Published online 28 November 2016

Nucleic Acids Research, 2017, Vol. 45, No. 2 ***951–967***
doi: 10.1093/nar/gkw1147

DNA structure directs positioning of the mitochondrial genome packaging protein Abf2p

Arka Chakraborty[1], Sébastien Lyonnais[1], Federica Battistini[2,3], Adam Hospital[2,3], Giorgio Medici[1], Rafel Prohens[4], Modesto Orozco[2,3,5], Josep Vilardell[6,7] and Maria Solà[1,*]

[1]Structural MitoLab, Department of Structural Biology, "Maria de Maeztu" Unit of Excellence, Molecular Biology Institute Barcelona (IBMB-CSIC), Barcelona 08028, Spain, [2]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac 10-12, Barcelona 08028, Spain, [3]Joint BSC-IRB Research Program in Computational Biology, Baldiri Reixac 10-12, Barcelona 08028, Spain, [4]Unitat de Polimorfisme i Calorimetria, Centres Científics i Tecnològics, University of Barcelona, Barcelona 08028, Spain, [5]Department of Biochemistry and Biomedicine, University of Barcelona, Barcelona 08028, Spain, [6]Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluis Companys 23, Barcelona 08010, Spain and [7]Molecular Genomics Department, Molecular Biology Institute Barcelona (IBMB-CSIC), Barcelona, 08028, Spain

Received June 1, 2016; Revised October 16, 2016; Editorial Decision October 31, 2016; Accepted November 01, 2016

ABSTRACT

The mitochondrial genome (mtDNA) is assembled into nucleo-protein structures termed nucleoids and maintained differently compared to nuclear DNA, the involved molecular basis remaining poorly understood. In yeast (*Saccharomyces cerevisiae*), mtDNA is a ∼80 kbp linear molecule and Abf2p, a double HMG-box protein, packages and maintains it. The protein binds DNA in a non-sequence-specific manner, but displays a distinct 'phased-binding' at specific DNA sequences containing poly-adenine tracts (A-tracts). We present here two crystal structures of Abf2p in complex with mtDNA-derived fragments bearing A-tracts. Each HMG-box of Abf2p induces a 90° bend in the contacted DNA, causing an overall U-turn. Together with previous data, this suggests that U-turn formation is the universal mechanism underlying mtDNA compaction induced by HMG-box proteins. Combining this structural information with mutational, biophysical and computational analyses, we reveal a unique DNA binding mechanism for Abf2p where a characteristic N-terminal flag and helix are crucial for mtDNA maintenance. Additionally, we provide the molecular basis for A-tract mediated exclusion of Abf2p binding. Due to high prevalence of A-tracts in yeast mtDNA, this has critical relevance for nucleoid architecture. Therefore, an unprecedented A-tract mediated protein positioning mechanism regulates DNA packaging proteins in the mitochondria, and in combination with DNA-bending and U-turn formation, governs mtDNA compaction.

INTRODUCTION

Mitochondria, the organelles to which aerobic eukaryotes owe their existence, possess their own genome, the mitochondrial DNA (mtDNA). This mtDNA codes for components of the electron transport system, which is essential for adenosine triphosphate (ATP) production (1–3). In addition, it encodes ribosomal proteins, ribosomal RNA and tRNAs (3,4). A species-specific High-Mobility Group-box (HMG-box) protein encoded in the nucleus is the main architectural component that compacts and packages mtDNA into nucleoprotein structures, termed nucleoids (5). In the budding yeast *Saccharomyces cerevisiae*, y-mtDNA is a predominantly linear molecule (6–8) of 70–85 kilo base pairs (kbps), which is packaged by the ARS (autonomously replicating sequence)-binding factor 2 protein, Abf2p (9–12).

Saccharomyces cerevisiae is a facultative aerobe that grows in both fermentable (e.g. glucose) and non-fermentable (e.g. glycerol) media. In non-fermentable media, the oxidative phosphorylation pathway involves the electron transport system (13), thus requiring intact, transcriptionally active y-mtDNA. Deletion of the *Abf2* gene causes loss of mtDNA, resulting in yeast cells that grow in fermentable media but cannot subsequently survive in non-fermentable conditions (9). This suggests that Abf2p holds an essential role in mtDNA maintenance (12,14). Moreover, a 2- and 3-fold increase in Abf2p levels leads to a 2-fold increase in y-mtDNA copy number, while a 8- to 10-fold Abf2p overexpression causes rapid loss of y-mtDNA and generation of mtDNA-depleted (petite ρ^0)

*To whom correspondence should be addressed. Tel: +34 9340 34950; Fax: +34 9340 34979; Email: maria.sola@ibmb.csic.es

© The Author(s) 2016. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

mutant cells (14). The latter has been related to excessive y-mtDNA compaction that would impair the access of the replicating machinery to the DNA. Thus, mtDNA transactions and compaction need to be in close balance and are regulated by Abf2p levels.

Abf2p organizes y-mtDNA into nucleoids via sequence-non-specific interactions and DNA bending (15,11). However, the protein shows distinct patterns of DNA 'phased binding' at ARS, rich in A-T bases and poly-adenine tracts, such as the ARS1 found in the nucleus, and the 64 bp region downstream of the mitochondrial origin of replication *ori5*/rep2 (9,10) (Figure 1A). This phased binding was attributed to Abf2p exclusion from a 11 bp consensus motif 5′-(A/T)AAA(T/C)ATAAA(A/T)-3′ (ARS-c) found in ARS1, and from its near matches (ARS-m, 9 or 10 out of 11 bp) in the 64 bp sequence (9,10). From a functional point of view, ARS-c promotes DNA replication while ARS-m, though not as efficient, can maintain similar ARS activity when present in sufficient number of copies (16). Such ARS activity in the mitochondria has been related to amplification of residual fragments of wild-type mtDNA that, in the yeast ρ^- cells (cells that have lost large portions of the wild-type mtDNA), are reiterated as concatamers, restoring the total amount of mtDNA to that found in ρ^+ (wild-type) cells (17). Among these fragments, the mentioned 64 bp region and a derived 35 bp sub-region are the shortest reiterated DNA sequences described so far (17); Figure 1A. The mechanism by which these sequences induce Abf2p phased binding remains to be elucidated.

Abf2p is predicted to consist of two HMG-box domains connected in tandem by a 10-aminoacid (aa) linker (9). In general terms, HMG-box domains are L-shaped and bind and bend the DNA by their concave surface. At this surface, amino acid side chains insert between DNA steps disrupting base-pair stacking and thus contributing to DNA distortions. Although both Abf2p and its human counterpart, the mitochondrial transcription factor A (TFAM), contain two HMG-boxes and show sequence-non-specific DNA binding, the respective DNA binding mechanisms must differ since TFAM has a longer (30aa) linker that allows intertwining with the DNA (18,19). Thus, the molecular basis of the Abf2p/DNA interaction and the corresponding recognition mechanism cannot be extrapolated *a priori*. We present here two crystal structures of Abf2p in complex with AT-rich DNA sequences possessing ARS-m motifs. Furthermore, we analyze the characteristics of DNA binding and respective functions of the Abf2p domains *in vitro*, *in vivo* and *in silico*. By combining all these data, we unveil the DNA-binding mechanism of Abf2p and the structural properties of the DNA that induce phased binding.

MATERIALS AND METHODS

Protein production

The yeast Abf2p gene was cloned from genomic DNA using standard polymerase chain reaction (PCR) with a proof-reading DNA polymerase (Pfu Ultra, from *Agilent*) and inserted into the pCri7a (20) expression vector to produce a 6-His-tagged fusion protein. This construct (residues 27–183, Saccharomyces Genome Database ID S000004676) corresponds to full length mature Abf2p without the N-terminal mitochondrial signalling sequence (21). The deletion mutants Mut1 to Mut6 (Figure 3) were generated by 'round the horn PCR' (22) with KOD DNA polymerase (Novagen). Subsequent DNA sequencing was performed for all constructs (native and mutant proteins). Plasmids were transformed into BL21 (*pLys*) *Escherichia coli* (Merck Millipore) strain. Cells were grown in LB medium for 2 h at 37°C until the optical density (O.D) at 600 nm reached 0.6. After subsequent induction with 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG), the culture was grown for 4 h at 37°C and shaking at 225 rpm. Cells were sonicated for 10 min (cycle 2 s on/6 s off) on ice in 50 mM Tris–HCl pH 7.5 and 1M NaCl. The lysate was injected in a Ni-NTA-affinity column (HisTrap HP, GE Healthcare) mounted on an AKTA Purifier (GE Healthcare) system. The proteins were eluted with a linear gradient from buffer A (20 mM imidazole, 50 mM Tris–HCl pH 7.5, 750 mM NaCl) to buffer B (500 mM imidazole, 50 mM Tris–HCl pH 7.5, 750 mM NaCl). Quality of the protein fractions was assessed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and the purest fractions pooled and concentrated for gel filtration chromatography using a Superdex 75 10/300 column (GE Healthcare), pre-equilibrated with running buffer (50 mM Tris–HCl pH 7.5, 750 mM NaCl). The purity of the peak fractions was analyzed by SDS-PAGE. Regarding Se-met derivatives, since the wild-type mature Abf2p contains only Met147, Leu52 was mutated to methionine (Abf2p-Semet protein variant) to have sufficient phasing power (23). Non-auxotrophic BL-21 *pLys* bacterial strain was grown at 37°C to an O.D of 0.6 and methionine synthesis inhibition (23) was carried out (inhibition mix was prepared with amino acids purchased from Sigma) prior to induction with 1mM IPTG (Base medium, amino acid nutrient mix and seleno-methionine were purchased from Molecular Dimensions).

Nucleic acids

The oligonucleotides used for this study are listed in Supplementary Table S1. Duplex DNA fragments for crystallization and isothermal titration calorimetry (ITC) were prepared by annealing complementary oligonucleotides (*Sigma*) Af2_22fw and Af2_22rev, Af2_shift22fw and Af22_shift22rev, GC_22fw and GC_22rev for Af2_22, Af2_shift22 and GC_22 respectively. The oligonucleotides (100 μM) were resuspended in 20 mM Tris–HCl, pH 7.5 and 100 mM NaCl, and heated at 85°C for 30 min followed by slow cooling to room temperature overnight. The (25 bp × 4) four-way DNA junctions were assembled by equimolar hybridization of the 50-mer oligonucleotides (Seq50.1-4, see Supplementary Table S1; 40 μM each) in 20 mM Tris–HCl pH 7.8, 50 mM NaCl and 5 mM $MgCl_2$. Oligonucleotides were heated for 5 min at 95°C and cooled down to room temperature overnight. M13mp18 dsDNA plasmids (*Bayou Biolabs*) were linearized with BsrB I (New England Biolabs) and purified with the illustra GFX PCR DNA purification kit (GE Healthcare).

Crystallization, structure solution and analysis

Crystallization trials were attempted with DNA fragments of different sequences and lengths, both GC and AT rich,

