



## OPEN

## Allostery through protein-induced DNA bubbles

## SUBJECT AREAS:

COMPUTATIONAL  
BIOPHYSICS

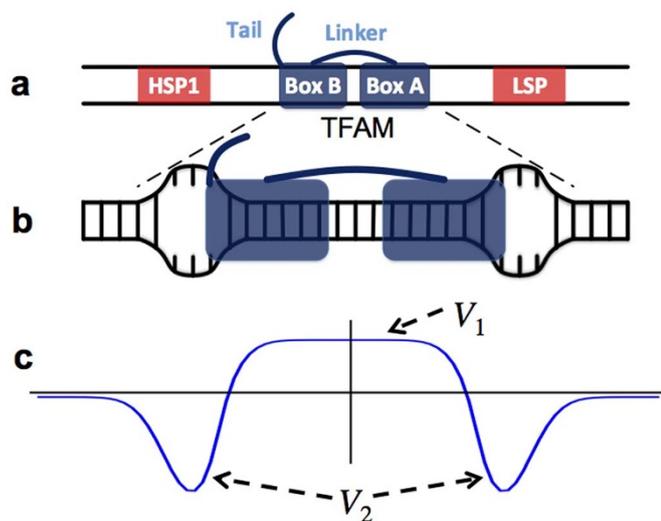
BIOLOGICAL PHYSICS

Joseph J. Traverso<sup>1,2</sup>, Valipuram S. Manoranjan<sup>3</sup>, A. R. Bishop<sup>4</sup>, Kim Ø. Rasmussen<sup>4</sup>  
& Nikolaos K. Voulgarakis<sup>1</sup><sup>1</sup>Department of Mathematics, Washington State University, Richland, WA 99354, USA, <sup>2</sup>Department of Mechanical Engineering, Washington State University, Richland, WA 99354, USA, <sup>3</sup>Department of Mathematics, Washington State University, Pullman, WA 99163, USA, <sup>4</sup>Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.Received  
27 October 2014Accepted  
13 February 2015Published  
12 March 2015Correspondence and  
requests for materials  
should be addressed to  
N.K.V. (n.  
voulgarakis@wsu.edu)

Allostery through DNA is increasingly recognized as an important modulator of DNA functions. Here, we show that the coalescence of protein-induced DNA bubbles can mediate allosteric interactions that drive protein aggregation. We propose that such allostery may regulate DNA's flexibility and the assembly of the transcription machinery. Mitochondrial transcription factor A (TFAM), a dual-function protein involved in mitochondrial DNA (mtDNA) packaging and transcription initiation, is an ideal candidate to test such a hypothesis owing to its ability to locally unwind the double helix. Numerical simulations demonstrate that the coalescence of TFAM-induced bubbles can explain experimentally observed TFAM oligomerization. The resulting melted DNA segment, approximately 10 base pairs long, around the joints of the oligomers act as flexible hinges, which explains the efficiency of TFAM in compacting DNA. Since mitochondrial polymerase (mitoRNAP) is involved in melting the transcription bubble, TFAM may use the same allosteric interaction to both recruit mitoRNAP and initiate transcription.

Precise communication between DNA-binding proteins is critical for many life processes, including the transcription, replication, and organization of DNA. In all of these cases, appropriate proteins form clusters, required to either initiate or execute the entire process. Although the origin of such protein assemblies is unclear, they are often assumed to be driven by direct protein–protein interactions. This assumption limits the role of DNA to simply facilitating the presence of proteins through protein–DNA interactions. Very recently, however, it has been shown that DNA may play a more active role in its own functions<sup>1</sup>. It has been demonstrated both experimentally<sup>1,2</sup> and computationally<sup>3,4</sup> that DNA deformations induced by binding proteins affect the affinity of other nearby proteins. In other words, allosteric signaling through DNA is also possible<sup>5</sup>. Most related studies have been restricted to two types of conformational DNA changes: stretching and bending<sup>1,2,4,6,7</sup>. In this work, we examine how local protein-induced unwinding of the double strand (bubbles) can also facilitate a different type of allosteric signaling. Since the local melting of DNA increases its flexibility<sup>8</sup> and also exposes the genetic code to RNA polymerase, such an allosteric signal potentially regulates both transcription and gene compaction.

Mitochondrial transcription factor A (TFAM) is an excellent example to test such a hypothesis since there is strong evidence that it locally unwinds mtDNA<sup>9–11</sup>. Structurally, TFAM consists of two high mobility group (HMG) box domains A and B connected with a linker and ending with a C-terminal tail attached to Box B<sup>12</sup> (Fig. 1a). TFAM binds specifically close to the light strand promoter (LSP) and heavy strand promoter (HSP1) to form the transcriptional machinery by recruiting transcription factor B2 (TFB2M) and mitochondrial polymerase (mitoRNAP)<sup>9,13–16</sup>. TFAM also binds nonspecifically and plays a critical role in mtDNA compaction<sup>17,18</sup>. The physical mechanism behind the dual function of TFAM is still unclear. Very recent experimental studies have shown that both specific and non-specific binding introduce a sharp U-turn in the mtDNA<sup>18–23</sup>, which, although seemingly vital in forming and appropriately orienting the transcription machinery, does not explain the high efficiency of DNA compaction in the presence of TFAM<sup>18</sup>. Instead, the ability of TFAM to slide rapidly on mtDNA and, upon colliding, to form stable and immobile oligomers seems to be directly related to mtDNA compaction<sup>17</sup>. It has been proposed that such an aggregation of two TFAM proteins melts a region of two to three base pairs (bp) at the point of contact, thus creating fixed flexible hinges that enhance mtDNA flexibility<sup>17</sup>. However, very recent high-resolution experiments have revealed that TFAM oligomers are neither stable nor immobile<sup>24</sup>. It is thus unclear how such highly diffusive hinges of limited lifetime could effectively compact mtDNA molecules.



