

Biophysical Journal

Supporting Material

Phi29 Connector-DNA Interactions Govern DNA Crunching and Rotation, Supporting the Check-Valve Model

Rajendra Kumar¹ and Helmut Grubmüller^{1,*}

¹Department of Theoretical and Computational Biophysics, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

Contents

1. SECTIONS	3
1.1. Equilibration of the Phi 29 Connector-DNA complex.....	3
1.2. Impact of the DNA on the connector	4
2. TABLES	7
Table S1.....	7
Table S2.....	7
Table S3.....	8
3. FIGURES.....	9
Figure S1	9
Figure S2	9
Figure S3	10
Figure S4	11
Figure S5	12
Figure S6	12
Figure S7	13
Figure S8	14
Figure S9	15
Figure S10	16
Figure S11	17
4. REFERENCES	18

1. SECTIONS

1.1. Equilibration of the Phi 29 Connector-DNA complex

Structural and energetic equilibration of the Phi 29 connector-DNA complex was monitored by quantifying RMSDs and interaction energies as a function of time (linear and logarithmic scale) derived from three ~370 ns MD simulations (SimA, SimB, and Sim C) as shown in Fig. S3.

At first, RMSDs for different regions of the connector and the DNA were calculated using backbone atoms with reference to the starting structure. As seen in Fig. S3A–B, the RMSD of the whole connector rapidly increased to a maximum value of 0.24 nm during the initial 50 ns simulation time and subsequently remained at less than 0.27 nm during the three MD simulations. Similarly, the RMSD of the channel comprising middle, hinge and bottom region rapidly increased to a maximum value of 0.18 nm during the first 50 ns simulation time and subsequently remained at less than 0.2 nm (Fig. S3C–D). The RMSD of the loop-DNA region fluctuated around values of 0.3, 0.38, and 0.42 nm after a gradual increase during the initial ~120 ns simulation time (Fig. S3E–F). The RMSD of the DNA fluctuated within a range of 1.0 to 2.0 nm after a sharp increase during the initial 10 ns simulation time (Fig. S3G–H).

These results show that the whole connector and its channel region sharply deviate from the starting structure during the first 50 ns simulation time. Subsequently, the deviation rates markedly slowed down. During the following 320 ns of the simulations, the obtained maximum RMSD values of 0.27 and 0.2 nm for the whole connector and the channel structures, respectively, suggest a small structural deviation for this large structure. The loop-DNA region changed (0.3 to 0.42 nm) from its earlier refined conformation during the three simulations. However, its conformation is not crucial to test the DNA packaging models because of the expected extreme flexibility of the loops.

To monitor the energetic equilibration of the Phi29 connector-DNA complex, electrostatic (Fig. S3I–J) and van der Waals (Fig. S3K–L) interaction energies were calculated separately for the connector and the DNA as a function of time from the three above mentioned simulations. After rapid decrease in energies during the initial ~50 ns simulation time, the electrostatic and van der Waals energies either slowly decreased or fluctuated within ranges of -6000 to -4500 and -1100 to -700 kJ mol⁻¹, respectively. The obtained energy values

suggest that the DNA interacts with the connector primarily through attractive electrostatic forces.

Overall, the obtained RMSDs for both the whole connector and the middle region were small with respect to the molecule size (~3200 residues). Additionally, the interactions between DNA and connector were strengthened during the respective simulations. The simulations to study the DNA packaging models were considered to be equilibrated sufficiently because of the observed very small structural changes and increased interactions between the connector and the DNA during the last 320 ns. Therefore, all subsequent analyses were done using the last 320 ns MD trajectories; the initial 50 ns were discarded.

1.2. Impact of the DNA on the connector

The impact of the DNA on the connector structure during the packaging process is so far unclear. To study potential structural changes of the connector in presence of DNA, we calculated two structural descriptors, the twist angle and the length of the whole connector (shown in Fig. S4), quantifying the untwisting-twisting and stretching-compression motions, respectively. Additionally, previous studies suggested the middle region to be one of the stiffest protein regions (1). To study the impact of the DNA on this region, these descriptors were calculated for this particular region as well.

For the whole connector, fluctuations in its twist angle were calculated as a function of its length (Fig. S5A). As seen in Fig. S5A, a two-peak distribution was obtained instead of a single-peak distribution, which was previously seen for the connector in absence of the DNA (1). This diversity in fluctuation was only observed for the twist angle. For a detailed analysis, the change in twist angle as a function of time was plotted and is shown in Fig. S6. The twist angle deviates by 2.3° from the starting crystal structure value of 74.2° during the first 25 ns (Fig. S6) simulation time as previously observed in simulations of the connector in absence of the DNA (1). However, after 180 ns of the SimA simulation time, the connector twist angle recovered to $\sim 74.5^\circ$, and remained within the range of 74° to 75° during the course of the subsequent simulation (Fig. S6). The observed diverse fluctuation was caused by the recovery of the twist angle during the SimA simulation. This result suggests that the connector untwists and recovers its starting twist angle value in the presence of the DNA; however, it is unclear whether conformations also recover with the twist angle, which will be further discussed in the following paragraphs.

