Biophysical Journal
Article

Biophysical Society

Protein Flexibility and Synergy of HMG Domains Underlie U-Turn Bending of DNA by TFAM in Solution

Anna Rubio-Cosials,[1] Federica Battistini,[2,3] Alexander Gansen,[4] Anna Cuppari,[1] Pau Bernadó,[5] Modesto Orozco,[2,3,6] Jörg Langowski,[4] Katalin Tóth,[4,*] and Maria Solà[1,*]

[1]Structural MitoLab, Department of Structural Biology, Molecular Biology Institute Barcelona (IBMB-CSIC), Barcelona, Spain; [2]Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology and [3]Joint BSC-IRB Program in Computational Biology, Institute for Research in Biomedicine, Barcelona, Spain; [4]Deutsches Krebsforschungszentrum, Division Biophysics of Macromolecules, Heidelberg, Germany; [5]Centre de Biochimie Structurale (CBS), Inserm, CNRS and Université de Montpellier, France; and [6]Department of Biochemistry and Biomedicine, Faculty of Biology, University of Barcelona, Barcelona, Spain

ABSTRACT Human mitochondrial transcription factor A (TFAM) distorts DNA into a U-turn, as shown by crystallographic studies. The relevance of this U-turn is associated with transcription initiation at the mitochondrial light strand promoter (LSP). However, it has not been yet discerned whether a tight U-turn or an alternative conformation, such as a V-shape, is formed in solution. Here, single-molecule FRET experiments on freely diffusing TFAM/LSP complexes containing different DNA lengths show that a DNA U-turn is induced by progressive and cooperative binding of the two TFAM HMG-box domains and the linker between them. SAXS studies further show compaction of the protein upon complex formation. Finally, molecular dynamics simulations reveal that TFAM/LSP complexes are dynamic entities, and the HMG boxes induce the U-turn against the tendency of the DNA to adopt a straighter conformation. This tension is resolved by reversible unfolding of the linker, which is a singular mechanism that allows a flexible protein to stabilize a tight bending of DNA.

INTRODUCTION

Mitochondrial transcription factor A (TFAM) regulates the transcription and transcription-dependent replication of human mitochondrial DNA (mtDNA) (1). In this genome, both the light and the heavy DNA strands code for proteins. In addition, there is a noncoding region where most of the cis-regulatory elements of mtDNA are found, including the heavy and light strand promoters (HSP1 and light strand promoter (LSP), respectively). By binding to specific sequences at LSP and HSP1, TFAM recruits the RNA polymerase (2). Another equally important function of this abundant protein is the compaction of mtDNA into nucleoprotein structures named nucleoids (1). TFAM activities are essential for mtDNA maintenance, mitochondrial biogenesis, and organism viability (1). TFAM belongs to the High Mobility Group B protein family. It contains two HMG box domains (HMG1 and HMG2) separated by a linker and followed by a C-terminal tail. In general terms, the HMG box domains consist of three α-helices arranged in an L-shape. The concave surface of this L-shape binds to the DNA minor groove, inducing pronounced bending of the DNA double helix (3).

The structure of TFAM in complex with its cognate binding site at LSP was determined by crystallographic studies, using either 22-bp (4) or 28-bp (5) DNA fragments (PDB: 3TQ6 and 3TMM, respectively). Additional structures include TFAM in complex with 22-bp fragments that contain either mitochondrial heavy strand promoter HSP1 (PDB: 4NNU) or nonspecific DNA (nsDNA; PDB: 4NOD) (6). In all these structures, each HMG-box domain of TFAM induces 90° bending to the DNA minor groove and the insertion of Leu58 (from HMG1) and Leu182 (from HMG2) into two 10 bp-separated sites, thus inducing an overall U-turn to the double helix (Fig. 1, *A* and *B*; Fig. S1 *A*). Additional residues from each HMG box domain contact the DNA and further stabilize the interaction (Fig. S1 *B*). Both HMG domains are joined by a positively charged 30-residue linker. In the complex, the protein and the DNA are intertwined. The HMG-boxes contact the minor groove on the same face of the DNA, whereas the

Submitted July 12, 2017, and accepted for publication November 15, 2017.

*Correspondence: kt@dkfz-heidelberg.de or maria.sola@ibmb.csic.es

Editor: David Rueda.

https://doi.org/10.1016/j.bpj.2017.11.3743

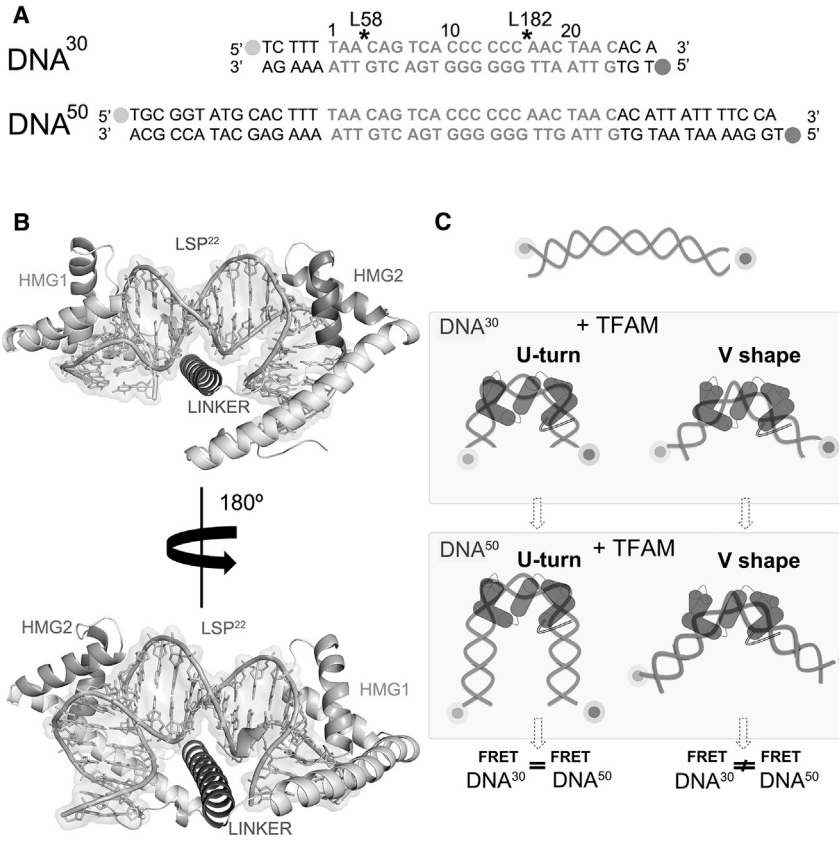

FIGURE 1 Experimental design for TFAM-LSP FRET experiments. (*A*) Sequences of DNA^{30} and DNA^{50} used for FRET assays were designed based on the TFAM/LSP22 crystal structure (4). DNA bases in light gray correspond to the crystallized fragment LSP22 (bent by 180°; see *B*), the additional bases in black correspond to the mtDNA sequence. Alexa 488 and Alexa 594 fluorophores are shown as light and dark gray dots, respectively. Black asterisks indicate the insertion sites of Leu58 and Leu182. (*B*) 180°-rotated views are given of the TFAM/LSP22 crystal structure (4). The DNA (LSP^{22}, 22 bp) is bent in a U-turn by the protein domains HMG1 and HMG2 (labeled in *light* and *dark gray*, respectively), stabilized by the linker domain. (*C*) Shown here is a scheme of the FRET strategy based on DNA^{30} and DNA^{50}. A U-turn (a V-shape) results in similar (different) FRET efficiencies for DNA^{30} and DNA^{50}. In a V-shape, the distance between dyes increases dramatically with DNA length (thus energy transfer decreases), whereas in the U-turn the distance is similar (and so is the FRET efficiency).

linker, in an α-helix conformation, nonspecifically interacts with the minor groove at the opposite side of the double helix. The C-terminal tail of TFAM contacts the mitochondrial RNA polymerase at the promoters (2). Both linker and C-terminal tail are unstructured in unbound TFAM, whereas, in a complex with DNA, the linker folds into the α-helix structure (4). These crystallographic studies have provided enormous insight into the atomic details of complex formation, but they present a static structure with potential lattice artifacts. This raises two questions: whether the x-ray structure represents the conformational ensemble of the TFAM/LSP complex in solution, and what underlying mechanism is used by the protein to bend the DNA.

Previous FRET studies on TFAM/LSP complexes showed consistency between the distance of the DNA ends in solution (~60 Å) (5–7) and the distance found in the 28-bp DNA crystal structure (~55 Å) (5). However, these and other studies (8,9) showed weak or no binding of isolated HMG2 constructs to the DNA, and suggested that HMG1 and the linker direct DNA binding and bending. Furthermore, these FRET studies failed to confirm the double kink shown by the crystal structure, and did not provide any information on the mechanism of TFAM-induced bending. Single-molecule FRET (smFRET) reports on the mutual distance between dyes and can resolve subpopulations within a heterogeneous sample to distinguish molecules in different dynamic states. Pioneered in the late 1990s on immobilized DNA (10) and freely diffusing molecules in solution (11), the power of this technique has been shown by many subsequent studies. smFRET has proven particularly suitable for the analysis of DNA distortion on protein binding in, for example, the yeast Nhp6 protein (12) or the eukaryotic transcription factor TATA-box binding protein (13).

In this study, we combined data from smFRET on freely diffusing molecules with results from SAXS and atomistic MD simulations to obtain structural information on the TFAM/LSP complex in solution, and to understand the mechanism of DNA binding and bending by TFAM. Our first aim was to discern whether TFAM bends DNA into a U-shape, which would have two kinks, or in a V-form with a single kink. We then studied the role of each protein domain in DNA binding and bending and derived a stepwise model for the TFAM/LSP interaction. SAXS studies provided a further biophysical characterization of the complex, revealing intrinsic structural dynamics of the complex. Finally, the structural properties underlying these dynamics were investigated by MD simulations. These properties were compared for a wild-type TFAM/LSP complex and an in silico nine-residue mutant, to analyze the contribution of protein and DNA components to complex dynamics and flexibility. Our study reveals the

molecular mechanism underlying LSP bending in which TFAM bends the DNA into a U-turn by synergic cooperation between its HMG box domains that kink the DNA, and the flexible linker that maintains the U-turn in a dynamic complex.

MATERIALS AND METHODS

The Supporting Material provides further details, adding to the succinct methods below.

Design of DNA constructs for smFRET experiments

Two DNA constructs of 30-bp (DNA^{30}) and 50-bp (DNA^{50}) length containing the cognate binding sequence at LSP were labeled with donor (Alexa 488) and acceptor (Alexa 594) dyes attached to the 5′ ends of complementary strands. After purification, single-stranded DNAs were annealed and the duplex was separated from nonannealed DNA by HPLC. Control samples were prepared with the same DNA sequences carrying only the donor (donor-only) or acceptor (acceptor-only) fluorophore.

Protein preparation

TFAM (residues 43–246; UniProt Q00059) and TFAM domains HMG1 (aa 43–125) and HMG2-Cter (aa 149–246) were cloned into pET28b(+) (Novagen (EMD) Biosciences, Madison, WI). The HMG1-L domain (aa 43–152) was cloned in pOPINF vector. All proteins and domain constructs were prepared as previously reported (4). Their folding was assessed by circular dichroism at the Centre for Genomic Regulation (Barcelona, Spain).

Protein and DNA complex formation

DNA and protein (or protein domains) were mixed at different molar ratios (1:1, 2:1) and dialyzed stepwise to a final buffer containing low salt (100 mM NaCl). smFRET measurements were made at a 10-fold excess of unlabeled DNA (from 500 pM total DNA, 50 pM were labeled). Complex formation was assessed by electrophoretic mobility shift assay (EMSA).

Single-molecule FRET experiments

smFRET experiments were performed on a home-built confocal microscope as described in (14,15). smFRET was measured on DNA and complexes in solution, in buffers containing 1 mM ascorbic acid and 0.01% surfactant Nonidet P40 to avoid photodamage and adhesion, by using multiplates (Greiner Bio-One, Kremsmünster, Austria) passivated with Sigmacote (Sigma-Aldrich, St. Louis, MO). For each fluorescence burst, the proximity ratio was analyzed by the sensitized emission of the acceptor upon selective donor excitation (14).

SAXS measurements

SAXS data of TFAM/LSP22 from 1.7 to 10.5 mg mL^{-1} were measured at the BioSAXS beamline from European Synchrotron Radiation Facility (Grenoble, France). Parameters were calculated from merged curves using appropriate programs. Models for comparison were generated with specific methods as detailed in the Supporting Material.

Molecular dynamics simulations

MD simulations of free DNA^{30} and DNA^{30} in complex with wild-type or mutant TFAM were performed using the AMBER 12 package. The starting coordinates of the TFAM/LSP complex were taken from the LSP22 crystal structure (PDB: 3TQ6). We created the mutant variant by mutating all residues of the HMG-boxes that directly interact with the DNA bases to alanine (L58A, Y57A, T77A, T78A, and I81A from HMG1, and Y162A, N163A, V166A, and L182A from HMG2-Cter). Details of preparation, equilibration, production, and analysis of the simulations can be found in the Supporting Material.

