Supporting Information for
Dynamics of terminal fraying-peeling and hydrogen bonds dictate the sequential vs co-operative melting pathways of nanoscale DNA and PNA triplexes

Sandip Mandal[a], Krishna N. Ganesh[b] and Prabal K. Maiti[a,1]
[a]*Center for Condensed Matter Theory, Department of Physics, Indian Institue of Science, Bangalore 560012, India*
[b]*Jawaharlal nehru Center for advanced Scientific Research (JNCASR), Jakkur, Bengaluru 560064, India*

1 Sequence dependent triplex melting with 18 T-A-T base triples

To investigate the influence of alternative sequences on the melting mechanism and TFO binding affinity, we substituted all of the protonated cytosine residues in the TFO third strand with Thymine (T) bases. As a result, all 18 base triples consist of T-A-T combinations, shown in Table 1. Because the base triples lack cytosine, the absence of protonation diminishes the electrostatic attraction and repulsion dominated within the negatively charged DNA-TFO and DNA:DNA duplexes. However, because TAT/TAT has stronger stacking interactions than TAT/CGC$^+$, all TAT triples triplexes are more stable.

Table 1: Details of the 18 TAT base triple systems simulated in this work

System	Sequence (18bp)	Buffer
DNA-DNA-DNA	5′-TTTTTTTTTTTTTTTTTT-3′ 3′-AAAAAAAAAAAAAAAAAA-5′ 5′-TTTTTTTTTTTTTTTTTT-3′	20 Å
PNA-DNA-PNA	Nter-TTTTTTTTTTTTTTTTTT-Cter 5′-AAAAAAAAAAAAAAAAAA-3′ Cter-TTTTTTTTTTTTTTTTTT-Nter	20 Å

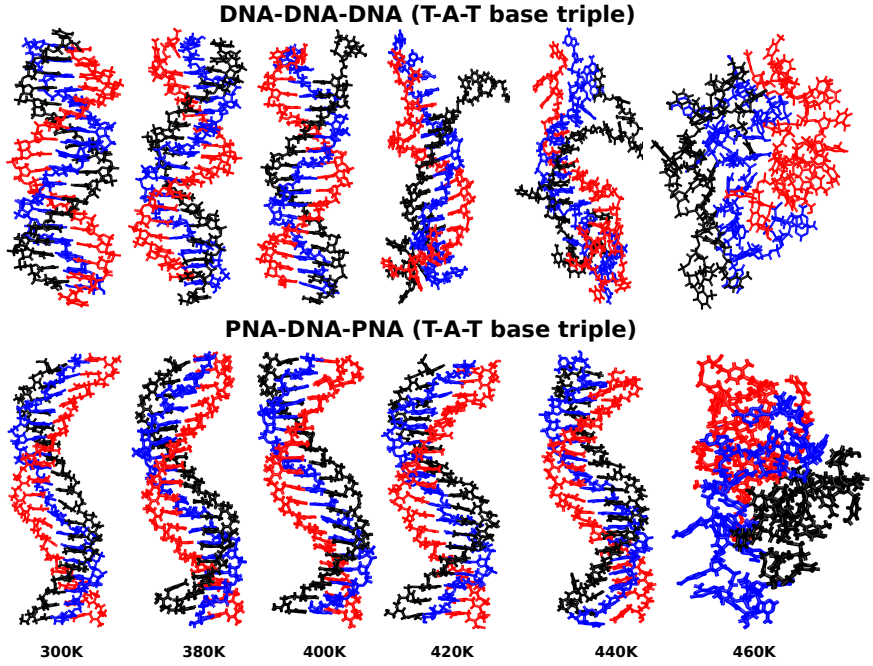

Figure S1: Snapshots of final structures of DNA-DNA-DNA and PNA-DNA-PNA triplexes from 200 ns MD simulation represented by upper and lower panel, respectively, at ambient and elevated temperature with all T-A-T base triples.

