Supporting Information

Manifold of Polyampholyte Necklaces: From Charge Migration to Hierarchical Structures

Yiheng Wu,[†] Artem M. Rumyantsev,[*,‡] and Juan J. de Pablo[*,¶]

†*Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, United States*

‡*Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina 27695-7905, United States*

¶*Departments of Chemical and Biological Engineering, Computer Science, and Physics, Tandon School of Engineering and Courant Institute of Mathematical Sciences, New York University, NY 11201*

E-mail: rumyantsev@ncsu.edu; depablo@uchicago.edu

I Charge Partition Between Head and Tail of Tadpole

Consider an asymmetric diblock polyampholyte chain in a good/poor solvent. We denote N_+ the number of monomers in the positively charged block, and N_- is the number of monomers in the negatively charged block (for definiteness, we assume $N_+ > N_-$). The fraction of charged monomers f is the same in both blocks. The number of positively charged monomers is $N_+ f$. The scaling theory and free energy analysis in the Θ solvent have been constructed in ref. 1.

First, we consider a good solvent case. In tadpole conformations, the almost neutral head is formed by oppositely charged blocks, and the tail is formed by an extended polycationic block. To derive the equal electrostatic potential criterion used in the main text

$$\frac{Q_{\rm head}}{D_{\rm head}} \simeq \frac{Q_{\rm tail}}{l_{\rm tail}}, \tag{S1}$$

we carry out the free energy analysis as in ref. 1. We assume that the head contains $2N_- + N_{\rm h}$ monomers (N_- monomers from the negatively charged block and $N_- + N_{\rm h}$ monomers from the positively charged block), and the tail contains $N_+ - N_- - N_{\rm h}$ monomers. The free energy of the head can be written as the sum of the energy of Coulomb correlation (fluctuation) attractions between the oppositely charged monomers in the head, the surface free energy of the head, and the Coulomb self-energy of the head due to its net nonzero charge

$$F_{\rm head} \simeq -v^{-1/7}(uf^2)^{5/7}(2N_- + N_{\rm h}) + v^{-2/21}(uf^2)^{10/21}(2N_- + N_{\rm h})^{2/3} + \frac{(uf^2)^{25/21}N_{\rm h}^2}{v^{5/21}(2N_- + N_{\rm h})^{1/3}} \tag{S2}$$

The free energy of the tail is the free energy of a stretched polyelectrolyte chain in a good solvent:

$$\frac{F_{\rm tail}}{k_{\rm B}T} \simeq v^{-1/7}(uf^2)^{5/7}(N_+ - N_- - N_{\rm h}) \tag{S3}$$

The equilibrium condition for the charge partitioning can be found by minimizing the free

energy of the tadpole, $F_{\text{tadpole}} \simeq F_{\text{head}} + F_{\text{tail}}$, with respect to the number of excess monomers in the head N_{h}:

$$\frac{\partial F_{\text{tadpole}}}{\partial N_{\text{h}}} \simeq -v^{-1/7}(uf^2)^{5/7} + v^{-2/21}(uf^2)^{10/21}(2N_- + N_{\text{h}})^{-1/3} + \frac{(uf^2)^{25/21}N_{\text{h}}}{v^{5/21}(2N_- + N_{\text{h}})^{1/3}} - \frac{(uf^2)^{25/21}N_{\text{h}}^2}{v^{5/21}(2N_- + N_{\text{h}})^{4/3}} - v^{-1/7}(uf^2)^{5/7} \quad \text{(S4)}$$

By setting the derivative to 0, the free energy is minimized, and (in the limit of $N_h \ll N_-$) the N_h value is approximately equal to

$$N_{\text{h}} \simeq v^{2/21}(uf^2)^{-10/21}N_-^{1/3} \simeq v^{2/21}(uf^2)^{-10/21}N^{1/3} \quad \text{(S5)}$$

The above free energy analysis shows that the dominant contributions to the chemical potential of the monomers in the head and tail come from the Coulomb self-energy of the head and the tail, respectively. As a result, eq. S5 from free energy analysis aligns with the requirement of the equal electrostatic potentials of the tail and head given by eq. S1 with $Q_{\text{head}} \simeq N_{\text{h}}f$.

In a poor solvent, the PE tail becomes unstable with respect to the classical PE necklace formation. Thus, the diblock PA forms one major head, and its tail represents the PE necklace. We continue assuming that the head and the tail contain $2N_- + N_h$ and $N_+ - N_- - N_h$ monomers, respectively. The free energy of the head is analogous to that in a good solvent:

$$F_{\text{head}} \simeq -v^2(2N_- + N_h) + v^{4/3}(2N_- + N_h)^{2/3} + \frac{v^{1/3}(uf^2)N_h^2}{(2N_- + N_h)^{1/3}} \quad \text{(S6)}$$

