The EMBO Journal Vol.16 No.11 pp.3198–3206, 1997

The HMG-box mitochondrial transcription factor xl-mtTFA binds DNA as a tetramer to activate bidirectional transcription

Igor Antoshechkin, Daniel F.Bogenhagen[1] and Iris A.Mastrangelo[2]

Department of Pharmacological Sciences, State University of New York at Stony Brook, Stony Brook, NY 11794-8651 and [2]Department of Biology, Brookhaven National Laboratory, Upton, NY 11973-5000, USA

[1]Corresponding author

The mitochondrial HMG-box transcription factor xl-mtTFA activates bidirectional transcription by binding to a site separating two core promoters in *Xenopus laevis* mitochondrial DNA (mtDNA). Three independent approaches were used to study the higher order structure of xl-mtTFA binding to this site. First, co-immunoprecipitation of differentially tagged recombinant mtTFA derivatives established that the protein exists as a multimer. Second, *in vitro* chemical cross-linking experiments provided evidence of cross-linked dimers, trimers and tetramers of xl-mtTFA. Finally, high resolution scanning transmission electron microscopy (STEM) established that xl-mtTFA binds to the specific promoter-proximal site predominantly as a tetramer. Computer analysis of several previously characterized binding sites for xl-mtTFA revealed a fine structure consisting of two half-sites in a symmetrical orientation. The predominant sequence of this dyad symmetry motif shows homology to binding sites of sequence-specific HMG-box-containing proteins such as Sry and Lef-1. We suggest that bidirectional activation of transcription results from the fact that binding of a tetramer of xl-mtTFA permits symmetrical interactions with other components of the transcription machinery at the adjacent core promoters.

Keywords: HMG box/mitochondria/mtTFA/STEM/transcription

Introduction

Transcription of *Xenopus laevis* mitochondrial DNA requires two factors for efficient promoter utilization *in vitro* in addition to mitochondrial RNA polymerase (Antoshechkin and Bogenhagen, 1995). One of them, xl-mtTFB, is necessary for basal transcription. It acts as a specificity factor, which allows RNA polymerase to locate and initiate transcription from a minimal octanucleotide promoter encompassing the transcriptional start site (Bogenhagen and Insdorf, 1988; Bogenhagen and Romanelli, 1988; Bogenhagen, 1996). The second protein, an HMG box-containing transcription factor xl-mtTFA, is responsible for activated transcription. Xl-mtTFA interacts with multiple DNA sequences surrounding *Xenopus* mitochondrial promoters and causes up to a 10-fold stimulation of promoter utilization *in vitro*. Deletion analysis of the *Xenopus* mitochondrial control region has shown that a single binding site located between two major bidirectional promoters is sufficient for activation of divergent transcription from heavy strand promoter 1 (HSP1) and light strand promoter 2 (LSP2) (Antoshechkin and Bogenhagen, 1995). This rules out the possibility that formation of a large nucleoprotein array with xl-mtTFA bound at several sites surrounding the promoters is necessary for transcription stimulation.

The mechanism by which xl-mtTFA activates transcription is not well understood. The amphibian factor is closely related to human and yeast mtTFA proteins. All three are abundant mitochondrial proteins with two HMG boxes and similar ability to bind to non-specific DNA (Diffley and Stillman, 1992; Fisher *et al.*, 1992; Antoshechkin and Bogenhagen, 1995). It is likely that all three are involved in packaging and maintaining their respective mitochondrial genomes, although a genetic demonstration that mtTFA is required for mtDNA maintenance has only been possible in *Saccharomyces cerevisiae* (Diffley and Stillman, 1991). The yeast mtTFA lacks a C-terminal tail found in vertebrate mtTFA proteins and does not efficiently stimulate transcription (Xu and Clayton, 1992). This C-terminal tail is required for the transcription stimulation activity of human (Dairaghi *et al.*, 1995) and *Xenopus* mtTFA (I.Antoshechkin, unpublished observation). Dairaghi *et al.* (1995) used recombinant DNA methodology to transfer the C-terminal tail of human mtTFA to yeast mtTFA, producing a chimeric protein with the ability to stimulate transcription. These experiments suggest that the C-terminal tail of mtTFA interacts with the basal transcription machinery to activate transcription.

Bidirectional transcription activation by xl-mtTFA is consistent with the observation that the human homolog, h-mtTFA, can activate transcription when the orientation of its binding site is reversed (Fisher *et al.*, 1987). We envisage two alternative models to explain how xl-mtTFA may be able to activate transcription in a bidirectional manner. One possibility is that a single xl-mtTFA molecule may be able to bind the promoter activation site in two orientations to interact with the basal transcription apparatus at either one or the other promoter. Alternatively, mtTFA may bind as a symmetrical multimer to DNA to interact with the basal transcription machinery on both sides. Previous physical studies of mtTFA reported a sedimentation coefficient of 2S for both *Xenopus* and human mtTFAs, suggesting that both proteins were monomers (Mignotte and Barat, 1986; Fisher and Clayton, 1988). However, we found that xl-mtTFA had anomalous sedimentation behavior and found it necessary to employ other methods to study its quaternary structure. In this paper, we use high resolution scanning transmission electron microscopy (STEM) and complementary biochemical analyses to elucidate the structure of xl-mtTFA

© Oxford University Press

bound to mitochondrial DNA. Our results demonstrate that xl-mtTFA binds to the promoter activation site preferentially as a tetramer. We have also identified a bipartite motif with partial homology to binding sites of such sequence-specific HMG-box proteins as Sry and Lef-1 that may be a preferred recognition sequence for the xl-mtTFA tetramer.

Results

Co-immunoprecipitation of epitope-tagged xl-mtTFA

To test whether xl-mtTFA forms a dimer or a higher order structure, we co-expressed two mtTFA derivatives with different peptide tags to determine whether both derivatives would co-precipitate with an antiserum directed against one of the two epitopes. Xl-mtTFA was cloned into two expression vectors which allow production of the protein in soluble form fused at the N-terminus to either T7-Tag or S-Tag peptides (Figure 1A). T7-Tag is an 11 amino acid peptide representing the natural N-terminal end of the T7 major capsid protein. A mouse monoclonal antibody directed against this peptide is commercially available. S-Tag is a 15 amino acid peptide derived from pancreatic ribonuclease A. It is able to associate with a 104 amino acid S-protein to form an enzymatically active complex. This interaction makes it possible to detect recombinant proteins carrying the S-Tag using the S-protein (Kim and Raines, 1993).

The two expression vectors carrying different selectable markers for ampicillin and kanamycin resistance were co-transfected into *Escherichia coli* BL21 cells and bacterial colonies harboring both plasmids were selected. Cells carrying both constructs were treated with IPTG to induce protein expression and lysed. Protein complexes were precipitated with the antibody directed against the T7-Tag, as described in Materials and methods. Proteins were fractionated by SDS–PAGE, transferred onto a membrane, and S-Tag fusion proteins were visualized with alkaline phosphatase-conjugated S-protein. As shown in Figure 1B, S-tagged xl-mtTFA was detected in the precipitate (lane 1). No S-tagged protein was precipitated in control experiments when the T7 antibody was omitted (lane 2) or when the cells expressed only the S-Tag construct (lane 3). This shows that precipitation is dependent on the presence of both the anti-T7-Tag antibody and the epitope tag. These results argue that xl-mtTFA associates into multimers under physiological conditions. However, they do not address the question of the number of subunits in the native molecule.