**Figure 1 | TFAM-DNA interaction model.** (a) Schematic representation of the TFAM structure. Note the orientation of TFAM on DNA relatively to HSP1 and LSP promoters. (b) Illustration of TFAM-DNA complex and the resulting local unwinding of the double helix at the end of each HMG Box domain. (c) Representation of the DNA-TFAM interaction potential (see Eq. 2).  $V_1$  suppresses the double strand at the core of the TFAM-DNA complex and  $V_2$  unwinds it at the two ends of the TFAM.

The ability of TFAM to unwind mtDNA at the end of each HMG box (see Fig. 1b) has two critical consequences. First, it creates two flexible hinges that can potentially increase the flexibility of the DNA. Second, it effectively generates an attractive interaction that drives TFAM oligomerization. We show that the mechanism underlying this allosteric interaction can be an unbalanced force created by the coalescence of two TFAM-induced bubbles. The role of *thermally* induced local openings of the double strand appears to be critical, since it affects both the transmission of the allosteric signal and the stability of the aggregations. The main result of TFAM oligomerization is excitation of a considerably larger bubble (hinge) at the point of contact of two TFAMs, which increases mtDNA flexibility even further and regulates compaction. Interestingly, TFAM binds specifically about 20 bp away from the transcription starting point, which, as we show below, is within the range of the allosteric attraction of two bubbles. Since both TFB2M and mitoRNAP are involved in exciting the transcription bubble<sup>25</sup>, TFAM can help the two proteins excite the transcription bubble and then a coalescence of the transcription bubble and a TFAM bubble could stabilize the transcription machinery.

## Results

To avoid computationally expensive atomistic molecular dynamics simulations, we use the extended Peyrard–Bishop–Dauxois (EPBD) model to describe the local melting dynamics of DNA. EPBD is a one-dimensional (1D) mathematical model with a demonstrated capability for reproducing experimental results on both the mechanical and thermal denaturation of DNA<sup>26–31</sup>. The potential energy of the EPBD model is:

$$V_{\text{DNA}}(\{y_i\}) = \sum_i \left[ D_i (1 - e^{-a_i y_i})^2 + k_{i,i-1} \left( 1 + \rho e^{-\beta(y_i + y_{i-1})} \right) (y_i - y_{i-1})^2 \right], \quad (1)$$

where  $y_i$  describes the distortion of the  $i$ th base pair from its equilibrium position. The hydrogen bonds of a base pair are modeled by Morse potentials (first term in Eq. (1)), while the stacking interactions are described by nonlinear springs (second term in Eq. (1)). The

model, although simple, takes into account the sequence specificity that is reflected in the parameters  $D_i$ ,  $a_i$ ,  $k_{i,i-1}$ ,  $\rho$ , and  $\beta$ . In this study, we will use the values of the parameters in Ref. 27, which have been adjusted to reproduce a variety of experimental observations.

The sliding of TFAMs on DNA is assumed to be purely 1D. The interaction between the protein and DNA has two parts:

$$V_{\text{int}}(\{y_i, R_{ij}\}) = \sum_{i,j} C(y_i) [V_1(R_{ij}) + V_2(R_{ij})], \quad (2)$$

where  $R_{ij} = r_j - ia$  is the distance of the center of the  $j$ th protein from the  $i$ th base pair and  $a$  is the distance between two consecutive base pairs. The first part of the equation,  $V_1(R_{ij}) = A_1 \{ \tanh[\gamma_1(\sigma/2 - R_{ij})] + \tanh[\gamma_1(\sigma/2 + R_{ij})] \}$ , describes the interaction of the binding protein with the DNA backbone, which slightly suppresses the base pair<sup>32</sup>. Here,  $\sigma$  denotes the size of the protein. The second part,  $V_2 = -A_2 \{ \exp[-\gamma_2(\sigma/2 - r)^2] + \exp[-\gamma_2(\sigma/2 + r)^2] \}$ , models the ability of the TFAM to unwind the DNA at the end of the two HMG box domains<sup>9–11</sup>. The coefficient  $C = \tanh[\gamma y_i]$ , where  $\gamma = 1 \text{ \AA}^{-1}$ , controls the strength of the interaction. It increases linearly until the base pair opens ( $y_i \geq 2 \text{ \AA}$ ) and then it plateaus. A schematic representation of the interaction potential is presented in figure 1c.

Since experiments indicate that TFAM proteins do not form oligomers in the absence of DNA (see Ref. 18, for instance), we neglect any possible direct attraction and use a Weeks–Chandler–Andersen (WCA) potential<sup>33</sup> to describe the repulsion (soft sphere) between two TFAMs:

$$V_{\text{WCA}}(r_{ij}) = \begin{cases} 4\epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right] + \epsilon, & r_{ij} \leq 2^{1/6}\sigma, \\ 0, & r_{ij} > 2^{1/6}\sigma \end{cases} \quad (3)$$