The free energy of the PE necklace tail consists of 5 parts:[2] the additional energy of correlation attraction in the PE beads $F_{\text{att}}^{\text{bead}} \simeq -v^2(N_+ - N_- - N_h)$, the Coulomb self-energy of the beads and the surface energy of the beads $F_{\text{Coul}}^{\text{bead}} \simeq F_{\text{surf}}^{\text{bead}} \simeq |v|(uf^2)^{1/3}(N_+ - N_- - N_h)$, and the surface energy of the strings and the Coulomb self-energy of the entire necklace tail $F_{\text{surf}}^{\text{str}} \simeq F_{\text{Coul}}^{\text{tail}} \simeq |v|^{1/2}u^{1/2}f(N_+ - N_- - N_h)$. We note that the energy of correlation attractions

is not explicitly expressed in ref 2 because the reference state is a melt of thermal blobs. We include the energy of correlation attraction in our analysis to stay consistent with ref 1, where the reference state is a linear array of (electrostatic/thermal) blobs. We are now in a position to minimize the total free energy of the tadpole F_{tadpole} with respect to the number of excess monomers N_h:

$$\frac{\partial F_{\text{tadpole}}}{\partial N_{\text{h}}} \simeq v^{4/3}(2N_- + N_{\text{h}})^{-1/3} + \frac{v^{1/3}(uf^2)N_{\text{h}}}{(2N_- + N_{\text{h}})^{1/3}} - \frac{v^{1/3}(uf^2)N_{\text{h}}^2}{(2N_- + N_{\text{h}})^{4/3}} - |v|(uf^2)^{1/3} - |v|^{1/2}u^{1/2}f \quad \text{(S7)}$$

The free energy is minimized by setting the derivative to 0, and the resulting N_h scales as

$$N_h \simeq v^{2/3}(uf^2)^{-2/3}N^{1/3} \quad \text{(S8)}$$

The above free energy also shows that the Coulomb self-energy of the major head and the minor PE beads dominates the contribution to the chemical potential of the monomers in the major head and the tail, respectively. Consequently, eq. S8 is consistent with the criterion of equal electrostatic potentials of the major head and the minor beads:

$$\frac{Q_{\text{head}}}{D_{\text{head}}} \simeq \frac{Q_{\text{bead}}^{\text{min}}}{D_{\text{bead}}^{\text{min}}} \quad \text{(S9)}$$

II Free Energy Comparison for Charge-in-beads Necklace III−a and Hierarchical Necklace III−b

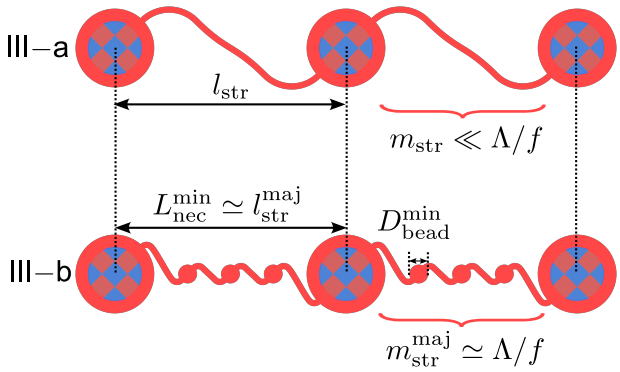

Figure S1: Schematic representation of necklaces of the type III−a and III−b. The necklace III−a here represents the structure if the necklace of the type III− formed under condition $\Lambda^* < \Lambda < \Lambda^{**}$, i.e. in regime III−b.

Here, by means of free energy comparison, we demonstrate that the charge migration drives the formation of hierarchical necklace III−b. If the necklace of the type III−a also forms above the boundary $\Lambda > \Lambda^*$ (or equivalently $|v| > |v|_{\rm III-a/III-b}$), the string would become PE because the PE string condition is satisfied, as shown in Figure S1. Comparing necklace III−a with III−b, the distances between the major beads in them are the same, $l_{\rm str}^{\rm III-a} \simeq l_{\rm str}^{\rm maj,\,III-b}$, because the string length is controlled by the major beads carrying most of the charge imbalance. The difference between necklace III−a and III−b is that necklace III−b contains more monomers in the major strings due to the formation of the minor PE necklace within the major necklace strings, see Figure S1:

$$m_{\rm str}^{\rm maj,\,III-b} \simeq \frac{\Lambda}{f} \gg m_{\rm str}^{\rm III-a} \simeq (|v|uf\Lambda)^{-1/2} N^{1/2} \simeq \left(\frac{\Lambda^*}{\Lambda}\right)^{3/2} \frac{\Lambda}{f} \tag{S10}$$

For both types of necklaces, a marginal charge migration takes place: the net charge in the major beads prefers to move to the strings of necklace III−a or the major strings of necklace III−b. However, for necklace III−b, the associated gain in the Coulomb energy is

higher because the major strings contain more monomers. If calculated per one major bead, the extra free energy gain in necklace III−b compared to III−a equals

$$\Delta F_{\rm Coul}^{\rm gain} \simeq F_{\rm Coul}^{\rm gain,\,III-b} - F_{\rm Coul}^{\rm gain,\,III-a} \tag{S11}$$

$$\simeq \frac{ua[(Q_{\rm bead}^{\rm maj})^2 - (Q_{\rm bead}^{\rm maj} - \Lambda)^2]}{D_{\rm bead}^{\rm maj}} - \frac{ua[(Q_{\rm bead}^{\rm maj})^2 - (Q_{\rm bead}^{\rm maj} - m_{\rm str}^{\rm III-a} f)^2]}{D_{\rm bead}^{\rm maj}} \tag{S12}$$

$$\simeq \frac{uaQ_{\rm bead}^{\rm maj}\Lambda}{D_{\rm bead}^{\rm maj}} - \frac{uaQ_{\rm bead}^{\rm maj}\sqrt{\frac{Nf}{u\Lambda|v|}}}{D_{\rm bead}^{\rm maj}} \tag{S13}$$

$$\simeq |v| u^{1/3} f^{-1/6} \Lambda^{5/6} N^{1/6} \left[ ]1 - \left(\frac{\Lambda^*}{\Lambda}\right)^{3/2} \right] \tag{S14}$$