RESULTS

Design of the DNA constructs for smFRET experiments

Until now, not much is known about the structure and the stability of the TFAM/LSP complex in solution. Thus, we first determined whether TFAM bends LSP into a U-shape—similar to the crystal structure (Fig. 1 *B*) (4,5)—or into a V-shape, as previously suggested (7). The rationale behind our approach is shown in Fig. 1. We designed two DNA constructs of 30 bp (DNA^{30}) and 50 bp (DNA^{50}) length, whose central part includes the cognate binding site of TFAM at LSP (LSP22) (Fig. 1 *A*). The DNA ends were labeled with a FRET donor (Alexa 488) and acceptor (Alexa 594). Assuming a canonical B-DNA double helix in the absence of TFAM, the distance between dyes would exceed 10 nm for both constructs (102 Å for DNA^{30}, 170 Å for DNA^{50}), hence we expected very little or no FRET in unbound DNA. Upon binding to the central cognate sequence (hereafter denoted as TFAM/DNA^{30} and TFAM/DNA^{50} complexes), TFAM is expected to bend the DNA, reducing the distance between both dyes and generating a FRET signal. If TFAM bends the DNA at a single site, the DNA would assume a V-form and the FRET signal for TFAM/DNA^{50} would be much smaller than for TFAM/DNA^{30}, because the distance between both fluorophores would increase with DNA length (Fig. 1 *C*, *right column*). However, if TFAM kinks LSP at two sites, a U-turn would be induced, and the distance between both fluorophores would be similar in TFAM/DNA^{30} and TFAM/DNA^{50} (Fig. 1 *C*, *left column*). Therefore, to probe the mechanism of TFAM-induced DNA bending, we performed smFRET experiments for DNA^{30} and DNA^{50} incubated with either full-length TFAM or individual TFAM domains.

TFAM binds LSP in a U-turn

smFRET histograms of TFAM/DNA^{30} and TFAM/DNA^{50} complexes are shown in Fig. 2. The concentration of labeled DNA was set to 50 pM (corresponding to 500 pM DNA in total; see the Supporting Material). Data were acquired for 600 s, which is sufficient to achieve a good separation of subpopulations in the

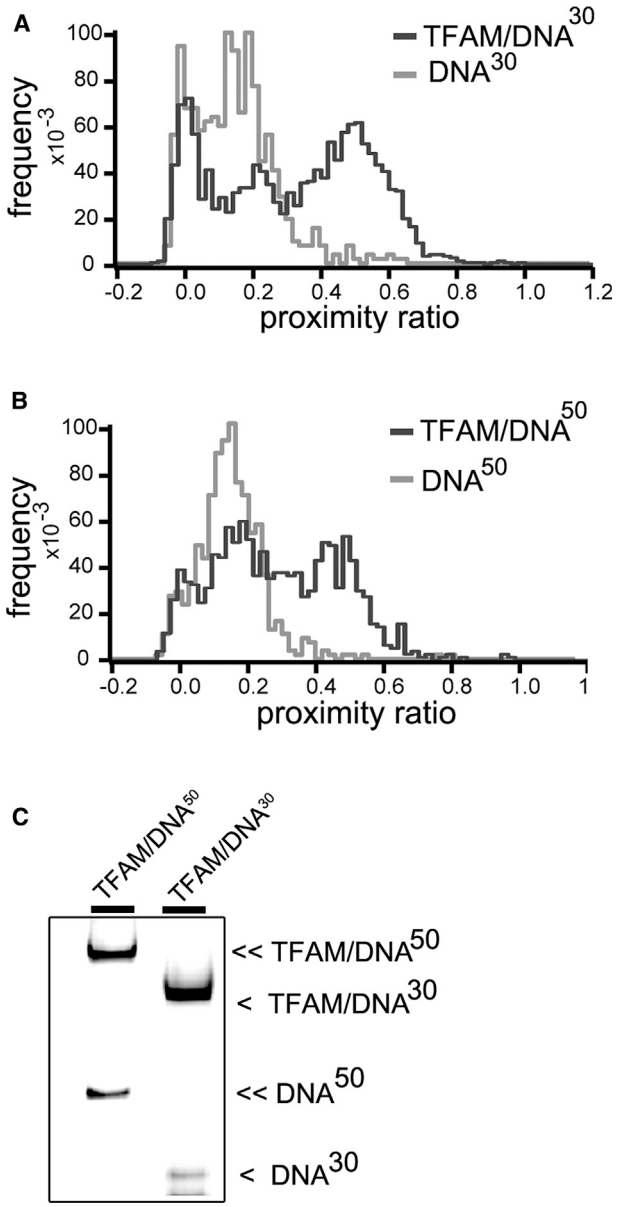

FIGURE 2 Representative smFRET histograms for TFAM/DNA^{30} (*A*) and TFAM/DNA^{50} (*B*) complexes. The distributions of free DNA^{30} and DNA^{50} are shown in light gray, whereas complexes with TFAM (TFAM/DNA) are shown in dark gray. The donor-only peak at $P = 0$ represents molecules that lack an active acceptor. The *p* values of high FRET peaks are indicated in Table 1. (*C*) Shown here is nondenaturing polyacrylamide gel (5%) with TFAM/DNA^{30} and TFAM/DNA^{50} samples. The gel bands were visualized using the fluorescence signal from Alexa 594 (excitation at 532 nm, detection at 595–625 nm). The symbol (<) shows migration of DNA^{30} (free and in complex) and (<<) shows DNA^{50} (free and in complex). Note that the lower band below free DNA^{30} is due to a minor impurity of single-stranded DNA in the sample.

histogram. After fitting with multiple Gaussian distributions, the characteristic proximity ratio value (*P*) of each peak was determined, as well as their width (FWHM). Free DNA showed two peaks; a donor-only species, which represents molecules that lack an active acceptor dye and which is centered around $P = 0$, and a second population that corresponds to the donor-acceptor duplex DNA (Fig. 2 *A*, *light gray line*, Table 1). The nonzero proximity ratio of the second peak is a result of residual acceptor excitation and does not arise from energy transfer. In contrast, TFAM/DNA^{30} formed an additional species with higher proximity ratio at $P \sim 0.5$ (Fig. 2 *A*, in *dark gray*). This population was broader (FWHM = 0.22 ± 0.08) than free DNA (FWHM = 0.13 ± 0.01). The excess width suggests either static heterogeneity, due to the presence of differently bent subconformations, or fast dynamics of the complex.

Importantly, the smFRET distribution of TFAM/DNA^{50} was similar to that of TFAM/DNA^{30}, with a slightly lower *P* for the TFAM/DNA^{50} complex ($P \sim 0.43$), as shown in Fig. 2, *A* and *B*. This small difference could arise from electrostatic repulsion of the longer DNA arms of DNA^{50}. For both DNA^{30} and DNA^{50}, the complex formation was previously assessed by EMSA (Fig. 2 *C*). The fact that DNA^{30} and DNA^{50} yielded comparable *P* values upon TFAM binding strongly supports the U-turn model for TFAM-induced DNA bending.

The individual protein domains orchestrate a stepwise DNA bending

Next, we studied the role of the individual TFAM domains in LSP bending. To do so, we generated a construct containing HMG1 domain alone (residues 43–125, called HMG1), HMG1 with the linker region (residues 43–152, HMG1-L), and HMG2 with the C-terminal tail (residues 149–246, HMG2-Cter) (Fig. S1 *A*). We first tested the ability of these protein fragments to bind fluorescently labeled DNA^{30} and DNA^{50} at nanomolar concentrations by EMSA. Increasing concentrations of full-length TFAM (>200 nM) resulted in a two-band shift (Fig. S2), which might be related to oligomerization of TFAM in the presence of DNA, as previously reported (8). Note that in subsequent smFRET studies, sample concentrations were in the subnanomolar range, so protein oligomerization should be negligible. Regarding the EMSAs, full-length TFAM showed higher DNA binding ability than the individual domains.

TABLE 1 Characteristic Proximity Ratio Values of the High FRET Peak for TFAM and TFAM Domains in Complex with DNA^{30} and DNA^{50}

Sample	Proximity Ratio DNA^{30}	Proximity Ratio DNA^{50}
Free DNA	0.160 ± 0.005	0.150 ± 0.003
TFAM/DNA	0.481 ± 0.011	0.437 ± 0.007
HMG1-L/DNA	0.277 ± 0.006	0.152 ± 0.004
HMG1/DNA	0.217 ± 0.007	0.156 ± 0.001
HMG2-Cter/DNA	0.175 ± 0.003	0.156 ± N/A

For each peak, the proximity ratio (i.e., mean value over replicates) and the associated SE ($\sigma/\sqrt{n}$) are shown. Note that the free DNA sample is described by the peak of the donor-acceptor duplex DNA, as the donor-only peak at $P = 0$ does not convey any structural information.

HMG1 showed blurred band shifts, which suggests that it is a less stable complex. In contrast, the presence of the linker in HMG1-L resulted in better DNA binding, whereas HMG2-Cter showed very weak (but still detectable) binding, in agreement with previous reports (8,9). The fact that bands shifted by full-length TFAM were better defined than those of HMG1-L indicates that HMG2 helps to stabilize the position of TFAM on DNA^{30}, even though it has a low affinity for DNA on its own.

Whereas EMSA data provide information on protein binding, FRET serves as a local probe for DNA bending. smFRET histograms of the domain constructs in complex with DNA were asymmetric at proximity ratios higher than those of free DNA^{30} (Fig. 3 *A*; Table 1). Binding of HMG1 led to a broader distribution with a slightly higher proximity ratio ($P \sim 0.22$) compared to the free DNA peak ($P \sim 0.16$). For HMG1-L, the proximity ratio further increased ($P \sim 0.28$) and showed a somewhat sharper distribution, which suggests that the linker stabilizes binding of HMG1 to DNA and enhances DNA bending. Still, the proximity ratio was lower than that observed for full-length TFAM, which indicates that the second HMG box is required to induce the final U-turn. Binding of HMG2-Cter to DNA^{30} showed a different picture. Although the center of the proximity ratio for the complex was similar to that of HMG1/DNA^{30}, its relative area was much lower, which suggests weaker binding. Such a weaker binding is consistent with the marginal DNA shift in EMSA (Fig. S2 *A*). A hypothetical representation of DNA bending within each complex is shown beside the corresponding FRET histogram in Fig. 3 *B*. Notably, the complex of any of these TFAM fragments with longer DNA^{50} resulted in histograms similar to that of unbound DNA^{50} (see EMSA and FRET data in Fig. S2 and FRET values in Table 1), which further indicates that each subdomain induces only a single kink and bends DNA into a V-shape.

TFAM forms a compact and heterogeneous complex with DNA

To further characterize the TFAM/LSP complex in solution, we performed SAXS measurements. From the TFAM/LSP scattering curve we could estimate a molecular mass of ~30 kDa (16), which agrees with one protein bound to one DNA molecule (25.6 KDa TFAM + 13.5 kDa DNA = 39.1 kDa). The estimated radius of gyration (R_g) of the complex ($R_g = 26.0 \pm 0.1$ Å) is smaller than that of the free protein ($R_g = 32.0 \pm 0.3$ Å) (4). This is consistent with previous single-molecule (17) and multiangle light scattering (5) studies, which indicated that the protein/DNA particle is more compact than the free protein. The Kratky plot derived from the scattering curve has a bell-shaped profile consistent with an overall globular shape in solution (Fig. 4 *A*). This is in contrast to the Kratky plot for the free protein, which does not show a maximum and indicates a high degree of conformational flexibility that we attributed to the structural variability, due to both the linker and C-terminal tail (Fig. 4 *A*; (4)). The pairwise distance distribution $p(r)$, which provides the overall shape of the particle in solution, displays a slightly asymmetric profile with a maximum intramolecular distance (D_{max}) of 82 Å. Notably, this value is smaller than D_{max} of unbound TFAM ($D_{max} = 135 \pm 5$ Å) (Fig. 4 *B*), which further demonstrates the compaction of the protein upon DNA binding.

We used the crystallographic structure of the complex as a model to describe the SAXS curve. Despite the overall similarity of both curves, the fit is relatively poor ($\chi^2 = 2.93$) and shows systematic deviations throughout the complete momentum transfer range (Fig. S3 *A*). This suggests that

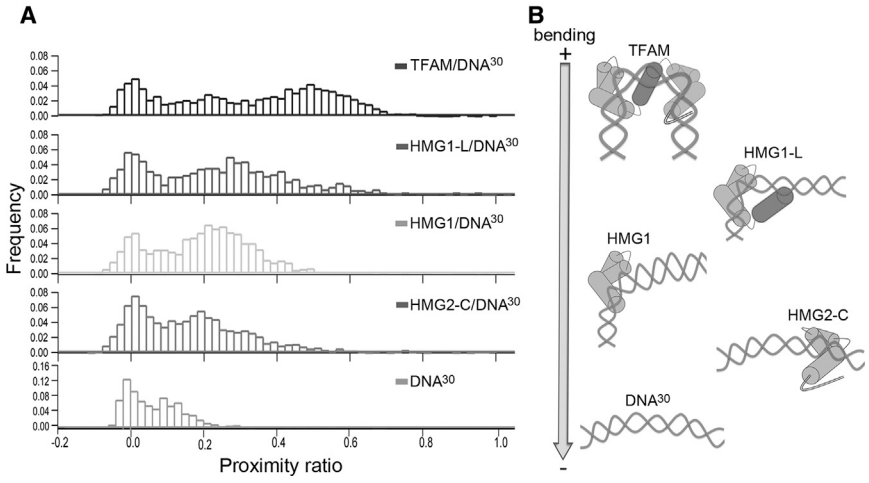

FIGURE 3 The individual protein domains orchestrate stepwise DNA bending. (*A*) Representative smFRET histograms for TFAM and TFAM domains in complex with DNA^{30} are shown. The complex with full-length TFAM is shown at the top, followed by HMG1-L, HMG1, HMG2-Cter, and free DNA (*bottom*). (*B*) A model of bending is represented aligned with each histogram and ordered from top (sharpest) to bottom (lowest) according to the extent of DNA bending. The TFAM/DNA model is based on the U-turn structure, whereas the complexes with domains are hypothetical representations based on the proximity ratio distributions. The DNA is represented as prebent, in accordance with our MD results. From bottom to top, the free DNA and HMG2 complex show similar bending due to the marginal effect of HMG2. Above, HMG1 introduces the first kink, which is enhanced by the linker. At the top, full-length TFAM induces a U-turn, suggesting cooperativity, because HMG2 kinks the DNA only in the presence of the N-terminal protein fragment.

