Supporting information for: "Probing the Effects of Lipid Substitution on Polycation Mediated DNA Aggregation: a Molecular Dynamics Simulations Study"

Chongbo Sun, Tian Tang,* and Hasan Uludağ

S1 Evidence of convergence of the simulation trajectories

Figure S1 shows the charge neutralization curves (cumulative net charge of PEI and ions as a function of the distance from any DNA C1' atom) based on four time windows (last 10 ns, 2nd last 10 ns, 3rd last 10 ns and 4th last 10 ns) at the late stage of the simulations. It can be seen that the curves almost overlap with one another, demonstrating the convergence of the simulation trajectories.

*To whom correspondence should be addressed

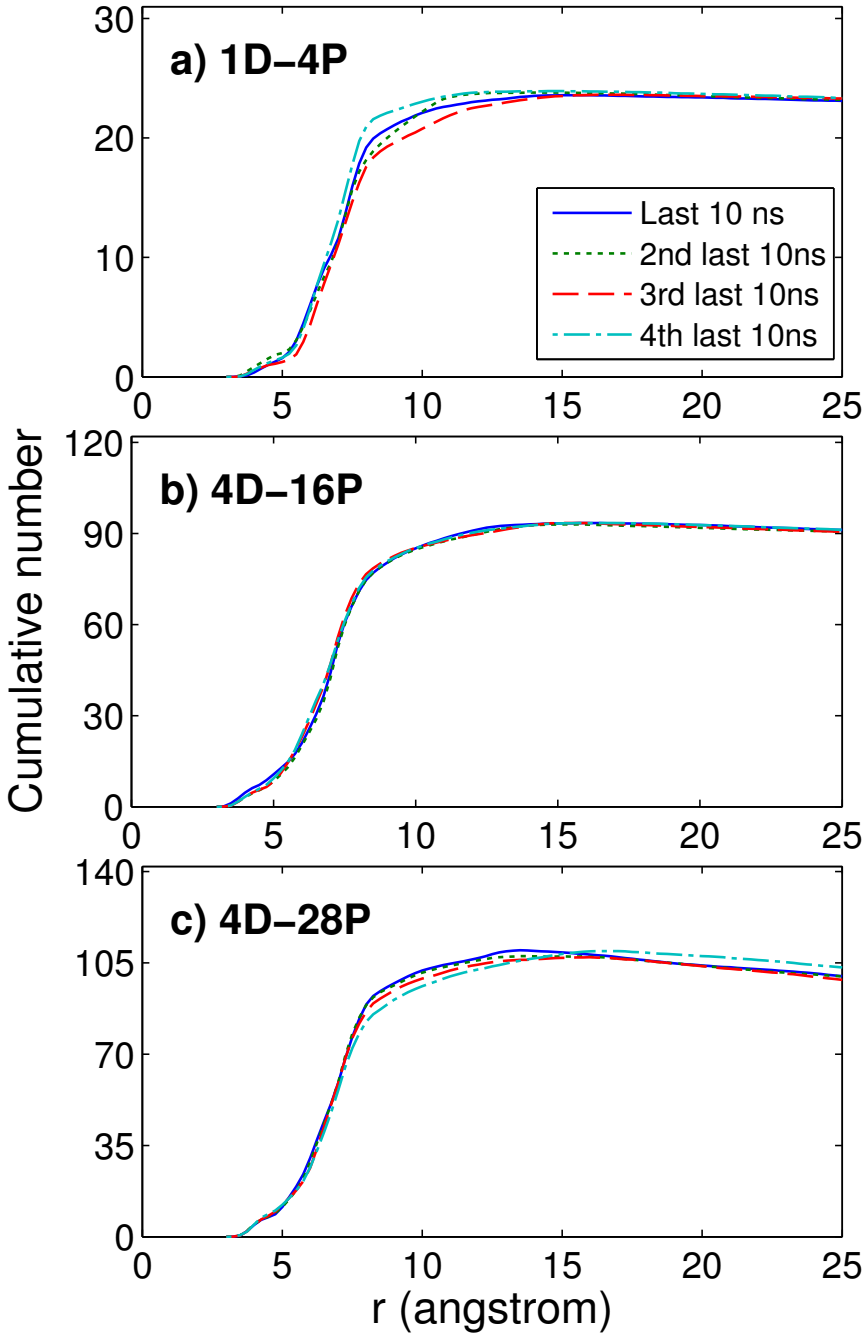

Figure S1: Cumulative numbers of net charge of PEI and ions as a function of the distance from any DNA C1' atom based on four time windows (last 10 ns, 2nd last 10 ns, 3rd last 10 ns and 4th last 10 ns) at the late stage of the simulation. (a) D-4P, (b) 4D-16P, (c) 4D-28P.

S2 Binding of individual lmPEIs to each DNA

A PEI N is said to be 'in close contact with the DNA' if it falls within 4 Å of any N/O atoms of the DNA. We chose 4 Å because this is the distance within which the PEI amine groups can form direct hydrogen bond with the DNA.