Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Protein droplets in systems of disordered homopeptides and the amyloid glass phase – Electronic Supplementary Information

Łukasz Mioduszewski and Marek Cieplak*

Institute of Physics, Polish Academy of Sciences,

Al. Lotników 32/46, 02-668 Warsaw, Poland

Coarse-grained model

The model we use is described in ref. [1]. A brief outline is as follows. We mark interacting pseudoatoms by indices i and j; $\vec{r}_i$ and $\vec{r}_j$ are their positions and $r_{i,j} = |\vec{r}_i - \vec{r}_j|$ is the relative distance. The mass of each residue is set to the average amino acid mass m. The positions evolve according to the Langevin equation with the time unit $\tau \approx 1\ ns$, damping coefficient $\gamma = 2m/\tau$ (corresponding to overdamped dynamics) and thermal white noise $\vec{\Gamma}_i$ with the variance $\sigma^2 = 2\gamma k_B T$:

$$m_i \frac{d^2\vec{r}_i}{d^2t} = \vec{F}_i - \gamma \frac{d\vec{r}_i}{dt} + \vec{\Gamma}_i \tag{1}$$

We solve the equations of motion by using the 5th order predictor-corrector algorithm [2]. We use energy unit ε where $0.35\varepsilon \approx k_B T$ for room temperature [3].

The existence of the excluded volume is ensured by the Lennard-Jones potential that is cut at $r_o = 5$ Å (so that $V_r(r_{i,j} \geq r_o) = 0$):

$$V_r(r_{i,j} \leq r_o) = \varepsilon \left(4 \left(\frac{\sigma_0}{r_{i,j}} \right)^{12} - 4 \left(\frac{\sigma_0}{r_{i,j}} \right)^{6} + 1 \right) \tag{2}$$

where $\sigma_0 = r_o \cdot (0.5)^{\frac{1}{6}}$, so that $V_r(r_o) = 0$. $i,\ j = i+2$ interactions are always described by that repulsive potential. The consecutive beads interact via the harmonic potential with the spring constant $k = 50\ \text{Å}^{-2} \cdot \varepsilon$ (consistent with refs. [4, 5]) and minimum at $r_b = 3.8$ Å.

Here we just briefly describe how we distinguish between backbone-backbone (bb), backbone-sidechain (bs) and sidechain-sidechain (ss) contacts in a one-bead-per-residue model: when two beads come close to each other, we compute two vectors based on positions of three consecutive beads $(i-1,i,i+1)$: $-\vec{n}_i$ is the negative normal vector, approximately pointing in the sidechain direction, and $\vec{h}_i$ is the binormal vector, pointing in the direction of a possible backbone hydrogen bond. Directions of these vectors are checked, and if, let's say, h_i and h_j point towards each other, a bb contact is made. Each type of amino acid has a specific number of contacts it can make with its virtual „backbone" and „sidechain". Each pair of amino acids has a different distance at the minimum of the Lennard-Jones potential corresponding to an ss contact between them (8.6 Å for a pair of glutamine residues, 6.4 Å for alanine, 5 Å for bb contacts and 6.8 Å for bs contacts). Each type of contact is modeled by the same potential (with the exception of bb interactions that are twice as strong):

$$V_{attr}(r_{i,j}) = 4\varepsilon \left(\left(\frac{\sigma_{attr}}{r_{i,j}} \right)^{12} - \left(\frac{\sigma_{attr}}{r_{i,j}} \right)^{6} \right) \quad (3)$$

Where σ_{attr} depends on the type of interaction. If criteria for creation of a given type of contact are met, attractive part of the potential is turned on quasi-adiabatically (during 10 τ). When the beads are more than $1.5\sigma_{attr}$ apart from each other, it is turned off in the same way.

The criterium for two chains to belong to the same cluster is that they should have at least one pair of residues with distance smaller than the corresponding ss contact distance (e.g., 8.6 Å for glutamine and 6.4 Å for alanine).

There are two changes relative to the description provided in ref. [1]:

- attractive electrostatic interactions are now treated as other sidechain-sidechain contacts (if the proper conditions are met) instead of being described by the modified Debye-Huckel potential. This correction enhances the agreement with experimental data for charged proteins. It has no impact on the results presented here, as polyglutamine and polyalanine are not charged.
- To achieve the best agreement with the data, the cutoff distance above which a contact is quasi-adiabatically turned off needs to be set to 1.3 σ, where σ is a parameter of the LJ potential. It is smaller than the previous value of 1.5 σ to account for the rapid breaking and reforming of the contacts in hydrophilic residues like glutamine, thus recreating the effect of water molecules disrupting the contacts.

* Electronic address: mc@ifpan.edu.pl

[1] Ł. Mioduszewski and M. Cieplak, Phys. Chem. Chem. Phys. **20**, 19057-19070 (2018).

[2] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids*, Oxford University Press, New York, 1987.

[3] J. I. Sułkowska and M. Cieplak, Biophysical Journal **95**, 317491 (2008).

[4] A. Korkut and W. A. Hendrickson, Proc. Natl. Acad. Sci. **106(37)**, 15667-15672 (2009).

[5] A. B. Poma, M. Chwastyk, and M. Cieplak. J. Phys.Chem. B. **119**, 12028-12041 (2015).