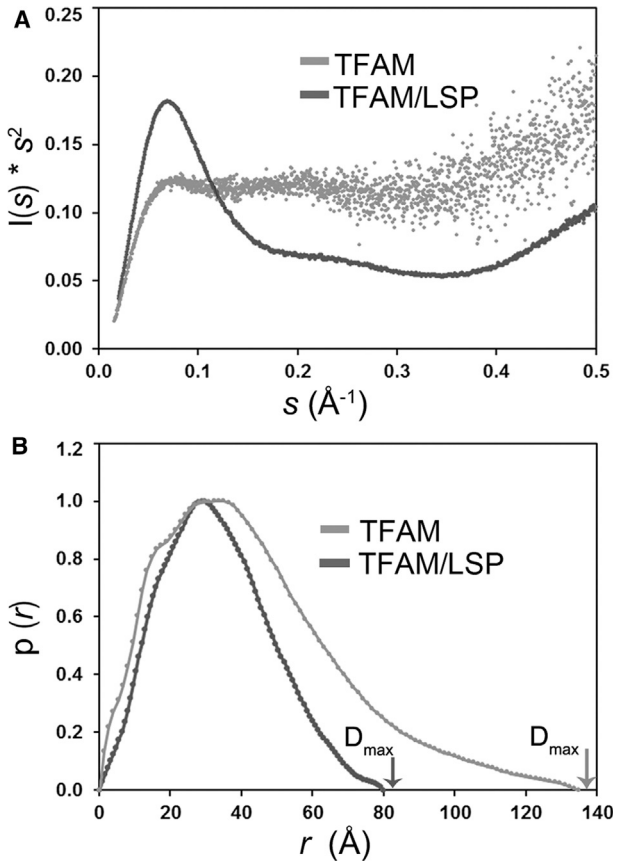

FIGURE 4 SAXS analysis of TFAM/LSP complexes. (*A*) Kratky representation of experimental SAXS curves. Values have been normalized by I_0. The TFAM/LSP complex shows the profile of a globular particle (*dark gray*), whereas free TFAM (in *light gray*) clearly shows the profile of a flexible protein. (*B*) Shown here is the pairwise distance distribution function $p(r)$. Maximal particle dimensions are arrowed for both free and DNA-bound TFAM.

TFAM/LSP is not a static complex. The aforementioned broadening of the TFAM/DNA peak in the smFRET distribution also hints at the conformational flexibility of the TFAM/LSP complex. A possible source of flexibility could be the intrinsic disorder of the C-terminal tail, as suggested by the weak electron density between Lys237 and Cys246 in both TFAM/LSP crystal structures. To check this, we applied the ensemble optimization method (18) using a large pool of TFAM/LSP structures with different C-terminal conformations added to the crystallographic structure (Fig. S3 *B*). However, the ensemble optimization method did not yield a better description of the experimental curve, which suggests that additional sources of flexibility apart from the C-terminal tail are present in TFAM/LSP.

Computational analysis shows the intrinsic curvature of the TFAM binding site at LSP

To better understand the flexibility of TFAM/LSP complexes at the molecular level, we characterized the structural and dynamic properties of free LSP-DNA and the protein/DNA complex by MD simulations. First, we assessed the sequence-dependent intrinsic curvature and flexibility of each base pair step of free DNA[30] in water. During the simulation, the initially straight DNA[30] spontaneously assumed a more curved conformation, with an average total bend of 72 ± 18° (Figs. 5 *A* and S4 *A*). In particular, the region specifically contacted by TFAM in the bound structure (represented by LSP22 in the crystal structure; see Fig. 1 *A*), showed an average curvature of 46 ± 12°, with an occasional high curvature conformation (113°) closer to that of the protein-bound DNA (169°) (Fig. S4 *B*). An analysis of the average minor groove width along the TFAM binding site showed sequence-dependent variability. In particular, the minor groove around the Leu58 insertion site is narrower than that around Leu182 in the free DNA (Fig. S4 *C*).

An analysis of the basepair parameters shows that step $A_3C_4/G_{19}T_{20}$, into which HMG1 Leu58 inserts, is moderately flexible (*left vertical line* in stiffness plot in Fig. 5 *B*) and has roll values close to those of an ideal B-DNA (−0.84° on average; *green line* in Fig. 5 *C*). In contrast, the neighboring step $C_4A_5/T_{18}G_{19}$ and, importantly, the insertion site of HMG2 Leu182, step $C_{15}A_{16}/T_7G_8$, have a higher, positive roll (10.4 and 10.5°, respectively) and are more flexible (Fig. 5, *B* and *C*). In summary, DNA[30] deviates moderately from the canonical B-DNA at the Leu58 insertion site, whereas at the following step and Leu182 site the deviations are much more pronounced.

LSP forms a nonstatic complex with TFAM

The influence of TFAM on DNA bending was analyzed by MD simulations of LSP in complex with wild-type protein. We monitored structural changes during the MD by secondary structure analysis (Fig. 6 *A*) and time-dependent root mean square deviation (RMSD) relative to the equilibrated initial structure (Fig. 6 *B*). During the MD simulation, the complex maintained the DNA-protein contacts, and converged to a structure similar to the initial one at the end of the simulation.

The most significant motion as captured by principal component analysis of the MD trajectory was a reversible dynamic separation between the two HMG domains, which moved apart back and forth in a butterfly or "breathing" movement (Fig. S5 *A*). At the center of this movement, the linker is predicted to locate the hinge point of the entire complex (*arrow* in Fig. S5 *A*). The breathing of the protein was coupled to a bending-relaxation movement of DNA, but no changes in protein/DNA contacts were detected. The largest fluctuations of atomic positions took place at HMG1 and, remarkably, in the linker region (the average atomic mobility during MD is represented by the root mean square fluctuations in Fig. S5 *B*, *black curve*). Interestingly, during the simulation, an event with the highest RMSD occurred (indicated with a *gray frame* in Fig. 6 *B*).

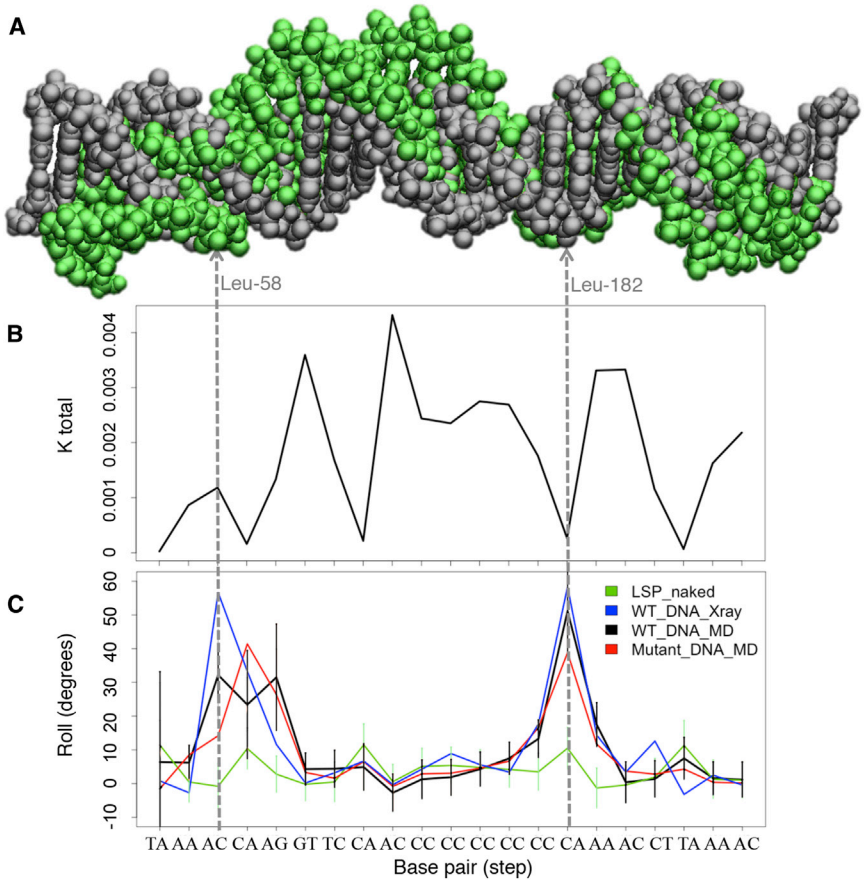

FIGURE 5 Structural features of the TFAM binding site at LSP. (*A*) Given here is a DNA[30] LSP sequence in the ideal B-DNA form (in *gray*) and a frame from the MD simulation (in *green*) that shows considerable distortions. (*B*) Values of the stiffness parameter K_{total} are for the naked LSP sequence, calculated from the MD simulation for each base pair step. (*C*) Given here is the base pair parameter roll (in degrees) for the DNA[30] LSP sequence in the x-ray crystal structure (in *blue*) and averaged during the MD simulations for the naked DNA (in *green*), for the DNA in complex with the wild-type protein (in *black*), and in complex with the MD-mutant protein (in *red*).

This corresponded to DNA straightening toward its ideal naked conformation (DNA bending ranges from 169 to 61°; Fig. 6 *C*), together with partial unfolding of the linker region (*black square* in Fig. 6, *A* and *D*, *left column*). Thus, the linker is the protein region that adapts most to DNA movements, and it unfolded precisely at the region that contacts the DNA. This conformational change suggests that the protein accommodates the DNA motions by virtue of its linker flexibility, without disrupting the contacts of the HMG boxes with the DNA. To relate these conformational changes to the smFRET broadening, we calculated interdye distances for those MD models that diverged most from the x-ray structure after relaxation (Fig. S3 *C*). This led to theoretical, MD-based FRET efficiency values, from 0.31 to 0.65 (Table 2). This broad range is in agreement with the increased peak width in the smFRET distributions, which indicates that the flexibility of the TFAM-LSP complex is reflected in the peak broadening.

We next took a closer look at the base pair parameters of the DNA. In the crystal structure, the protein-bound DNA displays the highest roll angle at the base pair steps contacted by Leu58 (at HMG1) or Leu182 (at HMG2) (4) (Fig. 5 *C*, *blue line*). During the MD simulation of the wild-type (WT) complex, the high roll at the Leu182 step was maintained, whereas the roll angle at the Leu58-contacted step $A_3C_4/G_{19}T_{20}$ diminished (Fig. 5 *C*, *black line*). The angle at the neighboring steps $C_4A_5/T_{18}G_{19}$ and $A_5G_6/C_{17}T_{18}$ was maintained or increased. Thus, in the HMG1 region, the stress at one point with high roll angle redistributes to the neighboring steps, keeping the bending of the DNA. These changes in roll at the HMG1 binding site may correspond to release of crystal packing constraints, which results in a less kinked and less stressed DNA than that found in the x-ray structure. Interestingly, the comparison of DNA in bound and unbound states shows that the major structural distortions found in the complex (such as the roll angle) are already present in the unbound form, yet less pronounced. In addition, the roll redistribution in the complex during the MD leads to a pattern that is more similar to that of naked DNA than the x-ray one. These similarities suggest that the free 22 bps in LSP that contact TFAM have a structural propensity toward the protein-bound conformation.

The dynamic tension between TFAM and LSP causes reversible unfolding of the linker

To further understand the influence of the DNA-contacting residues in HMG1 and HMG2 on DNA bending, we ran an MD simulation of an in silico TFAM mutant

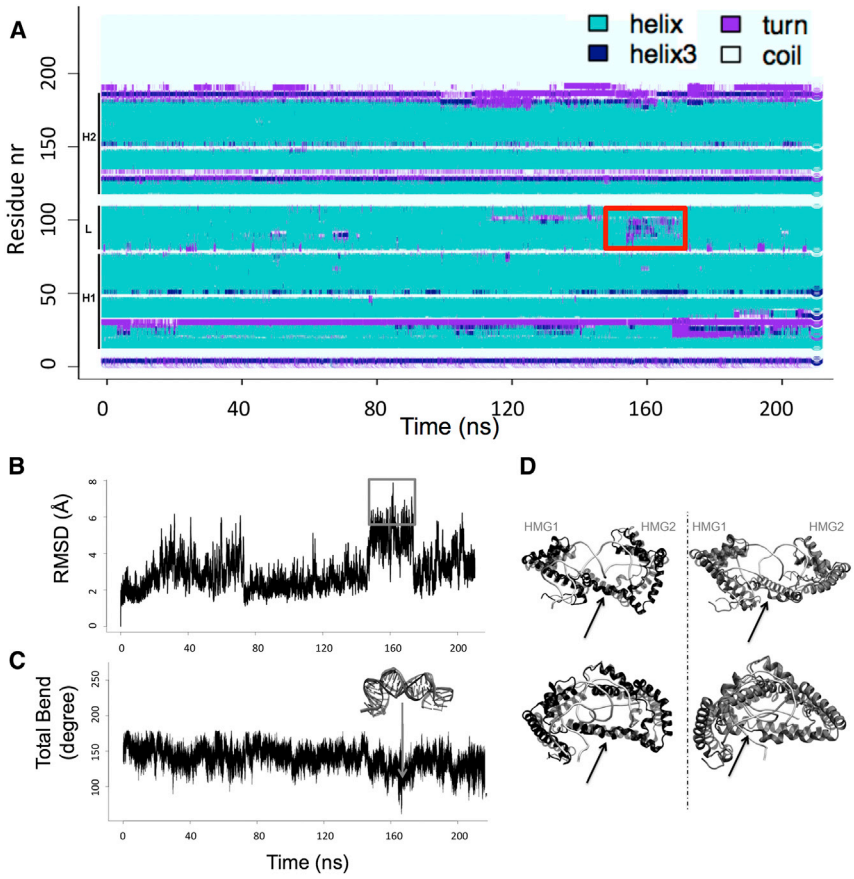

FIGURE 6 Structural variability of the TFAM/LSP complex. (*A*) Shown here is a secondary structure plot of WT protein in complex with LSP during the simulation (every 100th snapshot). The secondary structure of the WT protein, i.e., the HMG1, HMG2, and linker domains (marked as H1, H2 and L, respectively), are maintained during the simulation, except for an unfolding event at the linker region at ~160 ns (*red box*). Helical conformation is labeled in cyan (helix 3_{10} in *dark blue*), turns in violet, and coils in white. (*B*) Shown here is the time-dependent RMSD of Cα atoms in the WT protein in complex with DNA. The reference structure is the x-ray crystal structure after MD relaxation. The box highlights the highest values, which coincide with unfolding of the linker (compare with (*A*)). (*C*) Given here is the total bend (in degrees) of the DNA in complex with the WT protein during the MD simulation. The arrow points at the event during which the DNA significantly reduces its bending, as shown in the overlay between the starting DNA structure (*light gray*) and the DNA at this time step (*dark gray*). (*D*) Shown here are plots showing the unfolding of the linker compared to the initial structure (in *light gray*) for the WT protein (in *black*, *left panel*) and the MD-mutant protein (in *dark gray*, *right panel*) throughout the simulation. In the images (*front* and *side views* of the complexes), the black arrows point to the linker region that unfolds during the MD. To see this figure in color, go online.