[S1] A PEI is said to be 'bound' to a DNA molecule if it has one or more Ns in close contact with this DNA. If a PEI is 'bound' to two or more DNA molecules simultaneously, we say that this PEI form a polyion bridge between the DNAs. To quantify the ability of lmPEI to bridge the DNA molecules, we plotted the binding state of individual lmPEIs to each DNA in terms of the number of Ns from each lmPEI in close contact with each DNA, as shown in Figure S2 for the 4D-16P system and in Figure S3 for the 4D-28P system. In Figure S2, each subfigure corresponds to one of the 16 lmPEIs in the 4D-16P system, and it contains 4 curves each of which describes the number of Ns of this lmPEI in close contact with a particular DNA. Similarly, the 28 subfigures in Figure S3 correspond to the 28 lmPEIs in the 4D-28P system and the 4 curves in each subfigure describe the binding state of a lmPEI with the four DNAs. In Figure S2, out of the 16 lmPEIs, five lmPEIs (2, 6, 8, 10, and 13) participate in bridging two or three DNAs for longer than 50% of the simulation time. In Figure S3, there are also five lmPEIs (2, 6, 10, 13 and 18) participating in bridging two or three DNAs for longer than 50% of the simulation time, and lmPEI 10 bridges DNAs A, C and D during most time of the simulation. Six out of the added 12 lmPEIs (17, 18, 20, 22, 26 and 27) bind with at least one DNA for significantly long periods. All the original 16 lmPEIs keep binding to the DNAs and none of them was 'replaced' by the added lmPEIs.

