The EMBO Journal vol.11 no.10 pp.3695-3704, 1992

Mouse rRNA gene transcription factor mUBF requires both HMG-box1 and an acidic tail for nucleolar accumulation: molecular analysis of the nucleolar targeting mechanism

Yasushi Maeda, Koji Hisatake, Takashi Kondo, Ken-ichi Hanada, Chao-Zhong Song, Tetsuji Nishimura and Masami Muramatsu[1,2]

Department of Biochemistry, Faculty of Medicine, The University of Tokyo, Hongo, Bunkyo-Ku, Tokyo 113, Japan
[1]Corresponding author
[2]Present address: Department of Biochemistry, Saitama Medical School, 38 Moro-hongo, Moroyama, Iruma-gun, Saitama 350-04, Japan

Communicated by F.Amaldi

RNA polymerase I requires at least two nucleolar transcription factors, UBF and SL-1, for ribosomal RNA gene (rDNA) transcription. UBF requires SL-1 for the formation of a stable initiation complex on the rDNA promoter region. We have determined the region of mouse UBF (mUBF) required for nucleolar targeting. Although mUBF has a nuclear localization sequence, this sequence alone is not sufficient for mUBF to accumulate in the nucleolus. Deletion analyses show that mUBF requires a wide region except for the N-terminal 101 amino acids for nucleolar targeting. Deletion of either the HMG-box1, a region crucial for rDNA binding, or the acidic tail, a region that may interact with SL-1, results in the loss of nucleolar targeting. We show by DNA affinity analysis that the HMG-box1 is absolutely necessary for mUBF to bind to the upstream control element of the rDNA. We also show that mUBFs with various internal deletions retain both nucleolar targeting and DNA binding ability. A clear correlation was demonstrated between the DNA binding and nucleolar targeting ability. These results suggest that UBF is transferred to the nucleus by its NLS and is sequestered in the nucleolus by its specific and stable binding to the rDNA promoter via HMG-boxes and the acidic tail.

Key words: acidic tail/HMG-box/mUBF/nucleolar targeting/ribosomal RNA gene

Introduction

It has recently been shown that the transition of transcription factors, such as those of the *rel* family, from the cytoplasm to the nucleus is an important step for the decision of cell fate (Ghosh *et al.*, 1990; Kieran *et al.*, 1990; Ballard *et al.*, 1991). This is an example of the extreme significance of the mechanism that sorts nuclear proteins to their working compartments. Some nuclear proteins, including RNA polymerase I, nucleolin (Lapeyre *et al.*, 1987), p120 (Freeman *et al.*, 1988), NO38 (Schmidt *et al.*, 1987) and some ribosomal proteins accumulate in the nucleolus. Some viral proteins, such as HTLV-1 Rex and HIV Rev and Tat also accumulate in the nucleolus of infected cells.

Nuclear proteins are known to have a so-called nuclear localization sequence (NLS), which consists of an array of basic amino acids. Although nucleolar targeting signals have been proposed for the above-mentioned viral proteins (Simoi *et al.*, 1988; Cochrane *et al.*, 1990; Rosen, 1991), such signal sequences have not been demonstrated in cellular nucleolar proteins.

In this study, the mouse upstream binding factor (mUBF; Hisatake *et al.*, 1991), a transcription factor for rRNA genes (rDNA), was used to analyze the nucleolar targeting signal of a physiological protein. As the UBF forms part of the transcription initiation complex at the nucleolar organizer region (NOR) and as rRNA is exclusively synthesized there, UBF must first go to the nucleolus to function. It has been known that UBF specifically binds to the upstream control element (UCE) as well as to the core element of the rDNA promoter, forming a stable complex with the other crucial transcription factor, SL-1, containing several peptides (Comai *et al.*, 1992); it then activates rDNA transcription (Learned *et al.*, 1986; Bell *et al.*, 1988; Jantzen *et al.*, 1990; Bachvarov and Moss, 1991; O'Mahony and Rothblum, 1991). rDNA from one species can be transcribed by an evolutionarily closely related species (e.g. human and monkey, or mouse and rat), but not by more distantly related species (e.g. human and mouse) (Grummt *et al.*, 1982). Since the compatibility of UBF but not of SL-1 is demonstrated *in vitro* between mouse and human, SL-1 contains the species-specific factor that directs transcription only of the cognate template (Bell *et al.*, 1989). SL-1 is termed TFID, factor D, TIF1B or Rib1 in different laboratories (Mishima *et al.*, 1982; Clos *et al.*, 1986; Kato *et al.*, 1986; Tower *et al.*, 1986; Tanaka *et al.*, 1990; McStay *et al.*, 1991a).

The mUBF has six HMG-boxes (Hisatake *et al.*, 1991) which have amino acid sequences homologous to the relatively abundant non-histone nuclear protein, HMG protein 1, and it has an acidic C-terminal tail. It has been shown that human UBF (hUBF) binds to DNA via its HMG-boxes and it was speculated that it interacts with SL-1 at the acidic tail, because SL-1 is highly positively charged (Jantzen *et al.*, 1990). In this study we have prepared a variety of deletion mutants of mUBF and tested them for their nucleolar targeting capability as well as for their ability to bind to the rDNA promoter (UCE). The results indicate that the HMG-box1 is absolutely required for rDNA binding, but both the HMG-box1 and the acidic tail are required for nucleolar accumulation of mUBF.

Results

Expression and subcellular localization of an mUBF-β-galactosidase fusion protein in COS7 cells

Because the human rDNA promoter sequence is almost identical to that of monkey, and because human rDNA can be transcribed efficiently in COS7 monkey cells, we used COS7 cells for this analysis. A mUBF-β-galactosidase

