The EMBO Journal Vol.17 No.7 pp.2019–2032, 1998

UTF1, a novel transcriptional coactivator expressed in pluripotent embryonic stem cells and extra-embryonic cells

Akihiko Okuda, Akiko Fukushima, Masazumi Nishimoto, Akira Orimo, Toshiyuki Yamagishi[1], Yoko Nabeshima[2], Makoto Kuro-o[2], Yo-ichi Nabeshima[2], Kathy Boon[3], Marie Keaveney[3,4], Hendrik G.Stunnenberg[3,5] and Masami Muramatsu[6]

Department of Biochemistry and [1]Department of Anatomy, Saitama Medical School, 38 Morohongo, Moroyama, Iruma-gun, Saitama 350-0495, [2]Department of Molecular Genetics, National Institute of Neuroscience, NCNP 4-1-1, Ogawahigashi Kodaira, Tokyo 187, Japan and [3]EMBL Gene Expression Program, Meyerhofstrasse 1, D-6900 Heidelberg, Germany

[4]Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA
[5]Present address: Department of Molecular Biology and Cell Biology, University of Nijmegen Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

[6]Corresponding author

A.Okuda and A.Fukushima contributed equally to this work

We have obtained a novel transcriptional cofactor, termed undifferentiated embryonic cell transcription factor 1 (UTF1), from F9 embryonic carcinoma (EC) cells. This protein is expressed in EC and embryonic stem cells, as well as in germ line tissues, but could not be detected in any of the other adult mouse tissues tested. Furthermore, when EC cells are induced to differentiate, UTF1 expression is rapidly extinguished. In normal mouse embryos, UTF1 mRNA is present in the inner cell mass, the primitive ectoderm and the extra-embryonic tissues. During the primitive streak stage, the induction of mesodermal cells is accompanied by the down-regulation of UTF1 in the primitive ectoderm. However, its expression is maintained for up to 13.5 days post-coitum in the extra-embryonic tissue. Functionally, UTF1 boosts the level of transcription of the adenovirus E2A promoter. However, unlike the pluripotent cell-specific E1A-like activity, which requires the E2F sites of the E2A promoter for increased transcriptional activation, UTF1-mediated activation is dependent on the upstream ATF site of this promoter. This result indicates that UTF1 is not a major component of the E1A-like activity present in pluripotent embryonic cells. Further analyses revealed that UTF1 interacts not only with the activation domain of ATF-2, but also with the TFIID complex *in vivo*. Thus, UTF1 displays many of the hallmark characteristics expected for a tissue-specific transcriptional coactivator that works in early embryogenesis.

Keywords: ATF-2/E1A-like activity/embryonic stem cells/transcriptional coactivator

Introduction

Embryonic development in mammals begins with fertilization, and while the uterus subsequently prepares to receive the embryo, the early development process proceeds at a relatively slow cellular growth rate. Upon implantation, however, a dramatic increase in the growth rate of embryonic cells occurs, especially in a small number of cells known as the primitive ectodermal cells. The primitive ectoderm (also known as epiblast) is characterized as a pluripotent tissue, as this comprises the sole founder tissue of the fetus, from which both the somatic and germ line tissues are derived (for details see Gilbert, 1991). Therefore, a key issue in elucidating the complex process of early embryogenesis in mammals requires a detailed molecular understanding of several key developmental steps, including the necessary events for maintaining cells in the pluripotent state as well as the factors essential for the promotion of rapid growth of the primitive ectoderm leading to cellular differentiation. Embryonic carcinoma (EC) cell lines such as F9 and P19, are widely used as model systems for studying the regulation of gene expression during the early developmental stages of mammals. These cells retain pluripotent properties and can easily be induced to differentiate into several cellular types upon addition of retinoic acids and/or by aggregate formation (Martin, 1981).

Early studies carried out in EC cells clearly indicated that these cells possess a number of interesting and unique gene regulatory features, for example, Imperiale *et al.* (1984) observed that the heat-shock protein (*hsp*) 70 gene was relatively highly expressed in these cells even when they were grown under non-heat-inducing conditions. As the adenovirus E1A gene product was known to induce *hsp70* gene expression, it was speculated that EC cells may possess a factor similar, or equivalent, to E1A. This idea was further supported by experiments demonstrating that when an adenovirus mutant lacking the E1A region was used to infect EC cells, significantly high levels of expression were observed from the *E2A* gene. Under normal conditions E1A is required for such elevated levels of expression. The factor responsible for these effects was termed the E1A-like activity (E1A-LA); however, to date the identification and isolation of this activity remains to be carried out. Interestingly, the expression of the E1A-LA was shown to be down-regulated upon differentiation of EC cells and it was therefore postulated that it may play a role in maintaining the undifferentiated state of these cells (La Thangue and Rigby, 1987).

In addition to the gene-specific effects seen with *hsp70* and *E2A*, a number of other distinct transcription factors including Oct-3/4 (Schöler *et al.*, 1991) and retinoic acid receptor β2 (RARβ2) (Berkenstam *et al.*, 1992; Keaveney *et al.*, 1993) are known to be differentially regulated

during EC cellular differentiation, again suggesting the involvement of an embryonic stem-cell-specific cofactor(s) in the control of their transcriptional regulation. Another unique feature of embryonic stem cells is their inability to support the replication of SV40 and polyomaviruses upon infection (reviewed in Kelly and Condamine, 1982). This lack of virus replication has been demonstrated to be due to the presence of a factor that represses the activity of virus enhancers (Gorman *et al.*, 1985; Hen *et al.*, 1986). Although these miscellaneous activities represent unique aspects of gene regulation in EC cells, to date, the factors representing any one of these activities have remained elusive.

It has also been demonstrated that the expression levels of numerous transcription factors including the HOX family members are up-regulated during differentiation of EC cells (Colberg-Poley *et al.*, 1985; LaRosa and Gudas, 1988; Murphy *et al.*, 1988). However, only a small number of factors down-regulated upon differentiation have been identified. The Oct-3/4 (Okamoto *et al.*, 1990; Rosner *et al.*, 1990; Schöler *et al.*, 1990), REX-1 (Hosler *et al.*, 1989), SOX-2 (Yuan *et al.*, 1995) and PEA3 (Xin *et al.*, 1992) represent a few such examples. Therefore, the identification and subsequent cloning of transcriptional (co)factors expressed specifically in early embryonic tissue represent the first crucial step in unraveling the elusive transcriptional regulatory mechanisms in the early developmental stages in mammals.

In this report, we describe the molecular cloning of a novel transcription factor termed UTF1. This factor was isolated using a novel genetic screen in yeast. UTF1 augments adenovirus E2A promoter activity via a direct interaction with the activation domain of the promoter-bound ATF-2 protein. Furthermore, UTF1 can associate with the TFIID complex *in vivo*. Northern blot and *in situ* hybridization analyses have revealed that UTF1 expression is restricted mainly to pluripotent cells, and furthermore, that this expression is rapidly down-regulated upon differentiation. Therefore, UTF1 displays a number of properties expected for a pluripotent tissue transcriptional coactivator.

Results

UTF1 cloning

We are interested in understanding the unique transcriptional regulatory mechanisms that have been observed in pluripotent embryonic stem cells. As a means of isolating pluripotent cell-specific transcription factors, one approach we utilized was a genetic screen in yeast, which selects for proteins that augment transcription via the basal transcriptional machinery. A yeast strain, W303-1a, deleted for the *HIS3* gene was transformed with a CEN plasmid containing the *HIS3* coding sequence and the minimal promoter elements. Yeast cells transformed with this plasmid are unable to grow on media containing 3-aminotriazole (AT), a substance which acts as a competitive inhibitor of the *HIS3* gene product. Our screen selected for factors that elevate the level of transcription from the minimal *HIS3* promoter, thus enabling growth in the presence of AT. A cDNA library derived from F9 teratocarcinoma cells was cloned into the pKT10GAL vector carrying the inducible *GAL 1.10* promoter (Tanaka *et al.*, 1990). This library was introduced into yeast cells harboring the reporter plasmid and from a total of 10^5 yeast transformants only 47 colonies were able to grow in the presence of 10 mM AT. Of these 47 AT-resistant colonies, only three displayed a galactose-dependent AT-resistant phenotype and, as predicted, all three were unable to grow on AT in the presence of glucose. These results indicate that their His^+ phenotype is indeed dependent on the expression of the exogenous protein encoded by the introduced cDNAs. Upon plasmid rescue and subsequent sequence analysis it was revealed that all three independently recovered cDNAs were derived from the same gene. They differed only in the 5′ noncoding region and the length of the poly(A) stretch. We designated the gene undifferentiated embryonic cell transcription factor 1 (*UTF1*) because of its unique expression profile (see below).

To verify that the His^+ phenotype was indeed due to the expression of UTF1, we transformed the original yeast strain, W303-1a, carrying the reporter plasmid with the cDNA-encoding UTF1. As shown in Figure 1A, a large number of yeast colonies bearing the UTF1-expression plasmid grew in the presence of 10 mM AT. In contrast, when the same strain was transformed with the empty vector, only a few very slow-growing colonies appeared. To determine whether this phenotypic change was due to elevated levels of transcription from the UAS-deleted *HIS3* gene, we quantified the amount of *HIS3* mRNA by RNase protection assays. As shown in Figure 1B, *HIS3* mRNA was barely detectable in yeast cells carrying the reporter plasmid, irrespective of the carbon source used (lanes 1 and 2). However, the level of transcription was significantly elevated in yeast cells expressing UTF1. More importantly, the level of *HIS3* mRNA was further boosted when the cells were cultured using galactose instead of glucose as the carbon source (compare lanes 3 and 4). Thus, we conclude that UTF1 expression augments the level of transcription of the minimal UAS-deleted *HIS3* promoter.

The reporter plasmid used in these experiments contains the *cis*-element (GPEI) of the glutathione *S*-transferase P gene (Okuda *et al.*, 1989) which functions as a strong repressor in yeast cells (A.Okuda, unpublished data). This GPEI-element was cloned upstream of the *HIS3* gene to minimize readthrough transcription (see Materials and methods for details). To exclude the possibility that the observed increase in resistance to AT caused by UTF1 expression was due to alleviation of repression, the assays were repeated in the absence of the GPEI-repressor binding sites. As shown in lanes 5 and 6, UTF1 was again able to augment the level of transcription, suggesting that UTF1-dependent activation of transcription occurs through the minimal promoter of the *HIS3* gene. Unexpectedly, however, we subsequently found that UTF1 does not function as a basal transcriptional activator in mammalian cells, but instead, boosts the levels of transcription specifically in an upstream activator, ATF-2-dependent manner (see below). Therefore, it is unclear why UTF1 exhibits such activity specifically in yeast cells (see discussion).