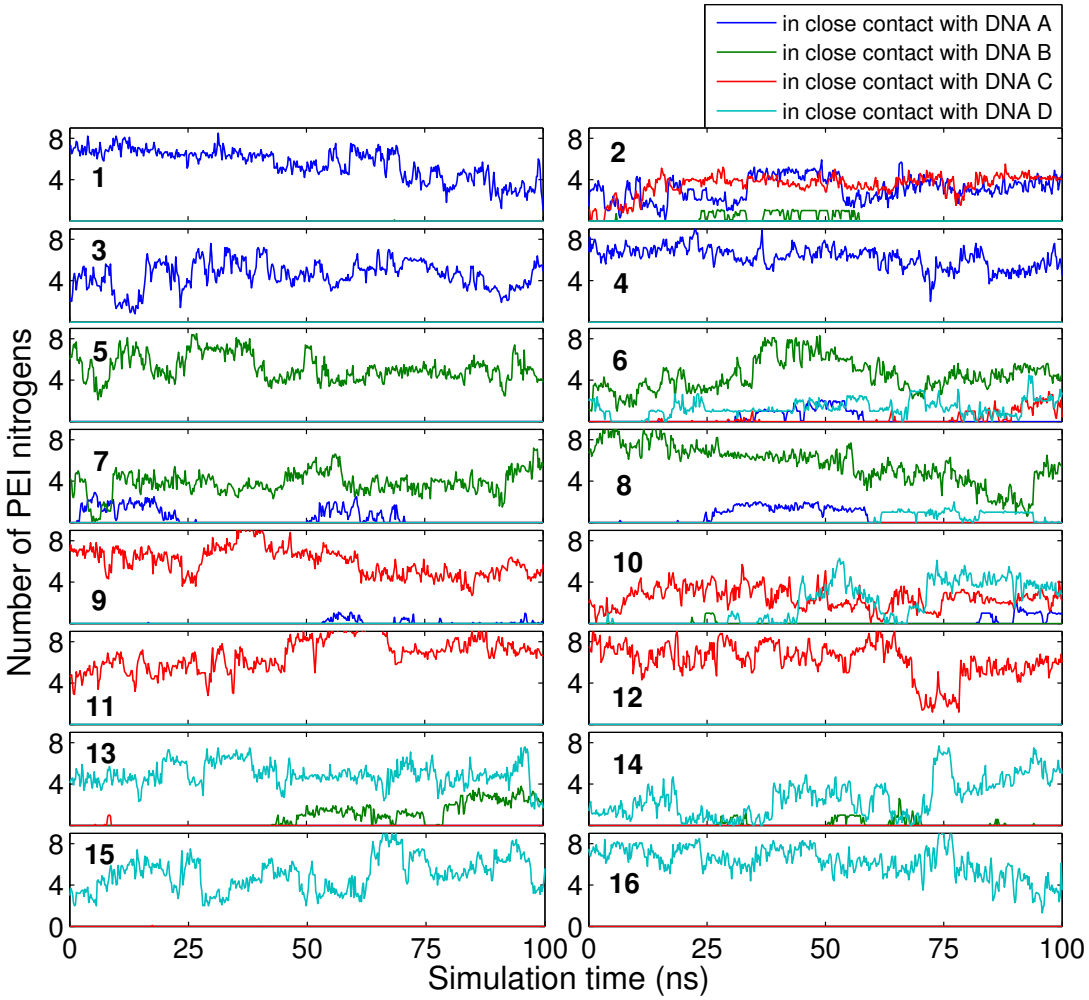

Figure S2: Number of nitrogens for each lmPEI that are within 4 Å of any N/O atom of each DNA as a function of the simulation time for the 4D-16P system.

