Volume 288, number 1,2, 55–59 FEBS 10041 August 1991

ADONIS 001457939100765S

Heterogeneity in the *Xenopus* ribosomal transcription factor xUBF has a molecular basis distinct from that in mammals

Dimcho Bachvarov, Marc Normandeau 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 5 June 1991; revised version received 24 June 1991

The *Xenopus* polymerase I transcription factor xUBF is an HMG-box protein which has been purified as two polypeptides of ~82 and 85 kDa. Recently a cDNA sequence predicted an xUBF protein (xUBF1) of 677 amino acids (79 kDa) containing five tandem HMG-boxes. Here a second and distinct xUBF cDNA has been isolated and characterised. This cDNA codes an xUBF protein (xUBF2) of 701 amino acids (82 kDa), having 93% homology with xUBF1 but containing an insertion of 22 amino acids between HMG-boxes 3 and 4. In vitro translation of synthetic mRNAs derived from the xUBF1 and 2 cDNAs was used to show that the electrophoretic mobility of the gene products accounted for the major xUBF molecular weight heterogeneity noted in vivo. It is also shown that the *Xenopus laevis* genome contains 3 or 4 distinct loci which hybridise with xUBF coding sequences, leaving open the possibility of yet further unrecognised heterogeneity in xUBF.

RNA polymerase I; Transcription factor; Ribosomal; xUBF; *Xenopus laevis*

1. INTRODUCTION

UBF is an RNA polymerase I transcription factor which has been isolated and cloned from both mammals [1–5] and *Xenopus* [3,6–8]. The UBFs are characterised by tandemly repeated homologies to the DNA binding domains of the chromosomal proteins HMG 1 and 2, the so-called HMG-boxes as well as a highly acidic and serine-rich C-terminal 'tail'. They form part of the growing family of HMG-box transcription factors, which includes the mammalian sex-determination genes [9,10], yeast proteins [11], lymphoid specific factors [12,13] and mitochondrial factor 1 [14]. It was recently shown that the major difference between the *Xenopus* and mammalian UBF proteins is a deletion in the form of a complete HMG-box. This may explain the observation that neither the mammalian UBF can function in *Xenopus* ribosomal transcription nor the *Xenopus* UBF (xUBF) in mammalian transcription [3,7].

The UBFs of different organisms have without exception been isolated as two polypeptides of differing mobility on SDS gels [1–3,6]. Recently it was shown that in mammals this is probably the result of the expression of two mRNAs coding different molecular weight forms of UBF [5]. Whether these mRNAs originate from two distinct genes or were the result of differential splicing of transcripts from a single gene was however not demonstrated. Here we show that in *Xenopus*, xUBF is expressed from two genes whose products differ in molecular weight and that this xUBF heterogeneity has a molecular basis distinct from that found in mammals.

Correspondence address: T. Moss, Centre de Recherche en Cancérologie de l'Université Laval, Hôtel-Dieu de Québec, 11 Côte du Palais, G1R 2J6 Québec, Canada

2. MATERIALS AND METHODS

λXlUBF4b was isolated as previously described [8] from a λgt10 stage 17 *X. laevis* cDNA bank, using the human UBF cDNA as probe [4]. The XlUBF4b cDNA was subcloned into pT3T7-U19 (Pharmacia) to produce pA68 which was then sequenced [8,15,16]. In vitro transcription of both pA68 (xUBF2) and a subclone of the xUBF1 coding sequence, p451 [8] with T3 polymerase [17] yielded synthetic mRNAs which were then translated in the presence of [^{35}S]methionine (NEN) in a reticulocyte lysate (Novagen) as recommended by the manufacturer. Protein gel electrophoresis was either performed on 5 to 15% gradient gels (Mini-protean pre-cast gels, Bio-Rad) as described by the manufacturer or on 10% gels as described [18] using SDS-PAGE molecular weight markers (Bio-Rad). The gels were either dried and autoradiographed overnight or silver-stained [19]. XUBF was partially purified from a *X. laevis* tissue culture line as described (Read et al., submitted). For Southern analyses, 5 μg DNA from a *X. laevis* individual was digested with the appropriate enzymes and separated on 1% agarose. After transfer, the DNA blot was probed in 6×SSC at 65°C, the final wash being made in 0.1×SSC also at 65°C.

