Trypanosome nuclear factors which bind to internal promoter elements of tRNA genes

Stephen D. Bell* and J. David Barry

Wellcome Unit of Molecular Parasitology, Anderson College, 56 Dumbarton Road, Glasgow G11 6NU, UK

Received July 3, 1995; Accepted July 18, 1995

ABSTRACT

Expression of eukaryotic nuclear encoded tRNA genes requires two transcription factors, TFIIB and TFIIC. To determine if the highly evolutionarily diverged parasitic protozoan *Trypanosoma brucei* possesses any analogous factors, trypanosome nuclear extracts were prepared. Using these extracts in gel-retardation assays on trypanosome tRNA genes we detected three specific protein–DNA complexes, a highly retarded species and a less retarded pair of complexes. Introduction of mutations into the B box of a lysyl tRNA gene greatly reduced formation of all three complexes. Footprinting studies indicated that all complexes protected the B box of the tRNA gene from cleavage. The most highly retarded complex protected both the A and B boxes from DNaseI cleavage. The less retarded pair of complexes showed footprints on the B box alone and the lowest B box specific complex is shown to be resistant to the polyanion heparin. All three complexes are demonstrated to induce bends in the DNA on binding.

INTRODUCTION

The initiation of transcription of eukaryotic nuclear encoded tRNA genes by RNA polymerase III (RNAP III) depends on *cis* and *trans* acting elements (reviewed in ref. 1). In all examples known thus far tRNA gene transcription requires two internal promoter elements, the A and B boxes. In a variety of organisms these elements have been demonstrated to bind a complex transcription factor, TFIIC. In yeast TFIIC is a protein complex of >600 kDa comprising six polypeptides. Different constituent protein subunits of TFIIC mediate binding to the A and B boxes; indeed, limited proteolysis of the yeast TFIIC liberates a sub-factor which recognises the B box only (2). No independent binding of any sub-factor to the A box alone has been observed. These data, combined with analysis of mutations in both A and B boxes (3), have led to the proposal that the A and B boxes play distinct roles, the B box being the primary contact point of TFIIC while the A box plays a secondary role in DNA binding of TFIIC, but defines the subsequent positioning of TFIIB, RNAP III and hence the transcription initiation site (4). In contrast, the human TFIIC appears to be comprised of two components which are

separable by B box DNA affinity chromatography. Together these components, TFIIC1 and TFIIC2, bind to A and B boxes and TFIIC2 alone can bind to the B box, but neither TFIIC1 nor TFIIC2 alone is capable of binding to the A box (5). However, different methods of preparation of this factor have yielded contradictory results, possibly because the two activities have co-fractionated during preparation (6–8). In addition to the A and B boxes, in a number of cases sequences to the 5' side of the transcription initiation site have been demonstrated to modulate the levels of expression of tRNA genes (9 and references therein).

Recent work in both yeast and human systems has led to the identification and cloning of genes encoding various TFIIC subunits (10–13). The genes encoding the principal B box binding protein have been cloned from both species but demonstrate little sequence homology. The mode of interaction of these proteins with the DNA substrate is unclear, although it is known that the yeast protein contains two potential high mobility group (HMG) domains. However, these regions are not well conserved in the human protein (10,11). Furthermore, recombinant yeast B box binding subunit protein expressed in *Escherichia coli* does not bind to DNA on its own, suggesting other components of TFIIC may be required for this interaction (10).

The eventual recruitment of RNAP III to the transcription initiation site of yeast tRNA genes has been shown to require a sequential assembly of transcription factors on the promoter (14). First, TFIIC binds to the A and B boxes, with the latter playing a dominant role. This facilitates binding of TFIIB to regions upstream of the initiation site. TFIIB, in turn, then recruits RNAP III to the initiation site. An interesting feature of the yeast system is that TFIIB incorporation is absolutely dependent on the prior assembly of TFIIC. However, once TFIIB is bound TFIIC can be stripped from the complex by treatment with high salt or polyanions such as heparin, leaving TFIIB stably bound (14).

Very little is known about the processes of transcription initiation in trypanosomes. The only promoters that have been characterised are those of the rRNA (15), variant surface glycoprotein (VSG) (15,16), procyclic acidic repetitive protein (PARP) (17,18) and U2 snRNA (19,20) genes and these demonstrate a number of features possibly peculiar to this highly diverged organism. The majority of genes are transcribed in polycistronic units, mRNAs have a common 5' end added by trans-splicing and certain mRNAs (encoding VSG and PARP) are thought to be transcribed by RNAP I (21). All trypanosome U-snRNAs are transcribed by RNAP III and a recent report

* To whom correspondence should be addressed

Figure 1. Structure of 7SL locus. The structure of the 7SL locus is diagrammed. Regions coding the 7SL, tRNA^{lys}, tRNA^{arg} and U3 snRNA genes are shown as shaded boxes. The orientation of transcription of each gene is shown by an arrow beneath the genes name. The spacing between genes is also shown. In addition the regions of the locus present in plasmids pUP-1 and pL3 are indicated by horizontal bars.

indicates that U2 snRNA transcription requires sequences upstream of the gene that have a high degree of similarity to A and B boxes (20).

