Elastic Properties and Heterogeneous Stiffness of the Phi29 Motor Connector Channel

Rajendra Kumar and Helmut Grubmüller*
Max-Planck-Institute for Biophysical Chemistry, Department of Theoretical and Computational Biophysics, Göttingen, Germany

ABSTRACT The DNA packaging motor of the bacteriophage ϕ29, comprising head-tail connector, ATPase, and pRNA, transports the viral DNA inside the procapsid against pressure differences of up to ~60 atm during replication. Several models for the DNA packaging mechanism have been proposed, which attribute different roles to the connector, and require specific mechanical properties of the connector. To characterize these properties at the atomic level, and to understand how the connector withstands this large pressure, we have carried out molecular dynamics simulations of the whole connector both in equilibrium and under mechanical stress. The simulations revealed a quite heterogeneous distribution of stiff and soft regions, resembling that of typical composite materials that are also optimized to resist mechanical stress. In particular, the conserved middle α-helical region is found to be remarkably stiff, similar only to structural proteins forming viral shell, silk, or collagen. In contrast, large parts of the peripheral interface to the ϕ29 procapsid turned out to be rather soft. Force probe and umbrella sampling simulations showed that large connector deformations are remarkably reversible, and served to calculate the free energies required for these deformations. In particular, for an untwisting deformation by 12°, as postulated by the untwist-twist model, more than four times' larger energy is required than is available from hydrolysis of one ATP molecule. Combined with previous experiments, this result is incompatible with the untwist-twist model. In contrast, our simulations support the recently proposed one-way revolution model and suggest in structural terms how the connector blocks DNA leakage. In particular, conserved loops at the rim of the central channel, which are in direct contact with the DNA, are found to be rather flexible and tightly anchored to the rigid central region. These findings suggest a check-valve mechanism, with the flexible loops obstructing the channel by interacting with the viral DNA.

INTRODUCTION

The bacteriophage ϕ29 infects the bacterium *Bacillus subtilis* and uses its molecular machinery for replication (1). During this process, the precursor capsid (procapsid) forms by the incorporation of newly synthesized proteins inside the bacteria. Subsequently, a DNA packaging motor assembles at the base of the procapsid and transports viral DNA into it (1,2). The motor comprises the connector (shown as *red* in Fig. 1 *A*), prohead-RNA (pRNA), and an ATPase (3–9). Both the pRNA as well as ATPase detach from the connector after completion of DNA packaging, and the tail proteins of the phage particle join at the bottom region of the connector, completing the formation of new phage particles.

The DNA packaging motor transports viral DNA against an internal-to-external pressure difference of ~60 atm (10) generated by the densely filled viral DNA inside the procapsid. Accordingly, it is one of the strongest biological molecular motors, which renders it a promising choice for a motor in nanodevices (2). Further potential applications have been suggested, such as a molecular sorter in nanopore-based DNA sequencing devices, or as a model system to develop antiviral drugs to treat infections caused by herpes simplex virus, adenovirus, parvovirus, or pox virus (2). Additionally, motor pRNA is used as a carrier for ribozyme and antisense RNA to inhibit the Hepatitis B virus, which renders it a potential gene delivery system (11).

With a view to these diverse emerging applications, several studies were aimed at elucidating the underlying molecular DNA packaging mechanism, and in particular the properties and the role of the connector (2,12,13). The average viral DNA translocation rate is ~165 basepairs per second at the initial phase, with a gradual decrease as packing progresses (14). The motor generates up to ~110 pN at the final stage of the filling (14). Translocation of 10 basepairs of viral DNA into the procapsid requires hydrolysis of five ATP molecules by the ATPase (13).

Here, we focus on the connector as one of the motor components, for which the crystal structure has been determined (9,15). The connector is a truncated cone-shaped dodecameric protein complex that forms a channel through which viral DNA is transported inside the procapsid (9) (Fig. 1, *A* and *B*, see also Fig. S1 in the Supporting Material). The connector can be divided into three major regions—an upper, a middle, and a bottom region (*blue*, *brown*, and *green regions* in Fig. 1 *B*). The upper and middle regions are in direct contact with the procapsid, whereas the bottom region is connected to the middle region by a hinge region (*red*) and protrudes toward the outside of the procapsid. The central DNA transport channel is mainly formed by the middle and bottom region.

To uncover the mechanism by which the connector withstands the large pressure difference of up to ~60 atm, we focus on its mechanical properties. Because the connector

Submitted July 26, 2013, and accepted for publication January 23, 2014.

*Correspondence: hgrubmu@gwdg.de

Editor: Fazoil Ataullakhanov.

0006-3495/14/03/1338/11 $2.00

http://dx.doi.org/10.1016/j.bpj.2014.01.028

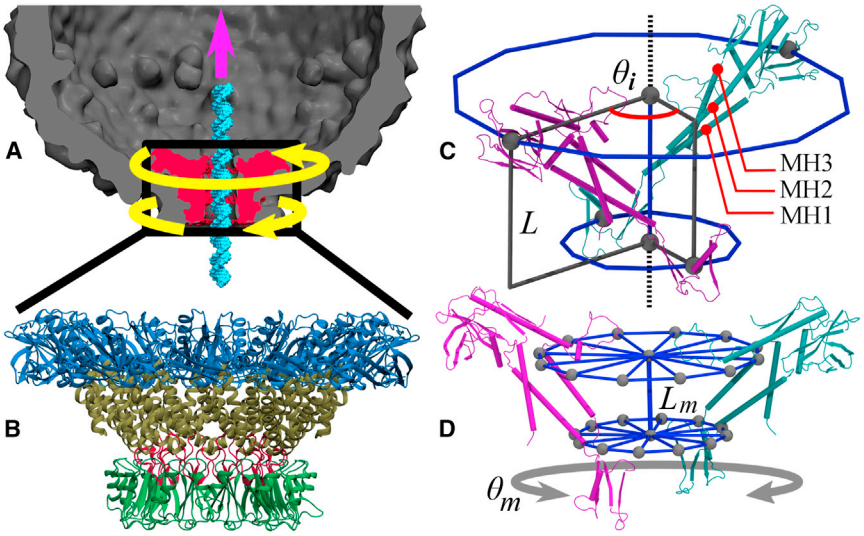

FIGURE 1 Illustration of the untwist-twist DNA packaging model; different regions of the connector (*left*), and definition of length and twist deformations (*right*) used for analysis. (*A*) The connector (*red*) and the DNA (*blue*) are aligned with the cryo-EM density map (obtained from the EM data bank, code EMD-1117 (7,8)) of the ϕ29 procapsid (*gray*). According to the untwist-twist DNA packaging model (9), the connector untwists and twists by 12° (*yellow arrows*) while the DNA is transported in two basepairs steps toward the procapsid interior (*magenta arrow*). (*B*) The structure of the connector reveals four distinct regions. The upper (*blue*) region faces the procapsid interior; the middle (*brown*) and bottom (*green*) regions, which form the channel, are connected by the hinge (*red*) region. (*C* and *D*) Two opposite (*purple*, *cyan*) out of the 12 connector subunits are illustrated in cartoon representations. (*C*) The whole connector's compression-stretching motions are characterized by the length *L* between the centers of mass (*dark gray spheres*) of the upper and bottom regions along the channel axis (*black dotted line*). Twist deformation is defined via relative rotation of the bottom rim with respect to the top rim (*blue circles*) and determined by the average angle θ of all torsion angles θ_i of the 12 individual subunits. The middle region consists of the three helices MH1, MH2, and MH3, referred to in the text. (*D*) Similarly, compression and twist deformations of the middle region are characterized by the length L_m between the center (*gray spheres*) of the upper and the lower disk, and the respective twist angle θ_m. The disks (*blue*) are defined via the Cα atoms of the MH1 and MH2 terminal residues, respectively (*gray spheres*) and the centers of mass of the Cα atoms of the respective disk (*axial gray spheres*).

is part of the procapsid (Fig. 1 *A*), its mechanical properties should also be similar to those of the procapsid, which have been probed earlier by atomic force microscopy (AFM) (16,17). In particular, a very large elastic modulus of ~1.8 GPa (i.e., similar to that of hard plastic) was measured, which was therefore suggested to be crucial for withstanding such high pressure difference (17). For the procapsid, a spring constant of ~0.07 N/m has been determined, whereas the connector region was found to be softer, with a spring constant of ~0.04 N/m (16). These findings render the question of how such a channel protein resists pressure differences of up to 60 atm even more puzzling.

The role of the connector in the DNA packaging mechanism is also unclear. One model, from Hendrix (18), proposed that the connector performs full rotations with respect to the procapsid, progressing by 12° in each step, thereby translocating two basepairs of the viral DNA inside the procapsid by consuming energy released from the hydrolysis of one ATP. However, more recent single-molecule florescence spectroscopic studies performed by Hugel et al. (19) ruled out such complete rotation of the connector.

A second model, which is compatible with the above results, has been proposed by Simpson et al. (9) and assumes partial rotations, with springlike coupled compression-stretching and twisting-untwisting motions (indicated in Fig. 1 *A*) of the middle helical region. According to this model, which is subsequently referred to as untwist-twist model, the whole connector first untwists by 12° through untwisting of its middle region (*yellow arrows*) under consumption of one ATP molecule; simultaneously, it stretches by 0.68 nm and grasps the next two basepairs of DNA. The connector then relaxes by twisting and contracting into its original configuration, thereby generating the force required to translocate two basepairs of the viral DNA toward the interior of the procapsid. Accordingly, this model requires the untwisting of the connector to be coupled by 17.6°/nm to an increase of its length.

A third model, the push-and-roll model, from Yu et al. (12), Moffitt et al. (13), and Aathavan et al. (20), proposes that a pentameric ATPase pushes the viral DNA by a "lever" into the procapsid in a fashion that the DNA rotates and rolls inside the motor channel during packaging process.

Recently, a fourth, the one-way revolution model, has been proposed by Guo, Schwartz, and co-workers (21–27), according to which the hexa-meric ATPase actively pushes the DNA into the procapsid, and the four electropositive lysine rings of the connector channel facilitate the DNA revolution, albeit without DNA rotation inside the channel during the packaging process. Simultaneously, the connector acts like a one-way valve and prevents leakage of the viral DNA against large counterpressure.

To reveal the molecular determinants of the exceptional mechanical stability and pressure resistance, we first probed the mechanical properties of the connector by equilibrium and nonequilibrium (force probe) molecular dynamics (MD) simulations (28–30). Remarkably, it turned out that the channel formed by the middle region of the connector is one of the stiffest known protein materials. To test which of the above packaging models is compatible with the properties of the connector, we next computed the deformation free energies required for twisting-untwisting as well as for compression-stretching motions both by a fluctuation

analysis of equilibrium MD simulations and by free energy umbrella sampling simulations. Our results demonstrate that the untwisting of the connector by 12°, as required by the untwist-twist model, would require much more energy than the ~50 kJ/mol provided by hydrolysis of one ATP molecule. Additionally, in recent voltage-ramping experiments (27), the connector facilitated DNA transport across the membrane even in the absence of the ATPase. These results are difficult to reconcile with this model, and thus render either the push-and-roll or the one-way revolution model more likely.

METHODS

Modeling and structure refinement

Four x-ray crystal structures of the connector at different resolutions have been published (9,15,31). We used the highest-resolution structure (PDB:1H5W; 2.10 Å) as a starting structure for all subsequent simulations. In none of the above structures, the residues ranging from A230 to S244 were resolved. These form a loop pointing toward the DNA and most likely interact with the DNA. (These were modeled and further refined as described in Section S1 in the Supporting Material, including Table S1, Fig. S1, and Fig. S2.)

Equilibrium MD simulations

The connector was extracted from the refined connector-DNA complex and taken as starting structure for subsequent equilibrium MD simulations. The simulation system was set up as follows: the connector was placed within a dodecahedron periodic box, solvated by addition of 91,559 water molecules, and neutralized by 84 sodium ions. The complete simulation system comprised 326,925 atoms. This system was energy-minimized and then heated up to 300 K during an NVT simulation with 1 fs time step. During this period, all heavy atoms were restrained with a harmonic force constant of 1000 kJ/(mol nm^2). During a subsequent 3 ns NPT simulation, the restraints for the heavy atoms were gradually removed. Finally, a 200 ns free equilibrium MD simulation was performed using a 2-fs time step.

All simulations were carried out using the GROMACS 4.0.7 package (32) with the AMBER ff99SB force field (33) and TIP3P water model (34). Long-range electrostatic interactions were computed by the particle-mesh Ewald method with a grid spacing of 1.2 Å and fourth-order cubic interpolation (35). Short-range nonbonded interactions were computed for all atom pairs within a 10 Å cutoff. The temperature was maintained at 300 K using the V-rescale algorithm (36) with a coupling time constant of 0.1 ps; during the equilibration and free equilibrium MD simulations, the pressure was maintained at 1 atm using Berendsen and Parrinello-Rahman pressure couplings (37,38) with a 1 ps coupling time constant, respectively. All bonds were constrained using the parallel LINCS implementation (39,40). The VMD program was used for visualization (41).

Umbrella sampling simulations

Deformation free energy profiles for twisting-untwisting and compression-stretching motions of the connector were calculated from umbrella sampling simulations. To obtain the reaction coordinates for umbrella sampling simulations, we performed several force-probe (FP-1, FP-2, FP-3, FP-4, FP-5, and FP-6) and relaxation (Relax-1, Relax-2, Relax-3, and Relax-4) simulations (see Section 3 and Table S2 of the Supporting Material for more details.) For the compression-stretching motion, 14 sampling windows were prepared, with umbrella potential minima between L_m = 2.725 and 3.149 nm (for details, see Table S3). Appropriate starting structures for these simulations were taken from snapshots at various times from the Relax-1, Relax-2, Relax-3, Relax-4, and equilibrium simulations. Each umbrella sampling simulation was performed for 100 ns (total 14 × 100 ns), and otherwise identical to the equilibrium simulations described above. The first 50 ns of each umbrella sampling simulation were discarded as an equilibration phase; the last 50 ns were used for collecting histograms (see Fig. S4), from which deformation free energy was calculated using by weighted histogram analysis method (WHAM) (42–44).

Deformation free energy for the twisting-untwisting motion was obtained in a similar manner, using the twist angle θ_m as a reaction coordinate (see Fig. 1 *E*). Fourteen umbrella windows were defined for which starting structures were taken from the FP-5 and FP-6 simulations. Because the conformations were close to the minimum in the energy landscape during these simulations compared with other force probe simulations, we have selected snapshots at different times from these two simulations as starting structures (for details, see Table S4). The umbrella sampling simulations were preceded by an equilibration simulation for 25 ns to equilibrate the connector structure at each respective umbrella potential minimum (total 14 × 25 ns). In these simulations, the atoms of the reference and rotational group (*upper* and *lower disk* in Fig. 1 *E*, respectively) were harmonically restrained by a 10,000 kJ/(mol nm^2) of force constant at the starting position such that the structures equilibrated at the starting twist angle of each umbrella window. Subsequently, two sets of umbrella sampling simulations were performed (15 ns for each window, total 28 × 15 ns). For the first set, the positions of reference atoms in the upper disk (shown in Fig. 1 *E*) were restrained while those of the bottom disk were subjected to an isotropic pivotal free rotational potential (29), and vice versa for the second set of simulations. The harmonic force constants used in simulations are listed in Table S4. To calculate the deformation free energy of the twisting-untwisting motion, a torsional spring constant for each window was calculated from the applied torque and resulting twist angle (for details, see Fig. S5). Histograms were collected from the umbrella sampling simulations and the deformation free energy was calculated using weighted histogram analysis method (WHAM) (42–44).

