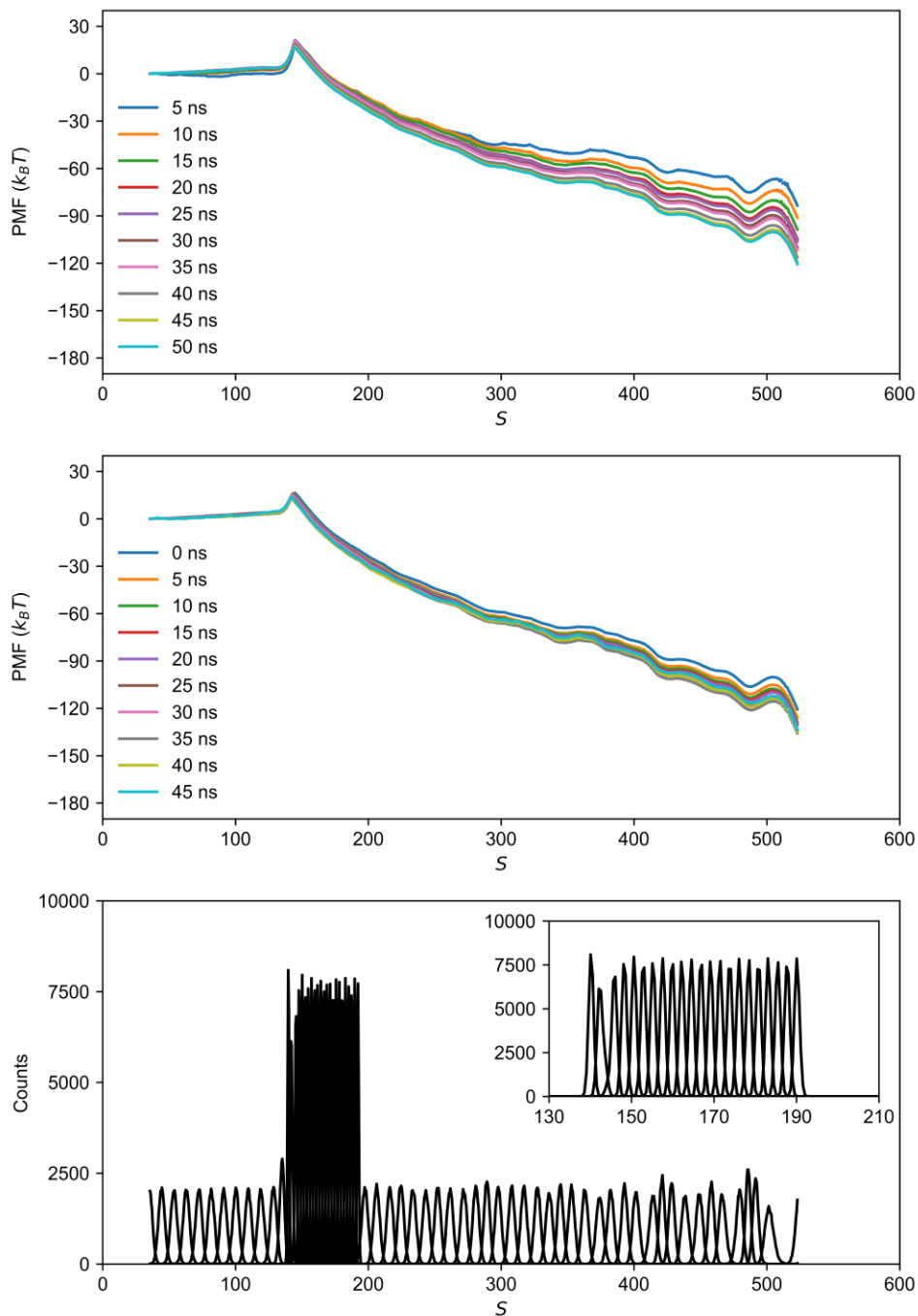
Figure S53. Potential of mean force (PMF) convergence for Replica #4 of the triple-branched 27-C ligand. Forward sampling (top row) and backward sampling (middle row) computed for every additional 5 ns of sampling. Histograms (bottom row) with inset plot of regions where a tighter spring constant was used to ensure sufficient overlap for regions of S of interest.