As seen in Fig. S5A, the obtained fluctuations of the whole connector length are in line with the previously observed behavior of the connector without DNA (1). This result suggests that the DNA does not affect the length of the whole connector (Fig. S5A). Next, fluctuations in twist angle and length of the middle region were calculated; the obtained distribution is presented in Fig. S5B. The middle region's angle and length are comparable in presence and absence of the DNA (1). Therefore, we conclude this region was unaffected by the DNA during the simulations.

The above presented results show that the whole connector twist angle regains its starting value during the course of the SimA simulation. However, monitoring twist angle and length are not sufficient to capture all conformational fluctuations of the connector. Consequently, the observed change in twist angle may not correspond to a full recovery of the crystal structure conformation during the simulations. To probe and characterize the conformational fluctuations during the simulations, principal component analyses (PCA) were performed on the middle and the bottom connector region with reference to the two most rigid helices of the middle region (Fig. 1). For the middle region, projections of the first two principal components (PC), which capture the largest conformational fluctuations, were calculated for the free and the bound connector. These fluctuations are compared in Fig. S7A.

As shown in Fig. S7A, the middle region conformations were similar in presence and absence of the DNA during simulation SimB and differ along PC-1 and PC-2 during simulations SimA and SimC, respectively. Despite the difference in the conformational distribution during simulations SimB and SimC, this distribution partially overlapped with the conformations of the connector in absence of the DNA. To compare these deviations, three average structures were obtained from the filtered trajectories containing conformational fluctuations along only the first and the second PCs (Fig. S7B). These average structures appeared well aligned to each other upon visual inspection. RMSDs between the three structures were calculated pair-wise to quantify this alignment (Table S2). The obtained values are deviated less than 0.8 Å implying that the observed conformational differences are rather minuscule (Table S2).

Similarly, for the bottom region, projections of the first two PCs were calculated both for the free and the bound connector and were compared to each other (Fig. S7C). As seen in Fig. S7C, the bottom region conformations are similar to each other in presence and absence of the DNA with the exception of PC-1 during simulation SimA. To compare structural

deviations, three average structures were obtained from the three trajectories that were filtered for PC-1 and PC-2 (Fig. S7D). These structures appeared to be aligned to each other except for the six subunits during simulation SimA. Further, RMSDs were calculated pairwise for these structures (Table S2). A very small RMSD of 0.3 Å was obtained between the average structure of simulations SimB and SimC. In contrast, the RMSD of SimA with respect to SimB and SimC was comparatively large (2.2 and 2.1 Å, respectively).

These results show that the conformation of the bottom region deviated from its crystal structure despite the above observed recovery of the twist angle during simulation SimA (Fig. S7). The angle was recovered by the slight shift in the subunit centers of mass used to calculate the angle.

In summary, the observed structural deviation of the middle connector region is small in presence of the DNA (Table S2) compared with the DNA free connector conformation. In contrast, the conformation of the bottom region changes in DNA presence significantly.

2. TABLES

Table S1: Heating-cooling cycles performed in SAMD simulations. The connector loops were heated up and cooled down consecutively 40 times during 10 ns of MD simulations. Each cycle consisted of 252 ps with two and eight steps of heating and cooling, respectively.

Temperature (K)	Time period (ps)	Cumulative time (ps)	Temperature (K)	Time period (ps)	Cumulative time (ps)
300	20	20	600	20	162
300-600	2	22	600-500	2	164
600	22	44	500	20	184
600-1000	2	46	500-450	2	186
1000	50	96	450	20	206
1000-800	2	98	450-400	2	208
800	20	118	400	20	228
800-700	2	120	400-350	2	230
700	20	140	350	20	250
700-600	2	142	350-300	2	252

Table S2: RMSDs in Å between the average structures of the middle and the bottom region, which are shown in Figs. S7 B and D.

	SimA	SimB	SimC	
SimA		0.8	0.8	Middle
SimB	2.2		0.4	
SimC	2.1	0.3		
Bottom region				

Table S3: Elastic constants of the whole connector and the middle region. These constants were calculated using three equilibrium MD trajectories, namely SimA, SimB and SimC. Standard errors were calculated using the block-averaging method with non-overlapping blocks.