Analysis

Structural descriptors for the elastic properties

To probe the elastic properties of the connector, two structural descriptors, namely the twist angle and the length, were considered that describe the twisting-untwisting and compression-stretching motion of the connector, respectively. The connector length L is the distance between the center of masses of the upper and the bottom region (see Fig. 1 *C*). The whole connector's twist angle

$$\theta = \frac{1}{12}\sum_{i=1}^{12}\theta_i$$

has been calculated by averaging over the twist angle θ_i of each subunit (see in Fig. 1 *C*). Similarly, the twist angle θ_m and length L_m of the middle region were calculated to investigate the elasticity of the channel, where θ_m denotes the rotational angle between the upper and the lower disk and L_m the distance between the two disks shown in Fig. 1 *D*. The elastic properties were calculated using the snapshots taken at 2 ps time difference during the last 100 ns of equilibrium MD simulations.

Elastic properties and spring constants from equilibrium simulations

The free energy landscape $G(\theta,L)$ as a function of the connector length L and angle θ as defined above was determined independently from equilibrium simulations in harmonic approximation from the equilibrium fluctuations of the connector via

$$G(\theta,L) = -k_BT\ln p(\theta,L),$$

where k_B and T are the Boltzmann constant and temperature, respectively. For a Gaussian approximation of the two degrees of freedom, θ and L, probability density is given as

$$p(\theta,L) \propto \exp\left(-\frac{1}{2}\left[(\theta-\overline{\theta}),(L-\overline{L})\right]\mathbf{C}^{-1}\begin{bmatrix}(\theta-\overline{\theta})\\(L-\overline{L})\end{bmatrix}\right), \quad (1)$$

where $\overline{\theta}$ and $\overline{L}$ are the average twist angle and length, respectively, and $\mathbf{C}$ is the covariance matrix. Because the energy landscape is assumed to be harmonic, its expression is

$$G(\theta,L) = \frac{1}{2}\left[K_\theta(\theta-\overline{\theta})^2 + K_L(L-\overline{L})^2\right] + K_c(\theta-\overline{\theta})(L-\overline{L}),$$

where K_θ is the torsional spring constant, K_L is the stretching spring constant, and K_c is the coupling constant between two motions. By comparing the above three equations, these constants can be written as

$$\begin{bmatrix}K_\theta & K_c\\K_c & K_L\end{bmatrix} = k_bT\mathbf{C}^{-1}. \quad (2)$$

This equation was used to calculate the spring constants from the equilibrium fluctuation for the whole connector and the middle region-containing channel.

Young's modulus of elasticity

To translate the obtained elastic constant into a Young's modulus of the connector, we described the connector as a homogeneous elastic material with elasticity Y and with the shape of a hollow truncated cone of length L (see Fig. S6). From the observed length change of

$$\delta = \int_0^L \frac{F}{YA(x)}dx$$

upon axial stress force F (45), where $A(x)$ is the cross-sectional area perpendicular to the symmetry axis, the elasticity Y in terms of stretching spring constant K_L is

$$Y = \int_0^L \frac{K_L}{A(x)}dx, \quad (3)$$

which yields, for the assumed hollow truncated cone having variant diameters,

$$Y = \frac{2K_L}{\pi}\frac{L}{D_bd_a - D_ad_b}\ln\left|\frac{(D_a+d_a)(D_b-d_b)}{(D_a-d_a)(D_b+d_b)}\right|, \quad (4)$$

where D_a and d_a are narrow-end exterior and interior diameter, respectively, and D_b and d_b are wide-end exterior and interior diameter of the truncated cone, respectively (*sketch* shown in Fig. S6). The dimensions are calculated for the whole connector and the middle region from equilibrium MD simulations and provided in Table S5.

Calculation of interface area

To quantify inter- and intrasubunit residue packing, we have calculated surface area (SA) accessible by a probe with a radius of 1 Å using the *g_sas* module of the GROMACS package (46). Intrasubunit packing was determined via helices MH1, MH2, and MH3 of the middle region (Fig. 1 *C*),

$$SA^{\text{intra}} = [(SA_{MH2} + SA_{MH3}) - SA_{MH2\text{-}MH3}]/2 + [(SA_{MH2\text{-}MH3} + SA_{MH1}) - SA_{MH1\text{-}MH2\text{-}MH3})]/2, \quad (5)$$

where subscripts denote either separate helix or attached pair helices. Similarly, intersubunit packing was calculated via

$$SA^{\text{inter}} = \left[\left(SA_i + SA_j\right) - \left(SA_{ij}\right)\right]/2, \quad (6)$$

where SA_i and SA_j refer to separated adjacent subunits, and SA_{ij} to the attached subunits.

RESULTS AND DISCUSSION

We first monitored the structural stability of the connector during equilibrium MD simulations via root mean-square deviation calculated for all C-α atoms with respect to the crystal structure (see Section S2, Fig. S7, and Fig. S8 in the Supporting Material). Overall, the structural changes during 200-ns equilibration were rather small, with the root mean-square deviation stabilizing at 0.25 nm. The last 100-ns part of the trajectory was used for further analysis.

How does the connector withstand the high pressure difference?

Elasticity of the connector from equilibrium fluctuations

To study the structural determinants that enable the connector and its channel to withstand large relative pressure, we first calculated the elastic properties of the whole connector as well as its middle region (see Table S6). The stiffness opposing twisting-untwisting and compression-stretching motions (see Fig. 1, *C* and *D*), described by torsional (K_θ) and stretching (K_L) spring constants (see Table S6) along with coupling constants (K_c) between motions, was determined using Eq. 2 from the equilibrium fluctuations shown in Fig. 2, *A* and *B*. As can be seen, the calculated torsional and stretching spring constants of the middle region are approximately two and eight times larger than that of the whole connector, respectively; therefore, most structural changes under pressure are expected for the upper and bottom regions.

Using the stretching spring constants and taking into account the geometry of the connector (see Eq. 4), Young's moduli of 0.36 ± 0.06 GPa and 3.4 ± 0.6 GPa for the whole connector and for the middle region, respectively, were obtained. Furthermore, the obtained moduli values were almost unchanged during the last 25 ns of the simulations that suggest these values are sufficiently converged (see Fig. S9). A similar convergence behavior has been seen also for a similar-sized molecular system, i.e., the central shaft γ-subunit of the F1-ATPase motor, for which, additionally, good agreement with experimental results has been shown (47).

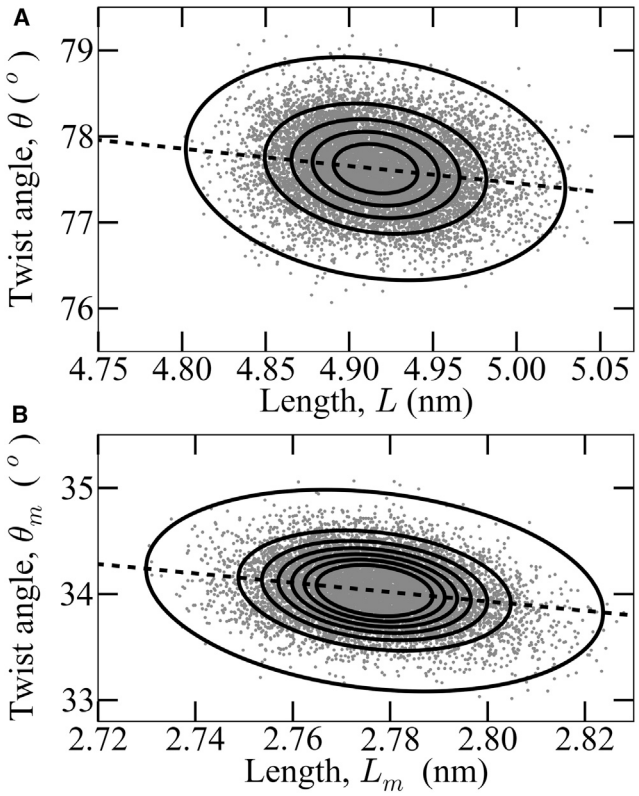

FIGURE 2 Equilibrium fluctuations of twist angles and lengths. (*A*) Fluctuations of twist angle θ and length L of the whole connector, derived from equilibrium simulations (*shaded dots*). The probability density was obtained using Eq. 1 (depicted as *solid lines* ($2^{\circ^{-1}}$ nm^{-1} spacing)). (*B*) Same representation for twist angle θ_m and length L_m and $5^{\circ^{-1}}$ nm^{-1} contour line spacing. Linear least-square fits (*dashed lines*) quantify the coupling between compression-stretching and twisting-untwisting motions.

For the whole ϕ29 procapsid, a Young's modulus of ~1.8 GPa was measured in AFM experiments (16,17); thus the connector as a whole appears to be softer than the procapsid, with a stiffer "core" inside. The elasticity of this middle region is similar to that of other structural proteins that withstand mechanical forces, such as collagen fibrils (0.2–11.5 GPa) (48), single-brin silkworm silk (5–17 GPa) (49), or dragline spider silk (11–13 GPa) (50). We now ask what are the structural determinants and the function of this heterogeneity.

Heterogeneous flexibility of the connector

To characterize the spatial distribution of the observed elastic heterogeneity as well as its relation to molecular fluctuations, we calculated, residuewise, root mean-square fluctuations (RMSF) for each subunit. Fig. 3 *A* (*green line*) depicts the obtained values averaged over the 12 subunits. As can be seen, overall the crystallographic temperature factors (*blue lines*, taken from Guasch et al. (15)) agree well with the RMSF (correlation coefficient: 0.83), which suggests that these fluctuations and the derived elastic properties are captured sufficiently accurately by our simulations.

To characterize the spatial distribution of the elastic properties, we identified rigid regions from RMSF depicted in Fig. 3 *A*. Residues with small fluctuations (RMSF $<$ 0.065 nm) were considered rigid (*red region* in Fig. 3, *B–D*), whereas those residues with larger fluctuations were considered flexible (*blue region*). In the upper region, rigid residues are embedded within flexible parts, very much like in composite materials. The middle region is the stiffest part and mainly consists of rigid residues. In contrast, the residues of the hinge and the bottom regions are flexible (Fig. 3 *B*). This finding suggests that the relatively low stiffness probed by AFM (16) is mainly due to the flexible and elastic bottom region of the connector.

Obviously, the regions defined in Fig. 1 *B* purely based on structural features also reflect different elastic properties of the connector. This mechanic heterogeneity suggests that these regions even fulfill different functions. As visible from Fig. 3 *C*, the pressure acts primarily on the rigid core (*red*), which defines the channel that is in direct contact to the DNA. To maintain a constant diameter of the channel, this region, therefore, has to withstand deformations due to the strong pressure gradient between the capsid interior and exterior as well as mechanical stress via the contacting parts of the procapsid, both in the presence and absence of DNA. Due to the functional relevance of this region, one would expect the structure of this region to be highly conserved. Indeed, very similar rigid helical scaffolds are also seen in other tailed phage connectors such as T7, SPP1, and P22 (51–57). Rigid residues are located inside the procapsid whereas the flexible bottom region is located outside of the procapsid.

Interestingly, the rigid core is directly exposed to the solvent of the capsid interior (*arrows* pointing high pressure in Fig. 3 *C*); here, the exposure to solvent evenly distributes the pressure over the surface of the core region. In contrast, contact of the upper region to the procapsid is buffered by flexible regions, which absorb local force peaks and distribute external forces evenly over the outer surface of the core region. In that respect, and also from the compartmented substructure of the outer rim visible in Fig. 3 *B*, this region resembles a composite material, very similar to the protein nanocrystals that are embedded within flexible unstructured peptide regions in silk (58) or the calcium carbonate (calcite or aragonite) crystals layers within a protein matrix in sea shells (59,60). From closer analysis of the direct contact region to the capsid (sketched in Fig. 3 *D*) we speculate that this structure particularly serves to withstand displacement of the connector due to vertical forces from the capsid interior, while tolerating horizontal motions or forces that the capsid channel rim may exert onto the middle region of the connector.

The flexible bottom region points toward the capsid exterior and, therefore, does not need to withstand large pressure

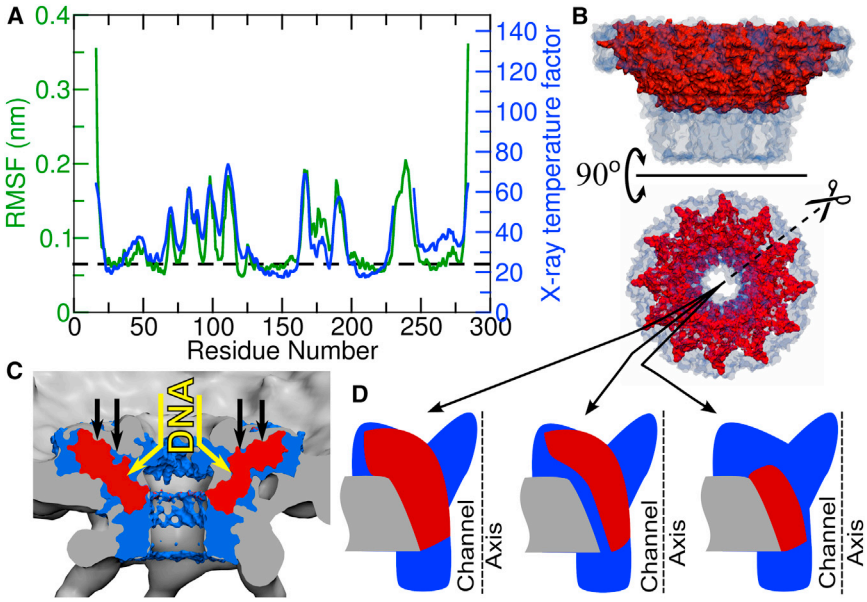

FIGURE 3 Heterogeneous distribution of rigid and flexible connector regions. (*A*) Root mean-square fluctuations (RMSF) (*green*) and temperature factors (*blue*) obtained from crystallography (15). Residues with an RMSF below 0.065 nm (*horizontal dashed line*) are considered rigid, otherwise flexible. (*B*) Spatial distribution of rigid (*red*) and flexible (*blue*) connector regions (*side* and *top view*) and (*C*) shown embedded within the procapsid (*gray*). Mainly rigid residues of the middle region (*red*) are exposed to pressure (*black arrows*) due to compacted DNA. This rigid region also supports loops formed by residues A230–S244 that directly interact with the DNA (*yellow arrows*). (*D*) Three different types of the connector-procapsid interfaces are illustrated by three schematic radial cross-sections (indicated by *three arrows*) of the connector (flexible regions in *blue*, rigid regions in *red*) and the part of the procapsid (*gray*) that is in direct contact with the connector. An alternating pattern of direct contacts is seen between the central rigid region (*red*), the rim of the procapsid (*gray*), and regions where flexible regions (*blue*) act as an interface.

differences or mechanical stress. Accordingly, its high flexibility might facilitate binding of the pRNA and ATPase at the start of the packaging process. Overall, the heterogeneous flexibility pattern seen in our simulations of the ϕ29 connector closely matches the mechanical properties that are required to perform its biological function within the extreme mechanical and pressure conditions of its procapsid environment.

