Nucleic Acids Research, 1995, Vol. 23, No. 22 4583–4590

HMG box 4 is the principal determinant of species specificity in the RNA polymerase I transcription factor UBF

Carol Cairns and Brian McStay*

Biomedical Research Centre, Ninewells Hospital and Medical School, The University of Dundee, Dundee DD1 9SY, UK

Received September 7, 1995; Revised and Accepted October 12, 1995

ABSTRACT

Transcription of ribosomal genes requires, in addition to RNA polymerase I, the *trans*-acting factors UBF and Rib1 in *Xenopus* or SL1 in humans. RNA polymerase I transcription is remarkably species specific. Between closely related species SL1 is the sole determinant of this specificity. Between more distantly related species, however, UBF is also a component of this species specificity. *Xenopus* UBF cannot function in human RNA polymerase I transcription and human UBF cannot function in *Xenopus* RNA polymerase I transcription. *Xenopus* and human UBFs are remarkably similar at the amino acid sequence level, both containing multiple HMG box DNA binding motifs. The only major difference between xUBF and hUBF is the lack of a HMG box 4 equivalent in xUBF. Utilizing a series of hybrid UBF molecules we have identified HMG box 4 as the principal determinant of species specificity. Addition of human HMG box 4 to xUBF converts it to a form that functions in human RNA polymerase I transcription. Deletion of HMG box 4 from hUBF converts it to a form that functions in *Xenopus* RNA polymerase I transcription. Furthermore, mutations within *Xenopus* UBF demonstrate that UBF requires a precise arrangement and number of HMG boxes to function in RNA polymerase I transcription.

INTRODUCTION

RNA polymerase I (pol I) transcription requires, in addition to pol I, the *trans*-acting factor UBF and a second factor termed Rib1 in *Xenopus* (1), SL1 in humans (2), TFIB (3) or Factor D (4,5) in the mouse and rSL1 (6) in the rat. For the sake of clarity we will refer to the factors, SL1, TFIB and Factor D collectively as SL1. Human SL1 is comprised of the TATA binding protein (TBP) and three TBP-associated factors (TAFs) of 110, 63 and 48 kDa (7,8). Recently the entire peptide sequences of all three human SL1 TAFs have been published and it has been demonstrated that transcriptionally active SL1 can be reconstituted from recombinant components (9,10). Rib1 is less well characterized, but it does appear to be the *Xenopus* equivalent of SL1 by virtue of the fact that it contains TBP (M.Bodeker and B.McStay, unpublished observation). One of the characteristics of RNA pol I transcription is that it is remarkably species specific. It appears that between more closely related species, for example mouse and humans, SL1 is the determining component in species specificity (2,11,12).

UBF has been cloned from human (13), rat (14), mouse (15,16) and *Xenopus* (1,17). The availability of recombinant UBF and a UBF responsive transcription system has led to a detailed understanding of the functional domains of UBF (18,19). UBF exists in solution and binds to DNA as a dimer. Sequences between amino acids 22 and 98 in xUBF appear to be necessary and sufficient for dimerization. An intact dimerization domain is essential for UBF to function in pol I transcription. The DNA binding activity of UBF is mediated by the high mobility group (HMG) boxes, so called because of their homology to the DNA binding domain of the small non-histone chromatin-associated proteins HMG1 and HMG2 (13). In human UBF there are six identifiable HMG boxes, while in *Xenopus* UBF there are five (17). In the case of xUBF it is clear that the first three HMG boxes are essential for transcription activity and that deletion of the two C-terminal HMG boxes has little effect on transcription. Likewise, deletion of any of the first four HMG boxes in hUBF has a negative effect on transcription, but deletion of the two C-terminal HMG boxes has little effect. Each of the first three HMG boxes in xUBF and each of the first four HMG boxes in hUBF constitutes an individual DNA binding domain (18,19). The C-terminus of UBF is highly acidic. This domain is required for full activity of UBF in transcription. Furthermore, it appears that phosphorylation of serine residues within the acidic tail of UBF by casein kinase II is a pre-requisite for full transcription activity (16,20–22).

As a consequence of alternate splicing it appears that in all mammalian species so far studied there are two forms of UBF (UBF1 and UBF2) (14). In UBF2 35 amino acids are removed from HMG box 2, rendering it inactive as a pol I transcription factor (19,23). The role of UBF2 remains unclear. In *Xenopus* there are also alternate forms of UBF, which arise as a result of there being two genes for UBF (24). However, unlike mammalian UBF, differences between both forms of xUBF are confined to the C-terminus. The result of this is that all forms of xUBF so far

* To whom correspondence should be addressed

cloned are functional in transcription (B.McStay, unpublished observation).

The absolute requirement for SL1 and Rib1 for pol I transcription in mammals and *Xenopus* respectively has been well documented. The role of UBF in pol I transcription *in vitro* is more equivocal. In the *Xenopus* system it is clear that UBF is an essential factor (1,18). In mammalian systems it appears that UBF is an auxilliary transcription factor (25–27). Indeed, some experiments have indicated that the role of UBF is to negate the effect of pol I transcription inhibitors (26).

As mentioned above, between closely related species SL1 is the only determinant of species specificity. Between more distantly related species, for example *Xenopus* and human, UBF is also species specific. It has been shown that purified hUBF cannot substitute for xUBF in *Xenopus* pol I transcription and that purified xUBF cannot substitute for hUBF in human pol I transcription (28). This result is all the more surprising since both species UBF give rise to identical DNase I footprints on the *Xenopus* rDNA promoter and the human rDNA promoter.