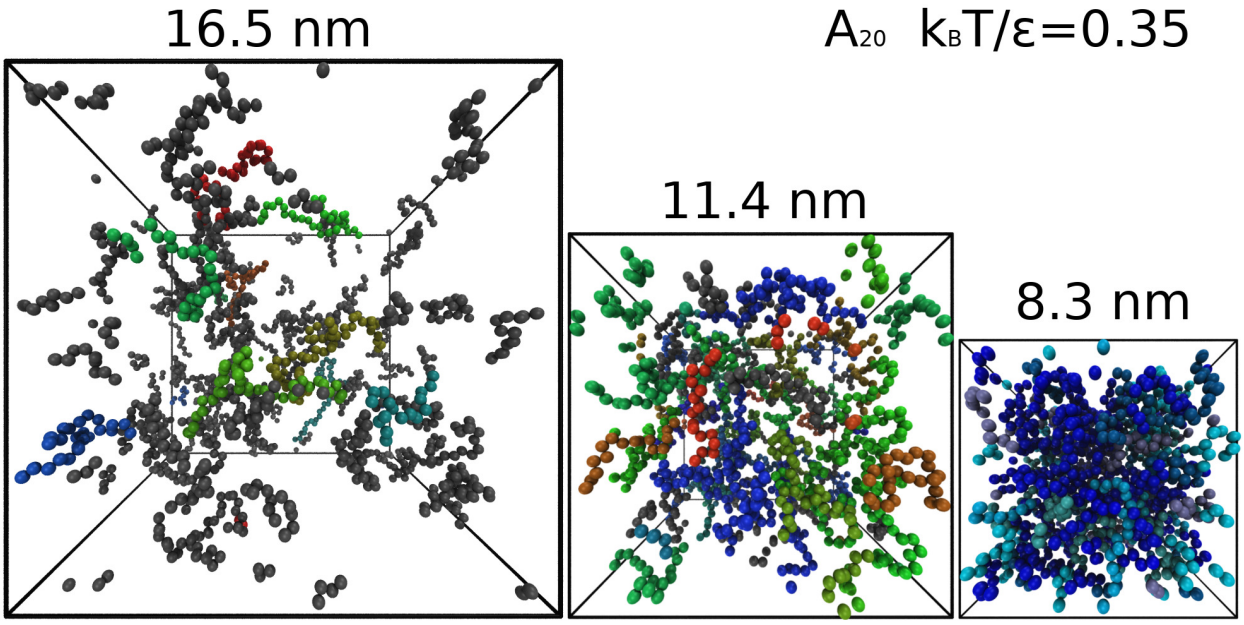

FIG. S1: Snapshots of 90 A_{20} chains at three densities. The panel on the left corresponds to S=16.5 nm (the G phase), in the center to 11.4 nm (the droplet phase), and on the right to 8.3 nm (the L phase). The corresponding densities are 0.4 1/nm^3, 1.2 1/nm^3 and 3.2 1/nm^3. The temperature is 0.35 ε/k_B.

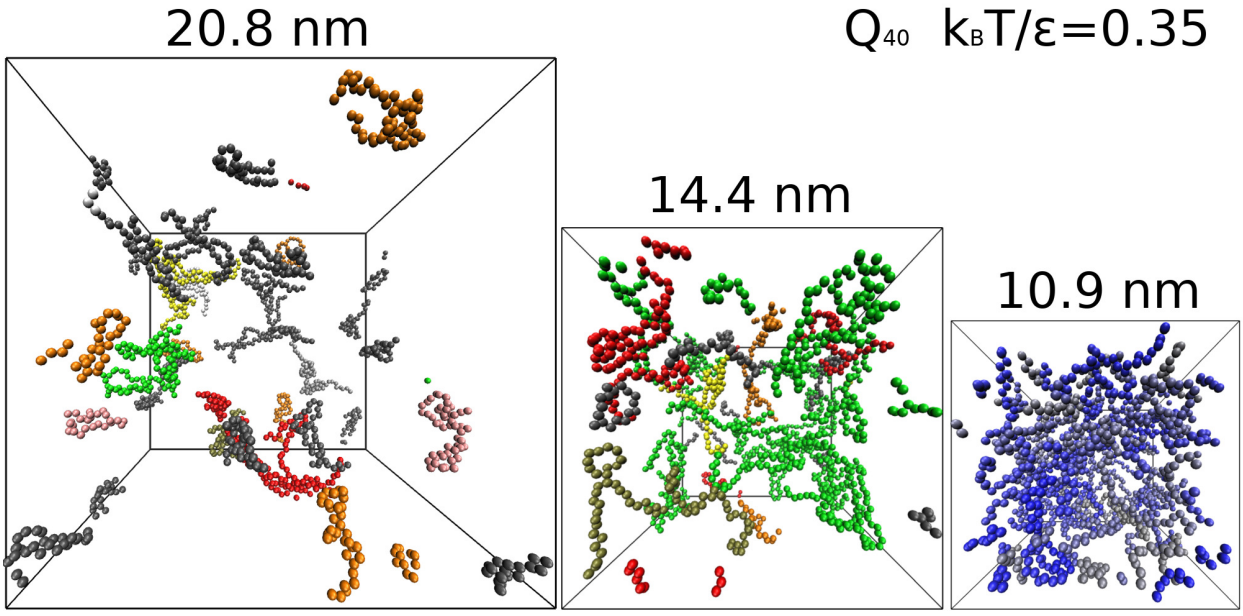

FIG. S2: Snapshots of 45 Q_{40} chains at three densities. The panel on the left corresponds to S=20.8 nm (the G phase), in the center to 14.4 nm (the droplet phase), and on the right to 10.9 nm (the L phase). The corresponding densities are 0.2 1/nm^3, 0.6 1/nm^3 and 1.4 1/nm^3. The temperature is 0.35 ε/k_B.

