xUBF contains a novel dimerization domain essential for RNA polymerase I transcription

Brian McStay, Mark W. Frazier, and Ronald H. Reeder
Basic Sciences Division, Hutchinson Cancer Research Center, Seattle, Washington 98104 USA

Xenopus laevis upstream binding factor (xUBF) is an RNA polymerase I transcription factor that is required for formation of the stable initiation complex. The 701-amino-acid protein contains three regions of homology to the chromosomal protein HMG1 (the HMG boxes), which act in comparative independence to cause DNA binding. DNA binding is augmented by a 102-residue amino-terminal domain that causes xUBF to form dimers. The dimerization domain is bipartite in structure, consisting of two regions with the potential to form amphipathic helices, separated by a gap of at least 22 amino acids. The carboxyl half of xUBF is relatively dispensable for transcription (including an 87-residue acidic tail). However, either altering the number of HMG boxes or interfering with dimerization eliminates transcription. The gap region of the dimerization domain is dispensable for dimerization but is absolutely required for transcription. This suggests that the gap region has a critical function in transcription distinct from any effect on dimerization or DNA binding.

[Key Words: Ribosomal genes; HMG box protein; transcription factor, DNA-binding protein]

Received July 16, 1991; revised version accepted August 27, 1991.

Fractionation of vertebrate transcription systems indicates that at least two distinct protein factors are required for accurate initiation by RNA polymerase I in vitro, in addition to the polymerase itself. The first of these fractions to be recognized is an activity that elutes at high salt from negatively charged ion exchange columns, is involved in formation of the stable initiation complex, and appears to be a major determinant of the species specificity of polymerase I transcription. A factor with these characteristics has been given different names in different laboratories, including factor D (Mishima et al. 1982; Tower et al. 1986), TIF-IB (Schnapp et al. 1990), and SL1 (Learned et al. 1985; Smith et al. 1990). In the Xenopus system we have given it the designation Rib1 (McStay et al. 1991). At present it is a reasonable hypothesis that all of these factors are functionally related to each other, however, it remains to be determined how many polypeptides are actually covered by these designations.

The second factor known to be required for initiation by polymerase I is upstream binding factor (UBF) [Bell et al. 1988; Pikaard et al. 1989]. cDNAs for this factor have been cloned from human [Jantzen et al. 1990], frog [Bachvarov and Moss 1991; McStay et al. 1991], and rat [O'Mahony and Rothblum 1991], and protein expressed from the human and frog cDNAs has been shown to completely replace naturally purified UBF for in vitro transcription. Xenopus laevis UBF (xUBF) has been shown further to be a necessary component of the stable initiation complex [McStay et al. 1991]. This agrees with other data showing that human UBF (hUBF) and SL1 cooperate to produce an extended footprint on the human ribosomal gene promoter (Bell et al. 1988, 1990). Sequencing a cDNA for hUBF revealed the presence of multiple domains, each ~80 amino acids long, which are related to sequences present in the chromosomal proteins HMG1 and HMG2 (Jantzen et al. 1990). Analysis of a series of progressive carboxy-terminal deletions of hUBF indicated that these HMG boxes are involved in the DNA-binding activity of hUBF, and the proposal was made that they are an example of a hitherto unrecognized type of DNA-binding motif. Since that proposal, the HMG box motif has been recognized in several additional proteins, all of which appear to be DNA-binding proteins [Kelly et al. 1988; Gubbay et al. 1990; Kolodrubetz 1990, Sinclair et al. 1990; Parisi and Clayton 1991; Travis et al. 1991; van de Wetering et al. 1991; Waterman et al. 1991].

Sequencing cDNAs for xUBF revealed that it is closely related to the sequence of hUBF, with the major exception of a deletion in xUBF that removes HMG box 3 and the insertion of 22 unrelated amino acids in its place [Bachvarov and Moss 1991, McStay et al. 1991]. This remodeling of the xUBF cDNA can account for the ~12-kD difference in size between the xUBF and hUBF proteins. It is possible that this size difference may also explain why xUBF and hUBF do not substitute for each other during in vitro transcription [Bell et al. 1989]. The xUBF cDNA that we have been studying codes for a protein of 701 amino acids, and its amino acid se-
McStay et al.

sequence is shown in Figure 1. xUBF contains three recognizable HMG box domains that we refer to as HMG boxes 1, 2, and 4 to indicate their relatedness to the corresponding HMG boxes in hUBF. xUBF also has an amino-terminal domain of 102 amino acids, a highly acidic carboxy-terminal tail (87 amino acids), and a region of 242 amino acids between the HMG boxes and the acidic tail. In this paper we show that DNA binding is restricted to the HMG boxes while the amino-terminal domain is essential for dimerization. Both DNA binding and dimerization are essential for transcriptional activity. The function of the carboxyl terminus is less clear because large portions of the carboxyl terminus can be deleted while retaining partial transcriptional activity.

Results

Construction of xUBF mutants

To mutagenize the xUBF cDNA in a systematic manner, we have used oligonucleotide-directed mutagenesis to insert novel BglII restriction sites into the coding region of the cDNA as described in Materials and methods. These BglII sites are designated A-L, and their relative locations are shown in Figure 1. Within the amino-terminal domain (prior to the first HMG box) BglII sites were spaced at ~20-amino-acid intervals. In the rest of the molecule the BglII sites were spaced more widely at ~100-amino-acid intervals.

Wild-type and mutant forms of xUBF were translated in vitro in a rabbit reticulocyte lysate. We have shown previously that in vitro translation of an xUBF cDNA yields active protein which, in turn, will stimulate a reconstituted, xUBF-dependent in vitro transcription system [McStay et al. 1991]. We have also shown that in vitro translation of active xUBF is stimulated 5- to 10-fold by insertion of a 586-nucleotide cap-independent translation enhancer (CITE) sequence from encephalomyocarditis virus [Parks et al. 1986; Elroy-Stein et al. 1989] immediately upstream of the first AUG of the xUBF open reading frame.