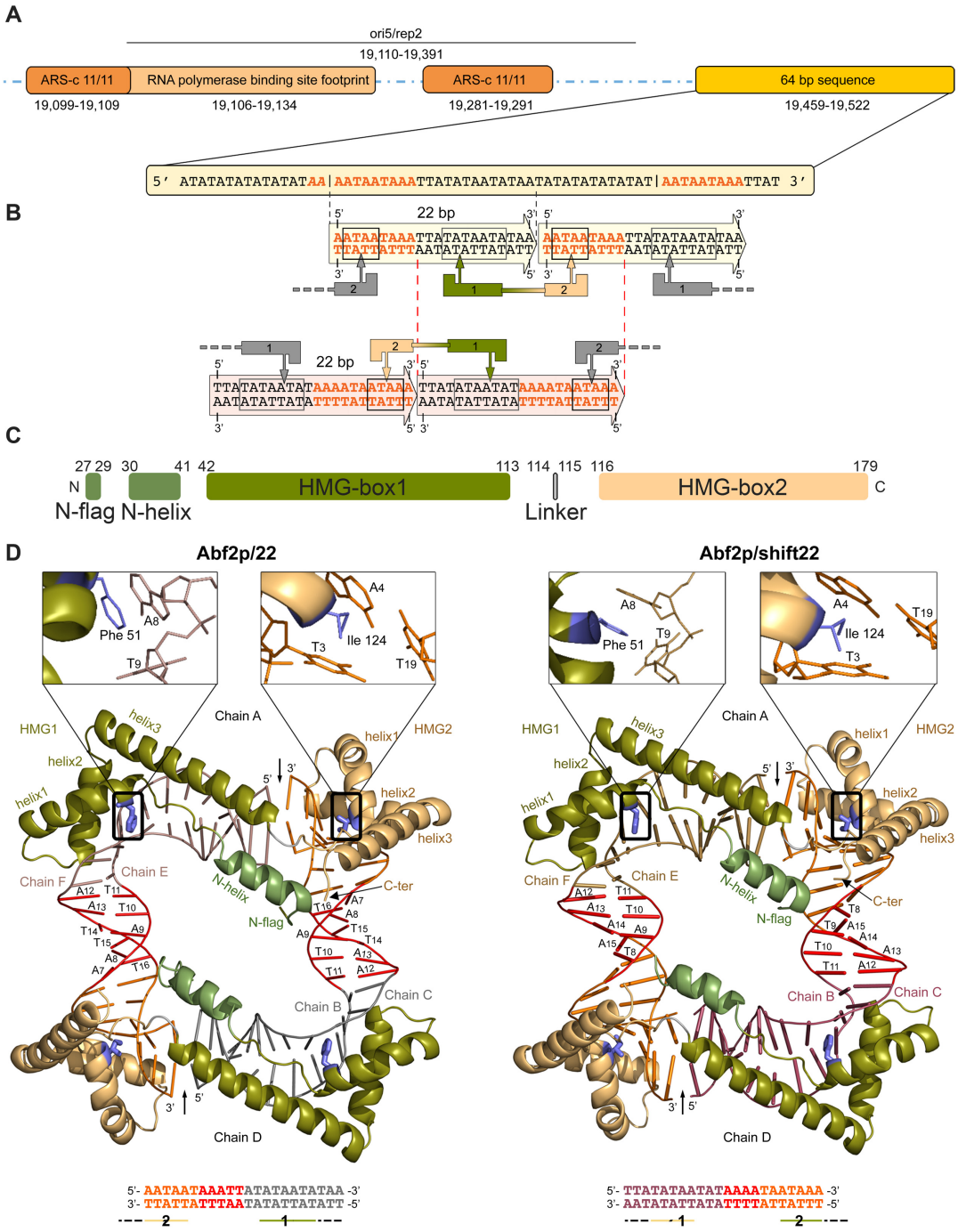

Figure 1. DNA design and Abf2p/DNA complex crystal structures. (**A**) General organization of the y-mtDNA ori5/rep2 origin of replication, including the location of ARS consensus sequences (ARS-c, 11/11 matches to 5′-(A/T)AAA(T/C)ATAAA(A/T)-3′, in orange) and the downstream 64 bp region (in yellow) in y-mtDNA[17]. Below, the 64 bp sequence is shown, the near matches (ARS-m; 9/11) to ARS-c are depicted in orange and the position of the 35 bp sequence (see Introduction) is demarcated by vertical bars. Vertical dashed lines indicate the DNA sequence 'Af2_22' within the 35bp fragment. (**B**) The dsDNA sequences used for crystallization are depicted inside arrows that show the head-to-tail arrangement found in the crystals; the DNA regions contacted by Abf2p are framed. Abf2p is represented as two connected L-shaped boxes 1 and 2 (HMG-box 1 and 2, respectively) with arrows representing DNA-inserting residues. The gray HMG-boxes represent the second protein molecule in the asymmetric unit; the dotted lines indicate the domains are connected. From the two 'Af2_22' sequences in tandem, a continuous 'Af2_shift22' dsDNA molecule was derived (vertical red dashed lines) that yielded the Abf2p/shift22 structure, where the protein again contacts two DNAs. (**C**) Schematic representation of the Abf2p domains. 'L' stands for Linker, 'N-flag' and 'N-helix' stand for N-terminal flag and helix, respectively. (**D**)The crystal structures of Abf2p/22 (left) and Abf2p/shift22 (right) are shown, with secondary structure elements and domains labeled. The intercalating residues Phe51 and Ile124 are highlighted in blue and their insertion sites are shown in insets. The DNA head-to-tail arrangement is indicated with a black arrow. The ARS-m-like regions (see main text) within Af2_22 and Af2_shift22 are shown in orange, the A-tract sub-region in red. Underneath, the crystallized dsDNA sequences and contacting protein domains are shown; note that the DNA sequences are in the same orientation while the Abf2p domains are in the opposite.

with and without A-tracts. Diffraction quality crystals were obtained only with Abf2p in complex with Af2_22 DNA. The protein/DNA complex for crystallization was prepared by mixing Abf2p and DNA fragment at a molar ratio of 2:1 (protein concentration 0.5 mg/ml) and performing stepwise overnight dialysis to reach a final buffer composition of 50 mM Tris–HCl, pH 7.5, 20 mM NaCl. Crystals were obtained in conditions of 21–25% w/v polyethylene glycol (PEG) 4000, 0.1 M tri-sodium citrate pH 4.5 and 0.2 M ammonium acetate at 20°C by sitting drop vapor diffusion method, at a protein concentration of 10–12 mg/ml. Crystals were cryo-protected with a solution of 15% glycerol, 21–25% PEG 4000, 0.1 M tri-sodium citrate pH 4.5, 0.2 M ammonium acetate and vitrified in liquid nitrogen. Data collection of intermediate crystals and optimized ones was performed at synchrotron ALBA (Cerdanyola del Vallès, Spain) and European Synchrotron Radiation Facility, ESRF (Grenoble, France). Specifically, diffraction datasets for Abf2p-Semet in complex with Af2_22 were collected at the seleno-methionine peak energy of 12.6623 keV, to a nominal resolution of 3 Å and with Friedel pairs collected on the same image. Crystals belonged to the space group C2. Diffraction data were processed with XDS (24) and Aimless (25,26). Single wavelength anomalous diffraction phasing was performed with anisotropy corrected data using Shelx C D E (27) followed by density modification using DM (28). A partial model could be built into the resulting Fourier map through iterative cycles of model building, visual inspection and automated refinement. Subsequently this partial model was used for molecular replacement with native Abf2p/22 crystal data collected at 12.658 eV with a nominal resolution of 2.18 Å, monoclinic space group C2 (see Table 1). Data processing was performed as mentioned above. For confirming sequence register of Abf2p/22, thymine 6 (T6) of Af2_22 was replaced with 5-bromo uracil (5-BrU) (oligonucleotides purchased from Biomers) and crystals were obtained under similar conditions as Abf2p/22, with space group $P2_1$ and a nominal resolution of 3.3 Å. The structure was solved by molecular replacement and the position of the Br atoms was located from an anomalous density map (visible peaks above 5 r.m.s.d) obtained using AnoDe (29). Abf2p/shift22 dataset was obtained with a nominal resolution of 2.6 Å, space group $P2_1$ (see Table 1). The structure was solved by molecular replacement and omit map was generated to identify the DNA ends.

CCP4i (28) and the tools therein were used for data manipulations such as truncation (Ctruncate) and merging (Aimless). Phenix (30) was used for automated structure refinement. Coot (31) was used for map visualization, model building and validation. PyMOL (DeLano Scientific) was used for figure generation. It is to be noted that for both the structures protein chainA and DNA chains B, C, E, F are treated as the reference as other protein chains in the asymmetric unit (a.u.) have regions with poor Fourier map definition (resulting in a relatively high number of RSRZ outliers), if not otherwise indicated.

Small angle X-ray scattering (SAXS)

Purified Abf2p was dialyzed overnight in 50 mM Tris–HCl, pH 7.5 and 500 mM or 100 mM NaCl. The protein samples were subsequently concentrated. The filtrate was used as a blank to obtain maximal buffer match between the blank and the sample. The Abf2p/22 complex was prepared identically as for crystallization. Small angle X-ray scattering (SAXS) measurements were performed at the ESRF. Data analysis was performed with Primus (32) in ATSAS package. Guinier approximation was used to calculate radius of gyration (*Rg*). Molecular weight of the protein was estimated from the Porod volume (32). Data from a concentration series (3.8 mg/ml, 5.4 mg/ml, 7.0 mg/ml, 9.5 mg/ml and 13.5 mg/ml) was collected at 50 mM Tris–HCl, pH 7.5, 500 mM NaCl. The estimated molecular masses are 14.3–19.0 kDa, 16.3–21.8 kDa, 15.1–20.1 kDa, 16.1–21.5 kDa and 16.3–21.8 kDa respectively (the lower and upper estimates for each concentration are obtained by dividing the Porod volume by 2 and 1.5 respectively). These values match closely with the 6-histidine-tagged monomer of Abf2p (19.5 kDa). The radii of gyration (*Rg*), calculated from the Guinier plot are 27.2 +/− 0.1, 27.3 +/− 0.1, 27.4 +/− 0.1, 27.3 +/− 0.1 and 27.8 (+/−0.1) Å, respectively. Additional data was collected at 50 mM Tris–HCl, pH 7.5, 100 mM NaCl (0.36 mg/ml; no reliable data could be obtained at other concentrations) for which the calculated molecular mass and *Rg* are 12.6−16.8 kDa and 28.4 (+/−0.6) Å respectively. The curve corresponding to the highest concentration (13.5 mg/ml) was used for all other subsequent analyses as it showed lesser noise at higher angles. Pair distribution ($p(r)$) plot and maximum particle dimension (D_{max}) were calculated from scattering data using GNOM (32). Ensemble analysis was performed using EOM 2.0 (33,34) with standard parameters for the genetic algorithm and for Crysol (32). For EOM, we generated a collection of 10, 000 models starting from the crystal structure coordinates. The N-flag+N-helix+HMG-box 1 assembly and HMG-box 2 were treated as two rigid domains while allowing flexibility to the linker (113-Lys-Leu-Pro-115, 113 added to include the unwinding of HMG-box 1 helix3 C-terminus as observed from MD simulations), the N-terminal His-tag and the last two C-terminal residues. Fitting the calculated scattering profile to the experimental SAXS data, yielded $\chi^2 = 1.379$. As an alternative, we used a frame from the MD simulation (extracted from simulation run after the r.m.s.d. had converged) as a starting model for EOM (Figure 4B). Using the same assignment of flexible regions, EOM generated a subset of five conformations which, collectively, were in better agreement ($\chi^2 = 0.957$) with the experimental data.

Molecular dynamics simulations

Molecular dynamics (MD) simulations for the protein, DNA and protein–DNA complexes were performed starting from the X-ray crystal structures using Amber 14 package. All simulations were carried out using state-of-the-art simulation conditions, the amber14SB with parmbsc1 refinement (35) force fields and explicit TIP3P water model. Models were prepared for simulation using tleap in Amber Tools 15. We added sodium and chlorine ions to neutralize the systems and give a realistic ionic environment (150

Table 1. Data collection and refinement statistics

	Abf2p/22	Abf2p/shift22
Wavelength	0.97950	0.97625
Resolution range	43.48 - 2.18 (2.258 - 2.18)	48.05 - 2.6 (2.693 - 2.6)
Space group	C 1 2 1	P 1 21 1
Unit cell	88.49 113.41 67.71 90 103.803 90	71.989 131.991 71.994 90 103.211 90
Total reflections	126763 (6246)	88589 (9083)
Unique reflections	33305 (3075)	38396 (3889)
Multiplicity	3.8 (2.0)	2.3 (2.3)
Completeness (%)	0.98 (0.92)	0.95 (0.97)
Mean I/sigma(I)	12.01 (1.99)	6.09 (1.05)
Wilson B-factor	49.68	55.15
R-merge	0.06649 (0.5319)	0.1022 (0.8523)
R-meas	0.07359 (0.6935)	0.1298 (1.085)
CC1/2	0.998 (0.632)	0.995 (0.207)
CC*	1 (0.88)	0.999 (0.585)
Reflections used in refinement	33305 (3074)	38396 (3885)
Reflections used for R-free	1881 (164)	1828 (166)
R-work	0.2090 (0.3061)	0.2267 (0.3656)
R-free	0.2364 (0.3213)	0.2601 (0.3623)
CC(work)	0.956 (0.754)	0.948 (0.446)
CC(free)	0.938 (0.772)	0.907 (0.339)
Number of non-hydrogen atoms	4520	8906
Macromolecule atoms	4445	8772
Protein residues	313	618
RMS(bonds)	0.008	0.008
RMS(angles)	0.95	1.00
Ramachandran favored (%)	99	98
Ramachandran allowed (%)	0.65	1.6
Ramachandran outliers (%)	0	0
Rotamer outliers (%)	0.7	1.3
Clashscore	8.45	12.79
Average B-factor	58.93	61.38
B-factor macromolecules	59.04	61.53
B-factor solvent	52.14	50.10

* Statistics for the highest-resolution shell are shown in parentheses.

mM NaCl). All simulations were performed using a truncated octahedron periodic box, periodic boundary conditions and particle-mesh-ewald for long range electrostatics. For Abf2p, Abf2p/DNA complex (1protein:2DNA; 22 bp each DNA), Af2_22 and Af2_shift22, we performed MD for 500 ns in each case. Analysis of trajectories were performed with cpptraj (36) in AmberTools 15. Visualization was performed using VMD (37). Analysis of DNA structural parameters was performed using Curves+ suite of programs (38,39).

DNA deformation energy and DNA stiffness calculations

DNA structures can be described in terms of base-pair and base-step parameters that consist of three translations (shift, slide and rise) and rotations (tilt, roll and twist) and the DNA deformability along these six directions can be described by the associated stiffness constant matrix (40). From the ensemble of MD simulations, the covariance matrix describing the deformability of the helical parameters for a given DNA fragment (e.g. a dinucleotide step) is computed and is inverted to generate the 6×6 stiffness matrix for each fragment. Pure stiffness constants corresponding to the six parameters mentioned above (kshift, kslide, krise, ktilt, kroll and ktwist) are extracted from the diagonal of the matrix while the total stiffness (*K*tot) is obtained as a product of these six constants and provides a rough estimate of the flexibility of each base pair step. The normalized Boltzmann-like probability distribution, $\exp(-\Delta E_{def}/k_bT)$, was derived from the deformation energy (ΔE_{def}) and was used to determine the relative probability of positioning the HMG-boxes on a given DNA segment. The deformation energy was calculated using a mesoscopic energy model (40,41), which is based on a harmonic approximation to describe deformability along DNA helical parameters. The equilibrium geometry and stiffness force constants were extracted from a dataset built from long all-atoms MD simulations of short oligonucleotides in water using the *parmbsc1* force field (35).