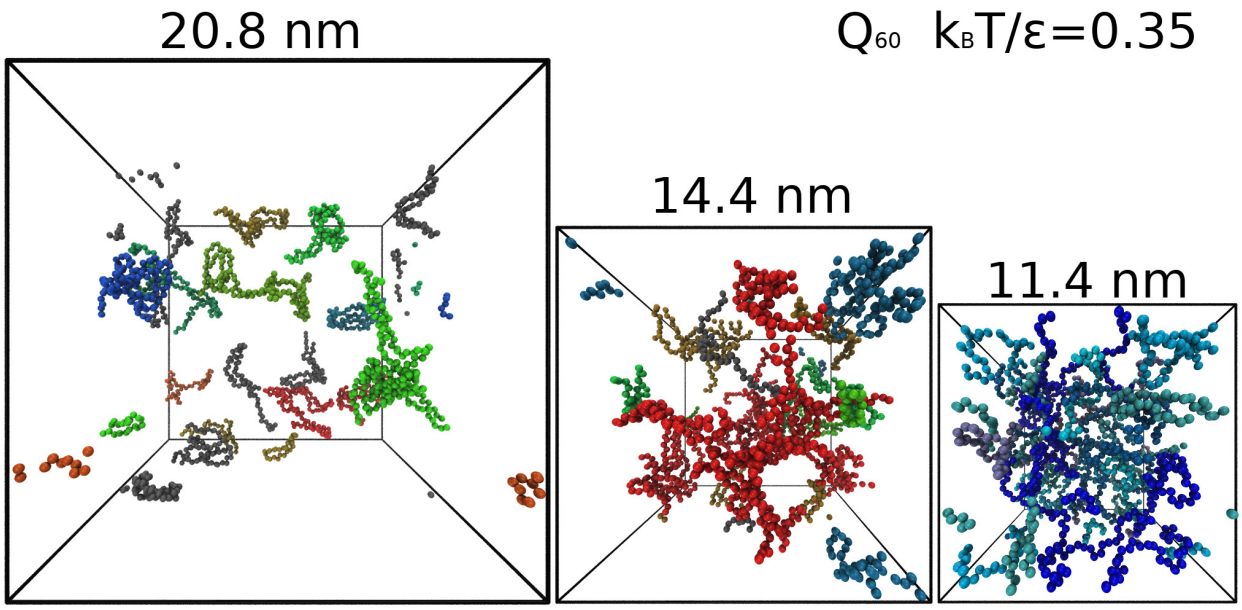

FIG. S3: Snapshots of 30 Q_{60} chains at three densities. The panel on the left corresponds to S=20.8 nm (the G phase), in the center to 14.4 nm (the droplet phase), and on the right to 11.4 nm (the L phase). The corresponding densities are 0.2 1/nm^3, 0.6 1/nm^3 and 1.2 1/nm^3. The temperature is 0.35 ε/k_B.

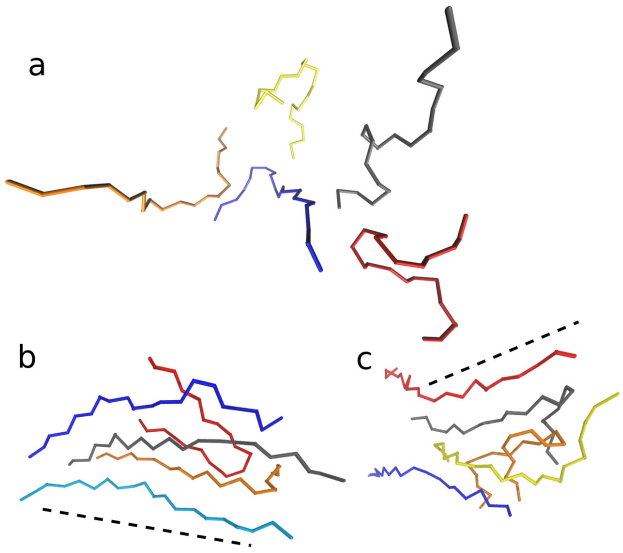

FIG. S4: The backbone alignment patterns observed in the A_{20} system. Panels a, b and c show 5 sample chains for a: $k_BT/\varepsilon = 0.35$, ρ=2 1/nm^3, b: $k_BT/\varepsilon = 0.2$, ρ=0.4 1/nm^3 and c: $k_BT/\varepsilon = 0.2$, ρ=2 1/nm^3.