© Oxford University Press

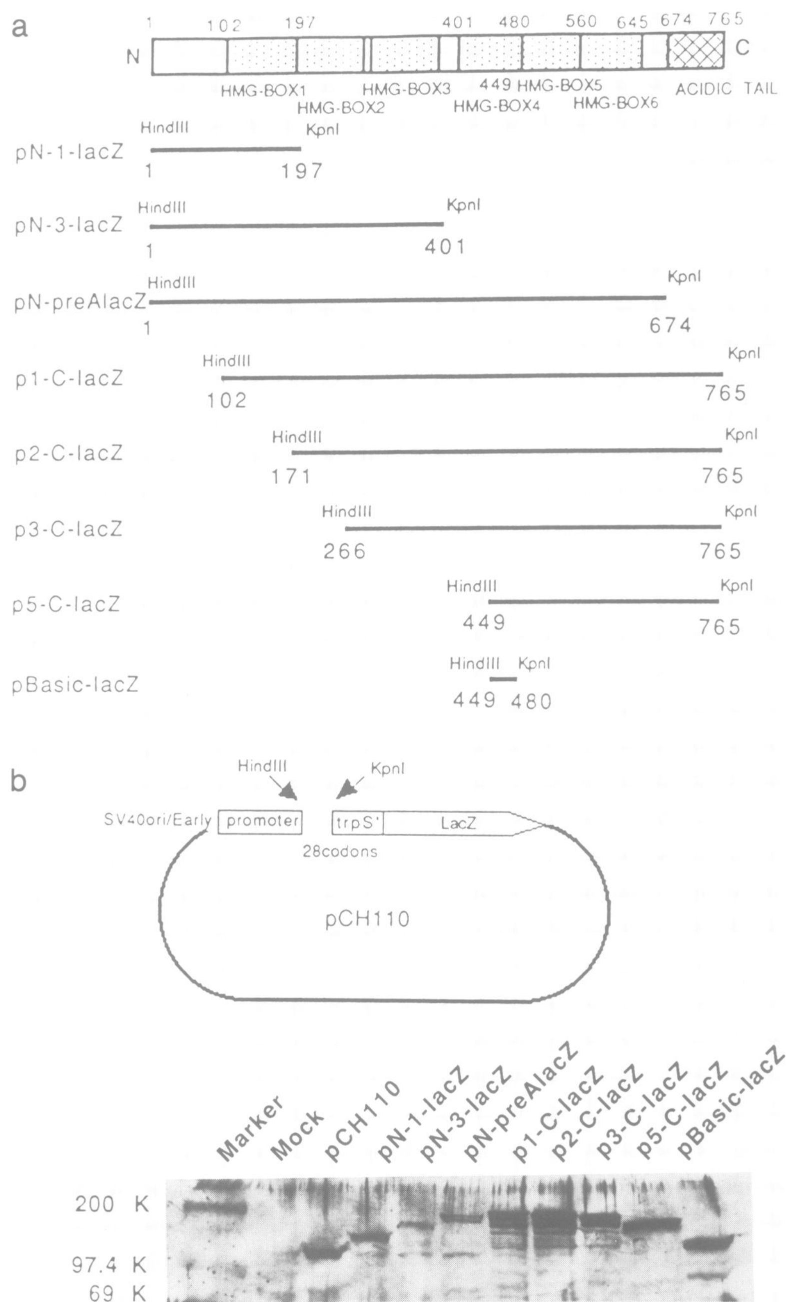

Fig. 1. mUBF–β-galactosidase expression plasmid constructs. (**a**) Eight kinds of expression plasmid were constructed to determine the region required for the nucleolar targeting of mUBF. The RNA expressed from pCH110 was translated from the ATG of intrinsic Ecogpt. pCH110 was digested with *Hin*dIII and *Kpn*I, and the translation start site eliminated. The N-terminal or C-terminal deletion mutants were ligated to the N-terminal portion of β-galactosidase. The shortest mUBF had only a part of HMG-box4 containing a nuclear localization sequence. (**b**) The production of expected fusion proteins in the COS7 cells. The expressed proteins were identified by Western blotting with a polyclonal antibody against β-galactosidase. pCH110 expresses an ~110 kDa protein. Mock transfection shows no detectable band. All the expression plasmids produce a protein having the expected molecular size.

fusion protein was expressed efficiently in COS7 cells. The expressed proteins were checked for their molecular sizes. Assuming that β-galactosidase has an M_r of 110 kDa, each fusion protein had the expected M_r as shown by Western blotting after SDS gel electrophoresis (Figure 1b). The expressed proteins also had β-galactosidase activity (data not shown). We tried to express the full length mUBF–β-galactosidase fusion protein without success. However, we were able to express a protein lacking only 101 amino acids from the N-terminus, which had a capacity to accumulate in the nucleolus as described below.

To determine the subcellular localization of the mUBF–β-galactosidase fusion protein, an antibody to β-galactosidase was used (see Materials and methods). The secondary antibody alone did not cross-react with any cellular component in this system (Figure 2A). After transformation with pCH110, the *lacZ* product was observed in the cytoplasm (Figure 2B). To determine the position of the NLS, we constructed C-terminal deletion mutants, pN-1-lacZ, pN-3-lacZ and pN-preA-lacZ (Figure 1a) and examined their cellular location after reacting with the β-galactosidase antibody. Fusion proteins produced from the

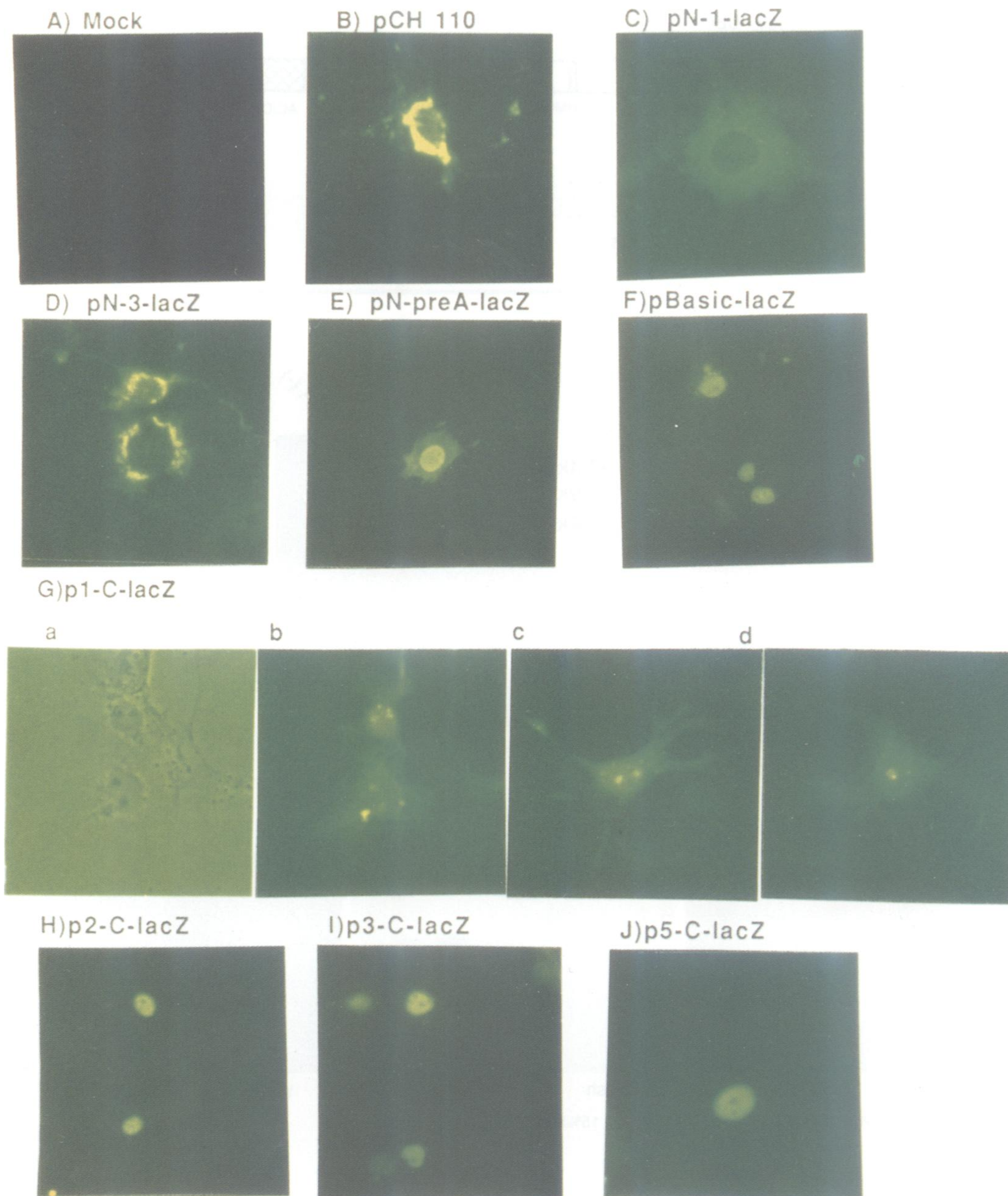

Fig. 2. Subcellular localization of mUBF−β-galactosidase fusion proteins. Transfected cells were stained by indirect immunofluorescence and pictures were taken. (**A**) No cross-reaction can be seen in the mock transfection. (**B**) β-galactosidase cannot enter the nucleus by itself. Panels B−F, panel G parts b, c and d and panel J show cells transfected with a construct containing (**C**) from the N-terminus to HMG-box1; (**D**) from the N-terminus to HMG-box3; (**E**) from the N-terminus to amino acid 674; (**F**) amino acids 449−480; (**G** parts b, c and d) from HMG-box1 to the C-terminus; (**H**) from HMG-box2 to the C-terminus; (**I**) from HMG-box3 to the C-terminus; (**J**) from HMG-box5 to the C-terminus. Panel G part a is a phase contrast micrograph of the field of part b.