[1]Corresponding author.
E-mail address: maiti@iisc.ac.in

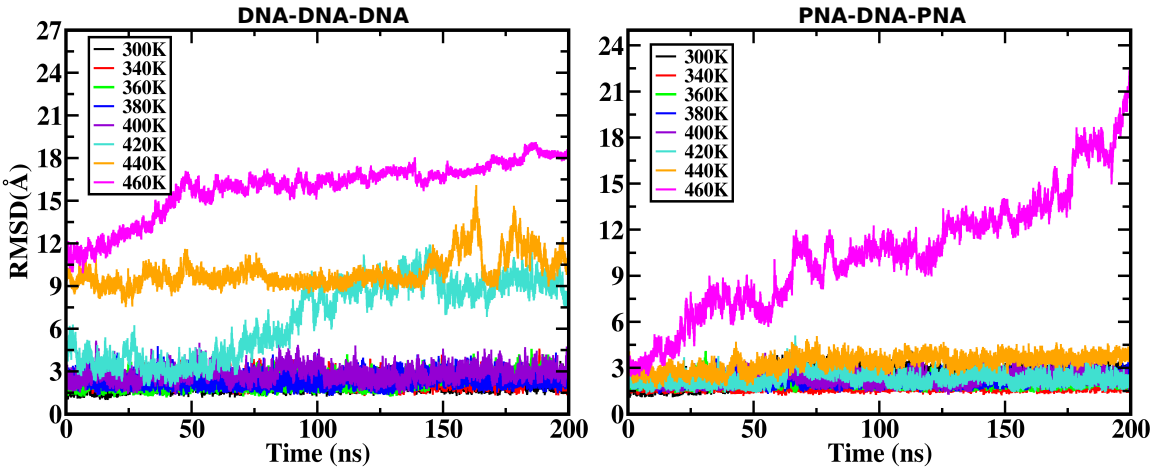

Figure S2: The time evolution of the Root-Mean-Square-Deviation(RMSD) from their initial energy-minimized structures at ambient and elevated temperatures are represented by left and right figures, respectively DNA-DNA-DNA and PNA-DNA-PNA triplex with 18 T-A-T base triples.

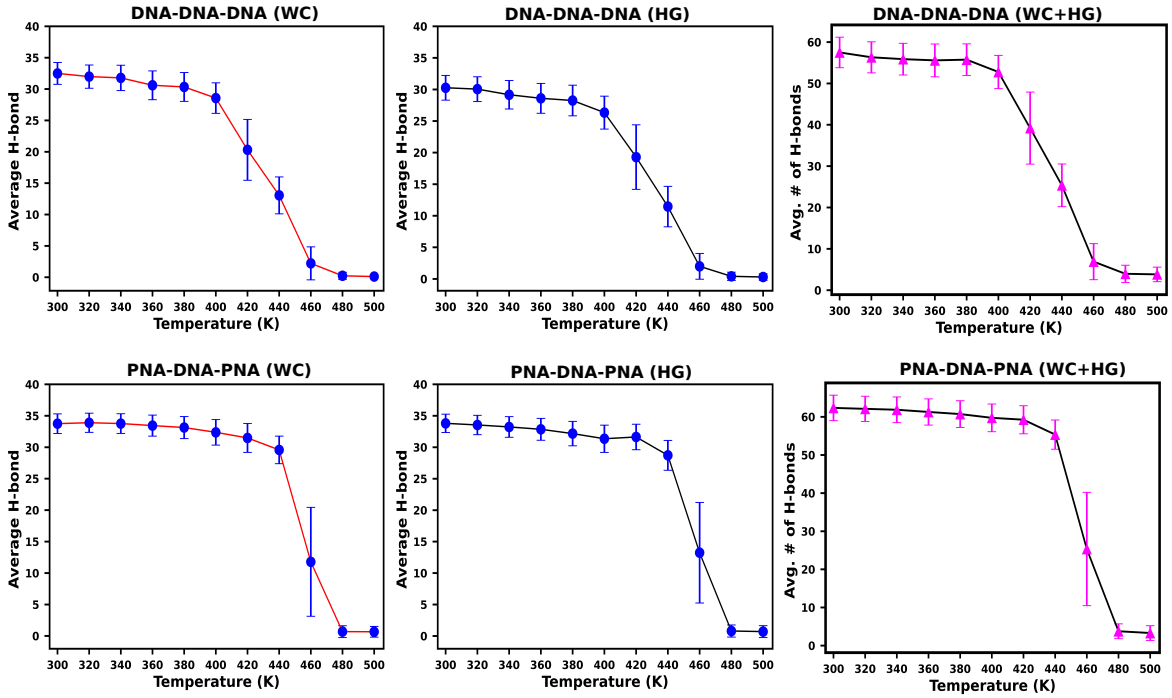

Figure S3: Time average of the total hydrogen bonds(H-bonds) at ambient and elevated temperatures for the DNA-DNA-DNA and PNA-DNA-PNA triplex structures with all T-A-T base triples, showing sequential(two-step) vs cooperative(one-step) melting are shown in upper and lower panel rightmost figures respectively.

Table 2: Melting Temperature T_m of the TFO strands from the fitting of two-state melting transition function for the DNA-DNA-DNA and PNA-DNA-PNA triplexes. The detailed description of the fitting function is in the 3.2 section in the main manuscript, using equation-5)

Melting Temp. [K]	DNA-DNA-DNA	PNA-DNA-PNA
T_m [f (T_m) = 0.5] [K] [Mix sequence]	396	420
T_m [f (T_m) = 0.5] [K] [TAT base Sequence]	415	453

Table 3: Binding free energy of the protonated DNA-TFO and PNA-TFO of the DNA-DNA-DNA and PNA-DNA-PNA triplexes calculated using last 100 ns of the 200 ns long MD trajectory at ambient and elevated temperature in (Kcal/mol) units.