In Figure 1 the amino-terminal six amino acids are underlined to indicate that they are changed from the sequence of wild-type xUBF as a result of the use of the CITE sequence. (Use of the CITE sequence causes the addition of four extra amino acids on the amino terminus and the mutagenesis necessary to attach the CITE sequence changes the second xUBF amino acid from asparagine to aspartic acid.) In vitro-translated xUBF is transcriptionally active in spite of these changes [McStay et al. 1991].

![Amino acid sequence of xUBF](image)

**Figure 1.** The amino acid sequence of xUBF. The sequence shown is the result of conceptual translation of a full-length xUBF cDNA [cloning was reported previously [McStay et al. 1991]]. The first six amino acids are indicated by single underlines to indicate that they derive from attachment of the CITE translation enhancer upstream of the xUBF-coding region. Pairs of amino acids with double underlines indicate the locations where novel BglII restriction sites were introduced into the cDNA (labeled A-L). Introduction of a BglII site causes the two underlined amino acids to be changed to RS (note that in many cases, one or both of the amino acids were not changed from the wild-type sequence by this alteration). Numbering of the amino acids is the same as for wild-type xUBF (and thus ignores the additional amino-terminal residues added by the CITE sequence). Regions labeled HMG box 1, 2, and 4 refer to regions of amino acid sequence similarity to domains in the chromosomal protein HMG1 [Jantzen et al. 1990]. The numbering indicates relatedness to corresponding HMG boxes in hUBF (xUBF is deleted for HMG box 3; McStay et al. 1991).
In the body of the xUBF sequence (Fig. 1) pairs of amino acids are indicated by double underlines at the locations where BglII sites were inserted into the cDNA. In all cases, insertion of these BglII sites changed the underlined pair of amino acids to arginine-serine (RS). Note that in some cases this change altered the wild-type sequence and in others it did not. With one exception, all of the amino acid changes introduced by the novel BglII sites are transcriptionally neutral (data not shown). The single exception is the BglII site introduced at location B. The effect of this mutation and its implications will be discussed below in more detail.

With the BglII restriction sites in place, a variety of amino-terminal, carboxy-terminal, and internal deletions of xUBF were constructed. The schematic structures of all of these constructs are summarized in Figures 7 and 8, below.

The amino-terminal domain of xUBF and the HMG boxes are essential for transcriptional activity

Figure 2A shows the effect of various amino- and carboxy-terminal deletions on the ability of xUBF to reconstitute transcription in an in vitro assay. Figure 2B shows that each of these mutants was translated at a similar efficiency by the reticulocyte lysate; and Figure 2C shows the schematic structure of each of the mutants tested. Because the reticulocyte translation system is partially inhibitory to transcription, the assay is only semiquantitative. However, by assaying two different volumes of translation reaction for their ability to stimulate in vitro transcription we were able to place each mutant on an activity scale ranging from — (no detectable activity) to +++++ (wild-type activity). The relative transcriptional activity of all the xUBF mutants we have tested is summarized in Figures 7 and 8, below.

The carboxyl half of xUBF is relatively dispensable for transcriptional activity. For example, removal of the acidic tail from the carboxyl terminus only reduces transcriptional activity by about fourfold (mutant 7, Fig. 2A). Deletion up to the HMG boxes (mutant 9, Fig. 2A) reduces transcription further, but ~10% residual activity is still observed at low xUBF input. Although the carboxyl half of xUBF (from BglII site I to L) clearly contributes to transcriptional activity, its presence is not absolutely essential. In contrast to the carboxyl half of xUBF, an amino-terminal deletion [mutant 1, Fig. 2A] eliminates transcription completely. Further deletions into the HMG boxes are also transcriptionally inactive [mutants 2–5, summarized in Fig. 7, below].

The effects of various internal deletions of xUBF on transcription are shown in Figure 3A (as shown previously, Fig. 3B illustrates that each of the mutants translated with similar efficiency and Fig. 3C shows the schematic structure of each mutant). Internal deletion of the entire region between the acidic tail and the HMG boxes (mutant 21) has at most a twofold down effect on transcription. Coupling this result with the result of deleting the acidic tail (mutant 7, Fig. 2A) leads us to conclude that the entire carboxyl half of the molecule, up to the HMG boxes, is relatively dispensable for transcription. In contrast, deletion of any one of the HMG boxes essentially abolishes transcriptional activity [Fig. 3A, mutants 13, 14, or 15].

DNA-binding activity is restricted to the HMG boxes

One possible mechanism by which deletion mutants can affect transcriptional activity adversely is by interfering with the ability to bind to DNA. We therefore assayed a variety of xUBF deletion mutants to determine which regions of the molecule are responsible for DNA-binding activity. The DNA-binding assay that we used involved labeling xUBF by in vitro translation with radioactive
crease in the DNA-binding ability of mutant 1. Further deletion into the HMG boxes (mutants 2, 3, and 4) abolishes DNA binding. Assay of various internal deletions (Fig. 4C) shows that DNA binding remains as long as one HMG box remains (mutants 13–17). However, removal

methionine as a precursor. The translation reaction was then passed over a DNA column, containing ribosomal gene promoter sequences, under standard conditions. xUBF is a relatively weak and nonspecific DNA-binding protein, and under optimal conditions only about half of the wild-type protein binds to the column. For purposes of the assay, this amount of binding was scored as plus [+]. In general, binding as detected by this assay was either plus (+) or minus (−) with little gradation in between. Examples of the DNA-binding assay as applied to various xUBF deletion mutants are shown in Figure 4.

For deletions from the carboxyl terminus (Fig. 4A) we obtain the same result as was reported previously for hUBF (Jantzen et al. 1990). DNA-binding activity remains as long as one HMG box is retained (mutant 11) and the amino-terminal 102 amino acids have no DNA-binding activity when all HMG boxes are removed (mutant 12). Starting from the other end, DNA-binding activity remains when the amino terminus is removed (mutant 1, Fig. 4B). However, in several repetitions of this experiment we have observed a reproducible de-
of all of the HMG boxes eliminates DNA binding [mutant 18].