Here we note that $Q_{\rm bead}^{\rm maj}$ and $D_{\rm bead}^{\rm maj}$ for necklace III−b are the same as $Q_{\rm bead}$ and $D_{\rm bead}$ for necklace III−a. For one major bead, the cost of forming $N_{\rm bead}^{\rm min}$ minor PE beads is

$$F_{\rm surf}^{\rm cost} \simeq N_{\rm bead}^{\rm min} \frac{(D_{\rm bead}^{\rm min})^2}{\xi_{\rm T}^2} \simeq |v| u^{1/3} f^{-1/3} \Lambda \left[ ]1 - \left(\frac{\Lambda^*}{\Lambda}\right)^{3/2} \right] \tag{S15}$$

The ratio between the free energy gain $\Delta F_{\rm Coul}^{\rm gain}$ and cost $F_{\rm surf}^{\rm cost}$ is

$$\frac{\Delta F_{\rm Coul}^{\rm gain}}{F_{\rm surf}^{\rm cost}} \simeq \left(\frac{Nf}{\Lambda}\right)^{1/6} \gg 1 \tag{S16}$$

Thus, the free energy comparison confirms that, in regime III−b, it is more favorable to form a hierarchical necklace III−b in comparison to III−a.

III Boundary III−b/IV− between Different Hierarchical Necklaces Derived from the Free Energy Analysis

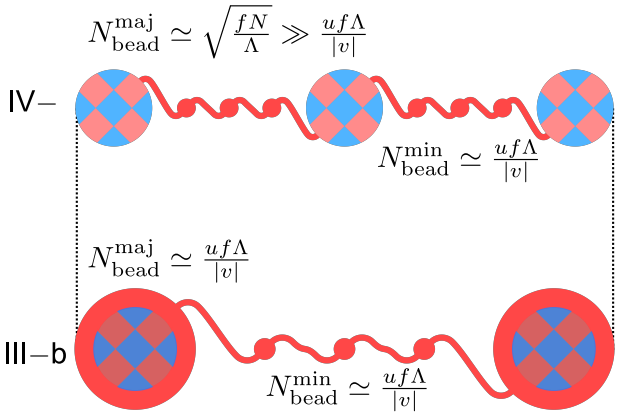

Figure S2: Schematic representation of necklaces of the type III−b and IV−. The necklace IV− here represents the structure if the necklace of the type IV− formed under condition $\Lambda^* < \Lambda < \Lambda^{**}$, i.e. in regime III−b.

If the hierarchical necklace of the type IV− forms below the boundary $\Lambda < \Lambda^{**}$ (or equivalently $|v| > |v|_{\text{III−b/IV−}}$), it would contain a higher number of major beads compared to the hierarchical necklace III−b, as shown in Figure S2. The dominant part of the free energy of necklace IV− can be written as the sum of the surface energy of the major beads and minor beads

$$F_{\text{IV−}} \simeq N_{\text{bead}}^{\text{maj, IV−}} \frac{(D_{\text{bead, IV−}}^{\text{maj}})^2}{\xi_{\text{T}}^2} + N_{\text{bead}}^{\text{maj, IV−}} N_{\text{bead}}^{\text{min, IV−}} \frac{(D_{\text{bead}}^{\text{min}})^2}{\xi_{\text{T}}^2} \tag{S17}$$

The Coulomb self-energy of the major beads is negligible for necklace IV−, and the Coulomb self-energy of the minor beads is on the same order as the surface energy; therefore, it does not change the scaling estimates. The dominant part of the free energy of the necklace III−b can be written as the sum of the Coulomb self-energy of the major beads and the surface

energy of the minor beads

$$F_{\mathrm{III-b}} \simeq N_{\mathrm{bead}}^{\mathrm{maj,\,III-b}} \frac{ua(Q_{\mathrm{bead}}^{\mathrm{maj,\,III-b}})^2}{D_{\mathrm{bead}}^{\mathrm{maj,\,III-b}}} + N_{\mathrm{bead}}^{\mathrm{maj,\,III-b}} N_{\mathrm{bead}}^{\mathrm{min,\,III-b}} \frac{(D_{\mathrm{bead}}^{\mathrm{min}})^2}{\xi_{\mathrm{T}}^2} \tag{S18}$$

The free energy gain due to forming necklace III−b instead of IV− is

$$\begin{aligned}\Delta F \simeq F_{\mathrm{IV-}} - F_{\mathrm{III-b}} \simeq v^{4/3} f^{1/6} \Lambda^{-1/6} N^{5/6} \left[1 - \left(\frac{\Lambda}{\Lambda^{**}}\right)^{1/2}\right] \\ + |v| u^{1/3} f^{1/6} \Lambda^{1/2} N^{1/2} \left[1 - \left(\frac{\Lambda}{\Lambda^{**}}\right)^{3/2}\right]\end{aligned} \tag{S19}$$

The free energy gain tends to zero when Λ approaches Λ^{**}, which gives the crossover boundary for regime IV− written as $\Lambda \simeq \Lambda^{**}$, or equivalently $|v|_{\mathrm{III-b/IV-}} \simeq u f^{1/2} \Lambda^{3/2} N^{-1/2}$.