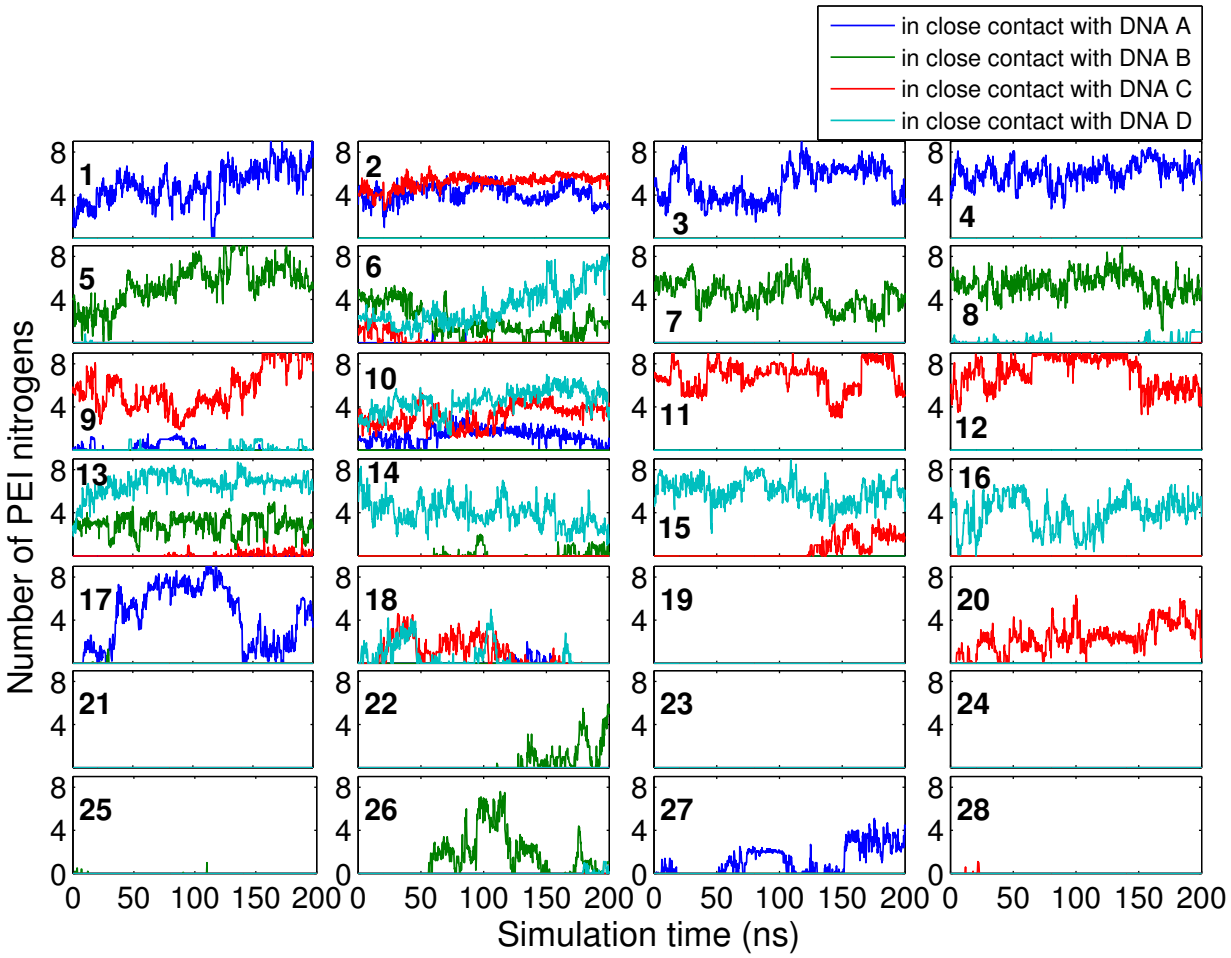

Figure S3: Number of nitrogens for each lmPEI that are within 4 Å of any N/O atom of each DNA as a function of the simulation time for the 4D-28P system.

S3 Calculation of water release

In order to calculate the number of water molecules released during the aggregation process, we counted the number of water molecules within 3 Å of the solutes in each system as summarized in Table S1. To determine the number of water molecules released for a particular system, we subtract the number of water molecules within 3 Å of the polyplex from the total number of water molecules within 3 Å of the individual molecules when they are separated. For example, for system 4D-16P with lmPEIs, the number of water molecules is calculated as '$330.8 \times 4(DNAs) + 99.4 \times 16(lmPEIs) - 1911.1 = 1002.5$'. For system 4D-28P with lmPEIs, as lmPEIs 19, 21, 23, 24 are associated in the solution and not bound to the polyplex, we subtract the number of released water molecules resulting from the association of these four lmPEIs from the number of released water molecules from the whole system.

Table S1: Number of water molecules within 3 Å of the solute at the late stage of each system

System	time window / entire simulation time	# of waters
DNA	last 5 ns / 20 ns	330.8
lmPEI	last 2 ns / 6 ns	99.4
native PEI	last 2 ns / 6 ns	77.8
4D-16P (lmPEI)	last 40 ns / 100 ns	1911.1
4D-16P (native PEI)	last 40 ns / 130 ns	1773.6
4D-28P (lmPEI)	last 40 ns / 200 ns	2700.9
lmPEI 19,21,23,24 in 4D-28P (lmPEI)	last 40 ns / 200 ns	337.0
4D-28P (native PEI)	last 40 ns / 200 ns	2657.5

