Supplementary Materials for The Aggregation Free Energy Landscapes of Polyglutamine Repeats

Mingchen Chen[1,2], MinYeh Tsai[1,3], Weihua Zheng[1,3], and Peter G. Wolynes*[1,3,4,5]

[1]Center for Theoretical Biological Physics, Rice University, Houston, TX
[2]Department of Bioengineering, Rice University, Houston, TX
[3]Department of Chemistry, Rice University, Houston, TX
[4]Department of Physics and Astronomy, Rice University, Houston, TX
[5]Department of Biosciences, Rice University, Houston, TX
*Email: pwolynes@rice.edu, Phone: (713)348-4101

correspondence to: pwolynes@rice.edu

Contents

S1 Extended Simulation Details **2**
S1.1 Details of Simulated Annealing. 2

S2 Aggregation Free Energy Profile for Q_{26} at 0.1mM at 300K. **3**

S3 Supplementary Figures for free energy profiles originally calculated at 370K. **5**
S3.1 Aggregation Free Energy Profile for Q_{20} at 0.1mM calculated at 370K. 5
S3.2 Aggregation Free Energy Profile for Q_{30} at 0.1mM calculated at 370K. 6
S3.3 Aggregation Free Energy Profile for $Q_{12}\mathrm{PG}Q_{12}$ at 0.1mM calculated at 370K. 7

S4 β-arch structures are observed for Q_{40} at 0.1mM calculated at 300K. **8**

S1 Extended Simulation Details

S1.1 Details of Simulated Annealing.

The Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) was used to carry out all the reported simulations. We first performed structure prediction for monomeric polyglutamine repeat peptide of different lengths, Q_{20}, Q_{24}, Q_{26}, Q_{30} and Q_{40}, using the force field detailed in the main text. Following this, the simulated annealing simulations were performed from 500K to 300K using Nose-Hoover thermostat under non-periodic boundary conditions. A 5 fs timestep was used, and the coordinates for the protein system were saved every 2000 steps. As shown in Fig 1, the predicted structures are usually β hairpins for Q_{30} and Q_{40}. However, predicted structure for Q_{50} is a three-strand β sheet (Fig. 1S).

Figure S1: Predicted 3-strand β sheet for Q_{50}.

S2 Aggregation Free Energy Profile for Q_{26} at 0.1mM at 300K.

We performed umbrella sampling simulations for six Q_{26} peptides in the simulation system (the nominal concentration is $\sim$0.1 mM) at 370K, and computed the free energy profiles for forming aggregates of Q_{26} extrapolated to the physiological temperature 300K. The free energy profile is almost downhill when we use the number of contacts in the simulation box as a reaction coordinate. There is a free energy barrier that occurs during the restructuring of the monomeric hairpin, suggesting that the critical nucleus size is 1 for Q_{26} at 0.1mM. But the most favorable final fibre form structure turns out to be a mixture of hairpins and extended chains (Fig. S2A), instead of the simple pure anti-parallel hairpins that are formed for Q_{30}. As shown in the 1D free energy plot as a function of the size of oligomer, the free energy of forming an oligomer is downhill before forming an anti-parallel β-hairpin hexamer, while forming an anti-parallel β-hairpin hexamer is uphill (Fig. S2B). The free energy profile found after making the correction for changing the concentration of free monomers in the simulation exhibits a downhill behavior (Fig. S2B). The predicted critical concentration (c*) is between 1μM and 10μM at the physiological temperature.