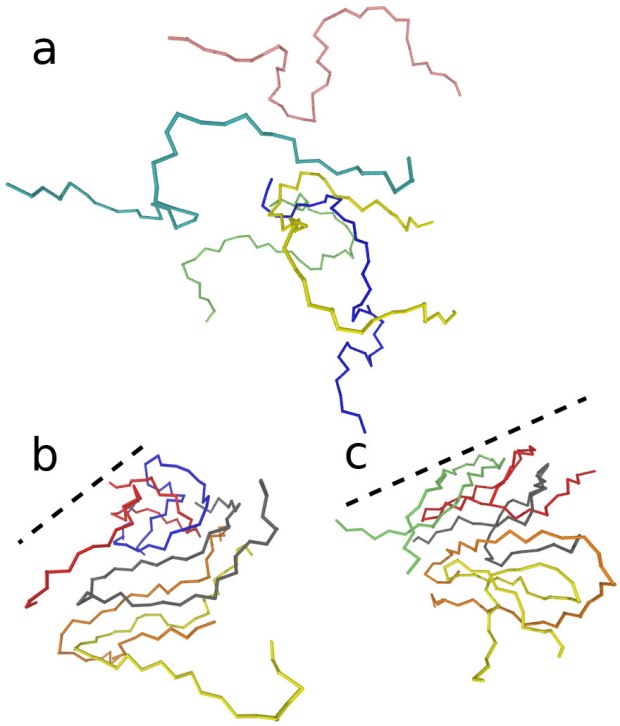

FIG. S5: The backbone alignment patterns observed in the Q_{40} system. Panels a, b and c show 5 sample chains for a: $k_BT/\varepsilon = 0.45$, ρ=1.2 1/nm^3, b: $k_BT/\varepsilon = 0.25$, ρ=0.2 1/nm^3 and c: $k_BT/\varepsilon = 0.25$, ρ=1.2 1/nm^3.

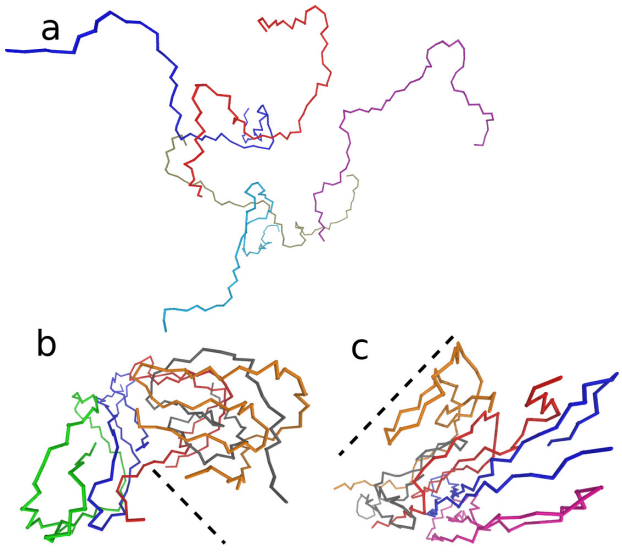

FIG. S6: The backbone alignment patterns observed in the Q_{60} system. Panels a,b and c show 5 sample chains for a: $k_BT/\varepsilon = 0.45$, ρ=1 1/nm^3, b: $k_BT/\varepsilon = 0.25$, ρ=0.2 1/nm^3 and c: $k_BT/\varepsilon = 0.25$, ρ=1 1/nm^3.