(MD-mutant) in complex with DNA. In the mutant, nine residues were substituted by alanines: L58A, Y57A, T77A, T78A, and I81A from HMG1; and Y162A, N163A, V166A, and L182A from HMG2 (Fig. S1 *B*). Comparison of the mutant and the WT trajectory should reveal the constraints imposed by the WT side chains to the bent DNA.

The MD trajectory of the MD-mutant/DNA complex showed higher deviation from the initial conformation (see the overall higher values of mutant RMSD, Fig. S5 *C*). Principal component analysis revealed a faster breathing motion of the HMG boxes around the linker hinge point than in the WT complex (Fig. S5 *A*). Compared to the WT complex, the root mean square fluctuation of DNA atoms increased significantly in regions that are contacted by the HMG boxes and, notably, by the linker (Fig. S5 *B*). During the simulation, the strain on the DNA diminished because the overall roll angles decreased, as did the global curvature (exemplified by *lower roll values*, Fig. 5 *C*). Most importantly, the linker irreversibly lost its helical structure due to unfolding and stretching at the region between residues 123–153 (Fig. 6 *D right column*, and Fig. S6), which on average increased its distance from the DNA (Fig. S5 *D*). Although these deformations were considerable, the contacts between the HMG boxes and DNA were preserved. Only a slight increase in the distance between HMG2 and the DNA was observed along the MD simulation, which did not result in disassembly of the complex at this timescale (Fig. S5 *D*).

The HMG2 DNA region has a kink characterized by the highest roll (*red graph* in Fig. 5 *C*), which only slightly diminished during the simulation, even though Leu182 was substituted by the much shorter alanine. On the other hand, the kink at the HMG1 region shifted from A_3C_4/$G_{19}T_{20}$ in the x-ray structure to the immediate neighbor C_4A_5/$T_{18}G_{19}$. Strikingly, the latter is also the step with the

TABLE 2 Donor-Acceptor Distance Calculation for the Initially Relaxed Crystal Structure and Four Representative Conformations from the MD Simulations

Sample	R_{DA}	σ_{DA}	E_{FRET}
0_pdb	48.1	10.3	0.31
121_pdb	52.3	11.8	0.43
31_pdb	56.6	7.2	0.54
26_pdb	53.3	10.5	0.45
00_pdb	61.2	8.9	0.65

The DNA from the MD structures was extended with the DNA[30] sequence. All parameters were computed using the FPS toolkit. 0_pdb indicates the initially relaxed structure, R_{DA} indicates the dye-to-dye distance in Å, σ_{DA} indicates the standard deviation, and E_{FRET} values are FRET efficiencies calculated from the weighted R_{DA}−s assuming a Förster radius of $R_0 = 55.6$ Å.

highest flexibility and roll in the naked DNA structure (compare *green* and *red lines* in Fig. 5 *C*). Thus, upon removal of the side chains that constrain the DNA in a kinked conformation, the DNA relocates the distortions to the steps that are intrinsically easier to distort, without affecting the binding or the bending. Therefore, in complex with the MD-mutant, the DNA tries to recover, even more closely, the intrinsic sequence-dependent structure of the naked double helix. This results in stronger distortions at the protein region that is most sensitive to structural variations in DNA, which is the linker.

DISCUSSION

TFAM is a transcription factor that induces strong bending to LSP and HSP1 promoters (19), enabling mtRNAP recruitment (2,5,7). In addition, it is an architectural protein that compacts human mtDNA by introducing strong bends to the double helix (5,6,20–22). In this study, we investigated the mechanism by which the protein binds and bends DNA in solution and whether it does so at a single point, leading to a V-kink deformation, or at two points, leading to the U-turn conformation found in the crystal. By combining single-molecule FRET and SAXS data, we unambiguously showed that TFAM compacts LSP by inducing a U-turn, and that each individual HMG box imposes a single kink to the DNA. Our data revealed that the linker between the HMG boxes enhances the bending exerted by HMG1 and, as unveiled by MD, confers considerable flexibility to the complex. In the TFAM/LSP crystal structure, an electropositive surface of the linker C-terminal region binds the concave surface of the U-turn, thus stabilizing the bending by electrostatic interactions (4,5). Similar contact might occur in the HMG1-L/DNA complex.

Interestingly, in the free protein the linker is disordered but gains α-helix character in complex with DNA, as seen in the crystal structure, by SAXS (4) and by CD spectroscopy (7,9). This transition is pivotal to the TFAM/LSP interaction. We found that the isolated HMG2-Cter domain has only low binding affinity to LSP. Previous studies showed that binding ability increases when HMG2 is fused to the linker C-terminal region (9), probably due to formation of a local hydrophobic core with the linker α-helix (4). This suggests that, for efficient DNA binding, HMG2 depends on the presence of the helical linker, whose fold in turn depends on HMG1 bound to the DNA. These findings support a unique stepwise binding mechanism, in which DNA binding is driven by HMG1, which initially bends the DNA into a V-shape. This reduces the accessible conformational space for the linker, which folds into a 29-residue α-helix that wraps around the DNA and reduces the distance between HMG boxes. This eventually facilitates HMG2 binding. The latter, in turn, introduces a second V-kink to the DNA, completing the U-shape conformation. The U-turn is thus a consequence of a double kink rather than a smooth DNA bending. To sum up, the system shows a sequential cooperative bending, in which the synergy between the individual domains and the folding of the linker mediate the correct alignment of protein and DNA.

An important question regarding the above model is whether binding follows a pure induced fit paradigm, in which the protein deforms a passive DNA, or if it has a certain degree of conformational selection. In the latter, the intrinsic deformability and conformation of the DNA help to achieve the distortion pattern required for binding. MD analysis shows that naked LSP has intrinsic overall curvature and conformation at the basepair level that may facilitate its bending into a U-turn. However, detailed analysis of trajectories shows the complexity of the contribution of the DNA to the protein binding. HMG1 inserts Leu58 into the DNA, but the strong distortion found in the crystal structure at the insertion site diminishes during the time course of the MD simulation and the neighboring steps, which have an intrinsic ability to curve, partially absorbed the distortion. Binding of HMG1 favors the formation of the linker α-helix that facilitates the insertion of HMG2 at a DNA step characterized by an extremely deformable and open base pair. Interestingly, the complex is not static due to the intrinsic flexibility of both LSP and the protein, as suggested by MD and distribution broadening in smFRET. This is backed up by the discrepancy between the experimental SAXS data and a model based on the static crystal structure. Additionally, an in silico mutant allowed us to gain information on such a flexible bending mechanism of TFAM. Mutations of key interacting residues in the two HMG domains lead to an increase in the tendency of DNA to recover a straighter conformation, which produces major corruption in the linker conformation and some departure of the HMG2 from the bound arrangement.

Taken together, all these observations strongly suggest that the TFAM/LSP complex is in a dynamic equilibrium governed by 1) protein-DNA contacts, especially those formed by the inserted leucines; 2) the equilibrium between disordered and α-helix conformation of the linker; and 3) the intrinsic tendency of DNA to fold back to a straighter structure. The global structure of the complex is subjected to opposing forces, leading to a flexible complex whose overall conformation varies considerably. The bending angles vary from 169° to 61°, a range that comprises the angles detected by previous studies using LSP (78° or 72°) or nsDNA (100° ± 20°) (7,21). Because U-turn bending is required to recruit the mitochondrial RNA polymerase to the LSP transcription initiation site (2), it can be speculated that the degree of bending may condition transcription activation efficiency. The isoform TFAMΔexon5 lacks the C-terminal part of the linker and the first helix of HMG2 (residues 148–179) and is unable to activate transcription in vitro (23). After the linker, such a polypeptide chain may fold only into two consecutive α-helices and a

disordered C-terminal tail. According to our data, this isoform might be unable to induce a U-turn on LSP, because only intact HMG2 might actively bend the DNA into the second V-shape. The C-terminal tail may contact the transcription machinery, but the partial bending may not be sufficient to bring it to the transcription initiation site. Such impaired transcription and transcription-dependent replication is expected to have a direct impact on mtDNA and mitochondria biogenesis.

TFAM is the main protein involved in mtDNA packaging, in which most of interactions are nonspecific. In this regard, it is highly plausible that variability in bending may depend on the properties of the sequence at the DNA binding sites. It was shown that HMG-box proteins increase DNA flexibility and decrease DNA persistence length (a parameter related to the stiffness of the DNA) by kinking the DNA. These HMG-box proteins include TFAM (17) and the *Saccharomyces cerevisiae* HMO1 (24) and Nhp6A (12) proteins. Two models have been proposed to explain the increase in flexibility of the kinked complexes ((12,25,26) and references therein): a static kink model, in which the protein binds to DNA creating fixed angle bends; and a flexible hinge model in which a highly dynamic complex has significant oscillations around a slightly preferred bending angle. Previous optical force experiments using long DNA molecules detected an increase in DNA flexibility upon TFAM binding (17). Based on the static crystal structure, the authors proposed that TFAM/DNA complexes were rigid entities and flexibility could arise from local denaturation of the DNA between protein/DNA kinks. However, our data clearly demonstrate that the TFAM/DNA complex is flexible and governed by dynamic tensions, as shown by our MD calculations, and suggests a rather flexible hinge model exhibiting variable bending. Thus, TFAM would introduce local flexible points at nsDNA during compaction, and such flexible assemblies should be compatible with formation of higher-order protein-DNA complexes.

In conclusion, our study provides important insights into the molecular mechanism underlying bending of LSP by TFAM. We show by FRET and SAXS that TFAM domains bind and bend LSP by a stepwise mechanism that induces a compact U-turn in solution. MD shows that the complex is dynamic and the DNA tends to recover its free state conformation, but this is counteracted by unfolding and refolding of the protein helical linker, which restores the U-turn. This tension results in variation of the bending angle. Our results point to a tight DNA bending mechanism based on the flexibility of TFAM. A similar mechanism may underlie DNA compaction in mitochondria, a process that is essential for cell life.

SUPPORTING MATERIAL

Supporting Materials and Methods and seven figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(17)34975-5.

AUTHOR CONTRIBUTIONS

A.R.-C. contributed to cloning. A.R.-C. and A.C. did protein production. A.R.-C. performed complex preparation and EMSA. A.R.-C., A.G., and J.L. performed the FRET studies. A.R.-C. and P.B. performed the SAXS studies. F.B. and M.O. performed molecular dynamics studies. M.S. and K.T. designed and supervised the project. All authors participated in writing the manuscript.

ACKNOWLEDGMENTS

Jörg Langowski sadly passed away on May 6, 2017.

We thank Sébastien Lyonnais for guidance with the EMSA gels. M.O. is an ICREA Academia researcher.

This work was supported by the Ministry of Economy and Competitiveness (MINECO) (BFU2012-33516 and BFU2015-70645-R to M.S., and BIO2012-32868 and BFU2014-61670-EXP to M.O.); Generalitat de Catalunya (SGR2009-1366 and 2014-SGR-997 to M.S., and SGR2009-1348, 2014 SGR-134 to M.O.); the Instituto Nacional de Bioinformática; the European Union (FP7-HEALTH-2010-261460, FP7-PEOPLE-2011-290246, and FP7-HEALTH-2012-306029-2 to M.S., and H2020-EINFRA-2015-1-675728 and H2020-EINFRA-2015-676556 to M.O.); and the European Research Council (ERC-2011-ADG_20110209-291433 to M.O.). A.R.-C. was awarded with a "Junta para la Ampliación de Estudios" (Programa JAE) fellowship from Consejo Superior de Investigaciones Científicas (CSIC). The Structural Biology Unit at IBMB-CSIC is a "Maria de Maeztu" Unit of Excellence awarded by the Ministry of Economy and Competitiveness (MINECO) under MDM-2014-0435. IRB Barcelona is the recipient of a Severo Ochoa Award of Excellence from the Ministry of Economy and Competitiveness (MINECO). The CBS is a member of the French Infrastructure for Integrated Structural Biology (FRISBI), a national infrastructure supported by the French National Research Agency (ANR-10-INBS-05).

REFERENCES

1. Gustafsson, C. M., M. Falkenberg, and N. G. Larsson. 2016. Maintenance and expression of mammalian mitochondrial DNA. *Annu. Rev. Biochem.* 85:133–160.
2. Morozov, Y. I., A. V. Parshin, …, D. Temiakov. 2015. A model for transcription initiation in human mitochondria. *Nucleic Acids Res.* 43:3726–3735.
3. Stros, M., D. Launholt, and K. D. Grasser. 2007. The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. *Cell. Mol. Life Sci.* 64:2590–2606.
4. Rubio-Cosials, A., J. F. Sidow, …, M. Solà. 2011. Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. *Nat. Struct. Mol. Biol.* 18:1281–1289.
5. Ngo, H. B., J. T. Kaiser, and D. C. Chan. 2011. The mitochondrial transcription and packaging factor TFAM imposes a U-turn on mitochondrial DNA. *Nat. Struct. Mol. Biol.* 18:1290–1296.
6. Ngo, H. B., G. A. Lovely, …, D. C. Chan. 2014. Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation. *Nat. Commun.* 5:3077.
7. Malarkey, C. S., M. Bestwick, …, M. E. Churchill. 2012. Transcriptional activation by mitochondrial transcription factor A involves preferential distortion of promoter DNA. *Nucleic Acids Res.* 40:614–624.
8. Gangelhoff, T. A., P. S. Mungalachetty, …, M. E. Churchill. 2009. Structural analysis and DNA binding of the HMG domains of the human mitochondrial transcription factor A. *Nucleic Acids Res.* 37:3153–3164.