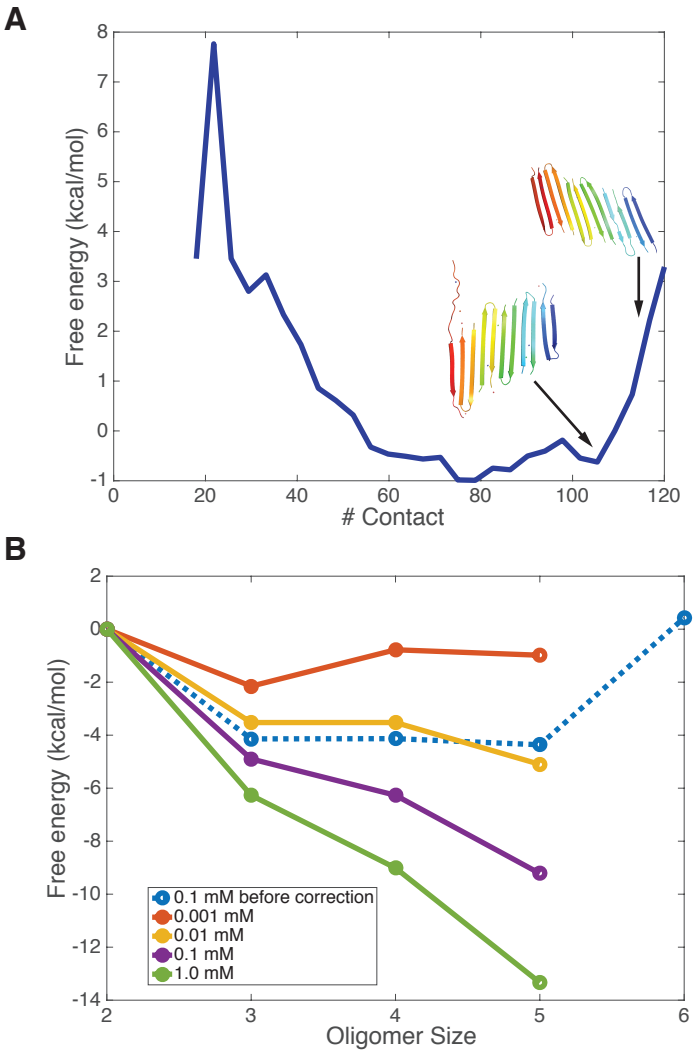

Figure S2: A): The free energy profile as a function of the number of total residue-residue contacts in the simulation system with six Q_{26} peptide chains at 300K. Representative structures found in each basin illustrate the progression through the different oligomeric states. (B): The grand canonical free energy for different oligomer states as corrected to have a fixed concentration of free monomers.

S3 Supplementary Figures for free energy profiles originally calculated at 370K.

S3.1 Aggregation Free Energy Profile for Q_{20} at 0.1mM calculated at 370K.