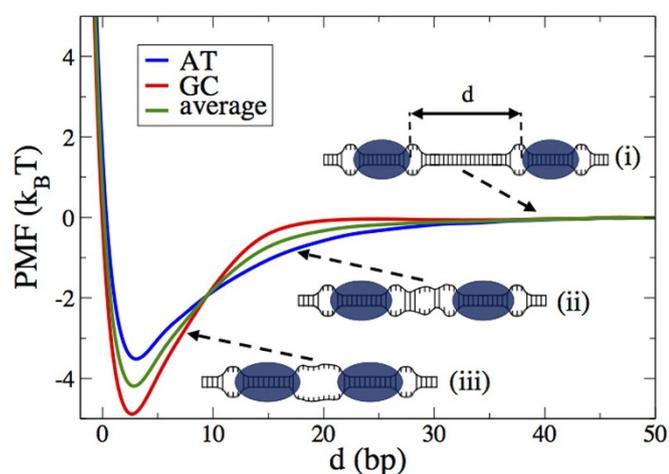
where  $r_{ij}$  is the distance between the centers of the  $i$ th and  $j$ th proteins, and  $\epsilon$  is the interaction strength. The total direct protein–protein interaction energy of a system of multiple proteins is

$$V_{\text{prot}}(\{r_{ij}\}) = \sum_{i>j} V_{\text{WCA}}(r_{ij}). \quad (4)$$

As explained below, the parameters  $A_1$ ,  $A_2$ ,  $\gamma_1$ ,  $\gamma_2$ , and  $\epsilon$  of equations (2) and (3) have been adjusted to reproduce the experimentally observed cooperative binding of TFAM<sup>17</sup> (see Methods). The protein size is assumed to be  $\sigma = 28 \text{ bp}$ , an estimate that is in good agreement with most experimental observations<sup>17,18</sup>. To study the behavior of this TFAM–DNA model, we perform Langevin dynamics simulations at a temperature  $T = 300 \text{ K}$  (see Methods). The potential energy of equation (2) melts a segment three to four bp long at the end of each HMG box.

**Bubble-mediated allosteric protein–protein interaction.** To test our hypothesis that the coalescence of bubbles drives protein aggregation, we perform standard potential of mean force (PMF) calculations (Methods). Figure 2 presents the PMF between two TFAMs in a homogeneous AT and GC molecule, which, as predicted, has an attractive structure. Protein aggregation is triggered by spontaneous thermal openings in the double strand. These openings (or thermal bubbles) exist even at temperatures well below the melting transition and are a result of the interplay between entropy, nonlinearity, and sequence specificity<sup>26,28,34</sup>. The communication between two proteins begins when they diffuse to positions where a spontaneous bubble nucleation of length approximately equal to their surface-to-surface distance is possible (Fig. 2 ii). This local thermal melting reduces the system’s total free energy and creates an unbalanced force that pushes the two proteins toward each other (Fig. 2 iii). This represents a new type of allostery initiated by protein-induced bubbles and transmitted through thermal bubbles.

The depth ( $\sim 4.23 \text{ k}_B T$ ), the average surface-to-surface distance ( $\sim 10 \text{ bp}$ ), and the range ( $\sim 20 \text{ bp}$ ) are three of the main characteristics of the allosteric potential presented in Figure 2. The parameters



**Figure 2 | Allosteric protein-protein interaction.** PMF of a TFAM dimer as a function of the surface-to-surface distance  $d$  for a homogeneous AT (blue) and GC (red) DNA molecule. The green line corresponds to the average of all three profiles. The minimum of all three profiles is located at  $d_0 = 3$  bp. Insets schematically show the role of thermally induced base pair openings in dimer formation. (i) For  $d > 40$  bp the two proteins practically do not feel the presence of each other. (ii) For  $d < 20$  bp thermally induced spontaneous base pair openings create a tunnel of partially open base pairs that connects the two proteins. This produces an unbalance force that drives proteins collapse. (iii) For  $d < 10$  bp the area between the two proteins is completely melted thereby creating a flexible hinge that increases DNA's flexibility.

of equations (2) and (3) were tuned so that the depth provides a cooperative factor of  $\sim 70$ , as estimated in the experimental work of Ref. 17. The coalescence of two small TFAM-induced DNA bubbles can be viewed as the elimination of two half-bubbles, i.e. two forks, from the system. The activation energy of such small forks is associated with the energy cost to unzip a base pair. Thus, the elimination of two forks lowers the free energy of the system by approximately the depth of the interaction potential presented in Figure 2. The average surface-to-surface area includes approximately 10 melted bps, which indicates that TFAM oligomerization provides an additional and significantly larger flexible hinge than the flexible hinge of a monomer would. Thus, TFAM oligomerization could potentially increase the flexibility of a DNA molecule and consequently regulate DNA compaction<sup>17,24,35,36</sup>. In the limit of maximum coverage of DNA by TFAM the energetically most favorable hinge is 3 bp, i.e. equal to the surface-to-surface distance,  $d_0$ , that corresponds to the minimum of PMF (see Fig. 2). Interestingly, this 3 bp melted segment in the limit of high TFAM concentration was also predicted by the authors of Ref. 17 using an independent calculation based on the counter length of DNA. Based on our analysis, the effective size of TFAM is  $\sigma_{\text{eff}} = \sigma + d_0$  or  $\sigma_{\text{eff}} = 31$  bp and the maximum number of TFAMs a DNA molecule can host is  $L_{\text{DNA}}/\sigma_{\text{eff}}$ , where  $L_{\text{DNA}}$  is the length of the DNA. According to Figure 2, a TFAM can attract another TFAM or other proteins from a distance of approximately 20 bp. This result is particularly important when we discuss below the role of TFAM in transcription initiation. The PMF is also sequence dependent. We see that, in homogeneous AT DNA molecules, the range of the potential is longer than in homogeneous GC molecules; however, GC regions support more stable dimerization. Thus, in a realistic DNA molecule, AT-rich regions can facilitate the long-distance transmission of allosteric signals, while GC regions provide a more stable aggregation.