Temperature(K)	DNA-DNA-DNA	PNA-DNA-PNA
300	63.91 (6.76)	-96.78 (12.30)
320	76.21 (7.66)	-90.19 (12.60)
340	85.86 (7.94)	-83.48 (11.65)
360	90.38 (8.47)	-77.43 (12.11)
380	103.77 (9.07)	-67.86 (12.21)
400	122.85 (11.62)	-60.09 (11.70)
420	291.64 (56.08)	-62.64 (12.37)
440	348.94 (29.79)	-45.65 (24.85)
460	680.60 (36.06)	99.54 (47.86)

2 Final structures of the protonated triplexes from independent runs

Three sets of statistically independent runs are carried out with different random seed values to ensure our results are reliable and reproducible. Here, we reported the final snapshots of two other independent runs. Instantaneous snapshots of the final structures from one of the runs are reported in the main manuscript.

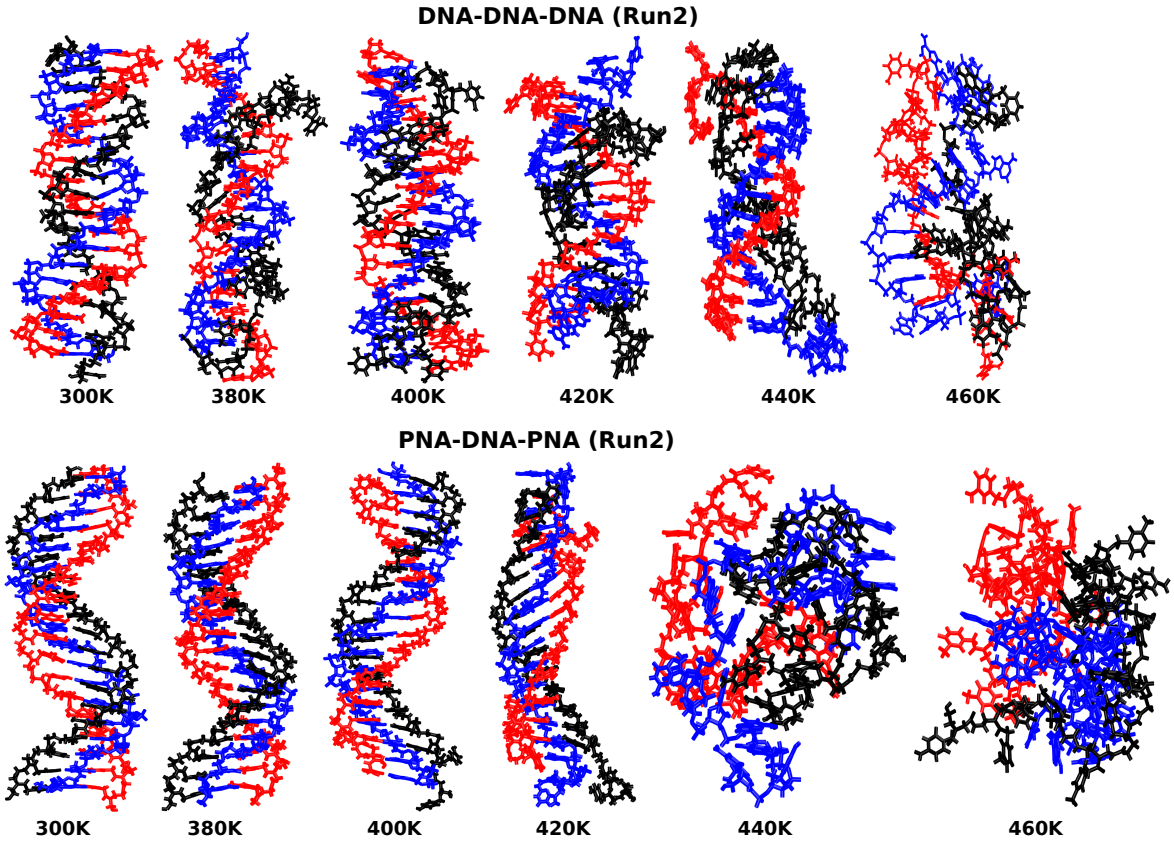

Figure S4: Instantaneous snapshots of final structures of the protonated DNA-DNA-DNA and PNA-DNA-PNA triplexes from 200 ns MD simulation represented by upper and lower panel, respectively, for one of the three independent runs, at ambient and elevated temperature. The HG TFO strands are shown in black, and the WC duplex strands are shown in red and blue, respectively