first two constructs did not enter the nucleus (Figure 2C and D), while the third construct, which lacked only the acidic tail, accumulated in the nucleus but did not enter the nucleolus (Figure 2E). These results suggest that an NLS exists between amino acids 401 and 674. A 32 amino acid region (positions 449−480) is extremely rich in basic amino acid and resembles the so-called NLS of other proteins. Indeed, as shown in Figure 2F, this basic region alone was sufficient for nuclear targeting when ligated to the β-galactosidase protein. The plasmid p1-C-lacZ, which lacks only 101 amino acids from the N-terminus, produced the predicted fusion protein and this protein did accumulate in the nucleolus (Figure 2G), indicating that the region from amino acid 102 to the C-terminus contains all the information necessary for nucleolar targeting. Because the construct pN-preA-lacZ, which contained the region from the N-terminus to amino acid 674, did not accumulate in the nucleolus (Figure 2E), the possibility of the nucleolar targeting signal being between amino acids 674 and 765 was considered. However, the fusion proteins produced by expression plasmids p2-C-lacZ, p3-C-lacZ and p5-C-lacZ, containing both the acidic tail and the NLS, but not the HMG-box1, did not accumulate in the nucleolus, although they did enter the nucleus (Figure 2H, I and J). These data indicate that both HMG-box1 and the acidic tail are required for mUBF to target the nucleolus.

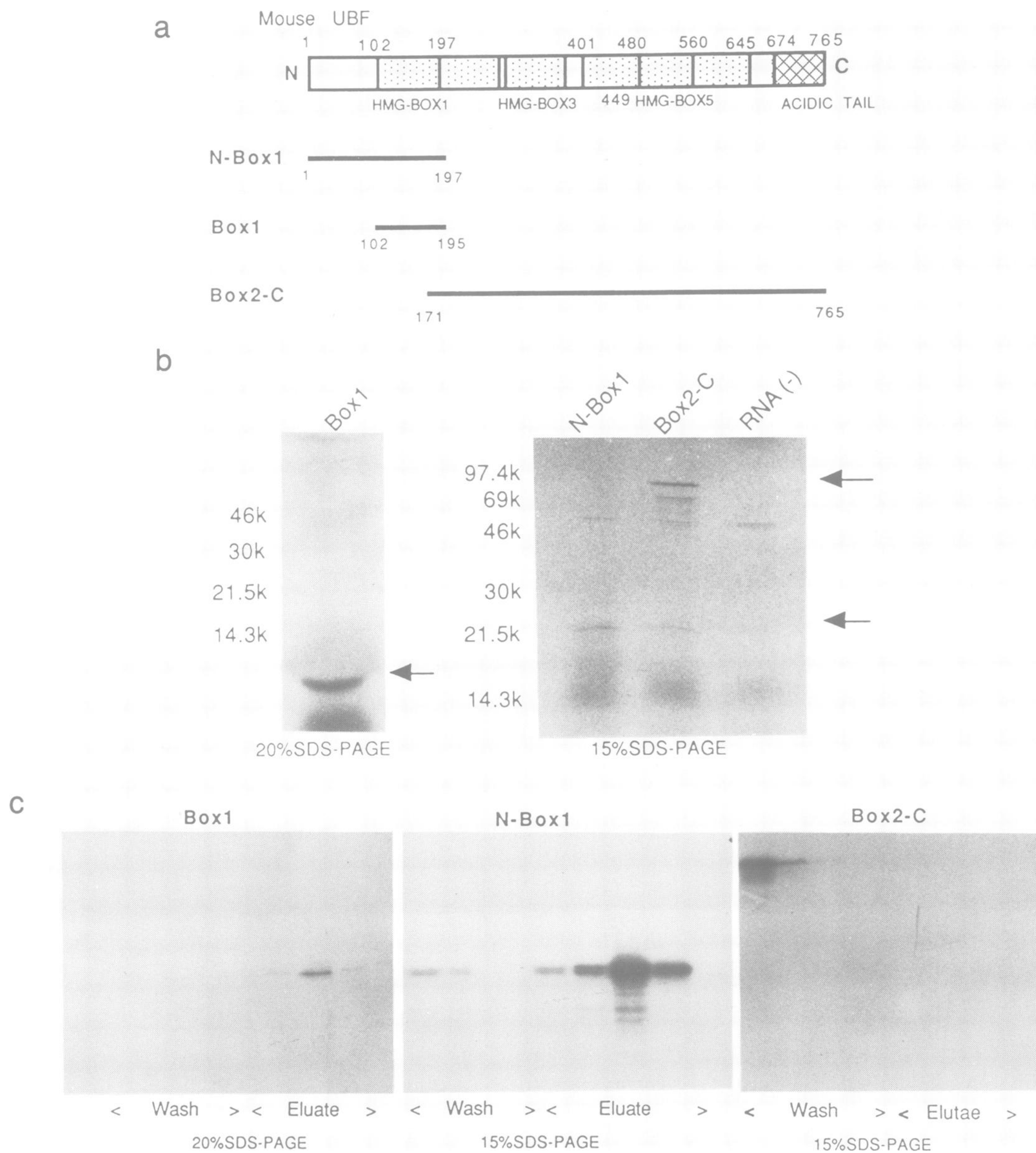

Fig. 3. Mapping the region of mUBF responsible for UCE-specific binding by UCE affinity chromatography. (**a** and **b**) Three kinds of protein were synthesized in an *in vitro* transcription−translation system. Each protein had the expected molecular size. N-Box1: from the N-terminus to the C-terminus of HMG-box1, containing 197 amino acids. Box1: only HMG-box1, containing 96 amino acids. Box2-C: from the N-terminus of HMG-box2 to the C-terminus, containing 595 amino acids. The latter lacks HMG-box1. This system produced proteins with the predicted sizes. Arrows indicate the synthesized proteins. (**c**) Part of the wash and elution fractions were electrophoresed on a 20% or 15% SDS−polyacrylamide gel. Because N-Box1 and Box1 proteins elute in the high salt (0.6 M KCl) eluate, they must have the capacity to bind DNA. In contrast, Box2-C protein lacking HMG-box1 cannot be seen in the eluate; it exists only in the wash fractions (0.1 M KCl).

HMG-box1 is absolutely necessary for UCE binding

To analyze the nucleolar targeting requirements further, we examined the role of the HMG-box1. Comparison of the results of p1-C-lacZ and p2-C-lacZ indicates that the HMG-box1 is essential for nucleolar targeting, as the deletion of HMG-box1 alone resulted in the loss of nucleolar targeting ability of mUBF (Figure 2G and H). Therefore, we investigated the role of the HMG-box1 in the specific binding of this molecule to the UCE sequence using sequence affinity chromatography as described by Jantzen *et al.* (1990). We have used the human UCE element for this binding study, because the mouse UCE has not yet been well defined, and because mUBF is known to bind to the human UCE as well as to the mouse promoter region. Various proteins were synthesized using an *in vitro* transcription−translation system (Figure 3a and b). Each methionine-labeled protein was applied to a UCE sequence affinity column and fractions that eluted with high salt were analyzed on an SDS−polyacrylamide gel. Two synthesized proteins containing the HMG-box1 bound to the UCE, whereas a protein lacking the HMG-box1 failed to bind. These results prove that the HMG-box1 is absolutely necessary for the mUBF to bind the UCE *in vitro*.