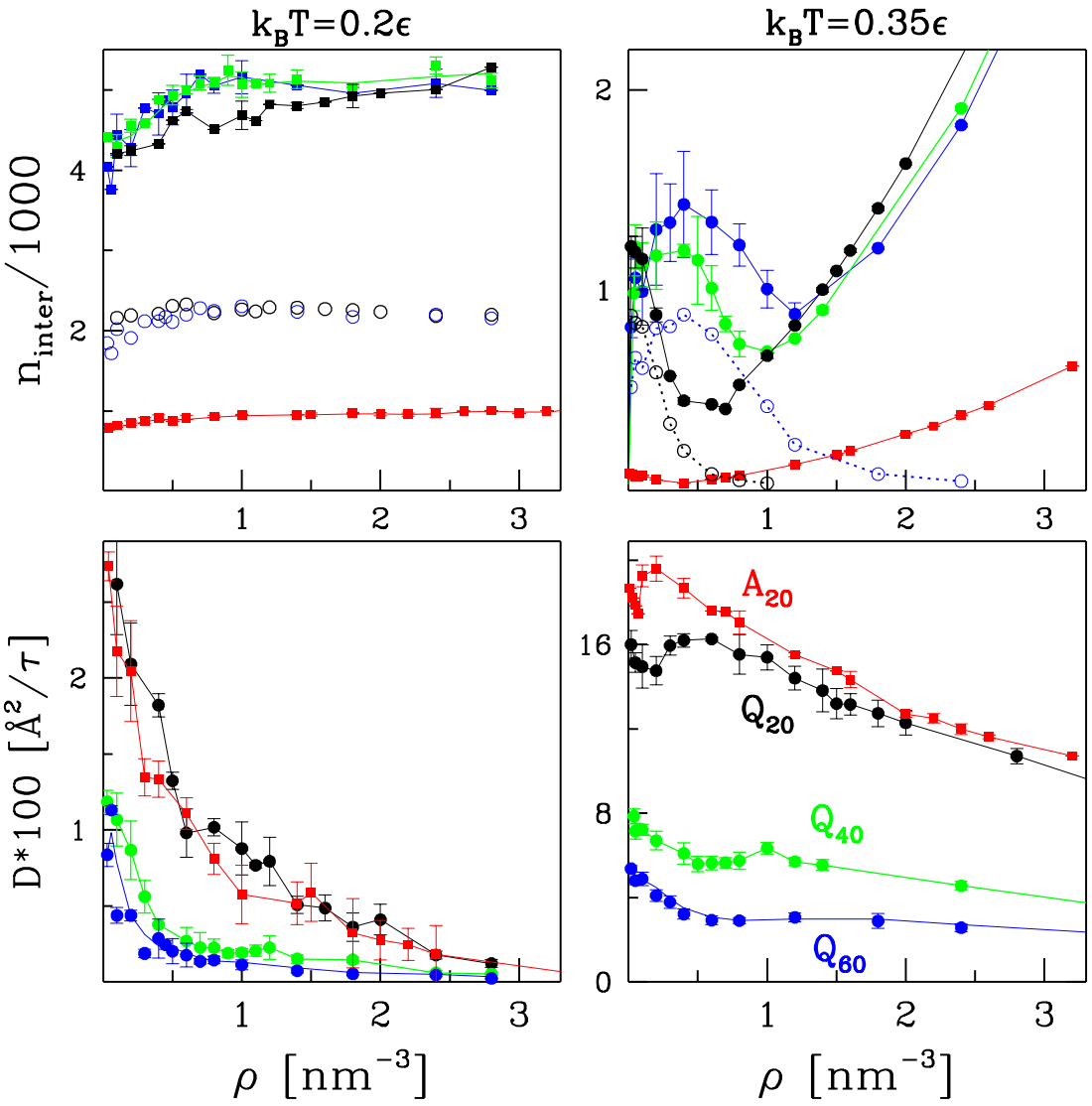

FIG. S7: The bottom panels show the diffusion coefficients as a function of ρ for A_{20} (red), Q_{20} (black), Q_{40} (green), and Q_{60} (blue). The left panel is for $T = 0.2\ \varepsilon/k_B$ and the right panel for $T = 0.35\ \varepsilon/k_B$. The upper panels show the corresponding numbers of the interchain contacts. The solid lines are for the distance-based criterion. The open circles (for N equal top 20 nd 60) correspond to contacts satisfying all necessary criteria, pertaining to the distance, angular characteristics and the allowed valence. For $t = 0.35\ \varepsilon/k_B$, the larger the system, the smaller the discrepancy between the two definitions at small densities. At high densities attractive interaction may arise through entanglements. The number of the interchain contacts initially rises with density, then it drops (as the growing ρ starts to disrupt formation of attractive contacts) and for large ρ's it rises again (because of the close packing). This nonmonotonic behavior is reflected in diffusion coefficients (more interchain contacts should hinder the diffusion), albeit in a noisy fashion.