3. RESULTS

During the screening of positive xUBF cDNAs, two of eight clones were found to lack an internal *Eco*RI site characteristic of xUBF [8]. One of these clones was therefore sequenced and found to encode an xUBF distinct from that previously identified [8]. Fig. 1 presents the deduced primary structure of this xUBF2 protein aligned with the previously published xUBF

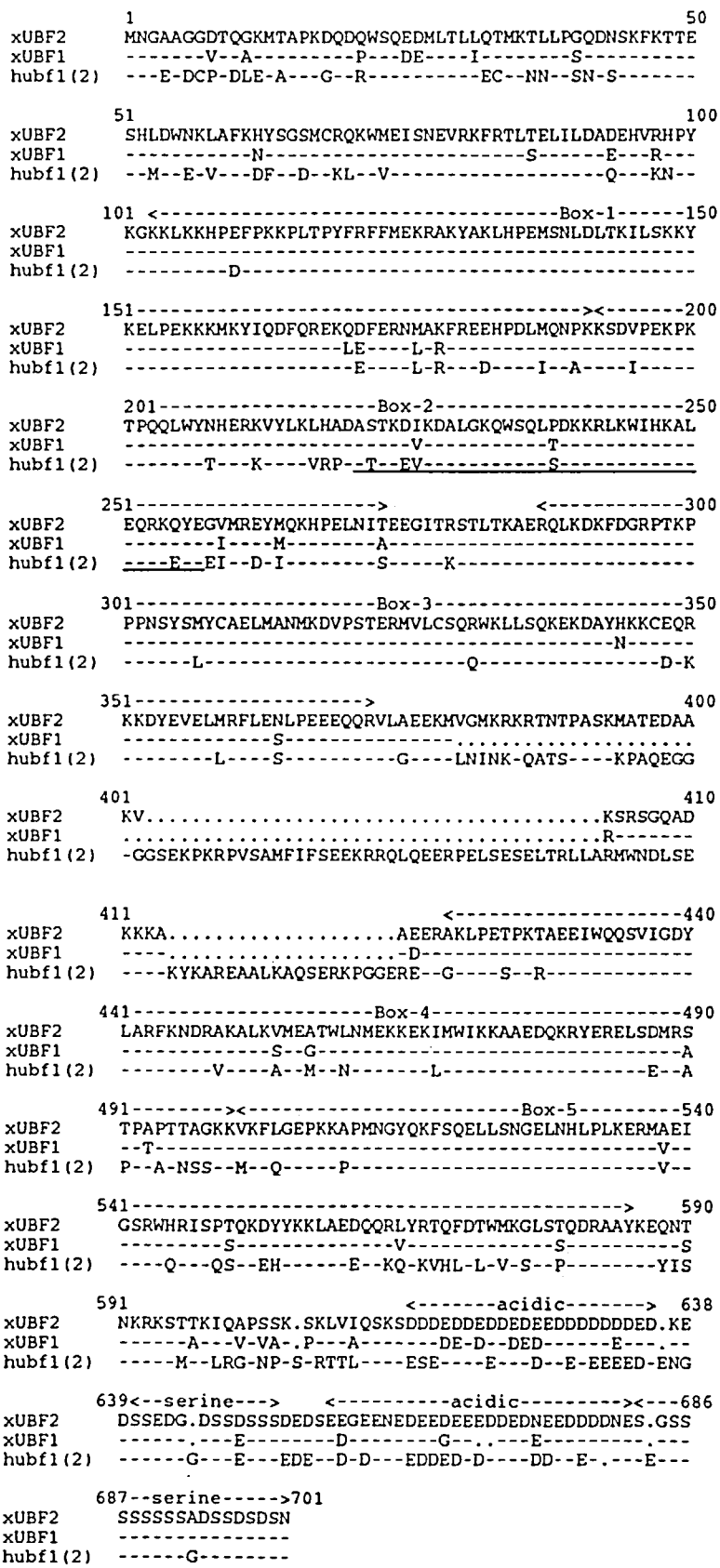

Fig. 1. The amino acid sequence of xUBF2 as deduced from its cDNA sequence, EMBL accession no. X59863. The sequence has been aligned with that of xUBF1 [8] and hUBF1 [4], only differences with xUBF2 are shown and gaps (·) have been introduced to improve sequence alignment. The region deleted in mammalian-UBF2 [5] is shown underlined. The HMG-box homologies are indicated for xUBF2 (<-Box-1-> etc) as are the highly acidic (<-acidic->) and serine-rich (<-serine->) sequences. Amino acid numering is given for xUBF2.