UTF1 protein

The nucleotide sequence of UTF1 revealed a potential open reading frame of 339 amino acids (Figure 2). As no stop codon was found in the 5′ flanking region, we could

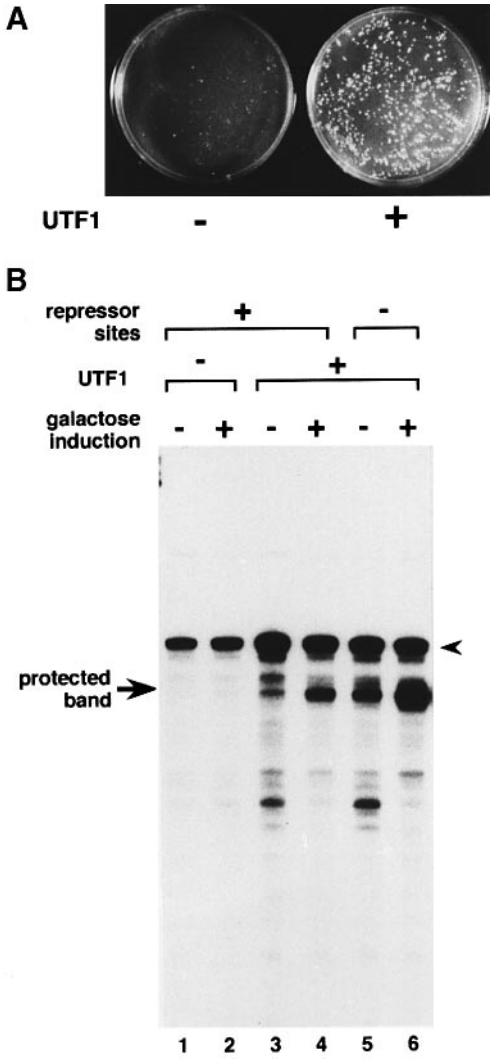


Fig. 1. Molecular cloning of UTF1. (**A**) UTF1 confers a His^+ phenotype on yeast. Plasmid DNA was recovered from His^+ yeast cells and used for the transformation of the original yeast strain carrying the UAS-deleted *HIS3* gene (right panel). Vector DNA without an insert was also transformed (left panel). These transformed yeast cells were plated on regular minimal media containing glucose, adenine, leucine and histidine, and, after 16 h incubation, plates were replica-plated on galactose–AT-containing plates. (**B**) Transcription of UAS-deleted *HIS3* is stimulated by UTF1. Radiolabeled anti-sense RNA was generated from SP64 carrying the *HIS3* gene fragment (–73 to +327), and *HIS3* mRNA levels were quantified by RNase protection assays. RNA was recovered from yeast cells carrying the reporter gene plasmid, with or without the GPEI repressor element, and the UTF1 expression vector or corresponding empty vector. Yeast cells were grown on glucose (–) or galactose (+) as a carbon source. The arrow indicates specific *HIS3* transcripts initiated from +13: transcription initiation is mainly specified by the TR TATA-box (Collart and Struhl, 1994). An arrowhead indicates the position of the undigested RNA probe.

```
   1 CCTCTACCTGGCTCAGGGATGCTGCTTCGTCCCCGGAGGTTGCCCGCTTTTTCTCCGCCG   60
   1                  M  L  L  R  P  R  R  L  P  A  F  S  P  P     14

  61 TCGCCAGCTAGCCCCGACGCTGAGCTGCGGTCGGCCGGGGATGTCCCGGTGACTACGTCT  120
  15  S  P  A  S  P  D  A  E  L  R  S  A  G  D  V  P  V  T  T  S     34

 121 GATGCCTTCGCTACTTCGGGCGGGATGGCTGAACCCGGCTCGCCCAAGGCTCCCGTGTCC  180
  35  D  A  F  A  T  S  G  G  M  A  E  P  G  S  P  K  A  P  V  S     54

 181 CCGGACTCGGCGCAGCGCACGCCCTGGAGTGCTCGAGAGACGGAGCTACTTCTGGGCACG  240
  55  P  D  S  A  Q  R  T  P  W  S  A  R  E  T  E  L  L  L  G  T     74

 241 CTGCTGCAGCCGGCCATGTGGCGCTCACTACTGCTGGACCGCAGGCAGACTCTGCCTACT  300
  75  L  L  Q  P  A  M  W  R  S  L  L  L  D  R  R  Q  T  L  P  T     94

 301 TACCGCCGCGTGTCCGCCGCACTGGCCCGTCAGCAAGTTCGGCGCACGCCCGCTCAGTGC  360
  95  Y  R  R  V  S  A  A  L  A  R  Q  Q  V  R  R  T  P  A  Q  C    114

 361 CGCCGTCGCTACAAGTTCCTCAAGGACAAACTCCGAGACTCCCAGGGCCAGCCGTCCGGA  420
 115  R  R  R  Y  K  F  L  K  D  K  L  R  D  S  Q  G  Q  P  S  G    134

 421 CCCTTCGATAACCAGATCCGCCAACTCATGGGGCTATTGGGGGACGATGGGCCGCCGAGG  480
 135  P  F  D  N  Q  I  R  Q  L  M  G  L  L  G  D  D  G  P  P  R    154

 481 GTTCGCCGCCGCTCTACTGGCCCTGGACGTCCCCAGCGCCGCGGCCGCTCGTCCCTCTCC  540
 155  V  R  R  R  S  T  G  P  G  R  P  Q  R  R  G  R  S  S  L  S    174

 541 GCGTTAGCACCCGCACCTGCACCGGTGGAGCAAGAGGCCGAGCTGCCGCTGGCTGCGGAA  600
 175  A  L  A  P  A  P  A  P  V  E  Q  E  A  E  L  P  L  A  A  E    194

 601 AATGACGAACCTGCCCCTGCGCTCAGATTCAGTTCTTCCACGACGAAGTCTGCAGGTGCC  660
 195  N  D  E  P  A  P  A  L  R  F  S  S  S  T  T  K  S  A  G  A    214

 661 CACCGCATTACCAGCTCTCCTCCTCTTACGAGCACCGACACTCTGCCTCCTGAGCCGGGC  720
 215  H  R  I  T  S  S  P  P  L  T  S  T  D  T  L  P  P  E  P  G    234

 721 CATACCTTCGAATCCTCCCCGACGCCAACCCCCGACCACGACGTGGAGACCCCGAATGAG  780
 235  H  T  F  E  S  S  P  T  P  T  P  D  H  D  V  E  T  P  N  E    254

 781 CCTCCTGGCCTTTCCCAGGGCCGTGCTTCTTCCCCGCAGGTTGCTCCCCAGTCGTTGAAT  840
 255  P  P  G  L  S  Q  G  R  A  S  S  P  Q  V  A  P  Q  S  L  N    274

 841 ACCGCGTTGCTGCAGACCTTGACGCACTTGGGCGACATCTCAACAGTTCTGGGCCCTCTG  900
 275  T  A  L  L  Q  T  L  T  H  L  G  D  I  S  T  V  L  G  P  P    294

 901 CGCGACCAGCTGTCGACCCTGAACCAGCACGTGGAGCATCTACGAGGTTCCTTCGACCAA  960
 295  R  D  Q  L  S  T  P  N  Q  H  V  E  H  P  R  G  S  F  D  Q    314

 961 ACCGTTTCTCTGGCGGTGGGGTTCATTCTGGGCAGTGCAGCCTCCGAGCGGGGCATCCTT 1020
 315  T  V  S  L  A  V  G  F  I  L  G  S  A  A  S  E  R  G  I  L    334

1021 GGGGACTTGCGCCAATAAAGCTTTGTTATCAGTCCTCTTCTCCAGAAACGACCCCCAGTC 1080
 335  G  D  L  R  Q  *                                              339

1081 TACTCCGACCCCTGTTACTGTCCCCATCCCCATCTCAAACCTAGGTTCTTACCAAAATAA 1140
1141 AGATTGTGTTTTCTACTTAA                                          1160
```

Fig. 2. Nucleotide and deduced amino acid sequences of UTF1. Translation of the open reading frame yields a protein of 339 amino acids with a molecular mass of 36 406 Da. The first and fourth residues within the seven amino acid repeat of the coiled-coil motif located at the C-terminal domain of the protein are indicated by bold letters underlined and black boxed, respectively.

not conclude definitely that the first ATG codon indeed represents the initiating codon. However, two lines of evidence predict that this is indeed the case. First, sequence comparison with human UTF1, which has been recently cloned, showed that this putative ATG and its downstream sequence were conserved between mouse and man, whereas no obvious similarity was evident in the upstream sequence (A.Fukushima and M.Muramatsu, unpublished data). Secondly, RNase protection experiments using genomic *UTF1* gene fragments revealed that the transcripts were generated from at least seven transcription start sites sustained by three GC boxes located at a distance of ~100 nucleotides upstream of the putative ATG initiating codon. Moreover, there are no stop codons or additional ATG codons in-frame in the transcribed region (M.Nishimoto and M.Muramatsu, unpublished data).

The UTF1 protein contains a typical leucine zipper motif located at its C terminus, which contains leucine or valine at the first position (Figure 2, bold letter with underline) and isoleucine or leucine at the fourth position (Figure 2, black box) within the seven-residue repeat (Figure 2). Comparison of the nucleotide and amino acid sequences of UTF1 with the entries in the database revealed no significant sequence homologies. Furthermore, no striking sequence motifs have been observed other than the leucine zipper. The amino acid composition of UTF1 reveals two striking features; first, UTF1 is rich in proline residues (41 out of 339 amino acids), which are scattered throughout the protein. Secondly, UTF1 is very rich in basic amino acids, particularly arginine (34 residues in total), which results in an unusually high isoelectric point (calculated isoelectric point of 10.46). As the leucine zipper motif is known to mediate subunit oligomerization (for review see Cohen and Parry, 1994), we have examined whether UTF1 has such an activity and found that UTF1 indeed interacts avidly with itself (data not shown).

Tissue distribution of UTF1 mRNA

To determine the expression pattern of UTF1, we performed RNase protection analyses of RNA from various cell lines and tissues. UTF1 mRNA was readily detectable in undifferentiated EC cells (F9 and P19) and embryonic stem (ES) cells, but no expression was detected in most of the adult mouse tissues examined except for germ line tissues. Indeed, low but appreciable levels of expression were detected in ovary (O) and testis (T) (Figure 3A). However, as whole organs of these tissues were used as RNA sources, it cannot be excluded that a small number of cells, for example germ cells, express UTF1 at relatively high levels. Figure 3B shows that the UTF1 mRNA rapidly disappears when P19 cells are induced to differentiate by the addition of retinoic acid, indicating that UTF1 expression is specific for the pluripotent state of EC cells.