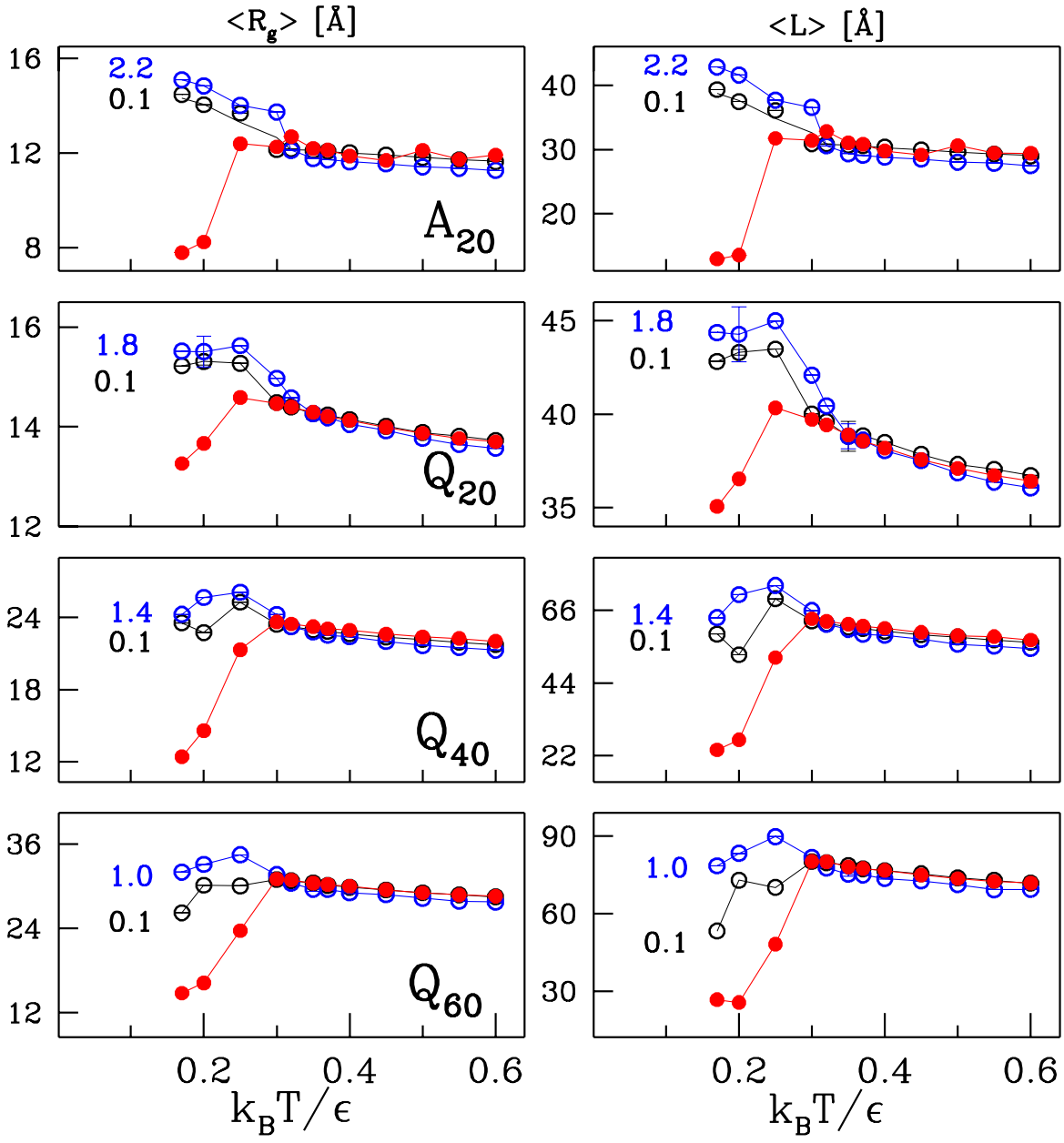

FIG. S8: The single-chain properties for the four systems studied as a function of T for the densities indicated (in units of 1/nm^3). The average radius of gyration on the left and the average end-to-end distance on the right.

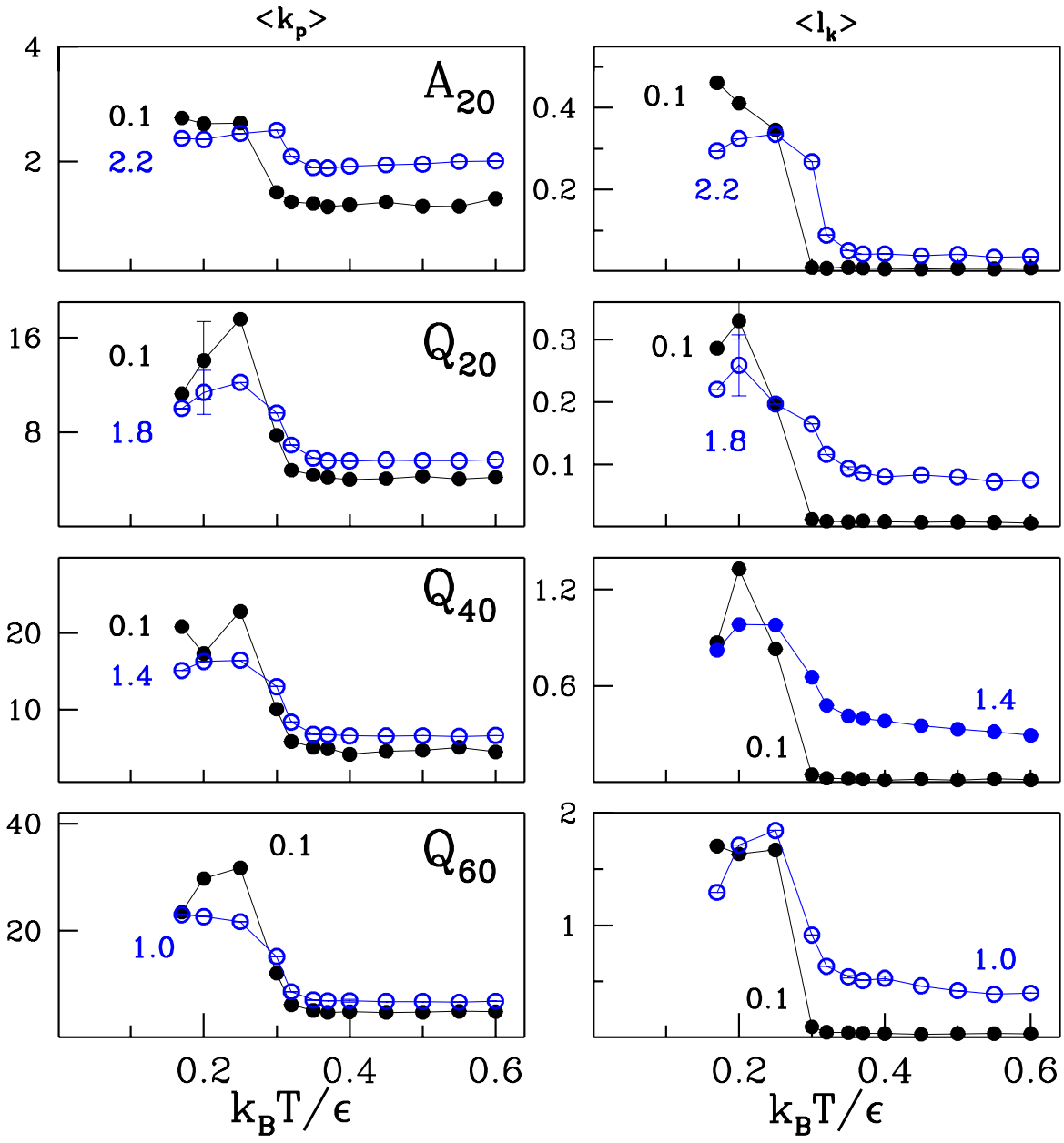

FIG. S9: The two-chain properties for the four systems studied: the average number of the distance-based contacts on the left and the average linking number on the right.