IV Details of Molecular Dynamics Simulations

Simulations are performed using the OpenMM package,[3] following our earlier work 4. The PAs are modeled as the Kremer-Grest chains[5] augmented with Coulomb interactions. Specifically, the particles are connected by finitely extensible nonlinear elastic (FENE) springs

$$U_{\mathrm{FENE}}(r) = -\frac{1}{2} K R_0^2 \ln\left(1 - \frac{r^2}{R_0^2}\right) \tag{S20}$$

and interact with each other through a truncated and shifted Lennard-Jones (LJ) potential (without any bond exclusions):

$$U_{\mathrm{LJ}}(r) = \begin{cases} 4\varepsilon_{\mathrm{LJ}}(r)\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} - c(R_{\mathrm{c}})\right], & r < R_{\mathrm{c}} \\ 0, & r \geq R_{\mathrm{c}} \end{cases} \tag{S21}$$

with $c(R_{\mathrm{c}}) = (\sigma/R_{\mathrm{c}})^{12} - (\sigma/R_{\mathrm{c}})^{6}$. To provide Θ solvent quality, the values of the parameters are set to $\varepsilon_{\mathrm{LJ}} = 0.34 k_{\mathrm{B}}T$, $R_{\mathrm{c}} = 2.5\sigma$, $K = 7 k_{\mathrm{B}}T/\sigma^2$, and $R_0 = 2\sigma$. For the case of hierarchical necklaces IV$-$, poor solvent quality is provided by $\varepsilon_{LJ} = 3.4 k_{\mathrm{B}}T$. Bare Coulombic interactions between ionic monomers are described by

$$U_{\mathrm{Coul}} = k_B T \frac{l_{\mathrm{B}} z_i z_j}{r_{ij}} \tag{S22}$$

with $z_{i,j} = \pm 1$ being the charge valencies.

Monomer sequences of PAs with $f = 0.5$ are generated using the Markov process with different Λ values ranging from 0.01 to 400. For each Λ, multiple sequence realizations are generated, and those carrying characteristic net charge values $Q_{\mathrm{tot}} = \sqrt{fN\Lambda}$ are selected. For different values of Λ, different numbers of sequence realizations are needed to provide reliable averages, as shown in Table S1.

The time step is set to $0.005\tau_{\mathrm{LJ}}$, and the temperature $T = 1$ is maintained by the Langevin

Table S1: Summary of the number of sequence realizations used for different blockiness and different Bjerrum lengths

u value	charge blockiness Λ	number of sequences
2	$0.22, 0.46$	10
2	$0.1, 1$	20
2	$1.49, 2.22, 3.31$	50
2	$4.94, 7.37, 10.99$	100
2	$16.38, 24.42, 36.41, 54.29, 80.94, 120.68, 179.94, 268.28, 400$	50
3	$0.1 - 400$	10
4	$0.1, 0.22, 0.46$	10
4	$1 - 400$	50
5	$0.1 - 370$	10

thermostat with damping factor damp $= 100\tau_{\mathrm{LJ}}$.[4] For each PA realization, the first half of the MD run represents equilibration, and only the second half is used to produce the data. For system with $u = 2$, $\Lambda = 36.41$, 54.29, and 80.94, the simulation run lenght is equal to 2×10^9 steps. For other systems, it is set to 2×10^8 steps. In total, the simulations' cost is equal to $15,000$ GPU hours (1.75 GPU years), which approaches the limit of our computational resources.

The convergence of simulations is checked by comparing the radius of gyrations and the number of beads between the runs with the different initial configurations (chain conformations): rod and globule. As shown in Figure S3-S10, most of the R_{g} and N_{bead} values for these two starting configurations coincide within the statistical error. For intermediate Λ values ranging between 10 and 100, the PA necklace contains many beads, and its equilibration appears challenging. The reason is that the conformations of polyampholyte necklaces are highly fluctuating, with the bead unwinding being the rate-limiting dynamical process. At the intermediate Λ values, the number of beads reaches its maximum, thus creating a free-energy landscape with numerous local minima corresponding to the necklaces with different numbers of beads. As a result, it is more challenging to equilibrate such a system, so the N_{bead} values for simulation runs starting with the rod or globule conformations can be considered as the upper and lower bounds of the equilibrium value of the N_{bead}. As shown in Figures S11 and S12, the relative errors for estimating the R_{g} value are always lower than

5%, and those for N_{bead} are lower than 25%. Thus, we conclude that the sampling challenges do not affect our primary conclusions regarding the transition between regimes IIIΘ and IVΘ due to charge migration. Specifically, the number of beads exhibits a non-monotonic behavior, and simulations successfully reproduce the theoretical scaling slopes. Better sampling of these challenging systems using enhanced sampling methods remains for the future work.

