The RNA polymerase I transcription factor xUBF contains 5 tandemly repeated HMG homology boxes

Dimcho Bachvarov and Tom Moss*
Centre de Recherche en Cancérologie de l'Université Laval, Hôtel-Dieu de Québec, 11 Côte du Palais, G1R 2J6 Québec, Canada

Received February 7, 1991; Revised and Accepted March 27, 1991
EMBL accession no. X57201

ABSTRACT

The RNA polymerase I transcription factor UBF has been identified in human, mouse, rat and *Xenopus* and the primary structure of the human protein has been determined. Human UBF was shown to contain four tandem homologies to the folding domains of the HMG1 and 2 proteins and hence to belong to a previously unrecognised familly of 'HMG-box' transcription factors. Here, cDNA clones encoding the *Xenopus laevis* UBF (xUBF) have been isolated and sequenced. Northern and Southern blots revealed that in tissue culture cells, xUBF is coded on a single major mRNA size species by a small number of genes. The deduced primary structure of xUBF is highly homologous with the human protein except for a central deletion which removes most of one HMG-box. This explains the major size difference between the *X. laevis* and human proteins and may well explain their different transcriptional specificities. It is shown that xUBF contains 5 tandemly repeated HMG-boxes and that by analogy the human protein contains 6.

INTRODUCTION

The isolation of the protein factors necessary for RNA polymerase I transcription has been undertaken in a wide range of organisms. However since most of these factors are defined only in terms of chromatographic fractions of differing purities, it is extremely difficult to compare the different systems. The human and acanthamoeba systems are probably the best characterised at present, followed closely by those of the mouse and rat (1–13). A few of these studies have recently shown that some factors are conserved at least among vertebrates (2,6,14). It would also appear that the common factors identified all show some degree of species specificity. Thus the assumption that a single factor in each system is solely responsible for the species specificity of ribosomal DNA (rDNA) transcription is probably no longer valid.

UBF and SL1 were first defined in human cell extracts (10,13), where together with the polymerase they reconstituted correct *in vitro* initiation on the human promoter. SL1 carries the major species selectivity observed between mammalian polymerase I promoters, while UBF was found necessary for its correct and efficient binding. In the presence of UBF1, SL1 binds to the upstream control (UCE) and the core elements of the human promoter. UBF has been purified to homogeneity as a protein doublet of about 94 and 97kd, but as yet SL1 has not (2,8). The mouse and rat equivalents of human UBF (hUBF) have also been identified and shown to have activities similar to and exchangeable with those of hUBF (2,14).

The *Xenopus laevis* equivalent of UBF was also recently purified to homogeneity as a doublet of about 82 and 85kd (15), i.e. significantly smaller than hUBF. It was shown to have footprinting activities like those of the human and rat equivalents (6,14). However xUBF will not functionally replace the human protein in human *in vitro* transcription assays and the converse is also true. Unlike the situation on the human promoter, both xUBF and hUBF footprint very extensively throughout the *X. laevis* promoter and XLUBF almost completely protects the spacer enhancers.

A cDNA coding for hUBF has been isolated and shown to code a protein of about 89.4kd having homologies to HMG1 and 2 (16). Along with the sex determination factors (17,18), hUBF defines a new familly of transcription factors. It has also been suggested that HMGs1 and 2 may be synonymous with the polymerase II transcription factor TFIIB (19). The hUBF has been shown to contain 4 tandem homologies to the folded domains of HMG 1 and 2, each HMG protein has two such domains. Further it has been shown that the HMG-boxes of hUBF are involved in DNA binding (16). As in the HMGs, a region of very predominantly aspartic and glutamic residues is present in the C-terminal segment of hUBF. By analogy this almost certainly forms a highly flexible if not totally random coil region in solution.

Here we present the structure of xUBF. We show it to be highly homologous with the human protein with the exception of one HMG-box. We also show that xUBF and by analogy hUBF contain respectively 5 and six tandem HMG-box repeats.