**Reversibly assembled protein aggregates.** Due to the finite depth of the interaction potential, the picture of multiple TFAMs sliding on mtDNA is expected to be a typical example of 1D reversible particle-

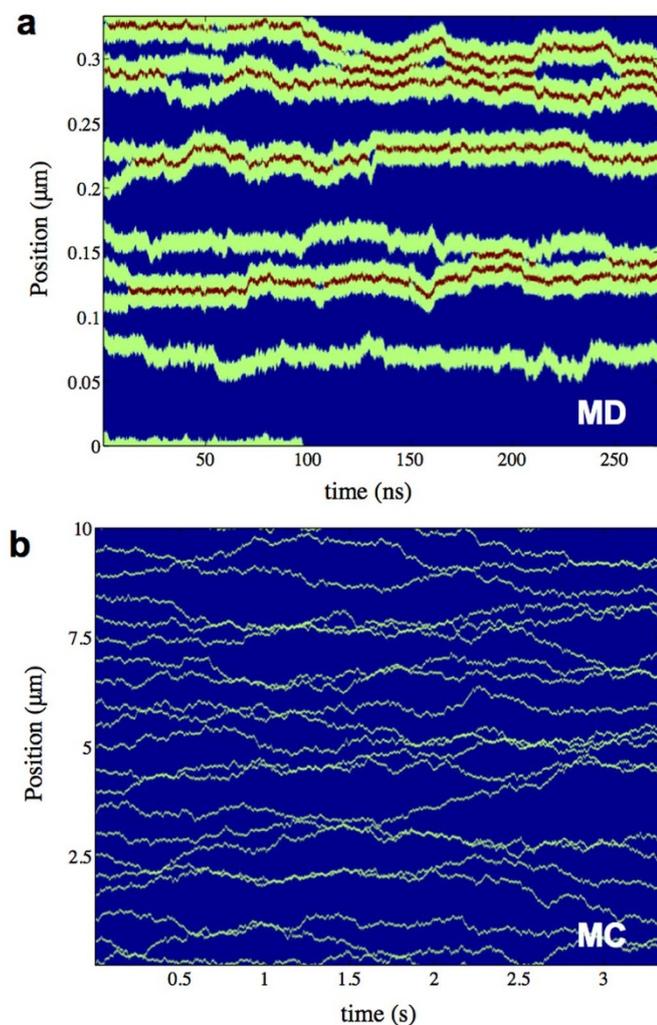
particle aggregation. In such systems, one expects oligomerization and dissociation events, as well as a reduction of mobility due to oligomerization, crowding, or even dynamically arrested states<sup>37</sup>. In general, large bubbles induced by TFAM oligomerization are expected to contribute more to DNA compaction than small hinges. However, their excitation, lifetime, and mobility ultimately determine their effectiveness. A large bubble with a short lifetime or high diffusivity, for instance, would have a very small probability of fully developing and melting DNA at a certain position. The question is, however, to what extent does our model agree with recent experimental observations and, in particular, the data presented in Refs. 17, 18, 24?

To obtain a qualitative picture of the dynamics of the system, we perform a standard Langevin simulation of 10 TFAMs in an 1000 bp ( $\sim 0.33 \mu\text{m}$ ) long mtDNA sequence. Figure 3a shows that the position of all TFAMs as a function of time is qualitatively similar to that in the experimental work of Ref. 24. In agreement with these authors, protein oligomerization and dissociation events, oligomer/monomer diffusion, and entrapment due to sequence specificity are also present in our numerical simulations. However, even the Langevin dynamics of a 1D model cannot access scales similar to the experimental ones ( $\sim \text{sec}$  and  $\sim 10 \mu\text{m}$ ). To overcome this limitation, we implement standard Monte Carlo (MC) simulations of multiple TFAMs interacting with the average PMF shown in Figure 2 (Methods). Figure 3b shows the MC time evolution of a similar system to that in Figure 3a, but for scales directly related to the experimental ones. We emphasize, however, that MC simulations do not take into account the sequence of the DNA; thus, entrapment due to sequence specificity is not observed. Figure 4a shows the distribution of the oligomer size,  $n$ , for different values of the coverage,  $c$ , of DNA by TFAM. The number of large flexible hinges is simply  $n - 1$ . In Figure 4b we show the mean square displacement (MSD) of the TFAMs as a function of time for the same values of  $c$ . We observe three distinct regions. The first region (I) describes the cluster diffusivity prior to collisions. It is purely linear and the slope determines the diffusivity of the system, which scales inversely with the oligomer size appropriately weighted by the distribution of sizes presented in Figure 4a. For intermediate times (region II) the system considerably slows down due to caging effects, i.e. clusters are arrested by nearby clusters and thus only the dynamics within the cage is described. The long time limit (region III) also shows linear behavior that is due to crowding effects<sup>37</sup>. A similar transition from the ballistic regime of individual proteins to region I is also observed but not shown in this plot. All three regimes affect DNA packaging. The slower a TFAM oligomer is, the more stable and long-lived are the developed large hinges.

## Discussion

In this work, we showed that protein-induced local melting of DNA is an alternative allosteric mechanism to drive protein assembly. Below, we discuss how TFAM may use such a mechanism to regulate DNA compaction and transcription initiation. Although we focus on TFAM, we believe that other proteins of the HMG family may also use the same allosteric signaling to control DNA functions [in preparation].

Our simulations support the hypothesis that the mechanism of flexible hinges induced by TFAM oligomerization underlies the compaction of mtDNA by TFAM (see Fig. 5a). More accurately, as we show in this work, the concept of spontaneously generated *diffusive flexible hinges* with *finite lifetimes* is closer to the experimental picture of Ref. 24. Although both small and large hinges contribute to DNA compaction, large hinges are expected to have a more significant impact, since they are energetically more favorable and diffuse much more slowly. Small hinges are more effective in specific binding or entrapment due to sequence specificity. Assuming that the compaction of DNA is primarily regulated by large hinges, one can