(xUBF1) sequence and that of the human UBF (hUBF1 and 2) [4,5]. The two *Xenopus* UBFs are 93% identical, showing some 50 amino acid changes, often conservative, distributed throughout the protein. A more striking difference between xUBFs 1 and 2 was the insertion/deletion of 22 amino acids between HMG-boxes 3 and 4. In comparison with the human UBF, xUBF1 was shown to lack a stretch of 82 amino acids between HMG-boxes 3 and 4. This essentially removes a complete HMG-box and 20 amino acids to its N-terminal side, which are present in the human protein. The extra 22 amino acids in xUBF2 as compared to xUBF1, show no homology with the missing HMG-box but are somewhat reminiscent of the sequences flanking it in hUBF.

3.1. *The xUBF1 and 2 explain the major heterogeneity in purified xUBF proteins*

XUBF is purified as a doublet of polypeptides of approximately 82 and 85 kDa [6]. The molecular weights of xUBF1 and 2, predicted from Fig. 1 were 79.2 and 82.0 kDa, i.e. well in agreement with the molecular weights estimated by gel electrophoresis. To demonstrate that the xUBF heterogeneity in Fig. 1 did in fact explain the two protein forms found in vivo, xUBF1 and 2 were produced by in vitro translation of synthetic message derived from the two different cDNAs, Fig. 2. The two translation products could be distinguished in size on gradient gel electrophoresis, and when mixed, gave a doublet closely reminiscent of that noted for xUBF purified from tissue culture, (Fig. 2a). The in vitro translation products also migrated at ~85 kDa relative to the molecular weight markers and identically with xUBF purified from tissue culture, detected by silver staining, (Fig. 2b). Thus it would appear that the products of the two xUBF cDNAs identified account for the major heterogeneity in xUBF proteins found in vivo.

3.2. *How many xUBF genes exist in Xenopus?*

The isolation of two distinct cDNAs for xUBF and the distribution of sequence differences between these cDNAs and also between their protein products (Fig. 1), strongly suggested the existence of two distinct xUBF genes. Southern blotting with a range of enzymes was therefore used to determine the minimum number of xUBF genes in the *X. laevis* genome. After digestion of the genome with *Hin*dIII, which cuts in neither the xUBF1 or 2 cDNA, the same three fragments were detected using two adjacent probes from the xUBF coding region (Fig. 3). A similar result was obtained for the enzyme *Pst*I, which does cut in the two cDNAs (Fig. 3c). In neither case could the three fragments be explained in terms of only one or even two genes. To do this it would have been necessary to postulate that one or both of the probes overlapped a *Hin*dIII or *Pst*I site respectively. This in turn would have excluded the two probes detecting the same three fragments. Thus it must