MATERIALS AND METHODS

The EcoRI–BstEII fragment from pSUBF1 (16), kindly provided by M.-H. Jantzen, was used to screen a λgt10 cDNA bank

* To whom correspondence should be addressed

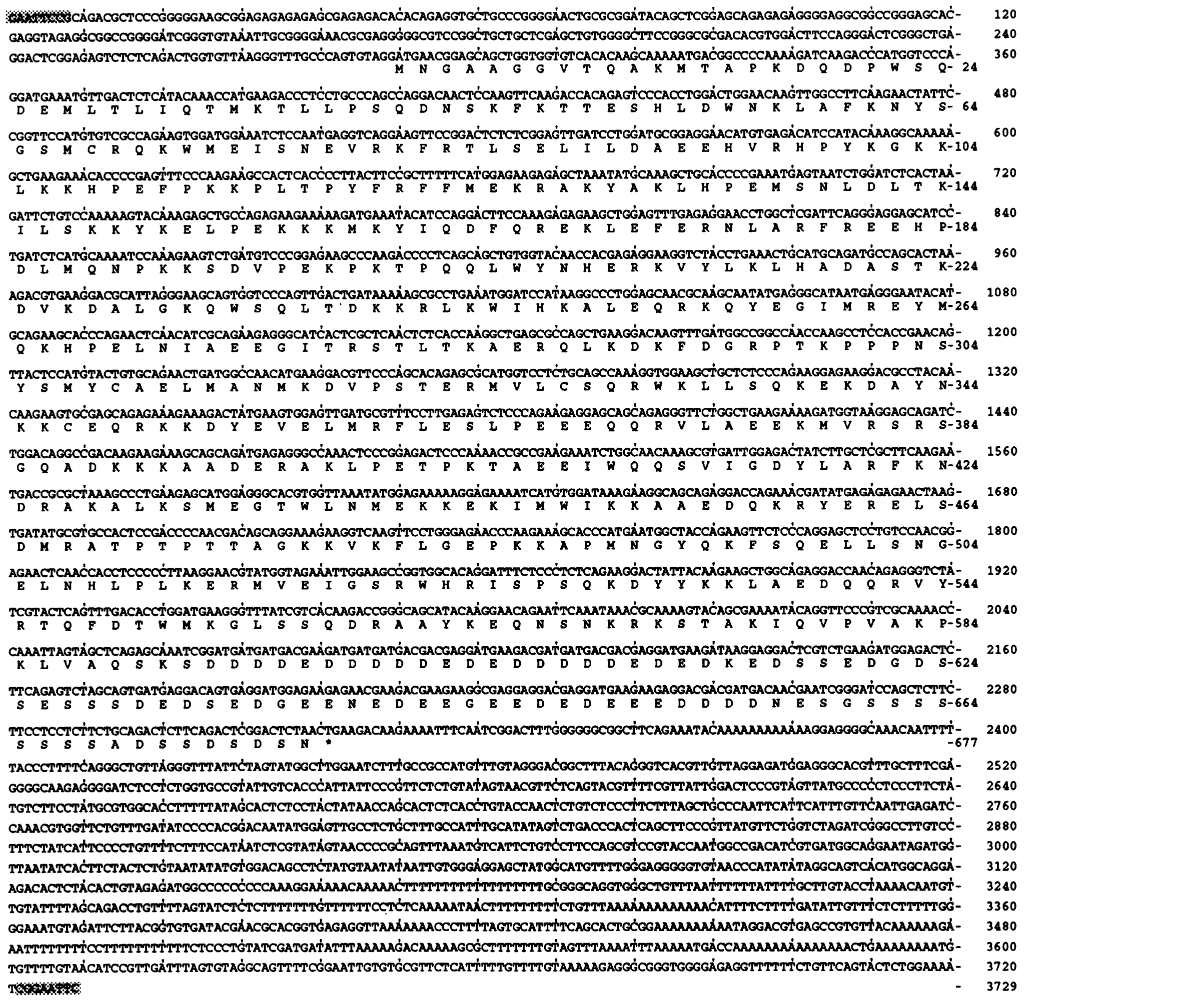

Figure 1. Sequence of the xUBF cDNA, (λXlUBF2c). The complete cDNA sequence including the EcoRI linkers (shaded) used in cloning is shown as well as the deduced amino acid sequence for xUBF. The sequence has been submitted to EMBL and given the accession number X57201.

prepared from *X. leavis* stage 17 to 19 polyA$^+$ RNA (kind gift of D. Melton), essentially following published procedures (20). The EcoRI insert of λXlUBF2c, which contains an internal EcoRI site, was subcloned as two fragments into pT3T7-U18 (Pharmacia). The sequence of each fragment was determined by the dideoxy-method (21,22) using a combination of Exo III deletions (Promega, Erase-a-base) and specific oligonucleotide primers. The sequence across the internal EcoRI site of λXlUBF2c was also determined by direct double strand sequencing (23).