UTF1 expression in early mouse embryos

We also examined the expression pattern of the UTF1 protein during early mouse development by whole mount *in situ* hybridization techniques. As shown in Figure 4A and B, expression of the UTF1 mRNA is not detected in the morula [corresponding to ~3.0 days post-coitum (d.p.c.) embryo] (Figure 4A), but appears at the blastocyst stage (corresponding to ~3.5 d.p.c. embryo) (Figure 4B), indicating that there is a rapid induction of expression of the *UTF1* gene between these two stages. In the blastocyst, UTF1 is specifically expressed in the inner cell mass (ICM), whereas no expression is detected in the trophectoderm. At later stages (6.5–7.5 d.p.c.), UTF1 mRNA is present in the primitive ectoderm which is derived from the ICM (Figure 4C, D and G). However, no expression is detected in the mesoderm. These results indicate that the differentiation of primitive ectoderm into mesoderm starting at the early primitive streak stage is accompanied by the extinction of expression of UTF1. We also note that part of the extra-embryonic ectodermal cells are stained faintly at the egg cylinder stage (data not shown), but the expression becomes very prominent at the early primitive streak stage (Figure 4C, D and G). Figure 4E and F shows longitudinal and transverse sections of embryos at an early primitive streak stage (~7.0 d.p.c.), respectively. From these sections, UTF1 expression in the primitive ectoderm and the extra-embryonic ectoderm is unequivocally identified. Figure 4G and H show ~7.5 d.p.c. embryos stained for UTF1 (Figure 4G) and Brachyury (Figure 4H), one of the markers of early mesoderm, while Figure 4I shows a similar stage embryo double-stained for these transcripts. From Figure 4G–I, it is obvious that UTF1 is indeed down-regulated from the posterior to the anterior part of the embryo during the primitive streak stage. At the early somite stage (~8.0 d.p.c.), expression of the UTF1 mRNA is significantly down-regulated in the ectoderm, although a low level of expression is still observed in the neural fold (Figure 4J and K). Rather strong expression of the UTF1 mRNA is detected in the chorion, and we suppose that UTF1-expressing cells in this tissue are derived from extra-embryonic ectodermal cells identified at the early primitive streak stage. While expression in the neural fold is extinguished by 9.5 d.p.c., UTF1 expression in the chorion continues for up to 13.5 d.p.c., but becomes undetectable by 14 d.p.c. (data not shown).

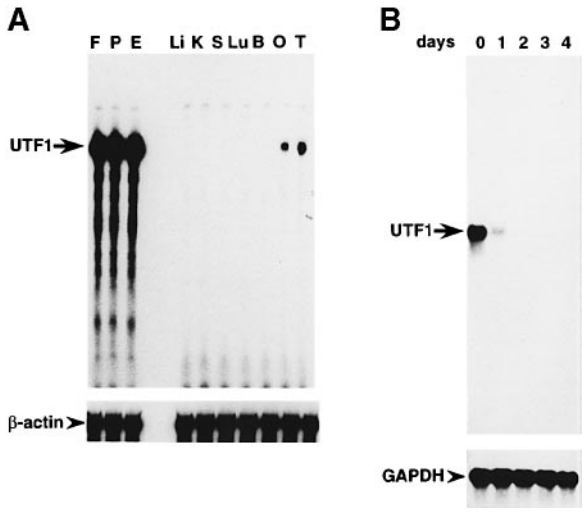
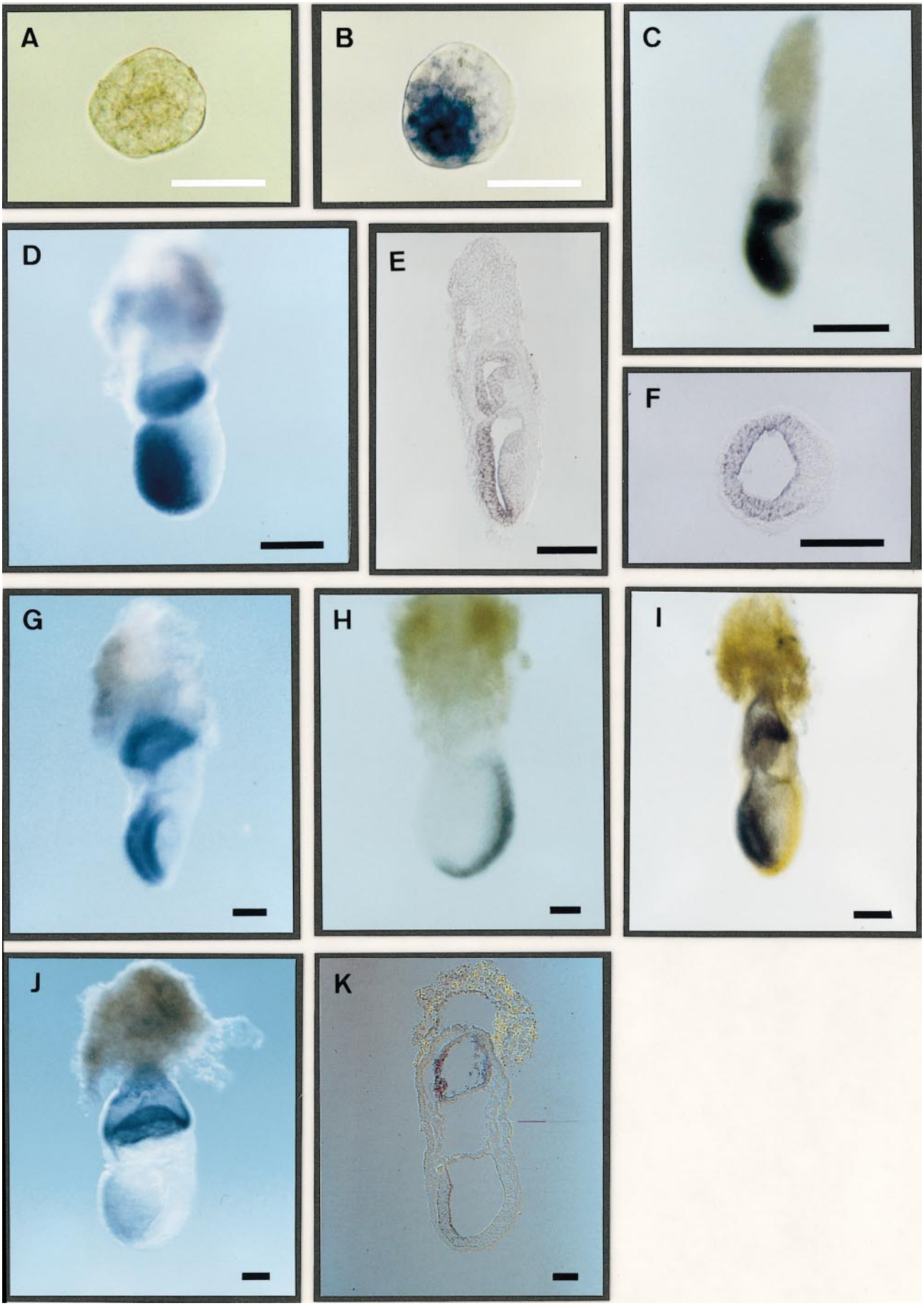


Fig. 3. The UTF1 is expressed in undifferentiated embryonic stem cells and germ line tissues. (**A**) The UTF1 mRNA levels in various tissues and cells. RNA isolated from ~8-week-old mouse liver (Li), kidney (K), spleen (S), lung (Lu), brain (B), ovary (O), testis (T) and undifferentiated F9 (F) and P19 (P) EC cells, and ES cells, was analyzed by RNase protection analysis using radiolabeled anti-sense UTF1 or β-actin RNA probe. Four micrograms of poly(A)$^+$ RNA were used for the detection of UTF1 mRNA in adult mouse tissues including ovary and testis, while 0.4 μg of RNA was used for the UTF1 expression in F9, P19 and ES cells. For the detection of β-actin mRNA, 0.2 μg of poly(A)$^+$ RNA was used from all sources. (**B**) Regulation of the *UTF1* gene expression during differentiation of P19 cells. P19 cells were induced to differentiate with 0.5 μM of retinoic acid. RNA was prepared at the indicated days after induction. After fractionation by agarose gel electrophoresis, RNA [0.4 μg of poly(A)$^+$ RNA in each lane] was transferred to a filter and hybridized with glyceraldehyde 3-phosphate dehydrogenase (GAPDH) cDNA as a probe to assess the integrity of the RNA samples. The filter was subsequently stripped of probe and rehybridized with radiolabeled UTF1 cDNA. The size of UTF1 mRNA was estimated to be ~1.2 kb using standard RNA size markers (not shown).

UTF1 boosts the level of transcription in an ATF-2-dependent manner

The pluripotent embryonic stem-cell-specific expression of the UTF1 protein described above prompted us to investigate whether UTF1 has any of the activities attributed to the elusive E1A-LA. We carried out the classical E2A transfection assay in Chinese hamster ovary (CHO) cells (Helin *et al.*, 1992) using the two different reporter plasmids shown in Figure 5A. Using the CAT reporter plasmid bearing the ATF and E2F sites, a significant increase in transcriptional activation (up to 5-fold) was observed upon transfection of UTF1. However, the protein

Fig. 4. Localization of the UTF1 transcript in mouse embryo. *In situ* hybridization was performed using UTF1 cDNA as a probe (see Materials and methods for details). (**A**) Morula. No expression was detected in this stage. (**B**) Blastocyst. UTF1 mRNA is detected in the ICM, but not in the trophectoderm. (**C**) 6.5 d.p.c. embryo. (**D**) 7.0 d.p.c. embryo. (**E**) Longitudinal and (**F**) transverse sections (5 μm) of 7.0 d.p.c. embryos. At this early primitive streak stage, UTF1 expression is detected in the primitive ectoderm and the extra-embryonic ectoderm, whereas no expression is detected in the mesoderm. 7.5 d.p.c. embryos stained for UTF1 (**F**) or Brachyury (**H**), or double-stained for these transcripts (**I**). Whole (**J**), or thin section (5 μm) (**K**) of 8.0 d.p.c. embryo. At this stage, expression of UTF1 in the ectoderm is significantly down-regulated. Conversely, strong expression is detected in the chorion. White and black bars correspond to 0.05 and 0.25 mm, respectively.