Elastic properties	Whole connector	Middle region
Stretching spring constant	$2900 \pm 500 \text{ pN nm}^{-1}$	$20000 \pm 3200 \text{ pN nm}^{-1}$
Torsional spring constant	–	$2300 \pm 150 \text{ pN nm deg}^{-2}$
Coupling constant	–	$7800 \pm 4000 \text{ pN deg}^{-1}$
Young's modulus elasticity	$0.4 \pm 0.08 \text{ GPa}$	$3.6 \pm 0.6 \text{ GPa}$

3. FIGURES

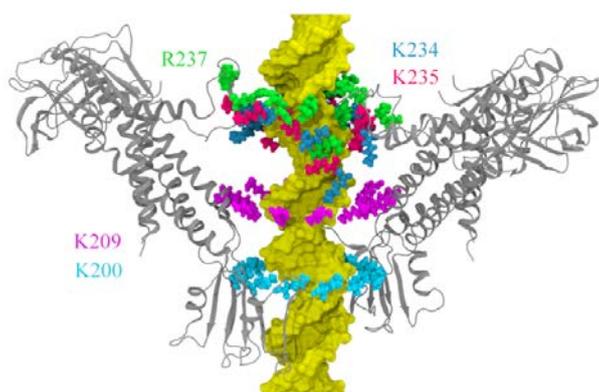


Figure S1: Positions of five electropositive residues of the connector channel (*grey*) in presence of DNA (*yellow*). Three residues, K234 (*blue*), K235 (*red*), and R237 (*green*), are part of the loop region. Two residues, K200 (*cyan*) and K209 (*magenta*), form rings inside the channel. Only four opposite subunits of the connector dodecamer are shown for visibility.

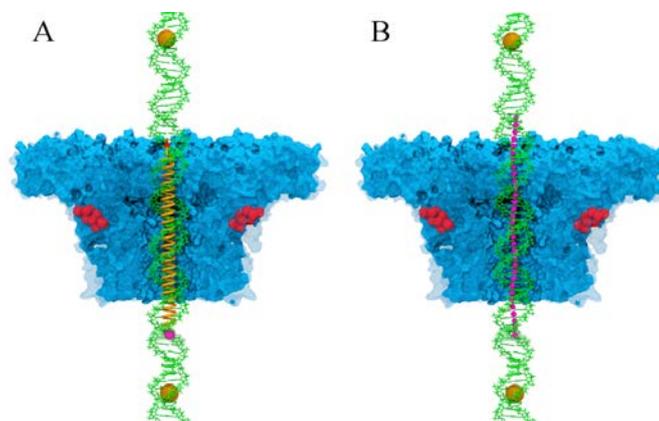


Figure S2: Two types of force-probe simulations were performed to move the DNA (*green*) through the connector channel (*blue*). Brown spheres depict the location of virtual reference particles. Red spheres represent C-alpha atoms, which were restrained during simulations to restrict translational and rotational motions of the connector. (A) The center of mass (*single magenta sphere*) of one base-step was pulled towards the upper virtual particle with a “virtual” spring (*brown*). (B) The centers of mass (*line of magenta spheres*) of 32 base-steps were pulled simultaneously towards both the upper and the bottom virtual particle with “virtual” springs (not depicted).