The results presented in Figure 4 leave us with a paradox. When all three HMG boxes are retained, the amino terminus is not needed for DNA binding. However, if one or two of the HMG boxes is deleted, the amino-terminal 102 amino acids must also be present for DNA binding to occur (cf. mutant 2 with mutants 13–17). This implies that the amino terminus aids DNA binding but is not essential if all three HMG boxes are retained. To explain these results we hypothesize that the amino terminus contains a dimerization domain. According to this view, one or two HMG boxes, by themselves, have a DNA affinity too weak to be detected by our assay. Dimerization, however, brings together enough HMG boxes so that the overall DNA affinity is above the threshold that our method can detect.

We present evidence below that this model is correct. These results also indicate that HMG boxes 1 and 4 are approximately equivalent in DNA-binding potential (HMG box 2 has not been tested separately as yet).

**xUBF forms homodimers**

xUBF, purified from *X. laevis* kidney cells, is readily cross-linked by glutaraldehyde to form an aggregate approximately the size of a dimer [Fig. 5A]. As we have noted previously, at least three closely spaced bands are visible when natural xUBF is electrophoresed on an SDS–polyacrylamide gel with apparent molecular masses ranging from ~82 to 87 kD [McStay et al. 1991]. These three bands are also visible in the untreated xUBF electrophoresed in Figure 5A, lane 6. At the lowest level of cross-linking (Fig. 5A, lane 2) some monomer xUBF is still visible, but most of the xUBF appears as a broad band of nearly 190 kD apparent size. With increasing times of cross-linking (lanes 2–5) the monomer bands disappear and all of the material appears in the broad high-molecular-weight band. The apparent size of this high-molecular-weight band decreases progressively with increasing times of cross-linking, running at ~160 kD after the longest time of glutaraldehyde treatment. We presume that the decrease in mobility of the high-molecular-weight material is a result of increased compaction as more cross-links are introduced into the protein. Both the size of the material cross-linked by glutaraldehyde and the fact that nothing of intermediate size is visible at low levels of cross-linking suggest that xUBF associates into dimers. We also do not detect any significant amount of material indicative of higher order multimers.

To verify that the amino-terminal domain of xUBF is required for dimerization, as suggested by DNA-binding studies, both wild-type xUBF and mutant 1 (lacking most of the amino terminus) were separately translated in vitro, renatured after isolating them from an SDS–polyacrylamide gel, and re-electrophoresed either with or without glutaraldehyde treatment. As shown in Figure 5B, wild-type xUBF electrophoreses as a single band in the absence of cross-linking while about half the material runs as a dimer following cross-linking (cf. lanes 1 and 2). When most of the amino terminus is deleted, as in mutant 1, only a single band is seen with or without cross-linking (cf. lanes 3 and 4). In this experiment we were able to drive only about half of the wild-type xUBF into dimers, and the residual monomer material of both wild-type and mutant electrophoreses in a faster broad band after glutaraldehyde treatment. Lack of 100% dimerization is probably due to a failure of all of the protein to renature properly following denaturation. The higher mobility of the monomer material is probably due to the same compaction phenomenon that causes the dimer material to change mobility with increasing glutaraldehyde treatment [Fig. 5A].

*The amino terminus of xUBF contains a dimerization motif*

To further characterize the region of xUBF responsible for dimerization we used an immunoprecipitation assay for dimerization. Two copies of an epitope tag from the *myc* oncogene [Munro and Pelham 1987] were inserted tandemly into the *BgIII* site at position I (on the carboxy-
terminal side of HMG box 4, see Fig. 6B). The particular myc epitope used as a tag has the amino acid sequence MEQKLISEEDLN. Insertion of the myc tag at this position has no effect on either transcriptional activity or DNA binding (data not shown), and the native molecule is readily precipitated with a monoclonal antibody (Evan et al. 1985) directed against the epitope [shown in Fig. 6].

Synthetic mRNA for myc-tagged xUBF was mixed with a twofold excess of mRNA for xUBF that had no myc tag but was deleted for various regions. The two mRNAs were cotranslated in vitro, the labeled translation products were precipitated with antibody against the myc tag, and the precipitated material was electrophoresed on an SDS–polyacrylamide gel. As shown in Figure 6, approximately equal amounts of tagged and untagged xUBF were coprecipitated in every case except when the untagged version was lacking the amino-terminal domain (mutant 1). In this case, only the tagged xUBF was precipitated. We have also determined that a mutant lacking all three HMG boxes (Fig. 4, mutant 18) precipitated along with tagged xUBF (data not shown). We conclude that the amino-terminal domain is necessary and sufficient for dimerization of xUBF.

Figure 6. Assay of mutants of xUBF for the ability to dimerize using a communoprecipitation assay. [A] Two copies of an epitope tag from the myc oncogene were inserted at position 1 of wild-type xUBF. The tagged xUBF was translated in vitro and was shown to be readily precipitated by a monoclonal antibody directed against the myc epitope [lane labeled none]. In each of the succeeding lanes, myc-tagged wild-type xUBF was cotranslated along with a different internal deletion mutant of xUBF. All of the mutants coprecipitated with wild-type xUBF except for mutant 1 [lacking the amino-terminal domain]. In other experiments [data not shown] neither wild-type xUBF nor any of the mutants was precipitated by the anti-myc antibody when translated in the absence of tagged xUBF. [B] Structure of the mutants tested in A.

