Supporting Information for:

Inositol hexakisphosphate (IP6) accelerates immature HIV-1 Gag protein assembly towards kinetically-trapped morphologies

Alexander J. Pak,[1,⊥] Manish Gupta,[1] Mark Yeager,[2,3,4,5,†] and Gregory A. Voth[1,*]

[1]Department of Chemistry, Chicago Center for Theoretical Chemistry, Institute for Biophysical Dynamics, and James Franck Institute, The University of Chicago, Chicago, IL 60637, United States
[2]Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, Charlottesville, VA 22908, United States
[3]Center for Membrane Biology, University of Virginia School of Medicine, Charlottesville, VA 22908, United States, United States
[4]Cardiovascular Research Center, University of Virginia School of Medicine, Charlottesville, VA 22908, United States
[5]Department of Medicine, Division of Cardiovascular Medicine, University of Virginia School of Medicine, Charlottesville, VA 22908, United States

[⊥] Present address: Department of Chemical and Biological Engineering, Colorado School of Mines, Golden, CO 80401, United States
[†] Present address: Frost Institute for Chemistry and Molecular Science, University of Miami, Coral Gables, FL 33124, United States

[*] Corresponding author: gavoth@uchicago.edu

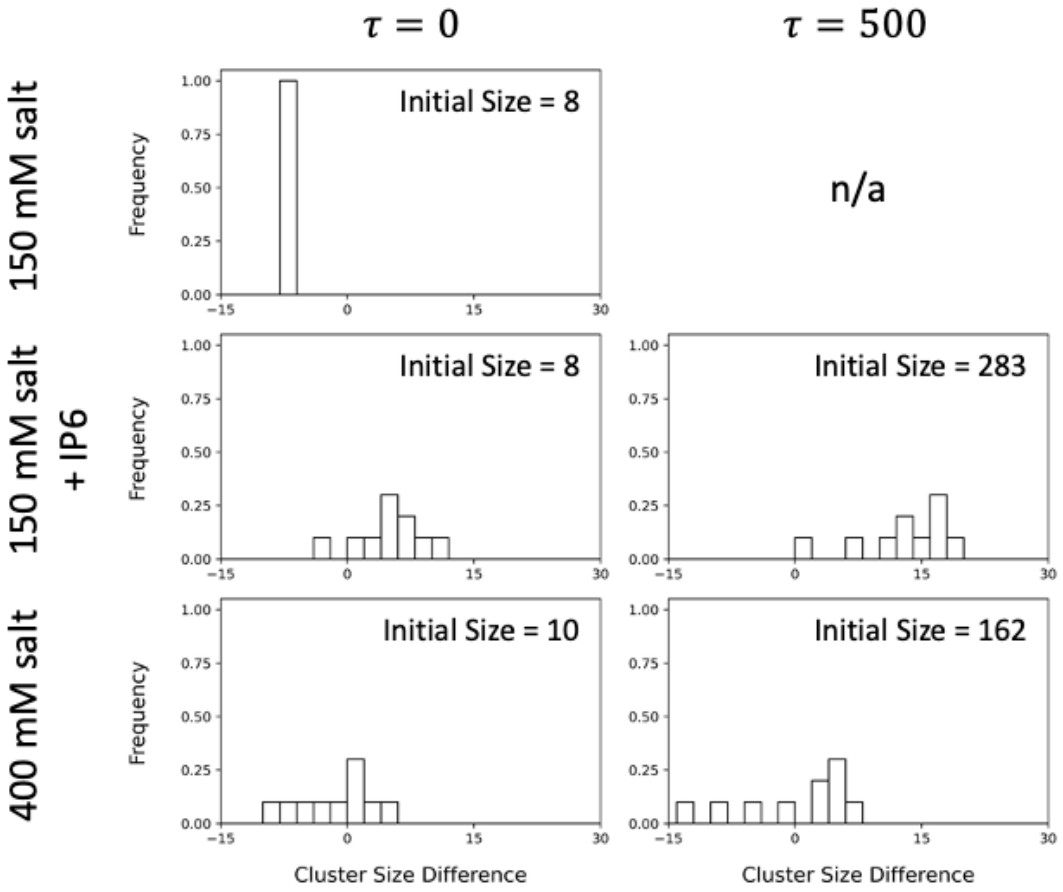

Figure S1. Distributions of cluster size differences (i.e. final cluster size – initial cluster size) from 10 independent replicas of one trajectory segment of length $\Delta\tau = 50$ x 10^6 time steps. All panels are labeled with their corresponding "*in vitro*" system conditions (detailed in the main text), their associated starting CG time step (x 10^6), and their initial cluster sizes.

Additional discussion on the path sampling algorithm. From Fig. S1, it is evident that the propensity for CA/SP1 assembly varies between strongly favored and strongly disfavored. It can be inferred that while assembly is strongly disfavored when IP6 is absent, assembly is weakly favored when salt concentrations are increased and further favored when IP6 is introduced instead. The shown distributions also suggest that assembly outcomes follow a one-sided heavy-tailed distribution with the tail in the left direction. In the case of CA/SP1 assembly under 400 mM salt conditions (and to a lesser extent in the 150 mM salt with IP6 condition), the tail of the distribution persisted in the negative lattice growth regime. In other words, we should expect a fraction of our sampled trajectories to undergo net disassembly within a time scale of 50 x 10^6 CG time steps although net assembly is still favored overall. Due to the computational cost of each trajectory, we adopted this path sampling approach to ensure that a representative trajectory of CA/SP1 assembly (in cases where assembly is weakly favored at minimum) is sampled despite using a reduced number of trajectories.

In this work, we modeled the distribution using a log-normal function but we expect that other heavy-tailed distributions may also be appropriate. We also cannot discount that our choice of 50 x 10^6 CG time steps per segment may be imposing a bias to our observed effective assembly pathway. In particular, if assembly via error correction were to proceed using a phase of net disassembly longer than 50 x 10^6 CG time steps, we would not be able to capture that using the current algorithm. However, this scenario would likely only emerge in the 400 mM salt case, and would effectively result in greater defect annealing throughout the assembled lattice.