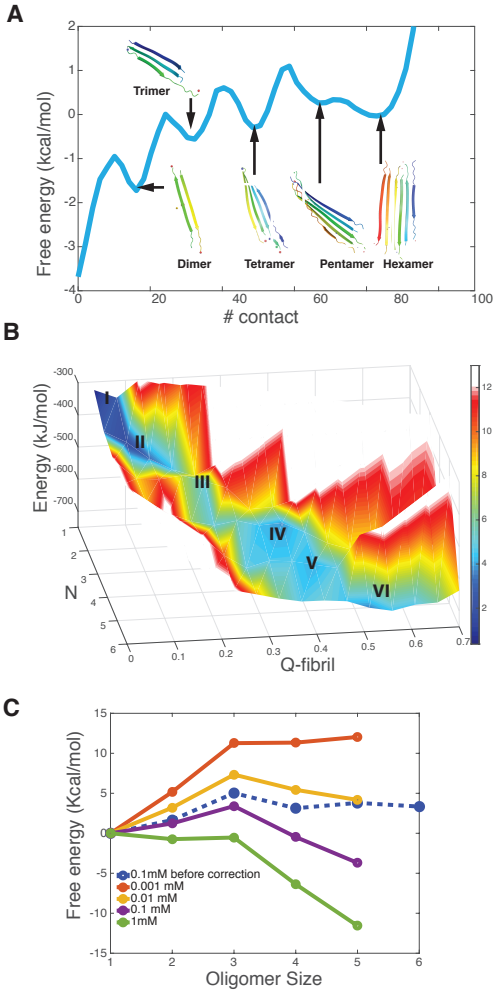

Figure S3: The aggregation free energy landscape for Q_{20} directly calculated at 370K. (A): The free energy profile as a function of the number of total residue-residue contacts in the simulation system with six Q_{20} peptide chains. Representative structures of different oligomers are shown in each free energy basin of the progress in different oligomeric states. (B): The energy and free energy surfaces for aggregation of Q_{20} are plotted as a function of the oligomer size (N), along with its structural similarity compared to the final fibre form (Q_{fibril}). The z-axis is the energy of the system, which decreases monotonically as the oligomer size increases. The color indicates the free energy. The local basins for different oligomer states are labelled by the size. A free energy barrier appears around N=3. (C): The grand canonical free energy profiles for different oligomer states as corrected for the monomer concentration changes in the fixed number simulation show the saturation value of the concentration.

S3.2 Aggregation Free Energy Profile for Q_{30} at 0.1mM calculated at 370K.

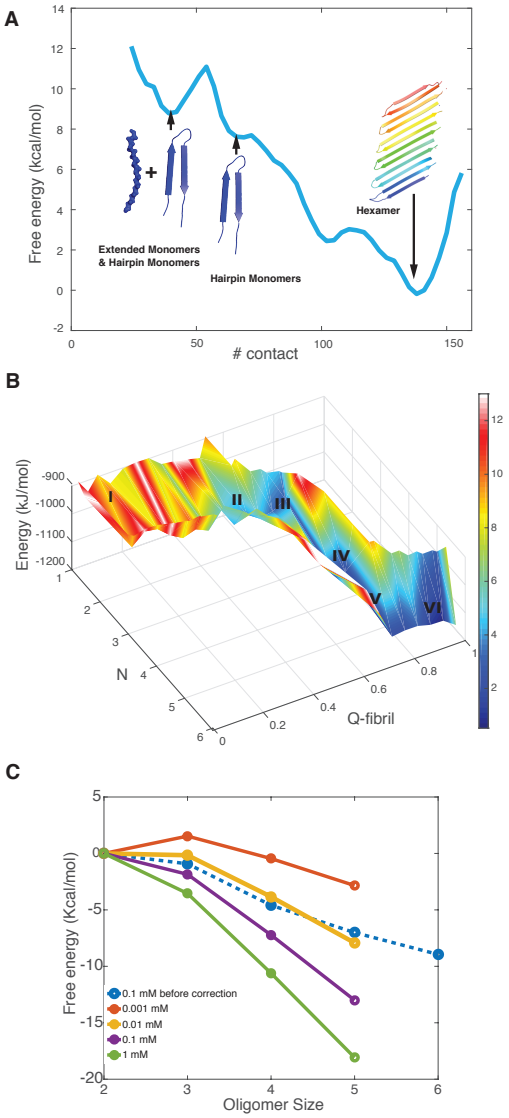

Figure S4: The aggregation free energy landscape for Q_{30} directly calculated at 370K. (A): The free energy profile as a function of the number of total residue-residue contacts in the simulation system with six Q_{30} peptide chains. Representative structures shown in each basin illustrate the progression through the different oligomeric states. B): The energy and free energy surfaces for aggregation of Q_{30} are plotted as a function of the oligomer size (N), and its structural similarity compared with the final fibre form (Q_{fibril}). The z-axis is the energy of the system, which decreases monotonically as the oligomer size increases. The color indicates the free energy, which includes the entropy cost of addition at concentration 0.1mM. The local basins for different oligomer states are indicated by their size. (C): The grand canonical free energy for different oligomer states as corrected to have a fixed concentration of free monomers.

S3.3 Aggregation Free Energy Profile for Q_{12}PGQ_{12} at 0.1mM calculated at 370K.