We have previously described the characterisation of two loci encoding very tightly linked genes likely to be RNAP III transcription units (22). One locus encodes the 7SL RNA, tRNA^{lys}, tRNA^{arg} and U3 snRNA homologue, the other tRNA^{tyr}, tRNA^{thr} and U6 snRNA. The genes in these loci are organised in alternating orientation of transcription. The spacing between the genes is highly conserved, in the 7SL locus being 95 nt between the 5' ends of the 7SL and lysyl tRNA genes and between the U3 snRNA and arginyl tRNA genes and 90 nt between the 3' ends of the two tRNA genes (Fig. 1). The tRNA genes both possess classical A and B boxes, elements that have recently been shown to be required for expression of trypanosomatid tRNA genes *in vivo* (23). Intriguingly the A and B boxes of the tRNA genes are also required for expression of the adjacent, divergently transcribed, small RNA genes *in vivo* in transient transfection assays (23). Two models have been proposed to account for the observation. In the first, the A and B boxes may be serving as bi-directional promoter elements, recruiting RNAP III to the intergenic region serving as upstream sequences for both tRNA and the adjacent small RNA gene. A second proposal has been that the tRNA A and B boxes may be playing a role analogous to the extragenic B box of the yeast U6 snRNA gene by facilitating TFIIC mediated derepression of chromatin on the gene (24).

As a first step towards characterisation of RNAP III transcription initiation in *T.brucei* the present work demonstrates the presence of activities in nuclear extracts which bind specifically to the A and B boxes of trypanosome tRNA genes. The binding of one of the activities has unusual heparin resistance, in marked contrast with the observations of the yeast TFIIC heparin sensitivity. Binding of these factors is shown to induce a bend, unusually, at the B box. We discuss the possible significance of this bending to the observed interaction between tRNA genes and neighbouring small RNA genes.

MATERIALS AND METHODS

Trypanosomes

Trypanosoma brucei EATRO 795 bloodstream trypanosomes grown from frozen cloned stabilate in Swiss TO mice and Wistar rats were purified from blood by anion exchange chromatography (25).

Plasmids

The plasmids containing the tRNA genes were pUP-1 and pL-3, as described previously (22).

Oligonucleotide mediated site directed mutagenesis of the lysyl tRNA gene B box was performed essentially as described (26) using an oligonucleotide BBMUT2 5'-CCCCCGTGGGGCTC-CATCCACGACCACACG-3'. The mutations were confirmed by DNA sequencing and the plasmid named pUP1-bb⁻.

For circular permutation analysis of protein-induced DNA bending the vector pBend-2 was used (27). Oligonucleotides LYSBBU (5'-TCTTAACCGTGTGGTCGTGGGTTTCGAGCCC-CACG GGGGGTGCTTTTCTTTT-3' and LYSBBL 5'-AAG-AAAAAAGAAAAGCACCCTGGGGCTCGAACC CACGACCACGGTT-3') corresponding to the region protected from DNaseI digestion (Fig. 5) were synthesised using an Applied Biosystems PCR-mate, annealed and ligated to blunt ended *Xba*I digested pBend-2. The resultant plasmid was named pBend-lys.

DNA probes

The arginyl tRNA gene encoding probe used in Figure 2A was generated by *Xba*I *Xho*I double digestion of pL3. The 300 nt lysyl tRNA gene encoding fragment was excised from pUP-1 by digestion with *Kpn*I and *Eco*RI for top strand labelling or *Xho*I and *Sst*I for bottom strand labelling. The 300 nt fragments were gel purified and radiolabelled using the Klenow fragment of *E.coli* DNA polymerase I (Stratagene) and 1 μ l (10 μ Ci) of the appropriate [α -³²P]dNTP (3000 Ci/mmol). Radiolabelling of blunt ended fragments of pBend-lys was performed by incubating the products of the appropriate restriction digest with Klenow fragment for 10 min at 37°C prior to addition of the dNTP mix.

Nuclear extract preparation

Female Wistar rats were infected with ILTat 1.2 *T.brucei* and sacrificed by cardiac puncture at parasitaemia of 10⁹ trypanosomes/ml. Trypanosomes were purified from blood cells by DE52 anion exchange chromatography. The nuclear extract was prepared essentially as described (28). Trypanosomes were pelleted by centrifugation (1600 g) and washed twice in PBS. Following the final wash trypanosomes were resuspended in two packed cell volumes of Buffer A. Cells were then lysed by 25 strokes of a Dounce homogeniser (A type pestle). Nuclei were pelleted for 5 min at 3700 g in a Beckman JS-7.5 rotor, resuspended in Buffer C at 10¹⁰ cells/ml using a Dounce homogeniser (A type pestle) and mixed gently at 4°C for 30 min. The resulting nuclear lysate was centrifuged for 30 min at 4°C, 25 000 g (Beckman JA21 rotor) and the supernatant dialyzed against 50 vol of Buffer D for 5 h at 4°C. The dialysate was centrifuged at 25 000 g for 25 min at 4°C (Beckman JA21 rotor) and the supernatant snap frozen in liquid nitrogen as 100 μ l aliquots. Extracts were stored at -80°C.