<u>Earlier onset of contacts increases free energy barriers for branched R groups.</u>

To identify how the onset of ligand-lipid contacts (N_c) may influence thermodynamic barriers and the transition state for nanoparticle adsorption, we compare the potential of mean force (PMF) as a function of the distance along the path (S) and compare it to the minimum free energy paths in CV space (D_z and N_c). As S is difficult compare directly to both collective variables, we plot the PMF curve as a function of S alongside N_c as a function of D_z. For branched ligands (dashed curves), we observe that an earlier onset of contacts (*i.e.*, ligands reaching the bilayer farther away from it) leads to higher free energy barriers for adsorption. On the contrary, when contacts are initiated when the nanoparticle is closer to the bilayer, the free energy barrier is lower. This is consistent with the fact that branched ligands are able to protrude more (Figure 4 from the main text), and that adsorption (or the transition state) is initiated by lipophilic contacts between the R group and the bilayer mediated by large exposed lipophilic groups. As the nanoparticle is farther away from the bilayer (with larger D_z), protrusions that generate contacts (between ligands and lipids) may be thermodynamically unfavorable since those contacts would simultaneously impose a force on D_z (and S in general). This then increases the PMF as a higher potential is required to maintain the nanoparticle at the corresponding z-distance in S.

Figure S54. Free energy barriers compared to onset of N_c for 11-C ligand pair. Left: potential of mean force curves zoomed into the region of S near the free energy barrier for adsorption. Right: Minimum free energy path in CV space zoomed into the region of S near the onset of N_c. Dashed lines correspond to the linear R group and solid lines correspond to the double-branched R group.
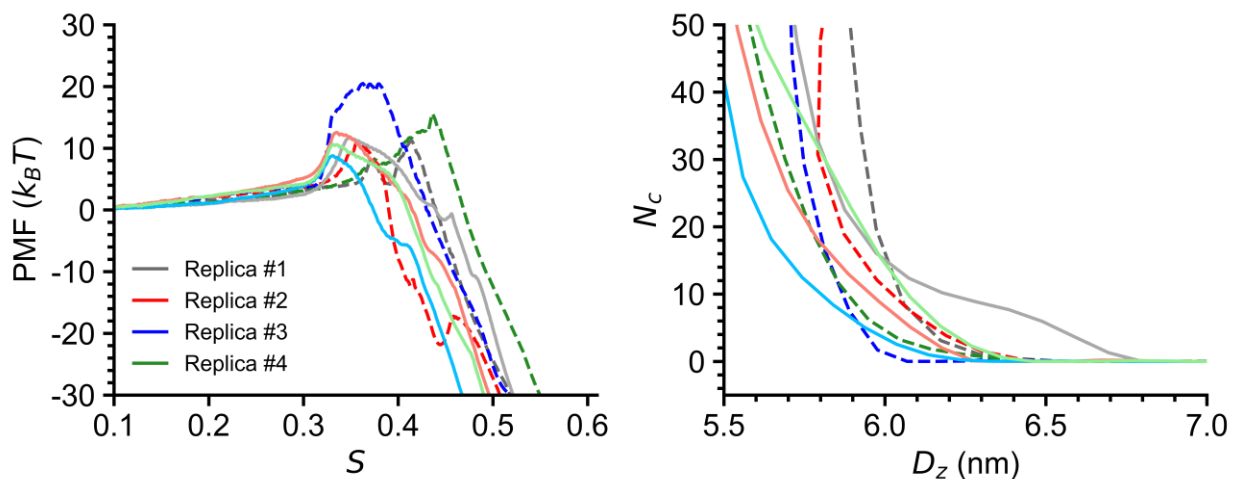
Figure S55. Free energy barriers compared to onset of N_c for 15-C ligand pair. Left: potential of mean force curves zoomed into the region of S near the free energy barrier for adsorption. Right: Minimum free energy path in CV space zoomed into the region of S near the onset of N_c. Dashed lines correspond to the linear R group and solid lines correspond to the triple-branched R group.