9. Wong, T. S., S. Rajagopalan, …, A. R. Fersht. 2009. Biophysical characterizations of human mitochondrial transcription factor A and its binding to tumor suppressor p53. *Nucleic Acids Res.* 37:6765–6783.
10. Ha, T., T. Enderle, …, S. Weiss. 1996. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. USA.* 93:6264–6268.
11. Deniz, A. A., M. Dahan, …, P. G. Schultz. 1999. Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Förster distance dependence and subpopulations. *Proc. Natl. Acad. Sci. USA.* 96:3670–3675.
12. Coats, J. E., Y. Lin, …, I. Rasnik. 2013. Single-molecule FRET analysis of DNA binding and bending by yeast HMGB protein Nhp6A. *Nucleic Acids Res.* 41:1372–1381.
13. Hieb, A. R., A. Gansen, …, J. Langowski. 2014. The conformational state of the nucleosome entry-exit site modulates TATA box-specific TBP binding. *Nucleic Acids Res.* 42:7561–7576.
14. Clegg, R. M. 1992. Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol.* 211:353–388.
15. Gansen, A., K. Tóth, …, J. Langowski. 2009. Structural variability of nucleosomes detected by single-pair Förster resonance energy transfer: histone acetylation, sequence variation, and salt effects. *J. Phys. Chem. B.* 113:2604–2613.
16. Rambo, R. P., and J. A. Tainer. 2013. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature.* 496:477–481.
17. Farge, G., N. Laurens, …, G. J. Wuite. 2012. Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor A. *Nat. Commun.* 3:1013.
18. Bernadó, P., E. Mylonas, …, D. I. Svergun. 2007. Structural characterization of flexible proteins using small-angle x-ray scattering. *J. Am. Chem. Soc.* 129:5656–5664.
19. Bonawitz, N. D., D. A. Clayton, and G. S. Shadel. 2006. Initiation and beyond: multiple functions of the human mitochondrial transcription machinery. *Mol. Cell.* 24:813–825.
20. Farge, G., M. Mehmedovic, …, M. Falkenberg. 2014. In vitro-reconstituted nucleoids can block mitochondrial DNA replication and transcription. *Cell Reports.* 8:66–74.
21. Kaufman, B. A., N. Durisic, …, E. A. Shoubridge. 2007. The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Mol. Biol. Cell.* 18:3225–3236.
22. Kukat, C., K. M. Davies, …, N. G. Larsson. 2015. Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proc. Natl. Acad. Sci. USA.* 112:11288–11293.
23. Dairaghi, D. J., G. S. Shadel, and D. A. Clayton. 1995. Addition of a 29-residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator. *J. Mol. Biol.* 249:11–28.
24. Murugesapillai, D., M. J. McCauley, …, M. C. Williams. 2014. DNA bridging and looping by HMO1 provides a mechanism for stabilizing nucleosome-free chromatin. *Nucleic Acids Res.* 42:8996–9004.
25. McCauley, M. J., J. Zimmerman, …, M. C. Williams. 2007. HMGB binding to DNA: single and double box motifs. *J. Mol. Biol.* 374:993–1004.
26. van Noort, J., S. Verbrugge, …, R. T. Dame. 2004. Dual architectural roles of HU: formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. USA.* 101:6969–6974.

Biophysical Journal, Volume *114*

Supplemental Information

Protein Flexibility and Synergy of HMG Domains Underlie U-Turn Bending of DNA by TFAM in Solution

Anna Rubio-Cosials, Federica Battistini, Alexander Gansen, Anna Cuppari, Pau Bernadó, Modesto Orozco, Jörg Langowski, Katalin Tóth, and Maria Solà

Supporting Figures

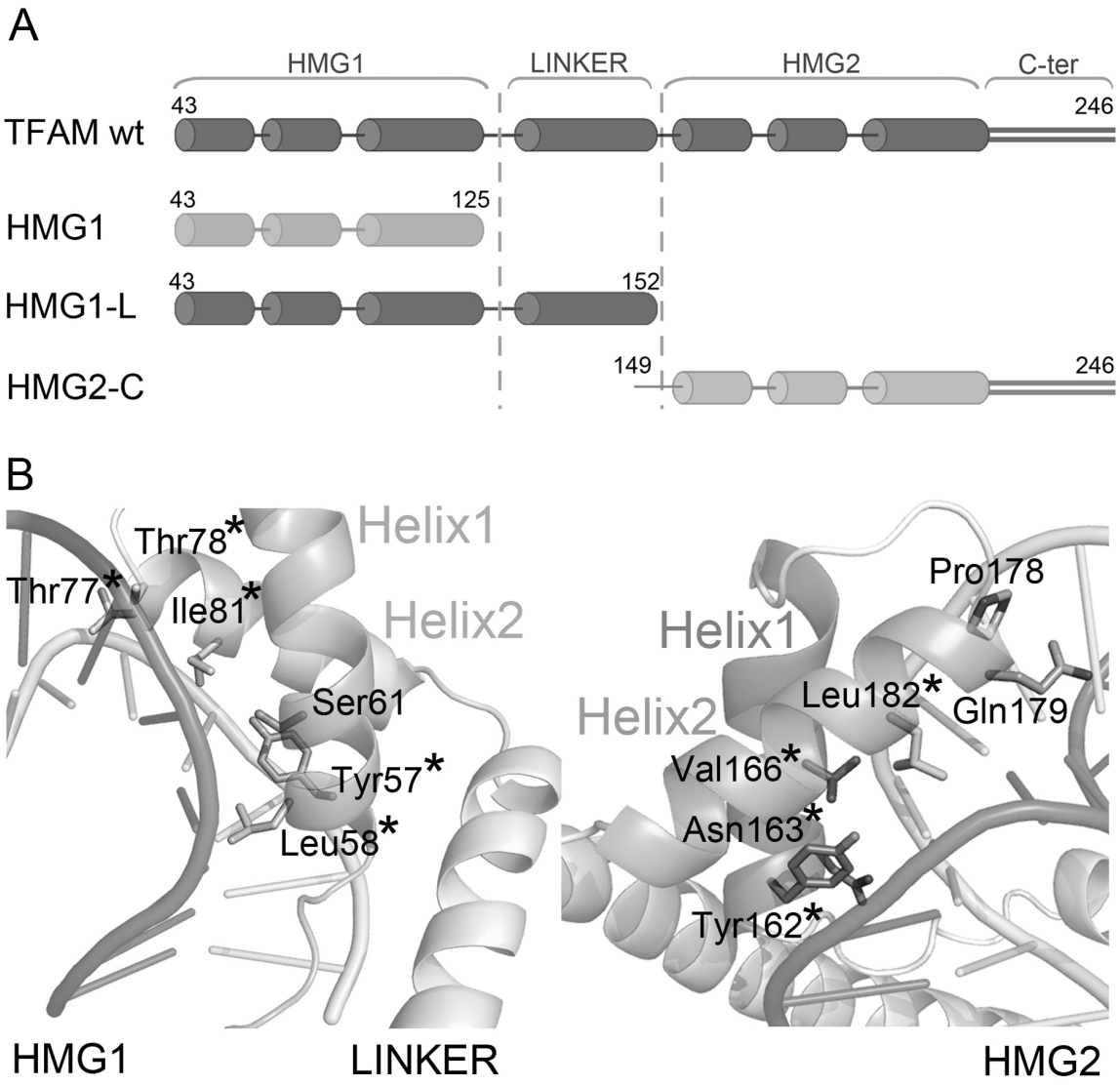

Figure S1. TFAM constructs and relevant regions. (A) Schematic representation of full length TFAM and TFAM domains HMG1 (residues 44-125), HMG1-L (43-152) and HMG2-Cter (149-246) used in smFRET experiments. **(B)** Local interactions of TFAM with DNA. Residues that are in direct contact with DNA bases are shown as sticks in HMG1 (Ser61, Tyr57, Leu58, Thre77, Thr78 and Ile81) and HMG2 (Tyr162, Asn163, Val166, Pro178, Gln179 and Leu182). Residues labelled with black asterisks were mutated to alanine for MD simulations.

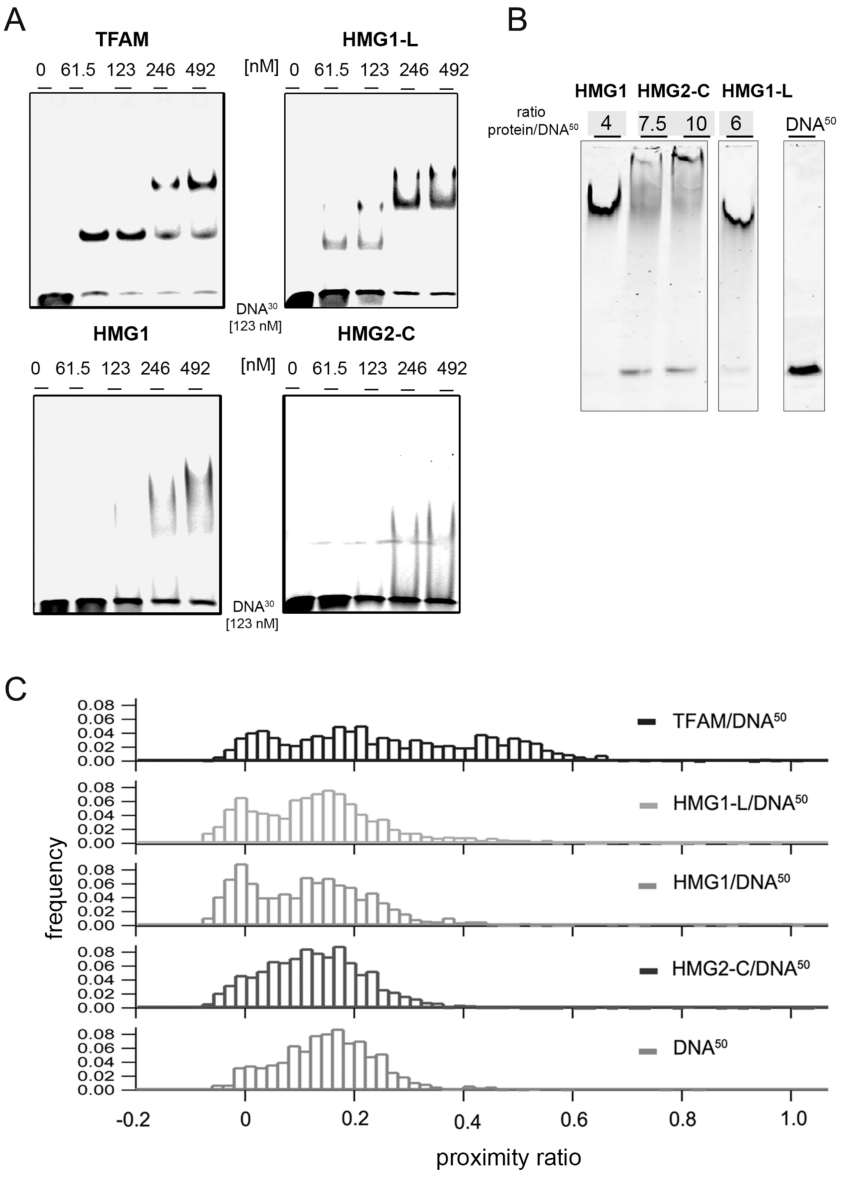

Figure S2. TFAM and TFAM domains do not bind equally strong to DNA. (A) EMSA assays for full-length TFAM and TFAM domains (HMG1-L, HMG1 and HMG2-Cter) with Alexa 594-labelled DNA^{30}. From left to right, protein concentration is increased from 0 to 492 nM versus constant DNA amounts of 123 nM. The gel bands were visualized using the fluorescence signal from Alexa 594 (excitation at 532 nm, detection at 595-625 nm). TFAM shows a higher binding ability than the isolated domains (HMG1 and HMG2-Cter) as it generates well-defined bands with no smear. (**B**) EMSA of TFAM domains HMG1, HMG2-Cter and HMG1-L to DNA^{50} labelled with Alexa 488 and Alexa 594. Complexes were run on a 5% polyacrylamide native gel. The gel bands were visualized by Alexa 594 emission. All TFAM domains bind to labelled-DNA^{50}, and HMG2-C shows the weakest binding. (**C**) Exemplary smFRET histograms for TFAM and TFAM domains in complex with DNA^{50}. Full-length TFAM is shown on top, and HMG1-L, HMG1 and HMG2-Cter in complex with DNA^{50} are shown below. The proximity ratio distribution for free DNA is shown for comparison at the bottom of the panel.