Buffers

(A) 10 mM Hepes (pH 7.9 at 4°C) 10 mM KCl, 0.5 mM DTT. (C) 20 mM Hepes (pH 7.9), 25% glycerol, 0.42 M NaCl, 0.2 mM EDTA, 0.5 mM DTT. (D) 20 mM Hepes (pH 7.9), 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT. All buffers contained proteinase inhibitors: 0.5 mM PMSF, 10 μ g/ml leupeptin, 10 μ g/ml aprotinin, 10 μ g/ml pepstatin A.

Figure 2. Sequence specific protein–DNA interactions on tRNA genes. (A) Gel retardation assays were performed using 5 fmol of ^{32}P -labelled *Xba*I–*Xho*I fragment of pL-3, containing the arginyl tRNA gene and various competitors. Reactions contain 8 μg crude nuclear extract. Binding reactions were performed as described in Materials and Methods and electrophoresed on a 5% polyacrylamide gel. Lanes 1–4 contain 0, 1, 10 and 100 $\mu\text{g}/\text{ml}$ poly (dI.dC). All subsequent lanes contain 10 $\mu\text{g}/\text{ml}$ poly (dI.dC). In addition lanes 5–8 contain 0, 15, 150, 1500 fmol pL-3 (arginyl tRNA gene); lanes 9–12 contain 0, 15, 150, 1500 fmol pUP-1 (lysyl tRNA gene) and lanes 13–16 contain 0, 10, 100, 1000 fmol pBluescript. Arrows indicate unbound DNA (F) and protein–DNA complexes (cIa, cIb and cII). (B) Reactions as described for A but using a ^{32}P radiolabelled 300 nt *Kpn*I–*Eco*RI tRNA lys encoding fragment of pUP-1 as probe with cold competitors as below. Lanes 1–5: 0, 6, 13, 25 and 50 fmol pUP1-bb $^{-}$ as competitor. Lanes 6–10: 0, 6, 13, 25 and 50 fmol pUP1 as competitor.

Fractionation of extract

Crude nuclear extract was fractionated using phosphocellulose P11 affinity chromatography. Twenty milligrams of nuclear extract were applied to a 20 ml phosphocellulose P11 column at room temperature in Buffer A (20 mM Hepes pH 8, 10% glycerol, 2 mM MgCl_2) + 100 mM KCl column at a flow rate of 0.5 ml/min and material eluted at salt steps of Buffer A + 0.3 M KCl and Buffer A + 1 M KCl. Fractions were assayed for DNA binding activity using the gel retardation assay and specific activity was enriched in positive fractions by spin dialysis using Centricon C100 filters. The DNA binding activity was retained by these filters.

Gel retardation assays

Apart from the exceptions noted in the figure legends all reactions were carried out under the following conditions. Varying amounts of either crude nuclear extract or phosphocellulose P11 chromatography fraction were incubated in 20 mM Hepes pH 7.9, 100 mM KCl, 5 mM MgCl_2 , 10% glycerol, 10 $\mu\text{g}/\text{ml}$ poly dI.dC (Boehringer Mannheim) with 3 fmol radiolabelled probe for 20 min at room temperature. Reactions were loaded onto a 5% polyacrylamide gel in 1 $\times$ Tris–borate EDTA (pH 8.3) and electrophoresed for 2.5 h at 200 V. Gels were then dried at 80°C under vacuum and autoradiographed at -70°C .

1-10 Copper phenanthroline footprinting

In-gel copper phenanthroline (Cu-OP) footprinting was performed essentially as described (29). Standard gel retardation conditions were scaled up 10-fold and run over three lanes of a standard retardation gel. Following electrophoresis the gel was soaked in 400 ml of 20 mM Tris–HCl pH 8 for 10 min at room temperature. Fresh Cu-OP was prepared by mixing 2 ml 9 mM

CuSO_4 and 2 ml 40 mM 1,10 phenanthroline and allowing the mix to stand for 1 min at room temperature prior to the addition of 36 ml of water. This Cu-OP solution was then added to the equilibrated gel followed by addition of 40 ml of 0.5% mercapto-propionic acid. Cleavage reactions were allowed to proceed for 10 min and terminated by the addition of 20 ml 1.2% 2,9-dimethyl–1-10 phenanthroline. Regions of gel containing free and complexed DNA were detected by autoradiography of the wet gel and gel slices containing the appropriate species excised. DNA was recovered from the gel, ethanol precipitated and resuspended in 90% formamide, 1 $\times$ TBE. The yield of resuspended material was determined by Cerenkov scintillation counting and 2500 c.p.m. of each sample electrophoresed on a 6% sequencing gel.