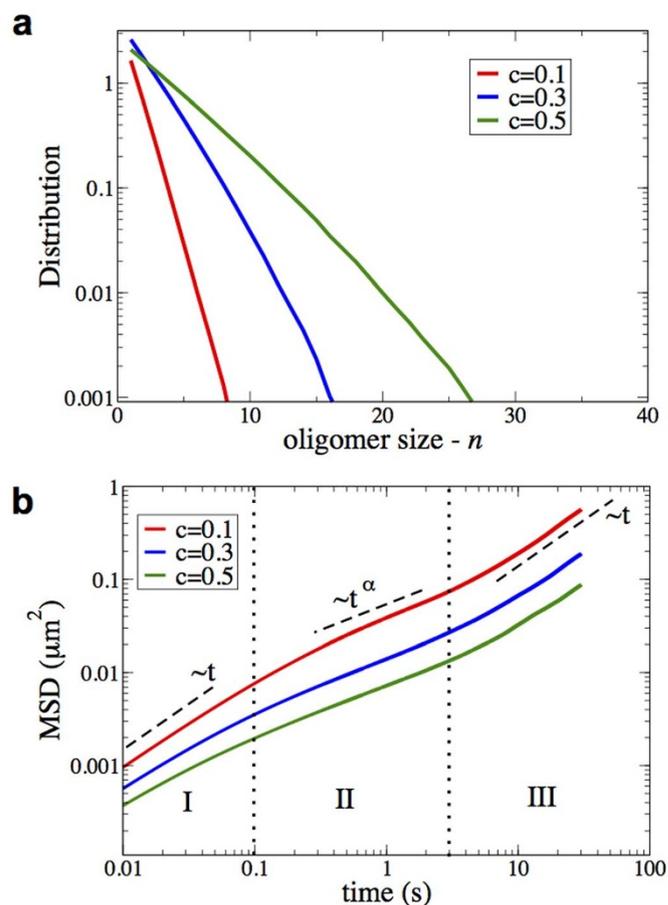


**Figure 3** | TFAM diffusion and oligomerization on DNA. (a) Langevin dynamics of 10 TFAMs diffusing in a 0.33  $\mu\text{m}$  long mtDNA. Green color shows the envelope and position of TFAMs as a function of time. Red areas correspond to large hinges induced by oligomerization. Spontaneous aggregation and dissociation of TFAM oligomers as well as entrapment of oligomers/monomers due to sequence specificity are present in this example. (b) Monte Carlo simulation of 20 TFAMs interacting with the average PMF presented in Figure 2. Picture is similar to a, but for considerably larger length and time scales, thus allowing direct comparison with experiments<sup>17,24</sup>.

use the distribution of hinges presented in Figure 4a to estimate the persistence length of mtDNA for different concentrations of TFAM. If  $P_0$  and  $P_p$  are the persistence lengths of the mtDNA in the absence of TFAM and fully covered by TFAMs, respectively, then, the persistence length for any value of coverage  $c$  can be estimated by<sup>38</sup>

$$P = P_0(1 + qP_0/P_p)^{-1}, \quad (5)$$

where  $q$  is the number of hinges for a given  $c$  normalized by the maximum number of hinges,  $L_{\text{DNA}}/\sigma_{\text{eff}}$ , in a DNA molecule. According to Ref. 17  $P_0 = 45 \text{ nm}$  and  $P_p = 3.9 \text{ nm}$ . The value of  $q$  can be calculated from the distribution of hinges presented in Figure 4a. Figure 5a compares the persistence length estimated by equation (5) with the experimental result of Ref. 17. TFAM coverage is converted to TFAM concentration through the McGhee-von Hippel formula<sup>39</sup> using a cooperativity factor of  $\omega = 70$ , equilibrium protein–DNA binding constant  $K = 1.6 \times 10^6 \text{ M}^{-1}$  (see ref. 17) and  $\sigma_{\text{eff}} = 31 \text{ bp}$  as the protein footprint. Although one should consider the full 3D problem [work in progress], we see that a simple mathemat-

