Supporting Information for "Liquid-liquid Phase Separation as the Second Step of Complex Coacervation"

Aditya N. Singh and Arun Yethiraj*

Theoretical Chemistry Institute and Department of Chemistry, 1101 University Avenue, University of Wisconsin-Madison, Madison, Wisconsin 53703

E-mail: yethiraj@wisc.edu

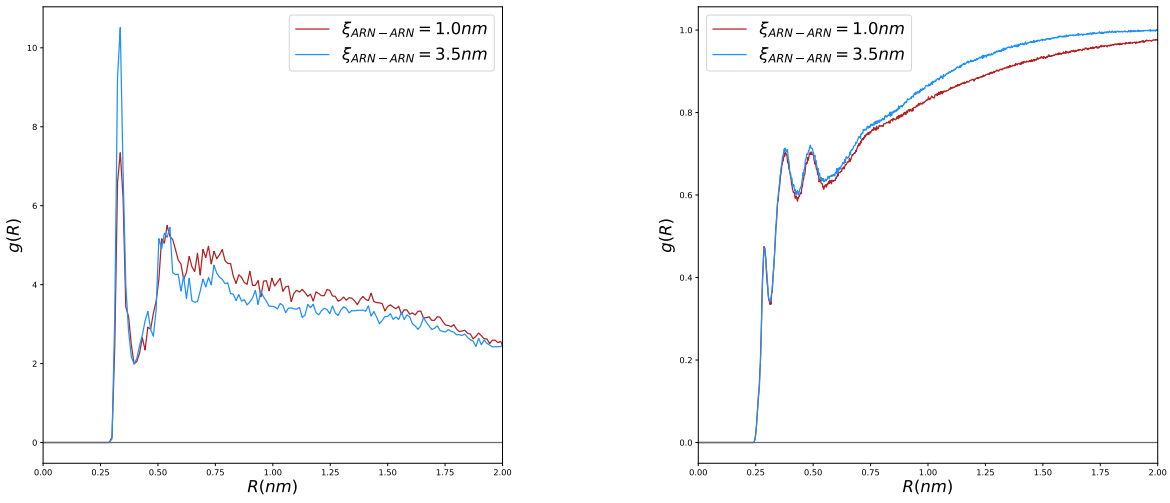


Figure S1: Average pair correlation function between the charged residues of the Arginine and (a) Cholride ion, and (b) Oxygen atoms of the water molecules, in the pR-pR system.

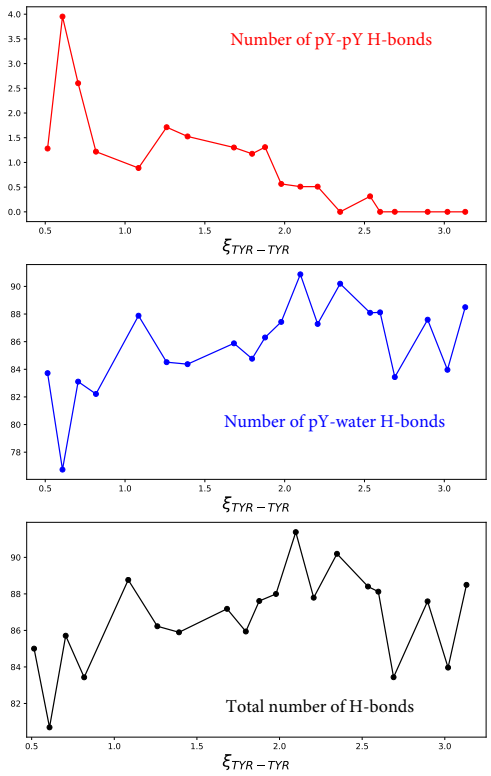


Figure S2: Hydrogen bonds between pY residues, between pY residues and water, and total hydrogen bonds as a function of the reaction co-ordinate.

To ensure the convergence of the umbrella sampling calculations, an extended simulation is run for the pY-pR system in the absence of salt at temperature $T = 280$ K. As before, a total of 30 windows spaced 0.1 nm apart from 0.5 to 3.5 nm are used along the reaction coordinate $\xi_{ARN-TYR}$. Each window is run for a total of 90 ns, which equals to a total simulation time of 2.7 μs. The simulations in all windows are then divided into three blocks (0-30 ns, 30-60 ns, 60-90 ns) and the PMF is calculated for each one of these blocks. The standard deviation is calculated by the histogram bootstrapping analysis (a total of 200 bootstraps are used). The results depicted in Fig 3 show a good quantitative agreement between the PMFs at different time blocks. The PMFs at different times are within one standard deviation of each other, from which we conclude that the umbrella sampling simulations have converged

within the first 30 ns.