DNaseI footprinting

Binding conditions were as described for the gel retardation assays except reactions were scaled up 2.5-fold, to a volume of 50 μl . Binding reactions were allowed to proceed for 20 min prior to the addition of 50 μl of 10 mM MgCl_2 , 5 mM CaCl_2 containing 12 ng/ml DNaseI. In the experiment shown in Figure 3B DNaseI digestion was terminated after 2 min by addition of EDTA to 20 mM. Protein–DNA complexes were separated on a 6% non-denaturing polyacrylamide gel. Complexes were visualised by autoradiography and recovered from the gel as for Cu-OP footprinting described above with the exception that 1000 c.p.m. of each sample was loaded on a 5% sequencing gel. In the experiment shown in Figure 5, after a 2 min incubation at room temperature cleavage reactions were terminated by the addition of 100 μl of 0.2 M NaCl, 20 mM EDTA, 1% SDS and 300 $\mu\text{g}/\text{ml}$ *E.coli* tRNA as carrier. Nucleic acids were recovered by phenol–chloroform extraction and ethanol precipitation, washed twice in 70% ethanol, dried and resuspended and treated as

Figure 3. DNaseI and Copper 1-10 phenanthroline footprinting of protein-DNA complexes. (A) Copper 1-10 phenanthroline footprints of individual protein-DNA complexes. Cu-OP DNA cleavage reactions were performed on separated protein-DNA complexes as described in Materials and Methods. Footprinting was carried out for the bottom, coding strand of the lysyl tRNA gene using ^{32}P radiolabelled 300 nt *SstI-XhoI* fragment of pUP-1. Lane 1, Maxam-Gilbert A+G sequence ladders, lane 2, Cu-OP cleavage of unbound DNA, lane 3, Cu-OP cleavage of complexes Ia and Ib and lane 4, Cu-OP cleavage pattern of complex II. Regions showing protection from cleavage are indicated by vertical bars. (B) DNaseI footprints of Complex Ia/b and Complex II. DNaseI footprinting of the top, non-coding, DNA strand of the lysyl tRNA gene was performed as detailed in Methods. A ^{32}P -radiolabelled 300 nt *KpnI-EcoRI* tRNA^{Lys} encoding fragment of pUP-1 as probe and complexes separated on a 5% non-denaturing polyacrylamide gel. Unbound DNA and DNA associated in protein-DNA complexes was recovered from the gel and electrophoresed on a 6% sequencing gel alongside a Maxam Gilbert A+G sequence ladder for reference. Lane 1, Maxam Gilbert A+G ladder, lane 2, Unbound DNA, lane 3, DNA associated with complexes Ia and Ib, lane 4, DNA associated with complex II. Regions showing protection from cleavage are indicated by vertical bars and protein induced DNaseI hypersensitivity indicated by an arrowhead.

described for Cu-OP footprinting above, with the exception that 15 000 c.p.m. were loaded per lane of a 6% sequencing gel.

Maxam-Gilbert sequencing of DNA

The Maxam-Gilbert (30) A+G sequencing reaction was carried out on Hybond M-G paper (Amersham) following the manufacturer's instructions.

RESULTS

Detection of DNA binding activities present in *T.brucei* nuclear extracts

In order to determine if there are any factors present within trypanosome nuclei which bind specifically to tRNA genes, nuclear extract was prepared as described in Materials and Methods. This crude nuclear extract was used in gel retardation assays using radiolabelled tRNA genes as DNA probes. The results shown in

Figure 2A demonstrate the presence of specific DNA binding activities. In the presence of the non-specific competitor, poly dI.dC, three complexes are detected: a closely migrating doublet (complexes Ia and Ib) and a more highly retarded species (complex II). The formation of these complexes can be abolished by addition of SDS to 0.1% prior to electrophoresis (data not shown). In the experiment shown in Figure 2A, radiolabelled arginyl tRNA gene was used as a probe in a series of competition assays, to determine the sequence specificity of DNA binding. Varying amounts of unlabelled supercoiled plasmid, either pL-3 (arginyl tRNA), pUP-1 (lysyl tRNA) or pBluescript KS+ were included in the binding reactions as competitor. The results indicate that both tRNA genes compete for binding to the labelled probe at between 3- and 30-fold molar excess, however pBluescript competes only partially, at 300 molar excess. It appears, therefore, that sequence-specific DNA binding activities are being detected, presumably directed against sequences conserved between the two tRNA genes. The likeliest candidates for these conserved regions are the A and B boxes.

Factor binding *in vitro* is reduced by mutations in the B box

Site directed mutagenesis was used to create a mutant tRNA lys gene with an altered B box. The universally conserved T54 and C56 were mutated to A and G respectively. These B box mutations correspond to mutations known to reduce expression of yeast tRNAs by at least 50-fold (3).

Figure 2B shows the results of gel retardation assays performed using both wild type and mutant tRNA genes as competitors for factor binding to the wild type gene. These results demonstrate that these mutations in the B box reduce binding *in vitro* by at least 10-fold.