Jantzen *et al.* (19) have tested hybrids between xUBF and hUBF for their ability to function in human pol I transcription. These hybrids were fused within their common HMG box 3. When the N-terminus is derived from xUBF and the C-terminus from hUBF the resulting hybrid UBF functions in human pol I transcription. The reciprocal of this (hUBF N-terminus and *Xenopus* C-terminus) does not function in human pol I transcription. Thus sequences present in hUBF C-terminal to HMG box 3 can convert xUBF into a form that functions in human pol I transcription. These hybrid molecules were not, however, tested in the *Xenopus* system.

In this study we have used a more systematic approach to identify the sequences in UBF that confer species specificity. We have constructed an extensive series of hybrid UBF molecules and tested their ability to stimulate pol I transcription in both *Xenopus* and human systems. We demonstrate that the principal determinant of UBF species specificity is the human HMG box 4, which is absent in xUBF. Furthermore, we demonstrate that the precise nature of the HMG boxes is important for the functioning of UBF in pol I transcription.

MATERIALS AND METHODS

Antibodies

Anti-xUBF antibodies were prepared by immunizing rabbits with full-length xUBF produced using a baculovirus expression system (29). Anti-human UBF antibodies were prepared by immunizing rabbits with a truncated form of hUBF1 expressed in bacteria. The hUBF dimer domain and HMG boxes 1–4 were cloned as a *Nco*I–*Bam*HI restriction fragment into a pET vector (Novagen) derivative (pET6H) that contained a six histidine tag. This truncated form of hUBF was expressed in the bacterial strain BL21 DE3/pLys S, purified by chromatography on nickel–agarose and refolded as described previously (29).

Detection of UBF in transcription extracts by Western blotting was performed using the above antibodies with secondary antibodies coupled to horseradish peroxidase (BioRad). Western blots were visualized using ECL chemiluminesence (Amersham).

Plasmids

pGem40. The plasmid pGem40 contains the *Xenopus* rDNA promoter sequences –245 to +50 cloned as a *Sal*I–*Bam*HI restriction fragment in the vector pGem3 (Promega). This plasmid was used as the template for *in vitro* transcription in the *Xenopus* system.

pHENA. The plasmid pHENA has been previously described (30); it contains the human rDNA promoter sequences from –500 to +80 cloned as an *Eco*RI restriction fragment. This plasmid was used as the template for *in vitro* transcription in the human system.

pCITE xUBF. An *Nco*I restriction site was created at the site of translation initiation in xUBF by site-directed mutagenesis. A 586 nt *Eco*RI–*Nco*I restriction fragment from the vector pCITE-I (Novagen), which contains the IRES element from EMCV, was fused to the newly created *Nco*I site in xUBF and the resulting CITE xUBF fusion was cloned as an *Eco*RI restriction fragment into the vector pBluescript SK$^+$ (Stratagene) such that the phage T7 RNA polymerase promoter can direct transcription of synthetic xUBF message. This plasmid is described in detail elsewhere (1).

pCITEhUBF1. A complete cDNA encoding hUBF1 was obtained by screening a λgt11 cDNA library prepared from HeLa cell mRNA. In order to obtain a full-length cDNA duplicate filters were probed with oligonucleotides derived from the 5′- and 3′-ends of the published cDNA sequence. The site of translation initiation of the hUBF1 open reading frame was converted to an *Nco*I restriction site, fused to the EMCV IRES element and cloned into pBluescript SK$^+$ as described above for xUBF.

Mutant 1. PCR was used to generate a DNA fragment encoding amino acids 388–474 (HMG box 4) of hUBF 1 with *Bgl*II restriction sites at both ends. This fragment was then cloned into a version of pCITExUBF in which amino acids 387 and 386 were mutated to encode a *Bgl*II restriction site (LS I in 18).

Mutant 2. Oligonucleotide-directed mutagenesis was used to precisely delete the DNA sequence that encodes amino acids 394–475 (HMG box 4) from the hUBF1 open reading frame in plasmid pCITEhUBF1.

Mutant 3. Oligonucleotide-directed mutagenesis was used to convert the sequence encoding amino acids 483 and 484 of hUBF1 in the plasmid pCITEhUBF1 into a novel *Bam*H1 restriction site. The *Nco*I–*Bam*HI restriction fragment encoding hUBF amino acids 1–482 was replaced with the *Nco*I–*Bgl*II fragment from the pCITExUBF derivative LS I, which encodes amino acids 1–386 of xUBF.

Mutant 4. The plasmid used to generate mutant protein number 4 is in fact the pCITExUBF derivative LS I. Truncated protein is produced by linearizing the plasmid with *Bgl*II (see below).

Mutant 5. This plasmid has been previously described. It is a derivative of pCITExUBF in which the DNA encoding amino acids 22 to 98 have been deleted and replaced with a novel *Bgl*II restriction site.

Mutant 6. PCR was used to generate a DNA fragment encoding amino acids 1–95 of hUBF1 with *Nco*I and *Bgl*II restriction sites at the 5′- and 3′-ends respectively. This fragment was then used to replace the *Nco*I–*Bgl*II restriction fragment of mutant 5 above.

Mutants 7–11. The pCITExUBF derivatives LS F, J, H and I have been described elsewhere (18). They contain novel *Bgl*II restriction sites that define the boundaries between the dimer domain and HMG box 1 (F), HMG box 1 and HMG box 2 (G), HMG box 2 and HMG box 3 (H) and the C-terminus of HMG box 3 (I). These mutants were used in various combinations to create a mutant xUBF in which HMG box 2 is replaced by a duplicate HMG box 1 (mutant 7), HMG box 2 is replaced by a duplicate HMG box 3 (mutant 8), a duplicate HMG box 1 has been inserted between HMG boxes 1 and 2 (mutant 9), a duplicate HMG box 2 has been inserted between HMG boxes 2 and 3 (mutant 10), a duplicate HMG box 3 has been inserted on the C-terminal side of HMG box 3 (mutant 11).

