The EMBO Journal vol.10 no.8 pp.2297 - 2303, 1991

xUBF and Rib1 are both required for formation of a stable polymerase I promoter complex in *X.laevis*

Brian McStay, Chin Hwa Hu, Craig S.Pikaard[1] and Ronald H.Reeder

Hutchinson Cancer Research Center, Basic Sciences Division, 1124 Columbia Street, Seattle, WA 98104, USA

[1]Present address: Department of Biology, Washington University in St Louis, St Louis, MO 63130-6860, USA

Communicated by H.Eisen

We show that three protein fractions are required for accurate transcription initiation at a *Xenopus laevis* ribosomal gene promoter *in vitro*: RNA polymerase I, Rib1 and xUBF. The Rib1 and xUBF fractions are both necessary and sufficient for formation of a stable initiation complex. The xUBF fraction can be completely replaced by recombinant xUBF. We also report the sequence of a cDNA clone for xUBF. xUBF is 701 amino acids in length, contains domains which are related to a domain found in chromosomal proteins HMG 1 and 2, and has an acidic carboxy terminus of 87 amino acids. xUBF is closely similar in amino acid sequence to its previously reported human homolog, hUBF, except that xUBF has only three of the HMG-related domains while hUBF has four and therefore is 63 amino acids longer than xUBF.

Key words: Rib1/RNA polymerase/stable complex formation/xUBF

Introduction

Fractionation of various mammalian RNA polymerase I transcription systems has shown the requirement for a fraction that elutes from negatively charged ion exchange columns at relatively high salt, and which in some cases has been shown to transfer species specificity. This fraction has been variously called SL1 (human, Learned *et al.*, 1985; mouse, Bell *et al.*, 1990; and rat, Smith *et al.*, 1990), Factor D (mouse, Mishima *et al.*, 1982; Tower *et al.*, 1986), and TIF-IB (mouse, Schnapp *et al.*, 1990). In this paper we show that an analogous fraction, which we call Rib1, is a required component of a polymerase I transcription system derived from an amphibian, *Xenopus laevis*.

Previous work on human (Bell *et al.*, 1988), frog (Pikaard *et al.*, 1989; Dunaway, 1989), mouse (Bell *et al.*, 1990) and rat (Smith *et al.*, 1990) systems has also shown the involvement of a second factor, called UBF, in polymerase I transcription. The sequence of a cDNA for hUBF has been reported (Jantzen *et al.*, 1990), and in this paper we report and compare the homologous sequence for xUBF.

Having the cDNA for xUBF has allowed us to translate this protein *in vitro* and show that recombinant xUBF can completely replace the xUBF fraction in a reconstituted transcription system. Furthermore, we have been able to show that both xUBF and Rib1 are necessary and sufficient for formation of the stable initiation complex at a ribosomal gene promoter.

These experiments contribute to the growing consensus that RNA polymerase I transcription machinery has been strongly conserved, at least throughout the vertebrates.

Results

Fractionation of the S-100 extract

An S-100 extract from *X.laevis* kidney cells was fractionated by ion exchange chromatography according to the scheme described in Figure 1 and in Materials and methods. Several points about this fractionation scheme are worth noting. The extract was first applied to a DEAE-Sepharose column and all of the transcription initiation machinery was eluted with a single step of 350 mM KCl (McStay and Reeder, 1990b). From previous work we know that Rib2, a protein factor required for termination of transcription by *X.laevis* polymerase I, elutes in the flow through of such a DEAE Sepharose column. Since all of the transcription templates used in this study were closed circles, they were UV irradiated as previously described, to provide artificial termination sites and prevent promoter occlusion due to polymerase read through around the plasmid circle (McStay and Reeder, 1990a).

The second ion exchange column was Heparin Ultrogel which divides the transcription machinery into two complementing fractions (column profile shown in Figure 2A). The fraction which we call Rib1 elutes from Ultrogel

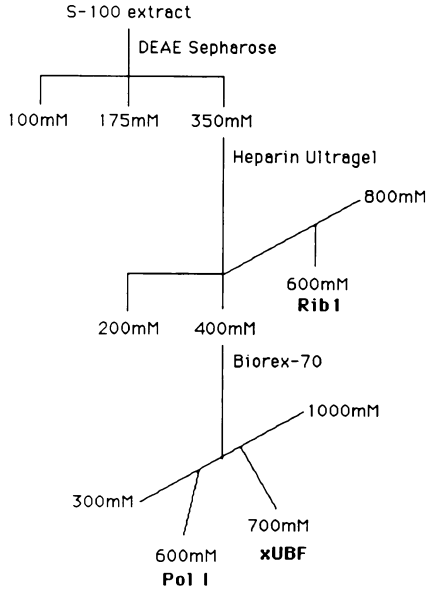

Fig. 1. Scheme for fractionating RNA polymerase I transcription factors from an *X.laevis* extract. Details are described in Materials and methods.

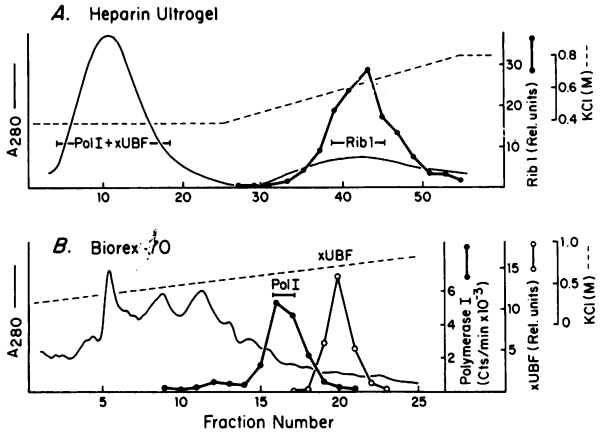

Fig. 2. Chromatographic separation of polymerase I transcription factors. **A.** Heparin Ultrogel separates RNA polymerase I and xUBF from Rib1. The CB350 eluate from DEAE−Sepharose was chromatographed on Heparin Ultrogel and Rib1 activity was assayed as described in Materials and methods. Fractions 39−45 were pooled and used as Rib1 in all later experiments. **B.** RNA polymerase I and xUBF are resolved on Biorex 70. The fraction that eluted from Heparin with 400 mm KCl was chromatographed on Biorex 70 and assayed for either polymerase I or xUBF as described in Materials and methods. Fractions 16 and 17 were pooled and used as polymerase I in all later experiments.

at relatively high salt and is free of any detectable polymerase I or xUBF activity. This is the Rib1 fraction which was used for all other experiments shown in this paper. In other experiments (not shown) we have found that this Rib1 fraction can be further chromatographed on Mono Q and still elutes as a single peak. However, the Rib1 fraction still contains multiple DNA binding activities at this point and we have no direct evidence as to how many polypeptides comprise the Rib1 activity. Elution of human extracts from heparin−agarose with high salt yields a fraction that has been called SL1 (Learned *et al.*, 1985). Evidence presented below suggests that Rib1 is the frog homolog of human SL1 although SL1 appears to elute from heparin at a higher salt than does Rib1.