Scatchard analysis with the expressed fusion proteins

In order to analyze the binding ability of mUBF in more detail, we made MBP fusion proteins using a bacterial expression system that can produce sufficiently high amounts

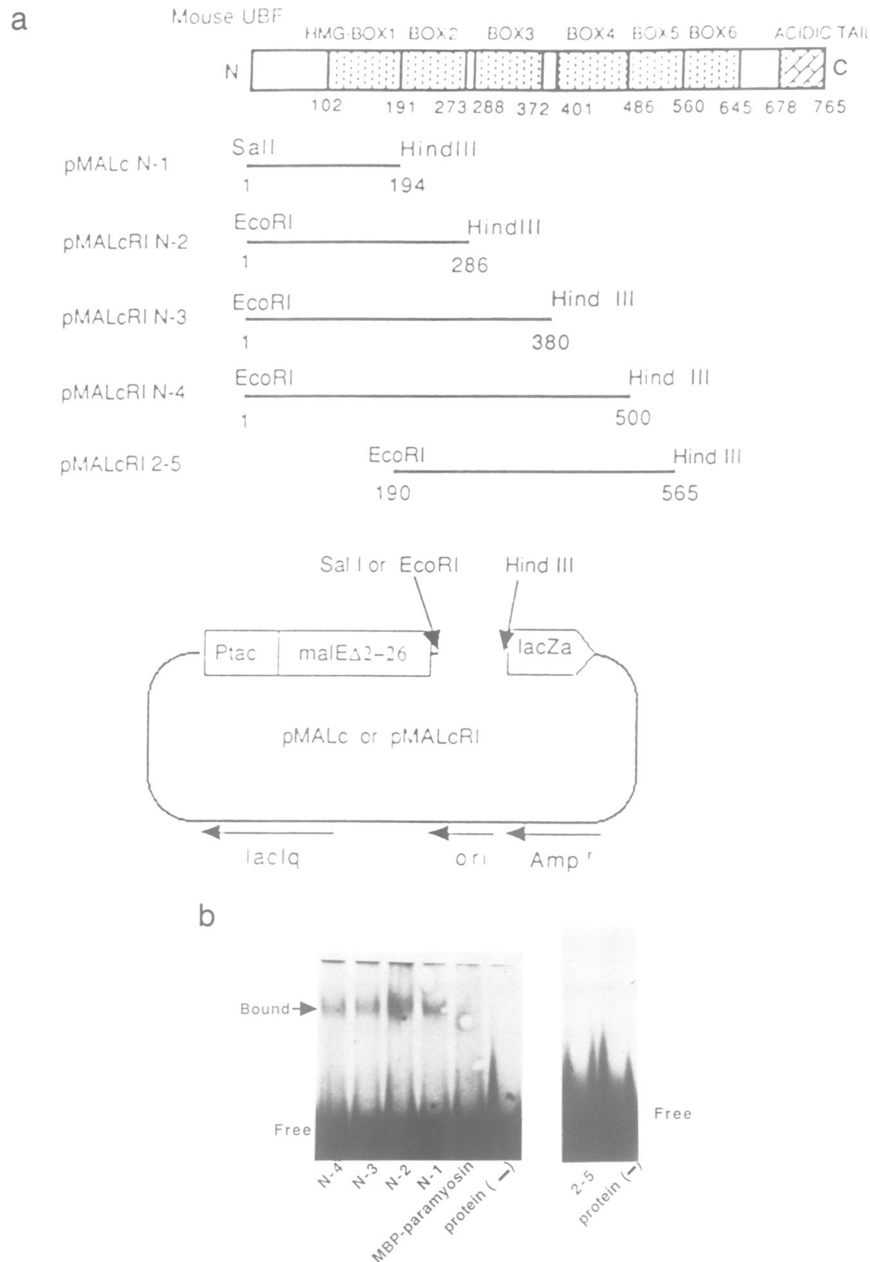

Fig. 4. Constructs of MBP−mUBF fusion protein expression vectors and capacity of fusion proteins to bind to human UCE. (**a**) Four types of expression plasmid, pMAL-c N-1, pMAL-cRI N-2, pMAL-cRI N-3 and pMAL-cRI N-4, are C-terminal deletions that contain HMG-box1. Only pMAL-cRI 2-5 lacks HMG-box1. The mol. wt of MBP is 42 kDa. Five kinds of fusion protein with mol. wts of 70, 80, 90, 95 and 85 kDa were expressed. (**b**) Results of the gel retardation assay. Four kinds of UBF fusion protein, N-1, N-2, N-3 and N-4, produced retarded bands but MBP−paramyosin did not. The specificity was confirmed by using a specific antibody against MBP. The fusion protein without HMG-box1 did not bind to the human UCE probe. Human UCE is the upstream region from −106 to −70 of human rDNA, the sequence of which is shown in Materials and methods.

of protein for a quantitative binding assay. The five fusion proteins (Figure 4a) produced with the expected M_rs were purified on amylose affinity resin and their capacity to bind UCE was assessed by gel retardation assay. As shown in Figure 4b, four proteins, N-1, N-2, N-3 and N-4, were able to bind. That this binding was specific was confirmed using a specific antibody against the MBP (data not shown). No detectable retarded bands were found on an MBP−paramyosin column, suggesting that binding did not occur via the MBP portion of the fusion proteins. Moreover, Box2−Box5 protein did not bind to the human UCE (Figure 4b). Although we did not try to express a protein having only HMG-box1 to HMG-box5, the binding ability of a protein corresponding to this was shown recently in *Xenopus* UBF using the MBP fusion protein system and South-western blotting (O'Mahony *et al.*, 1992b). These data also confirm that the HMG-box1 is essential for the binding to UCE.

We next determined the affinity of these proteins for the UCE by Scatchard analysis. All binding reactions showed saturation curves, with the K_ds of N-1, N-2, N-3 and N-4 being 61, 38, 73 and 1.2 nM, respectively (Figure 5). The affinity of N-4 for the UCE is more than 50 times higher than that of N-1. This is in agreement with the recent report by O'Mahony *et al.* (1992b). The affinities of N-1, N-2 and N-3 are very similar and very low. Although the Box2−Box5 protein has four HMG-boxes, it did not bind

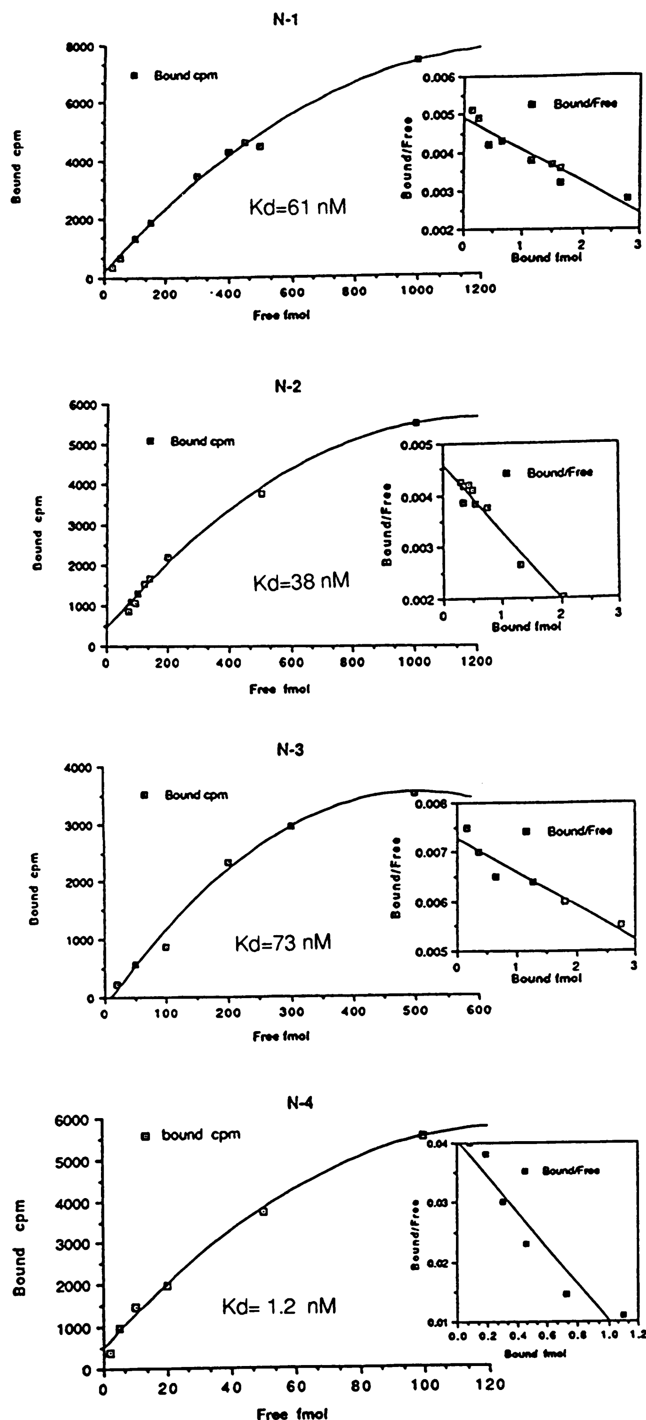

Fig. 5. Scatchard analysis. Scatchard analysis was carried out with four kinds of fusion protein that were able to bind DNA. Each protein shows a saturating curve when the probe concentration increases. The K_ds of N-1, N-2, N-3 and N-4 are 61, 38, 73 and 1.2 nM, respectively.