Abf2p *in vivo* assays

Diploid *ABF2* (YMR072W) hemizygous strain Y26205 (BY4743; *MAT*a/*MAT*α; *ura3Δ0/ura3Δ0*; *leu2Δ0/leu2Δ0*; *his3Δ1/his3Δ1; met15Δ0/MET15; LYS2/lys2Δ0*; *YMR072w/YMR072w*::*KANMX4*) was obtained from the European *S. cerevisiae* Archive for Functional Analysis (*Euroscarf*). A plasmid-borne *ABF2* was made by cloning a PCR fragment including Abf2 CDS plus 500 nt on both sides into pRS316 (42). This construct was introduced into Y26205 diploid cells, which were grown on uracil deficient synthetic dextrose (SDC) medium

and subsequently sporulated. Kan$^+$ (*ABF2* disrupted) Ura$^+$ (pRS316 transfected) cells were isolated and the genotype was verified by PCR (43). *ABF2* truncations were made by PCR while preserving the 500 nt flanks, and were cloned into pRS315 (a plasmid providing leucine prototrophy (42)). To test *ABF2* truncations, we followed a plasmid-shuffling strategy using 5-fluroorotic acid (5-FOA), toxic for Ura$^+$ cells (44,45). The mitochondrial DNA integrity of the corresponding Kan$^+$ Ura$^-$ Leu$^+$ cells was tested by growth in glycerol (respiration-only conditions). For this purpose, a spot assay was performed with 1/5 serial dilutions, starting from cultures grown in leucine deficient SDC medium to an O.D of 0.5. All cultures and plates were grown at 30°C.

Isothermal titration calorimetry (ITC)

Samples for ITC were prepared by simultaneously dialyzing the protein and DNA in the same buffer (25 mM Hepes pH 7.5, 150 mM NaCl) to obtain maximal buffer match between samples. Titrations were performed with a *VP ITC* instrument with 1400 μl cell volume and 300 μl syringe volume, at a temperature of 25°C. The general setup was designed with protein in the cell and DNA in the syringe. The protein was used at a concentration of 8–16 μM and the DNA at a concentration of 100–150 μM. DNA into buffer titrations were performed as control and since the profile was flat, subtraction of the control was not performed. The resulting data were analyzed using Origin (Origin Lab, Northampton, MA, USA). The dsDNA fragment GC_22 (complementary strands 5′ GAAGATATCCGGGTCCCA ATAA 3′ and 5′ TTATTGGGACCCGGATATCTTC 3′) was used as a control.

Electrophoretic mobility shift assays (EMSA)

Two-fold serial dilutions of Abf2p and deletion mutants were performed on ice in 20 mM Tris–HCl pH 7.5 and 750 mM NaCl. Reactions were initiated by addition of proteins as appropriate and incubated for 30 min at 25°C. Double-stranded (ds) DNA binding experiments contained 2ng·μl^{-1} linearized M13mp18 and were performed in a binding buffer with a final composition of 20 mM Tris–HCl pH 7.5 and 100 mM NaCl (20μl final volume). Subsequently, samples were resolved on 25 cm long, 1.2% (w/v) agarose gels (SeaKem LE Agarose, Lonza) in 0.5× TBE (Sigma). Gels were run in a Sub-Cell GT cuvette (BioRad) for 17–18 h at room temperature at 3V/cm.The reactions with the four-way junctions contained 100 nM of the (25 bp × 4) DNA and the buffer was supplemented with 1mM $MgCl_2$. After addition of a loading dye, the binding reactions were loaded on a 7% polyacrylamide gel (Sigma) in 0.5 × TBE. Gels were pre-run for 15 min and run at 11V/cm on a HoefferMiniVE cuvette placed on ice. Agarose and polyacrylamide gels were stained with SybrSafe (Molecular Probes) and scanned for fluorescence using a Typhoon 8600 scanner (GE Healthcare).

Fluorescence confocal microscopy

Kan$^+$ Ura$^-$ Leu$^+$ cells harboring different plasmid borne Abf2p truncations were grown in leucine deficient SDC media to O.D. of 0.5–1. Positive and negative controls were grown in glycerol and YPD respectively. Positive control consisted of wild-type (*ABF2* intact) haploids and negative control consisted of Kan$^+$ (*ABF2* disrupted) haploids. Cells were fixed with 4.5% formaldehyde for 1 h, treated with zymolyase Z100T (Ecogen) for 20 min at 30°C, deposited on poly-lysine coated slides and stained with 300 nM solution of 4′, 6-Diamidine-2′-phenylindole dihydrochloride (DAPI) for 5 min prior to mounting. Fluorescence confocal microscopy images were acquired on a *Zeiss LSM780*, with 64X oil immersion objectives, numerical aperture 1.4. Excitation wavelength was 405 nm (Diode laser) and emission range was 419.47–535.16 nm. Pixel size for acquired images was 0.05 μm. For each sample, maximum intensity projection of the stacks was performed with Fiji (46). Counting of cytoplasmic DAPI foci (proxy for mtDNA content) was performed manually for 50 cells per sample type and an average was obtained.

Circular dichroism

The protein samples were dialyzed overnight in 10 mM potassium phosphate pH 7.5, 100 mM ammonium sulfate. A quartz sample cell of 1 mm path length was used for data collection on a JascoJ-815CD spectrophotometer, with a scan rate of 50 nm/min and data pitch of 0.5 nm resulting in 150 data points between 260–185 nm. Protein concentration was kept constant at 0.1mg/ml for all samples.

RESULTS

Well-diffracting crystals were obtained for Abf2p (mature form, without the mitochondrial targeting sequence) bound to an A-T rich double-stranded (ds) DNA, Af2_22 (5′-AATAATAAATTATATAATATAA-3′, Figure 1B), spanning 22 of the 35 bp sequence (19477 to 19511bp in ρ$^+$ y-mtDNA (4), Figure 1A) that is present in the wild-type *S. cerevisiae* ρ$^+$ mtDNA and additionally amplified as concatamers in ρ$^-$ yeast (*S. cerevisiae*). This sequence is expected to be recognized by Abf2p *in vivo*. The crystal structure (hereafter termed Abf2p/22; PDB ID 5JH0; Table 1; Supplementary Figure S1) reveals an asymmetric unit (a.u) composed of two Abf2p molecules (polypeptide chains Abf2p-A and -D) and dsDNAs (chains-BC and -EF, 22 bp each) (Figure 1B and D). Abf2p consists of two HMG-boxes in tandem, HMG-box 1 and 2. At the N-terminus of the protein, three residues in extended conformation (27-Lys-Ala-Ser-29, henceforth referred to as 'N-flag') and a 12 amino-acid (aa) N-terminal helix (N-helix) (Figures 1C, D and 2A) are found. Notably, such a helix is not present in any structure containing HMG-boxes reported so far. The N-helix leads to HMG-box 1 (42–113) (Figures 1C and 2A), which displays a typical HMG-box 'L-shape' consisting of an extended segment (42–48), helix1 (49–64, with a kink at position 57), a four-residue turn, helix2 (70–83, which with helix1 forms the short L-arm), a short two-residue connecting loop and finally, helix3 (86–113, antiparallel to the extended segment and forming the long L-arm of the HMG-box L-shape). A Linker of two residues (Leu-Pro) follows, which is much shorter than the predicted length of 10 aa

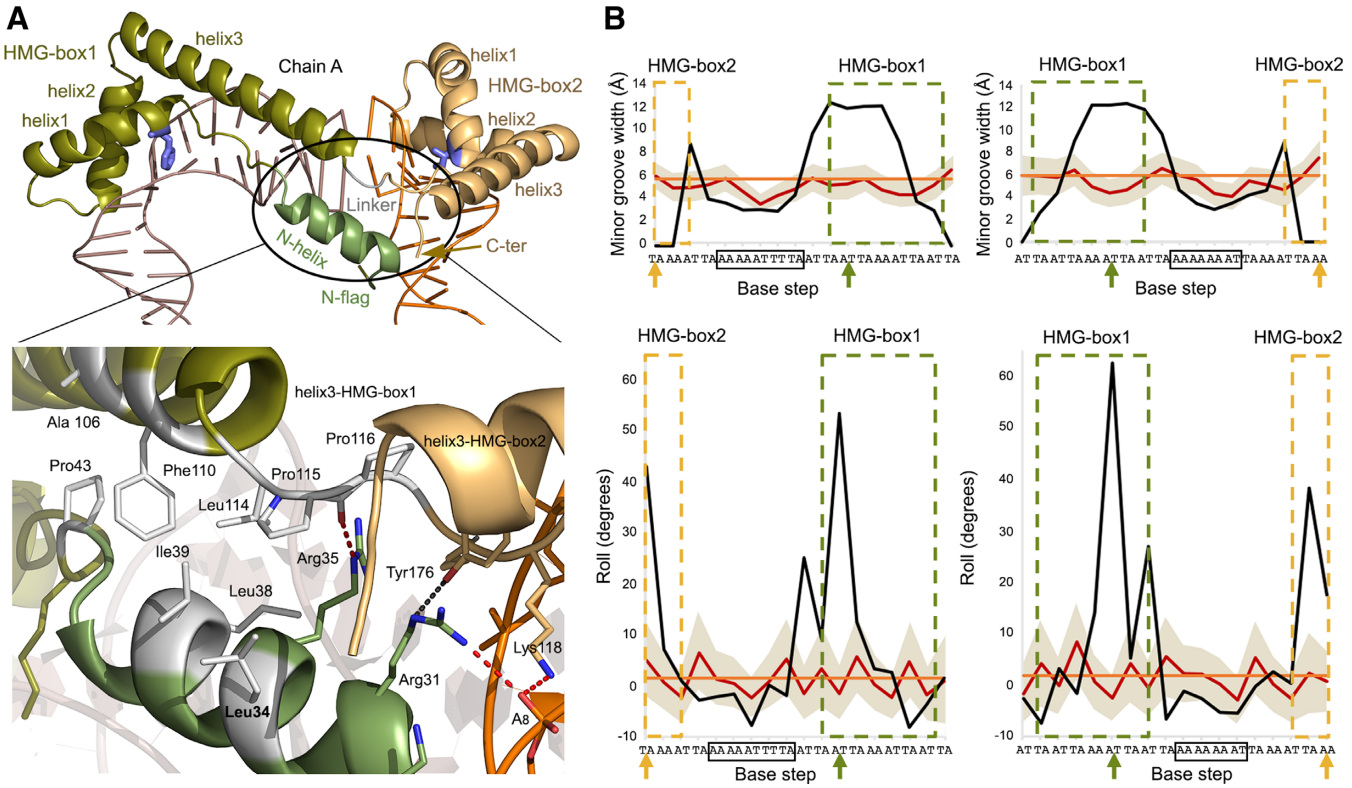

Figure 2. Interactions within Abf2p and DNA dynamics. (**A**) Above, side view of Abf2p (chain A) from the Abf2p/22 crystal. The encircled secondary structure elements (N-helix, helix3 of HMG-box 1, the Linker and helix3 of HMG-box 2) participate in the interactions that connect the Abf2p N- and C-terminal regions, shown in detail below. Hydrophobic side chains are shown in white, electrostatic contacts are shown in red (the weak ones in gray). (**B**) The DNA parameters 'minor groove width' (top graphs) and 'roll' (bottom) are shown for each base-pair step along the Af2_22 (left) and Af2_shift22 (right) sequences. Values that correspond to the crystal structures (in black), to the MD simulations for unbound Af2_22 and Af2_shift22 DNAs (averaging over the individual trajectories, in red; spread of the values, in grey), or to an ideal B-DNA (in orange) are shown. Values corresponding to the DNA contacted by Abf2p is framed in olive green for HMG-box 1 and light orange for HMG-box 2. The arrows, colored according to the HMG-boxes, indicate the insertion sites. The A-tracts are framed in black. Note the switch of the relative orientations of DNA and protein between complexes (also shown in Figure 1B and D).

(9). Similar to HMG-box 1, HMG-box 2 (116–179) consists of an extended segment (116–121), helix1 (122–137), a five-residue turn, helix2 (143–156, with a kink at position 130) and a two-residue connecting loop leading to helix3 (159–179, Figures 1C, D and 2A). Finally, the last four residues 180-Tyr-Pro-Leu-Asn-183 form a short extended C-terminal segment.