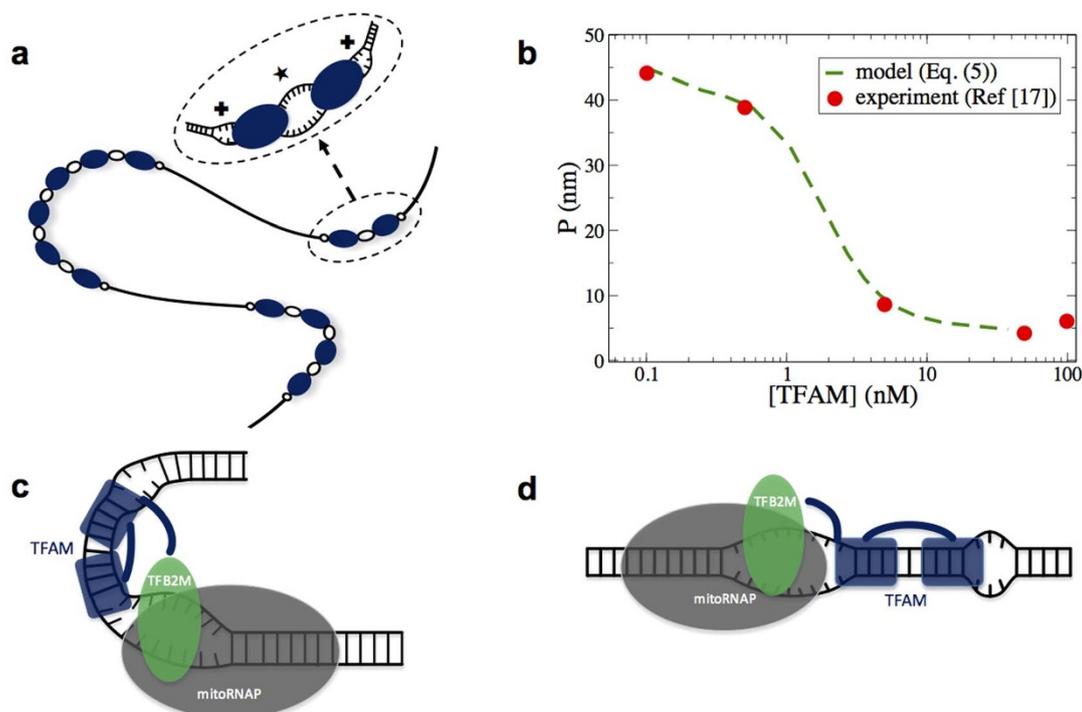


**Figure 4** | Statistics and dynamics of TFAM oligomerization.

(a) Distribution of oligomer size for three different values of fractional coverage of DNA by TFAM,  $c$ . Red line corresponds to  $c = 0.1$ , blue to  $c = 0.3$  and green to  $c = 0.5$ . The number of large flexible hinges in each oligomer is  $n - 1$ . (b) Log-Log plot of the mean square displacement (MSD) as a function of time for the same values of coverage as in a. The vertical lines define three areas of different diffusivity. The linear behavior at short (I) and long times (III) is due to protein oligomerization and crowding effects, respectively<sup>37</sup>. At intermediate times (II), transient arrested states (caging) result in sub-diffusive behavior ( $\alpha < 1$ ). Results for both a and b are obtained by MC simulations of protein dynamics in a 3000 bp long DNA.

ical model can still provide estimates in good agreement with experimental observations.

Appropriate modifications of equations (2) and (3) can also describe the effect of TFAM mutants on compaction efficiency<sup>18</sup>. TFAM mutants missing either Box A or Box B exhibit significantly lower compaction efficiency. Based on our model, with such mutants, which can be described by eliminating one of the two terms of  $V_2$ , only dimerization is possible. This leads to a significantly lower number of flexible hinges, which ultimately reduces the flexibility of the DNA. Mutants with a modified linker (L6) present similar behavior and reduce the efficiency of compaction by approximately the same amount. It appears, as we explain in more detail below, that the L6 mutant reduces only the ability of Box A to unwind the DNA molecule, which can be interpreted in our model by using an asymmetric strength in the expression of  $V_2$ . This modification also provides only dimers, thus reducing the mutant's ability to compact the DNA in a way similar to that of a mutant that is missing Box A. Finally, the dimer mutants presented in Ref. 18 also show a significant reduction in compaction efficiency. Dimer mutants do not interact strongly with each other because their surface has been



**Figure 5 | The Role of TFAM-induced bubbles on DNA compaction and transcription initiation.** (a) Effect of large hinge formation (star) due to TFAM oligomerization on DNA compaction. Large oligomers bend DNA more effectively. Note that small hinges (crosses) develop only when the oligomers diffuse extremely slowly. For monomers, small hinges are evident only upon specific binding (see below). (b) Persistent length of the DNA as a function of TFAM concentration, calculated using equation (5), in comparison with experimental results of Ref. 17. (c) Schematic illustration of LSP activation by TFAM. TFAM first assists TFB2M and mitoRNAP to excite the transcription bubble and then a nucleation of the bubble induced by Box A and the transcription bubble stabilizes the transcription machinery. The small bubble at the end of Box B and the large bubble at the end of Box A act as flexible hinges allowing the DNA to rotate. The interaction between the tail and TFB2M can stabilize a U-turn in agreement with the experimental works of Ref. 18. (d) Same as in c but for the HSP activation. Note that in this case the interaction of the tail with TFB2M prevents rotation around the large flexible hinge and thus a U-turn is not possible, as also observed in Ref. 18.

modified. Their ability, however, to locally unwind the DNA on both sides of the TFAM is preserved. In our model, dimer mutants can be interpreted by increasing the repulsive potential of equation (3). The resulting PMF will have a smaller depth, which will finally reduce the lifetime of TFAM oligomerization and, as a result, reduce the bendability of the DNA. However, it has to be mentioned that, even in the case of a perfect dimer mutant, which is the limit of hard spheres, caging effects can also provide some large flexible hinges that could contribute to DNA compaction<sup>37</sup>.

According to our hypothesis, in an intermediate step, TFAM-induced bubbles first assist TFB2M and mitoRNAP to excite the transcription bubble and then a coalescence of the two bubbles stabilizes the transcription machinery. The size of the resulting bubble (~10 bp) is consistent with the typical size of transcription bubbles<sup>25,40–42</sup>. Since TFAM can melt DNA in both Boxes A and B, the same mechanism can be used to activate transcription at LSP and HSP1, as presented in Figures 5c and 5d, respectively. This leads to the creation of a large hinge on the promoter's side and a small hinge on the other side of TFAM. These two hinges in combination with the strong interaction between the TFAM tail and TFB2M can also explain why a U-turn is present in LSP and not necessarily present in HSP1 (see Figs. 5c and 5d)<sup>18</sup>. It is worth noting that dimer mutants do not affect transcriptional activity, which is in accordance with our hypothesis, mentioned above, that dimer mutants preserve the ability to locally melt the double strand. Additionally, it indicates that a dimer mutant modifies the repulsive interaction between TFAMs but not necessarily the repulsion between a TFAM and other proteins. In Ref. 18, the L6 mutant appears unable to activate transcription in LSP. According to our hypothesis, this observation implies that L6 does not melt the DNA in Box A and consequently cannot

recruit TFB2M and mitoRNAP by using the mechanism described above. That L6 can unwind mtDNA only at Box B is further supported by the fact that L6 activates HSP1. It can also activate LSP only upon interchanging the box domains of TFAM<sup>18</sup>.