The xUBF dimerization domain consists of two potential amphipathic helices, separated by a region essential for transcriptional activation

DNA-binding, transcriptional, and dimerization activity of xUBF mutants is summarized in Figure 7. Because deletion of most of the amino terminus eliminates both dimerization and transcriptional activity (see mutant 1 in Figs. 2A, 5B, and 6A), we wished to examine this region in more detail. For this purpose we utilized BglII sites A–F [described in Fig. 1] to construct a finer set of deletions and assayed them for both dimerization and transcriptional activity. The transcriptional activity of various mutants of the amino terminus is shown in Figure 8A, the translation efficiency of each of the mutants is shown in Figure 8B, the ability of selected mutants to coprecipitate with myc-tagged xUBF is shown in Figure 8C, and Figure 8D gives the structure of each mutant plus a summary of the transcription and dimerization assays.

Examination of the coprecipitation data in Figure 8C shows that mutant 24 defines one region required for dimerization while the contiguous mutants 26 and 27 define another required region. Separating these two regions are at least 22 amino acids (defined by mutant 25), which are dispensable for dimerization. The fact that two patches of amino acids are required and are separated by a dispensable stretch of 22 amino acids immediately calls to mind the helix–loop–helix [HLH] dimerization motif that has been observed in a large family of RNA polymerase II transcription factors (Murre et al. 1989), as well as the helix–span–helix motif that has been reported in the polymerase II enhancer-binding pro-
Figure 8. Transcription and dimerization activities in the amino terminus of xUBF. (A) Assay of amino-terminal mutations for transcriptional activity. A series of finer deletions within the amino terminus of xUBF was constructed (shown in part D, below). Each mutant was translated in vitro and tested for its ability to reconstitute in vitro transcription as described in Fig. 2A. A–F refer to assays of mutants created when each of the novel BgII sites was inserted into the amino-terminal domain. Note that all of the BgII insertions are transcriptionally neutral except for insertion at site B, which reduces transcription ~50%. Of the deletion mutants, deletion of the extreme amino terminus (mutant 23) also reduces transcription by 50%. All of the other deletions abolish transcriptional activity. (B) Translation efficiency of each of the mutants assayed in A. (C) Coimmunoprecipitation assay of each of the amino-terminal deletion mutants. Each mutant was cotranslated with myc-tagged wild-type xUBF and tested for its ability to coprecipitate as described in Fig. 6A. Note that mutants 24, 26, and 27 fail to precipitate and are therefore presumably unable to dimerize with wild-type xUBF. (D) Amino acid sequence of the mutants assayed in A and C. The transcriptional and dimerization activity of the mutants is summarized at right.

There is no apparent sequence similarity between the xUBF dimerization domain and either of the aforementioned polymerase II factor motifs. However, we do note that amino acids between BgII sites B and C, as well as those between sites D and F, have the potential of forming amphipathic helices (see Fig. 9A,B), a type of structure that has been implicated in the HLH and helix–span–helix motifs as well as the bZIP family of leucine zipper proteins. We will return to this point in the Discussion (below).

Assaying mutations in the A–F region for transcriptional activity leads to the conclusion that dimerization is essential for transcriptional activity (Fig. 8A). Deletion of any part of the dimerization domain [mutant 24, 26, or 27] completely eliminates transcription. Mutant 23, which is missing the amino-terminal 20 amino acids, exhibits about a 50% decrease in transcription. However, insertion of a BgII site alone at position B also reduces transcription by the same amount. Thus, it is likely that the BgII site inserted at position B lies on the boundary of the dimerization motif, and most (perhaps all) of the first 20 amino acids are not required for either transcription or dimerization.

The surprising result is that mutant 25, which lacks 22 amino acids within the dimerization domain, is completely competent for dimerization with wild-type xUBF [Fig. 8C] or for formation of homodimers (data not shown) but lacks all transcriptional activity (Fig. 8A). This result suggests that the 22 amino acids in the C–D region might affect transcriptional activation directly by some mechanism other than dimerization. We have performed two additional experiments that support this possibility. In mutant 28 (Fig. 8D), 10 amino acids between BgII sites C and D were altered, but the spacing between the dimerization regions was left unchanged. Mutant 28 dimerizes as well as wild-type xUBF but is completely inactive in transcription (data not shown). In a different approach we reasoned that if the only function of the terminal 102 amino acids was to cause dimer formation, we might be able to replace the amino terminus with an unrelated dimerization motif and still retain transcriptional activity. Accordingly, we removed the
It is becoming clear that many transcription factors are DNA binding or dimerization or both. Bachvarov and Moss (1991) recently published an analysis of the amino acid sequence of xUBF in which they showed that the region between HMG box 4 and the acidic tail (between BglII sites I and K in Fig. 1) contains two more elements with detectable similarity to the HMG box motif. Thus, they interpret xUBF as having a total of five HMG boxes. It seems possible that the entire body of xUBF (exclusive of the amino-terminal dimerization domain and the carboxy-terminal acidic tail) arose via tandem repetition of HMG box motifs. However, it appears that the two elements located in region I–K have diverged so far as to lose any detectable function.

The acidic tail of xUBF contains a novel helix–gap–helix dimerization motif

The role of the amino terminus of xUBF can be seen more clearly. Within the first 102 amino acids, prior to the first HMG box, we have identified a bipartite motif that functions to associate the protein into dimers. The acidic tail performs a function that is either not required or is not rate limiting in vitro. For example, we know that the repetitive enhancer elements located in the intergenic spacer region strongly influence transcription in vivo but their requirement does not show up in most in vitro assays (Kuhn et al. 1990, Pikaard et al. 1990a). xUBF binds to the enhancer elements (Pikaard et al. 1989), and a plausible speculation is that the acidic tail has some role in enhancer function that is not absolutely required for basal promoter activity.

The function of the region between the HMG boxes and the acidic tail is even less clear, because deletion of this entire region (Fig. 7, mutant 21) only decreases transcription by twofold and has no effect on either dimerization or DNA binding. Bachvarov and Moss (1991) recently published an analysis of the amino acid sequence of xUBF in which they showed that the region between HMG box 4 and the acidic tail (between BglII sites I and K in Fig. 1) contains two more elements with detectable similarity to the HMG box motif. Thus, they interpret xUBF as having a total of five HMG boxes. It seems possible that the entire body of xUBF (exclusive of the amino-terminal dimerization domain and the carboxy-terminal acidic tail) arose via tandem repetition of HMG box motifs. However, it appears that the two elements located in region I–K have diverged so far as to lose any detectable function.