Abf2p binds to two separate DNAs via its two domains

In the a.u. of the Abf2p/22 crystal, two curved dsDNA molecules of 22 bp are arranged head-to-tail by two protein molecules, forming a pseudo-continuous DNA circle exhibiting stacking interactions between respective 3′-5′ DNA ends (Figure 1D). Thus, each protein holds together and bends two different DNA molecules via its HMG-boxes, much like a staple (hereafter referred to as dual-binding) (Figure 2A). Specifically, HMG-box 1 from Abf2p-A contacts dsDNA EF chains (from T_{13} to A_{19} of chain-E), while HMG-box 2 contacts the other DNA molecule BC (A_2-A_5 chain-B), with the DNA ends participating in perfect stacking interactions between HMG-box binding sites (Figures 1B, D and 2A). A similar scheme applies for the second protein molecule in the a.u. where HMG-box 1 and HMG-box 2 contact dsDNA chains BC and EF respectively. Each HMG-box engages in mostly electrostatic interactions with the DNA minor groove (Supplementary Figure S2), opens it up to 12 Å (ideal B-DNA standard minor groove width 5.9Å, Figure 2B) and concomitantly bends the DNA by ∼90°. Additionally, each HMG-box inserts hydrophobic residues between base steps, thus causing considerable alterations of DNA parameters (Figure 2B). Specifically, HMG-box 1 Phe51 inserts at step $A_8T_9/A_{14}T_{15}$ (chains F/E), the phenyl ring stacking with the six-membered ring of A_8 (Figure 1B and D). This base-step shows a high positive roll and the involved base pairs show a prominent buckle (Figures 1D, and 2B). In comparison, Ile124 from HMG-box 2 inserts partially at $T_3A_4/T_{19}A_{20}$ (B/C), and induces a positive roll (Figures 1D and 2B). Superposition of both HMG-boxes indicates high structural similarity (r.m.s.d. of 1.1Å for 66 C^{α}), both Phe51 and Ile124 being at equivalent positions in helix1. Thus, the two DNAs that are contacted by a single protein together form a U-turn (or a half DNA circle, Figure 2A). Importantly, the Abf2p mutant K44A, R45A, K117A and K118A causes 80% of *S. cerevisiae* cells to lose mtDNA (14,47) (and thus only survive in fermentable conditions). These amino acid residues correspond to a conserved Pro-B-B-Pro (B, basic aa) motif found

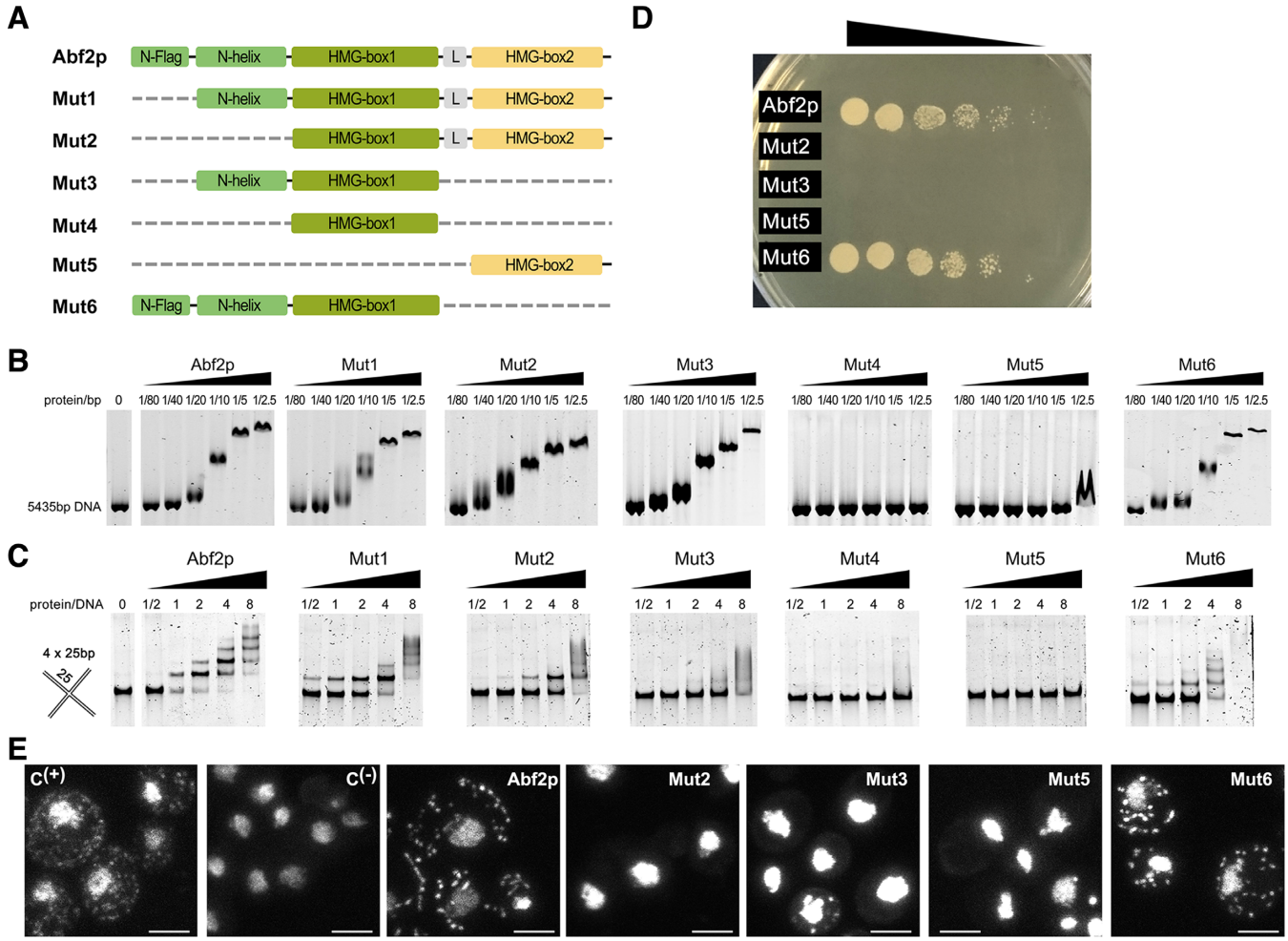

Figure 3. Abf2p truncations and their effect on DNA binding. (**A**) Schematic representation of the Abf2p deletion mutants. The domains are designated as in the text except for the linker, indicated with an 'L'. Dashed lines indicate deleted regions. (**B**) Titration of Abf2p and Mut1 to Mut6 constructs binding on linearized M13mp18 dsDNA (8 fmol) at 25°C, by EMSA. The protein/bp ratios are indicated and correspond respectively to 0.37, 0.75, 1.5, 3, 6 and 12 pmol of Abf2p. (**C**) Titration of Abf2p and Mut1 to Mut6 binding to four-way junction DNA (100 nM) by EMSA. (**D**) Effect of Abf2p truncations *in vivo*. Spot assay of cells with 1/5 serial dilutions on plate containing glycerol as carbon-source ('YPG plate') shows the effect of different Abf2p truncations on the ability to maintain y-mtDNA and thus on capability of yeast (*Saccharomyces cerevisiae*) cells to use non-fermentable carbon source. Full-length Abf2p and Mut6 (Abf2p without HMG-box 2) are capable of protecting mtDNA whereas Mut2 (Abf2p without N-flag and N-helix), Mut3 (N-helix+HMG-box 1) and Mut5 (HMG-box 2 alone) are not functional. (**E**) Visualization of mitochondrial nucleoids in cells harboring different Abf2p truncations by DAPI staining of DNA and fluorescence confocal microscopy. Abf2p and Mut6 maintain mtDNA whereas Mut2, Mut3 and Mut5 cannot. $C^{(+)}$ corresponds to wild-type haploids, and $C^{(-)}$ to *Abf2*Δcells. The scale bars correspond to 2.5 μm.

in the N-terminal extended segment of either HMG-box domains, 43-Pro-Lys-Arg-Pro-46 in HMG-box 1 and 116-Pro-Lys-Lys-Pro-119 in HMG-box 2. The structure shows the NZ at the tip of these lysine side-chains being involved in interactions with the DNA phosphate backbone, while the Arg45 guanidinium is positioned close to the O2 base atoms of T_6 (chainF) and T_{18} (chainE) at the minor groove, causing a negative roll (Figure 2B, Supplementary Figure S2). Thus, absence of these contacts likely weakens DNA binding by Abf2p, eventually leading to y-mtDNA loss. Additional residues that are highly conserved among the rather divergent yeast mitochondrial HMG-box proteins are found in HMG-box 2, which shows higher conservation than HMGbox 1 (48,49). Some of these conserved positions are located at the hydrophobic core within the L-shape elbow and include, from helix1, Phe123 that contacts the DNA backbone (corresponds to Tyr50 in HMG-box 1 and contacts both the DNA backbone and N3 from A14) and Tyr126 (Tyr53 in HMG-box 1); from helix2, Ile150 (Ala77 in HMG-box 1) and Trp154 whose N^{ϵ} contacts the A5 phosphate (Trp81 in HMG-box 1 whose N^{ϵ} contacts A16 phosphate); and from helix3 the methylene groups of Lys162 (Lys89) together with Tyr165 (Tyr192). Important for the forthcoming results and the proposed mechanism of protein/DNA interaction is that the DNA binding induces contacts between protein regions that are distant in the sequence. These include the N-helix, the HMG-box 1 helix3 C-terminus, and the Pro-Leu Linker, which together form a second, 'N-terminal' hydrophobic core ('N-hydrophobic core', Figure 2A). Further, the N-helix forms electrostatic contacts with the protein C-terminus (N^{ϵ} of Arg31 with conserved Tyr176 OH, Figure 2A). In this arrangement,

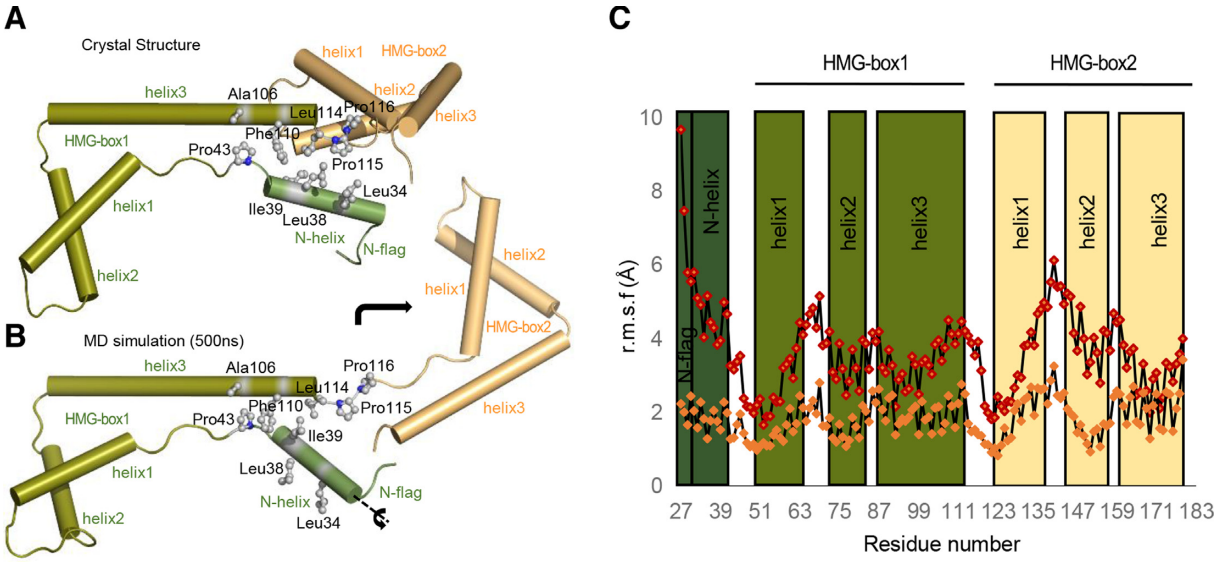

Figure 4. Abf2p flexibility and dynamics. (**A**) The Abf2p domains in the Abf2p/22 crystal structure are shown (same color-code as in Figure 1, the DNA molecule is not displayed for clarity) together with the amino-acid side-chains (depicted in gray ball-and-sticks and labeled) that participate in the N-hydrophobic core that stabilizes the contact between the N-helix, the C-terminal region of HMG-box 1 helix3 and the linker. (**B**) A frame from the 500 ns MD simulation of Abf2p that shows the hydrophobic residues from the N-helix exposed to the solvent, the changes in the orientation of N-helix and HMG-box 2 (HMG2) (relative to HMG-box 1 in the crystal structure above) indicated by arrows. (**C**) Root mean square fluctuation (r.m.s.f, in Å) of individual residues during MD simulation (500 ns) of Abf2p alone (red dots) and bound to DNA (orange dots). The different regions of the protein are schematically represented along the sequence. Note that the Abf2p flexibility reduces in the protein/DNA complex.

residues from the N-flag are close to the minor groove region immediately downstream of HMG-box 2 DNA binding site. Therefore, the N- and C-terminal regions are in close proximity in the protein/DNA complex.