To study the quaternary structure of xl-mtTFA, we first used gel filtration and glycerol gradient centrifugation. Gel filtration analysis performed at various salt concentrations, using both Superdex 75 and Superose 6 matrices (Pharmacia), showed that xl-mtTFA interacted strongly with the column resins, preventing unambiguous determination of the Stokes' radius. Glycerol gradient sedimentation was also used, since this is a conventional method that has been used to suggest that HMG1 proteins can associate as tetramers under some conditions (Alexandrova and Beltchev, 1987). We performed sedimentation experiments with xl-mtTFA at salt concentrations ranging from 50 to 500 mM NaCl. In all cases, xl-mtTFA sedimented with a coefficient of ~2S, expected for a globular protein with a molecular weight of ~30 kDa. To explore the possibility that a multimeric form of xl-mtTFA might dissociate to smaller subunits during sedimentation, we also subjected cross-linked xl-mtTFA to sedimentation analysis. Cross-linking did not affect the sedimentation properties of xl-mtTFA but did show that complexes as large as tetramers sedimented at 2S. This suggested that the apparent slow sedimentation is caused by the irregularity of the protein shape (data not shown). We concluded that the anomalous behavior of xl-mtTFA does not permit a rigorous determination of the quaternary structure of xl-mtTFA using these techniques.

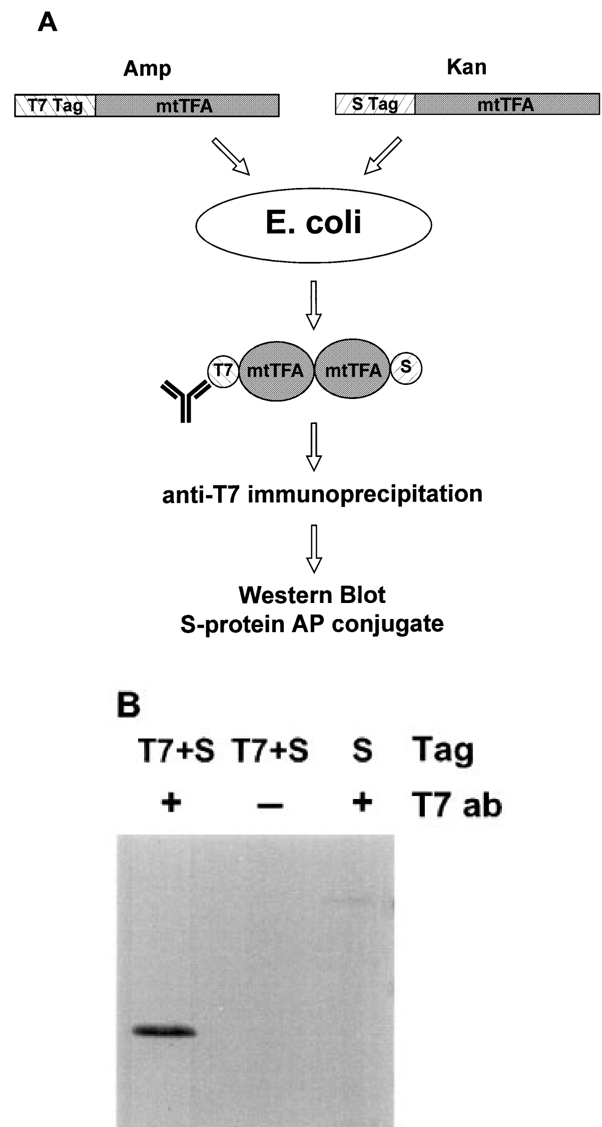

Fig. 1. Co-immunoprecipitation of differentially tagged xl-mtTFA proteins. (**A**) Scheme of the experiment (details in text). (**B**) Western blot analysis of the protein precipitated with anti-T7-Tag antibody. The blot was probed for xl-mtTFA labeled with the S tag using alkaline phosphatase-conjugated S protein. Reactive species were detected with colorimetric reagents, as described in Materials and methods.

In vitro cross-linking of xl-mtTFA

Chemical cross-linking is a widely used technique for protein structural studies. One of the most common

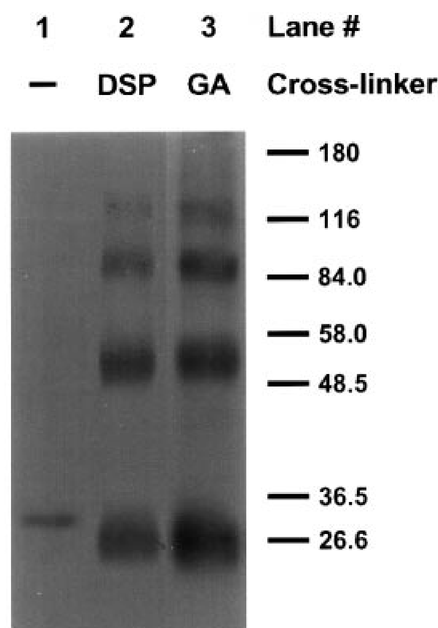

Fig. 2. *In vitro* cross-linking reveals xl-mtTFA multimers. Purified xl-mtTFA (mol. wt 28.1 kDa) was treated with either DSP (lane 2) or glutaraldehyde (lane 3), run on an SDS protein gel, and analyzed by Western blotting, as described in Materials and methods. Untreated xl-mtTFA is shown in lane 1. Positions of molecular weight markers (kDa) are indicated on the right. Bands produced by cross-linked species are not as distinct as the band of untreated xl-mtTFA due to heterogeneous intra- and intermolecular cross-linking.

cross-linking reagents is dithio*bis*(succinimidylpropionate) (DSP), also known as Lomant's reagent (Park *et al.*, 1986; Tarvers *et al.*, 1982). DSP is a homobifunctional *N*-hydroxysuccinimidyl ester which reacts with primary amines commonly found in proteins at physiological pH. It has a relatively short spacer arm that positions the two reactive groups 12 Å apart and minimizes non-specific cross-linking of proteins. We used this reagent to study the multimeric structure of xl-mtTFA *in vitro*.