RESULTS

The human UBF1 (hUBF) cDNA clone (16) was used to screen a *Xenopus laevis* λgt10 cDNA library. From 30 or so positives, 12 were purified and of these 8 were found to have inserts which hybridised specifically on a southern blot. The sequences of 3 of these clones were determined from the internal EcoRI site (position 1987 in figure 1) and all found to be identical. One of these clones, λXlUBF2c, was selected and the insert of 3729 b.p. sequenced in full, fig. 1. The sequence showed only one open reading frame (290b.p. – 2320b.p.) long enough to encode the *X. laevis* UBF, (xUBF) (15). This open reading frame encoded a polypeptide of 677 amino acids (a.a.) and was proceded by an inframe stop codon at 269, suggesting that the complete coding sequence of xUBF had been isolated. The encoded protein would have a molecular weight of 79.2kd, in reasonable agreement with the relative molecular weight of 85kd determined by gel electrophoresis (15).

Of the eight purified xUBF clones, only two lacked an internal EcoRI site. One of these was also completely sequenced and found to encode a highly homologous but distinct xUBF species, thus demonstrating the existence of two xUBF genes, (data not presented, but see also Discussion).

Northern blot analysis of *X. laevis* RNA probed with XlUBF2c cDNA showed a major RNA species migrating at about 4kb. and minor hybridisation at about 4.7kb., fig. 2A. Thus if the cDNA represented in figure 1 corresponded to the major mRNA species, it was probably 200 to 300 b.p. short. Since no 3'-terminal polyA tract and no consensus polyA addition signal was found in the cDNA, some of the missing sequence probably lies in the already very extensive 3'-untranslated region. Southern blot analysis of DNA from a *X. laevis* individual probed with XlUBF2c at high stringency, showed several major and minor hybridising fragments, fig. 2B. The distribution of major fragments was consistent with the presence of two or more UBF genes. The

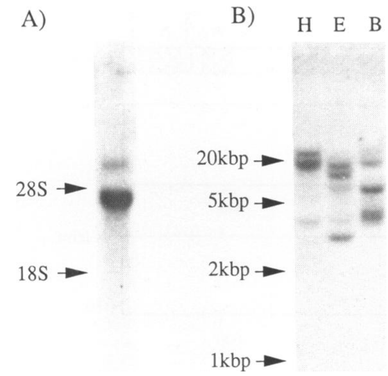

Figure 2. A) Northern analysis of the xUBF mRNA. 3μg of poly-A$^+$ RNA from X. laevis tissue culture cells was gloxylated and electrophoresed on 1% agarose in 10mM Na-phosphate (34). After transfer the RNA blot was probed with a EcoRI fragment (positions 1 to 1987) labelled with ^{32}P by random priming (35,36). Hybridisation was at 42°C in 50% formamide, 6×SSC and the last wash was in 0.2×SSC at 42°C, (37). B) Southern analysis of the xUBF gene. 5μg DNA from a *X. laevis* inividual was digested with BamHI (B), EcoRI (E) and HindIII (H) and separated on 1% agarose. After transfer the DNA blot was probed as in A) excepting that hybridisation was in 6×SSC at 65°C and the final wash in 0.1×SSC also at 65°C. To assure complete digestion, after an initial digestion with a 10× excess of enzyme for 16 hr, the DNA was extracted with phenol/chloroform, precipitated and redigested with the same amount of enzyme for an additional period of 2hr. No difference in hybridisation was noted before and after the second digestion.

minor bands could represent genes carrying related sequences, e.g. closely related HMG-boxes.