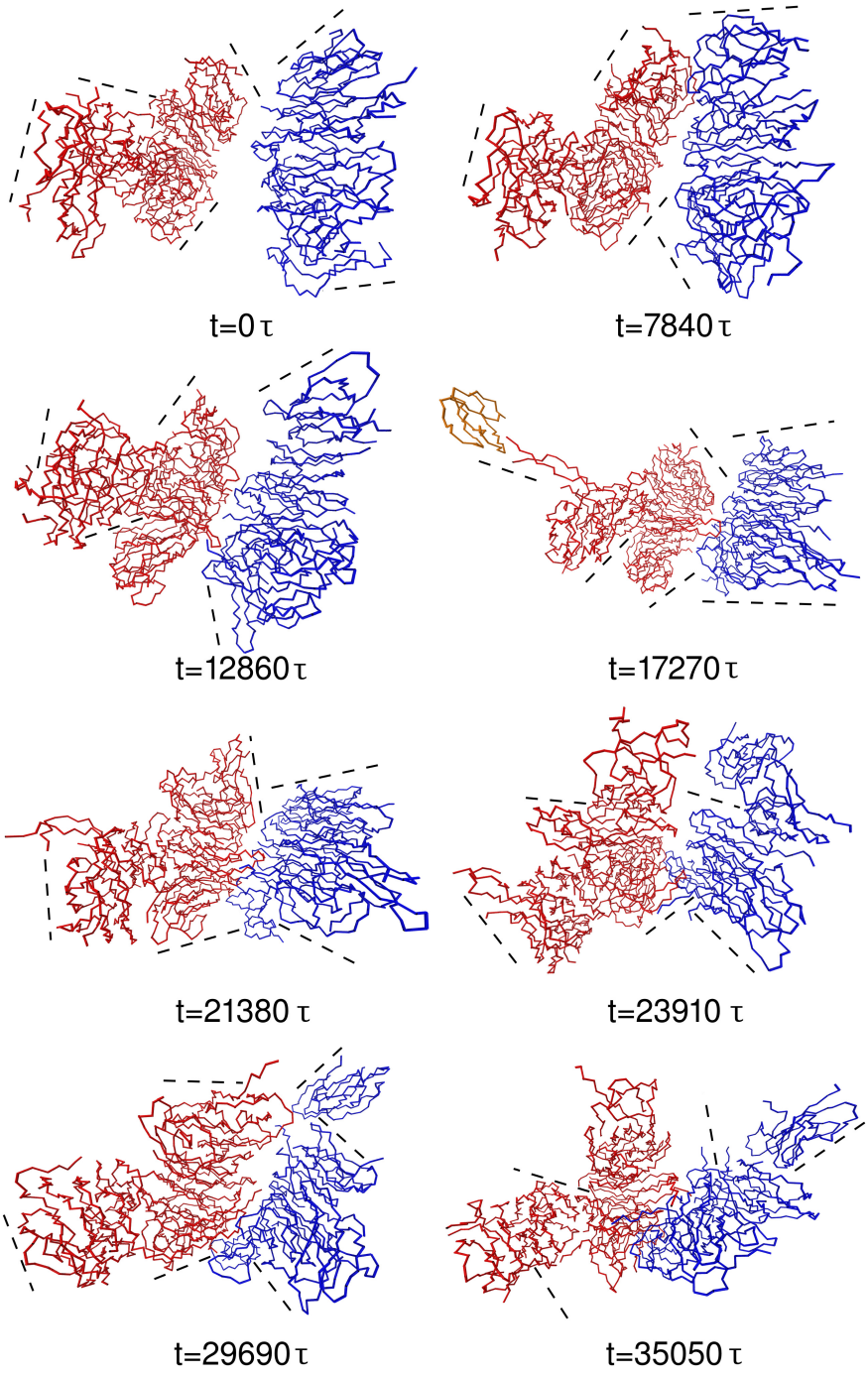

FIG. S10: An example of the time evolution of two clusters for Q_{60} at $T = 0.2\ \varepsilon/k_B$ in the span of 35 050 τ. In the first panel in the series of snapshots shown, there are two non-interacting clusters. In the final state, the two clusters merged into a single object. Formation of the first inter-cluster contact occurs between the first two panels and is visualized by Fig. 7 in the main text. In the 4th panel (on the right), one small cluster, shown in orange, temporarily separates from the cluster shown in red, but it returns later.