Transcription extracts and UBF immunodepletion

Xenopus S100 transcription extracts were prepared from XlK-2 cells as previously described (31). *Xenopus* UBF was immunodepleted from S100 extract by incubating 2.0 ml S100 extract with 10 µl α-xUBF antiserum on ice for 30 min then chromatographed three times over a 0.5 ml protein A–Sepharose fast flow column (Pharmacia). Depletion of UBF was monitored by Western blotting as described above.

Human S100 transcription extracts were prepared from HeLa cells as previously described for *Xenopus* cells. HeLa cells were grown in spinner culture in RPMI media with 10% fetal bovine serum to a density of 5×10^5 cells/ml. Human UBF was immunodepleted as described above for xUBF except that 25 µl α-hUBF antiserum were added to 2 ml S100 extract.

In vitro transcription/translation

The plasmids pCITExUBF, pCITEhUBF1 and their mutant derivatives were digested with the restriction enzyme *Xba*I, which cuts in the polylinker downstream of the UBF open reading frame. The one exception was that to produce the xUBF C-terminal deletion (mutant 4), the pCITExUBF derivative LS I being digested with *Bgl*II. Each of the digested plasmids was then transcribed with phage T7 RNA polymerase and the resulting transcript was translated *in vitro* in a rabbit reticulocyte lysate. Translation reactions were in a 25 µl volume and contained 17.5 µl rabbit reticulocyte lysate (Promega), 20 µM amino acids (including methionine), 20 µCi [^{35}S]methionine (1000 Ci/mM; Amersham), 20 U RNasin (Promega) and 0.5 µg synthetic UBF message. Reactions were incubated at 30°C for 90 min. Aliquots of each reaction were electrophoresed in 10% SDS–polyacrylamide gels. Following electrophoresis gels were fixed in 40% methanol, 10% acetic acid, dried and autoradiographed.

Transcription assays

Transcription extracts (20 µl/reaction) on their own or together with α-UBF antibodies or *in vitro* translated UBF were combined with 400 ng supercoiled template DNA and incubated on ice for 10 min. All the plasmid templates used here were irradiated with UV as described previously (31). This was to prevent RNA polymerase from reading through and disrupting the transcription complex on the promoter. Reactions were initiated by the addition of 20 µl buffer containing $MgCl_2$ and nucleotide triphosphates (NTPs). The final reaction conditions were 25 mM HEPES, pH 7.5, 90 mM KCl, 6 mM$MgCl_2$, 10% glycerol, 1 mM DTT, 0.5 mM NTPs, 10 mM creatine phosphate and 100 µg/ml α-amanatin. *Xenopus* transcription reactions were incubated at 25°C for 180 min and human transcription reactions were incubated at 30°C for 60 min. Transcription reactions were terminated and analysed by S1 nuclease protection as described previously (31). The probe used for detection of transcripts from *Xenopus* templates was the 5′-end-labelled coding strand of the –245 to +50 *Sal*I–*Bam*HI insert of pGem40. The probe used for detection of transcripts from human templates was the 5′-end-labelled coding strand of the –55 to +80 *Eco*RI fragment from p5′Δ55HENA (a 5′ deletion mutant derived from pHENA). Quantitation of transcription signals was performed using a phosphorimager (BioRad).

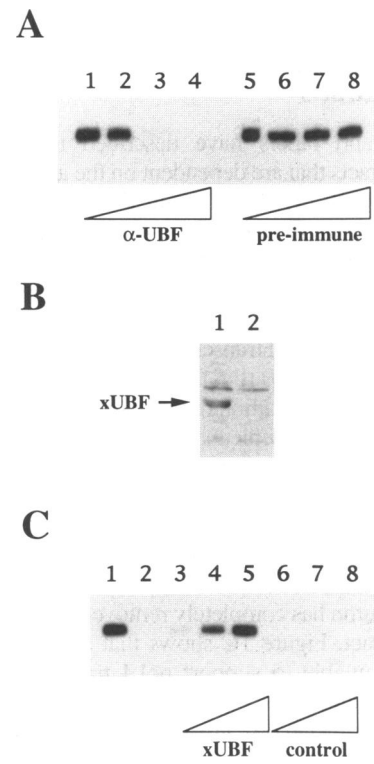

Figure 1. Immunodepletion of xUBF from *Xenopus* S100 extract results in a xUBF responsive pol I transcription extract. (**A**) Transcription reactions were performed in a *Xenopus* S100 transcription extract in the presence of 0.05 (lane 1), 0.1 (lane 2), 0.5 (lane 3) or 1.0 µl (lane 4) α-xUBF antiserum or in the presence of 0.05 (lane 5), 0.1 (lane 6), 0.5 (lane 7) or 1.0 µl (lane 8) pre-immune serum. (**B**) A Western blot of a 5 µl aliquot of untreated (lane 1) and xUBF-immunodepleted (lane 2) *Xenopus* S100 extract. The position of xUBF is indicated with an arrow. (**C**) Transcription reactions were performed in the untreated (lane 1) and immunodepleted (lanes 2–8) *Xenopus* S100 extracts. The transcription reaction shown in lane 2 contained no added xUBF. Transcription reactions shown in lanes 3–5 contained respectively 0.5, 1.5 and 2.5 µl *in vitro* translated xUBF. Transcription reactions shown in lanes 6–8 contained respectively 0.5, 1.5 and 2.5 µl control translation reaction.