We reasoned that if the dual binding mode was caused by a specific preference of the HMG-boxes for the contacted DNA sequence patches, then, joining both patches into a single 22 bp DNA molecule should result in a complex of one protein bound to a single, continuous-DNA. Accordingly, we designed the sequence 'Af2_shift22' (5′-TTATATAATATAAAATAATAAA-3′) (Figure 1B). To our surprise, the resulting structure (hereafter referred to as Abf2p/shift22; PDB ID 5JGH; Table 1) again showed dual binding in an arrangement highly similar to the previous Abf2p/22 complex (Figure 1B and D). To explain this, we performed a thorough analysis of both DNA sequences, and noticed the presence of an ARS-m-like sequence containing poly-adenine tracts (A-tracts) in both DNAs (chainB $(A_7)A_8A_9T_{10}T_{11}$ in Af_22, and chainC $A_{12}A_{13}A_{14}A_{15}$ in Af2_shift22, Figure 1D). Noticeably, the A-tracts are not contacted by the protein in either crystal. Moreover, crystallization trials with additional DNA fragments devoid of A-tracts (see 'Materials and Methods' section), yielded poorly diffracting crystals, suggesting high disorder. Thus, the A-tracts are presumably delimiting the potential protein binding sites on the DNA and thus promoting formation of crystals with sufficient diffraction quality. The above findings are strongly reminiscent of the phased binding pattern associated with Abf2p (see 'Introduction' section). Such patterns have been speculated to be due to the presence of A-tracts (10,11) and could have relevant functional implications. Thus, we decided to characterize the properties of the protein and the DNA and their possible role in modulating DNA binding by Abf2p.

Role of different domains of Abf2p in DNA binding

To discern the role of the different Abf2p regions (N-flag, N-helix, HMG-box 1 and HMG-box 2) in binding DNA we produced deletion mutants and analyzed their assembly on linearized M13mp18 dsDNA plasmids (5435bp) by electrophoretic mobility shift assays (EMSA).We constructed mutants Mut1 (no N-flag), Mut2 (HMG-box 1+HMG-box 2), Mut3 (N-helix+HMG-box 1), Mut4 (HMG-box 1), Mut5 (HMG-box 2) and Mut6 (N-flag+N-helix+HMG-box 1) (Figure 3A). Circular-dichroism experiments confirmed that all mutants were properly folded (Supplementary Figure S3A). The EMSA analyses (Figure 3B) showed that, under our experimental conditions, increasing concentrations of full length protein, in presence of a constant concentration of DNA, progressively up-shifted the band with a clear shift at a ratio of 1Abf2p:10 bp. Removal of the N-flag in Mut1 did not modify DNA binding, although DNA migration showed some smearing effect, suggesting some destabilization of the nucleoprotein complex. A similar effect was observed when the entire N-terminal segment was removed (N-flag+N-helix in Mut2). In contrast, neither HMG-box 1 (Mut4) nor HMG-box 2 (Mut5) showed any appreciable binding in this assay. However, HMG-box 1, in presence of N-flag+N-helix (Mut6), surprisingly showed a DNA binding efficiency similar to the full-length protein. Dissecting the latter, Mut3 (N-helix+HMG-box 1) behaved identically to Mut6, apparently suggesting that the N-flag does not have an additional, predominant functional role in DNA binding. Thus, the N-helix and HMG-box 2 individually endow DNA binding ability to HMG-box 1 *in*

vitro. On the other hand, Abf2p levels have been additionally shown to be positively correlated to the level of recombination intermediates in mtDNA (14,47) and Abf2p has been shown to possess high affinity for four-way junctions *in vitro* (50). Thus, we additionally performed EMSA assays with four-way junction templates (Figure 3C). We observed that full length Abf2p binds a four-way junction efficiently (Figure 3C). All of the DNA was bound at a protein:DNA ratio of 4:1, which possibly involves oligomerization of the protein on DNA. However, binding efficiency deteriorates slightly for Mut1 and significantly for Mut2, indicating that the absence of the N-flag or the entire N-module affects binding. The HMG-boxes alone i.e. Mut4 and Mut5 are inefficient binders. In contrast, HMG-box 1 bound to the N-helix (Mut3) slightly improves binding while presence of Nflag+Nhelix (Mut6) restores binding markedly. Summarizing, while all mutants except the HMG boxes alone have a similar behavior in binding dsDNA, they present differences when binding to four-way junctions, suggesting different binding modes.

In order to observe effects of the deletions *in vivo*, we subsequently generated haploid yeast *Abf2*Δ cells, bearing plasmids encoding the full length or mutant Abf2p constructs described above. Cells expressing full length, wild-type Abf2p grew normally on non-fermentable carbon source (e.g. glycerol), indicating a functional respiratory chain (Figure 3D). Strikingly, cells with Abf2p lacking the N-flag+N-helix (Mut2) were able to ferment glucose but failed to grow on glycerol. This indicates that the two HMG-box domains, coordinated by the Leu-Pro linker enable Mut2 to bind DNA (see above) but this is not sufficient for y-mtDNA stability and maintenance, which underlines the functional importance of the N-module. In contrast, and as previously shown (51), cells carrying the N-flag+N-helix+HMG-box 1 construct (Mut6) grew normally on glycerol indicating that this Abf2p mutant is sufficient for maintaining y-mtDNA. Removal of the N-flag (Mut3) from Mut6 though, results in loss of viability on glycerol. In light of the crystal structure, this is attributable to a role of the N-flag in establishing additional contacts with the DNA. Previous work with an Abf2p construct spanning amino acids 54–183, thus lacking the N-flag+N-helix module and an N-terminal segment of HMG-box 1, did not shift DNA *in vitro* but did not impair cell growth in YPG media containing traces of galactose (51). The different length of this and our construct together with the different experimental conditions tested (use of an inducible (51) or endogenous (this work) promoter; different strain background, temperature and media of growth) may explain the different results. To get further insight into the effect of our mutants on mtDNA, we visualized mitochondrial nucleoids by fluorescence confocal microscopy (Figure 3E, Supplementary Figure S3B). This revealed that *Abf2*Δ cells expressing plasmid-borne full length Abf2p or Mut6 have an average count of cytoplasmic y-mtDNA foci (15.5 and 10.9 respectively, see 'Materials and Methods' section) comparable to wild-type haploids ($C^{(+)}$, 22.2 foci). In contrast, *Abf2*Δ cells expressing Mut2, Mut3 or Mut5 lose almost all of their mtDNA (average foci count of 0.42, 0.48 and 0.9 respectively), similar to *Abf2*Δcells ($C^{(-)}$, average count of 0.4). Thus, combining our *in vitro* and *in vivo* results, we conclude that an intact N-flag+N-helix module is essential for mtDNA maintenance.

Abf2p dynamics provides clues to DNA binding mechanism

Our EMSA and *in vivo* assays with different deletion mutants pointed to the relevance of the N-flag and helix in y-mtDNA binding and metabolism. On the other hand, the crystal structure shows a detailed conformation of the final DNA bound state but does not provide mechanistic information about the involved binding process. To obtain this information, we performed MD simulations on the isolated protein and the complex in aqueous solution. A 500 ns MD simulation of the free protein shows that, in absence of bound DNA, Abf2p adopts a rather extended conformation where the two HMG-boxes separate from each other and the contacts between the N- and C-termini are lost (Figure 4B). This movement is allowed by the flexibility of the short two-residues Linker (114-Leu-Pro-115) and can be aided by occasional unwinding of the C-terminal end of HMG-box 1 helix3 (specifically Lys-113), which undergoes helix-turn-coil transitions (Supplementary Figure S4A). In addition, the N-helix shows a great variability in its orientation (Supplementary Figure S4B) and turns along its main axis resulting in Leu34, Leu38 and Ile39 side chains being exposed to the solvent (Figure 4B) and a consequent disruption of the N-hydrophobic core. The calculated root mean square fluctuation (r.m.s.f) along the simulation demonstrates that the N-helix is highly flexible (Figure 4C). In contrast, a 500ns all-atom MD simulation of Abf2p/DNA complex (1protein:2DNA) shows that the protein remains tightly bound to the DNA (Figure 4C and Supplementary Figure S5A), with overall lesser fluctuation. In this bound state, the stacking between the DNA ends is preserved (Supplementary Figure S5A) and the Phe51 and Ile124 insertions are maintained. The N-flag, N-helix and the N-hydrophobic core are also stably maintained, as in the crystal structures (Figure 4C). Therefore, the orientation of the N-flag and helix are fixed during complex formation with simultaneous consolidation of the N-hydrophobic core.

Abf2p compacts upon DNA binding in solution

The inter-domain flexibility observed from MD simulations was analyzed experimentally by SAXS, a technique that allows extraction of information about the structure and dynamics of macromolecules in solution (52). The data indicated that Abf2p is a monomer in solution (estimated 16.3–21.8 kDa, which match closely with a Abf2p his-tagged monomer of 19.5 kDa, see 'Materials and Methods' section). The calculated radius of gyration (Rg = 27.8 Å) is larger than that expected for a ∼20 kDa protein (17Å according to Flory equation $Rg = 3\ x\ Ne0.33$ (53)) and agrees well with that calculated from MD simulation of the free protein (Rg = 29.9Å).These results indicate an extended state of the protein. Accordingly, the pair-wise distribution function ($P(r)$), which reflects the distribution of intra-particle distances (52), yielded a large maximal particle dimension D_{max} of 130Å (Supplementary Figure S6A). Additional information regarding the conformational variability of macromolecules can be obtained from the Kratky

representation (52). In this case, the Kratky plot showed a flat profile that is typical of a flexible particle (Supplementary Figure S6B). To further describe Abf2p flexibility and the expanse of its conformational variability, we applied the ensemble optimization method (33,34) (see 'Materials and Methods' section) which, from a large pool of static structures (thousands, generated from an initial starting model), selects a sub-ensemble of conformations that best describes the SAXS data. In the present case, the sub-ensemble selected from this calculation describes the experimental data with significant accuracy ($\chi^2 = 0.957$, Figure 5A), and shows an important variability in the relative orientations of the HMG-boxes (Figure 5B). The *Rg* estimates for the sub-ensemble (Figure 5C) show a bimodal distribution suggesting two alternative predominant conformations of the free protein in solution, one being more compact than the other. The Abf2p/Af2_22 complex, in contrast to the free protein, showed characteristics of a more compact species, as reflected in the respective radii of gyration and maximal particle dimensions (*Rg* = 13.8Å versus 27.8Å; D_{max} = 65Å versus 130Å, Supplementary Figure S6A). This is supported by the corresponding Kratky plot (Supplementary Figure S6B), which shows a bell-shaped profile, characteristic of a compact particle (52). Therefore, free Abf2p is highly flexible and extended in solution while DNA binding reduces its conformational freedom, leading to a more compact overall arrangement.

Poly-adenine tracts direct Abf2p binding on DNA

In order to unveil the features of the DNA molecules responsible for the dual-binding mode observed in the crystal structures (one protein binding to two DNA molecules), we analyzed the structural properties of the Af2_22 and Af2_shift22 DNAs by all-atoms MD simulations, starting from ideal B-DNA models. Averages of the intra- and inter-base pair parameters and groove parameters from the trajectories revealed that the A-tracts showed conformational features distinct from standard B-DNA (40,54). In general terms, there are two types of A-tracts. *Asymmetric* A-tracts are a contiguous stretch of adenine bases (A_n, where n ≥ 4), which show a progressive narrowing of the DNA minor groove going from the 5′ to the 3′ end, while *symmetric* A-tracts are of the form A_nT_n (where n≥2) and exhibit narrowest minor groove at the central AT junction (54). Both A-tract types are represented in our crystal structures. The Af2_22 sequence has the symmetric tract $(A_7)A_8A_9T_{10}T_{11}$ (chainB), whereas Af2_shift22 DNA possesses the asymmetric tract $A_{12}A_{13}A_{14}A_{15}$ (chainC). The MD simulations showed that the symmetric A-tract at Af2_shift22 narrows from both ends to a minimum width of 3.75 Å (compared to the 5.91 Å in an ideal B-form), precisely at the A_9T_{10} junction (Figure 2B). For $A_{12}A_{13}A_{14}A_{15}$, the minor groove width shows a prominent decrease from 5′ to 3′, reaching a minimum of 4.0 Å. Notably, a MD simulation of the Abf2p/22 complex (1protein:2DNA) showed that the A-tract maintains the structural characteristics of the unbound form (Supplementary Figure S5B) and is not affected by the DNA distortions due to protein binding at adjacent non-A-tract regions. Furthermore, the crystallographic structures show remarkably narrow minor grooves precisely at the A-tracts (3.08 Å and 2.8 Å for Abf2p/22 and /shift22, respectively; Figure 2B). Thus, it can be deduced that Abf2p, a protein that binds to and opens up DNA minor grooves, finds A-tracts inhospitable. Values for the DNA stiffness descriptor (40) (*Ktot*, see 'Materials and Methods' section, Supplementary Figure S5C), revealed that the above A-tract regions possess high intrinsic rigidity (Supplementary Figure S5C and 'Materials and Methods' section) that make them difficult to be distorted. All these observations suggest that the unusually narrow minor groove of A-tracts together with their inherent rigidity impede Abf2p binding, diverting the protein towards more permissive DNA sequences. This presumably causes the unexpected dual-binding observed in both the crystals. Specifically, in the crystal structures, HMG-box 2 binds inside the ARS-m-like sequence, but avoids the A-tracts (Figure 1B and D), further indicating that the A-tracts are the key elements responsible for positioning Abf2p on DNA.