Copper phenanthroline and DNaseI footprinting assay

Copper-phenanthroline footprinting was performed to determine which sequences within the tRNA genes were responsible for mediating formation of the protein-DNA complexes described above. This technique couples the simplicity of the gel retardation assay with the detailed sequence information available from footprinting, allowing footprinting of individual retarded complexes within the gel matrix (29). Although this was straightforward for the highly retarded complex, the extremely close migration of the doublet complex prevented separation of these two complexes, hence the Cu-OP footprint (Fig. 3A, lane 3) is a composite of the footprints of these two species. The Cu-OP footprints of both upper and lower complexes indicate protein binding to regions encompassing the B box and T-rich putative transcription terminator immediately downstream of the tRNA gene. The region of altered Cu-OP cleavage in the upper complex (Fig. 3A, lane 4) extends towards the 5' end of the coding region with weak protection observed over the A box region of the tRNA gene. In addition to this A box protection, a region of enhanced sensitivity to Cu-OP cleavage is observed between the A and B boxes in complex cII.

The data derived from the Cu-OP footprinting are confirmed by DNaseI footprinting assays. Protein-DNA complexes were allowed to form under standard binding conditions, then following a 15 min binding reaction DNaseI was added. After 2 min, DNaseI digestion was terminated by addition of EDTA and complexes were separated on a 5% non-denaturing polyacryl-

Figure 4. Effect of heparin on complex formation. Standard binding reactions were performed using 8 µg crude nuclear extract and 5 fmol of 32 P-labelled 300 bp *EcoRI-KpnI* fragment of pUP-1 containing the lysyl tRNA gene. The binding reactions were electrophoresed on a 4.5% polyacrylamide gel and complexes visualised by autoradiography. Lane 1 contains a standard reaction with no heparin added. Lanes 2–6 contain reactions to which heparin was added to 300 µg/ml at 0, 7.5, 15, 22.5 and 30 min into a 30 min binding reaction. Lanes 7–10 contain heparin added to 75, 150, 225 and 300 µg/ml at 25 min into a 30 min binding reaction. Arrows indicate unbound DNA (F) and protein–DNA complexes (cla, clb and cII).

amide gel. Following electrophoresis individual complexes were identified by autoradiography and excised from the gel, DNA was then recovered and electrophoresed on a 5% sequencing gel. As for the Cu-OP footprinting, the extremely close migration of the doublet complex prevented separation of these two complexes. The resultant DNaseI cleavage patterns are shown in Fig. 3B. As with the Cu-OP footprinting, both complexes display protection of the B box region of the tRNA gene, however complex II shows considerable protection of the A box region of the gene. The 5' limit of protection coincides with the proposed 5' end of the mature tRNA.

Heparin sensitivity of complex formation

The DNA binding activities of the TFIIC and TFIIB transcription factors of yeast display differential sensitivity towards polyanions such as heparin (14). Yeast TFIIC is displaced from DNA by 300 µg/ml heparin whereas TFIIB remains bound under these conditions. To determine if the DNA binding complexes detected in trypanosome nuclear extracts possess similar properties, heparin was added to binding reactions. Nuclear extract was incubated with probe DNA under standard binding conditions for 30 min at room temperature. In the experiments shown in Figure 4, lanes 1–6, heparin was added to 300 µg/ml into the binding reaction at various time points. In the presence of heparin only two bands are visible, the predominant retarded species co-migrates with the lower complex of the doublet band in the untreated sample (Fig. 4, lane 1), and a second, very faint, complex of higher mobility is also present. It appears, therefore, that the upper complex of the doublet (complex Ib) and the highest complex (complex II) are sensitive to heparin. However, the effect of heparin is dependent on the time at which it is added to the reaction, with addition at an early time-point resulting in a far lower yield of resistant complex than addition at a late time point. This effect could be due to a degree of heparin sensitivity

Figure 5. DNaseI footprint of heparin resistant complex on the lysyl tRNA gene. Lanes 1–3, footprint on coding strand, lanes 4–6 footprint on non-coding strand. Lanes 1+4, Maxam–Gilbert A+G sequence ladder, lanes 2+5, DNaseI cleavage pattern of unbound DNA, lanes 3+6, DNaseI cleavage pattern of DNA incubated with 20 µl P11 fraction. Regions showing protection from cleavage are indicated by vertical bars and protein induced DNaseI hypersensitivity indicated by an arrowhead.

of the 'resistant' complex such that a long exposure to heparin may be sufficient to strip the factor from DNA. Alternatively, the effect of heparin may be to prevent initial recognition of DNA by the factor but not destabilise previously formed complex. To examine this further, heparin was added at varying concentrations at a late time point. The results shown in Figure 4, lanes 7–10, show that the yield of resistant complex is, in fact, dependent on the concentration of heparin added at 25 min into a 30 min binding reaction, suggesting a degree of heparin sensitivity in this complex.

The heparin resistant complex represents a B box mediated protein–DNA interaction

Crude nuclear extract was broadly fractionated using phosphocellulose chromatography. Positive fractions were detected by the gel retardation assay and concentrated using centricon C100 filters. These fractions were enriched for the activity which gives rise to only complex Ia, the heparin resistant lower species in the doublet complex (data not shown).