A
B
C
D
E
F
G
H
I
J
K

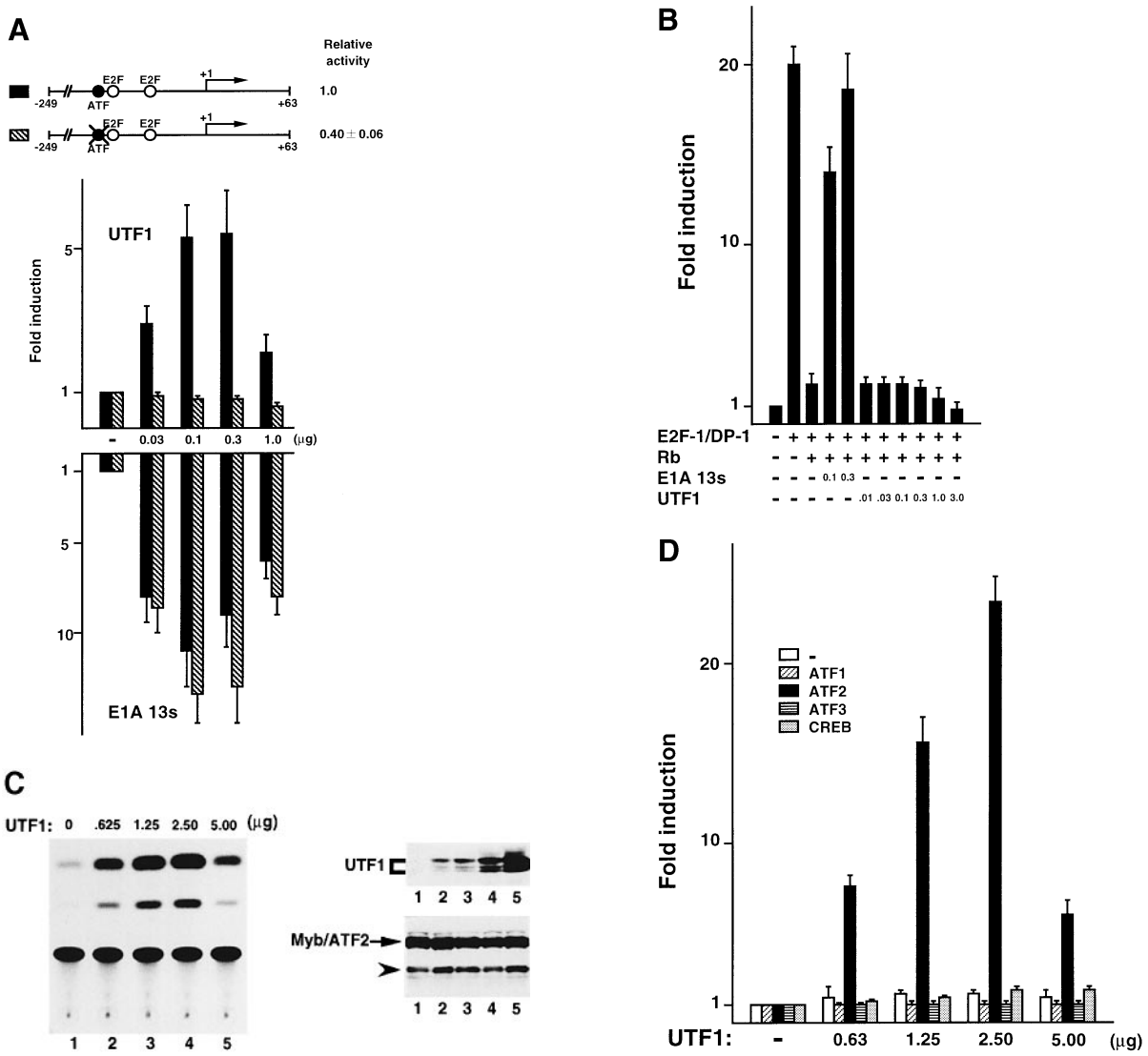


Fig. 5. UTF1-mediated activation of ATF-2. (**A**) UTF1 elevates the level of transcription from the adenovirus E2A promoter. Two different E2A promoter-CAT reporter plasmids were constructed as depicted. ATF-2 binding site (5′-TGACGTA-3′) was mutagenized to the sequence (5′-GGCCGTA-3) in the reporter gene denoted by the hatched box. Chinese hamster ovary (CHO) cells were cotransfected with a constant amount of either one of the E2A–CAT reporter plasmids (0.5 μg) and increasing amounts of E1A or UTF1 expression vector as indicated. RSV–β-gal (0.5 μg), which is insensitive to the expression of UTF1 or E1A was also transfected as an internal control. The total amount of transfected plasmid was adjusted to 20 μg with pUC18. CAT and β-galactosidase assays were done as previously described (Okuda *et al.*, 1990). The intrinsic level of transcription from each CAT reporter in the absence of transcriptional inducers was arbitrarily set as one and the fold induction due to the expression of UTF1 or E1A 13S was calculated. The data shown represent the average of five independent experiments. (**B**) UTF1 fails to counteract the negative effect of RB on the E2F-1–DP-1 complex. The SAOS-2 cell line which is deficient for functional RB (Shew *et al.*, 1990) was cotransfected as indicated using a constant amount (0.5 μg) of the CAT reporter plasmid, which contains the *E2A* gene promoter (–69 to +63) bearing two E2F sites, but no ATF-2 site. The following amounts of expression vectors or corresponding empty vector were transfected: pCMV–E2F-1, 0.01 μg; pCMV–DP-1, 0.01 μg; pCMV–RB, 0.15 μg; internal control RSV–β-gal, 0.5 μg. Constructions of these expression vectors were described elsewhere (Helin *et al.*, 1993). (**C**) Potentiation of transcriptional stimulating activity of Myb–ATF-2 fusion protein by UTF1 *in vivo*. CHO cells were cotransfected with a constant amount (4 μg) of collagen CAT–Myb, the reporter gene bearing 6×Myb binding site (diagram shown in Figure 6A), an Myb–ATF-2 fusion protein expression vector (10 μg) and increasing amounts of a UTF1 expression vector, in addition to the internal control RSV–β-gal. The total amount of transfected plasmid was adjusted to 20 μg with pUC18. CAT and β-galactosidase assays were carried out as in (A). The remainder of these extracts were used to detect the expression levels of UTF1 and ATF-2 by Western blot analyses. The arrowhead points to the endogenous ATF-2. (**D**) UTF1 elevates the levels of transcription specifically in an ATF-2-dependent manner. CHO cells were transfected with a constant amount of either one of the Myb fusion protein expression vectors indicated (10 μg) and increasing amounts of UTF1 expression vector in addition to the CAT reporter gene (4 μg) and an internal control RSV–β-gal (0.5 μg). Transfection and CAT and β-galactosidase assays were carried out as in (A). Intrinsic activities of the Myb fusion proteins relative to the Myb DNA-binding domain alone are as follows. Myb–ATF-1: 0.76; Myb–ATF-2: 1.13; Myb–ATF-3: 0.65; Myb–CREB: 1.05. All these activities were arbitrarily set to one and fold induction due to the expression of UTF1 was calculated. Data shown are the summary of at least five independent experiments with comparable results.

failed to augment the level of transcription from the CAT reporter plasmid when the ATF site was mutagenized. Thus, UTF1 stimulates transcription through the ATF site of the E2A promoter, but fails to do so through the E2F sites. In contrast, when the experiments were performed with E1A 13S, both promoter activities were up-regulated. We have also carried out the experiments in the osteosarcoma cell line, SAOS-2 which is deficient for the functional RB gene product (Shew *et al.*, 1990), and confirmed that E1A, but not UTF1, potentiated the activity of the

E2F–DP-1 complex by counteracting the negative effect of RB (Figure 5B). From these results, we conclude that UTF1 *per se* does not represent a major component of the E1A-LA.

It is known that E1A cooperates with ATF-2 to stimulate transcription (Lee *et al.*, 1987; Maekawa *et al.*, 1991; Liu and Green, 1994; Li and Green, 1996). To determine whether ATF-2 is also involved in the UTF1-mediated activation, we carried out further transfection experiments in CHO cells using fusion proteins consisting of the DNA-binding domain of c-Myb fused to ATF-2 as described by Maekawa *et al.* (1991) (see Figure 6A for a diagram of the reporter plasmids). This synthetic reporter system was employed to eliminate the contribution of endogenously expressed ATF-2 to the activity. As shown in Figure 5C, UTF1 activated the expression of the CAT reporter gene bearing Myb sites up to 23-fold in a dose-dependent manner (see Figure 5D for actual values of fold induction). However, large amounts of UTF1 expression (5 μg of expression vector transfected) resulted in an attenuation of the induction, probably due to the squelching effects of the protein. Increasing amounts of UTF1 expression vector, and a constant amount of MybDBD–ATF-2 fusion protein was used throughout. The detection of two distinct UTF1 protein bands appears to reflect different phosphorylation states rather than proteolysis of the protein, as extracts treated with λ-phosphatase gave a single protein band on SDS–PAGE with a mobility identical to that of the faster migrating polypeptide (A.Fukushima and M.Muramatsu, unpublished data). Similar to E1A, the activation by UTF1 is ATF-2 protein-dependent, because MybDBD alone, as well as other fusion proteins, failed to cooperate with UTF1 (Figure 5D and data not shown). Gel retardation analyses using a probe containing the Myb consensus site revealed comparable binding activities of fusion proteins, thus eliminating the possibility that differences in the response to UTF1 were due to the variations in their expression levels (data not shown).

Minimal activation domain is sufficient for UTF1-mediated activation of ATF-2

To map the region of ATF-2 required for activation by UTF1, we made truncated forms of ATF-2 fused to the MybDBD (Figure 6A) and examined whether the transcriptional activities of ATF-2 and its derivatives were modulated by the expression of either UTF1 or E1A (Figure 6B). Recently, Li and Green (1996) have demonstrated that the intrinsic activity of the activation domain of ATF-2 is suppressed by the intramolecular interaction between the activation and the bZIP domains of ATF-2 and, therefore, deletion of the latter domain renders the ATF-2 constitutively active even in the absence of transcriptional inducers. Consistent with this observation, ATF-2 derivatives lacking the bZIP domain showed very potent (1–109), or moderate (1–243, 1–350) transcriptional stimulating activities by themselves when Myb–TK CAT was used as a reporter plasmid (Figure 6A). Furthermore, we found that these highly active Myb–ATF-2 derivatives lacking the bZIP domain are not further activated by coexpression of either UTF1 or E1A (data not shown). However, when the Myb binding sites were positioned distal to the start site of transcription as in the collagen CAT–Myb reporter, all of these fusion proteins alone did not exhibit significant transcriptional stimulating activities. Only the minimal ATF-2 activation domain (1–109) showed ~3-fold higher activities than the background level.

The above results suggest that the activation domain of ATF-2 is able to directly contact components of the basal transcriptional machinery and consequently activates transcription when the fusion proteins are recruited in close proximity to the promoter, but fails to do so when the fusion protein binding sites are located at a distal position. However, it remains to be determined whether this promoter effect is dependent on the expression levels of the Myb–ATF-2 fusion proteins. We have, however, examined how E1A 13S and UTF1 influence the transcriptional stimulating activities of these fusion proteins using the collagen CAT–Myb reporter plasmid. As shown in the bottom half of Figure 6B, E1A is able to boost the level of transcription, only when the full length ATF-2 protein is used, and all of the other truncated ATF derivatives are insensitive to the expression levels of the E1A protein. These results are not unexpected, as all of the ATF-2 derivatives lack the bZIP domain, which is essential for the interaction with E1A (Liu and Green, 1994). Interestingly, when similar experiments were carried out with the UTF1 expression vector, the protein was found to require a domain completely distinct from that required by E1A for mediating the transcriptional induction. All of the ATF-2 derivatives, as well as the full-length protein, are transcriptionally highly active in the presence of UTF1, indicating that the minimal activation domain of ATF-2 is sufficient for mediating UTF1-inducible transcriptional stimulation (Figure 6B).