appreciably to the UCE, indicating that the number of HMG-boxes is not the key to high affinity binding. This result indicates that the mUBF has a high affinity for the UCE when HMG-box1 and a few more boxes are present.

Internally deleted mUBFs can accumulate in the nucleolus and also retain UCE-binding activity

To examine the possible effect of conformational change of UBF on its nucleolar targeting, four types of internally deleted mUBF were expressed in COS7 cells. In constructing the internal deletion mutants, the number of HMG-boxes seemed to be important for high affinity binding as mentioned in the previous section. We therefore deleted more than a half of each HMG-box using overlap-extension PCR (Figure 6a) but kept the other HMG-boxes intact. When these fusion proteins were expressed in COS7 cells, these UBF$-\beta$-galactosidase molecules had intact enzymatic activity and the expected M_r, and they were able to accumulate in the nucleolus (Figure 6b). The binding ability of these deleted proteins was also verified biochemically (Figure 7). These results indicate that the possible conformational change caused by internal deletion of the molecule did not affect either the nucleolar targeting or the DNA binding ability, provided that the mUBF retained the HMG-box1 and acidic tail (see Discussion).

Discussion

UBF has a nuclear localization sequence but no nucleolar targeting signal

We have demonstrated here that UBF has an NLS but not a nucleolar targeting signal like those found in viral proteins. All of the expression plasmids that contained amino acids 401–674 produced proteins that were transported effectively into the nucleus. The shortest peptide that could target β-galactosidase fusion protein into the nucleus was 31 amino acids long and located in HMG-box4. Although we could not find a sequence identical to any known NLS, a sequence rich in basic amino acid residues was found between amino acids 452 and 456 (KKKAK). This sequence is located at the center of an α-helix and is presumed to be exposed on the surface of the UBF molecule, which may facilitate the interaction of UBF with import machinery. Since these two features satisfy the general characteristics of an NLS (Kalderon *et al.*, 1984a,b; Richardson *et al.*, 1986; Loewinger and Mckeon, 1988; Hauber *et al.*, 1989; Ruben *et al.*, 1989; Silver, 1991), we conclude that this sequence is an mUBF NLS. It should be noted that the five amino acids surrounding this basic cluster are conserved between human, mouse and rat UBFs. This cluster is also found in *Xenopus*, although the neighboring regions are missing (Bachvarov and Moss, 1991; McStay *et al.*, 1991b).

We have also demonstrated that the N-terminal 101 amino acid sequence is not required for nucleolar targeting of mUBF. However, the absolute necessity of HMG-box1 as well as the C-terminal acidic tail was demonstrated. The results together indicate that the nucleolar targeting signal of UBF is not contained in a short, basic amino acid sequence similar to the sequences of viral proteins such as Rex, Rev and Tat, but resides in a wide region that includes both the DNA binding domain and the C-terminal acidic tail, which may be the domain that interacts with SL-1.

Specific DNA binding ability

Jantzen *et al.* reported the importance of the hUBF HMG-box1 for specific UCE binding (Jantzen *et al.*, 1990). We have demonstrated here that the HMG-box1 is absolutely necessary for nucleolar targeting as well as for the specific binding with UCE *in vitro* (Figures 3 and 4). When the HMG-box1 was removed from UBF, these UBF mutants lost the ability to bind to UCE. We were also able to show that the effect of other HMG-boxes was cooperative rather than additive. When HMG-box4 was added, its overall DNA

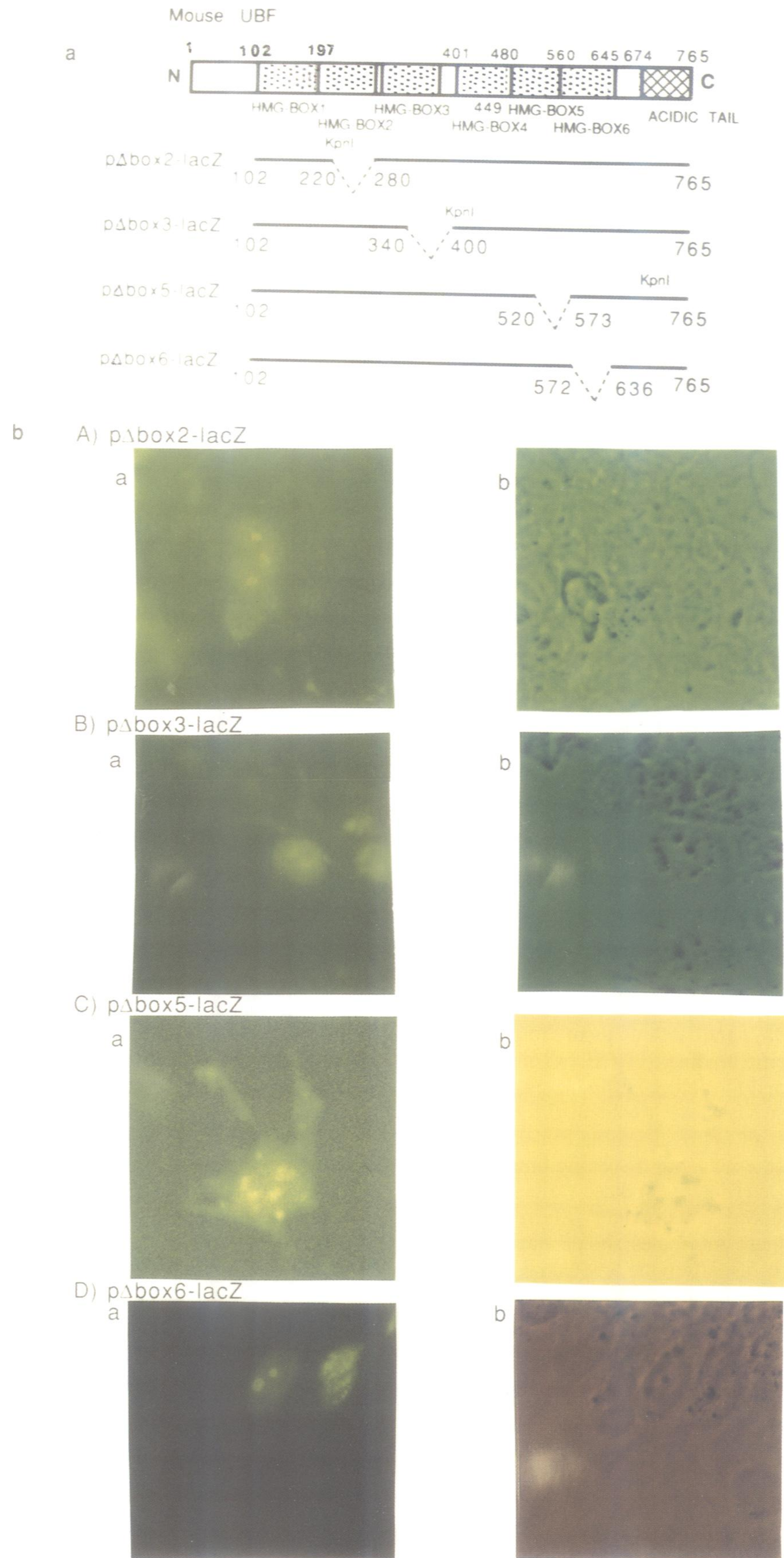

Fig. 6. Internal deletion constructs for the expression in COS7 cells and subcellular localization of fusion proteins. (**a**) The deleted regions of UBF in each construct. The fusion protein was expressed in COS7 cells and had an M_r of ~180 kDa. (**b**) Left panels are pictures of indirect immunofluorescence staining; right panels are pictures of the phase contrast micrograph of the left counterpart. Panels A, B, C and D show cells transfected with (**A**) pΔbox2–lacZ, (**B**) pΔbox3–lacZ, (**C**) pΔbox5–lacZ and (**D**) pΔbox6–lacZ. All four proteins exist in the nucleolus.