Elastic properties determined from force probe simulations

Above, the elastic properties of the connector were probed via equilibrium simulations. Accordingly, the probed length and twist angle changes only covered the range accessible to thermal fluctuations. Further, by calculating elastic constants, we assumed a harmonic (i.e., Hookean) and reversible behavior of the connector. However, the different models of the packaging mechanism, as discussed in the Introduction, require much larger connector deformations than those observed above; it is thus unclear if the connector also exhibits Hookean properties in this case, which also would require a fully elastic, i.e., reversible behavior. We therefore need to determine its elastic limit, i.e., to which extent it can be deformed reversibly. It is, finally, unclear if the potentially functionally relevant coupling between stretching and twisting of the connector, as calculated from the equilibrium fluctuations shown in Fig. 2 *A*, also extends into deformations that are thermally inaccessible.

To address these questions, we carried out several sets of force probe (FP) simulations, which are presented and discussed in Section S3 of the Supporting Material (see Table S2, Fig. S10, Fig. S11, Movie S1, and Movie S2 in the Supporting Material). In summary, deformations within the range of $L = 4.70$–5.3 nm and $\theta = 70$–79.5° seem to be fully elastic, and the twisting-untwisting motion is linearly coupled to the compression-stretching motion within this elastic range. The connector recovered its equilibrium structure even after structural breakdown within the obtained elastic limit (see Fig. S12 and Fig. S13). The obtained coupling of 18°/nm is remarkably close to the coupling of 17.6°/nm required in the untwist-twist DNA packaging model proposed by Simpson et al. (9). We therefore asked next if the other elastic and energetic properties of the connector are also compatible with this model.

Role of the connector in the DNA packaging process

Twisting-untwisting of the connector

Specifically, according to the untwist-twist model, the coupled untwisting and stretching motion of the connector by 12° and 0.68 nm, respectively, is driven by hydrolysis of one ATP molecule, which under physiological conditions releases an energy of ~50 kJ/mol.

As a first step, to check if this energy suffices to induce the above deformation, we estimated the required free energy in harmonic approximation from the equilibrium fluctuations (see *gray dots* in Fig. S14), as described in the Methods. Extrapolation of this harmonic approximation (depicted as *contour lines*) suggests, however, that a very large amount is actually required for the deformation of 12°/0.68 nm, >10 times larger than the 50 kJ/mol available from ATP hydrolysis.

However, the above force-probe simulations (see Fig. S10, *B* and *C*) suggest that the simple harmonic approximation may not hold for such extreme deformations.

In a second step, we therefore computed the potential of mean force (Fig. 4, *A* and *B*) along both the

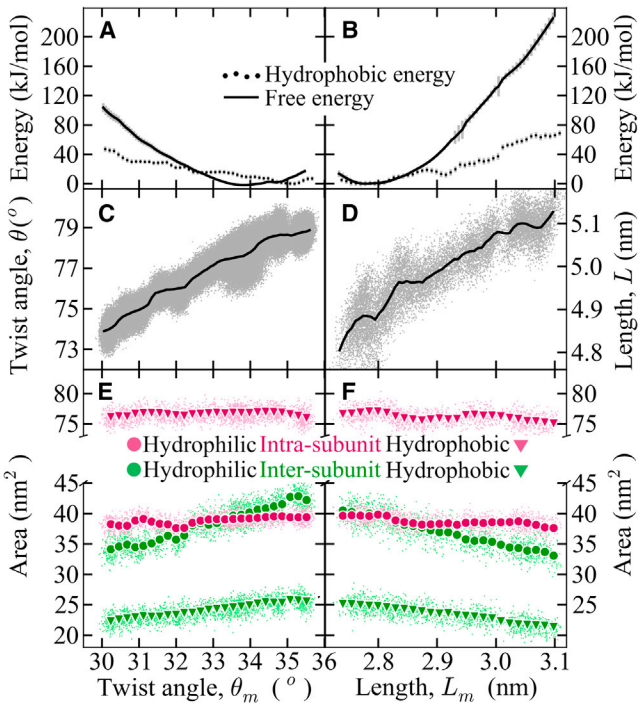

FIGURE 4 Deformation free energies and exposure of hydrophobic and hydrophilic residues of the connector from umbrella sampling simulations. Compressing-stretching deformations (L, L_m) and twisting-untwisting deformations (θ and θ_m) for the whole connector and, respectively, the middle region, are shown. (*A* and *B*) Deformation free energy (*black solid line*) as well as the contributions of hydrophobic solvent-accessible surface area to the free energies (*black dotted lines*) for (*A*) twisting-untwisting and (*B*) compression-stretching deformation of the connector middle region. (*C* and *D*) Coupling between deformations of the whole connector and respective middle region deformations (*small gray dots*), with running averages (*solid lines*). (*E* and *F*) Change of intersubunit (*green*) and intrasubunit (*red*) interface contact areas (*small dots*, as defined in the text) during (*E*) twisting-untwisting and (*F*) compression-stretching deformations of the middle region; decreasing hydrophobic (*triangle*) and hydrophilic (*circle*) contact areas imply increasing solvent exposure. Running averages are shown as larger symbols.

twisting-untwisting (θ_m) as well as the compression-stretching (L_m) coordinate from umbrella sampling simulations using the weighted histogram analysis method (see Methods for details). As can be seen from the first set of umbrella sampling simulations with length (L_m) of the middle region of the connector as reaction coordinate (Fig. 4, *B* and *D*), the available 50 kJ/mol allow for an extension of only 0.1 nm, thus confirming the above extrapolation. Similarly, using the twist angle (θ_m) as a reaction coordinate, already ~100 kJ/mol are required to untwist both the middle region and the whole connector by only 4° (Fig. 4, *A* and *C*), i.e., by one-third of the deformation required by the proposed untwist-twist model. Therefore, our umbrella calculations also strongly suggest that the free energy released by ATP hydrolysis is too small for the required deformations.

Furthermore, to characterize the coupling between the deformations of the middle region and the whole connector that was proposed in the untwist-twist model, we determined the whole connector's deformation, i.e., the change in the twist angle (θ) and the length (L) with respect to that (θ_m and L_m) of the middle region (Fig. 4, *C* and *D*). As can be seen, the deformation of middle region entirely translates to the whole connector, which suggests that the deformations of the whole connector and the middle region are indeed strongly coupled.

We note that our above estimate of the deformation free energy rests on the assumption that it does not decrease with stretching or untwisting beyond L_m =3.1 nm and θ_m =30°, as might happen, e.g., after complete structural breakdown of the connector. However, the reversible recovery of the equilibrium conformation after deformations of up to 5.1 and 5.3 nm, respectively (see Fig. S13, *A* and *B*), strongly speaks against a complete structural breakdown.

To assess the convergence of the umbrella sampling simulations, we compared the resulting untwisting-stretching path from the first set of simulations with the second set (shown in Fig. S15). These two paths are similar to each other (*red* and *black symbols*) and also coincide with the path obtained in FP-5/6 simulations (*green symbols*), underscoring that the free energy profiles are nearly converged.

All potentials of mean force were calculated in absence of the DNA, which may affect the energy required for the proposed untwisting/stretching. However, according to the above results, already half of the proposed deformation requires ~200 kJ/mol. In contrast, the maximum barrier height that can be overcome with the available 50 kJ/mol for an activated process is ~100 kJ/mol, assuming an additional barrier of 50 kJ/mol, estimated from a Kramer's attempt rate of 10^{10}/s and a plausible turnover rate of 165/s (14). We consider it therefore unlikely that the presence of DNA would decrease the required deformation free energy by such a large amount to drive the deformations required by the untwist-twist model. Interestingly, Jing et al. (27) observed in their voltage-ramping experiments that DNA was driven by the electrochemical gradient through the connector even in the absence of pRNA and ATPase. According to the Nernst equation, the applied potential of 75 mV (= 7.24 kJ/mol) translates, for the two basepairs (eight negative charges) by which the DNA is advanced in one cycle, to an energy of ~58 kJ/mol. In these experiments, this is thus the maximal available energy for possible connector deformations—again much smaller than required by the untwist-twist model. Combined with our above results, the observed DNA diffusion in the absence of an ATPase strongly suggests that in these experiments, the connector also facilitates DNA transport without performing the proposed twisting-untwisting or compression-stretching motion.

The connector as a one-way valve

As alternatives, the so-called push-roll and one-way revolution models have recently been proposed (12,21). Both

models postulate that the ATPase interacts with and actively transports the viral DNA. Additionally, the latter model proposes that the connector acts as a one-way valve, restricting leakage of the DNA during the packaging process (21). In fact, in the above-mentioned voltage-ramping experiments, selective one-directional diffusion of DNA across the lipid membrane through the connector was observed; for a reversed potential, no backdiffusion of the DNA was seen (27). Further, the connector's ability to restrict the leakage of DNA was probed through sedimentation assays in which the DNA was retained inside partially filled procapsids in the absence of pRNA/ATPase after ultracentrifugation (27).

Furthermore, in mutagenesis experiments, deletion/mutation of particular connector loop residues (K234–R237; loops depicted in Fig. S1 *B* and Fig. S2 *C*) severely affected the successful production of infectious bacteriophage (61). In sedimentation assays, the mutant K235A.E236A.R237A was unable to retain DNA inside completely filled procapsids in the presence of pRNA/ATPase. These results suggested that the connector loops are essential for the observed unidirectional DNA transport (26) and strongly interact with the DNA via three positively charged residues (K234, K235, and R237) per each subunit (61). One possible mechanism that would involve the connector loops as a key is that the connector acts as a check-valve, with the flexible loops bound to the DNA acting as movable parts that obstruct the channel and arrest the DNA against backmotion toward the outside, due to the strong pressure force. Such mechanism, to function, would require a rigid anchoring of the loops to withstand the backtracking force acting onto the DNA. Indeed, this particular loop region is solidly attached to and supported by the stiff region of the connector that has been identified above (*red region*, Fig. 3, *B* and *C*). Further, recent voltage-ramping experiments revealed that after deletion of these loops, the connector failed to restrict the backtransport of the DNA across the lipid membrane (21). Overall, the obtained heterogeneity in the mechanical properties of the connector seems to be compatible with the one-way revolution model.

Structure and energetic determinants for connector stiffness

To answer the question of which features of the connector determine its remarkable mechanical properties, we calculated and analyzed the extent to which selected structural and energetic quantities changed during the deformations induced by the above-described umbrella sampling simulations. As one possible source of the observed elastic restoring forces, we focused at the hydrophobic core (62–64) of the middle region of the connector. Should this core become partly exposed to the solvent during untwisting or stretching motions, the corresponding free energy cost might explain these forces. To test if this is actually the case, we calculated the hydrophobic solvent-accessible surface areas of the middle region (see Fig. S16) and estimated the resulting hydrophobic free energy from an energy cost of 18 cal/mol per unit surface area ($Å^{-2}$) (65) (*dotted lines* in Fig. 4, *A* and *B*). As can be seen, the hydrophobic energy indeed contributes approximately one-third to the total free energy, but clearly does not fully explain the connector's elastic properties.

We then asked if a possible loosening of the packing within the connector, and the resulting loss of intra- and intersubunit interactions within the middle regions, might also contribute to the mechanical properties. To this aim, we have quantified the packing within each subunit by the intrasubunit packing (Eq. 5), defined by the interface area (Fig. 4, *E* and *F*, *red symbols*) between each of the three helices (MH1, MH2, and MH3 shown in Fig. 1 *C*), as well as the packing between adjacent subunits (Eq. 6), defined by the interface area between two adjacent subunits (*green symbols*). Further, Fig. 4, *E* and *F*, distinguishes between hydrophobic and hydrophilic interface surface.

For the intrasubunit packing, the hydrophobic area was approximately two times larger than the hydrophilic area, suggesting the presence of an internal hydrophobic "core" within each subunit. However, no significant changes of the respective areas are present during both untwisting and stretching, such that their contribution to the deformation free energy is small. In contrast, both intersubunit hydrophilic and hydrophobic interface areas decreased during stretching and untwisting, indicating a loosening of their mutual packing. That this loosening actually implies significant interaction enthalpy changes is demonstrated in Fig. S17, which shows the respective van der Waals and electrostatic energies during deformation. Because, due to the size of the simulation system, the corresponding entropic part could not be quantified, we cannot decide whether this second free energy contribution, together with the hydrophobic surface changes quantified above, suffices to fully explain the deformation free energies derived from our umbrella sampling simulations. However, we consider it unlikely that the respective enthalpy/entropy reduces the resulting free energy to such a large extent as to render the intersubunit packing changes unimportant.

CONCLUSIONS

During assembly of many bacteriophages, strong motors transport the viral DNA inside the phage procapsid through a connector against pressure difference of up to ~60 atm. Here we have investigated the mechanical elastic properties of the ϕ29 head-tail connector (a component of the ϕ29 bacteriophage DNA packaging motor), and its possible role in the DNA packaging mechanism. A fluctuation analysis of equilibrium MD simulations revealed an exceptionally stiff α-helical channel region, with an elastic modulus that is only approached by proteins that have evolved to withstand strong mechanical forces, such as crystalline

silk domains. The functional relevance of this middle region is underscored by its high conservation level in head-tail connectors of other phages such as SPP1 and P22 (52,54–56). The high stiffness may serve to stabilize the channel and in particular to maintain its inner diameter in absence of DNA. Remarkably, soft regions were also seen within the connector, such that, overall, the connector displays an unexpected mechanical heterogeneity, resembling that of typical composite materials such as silk and sea shells—not only structurally but also, most likely, functionally.

Our force probe simulations revealed a large elastic regime of the connector, and showed that quite large untwisting-twisting and compression-stretching deformations are in fact reversible on very short timescales. The extent of fully reversible deformations is markedly larger than that typically seen for proteins such as immunoglobulin domains or ubiquitin (66,67), but comparable to proteins with functionally important elasticity such as importin-β (68,69) or viral capsid proteins (30,70). It will be interesting to study how mechanical properties of the connector might change in the presence of DNA.

Subsequent closer analysis of umbrella sampling trajectories revealed that the observed pronounced exposure of hydrophobic residues to solvent during deformation markedly contributes to the total free energy required to deform the connector, as also been observed for viral capsid proteins (70). Upon deformation of the connector, its hydrophobic residues core is exposed to the solvent, particularly at the intersubunit interface, which suggests that this core is a structural feature that contributes to the exceptional stiffness of the connector. Further, the interaction enthalpies between adjacent subunits were observed to weaken during the deformations, which suggests an additional, mainly enthalpic, contribution to the overall elasticity.

Our results also shine light on the four possible different packaging mechanisms that have been proposed (9,12,13,18–20,26,27,61). In particular, the free energy required to deform the connector to an extent required for the untwist-twist mechanism is much larger than the energy provided by ATP hydrolysis, which—combined with results from voltage-ramping experiments (27)—speak against that mechanism. Our results neither rule out nor support the first proposed mechanism, involving full rotations of the connector (18); however, this mechanism has already been ruled out experimentally (19).

The heterogeneous stiffness pattern of the connector and its elastic properties, as obtained from our simulations, are compatible with the fourth, recently proposed one-way revolution mechanism. According to this mechanism, the connector acts as a one-way valve and thereby restricts leakage of the viral DNA during and after the DNA packaging process. In particular, strong interactions are expected between flexible loops located at the inner rim of the channel formed by residues K234–R237, which can obstruct the channel similar to a check-valve mechanism, whereas the stiff regions of the connector provide a solid anchor for these flexible loops. We note that, because no structural information is available on the bound DNA, it has not been included in our simulations and, therefore, our results neither rule out nor support the DNA revolution mechanism that is proposed in the one-way revolution model. Similarly, our results neither rule out nor support the DNA rotation-roll mechanism that is proposed in the third proposed model, the push-roll model. To that aim, it remains to be tested how the connector interacts with the pRNA and ATPase to facilitate DNA translocation and how the viral DNA moves into the procapsid during the packaging process.