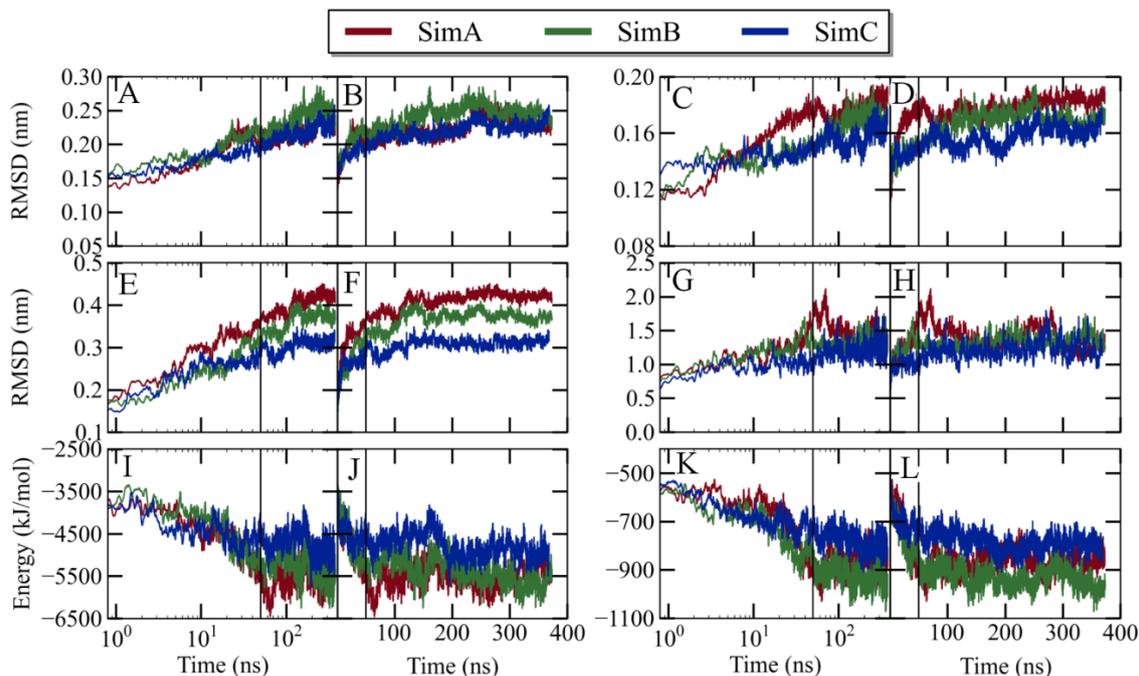


Figure S3: RMSDs and interaction energies during three equilibrium simulations, SimA (red), SimB (green), and SimC (blue). RMSD variations with respect to logarithmic and linear time are shown for (A-B) the whole connector, (C-D) the channel, (E-F) the loop-DNA region, and (G-H) the DNA. Variations in (I-J) electrostatic and (K-L) van der Waals interaction energies between the connector and the DNA with respect to the logarithmic and the linear time are shown. The vertical black line indicates 50 ns simulation time in all plots.

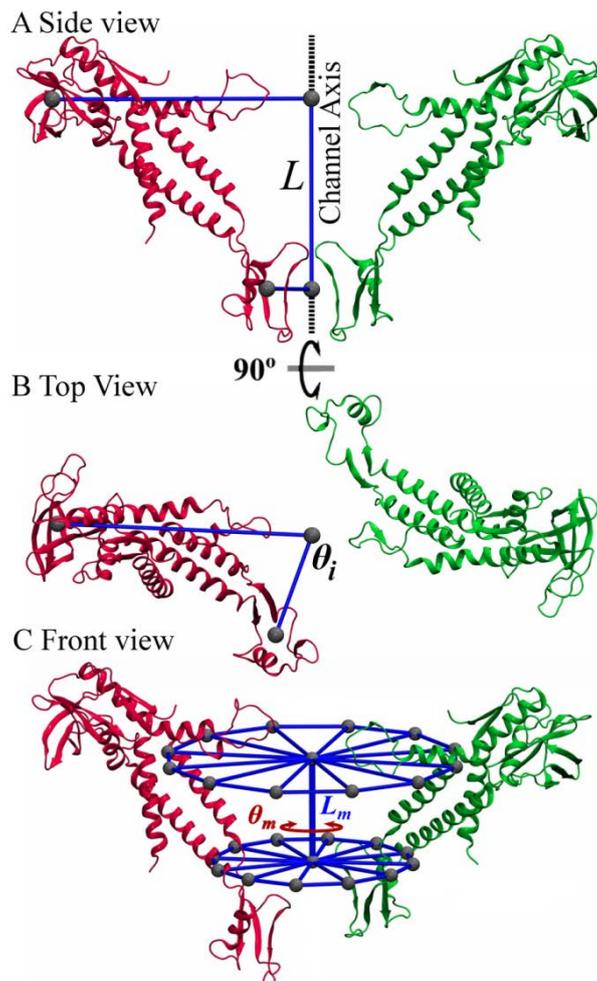


Figure S4: Definition of the twist angle and length of the whole connector and the middle region. (A) Side view, (B) top view, and (C) front view of two opposite subunits (*red* and *green*). (A) The length L between the centers of mass (*grey spheres*) of the upper and the bottom region of the connector was used to quantify compression-stretching motions of the connector. (B) Each subunit is tilted with respect to the principal axis by angle θ_i . The average of θ_i for twelve subunits was defined as twist angle to quantify untwisting-twisting motions. (C) Gray spheres depicted at the circumference of discs are C-alpha atoms from one of the helices of each subunit and the sphere at the center of the disc is the center of mass of these atoms. The length L_m of the middle region is quantified by the distances between the centers of mass of the upper and lower disc. The twist angle θ_m of the middle region is quantified by the rotation between the upper and lower disc.

