Supplementary Information for Topology of polymer chains under nanoscale confinement

Vahid Satarifard^{a,b‡}, Maziar Heidari^{a,c‡}, Samaneh Mashaghi^d, Sander J. Tans^e, Reza Ejtehadi^f, Alireza Mashaghi^{a,*}

^aLeiden Academic Centre for Drug Research, Faculty of Mathematics

, and Natural Sciences, Leiden University, Leiden, The Netherlands

^bMax Planck Institute of Colloids and Interfaces, Potsdam, Germany

^cMax Planck Institute for Polymer Research, Mainz, Germany

^dSchool of Engineering and Applied Sciences and Department of Physics,

Harvard University, Cambridge, Massachusetts, USA

^eFOM institute AMOLF, Amsterdam, The Netherlands ^fDepartment of Physics, Sharif University of Technology, Tehran, Iran and

tEqual contribution

(Dated: July 20, 2017)

I. CONTACT PROBABILITY ANALYSIS

In order to determine the conformational sampling, we calculate contact probability (CP) of monomers collisions. The contact probability among the beads of the polymer chains are calculated by [1, 2]

$$P_{i,j} = \frac{1}{N_c} \sum_{N=1}^{N_c} \Theta\left(d_c - d_{i,j}^N\right)$$
(S1)

Here N_c is total number of conformations, $d_{i,j}$ is the spatial separation of i^{th} and j^{th} monomers in the N^{th} conformation; Θ is the Heaviside step function which becomes 1.0 when the distance is lower than a specified cut-off radius $(r_c/\sigma = 1.5)$ and 0.0 otherwise.

The contact probability for 2×10^5 conformations in two confinement radii, $R_c/\sigma = 10$ and $R_c/\sigma = 5$ are depicted in Fig. S1. For the sake of clarity and comparability of CPs, we do not calculate the contact probabilities of identical residues (self contact) and bonded neighbour residues which are in a fully contacted state based on our analysis criteria. The nearly homogeneous probability of monomer collisions indicates that sampling is sufficient for ensemble analysis. Moreover, the narrow ribbon around diagonal elements having comparatively low probability, indicates that confinement creates long-ranged contacting regimes in the contact probability matrix which consequently influence topological state of polymer chain.

II. CHAIN WITH DIFFERENT PERSISTENCE LENGTH

The same confining procedures as described in the main text are performed on the chain with different persistence lengths (l_p) . The topological fractions of the results are sorted by monotonically increasing ratios of the persistence length to confining radius (l_p/R_c) and represented in Fig. S2.



FIG. S1: (color online) Contact probability matrix for 2×10^5 conformations in different confinement radii; (a) $R_c/\sigma = 10$ and (b) $R_c/\sigma = 5$.





III. CHAIN WITH FOUR BINDING SITES

Four monomers of a chain whose length is 200 monomers are randomly selected and the strength of their attractive interactions are set to $\epsilon = 5k_BT$. The number of the binding sites along the chain and their corresponding contact orders for ten different groups are shown in

TABLE S1: The Positions of the random binding sites in a chain having 200 beads. The groups are ordered in terms of increasing magnitude of contact order (CO)

found in topologies of series configurations.

	Binding Sites				СО		
	i	j	k	1	S	Р	Х
Group 1	21	39	136	172	0.135	0.62	0.62
Group 2	20	65	184	199	0.15	0.745	0.745
Group 3	61	110	124	135	0.15	0.22	0.22
Group 4	47	59	81	141	0.18	0.29	0.29
Group 5	27	69	137	193	0.245	0.59	0.59
Group 6	3	55	107	169	0.285	0.545	0.545
Group 7	51	113	139	198	0.302	0.43	0.43
Group 8	5	106	136	165	0.325	0.475	0.475
Group 9	21	118	149	193	0.352	0.51	0.51
Group 10	16	75	99	186	0.365	0.49	0.49

Table S1. In each group, binding sites are able to attach one another to construct different topologies named as series, parallel and cross topologies. The groups in Table S1 are sorted by monotonically increasing contact orders of series configurations. Afterwards, a confining sphere envelopes the chain and gradually decreases its radius. In these cases, the topological analyses are solely done over two contacts.