RESULTS

Immunodepletion of xUBF from *Xenopus* pol I transcription extracts

Previously we and others have described fractionated pol I transcription extracts that are dependent on the addition of purified or recombinant UBF (1,13,18,19). In this report we have used immunodepletion to make transcription extracts that are dependent on the addition of exogenous UBF. In order to generate α-xUBF antibodies rabbits were immunized with full-length xUBF produced in a baculovirus expression system. The resulting antiserum can specifically inhibit pol I transcription (Fig. 1A). Addition of 0.5 μl α-xUBF antiserum to 20 μl S100 extract can totally inhibit pol I transcription, whereas addition of up to 1.0 μl pre-immune serum has no effect. To immunodeplete the *Xenopus* transcription extract of xUBF S100 extract was incubated with α-UBF antiserum, then loaded repeatedly onto a protein A–Sepharose column. A Western blot of treated and untreated S100 extract shows that we have successfully immunodepleted xUBF (Fig. 1B). Since no rabbit IgG was detected on the Western blot we can conclude that the protein A column has completely removed all of the IgG added to the S100 extract. Figure 1C shows that this immunodepleted S100 extract is unable to support pol I transcription initiation (compare lanes 1 and 2).

Xenopus UBF translated *in vitro* in a rabbit reticulocyte lysate can substitute for purified xUBF in a fractionated *Xenopus* pol I transcription system (1). In Figure 1C we show that *in vitro* translated xUBF can restore transcription activity to the UBF-depleted extract. Addition of 2.5 μl xUBF translation reaction restores transcription activity to levels comparable with the non-depleted S100 extract (lane 5). Addition of 2.5 μl control translation reaction has no stimulatory effect on the depleted extract (lane 8). From this experiment we can also conclude that immunoprecipitation of UBF has not co-precipitated to a significant degree any other factor that is essential for pol I transcription initiation.

Immunodepletion of hUBF from human pol I transcription extracts

Attempts to immunodeplete a human pol I transcription extract of hUBF using α-xUBF antiserum proved unsuccessful. This was somewhat surprising given the similarity of *Xenopus* and human UBFs at the amino acid sequence level. As a consequence we have expressed a portion of human UBF1 (dimer domain plus HMG boxes 1–4) in *Escherichia coli* and used this to immunize rabbits (see Materials and Methods for details). The resulting antiserum was used to immunodeplete a human pol I transcription extract as described above for the *Xenopus* system. Again, Western blotting demonstrated that we had successfully immunodepleted the extract of hUBF (Fig. 2A). In contrast to the *Xenopus* system, immunodepletion of hUBF does not result in a complete inhibition of pol I transcription initiation (Fig. 2B, compare lanes 1 and 2). Nonetheless, addition of *in vitro* translated hUBF can stimulate transcription in this depleted extract. Quantitation of the stimulation shows that addition of 2.5 μl *in vitro* translated hUBF to the depleted system results in a 6-fold stimulation of transcription, whereas addition of 2.5 μl control *in vitro* translation reaction has no effect (lane 8). Addition of *in vitro* translated hUBF does not fully restore transcription to the level observed in the non-depleted extract. However, addition

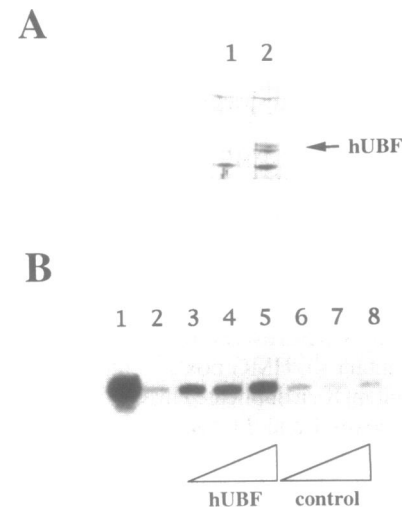

Figure 2. Immunodepletion of hUBF from HeLa S100 extract results in a hUBF responsive pol I transcription extract. (**A**) A Western blot of a 5 μl aliquot of hUBF-immunodepleted (lane 1) and untreated (lane 2) HeLa S100 extract. The position of hUBF is indicated with an arrow. (**B**) Transcription reactions were performed in the untreated (lane 1) and immunodepleted (lanes 2–8) HeLa S100 extracts. The transcription reaction shown in lane 2 contained no added hUBF. Transcription reactions shown in lanes 3–5 contained respectively 0.5, 1.5 and 2.5 μl *in vitro* translated hUBF. Transcription reactions shown in lanes 6–8 contained respectively 0.5, 1.5 and 2.5 μl control translation reaction.

of larger amounts of hUBF, produced using a baculovirus expression system, can fully restore transcription activity (data not shown). Thus we conclude that immunodepletion of hUBF from the HeLa cell extract does not remove to a significant degree any other essential pol I transcription factor.

There are a number of possible reasons that immunodepletion of UBF from human transcription extracts does not completely eliminate pol I transcription. Firstly, it is possible that immunodepletion from the human system is not as efficient as that from the *Xenopus* transcription extract. A second, more likely, explaination is that a fraction of the transcription observed in the human system is independent of UBF. This has already been observed in human and other mammalian systems (25–27). In support of this second explanation we have observed that a human pol I promoter deleted on its 5′-side to –55 is transcribed at 10–20% of the efficiency of a full promoter in a non-depleted extract. In a depleted extract the full promoter and core promoters are transcribed with equal, albeit lower, efficiency (C. Cairns and B. McStay, unpublished observation). Thus we conclude that the transcription observed in the UBF-depleted human extract (Fig. 2B, lane 2) is equivalent to that observed from a deleted promoter and is UBF independent.