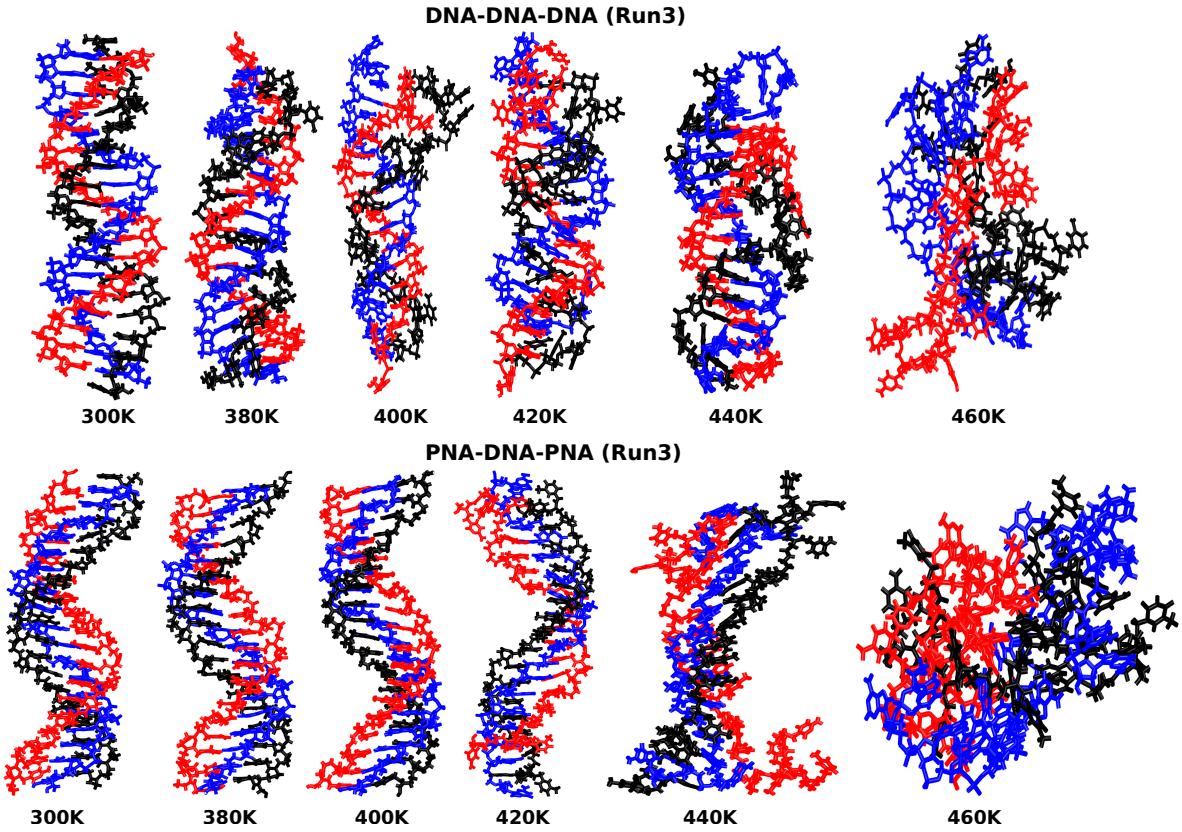

Figure S5: Instantaneous snapshots of final structures of the protonated DNA-DNA-DNA and PNA-DNA-PNA triplexes from 200 ns MD simulation represented by upper and lower panel, respectively, for one of the three independent runs, at ambient and elevated temperature. The HG TFO strands are shown in black, and the WC duplex strands are shown in red and blue, respectively

Table 4: Average RMSD from the three independent runs for PNA-DNA-PNA triplexes

Temp [K]	Run1	Run2	Run3
300	3.20 ± 0.60	2.93 ± 0.23	2.86 ± 0.47
320	3.25 ± 0.50	2.53 ± 0.19	3.07 ± 0.54
340	3.35 ± 0.33	2.41 ± 0.24	3.39 ± 0.34
360	3.03 ± 0.42	2.88 ± 0.30	3.65 ± 0.36
380	4.34 ± 0.70	2.50 ± 0.28	3.58 ± 0.46
400	4.28 ± 0.42	4.22 ± 0.85	3.92 ± 0.46
420	8.99 ± 0.65	7.47 ± 0.85	6.34 ± 0.84

3 Final structures of the non-protonated triplexes and Watson-Crick duplexes

Due to unfavorable electrostatic repulsion between the negatively charged Watson-Crick (WC) duplex's major grooves and DNA-TFO strands, the non-protonated DNA-TFO strands, as depicted in Fig.S6A, are unable to form a stable DNA-DNA-DNA triplex. Consequently, as the temperature rises, the weak Hoogsteen (HG) hydrogen bond breaks easily and dissociates from the major grooves of the WC duplex, followed by the sequential melting of the WC base pairs in the duplexes(S7A). Nevertheless, as Fig.S6B illustrates, the non-protonated PNA-DNA-PNA triplex exhibits a single-step cooperative melting at higher temperatures because of the strong binding affinity between the charge-neutral PNA-TFO strand and the PNA:DNA duplex.