Figure S56. Free energy barriers compared to onset of N_c for 19-C ligand pair. Left: potential of mean force curves zoomed into the region of S near the free energy barrier for adsorption. Right: Minimum free energy path in CV space zoomed into the region of S near the onset of N_c. Dashed lines correspond to the linear R group and solid lines correspond to the double-branched R group.
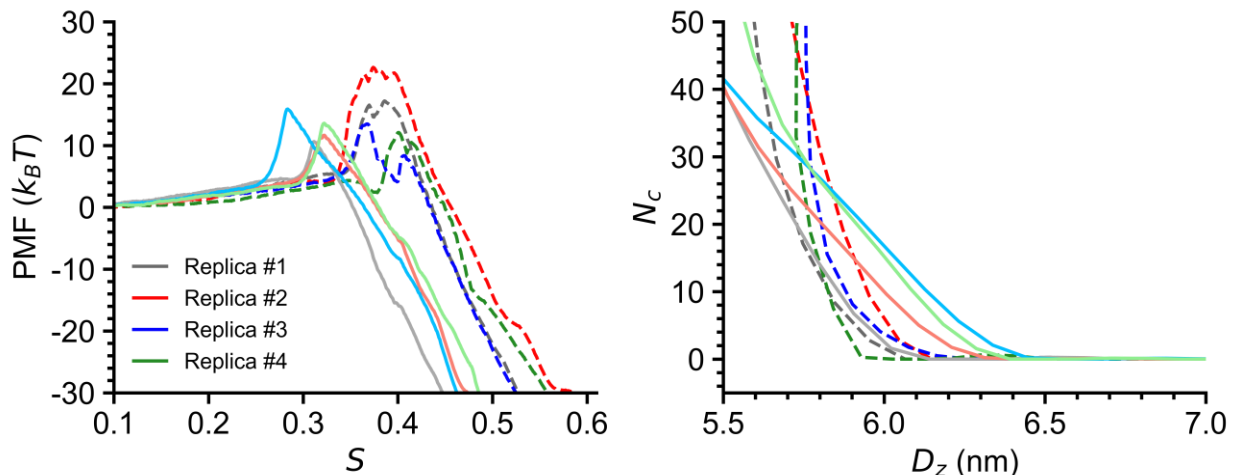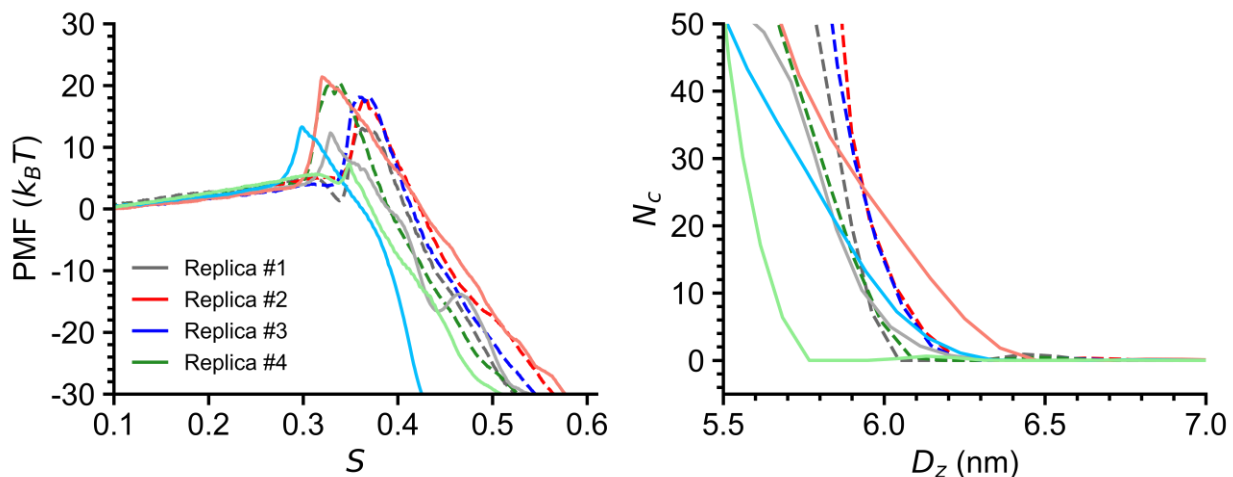
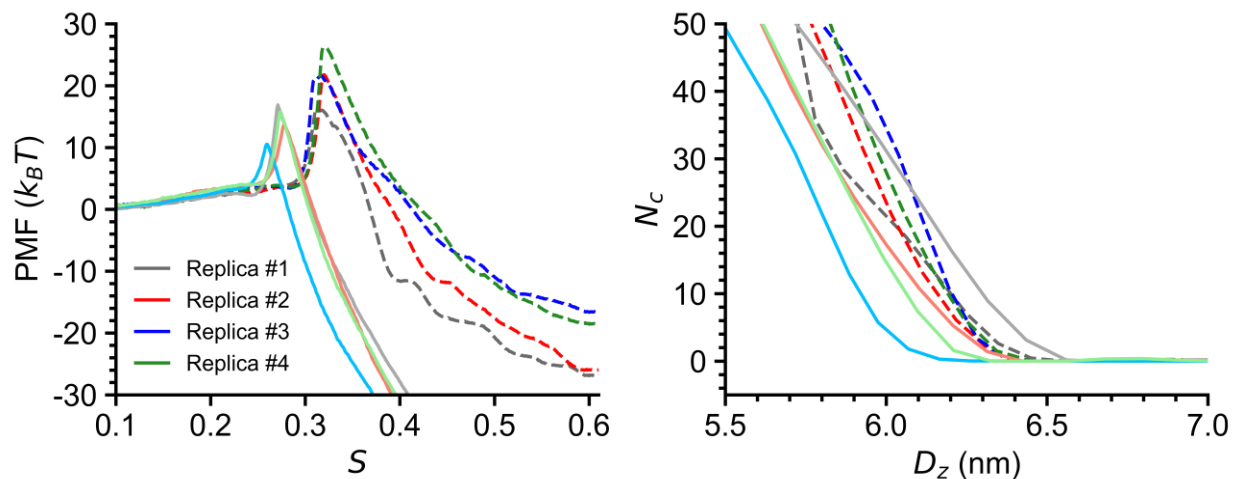Figure S57. Free energy barriers compared to onset of N_c for 27-C ligand pair. Left: potential of mean force curves zoomed into the region of S near the free energy barrier for adsorption. Right: Minimum free energy path in CV space zoomed into the region of S near the onset of N_c. Dashed lines correspond to the double-branched R group and solid lines correspond to the triple-branched R group.