The 400 mM step eluate from the heparin column was further chromatographed by eluting with a salt gradient from Biorex-70 to separate polymerase I and xUBF (shown in Figure 2B). RNA polymerase I elutes just before xUBF and by pooling the peak fractions, a polymerase preparation can be obtained that is essentially free of either xUBF or Rib1. This is the polymerase preparation that was used for all further reconstitution experiments.

xUBF from the Biorex-70 column can be purified to homogeneity by further chromatography on Mono Q (data not shown). Alternatively, homogeneous xUBF can be purified from nuclear extracts via a slightly different set of ion exchange columns (see Materials and methods and Figure 4). xUBF purified from either S-100 or nuclear extracts will completely replace the impure Biorex-70 fraction (data for xUBF from nuclear extracts is shown in Figure 3). In addition, the Biorex-70 fractions can be replaced by recombinant xUBF made by *in vitro* translation of a synthetic mRNA (Figure 5).

Reconstitution of specific promoter recognition from the separated fractions

Figure 3 shows the result when polymerase I, Rib1 and xUBF are tested, singly and in various combinations, for

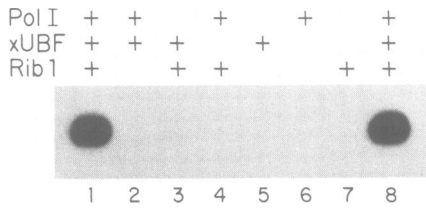

Fig. 3. Polymerase I, xUBF and Rib1 are all required for recognition of a ribosomal gene promoter. *In vitro* transcription reactions shown in lanes 1 and 8 contained 10 μl polymerase I, 10 μl Rib1, 2 μl xUBF (purified from nuclei) and 200 ng of pGEM40 as template. In the reaction shown in lanes 2−7, various of these reaction components have been omitted. In each case the same reaction volume was maintained by the addition of CB100. The products of each transcription reaction were analyzed by S1 nuclease protection.

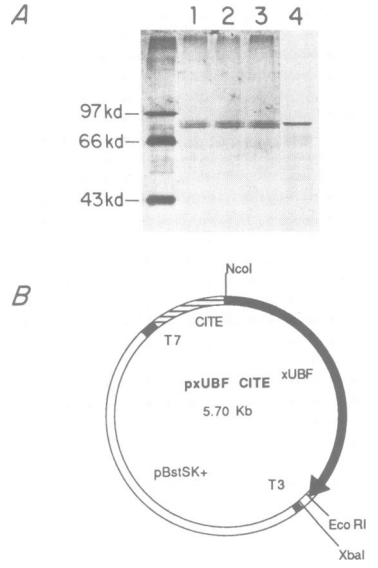

Fig. 4. *In vitro* translation of xUBF. **A.** Comparison of natural xUBF with the product of *in vitro* translation. About 100 ng of purified xUBF (lanes 1, 2 and 3) and 2 μl of an *in vitro* translation reaction that was programmed with transcripts of pxUBF CITE (lane 4) were electrophoresed side by side on a 7.5% SDS−polyacrylamide gel. The gel was silver stained and then autoradiographed. Molecular weight markers were used to align the silver stained gel and the autoradiograph. **B.** Structure of pxUBF CITE. Solid box with arrowhead represents xUBF coding sequences, open box is pBluescript vector, cross-hatched box is the cap independent translation enhancer (CITE). Promoters for phage T7 and T3 polymerases are shown as small solid boxes, and the positions of relevant restriction sites are shown.

the ability to initiate transcription at the *X.laevis* ribosomal gene promoter. For this experiment the xUBF was a homogeneous preparation (a silver-stained gel is shown in Figure 4) purified from nuclear extracts. The general conclusion to be drawn from Figure 3 is that none of the three fractions is significantly contaminated with the other fractions, and all three are required for specific initiation. In contrast to mammalian transcription systems, the reconstituted *X.laevis* system is strongly dependent upon the upstream promoter domain (extending to 5′−142), as we have previously observed in both whole oocytes and in crude *in vitro* extracts (Reeder *et al.*, 1987; McStay and Reeder, 1990a). Mutation of the upstream domain causes at least a 100-fold decrease in initiation (data not shown).

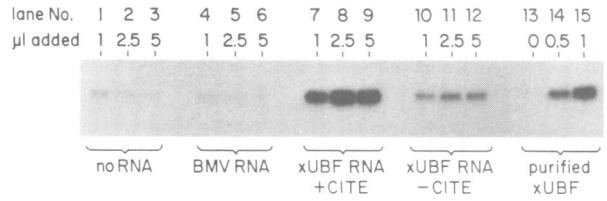

Fig. 5. xUBF translated *in vitro* can substitute in transcription reactions for xUBF purified from nuclei. *In vitro* transcription reactions contained PolI (10 µl), Rib1 (10 µl), 200 ng of pGEM40 as template and aliquots of various *in vitro* translation reactions or xUBF purified from nuclei. The volume of translation reaction (lanes 1−12) or purified xUBF (lanes 13−15) added to transcription reactions, is shown above each lane. Translation reactions contained no RNA (lanes 1−3), brome mosaic virus RNA (lanes 4−6), T7 RNA polymerase transcripts of xUBF CITE (lanes 7−9) or T7 RNA polymerase transcripts of pxUBF (lanes 10−12). A silver stained SDS−polyacrylamide gel of the xUBF used in lanes 14 and 15 is shown in Figure 4A.