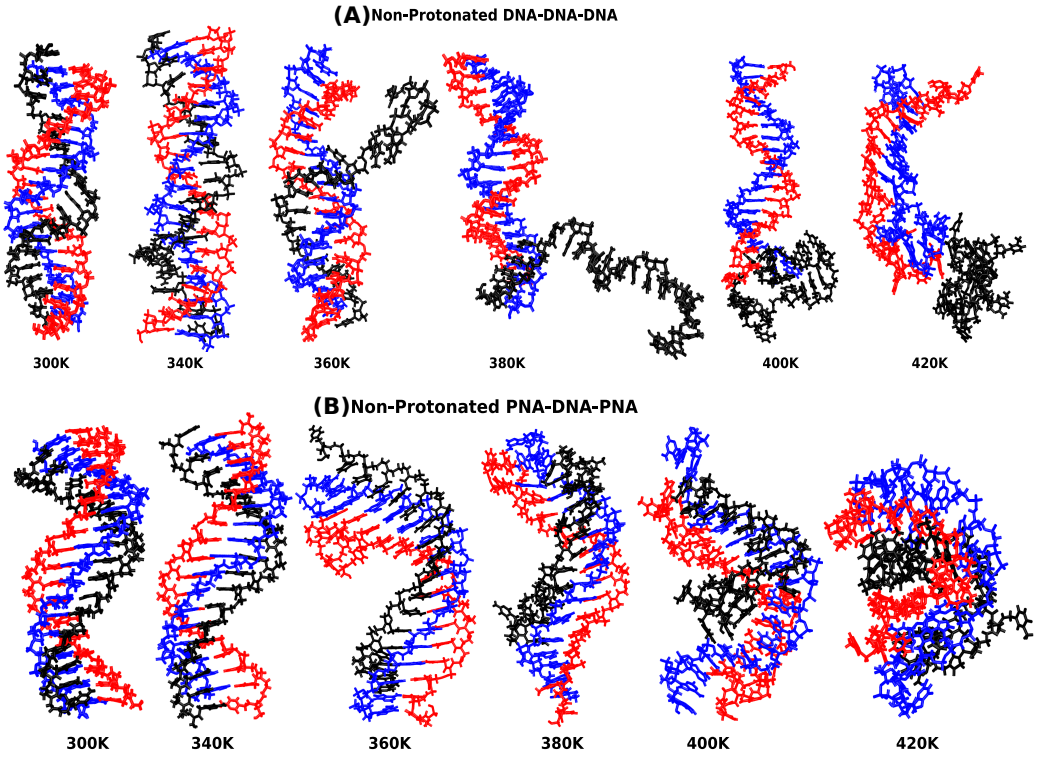

Figure S6: Snapshots of final structures of Non-Protonated DNA-DNA-DNA and PNA-DNA-PNA triplexes from 200 ns MD simulation represented by (A) and (B), respectively, at ambient and elevated temperature. The HG TFO strands are shown in black, and the WC duplex strands are shown in red and blue, respectively

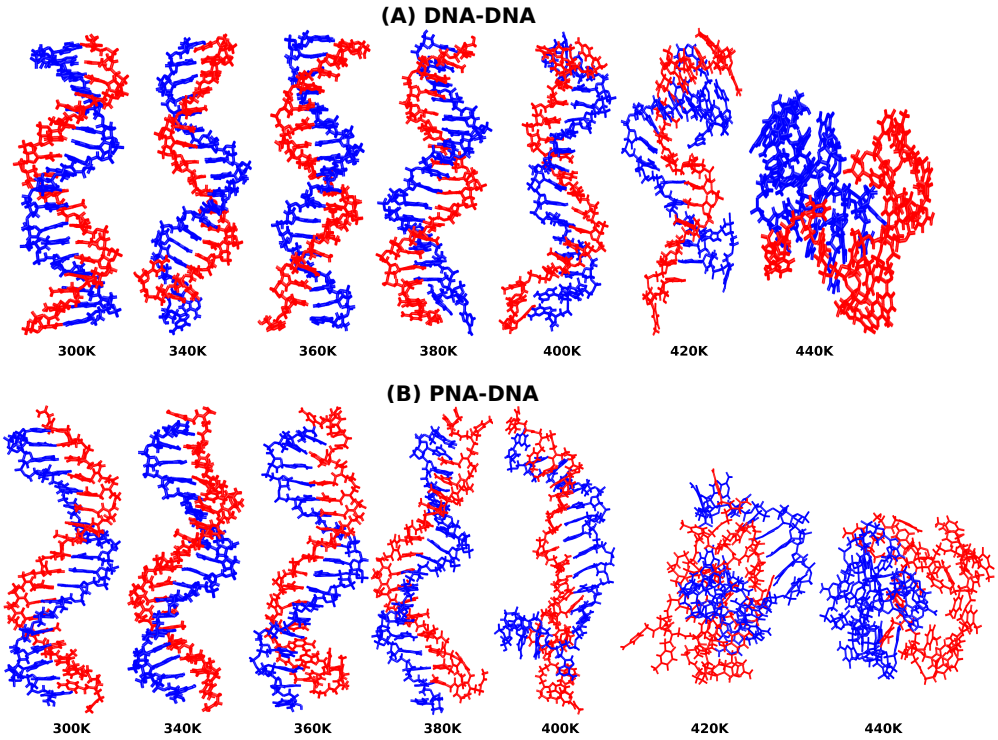

Figure S7: Snapshots of the final frames of the Watson-Crick DNA:DNA and PNA:DNA duplex structures after a 200 ns MD simulation are displayed by (A) and (B), respectively