Purified recombinant xl-mtTFA was treated with DSP, run on an SDS–polyacrylamide gel, and detected by Western blotting, as described in Materials and methods. The gel analysis of DSP-treated xl-mtTFA revealed a ladder of cross-linked species with apparent gel mobilities consistent with dimers, trimers and tetramers (lane 2 of Figure 2). The putative multimers as well as monomeric mtTFA migrated as diffuse bands presumably due to intramolecular cross-linking and formation of mono-adducts by DSP. Longer incubation times or higher concentrations of DSP did not significantly increase the efficiency of the reaction. As shown in lane 3 of Figure 2, the cross-linking pattern is not specific for DSP, since treatment with glutaraldehyde, another bifunctional cross-linker, yields a similar oligomeric ladder. Comparable results (data not shown) were obtained with other cross-linkers such as disuccinimidyl suberate (DSS) and *N*-succinimidyl-6(4′-azido-2′-nitrophenyl-amino)hexanoate (SANPAH). DSP treatment at salt concentrations as high as 1 M KCl did not change the relative abundance of cross-linked oligomers. The proportion of cross- linked tetramers was not increased when cross-linking was performed in the presence of a DNA fragment containing a binding site for xl-mtTFA (data not shown). In all of our cross-linking experiments we never observed complete cross-linking of protein to a single multimeric species (e.g. to a tetramer). It is frequently difficult to drive chemical cross-linking reactions to completion, since accessible amino groups on the protein may form mono-adducts with cross-linker molecules that are unable to react with a second amino group in the protein. We have observed a similar pattern of partially cross-linked forms of mitochondrial SSB, which is known to be a stable tetramer (Mikhailov and Bogenhagen, 1996; other unpublished data). However, the apparent incomplete cross-linking may also reflect an equilibrium distribution of xl-mtTFA in tetramers and smaller oligomeric species. In summary, these cross-linking experiments do not provide a clear model for the quaternary structure of xl-mtTFA but do show that the protein exists in complexes as large as tetramers in the absence of DNA.

EM and STEM show xl-mtTFA forms a specific complex at the promoter activation site

STEM allows direct visualization and molecular mass measurement of free proteins and protein–DNA complexes (Mastrangelo *et al.*, 1989; Stenger *et al.*, 1994; Blackwell *et al.*, 1996). Unlike conventional transmission EM (TEM), which relies on heavy metal shadowing to provide contrast, STEM employs highly efficient detectors operating in dark field mode to provide high resolution images of bare molecules scanned at 2.5 Å intervals (Wall, 1979). This method provides a direct measurement of the mass of macromolecules. We employed STEM to study xl-mtTFA binding to a site referred to as the promoter activation site, which is located between the two major mitochondrial promoters HSP1 and LSP2 in the control region of *Xenopus* mtDNA. Binding of xl-mtTFA to this site was identified previously by DNase I footprinting and shown to stimulate transcription from HSP1 and LSP2 promoters *in vitro* (Antoshechkin and Bogenhagen, 1995).

To devise conditions under which the specific binding of xl-mtTFA to the promoter activation site was preserved during EM, we carried out preliminary experiments using TEM. A pUC9 plasmid carrying the XLMT 895/988 construct, which contains both mitochondrial promoter regions separated by the single xl-mtTFA binding site, was cut with *Afl*III and *Nde*I to produce a 682-bp DNA fragment (Figure 3A). In preliminary experiments, we found that glutaraldehyde cross-linking was required to keep xl-mtTFA complexed with DNA. The reaction conditions were otherwise similar to those used in DNase I footprinting experiments. After 10 min incubation, reaction mixtures were treated with 0.1% glutaraldehyde for 2 min and aliquots were deposited on EM grids. Position histograms of xl-mtTFA bound to the 682-bp DNA fragment obtained by TEM and STEM are shown in Figure 3B and C respectively. The histograms demonstrate that, under the experimental conditions used in both TEM and STEM experiments, xl-mtTFA binds predominantly at the specific promoter activation site.

STEM identifies xl-mtTFA tetramers at the promoter activation site

STEM mass histograms of free and DNA-bound xl-mtTFA are shown in Figure 4A and B respectively. Free xl-mtTFA is distributed in prominent peaks corresponding

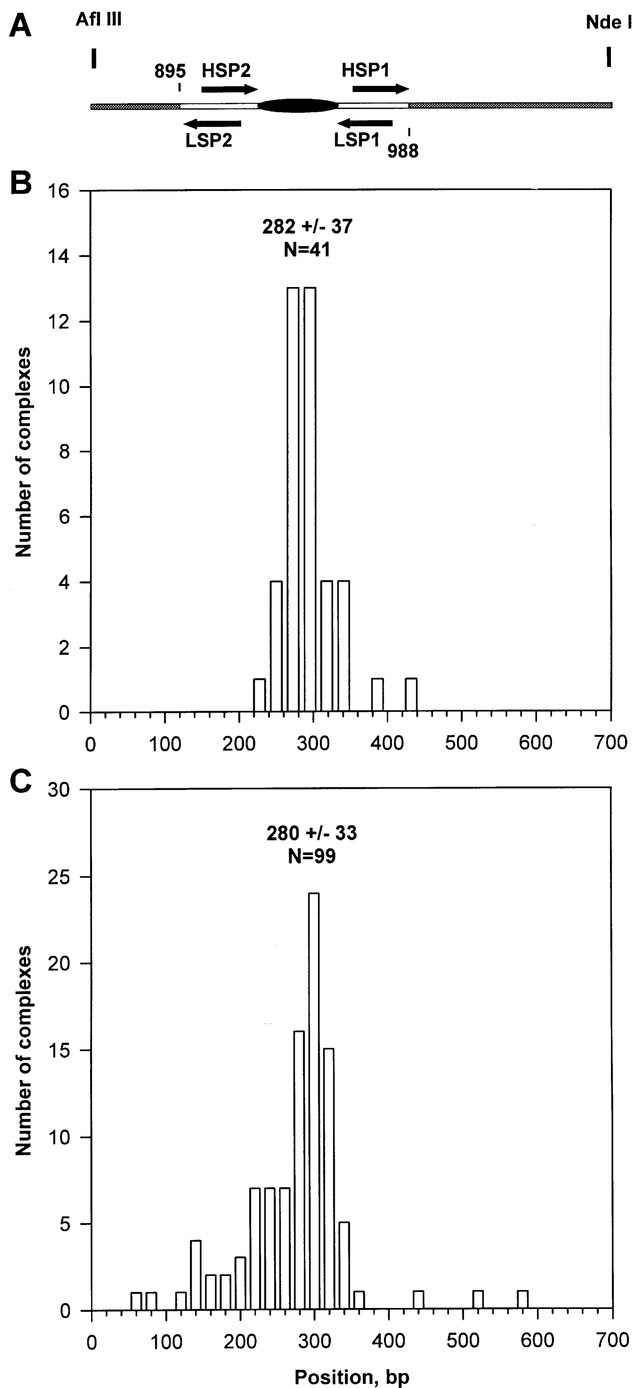

Fig. 3. EM analyses show binding specificity of xl-mtTFA. (**A**) Diagram (not drawn to scale) of the 682-bp L895/R988 DNA fragment, used in EM experiments, which contains two mitochondrial promoter regions separated by the promoter activation site that binds xl-mtTFA. Shaded regions outside the deletion end points depict vector DNA sequences. The dark oval represents mtTFA bound to the promoter activation site. (**B**) and (**C**) Histograms of the distributions of xl-mtTFA binding positions along the DNA fragment obtained in TEM and STEM experiments respectively. Binding positions were determined by measuring distances from the end of the short DNA arm to the center of the protein. Complexes in which DNA ends could not be unambiguously identified, as well as ones formed at intersections of DNA molecules and DNA ends, were not included in this analysis. Mean binding positions, standard deviations and numbers of complexes included in the calculations are indicated.