Positive fractions were used in DNaseI footprinting assays and gave complete protection of ~50 bp encompassing the B box of the lysyl tRNA gene (Fig. 5). In agreement with the DNaseI and Cu-OP footprinting of combined complexes Ia and Ib (Fig. 3A and B, lane 3), the DNaseI footprint extends beyond the 3' end of the coding region of the gene and does not cover the A box. These results suggest considerable deviation from the yeast TFIIC activity which binds to both A and B boxes. [However, it should be noted that limited proteolysis of yeast TFIIC liberates a proteolytic fragment with B box binding activity only (2). The use of a cocktail of protease inhibitors during preparation of the

trypanosome nuclear extract makes it unlikely that proteolysis has occurred in our extract.] Also, yeast TFIIC binding is heparin sensitive whereas one of the trypanosome B box binding activities is heparin resistant.

A feature of the trypanosome factor DNaseI footprint is that there is an increase in sensitivity to DNaseI cleavage towards the 3' boundary of the tRNA gene (Figs 3B and 5). Such a phenomenon is indicative of a protein-induced alteration in the DNA helical conformation, such that the minor groove is locally broadened, thereby facilitating recognition and cleavage by DNaseI. Yeast TFIIC binding induces a similar hypersensitive site but at the other end of the tRNA gene, near the transcription initiation site and has been demonstrated to induce a bend in the DNA at approximately this position (31).

The trypanosome B box binding factor induces a bend in DNA

In order to test the possibility that factor binding may be inducing a bend in the DNA, circular permutation assays were performed. These assays are based on the observation that DNA molecules of identical length but with bends at different positions within the molecule show differential migration on electrophoresis; a molecule with a central bend has a retarded mobility relative to a molecule with a terminal bend (32). Complementary 52 bp oligonucleotides (Materials and Methods) encompassing the observed DNaseI footprint on the lysyl tRNA B box were synthesised, annealed and cloned into unique restriction sites in the centre of the directly repeated polylinker of pBend-2 (27). Restriction sites in the direct repeats were used to generate several circularly permuted fragments of identical length. These fragments were then used in gel retardation assays which show that, while there is no detectable intrinsic bend in the DNA, complex formation does induce a bend in the DNA substrate (Fig. 6). Furthermore, it indicates that these cloned oligonucleotides contain sufficient sequence information to mediate specific binding of all three complexes.

DISCUSSION

We have demonstrated the presence of DNA binding activities in trypanosome nuclear extracts and revealed some of the requirements for their specific binding to tRNA genes. These activities are directed primarily against a conserved internal promoter element in tRNA genes, the B box and all of them are dependent on B boxes being intact. There is a previous report of a trypanosome DNA binding activity directed against a B box-like promoter element which lies 160 nt upstream of the trypanosome U2 snRNA gene, but it gave an extremely weak DNaseI protection of six bases over that core B box homology in this region (19). It is currently unclear which, if any, of the complexes we describe correlate with that U2 extragenic B box-binding activity. The A and B box elements of trypanosome tRNA genes have been demonstrated to be required for trypanosome tRNA gene expression *in vivo* (23). It appears, therefore, that the activities we have described may correspond to the TFIIC factor of other eukaryotes. However, the trypanosome factor has a number of peculiar features. In the crude nuclear extract, interactions with A and B boxes and the B box alone are observed and one complex has unusual heparin resistance. Like yeast TFIIC the trypanosome factor induces a bend on binding, but whereas yeast TFIIC bends DNA near the transcription initiation

Figure 6. Circular permutation analysis of protein-DNA complexes. pBend-lys was constructed as described in Methods. Seven identically sized restriction fragments containing the lysyl tRNA gene B box were generated by cleavage with various restriction enzymes in the directly repeated polylinker flanking the inserted oligonucleotides in this plasmid. These fragments were radiolabelled using Klenow fragment as detailed in Methods. The various fragments were then used in standard gel-retardation assays as described in Methods and electrophoresed on an 8% polyacrylamide gel in $1 \times$ TBE. Lane 1, *MluI* digestion; lane 2, *SpeI* digestion; lane 3, *XhoI* digestion; lane 4, *EcoRV* digestion; lane 5, *SruI* digestion; lane 6, *NruI* digestion; lane 7, *BamHI* digestion. Arrows indicate unbound DNA (F) and protein-DNA complexes (cIa, cIb and cII), the band indicates by VB is the 3000 nt vector backbone of pBend included as a control for electrophoresis.

site (31), the trypanosome factor induces a bend near the transcriptional terminator.

We propose that the three complexes we detect are interrelated. The lowest, heparin resistant, complex Ia may represent a core B box binding activity and the upper complexes may correspond to additional components to this core factor. This is supported by the observed extension of the Cu-OP and DNase footprints of the top complex II to cover the A box. The unique footprint of complex Ib is not known. Presumably, the action of heparin is to disrupt the interaction between the A and B box binding domains, interfering at either the level of accessory protein-core protein or accessory protein-DNA interaction. This effect of heparin may be analogous to the effect of single stranded DNA on the binding of yeast TFIIC. Addition of single stranded DNA displaces yeast TFIIC from DNA in binding reactions, this effect probably being mediated by the A box binding domain (33). Experiments using partially proteolysed TFIIC led to the proposal that the binding of single stranded DNA to the A box binding domain destabilized the interaction of the B box binding domain with B box DNA. It is possible, therefore, that the effect of heparin on binding of the trypanosome factor is to destabilise the A box binding domain. Consequently, because of a relatively weak interaction between the A and B box binding domains of the trypanosome factor the net effect may be to disrupt the interaction of the A and B box binding domains, stripping the A box binding domain from the

factor and leaving the B box binding domain bound. The relevance of these *in vitro* observations to the *in vivo* state is currently unknown. It is unclear whether the various complexes represent different functional stages in the stepwise assembly of a large factor or if they simply reflect relatively weak interactions within a large complex which have been disrupted during preparation of the extract.