Construction and expression of xUBF/hUBF domain swap mutants

In order to identify sequences in UBF that confer species specificity we have constructed a series of UBF deletion mutants and xUBF/hUBF hybrid molecules. The structure of these mutants is shown in Figure 3A. In mutant 1 hUBF HMG box 4 has been inserted into a position in xUBF equivalent to its normal location in hUBF1. This results in a xUBF that now has a similar configuration of HMG boxes to hUBF1. In mutant 2 hUBF has

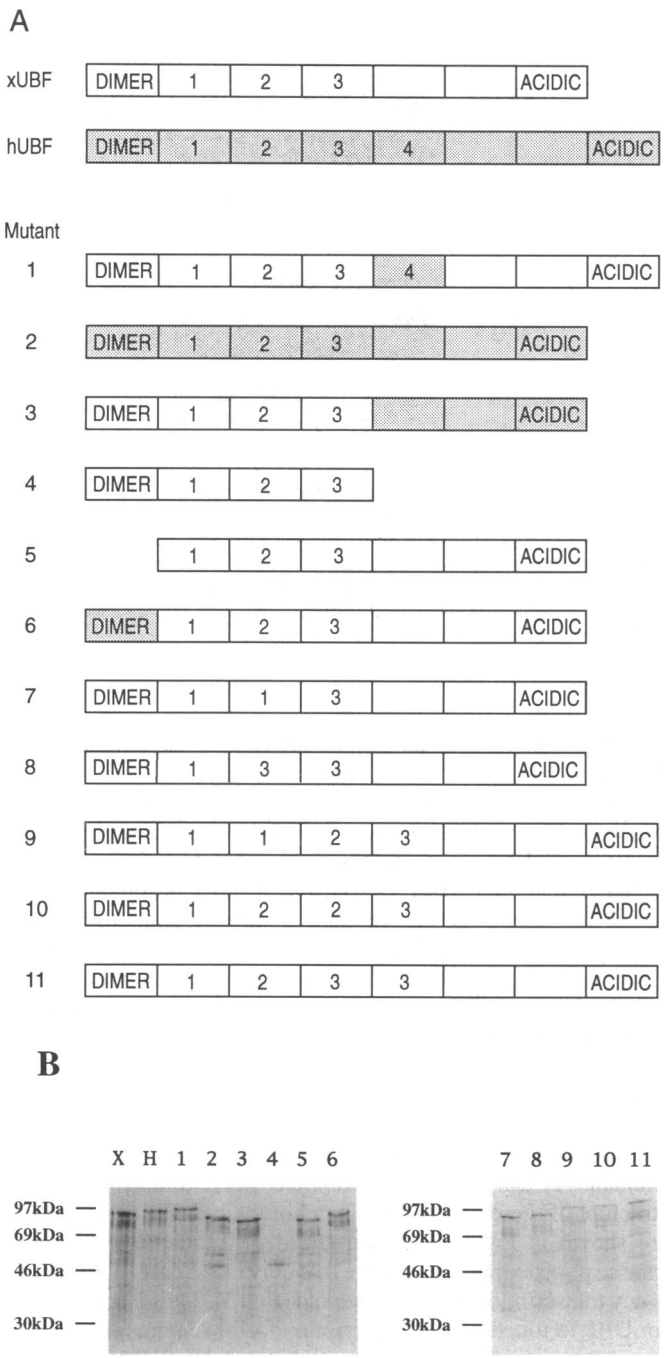

Figure 3. Structure and expression of xUBF/hUBF mutants. (**A**) The structure of xUBF, hUBF1 and their mutant derivatives are shown in diagrammatic form. Open boxes are used to denote sequences derived from xUBF and shaded boxes to denote sequences derived from hUBF. The dimerization domain is labelled DIMER, the C-terminal acidic tail is labelled ACIDIC and HMG boxes 1–3 in xUBF and 1–4 in hUBF are identified by a number in the box. For reasons of clarity the C-terminal two HMG boxes in xUBF and hUBF are not numbered. (**B**) xUBF, hUBF and all the mutant derivatives were transcribed by T7 RNA polymerase and the resulting transcripts translated in a rabbit reticulocyte lysate in the presence of [^{35}S]methionine. Aliquots of each translation reaction were electrophoresed in a 10% SDS–polyacrylamide gel alongside pre-stained molecular weight standards. The identity of each translation product is shown above the gel and the positions of the molecular weight standards are shown alongside.

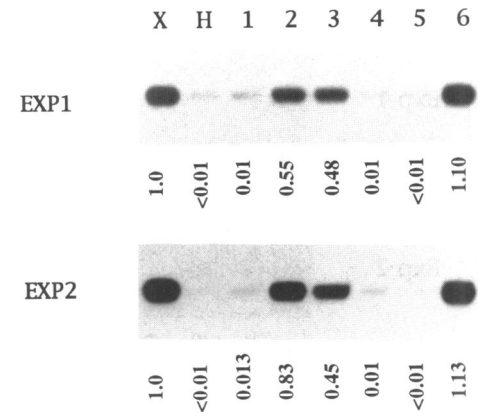

Figure 4. xUBF/hUBF domain swap mutants tested in the *Xenopus* pol I transcription system. *In vitro* translated xUBF, hUBF and mutants 1–6 were compared for their ability to stimulate transcription in the UBF immunodepleted *Xenopus* S100 extract. In experiment 1 (EXP1) 1.5 µl aliquots of each translation reaction were added and in experiment 2 (EXP2) 2.5 µl aliquots were added. In each experiment transcription signals for each UBF form were quantitated and expressed relative to that observed with xUBF set at 1.0.

been converted to a molecule that more closely resembles xUBF by deletion of HMG box 4. In mutant 3 the acidic tail and C-terminal HMG boxes in xUBF have been replaced with those from hUBF and in mutant 4 they have been deleted. Finally, in mutant 5 the dimerization domain of xUBF has been deleted and in mutant 6 replaced with that of hUBF. Mutants 7–11, also shown in Figure 3A, will be described below.