The predicted primary structure of xUBF was found to be 73% identical to that predicted for hUBF (16) and 50% of the amino acid changes found to be either conservative or semi-conservative, fig. 3. The choice of start codon for the xUBF was made on the basis of a) the largest open reading frame, b) the use of the first ATG of the cDNA sequence and c) the predicted amino acid sequence homology with hUBF. The first AUG of the message is known to be almost exclusively used as the start codon in eukaryotes (>90% of analysed mRNAs), e.g. see (24). The context of the first ATG in figure 1, (the choice of bases at −3 and +4), was also one of the three most common found in eukaryotic mRNAs (24). However, ATG codons occured in both the x- and hUBF sequences at +37, +79 and +100b.p. relative to the chosen start codon. Thus the use of these as start codons could not be excluded purely by comparison of the predicted and measured molecular weights for xUBF. However alignment of the predicted x- and hUBF a.a. sequences, fig. 3, showed very significant homology, i.e. 5 identical matches and 2 conservative replacements, within the 12 a.a. before the second methionine at residue 13. Therefore it is highly likely that the xUBF coding sequence starts with the ATG indicated in figure 1, but final confirmation of this must await N-terminal sequence analysis of the protein.

HUBF has been shown to contain four tandem primary structure homologies to the folded domains of the HMG 1 and 2 proteins (HMG-boxes 1, 2, 2a and 3, in figure 3) as well as having a highly acidic C-terminal region in common with these proteins (16). xUBF retains both these characteristics. The first three of the HMG-box homologies were found to be present and highly conserved in xUBF, (94, 83 and 93% sequence identity respectively). However the forth HMG-box of hUBF was found

```
xUBF -    1 MNGAAGGVTQAKMTAPKDQDPWSQDEMLTLIQTMKTLLPSQDNSKFKTTE  50
hUBF -    1 MNGEADCPTDLEMAAPKGQDRWSQEDMLTLLECMKNNLPSNDSSKFKTTE  50

         51 SHLDWNKLAFKNYSGSMCRQKWMEISNEVRKFRTLSELILDAEEHVRHPY 100
         51 SHMDWEKVAFKDFSGDMCKLKWVEISNEVRKFRTLTELILDAQEHVKNPY 100

x-Box 1 101 KGKKLKKHPEFPKKPLTPYFRFFMEKRAKYAKLHPEMSNLDLTKILSKKY 150
h-Box 1 101 KGKKLKKHPDFPKKPLTPYFRFFMEKRAKYAKLHPEMSNLDLTKILSKKY 150

        151 KELPEKKKMKYIQDFQREKLEFERNLARFREEHPDLMQNPKKSDVPEKPK 200
        151 KELPEKKKMKYIQDFQREKQEFERNLARFREDHPDLIQNAKKSDIPEKPK 200

x-Box 2 201 TPQQLWYNHERKVYLKLHADASTKDVKDALGKQWSQLTDKKRLKWIHKAL 250
hBox 2  201 TPQQLWYTHEKKVYLKVRPDATTKEVKDSLGKQWSQLSDKKRLKWIHKAL 250

        251 EQRKQYEGIMREYMQKHPELNIAEEGITRSTLTKAERQLKDKFDGRPTKP 300
        251 EQRKEYEEIMRDYIQKHPELNISEEGITKSTLTKAERQLKDKFDGRPTKP 300

x-Box 3 301 PPNSYSMYCAELMANMKDVPSTERMVLCSQRWKLLSQKEKDAYNKKCEQR 350
h-Box 2a 301 PPNSYSLYCAELMANMKDVPSTERMVLCSQQWKLLSQKEKDAYHKKCDQK 350

        351 KKDYEVELMRFLESLPEEEQQRVLAEEKMV.................... 380
        351 KKDYEVELLRFLESLPEEEQQRVLGEEKMLNINKKQATSPASKKPAQEGG 400

 -Δ-    381 ..........................................RSRSGQAD 388
h-Box 3 401 KGGSEKPKRPVSAMFIFSEEKRRQLQEERPELSESELTRLLARMWNDLSE 450

        389 KKKAA.....................DERAKLPETPKTAEEIWQQSVIGDY 418
        451 KKKAKYKAREAALKAQSERKPGGEREERGKLPESPKRAEEIWQQSVIGDY 500

x-Box 4 419 LARFKNDRAKALKSMEGTWLNMEKKEKIMWIKKAAEDQKRYERELSDMRA 468
h-Box 4 501 LARFKNDRVKALKAMEMTWNNMEKKEKLMWIKKAAEDQKRYERELSEMRA 550

        469 TPTPTTAGKKVKFLGEPKKAPMNGYQKFSQELLSNGELNHLPLKERMVEI 518
        551 PPAATNSSKKMKFQGEPKKPPMNGYQKFSQELLSNGELNHLPLKERMVEI 600

x-Box 5 519 GSRWHRISPSQKDYYKKLAEDQQRVYRTQFDTWMKGLSSQDRAAYKEQNS 568
h-Box 5 601 GSRWQRISQSQKEHYKKLAEEQQKQYKVHLDLWVKSLSPQDRAAYKEYIS 650

        569 NKRKSTAKIQVPVAKP...KLVAQSKSDDDDEDDDDDEDEDDDDDEDEDK 615
        651 NKRKSMTKLRGPNPKSSRTTLQSKSESEEDDEEDEDDEDEDEEEEDDENG 700

        616 EDSSEDSDSSESSSDEDSEDGEENEDEEGEEDEDEEED.DDDNES.SSSS 663
        701 DSSEDGSDSSESSSEDESEDGDENEEDDEDEDDDEDDDEDEDNESESSSS 750

        664 SSSSSADSSDSDSN 677
        751 SSSSSGDSSDSDSN 764
```