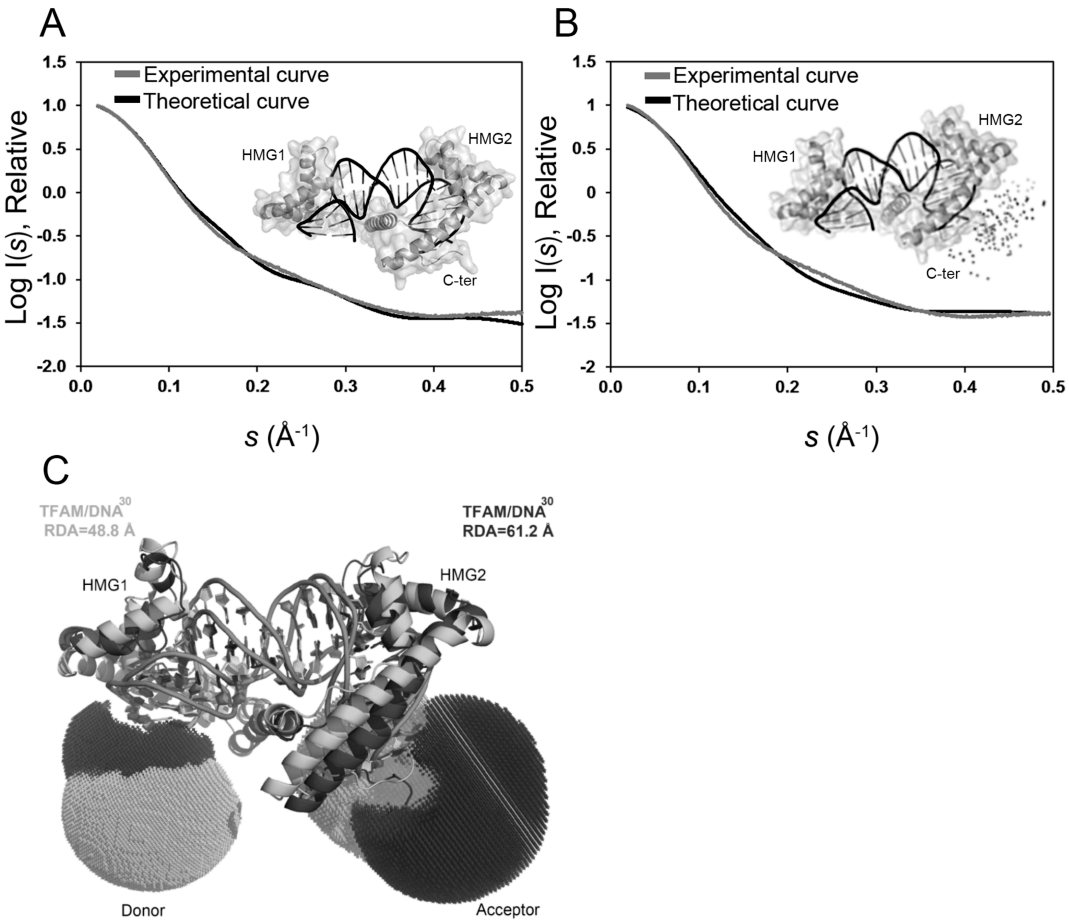

Figure S3. Structural flexibility of the TFAM/LSP complex. **(A)** SAXS analysis of the TFAM/LSP complex. Comparison of the experimental scattering profile (in black) and a theoretical curve based on the TFAM/LSP crystal structure (in grey). The crystallographic structure used to calculate the theoretical curve is shown. Both curves are represented in logarithmic scale as a function of the momentum transfer, $s = 4\pi\sin(\theta)/\lambda$ (2θ, scattering angle; X-ray wavelength $\lambda = 1.5$ Å). The theoretical curve was calculated using CRYSOL. Despite the similarity of both curves, the overall agreement is poor ($\chi 2 = 2.93$), which suggests the presence of intrinsic structural flexibility. **(B)** The same as in (A) but using a subensemble of 50 conformations with full flexibility at TFAM's C-terminal tail (residues Gly226-Cys246). The 50 conformations subensemble used to calculate the theoretical curve are also shown, flexible C-ter$^{\alpha}$ represented as dots. **(C)** Flexibility of TFAM-LSP complex leads to variations in interdye distance. Modelling of interdye distance for two representative TFAM/DNA^{30} structures: the LSP-22 crystal structure (in light grey) and the conformation from the MD simulation that showed the highest *rmsd* divergence from the crystal structure (in dark grey). The accessible volumes of the fluorophores were modelled using the FPS toolkit (1).

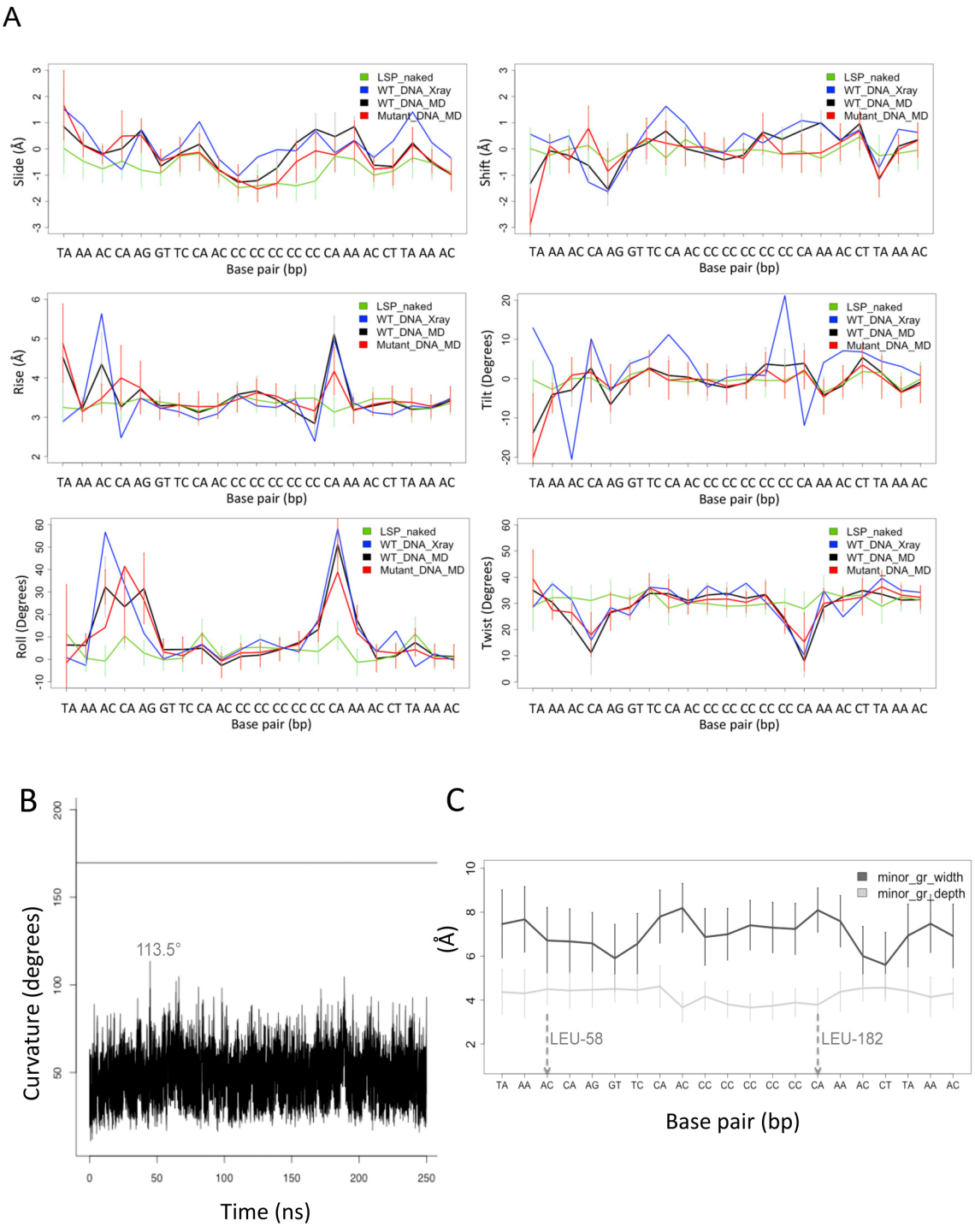

Figure S4. Properties of protein-bound and naked DNA. **(A)** Base pair translational (shift, slide and rise) and rotational (tilt, roll and twist) parameters for the LSP sequence in the X-ray crystal structure (blue) and averaged during the MD simulations for unbound DNA (green), DNA in complex with the wild type

(WT) protein (black) and DNA in complex with the MD-mutant (red). **(B)** Total curvature of the unbound LSP DNA[30] sequence during the MD simulation. The horizontal line at 169 degrees corresponds to the curvature of LSP in the X-ray structure of the complex. The average curvature of unbound DNA during the simulation is (47±12)°. Note that the curvature shows high peaks during the simulation, with the largest curvature reaching 113.5°. **(C)** Minor groove width and depth averaged during the MD simulation of unbound LSP DNA[30] (standard deviation represented as vertical bars). The left vertical arrow highlights the Leu58 intercalation site in the TFAM/LSP complex. This site is characterized by a narrower and slightly deeper minor groove than the Leu182 insertion site (right vertical arrow).

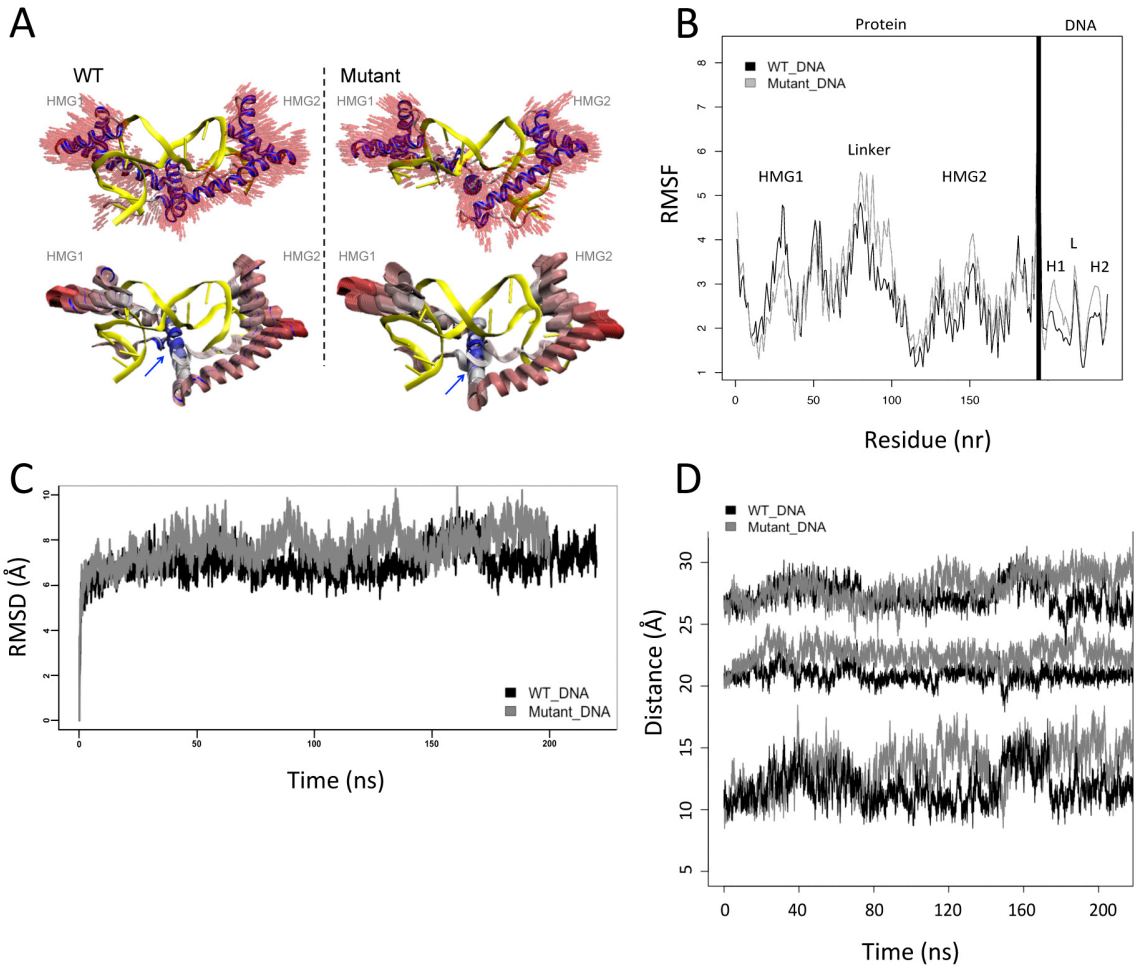

Figure S5. Comparison of the wild type and MD-mutant complexes during the MD simulation. **(A)** Representation of the movement along the PCA (principal component analysis) first normal mode for the WT (left panel) and MD-mutant (right panel) proteins within the protein/DNA complex. This first component describes a significant part of the motion during the simulation of the WT and MD-mutant (10% and 18%, respectively). The two top left and right panels show the directions of the first normal mode. In this case, the first eigenvector describes the two HMG boxes moving back and forth, in a "breathing movement". The bottom panels show the overlay of the protein structures during the MD (thicker structural segments indicate wider displacements). The highest mobility of the protein is represented in red, the lowest in blue. The helices and the HMG box L–shape elbows show highest movement; the hinge point is located in the middle of the linker region (indicated with blue arrows in both panels) around which the protein domains move. The relative movements of the DNA have not been represented for clarity. **(B)** Root mean square fluctuations (RMSF) values of each residue in WT (black line) and MD-mutant (grey line) complexes. The RMSF were computed along the MD simulations for both complexes. The HMG1, HMG2 and linker regions are indicated. H1, H2 and L indicate the DNA regions that contact the corresponding protein domains HMG1, HMG2 and linker, respectively. **(C)** Time-dependent root mean square deviation (RMSD) of all protein-DNA heavy atoms from the WT and the MD-mutant complexes (in black and grey, respectively). The RMSD values were computed relative to the initial crystal structure after relaxation. **(D)** Distance between the protein and the DNA, for the WT (in black) and MD-mutant (in grey) complexes. From top to bottom, distances related to HMG1, HMG2 and the linker are

shown, respectively. Note that for the MD-mutant, HMG2 slightly separates from the DNA at the beginning and during the whole simulation. HMG1 shows similar distances in both complexes. In contrast, the linker shows a prominent irreversible separation in the MD-mutant.