In order to facilitate translation in a rabbit reticulocyte lysate xUBF, hUBF and all of the above mutant UBFs were cloned downstream of the internal ribosome entry site (IRES) element from encephalomyocarditis virus (EMCV). We have demonstrated previously that this EMCV IRES element results in up to a 10-fold increase in translational efficiency compared with the natural UBF 5′-untranslated region (1). *Xenopus* and human UBFs and all of the mutant derivatives described above were translated with approximately equal efficiency, giving rise to products of the predicted size. This was determined by SDS–PAGE of aliquots of *in vitro* translation reactions that contained [^{35}S]methionine (Fig. 3B).

xUBF/hUBF domain swap mutants tested in *Xenopus* and human pol I transcription

In vitro translated xUBF, hUBF and mutants 1–6 were tested for their ability to stimulate pol I transcription in the immunodepleted *Xenopus* extract. Figure 4 shows the results of two independent experiments. In both experiments *in vitro* translated hUBF stimulates transcription at <1% of the efficiency of xUBF, thus confirming previous results with purified xUBF and hUBF (28). Insertion of the human HMG box 4 into xUBF results in a reduction of transcription signal equivalent to that observed with hUBF. Deletion of HMG box 4 from human UBF converts it to a form that approximates the transcription activity observed with unaltered xUBF (55% in experiment 1 and 83% in experiment 2). This leads to the conclusion that the main reason for hUBF not functioning in the *Xenopus* system is the presence of HMG box 4. This conclusion is further strengthened by the observation that deletion of sequences C-terminal to HMG box 3 in xUBF (mutant 4) results in its inactivation in transcription. Replacement of these

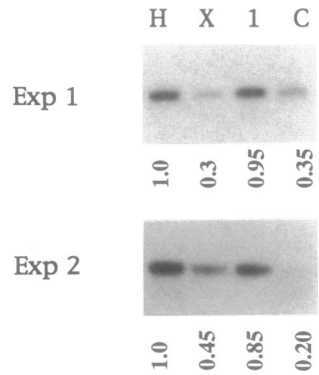

Figure 5. A xUBF mutant containing the human HMG box 4 functions in human pol I transcription. *In vitro* translated hUBF, xUBF and mutant 1 were compared for their ability to stimulate transcription in the UBF immunodepleted human S100 extract. In both experiments (Exp 1 and Exp 2) 2.5 µl aliquots of each translation reaction were added. In each experiment a control, addition of a 2.5 µl aliquot of a control translation reaction, was included. For each experiment the transcription signals observed are expressed relative to that observed with hUBF set at 1.0.

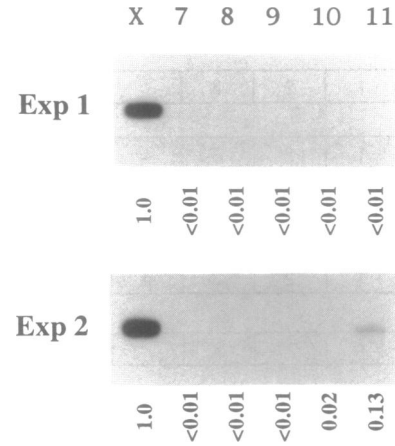

Figure 6. xUBF is sensitive to changes in number or identity of HMG boxes. *In vitro* translated xUBF and mutants 7–11 were compared for their ability to stimulate transcription in the UBF-immunodepleted *Xenopus* S100 extract. In experiment 1 (Exp 1) 1.5 µl aliquots of each translation reaction were added and in experiment 2 (Exp 2) 2.5 µl aliquots were added. In each experiment transcription signals for each UBF form were quantitated and expressed relative to that observed with xUBF, set at 1.0.

sequences with the equivalent sequences from hUBF (mutant 3) restores activity to 48 (experiment 1) or 45% (experiment 2) of that observed with unaltered xUBF. Similarly, deletion of the xUBF dimer domain (mutant 5) inactivates it as a transcription factor, but replacement of the xUBF dimer domain with that from hUBF (mutant 6) restores transcription activity to levels that are even higher than with normal xUBF (110% in experiment 1 and 113% in experiment 2).

In vitro translated xUBF, hUBF and mutant 1 (xUBF + HMG box 4) were also tested for their ability to stimulate transcripton in the depleted human transcription extract. As a result of background transcription and the relatively low level of stimulation in the UBF-depleted human transcription extract this experiment was performed numerous times. Two representative experiments are shown in Figure 5. Control reactions, in which aliquots of a non-UBF-programmed translation mix were added, demonstrate the level of UBF-independent transcription in the depleted extract. In both experiments shown hUBF stimulates transcription 3- to 4-fold, xUBF does not stimulate transcription and mutant 1 (xUBF + HMG box 4) stimulates transcription to a similar degree to hUBF. Thus it appears that the primary reason for the inability of xUBF to work in the human transcription extract is the absence of HMG box 4.

A precise order and number of HMG boxes is essential for UBF function in pol I transcription

One of the characteristics of the HMG boxes found in UBF is that a given HMG box has a far greater similarity to the equivalent box in another species than to adjacent HMG boxes within the same UBF molecule. This observation holds true even for more distantly related species, such as *Xenopus* and human UBF (1,17). This leads to the hypothesis that each HMG box has a discrete function. A test of this hypothesis is to delete an individual HMG box and replace it with an alternative HMG box. To this end we have constructed xUBF mutants 7 and 8 in which HMG box 2 has been deleted and replaced with a duplicate HMG box 1 (mutant 7) or HMG box 3 (mutant 8). The structure of these mutants and their translation products are shown in Figure 3.

Figure 6 shows the results of two independent experiments in which *in vitro* translated mutants 7 and 8 were compared with unaltered xUBF for their ability to stimulate transcription in the depleted *Xenopus* extract. Despite having an appropriate number of HMG boxes, neither mutant 7 nor 8 can function in transcription. We conclude from this result that an individual HMG box cannot be replaced with an alternative box from within the same species. This is in contrast to the experiments described above, which show that all HMG boxes can be replaced by the equivalent HMG boxes from even a distantly related species.