Figure 3. Alignment of the xUBF and hUBF amino acid sequences. The homologies with HMG 1 and 2 are indicated by open boxes and labelled as box 1 to 5. The nomenclature for the HMG boxes 1 to 3 of hUBF follows ref. (16) and M.-H. Jantzen (personal comm.). The highly acidic regions have been underlined and the conserved serine rich segments shaded. Amino acids conserved between x- and hUBF have been indicated by '|', ':', '.' according to their descending degree of conservation (38,39).

to be essentially absent from xUBF. This accounts in greater part for the molecular weight difference between the Human and *Xenopus* proteins of about 12kd (8,15). It may also be a major factor in determining the transcriptional specificities of these proteins (6), see discussion.

The region N-terminal to HMG-box 1 was found to be quite stringently conserved (70%) between the human and *X. laevis* proteins, but less so than the HMG-boxes. The region between HMG box 3 and the C-terminal acidic domains was however found to be 81% identical (xUBF residues 405−573) with hUBF. Comparison of these sequences with HMG1 identified a further two HMG-box homologies, fig.4A. These have been indicated in figure 3 and 4 and designated Boxes 4 and 5. HMG-box 4 of hUBF would overlap with the previous suggested boundary of box 3. This could be simply due to the difficulty in defining

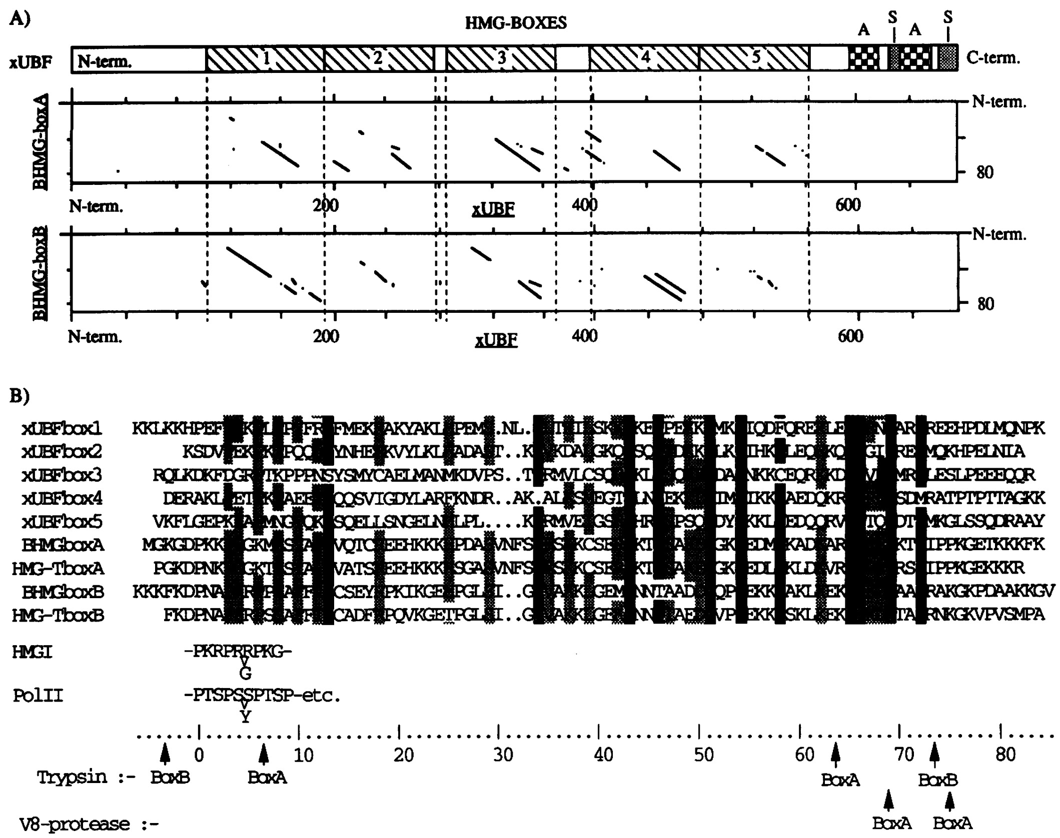

Figure 4. Alignment of HMG-boxes. A) Dotplot comparison of the amino acid sequence of xUBF (horizontal axes) with that of bovine HMG1 folding domains A and B (BHMG-boxA and B) (vertical axes). The programs COMPARE and DOTPLOT (38) were used with a window size of 30 and cutoff of 15, i.e. a 50% match over 30 a.a.. The schematic of the structure of xUBF is shown above the horizontal axes of the plots. 'A' and 'S' refer respectively to the highly acidic regions and the serine-rich segments. B) Alignment of the HMG-boxes in xUBF (xUBFbox1–5) with those of bovine HMG1 (BHMGboxA and B) and trout HMG-T (HMG-TboxA and B). Black and shaded positions indicate respectively 8 or 9 out of 9 and 6 or 7 out of 9 identical or closely related amino acid matches, i.e. basic K,R; acidic D,E; aromatic Y,F,W; hydrophobic I,L,V,F,M; serine/theonine. 