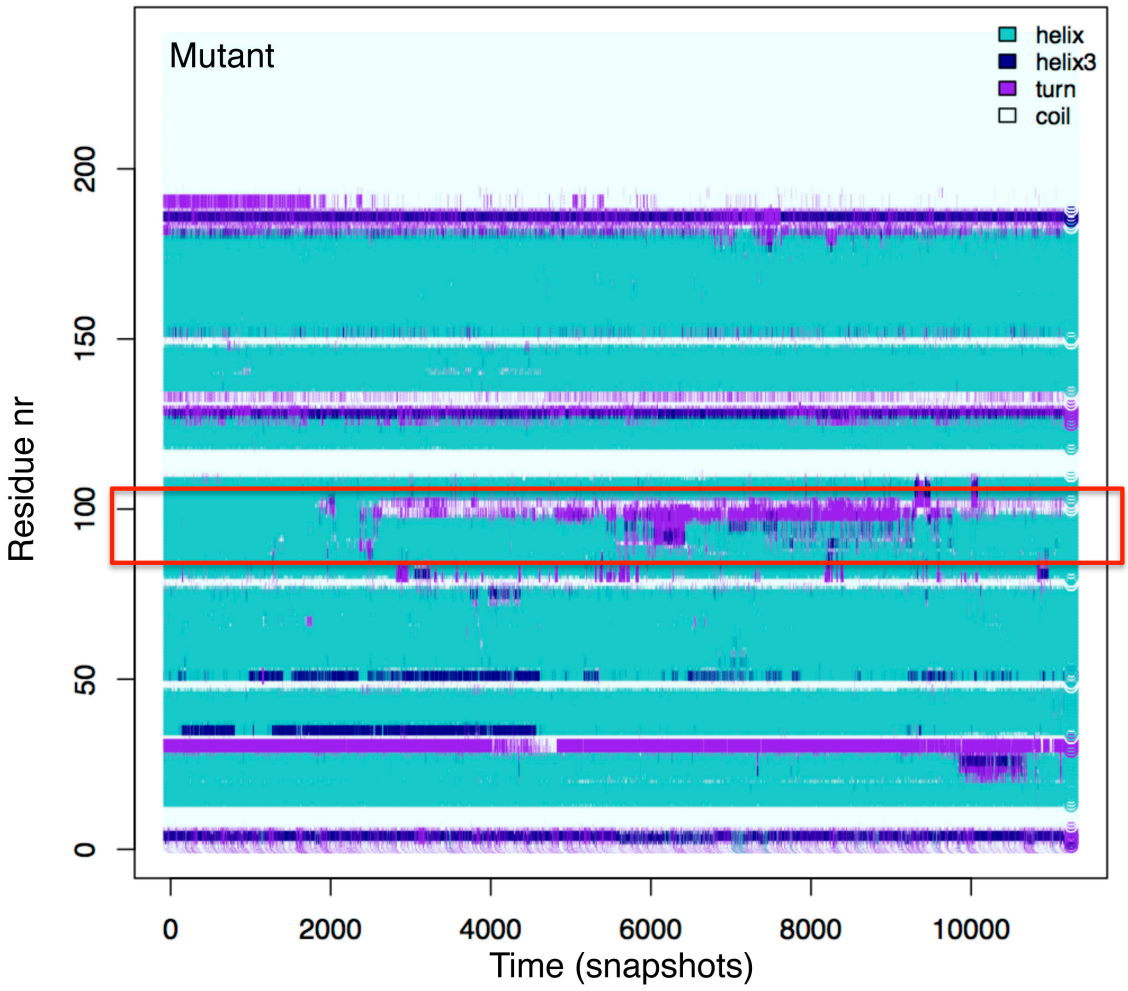

Figure S6. Variability of the MD-mutant secondary structure during the simulation (every 100th snapshot). The secondary structure is maintained except for the linker (red frame) that irreversibly unfolds from the helix conformation (in cyan) to turn or coil (in violet or white, respectively).

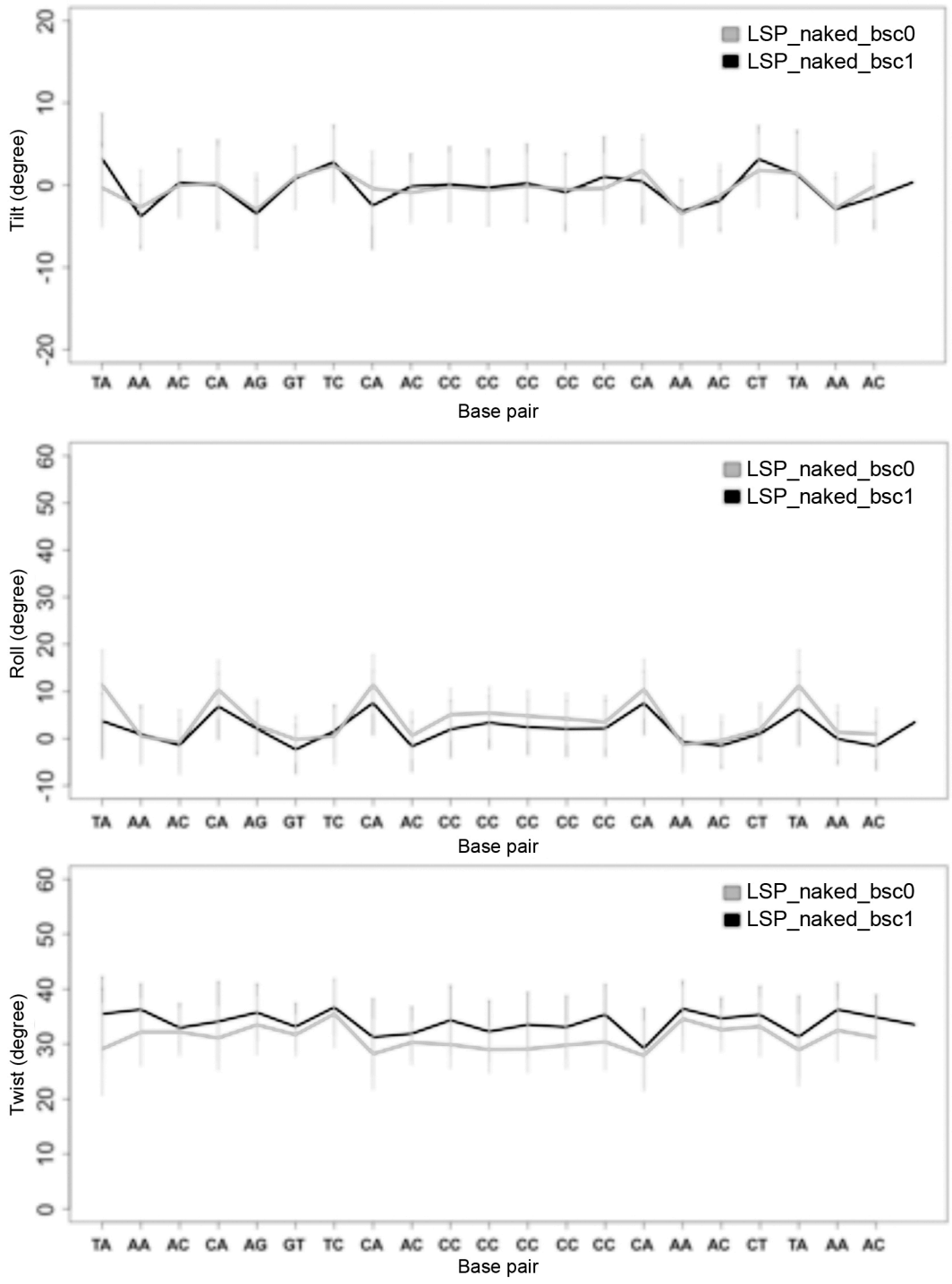

Figure S7. Comparison of the time-averaged base pair rotational parameters (tilt, roll and twist in degrees) for unbound DNA[30] (LSP_naked) during the MD simulations using the force fields parmBSC0 (in grey) and parmBSC1 (in black), respectively.

Supporting Materials and Methods

Design of DNA constructs for smFRET experiments

To determine whether TFAM bends DNA in a U- or a V-shape, we performed smFRET experiments on TFAM in complex with 30 bp (DNA^{30}) and 50 bp (DNA^{50}) long DNAs. Both constructs were centred on the TFAM cognate binding sequence LSP22 and flanked by the corresponding mtDNA sequences on either side, (see Figure 1). The donor fluorophore (Alexa 488) was attached to the 5' end of the cytosine-rich DNA strand, the acceptor (Alexa 594) to the 5' end of the complementary strand. The use of a hexamethylene linker ensured rotational freedom of the dyes. Since the crystal structure predicts that, at the inner U-turn face, the DNA ends are within 20 Å of each other, we intended to place the dyes at the outer face of the U- or V-turn, to minimize the risk of self-quenching at very short distances. The dyes were chosen because a) their Förster radius (R_0 = 55.6 Å) is of the order of the expected interdye distance after protein binding, enabling high sensitivity, and b) their emission peaks are separated by 100 nm (519 nm for Alexa 488 *vs.* 617 nm for Alexa 594), allowing spectral separation with little cross talk. Labelled single DNA strands were purchased from *IBA*, purified by reverse phase HPLC and annealed in TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH = 7.5) with 100 mM NaCl. Samples were heated to 90°C for 5 minutes and cooled to room temperature overnight; the duplex was then separated from non-annealed DNA by HPLC. The sequences of the oligonucleotides were: DNA^{30}: 5'-(Alexa 488)*TC TTT TAA CAG TCA CCC CCC AAC TAA CAC A-3' and 5'-(Alexa 594)*T GTG TTA GTT GGG GGG TGA CTG TTA AAA GA; DNA^{50}: 5'-(Alexa 488)*TGC GGT AG CAC TTT TAA CAG TCA CCC CCC AAC TAA CAC ATT ATT TTC CA-3' and 5'-(Alexa 594)*TC GAA AAT AA GTG TTA GTT GGG GGG TGA CTG TTA AAA GTG CAT ACC GCA. Additional control samples were prepared with the same DNA sequences carrying only the donor fluorophore (donor-only) or only the acceptor fluorophore (acceptor-only).

Protein preparation

TFAM (residues 43-246; UniProt Q00059) was produced as previously reported (2). TFAM domains (Supplementary Figure 1A) HMG1 and HMG2-Cter (residues 44-125 and 149-246, respectively) were cloned into the expression vector pET28b(+) (*Novagen*). The HMG1-L domain (residues 43-152) was cloned using the *In-Fusion*^{TM} system in a pOPINF vector that added an N-terminal 6-histidine tag. TFAM domains were prepared as the WT protein (2), and their folding assessed by circular dichroism at the Biomolecular Screening and Protein Technologies Unit at the Center for Genomic Regulation (CRG, Barcelona, Spain).

Protein and DNA complex formation

The high salt concentration in the protein purification buffer (50 mM HEPES pH 7.5, with 750 mM NaCl, and 5 mM DTT) (2) impairs TFAM/DNA interaction. Therefore, DNA and protein (or protein domains) were mixed at different molar ratios (1:1, 2:1) and dialyzed in three dialysis steps (the last one overnight) against buffers with lower salt down to 100 mM NaCl. After dialysis, the complex formation was checked on a native polyacrylamide gel (see below). For DNA binding assays, complexes at different protein:DNA ratios were incubated on ice for 20 minutes (10 µl final volume).

smFRET measurements require less than 100 pM of labelled sample to ensure single molecule discrimination. To minimize complex dissociation at such low concentrations we added a ten-fold excess of unlabelled complexes to the labelled ones. smFRET experiments were performed at 500 pM total DNA (50 pM of which were labelled). The smFRET buffer consisted of 50 mM Tris pH 7.5, with 100 mM NaCl, 0.01% NP40 to prevent sample aggregation and adsorption to the chamber walls and 1 mM ascorbic acid to improve the photostability of the dyes. The buffer was freshly prepared before each set of experiments and filtered through a polycarbonate membrane with a cut-off of 0.02 µm. Additionally, buffer conditions were optimized by lowering the DTT concentration from 5 mM to 1 mM for smFRET experiments, because biological reducing agents like DTT decrease dye stability (3).

Electrophoretic mobility shift assays

EMSAs were used to check complex formation after dialysis and for DNA binding assays with protein domains. After incubation of protein and DNA for complex formation (see above), samples were loaded on a non-denaturing polyacrylamide gel (5% polyacrylamide, 0.5 x Tris-Borate-EDTA [TBE] buffer). An electric field of 10 V cm^{-1} was applied for 1 h at 4 °C. DNA bands were visualized using either fluorescently labelled or non-labelled DNA. Gels were imaged on a Typhoon 9400 scanner with appropriate filter settings. Double-labelled samples were imaged in three spectral windows (donor channel: excitation at 488 nm, detection at 500–540 nm; acceptor channel: excitation at 532 nm, detection at 595–625 nm; transfer channel: excitation at 488 nm, detection at 595–625 nm). Acceptor-only labelled DNA was visualized using the acceptor channel only. For unlabelled DNA samples, gels were stained with SyBr Gold (Invitrogen) to visualize protein/DNA complexes.

Single molecule FRET experiments

smFRET experiments were performed on a home-built confocal microscope as described in (4, 5). The donor dye (Alexa 488) was excited at 491 nm with a continuous-wave DPSS-laser (Cobalt), whose beam was focused to a diffraction-limited spot by a 60x / NA = 1.2 objective (UPlanApo, Olympus). Donor and acceptor emission were detected in epifluorescence, separated from scattered laser light by a dichroic beam splitter (505DRLP) and imaged onto a 100 µm pinhole, which rejected any out-of-focus light. The effective

observation volume of this setup was about 1 fl. Transmitted fluorescence was split into two detection channels (donor signal: 500-540 nm, acceptor signal 610-700 nm) defined by an infrared blocking filter (700CFSP), a dichroic beam splitter (580DRLP) and appropriate interference filters (donor channel: 520DF40, transfer channel: 610ALP, all filters and dichroic beam splitters from Omega Optical). The signals from the two avalanche photodiodes (SPAD-AQ-14, Perkin-Elmer) were read by a time-correlated single photon counting (TCSPC) board (TimeHarp200, Picoquant) and processed by our own software. The single photon data stream was smoothed by a Lee filter and single molecules were identified as bursts of at least 50 photons with a mutual separation of less than 120 µs. For each burst several parameters were recorded, including the proximity ratio P, see equation S2, burst duration and photon intensity per time. Single molecule distributions of the proximity ratio and other parameters were built and further analysed with IGOR Pro software (WaveMetrics).