Interaction of UTF1 with ATF-2

The cotransfection assays in the previous section have demonstrated that the minimal activation domain of ATF-2 is sufficient for the UTF1-dependent transcriptional response *in vivo*. Hence, we performed GST pull-down assays to examine whether this domain is the interaction site for UTF1. ^{35}S-labeled UTF1 was incubated with either GST alone, GST–ATF-2 activation domain (1–109), or GST–ATF-2 bZIP domain (347–505) immobilized on glutathione–Sepharose beads. After extensive washing, the amount of the UTF1 protein bound to the beads was quantified. As shown in Figure 7A (lanes 1–4), GST–ATF-2 (1–109) bound UTF1 efficiently. By contrast, UTF1 did not bind to the control GST or GST–ATF-2 (347–505), verifying the specificity of the interaction between the activation domain of ATF-2 and UTF1. We also characterized the interaction between ATF-2 and the adenovirus E1A 13S. Consistent with the previous report (Liu and Green, 1994), E1A did not bind to the activation domain, but bound strongly to the bZIP domain of ATF-2 (Figure 7A, lanes 5–8). Thus, UTF1 and E1A bind to different domains of ATF-2 and this difference in binding specificity explains why these two transcriptional cofactors require distinct domains of ATF-2 for the induction shown in Figure 6B. We have also carried out these experiments using recombinant UTF1 to eliminate the possibility that certain factor(s) in reticulocyte lysate may be involved in the interaction. UTF1 protein was overexpressed in Sf9 cells infected with recombinant baculovirus, immunopurified from whole cell extracts (see Materials and methods) and used in the interaction assays. As shown in Figure

7B (lanes 2 and 3), UTF1 did not bind to GST beads, but bound weakly to the GST–ATF-2 (1–505) beads. This weak interaction is probably due to the aforementioned intramolecular interaction between the activation and the bZIP domains of ATF-2. Consistent with this idea, the ATF-2 mutant (1–109) lacking the bZIP portion bound to UTF1 more avidly (lane 4). To strengthen the above observation we also carried out yeast two-hybrid interaction assays. The yeast strain used for the two-hybrid assay is Y153 which contains an integrated GAL4-dependent lacZ reporter. GAL4 DNA-binding domain (DBD) and activation domain (AD) fusions were constructed and

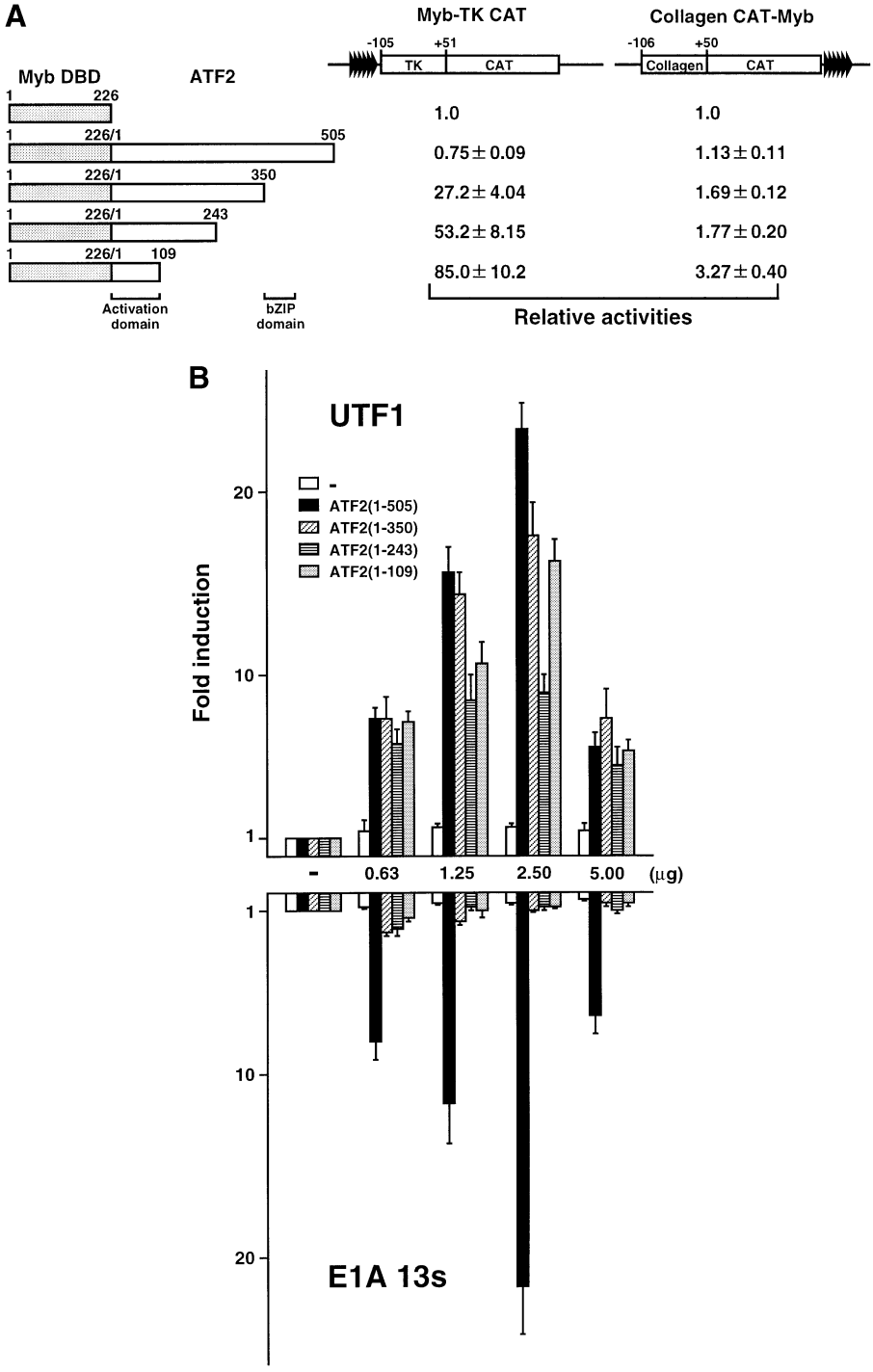


Fig. 6. UTF1 and E1A require distinct domains of ATF-2 for their transcriptional induction. (**A**) Schematic representation of ATF-2 and its C-terminally truncated derivatives fused to the Myb DNA-binding domain. Intrinsic transcriptional stimulating activities of these fusion proteins were obtained using two different reporter plasmids bearing six copies of the fusion protein binding sites as depicted. Transfection and CAT assays were carried out as in Figure 5, with 0.8 μg of Myb–TK CAT or 4 μg of collagen CAT–Myb and 10 μg of one of fusion expression vectors, in addition to an internal control RSV–β-gal (0.5 μg). The data represents the average from five independent experiments. (**B**) Delineation of ATF-2 domains required for UTF1- and E1A-mediated activations. Transfections were carried out as in Figure 5. The intrinsic activities of the Myb–ATF-2 fusion protein and its truncated derivatives were obtained as in (A) using the collagen CAT–Myb reporter plasmid. These values were arbitrarily set to one and the fold increase in transcription due to the expression of either UTF1 or E1A is presented.

assayed in different combinations for activation using the β-galactosidase filter lift assay (Yao *et al.*, 1996). As shown in Figure 7C, GAL4-DBD alone or GAL4-DBD–ATF2 (1–109) did not increase the level of β-galactosidase activity, indicating that the activation domain of ATF-2 does not function by itself in yeast cells. Likewise, a GAL4-AD–UTF1 fusion protein alone also failed to activate transcription, however, strong activation was observed when this plasmid was coexpressed with a GAL4-DBD–ATF-2 fusion vector. Thus, the yeast two-hybrid assay lends further evidence to a direct physical interaction between the activation domain of ATF-2 and UTF1.