'HMGI' and 'PolII' indicate respectively the consensus DNA binding sequence of human HMGI (27) and the consensus sequence of the CTD of RNA polymerase II, e.g. see (40). The rapid cleavage sites for trypsin and V8-protease in bovine HMG1 (BoxA and BoxB) are indicated by arrows.

the boundaries of an HMG-box domain or may indicate that box 3 of hUBF is incomplete. If the latter were true and box 3 of hUBF were therefore non-functional, its nearly complete removal in xUBF would be of less significance than it at first sight appears.

As in hUBF, the acidic C-terminal domain of xUBF was split into two segments by a serine rich sequence and both proteins terminated in another such sequence, figures 3 and 4A. Both serine rich sequences were almost perfectly conserved between x- and hUBF.

DISCUSSION

The predicted primary structure for the *X. laevis* UBF shows a high degree of homology with that of its human counterpart throughout most of its length. Such homology might be expected in a ribosomal transcription factor, especially one which binds in a very similar way to both the human and *Xenopus* promoters (6). The fact that the xUBF essentially lacks a complete putative DNA binding domain present in hUBF, (hUBF-box3, figure 3), is therefore somewhat of a surprise. It was however shown that the xUBF and the hUBF are not interchangeable in *in vitro* transcription assays (6). Thus, this species specificity could at least in part be due to the lack of this HMG-box. The predicted structure for xUBF has also been determined from a second distinct cDNA (unpublished data, D. Bachvarov). This second xUBF shows 95.9% homology with the sequence given in figure 3 except for an insertion of 22 a.a. in the region of the hUBF/xUBF deletion. This insertion probably explains why xUBF is purified as a doublet of ~82 and 85kd (15).