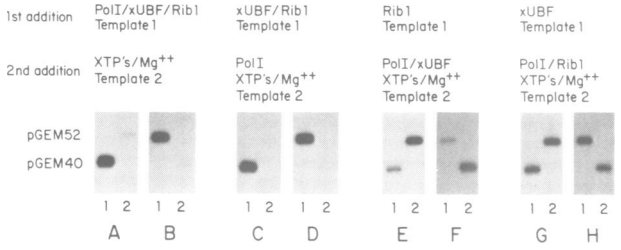

Fig. 6. xUBF and Rib1 are necessary and sufficient for formation of a stable promoter complex. **A.** RibI (10 µl), PolI (10 µl) and xUBF purified from nuclei (2 µl) were combined and incubated at room temperature for 10 min with 200 ng of template 1 (pGEM40). Then 20 µl of reaction buffer containing $MgCl_2$ and nucleoside triphosphates (see Materials and methods) and 200 ng of template 2 (pGEM52) were added and the mixture was incubated at 25°C for 60 min. Transcription was assayed by S1 protection using two probes of equal specific activity. One probe yields a shorter protected fragment from transcripts of pGEM40 (see lane 1 in panel A) while the other probe yields a longer protected fragment from transcripts of pGEM52 (see lane 1 in panel B). **B.** Same as A except that pGEM52 was added first and pGEM40 was added second. **C** and **D.** Same as A and B except that xUBF and Rib1 were added with the first template while PolI was added with the second template. **E** and **F.** Same as A and B except that Rib1 was added with the first template while xUBF and PolI were added with the second template. **G** and **H.** Same as A and B except that xUBF was added with the first template while PolI and Rib1 were added with the second template.

Recombinant xUBF is fully active in reconstituting specific promoter recognition

We have previously reported the purification of xUBF from extracts of *X. laevis* kidney cells (Pikaard *et al.*, 1989). The final step in that initial purification scheme was affinity chromatography on a column containing DNA sequences from the *X. laevis* ribosomal gene enhancers. Subsequently, we have found that the low yield DNA affinity step can be eliminated and pure xUBF can be isolated in much better yield by a series of ion exchange columns ending with chromatography on Mono Q (described in Materials and methods). This allowed isolation of enough xUBF for micro-sequencing. Based on the sequence of an 18 amino acid peptide, we prepared oligonucleotide probes, used them to screen a cDNA library, and isolated a cDNA containing a 2.21 kb insert which contains the entire translated region of xUBF (the nucleotide sequence of this cDNA is shown in Figure 8).

To prove that this cDNA codes for authentic xUBF, the cDNA was used to make a synthetic mRNA, this mRNA was translated *in vitro*, and the *in vitro* translation product was shown to replace authentic xUBF in a reconstituted transcription system. The cDNA insert in the original lambda vector was moved as an *Eco*RI fragment into a pBluescript SK^- plasmid. This plasmid contains a phage T7 promoter which was used to generate synthetic mRNA from the cDNA insert. The synthetic mRNA was then translated *in vitro* using a rabbit reticulocyte translation extract. Aliquots of the translation reaction were then tested for their ability to reconstitute transcription when mixed with RNA polymerase I and Rib1.

As shown in Figure 5, *in vitro* translated xUBF caused a modest stimulation of transcription initiation in the reconstituted system (compare lanes 10−12 with lanes 1−3). However, the amount of transcription stimulation was limited by negative effects of adding too much reticulocyte translation extract to the transcription reaction. To obtain a greater stimulation of transcription, we increased the efficiency of translation of the synthetic mRNA by inserting a 586 bp CITE (Cap Independent Translation Enhancer; Parks *et al.*, 1986; Elroy-Stein *et al.*, 1989) sequence, derived from encephalomyocarditis virus RNA, upstream of the xUBF coding sequence (see Figure 4B for the structure of this construct). By use of the CITE sequence we were able to obtain a level of transcription stimulation equal to or better than that obtained with authentic xUBF purified from cultured cells (Figure 5, compare lanes 7−9 with lanes 13−14).

In Figure 4A we compare the apparent size of *in vitro* translated xUBF with that of natural xUBF purified from an *X. laevis* kidney cell line. We have previously reported that natural xUBF migrates in electrophoresis as two closely spaced bands of apparent molecular weight 82 and 85 kd (Pikaard *et al.*, 1989). In more recent preparations, using a higher resolution gel, we also see a fainter third band (as shown in Figure 4A). The *in vitro* translation product migrates as a single band. Allowing for the fact that the CITE sequence adds four amino acids to the length of *in vitro* translated xUBF, our best estimate is that the *in vitro* product corresponds to the middle band of natural xUBF. We conclude that we are probably translating a full-length xUBF product.

The basis for the apparent size variation in natural xUBF is not yet established. Natural UBF from human, mouse and rat all migrate as two closely spaced bands of apparent size 94 and 97 kd (Bell *et al.*, 1988, 1990; Pikaard *et al.*, 1990a,b). O'Mahony and Rothblum (1991) have recently reported the cloning of two different UBF cDNAs from the rat which differ by an insertion/deletion that could completely account for the two bands of natural rUBF that are observed. It is possible that similar insertion/deletions are present in xUBF and we have cloned a cDNA for only one of the variants. Probing *Xenopus* genomic DNA with probes prepared from the region between HMG domain 4 and the acidic tail of the xUBF cDNA detects at least two copies of the xUBF gene which differ in their flanking sequences (data not shown).

These experiments demonstrate that the cDNA we have isolated codes for a full length, authentic xUBF. They further indicate that the recombinant xUBF is fully active and

requires no modifications that are not supplied by the *in vitro* system. This result also supplies the final proof that xUBF is the active component of the xUBF fraction produced in the fractionation scheme shown in Figure 1.

Requirements for stable promoter complex formation

A variety of direct and indirect experiments lead to the conclusion that ribosomal gene transcription proceeds via formation of a stable complex of transcription factors which remain bound to the promoter through successive rounds of initiation by polymerase I (reviewed recently in Reeder, 1991). Figure 6 shows an experiment in which we determined the fractions required for stable complex formation in the *X. laevis* system. In this experiment we used two ribosomal gene promoters, pGEM40 and pGEM52, which are identical in sequence except that one has a 40 bp linker inserted downstream of the promoter while the other has a 52 bp linker inserted at that site. Thus we can readily distinguish the transcripts from these two promoters by use of the appropriate S1 nuclease protection probe (McStay and Reeder, 1990b).