4 Principal Component Ananlysis

4.1 Eigenvalue Fraction and Cumulative Sum of Eigenvalues of the Principal Components

The First few eigenvectors and eigenvalues, critical for determining the overall mobility of the triplex structures during the temperature-induced melting transitions, are calculated using Principal component Analysis(PCA) analysis from the atomic coordinates of the simulation trajectory.

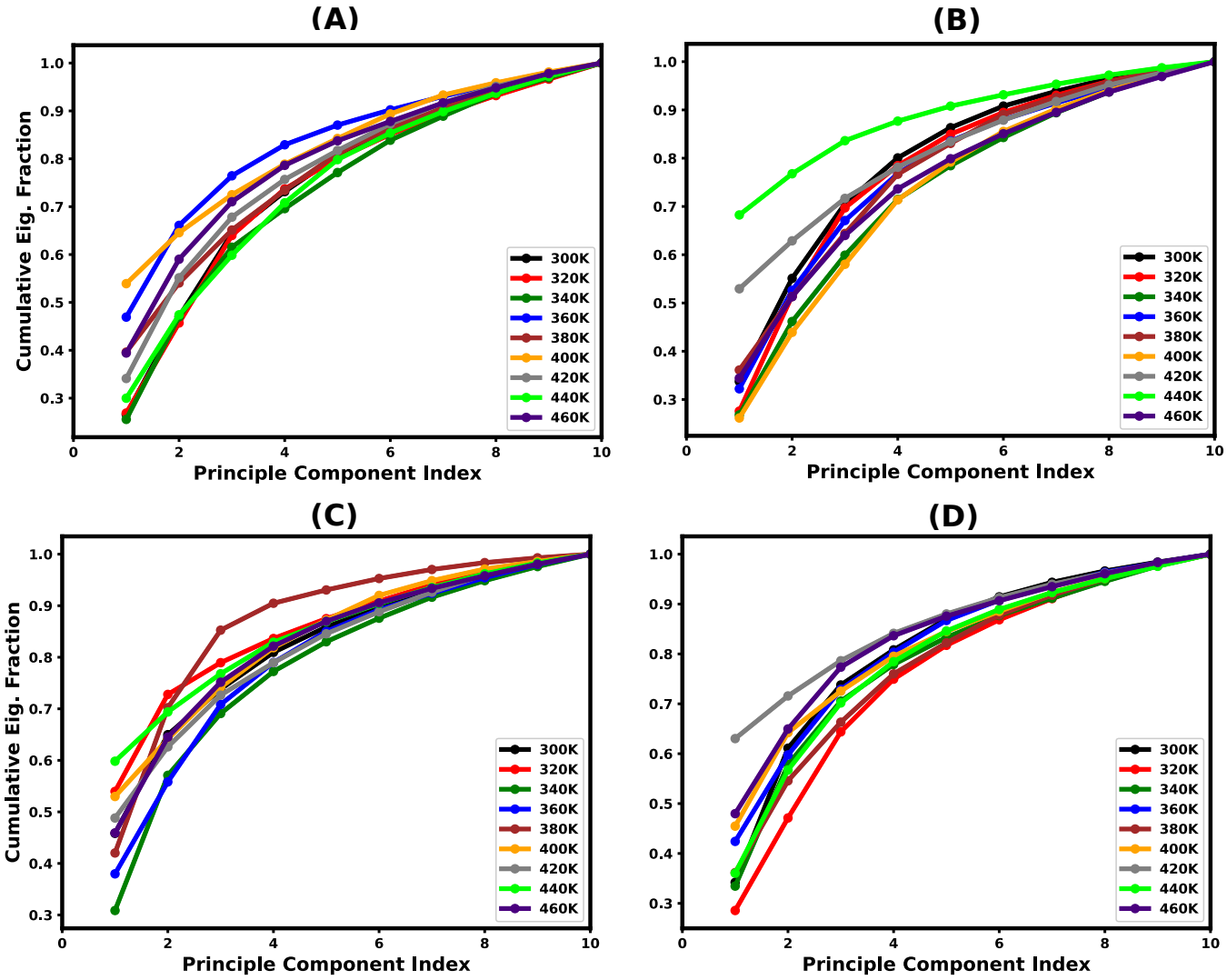

Figure S8: Cumulative sum of the first ten principal components from DNA and PNA triplexes, respectively, at ambient and elevated temperatures. (A) and (B) represents protonated DNA and PNA triplexes. (C) and (D) display the non-protonated DNA and PNA triplexes.