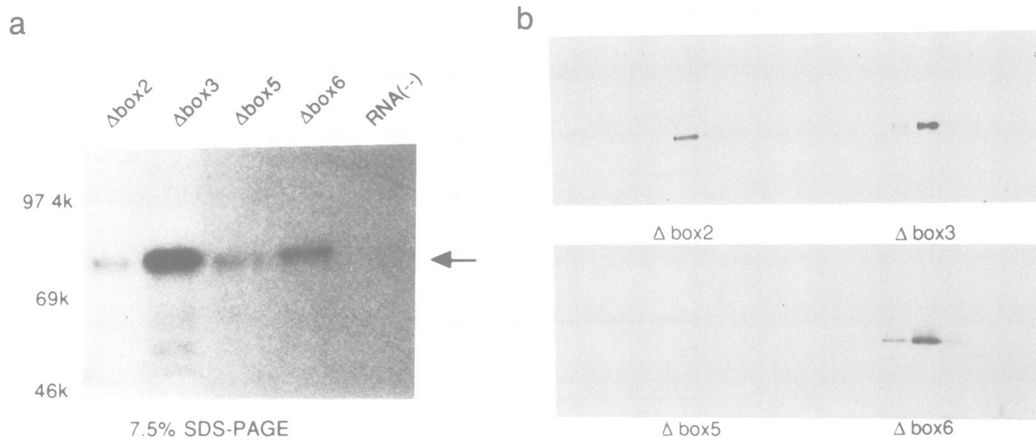

Fig. 7. UCE binding ability of internally deleted mUBF. (**a**) Four internally deleted mUBFs were synthesized by *in vitro* transcription–translation and electrophoresed on a 7.5% SDS–polyacrylamide gel. The arrow indicates synthesized protein. (**b**) Eluate fractions from UCE affinity column were electrophoresed on a 7.5% SDS–polyacrylamide gel. All four kinds of internally deleted protein can bind to the UCE affinity resin.

binding affinity became 50 times higher than that with less than three HMG-boxes. Although a dimerization motif was found in the N-terminal region of *Xenopus* UBF (xUBF) (McStay *et al.*, 1991b), no retarded band corresponding to a dimer complex could be seen in our gel retardation assay. The results of the gel retardation assay indicate that HMG-box1 is necessary for UCE binding and that other HMG-boxes act to increase the total DNA binding ability. The presence of multiple HMG-boxes in UBF may be an evolutionary adaptation for more efficient binding.

Nucleolar localization ability and DNA binding ability do not require conformational integrity

Two possibilities may be considered for the mechanism of nucleolar targeting. One (the active mechanism) is the presence of a carrier protein and the other (the passive mechanism) is that UBF is sequestered in the nucleolus just by virtue of its strong binding to rDNA. In an attempt to address this question, we made a series of internally deleted molecules of mUBF, and tested them for nucleolar localization and specific binding to UCE. All the deletions had the HMG-box1 and four other HMG-boxes; the HMG-box4 was not removed because it contained the NLS. As shown in Figure 6b, all of these internally deleted UBFs could accumulate in the nucleolus. Specific binding to the UCE was also preserved (Figure 7). Since various structures containing both deduced α-helix and β-turn had been removed, some of these molecules should have different three-dimensional structures from the original mUBF. These structural changes might be expected to disturb the ability of UBF to be recognized by some carrier proteins or chaperones, if such a mechanism existed. However, our results have demonstrated that nucleolar targeting is not inhibited by such deletions, strongly suggesting that such a mechanism may not exist.

It has been demonstrated that hUBF is bound to the rDNA promoter rather weakly, but SL-1 strongly enhances the binding, as evidenced by enhanced DNase I protection in the UBF binding site and the extension of the protected region (Learned *et al.*, 1986; Bell *et al.*, 1988). It has also been indicated that the interaction between xUBF and Rib1, the *Xenopus* counterpart of SL-1, is necessary for stable complex formation on the rDNA promoter (McStay *et al.*, 1991a). Furthermore, it is speculated that the acidic tail of UBF might serve as an activating domain, interacting with other polymerase I transcription factors, by analogy with the polymerase II example (Jantzen *et al.*, 1990). Although distinct regions in xUBF that interact with Rib1 have not yet been identified, it has also been shown that the C-terminal half of xUBF contributed to transcriptional activity (McStay *et al.*, 1991b). We have also shown that this region is important for UBF not only in transcription but also in nucleolar targeting. Interestingly, hamster UBF shows a dynamic transition between the nucleolus and nucleus according to the growth rate as a consequence of changing the serum concentration, and UBF is phosphorylated by casein kinase II *in vitro* in the C-terminal acidic tail (O'Mahony *et al.* 1992a). These results suggest that the nucleolar targeting of the UBF may also be regulated by its modification.

As a whole, there is a clear correlation between the UCE binding ability, the transcriptional activity and the nucleolar targeting ability of mUBF. This correlation indicates that the events necessary for transcription initiation are also required for nucleolar targeting. Therefore, we conclude that UBF–SL-1 complex formation and its strong binding affinity for rDNA as a result of this interaction are necessary for nucleolar targeting.

Since a nucleolus is not separated by any membranous structure from the surrounding nucleoplasm, any protein having specific affinity for the nucleolar components could easily reach and accumulate in the nucleolus. We suggest that there may not be a specific nucleolar targeting signal or machinery for some physiological nucleolar proteins such as UBF, but that these proteins accumulate in the nucleolus by virtue of the NLS and the specific binding capacity to the nucleolar components including rDNA.

Materials and methods

Expression constructs for mammalian cells

Serial mUBF–lacZ expression plasmids were constructed based on pCH110 (Pharmacia). mUBF deletion derivatives made by the polymerase chain reaction (PCR) were ligated to the N-terminus of the TrpS–LacZ region of pCH110 (Figure 1a). For efficient translation, we designed a 5′ primer with a Kozak consensus sequence (Kozak, 1989; Lotteau and Peterson, 1990) and a *Hin*dIII restriction site. For in-frame conjunction between mUBF and TrpS, we designed a 3′ primer with a *Kpn*I restriction site. The primers used for PCR amplification are shown in Table I.

Table I.