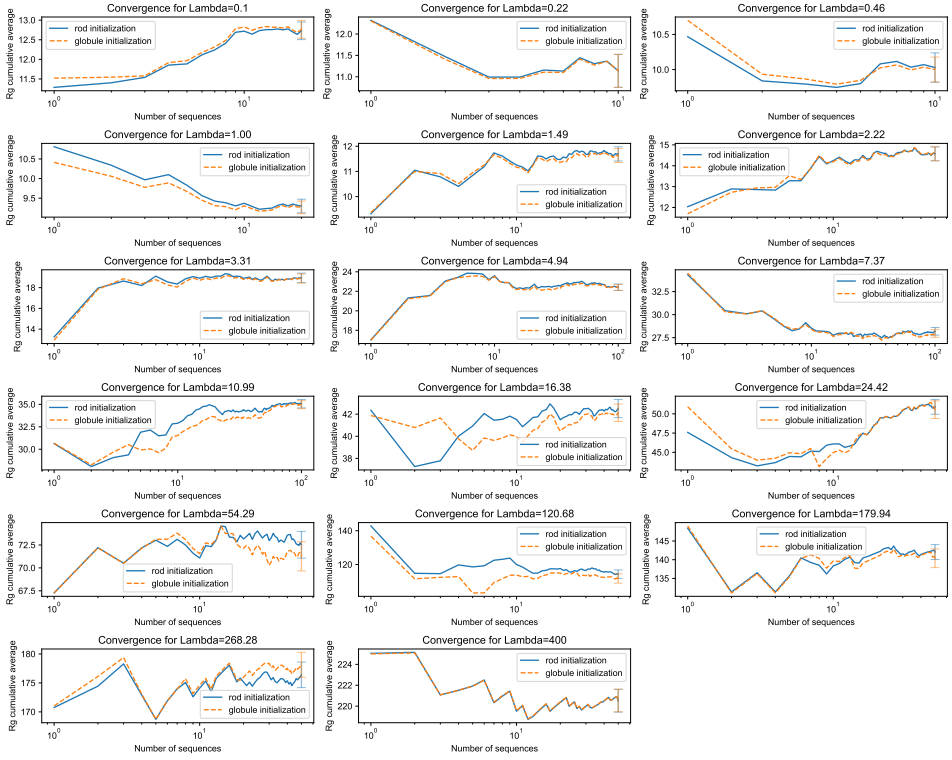

Figure S3: Convergence of radius of gyration R_{g} with respect to the number of sequence realizations at $u = 2$.