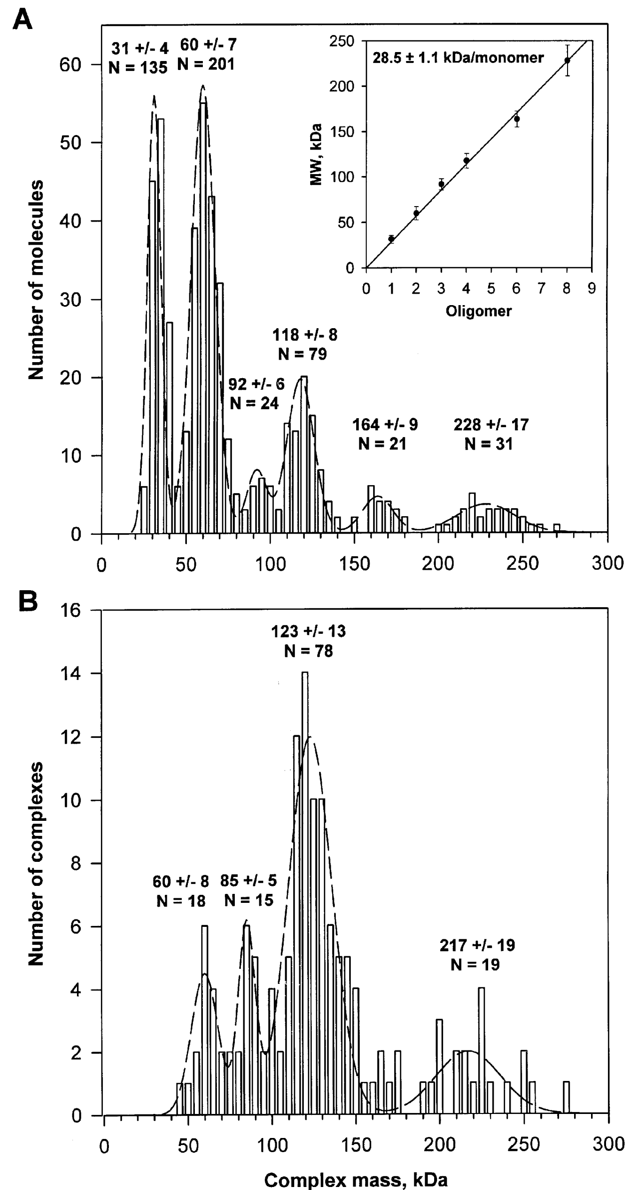

Fig. 4. Tetramers preferentially bind at the promoter activation site. Histograms show STEM mass measurements of (**A**) free protein deposited on EM grids and (**B**) xl-mtTFA bound to the promoter activation site. Mean masses and standard deviations of Gaussian curves fitted to peaks and the number of molecules are indicated. The inset in (A) plots mean mass versus the oligomer number and defines a mass ladder (Mastrangelo *et al.*, 1989). The slope of the regression line, 28.5 ± 1.1 kDa, is the monomer unit mass.

to monomers, dimers and tetramers, with minor peaks representing trimers, hexamers and octamers. In the inset of Figure 4A, the slope of the mass ladder is 28.5 ± 1.1 kDa, the monomer unit mass. This value coincides closely with the 28.1 kDa mass of xl-mtTFA predicted from the amino acid sequence. The mass ladder, therefore, confirms the identification of xl-mtTFA oligomers by STEM. The higher frequency of dimers, tetramers and octamers, compared with monomers, trimers and hexamers, suggests that xl-mtTFA dimers assemble into tetramers, and that tetramers can interact to form octamers.

STEM analysis of mtTFA bound at the promoter activation site is shown in Figure 4B. Protein mass was derived

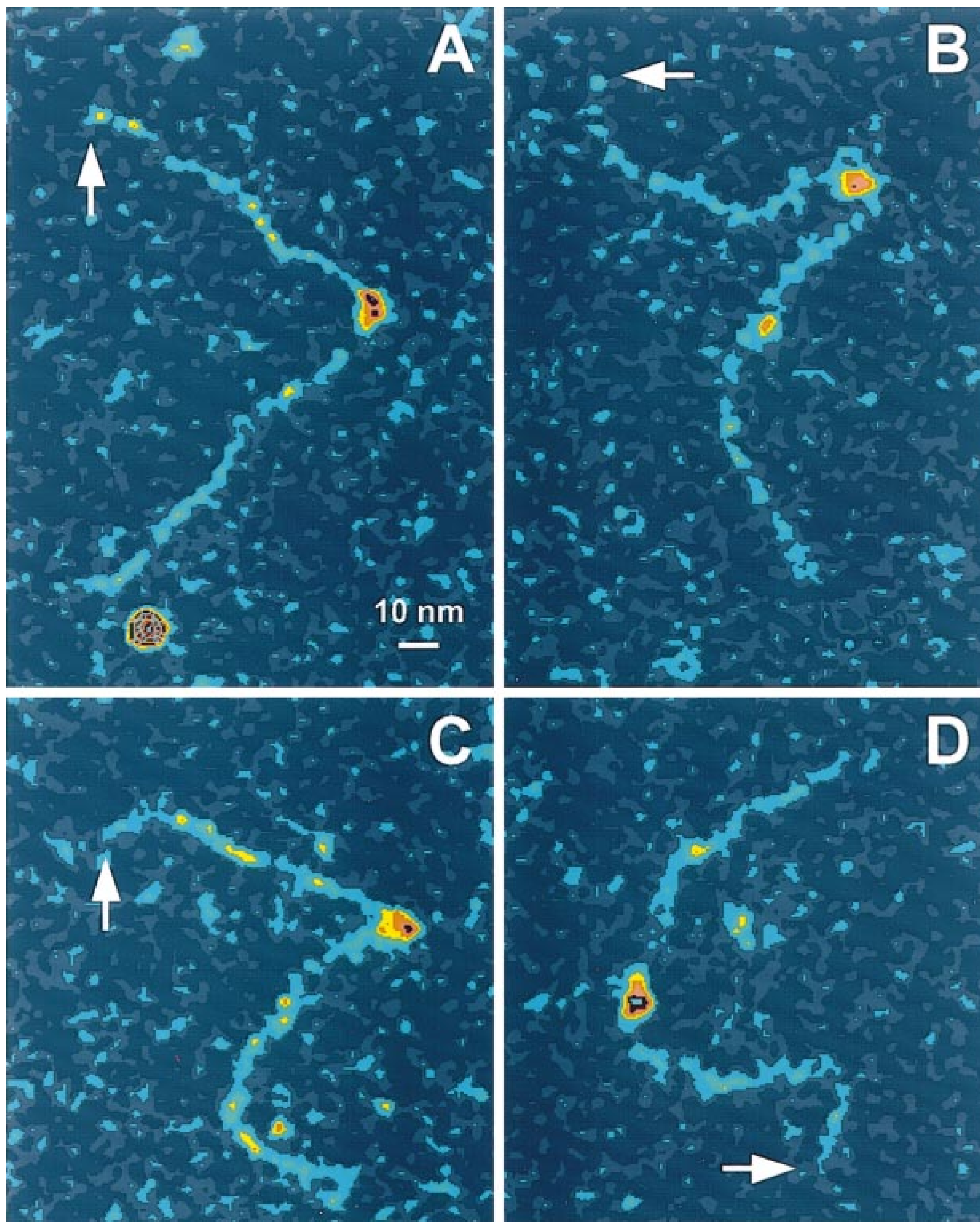

Fig. 5. Representative STEM images of xl-mtTFA tetramers bound to DNA. Differences in the mass density are visualized by color changes representing mass increases of 500 D/1.00 nm^2. As depicted, binding of xl-mtTFA tetramers to the promoter activation site in linear DNA induced a sharp bend in the DNA duplex in most (**A**, **B** and **C**), but not all (**D**) complexes. Arrows mark proximal ends of DNA fragments. The white bar in (A) represents 10 nm.