Previous work has demonstrated that the A and B boxes of trypanosomatid tRNA genes are required for their expression *in vivo*. Furthermore, these sequences appear to be important for the *in vivo* expression of neighbouring small RNA genes (23). A model has been proposed to explain this observation which invokes bidirectional promoter activities directed by the A and B boxes and proposes a role for factor mediated DNA bending within upstream intergenic regions. This proposal is in part based on the observation that yeast TFIIC induces a bend in DNA near the 5' end of tRNA genes and that yeast TFIIB bends DNA upstream of the gene. However, to date nothing is known about trypanosome factors which mediate tRNA gene expression.

In an attempt to detect factor induced conformational alterations in the trypanosome tRNA^{lys} gene we carried out circular permutation analysis and detected a modulation of the degree of retardation of the protein-DNA complexes. Such an effect can be explained by induction of a static bend on factor binding. However, it is theoretically possible that factor binding could induce a localised flexibility of DNA on binding. Such enhanced bendability of DNA would also cause differential migration on circular permutation. However, a recent study of flexible DNA molecules showed that in circular permutation analysis the gel mobilities of the most and least retarded species of a range of highly flexible molecules vary by <5% (34). In our experiment the reduction in mobility between terminally and centrally bent molecules is in excess of 14%.

The locus of bending correlates with the presence of DNaseI hypersensitive sites in the DNaseI footprint. The fact that sequences on either side of the hypersensitivity are protected from cleavage by both DNaseI and, more importantly, the high resolution footprinting reagent copper 1-10 phenanthroline, suggests that the protein is contacting DNA on both sides of the proposed bending locus. Because of a combination of the large size of the protein factor and the necessarily large length of the insert cloned into pBend-2, the circular permutation assay does not give a high enough resolution to map precisely the bend centre. However, the most highly retarded species has the hypersensitive site in the most central position (Fig. 6, lane 4). If it is assumed that the bend centre is at the hypersensitive site then this may have considerable structural significance in the locus in which the tRNA^{lys} gene is encoded. Our experiments have also shown a corresponding hypersensitive site in the DNaseI footprint on the tRNA^{arg} gene (data not shown). These genes are tightly linked in the genome with conserved length of spacing between genes, indeed the linkage between tRNA genes in this locus is conserved between several trypanosomatid species (22,23). The spacing between the two hypersensitive sites is 95 nt, corresponding to nine helical turns of B-form DNA double helix. This suggests that the sites are in phase, making it probable, therefore, that the bends induced by factor binding are also in phase and thus additive.

Studies of intrinsically bent DNA and bent DNA-protein complexes formed by a number of small bacterial proteins have revealed that in 8% polyacrylamide gels the mobility of a bent

DNA molecule is proportional to $L \cos \alpha / 2$ where L is the length of the unbent DNA and α the angle by which the DNA is bent from linearity (35). If it is assumed that DNA with a bend at the end has a length virtually identical to L then the ratio of the relative mobilities of terminally bent and centrally bent DNA approximates to $\cos \alpha / 2$. It is possible, therefore, to determine an approximate value for the bend angle induced by factor binding. Such calculations give a value for α induced by the trypanosome factor of at least 45°. Within the 7SL locus the additive effect of two such bends will be to deform greatly the helical axis of DNA within this region, conceivably enforcing a novel and functionally significant local chromatin structure. The recent observation that, in yeast, TFIIC is required to bind to a downstream B box in order to overcome chromatin repression of the U6 gene suggests that an important role of TFIIC is to alter chromatin conformation favouring transcription (24). It is possible, therefore, that the observations that the trypanosome, RNAP III transcribed, U2 snRNA gene requires upstream A and B boxes for expression *in vivo* (20) and that mutation of tRNA genes within the 7SL and U6 loci affects expression of the neighbouring, but divergently transcribed, 7SL, U3 and U6 genes (23) may be explained by the trypanosome TFIIC activity playing a structural role rather than purely a direct role in assembly of the pre-initiation complex.

ACKNOWLEDGEMENTS

We acknowledge financial support from the Medical Research Council and the Wellcome Trust. SDB is an MRC Research Training Fellow and JDB is a Wellcome Trust Senior Lecturer.