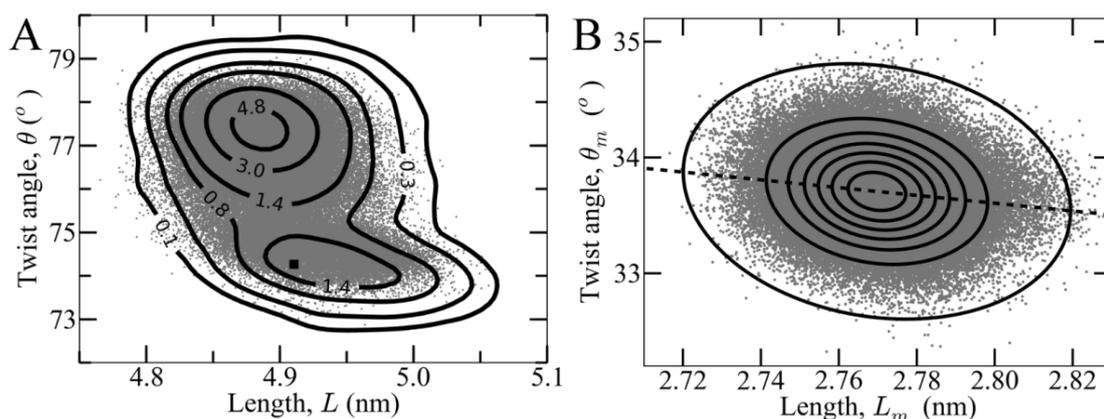


Figure S5: Twist angle and length fluctuation analysis. (A) Twist angle and length of the whole connector (*grey dots*). The solid square depicts the value obtained from the X-ray crystal structure. Contour lines with spacing labeled in $\text{deg}^{-1}\text{nm}^{-1}$ illustrate joint probability densities for angle and length. (B) Twist angle and length of the connector middle region (*grey dots*). Contour lines with a spacing of $5 \text{ deg}^{-1}\text{nm}^{-1}$ show joint probability densities for angle and length. The dashed straight line shows the coupling between twist angle and length.

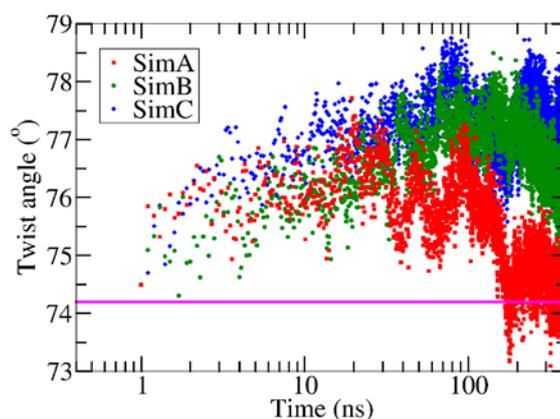


Figure S6: Deviations in twist angle of the whole connector (*red, green and blue dots*) with respect to the crystal structure value (*magenta line*) during the three individual simulations (SimA, SimB, and SimC). Note that the simulation time is shown in logarithmic scale.