In all simulations, the system is equilibrated for 10^5 MD time-steps and then the spherical confinement, whose radius (R_c) gradually decreases, is introduced. For each confinement radius, 2×10^5 polymer conformations are sampled for 5×10^7 MD time-steps. The fraction of topologies averaged over all ten groups are shown in Fig. S3. The trends of fractional changes are similar to the self-avoiding polymer chain and interestingly, all topological fractions become equal at the confining radius $(R_c = 10\sigma)$. The averaged radius of gyration of each group when the chain does not experience the confinement are shown in Fig. S4. As shown, when the binding sites are connected through topological series configurations, the chain has a larger radius of gyration than the other two topological circuits. The entropy loss and folding rates of free chains are calculated via the methods introduced by Makarov and Rodin [3, 4] and are shown in Fig.S5 and S6. Chains with contacts of series topology have less entropy loss in comparison with the other two topologies and accordingly this leads to enhancement in folding rates of the loops categorized in this class of topology. Although the loops of free chain which are connected with series topological circuits have higher folding rates than the other topologies, our confined chain results reveal that the loops with cross or parallel topology would be more probable in highly packed environments. Thus, it is necessary to investigate the kinetics and folding rates of the loops with different topologies when the chain is



FIG. S3: (color online) Fraction of topologies observed in MD simulations of linear chains with ten sets of randomly positioned binding sites (Table S1).



FIG. S4: (color online) Average of radius of gyration for three topological arrangements, by positioning four random binding sites on the chain (Table S1). The simulations were performed in the absence of confinement.

confined.

IV. SOLVENT QUALITY EFFECT

To study the effect of solvent quality on the topological circuits of the chain, we firstly investigated the scaling behaviour of mean radius of gyration with chain length; i.e. $R_g \sim N^{\nu}$, where N is the number of polymer beads, R_g is the radius of gyration and ν is the Flory exponent. It has been shown that for good solvent $\nu = 0.588$ and for poor solvent $\nu = 0.33$ [5]. The quality of implicit solvent is set by changing the potential strength of the Lennard-Jones potential (ϵ) and cut-off radius (r_c). The values $\epsilon = k_B T$ and $r_c = 2^{1/6} \sigma$ have been used



FIG. S5: (color online) Entropy loss for formation of three topological arrangements, by positioning four random binding sites on the chain (Table S1); they could form two contacts with distinct topologies.



FIG. S6: (color online) Folding rate of three topologies in polymers with four randomly positioned binding sites (Table S1). All four binding sites are able to form two contacts with distinct topologies.

to simulate good solvent conditions. The poor solvent regimes are simulated by $\epsilon = 1.25, 1.5, 1.75, 2, 2.25k_BT$ and $2.5k_BT$ and $r_c = 2.5\sigma$. In order to obtain an equilibrium polymer structure, after each incremental increase in the strength of the potential ($\Delta \epsilon = 0.1k_BT$), the system is equilibrated for 10^6 steps. Then, 1×10^5 polymer conformations are sampled for 2.5×10^7 MD timesteps to calculate the mean radius of gyration.

As shown in Fig.S7 a, the exponent $\nu = 0.578$ is obtained for the strength of $\epsilon = k_B T$ and the exponent $\nu = 0.325$ is obtained for $\epsilon = 1.75k_B T$. The resulting exponents are very close to the expected values for good solvent and poor solvent, respectively [5]. Fig. S7 b shows that in good solvent condition, polymer chain

TABLE S2: Mean radius of gyration in different packing ratios and different spherical confinement radii.

	ľ	N=200	N=400		
PR	R_c/σ	R_g/σ	R_c/σ	R_g/σ	
1.600	5	3.27 ± 0.02	6.299	4.27 ± 0.02	
0.390	8	5.08 ± 0.12	10.079	6.59 ± 0.12	
0.200	10	$6.16 \ {\pm} 0.22$	12.599	$8.01 \ {\pm} 0.22$	
0.025	20	$10.29{\pm}1.04$	25.198	13.74 ± 1.08	
0.007	30	$12.57 {\pm} 1.90$	37.798	17.43 ± 2.30	

mostly forms independent series loop similar to the noconfinement and weak confinement condition. However, in the poor solvent condition, the polymer chain starts to collapse and forms a globular structure. Similar to the strong spherical confinement, the poor solvent condition brings the topological state of the chain into an equally distributed fraction.



FIG. S7: (color online) (a) The mean radius of gyration as a function of chain length for different values of attraction parameter ϵ , the dashed lines are the linear fit, (b) the topology fraction as a function of attraction parameter ϵ . The obtained values for the Flory exponent are approximately 0.578 and 0.325 for good and poor solvent, respectively.