## Methods

**Langevin Dynamics.** The Langevin equations of motion for the base pairs and proteins are, respectively,

$$m \partial^2 y_i / \partial t^2 = -\partial[V_{\text{DNA}} + V_{\text{int}}] / \partial y_i - \eta m \partial y_i / \partial t + \tilde{\xi}_i(t), \quad (6)$$

and

$$m_p \partial^2 r_j / \partial t^2 = -\partial[V_{\text{prot}} + V_{\text{int}}] / \partial r_j - \eta_p m_p \partial r_j / \partial t + \tilde{\Xi}_j(t), \quad (7)$$

where  $i = 1, 2, \dots, N_b$  and  $j = 1, 2, \dots, N_p$ . Here,  $N_b$  represents the total number of base pairs and  $N_p$  the number of proteins sliding on the DNA. The molecular weight of a base pair is  $m = 600$  Da and for a TFAM is  $m_p = 29$  kDa. The potential energy  $V_{\text{DNA}}$ ,  $V_{\text{int}}$ , and  $V_{\text{prot}}$  are given by equations (1)–(4). The parameters of  $V_{\text{int}}$  and  $V_{\text{prot}}$  were fitted to reproduce the experimentally observed cooperativity factor of TFAM binding affinity. Specifically  $A_1 = 0.025$  eV,  $A_2 = 0.13$  eV,  $\gamma_1 = 2 \text{ \AA}^{-1}$ ,  $\gamma_2 = 0.225 \text{ \AA}^{-2}$ , and  $\varepsilon = 0.125$  eV. The phenomenological Langevin friction coefficients are  $\eta = 0.1 \text{ ps}^{-1}$  (for the base pairs), and  $\eta_p = 0.1 \text{ ps}^{-1}$  (for the proteins). The stochastic forces  $\tilde{\xi}_i(t)$  and  $\tilde{\Xi}_j(t)$  are modeled as Gaussian random noise with covariances of  $\langle \tilde{\xi}_i(t) \tilde{\xi}_i(t') \rangle = 2\eta m k_B T \delta_{i,i'} \delta(t-t')$  and  $\langle \tilde{\Xi}_j(t) \tilde{\Xi}_j(t') \rangle = 2\eta_p m_p k_B T \delta_{j,j'} \delta(t-t')$ , respectively, where  $T$  is the temperature and  $k_B$  is the Boltzmann constant. The equations of motion were integrated numerically using a second order Runge-Kutta method<sup>43</sup> with periodic boundary conditions. The time step  $dt = 0.001$  ps ensured stable and accurate simulations. For each simulation the system was initially thermalized for 50 ns before starting to monitor the trajectories.

**Potential of Mean Force.** The effective force between two proteins was probed through a harmonic spring of strength  $k$  connecting the centers of the two proteins.



The equilibrium length  $L_0$  of the spring varied between  $\sigma - 3$  bp and  $\sigma + 100$  bp. For each  $L_0$ , we performed 100 independent Langevin simulations of 1  $\mu$ s duration to compute the average inter-protein distance  $\langle r_{ij} \rangle$ . The mean force between the two proteins was estimated by  $F = k(L_0 - \langle r_{ij} \rangle)$ . The potential of mean force (PMF) was then calculated by a simple integration of the computed force in space.

**Monte Carlo Simulations.** TFAM proteins interact through the average PMF presented in Figure 2 (green line). In each MC step all proteins are moved by  $dx = \sqrt{2D_0 dt}$ , where  $D_0 = 0.08 \mu\text{m}^2/\text{s}$  is the reference diffusivity as calculated in Ref. 24 and  $dt = 0.003$  s is the time step of an MC step. Each trial move of a protein is accepted/rejected based on the standard Metropolis algorithm. Before each MC sampling the system was thermalized for 1 sec. All results were obtained by averaging 100 independent MC simulations.

- Kim, S. *et al.* Probing Allostery Through DNA. *Science* **339**, 816–819 (2013).
- Chenoweth, D. M. & Dervan, P. B. Allosteric modulation of DNA by small molecules. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13175–13179 (2009).
- Xu, X., Ge, H., Gu, C., Gao, Y. Q. & Wang, S. S. Modeling Spatial Correlation of DNA Deformation: DNA Allostery in Protein Binding. *J. Phys. Chem. B* **117**, 13378–13387 (2013).
- Brackley, C. A., Taylor, S., Papantonis, A., Cook, P. R. & Marenduzzo, D. Nonspecific bridging-induced attraction drives clustering of DNA-binding proteins and genome organization. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E3605–E3611 (2013).
- Chaires, J. B. Allostery: DNA does it, too. *ACS Chem. Biol.* **3**, 207–209 (2008).
- Koslover, E. & Spakowitz, A. Twist- and Tension-Mediated Elastic Coupling between DNA-Binding Proteins. *Phys. Rev. Lett.* **102**, 178102 (2009).
- Zhang, H. & Marko, J. F. Intrinsic and force-generated cooperativity in a theory of DNA-bending proteins. *Phys. Rev. E* **82** (2010).
- Forties, R. A., Bundschuh, R. & Poirier, M. G. The flexibility of locally melted DNA. *Nucleic Acids Res.* **37**, 4580–4586 (2009).
- Litonin, D. *et al.* Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro. *J. Biol. Chem.* **285**, 18129–18133 (2010).
- Giese, K., Pagel, J. & Grosschedl, R. Functional analysis of DNA bending and unwinding by the high mobility group domain of LEF-1. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12845–12850 (1997).
- Werner, M. H., Gronenborn, A. M. & Clore, G. M. Intercalation, DNA Kinking, and the Control of Transcription. *Science* **271**, 778–784 (1996).
- Kanki, T. *et al.* Architectural Role of Mitochondrial Transcription Factor A in Maintenance of Human Mitochondrial DNA. *Mol. Biol. Cell* **24**, 9823–9834 (2004).
- Shi, Y. *et al.* Mammalian transcription factor A is a core component of the mitochondrial transcription machinery. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16510–16515 (2012).
- Malarkey, C. S., Bestwick, M., Kuhlwlilm, J. E., Shadel, G. S. & Churchill, M. E. A. Transcriptional activation by mitochondrial transcription factor A involves preferential distortion of promoter DNA. *Nucleic Acids Res.* **40**, 614–624 (2012).
- Kim, H., Tang, G.-Q., Patel, S. S. & Ha, T. Opening-closing dynamics of the mitochondrial transcription pre-initiation complex. *Nucleic Acids Res.* **40**, gkr736–380 (2011).
- Falkenberg, M. *et al.* Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nat. Genet.* **31**, 289–294 (2002).
- Farge, G. *et al.* Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor A. *Nat. Commun.* **3**, 1013 (2012).
- Ngo, H. B., Lovely, G. A., Phillips, R. & Chan, D. C. Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation. *Nat. Commun.* **5**, 3077 (2014).
- Rubio-Cosials, A. *et al.* Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. *Nat. Struct. Mol. Biol.* **18**, 1281–1289 (2011).
- Kukat, C. & Larsson, N.-G. mtDNA makes a U-turn for the mitochondrial nucleoid. *Trends Cell Biol.* **23**, 457–463 (2013).
- Rubio-Cosials, A. & Solá, M. U-turn DNA bending by human mitochondrial transcription factor A. *Curr. Opin. Struct. Biol.* **23**, 116–124 (2013).
- Ngo, H. B., Kaiser, J. T. & Chan, D. C. The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. *Nat. Struct. Mol. Biol.* **18**, 1290–1296 (2011).
- Hallberg, B. M. & Larsson, N.-G. TFAM forces mtDNA to make a U-turn. *Nat. Struct. Mol. Biol.* **18**, 1179–1181 (2011).
- Heller, I. *et al.* STED nanoscopy combined with optical tweezers reveals protein dynamics on densely covered DNA. *Nat. Methods* **10**, 910–916 (2013).