In Fig.S8(A-D) and Fig.S9(A-D), we showed cumulative sums for the first ten principal component indices and their eigenvalue fractions. The cumulative sum of eigenvalues is calculated by adding the eigenvalues in decreasing order. The PCs represent the most considerable conformational alterations in the structure with the highest cumulative eigenvalues. Figure S8 illustrates that, for DNA-involved triplexes (A and C), the cumulative-eigen-fraction occupied by the first two principal components likewise increases sequentially as temperatures rise. However, the sudden increase in the cumulative-eigen-fraction is only near melting temperature (B and D) for PNA-involved triplex structures, which reveals a co-operative melting signature.

The eigenvalue fractions of protonated DNA and PNA triplexes are shown in Fig.S9. For DNA triplexes (A and C), these proportions increase gradually as the temperature rises, indicating a sequential versus cooperative melting nature; however, for PNA triplexes (B and D), a sudden increment around 420K indicates a cooperative melting nature.

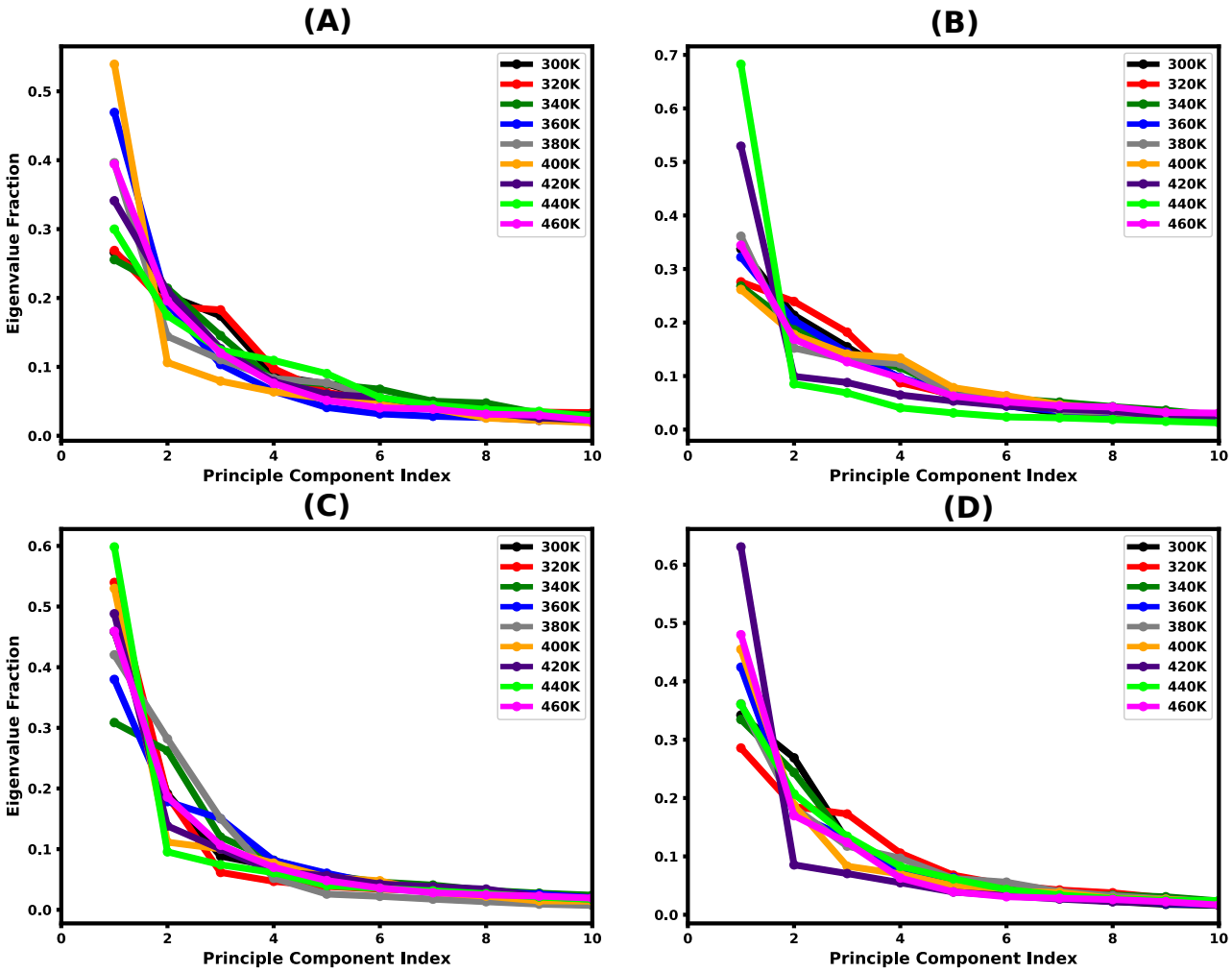

Figure S9: Eigenvalue fraction of the first ten principal modes of DNA and PNA triplexes respectively at ambient and elevated temperatures is represented by (A) protonated DNA-DNA-DNA (B) protonated PNA-DNA-PNA triplex, (C) non-protonated DNA-DNA-DNA and (D) non-protonated PNA-DNA-PNA triplex.