The acidic tail of xUBF is not required for in vitro transcription

It is becoming clear that many transcription factors are arranged in a modular fashion with multiple distinct domains that function in relative independence from each other. As we have shown in this paper, xUBF is constructed with similar modularity, and we have identified domains for several discrete functions including dimerization and DNA binding.

Figure 9. Models of xUBF structure. (A) Helical wheel model of amino acids 23–43. The drawing presents the view down the axis of the presumed α-helix. On the assumption that the amino acids in this region form α-helices, which assume a coiled-coil configuration in the dimer, every seventh amino acid is shown lining up atop one another. The shaded area calls attention to a face of the helix that is rich in hydrophobic amino acids and presumably could be involved in dimer interactions. (B) Helical wheel model of amino acids 66–100, drawn with the same assumptions used in A. The shaded area calls attention to a hydrophobic ridge of amino acids that could be involved in dimer formation. (C) Model of xUBF drawn as a dimer and showing the various functional domains. The dimerization motifs are shown interacting in a parallel fashion, but no direct evidence for or against this arrangement is yet available.

amino terminus from xUBF and fused the HLH motif from the mouse protein, E47 (Murre et al. 1989), to the BglII site at position F. This chimeric protein readily forms dimers but is completely inactive in transcription (data not shown).

Considering all of the data in Figure 8 we come to two conclusions: (1) Dimerization is essential for transcriptional activity of xUBF; (2) the 22 amino acids in the amino-terminal C–D region are essential for transcriptional activation via a mechanism distinct from either DNA binding or dimerization.

Discussion

The acidic tail of xUBF is not required for in vitro transcription

It is becoming clear that many transcription factors are
xUBF dimerization motif for possible relationships with any of the protein dimerization motifs described previously such as the bZIP [Landschulz et al. 1988] or HLH families [Murre et al. 1989], or the helix–span–helix motif suggested for AP-2 [Williams and Tjian 1991], and can detect no obvious similarities with the primary sequence. However, if we assume that the dimerization domains of xUBF are arranged as coiled coils, both dimerization domains have the potential of forming an amphipathic helix (diagramed in Fig. 9A,B). For amino acids 23–43, three ridges of the helix are composed predominantly of hydrophobic amino acids and are all grouped on one side [Fig. 9A]. For amino acids 66–100, one ridge of the helix is hydrophobic throughout four heptad repeats. Because amphipathic helices with hydrophobic ridges have been implicated in the bZIP, HLH, and helix–span–helix dimerization motifs, it is possible that amphipathic helices play a similar role in xUBF. We propose that this new class of dimerization domain be called the helix–gap–helix motif.

The gap in the helix–gap–helix motif is essential for transcriptional activation

Transcriptional activity of xUBF appears to be completely dependent on dimerization. Any of the deletions that interfere with dimerization completely abolish the ability to stimulate a reconstituted in vitro transcription system [see Fig. 8]. In this regard, xUBF is similar to the bZIP and HLH classes of transcription factors whose activity also depends upon dimerization. The unexpected result, however, is that mutations in the gap region of xUBF also eliminate transcription but have no detectable effect on dimer formation. This result is obtained whether 22 amino acids are deleted from the gap [Fig. 8D, mutant 25] or 10 amino acids within the gap are altered while leaving the gap spacing unchanged [Fig. 8D, mutant 28]. The amino acids in the gap region must be quite flexible for a heterodimer to form the deletion [mutant 25] and the wild-type partner.

We have tried to replace the dimerization function of xUBF by substituting the HLH motif from the mouse protein E47 [Murre et al. 1989; the HLH motif was fused to xUBF at position F and replaced the first 102 amino acids]. This chimeric protein readily forms dimers and binds to DNA but is completely inactive in transcription [Fig. 7]. Protein made from this mutant was preincubated in a complete transcription reaction [lacking only nucleotides and Mg2+] for 10 min before adding wild-type xUBF and beginning transcription. We know from previous work that a nonexchangeable stable initiation complex can form between wild-type xUBF, Rib1, and the promoter under these preincubation conditions [McStay et al. 1991]. Nevertheless, preincubating with mutant xUBF had no inhibitory effect on transcription [data not shown]. Deletion mutants of xUBF, deficient in other parts of the molecule, were similarly tested with no inhibitory effect on transcription. There are several possible explanations for such negative results, and they do not rule out that interaction between the gap region and other components of the transcription machinery.

HMG boxes and DNA binding

The HMG boxes constitute a second modular component of xUBF. The deletion experiments summarized in Figure 7 suggest that each HMG box is probably an independent DNA-binding entity, but several boxes must be present on the same molecule to achieve a binding affinity above the threshold that our DNA-binding assay can detect. To a first approximation, a sufficient number of HMG boxes can be gathered together either by arranging them in tandem in one primary amino acid chain or by linking them together via the helix–gap–helix dimerization motif. Again, there is a parallel between xUBF and the bZIP or HLH proteins in that dimerization serves to bring DNA-binding motifs together. However, xUBF differs from the bZIP and HLH proteins in that DNA binding is readily detected even in the absence of dimerization as long as sufficient HMG boxes remain in the monomer.