That both x- and hUBF give identical footprints, (ref. (6) and unpublished observations, B. Leblanc), might be interpreted to mean that the hUBF-box3 does not contact DNA in any significant way. This interpretation would also conform with the apparent overlap between hUBF-boxes 3 and 4, (figure 3, but see also below), since box 3 would not need to retain its DNA binding function.

Analysis of the x- and hUBF sequences has allowed us to identify two further HMG1 and 2 homologies, HMG-boxes 4 and 5, fig. 4A. Thus, apart from the 100 or so residues N-terminal to box 1 and the acidic tail, the UBFs appear to consist of 5 or 6 direct repeats of the HMG folding domain. Assuming that all or most of these HMG-boxes constitute a DNA binding domain, it is relatively easy to understand why xUBF gives such very extensive footprints on the *Xenopus* rDNA promoters and enhancers. Each domain when folded would be nearly 3nm in diameter and five such domains strung out along the DNA could then occupy a DNA site more than 50 b.p. long, e.g. a complete 60b.p. enhancer repeat.

It is clear from figures 3 and 4B) that the degree of homology between the HMG boxes of human and *Xenopus* UBF is much greater than the homologies between the boxes within a given UBF. This suggests that each box evolved a distinct role at a very early stage in evolution. The same arguement holds for the HMGs, boxes A or boxes B of HMG1 and HMG-T being very similar, but box A and B of the same HMG being quite dissimilar, (figure 4B).

We have noted a proline repeat, xPxxPxxPx where x is often a basic, theonine or serine residue, which occurs at the N-terminal of each HMG-box, (figure 4B). In bovine HMG 1 this sequence is cleaved from the rest of box A by trypsin with apparently little or no effect on the folded structure of the box (25,26). The sequence of this motif has similarities with the DNA binding motifs of HMGI (27), the C-terminal domain (CTD) of RNA polymerase II (28) and other DNA binding proteins (29–31), (figure 4B) which are believed to bind in the DNA minor groove.

The conservation between *Xenopus* and human of the acidic tail and flanking serine rich segments of UBF (fig. 3 and 4A)) suggests they have conserved functional roles. The acidic residues may play a part in transcription by interacting with other factors in a relatively non-specific way, as has been described for some RNA polymerase II factors, e.g. ref. (32). On the other hand these residues may be important to displace histones from the chromatin in order to allow access of other factors to the DNA, e.g. see ref. (33). If phosphorylated the adjacent serine rich segments could aid in either role.

ACKNOWLEDGEMENTS

The authors wish to thank Dr M.-H. Jantzen and Dr C. Crane-Robinson for advice and many useful discussions during the course of this work. The work was supported by the Medical Research Council of Canada (MRC), T.M. was supported by an MRC scholarship and D.B. was partially supported by a postdoctoral fellowship from Laval University. T.M. is a member of the Centre de Recherche en Cancérologie de l'Université Laval which is supported by the FCAR of Québec.