SUPPORTING MATERIAL

Six tables, 17 figures, two movies and References (71–80) are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(14)00126-X.

This work has been supported by the European Commission under grant Nos. 211800 SBMPS and FP7/2007-2013.

REFERENCES

1. Meijer, W. J., J. A. Horcajadas, and M. Salas. 2001. ϕ29 family of phages. *Microbiol. Mol. Biol. Rev.* 65:261–287.
2. Guo, P. 2005. Bacterial virus ϕ29 DNA-packaging motor and its potential applications in gene therapy and nanotechnology. *Methods Mol. Biol.* 300:285–324.
3. Guo, P. X., S. Bailey, …, D. Anderson. 1987. Characterization of the small RNA of the bacteriophage ϕ29 DNA packaging machine. *Nucleic Acids Res.* 15:7081–7090.
4. Guo, P. X., S. Erickson, and D. Anderson. 1987. A small viral RNA is required for in vitro packaging of bacteriophage ϕ29 DNA. *Science.* 236:690–694.
5. Lee, T. J., and P. Guo. 2006. Interaction of gp16 with pRNA and DNA for genome packaging by the motor of bacterial virus ϕ29. *J. Mol. Biol.* 356:589–599.
6. Valpuesta, J. M., and J. L. Carrascosa. 1994. Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. *Q. Rev. Biophys.* 27:107–155.
7. Lawson, C. L., M. L. Baker, …, W. Chiu. 2011. EMDataBank.org: unified data resource for CryoEM. *Nucleic Acids Res.* 39 (Database issue):D456–D464.
8. Morais, M. C., K. H. Choi, …, M. G. Rossmann. 2005. Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of ϕ29. *Mol. Cell.* 18:149–159.
9. Simpson, A. A., Y. Tao, …, M. G. Rossmann. 2000. Structure of the bacteriophage ϕ29 DNA packaging motor. *Nature.* 408:745–750.
10. Smith, D. E., S. J. Tans, …, C. Bustamante. 2001. The bacteriophage straight ϕ29 portal motor can package DNA against a large internal force. *Nature.* 413:748–752.
11. Hoeprich, S., Q. Zhou, …, P. Guo. 2003. Bacterial virus ϕ29 pRNA as a hammerhead ribozyme escort to destroy hepatitis B virus. *Gene Ther.* 10:1258–1267.
12. Yu, J., J. Moffitt, …, G. Oster. 2010. Mechanochemistry of a viral DNA packaging motor. *J. Mol. Biol.* 400:186–203.
13. Moffitt, J. R., Y. R. Chemla, …, C. Bustamante. 2009. Intersubunit coordination in a homomeric ring ATPase. *Nature.* 457:446–450.

14. Rickgauer, J. P., D. N. Fuller, …, D. E. Smith. 2008. Portal motor velocity and internal force resisting viral DNA packaging in bacteriophage ϕ29. *Biophys. J.* 94:159–167.
15. Guasch, A., J. Pous, …, M. Coll. 2002. Detailed architecture of a DNA translocating machine: the high-resolution structure of the bacteriophage ϕ29 connector particle. *J. Mol. Biol.* 315:663–676.
16. Carrasco, C., A. Luque, …, P. J. de Pablo. 2011. Built-in mechanical stress in viral shells. *Biophys. J.* 100:1100–1108.
17. Ivanovska, I. L., P. J. de Pablo, …, G. J. Wuite. 2004. Bacteriophage capsids: tough nanoshells with complex elastic properties. *Proc. Natl. Acad. Sci. USA.* 101:7600–7605.
18. Hendrix, R. W. 1978. Symmetry mismatch and DNA packaging in large bacteriophages. *Proc. Natl. Acad. Sci. USA.* 75:4779–4783.
19. Hugel, T., J. Michaelis, …, C. Bustamante. 2007. Experimental test of connector rotation during DNA packaging into bacteriophage ϕ29 capsids. *PLoS Biol.* 5:e59.
20. Aathavan, K., A. T. Politzer, …, C. Bustamante. 2009. Substrate interactions and promiscuity in a viral DNA packaging motor. *Nature.* 461:669–673.
21. Zhao, Z., E. Khisamutdinov, …, P. Guo. 2013. Mechanism of one-way traffic of hexameric ϕ29 DNA packaging motor with four electropositive relaying layers facilitating antiparallel revolution. *ACS Nano.* 7:4082–4092.
22. Schwartz, C., and P. Guo. 2013. Ultrastable pRNA hexameric ring gearing hexameric ϕ29 DNA-packaging motor by revolving without rotating and coiling. *Curr. Opin. Biotechnol.* 24:581–590.
23. Schwartz, C., G. M. De Donatis, …, P. Guo. 2013. Revolution rather than rotation of AAA+ hexameric ϕ29 nanomotor for viral dsDNA packaging without coiling. *Virology.* 443:28–39.
24. Schwartz, C., G. M. De Donatis, …, P. Guo. 2013. The ATPase of the ϕ29 DNA packaging motor is a member of the hexameric AAA+ superfamily. *Virology.* 443:20–27.
25. Zhang, H., C. Schwartz, …, P. Guo. 2012. Chapt. 9. "Push through one-way valve" mechanism of viral DNA packaging. *In* Advances in Virus Research. Ł. Małgorzata and S. Wacław, editors. Academic Press, New York, pp. 415–465.
26. Fang, H., P. Jing, …, P. Guo. 2012. Role of channel lysines and the "push through a one-way valve" mechanism of the viral DNA packaging motor. *Biophys. J.* 102:127–135.
27. Jing, P., F. Haque, …, P. Guo. 2010. One-way traffic of a viral motor channel for double-stranded DNA translocation. *Nano Lett.* 10:3620–3627.
28. Grubmüller, H., B. Heymann, and P. Tavan. 1996. Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science.* 271:997–999.
29. Kutzner, C., J. Czub, and H. Grubmüller. 2011. Keep it flexible: driving macromolecular rotary motions in atomistic simulations with GROMACS. *J. Chem. Theory Comput.* 7:1381–1393.
30. Zink, M., and H. Grubmüller. 2009. Mechanical properties of the icosahedral shell of Southern bean mosaic virus: a molecular dynamics study. *Biophys. J.* 96:1350–1363.
31. Simpson, A. A., P. G. Leiman, …, M. G. Rossmann. 2001. Structure determination of the head-tail connector of bacteriophage ϕ29. *Acta Crystallogr. D Biol. Crystallogr.* 57:1260–1269.
32. Hess, B., C. Kutzner, …, E. Lindahl. 2008. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4:435–447.
33. Hornak, V., R. Abel, …, C. Simmerling. 2006. Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins.* 65:712–725.
34. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
35. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald—an *N*.log(*N*) method for Ewald sums in large systems. *j. Chem. Phys.* 98:10089–10092.
36. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126:014101.
37. Berendsen, H. J. C., J. P. M. Postma, …, J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.
38. Nose, S., and M. L. Klein. 1983. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* 50:1055–1076.
39. Hess, B. 2008. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* 4:116–122.
40. Hess, B., H. Bekker, …, J. G. E. M. Fraaije. 1997. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18:1463–1472.
41. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:27–38.
42. Hub, J. S., B. L. de Groot, and D. van der Spoel. 2010. G_WHAM-A free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* 6: 3713–3720.
43. Kumar, S., D. Bouzida, …, J. M. Rosenberg. 1992. The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. *J. Comput. Chem.* 13:1011–1021.
44. Roux, B. 1995. The calculation of the potential of mean force using computer-simulations. *Comput. Phys. Commun.* 91:275–282.
45. Gere, J. M. 2004. Mechanics of Materials. Brooks/Cole-Thomson/ Cengage Learning, Stamford, CT, pp. 78–83.
46. Eisenhaber, F., P. Lijnzaad, …, M. Scharf. 1995. The double cubic lattice method—efficient approaches to numerical-integration of surface-area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* 16:273–284.
47. Czub, J., and H. Grubmüller. 2011. Torsional elasticity and energetics of F_1-ATPase. *Proc. Natl. Acad. Sci. USA.* 108:7408–7413.
48. Wenger, M. P. E., L. Bozec, …, P. Mesquida. 2007. Mechanical properties of collagen fibrils. *Biophys. J.* 93:1255–1263.
49. Perez-Rigueiro, J., C. Viney, …, M. Elices. 2000. Mechanical properties of single-brin silkworm silk. *J. Appl. Polym. Sci.* 75:1270–1277.
50. Cunniff, P. M., S. A. Fossey, …, D. L. Vezie. 1994. Mechanical and thermal properties of dragline silk from the spider *Nephila clavipes*. *Polym. Adv. Technol.* 5:401–410.
51. Agirrezabala, X., J. Martín-Benito, …, J. L. Carrascosa. 2005. Structure of the connector of bacteriophage T7 at 8Å resolution: structural homologies of a basic component of a DNA translocating machinery. *J. Mol. Biol.* 347:895–902.
52. Cuervo, A., and J. L. Carrascosa. 2012. Viral connectors for DNA encapsulation. *Curr. Opin. Biotechnol.* 23:529–536.
53. Cuervo, A., M. C. Vaney, …, L. Oliveira. 2007. Structural rearrangements between portal protein subunits are essential for viral DNA translocation. *J. Biol. Chem.* 282:18907–18913.
54. Lebedev, A. A., M. H. Krause, …, A. A. Antson. 2007. Structural framework for DNA translocation via the viral portal protein. *EMBO J.* 26:1984–1994.
55. Lhuillier, S., M. Gallopin, …, S. Zinn-Justin. 2009. Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating. *Proc. Natl. Acad. Sci. USA.* 106:8507–8512.
56. Olia, A. S., P. E. Prevelige, Jr., …, G. Cingolani. 2011. Three-dimensional structure of a viral genome-delivery portal vertex. *Nat. Struct. Mol. Biol.* 18:597–603.
57. Veesler, D., and C. Cambillau. 2011. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiol. Mol. Biol. Rev.* 75:423–433.
58. Römer, L., and T. Scheibel. 2008. The elaborate structure of spider silk: structure and function of a natural high performance fiber. *Prion.* 2:154–161.
59. Addadi, L., and S. Weiner. 1992. Control and design principles in biological mineralization. *Angew. Chem. Int. Ed. Engl.* 31:153–169.

60. Luz, G. M., and J. F. Mano. 2010. Mineralized structures in nature: examples and inspirations for the design of new composite materials and biomaterials. *Compos. Sci. Technol.* 70:1777–1788.

61. Grimes, S., S. Ma, …, P. J. Jardine. 2011. Role of ϕ29 connector channel loops in late-stage DNA packaging. *J. Mol. Biol.* 410:50–59.

62. Honig, B., and A. S. Yang. 1995. Free energy balance in protein folding. *Adv. Protein Chem.* 46:27–58.

63. Eisenberg, D., and A. D. McLachlan. 1986. Solvation energy in protein folding and binding. *Nature.* 319:199–203.

64. Dill, K. A. 1990. Dominant forces in protein folding. *Biochemistry.* 29:7133–7155.

65. Eisenhaber, F. 1996. Hydrophobic regions on protein surfaces. Derivation of the solvation energy from their area distribution in crystallographic protein structures. *Protein Sci.* 5:1676–1686.

66. Gräter, F., and H. Grubmüller. 2007. Fluctuations of primary ubiquitin folding intermediates in a force clamp. *J. Struct. Biol.* 157:557–569.

67. Fernandez, J. M., and H. Li. 2004. Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science.* 303:1674–1678.

68. Kappel, C., U. Zachariae, …, H. Grubmüller. 2010. An unusual hydrophobic core confers extreme flexibility to HEAT repeat proteins. *Biophys. J.* 99:1596–1603.

69. Kappel, C., N. Dölker, …, H. Grubmüller. 2012. Universal relaxation governs the nonequilibrium elasticity of biomolecules. *Phys. Rev. Lett.* 109:118304.

70. Zink, M., and H. Grubmüller. 2010. Primary changes of the mechanical properties of Southern Bean Mosaic Virus upon calcium removal. *Biophys. J.* 98:687–695.

71. Amadei, A., A. B. Linssen, and H. J. Berendsen. 1993. Essential dynamics of proteins. *Proteins.* 17:412–425.

72. Berendsen, H. J., and S. Hayward. 2000. Collective protein dynamics in relation to function. *Curr. Opin. Struct. Biol.* 10:165–169.

73. Eswar, N., B. Webb, …, A. Sali. 2007. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein. Sci.* Chapter 2:Unit 2.9. http://dx.doi.org/10.1002/0471140864.ps0209s50.

74. Fernandez-Fuentes, N., J. Zhai, and A. Fiser. 2006. ArchPRED: a template based loop structure prediction server. *Nucleic Acids Res.* 34 (Web Server issue):W173–W176.

75. García, A. E. 1992. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* 68:2696–2699.

76. Hayward, S., A. Kitao, …, N. Go. 1993. Effect of solvent on collective motions in globular protein. *J. Mol. Biol.* 234:1207–1217.

77. Hess, B. 2000. Similarities between principal components of protein dynamics and random diffusion. *Phys. Rev. E.* 62 (6 Pt B):8438–8448.

78. Kitao, A., F. Hirata, and N. Gō. 1991. The effects of solvent on the conformation and the collective motions of protein: normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum. *Chem. Phys.* 158:447–472.

79. Macke Thomas, J., and A. Case David. 1997. Modeling unusual nucleic acid structures. *In* Molecular Modeling of Nucleic Acids American Chemical Society, Washington, DC, pp. 379–393.

80. Pérez, A., I. Marchán, …, M. Orozco. 2007. Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophys. J.* 92:3817–3829.

Elastic properties and heterogeneous stiffness of the Phi29 motor connector channel

Rajendra Kumar and Helmut Grubmüller
Max Planck Institute for Biophysical Chemistry,
Department of Theoretical and Computational Biophysics,
Göttingen, 37077, Germany

Supporting Material

Contents

Sections

Tables

Figures

References

1. Description of connector loops modeling and refinement

The connector loops A230-S244 were unresolved in the X-ray crystal structure (Fig. S1A) and therefore modeled as follows. First, an initial structure model was obtained for one of the twelve symmetry-related missing loops using the ArchPred server (1). This seed structure was then replicated (according to the 12-fold symmetry of the connector) and manually integrated within all other remaining sub-units after aligning the seed structure (see Fig. S1B). In addition, four of the twelve subunits of the connector lacked residues (Q166-L169) in the bottom region. These residues were modeled using Modeller program (2) by taking a template subunit in which these residues are present.

Also, in a structure of the connector in complex with DNA, the diversity in DNA-loop interactions will cause the 12 connector loop structures to slightly differ from each other because of the lack of 12-fold symmetry in the DNA helical grooves. To model this structural heterogeneity, all 12 connector loops were refined by a combined simulated annealing and molecular dynamics (SA/MD) protocol as described in following.

First, the viral DNA of the Phi29 bacteriophage gp10 gene (NCBI Reference Sequence: NC_011048.1) was modeled into the connector channel center. The first 60 nucleotides (ATG GCA CGT AAA CGC AGT AAC ACA TAC CGA TCT ATC AAT GAG ATA CAG CGT CAA AAA CGG) were modeled as B-DNA strand using the Nucleic Acids Builder NAB (3) and is depicted as *yellow* and *blue* ring (top view) in Fig. S1B.