S4 Radii of gyration of the DNAs in systems 4D-16P and 4D-28P and comparison with their counterparts involving native PEIs

Figure S4 shows the radii of gyration R_g of the four DNAs as a function of simulation time in systems 4D-16P (lmPEI), 4D-16P (native PEIs), 4D-28P (lmPEI) and 4D-28P (native PEIs). It can

be seen that R_g of the four DNAs aggregated by lmPEIs in 4D-16P (average R_g over last 40 ns = 25.7 Å) is nearly identical to that of the four DNAs aggregated by 16 native PEIs (average R_g over last 40 ns = 26.3 Å). So is R_g of the four DNAs in 4D-28P with lmPEIs (average R_g over last 40 ns = 24.6 Å) compared with R_g of the four DNAs aggregated by 28 native PEIs (average R_g over last 40 ns = 24.9 Å).

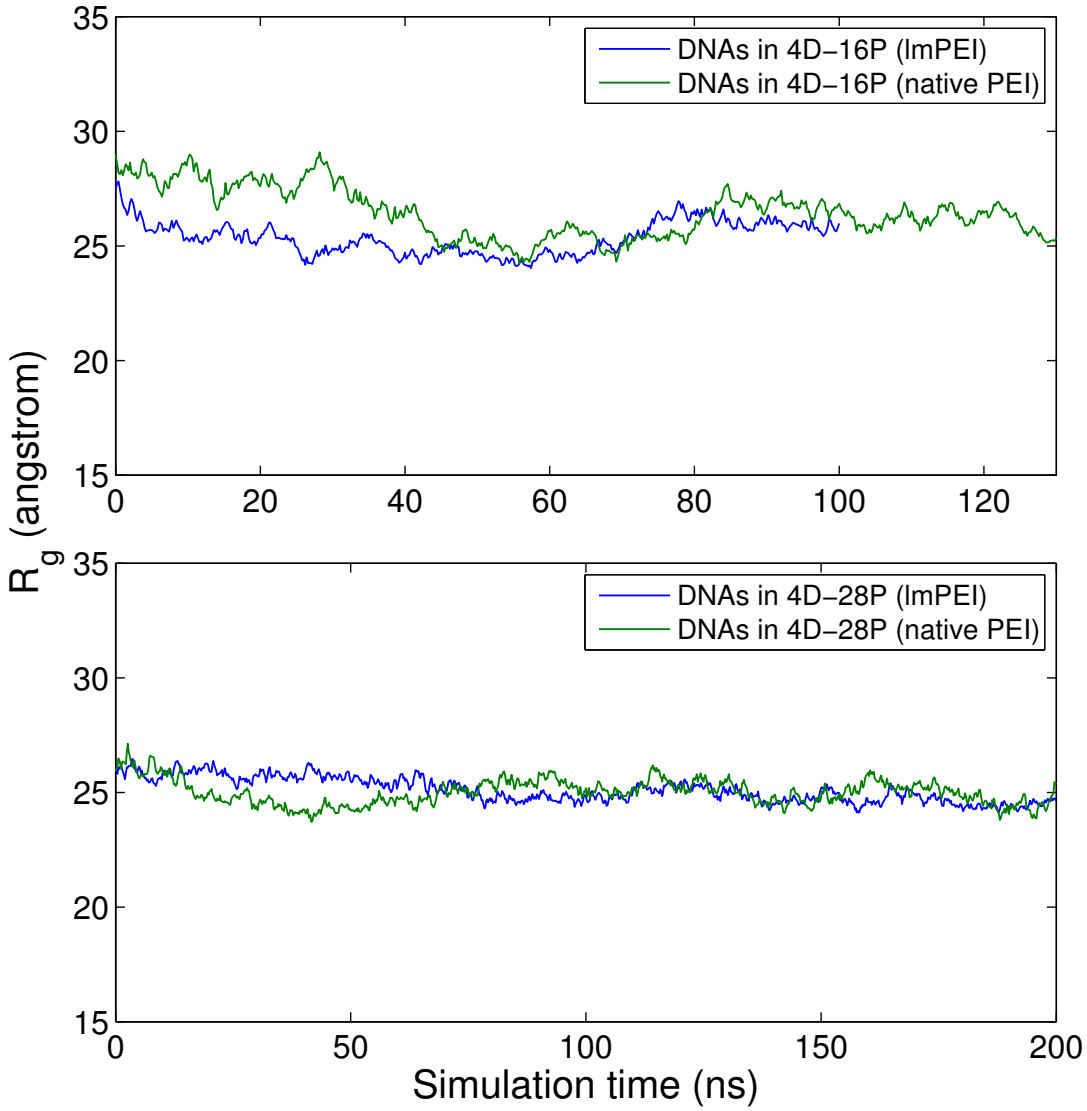

Figure S4: Radius of gyration of the four DNAs in each system as a function of simulation time.

S5 DNA-DNA spacing

DNA-DNA spacing is an important parameter to gauge how compact the DNAs are condensed. To investigate the DNA-DNA spacing in the aggregates, we have defined the 'shortest distance' and 'root mean square (RMS) distance' between two DNA molecules in Ref. S2. We first represent each DNA as a series of points each being the COM of a Watson-Crick DNA base pair. For each dodecamer studied in this work, there are 12 such points, and connecting neighboring points results in 11 segments. For a pair of segments from two different DNAs, we can calculate their shortest distance, and there are 121 such distances ($d_i, i = 1 \ldots 121$) between all pairs of segments