The structural information presented in this paper suggests explanations for some previously puzzling observations concerning the way in which xUBF binds to DNA. hUBF was initially presented as a sequence-specific DNA-binding protein that protected a consensus sequence in the upstream region of the human ribosomal gene promoter [Bell et al. 1988]. Since that time, as more and more binding sites for UBF have been identified, it has become apparent that any sequence specificity that UBF may have is very relaxed. It is clear that UBF strongly prefers GC-rich DNA over AT-rich DNA. Beyond that, a convincing consensus is difficult to derive. Furthermore, on X. laevis rDNA, xUBF binds over a very extensive region. We have observed DNase I protection extending from around the transcription start site at +1 back to the S' boundary of the gene promoter. In addition, xUBF protects the region between the gene promoter and the T3 terminator, on both sides of the terminator and over all of the enhancer elements [Pikaard et al. 1989]. Thus, xUBF footprints over several kilobases, at least, on X. laevis rDNA.

In this paper we have shown that an xUBF dimer contains six DNA-binding domains [the HMG boxes] that
function with considerable autonomy. The HMG boxes differ from each other in amino acid sequence [these differences have been largely conserved between *Xenopus* and human UBFs], and it seems possible that each type has a different preferred DNA-binding sequence. If this speculation is correct, the DNA sequence specificity of intact xUBF is the sum of the specificities of each HMG box and it is no surprise that an overall consensus has been difficult to detect.

We have also noted that the DNase I footprint of xUBF is relatively resistant to mutagenesis of the DNA (Piakaard et al. 1989). Clustered point mutations usually alter the footprint just at the site of mutation with little or no effect on either side of the mutation. This result is readily understood if we view xUBF as a molecule with six independent DNA-binding domains. A mutation that alters the ability of one HMG box to bind probably has little effect on the binding of the other five HMG boxes.

**HMG boxes and transcription**

It is puzzling that removal of a single HMG box abolishes transcription even though dimerization and DNA binding appear unaffected [Fig. 7, mutants 13–15]. One possible explanation is that each HMG box contains two functions, one for DNA binding and one causing transcriptional activation. If this explanation were correct we would have to hypothesize further that the transcription function is much more sensitive to deletion than is DNA binding.

A second possibility, which seems more likely to us, is that the DNA-binding activity of the HMG boxes determines a critical spacing of domains within the promoter. If one or two HMG boxes are removed, DNA binding remains but xUBF can no longer position the promoter so that Rlb can recognize and interact with it; consequently, transcription initiation does not occur. As evidence for this second possibility, we note that duplicating one of the HMG boxes, so that a total of four boxes are present, is as deleterious to transcription as deleting a box. [The data for this experiment are not shown; however, the schematic structure of the mutant, 22, is included in the summary shown in Fig. 7]. The mammalian UBFs that have been examined so far all have four HMG boxes [Jantzen et al. 1990; O’Mahony and Rothblum 1991] and are relatively interchangeable [Bell et al. 1990]. In contrast, xUBF has only three HMG boxes [Bachvarov and Moss 1991; McStay et al. 1991] and appears to substitute either not at all [Bell et al. 1989] or only partially [Piakaard et al. 1990b] for mammalian UBF.

We conclude from this analysis that frog and mammalian ribosomal gene promoters have different spacing of their promoter domains and their respective UBFs are adjusted to this spacing by virtue of having different numbers of HMG boxes. In support of this interpretation we note that a frog ribosomal gene promoter can be transformed into a super mouse promoter by altering the spacing between its upstream and core promoter domains [Pape et al. 1990].

**Materials and methods**

**Extracts, fractionation, and transcription**

S-100 extracts were prepared from the *X. laevis* cell line, XlK-2, as described previously [McStay and Reeder 1990]. S-100 extract was chromatographed successively on DEAE–Sephacel CL6B [Pharmacia], heparin Ultrogel (IBF Biotech), and Bio-Rex 70 [200–400 mesh, Bio-Rad] exactly as described [McStay et al. 1991], to generate Rlb, RNA polymerase I, and xUBF fractions that were not cross-contaminated. Transcription reactions were programmed with [McStay et al. 1991]. Each reaction contained 10 μl of RNA polymerase 1, 10 μl of Rlb1, and 1 or 3 μl of in vitro translation reactions programmed with either wild-type or mutant xUBF RNAs. The template in each reaction was 200 ng of UV-treated pGEM-40 [McStay and Reeder 1990].

**Mutagenesis**

The plasmid used for in vitro translation of xUBF and for subsequent mutagenesis was pxUBF CITE. pxUBF CITE contains the CITE from encephalomyocarditis virus fused upstream of the complete xUBF open reading frame. This fusion was cloned into the vector pBluescript SK + [Stratagene] in the orientation that the promoter for phage T7 RNA polymerase can be used to generate synthetic xUBF mRNA for in vitro translation. A detailed description of this plasmid is given elsewhere [McStay et al. 1991]. Each reaction contained 10 μl of RNA polymerase 1, 10 μl of Rlb1, and 1 or 3 μl of in vitro translation reactions programmed with either wild-type or mutant xUBF RNAs. The template in each reaction was 200 ng of UV-treated pGEM-40 [McStay and Reeder 1990].

**In vitro transcription and translation**

The plasmid pxUBF–CITE and its mutant derivatives were linearized at the unique *XbaI* restriction site in the polylinker downstream of the xUBF, transcribed with phage T7 RNA polymerase, and translated in a rabbit reticulocyte lysate as described previously [McStay et al. 1991]. For carboxy-terminal deletions, xUBF CITE derivatives were linearized with *BglII*. We have calculated that ~100 ng of full-length xUBF is synthesized in a standard translation reaction.