4.2 Principal Component Projection

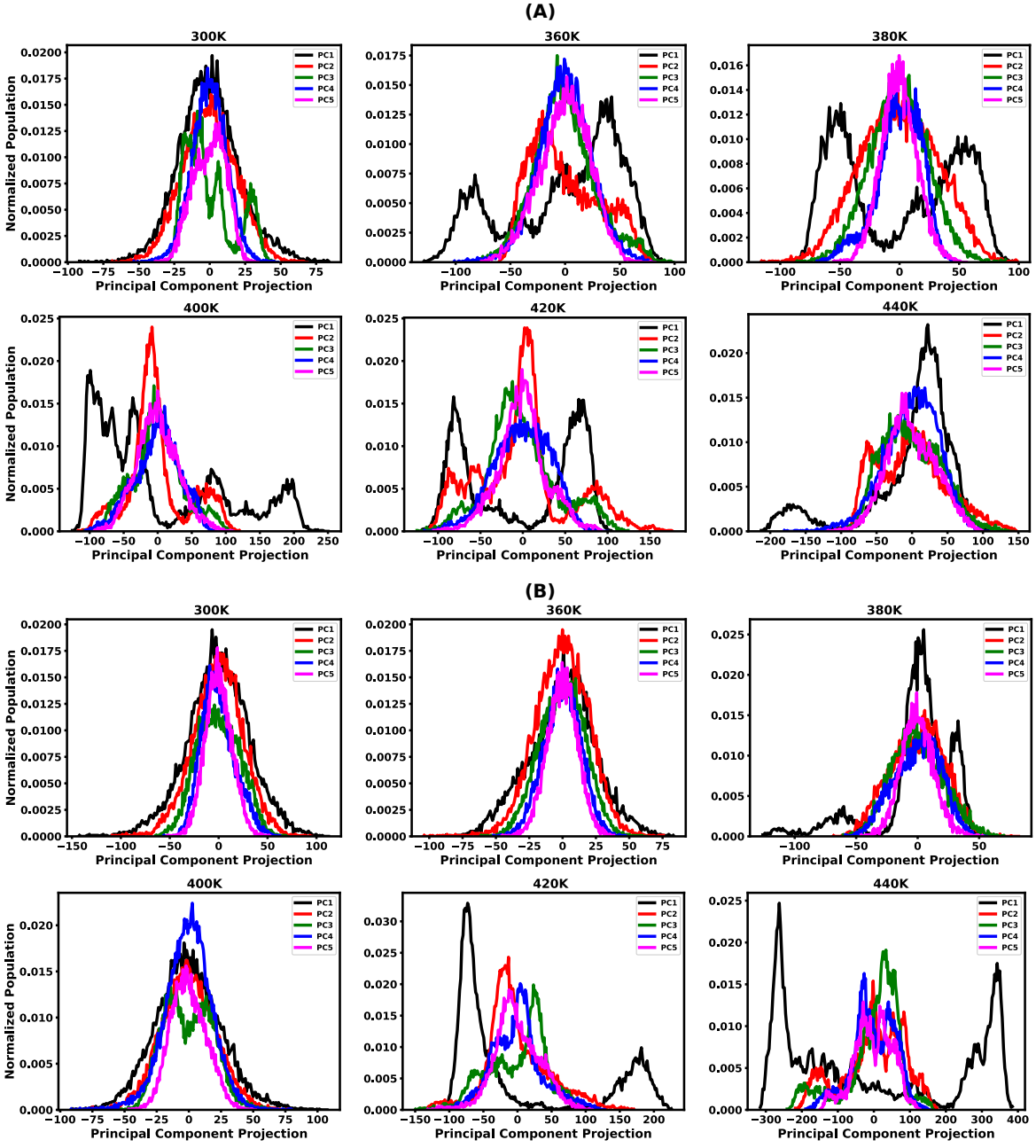

Figure S10: Probability distribution of the first five principal component projections for the protonated DNA and PNA triplexes at ambient and elevated temperatures. The normalized projection population along the first principal component (PC1, displayed in black colour) shows a significant change along the principal component projection (x-axis) with a wider span and bimodal probability distribution sequentially from 360 K onwards for the DNA triplex, indicating sequential melting as shown in (A), whereas for protonated PNA triplex the probability distribution is unimodal (intact structure) with less spreading along the projection axis up to 400 K but sudden increase in the projection span with bimodal distribution near melting temperature (∼420 K) indicates cooperative melting transitions as shown in (B).