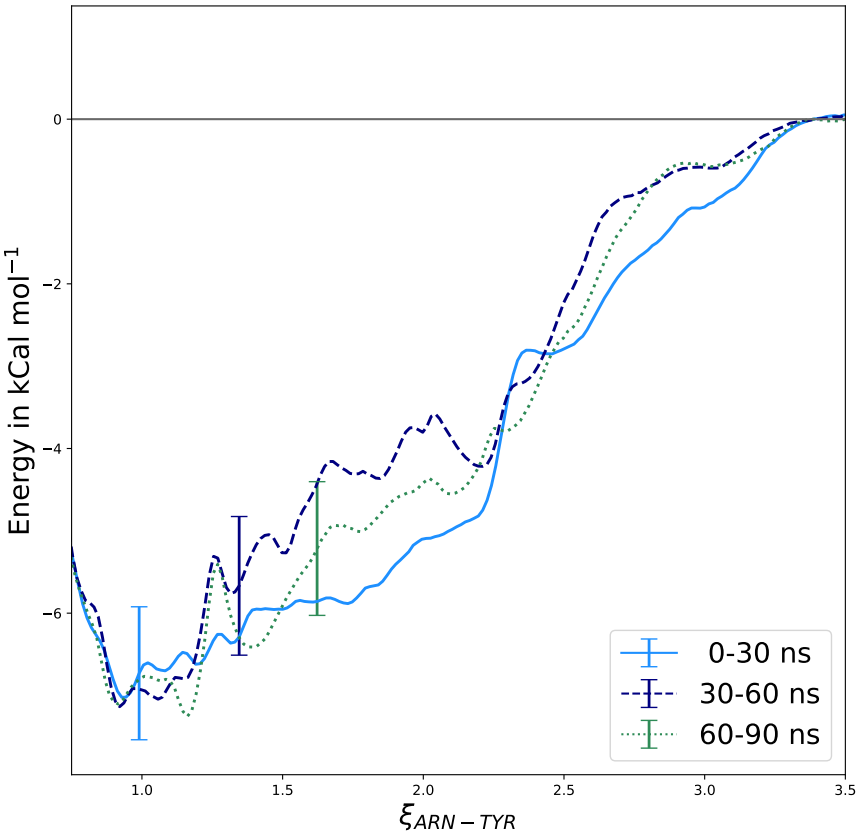
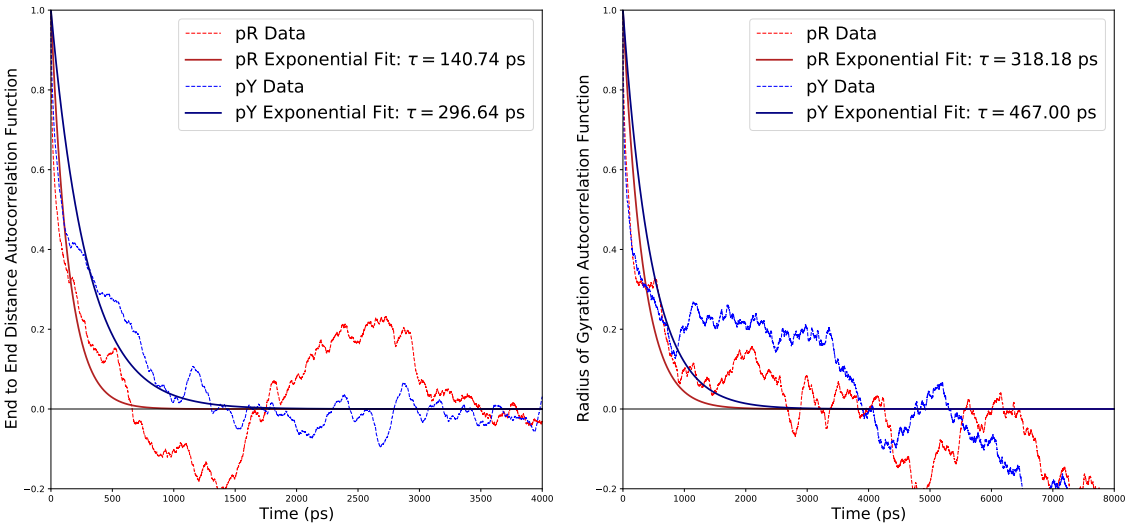


Figure S3: Potential of mean force obtained from an extended simulation of the pY-pR system in the absence of excess salt at temperature $T = 280$ K. Each window is run for **90 ns**, and the total simulation time is $90 \times 30 =$ **2.7 μs**. The PMF is calculated by block-averaging – each window is divided into 3 blocks (0-30, 30-60, and 60-90 ns) and the PMF is calculated for each of these blocks. A total of 200 bootstraps are used for the histogram bootstrapping analysis to calculate the standard deviation.