In Figure 6, panel A, all three factors (pol I, xUBF and Rib1) were added to an excess of template 1 (pGEM40), then template 2 (pGEM52) was added together with the nucleotides and transcription was allowed to proceed for 60 min before assaying the reaction. Transcription only occurred from template 1, indicating that stable complexes had formed on template 1 and no exchange occurred onto template 2. In panel B the same experiment was performed except that pGEM52 was used as template 1 and pGEM40 was used as template 2. Again, the first template to be added pre-empted all of the transcription machinery and essentially no transcription occurred from the second template.

Analogous experiments shown in panels C and D indicate

```
                              |                             |                             |
GCACATTCCTGACCCAGCGTCGCCAGGAGAGACAGAGGAGAGCCGCACTGGGTGTTAAGGGTCTGCTCAGCAGAGGATGAACGGAGCTGCTGGGGGTGATACACAAGGAAAAATGACTGC   120
                                                                            M  N  G  A  A  G  G  D  T  Q  G  K  M  T  A     15

CCCAAAGGATCAAGATCAATGGTCCCAGGAGGACATGCTGACTCTCCTCCAGACCATGAAGACTCTCCTCCCTGGCCAGGACAACTCCAAGTTTAAGACCACCGAGTCCCACCTGGACTG   240
 P  K  D  Q  D  Q  W  S  Q  E  D  M  L  T  L  L  Q  T  M  K  T  L  L  P  G  Q  D  N  S  K  F  K  T  T  E  S  H  L  D  W     55

GAACAAACTGGCCTTCAAACATTATTCTGGTTCCATGTGTCGCCAGAAGTGGATGGAGATTTCCAACGAGGTAAGGAAGTTCCGGACTCTGACAGAGTTGATCCTGGATGCAGACGAACA   360
 N  K  L  A  F  K  H  Y  S  G  S  M  C  R  Q  K  W  M  E  I  S  N  E  V  R  K  F  R  T  L  T  E  L  I  L  D  A  D  E  H     95

CGTGAGACATCCCTACAAAGGCAAGAAGCTGAAGAAACACCCAGAGTTCCCCAAGAAGCCGCTCACCCCTTACTTCCGCTTCTTCATGGAGAAGAGAGCTAAATATGCCAAGCTGCACCC   480
 V  R  H  P  Y  K  G  K  K  L  K  K  H  P  E  F  P  K  K  P  L  T  P  Y  F  R  F  F  M  E  K  R  A  K  Y  A  K  L  H  P    135
                      [- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

CGAGATGAGTAATCTGGATCTCACCAAGATTCTGTCCAAGAAATACAAAGAGCTGCCGGAGAAGAAAAAGATGAAATACATCCAGGACTTCCAGAGGGAGAAGCAGGACTTTGAGAGGAA   600
 E  M  S  N  L  D  L  T  K  I  L  S  K  K  Y  K  E  L  P  E  K  K  K  M  K  Y  I  Q  D  F  Q  R  E  K  Q  D  F  E  R  N    175
- - - - - - - - - - - - HMG domain 1 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

CATGGCTAAATTCAGGGAGGAGCATCCTGATCTCATGCAAAATCCAAAGAAATCTGATGTCCCGGAGAAGCCCAAGACCCCTCAGCAGCTGTGGTACAACCATGAGAGGAAGGTCTACCT   720
 M  A  K  F  R  E  E  H  P  D  L  M  Q  N  P  K  K  S  D  V  P  E  K  P  K  T  P  Q  Q  L  W  Y  N  H  E  R  K  V  Y  L    215
- - - - - - - - - - - - - - - - - - - - - - - - - ][ - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

GAAACTGCATGCTGACGCCAGCACAAAAGACATTAAAGATGCATTAGGGAAGCAGTGGTCCCAGTTACCCGATAAGAAGCGCCTGAAATGGATCCATAAGGCACTGGAACAACGCAAGCA   840
 K  L  H  A  D  A  S  T  K  D  I  K  D  A  L  G  K  Q  W  S  Q  L  P  D  K  K  R  L  K  W  I  H  K  A  L  E  Q  R  K  Q    255
- - - - - - - - - - - - - - - - - - - - - - - HMG domain 2 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

ATATGAGGGTGTAATGAGAGAATACATGCAGAAGCATCCGGAACTGAACATCACAGAAGAGGGCATCACTCGCTCCACTCTCACTAAGGCTGAGCGCCAGCTGAAGGACAAGTTTGATGG   960
 Y  E  G  V  M  R  E  Y  M  Q  K  H  P  E  L  N  I  T  E  E  G  I  T  R  S  T  L  T  K  A  E  R  Q  L  K  D  K  F  D  G    295
- - - - - - - - - - - - - - - - - - - - - - - - - - - -]                                        [- - - - - - - - - - - -

TCGGCCAACCAAGCCTCCACCGAACAGTTACTCGATGTACTGTGCAGAACTAATGGCCAATATGAAGGACGTTCCCAGCACGGAGCGCATGGTCCTCTGCAGCCAAAGGTGGAAGCTGCT   1080
 R  P  T  K  P  P  P  N  S  Y  S  M  Y  C  A  E  L  M  A  N  M  K  D  V  P  S  T  E  R  M  V  L  C  S  Q  R  W  K  L  L    335
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - HMG domain 4 - - - - - - - - - - - - - - - - -

CTCCCAGAAGGAGAAGGACGCCTACCACAAGAAGTGCGAGCAGAGAAAGAAAGACTATGAAGTGGAGTTGATGCGTTTCCTTGAGAATCTCCCAGAAGAGGAGCAGCAGAGGGTTCTGGC   1200
 S  Q  K  E  K  D  A  Y  H  K  K  C  E  Q  R  K  K  D  Y  E  V  E  L  M  R  F  L  E  N  L  P  E  E  E  Q  Q  R  V  L  A    375
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -]

GGAAGAAAAGATGGTGGGTATGAAAAGAAAAAGAACTAACACCCCAGCATCAAAGATGGCCACCGAAGATGCTGCAAAGGTAAAGAGCAGGTCTGGCCAGGCAGACAAGAAGAAAGCAGC   1320
 E  E  K  M  V  G  M  K  R  K  R  T  N  T  P  A  S  K  M  A  T  E  D  A  A  K  V  K  S  R  S  G  Q  A  D  K  K  K  A  A    415

AGAGGAGAGGGCTAAACTGCCAGAGACCCCCAAAACTGCCGAAGAAATCTGGCAGCAAAGTGTGATTGGAGACTATCTTGCTCGATTTAAGAATGATCGCGCTAAAGCCCTGAAGGTCAT   1440
 E  E  R  A  K  L  P  E  T  P  K  T  A  E  E  I  W  Q  Q  S  V  I  G  D  Y  L  A  R  F  K  N  D  R  A  K  A  L  K  V  M    455

GGAGGCCACATGGTTAAATATGGAGAAGAAGGAGAAAATCATGTGGATAAAGAAGGCCGCCGAGGACCAGAAACGCTACGAGAGAGAACTAAGTGATATGCGCTCTACTCCAGCCCCAAC   1560
 E  A  T  W  L  N  M  E  K  K  E  K  I  M  W  I  K  K  A  A  E  D  Q  K  R  Y  E  R  E  L  S  D  M  R  S  T  P  A  P  T    495

GACAGCAGGAAAAAAGGTCAAGTTTTTGGGAGAACCCAAGAAAGCGCCCATGAATGGCTACCAGAAGTTCTCTCAGGAGCTCCTGTCCAACGGAGAACTCAATCACCTCCCCCTTAAGGA   1680
 T  A  G  K  K  V  K  F  L  G  E  P  K  K  A  P  M  N  G  Y  Q  K  F  S  Q  E  L  L  S  N  G  E  L  N  H  L  P  L  K  E    535

ACGTATGGTAGAAATTGGAAGCCGGTGGCACAGGATTTCTCCCACTCAAAAGGATTATTACAAGAAGCTGGCAGAGGACCAACAGAGGCTCTATCGTACTCAGTTTGACACCTGGATGAA   1800
 R  M  V  E  I  G  S  R  W  H  R  I  S  P  T  Q  K  D  Y  Y  K  K  L  A  E  D  Q  Q  R  L  Y  R  T  Q  F  D  T  W  M  K    575

GGGGTTATCGACACAAGACCGTGCAGCATACAAGGAACAGAATACAAATAAACGCAAAAGTACAACAAAAATACAGGCTCCTAGCAGCAAATCCAAATTAGTAATTCAGAGCAAATCGGA   1920
 G  L  S  T  Q  D  R  A  A  Y  K  E  Q  N  T  N  K  R  K  S  T  T  K  I  Q  A  P  S  S  K  S  K  L  V  I  Q  S  K  S  D    615
                                                                                                                      [ -

TGATGACGAAGATGATGAGGATGATGAGGATGAAGAAGACGATGATGACGACGATGATGAAGATAAGGAGGACTCGTCTGAAGACGGAGATTCTTCAGATTCTAGCAGTGATGAGGACAG   2040
 D  D  E  D  D  E  D  D  E  D  E  E  D  D  D  D  D  D  E  D  K  E  D  S  S  E  D  G  D  S  S  D  S  S  S  D  E  D  S    655
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Acidic Tail - - - - - - - - - - - - -

CGAGGAAGGCGAAGAGAATGAAGATGAAGAAGATGAGGAGGACGATGATGAGGACAACGAAGAGGACGATGATGACAATGAATCAGGATCCAGCTCTTCATCATCCTCTTCTGCCGACTC   2160
 E  E  G  E  E  N  E  D  E  E  D  E  E  D  D  D  E  D  N  E  E  D  D  D  D  N  E  S  G  S  S  S  S  S  S  S  S  A  D  S    695
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

TTCTGACTCGGACTCAAACTGAATACATGAGACTTGGGGATGGCTTTG
 S  D  S  D  S  N
- - - - - - - - - ]
```