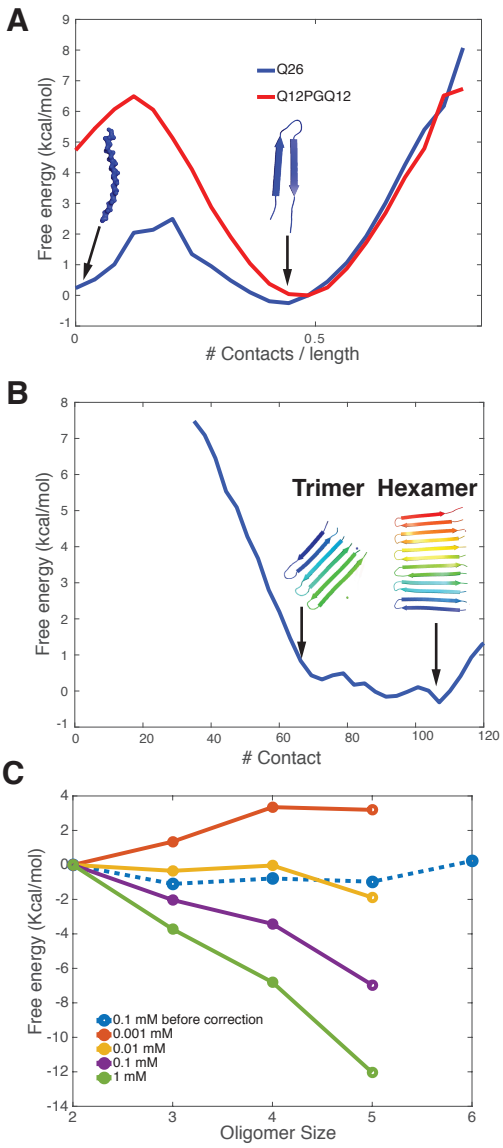

Figure S5: Changing the structural preference for 26 length repeats by making point mutations alters the nucleation behavior. (A) An inserted $L-PG$ mutation makes the hairpin conformation more favorable; (B) The free energy profile as a function of the total number of residue-residue contacts in the simulation system with six Q_{12}PGQ_{12} peptide chains calculated at 370K. Representative structures found in each basin illustrate the progression through the different oligomeric states. (C): Grand canonical free energy profile for different oligomer states as corrected to have fixed concentration of free monomers at 370K.

S4 β-arch structures are observed for Q_{40} at 0.1mM calculated at 300K.

For longer repeats (Q_{40}), we did observe some formation of β-arch structures. The observed configurations, however, are not purely composed of β-arch structures.

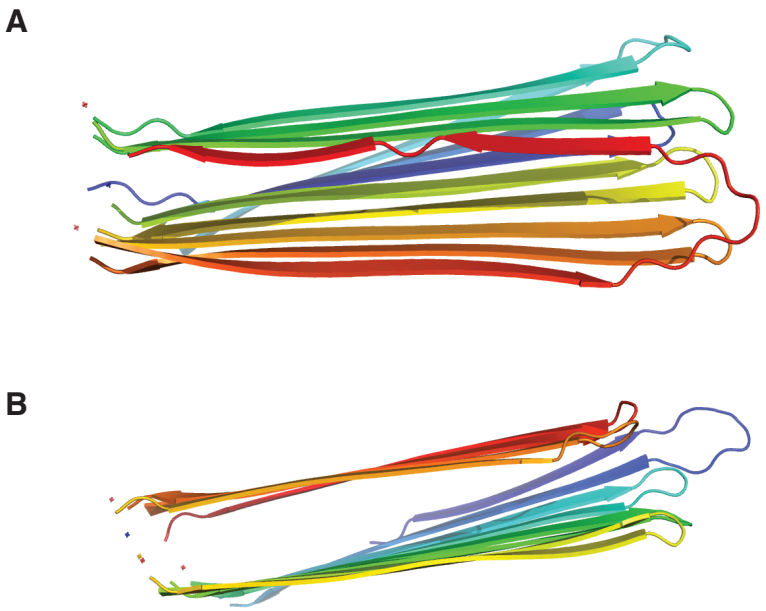

Figure S6: Selected intermediate structures for Q_{40} having different topologies. (A) A two-layered fibre form having both β-arches and β-hairpins. (B) A two-layered configuration composed solely of β-hairpins. The different component monomer are color-coded.