Next, the initial molecular system was prepared for the Simulated Annealing Molecular Dynamics (SA MD) simulations: The connector-DNA complex was set in the center of a dodecahedron periodic box, solvated with 241,067 water molecules, and neutralized by addition of a sufficient number of Sodium ions (in total 202). The system was energy-minimized and subsequently heated during a NVT simulation of 500 ps with a 1 fs time step. Pressure was equilibrated during a subsequent NPT simulation of 1 ns with a 2 fs time step. In both simulations, the positions of all heavy atoms were restrained by an harmonic force constant of 1000 kJ/(mol nm^2). 40 SA cycles of 252 ps length were performed for 10 ns during the SA MD simulations. In each SA cycle, the loops were heated up in two steps to a temperature of 1000 K and subsequently cooled down in seven steps to 300 K (Table S1). Only the loops were free to

move during the SA MD simulations, whereas other heavy atoms of the connector-DNA complex were restrained at the starting position by a force constant of 1000 kJ/(mol nm^2) to preserve the structure of the complex at high temperatures.

The SA MD and below described cooling simulations were performed using the GROMACS 4.0.7 package (4), in which AMBER ff99SB (5) and parmbsc0 (6) force fields were used for connector and DNA, respectively. The temperature was regulated during heating and cooling by Berendsen temperature coupling, whereas the pressure was maintained at 1 atm by the Berendsen pressure coupling (7). Long range electrostatic interactions were computed by applying the PME method with a grid spacing of 1.2 Å and a 4^{th} order of cubic interpolation (8). Short range non-bonded interactions were computed for the atom pairs within a distance of 10 Å. The temperature used in SA MD simulations flattened the energy landscape and facilitated random changes in loop conformations that were collected during 10 ns. To assess the convergence of loop conformations during MD simulations, principal component analyses (PCA) were performed (9-13) on the last 6 ns of each SA MD trajectory. The first two principal components (PC) were used because these have the largest variance and hence show the largest conformation changes.

Figure S2A shows the two-dimensional plane of the first PC with respect to the second PC. The semi-circle shape of the plane reveals randomness in conformations along the first two PCs (14). The obtained high temperature conformations (red dots in Fig. S2A) are most likely located at a high energy region on the free energy landscape and were not used for MD simulations. However, cooling of these conformations is expected to allow achieving nearest free energy minimum loop conformations. To this aim, five different high temperature conformations were chosen from the projection planes of PC1 and PC2 (A, B, C, D and E in Fig. S2A). These conformations were cooled from 1000 K down to 300 K in 6 ns and subsequently equilibrated for 1 ns at 300 K in a NPT simulation (cooling simulations). Finally, five clusters of structures were obtained (A′, B′, C′, D′ and E′ in Fig. S2A) at 300 K. These were compared by computing root mean square deviations (RMSD) of C-alpha atoms of the loops with respect to the initial structure model (five different blue symbols in Fig. S2B) and found to be within 0.2 nm difference. The deviation is not significant for these highly flexible loops. Consequently, the central structure from this cluster was used for further studies. The final loop conformations and DNA-loop interactions are illustrated in Fig. S2C.

2. Structural and conformational equilibration of the connector

We monitored the structural stability of the connector during equilibrium MD simulations via root mean square deviation (RMSD) calculated for C-alpha atoms with respect to the crystal structure (Fig. S7). After an initial sharp rise during the first 10 ns, the RMSD-values stabilize at about 0.25 nm during the subsequent 190 ns. Averaged over the last 190 ns of the equilibration phase, a twist angle of $\theta = 77.7°$ and a length of L = 4.89 nm was obtained, with a standard deviation of 0.4° and 0.04 nm, respectively (Fig. 2A). Because of the RMSD drift observed during the first 100 ns, average values were also computed for the last 100 ns, and a similar twist angle ($\theta = 77.6°$) and length (L = 4.91 nm) were obtained. Whereas the connector length remained at that of the crystal structure (L = 4.91 nm), the twist angle increased by ca. 3.5° with respect to the crystal structure ($\theta = 74.2°$) during the first 10 ns. Closer inspection showed that the deviation was actually due to an increased tilt of the helices as depicted in Fig. S8, and in particular those helices which are in contact with symmetry mates in the crystal structure. This increased tilt of those helices slightly shifted the center of mass of the bottom region of each subunit that was used to calculate the twist angle (see Figs. 1C and D). We therefore assume that this tilt in the crystal structure is due to crystal contacts between these alpha helices and the upper region. The structures of the middle and upper regions are largely unaffected. Given the size of the connector, we consider the observed RMSD of ca. 0.25 nm and the underlying structural changes during equilibration rather small. Because the overall structure and the RMSD remained unchanged during the last 100 ns, this part of the trajectory was used for further analysis.

3. Elastic properties determined from force-probe simulations

3.1. Methods

3.1.1. Force probe MD simulations

To study the mechanical properties of the connector by force probe simulations, the structure was subjected to external forces via two structural descriptors (see Fig. 1D), the twist angle θ_m which describes the twisting-untwisting motion of the middle region of the connector, and its length L_m, which quantifies the compression-stretching motion of the middle region. The upper and lower disk in Fig. 1E represents Cα atom from the upper- and lower-terminus of MH1 and MH2 helices of each subunit, respectively. We will refer to the atoms of the lower and upper disk as pull or rotational group and reference group in the subsequent sections, respectively. Rotation between two disks defines the twisting-untwisting motion while motion of disks along channel axis shows the compression-stretching motion.

Six force-probe MD simulations (FP-1, FP-2, FP-3, FP-4, FP-5 and FP-6) were performed to investigate the twisting-untwisting and compression-stretching motions of the connector beyond the equilibrium fluctuations (see details in Table. S2). All force probe simulations started from the equilibration trajectory after 25 ns. To drive the compression and stretching motion, the center of mass of the pull group (lower disk in Fig. 1D) was moved away (stretching) and towards (compression) the center of mass of the reference group (upper disk in Fig. 1D). To drive the untwisting and twisting motions, the atoms of the rotational group (lower disk in Fig. 1D) were forced to move along circles centered at the z-axis by applying a torque potential in both anti-clock wise (untwist) and clock wise (twist) direction, respectively (grey arrow in Fig. 1D). At the same time, the center of mass of the reference group was kept fixed by a harmonic restraint of 1000 kJ/(mol nm^2). As defined in Ref. (15) the variant *rm2-pf* was chosen as rotational potential, which allowed the radial motion of the rotational group atoms (15). All simulations were carried out using a modified version of GROMACS in which rotational pulling was implemented (15). The total simulation time was ~160 ns.

3.1.2. Relaxation simulations

To study within which regime the observed deformation is reversible and hence elastic, two intermediate structures of 5.1 and 5.3 nm length, were chosen from the FP-6 simulation (see

Figs. S3B and C) for subsequent relaxation simulations (Relax-1 and Relax-2, respectively). To equilibrate the structure at the starting twist angle and length, all atoms of pull and reference group were restrained by force constant of 1000 kJ/(mol nm^2) for the first 4.4 ns. The force was then removed, and the connector was allowed to relax freely, during which period the approach to its equilibrium length and twist angle was monitored. Furthermore, to check the structural reversibility of the compressed connector, an intermediate representative structure of the connector with a middle region's length L_m=2.7 nm was taken from the 7.52 ns of the FP-5 simulation (see Fig. S3A) for a subsequent relaxation simulation (Relax-3). At this specific snapshot of the FP-5 trajectory, the length L_m=2.699 nm was nearest to the required value of L_m=2.7 nm. In this relaxation simulation, the connector was kept at its enforced twist angle and length for 3.3 ns before the pulling and restraint potentials were removed in order to allow relaxation towards its equilibrium length. An additional relaxation simulation (Relax-4) was performed from FP-6 simulation to obtain relaxed structures of the connector which was later used for the umbrella sampling simulations.

3.2. Results and discussion

In the first set of simulations (FP-1 and FP-2), the connector was compressed and stretched by applying linear pulling forces as described in the methods section (Fig. 1D and Table S2). The required forces as well as the resulting connector twist angle θ as a function of connector length L was observed (Fig. S10A, red symbols). Indeed, for the range accessible to equilibrium fluctuations (black ellipse) as well as for larger compressions beyond the equilibrium fluctuations (left branch of the curve), the connector's twist angle changes by 2°/nm (Table S6), and is herewith similar to the one determined above (dashed line in Fig. S10A). Up to the maximum exerted force of 5000 pN, and at a compression by ca. 4%, the force extension curve (Fig. S10B, red line) is linear, indicating Hookean behavior within this range.

For stretching beyond the equilibrium range (outside of ellipse in Fig. S10A), an unexpectedly large untwisting is observed, by 13.5°/nm up to 5.1 nm length, with a continued linear (Hookean) force. Beyond that critical force, the structure started to break down, which is also reflected in the decreased slope of the force extension curve and may indicate non-elastic behavior. Interestingly, the coupling between extension and untwisting remains linear also beyond the

critical force. Within the Hookean regime, a spring constant of ~27200 pN/nm (red line in Fig. S10B) is observed.

The large untwisting motion beyond L=5.05 nm is unexpected because non-equilibrium relaxation would likely make the untwisting lag behind its equilibrium pathway (dashed curve in Fig. S10A), in contrast to the observed increased slope. Note, however, that this additional untwisting is solely due to structural re-arrangements of the bottom region of the connector, which is not subjected to the pulling force (cf. inset in Fig. S10A). Indeed, the untwisting motion of the middle region alone displays linear coupling up to full extension (see red line Fig. S11A). It remains to be tested whether or not this linear pathway resembles the equilibrium path.

In a second set of FP simulations (FP-3 and FP-4), the connector was twisted and untwisted by applying a torque (see Methods Section), and the required torque as well as the length change as a function of twist angle was monitored (blue symbols, Figs. S10A and C). Beyond the equilibrium fluctuations, the connector twist angle θ changed linearly with length L by 27.9°/nm, markedly larger than that calculated from equilibrium fluctuations (Table S6). By exceeding an untwisting of 70^{o} the structure started to break down. The torque-angle curve (blue line, Fig. S10C) is Hookean between 75^{o} and 78^{o}, with a torsional spring constant of ~3500 pN nm/^{o}.

In all of the above simulations (FP-1 to FP-4), the obtained stretching and torsional spring constants were larger than those derived from equilibrium simulations. We assume this discrepancies to be caused by non-equilibrium effects in force probe simulations, which are absent in equilibrium simulations. Specifically, during the relatively short time of about ten nanoseconds for each simulation, relaxation is likely to be incomplete for several of the internal degrees of freedom of the connector, as previously described and quantified in terms of a two-dimensional time-dependent transition state theory also for the connector (16). In this scenario, as a result of this incomplete relaxation, the twist angle θ should 'lag behind' (i.e., larger θ values) a fully equilibrated pathway along the minimum ('valley') of the underlying untwisting-stretching free energy landscape. Vice versa, when untwisting the connector (blue symbols, Fig. S10A), the length expansion is expected to lag behind, which explains why this untwisting-stretching curve falls below that of the stretching simulations. Accordingly, the equilibrium pathway is expected in between these two extremes.

We tested this hypothesis by performing several relaxation simulations, starting from different points along the stretching and untwisting simulations, in which the pulling force or torque, respectively, was removed. If relaxation effects dominate the observed differences between stretching and untwisting, one would expect the relaxation trajectories to deviate from the force probe trajectories and to converge towards their common equilibrium pathway. Indeed, Fig. S12 clearly shows such convergence. Both twist angle and length changed rapidly towards an intermediate pathway, and almost reach an equilibrium region within ca. 10 to 20 ns. Notably, the slope of the converged line is larger than that of the equilibrium coupling, which suggests that the coupling mechanism between length change and untwisting outside the equilibrium region differs from that in thermal equilibrium.

To approximate this equilibrium path, which will also be used for subsequent deformation free energy calculations, two further force probe simulations (FP-5 and FP-6) were performed in which both stretching forces and torque were applied simultaneously. Because the above relaxation simulations suggest that the equilibrium path is located approximately right in between the paths probed by simulations FP 1-2 and 3-4, respectively, the same ratio between pulling and rotation speed was chosen here. To allow for more complete relaxation, a ten times reduced pulling force and torque rate (Table S2) was used.

The obtained length and twist angle changes are shown in Fig. S10A (green symbol). As expected, the resulting deformation pathway falls right in between the paths obtained in the stretching-only, untwisting-only and relaxation simulations. As an example, movie S1 shows an animation of trajectory FP-6. Linear coupling is observed, with a coupling coefficient of 18°/nm throughout the whole stretching/untwisting process (green line in Fig. S10A). Also in these simulations, structural break-down occurred at L = 5.15 nm (shown in movie S2). No linear coupling between compressions and twisting is observed; rather, the twist angle saturates at 79.5° during compression. Within the regime of equilibrium fluctuations, Hookean behavior of both the stretching and untwist is observed (Figs. S10B and C, green symbols), with stretching and torsional spring constants similar (Table S6) to those obtained from simulations FP-1/2 and FP-3/4, respectively. For extensions exceeding 5.0 nm length and twist angle below 76°, respectively, non-Hookean behavior sets in. Remarkably, the change of the connector's

mechanical properties is very abrupt at this critical deformation, with no additional torque being required for further large connector untwisting.

To test whether the observed non-Hookean behavior also implies non-elastic behavior for both the untwisting-stretching motion as well as the observed structural break-down, we have carried out further relaxation simulations (see methods section), starting from non-equilibrium conformations extracted from simulations FP-5 and FP-6 at lengths L of 4.75, 5.1, and 5.3 nm, respectively (the corresponding structures are shown in Fig. S3). Figure S13A shows the obtained relaxation motions in terms of both length L and twist angle θ during the final part of the restrained simulation (left of the dashed line) and after removal of the restraining potential (right of dashed line). The equilibrium values are regained in all three simulations. Even the extreme non-Hookean deformation of up to L = 5.3 nm length approaches the equilibrium value after 100 ns, despite the structural deformations shown in Fig. S3C (shaded box) and Movie S2, which demonstrates that the connector is able to recover its equilibrium structure even after structural break down. As can be seen in Figs. S13B and C, all three relaxation trajectories remain close to the untwisting/stretching paths of simulations FP-5 and FP-6, which underscores that these simulations remained close to the 'valley' of minimal perturbation.

Overall, deformations within the range of L = 4.70-5.3 nm and θ = 70°-79.5° seem to be fully elastic, and the twisting-untwisting motion is linearly coupled to the compression-stretching motion within this elastic range. The obtained coupling of 18°/nm is remarkably close to the coupling of 17.6°/nm required in the untwist-twist DNA packaging model proposed by Simpson et al. (17). We therefore asked next if also the other elastic and energetic properties of the connector are compatible with this model.

Tables

Table S1: Heating-cooling cycles performed in SA MD 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)
300	20	20
300-600	2	22
600	22	44
600-1000	2	46
1000	50	96
1000-800	2	98
800	20	118
800-700	2	120
700	20	140
700-600	2	142
600	20	162
600-500	2	164
500	20	184
500-450	2	186
450	20	206
450-400	2	208
400	20	228
400-350	2	230
350	20	250
350-300	2	252

Table S2: Force rates and torque/force constants applied in FP simulations. Corresponding simulation names are used throughout the main article and the supplementary text.