Fig. 7. Nucleotide sequence and conceptual translation of a cDNA clone for xUBF. The entire nucleotide sequence (2208 nucleotides) of the xUBF cDNA clone (obtained as described in Materials and methods) is shown. Underneath the DNA sequence the deduced amino acid sequence of the largest open reading frame (701 amino acids, single letter code) is shown. Also shown is the location of the HMG boxes 1, 2 and 4 and the acidic tail (see text for details).

that xUBF and Rib1 are able to form the stable complex when added together in the absence of polymerase. Thus we can conclude that nothing hidden in the polymerase fraction is required for stable complex formation and xUBF plus Rib1 are sufficient. Panels E–H show that neither xUBF nor Rib1 are able to cause stable template commitment when added by themselves. We note, in panels E and F, that Rib1 actually depressed transcription when added alone to the first template. This is probably due to the presence of other non-specific DNA-binding proteins still present in the Rib1 fraction. In other experiments (data not shown) we have found that Rib1 further chromatographed on Mono Q (eluting as a single peak) is still competent for stable complex formation.

Comparison of xUBF with hUBF

The nucleotide sequence of the xUBF cDNA and its conceptual translation is shown in Figure 7. The cDNA contains a large open reading frame coding for a protein of 701 amino acids in length with a predicted molecular weight of 81 963 daltons. This compares well with the apparent molecular weight of authentic xUBF as well as the size of the *in vitro* translation product directed by the xUBF cDNA (82–85 kd, see Figure 5).

We have previously observed that xUBF has DNA binding and DNase I footprinting characteristics that are essentially identical to those of UBF from human (Bell *et al.*, 1989), mouse (Pikaard *et al.*, 1990a) and rat (Pikaard *et al.*, 1990b). However, xUBF is ~12 kd smaller than any of the mammalian UBFs. Comparison of the amino acid sequence of xUBF with the previously published sequence of hUBF suggests an explanation for this conservation of DNA binding characteristics. Jantzen *et al.* (1990) noted that hUBF contains four repeats of a domain which is related to a domain present in the chromatosomal proteins HMG1 and 2. In hUBF it appears to be these HMG domains which are responsible for DNA binding. In addition, hUBF has a highly acidic tail at its carboxy terminus similar to the HMG proteins. xUBF is closely related in sequence to hUBF and contains similar domains, as summarized in Figure 8. The major difference is that xUBF lacks one of the HMG domains and has 22 amino acids of unrelated sequence in its place, thus accounting for the fact that xUBF is ~12 kd smaller in size than hUBF. Apparently, deleting one out of four HMG domains is not enough to change the DNA-binding characteristics of the protein significantly.

Except for the deletion of one HMG domain, xUBF is very closely related in sequence to hUBF. In HMG domains 1, 2 and 4, the homology between hUBF and xUBF is 90–97%, if conservative amino acid replacements are allowed. Another striking fact is that, excluding the acidic tail and the deleted HMG domain, the spacing of the rest of the molecule has been precisely preserved. Spacing may be less critical within the acidic tail since in this region there are five individual amino acid insertions or deletions.