from the two DNA dodecamers. We defined the minimum of these 121 distances as the 'shortest distance' d_{shortest} and the root mean square of these 121 distances as the 'RMS distance' d_{RMS} ($d_{\text{RMS}} = \sqrt{(\Sigma_{i=1}^{121} d_i^2)/121}$). Figures S5 and S6 plot d_{shortest} and d_{RMS} for systems 4D-16P and 4D-28P, respectively. The average values of d_{shortest} and d_{RMS} for each pair of DNAs over the last 40 ns of the simulations are summarized in Table S2. Among the 6 pairs, the aggregation of three of them (A-C, B-D, C-D) involves direct bridging by the same lmPEIs, while the other three (A-B, A-D, B-C) are aggregated only through the lipid association among different lmPEIs. Overall, the directly bridged DNA pairs have smaller d_{shortest} and d_{RMS}. In fact, for the (A-C, B-D, C-D) pairs, the average values are 18.5 Å for d_{shortest} and 30.1 Å for d_{RMS}, which are not significantly different from the respective values of 21.4 Å and 29.0 Å for native PEI mediated DNA aggregation. Due to steric effect, DNA pairs brought together by lipid association (A-B, A-D, B-C) show much larger d_{shortest} and d_{RMS} (except for A-B in 4D-16P system which has slightly smaller d_{RMS} than that of C-D).

Table S2: DNA-DNA distance (Å) averaged over the last 40 ns of the simulations.

System	4D-16P						4D-28P					
DNA-DNA	A-B	A-C	A-D	B-C	B-D	C-D	A-B	A-C	A-D	B-C	B-D	C-D
d_{shortest}	29.0	18.2	36.5	26.0	20.3	20.8	36.1	13.6	27.2	32.8	18.5	19.7
d_{RMS}	36.3	30.7	44.7	43.3	28.4	37.0	41.5	31.2	41.0	41.1	27.1	26.0