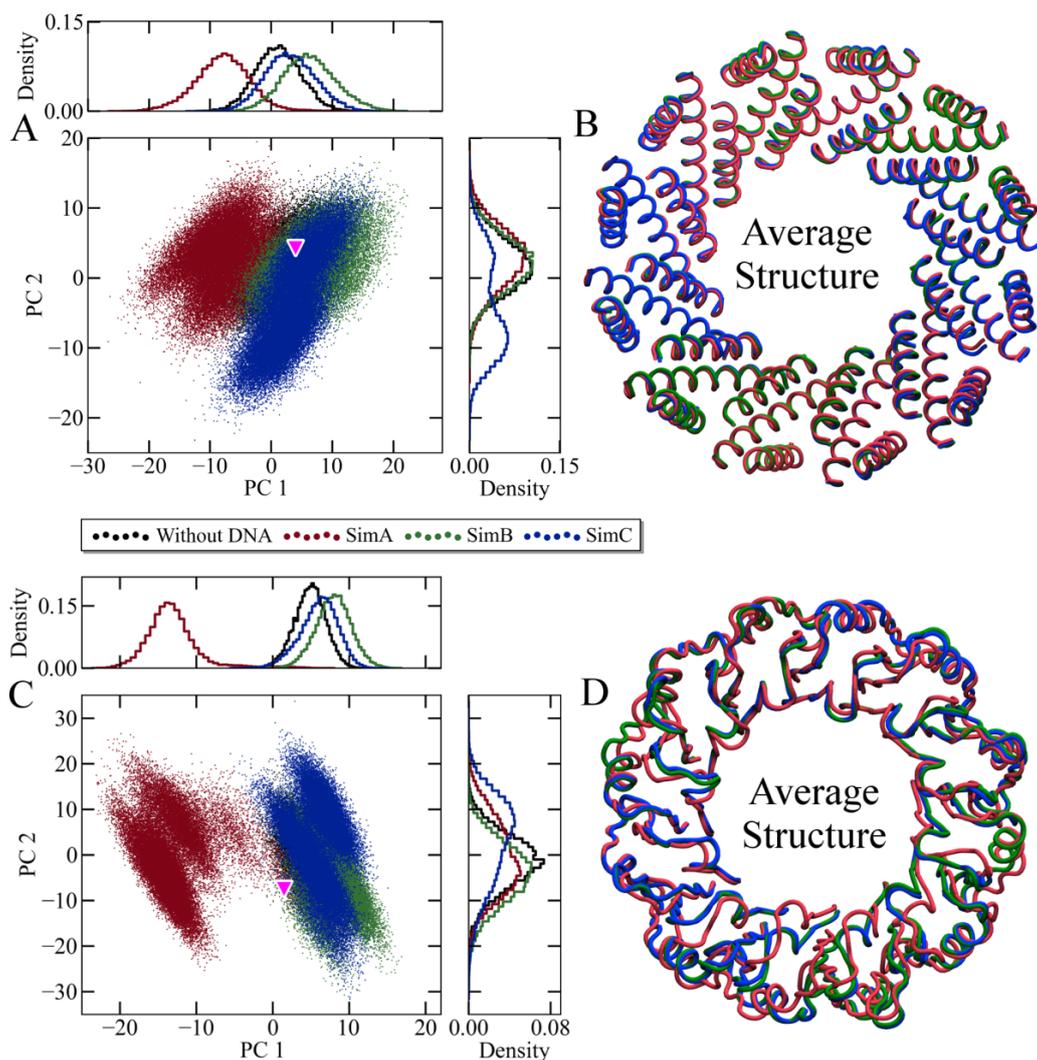


Figure S7: Comparison of conformational sampling during three and one equilibrium simulations of the connector with (*red, green and blue*) and without (*black*) DNA, respectively, using PCA. For reference, the projection of the X-ray crystal structure (*magenta triangle*) is shown. (A) Two-dimensional projection of first and second PCs that were obtained for the connector's middle region. (B) Illustration of the average middle region conformations obtained from three trajectories (SimA, SimB, and SimC) after filtering for the first and second PC. (C) Projection plane for the connector bottom region. (D) Illustration of the average bottom region conformations that were obtained from respective three trajectories after filtering for the first and second PC.

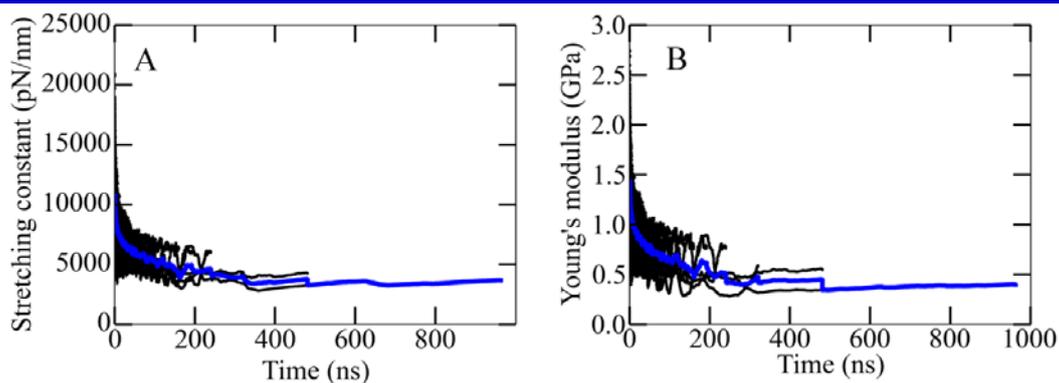


Figure S8: Convergence in elastic properties of the whole connector during the simulations. Both the time-blocks (*black dots*) as well as average (*blue line*) values of the (A) Stretching spring constant and (B) Young's moduli with respect to the time are shown for the whole connector. The property values were obtained from non-overlapping time-blocks of the MD trajectory. The obtained value from each block is shown as a black dot. The average value over the respective block size is shown as a blue solid line.