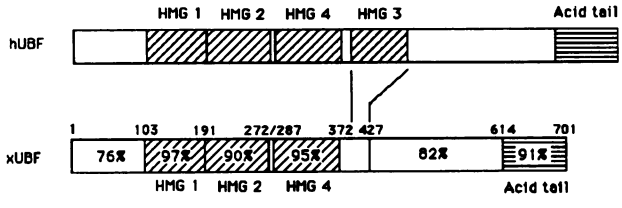

Fig. 8. Comparison of the structures of xUBF and hUBF. hUBF contains four domains related in sequence to a domain present in the chromosomal proteins HMG1 and HMG2 (Jantzen *et al.*, 1990). xUBF lacks one of these domains. With the exception of this deletion, hUBF and xUBF are very closely related in sequence, as shown by the percentages written over each xUBF domain. These percentages were calculated allowing conservative amino acid replacements as follows: (K,R) (E,D) (Q,N) (S,T) (F,Y,W,H) (L,I,V,A) (P,G) (M,C).

Discussion

In this report we show that at least three fractions are required for accurate *in vitro* initiation at an *X.laevis* ribosomal gene promoter: RNA polymerase I, xUBF and Rib1. The active factor in the xUBF fraction is firmly established since it can be replaced by the *in vitro* translation product of xUBF cDNA. The Rib1 fraction is still too impure to say anything about its molecular composition. However, the ability of Rib1 to complement xUBF and polymerase, its involvement in formation of the stable initiation complex, and its chromatographic properties all indicate that Rib1 is the frog homolog of the factors called SL1 (Learned *et al.*, 1985; Bell *et al.*, 1990; Smith *et al.*, 1990). TIF-IB (Schnapp *et al.*, 1990) and Factor D (Mishima *et al.*, 1982; Tower *et al.*, 1986) which have been previously identified in mammalian systems. These results support the conclusion that all of the vertebrates utilize a set of *trans*-acting factors for RNA polymerase I with similar chromatographic properties.

Despite their chromatographic relatedness, it is well known that RNA polymerase I transcription machinery exhibits considerable species specificity (Grummt *et al.*, 1982). Between rodent and human this specificity appears to reside primarily in the SL1–Factor D fraction since transferring that factor from one extract to another can also transfer species specific promoter recognition (Mishima *et al.*, 1982; Bell *et al.*, 1990). Between more distantly related species, such as *X.laevis* and human, it appears that UBF is also species specific. The footprinting characteristics of frog and human UBF cannot be distinguished but xUBF does not exhibit transcription activity in a human system nor vice versa (Bell *et al.*, 1989). Comparison of xUBF with hUBF immediately suggests that the basis for this species difference might be the fact that xUBF and hUBF have different numbers of HMG domains (this paper, Figure 8, and Jantzen *et al.*, 1990). It will be of interest to make chimeras between xUBF and hUBF and see if this prediction is verified. Despite the differences between polymerase I transcription systems from amphibians and mammals, Pape *et al.* (1990) have shown that a frog promoter can be turned into an excellent mouse promoter by the simple expedient of removing five base pairs from the middle of the frog promoter. This suggests that all vertebrate polymerase I transcription systems contain a set of closely related, highly conserved components. However, in each species, the components may have evolved to recognize a different critical spacing of promoter domains.

We also show in this report that xUBF is required for formation of the stable initiation complex at the ribosomal gene promoter (Figure 6). UBF homologs were initially identified in human and frog extracts on the basis of their characteristic DNase I footprints (Bell *et al.*, 1988; Pikaard *et al.*, 1989). However, since UBF is relatively abundant in extracts and chromatographs close to polymerase I, assays

of its transcription activity only showed stimulation from a fairly high basal level. This left open the possibility that UBF might be a stimulatory factor but not be absolutely required for polymerase I transcription. With further refinement of the assay (i.e. making polymerase I more free of UBF) it is now possible to obain 100-fold or greater dependence upon UBF in reconstituted systems, both in the human (Jantzen *et al.*, 1990) and the frog (Figure 3). The further demonstration that UBF is required for stable complex formation provides additional, conclusive evidence that UBF is a required *trans*-acting factor for RNA polymerase I.

It has been previously reported that the mouse homolog of Rib1 (Factor D or TIF-IB) was capable of forming the stable complex unaided (Tower *et al.*, 1986; Clos *et al.*, 1986). It now seems likely that this result was due to the unsuspected contamination of these factor preparations by mUBF.

UBF is a relatively abundant protein with a structure related to the chromosomal proteins HMG 1 and 2. By itself UBF has a relatively low affinity for DNA and binds rather promiscuously. In the *X.laevis* ribosomal DNA we have identified footprinting sites at several locations within the promoter, on both sides of the T3 transcription terminator, and within each of the repetitive enhancer elements in the intergenic spacer (Pikaard *et al.*, 1989). We have also observed xUBF footprinting sites in both the human (Bell *et al.*, 1989) and mouse (Pikaard *et al.*, 1990a) ribosomal gene promoters, in the mouse ribosomal gene enhancers (Pikaard *et al.*, 1990a) and at several locations within the bacterial plasmid, pBR322 (C.S.Pikaard, unpublished). Aligning all of these sites reveals that UBF prefers to bind to GC-rich DNA but the consensus recognition sequence is sufficiently degenerate that it is not readily apparent to visual inspection. It is thus all the more striking that when UBF interacts with Rib1, the combination is able to recognize specifically ribosomal gene promoters to the exclusion of all other promoters in the genome. It will be of great interest to eventually understand how this high degree of promoter specificity is achieved.

Materials and methods

Extracts

S-100 extracts were prepared from a *Xenopus laevis* kidney cell line, X1K-2, as previously described (McStay and Reeder, 1990b). Nuclear extracts were prepared from XlK-2 cells as described in reference (Pikaard *et al.*, 1989).