from the total complex mass by subtracting the DNA mass, assuming a direct DNA pathway through the complex. The mass of protein complexes was not changed significantly by tracing a curved path for the DNA through the complex. Furthermore, the length of DNA covered by mtTFA tetramers averaged 41 $\pm$ 8 bp, in good agreement with the 35 bp DNase I footprint of mtTFA bound to this site. The distribution presents a striking contrast to that observed for free xl-mtTFA in that it is dominated by a 123 kDa peak, which corresponds to an xl-mtTFA tetramer. Three other peaks can be discerned with mean masses of 60, 85 and 217 kDa. These data indicate that xl-mtTFA binds to the promoter activation site primarily as a tetramer.

STEM micrographs show in the majority of cases that binding of xl-mtTFA at the promoter activation site produces a sharp bend of the DNA duplex (Figure 5). This bending is consistent with the known DNA-binding properties of human and yeast mtTFA (Diffley and Stillman, 1992; Fisher *et al.*, 1992), as well as with the general model of interaction between HMG-box proteins and DNA (Haqq *et al.*, 1994; Love *et al.*, 1995). Xl-mtTFA molecules bound to DNA ends in 10 out of 138 complexes. Binding to intersections of two DNA molecules outside the promoter activation site was seen in five cases. These observations are in agreement with findings that HMG-box proteins bind to irregular structures such as four-way junctions, kinked DNA and broken DNA with high affinity regardless of sequence (Grosschedl *et al.*, 1994; Baxevanis and Landsman, 1995; Churchill *et al.*, 1995; Peters *et al.*, 1995; Teo *et al.*, 1995). This binding represented a minor fraction of all complexes and was not included in the histogram in Figure 3C.

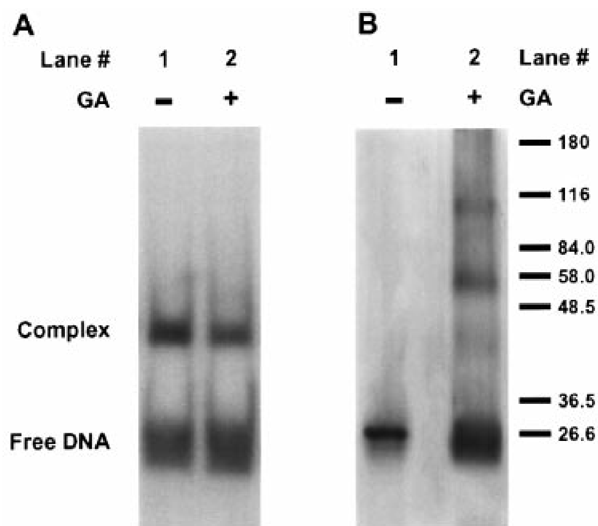

Fig. 6. Protein–DNA complexes fixed with glutaraldehyde retain native structure. (**A**) A ^{32}P-labeled duplex oligonucleotide containing the xl-mtTFA promoter activation site was incubated with xl-mtTFA. The sample in lane 2 was treated with 0.1% glutaraldehyde, while the sample in lane 1 was not. Both were run on a native polyacrylamide gel, as described in Materials and methods. Positions of free and complexed DNA are indicated on the left. (**B**) Western blot analysis of protein–DNA complexes excised from the gel shown in (A) after electrophoresis on an SDS–polyacrylamide gel. Positions of molecular weight markers (kDa) are indicated on the right.

The native structure of the protein–DNA complex is preserved in fixed complexes

Treatment of protein–DNA complexes with glutaraldehyde is frequently necessary to preserve them during harsh preparation procedures for EM (Mastrangelo *et al.*, 1989, 1991, 1993; Stenger *et al.*, 1994). However, it is conceivable that fixed complexes differ from the native ones. To address this possibility, we compared structures of xl-mtTFA–DNA complexes in the presence and absence of glutaraldehyde using a gel mobility shift assay.

A ^{32}P-labeled 42-bp duplex oligonucleotide containing the promoter activation site was incubated with xl-mtTFA in two identical reactions. Glutaraldehyde was added to one reaction at a final concentration of 0.1%. Both binding reactions were subjected to electrophoresis on a native polyacrylamide gel. Free DNA and DNA–protein complexes were detected by autoradiography (Figure 6A). The electrophoretic mobility of protein–DNA complexes was not altered by glutaraldehyde treatment. This strongly argues that the structure of the complexes is not perturbed by the fixation procedure. The bands that correspond to the protein–DNA complexes were excised, and their protein content was analyzed by SDS gel electrophoresis followed by Western blotting, as described in Materials and methods. The complexes formed in the presence of glutaraldehyde contain cross-linked xl-mtTFA species consistent with dimers and tetramers of xl-mtTFA (Figure 6B). Given the preponderance of tetramers observed by STEM at the same binding site, we conclude that xl-mtTFA binds to this site as a tetramer which is incompletely cross-linked by glutaraldehyde. However, we cannot rule out the possibility that the tetramer has a substructure consisting of two dimers of xl-mtTFA.

Discussion

We used three different approaches to resolve the quaternary structure of xl-mtTFA. First, immunoprecipitation experiments with two differentially tagged xl-mtTFA variants showed that the protein molecules interacted with each other in solution under physiological conditions. Second, chemical cross-linking with various reagents identified multimeric forms of xl-mtTFA, with a significant fraction of cross-linked species having a gel mobility expected for a tetramer. Third, STEM mass measurements demonstrated that the protein binds to the promoter activation site predominantly as a tetramer. We propose that binding of the xl-mtTFA tetramer to the promoter activation site presents identical surfaces for interactions with the basal mitochondrial transcription machinery on both sides. Bidirectional transcription activation would be provided equally well by either a stable xl-mtTFA tetramer or a combination of two dimers. It is likely that bending of DNA induced by binding of xl-mtTFA tetramers may also contribute to promoter activation.