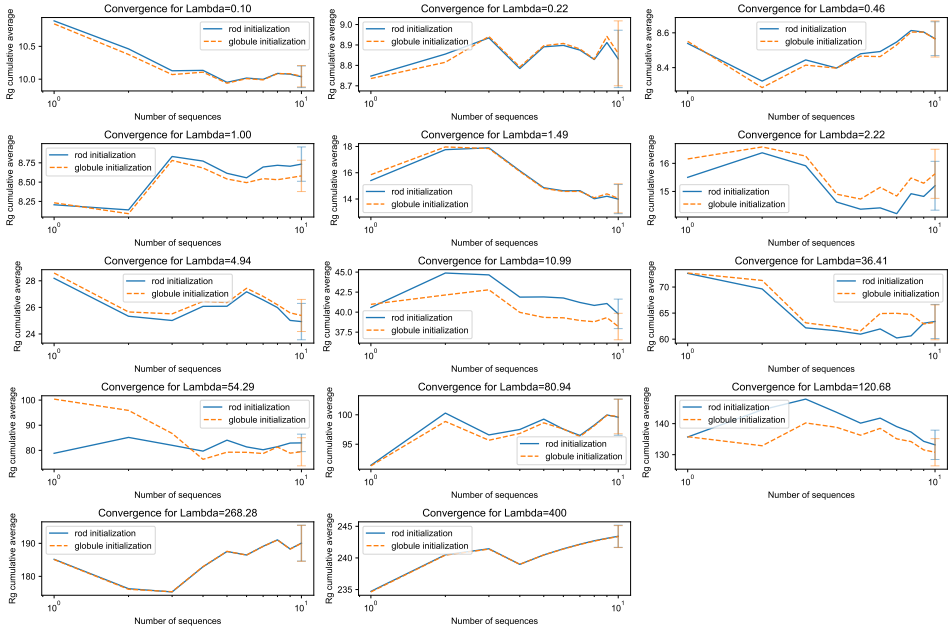

Figure S4: Convergence of radius of gyration R_g at $u = 3$

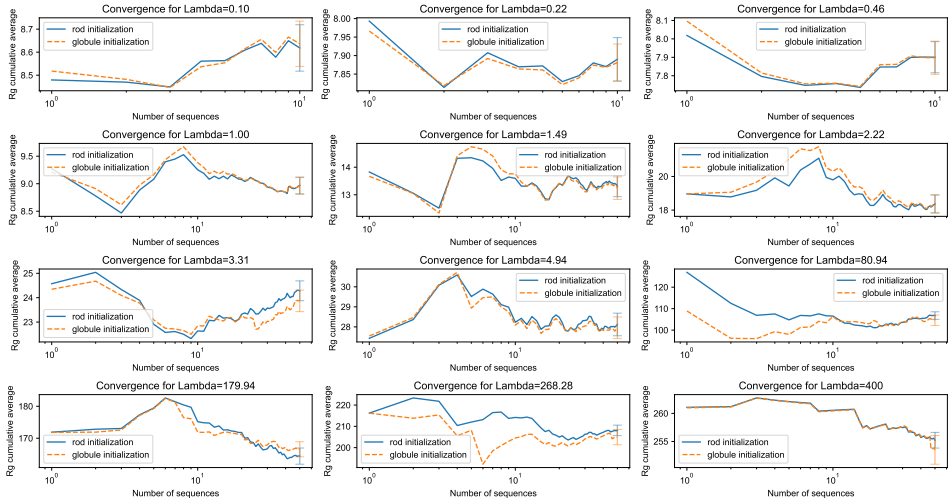

Figure S5: Convergence of radius of gyration R_g at $u = 4$

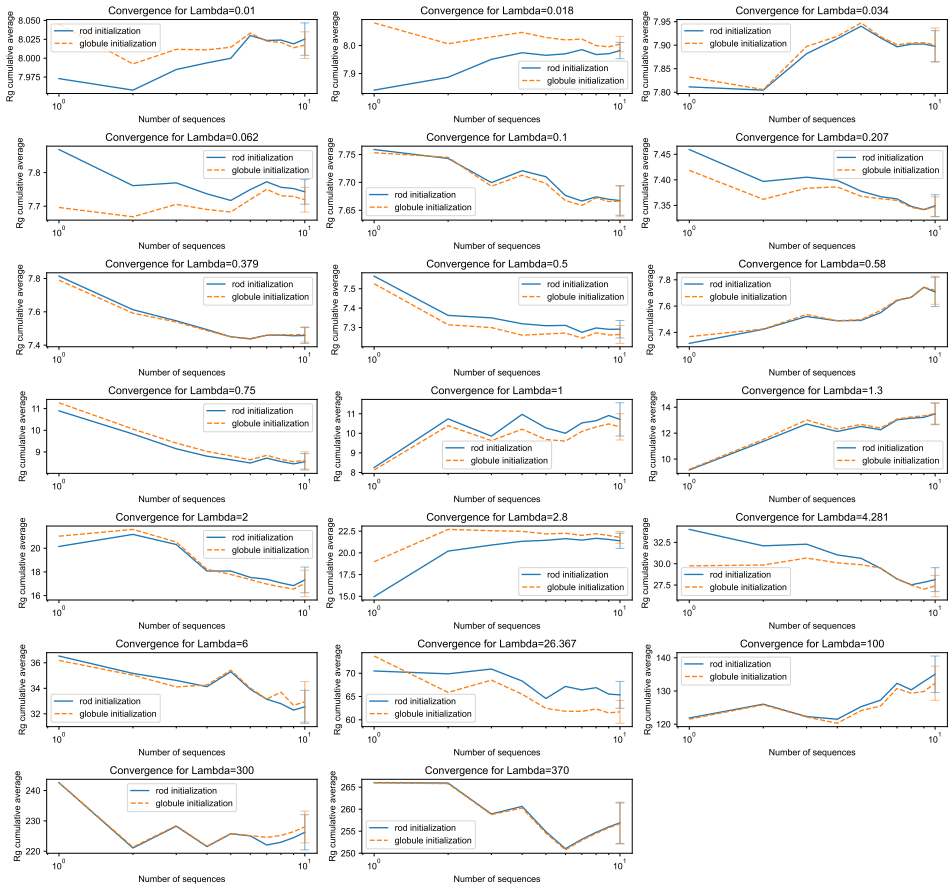

Figure S6: Convergence of radius of gyration R_{g} at $u = 5$

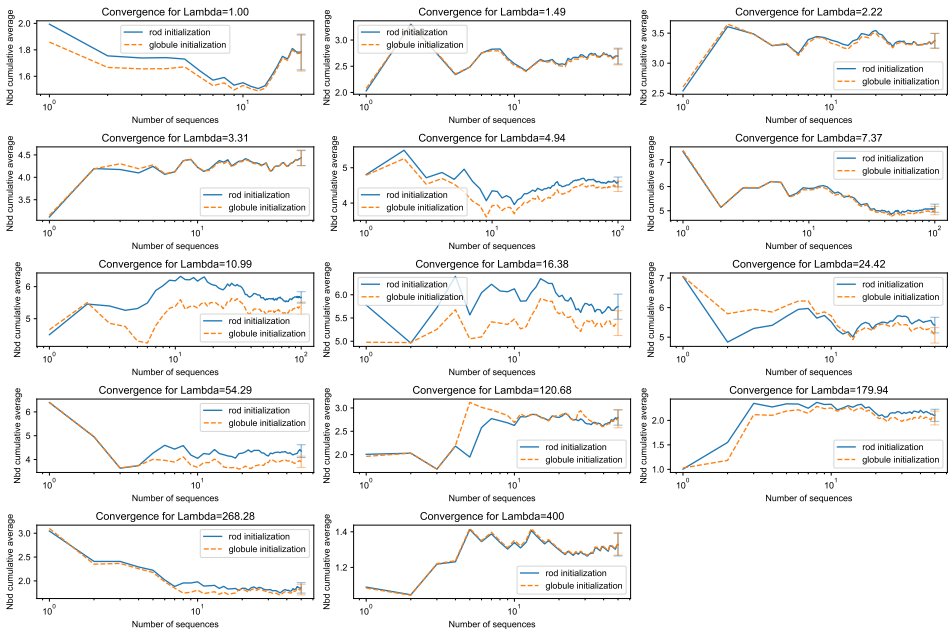

Figure S7: Convergence of number of beads N_{bead} at $u = 2$

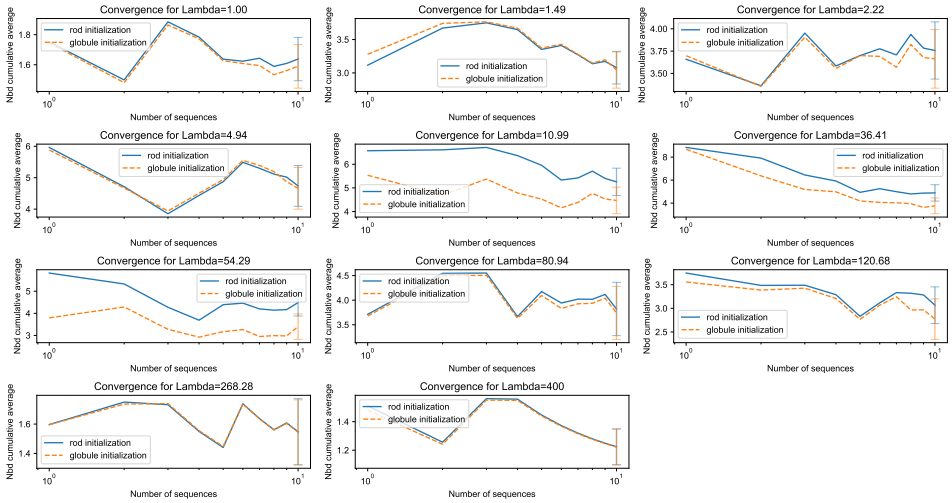

Figure S8: Convergence of number of beads N_{bead} at $u = 3$

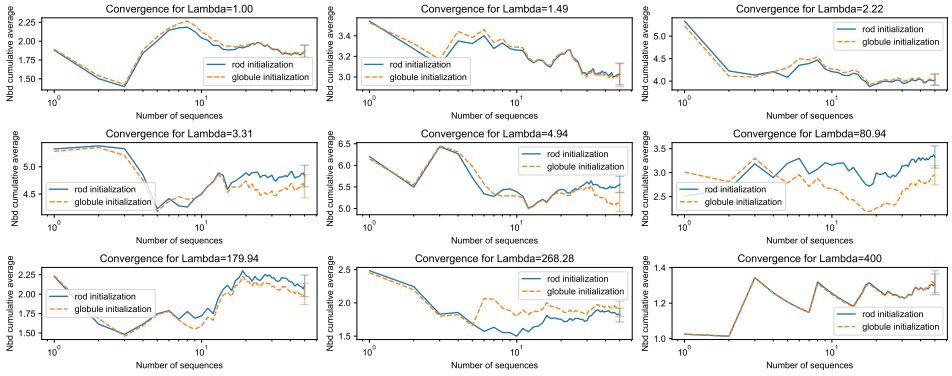