In another class of xUBF mutants we have asked the question whether xUBF that contains intact HMG boxes 1, 2 and 3, but with a duplicate box 1 (mutant 9), box 2 (mutant 10) or box 3 (mutant 11) can function in *Xenopus* pol I transcription. The structure of these mutants and a gel of their translation products are shown in Figure 3. When tested in the depleted *Xenopus* transcription extract we observe that only in the case of mutant 11 is any transcription activity observed (Fig. 5). Thus it appears that for UBF to function as a transcription factor requires not only an appropriate number of HMG boxes, but that these HMG boxes be of a specific type and in a specific order.

DISCUSSION

A comparison of the primary sequence of hUBF and xUBF reveals that they are strikingly similar (1,17). Individual *Xenopus* HMG boxes have between 90 and 97% similarity (allowing for conservative changes) with their counterpart in hUBF. The dimer domain and the acidic C-terminal tail have 76 and 91% homology respectively. The only major difference between the two UBFs is the absence of a HMG box 4 equivalent in xUBF. In the experiments described here we have clearly demonstrated that it is the absence or presence of this HMG box that is the main determinant of species specificity between *Xenopus* and humans.

A hUBF1 mutant protein in which HMG box 4 has been deleted behaves in all respects like xUBF and a xUBF mutant protein in which human HMG box 4 has been inserted at an appropriate position behaves in all respects like hUBF. Other hybrid UBF proteins confirm that the dimer domain and the acidic tails of both species UBF are fully interchangeable.

As stated above, one of the characteristics of UBF HMG boxes is that they are far more closely related to their conterparts in a distantly related species, such as between *Xenopus* and human, than they are to adjacent HMG boxes within the same UBF molecule. This suggests that not only does each HMG box have a discrete function, but that this function has been conserved throughout evolution. A demonstration that each HMG box has a discrete role in UBF is that xUBF mutants in which one of the required HMG boxes has been replaced by a duplicate version of one of the remaining HMG boxes cannot function in pol I transcription. Furthermore, xUBF derivatives that contain four HMG boxes, as a result of duplicating any one of the functional HMG boxes, are non-functional in pol I transcription.

We can therefore come to the following conclusions: (i) the number of functional HMG boxes is critical to UBF function in transcription; (ii) the nature of the HMG boxes is critical for transcription; (iii) the function of each HMG box is highly conserved throughout evolution.

DNase I footprinting experiments with human SL1 have shown that it cannot bind to the promoter on its own, but in the presence of hUBF it can bind to promoter DNA (11,25). Template commitment experiments in *Xenopus* have demonstrated that Rib1 and xUBF combine to form a stable transcription complex on the *Xenopus* promoter (1). This is suggestive of an interaction between xUBF and Rib1 with the promoter that is similar to SL1 and hUBF.

Jantzen *et al.* (19) demonstrated that a xUBF/hUBF hybrid could function in human pol I transcription. DNase I footprinting showed that this hybrid UBF had the ability to potentiate SL1 binding to the human promoter. This observation led to an interpretation of UBF species specificity in terms of protein–protein interactions between UBF and SL1. It was suggested that when xUBF is bound to the human pol I promoter it cannot interact with SL1 through protein–protein contacts and thus potentiate SL1 binding to DNA sequences within the promoter. A number of observations now make this explaination unlikely. Firstly, it is difficult to understand how inserting human HMG box 4 into an otherwise intact xUBF could abrogate potential xUBF–Rib1 interactions. Secondly, duplication of individual HMG boxes within xUBF would not be expected to disrupt xUBF–Rib1 interactions. Yet such mutants are completely inactive in *Xenopus* pol I transcription.

More recent work with UBF and other HMG box proteins suggests that UBF may be performing an architectural role in pol I transcription initiation. HMG boxes have been identified in a number of other eukaryotic DNA binding proteins, most notably Sry (32) and LEF (33). One of the major characteristics of these proteins is their ability to bend DNA. It has been demonstrated that the single HMG box present in LEF can introduce a 130° bend into DNA (34). Similar bending studies with the HMG boxes in UBF have been thwarted by the inability to identify a concensus binding sequence. However, it is expected that the individual HMG boxes in UBF have a similar ability to bend DNA. Two sets of observations suggest that this is indeed the case. In one study analysis of xUBF–DNA interactions by electron spectroscopic imaging demonstrated that a single dimer of xUBF can organize ~180 bp of DNA into a loop of almost 360° (35). This leads to an estimation of a bend angle associated with each HMG box of 60°. In a second study a ligase-mediated probe circularization assay was used to show the ability of xUBF to loop short DNA fragments (36). It seems likely, therfore, that the role of UBF is to precisely organize promoter DNA such that SL1/Rib1 can bind to DNA sequences within the promoter and thus form a stable transcription complex.

These observations, combined with the results presented here, suggest an alternative explaination of UBF species specificity. If the promoter-bound UBF contained an inappropriate number or organization of HMG boxes (bending elements), the promoter DNA would be organized in a form unrecognizable by SL1/Rib1, resulting in failure to produce a productive transcription complex. Thus we believe that hUBF bound to the *Xenopus* promoter overbends DNA such that Rib1 cannot bind and that xUBF bound to the human promoter underbends DNA such that SL1 cannot bind.

ACKNOWLEDGEMENTS

We would like to thank Ron Reeder in whose laboratory this work was initiated, Judy Roan for obtaining the hUBF1 cDNA and both Iain Goldsmith and Del Watling at ICRF for oligonucleotide synthesis and antibody production respectively. Finally we would like to thank Gareth Sullivan for critical reading of the manuscript. This work was supported by a grant from the Wellcome Trust (BMcS).