Name	Force probe simulations (along motion type)	Rate		Force constant [kJ/(mol nm^2)]	
		Rotational (°/ps)	Pulling (nm/ps)	Rotational	Pulling
FP-1	Compression*	—	0.010	—	100
FP-2	Stretching*	—	0.010	—	100
FP-3	Twisting*	0.0040	—	500	—
FP-4	Untwisting*	0.0040	—	500	—
FP-5	Twisting and compression	0.0004	0.001	500	100
FP-6	Untwisting and stretching	0.0004	0.001	500	100

*These simulations were performed three times in parallel.

Table S3: Details of compression-stretching umbrella sampling simulations. The simulation type from which the reference starting conformation for each sampling window was extracted is denoted accordingly. The frame time corresponds to the time of the reference simulation from which the reference structure was taken. The relaxation simulations Relax-1, Relax-2, Relax-3, and Relax-4 were performed using the non-equilibrium starting structures from simulations FP-5 and FP-6.

Sampling window	Reference simulation type	Frame time (ps)	Reference reaction coordinate, L_m (nm)	Force constant [kJ/(mol nm^2)]
1	Relax-3	3300	2.725	2500
2	Equilibrium	107000	2.750	500
3	Equilibrium	74200	2.800	500
4	Relax-1	16940	2.850	2500
5	Relax-1	4460	2.876	5000
6	Relax-2	16420	2.900	7500
7	Relax-2	8000	2.950	7500
8	Relax-2	5320	2.997	10000
9	Relax-2	4720	3.025	15000
10	Relax-2	4580	3.052	15000
11	Relax-4	5200	3.075	17500
12	Relax-4	5140	3.100	20000
13	Relax-4	5060	3.125	20000
14	Relax-4	5040	3.149	20000

Table S4: Details of twisting-untwisting umbrella sampling simulations. The simulation type from which the reference starting conformation for each sampling window was extracted is denoted accordingly. The frame time corresponds to the time of the reference simulation from which the reference structure was taken.

Sampling window	Reference simulation type	Frame time (ps)	Reference reaction coordinate, θ_m°	Force constant [kJ/(mol.nm^2)]
1	FP-6	24700	29.55	10000
2	FP-6	24090	30.02	7500
3	FP-6	22880	30.50	6000
4	FP-6	21570	31.01	5000
5	FP-6	18520	31.57	5000
6	FP-6	18310	32.02	2500
7	FP-6	12522	32.44	2500
8	FP-6	11362	33.13	1000
9	FP-6	8772	33.45	500
10	FP-6	7042	33.90	500
11	FP-5	1576	34.50	1000
12	FP-5	5566	35.00	6500
13	FP-5	9748	35.50	6500
14	FP-5	11030	36.00	10000

Table S5: Dimensions of the connector and its middle region used to calculate the Young's modulus of elasticity. Values were obtained from equilibrium MD simulations using Eq. 4. Notations are in accordance with labels used in the truncated cone sketch shown in Fig. S6. Given error represents the SE obtained using the block-averaging method (18).

Connector region	D_a [nm]	D_b [nm]	d_a [nm]	d_b [nm]	L [nm]
Whole connector	6.259 ± 0.015	14.720 ± 0.012	4.229 ± 0.103	5.866 ± 0.011	6.578 ± 0.025
Middle region	5.194 ± 0.044	9.268 ± 0.012	3.668 ± 0.028	7.188 ± 0.005	2.861 ± 0.002

Table S6: Mechanical properties of the whole connector and its middle region. Equilibrium mechanical properties were obtained using Eq. 2. Best-fit lines within the linear regime are depicted in Figs. S10B, S10C, S11B and S11C. Couplings between untwisting-stretching motions were obtained in a similar manner and are shown in Figs. 2, S10A and S11A.

Simulation type	Connector region	K_θ [(pN nm)/Deg^2]	K_L [pN/nm]	K_c [pN/Deg]	Coupling (Deg/nm)
Equilibrium	Whole connector	1347	3073	2715	2.0
	Middle region	3374	24109	14848	4.4
FP 1 and 2	Whole connector	—	27204	—	2.0 and 13.5
	Middle region	—	33437	—	4.2
FP 3 and 4	Whole connector	3558	—	—	27.9
	Middle region	3404	—	—	51.3
FP 5 and 6	Whole connector	3033	27747	—	18.1
	Middle region	3428	28282	—	17.7

Figures

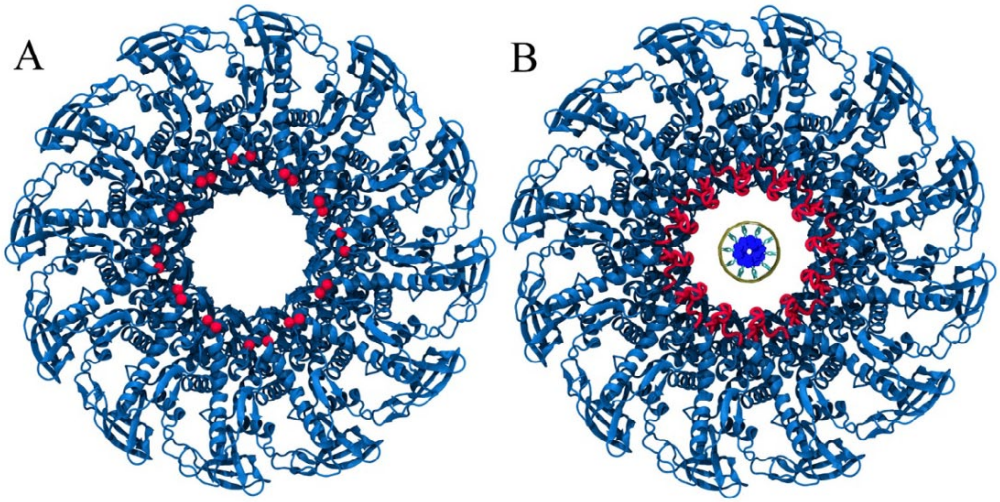

Figure S1: Crystal structure of the connector in top view. (A) Red spheres denote the location of the missing loops. (B) Modeled loops (*red*) are not in contact with the DNA (*yellow* and *blue* ring at the center of channel).

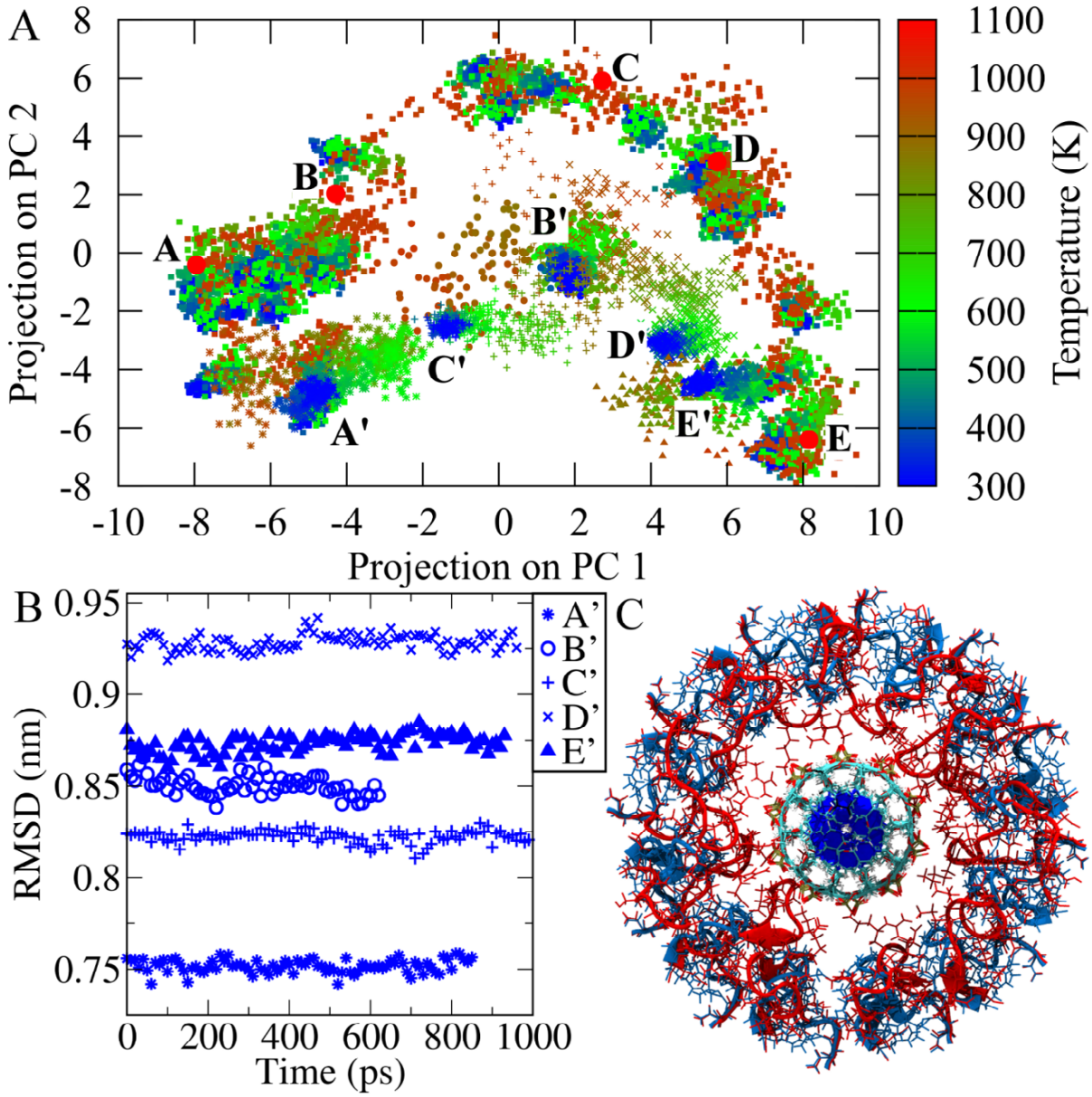

Figure S2: Refinement of the connector loops. (A) Plane of projection between first and second principal components (PC1 and PC2) obtained after performing PCA on the SA MD trajectories (*squares*). Five high temperature conformations A, B, C, D, and E (*red circles*) were selected from the projection plane and cooled down to 300 K. Dense clusters (*blue symbols*) labeled with A', B', C', D', and E' were obtained after cooling. (B) The structures of these five clusters (*blue symbols*) were compared by computing RMSDs of Cα atoms with reference to the starting loop structure models. (C) Comparison of modeled (*blue*) and refined loops (*red*) interacting with the DNA in the channel center.

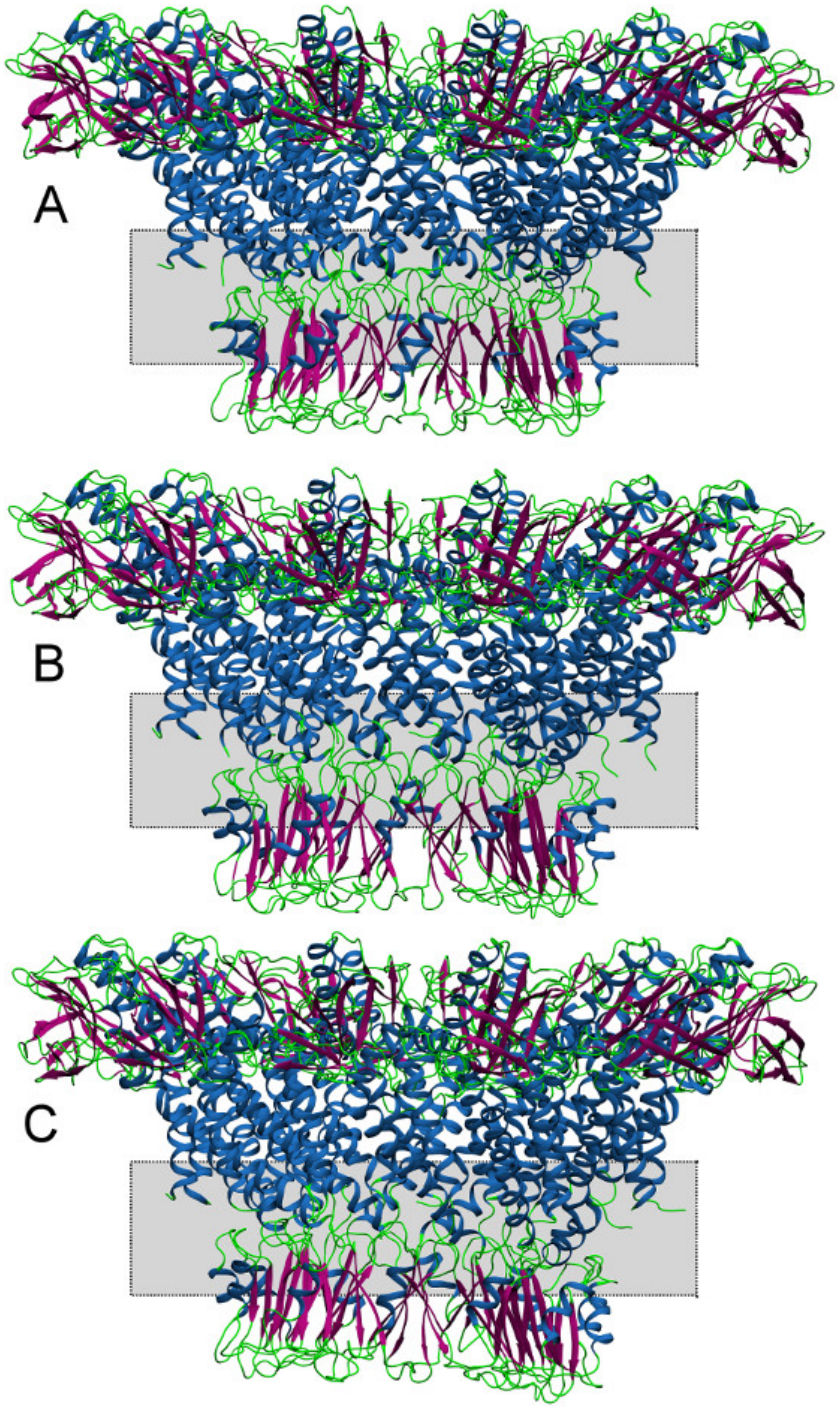

Figure S3: Starting non-equilibrium conformations used to characterize relaxation behaviors. Three non-equilibrium connector conformations of (A) 4.72 nm, (B) 5.1 nm, and (C) 5.3 nm length were selected for relaxation simulations. Region of structural deformations is shown as *shaded area*.

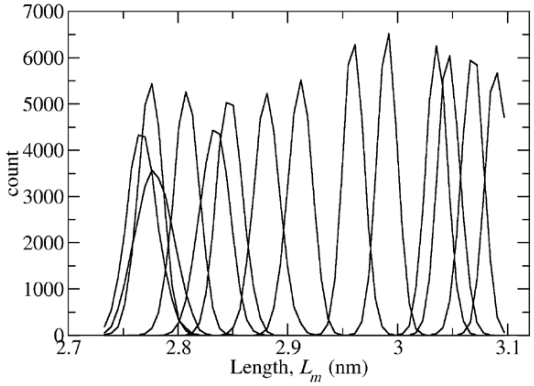

Figure S4: Collected histograms from compression-stretching umbrella sampling simulations. These histograms show the population of connector conformations along the reaction coordinate L_m during umbrella sampling simulations and were used to calculate the deformation free energy using WHAM (19-21).