Structure of mtDNA nucleoprotein complexes

The unexpected finding that native mtTFA binds DNA as a tetramer has important implications for the structure of mtDNA nucleoprotein complexes. We reported previously that *X.laevis* ovary mitochondria contain ~200 molecules of xl-mtTFA per molecule of mtDNA. This corresponds to ~90 bp of DNA per mtTFA monomer. If mtTFA bound as a monomer with ~35 bp for each binding site, more than one-third of mtDNA would be covered by mtTFA. The observation that mtTFA actually binds as a tetramer permits a more accurate calculation of the fraction of mtDNA covered by mtTFA. With the tetramer model, 200 molecules of mtTFA would bind as ~50 complexes, each occupying ~35 bp. This arrangement leaves a much larger fraction of mtDNA free of mtTFA. It is interesting to note that this number of mtTFA complexes agrees well with the results of EM studies of mtDNA nucleoprotein complexes isolated from *Xenopus* oocytes, in which an average of 48 globular particles interconnected by bare DNA were found per mtDNA molecule (Pinon *et al.*, 1978). It remains to be established whether these complexes represent mtTFA tetramers. It is possible that vertebrate mitochondria contain other DNA binding proteins similar in abundance to mtTFA, as has been reported for yeast mitochondria (Newman *et al.*, 1996).

Binding of mtTFA may be important for specific structuring of regulatory regions of mtDNA. A 500 bp segment of the D-loop region, containing major promoters for transcription of both light and heavy mtDNA strands, as well as the heavy strand origin of replication and conserved sequence blocks 2 and 3, has at least six binding sites for xl-mtTFA (Antoshechkin and Bogenhagen, 1995). Two of these documented binding sites provide DNase I footprints twice as large as the ~35 bp footprint of a single xl-mtTFA tetramer and presumably result from binding of two tetramers. The clustering of as many as eight xl-mtTFA binding sites in the control region suggests that xl-mtTFA may regulate other processes involved in expression and replication of the mitochondrial genome. This is consistent with the observation that deletion of the mtTFA gene in *S.cerevisiae* causes loss of mtDNA, even though

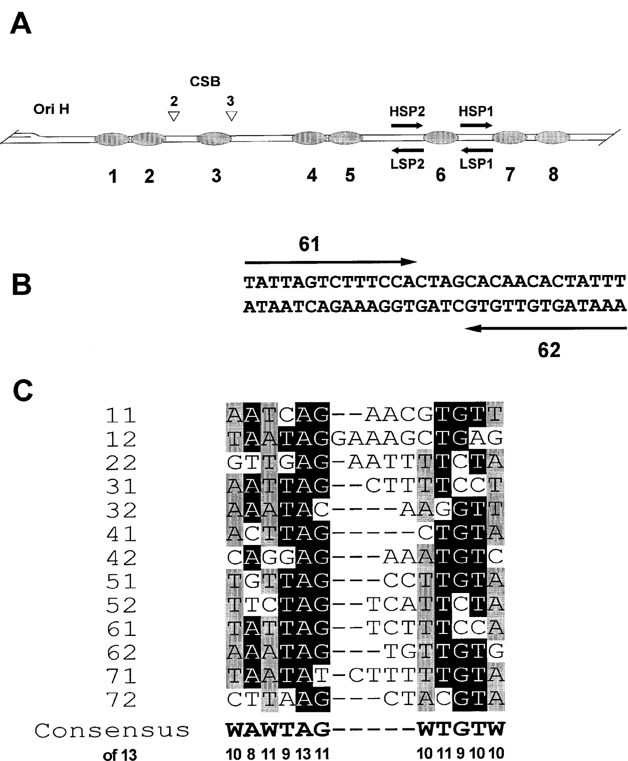

Fig. 7. Analysis of xl-mtTFA binding site sequences. (**A**) Diagram of the control region of *Xenopus* mtDNA. Bidirectional promoters 1 and 2, conserved sequence blocks 2 and 3, and the origin of heavy strand replication are indicated. Binding sites for xl-mtTFA are represented by shaded numbered ovals. (**B**) The dyad orientation of the half-sites is illustrated for site #6, the promoter activation site. (**C**) Alignment of half-sites of xl-mtTFA. Seven of eight mapped binding sites were considered in the analysis. Site #8 and half-site #21 represent poor matches possibly due to co-operative interactions with adjacent sites and are excluded from the alignment. Numbers of the half-sites are shown on the left with the first digit corresponding to the binding site number as defined in (A) followed by the half-site number. Identical residues are printed in white on a black background, similar ones are highlighted with gray. The consensus sequence for the half-site is indicated. W stands for A or T degenerate positions. Numbers on the bottom show how many residues out of 13 match the consensus sequence.

sc-mtTFA does not appear to function as a transcription factor (Diffley and Stillman, 1991).

Dyad symmetry in xl-mtTFA binding sites

Previous analysis of the binding specificity of xl-mtTFA did not identify a clear consensus sequence for protein recognition. Our observation of the tetrameric nature of xl-mtTFA led us to re-examine the structure of its binding sites for the presence of symmetrical sequence motifs. The tetrameric nature of xl-mtTFA suggests that a binding site might be composed of short repetitive sequences which are recognized by individual HMG boxes, or pairs of boxes, arranged in a symmetrical fashion. Since binding of mtTFA results in bidirectional activation of transcription, it is reasonable to expect dyad symmetry within the binding site. Therefore, we divided each previously identified binding site in the control region of mtDNA into two half-sites and attempted to align the opposite strands. The analysis in Figure 7 shows that each half-site consists of a WAWTAG (W = A or T) motif separated by two to three bases (at most, five bases) from a WTGTW motif. The half-sites are separated by approximately three to seven base pairs and are inverted with respect to one another, thus forming a binding site with dyad symmetry. Both motifs are not always present and their combination may differ in any particular binding site, consistent with the hypothesis that each of these elements is recognized separately by an individual HMG box, or pair of boxes, and that not all HMG boxes make specific DNA contacts.

The degenerate A:T or T:A base pairs included as W residues in the motif sequences present similar interacting surfaces in the minor groove of the DNA helix. Methylation interference analysis of the promoter activation site shows that minor groove modification of practically all adenines affects xl-mtTFA binding affinity (I.Antoshechkin, data not shown), as expected since HMG-box proteins are known to interact with the minor groove (van de Wetering and Clevers, 1992; Werner *et al.*, 1995). Further experiments will be required to determine the contribution of individual base pairs and the spatial relationship of motif sequences to the relative strength of interaction of xl-mtTFA with DNA. It is unlikely that all eight HMG boxes present in a tetramer of xl-mtTFA will make similar contacts to the DNA. In this respect, xl-mtTFA may resemble the RNA polymerase I transcription factor UBF, in which some HMG boxes do not appear to contact DNA (Jantzen *et al.*, 1992).

It is interesting to note that one of the proposed recognition sequence elements, WTGTW, is almost identical to the recognition sequences of human Sry, LEF-1 and TCF-1 (van de Wetering and Clevers, 1992; Harley *et al.*, 1994; Love *et al.*, 1995). The presence of a TTG motif has been demonstrated in binding sites of Rox1, Ste11, Sox-5 and some other proteins of this family (Landsman and Bustin, 1993; Grosschedl *et al.*, 1994). Studies of *Drosophila* HMG-D protein argue that easily deformable or stably bent sequences such as a TG dinucleotide embedded in A/T-rich regions constitute a common core element of sequence-specific HMG-box proteins (Churchill *et al.*, 1995). These data argue that at least some proteins with multiple HMG-box domains are capable of sequence-specific binding, and that the mechanism of sequence recognition is similar to that of the single HMG-box-containing proteins. For some proteins, the combined interaction energy of several HMG boxes with non-specific DNA may be sufficient to permit stable binding. For others, a single HMG box may provide a strong enough interaction with specific DNA to permit sequence-specific binding.