Mutant name	Forward primer	Backward primer
N-1-lacZ	primer 1	primer 6
N-3-lacZ	primer 1	primer 7
N-preA-lacZ	primer 1	primer 8
1-C-lacZ	primer 2	primer 9
2-C-lacZ	primer 3	primer 9
3-C-lacZ	primer 4	primer 9
5-C-lacZ	primer 5	primer 9
Basic-lacZ	primer 5	primer 10

primer 1: 5'-GGTAAGCTTTCTAGACCATGAACGGAGAAGCG-3'
primer 2: 5'-CGGAAGCTTTCTAGACCATGGGCAAAAAA-CTCAAG-3'
primer 3: 5'-CGGAAGCTTTCTAGACCATGGAGTTCGAG-CGAAAC-3'
primer 4: 5'-CGGAAGCTTTCTAGACCATGAAGCACCCT-GAGCTG-3'
primer 5: 5'-GGTAAGCTTTCTAGACCATGACCGAGAAG-AAGAAG-3'
primer 6: 5'-GATTGGTACCTCGGGGATGTCCGA-3'
primer 7: 5'-GATTGGTACCCTGCCACCTTCCTG-3'
primer 8: 5'-CTTGGTACCTTTGACTGCAGGGTG-3'
primer 9: 5'-GAGGTACCGAGGCTGAGCCTAAGTTGGAGTC-3'
primer 10: 5'-GATTGGTACCTTGCCCCTATCTTC-3'

Table II. Primers used for overlap-extension PCR

Mutant name	Inside backward primer	Inside forward primer
Δbox2-lacZ	primer 11	primer 15
Δbox3-lacZ	primer 12	primer 16
Δbox5-lacZ	primer 13	primer 17
Δbox6-lacZ	primer 14	primer 18

primer 11: 5'-CTCAAAGTGCGGCCGGATTCCACCCTCACC-AAGGCC-3'
primer 12: 5'-CTTCTCGGAGCCACCCCTCTTCTCCTTCTG-AGAGAG-3'
primer 13: 5'-GAACTTCTGGTAACCGTTGTTCCAAGTCAT-CTCCAT-3'
primer 14: 5'-GTCCTGGGGAGACAGGCTCATAGG-AGGTTTCTTGGG-3'
primer 15: 5'-ATCCGGCCGCACTTTGAG-3'
primer 16: 5'-AGGGGTGGCTCCGAGAAG-3'
primer 17: 5'-AACGGTTACCAGAAGTTC-3'
primer 18: 5'-AGCCTCGAACCAGTACAA-3'

All C-terminal primers contain TAA instead of the intrinsic stop codon TCA and a part of the non-translated sequence of mUBF for efficient and precise PCR. Amplification was carried out as follows; first two cycles (30 s denaturation at 95°C, 90 s annealing at 45°C, 3 min elongation at 72°C), then 25 cycles (30 s denaturation at 95°C, 90 s annealing at 65°C, 3 min elongation at 72°C) and finally 5 min elongation at 72°C.

In the case of the construction of internally deleted UBF, we devised a modified PCR method, using overlap-extension PCR. To delete an internal region, two neighboring fragments encompassing a deleted region were first amplified. One of the inside primers overlapped another inside primer. Each fragment was then purified to eliminate the primers. The two fragments were mixed and a PCR reaction was carried out without primers once only. This step creates an internally deleted fragment. The final PCR was performed under usual PCR conditions between N-terminal and C-terminal primers. A part of the mixture containing the internal deletion fragment was added to this final PCR mixture as template. The amplified internal deletion fragment was digested with *Hin*dIII and *Kpn*I, then ligated to pCH110. The sequence was verified by T7 polymerase sequencing. PCR amplifications were carried out with the primer combinations shown in Table II. Primers 2 and 9 were used among four constructs as N-terminal and C-terminal primers, respectively.

Identification of the molecular weight and the subcellular localization of mUBF–β-galactosidase fusion proteins in COS7 cells

Fusion proteins were transiently expressed in COS7 cells by the Chen–Okayama procedure (Chen and Okayama, 1987). To determine the M_r and the subcellular localization of the expressed protein, Western blotting and immunohistostaining were carried out by conventional methods (Harlow and Lane, 1988).

Binding assay using sequence affinity chromatography (Turner and Tjian, 1989)

To synthesize N-Box1, Box1, Box2-C, Δbox2, Δbox3, Δbox5 and Δbox6 proteins *in vitro*, pN-1-lacZ, p1-C-lacZ, p2-C-lacZ (Figure 1a), pΔbox2-lacZ, pΔbox3-lacZ, pΔbox5-lacZ and pΔbox6-lacZ constructs (Figure 6a) were digested with *Xba*I and *Kpn*I, then UBF fragments were ligated to pBluescript KS$^+$. N-Box1, Box2-C and four kinds of internal deletion in pBluescript were digested with *Kpn*I and then treated with T4 DNA polymerase to blunt the 3'-protruding ends. For Box1 protein, cloned 1-C fragment was digested with *Ava*I immediately downstream of HMG-box1. These templates were transcribed by T7 RNA polymerase *in vitro*. The reaction was carried out according to the instruction manual for the Stratagene mRNA capping kit. Each synthesized RNA was translated in rabbit reticulocyte lysate (Promega) with [^{35}S]methionine according to the manufacturer's instructions, and the synthesized labeled protein was electrophoresed on a 7.5, 15 or 20% SDS–polyacrylamide gel (Figures 3b and 7). These proteins were mixed with 1 μg poly[d(AT)d(AT)] and 0.05 μg sonicated salmon sperm DNA as nonspecific competitor. Then they were loaded on a 200 μl UCE affinity chromatography column. Because mUBF has 98% similarity to human UBF, its ability to bind human UCE reflects that of hUBF to human UCE. The column buffer was TM$^+$ (Bell *et al.*, 1988: 50 mM Tris–HCl, pH 7.9, 12.5 mM $MgCl_2$, 1 mM EDTA, 1 mM dithiothreitol and 20% glycerol) containing 0.1 M KCl (TM$^+$/0.1 M KCl). The affinity column was washed with 1 ml TM$^+$/0.1 M KCl and the bound proteins were eluted with TM$^+$/0.6 M KCl. Aliquots of each wash and eluate fraction were analyzed on a 7.5, 15 or 20% SDS–polyacrylamide gel. After fixation with 10% acetic acid and 10% methanol for 15 min, gels were soaked in Enlightening (NEN Research Products) for another 15 min, dried and subjected to autoradiography.

Table III.

Mutant name	Forward primer	Backward primer
N-1	primer 19	primer 22
N-2	primer 20	primer 23
N-3	primer 20	primer 24
N-4	primer 20	primer 25
Box2–Box5	primer 21	primer 26

primer 19: 5'-GTTGTCGACATGAACGGAGAAGCGGAC-3'
primer 20: 5'-GGTGAATTCATGAACGGAGAAGCG-3'
primer 21: 5'-GGTGAATTCATGAACGGAGAAGCG-3'
primer 22: 5'-CCGAAGCTTTTAGTCCGACTTCTTGGC-3'
primer 23: 5'-CCTAAGCTTTTACTCGGCCTTGGTGAG-3'
primer 24: 5'-CCTAAGCTTTTAGCCACCTTCCTGTGA-3'
primer 25: 5'-CGGAAGCTTTTAATAGTCTCCGATGAC-3'
primer 26: 5'-CCGAAGCTTTTATCCCTGGAACTTCAT-3'

Construction of expression plasmids for bacteria

We used the maltose binding protein (MBP) fusion system for expressing mUBF in bacteria. For this purpose, we used the pMAL-c or pMAL-cRI expression vectors and TB-1 host cells (New England Biolabs). Five kinds of mUBF derivative were made by PCR (Figure 4a). Primer combinations were as shown in Table III. All 3' primers contained the stop codon. Each PCR product was ligated to the C-terminal portion of MBP (Figure 4a). PCR amplification was carried out as described above. N-1 fragment was digested with *Sal*I and *Hin*dIII, and then ligated between the *Sal*I and *Hin*dIII sites of pMAL-c. N-2, N-3, N-4 and Box2-Box5 fragments were digested with *Eco*RI and *Hin*dIII and then ligated to pMAL-cRI between the *Eco*RI and *Hin*dIII sites.

Preparation of the purified expressed MBP–mUBF fusion protein

Expression of the MBP–mUBF fusion protein was carried out according to the manufacturer's instructions. The culture size was 200 ml. During extraction of the fusion protein, a protease inhibitor cocktail (1 mM PMSF, 2 μM pepstatin A, 0.6 μM leupeptin, 2 mM benzamidine, 2 μg/ml chymostatin) was added to the extraction buffer. 10 ml crude extract was finally prepared, part of which was electrophoresed on an SDS–polyacrylamide gel.