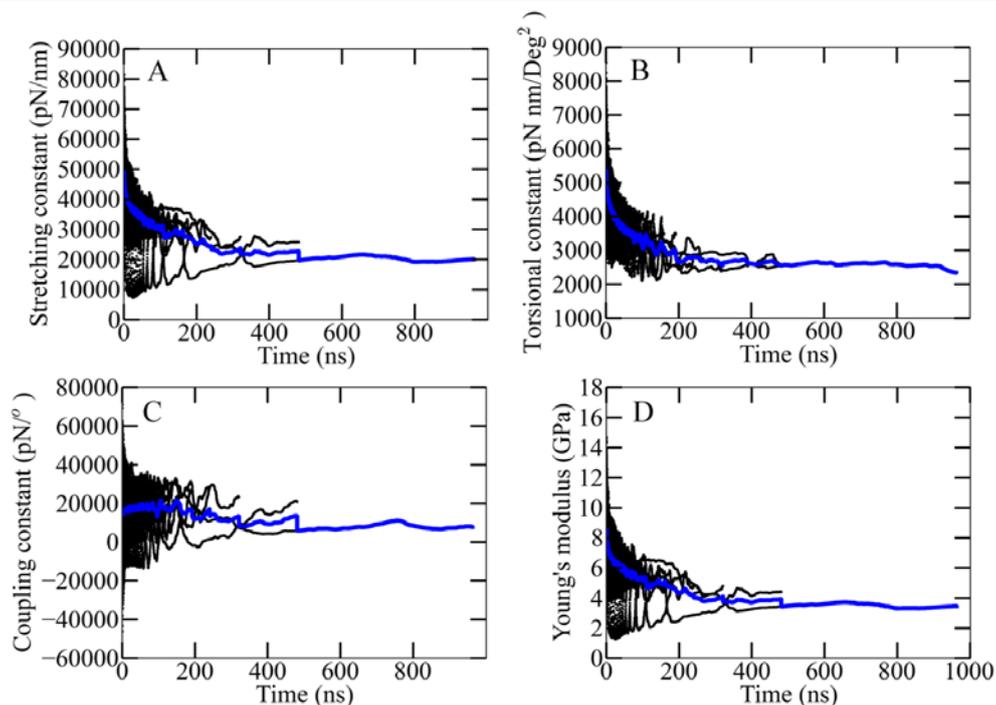


Figure S9: Convergence in elastic properties of the middle region during the simulations. Both the time-blocks (*black dots*) as well as average (*blue line*) values of the (A) Stretching spring constant, (B) Torsional spring constant, (C) Coupling constant and (D) Young's moduli with respect to the time are shown for the middle region. The property values were obtained from non-overlapping time-blocks of the MD trajectory. The obtained value from each block is shown as a black dot. The average value over the respective block size is shown as a blue solid line.