R group protrusions quantified from radial distribution functions (RDFs).

The RDFs for the quaternary ammonium group (TTMA) and the R groups of the ligands were computed from unbiased simulations of AuNPs in solution without the bilayer. The *gmx rdf* tool was used for RDF calculations using a bin size of 0.01 nm. The radial distance of the TTMA (r_{TTMA}) bead (Q0) was used as a reference cut-off to quantify the number of R group protrusions. For the ligand R groups, we integrated $g(r)$ for $r > r_{TTMA}$. In the case of the triple-branched 27-C, the RDF of the quaternary ammonium showed a binodal distribution, so we used the mean radial distance instead as shown in Figure 6 of the main text. The integral represents a quantitative metric of the number of protrusions of the R groups. Since we compare across pairs of ligands that have the same number of beads on their R groups, per-bead normalization is not needed and is consistent for discerning the probability for each R group to protrude. The integrated areas under the curves are tabulated in Table S2 and we can see that for each ligand pair, the R groups with the higher number of branches protrude more. The double-branched 15-C and triple-branched 19-C show the highest increase in protrusions (82× and 10× higher), consistent with a change in mechanism for adsorption from lipid tail protrusions to ligand R group intercalation.

Table S2. Integrals of the area under the curve for R groups in reference to the quaternary ammonium group of ligands. Only area under the curve for radial distances larger than the radial distance of the quaternary ammonium was integrated.

	11-C	15-C	19-C	27-C
Lower-Branched	203.88	16.26	17.15	128.83
Higher-Branched	1024.89	1325.04	169.26	676.32

References

1. A. Grossfield, *WHAM: The Weighted Histogram Analysis Method, 2.0.11*, Grossfield Lab, Rochester, NY, 2013.