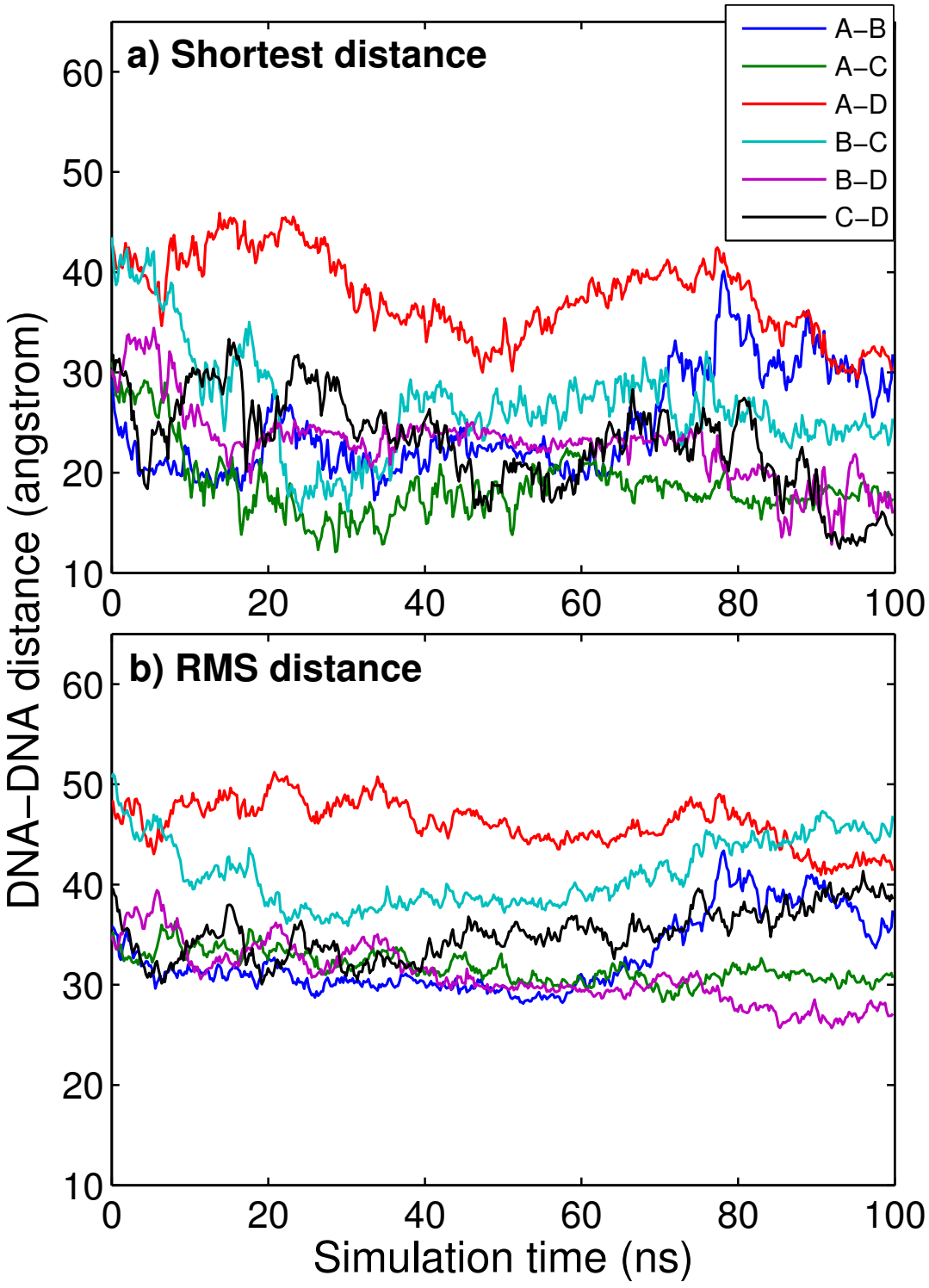

Figure S5: Distance between DNAs for the 4D-16P system. (a) Shortest distance. (b) Root mean square distance.