In order to explore how the DNA conformational properties affect Abf2p binding, we used our Elastic Deformation Energy model (40,41,55) to calculate the energy required to contort segments of the natural y-mtDNA 64bp sequence into the conformations induced by HMG-box 1 or HMG-box 2 (see 'Materials and Methods' section). Based on the rationale that binding propensity will be higher for those segments that require less energy to be distorted, we derived a sequence-dependent probability of binding (see 'Materials and Methods' section, Figure 6). The results show exclusion of the HMG-boxes from the A-tracts (Figure 6). Additionally, HMG-box 1 binding probability (Figure 6A) shows a more restrained pattern compared to HMG-box 2 (Figure 6B), indicating that the former has more stringent sequence requirements and a higher tendency to avoid A-tracts. The binding probability of full-length Abf2p can be expected to be a combination of that observed for the individual boxes. In summary, the narrow minor groove and rigidity of the A-tracts make them unsuitable as binding sites for Abf2p, while the adjacent sequence stretches are well suited for insertion and are favored for positioning the HMG boxes. This points to a mechanism of protein positioning mediated by DNA structure, which somehow resembles that suggested for nucleosome placements in nuclear chromatin (56–59). Interestingly, we calculated that 33.3% of the y-mtDNA is composed of A-tracts (3) (symmetric and asymmetric; Supplementary Table S2 and Supplementary 'A-tracts in y-mtDNA'), consistent with previous estimates (60). According to the above results, Abf2p DNA binding would follow a similar trend at other y-mtDNA regions possessing A-tracts.

The thermodynamics of Abf2p DNA binding is altered by A-tracts

Analysis of the thermodynamics of Abf2p DNA recognition by ITC showed that the presence of A-tracts modifies the mode of binding. Titration of Abf2p with Af2_22 DNA (protein in cell, DNA in syringe), produced a thermogram with an exothermic phase until a DNA:protein molar ratio of 1:1 was reached. Notably, a second endothermic phase followed (Figure 7A, top left). The isotherm was fitted to a model for two independent sets of sites (Fig-

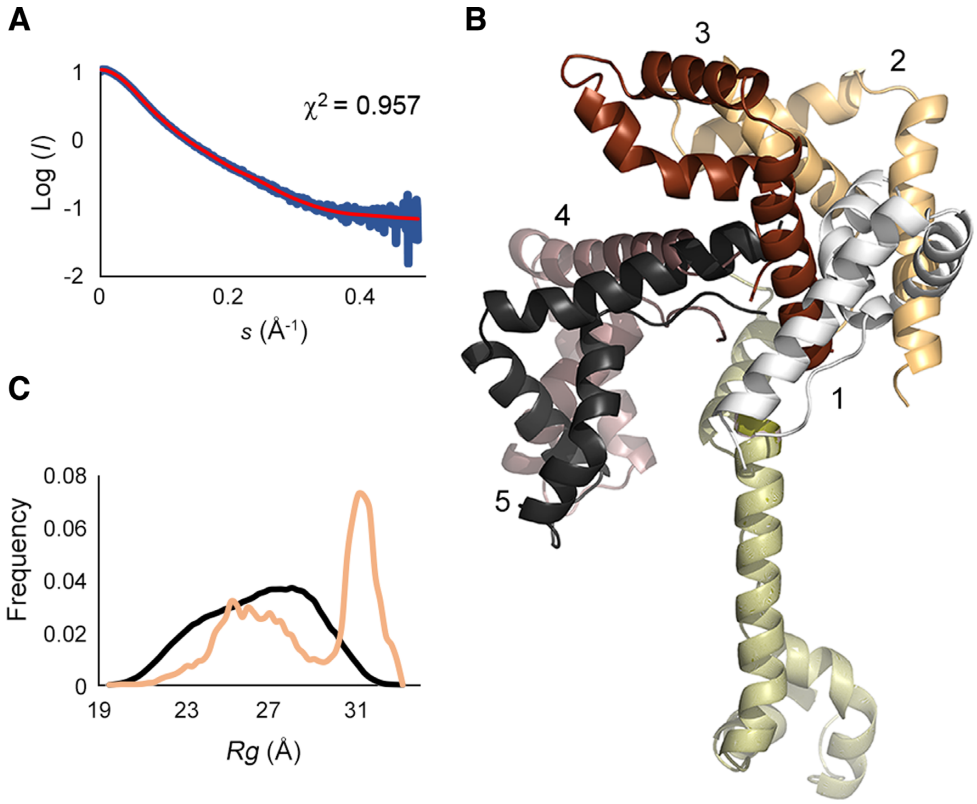

Figure 5. Conformational flexibility of Abf2p in solution. (**A**) Plot of the experimental SAXS intensity (*Log(I)*) along the momentum transfer *s* (blue curve) and fit ($\chi^2 = 0.957$) of the calculated average intensity from the final selected sub-ensemble (red curve) (**B**). Final sub-ensemble of five models selected by EOM, superposed by the HMG-box 1 domain and showing the variability of the HMG-box 2 position in the different models. (**C**) Radius of gyration (Rg) distribution of the EOM selected pool (light brown) compared to the initial random pool (black). The former shows a bimodal distribution.

ure 7A, bottom left), which would correspond to two binding sites on the protein. For site1, a stoichiometry of 0.86 was obtained corresponding to a ∼1:1 DNA:site1 ratio. For site2, a binding stoichiometry of 0.49 was obtained, corresponding to a 1:2 (DNA:site2) ratio. In addition, the derived affinity for site1 (K_d = 52.1 nM) was higher than for site2 (K_d = 166.4 nM). Our EMSA assays demonstrate that the Nflag+Nhelix+HMG-box 1 construct has a much higher DNA binding efficiency than HMG-box 2. Additionally, the crystal structures and the predicted probability of binding to the 64bp sequence indicate that the N-flag+Nhelix+HMG-box 1 module is constrained to be positioned outside the A-tracts and has a single preferred binding site on the Af2_22 sequence (in grey in Figure 7B, top, Figure 6). In contrast, HMG-box 2 possesses two putative binding sites on Af2_22. Therefore, we rationalized that the first exothermic phase is mostly attributable to the binding of N-helix+HMG-box 1, until a 1:1 (DNA:site1) molar ratio is reached. After the N-flag+N-helix+HMG-box 1 module is saturated, the HMG-box 2 domain has access to two sites on each new injected DNA molecule, yielding a1:2 (DNA:site2) stoichiometry. Thus, the two binding events together lead to an overall molar ratio of 3:2 (DNA:protein).

Additional ITC experiments were performed with a 22 bp DNA sequence that did not contain any A-tract (see 'Materials and Methods' section and Figure 7A, right). These showed an exothermic phase until a molar ratio close to 1:1 was reached, followed by a slightly endothermic phase (Figure 7A, top right). Fitting the data to the two-sites model yielded stoichiometry values of 0.55 and 0.50 for sites 1 and 2 respectively (Figure 7A, bottom right). Thus, both sites correspond to a 1:2 (DNA:site) molar ratio. In this case we propose that the absence of A-tracts on GC_22 allows the Nflag+Nhelix+HMGbox 1 module (site 1) to bind freely, giving a 1:2 (DNA:site1) complex (Figure 7B, bottom). On near-saturation of the Nflag+Nhelix+HMGbox 1 fragment, HMG-box 2 (site 2) binding becomes predominant, giving a 1:2 ratio for site 2. Therefore, for the GC_22 complex the overall molar ratio (DNA:protein) is 1:1, whereas for the Af2_22 it is 3:2. Thus, the ITC experiments provide further experimental evidence that the A-tracts dictate the mode of DNA binding by Abf2p.

DISCUSSION

Abf2p functions like a staple on DNA with each of the two HMG-boxes bending the DNA by 90°, thus generating a U turn. In this arrangement, the N- and C-terminus of Abf2p are close to each other. In the free protein, the two ends are far apart suggesting that the protein conformation observed in the crystal is induced by DNA binding. We also show that the HMG-boxes alone have very low DNA binding efficiency, and the presence of the N-flag+N-helix module is crucial for mtDNA maintenance, as *in vivo* and microscopy results clearly demonstrate. Our findings sug-

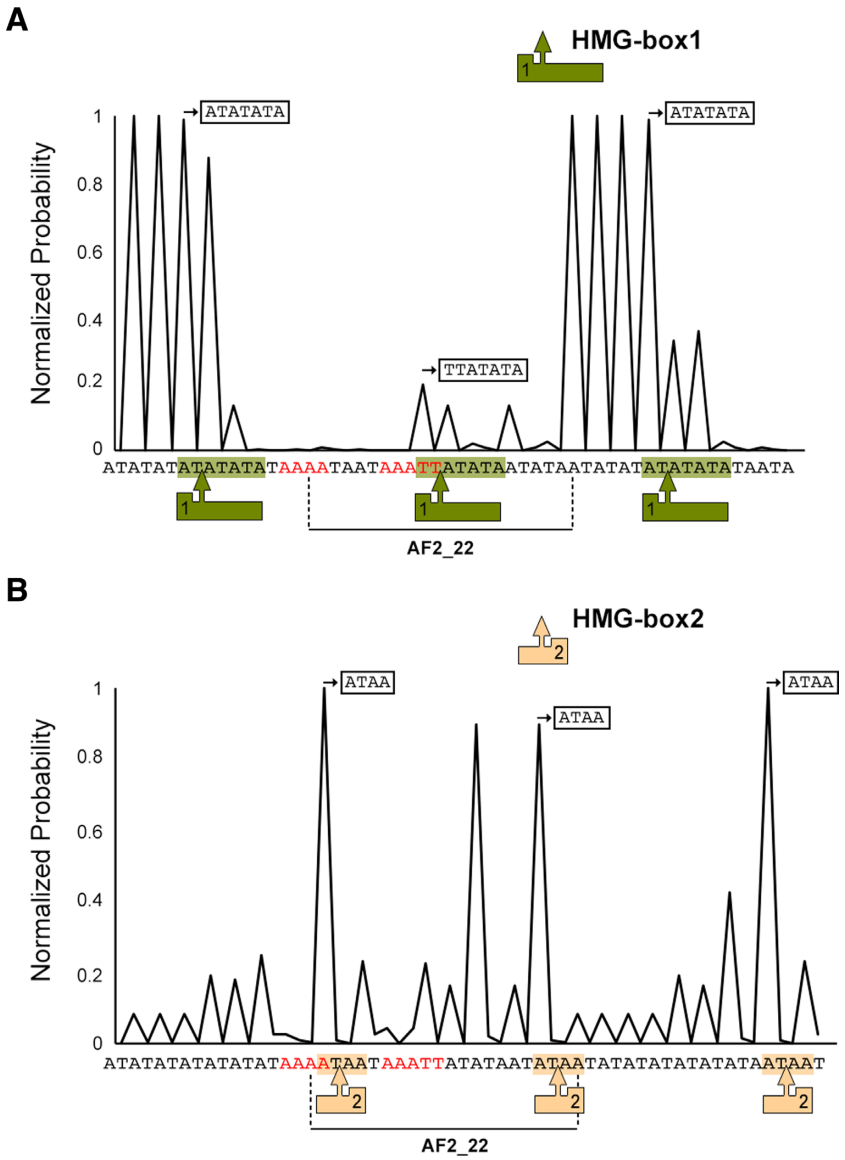

Figure 6. Binding probability of HMG-boxes on the 64 bp sequence. (**A**) Normalized probability for HMG-box 1 to bind and distort a sequence stretch (see 'Materials and Methods' section). Each peak corresponds to the probability of HMG-box 1 binding to the 7 bp stretch starting from the peak position. The said sequence stretches are indicated in boxes above each relevant peak and additionally highlighted in green on the 64 bp sequence. HMG-box 1 is represented by a green L shape (1) and the site of insertion is indicated by an arrow in each case. The location of the Af2_22 sequence within the 64 bp sequence is demarcated. The A-tracts within the 64 bp sequence are shown in red. (**B**) Normalized probability for HMG-box 2. Features are represented similar to (A). The 4 bp HMG-box 2 binding sites are shown for representative peaks.

gest a sequential DNA binding model where the Nflag+N-helix+HMG-box 1 module initiates DNA binding. In this event, the N-helix functions as a pin-lock that consolidates the N-hydrophobic core between the N-helix, HMG-box 1 helix3 and the Linker. This arrangement might facilitate interaction of HMG-box 2 with downstream DNA and consequent bending. The N-flag and the first turn of the N-helix make contacts with DNA minor groove further downstream, which presumably contributes to the molecular events leading to consolidation of the N-hydrophobic core and to the stability of the U-turn. The final outcome is a stable protein/DNA complex where the N- and C-termini of Abf2p are in close proximity.