**DNA-binding assay**

A promoter DNA affinity column was prepared as follows. The plasmid pGEM-40-16, which contains 16 copies of the *X. laevis* promoter (~245 to +40) in a tandem array [gift of M. Schultz] in the vector pGEM-4 [Promega], was linearized with HindIII and cross-linked with cyanogen bromide to Sepharose CL-2B [Pharmacia] by using the method of Kadonaga and Tjian [1986]. [3H]Methionine-labeled xUBF was prepared by using a standard in vitro translation reaction [50 μl]; then 200 μl of affinity column buffer containing 100 mM KCl [ACB 100; 25 mM HEPES [pH 7.9], 100 mM KCl, 5 mM MgCl2, 1 mM DTT, 20% glycerol] was added along with 1 μg of poly[d(A-T)] and incubated on ice for 15 min. Then 200 μl of the diluted translation reaction was applied to a 200-μl promoter–DNA affinity column. The column was washed successively with ACB 100 [2× 1 ml], and ACB 200 [500 μl], and the bound protein was eluted from the column with ACB 600 [2× 200 μl]. A 1-μl aliquot of the diluted translation reaction was electrophoresed on an SDS–polyacrylamide gel [Laemmli 1970] alongside a 4-μl aliquot of the eluted
material. After fixing in 40% methanol/10% acetic acid, the gels were dried and autoradiographed. Note that equality in signal between bound and unbound material reflects 50% binding to the column.

Eptope tagging and immunoprecipitation

The following oligonucleotide was synthesized: 5'-GATC-CATGGAGCATAAGGCTTCTTTACGAGACCTTGAAT-A-3'. This was then annealed to a second oligonucleotide with a complementary sequence designed to make the entire sequence double stranded except for a 5' overhang (GATC) at both ends. The double-stranded oligonucleotide codes for the amino acid sequence of an epitope from the c-myc proto-oncogene that is recognized by monoclonal antibody 9E10 (Evans et al. 1985). Two tandem copies of this oligonucleotide were cloned into the BglII site at position I (see Fig. 6B). Immunoprecipitations were performed with 10 μg of purified mAb 9E10 per 50 μl of in vitro translation reaction and 15 μl of protein G-Sepharose Fast Flow (Pharmacia). Immune complexes were washed four times in 0.5-ml aliquots of PBS and were resolved on a 10% SDS-polyacrylamide gel (Laemmli 1970).

Glutaraldehyde cross-linking

In vitro-translated, 35S-labeled xUBF was partially purified after translation by electrophoresis on a 7.5% SDS–polyacrylamide gel. Following autoradiography of the wet gel, the protein band was excised and crushed with a Teflon pestle in a siliconized 2-ml microcentrifuge tube. The protein was eluted from the gel at room temperature for 3 hr in a buffer containing 50 mM Tris-HCl (pH 7.9), 0.1 mM EDTA, 5 mM DTT, 0.1% SDS, and 150 mM NaCl. After removing the polyacrylamide by centrifugation, protein was precipitated with 5 volumes of acetone at −20°C for 10 min and the precipitate was collected by centrifugation at 14,000 g for 10 min. Protein was resuspended in 50 μl of 20% glycerol, 25 mM Hepes (pH 7.9), 0.2 mM EDTA, 0.2 mM PMSF, 100 mM KCl, 1 mM DTT, and 6 mM guanidinium–HCl for 30 min at room temperature. The sample was then passed over a 400-μl column of Sephadex G-25 equilibrated in the same buffer but lacking guanidinium–HCl to remove the denaturant. Protein was allowed to renature overnight at 4°C. Nine microliters of 35S-laabeled xUBF purified by this method or 9 μl of xUBF purified by column chromatography from tissue culture cells (McStay et al. 1991) was incubated in either 0.05% or 0.01% glutaraldehyde (Sigma) at room temperature by adding 1 μl of a 10× stock in water. Reactions were quenched by adding lysine to 30 mM. An equal volume of 2× sample buffer was added, and the sample was boiled 5 min before separation on a 10% SDS–polyacrylamide gel.

Acknowledgments

We thank Judith Roan for technical assistance and Keith Blackwell, Mike Krause, Steve Hahn, and Hal Weintraub for helpful discussions. Mark Roth provided materials for epitope tagging. This work was supported by grant GM26624 to R.H.R. and a National Research Service Award [NRSA] fellowship to MWF.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

Bachvarov, D. and T. Moss. 1991. The RNA polymerase I transcription factor xUBF contains 5 tandemly repeated HMG homology boxes. Nucleic Acids Res. 19: 2331–2335.

Bell, S.P., R.M. Learned, H.-M. Jantzen, and R. Tjian. 1988.

xUBF domains

Functional cooperation between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. Science 241: 1192–1197.

Bell, S.P., C.S. Pikaard, R. H. Reeder, and R. Tjian. 1989. Molecular mechanisms governing species specific transcription of ribosomal RNA. Cell 69: 489–497.

Bell, S.P., H.-M. Jantzen, and R. Tjian. 1990. Assembly of alternative multiprotein complexes directs RNA promoter selectivity. Genes & Dev. 4: 943–954.

Elroy-Stein, O., T.R. Fuerst, and B. Moss. 1989. Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5’sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. Proc. Natl. Acad. Sci. 86: 6126–6130.

Evans, G.I., G.K. Lewis, G. Ramsay, and J.M. Bishop. 1985. Isolation of a monoclonal antibody specific for human c-myc proto-oncogene product. Mol. Cell. Biol. 6: 3610–3616.

Gubbay, J., J. Collignon, P. Koopman, B. Capel, A. Economou, A. Munsterberg, N. Vivan, P. Goodfellow, and R. Lovell-Badge. 1990. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature 346: 245–250.

Jantzen, H.-M., A. Admon, S.P. Bell, and R. Tjian. 1990. Nucleolar transcription factor xUBF contains a DNA-binding motif with homology to HMG proteins. Nature 344: 830–836.

Kadonaga, J.T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. 83: 5889–5893.

Kelly, M., J. Burke, M. Smith, A. Klar, and D. Beach. 1988. Four mating-type genes control sexual differentiation in the fission yeast. EMBO J. 7: 1537–1547.

Kolodrubetz, D. 1990. Consensus sequence for HMGI-like DNA binding domains. Nucleic Acids Res. 18: 5565.

Kuhn, A., U. Deppert, and L. Grumment. 1990. A 140-base-pair repetitive sequence element in the mouse rRNA gene spacer enhances transcription by RNA polymerase I in a cell-free system. Proc. Natl. Acad. Sci. 87: 7527–7531.