Materials and methods

Plasmids and constructs

The DNA fragment used in EM studies was derived from a plasmid generated during deletion analysis of the promoter region of *Xenopus* mtDNA, designated pXLMTΔL895ΔR988 (Antoshechkin and Bogenhagen, 1995). The plasmid was cut with *Afl*III and *Nde*I and the 682-bp fragment containing the mtDNA sequences flanked by vector sequences was fractionated by PAGE and recovered by electroelution on an IBI model UEA electrophoresis apparatus. Two complementary 42mer oligonucleotides that contain the promoter activation site sequence (the xl-MTPR fragment) were synthesized using a Beckman automatic DNA synthesizer, gel purified, and annealed for use in cross-linking experiments. The sequence of the top strand oligonucleotide is 5′-GATCAAAACTATATTAGTCTTTCCACTAGCACAACACTATTT-3′.

Chemical cross-linking

Approximately 13 pmol of recombinant xl-mtTFA purified to homogeneity was incubated with 100 to 200 μM DSP or 0.1% glutaraldehyde in 25 μl reactions containing 40 mM HEPES pH 8.0, 50 mM KCl and 0.02% Triton X-100 for 30 min at room temperature. Cross-linking reactions were stopped by adding Tris pH 8.0 to a concentration of 100 mM. 60 μl of sample loading buffer lacking DTT was added. The reactions were boiled for 4 min and loaded onto a 10% SDS–polyacrylamide gel. Protein was detected by Western blotting using rabbit polyclonal antibody against recombinant xl-mtTFA and an alkaline phosphatase-conjugated goat anti-rabbit antibody.

Immunoprecipitation

Xl-mtTFA cDNA lacking the N-terminal mitochondrial targeting sequence was cloned into T7-based expression vectors pET-21a(+) and pET-29a(+) (Novagen) to produce protein fused at the N-terminus to either T7-Tag or S-Tag peptides respectively. *E.coli* BL21(DE3), which carries a chromosomal copy of the T7 RNA polymerase gene under the control of the lac promoter, was transformed with two constructs simultaneously, and cells resistant to both ampicillin and kanamycin were selected. Cells from a single colony were grown in culture at 37°C until OD_{600} reached ~0.4 when 1 mM IPTG was added to induce protein expression. After 4 h of incubation at 37°C bacteria were lysed by sonication in lysis buffer (50 mM HEPES pH 7.5, 200 mM KCl, 5 mM EDTA, 2 mM DTT, 0.5 mM benzamidine–HCl, 1 μM pepstatin, 0.2 mM PMSF, 50 μg/ml leupeptin). Protein complexes were precipitated from 1 ml of lysates with the monoclonal antibody directed against the T7-Tag (Novagen) and protein A-conjugated Sepharose beads (Pharmacia) as described (Harlow and Lane, 1988). The precipitated proteins were run on an SDS–polyacrylamide gel, transferred onto a PVDF membrane (Immobilon P, Millipore), and S-Tag fusion proteins were detected with an alkaline phosphatase-conjugated S-protein as suggested by Novagen using BCIP/NBT (5-bromo-4-chloro-3-indolyl-phosphate/nitroblue tetrazolium) colorimetric substrates (Kirkegaard and Perry Laboratories).

DNA binding reactions

Typical DNA binding reactions contained 40 ng of the 682-bp *Afl*III–*Nde*I DNA fragment described above and 25–30 ng of recombinant xl-mtTFA in a 20 μl volume containing 40 mM HEPES pH 8, 50 mM KCl, 10 mM $MgCl_2$, 1 mM DTT. For applications other than EM, the binding buffer included 10% glycerol and 0.02% Triton X-100. Reactions were incubated at room temperature for 10 min before glutaraldehyde was added to a final concentration of 0.1% for 2 min of continued incubation.

TEM and STEM analysis

Specimens for both types of EM were prepared and scanned as described (Mastrangelo *et al.*, 1989, 1991; Stenger *et al.*, 1994) with the following modifications: the DNA fragments were electroeluted from acrylamide gels; 0.1% glutaraldehyde was used for stabilization of protein–DNA complexes; the complexes were not column purified but deposited directly onto grids. Photographic negatives obtained in TEM experiments were scanned using a Bio-Rad densitometer to generate files suitable for further analysis. STEM complexes were scanned in 0.5 μm fields. Large angle electron count data from STEM were subjected to local weighted averaging and interpolation prior to analysis. Data analysis was performed using SigmaScan Pro software (Jandel Scientific). STEM data display and analysis procedures are available upon request.

Fixation control

8 pmol of ^{32}P-labeled 42mer duplex oligonucleotide xl-MTPR was incubated with purified recombinant xl-mtTFA in a 20 μl reaction containing 40 mM HEPES pH 8.0, 10 mM $MgCl_2$, 50 mM KCl, 10% glycerol and 0.02% Triton X-100 for 15 min at 20°C. Glutaraldehyde was then added to a final concentration of 0.1% where appropriate and the incubation was continued for 10 min. The reactions were then run on a native polyacrylamide gel in 1× Tris-borate buffer supplemented with 0.01% Triton X-100. The gel was exposed to Kodak XAR-5 film. Bands corresponding to protein–DNA complexes were excised using the autoradiogram as a guide. Gel fragments were incubated in 1× SDS loading buffer for 30 min at 20°C followed by boiling for 2 min. Protein content of the complexes was analyzed by SDS gel electrophoresis by placing the gel pieces directly into wells of a 10% SDS–polyacrylamide gel. Proteins were detected by Western blotting using polyclonal antibodies against recombinant xl-mtTFA.

Acknowledgements

We thank Dr Martha Simon for assistance with use of the STEM and Dr Peter Tegtmeyer for helpful comments on the manuscript. This research was supported by grant NIHGM29681 (D.F.B.) and by the Department of Energy, Office of Health and Environmental Research (I.A.M.). The Brookhaven STEM is an NIH Biotechnology Resource, RR0177.

References

Alexandrova,E.A. and Beltchev,B.G. (1987) Differences between HMG1 proteins isolated from normal and tumor cells. *Biochim. Biophys. Acta*, **915**, 399–405.

Antoshechkin,I. and Bogenhagen,D.F. (1995) Distinct roles for two purified factors in transcription of *Xenopus* mitochondrial DNA. *Mol. Cell. Biol.*, **15**, 7032–7042.

Baxevanis,A.D. and Landsman,D. (1995) The HMG-1 box protein family: classification and functional relationships. *Nucleic Acids Res.*, **23**, 1604–1613.