Figure S9: Convergence of number of beads N_{bead} at $u = 4$

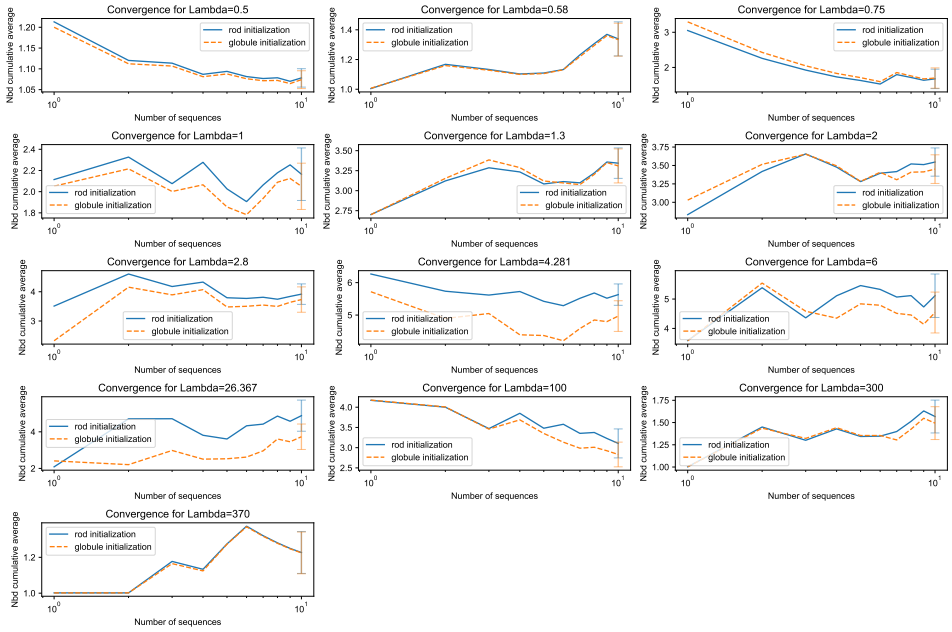

Figure S10: Convergence of number of beads N_{bead} at $u = 5$

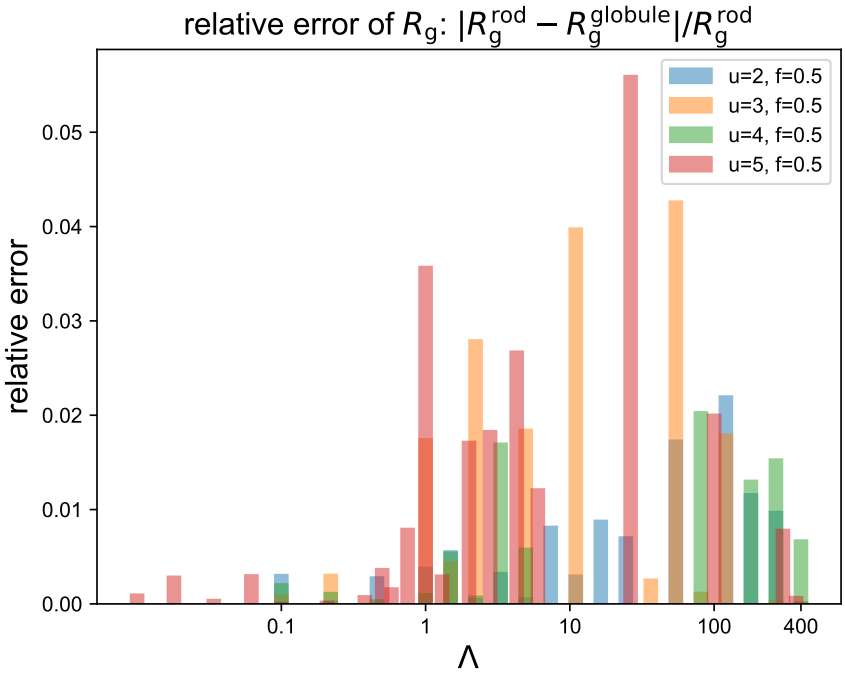

Figure S11: Relative error of the estimation of R_{g} value

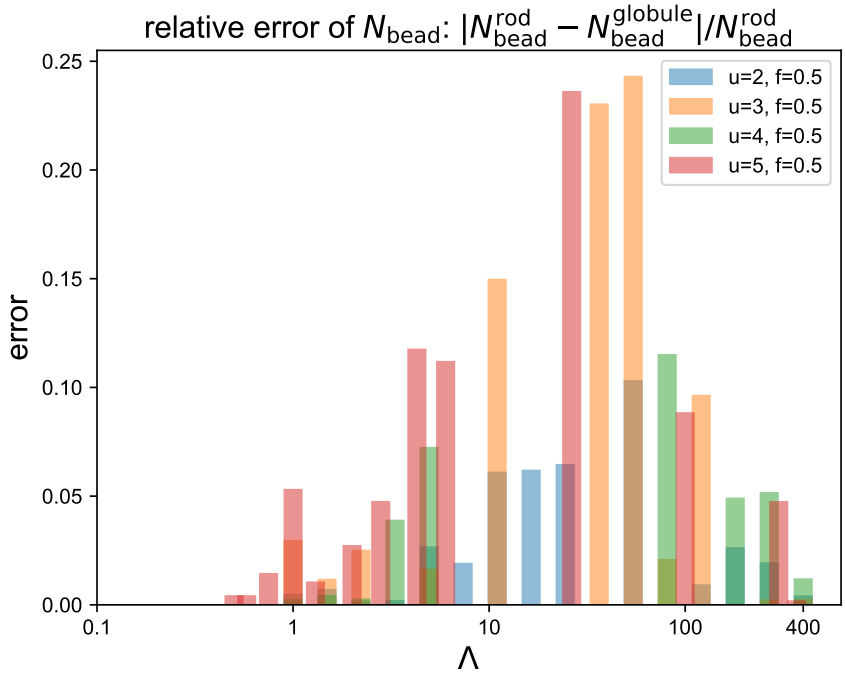

Figure S12: Relative error of the estimation of N_{bead} value

References

(1) Shusharina, N.; Zhulina, E.; Dobrynin, A.; Rubinstein, M. Scaling theory of diblock polyampholyte solutions. *Macromolecules* **2005**, *38*, 8870–8881.

(2) Dobrynin, A. V.; Rubinstein, M.; Obukhov, S. P. Cascade of transitions of polyelectrolytes in poor solvents. *Macromolecules* **1996**, *29*, 2974–2979.

(3) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D., et al. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **2017**, *13*, e1005659.

(4) Rumyantsev, A. M.; Johner, A.; de Pablo, J. J. Sequence Blockiness Controls the Structure of Polyampholyte Necklaces. *ACS Macro Lett.* **2021**, *10*, 1048–1054.

(5) Kremer, K.; Grest, G. S. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.* **1990**, *92*, 5057–5086.