V. POLYMER LENGTH EFFECT

We studied the effect of polymer chain length by doubling the number of beads N=400. As shown in Fig. S8 the topology fraction and average contact order does not change significantly when we compare it with the polymer chain with N=200 beads. We note that, despite the increase in the length, we are still in the short polymer chain regime. We will address the effect of polymer length on scaling behaviour in our future studies.

As a well-defined measure of confinement, we used the packing ratio (PR) which is defined as PR=(total bead volume)/(spherical confinement volume). This enables us to scale and compare the different levels of confinement for the long and short polymer lengths. For instance, the packing ratio of a polymer with N = 400beads confined in $R_c/\sigma = 6.299$ is equal to a polymer having N = 200 beads where confined in $R_c/\sigma = 5.0$, Table S2.. We used confinement radius rather than the radius of gyration in order to present confinement condition. Confinement radius has the benefit that one can simply calculate PR to shift the confinement condition to the comparable state for longer as well as shorter polymer chains. Radius of gyration R_q can, however, be used as a measure for evaluation of strong confinement versus weak confinement. Table S2 shows the results of R_q in different packing ratios and different spherical confinement radii.

VI. BROWNIAN DYNAMICS SIMULATIONS

If the sampling is not enough one would expect that the polymer chain would be trapped in different available states with specific and distinct topologies. Then by changing the simulation methodology and initial conditions the outcomes will not remain reproducible. In order to investigate the effect of sampling in strong confinement condition we performed Brownian Dynamics simulations. We used the new time-discretization algorithm of the Langevin model [6, 7] which enabled us to take longer timesteps without loosing stability. To maintain the stability of BD simulations, the time step and damping parameter for Langevin thermostat are chosen to be $\Delta t = 0.02\tau$ and $\lambda = 10\tau^{-1}$, respectively. The time step is four times larger than the one used in MD simulations. In each simulations we started from different initial conditions and introduced spherical confinement by gradually reducing the radius of confining sphere. Before sampling, the number of timesteps $10^4 \tau$, $5 \times 10^4 \tau$, $10^5 \tau$, $2 \times 10^5 \tau$ and $10^6 \tau$ are taken to confine a free linear polymer chain N=200 inside a sphere $R_c/\sigma = 5$, the legend in Fig. S9 indicates the number of these steps. Afterward, 10^4 polymer conformations are sampled for



FIG. S8: (color online) Fraction of topological circuits and corresponding averaged contact order as a function of confining radius for polymer chain N=400. The cut-off radius of $r_c = 1.5\sigma$ is used to evaluate loop formation and to estimate the fraction of topological structures.

 2.5×10^6 MD time-steps. The sampling is done after 250 steps (Fig .S9 a) and after 500 steps (Fig .S9 b). As shown in Fig .S9 the topology fractions are within the error bars and are similar to the results from MD simulations with much smaller timesteps. Thus, one can conclude that the simulation methodology, sampling scheme and initial conditions do not have significant effect on the statistical analysis of outcomes.

- O. Mashinchian, S. Bonakdar, H. Taghinejad, V. Satarifard, M. Heidari, M. Majidi, S. Sharifi, A. Peirovi, S. Saffar, M. Taghinejad *et al.*, ACS applied materials & interfaces, 2014, 6, 13280–13292.
- [2] D. Meluzzi and G. Arya, Nucleic acids research, 2012, 41, 63–75.
- [3] D. E. Makarov and G. J. Rodin, *Physical Review E*, 2002, 66, 011908.
- [4] A. Szabo, K. Schulten and Z. Schulten, The Journal of chemical physics, 1980, 72, 4350–4357.
- [5] P. Flory, New York, 1988, 401–403.
- [6] N. Grønbech-Jensen and O. Farago, *Molecular Physics*, 2013, **111**, 983–991.
- [7] N. Grønbech-Jensen, N. R. Hayre and O. Farago, Computer Physics Communications, 2014, 185, 524–527.



FIG. S9: (color online) Fraction of topological circuits in confining radii $R_c/\sigma = 5$ for linear polymer chain N=200. Sampling is done in (a) every 250 steps and (b) every 500 steps to sample 10000 polymer conformations. The legend indicates the number of steps in order to gradually confine a free chain inside a sphere $R_c/\sigma = 5$. The cut-off radius of $r_c = 1.5\sigma$ is used to evaluate loop formation and to estimate the fraction of topological

structures.