REFERENCES

1 McStay,B., Hu,C.H., Pikaard,C.S. and Reeder,R.H. (1991) *EMBO J.*, **10**, 2297–2303.
2 Learned,R.M., Cordes,S. and Tjian,R. (1985) *Mol. Cell. Biol.*, **5**, 1358–1369.
3 Schnapp,A., Clos,J., Hadelt,W., Schreck,R., Cvekl,A. and Grummt,I. (1990) *Nucleic Acids Res.*, **18**, 1385–1393.
4 Mishima,Y., Financsek,I., Kominami,R. and Muramatsu,M. (1982) *Nucleic Acids Res.*, **10**, 6659–6670.
5 Tower,J., Culotta,V. C. and Sollner,W.B. (1986) *Mol. Cell. Biol.*, **6**, 3451–3462.
6 Smith,S.D., Oriahi,E., Lowe,D., Yang,Y.H., O'Mahony,D., Rose,K., Chen,K. and Rothblum,L.I. (1990) *Mol. Cell. Biol.*, **10**, 3105–3116.
7 Comai,L., Tanese,N. and Tjian,R. (1992) *Cell*, **68**, 965–976.
8 Eberhard,D., Tora,L., Egly,J.M. and Grummt,I. (1993) *Nucleic Acids Res.*, **21**, 4180–4186.
9 Comai,L., Zomerdijk,J.C., Beckmann,H., Zhou,S., Admon,A. and Tjian,R. (1994) *Science*, **266**, 1966–1972.
10 Zomerdijk,J.C., Beckmann,H., Comai,L. and Tjian,R. (1994) *Science*, **266**, 2015–2018.
11 Bell,S.P., Jantzen,H.M. and Tjian,R. (1990) *Genes Dev.*, **4**, 943–954.
12 Rudloff,U., Eberhard,D., Tora,L., Stunnenberg,H. and Grummt,I. (1994) *EMBO J.*, **13**, 2611–2616.
13 Jantzen,H.M., Admon,A., Bell,S.P. and Tjian,R. (1990) *Nature*, **344**, 830–836.
14 O'Mahony,D.J. and Rothblum,L.I. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 3180–3184.
15 Hisatake,K., Nishimura,T., Maeda,Y., Hanada,K., Song,C.Z. and Muramatsu,M. (1991) *Nucleic Acids Res.*, **19**, 4631–4637.
16 Voit,R., Schnapp,A., Kuhn,A., Rosenbauer,H., Hirschmann,P., Stunnenberg,H.G. and Grummt,I. (1992) *EMBO J.*, **11**, 2211–2218.
17 Bachvarov,D. and Moss,T. (1991) *Nucleic Acids Res.*, **19**, 2331–2335.
18 McStay,B., Frazier,M.W. and Reeder,R.H. (1991) *Genes Dev.*, **5**, 1957–1968.
19 Jantzen,H.M., Chow,A.M., King,D.S. and Tjian,R. (1992) *Genes Dev.*, **6**, 1950–1963.

20 O'Mahony,D.J., Smith,S.D., Xie,W. and Rothblum,L.I. (1992) *Nucleic Acids Res.*, **20**, 1301–1308.
21 O'Mahony,D.J., Xie,W.Q., Smith,S.D., Singer,H.A. and Rothblum,L.I. (1992) *J. Biol. Chem.*, **267**, 35–38.
22 Voit,R., Kuhn,A., Sander,E.E. and Grummt,I. (1995) *Nucleic Acids Res.*, **23**, 2593–2599.
23 Kuhn,A., Voit,R., Stefanovsky,V., Evers,R., Bianchi,M. and Grummt,I. (1994) *EMBO J.*, **13**, 416–424.
24 Bachvarov,D., Normandeau,M. and Moss,T. (1991) *Febs Lett.*, 288, 55–59.
25 Bell,S.P., Learned,R.M., Jantzen,H.M. and Tjian,R. (1988) *Science*, **241**, 1192–1197.
26 Kuhn,A. and Grummt,I. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 7340–7344.
27 Smith,S.D., O'Mahony,D.J., Kinsella,B.T. and Rothblum,L.I. (1993) *Gene Expression*, **3**, 229–236.
28 Bell,S.P., Pikaard,C.S., Reeder,R.H. and Tjian,R. (1989) *Cell*, **59**, 489–497.
29 Hu,C.H., McStay,B., Jeong,S.W. and Reeder,R.H. (1994) *Mol. Cell. Biol.*, **14**, 2871–2882.
30 Palmer,T.D., Miller,A.D., Reeder,R.H. and McStay,B. (1993) *Nucleic Acids Res.*, **21**, 3451–3457.
31 McStay,B. and Reeder,R.H. (1990) *Genes Dev.*, **4**, 1240–1251.
32 Sinclair,A.H., Berta,P., Palmer,M.S., Hawkins,J.R., Griffiths,B.L., Smith,M.J., Foster,J.W., Frischauf,A.M., Lovell,B.R. and Goodfellow,P.N. (1990) *Nature*, **346**, 240–244.
33 Giese,K., Amsterdam,A. and Grosschedl,R. (1991) *Genes Dev.*, **5**, 2567–2578.
34 Giese,K., Cox,J. and Grosschedl,R. (1992) *Cell*, **69**, 185–195.
35 Bazett,J.D., Leblanc,B., Herfort,M. and Moss,T. (1994) *Science*, **264**, 1134–1137.
36 Putnam,C.D., Copenhaver,G.P., Denton,M.L. and Pikaard,C.S. (1994) *Mol. Cell. Biol.*, **14**, 6476–6488.