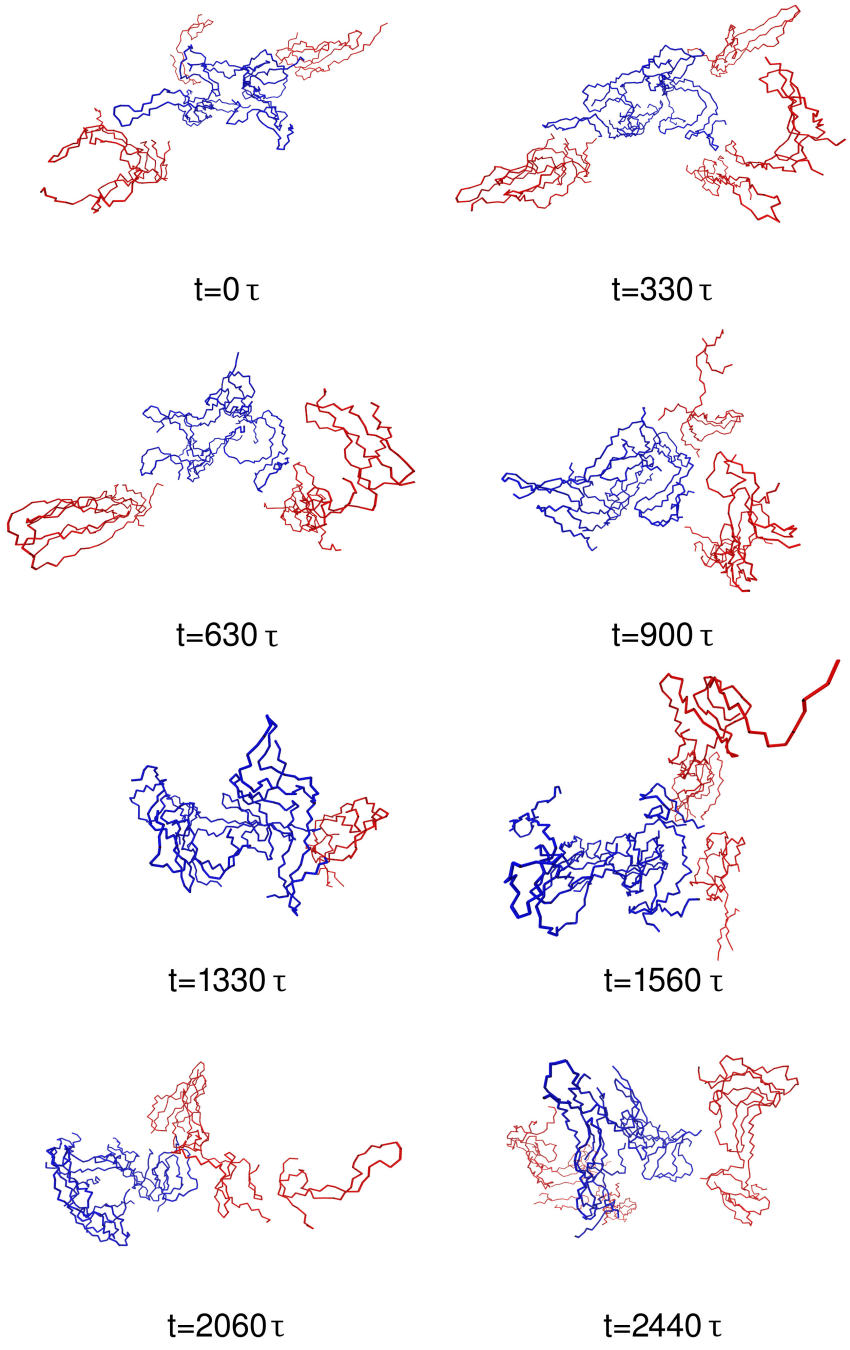

FIG. S11: An example of the time evolution of a cluster for Q_{60} at $T = 0.35\ \varepsilon/k_B$ in the span of 2 440 τ. The core of the cluster (blue) is made from 9 chains that stay together, while other chains (red) leave and reattach to the cluster in a sequence of merges and partitions.

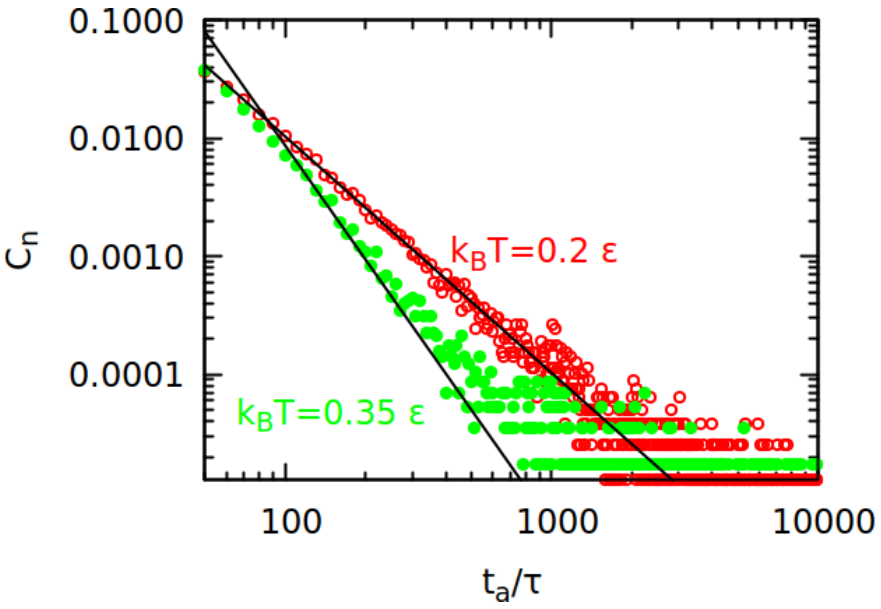

FIG. S12: Histograms of inter-chain contact duration time t_a with 10 τ bin size for $k_BT = 0.2\ \varepsilon$ (open circles) and $k_BT = 0.35\ \varepsilon$ (full circles). Two chains are in contact if there is at least one residue-level contact between them. Histograms were normalized by the total number of counts (78867 for $k_BT = 0.2\ \varepsilon$ and 57268 for $k_BT = 0.35\ \varepsilon$). Distributions were gathered from the last 90 000 τ of the non-equilibrium simulations of 80 Q_{60} chains (equilibrium distributions for 30 chains are similar, but finite-size effects are more pronounced). Functions with power law exponents -2 for $k_BT = 0.2\ \varepsilon$ and -3.2 for $k_BT = 0.35\ \varepsilon$ were fitted.