To ensure that the polypeptides are sufficiently sampling all configurations in 30 ns, the end to end distance and the radius of gyration autocorrelation functions are computed for both pY and pR in the window at temperature $T = 280$ K in which the distance of separation between them ($\xi_{ARN-TYR}$) is 0.8 nm. An exponential curve is fit to the data to get an estimate of the relaxation time. The result presented in Fig 4 show that the autocorrelation decays quickly relative to the simulation time. The relaxation time for the

end to end distance is found to be 296.64 ns for pY and 140.74 ns for pR. The relaxation time for the radius of gyration is found to be 467.00 ps for pY and 318.18 ps for pR. Hence, in the worst case scenario, the simulations are run for ≈ 64 times the relaxation time of the polypeptides, which is more than sufficient to give a proper quantitative estimates of the quantity of interest. These results backed by the extended PMF calculations strongly suggest that 30 ns is adequate to sample all the relevant conformations of polypeptides and obtain a converged umbrella sampling calculation.



(a) End to End Distance **(b)** Radius of Gyration

Figure S4: Autocorrelation function for pY and pR at temperature $T = 280$ K and distance of separation $\xi_{ARN-TYR} = 0.8$ nm. The system has no excess salt.

Similar simulation times are used in other simulations for the PMF in biological[1–8] For instance, Razzokov et al.[1] perform a 20 ns (per window) umbrella sampling calculation to study the aggregation of $\alpha\beta$ pentamer. Fabian & Zacharlas[2] estimate binding energies of proteins by performing umbrella sampling calculations for 32 ns per window. Mills and Andricioaei[3] develop an experimentally guided methodology to efficiently sample biomolecules, in which they use a normal umbrella sampling calculation run for 20 ns per window as a reference. Bagai et al.[4] study the aggregation of DNA molecules by performing US calculations of 20 ns per window.

References

(1) Razzokov, J.; Yusupov, M.; Bogaerts, A. Oxidation destabilizes toxic amyloid beta peptide aggregation. *Scientific reports* **2019**, *9*, 1–7.

(2) Zeller, F.; Zacharias, M. Efficient calculation of relative binding free energies by umbrella sampling perturbation. *Journal of computational chemistry* **2014**, *35*, 2256–2262.

(3) Mills, M.; Andricioaei, I. An experimentally guided umbrella sampling protocol for biomolecules. *The Journal of chemical physics* **2008**, *129*, 114101.

(4) Bagai, S.; Sun, C.; Tang, T. Potential of mean force of polyethylenimine-mediated DNA attraction. *The Journal of Physical Chemistry B* **2013**, *117*, 49–56.

(5) Sedghi, M.; Goual, L.; Welch, W.; Kubelka, J. Effect of asphaltene structure on association and aggregation using molecular dynamics. *The Journal of Physical Chemistry B* **2013**, *117*, 5765–5776.

(6) Arsiccio, A.; Pisano, R. Surfactants as stabilizers for biopharmaceuticals: an insight into the molecular mechanisms for inhibition of protein aggregation. *European Journal of Pharmaceutics and Biopharmaceutics* **2018**, *128*, 98–106.

(7) Levine, Z. A.; Rapp, M. V.; Wei, W.; Mullen, R. G.; Wu, C.; Zerze, G. H.; Mittal, J.; Waite, J. H.; Israelachvili, J. N.; Shea, J.-E. Surface force measurements and simulations of mussel-derived peptide adhesives on wet organic surfaces. *Proceedings of the National Academy of Sciences* **2016**, *113*, 4332–4337.

(8) Babakhani, A.; Gorfe, A. A.; Kim, J. E.; McCammon, J. A. Thermodynamics of peptide insertion and aggregation in a lipid bilayer. *The Journal of Physical Chemistry B* **2008**, *112*, 10528–10534.