REFERENCES

1. Tanaka,N., Kato,H., Ishikawa,Y., Hisatake,K., Tashiro,K., Kominami,R. and Muramatsu,M. (1990) J. Biol. Chem., **265**, 13836–13842.
2. Bell,S.P., Jantzen,H.-M. and Tjian,R. (1990) Genes. Dev., **4**, 943–954.
3. Smith,S.D., Oriahi,E., Lowe,D., Yang-Yen,H.-F., O'Mahony,D., Rose,K., Chen,K. and Rothblum,L.I. (1990) Mol. Cell Biol., **10**, 3105–3116.
4. Smith,S.D., Oriahi,E., Yang-Yen,H.-F., Xie,W., Chen,C. and Rothblum,L.I. (1990) Nucleic. Acids. Res., **18**, 1677–1685.
5. Schnapp,A., Clos,J., Hädelt,W., Schreck,R., Cvekl,A. and Grummt,I. (1990) Nucleic. Acids. Res., **18**, 1385–1393.
6. Bell,S.P., Pikaard,C.S., Reeder,R.H. and Tjian,R. (1989) Cell, **59**, 489–497.
7. Bateman,E. and Paule,M.R. (1988) Mol. Cell Biol., **8**, 1940–1946.
8. Bell,S.P., Learned,R.M., Jantzen,H.M. and Tjian,R. (1988) Science, **241**, 1192–1197.
9. Bateman,E. and Paule,M.R. (1986) Cell, **47**, 445–450.
10. Learned,R.M., Learned,T.K., Haltiner,M.M. and Tjian,R.T. (1986) Cell, **45**, 847–857.
11. Tower,J., Culotta,V.C. and Sollner-Webb,B. (1986) Mol. Cell Biol., **6**, 3451–3462.
12. Bateman,E., Iida,C.T., Kownin,P. and Paule,M.R. (1985) Proc. Natl. Acad. Sci. USA, **82**, 8004–8008.
13. Learned,R.M., Cordes,S. and Tjian,R. (1985) Mol. Cell Biol., **5**, 1358–1369.
14. Pikaard,C.S., Smith,S.D., Reeder,R.H. and Rothblum,L. (1990) Mol. Cell Biol., **10**, 3810–3812.
15. Pikaard,C.S., McStay,B., Schultz,M.C., Bell,S.P. and Reeder,R.H. (1989) Genes Dev., **3**, 1779–1788.
16. Jantzen,H.-M., Admon,A., Bell,S.P. and Tjian,R. (1990) Nature., **344**, 830–836.
17. Sinclair,A.H., Berta,P., Palmer,M.S., Hawkins,J.R., Griffiths,B.L., Smith,M.J., Foster,J.W., Frischauf,A.-M., Lovell-Badge,R. and Goodfellow,P.N. (1990) Nature., **346**, 240–244.
18. Gubbay,J., Collignon,J., Koopman,P., Capel,B., Economou,A., Münsterberg,A., Vivian,N., Goodfellow,P. and Lovell-Badge,R. (1990) Nature., **346**, 245–250.
19. Singh,J. and Dixon,G.H. (1990) Biochemistry, **29**, 6295–6302.
20. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) Molecular Cloning, a laboratory manual. Cold Sring Harbor Laboratory., Cold Spring Harbor, New York.
21. Sanger,F., Nicklen,S. and Coulson,A. (1977) Proc. Natl. Acad. Sci. USA, **74**, 5463–5467.
22. Riley,D.E. (1989) Gene, **75**, 193–196.
23. Manfioletti,G. and Schneider,C. (1988) Nucleic Acids Res., **16**, 2873–2884.
24. Kozak,M. (1987) Nucleic. Acids. Res., **15**, 8125–8148.
25. Cary,P.D., Turner,C.H., Leung,I., Mayes,E. and Crane-Robinson,C. (1984) Eur. J. Biochem., **143**, 323–330.
26. Cary,P.D., Turner,C.H., Mayes,E. and Crane-Robinson,C. (1983) Eur. J. Biochem., **131**, 367–374.
27. Reeves,R. and Nissen,M.S. (1990) J. Biol. Chem., **265**, 8573–8582.
28. Suzuki,M. (1990) Nature., **344**, 562–565.
29. Churchill,M.E. and Suzuki,M. (1989) EMBO. J., **8**, 4189–4195.
30. Suzuki,M. (1989) J. Mol. Biol., **207**, 61–84.
31. Suzuki,M. (1989) EMBO. J., **8**, 797–804.
32. Ma,J. and Ptashne,M. (1987) Cell, **48**, 847–853.
33. Carballo,M., Puigdom:enech,P. and Palau,J. (1983) EMBO. J., **2**, 1759–1764.
34. McMaster,G.K. and Carmichael,G.C. (1977) Proc. Natl. Acad. Sci. USA, **11**, 4835–4838.
35. Feinberg,A.P. and Vogelstein,B. (1984) Anal. Biochem., **137**, 266–267.
36. Feinberg,A.P. and Vogelstein,B. (1983) Anal. Biochem., **132**, 6–13.
37. (1987) In Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds.), Current protocols in Molecular Biology. Greene Publishing Associates and Wiley-Interscience, New York,
38. Devereux,J., Haeberli,P. and Smithies,O. (1984) Nucleic Acids Res., **12**, 387–395.
39. Gribskov,M. and Burgess,R.R. (1986) Nucleic Acids Res., **14**, 6745–6763.
40. Corden,J.L. (1990) TIBS., **15**, 383–387.