- Sologub, M., Litonin, D., Anikin, M., Mustaev, A. & Temiakov, D. TFB2 Is a Transient Component of the Catalytic Site of the Human Mitochondrial RNA Polymerase. *Cell* **139**, 934–944 (2009).
- Dauxois, T., Peyrard, M. & Bishop, A. Entropy-driven DNA denaturation. *Phys. Rev. E* **47**, R44–R47 (1993).
- Alexandrov, B. S. *et al.* A nonlinear dynamic model of DNA with a sequence-dependent stacking term. *Nucleic Acids Res.* **37**, 2405–2410 (2009).
- Campa, A. & Giansanti, A. Experimental tests of the Peyrard-Bishop model applied to the melting of very short DNA chains. *Phys. Rev. E* **58**, 3585–3588 (1998).
- Voulgarakis, N. K., Redondo, A., Bishop, A. & Rasmussen, K. Probing the Mechanical Unzipping of DNA. *Phys. Rev. Lett.* **96**, 248101 (2006).
- Voulgarakis, N. K., Redondo, A., Bishop, A. R. & Rasmussen, K. O. Sequencing DNA by Dynamic Force Spectroscopy: Limitations and Prospects. *Nano Lett.* **6**, 1483–1486 (2006).
- Alexandrov, B. S. *et al.* Specificity and Heterogeneity of Terahertz Radiation Effect on Gene Expression in Mouse Mesenchymal Stem Cells. *Sci. Rep.* **3**, 1184; DOI:10.1038/srep01184 (2013).
- Farge, G. *et al.* In vitro-Reconstituted Nucleoids Can Block Mitochondrial DNA Replication and Transcription. *Cell Rep.* **8**, 66–74 (2014).
- Weeks, J. D. Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids. *J. Chem. Phys.* **54**, 5237 (1971).
- Voulgarakis, N. K., Kalosakas, G., Rasmussen, K. O. & Bishop, A. R. Temperature-dependent signatures of coherent vibrational openings in DNA. *Nano Lett.* **4**, 629–632 (2004).
- Yan, J. & Marko, J. Localized Single-Stranded Bubble Mechanism for Cyclization of Short Double Helix DNA. *Phys. Rev. Lett.* **93**, 108108 (2004).
- Vafabakhsh, R. & Ha, T. Extreme Bendability of DNA Less than 100 Base Pairs Long Revealed by Single-Molecule Cyclization. *Science* **337**, 1097–1101 (2012).
- Babu, S., Gimel, J. C. & Nicolai, T. Self-diffusion of reversibly aggregating spheres. *J. Chem. Phys.* **127**, 054503 (2007).
- Rouzina, I. & Bloomfield, V. A. DNA bending by small, mobile multivalent cations. *Biophys. J.* **74**, 3152–3164 (1998).
- Berg, O. G., Winter, R. B. & Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. I. Models and theory. *Biochemistry* **20**, 6929–6948 (1981).
- Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. Structural Basis of Transcription Initiation: An RNA Polymerase Holoenzyme-DNA Complex. *Science* **296**, 1285–1290 (2002).
- Korzheva, N. *et al.* A Structural Model of Transcription Elongation. *Science* **289**, 619–625 (2000).
- Zhang, G. *et al.* Crystal Structure of Thermus aquaticus Core RNA Polymerase at 3.3 Å Resolution. *Cell* **98**, 811–824 (1999).
- Honeycutt, R. Stochastic Runge-Kutta algorithms. I. White noise. *Phys. Rev. A* **45**, 600–603 (1992).

## Acknowledgements

J.J.T. acknowledges support from Auvil and Carson fellowships of Washington State University.

## Author contributions

N.K.V. conceived and designed the research, carried out the simulations, and analyzed the results. J.J.T. carried out part of PMF calculations. V.S.M., A.R.B. and K.O.R. contributed with the interpretation of the data and gave conceptual advice. All authors wrote the manuscript.

## Additional information

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Traverso, J.J., Manoranjan, V.S., Bishop, A.R., Rasmussen, K.O. & Voulgarakis, N.K. Allostery through protein-induced DNA bubbles. *Sci. Rep.* **5**, 9037; DOI:10.1038/srep09037 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>