Blackwell,L.J., Borowiec,J.A. and Mastrangelo,I.A. (1996) Single-stranded DNA binding alters human replication protein A structure and facilitates interaction with DNA-dependent protein kinase. *Mol. Cell. Biol.*, **16**, 4798–4807.

Bogenhagen,D.F. (1996) Interaction of mtTFB and mtRNA polymerase at core promoters for transcription of *Xenopus laevis* mtDNA. *J. Biol. Chem.*, **271**, 12036–12041.

Bogenhagen,D.F. and Insdorf,N.F. (1988) Purification of *Xenopus laevis* mitochondrial RNA polymerase and identification of a dissociable factor required for specific transcription. *Mol. Cell. Biol.*, **8**, 2910–2916.

Bogenhagen,D.F. and Romanelli,M.F. (1988) Template sequences required for transcription of *Xenopus laevis* mitochondrial DNA from two bidirectional promoters. *Mol. Cell. Biol.*, **8**, 2917–2924.

Churchill,M.E., Jones,D.N., Glaser,T., Hefner,H., Searles,M.A. and Travers,A.A. (1995) HMG-D is an architecture-specific protein that preferentially binds to DNA containing the dinucleotide TG. *EMBO J.*, **14**, 1264–1275.

Dairaghi,D., Shadel,G. and Clayton,D. (1995) Addition of a 29 residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator. *J. Mol. Biol.*, **249**, 11–28.

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. USA*, **88**, 7864–7868.

Diffley,J.F. and Stillman,B. (1992) DNA binding properties of an HMG1-related protein from yeast mitochondria. *J. Biol. Chem.*, **267**, 3368–3374.

Fisher,R.P. and Clayton,D.A. (1988) Purification and characterization of human mitochondrial transcription factor 1. *Mol. Cell. Biol.*, **8**, 3496–3509.

Fisher,R.P., Topper,J.N. and Clayton,D.A. (1987) Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. *Cell*, **50**, 247–258.

Fisher,R.P., Lisowsky,T., Parisi,M.A. and Clayton,D.A. (1992) DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. *J. Biol. Chem.*, **267**, 3358–3367.

Grosschedl,R., Giese,K. and Pagel,J. (1994) HMG domain proteins; architectural elements in the assembly of nucleoprotein structures. *Trends Genet.*, **10**, 94–99.

Haqq,C.M., King,C.Y., Ukiyama,E., Falsafi,S., Haqq,T.N., Donahoe,P.K. and Weiss,M.A. (1994) Molecular basis of mammalian sexual determination: activation of Mullerian inhibiting substance gene expression by SRY. *Science*, **266**, 1494–1500.

Harley,V.R., Lovell-Badge,R. and Goodfellow,P.N. (1994) Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res.*, **22**, 1500–1501.

Harlow,E. and Lane,D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Jantzen,H.M., Chow,A.M., King,D.S. and Tjian,R. (1992) Multiple domains of the RNA polymerase I activator hUBF interact with the TATA-binding protein complex hSL1 to mediate transcription. *Genes Dev.*, **6**, 1950–1963.

Kim,J.S. and Raines,R.T. (1993) Ribonuclease S-peptide as a carrier in fusion proteins. *Protein Sci.*, **2**, 348–356.

Landsman,D. and Bustin,M. (1993) A signature for the HMG-1 box DNA-binding proteins. *BioEssays*, **15**, 539–546.

Love,J.J., Li,X., Case,D.A., Giese,K., Grosschedl,R. and Wright,P.E. (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature*, **376**, 791–795.

Mastrangelo,I.A., Hough,P.V., Wall,J.S., Dodson,M., Dean,F.B. and Hurwitz,J. (1989) ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. *Nature*, **338**, 658–662.

Mastrangelo,I.A., Courey,A.J., Wall,J.S., Jackson,S.P. and Hough,P.V. (1991) DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement. *Proc. Natl Acad. Sci. USA*, **88**, 5670–5674.

Mastrangelo,I.A., Held,P.G., Dailey,L., Wall,J.S., Hough,P.V., Heintz,N. and Heintz,N.H. (1993) RIP60 dimers and multiples of dimers assemble link structures at an origin of bidirectional replication in the dihydrofolate reductase amplicon of Chinese hamster ovary cells. *J. Mol. Biol.*, **232**, 766–778.

Mignotte,B. and Barat,M. (1986) Characterization of a *Xenopus laevis* mitochondrial protein with a high affinity for supercoiled DNA. *Nucleic Acids Res.*, **14**, 5969–5980.

Mikhailov,V.S. and Bogenhagen,D.F. (1996) Effects of *Xenopus laevis* mitochondrial single-stranded DNA-binding protein on primer-template binding and 3′→5′ exonuclease activity of DNA polymerase γ. *J. Biol. Chem.*, **271**, 18939–18946.

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.

Park,L.S., Friend,D., Gillis,S. and Urdal,D.L. (1986) Characterization of the cell surface receptor for a multi-lineage colony-stimulating factor (CSF-2 alpha). *J. Biol. Chem.*, **261**, 205–210.

Peters,R., King,C.Y., Ukiyama,E., Falsafi,S., Donahoe,P.K. and Weiss,M.A. (1995) An SRY mutation causing human sex reversal resolves a general mechanism of structure-specific DNA recognition: application to the four-way DNA junction. *Biochemistry*, **34**, 4569–4576.

Pinon,H., Barat,M., Tourte,M., Dufresne,C. and Mounolou,J. (1978) Evidence for a mitochondrial chromosome in *Xenopus laevis* oocytes. *Chromosoma*, **65**, 383–389.

Stenger,J.E., Tegtmeyer,P., Mayr,G.A., Reed,M., Wang,Y., Wang,P., Hough,P.V. and Mastrangelo,I.A. (1994) p53 oligomerization and DNA looping are linked with transcriptional activation. *EMBO J.*, **13**, 6011–6020.

Tarvers,R.C., Noyes,C.M., Roberts,H.R. and Lundblad,R.L. (1982) Influence of metal ions on prothrombin self-association. Demonstration of dimer formation by intermolecular cross-linking with dithiobis(succinimidylpropionate). *J. Biol. Chem.*, **257**, 10708–10714.

Teo,S.H., Grasser,K.D., Hardman,C.H., Broadhurst,R.W., Laue,E.D. and Thomas,J.O. (1995) Two mutations in the HMG-box with very different structural consequences provide insights into the nature of binding to four-way junction DNA. *EMBO J.*, **14**, 3844–3853.

van de Wetering,M. and Clevers,H. (1992) Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson–Crick double helix. *EMBO J.*, **11**, 3039–3044.

Wall,J.S. (1979) Biological scanning transmission electron microscopy. In Hren,J., Goldstein,J. and Jay,D. (eds), *Analytical Electron Microscopy.* Plenum Press, New York, pp. 333–342.

Werner,M.H., Huth,J.R., Gronenborn,A.M. and Clore,G.M. (1995) Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell*, **81**, 705–714.

Xu,B. and Clayton,D.A. (1992) Assignment of a yeast protein necessary for mitochondrial transcription initiation. *Nucleic Acids Res.*, **20**, 1053–1059.

Received on December 12, 1996; revised on January 31, 1997