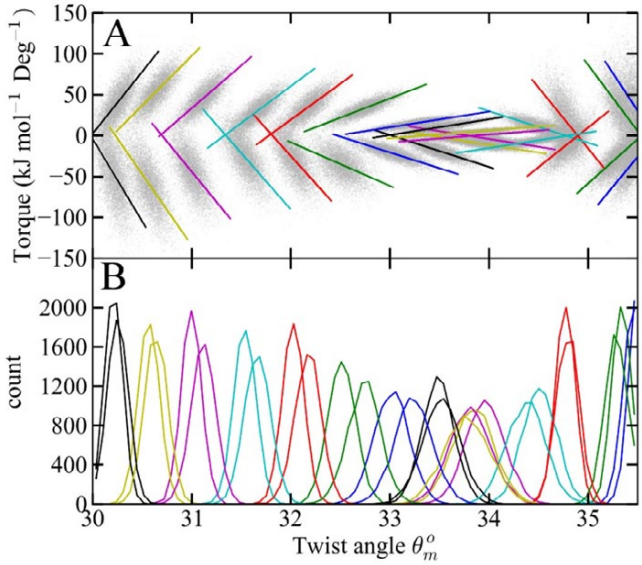

Figure S5: Torsional harmonic constants and collected histograms from twisting-untwisting umbrella sampling simulations. The group of atoms was harmonically restraint by applying a tangential harmonic force constant (kJ mol^{-1} nm^{-2}) during umbrella sampling simulations (15). Colors represent different sampling windows. The torsional harmonic force constants (kJ mol^{-1} deg^{-2}) were calculated from best line fits of the torque-twist angle values and were used to derive the deformation free energies for twisting-untwisting motions using WHAM (19-21). (A) Torque with respect to the twist angle θ_m during umbrella sampling simulations. Colored lines are the best fit lines of the respective sampling window. (B) Collected histograms showing the population of connector conformations along the reaction coordinate θ_m^o from each sampling window.

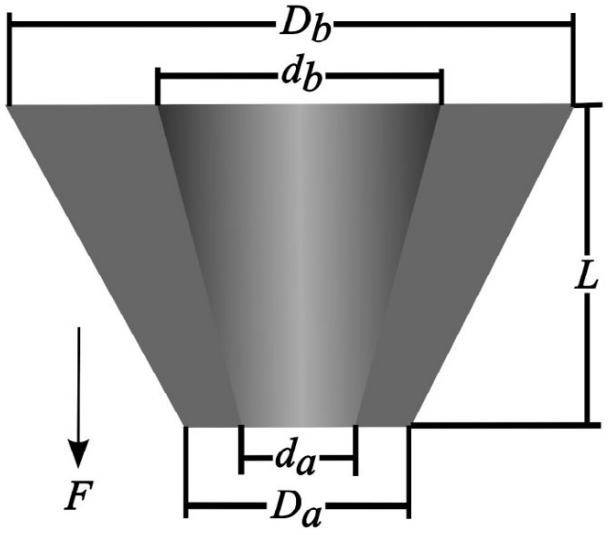

Figure S6: Geometrical characterization of the truncated hollow cone model used for calculating the Young's modulus of elasticity. The sketch shows the truncated hollow cone of length L, where D_a and d_a denote the narrow end exterior and interior diameter, respectively; D_b and d_b denote the wide end exterior and interior diameter, respectively. Equation 4 was derived on the basis of this model by assuming that the force F is acting in arrow direction.

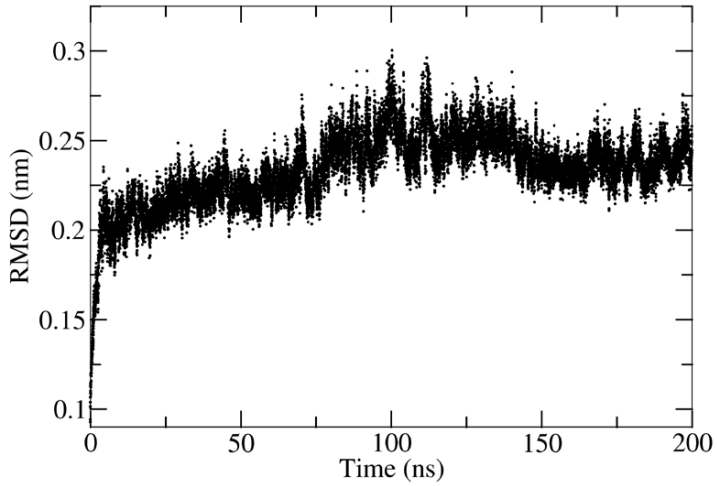

Figure S7: Root Mean Square Deviations of the connector from equilibrium simulations. RMSDs were calculated for Cα atoms of the connector with reference to the X-ray crystal structure (PDB ID: 1H5W).

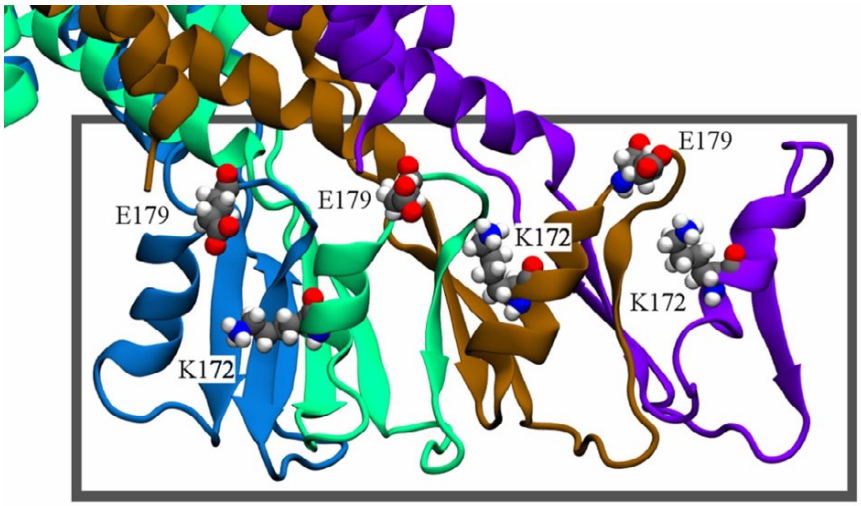

Figure S8: Deviations in twist angle during equilibrium simulations. During MD simulations, the twist angle of the connector deviated by 3.5° with respect to the crystal structure (PDB ID: 1H5W). The bottom connector region rotated with respect to the remaining part of the connector during the first 10 ns of the simulation. The grey box highlights the bottom region of four subunits that are depicted as colored cartoon representations. Residues E179 and K172 (*space-fill*) are mostly responsible for tilting of encasing α-helix.

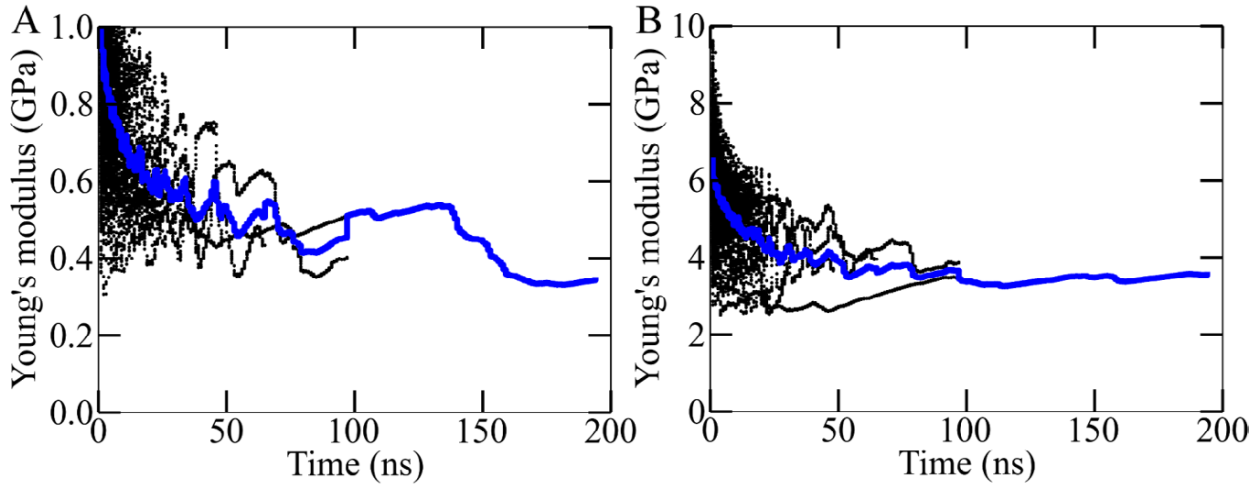

Figure S9: Convergence in the Young's modulus during the simulations. Both the time-blocks (black dots) as well as average (blue line) values of the Young's moduli with respect to the time are shown for (A) the whole connector and (B) the middle region. The moduli were calculated using Eq. 4 and, the required average of the dimensions (Fig. S6) over time and the stretching spring constant were obtained from non-overlapping time-blocks of the MD trajectory. The obtained modulus from each block is shown as a black dot. The average Young's modulus over the respective block size is shown as a blue solid line.

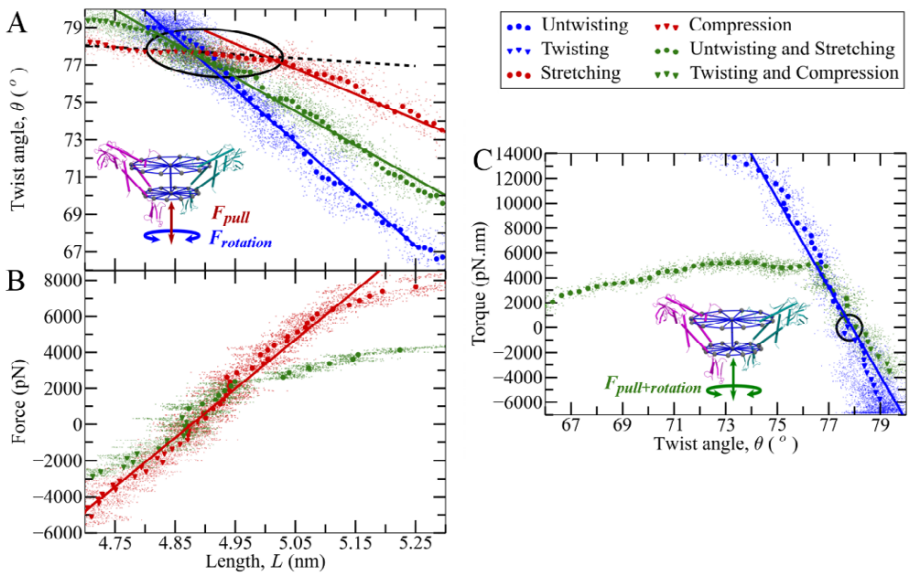

Figure S10: Elastic properties of the whole connector derived from force probe (FP) simulations; fluctuating instantaneous lengths and twist angles are shown as small dots, solid symbols represent averages over intervals in *x*-direction. The insets indicate pulling forces (F_{pull}) and torques ($F_{rotation}$) applied for different simulation sets FP-1/2 (*red arrow*), FP-3/4 (*blue arrow*), and FP-5/6 (*green arrows*) referred to in the text. (A) Enforced length change and/or twist: change in twist angle θ during enforced change of length L (*red*), length change during enforced twist/untwist (*blue*), twist angle and length change during simultaneous enforced twisting-compression and untwisting-stretching (*green*). The coupling between compression-stretching and untwisting-twisting motions during FP-1/2 (*red line*), FP-3/4 (*blue line*), and FP-5/6 (*green line*) is compared to the coupling (*black dashed line*) and the fluctuation range (*black ellipse*) obtained from equilibrium simulations. (B, C) Same data and coloring as (A), shown with applied force and torque and linear least square fits to the respective linear regime (lines). In (C), the black circle denotes the initial twist angle.

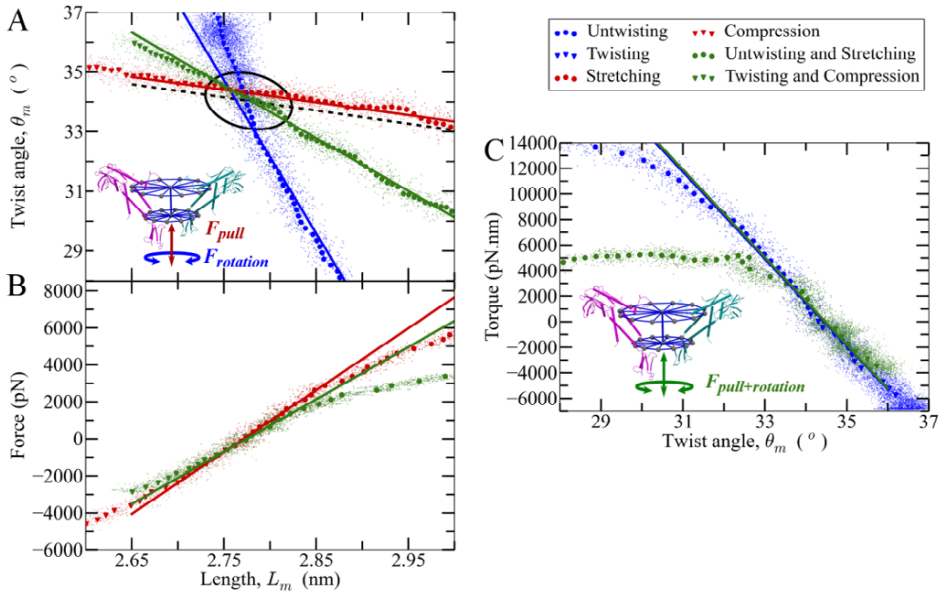

Figure S11: Elastic properties of the middle region derived from FP simulations; fluctuating instantaneous lengths and twist angles are shown as small dots, solid symbols represent averages over intervals in *x*-direction. Insets indicate applied pulling forces (F_{pull}) and torques ($F_{rotation}$) used in different simulation sets FP-1/2 (*red arrow*), FP-3/4 (*blue arrow*), and FP-5/6 (*green arrows*), as listed in Table S2. (A) Enforced length change and/or twist: change in twist angle θ during enforced change of length L (*red*), length change during enforced twist/untwist (*blue*), twist angle and length change during simultaneous enforced twisting-compression and untwisting-stretching (*green*). The coupling between compression-stretching and untwisting-twisting motions during FP-1/2 (*red line*), FP-3/4 (*blue line*), and FP-5/6 (*green line*) is compared to the coupling (*black dashed line*) and the fluctuation range (*black ellipse*) obtained from equilibrium simulations. (B, C) Same data and coloring as (A), shown with applied force or torque and linear least square fits to the respective linear regime (*lines*).