Calculation of the proximity ratio from the smFRET measurements

Energy transfer was analysed via the sensitized emission of the acceptor upon selective donor excitation (6). In smFRET, the proximity ratio P, as a measure of energy transfer, is estimated from the detected photon numbers in the donor and acceptor channel per burst, N^0_D and N^0_T. These raw intensities contain contributions from background (B_D and B_T) and donor crosstalk into the acceptor channel (α_{DT}) which have to be corrected for, giving corrected intensities N_D and N_T:

$$\begin{aligned} N_T &= \left(N_T^0 - B_T\right) - \alpha_{DT}\left(N_D^0 - B_D\right) \\ N_D &= \left(N_D^0 - B_D\right) \end{aligned} \qquad \text{(S1)}$$

The proximity ratio is then calculated as

$$P = \frac{N_T}{N_T + N_D} \qquad \text{(S2)}$$

All correction factors were determined in independent control experiments with a donor-only sample and pure buffer solution. For background correction, we first computed the average count rate of background photons for the buffer solution, B_D and B_T, as photons per millisecond. For the i-th single molecule event, b_D and b_T are then multiplied with the duration of the i-th burst, d_i, to yield

$$B_{D,T}{}^{i} = b_{D,T} \cdot d^{i}. \qquad \text{(S3)}$$

Donor crosstalk into the acceptor channel is corrected for by an independent measurement of a donor-only sample that is measured for 5-10 minutes to build a histogram of the proximity ratio P. After background subtraction, the peak value in the P histogram, $P_{D\text{-only}}$, yields the crosstalk factor as

$$\alpha_{DT} = \frac{1}{{}^{1}/_{P_{D-only}} - 1} \quad \text{(S4)}$$

In our analysis, we did not include an additional correction for direct acceptor excitation, since we were only interested in relative changes in P distribution between samples.

SAXS measurements

Small-angle X-ray scattering (SAXS) data of TFAM/LSP22 at 1.7, 3.9 and 10.5 mg mL^{-1} were measured at the BioSAXS beamline BM29 at the European Synchrotron Research Facility (ESRF) in Grenoble, France. Samples were prepared in 50 mM HEPES pH 7.5, 20 mM NaCl, 5 mM DTT, 10% glycerol. For each sample, data were collected in 10 frames of 1 s, at 20°C. Sample scattering curves were averaged and subtracted from the buffer scattering curves using standard procedures (7). Curves obtained at the three concentrations were merged into a single curve that was used for further analysis. The forward scattering intensity $I(0)$ and the radius of gyration R_g were calculated with the Guinier approximation assuming that, at very small angles ($s < 1.3/R_g$), intensity is represented as $I(s) = I(0)\exp(-(sR_g)2/3)$. The molecular weight of the particle was calculated with Scåtter (8). The maximum particle dimension, D_{max}, and the distance distribution were calculated from the scattering pattern with program GNOM (9) using the momentum transfer range $0.015 < s < 0.50$ $Å^{-1}$. The SAXS profile was compared with the crystallographic structure using the program CRYSOL (10). The Ensemble Optimization Method (EOM) was used to study the conformational variability of the TFAM/LSP complex (11). EOM selects from a large pool of random conformations (10,000) a subensemble (50 conformations) that collectively describes the experimental SAXS curve. A large pool of TFAM/LSP conformations was built by adding flexible C-terminal tails to the crystallographic structure. These flexible chains spanning from Gly226 to Cys246 and the his-tag were obtained with Flexible-Meccano (12) and side-chains were added with SCCOMP (13). For each of these models, a theoretical SAXS profile was computed with the program CRYSOL (10) using standard parameters.

Molecular dynamics simulations

MD simulations of free DNA[30] and DNA in complex with wild type (WT) or mutant TFAM were performed using the AMBER 12 package (14). The starting coordinates of the TFAM/LSP complex were taken from the LSP22 X-ray crystal structure (PDB entry 3TQ6). The mutant variant was created by mutating all the residues of the HMG-boxes that directly interact with the DNA bases to alanine (L58A, Y57A, T77A, T78A and I81A from HMG-box1, and Y162A, N163A, V166A, L182A from HMG-box2). The topology coordinates and parameter files of free DNA, mutant and WT complexes were built using the AMBER leap module. All trajectories were obtained using *state-of-the-art* simulation conditions (15) and

the parmbsc0 refinement (16) to the AMBER parm99 force field (17, 18). Each simulated structure was solvated by placing it at the centre of an octahedron box and filling with TIP3P (19) water molecules, at a minimum solute-wall distance of 10 Å. The system was neutralized with the AMBER leap module using Na^+ as counterions. The total system size for free DNA^{30} was 1905 DNA atoms, 58 Na^+ and 37140 water molecules; for wild type/LSP 4721 protein-DNA atoms, 26 Na^+ counter-ions, and 60383 water atoms; and for MD-mutant/LSP 4654 protein-DNA atoms, 26 Na^+ counter-ions, and 60350 water atoms. Each system was simulated in periodic boundary conditions. Electrostatic interactions were calculated with the particle mesh Ewald method (20). The SHAKE algorithm (21) was used to constrain all bond lengths involving hydrogen atoms at their equilibrium values.

Solvent and ions were initially optimized and relaxed by keeping the solute atoms constrained to their initial position with decreasing force constants of 25, 20, 15, 10 and 5 kcal/(mol Å), using our published multistep protocol (22),(23). The subsequent MD procedure consisted of 100 ps of system thermalization (at T = 298 K), minimization and equilibration without any constraint for 1 ns. This was followed by a 300 ns simulation at constant temperature (298 K) and at a constant pressure of one bar at 2 fs time steps. The atomic positions were saved every 1 ps for subsequent analysis.

Trajectory analysis

The AmberTools suite of programs was used to calculate the root-mean-square deviations (RMSD), distances and root-mean-square fluctuations (RMSF) from the trajectories. The VMD program was used for the analysis of the hydrogen bonds (HBs), secondary structure and principal component analysis (24). HBs were defined using a cut-off of 3.5 Å for the distance between H-donor and H-acceptor atoms and an angle of 120 degrees between them. Physical and geometrical descriptors from the MD simulations were used to study DNA deformability at the level of individual base pair steps. Each step is geometrically described by a set of six *helical movements*: three translations (rise [s], slide [l] and shift [f]) and three rotations (twist [w], roll [r] and tilt [t]). We computed the time-averaged helical parameters as well as the curvature of free and protein-bound DNA from the MD trajectories using the program Curves+ (25). The deformability during these movements is described by the stiffness constants (k_i) associated with the displacements with respect to each helical parameter at the equilibrium. The values for the parameters describing the equilibrium geometry and stiffness constants of naked and protein-bound DNA were derived from long atomistic MD simulations of DNA^{30} sequence in water, free (250 ns) or bound (200 ns) to wild type or to mutant TFAM. In detail, to obtain the equilibrium values and associated force constants of each helical parameter for each base pair step, DNA geometries extracted from the MD simulations were projected into a helical reference system using Curves (25). By collecting the values of these helical parameters during the MD, we built a covariance matrix, *C*h, in helical space for each unique base pair step. The inversion of the covariance matrix, multiplied by the Boltzmann constant *(k_B)* and the absolute temperature (*T), gives* stiffness matrices Ξ_h:

$$\Xi_{\mathrm{h}} = k_{\mathrm{B}} T C_{\mathrm{h}}^{-1} = \begin{bmatrix} k_{\mathrm{twist}} & k_{t-r} & k_{t-l} & k_{t-i} & k_{t-s} & k_{t-d} \\ k_{t-r} & k_{\mathrm{roll}} & k_{r-l} & k_{r-i} & k_{r-s} & k_{r-d} \\ k_{t-l} & k_{r-l} & k_{tilt} & k_{l-i} & k_{l-s} & k_{l-d} \\ k_{t-i} & k_{r-i} & k_{l-i} & k_{\mathrm{rise}} & k_{i-s} & k_{i-d} \\ k_{t-s} & k_{r-s} & k_{l-s} & k_{i-s} & k_{\mathrm{shift}} & k_{s-d} \\ k_{t-d} & k_{r-d} & k_{l-d} & k_{i-d} & k_{s-d} & k_{\mathrm{slide}} \end{bmatrix} \quad \text{(S5)}$$

These contain the elastic force constants associated with helical deformation at the base pair step level. The diagonal elements provide the stiffness constants associated with pure rotational (twist, roll and tilt) and translational (rise, shift and slide) deformations within the given step (26),(27). A rough global estimate of each step's stiffness (k_{total}) was determined by the product of the pure stiffness constants $k_w \cdot k_r \cdot k_t \cdot k_s \cdot k_l \cdot k_f$, appearing at the diagonal of the matrix.

Calculations for the naked DNA[30] were also performed using the recent force-field parmBSC1 (28). The comparison with parmBSC0 (Supplementary Figure S7) shows that average values follow the same pattern and fall inside mutual standard deviations as previously verified (28);(29).

Supporting Material References

1. Kalinin, S., T. Peulen, S. Sindbert, P. J. Rothwell, S. Berger, T. Restle, R. S. Goody, H. Gohlke, and C. A. Seidel. 2012. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat Methods 9:1218-1225.
2. Rubio-Cosials, A., J. F. Sydow, N. Jimenez-Menendez, P. Fernandez-Millan, J. Montoya, H. T. Jacobs, M. Coll, P. Bernado, and M. Sola. 2011. Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. Nature structural & molecular biology 18:1281-1289.
3. Aitken, C. E., R. A. Marshall, and J. D. Puglisi. 2008. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. Biophysical journal 94:1826-1835.
4. Gansen, A., K. Toth, N. Schwarz, and J. Langowski. 2009. Structural variability of nucleosomes detected by single-pair Forster resonance energy transfer: histone acetylation, sequence variation, and salt effects. J Phys Chem B 113:2604-2613.
5. Gansen, A., K. Toth, N. Schwarz, and J. Langowski. 2015. Opposing roles of H3- and H4-acetylation in the regulation of nucleosome structure--a FRET study. Nucleic acids research 43:1433-1443.
6. Clegg, R. M. 1992. Fluorescence resonance energy transfer and nucleic acids. Methods Enzymol 211:353-388.
7. Petoukhov, M. V., D. Franke, A. V. Shkumatov, G. Tria, A. G. Kikhney, M. Gajda, C. Gorba, H. D. Mertens, P. V. Konarev, and D. I. Svergun. 2012. New developments in the ATSAS program package for small-angle scattering data analysis. J Appl Crystallogr 45:342-350.
8. Rambo, R. P., and J. A. Tainer. 2013. Accurate assessment of mass, models and resolution by small-angle scattering. Nature 496:477-481.

9. Svergun, D. I. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. 25:495-503
10. Svergun, D. I., C. Barberato, and M. H. J. Koch. 1995. CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. . J. Appl. Crystallogr. 28:768-773.
11. Bernadó, P., E. Mylonas, M. V. Petoukhov, M. Blackledge, and D. I. Svergun. 2007. Structural characterization of flexible proteins using small-angle X-ray scattering. Journal of the American Chemical Society 129:5656-5664.
12. Bernado, P., L. Blanchard, P. Timmins, D. Marion, R. W. Ruigrok, and M. Blackledge. 2005. A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. Proceedings of the National Academy of Sciences of the United States of America 102:17002-17007.
13. Eyal, E., R. Najmanovich, B. J. McConkey, M. Edelman, and V. Sobolev. 2004. Importance of solvent accessibility and contact surfaces in modeling side-chain conformations in proteins. J Comput Chem 25:712-724.
14. Case, D. A., V. Babin, J. T. Berryman, R. M. Betz , Q. Cai , D. S. Cerutti, T. E. I. Cheatham, T. A. Darden, R. E. Duke, and H. Gohlke. 2014. AMBER. University of California, San Francisco.
15. Perez, A., F. J. Luque, and M. Orozco. 2007. Dynamics of B-DNA on the microsecond time scale. Journal of the American Chemical Society 129:14739-14745.
16. 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. Biophysical journal 92:3817-3829.
17. Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J. Am. Chem. Soc. 117:5179–5197.
18. Cheatham, T. E., 3rd, P. Cieplak, and P. A. Kollman. 1999. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. J Biomol Struct Dyn 16:845-862.
19. Jorgensen, W. L., J. Chandrasekhar, and J. D. Madura. 1983. J. Chem. Phys. 79:926-935.
20. 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.
21. Ryckaert, J.-P., G. Ciccotti, and H. J. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. 23:327–341.
22. Dickerson, R. E., and H. L. Ng. 2001. DNA structure from A to B. Proceedings of the National Academy of Sciences of the United States of America 98:6986-6988.
23. Shields, G. C., C. A. Laughton, and M. J. Orozco. 1997. Molecular Dynamics Simulations of the d(T·A·T) Triple Helix. J. Am. Chem. Soc. 119:7463–7469.
24. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. J Mol Graph 14:33-38, 27-38.
25. Lavery, R., M. Moakher, J. H. Maddocks, D. Petkeviciute, and K. Zakrzewska. 2009. Conformational analysis of nucleic acids revisited: Curves+. Nucleic acids research 37:5917-5929.

26. Lankas, F., J. Sponer, J. Langowski, and T. E. Cheatham, 3rd. 2003. DNA basepair step deformability inferred from molecular dynamics simulations. Biophysical journal 85:2872-2883.
27. Olson, W. K., A. A. Gorin, X. J. Lu, L. M. Hock, and V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. Proceedings of the National Academy of Sciences of the United States of America 95:11163-11168.
28. Ivani, I., P. D. Dans, A. Noy, A. Perez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goni, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpi, C. Gonzalez, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case, and M. Orozco. 2016. Parmbsc1: a refined force field for DNA simulations. Nat Methods 13:55-58.
29. Dans, P. D., L. Danilane, I. Ivani, T. Drsata, F. Lankas, A. Hospital, J. Walther, R. I. Pujagut, F. Battistini, J. L. Gelpi, R. Lavery, and M. Orozco. 2016. Long-timescale dynamics of the Drew-Dickerson dodecamer. Nucleic acids research 44:4052-4066.