REFERENCES

- Geiduschek, E. P. and Kassavetis, G.A. In McKnight, S.L. and Yamamoto, K.R. (eds) *Transcriptional Regulation*. Vol. 1, Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY, pp. 247-280.
- Marzouki, N., Camier, S., Ruet, A., Moenne, A. and Sentenac, A. (1986) *Nature*, **323**, 176-178.
- Geiduschek, E.P. and Tocchini-Valentini, G.P. (1988) Transcription by RNA polymerase III. *Annu. Rev. Biochem.*, **59**, 933-969.
- Baker, R. E., Camier, S., Sentenac, A. and Hall, B. D. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8768-8772.
- Dean, N. and Berk, A.J. (1987). *Nucleic Acids Res.*, **15**, 9895-9907.
- Cromlish, J.A. and Roeder, R.G. (1989) *J. Biol. Chem.*, **264**, 18 100-18 109.
- Schneider, H.R., Waldschmidt, R. and Seifart, K. (1990) *Nucleic Acids Res.*, **17**, 5003-5017.
- Schneider, H.R., Waldschmidt, R., Jahn, D. and K.H. Seifart. (1989) *Nucleic Acids Res.*, **18**, 4743-4750.
- Tapping, R.I., Syroid, D.E., Bilan, P.T. and Capone, J.P. (1993) *Nucleic Acids Res.*, **21**, 4476-4482.
- Lefebvre, O., Carles, C., Conesa, C., Swanson, R.N., Bouet, F., Riva, M. and Sentenac, A. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 10 512-10 516.
- L'Etoile, N. D., Fahnestock, M. L., Shen, Y., Aebersold, R. and Berk, A.J. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 1652-1656.
- Marck, C., Lefebvre, O., Carles, C., Riva, M., Chaussivert, N., Ruet, A. and Sentenac, A. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 4027-4031.
- Swanson, R. N., Conesa, C., Lefebvre, O., Carles, C., Ruet, A., Quemeneur, E., Gagnon J. and Sentenac, A. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4887-4891.
- Kassavetis, G.A., Braun, B.R., Nguyen, L.H. and Geiduschek, E.P. (1990) *Cell*, **60**, 235-245.
- Zomerijk, J.C.B.M., Kieft, R., Shiels, P.G. and Borst, P. (1991) *Nucleic Acids Res.*, **19**, 5153-5158.
- Zomerijk, J.C.B.M., Oullette, M., tenAsbroek, A.L.M.A., Kieft, R., Bommer, A.M.M., Clayton, C.E. and P. Borst, P. (1990) *EMBO J.*, **9**, 2791-2801.

- 17 Brown, S.D., Huang, J. and Van der Ploeg, L.H.T. (1992) *Mol. Cell. Biol.*, **12**, 2644–2652
- 18 Clayton, C.E., Fueri, J.P., Itzhaki, J.E., Bellofatto, V., Sherman, D.R., Wisdom, G.S., Vijayasarathy, S. and Mowatt, M.R. (1990) *Mol. Cell. Biol.*, **10**, 3036–3047
- 19 Tschudi, C., Williams, S.P. and Ullu, E. (1990) *Gene*, **91**, 71–77
- 20 Fantoni, A., Dare, A.O. and Tschudi, C. (1994) *Mol. Cell. Biol.*, **14**, 2021–2028
- 21 Clayton C.E. (1992) *Nucleic Acid Res. Mol. Biol.*, **43**, 37–66
- 22 Mottram, J. C., Bell, S.D., Nelson, R.G. and Barry, J.D. (1991) *J. Biol. Chem.*, **266**, 18 313–18 317
- 23 Nakaar, V., Dare, A.O., Hong, D., Ullu, E. and Tschudi, C. (1994) *Mol. Cell. Biol.*, **14**, 6736–6742
- 24 Burnol, A.F., Margottin, F., Huet, J., Almouzni, G., Prioleau, M.-N., Mechali, M. and Sentenac, A. (1993) *Nature*, **362**, 475–477
- 25 Lanham S.M. and Godfrey, D. (1970) *Exp. Parasitol.*, **28**, 521–535
- 26 Deng, W.P. and Nickoloff, J.A. (1992) *Anal. Biochem.*, **200**, 81–88
- 27 Kim, J., Zweib, C., Wu, C. and Adhya, S. (1989) *Gene*, **85**, 15–23
- 28 Dignam, J. D., Lebovitz, R.M. and Roeder, R. G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489
- 29 Sigman, D. S., Kuwabara, M. D., Chen, C.-H. B. and Bruice, T. W. (1991) *Methods Enzymol.*, **208**, 414–433
- 30 Maxam, A.M. and Gilbert W. (1980) *Methods Enzymol.*, **65**, 499–560
- 31 Leveillard, T., Kassavitis, G.A. and Geiduschek, E.P. (1991) *J. Biol. Chem.*, **266**, 5162–5168
- 32 Wu, H.-M. and Cruthers, D. (1984) *Nature*, **308**, 509–513
- 33 Stillman, D. J., Caspers, P. and Geiduschek, E.P. (1985) *Cell*, **40**, 311–317
- 34 Kahn, J.D., Yun, E. and Crothers, D.M. (1994) *Nature*, **368**, 163–166
- 35 Thompson, J.F. and Landy, A. (1988) *Nucleic Acids Res.*, **20**, 9687–9705