Fractionation of S-100 extract

All chromatography was performed in column buffer (CB; 25 mM HEPES [pH 7.9], 0.1 mM EDTA, 1mM DTT, 0.5 mM PMSF, 20% glycerol) with the concentration of KCl indicated. 100 ml of S-100 extract in column buffer with 100 mM KCl (CB100) was loaded onto a 50 ml DEAE−Sepharose CL6B column (Pharmacia). The column was then eluted with CB175 followed by CB350. Peak protein fractions that eluted with CB350 were pooled, diluted with an equal volume of CB containing no KCl then loaded onto a 10 ml Heparin Ultrogel column (IBF Biotech). Protein was eluted from the Heparin column with CB200 (20 ml), CB400 (20 ml) and then a 40 ml linear gradient between CB400 and CB800. Peak protein fractions that eluted from heparin with CB400 contained polymerase I and xUBF. These fractions were pooled, diluted to CB300 and loaded onto a 10 ml BioRex 70 column (200−400 mesh, BioRad). Elution of the BioRex 70 column with a 60 ml linear gradient between CB300 and CB1000 separated polymerase I from xUBF (shown in Figure 2B). Chromatography on Heparin Ultrogel and BioRex 70 was performed using FPLC (Pharmacia).

Rib1 activity eluted from the Heparin Ultrogel column in the CB400-CB800 gradient fractions (see Figure 2A). Rib1 activity was assayed by adding 10 μl from every second fraction (dialyzed against CB100) to a transcription reaction that contained 10 μl of the CB400 Heparin Ultrogel fraction (dialyzed against CB100). Peak Rib1 fractions were pooled, dialyzed against CB100 and stored in aliquots at −70°C.

Fractions from the BioRex column were assayed for RNA polymerase I activity using nicked calf thymus DNA template as described (Roeder, 1974). Assays contained 5 μl of each test fraction plus 100 μg/ml α-amanitin. Fractions containing the peak of RNA polymerase I activity were pooled, dialyzed against CB 100 and stored in aliquots at −70°C. xUBF was assayed by adding 5 μl of each fraction (dialyzed against CB100) to a transcription reaction that contained 10 μl polymerase I and 10 μl Rib1.

Purification of xUBF from nuclear extracts

The xUBF used in most transcription reactions was purified from nuclear extracts (Pikaard *et al.*, 1989). 60 ml of extract in CB100 was loaded onto a 10 ml Q Sepharose fast flow column (Pharmacia). The column was eluted with CB250 (20 ml) followed by CB600 (20 ml). Protein-containing fractions which eluted with CB600 were pooled, diluted with CB300, then loaded onto a 10 ml Biorex 70 column, and eluted with a CB300−CB1000 linear gradient. xUBF was located in gradient fractions by DNase I footprinting as described previously (Pikaard *et al.*, 1989). Fractions containing xUBF were pooled and dialyzed against CB100, then loaded onto a Mono Q column (HR 5/5, Pharmacia). The column was washed with CB350 followed by elution of xUBF with aCB350−CB600 linear gradient. Mono Q fractions were analyzed for xUBF by SDS−PAGE (Laemmli, 1970) followed by silver staining. Peak xUBF fractions were pooled and dialyzed against 50 mM HEPES (pH 7.9), 200 mM KCl, 2 mM DTT, 0.5 mM PMSF, 0.2mM EDTA, 40% glycerol, and stored at −20°C. All columns used in the purification of xUBF were run with FPLC.

Overall yield of xUBF starting from 1.12×10^{10} cells (60 ml of nuclear extract) was about 17 μg. With no allowance for losses, this is equivalent to 10^4 molecules per cell.

Plasmids

The plasmids pGEM40 and pGEM52 which were used as templates in transcription reactions have been described elsewhere (McStay and Reeder, 1990a).

The plasmid pxUBF was the substrate for sequencing the xUBF cDNA and for the initial *in vitro* transcription−translation experiment. It contains a full-length xUBF cDNA clone (2.2 kb), sub-cloned as an *Eco*RI fragment into the plasmid vector pBluescript SK$^-$ (Stratagene) oriented such that the 5′ end of the cDNA is adjacent to the promoter for phage T7 RNA polymerase.

The plasmid pxUBF CITE was used for high efficiency *in vitro* translation of xUBF. It was constructed as follows: a unique site for the restriction enzyme *Nco*I (5′-CCATGG-3′) was created at the translation start site of xUBF in the plasmid pxUBF by site-directed mutagenesis with the oligonucleotide 5′-CAGCAGCTCCGTCCATGGTCTGCTGAGCAG-3′. A 586 nucleotide *Eco*RI−*Nco*I fragment from the vector pCITE-I (Novagen) that contains a cap independent translation enhancer (CITE) from encephalomyocarditis virus was fused to the newly created *Nco*I site in xUBF. This CITE−xUBF fusion was then cloned as an *Eco*RI fragment into the vector pBluescript SK$^+$ (Stratagene) oriented such that CITE and the 5′ end of the cDNA are adjacent to the promoter for phage T7 RNA polymerase (see Figure 5B). The first four amino acids (Met-Ala-Thr-Thr) of the predicted translation product from pxUBF CITE are coded for by sequences in CITE. These are followed by the entire coding sequence of xUBF. Note also that due to the creation of the *Nco*I site the second xUBF codon in this xUBF−CITE fusion codes for aspartic acid in place of asparagine.

Isolation and sequencing of xUBF cDNA clone

Approximately 350 pmol (30 μg) of purified xUBF was dialyzed into 50 mM ammonium bicarbonate, lyophilized to a final volume of 200 μl, and sent to Dr William S.Lane at the Harvard Microchemistry Facility for tryptic digestion, HPLC separation of peptides, and amino acid sequence determination. One of the longest peptides analyzed had the squence; Thr-Ala-Glu-Glu-Ile-Trp-Gln-Ser-Val-Ile-Gly-Asp-Tyr-Leu-Ala-Arg (this sequence precisely matches amino acids 427−443 in the conceptual translation of pxUBF shown in Figure 7). Guided by this amino acid sequence and codon usage tables an oligonucleotide with the sequence 5′-GCTGAGGAGATCTGGCAGCAGTCTGTGATTGGCGACTACCTG-GGC-3′ was synthesized and used to screen 4×10^5 primary clones from an *X.laevis* cDNA libary in the vector λgt11. The library was prepared by the method of Gubler and Hoffman (1983) from mRNA isolated from the *X.laevis* culture cell line XlK-2. *Eco*RI adaptors (5′-AATTCGGCACGAG-3′ annealed to 5′-CTCGTGCCG-3′) were ligated onto the blunt ended cDNA to facilitate cloning into the *Eco*RI site of λ-gt11.

The phage giving the strongest hybridization signal was plaque purified and shown to contain an insert of 2.2 kb. This cDNA insert was sub-cloned into the vector pBluescript SK$^-$ (Stratagene) for DNA sequence analysis.