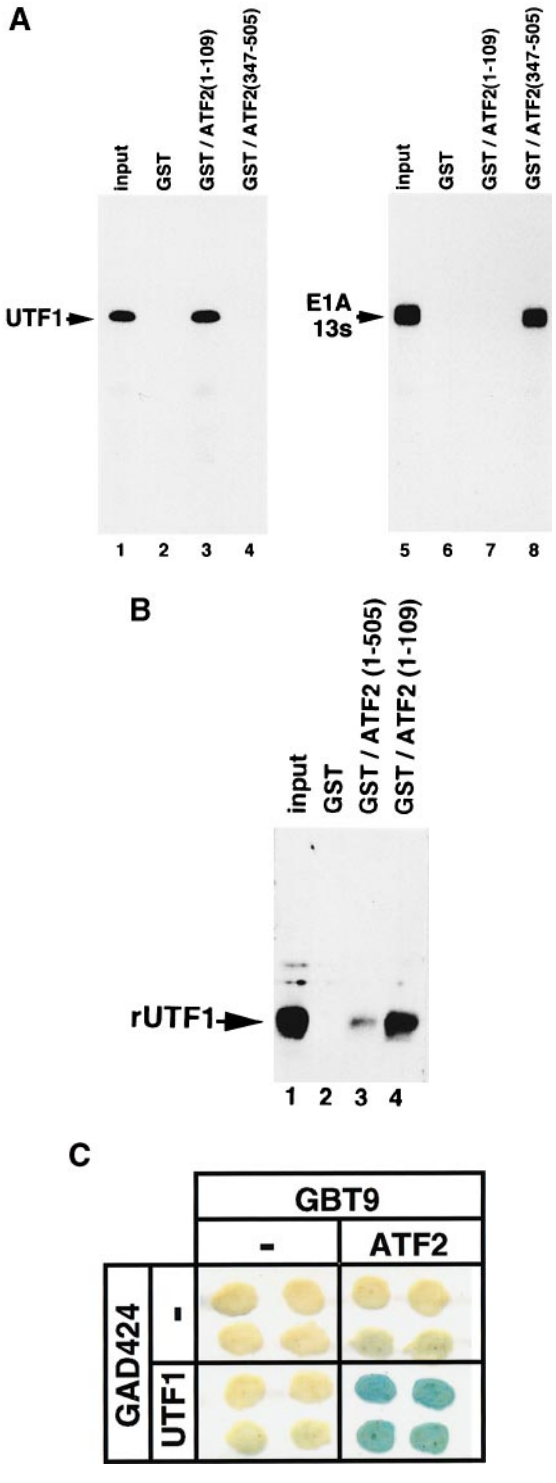


Fig. 7. Interaction of UTF1 with ATF2. (**A**) UTF1 and E1A bind to distinct domains of ATF-2. UTF1 and E1A were produced *in vitro* using rabbit reticulocyte lysate in the presence of ^{35}S-labeled methionine. GST alone, GST–ATF-2 (1–109), or GST–ATF-2 (347–505) was mixed with glutathione–Sepharose beads (30 µl) and incubated at 4°C for 30 min. These protein-bound beads were then incubated with radiolabeled *in vitro* translated UTF1 (20 µl of lysate) and washed extensively with the following buffer (20 mM Tris–HCl pH 7.9, 10% glycerol, 100 mM KCl, 1 mM $MgCl_2$, 0.2 mM EDTA, 10 µM $ZnCl_2$, 0.5 mM DTT). Subsequently, proteins bound to the beads were eluted and analyzed by SDS–PAGE followed by autoradiography. Ten percent of the input sample was applied to the first lane of the gel. (**B**) Recombinant UTF1 protein (rUTF1) is able to bind to ATF-2 directly. The interaction assay was carried out as above using UTF1 protein (0.3 µg) immuno-purified from baculovirus infected Sf9 cells and the GST fusion proteins as indicated. Fifty micrograms of BSA were added to each reaction mixture. The bound protein was visualized by Western blot analysis using an anti-UTF1 antibody. Ten percent of the input sample was applied to the first lane of the gel. (**C**) Yeast two-hybrid analysis with ATF-2 activation domain and UTF1. The activation domains of ATF-2 and the full length UTF1 were subcloned into pGBT9 (GAL4 DNA-binding domain fusion construct) and pGAD424 (GAL4 activation domain fusion construct) (Chien *et al.*, 1991), in the correct orientation and reading frame. The yeast strain, Y153 (*Matα leu2-3, 112, ura3-52, trp1-901, his3-Δ200, ade2-101, gal4Δgal80ΔURA3::GAL-LacZ, LYS2::GAL-HIS3*) was transformed with pGBT9–ATF-2(1–109) or pGBT9 vector alone. The resultant yeast strains were subsequently transformed with either pGAD424–UTF1 or pGAD424 vector alone. Transformants were restreaked and the amounts of reconstituted GAL4 protein in yeast were quantified by filter lifting assay, as previously described (Yao *et al.*, 1996).

Association of UTF1 with ATF-2 and TFIID in EC cells

Although it is clear that UTF1 stimulates ATF-2-dependent transcription, it has not been established whether UTF1 and ATF-2 are indeed coexpressed in embryonic stem cells. To this end, whole cell extracts were prepared from F9, P19 and ES cells and subjected to Western blot analyses using anti-UTF1 or ATF-2 antibodies. These analyses revealed that all three different embryonic cells express significant amounts of both proteins (Figure 8A). A number of UTF1 protein bands were observed on Western analyses; however, treatment with phosphatase again resulted in the conversion of these bands into the faster migrating species, indicating that they were due to phosphorylation of the UTF1 protein in EC cells (A.Fukushima and M.Muramatsu, unpublished data). Unlike UTF1 which is down-regulated upon pluripotent cell differentiation, it appears that the level of ATF-2 expression remains constant (data not shown). As both UTF1 and ATF-2 proteins are coexpressed in pluripotent cells, it is reasonable to assume that under physiological conditions they also interact directly. To assess this we performed co-immunoprecipitation analyses using nuclear extracts from undifferentiated F9 cells. Although the anti-UTF1 antibody failed to co-immunoprecipitate ATF-2, in the reciprocal experiment, anti-ATF-2 antibody clearly co-immunoprecipitated the UTF1 protein (Figure 8B, lane 2). The former negative result is probably because anti-UTF1 antibody disrupted the interaction, as this antibody was generated using the domain including the ATF-2-interacting leucine zipper portion of UTF1 (see Materials and methods). Using two unrelated antibodies, anti-E2F-

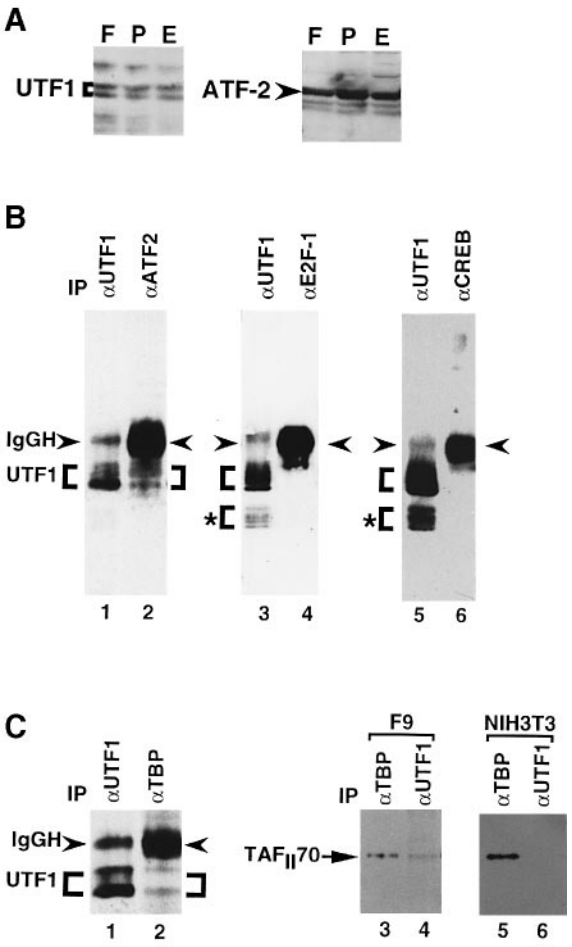


Fig. 8. Interaction of UTF1 with ATF-2 and TFIID complex *in vivo*. (**A**) UTF1 and ATF-2 were coexpressed in EC and ES cells. Whole cell extracts were prepared from ~10^7 cells of the indicated cell lines and were subject to Western blot analyses using anti-UTF1 (left panel) or anti-ATF-2 antibody (right panel). (**B**) UTF1 interacts with ATF-2 in EC cells. Undifferentiated F9 cell nuclear extracts (200 µl) were immunoprecipitated with anti-UTF1, ATF-2, E2F-1 or CREB antibodies which were coupled to CNBr-activated Sepharose beads. Immunoprecipitates were separated by SDS–PAGE, transferred to PVDF membrane and probed with the anti-UTF1 antibody. Ten and 100% of immunoprecipitated samples were applied to odd and even numbered lanes, respectively. The asterisk indicates partially degraded UTF1 protein. (**C**) UTF1 associates with TFIID complex *in vivo*. F9 (left and middle panels) or NIH 3T3 cell (right panel) nuclear extracts (200 µl) were immunoprecipitated with either polyclonal anti-TBP or anti-UTF1 antibody. Immunoprecipitates were fractionated as above. The filters were probed with a polyclonal anti-UTF1 antibody (left panel) or a monoclonal anti-$TAF_{II}70$ antibody obtained from Transduction Laboratories (middle and right panels). Ten percent of the samples were applied to lanes 1, 3 and 5, while the complete samples were applied to lanes 2, 4 and 6.

1 and anti-CREB, we failed to co-immunoprecipitate UTF1 (Figure 8B, lanes 4 and 6), verifying the specificity of the UTF1–ATF-2 interaction. In this regard, it should be noted that under identical conditions, the anti-E2F-1 antibody efficiently co-immunoprecipitated DP-1 (data not shown). From these results, we conclude that UTF1 and ATF-2 are coexpressed in F9 embryonic cells and that, furthermore, these two proteins are physiologically associated in these cells.

Finally, we examined the possible interaction of UTF1 with the TFIID complex in EC cells. Again using co-immunoprecipitation assays with extracts derived from F9 cells, we show that a polyclonal anti-TBP antibody efficiently co-immunoprecipitates UTF1 (Figure 8C, lane 2). We also show that both anti-TBP and anti-UTF1 antibodies immunoprecipitate $TAF_{II}70$ (Figure 8C, lanes 3 and 4), one of the subunits of the TFIID complex. It should be noted, however, that the anti-UTF1 antibody was much less efficient in this procedure, immunoprecipitating only 5% of $TAF_{II}70$ compared with the anti-TBP antibody. This result will be discussed in more detail later (see Discussion). Nevertheless, the observed band appears to reflect a specific interaction between UTF1 and the TFIID complex, because, as shown in lane 6 of Figure 8C, the anti-UTF1 antibody did not co-immunoprecipitate $TAF_{II}70$ in nuclear extracts derived from NIH 3T3 cells, a cell line that does not express UTF1. Thus, we conclude that UTF1 interacts with the TFIID complex in EC cells.

Discussion

Understanding the molecular mechanism underlying cellular differentiation is paramount to the advancement of our knowledge of developmental biology. In this regard, elucidation of the molecular mechanism that governs early mouse developmental stages, such as the blastocyst and egg cylinder stages, is an issue of prime importance. To this end, it is important to isolate novel transcription factors which are specifically expressed in these early stages of development, and the UTF1 protein, reported here, represents one such factor. In tissue culture, differentiation of EC cells induced by retinoic acid results in a rapid extinction of UTF1 mRNA. We have recently found that this expression profile is also conserved in man; the human teratocarcinoma cell line N-tera 2 expresses high levels of the UTF1 mRNA, whereas no expression is detected in HeLa, HL60 or Jurkat cells (A.Fukushima and M.Muramatsu, unpublished data). Analyses of *in situ* staining of mouse embryos corroborate and extend the results obtained with ES and EC cells. UTF1 mRNA is readily detectable in the ICM and primitive ectoderm, but not in the trophectoderm. During the primitive streak stage, UTF1 expression is inversely correlated with the differentiated state of the embryonic cells. These analyses show that the differentiation process in embryonic cells is accompanied by the down-regulation of UTF1 protein expression. This UTF1 expression profile is quite similar to that of Oct-3/4 (Rosner *et al.*, 1990), although there are certain distinct differences, such as the lack of Oct-3/4 expression in the chorion. Therefore, it initially appeared plausible that UTF1 may function as a coactivator of Oct-3/4, but further experimental analyses revealed that this was not the case (A.Fukushima and M.Muramatsu, unpublished data). Interestingly, however, results from more recent experiments suggest that UTF1 may act downstream of Oct-3/4 in the regulatory cascade, which may explain, at least in part, the similar expression profiles of these two genes.

It has long been known that EC cells harbor a factor termed the E1A-LA which has similar properties to those of the adenovirus E1A protein. This E1A-LA had originally been postulated to explain the specific activation of the adenovirus early *E2A* gene in the absence of viral E1A in EC cells (Imperiale *et al.*, 1984). This increase in transcriptional activation occurs via the E2F–DP-1

complex(es) which interact at the E2F sites of the E2A promoter (Bandara and La Thangue, 1991; Murray *et al.*, 1991). In contrast to the E1A-LA, UTF1 does not exert its stimulatory effect through this complex but rather functions through the upstream ATF-2 site. Therefore, it appears that UTF1 does not represent a major component of the E1A-LA. We further examined whether UTF1 displayed any of the characteristics of the other identified EC cell-specific activities such as cooperating activities with Oct3/4 and RARβ2. However, UTF1 protein function did not mimic any of these activities (data not shown). Furthermore, we found that UTF1 fails to activate any minimal (TATA box only) gene promoters such as that of the *hsp70* gene. These results were somewhat unexpected because, as shown in Figure 1, UTF1 was isolated based on its ability to boost the levels of transcription from a minimal promoter in yeast cells. In this regard, it is interesting that we have recently shown that a UTF1 mutant lacking the leucine zipper portion augments the level of transcription preferentially from E1B and *hsp70* minimal promoters such as E1A and E1A-LA (M.Keaveney, K.Boon, H.G.Stunnenberg, A.Fukushima and M.Muramatsu, unpublished data). Therefore, it is intriguing to speculate that UTF1 also displays these activities in the context of full-length protein, together with a certain factor(s) which attenuate the self-association properties of UTF1.