Abf2p and its human counterpart TFAM are the best characterized mtDNA packaging proteins. Both show universal features of mtDNA bending and compaction through their HMG-box domains, but present very low sequence similarity. We don't know if this reflects different roles played by the two proteins in nucleoid organization and segregation or possibly, differences in mtDNA base composition and topology. Indeed, striking differences are found between respective mtDNAs: human mtDNA is a small G-C rich circular molecule of 16.5 kbp, whilst *S. cerevisiae* mtDNA is an A-T rich linear molecule of ∼80 kbp. Both proteins generate a U-turn on the DNA (61,19). However, TFAM, by virtue of its long and flexible 30aa linker, intertwines and bends the DNA, positioning the HMG-boxes

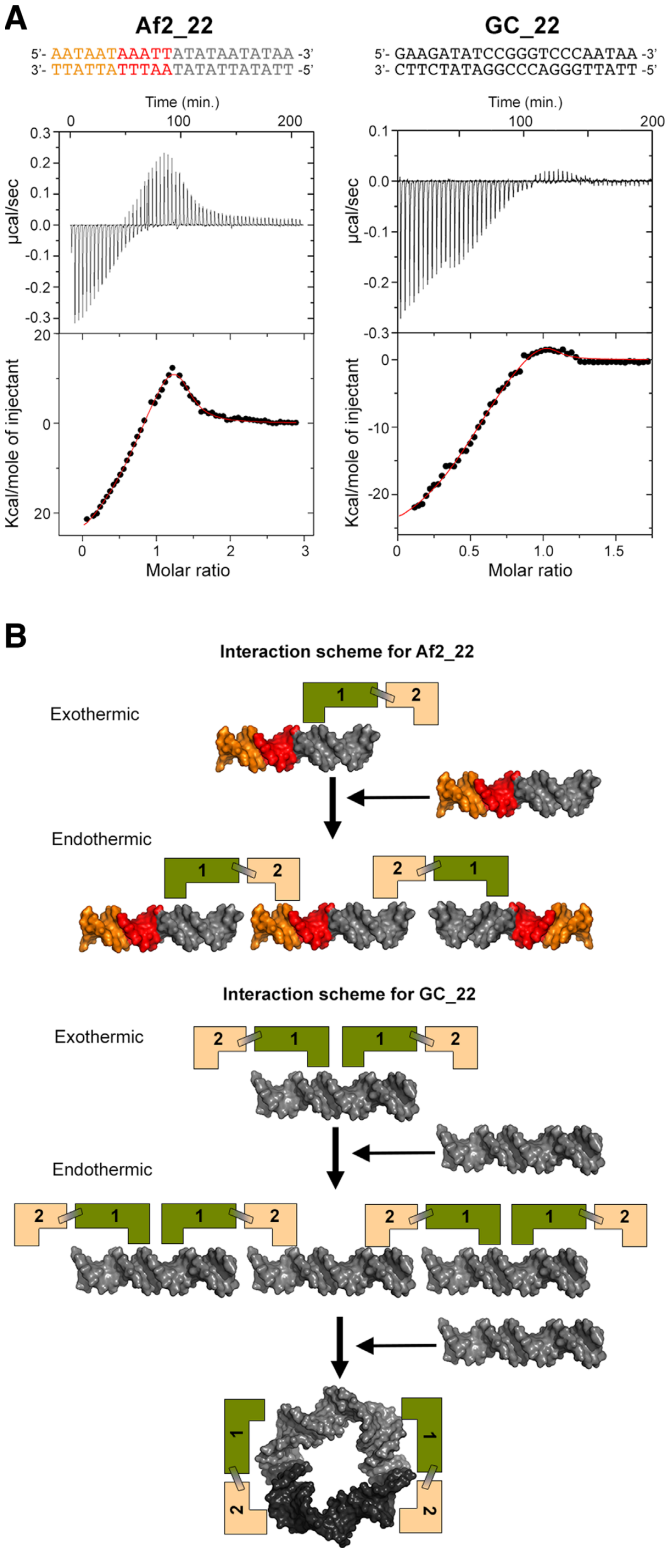

Figure 7. Thermodynamics of interaction of Abf2p with DNA sequences with and without A-tracts (**A**) Isothermal titration calorimetry thermograms (above) and fitting of the binding isotherms (below) to a model with two independent binding sites for Af2_22 (left) and GC_22 (right) DNA. The corresponding DNA sequences are shown at the top of each thermogram. (**B**) Schemes of the binding model with two independent binding sites for Af2_22 (top) and GC_22 DNA (bottom). Abf2p is depicted schematically with the HMG-boxes 1 and 2 represented in green (1) and orange (2) respectively. The non A-tract part of the ARSm-like sequence in Af2_22 is represented in orange whereas the A-tract is depicted in red. At the very bottom, the 1:1 stoichiometry is represented as a plausible circle in order to show its compatibility with the crystal structure (the two dsDNA molecules in the circle are depicted in different gray tones).

at opposite faces on the DNA. In contrast, Abf2p possesses a short linker and thus the two vicinal HMG-boxes interact with the DNA from one side. As a result, the N- and C-termini are distant in space in TFAM, but come close in Abf2p. Therefore, the two homologs utilize different binding strategies to achieve a similar DNA bending. This suggests that U-turns constitute the basic structural unit of mitochondrial nucleoid organization in both systems.

Our EMSA assays (Figure 3B) reveal that Abf2p shifts dsDNA at a ratio of 1Abf2p:10-20 bps, which agrees with the number of DNA bps (∼19 bps) contacted by the protein in the crystal structures. Additionally, previous atomic force microscopy studies showed that Abf2p induces compaction at a 1Abf2p:20 bps ratio (15), being more prominent at 1:10. At these ratios TFAM showed higher compaction, which correlated with a reduction in *in vitro* DNA replication and transcription (62). A similar effect was observed for Abf2p, where Abf2p overexpression resulted in loss of y-mtDNA, suggesting excessive DNA compaction and consequent reduced access for the replication machinery (14). The dual-binding found in the crystal structures and the relative conformational freedom of the two HMG-boxes observed from SAXS-EOM analyses demonstrates the capability of the protein to bind two DNA molecules simultaneously. This ability would be compatible with 'cross-strand binding' or DNA looping by Abf2p, as also described for TFAM (63,64). Thus, both DNA bending and cross-strand binding might be underlying the nucleoid packaging mechanism of the two proteins.

We have also demonstrated that the A-tracts direct positioning of Abf2p on DNA. Since Abf2p has been shown to bind GC rich sequences with relatively higher preference (11), additional crystallization trials were made with DNA fragments of different lengths that are either GC-rich or devoid of A-tracts. Both trials however yielded poorly diffracting crystals, suggesting high disorder. Presumably, by specifying positioning of the protein, the A-tracts induced a complex that formed the ordered crystals reported here. Accordingly, previous competition experiments showed that binding of Abf2p to an ARS1 probe could not be competed by a poly(dA)·poly(dT) double-stranded homo-polymer (10,11), indicating the incapability of this protein to bind to poly-adenine DNA. A-tract guided nucleoprotein interactions have also been proposed in the nucleus, where biochemical studies have shown that A-tracts help in positioning nucleosomes (56–59). This has been further corroborated by computational approaches (41,55). Additionally, other biochemical experiments have demonstrated that the A-tracts are involved in DNA synthesis termination in the human immunodeficiency virus (65,66) and in regulation of transcription in pathogenic bacteria (67–69). Thus, A-tract mediated control of protein/nucleic acid interactions appears to be a general strategy across species. We and others detected that y-mtDNA has a high content of A-tracts

(3,60). Thus, location of these tracts might determine strategic positioning of Abf2p, suggesting a DNA-mediated control of nucleoid architecture and of mtDNA accessibility.

Y-mtDNA undergoes frequent recombination ((7) and references therein). Several studies support that recombination is involved in y-mtDNA replication and maintenance, and a mechanism of recombination-dependent replication has been proposed (70,71). In mitochondria, Abf2p levels correlate with the level of recombination intermediates (14,47) and *in vitro* studies showed that Abf2p and its orthologs from different yeast species show preference for such DNA structures (50). This suggested that Abf2p could be involved in maintenance of y-mtDNA copy number via interaction with recombination intermediates (14,47). Our studies confirm the binding of Abf2p to four-way junctions. The different Abf2p mutants additionally show that such binding involves cooperativity between different protein regions and indicate an important role of the N-flag+N-helix module. Mut2 and 3 bind weakly to four-way junction and cannot support growth on glycerol plates, whereas growth is observed for Mut6 which shows better binding. This is consistent with the hypothesis that Abf2p influences the levels of recombination intermediates *in vivo,* thus putatively playing a role in recombination-dependent replication and DNA maintenance. The preference of Abf2p for four-way junctions over dsDNA (50) suggests a second level of regulation in which binding to replication intermediates is favored over DNA compaction that, in turn, is regulated by A-tracts. Binding to four-way junctions, in conjunction with DNA packaging, might provide a protective effect against mutagens through involvement in DNA repair. We can speculate that such a mechanism might be responsible for similar effects observed for HMG-box proteins in general (72,73). In connection to this, a novel and unexpected feature observed in the crystals is DNA end-joining by Abf2p. This could be relevant for cellular processes where DNA ends are available, such as in DNA breaks or recombination. However, any potential biological implication of Abf2p DNA end-joining in recombination or DNA repair remains to be analyzed by further studies.

To conclude, the binding of Abf2p to DNA is modulated by the positioning of poly-adenine tracts, posing a molecular strategy for y-mtDNA regulation and nucleoid organization. Our results demonstrate how the intrinsic structural properties of DNA can play a key role in orchestrating localization of non-specific DNA binding proteins that are essential for mitochondrial maintenance, and thus for cell life.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Javier Bermúdez for technical support; Pau Bernardó (CSB, Montpellier, France) for valuable SAXS related discussions and Raimundo Gargallo (University of Barcelona) for help with CD experiments. We thank Dr Elena Rebollo Arredondo (IBMB, Barcelona) for her extensive help in collection and analysis of fluorescence confocal microscopy data. We thank the Automatic Crystallization Platform (Barcelona Science Park), the European Synchrotron Radiation Facility (ESRF, France) and XALOC beamline at ALBA Synchrotron (Cerdanyola, Spain) and their staff for excellent support. We thank Dr Andrew McCarthy (EMBL, Grenoble) for help with data collection.
Author Contributions: A.C., S.L. and G.M. contributed to cloning, protein production and biophysical analysis; A.C. in crystallization, X-ray structure solution, SAXS studies and fluorescence microscopy; A.C. and S.L. in EMSA studies; A.C., F.B., A. H. and M.O. in MD studies; A.C. and R.P. in ITC studies; A.C. and J.V. in yeast studies. M.S. designed and supervised the project. All authors participated in manuscript writing.

FUNDING

Ministry of Economy and Competitiveness (MINECO) [BFU2012-33516, BFU2015-70645-R to M.S; BIO2012-32868, BFU2014-61670-EXP to M.O.; BFU2011-25697, BFU2014-60550-P to J.V.; FEDER CSIC13-4E-2065 to the Molecular imaging platform IBMB-PCB; Biomolecular and Bioinformatics Resources Platform (ISCIII PT 13/0001/0030) cofunded by the Fondo Europeo de Desarrollo Regional (FEDER) to M.O.]; Generalitat de Catalunya [SGR2009-1366, 2014-SGR-997 to M.S., SGR2009-1348, 2014 SGR-134 to M.O.]; Instituto Nacional de Bioinformática; European Union [FP7-HEALTH-2010-261460, FP7-PEOPLE-2011-290246, FP7-HEALTH-2012-306029-2 to M.S.; H2020-EINFRA-2015-1-675728, H2020-EINFRA-2015-676556 to M.O.]; European Research Council [ERC-2011-ADG_20110209-291433 to M.O.]; ITN Fellowship [FP7-PEOPLE-2011-290246 to A.C.]; Structural Biology Unit at Molecular Biology Institute of Barcelona (IBMB) is a 'Maria de Maeztu' Unit of Excellence awarded by MINECO [MDM-2014-0435]; Severo Ochoa Award of Excellence from MINECO (to IRB Barcelona); ICREA Academia Research (to M.O.). Funding for open access charge: Ministry of Economy and Competitiviness [BFU2015-70645-R].
Conflict of interest statement. None declared.

REFERENCES

1. Nunnari,J. and Suomalainen,A. (2012) Mitochondria: in sickness and in health. *Cell*, **148**, 1145–1159.
2. Archibald,J.M. (2015) Endosymbiosis and eukaryotic cell evolution. *Curr. Biol.*, **25**, R911–R921.
3. Foury,F., Roganti,T., Lecrenier,N. and Purnelle,B. (1998) The complete sequence of the mitochondrial genome of Saccharomyces cerevisiae. *FEBS Lett.*, **440**, 325–331.
4. De Zamaroczy,M. and Bernardi,G. (1986) The primary structure of the mitochondrial genome of Saccharomyces cerevisiae–a review. *Gene*, **47**, 155–177.
5. Chen,X.J. and Butow,R.A. (2005) The organization and inheritance of the mitochondrial genome. *Nat. Rev. Genet.*, **6**, 815–825.
6. Maleszka,R., Skelly,P.J. and Clark-Walker,G.D. (1991) Rolling circle replication of DNA in yeast mitochondria. *EMBO J.*, **10**, 3923–3929.
7. Williamson,D. (2002) The curious history of yeast mitochondrial DNA. *Nat. Rev. Genet.*, **3**, 475–481.
8. Bendich,A.J. (2010) The end of the circle for yeast mitochondrial DNA. *Mol. Cell*, **39**, 831–832.
9. Diffley,J.F. and Stillman,B. (1991) A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 7864–7868.