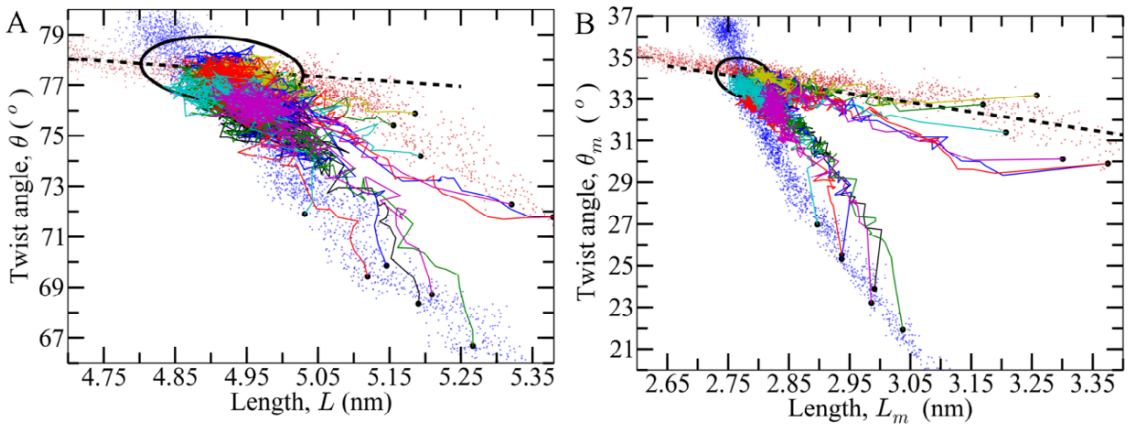

Figure S12: Convergence of relaxation paths towards an equilibrium. The comparison of relaxation paths (*colored lines*) with the obtained paths from enforced untwisting-only (*blue dots*) and stretching-only (*red dots*) motions for the (A) whole connector and (B) the middle region. Black dots denote starting non-equilibrium conformations taken from stretching simulation FP-2 (*red dots*) and untwisting simulation FP-4 (*blue dots*). To compare with equilibrium simulations, the range of fluctuations is denoted as black ellipse. Untwisting-stretching coupling is given by the slope of the best-fit line (*black dashed line*).

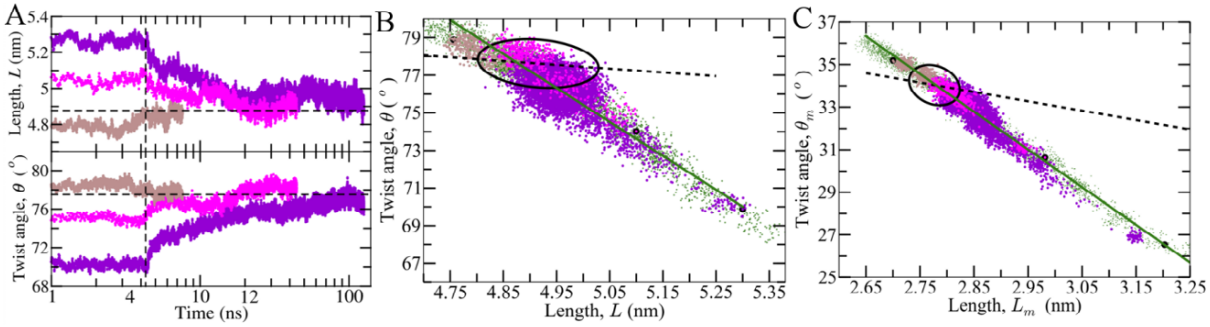

Figure S13: Reversible recovery of the equilibrium conformation. For three deformed connector structures with L=4.72, 5.1, and 5.3 nm (Figs. S3A-C), selected from simulations FP-5/6, their relaxation behavior (*purple*, *magenta,* and *light brown dots*) are illustrated. (A) The relaxation of connector's twist angle and length along the time towards the equilibrium structure (*horizontal dashed lines*) was observed for 120 ns (note the logarithmic scale). At 4.4 ns (*vertical dashed black line*) the restraint force was removed for complete relaxation. Relaxation pathways for (B) the whole connector and (C) the middle region towards the equilibrium structural fluctuations (*black ellipse*). Also, the comparisons of the relaxation paths with the obtained paths (green symbols and line) of untwisting-stretching motions from the same simulation set are shown. The coupling between compression-stretching and twisting-untwisting motions during FP-5/6 (*green line*) is compared to the coupling (*black dashed line*) obtained from the equilibrium simulations.

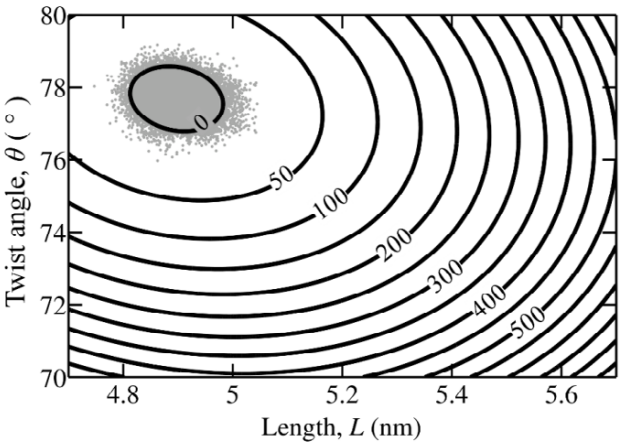

Figure S14: Free energy landscape from equilibrium fluctuations**.** The free energy of deformations for twisting-untwisting and compression-stretching motions was computed from equilibrium simulations under harmonic approximation of the energy landscape. The free energy (kJ/mol) was calculated from the probability density using the following equation $G(\theta, L) = -k_B T \ln[p(\theta, L)]$, where θ denotes the twist angle, L the length of the connector, k_B the Boltzmann constant, and T the temperature, respectively. The landscape was extrapolated beyond the equilibrium region to test the proposed untwist-twist DNA packaging mechanism.

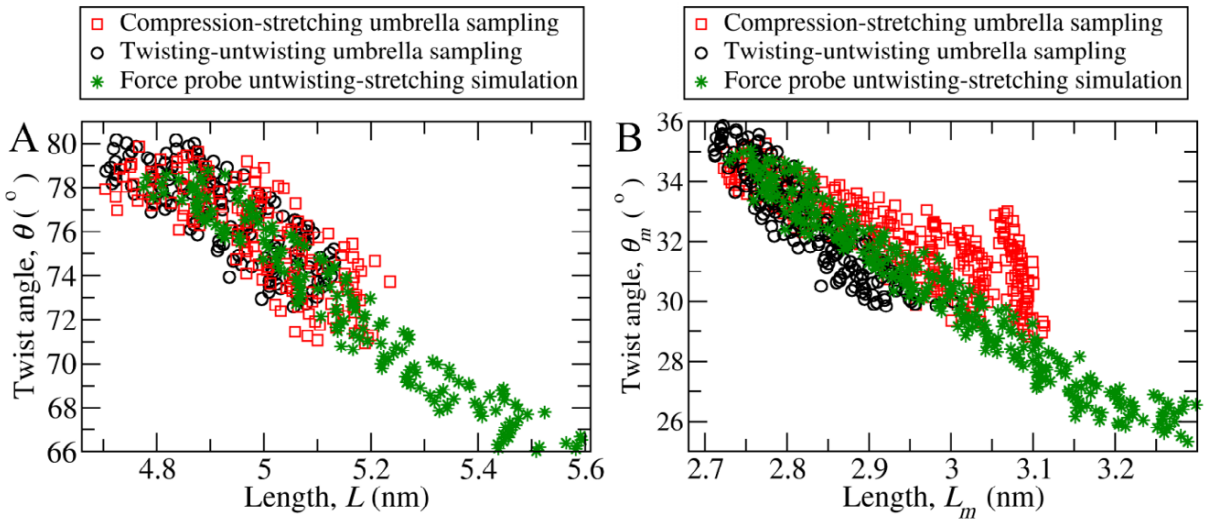

Figure S15: Convergence of deformation paths in umbrella sampling simulations. The pathways of untwisting-stretching motions were obtained from compression-stretching (*red squares*) and twisting-untwisting (*black circles*) umbrella sampling simulations and are compared to each other and to force probe simulation FP-6 (*green asterisks*) for (A) the whole connector and (B) the middle region.

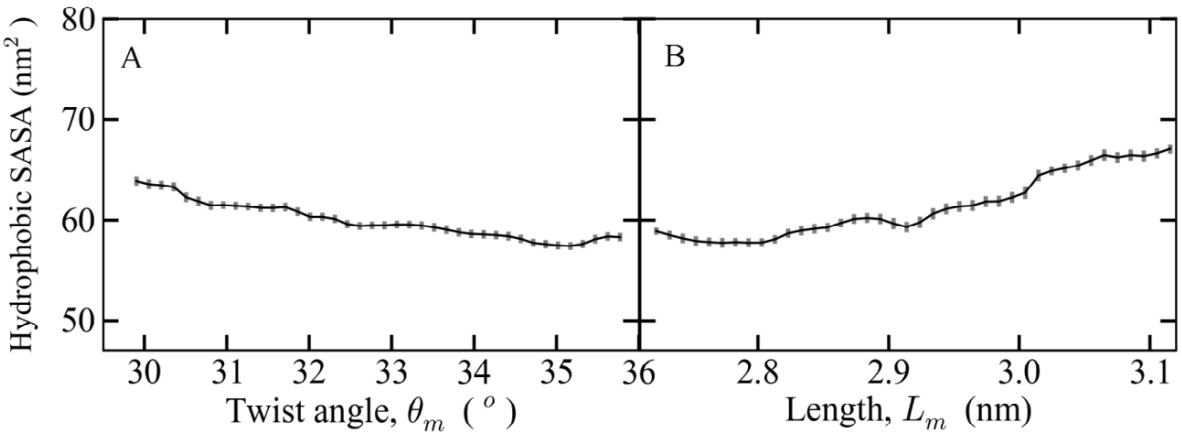

Figure S16: Deformation dependent changes in hydrophobic solvent accessible surface areas (SASAs) of the middle connector region derived from umbrella sampling simulations. Change in hydrophobic SASAs during (A) untwisting-stretching (twist angle θ_m) and (B) compression-stretching (length L_m) deformations. Hydrophobic SASAs were used to calculate the respective free energy. Error bars represent SE obtained from boot-strapping.

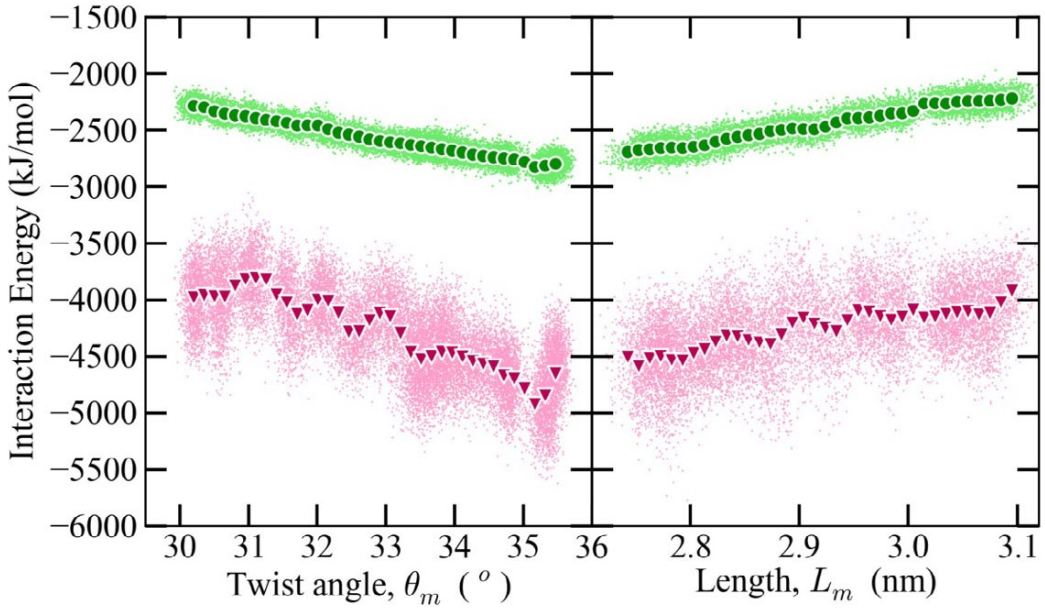

Figure S17: Residue packing dependent changes in interaction energies between deformed subunits of the middle connector region derived from umbrella sampling simulations. Fluctuating instantaneous interaction energies are shown as small dots; solid symbols represent averages over intervals in *x*-direction. Calculated van der Waals (*green circles*) and electrostatic (*red triangles*) interaction energies with respect to the deformation in (A) twist angle θ_m and (B) length L_m are depicted.

References

1. Fernandez-Fuentes, N., J. Zhai, and A. Fiser. 2006. ArchPRED: a template based loop structure prediction server. Nucleic. Acids. Res. 34:W173-176.

2. Eswar, N., B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper, and A. Sali. 2007. Comparative protein structure modeling using MODELLER. Curr. Protoc. Protein. Sci. Chapter 2:Unit 2 9.

3. Macke Thomas, J. and A. Case David. 1997. Modeling Unusual Nucleic Acid Structures. In Molecular Modeling of Nucleic Acids. American Chemical Society. 379-393.

4. Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl. 2008. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Chem. Theory. Comput. 4:435-447.

5. Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 65:712-725.

6. Perez, A., I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, 3rd, C. A. Laughton, and M. Orozco. 2007. Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. Biophys. J. 92:3817-3829.

7. Berendsen, H. J. C., J. P. M. Postma, W. F. Vangunsteren, A. Dinola, and J. R. Haak. 1984. Molecular-Dynamics with Coupling to an External Bath. J. Chem. Phys. 81:3684-3690.

8. Darden, T., D. York, and L. Pedersen. 1993. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. J. Chem. Phys. 98:10089-10092.

9. Amadei, A., A. B. Linssen, and H. J. Berendsen. 1993. Essential dynamics of proteins. Proteins 17:412-425.

10. Berendsen, H. J. and S. Hayward. 2000. Collective protein dynamics in relation to function. Curr. Opin. Struct. Biol. 10:165-169.

11. Garcia, A. E. 1992. Large-amplitude nonlinear motions in proteins. Phys. Rev. Lett. 68:2696-2699.

12. Hayward, S., A. Kitao, F. Hirata, and N. Go. 1993. Effect of solvent on collective motions in globular protein. J. Mol. Biol. 234:1207-1217.

13. Kitao, A., F. Hirata, and N. Gō. 1991. The effects of solvent on the conformation and the collective motions of protein: Normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum. Chem. Phys. 158:447-472.

14. Hess, B. 2000. Similarities between principal components of protein dynamics and random diffusion. Phys. Rev. E Stat. Phys. Plasmas. Fluids. Relat. Interdiscip. Topics 62:8438-8448.

15. Kutzner, C., J. Czub, and H. Grubmuller. 2011. Keep It Flexible: Driving Macromolecular Rotary Motions in Atomistic Simulations with GROMACS. J. Chem. Theory. Comput. 7:1381-1393.

16. Kappel, C., N. Dolker, R. Kumar, M. Zink, U. Zachariae, and H. Grubmuller. 2012. Universal relaxation governs the nonequilibrium elasticity of biomolecules. Phys. Rev. Lett. 109:118304.

17. Simpson, A. A., Y. Tao, P. G. Leiman, M. O. Badasso, Y. He, P. J. Jardine, N. H. Olson, M. C. Morais, S. Grimes, D. L. Anderson, T. S. Baker, and M. G. Rossmann. 2000. Structure of the bacteriophage phi29 DNA packaging motor. Nature 408:745-750.

18. Hess, B. 2002. Determining the shear viscosity of model liquids from molecular dynamics simulations. J. Chem. Phys. 116:209-217.

19. Hub, J. S., B. L. de Groot, and D. van der Spoel. 2010. g_wham-A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. J. Chem. Theory. Comput. 6:3713-3720.

20. Kumar, S., D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg. 1992. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules .1. The Method. J. Comput. Chem. 13:1011-1021.

21. Roux, B. 1995. The Calculation of the Potential of Mean Force Using Computer-Simulations. Comput. Phys. Commun. 91:275-282.