5 ml amylose resin (New England Biolabs, Inc.) was added to the crude extract in a 50 ml Falcon tube, which was rotated gently at 4°C for 5 h. PMSF was added every hour to keep the final concentration to 1 mM, and the resin was pelleted by centrifugation. The resin was washed thoroughly with 5 l column buffer (10 mM phosphate, 0.5 M NaCl, 1 mM EGTA), the first liter containing 0.25% Tween-20 and the subsequent 4 l without Tween-20. After this wash, 5 ml elution buffer (10 mM maltose-containing column buffer) was added to the resin and centrifuged, and the supernatant was collected. This elution was repeated twice. The eluate was concentrated using a Centricut mini-V20 (Kurabo) at 3000 *g* for 3 h. The concentrated samples were electrophoresed on an SDS–polyacrylamide gel. Purified fusion protein was stored frozen at −110°C.

Gel retardation analysis and Scatchard analysis

The oligonucleotides for the UCE probe, 5′-GGGGTCCGTGTCGCGC-GTCGCCTGGGCCGGCGGCG-3′ and 5′-GGGCGCCGCCGGCCCA-GGCGACGCGCGACACGGAC-3′, were synthesized using an Applied Biosystems 381A DNA synthesizer. 1 μl purified extract was incubated with 1 ng labeled UCE in 20 μl gel retardation mixture [12 mM HEPES–KOH (pH 7.9), 60 mM KCl, 12% glycerol, 0.3 mM DTT, 8 mM spermidine, 5 μg BSA] containing 0.5 μg poly(dA-dT)poly(dA-dT) as non-specific competitor. After a 1 h incubation at 25°C, the mixture was loaded on to a 6% polyacrylamide gel and electrophoresed in 6.7 mM Tris–HCl (pH 7.9), 1 mM EDTA, 3.3 mM sodium acetate at 4°C with buffer circulation.

In Scatchard analysis, the gel retardation assay was carried out as described above. Various concentrations of labeled UCE were added to 1 μl of each protein. The free and the specific retarded bands were cut out from the gel and counted in Econofluor scintillation fluid. The results were plotted according to the Scatchard method. The best-fitted line was obtained using a Cricket graph program (Figure 5).

Acknowledgements

We thank Dr Higashi for comments on the manuscript and Dr T.Sakurai for technical advice on photography.

References

Bachvarov,D. and Moss,T. (1991) *Nucleic Acids Res.*, **19**, 2331–2335.
Ballard,D.W., Walker,W.H., Doerre,S., Sista,P., Molitor,J.A., Dixon,E.P., Peffer,N.J., Hannink,M. and Greene,W.C. (1990) *Cell*, **63**, 803–814.
Bell,S.P., Learned,R.M., Jantzen,H.M. and Tjian,R. (1988) *Science*, **241**, 1198–1197.
Bell,S.P., Pikkard,C.S., Reeder,R.H. and Tjian,R. (1989) *Cell*, **59**, 489–497.
Chen,C. and Okayama,H. (1987) *Mol. Cell. Biol.*, **7**, 2745–2752.
Clos,J., Buttgereit,D. and Grummt,I. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 604–608.
Cochrane,A.W., Perkins,A. and Rosen,C. A. (1990) *J. Virol.*, **64**, 881–885.
Comai,L., Tanase,N. and Tjian,R. (1992) *Cell*, **68**, 965–976.
Freeman,J.W., Busch,R.K., Gyorkey,F., Gyorkey,P., Ross,B.E. and Busch,H. (1988) *Cancer Res.*, **48**, 1244–1251.
Ghosh,S., Gifford,A.M., Riviere,L.R., Tempst,P., Nolan,G.P. and Baltimore,D. (1990) *Cell*, **62**, 1019–1029.
Grummt,I., Roth,E. and Paule,M. (1982) *Nature*, **296**, 173–174.
Harlow,E. and Lane,D. (1988) *Antibodies, A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Hauber,J., Malim,M.H. and Cullen,B.R. (1989) *J. Virol.*, **63**, 1181–1187.
Hisatake,K., Nishimura,T., Maeda,Y., Hanada,K., Song,C.-Z. and Muramatsu,M. (1991) *Nucleic Acids Res.*, **19**, 4631–4637.
Jantzen,H.M., Admon,A., Bell,S.P. and Tjian,R. (1990) *Nature*, **344**, 830–836.
Kalderon,D., Richardson,W.D., Markham,A.F. and Smith,A.E.(1984a) *Nature*, **311**, 33–38.
Kalderon,D., Roberts,B.L., Richardson,W.D. and Smith,A.E. (1984b) *Cell*, **39**, 499–509.
Kato,H., Nagamine,M., Kominami,R. and Muramatsu,M. (1986) *Mol. Cell. Biol.*, **6**, 3418–3427.
Kieran,M., Blank,V., Logeat,F., Vandekerckhove,J., Lottspeich,F., Bail,O.L., Urban,M.B., Kourilsky,P., Baeuerle,P.A. and Israël,A. (1990) *Cell*, **62**, 1007–1018.
Kozak,M. (1989) *J. Cell. Biol.*, **108**, 229–237.
Lapeyre,B., Bourbon,H. and Amalric,F. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1472–1476.
Learned,R.M., Learned,T.K., Haltiner,M.M. and Tjian,R.T. (1986) *Cell*, **45**, 847–857.
Loewinger,L. and McKeon,F. (1988) *EMBO J.*, **7**, 2301–2309.
Lotteau,D. and Peterson,P. A. (1990) *Nature*, **348**, 600–605.
McStay,B., Hu,C.H., Pikkard,C.S. and Reeder,R.H. (1991a) *EMBO J.*, **10**, 2297–2303.
McStay,B., Frazier,W.M. and Reeder,R.H. (1991b) *Genes Dev.*, **5**, 1957–1968.
Mishima,Y., Financsek,I., Kominami,R. and Muramatsu,M. (1982) *Nucleic Acids Res.*, **10**, 6659–6670.
O'Mahony,D.J. and Rothblum,L.I. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 3180–3184.
O'Mahony,D.J., Xie,W., Smith,S.D., Singer,H.A. and Rothblum,L.I. (1992a) *J. Biol. Chem.*, **267**, 35–38.
O'Mahony,D.J., Smith,S.D., Xie,W. and Rothblum,L.I. (1992b) *Nucleic Acids Res.*, **20**, 1301–1308
Richardson,W.D., Roberts,B.L. and Smith,A.E. (1986) *Cell*, **44**, 77–85.
Rosen,C.A. (1991) *Trends Genet.*, **7**, 9–14.
Ruben,S., Perkins,A., Purcell,R., Joung,K., Sia,R., Burghoff,R., Haseltine,W. and Rosen,C. (1989) *J. Virol.*, **63**, 1–8.
Schmidt-Zachmann,M.S., Hugle-Dorr,B. and Franke,W.W. (1987) *EMBO J.*, **6**, 1881–1890.
Silver,P.A. (1991) *Cell*, **64**, 489–497.
Simoi,H., Shida,H., Nam,S.H., Nosaka,T., Maki,M. and Hatanaka,M. (1988) *Cell*, **55**, 197–209.
Tanaka,N., Kato,H., Ishikawa,Y., Hisatake,K., Tashiro,K., Kominami,R. and Muramatsu,M. (1990) *J. Biol. Chem.*, **256**, 13836–13842.
Tower,J., Culotta,V.C. and Sollner-Webb,B. (1986) *Mol. Cell. Biol.*, **6**, 3451–3462.
Turner,R. and Tjian,R. (1989) *Science*, **243**, 1689–1694.

Received on April 30, 1992, revised on June 11, 1992