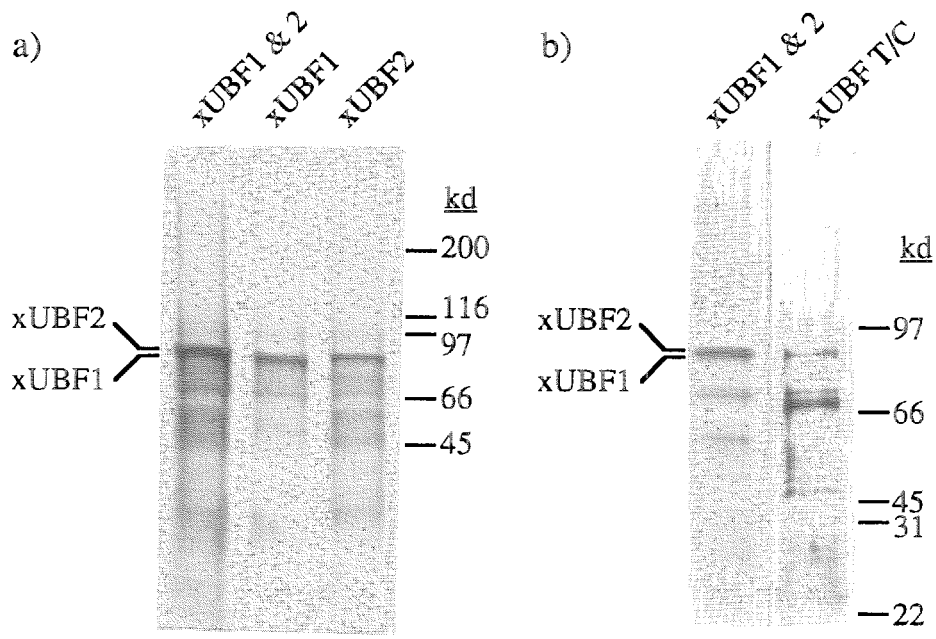

Fig. 2. In vitro translation products of xUBF1 and 2 synthetic RNAs. (a) The respective RNAs were translated, the products fractionated on a gradient gel and the dried gel was autoradiographed overnight. XUBF1 and 2 indicates a mixture of the two translation products. (b) The same mixture of translation products as in (a) was fractionated on a 10% gel in parallel with xUBF partially purified from tissue culture. The gel was stained with silver (xUBF T/C) to detect the purified xUBF was autoradiographed overnight to detect the in vitro translation products (xUBF 1 and 2). The yield of the two molecular weight forms of xUBF in different purified preparations varied. The preparation used here predominantly contained the lower molecular weight form.

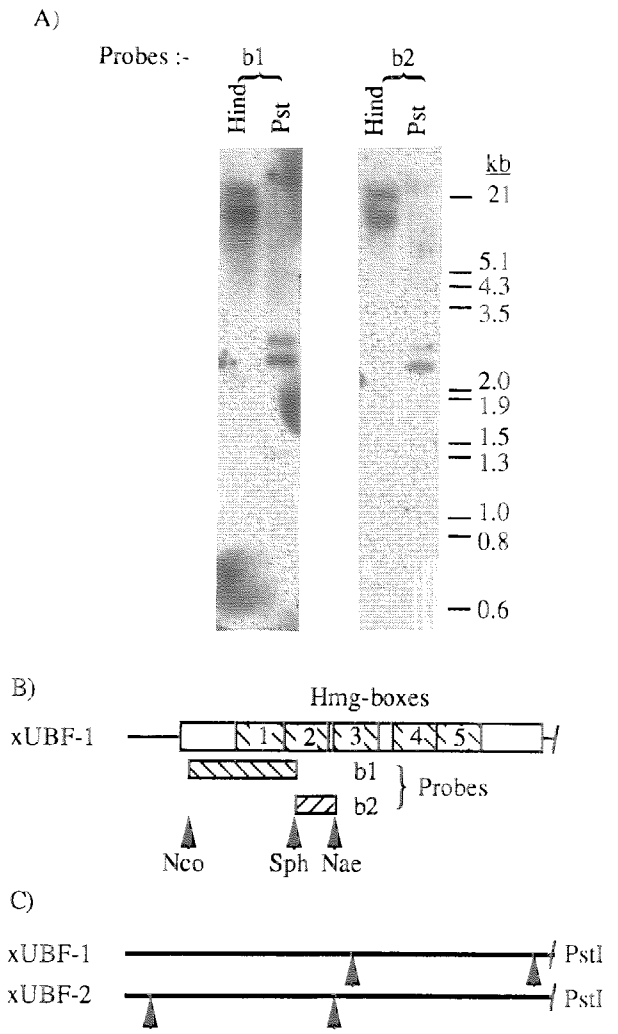

Fig. 3. (a) Southern hybridisation of *Hin*dIII (Hind) and *Pst*I (Pst) restricted genomic DNA with probes specific for the HMG-box-1 (b1) or -2 (b2) of xUBF. (b) The regions of the xUBF1 cDNA contained within probes b1 and b2, Nco, Sph and Nae refer to sites for the enzymes *Nco*I, *Sph*I and *Nae*I, respectively. (c) The positions of cleavage of *Pst*I within the xUBF1 and xUBF 2 cDNAs. Note that *Hin*dIII cuts in neither cDNA.

be concluded that the *X. laevis* genome contains a minimum of three xUBF genes. The second largest of the *Hin*dIII and of the *Pst*I fragments in Fig. 3 were found to be more intense than the other two, suggesting two or more co-migrating fragments. Since the xUBF1 and 2 cDNA sequences are highly homologous and both probes gave the same relative hybridisation intensities, it seems unlikely that this was due to different hybridisation efficiencies. Thus it can be concluded that a minimum of three and more probably four distinct xUBF genes or pseudogenes exist in *X. laevis*.