10. Diffley,J.F. and Stillman,B. (1992) DNA binding properties of an HMG1-related protein from yeast mitochondria. *J. Biol. Chem.*, **267**, 3368–3374.
11. Kucej,M., Kucejova,B., Subramanian,R., Chen,X.J. and Butow,R.A. (2008) Mitochondrial nucleoids undergo remodeling in response to metabolic cues. *J. Cell Sci.*, **121**, 1861–1868.
12. Newman,S.M., Zelenaya-Troitskaya,O., Perlman,P.S. and Butow,R.A. (1996) Analysis of mitochondrial DNA nucleoids in wild-type and a mutant strain of Saccharomyces cerevisiae that lacks the mitochondrial HMG box protein Abf2p. *Nucleic Acids Res.*, **24**, 386–393.
13. Bunoust,O., Devin,A., Avéret,N., Camougrand,N. and Rigoulet,M. (2005) Competition of electrons to enter the respiratory chain: a new regulatory mechanism of oxidative metabolism in Saccharomyces cerevisiae. *J. Biol. Chem.*, **280**, 3407–3413.
14. Zelenaya-Troitskaya,O., Newman,S.M., Okamoto,K., Perlman,P.S. and Butow,R.A. (1998) Functions of the high mobility group protein, Abf2p, in mitochondrial DNA segregation, recombination and copy number in Saccharomyces cerevisiae. *Genetics*, **148**, 1763–1776.
15. Friddle,R.W., Klare,J.E., Martin,S.S., Corzett,M., Balhorn,R., Baldwin,E.P., Baskin,R.J. and Noy,A. (2004) Mechanism of DNA compaction by yeast mitochondrial protein Abf2p. *Biophys. J.*, **86**, 1632–1639.
16. Zweifel,S.G. and Fangman,W.L. (1990) Creation of ARS activity in yeast through iteration of non-functional sequences. *Yeast (Chichester, England)*, **6**, 179–186.
17. Fangman,W.L., Henly,J.W., Churchill,G. and Brewer,B.J. (1989) Stable maintenance of a 35-base-pair yeast mitochondrial genome. *Mol. Cell. Biol.*, **9**, 1917–1921.
18. Rubio-Cosials,A., Sidow,J.F., Jiménez-Menéndez,N., Fernández-Millán,P., Montoya,J., Jacobs,H.T., Coll,M., Bernadó,P. and Solà,M. (2011) Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. *Nat. Struct. Mol. Biol.*, **18**, 1281–1289.
19. Ngo,H.B., Lovely,G.A., Phillips,R. and Chan,D.C. (2014) Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation. *Nat. Commun.*, **5**, 1–12.
20. Goulas,T., Cuppari,A., Garcia-Castellanos,R., Snipas,S., Glockshuber,R., Arolas,J.L. and Gomis-Rüth,F.X. (2014) The pCri System: a vector collection for recombinant protein expression and purification. *PLoS One*, **9**, e112643.
21. Diffley,J.F. and Stillman,B. (1991) A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 7864–7868.
22. Nishioka,M., Mizuguchi,H., Fujiwara,S., Komatsubara,S., Kitabayashi,M., Uemura,H., Takagi,M. and Imanaka,T. (2001) Long and accurate PCR with a mixture of KOD DNA polymerase and its exonuclease deficient mutant enzyme. *J. Biotechnol.*, **88**, 141–149.
23. Doublié,S. (2007) Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Methods in Mol. Biol.*, **363**, 91–108.
24. Kabsch,W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
25. Evans,P.R. and Murshudov,G.N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.*, **69**, 1204–1214.
26. Evans,P.R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 282–292.
27. Sheldrick,G.M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 479–485.
28. Winn,M.D., Ballard,C.C., Cowtan,K.D., Dodson,E.J., Emsley,P., Evans,P.R., Keegan,R.M., Krissinel,E.B., Leslie,A.G., McCoy,A. *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 235–242.
29. Thorn,A. and Sheldrick,G.M. (2011) ANODE: anomalous and heavy-atom density calculation. *J. Appl. Crystallogr.*, **44**, 1285–1287.
30. Adams,P.D., Afonine,P.V., Bunkóczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.-W.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
31. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
32. Petoukhov,M.V., Franke,D., Shkumatov,A.V., Tria,G., Kikhney,A.G., Gajda,M., Gorba,C., Mertens,H.D.T., Konarev,P.V. and Svergun,D.I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **45**, 342–350.
33. Bernadó,P. and Svergun,D.I. (2012) Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Mol. Biosyst.*, **8**, 151–167.
34. Tria,G., Mertens,H.D., Kachala,M. and Svergun,D.I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ*, **2**, 207–217.
35. Ivani,I., Dans,P.D., Noy,A., Pérez,A., Faustino,I., Hospital,A., Walther,J., Andrio,P., Goñi,R., Balaceanu,A. *et al.* (2016) Parmbsc1: a refined force field for DNA simulations. *Nat. Methods*, **13**, 55–58.
36. Roe,D.R. and Cheatham,T.E. (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.*, **9**, 3084–3095.
37. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 27–28.
38. Lavery,R., Moakher,M., Maddocks,J.H., Petkeviciute,D. and Zakrzewska,K. (2009) Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.*, **37**, 5917–5929.
39. Blanchet,C., Pasi,M., Zakrzewska,K. and Lavery,R. (2011) CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Res.*, **39**, W68–W73.
40. Geggier,P., Dršata,T., Špačková,N., Jurečka,P., Zgarbová,M., Šponer,J. and Lankaš,F. (2014) Mechanical properties of symmetric and asymmetric DNA A-tracts: implications for looping and nucleosome positioning. *Nucleic Acids Res.*, **42**, 7383–7394.
41. Portella,G., Battistini,F. and Orozco,M. (2013) Understanding the connection between epigenetic DNA methylation and nucleosome positioning from computer simulations. *PLoS Comput. Biol.*, **9**, e1003354.
42. Sikorski,R.S. and Hieter,P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics*, **122**, 19–27.
43. Huxley,C., Green,E.D. and Dunham,I. (1990) Rapid assessment of S. cerevisiae mating type by PCR. *Trends Genet.*, **6**, 236.
44. Boeke,J., Trueheart,J., Natsoulis,G. and Fink,G. (1987) [10]5-Fluoroorotic acid as a selective agent in yeast molecular genetics sciencedirect. *Methods in Enzymol.*, **154**, 164–175.
45. Lundblad,V. and Zhou,H. (2001) Manipulation of Plasmids from Yeast Cells.*Curr. Protoc. Mol. Biol.*, **39**, 13.9.1–13.9.6.
46. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.
47. MacAlpine,D., Perlman,P. and Butow,R. (1998) The high mobility group protein Abf2p influences the level of yeast mitochondrial DNA recombination intermediates in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 6739–6743.
48. Miyakawa,I., Okamuro,A., Kinsky,S., Visacka,K., Tomaska,L. and Nosek,J. (2009) Mitochondrial nucleoids from the yeast Candida parapsilosis: expansion of the repertoire of proteins associated with mitochondrial DNA. *Microbiology*, **155**, 1558–1568.
49. Visacka,K., Gerhold,J.M., Petrovicova,J., Kinsky,S., Joers,P., Nosek,J., Sedman,J. and Tomaska,L. (2009) Novel subfamily of mitochondrial HMG box-containing proteins: functional analysis of Gcf1p from Candida albicans. *Microbiology*, **155**, 1226–1240.
50. Bakkaiova,J., Marini,V., Willcox,S., Nosek,J., Griffith,J.D., Krejci,L. and Tomaska,L. (2015) Yeast mitochondrial HMG proteins: DNA-binding properties of the most evolutionarily divergent component of mitochondrial nucleoids. *Biosci. Rep.*, **36**, e00288.
51. Kao,L.R., Megraw,T.L. and Chae,C.B. (1993) Essential role of the HMG domain in the function of yeast mitochondrial histone HM: functional complementation of HM by the nuclear nonhistone protein NHP6A. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 5598–5602.
52. Kikhney,A.G. and Svergun,D.I. (2015) A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.*, **589**, 2570–2577.

53. Kohn,J.E. and Plaxco,K.W. (2005) Engineering a signal transduction mechanism for protein-based biosensors. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 10841–10845.
54. Haran,T.E. and Mohanty,U. (2009) The unique structure of A-tracts and intrinsic DNA bending. *Q. Rev. Biophys.*, **42**, 41–81.
55. Deniz,O., Flores,O., Battistini,F., Pérez,A., Soler-López,M. and Orozco,M. (2011) Physical properties of naked DNA influence nucleosome positioning and correlate with transcription start and termination sites in yeast. *BMC Genomics*, **12**, 489–500.
56. Rhodes,D. (1979) Nucleosome cores reconstituted from poly (dA-dT) and the octamer of histones. *Nucleic Acids Res.*, **6**, 1805–1816.
57. Kunkel,G.R. and Martinson,H.G. (1981) Nucleosomes will not form on double-stranded RNa or over poly(dA).poly(dT) tracts in recombinant DNA. *Nucleic Acids Res.*, **9**, 6869–6888.
58. Prunell,A. (1982) Nucleosome reconstitution on plasmid-inserted poly(dA) . poly(dT). *EMBO J.*, **1**, 173–179.
59. Puhl,H.L., Gudibande,S.R. and Behe,M.J. (1992) Poly[d(A.T)] and other synthetic polydeoxynucleotides containing oligoadenosine tracts form nucleosomes easily. *J. Mol. Biol.*, **222**, 1149–1160.
60. Ehrlich,S., Thiery,J.-P. and Bernardi,G. (1972) The mitochondrial genome of wild-type yeast cells III. The pyrimidine tracts of mitochondrial DNA. *J. Mol. Biol.*, **65**, 207–212.
61. Rubio-Cosials,A. and Solà,M. (2013) U-turn DNA bending by human mitochondrial transcription factor A. *Curr. Opin. Struct. Biol.*, **23**, 116–124.
62. Farge,G., Mehmedovic,M., Baclayon,M., van den Wildenberg,S.M., Roos,W.H., Gustafsson,C.M., Wuite,G.J. and Falkenberg,M. (2014) In vitro-reconstituted nucleoids can block mitochondrial DNA replication and transcription. *Cell Rep.*, **8**, 66–74.
63. Kukat,C., Davies,K.M., Wurm,C.A., Spåhr,H., Bonekamp,N.A., Kühl,I., Joos,F., Polosa,P.L., Park,C.B., Posse,V. *et al.* (2015) Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 11288–11293.
64. Kaufman,B.A., Durisic,N., Mativetsky,J.M., Costantino,S., Hancock,M.A., Grutter,P. and Shoubridge,E.A. (2007) The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Mol. Biol. Cell*, **18**, 3225–3236.
65. Lavigne,M., Roux,P., Buc,H. and Schaeffer,F. (1997) DNA curvature controls termination of plus strand DNA synthesis at the centre of HIV-1 genome. *J. Mol. Biol.*, **266**, 507–524.
66. Lavigne,M. and Buc,H. (1999) Compression of the DNA minor groove is responsible for termination of DNA synthesis by HIV-1 reverse transcriptase. *J. Mol. Biol.*, **285**, 977–995.
67. Falconi,M., Colonna,B., Prosseda,G., Micheli,G. and Gualerzi,C.O. (1999) Thermoregulation of Shigella and Escherichia coli EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of virF promoter to transcriptional repressor H-NS. *EMBO J.*, **17**, 7033–7043.
68. Prosseda,G., Falconi,M., Giangrossi,M., Gualerzi,C.O., Micheli,G. and Colonna,B. (2004) The virF promoter in Shigella: more than just a curved DNA stretch. *Mol. Microbiol.*, **51**, 523–537.
69. Hagerman,P.J. (1990) Sequence-directed curvature of DNA. *Annu. Rev. Biochem.*, **59**, 755–781.
70. Gerhold,J.M., Aun,A., Sedman,T., Jõers,P. and Sedman,J. (2010) Strand invasion structures in the inverted repeat of Candida albicans mitochondrial DNA reveal a role for homologous recombination in replication. *Mol. Cell*, **39**, 851–861.
71. Gerhold,J.M., Sedman,T., Visacka,K., Slezakova,J., Tomaska,L., Nosek,J. and Sedman,J. (2014) Replication intermediates of the linear mitochondrial DNA of Candida parapsilosis suggest a common recombination based mechanism for yeast mitochondria. *J. Biol. Chem.*, **289**, 22659–22670.
72. Sia,R.A., Carrol,S., Kalifa,L., Hochmuth,C. and Sia,E.A. (2009) Loss of the mitochondrial nucleoid protein, Abf2p, destabilizes repetitive DNA in the yeast mitochondrial genome. *Genetics*, **181**, 331–334.
73. Bakkaiova,J., Arata,K., Matsunobu,M., Ono,B., Aoki,T., Lajdova,D., Nebohacova,M., Nosek,J., Miyakawa,I. and Tomaska,L. (2014) The strictly aerobic yeast Yarrowia lipolytica tolerates loss of a mitochondrial DNA-packaging protein. *Eukaryot. Cell*, **13**, 1143–1157.