Cotransfection studies in mammalian cells revealed that the E1A and UTF1 proteins require different domains of ATF-2 for their transcriptional induction, i.e. UTF1 binds to the activation domain of ATF-2, whereas E1A binds to the bZIP domain of this same factor. It is known that the metal-binding motif of E1A is involved in the interaction with the bZIP domain of ATF-2 (Liu and Green, 1994). In addition to this particular case, there are a number of examples in which these two motifs are involved in protein–protein interaction (Diamond *et al.*, 1990; Yang-Yen *et al.*, 1990; Wagner and Green, 1993; Cheng *et al.*, 1997). Furthermore, ATF-2 itself has both metal-binding and leucine zipper-containing domains and intramolecular interactions between these two domains have previously been demonstrated (Li and Green, 1996). Coincidentally, the metal-binding motif is located in the N-terminal portion of ATF-2, which is also the binding region for UTF1. Therefore, we considered that these two motifs might also be involved in the direct interaction between UTF1 and ATF-2. We recently confirmed that the leucine zipper region of UTF1 and the metal-binding motif of ATF-2 are important for this interaction (A.Fukushima and M.Muramatsu, unpublished data). *In vitro*, UTF1 binds very weakly to full-length ATF-2 (Figure 7B). However, in transient transfection assays, both full-length and truncated forms of ATF-2 were shown to be equally activated by UTF1 expression. Although we do not know the molecular basis of these findings at present, these results may suggest the presence of an additional factor(s) or enzymes in mammalian cells which potentiate the magnitude of UTF1-dependent activation, especially in the context of full-length ATF-2 protein. Since the activity of the activation domain of ATF-2 is regulated by stress-activated kinases (Dam *et al.*, 1995; Livingstone *et al.*, 1995), it would be interesting to examine whether these kinases play such a role.

Co-immunoprecipitation analyses demonstrate that UTF1 interacts with the TFIID complex in EC cells. TFIID consists of TBP and several associated proteins called TAFs (for review see Verrijzer and Tjian, 1996). The TBP–TAF complex is known to be stable at high salt concentrations (Zhou *et al.*, 1993). In contrast, the interaction between UTF1 and TFIID is disrupted at moderately elevated KCl concentrations (>500 mM), indicating that UTF1 is not an integral component of the TFIID complex (data not shown). Furthermore, in comparable co-immunoprecipitation experiments, we observed that an anti-UTF1 antibody immnoprecipitates only ~5% of the amount of $TAF_{II}70$ brought down by an anti-TBP antibody (data not shown, see Figure 8B for reference). We observed essentially the same association pattern when we probed for $TAF_{II}250$ in immunoprecipitations performed using the same two antibodies (data not shown). Therefore, these results may indicate that UTF1 associates with just a subpopulation of the TFIID complex *in vivo*. However, as yet we cannot firmly eliminate the possibility that the reduced amount of TAF proteins co-immunoprecipitated with the UTF1 antibody simply reflects the weak interaction between UTF1 and TAFs and substantial amounts of UTF1 were dissociated from the TFIID complex during the immunoprecipitation procedure.

GST-pull-down analyses have revealed that the UTF1 protein binds to the evolutionarily conserved C-terminal half of TBP (A.Fukushima and M.Muramatsu, unpublished data). However, it remains to be determined whether UTF1 interacts directly with any of the TAF proteins or other general transcription factors. All of the biochemical studies, together with the expression profile, indicate that UTF1 has properties pertaining to a coactivator expressed in specific tissues during mouse early embryogenesis. To our knowledge, there are only a few other examples of coactivators that are highly restricted to specific cell types. One is OBF-1 (also called Bob1 or OCA-B), a B-cell-specific coactivator that stimulates transcription through its association with octamer-binding proteins (Gstaiger *et al.*, 1995; Luo and Roeder, 1995; Strubin *et al.*, 1995). The second is $TAF_{II}105$, which is also involved in B-cell-specific transcription (Dikstein *et al.*, 1996), although promoter-specific transcriptional activators that require the $TAF_{II}105$ have not yet been identified. In addition to these, tissue-specific TBP-related factor (TRF) has recently been shown to act as a functional homolog of TBP (Hansen *et al.*, 1997). Indeed, it was shown that TRF can form a stable complex with TFIIA and TFIIB on TATA box and substitute for TBP in directing RNA polymerase II-specific transcription *in vitro*. However, unlike TBP, TRF functions to direct tissue- and gene-specific transcription. These findings suggest that more than one TFIID complex exists in cells and some of them are tissue specific such as the one containing $TAF_{II}105$. Therefore, it may be worth systematically examining the possibility that UTF1 represents an embryonic cell-specific TAF, although data obtained from the immunoprecipitation analyses were somewhat against this possibility.

The regulation of gene expression during specific developmental events such as cell fate commitment is directed by a large number of promoter-specific transcription factors. To date, however, only a limited number of coactivators have been identified in cells. Hence, it may

be predicted that each coactivator is involved in multiple signal transduction pathways by cooperating with a large number of sequence-specific transcription factors. Indeed, CREB-binding protein has been shown to bind to numerous factors and to serve as an integrator of multiple regulatory circuits (Dai, 1996; Bhattacharya *et al.*, 1996; Chakravarti *et al.*, 1996; Kamei *et al.*, 1996; Avantaggiati *et al.*, 1997; Gu *et al.*, 1997; Lill *et al.*, 1997). Therefore, we anticipate that a more systematic screening may reveal additional promoter-specific factors which cooperate with UTF1. In any event, the isolation of UTF1 opens up new avenues for studying the unique regulation of transcription in pluri- and omnipotent cells.

Materials and methods

DNA manipulations

Site-directed mutagenesis was used to create an *Eco*RI site between –68 and –73 of the *HIS3* gene (Chen and Struhl, 1988). An *Eco*RI–*Xho*I fragment carrying the *HIS3* coding region in addition to the minimal promoter sequence was cloned into pRS414 (Sikorski and Hieter, 1989). The *TRP5* terminator (Miyajima *et al.*, 1984) was cloned upstream of the coding sequence into the *Sac*I–*Sac*II sites of pRS414 to prevent readthrough transcription. Four tandem arrays of the GPEI *cis*-element: 5′-GTAGTCAGTCACTATGATTCAGCAA-3′, which functions as a strong repressor element in yeast cells (Okuda *et al.*, 1989 and unpublished data), were inserted into the *Spe*I site of pRS414 to further reduce readthrough transcription. A cDNA library of F9 teratocarcinoma cells was prepared by conventional methods and cloned into the *Eco*RI site of pKT10GAL placing the cDNA downstream of the inducible *GAL 1.10* promoter (Tanaka *et al.*, 1990).

An E2A promoter fragment from position –249 to +63 was amplified by PCR using adenovirus DNA as a template, and substituted for the *Acc*I–*Hin*dIII fragment of pSV2CAT to create the E2A/CAT reporter plasmid shown in Figure 5A. The reporter carrying the mutated ATF-2 site (5′-GGCCGTA-3′), also shown in Figure 5A, was constructed according to the standard protocol for site-directed mutagenesis. The reporter plasmid used in Figure 6B carries the promoter region of *E2A* gene from –69 to +63. The collagen CAT–Myb reporter plasmid is described in Maekawa *et al.* (1991), in which they refer to this reporter as pMFcolCAT6MBS-I. Myb–TK CAT was constructed by cloning the six tandem repeats of the Myb binding site from collagen CAT–Myb into the *Bam*HI site of TK CAT (Luckow and Schütz, 1987). The Myb–ATF2 fusion protein expression vector is described in Maekawa *et al.* (1991) in which they refer to this expression vector as pact-c-myb–CRE-BP1. To construct other fusion protein expression vectors, the ATF-2 portion of the expression vector was replaced by either ATF-1, ATF-3, CREB or ATF-2 derivatives as a *Sph*I–*Xba*I fragment amplified by PCR. Expression vectors of E1A and UTF1 were constructed by subcloning the entire cDNAs into pCEP4 (Invitrogen).

Full-length and truncated forms of ATF-2 (activation and bZIP domains), as well as the C-terminal portion of UTF1 (206–339) were recovered as *Bam*HI–*Eco*RV fragments by PCR amplification and subcloned into the *Bam*HI–*Sma*I sites of pGEX-2T in-frame with the GST moiety.

Molecular cloning of the UTF1 cDNA

The F9 cDNA library was used to transform the yeast strain W303-1a (*MAT*α*, ade2-1, ura3-1, his3-11, trp1-1, leu2-3, 112, can1-100*) (Brill and Sternglanz, 1988), which contains a *HIS3* reporter on a CEN plasmid. Transformations were carried out using the lithium acetate method described by Ito *et al.* (1983). The resulting transformants were plated on minimal media containing glucose and supplemented with amino acids adenine, leucine and histidine. After 16 h of incubation at 30°C, the plates were replica plated onto minimal media containing galactose as the carbon source. All plates also contained AT at a concentration of 10 mM and were supplemented with adenine and leucine. After incubation for an additional 3 days, yeast colonies were picked up and grown at 30°C for 24 h in minimal media containing the required amino acids, including histidine. Three microliters of this overnight culture were transferred to AT-containing glucose or galactose plates. Only the yeast colonies which grew on the galactose-containing plates were used for further analyses. Plasmid DNAs were rescued from these yeast cells as described by Hoffman and Winston (1987) and reintroduced into the original recipient yeast strain.

DNA sequencing

The original UTF1 clone was digested with various restriction enzymes and the fragments were subcloned into Bluescript KSII+ derived vectors. Sequencing was carried out according to the dideoxy nucleotide procedure using the M13 universal or reverse primers.

Northern blot and RNase protection assays

Total RNA was prepared according to the guanidine hydrochloride procedure (Okayama *et al.*, 1987). Northern blot analysis was performed according to standard procedures. Poly(A)$^+$ RNA samples were isolated using an oligo-dT cellulose column, separated on a formaldehyde–agarose gel and transferred to nitrocellulose filters. The filters were hybridized with UTF1 or GAPDH (Maehara *et al.*, 1985) cDNA probes. The same poly(A)$^+$ RNA was used for the RNase protection analysis shown in Figure 3A. For preparing the riboprobe, the first 251 bp of the UTF1 cDNA were isolated by PCR using pCEP4/UTF1 as a template and subcloned into pSP64. Radiolabeled anti-sense UTF1 RNA was produced by the standard method according to the manufacturer (Promega) and hybridized with the RNAs as described previously (Okuda *et al.*, 1989).