4. DISCUSSION

Here it has been demonstrated that the purification of xUBF as two peptides of about 82 and 85 kDa relative molecular weight is the result of the expression of two distinct xUBF genes, referred to respectively as XUBF1 and 2. The major difference between the two xUBFs is the presence in xUBF2 of an extra stretch of 22 amino acids. These extra amino acids lie within the region of major length difference between the mammalian UBFs and xUBF1. Both xUBF1 and 2, however, lack one of the putative DNA binding domains (HMG-boxes) of the mammalian UBFs.

It has been shown in Fig. 3 that xUBF may be coded on more than two genes. The cDNA sequences presented here and in [8] were isolated from stage 17 embryos whereas the xUBF (Fig. 2b) has been purified from a tissue culture line. It is therefore possible that the genes expressed in the culture line and in the embryo are distinct and hence their products are also distinct. PCR amplification from tissue culture mRNA of the region of length variability between xUBF1 and 2 however detects only two distinct RNAs, whose sizes correspond with the cloned cDNAs, (A. Guimond, unpublished results). Thus it is likely that the same major forms of xUBF are expressed in embryos and in tissue culture.

Since the UBFs have been purified as a doublet of two polypeptides in all species so far studied, it was rather surprising to find that the UBF heterogeneity found in *Xenopus* had a distinct molecular basis to that in mammals. Two forms of UBF mRNA were recently found to exist in mammals [5]. Though it was not shown whether these two forms were coded on distinct genes, it was demonstrated that the encoded UBFs are related by a deletion lying within HMG-box2, (underlined in Fig. 1). Amplification by PCR has clearly shown no equivalent heterogeneity within *X. laevis* mRNA, (A. Guimond, unpublished results). The different molecular bases for UBF heterogeneity in *Xenopus* and in mammals suggests a high degree of evolutionary adaption. This is in good agreement with the species specific transcriptional properties of the UBFs, but not with their DNA-binding characteristics.

Acknowledgements: The authors wish to thank Dr M.-H. Jantzen for useful discussions during the course of this work. The work was supported by the Medical Research Council of Canada (MRC). 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] Bell, S.P., Learned, R.M., Jantzen, H.M. and Tjian, R. (1988) Science 241, 1192–1197.
[2] Bell, S.P., Jantzen, H.-M. and Tjian, R. (1990) Genes Dev. 4, 943–954.
[3] Pikaard, C.S., Smith, S.D., Reeder, R.H. and Rothblum, L. (1990) Mol. Cell Biol. 10, 3810–3812.
[4] Jantzen, H.-M., Admon, A., Bell, S.P. and Tjian, R. (1990) Nature 344, 830–836.

[5] O'Mahony, D.J. and Rothblum, L.I. (1991) Proc. Natl. Acad. Sci. USA 88, 3180-3184.
[6] Pikaard, C.S., McStay, B., Schultz, M.C., Bell, S.P. and Reeder, R.H. (1989) Genes Dev. 3, 1779-1788.
[7] Bell, S.P., Pikaard, C.S., Reeder, R.H. and Tjian, R. (1989) Cell 59, 489-497.
[8] Bachvarov, D. and Moss, T. (1991) Nucleic Acids Res. 19, 2331-2335.
[9] 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.
[10] 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.
[11] Kolodrubetz, D. and Burgum, A. (1990) J. Biol. Chem. 265, 3234-3239.
[12] Van de Wetering, M., Oosterwegel, M., Dooijes, D. and Clevers, H. (1991) EMBO J. 10, 123-132.
[13] Travis, A., Amsterdam, A., Belanger, C. and Grosschedl, R. (1991) Genes Dev. 5, 880-894.
[14] Parisi, M.A. and Clayton, D.A. (1991) Science 252, 965-969.
[15] Sanger, F., Nicklen, S. and Coulson, A. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
[16] Riley, D.E. (1989) Gene 75, 193-196.
[17] Neilsen, D.A. and Shapiro, D.J. (1986) Nucleic Acids Res. 14, 5936.
[18] Laemmli, U.K. (1970) Nature 227, 680-685.
[19] Wray, W., Boulikas, T., Wray, V.P. and Hancock, R. (1981) Anal. Biochem. 118, 197-203.