The entire DNA sequence from both strands of the cDNA was determined by first making ordered deletions (Henikoff, 1984, 1990) and then sequencing by the dideoxy method (Sanger *et al.*, 1977) with the enzyme Sequenase (USB).

Transcription assays

Fractions to be tested for transcription were combined on ice. Then 20 μl of a transcription reaction buffer (25 mM HEPES [pH 7.9], 80 mM KCl, 12 mM $MgCl_2$, 1 mM DTT, 1 mM nucleotide triphosphates, 10 mM creatine phosphate and 200 μg/ml α-amanatin) was added plus 20 units of RNasin (Promega) and 200 ng of supercoiled template (pGEM40) that had been irradiated with UV as described previously (McStay and Reeder, 1990a). In the stable complex experiment shown in Figure 6, a second template, pGEM52 was also used and the order of addition of reaction components was as described in the figure legend. All transcription reactions were incubated at 25°C for 1 h then terminated and analyzed by S1 nuclease protection as described previously (McStay and Reeder, 1990a).

In vitro transcription–translation

Both pxUBF and pxUBF CITE were digested with *Xba*I at a unique site in the polylinker downstream of the insert. Then 1 μg of each linearized plasmid was transcribed with phage T7 RNA polymerase (80 U, Boehringer Mannheim) for 1 h at 37°C in a 50 μl reaction that contained 40 mM Tris (pH 8.0), 25 mM NaCl, 8 mM $MgCl_2$, 2mM spermidine, 5 mM DTT, 1mM nucleotide triphosphates, and 40 U RNasin (Promega). The reaction was then extracted with phenol–chloroform and the RNA precipitated by the addition of one-tenth volume of 7.5 M ammonium acetate and 2.5 volumes of ethanol. Transcripts were resuspended in H_2O and stored at −70°C. The integrity of transcripts was monitored by gel electrophoresis.

Synthetic xUBF mRNAs were translated in the rabbit reticulocyte lysate system. Reactions were 50 μl in volume and contained 35 μl of reticulocyte lysate (Promega), 20 μM amino acids, 40 μCi [^{35}S]methionine (1000 Ci/mmol, NEN), 40 U RNasin (Promega) and ~1 μg of synthetic xUBF mRNA. Reactions were incubated at 30°C for 90 min. Aliquots of each reaction were electrophoresed in 10% SDS-polyacrylamide gels (Laemmli, 1970). Following fixing in 40% methanol/10% acetic acid, gels were dried and autoradiographed. Control translation reactions contained no RNA or brome mosaic virus RNA (supplied with the lysate).

Acknowledgements

We thank J.Roan for technical assistance. This work was supported by grant no. GM26624 to R.H.R. from the National Institutes of Health.

References

Bell,S.P., Learned,R.M., Jantzen,H.-M. and Tjian,R. (1988) *Science*, **241**, 1192–1197.

Bell,S.P., Pikaard,C.S., Reeder,R.H. and Tjian,R. (1989) *Cell*, **59**, 489–497.

Bell,S.P., Jantzen,H.-M. and Tjian,R. (1990) *Genes Dev.*, **4**, 943–954.

Clos,J. Buttgereit,D. and Grummt,I. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 604–608.

Dunaway,M. (1989) *Genes Dev.*, **3**, 1768–1778.

Elroy-Stein,O., Fuerst,T.R. and Moss,B. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6126–6130.

Grummt,I., Roth,E. and Paule,M.R. (1982) *Nature*, **296**, 173–174.

Gubler,U. and Hoffman,B.J. (1983) *Gene*, **25**, 263–269.

Henikoff,S. (1984) *Gene*, **28**, 351–359.

Henikoff,S. (1990) *Nucleic Acids Res.*, **18**, 2961–2966.

Jantzen,H.-M., Admon,A., Bell,S.P. and Tjian,R. (1990) *Nature*, **344**, 830–836.

Laemmli,U.K. (1970) *Nature*, **227**, 680–685.

Learned,R.M., Cordes,S. and Tjian,R. (1985) *Mol. Cell. Biol.*, **5**, 1358–1369.

McStay,B. and Reeder,R.H. (1990a) *Genes Dev.*, **4**, 1240–1252.

McStay,B. and Reeder,R.H. (1990b) *Mol. Cell. Biol.*, **10**, 2793–2800.

Mishima,Y., Finanscek,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.

Pape,L.K., Windle,J.J. and Sollner-Webb,B. (1990) *Genes Dev.*, **9**, 5093–5104.

Parks,G.D., Duke,G.M. and Plamenberg,A.C. (1986) *J. Virol.*, **60**, 376–384.

Pikaard,C.S., McStay,B., Schultz,M.C., Bell,S.P. and Reeder,R.H. (1989) *Genes Dev.*, **3**, 1779–1788.

Pikaard,C.S., Pape,L.K., Henderson,S.L., Ryan,K., Paalman,M.H., Lopata,M.A., Reeder,R.H. and Sollner-Webb,B. (1990a) *Mol. Cell. Biol.*, **10**, 4816–4825.

Pikaard,C.S., Smith,S.D., Reeder,R.H. and Rothblum,L. (1990b) *Mol. Cell. Biol.*, **10**, 3810–3812.

Reeder,R.H. (1991) In Yamamoto,K. and McKnight,S.L. (eds), *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbour, NY, in press.

Reeder,R.H., Pennock,D., McStay,B., Roan,J., Tolentino,E. and Walker,P. (1987) *Nucleic Acids Res.*, **15**, 7429–7441.

Roeder,R.G. (1974) *J. Biol. Chem.*, **249**, 241–248.

Sanger,F., Nicklen,S. and Coulsen,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.

Schnapp,A., Clos,J., Hadelt,W., Schreck,R., Cvekl,A. and Grummt,I. (1990) *Nucleic Acids Res.*, **18**, 1385–1393.

Smith,S.D., Oriahi,E., Lowe,D., Yang-Yen,H.-F., O'Mahoney,D., Rose,K., Chen,K. and Rothblum,L.I. (1990) *Mol. Cell. Biol.*, **10**, 3105–3116.

Tower,J., Culotta,V. and Sollner-Webb,B. (1986) *Mol. Cell. Biol.*, **6**, 3451–3462.

Received on February 19, 1991; revised on April 18, 1991

Note added in proof

The sequence data reported here are available from the EMBL/GenBank/DDBJ sequence databases under the accession number X57561.