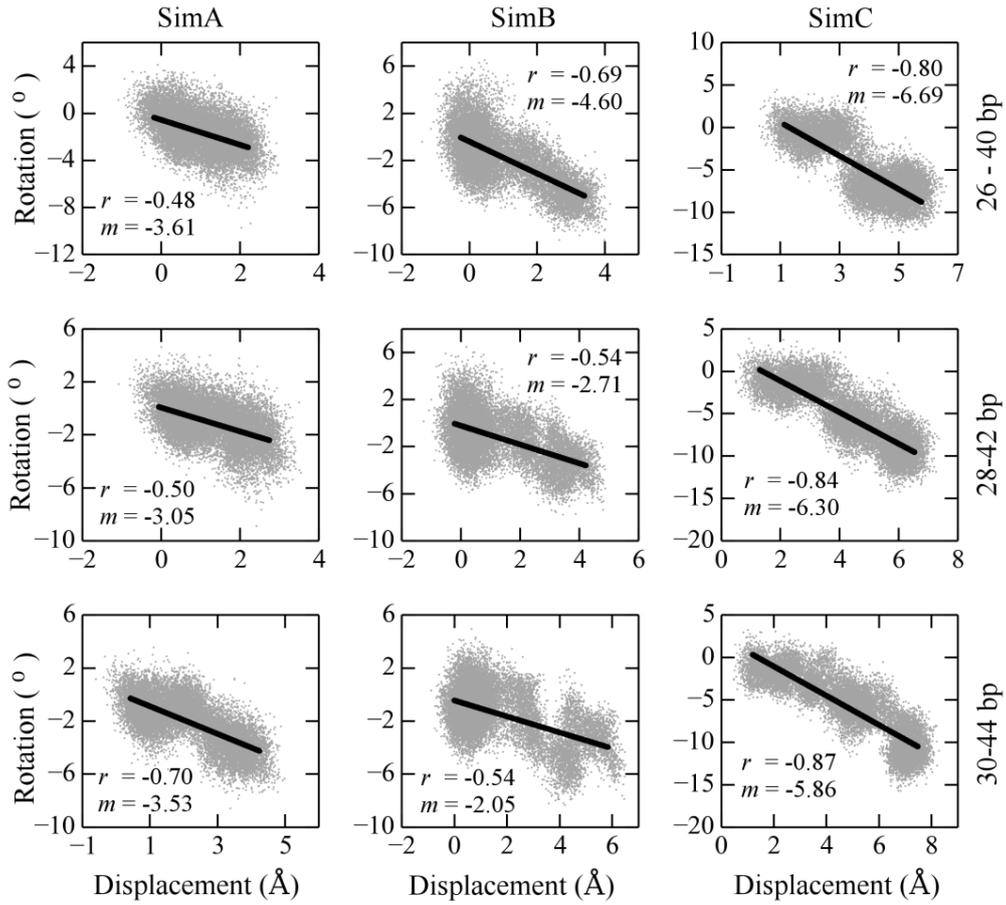


Figure S10: Rotation of the DNA during the inward transport. All figures show the DNA rotation with respect to the inward DNA displacement of three 15 bp DNA segments during three FP-T1 simulations (SimA, SimB, and SimC). The correlation coefficient r between rotation and displacement is shown for each DNA segment. The slope m of the fitted line, which represents the coupling between rotational and translational motion of the DNA, is also shown for each segment during the three simulations. Average values including standard error of the correlation and the coupling are $r = -0.67 \pm 0.08$ and $m = -4.2 \pm 0.9$ °/bp, respectively, as referred to in the main text.

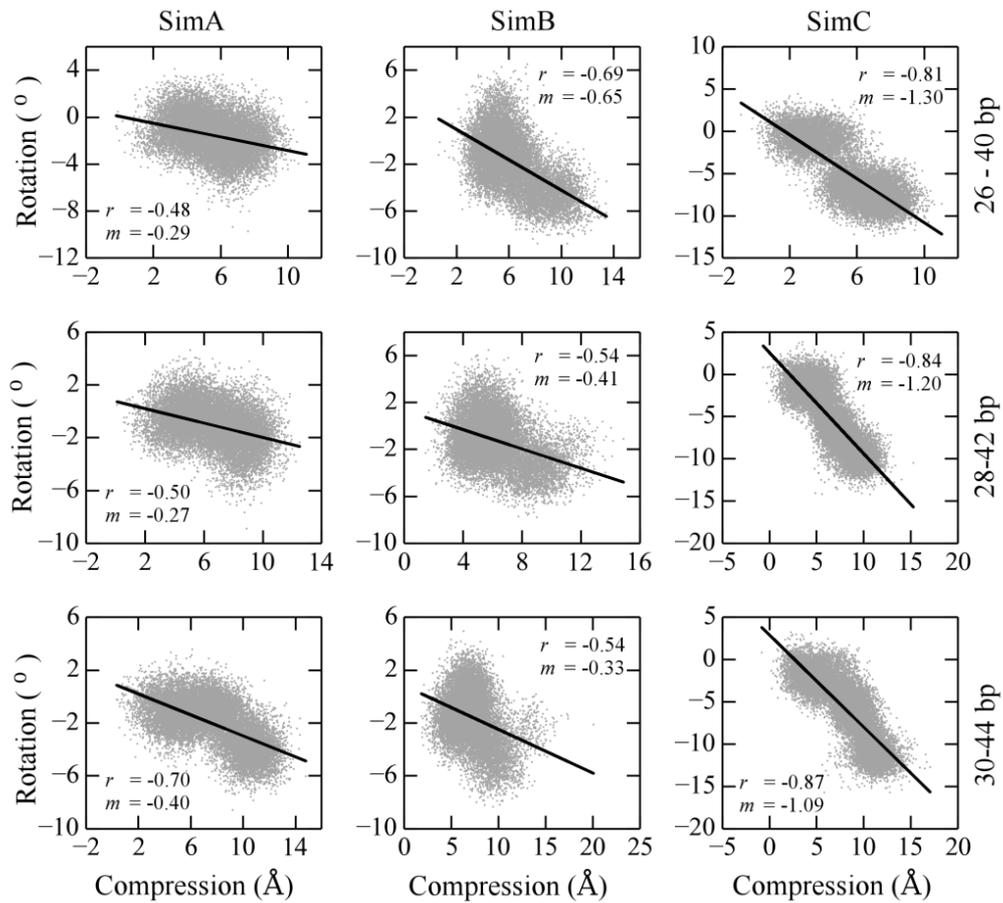


Figure S11: Rotation of the DNA with respect to its compression during inward transport. All figures show the DNA rotation with respect to the compression of three 15 bp DNA segments during three FP-T1 simulations (SimA, SimB, and SimC). The correlation coefficient r between rotation and compression is shown for each DNA segment. The slope m of the fitted line, which represents the coupling between rotation and compression of the DNA is also shown for each segment during the three simulations. Average values including standard error of the correlation and the coupling are $r = -0.67 \pm 0.08$ and $m = -0.66 \pm 0.23 \text{ } ^\circ/\text{\AA}$, respectively, as referred to in the main text.

4. REFERENCES

1. Kumar, R. and Grubmuller, H. (2014) Elastic properties and heterogeneous stiffness of the phi29 motor connector channel. *Biophys J*, **106**, 1338-1348.