Kunkel, T.A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. 82: 488–492.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685.

Landschulz, W.H., P.F. Johnson, and S.L. McKnight. 1988. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. Science 240: 1759–1764.

Learned, R.M., S. Cordes, and R. Tjian. 1985. Purification and characterization of a transcription factor that confers promoter specificity to human RNA polymerase I. Mol. Cell. Biol. 6: 1358–1369.

Ma, J. and M. Ptashne. 1987. A new class of yeast transcriptional activators. Cell 61: 113–119.

McStay, B., C.-H. Hu, C.S. Pikaard, and R.H. Reeder. 1991. xUBF and Rib1 are both required for formation of a stable polymerase I promoter complex in X. laevis. EMBO J. 10: 2297–2303.

McStay, B., and R.H. Reeder. 1990. An RNA polymerase I termination site can stimulate the adjacent ribosomal gene promoter by two distinct mechanisms in Xenopus laevis. Genes & Dev. 4: 1240–1252.

Mishima, Y., J. Finansec, K. Kominami, and M. Muramatsu. 1982. Fractionation and reconstitution of factors required for accurate transcription of mammalian ribosomal RNA genes: Identification of a species-dependent factor. Nucleic Acids Res. 10: 6659–6670.

Munro, S. and H.R.B. Pelham. 1987. A carboxy-terminal signal
provides secretion of luminal ER proteins. Cell 48: 899–907.
Murre, C., P.S. McCaw, and D. Baltimore. 1989. A new DNA
binding and dimerization motif in immunoglobin enhancer
binding, daughterless, MyoD, and myc proteins. Cell 61: 49–
59.
O’Mahony, D.J. and L.I. Rothblum. 1991. Identification of two
forms of the RNA polymerase I transcription factor UBF.
Proc. Natl. Acad. Sci. 88: 3180–3184.
Pape, L.K., J.J. Windle, and B. Sollner-Webb. 1990. Half-helical
turn spacing changes convert a frog into a mouse rDNA pro-
moter: A distant upstream domain determines the helix face
of the initiation site. Genes & Dev. 4: 52–62.
Parisi, M.A. and D.A. Clayton. 1991. Similarity of human mi-
 tochondrial transcription factor 1 to high mobility group
proteins. Science 252: 965–969.
Parks, C.D., G.M. Duke, and A.C. Palmenberg. 1986. Enceph-
alomycocarditis virus 3C protease: Efficient cell-free expres-
sion from clones which link viral S’ noncoding sequences to
the P3 region. J. Virol. 60: 376–384.
Pikaard, C.S., B. McStay, M.C. Schultz, S.P. Bell, and R.H.
Reeder. 1989. The Xenopus ribosomal gene enhancers bind
an essential polymerase I transcription factor, xUBF. Genes
& Dev. 3: 1779–1788.
Pikaard, C.S., L.K. Pape, S.L. Henderson, K. Ryan, M.H. Paal-
man, M.A. Lopata, R.H. Reeder, and B. Sollner-Webb. 1990a.
Enhancers for RNA polymerase I in mouse ribosomal DNA.
Mol. Cell. Biol. 10: 4816–4825.
Pikaard, C.S., S.D. Smith, R.H. Reeder, and L. Rothblum. 1990b.
rUBF, an RNA polymerase I transcription factor from rats,
produces DNase I footprints identical to those produced by
xUBF, its homolog from frogs. Mol. Cell Biol. 10: 3810–
3812.
Ptashne, M. 1988. How eukaryotic transcriptional activators
work. Nature 335: 683–689.
Schnapp, A., J. Clos, W. Hadel, R. Schreck, A. Cvekl, and I.
Grunmt. 1990. Isolation and functional characterization of
TIF-IB, a factor that confers promoter specificity to mouse
RNA polymerase I. Nucleic Acids Res. 18: 1385–1393.
Sinclair, A.H., P. Berta, M.S. Palmer, J.R. Hawkins, B.L. Grif-
iths, M.J. Smith, J.W. Foster, A.M. Frischchauf, R. Lovell-
Badge, and P.N. Goodfellow. 1990. A gene from the human
sex-determining region encoded a protein with homology to
a conserved DNA-binding motif. Nature 346: 240–244.
Smith, S.D., E. Oriahi, D. Lowe, H.-F. Yang-Yen, D. O’Mahony,
K. Rose, K. Chen, and L.I. Rothblum. 1990. Characterization
of factors that direct transcription of rat ribosomal DNA.
Mol. Cell Biol. 10: 3105–3116.
Tower, J., V. Culotta, and B. Sollner-Webb. 1986. The factors
and nucleotide sequences that direct ribosomal DNA tran-
scription and their relationship to the stable complex. Mol.
Cell Biol. 6: 3451–3462.
Travis, A., A. Amsterdam, C. Belanger, and R. Grosschedl. 1991.
LEF-1, a gene encoding a lymphoid-specific protein, with an
HMG domain, regulates T-cell receptor α enhancer function.
Genes & Dev. 5: 880–894.
van de Watering, M., M. Oosterwegel, D. Dooijes, and H. Clev-
ers. 1991. Identification and cloning of TCF-1, a T lympho-
cyte-specific transcription factor containing a sequence-spe-
cific HMG box. EMBO J. 10: 123–132.
Waterman, M.L., W.H. Fischer, and K.A. Jones. 1991. A thymus-
specific member of the HMG protein family regulates the
human T-cell receptor Ca enhancer. Genes & Dev. 5: 656–
669.
Williams, T. and R. Tjian. 1991. Characterization of a dimer-
ization motif in AP-2 and its function in heterologous DNA-
binding proteins. Science 251: 1067–1071.
xUBF contains a novel dimerization domain essential for RNA polymerase I transcription.

B McStay, M W Frazier and R H Reeder

*Genes Dev.* 1991, 5:
Access the most recent version at doi:10.1101/gad.5.11.1957