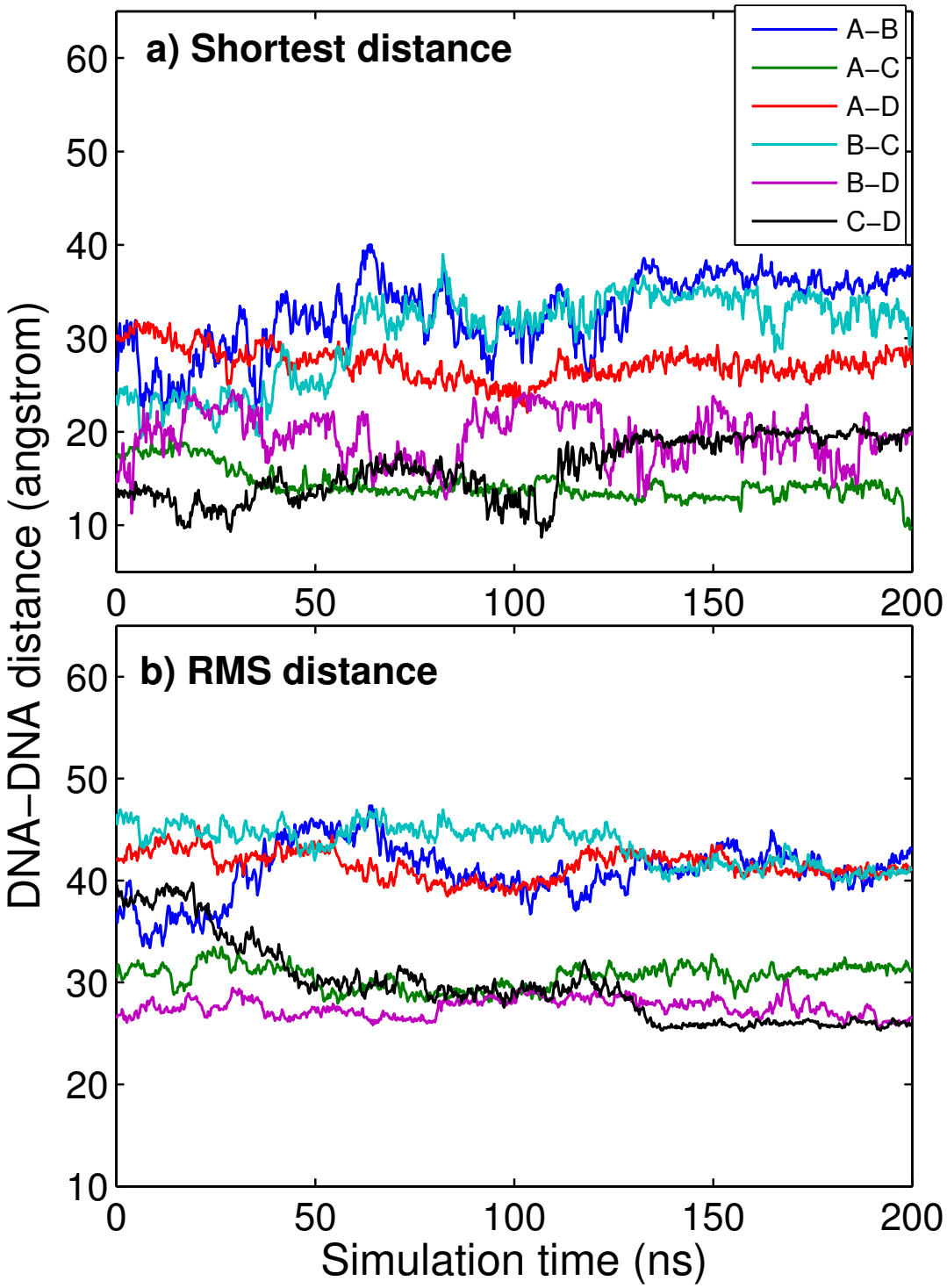

Figure S6: Distance between DNAs for the 4D-28P system. (a) Shortest distance. (b) Root mean square distance.

References

(S1) Sun, C.; Tang, T.; Uludag, H.; Cuervo, J. E. *Biophys. J.* **2011**, *100*, 2754–2763.

(S2) Sun, C.; Tang, T.; Uludag, H. *Biomacromolecules* **2011**, *12*, 3698–3707.