In situ hybridization

To collect a large number of embryos, 8–12-week-old female mice (C3H×C57BL/6JF1) were super-ovulated by the administration of gonadotropins. Briefly, the mice were injected intraperitoneally with 7 units of human chorionic gonadotropin and then mated with C3H males. Pre-implantation embryos were recovered on 2.5–3.5 d.p.c. by flushing the oviducts with modified Krebs–Ringer solution. Post-implantation embryos were recovered on 6.5–8.0 d.p.c. by dissecting the uterine horns in PBS. Collected embryos were immediately used for *in situ* hybridization studies. Whole mount *in situ* hybridization was carried out essentially according to Wilkinson (1992). However, methanol treatment was omitted and, furthermore, a 10-fold lower concentrated proteinase K (1 μg/ml) was used in the manipulation of morulas and blastocysts. Sense and anti-sense UTF1 RNA probes encompassing 136–909, and that of Brachyury (1397–1921) (Herrmann *et al.*, 1990), were labeled with digoxygenine according to the procedure recommended by the supplier (Boehringer) and hybridized at 70°C. For the double-stained embryo shown in Figure 4I, Brachyury RNA was labeled with fluorescein and stained using 2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyl tetrazolium choride and 5-bromo-4-chloro-3-indolyl phosphate.

Production of antibodies against UTF1 and TBP

The C-terminal portion (206–339) of UTF1 or full-length TBP was fused to GST and expressed in *Escherichia coli* for purification using a glutathione–Sepharose column. The protein (100 μg) was mixed with Freund's adjuvant and injected subcutaneously into rabbit. Antiserum was recovered after six boosts and passed through an *E.coli* protein–Sepharose column followed by passage through a GST–Sepharose column. As a final step, the serum was purified using a GST–UTF1 or GST–TBP–Sepharose column.

Purification of recombinant UTF1 protein

For the expression of UTF1 in Sf9 cells, UTF1 cDNA was cloned into the pVL1392 baculovirus expression vector. Sf9 cell infection and WCE preparation were performed according to O'Reilly *et al.* (1992). Subsequently, a cell extract prepared from a 2 l culture was applied onto a phosphocellulose column equilibrated with buffer A (20 mM Tris–HCl pH 7.9, 100 mM KCl, 20% glycerol). UTF1 was eluted by increasing the KCl concentration up to 1 M. The eluted protein was diluted four times with KCl-free buffer and applied onto a heparin–Sepharose column equilibrated with buffer B (20 mM HEPES–KOH pH 7.9, 250 mM KCl, 20% glycerol); UTF1 was eluted with 1.0 M KCl and dialyzed against Dignam buffer D (Dignam *et al.*, 1983). Finally, the sample was mixed with anti-UTF1 antibody covalently bound to Sepharose. After extensive washing, bound protein was eluted under mild acidic conditions (0.2 M glycine–HCl pH 2.3) directly into neutralizing 1.0 M Tris–HCl pH 8.0 and then dialyzed against Dignam buffer D.

In vitro protein–protein interaction assay

The UTF1 and its derivatives, as well as E1A 13S, were subcloned into the *Nco*I–*Bam*HI site of pTM1 (Lee *et al.*, 1991) and *in vitro* transcribed/translated in rabbit reticulocyte lysates according to the manufacturer (Promega). GST–UTF1 and GST–ATF-2 fusion proteins were expressed

in *E.coli* by induction with 0.5 mM isopropyl β-D-thiogalactopyranoside. Cells were harvested after 2 h of induction and lysates were prepared as described by Hoey *et al.* (1988). GST-fusion proteins were purified using a glutathione–Sepharose column. For the interaction assay, purified GST or GST fusion proteins were recoupled to glutathione–Sepharose beads. Thirty microliters of the beads, in which ~5 μg of one of the GST fusion proteins were bound, were incubated with 20 μl of *in vitro* translated protein samples. After extensive washing, the beads were mixed with sample buffer and fractionated on SDS–polyacrylamide gels. The gels were treated with Amplify according to the manufacturer (Amersham) and exposed to films to detect the bound proteins. The interaction assay using recombinant UTF1 was performed in a similar manner, except that 50 μg of BSA was added during the incubation of UTF1 with the GST fusion proteins. To detect the bound recombinant UTF1, Western blot analysis was performed using anti-UTF1 antibody.

Immunoprecipitation

For the analyses shown in Figure 8B, anti-UTF1, ATF-2, E2F-1 or CREB antibody (obtained from Santa Cruz) was coupled directly to CNBr-activated Sepharose according to the manufacturer (Pharmacia). The antibody-bound Sepharose beads (10 μl) were mixed with 200 μl of F9 EC cell extract and incubated with rotation for 3 h at 4°C. Beads were washed extensively with IP buffer (20 mM HEPES–KOH pH 7.9, 200 mM KCl, 12.5 mM $MgCl_2$, 0.1% NP-40, 10% glycerol, 0.5 mM PMSF) and mixed with sample buffer without reducing reagent. 2-mercaptoethanol was added to the sample after removal of the Sepharose beads. The samples were then heat treated and fractionated by SDS–polyacrylamide gel electrophoresis. Anti-UTF1 antibody was used for detection of the protein on Western blots. For the analysis shown in Figure 8C, the same protocol was followed using polyclonal anti-UTF1 and TBP antibodies coupled to CNBr-activated Sepharose.

Accession No.

The accession number for the sequence reported in this paper is D31647.

Acknowledgements

The authors are indebted to Drs Yasuhisa Nogi, Koji Hisatake, Hiroshi Hamada, Tetsu-ichiro Saito, Arnold Berk, Hiroshi Miwatani and Satoshi Inoue for manipulation of yeast, baculovirus infection, differentiation of P19 cells, double-staining of an embryo, providing pTM-1 and pKT10GAL vectors, and computer analyses of amino acid sequence, respectively. We are also grateful to Nobuko Tominaga, Miyuki Satoh and Namiko Hihara for their excellent technical assistance throughout this study. We wish to thank Francis Stewart and Markus Meyer for discussions and critical reading of the manuscript. This work was supported in part by the Ministry of Education, Science, Sports and Culture and the Ministry of Health and Welfare, Japan and by the European Union, Human Capital and Mobility Program. A.O. was also supported by Sankyo Foundation of Life Science and Ciba-Geigy Foundation (Japan) for the promotion of Science. M.Keaveney was a recipient of a EU-HCM post-doctoral fellowship.

References

Avantaggiati,M.L., Ogryzko,V., Gardner,K., Giordano,A., Levine,A.S. and Kelly,K. (1997) Recruitment of p300/CBP in p53-dependent signal pathways. *Cell*, **89**, 1175–1184.

Bandara,L.R. and La Thangue,N.B. (1991) Adenovirus E1a prevents the retinoblastoma gene product from complexing with a cellular transcription factor. *Nature*, **351**, 494–497.

Berkenstam,A., Vivanco Ruiz,M.d.M., Barettino,D., Horikoshi,M. and Stunnenberg,H.G. (1992) Cooperativity in transactivation between retinoic acid receptor and TFIID requires an activity analogous to E1A. *Cell*, **69**, 401–412.

Bhattacharya,S., Eckner,R., Grossman,S., Oldread,E., Arany,Z., D'Andrea,A. and Livingston,D.M. (1996) Cooperation of Stat2 and p300/CBP in signalling induced by interferon-α. *Nature*, **383**, 344–347.

Brill,S.J. and Sternglanz,R. (1988) Transcription-dependent DNA supercoiling in yeast DNA topoisomerase mutants. *Cell*, **54**, 403–411.

Chakravarti,D., LaMorte,V., Nelson,N.C., Nakajima,T., Schulman,I.G., Juguilon,H., Montminy,M. and Evans,R.M. (1996) Role of CBP/p300 in nuclear receptor signalling. *Nature*, **383**, 99–103.

Chen,W. and Struhl,K. (1988) Saturation mutagenesis of a yeast *HIS3* 'TATA element': Genetic evidence for a specific TATA-binding protein. *Proc. Natl Acad. Sci. USA*, **85**, 2691–2695.

Cheng,X., Reginato,M.J., Andrews,N.C. and Lazar,M.A. (1997) The transcriptional integrator CREB-binding protein mediates positive cross talk between nuclear hormone receptors and the hematopoietic bZip protein p45/NF-E2. *Mol. Cell. Biol.*, **17**, 1407–1416.

Chien,C.T., Bartel,P.L., Sternglanz,R. and Fields,S. (1991) The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl Acad. Sci. USA*, **88**, 1338–1343.

Cohen,C. and Parry,D.A. (1994) α-Helical coiled coils: more facts and better predictions. *Science*, **263**, 488–489.

Colberg-Poley,A.M., Voss,S.D., Chowdhury,K. and Gruss,P. (1985) Structural analysis of murine genes containing homeo box sequences and their expression in embryonic carcinoma cells. *Nature*, **314**, 713–718.

Collart,M. and Struhl,K. (1994) NOT1(CDC39), NOT2(CDC36), NOT3, and NOT4 encode a global-negative regulator of transcription that differentially affects TATA-element utilization. *Genes Dev.*, **8**, 525–537.

Dai,P., Akimaru,H., Tanaka,Y., Hou,D.X., Yasukawa,T., Kanei-ishii,C., Takahashi, T. and Ishii,S. (1996) CBP as a transcriptional coactivator of c-Myb. *Genes Dev.*, **10**, 528–540.

Dam,H.V., Wilhelm,D., Herr,I., Steffen,A., Herrlich,P. and Angel,P. (1995) ATF-2 is preferentially activated by stress-activated protein kinases to mediate c-*jun* induction in response to genotoxic agents. *EMBO J.*, **14**, 1798–1811.

Diamond,M.L., Miner,J.N., Yoshinaga,S.K. and Yamamoto,K.R. (1990) Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. *Science*, **249**, 1266–1272.

Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.*, **11**, 1475–1489.

Dikstein,R., Zou,S. and Tjian,R. (1996) Human $TAF_{II}105$ is a cell type-specific TFIID subunit related to $hTAF_{II}130$. *Cell*, **87**, 137–146.

Gilbert, S.F. (1991) *Developmental Biology.* Sinauer Associate, Inc., Sunderland, MA.

Gorman,C.M., Rigby,P.W.J. and Lane,D.P. (1985) Negative regulation of viral enhancers in undifferentiated embryonic stem cells. *Cell*, **42**, 519–526.

Gstaiger,M., Knoepfel,L., Georglev,O., Schaffner,W. and Hovens,C.M. (1995) A B-cell coactivator of octamer-binding transcription factors. *Nature*, **373**, 360–362.

Gu,W., Shi,X.-L. and Roeder,R.G. (1997) Synergistic activation of transcription by CBP and p53. *Nature*, **387**, 819–823.

Hansen,S.K., Takada,S., Jacobson,R.H., Lis,J.T. and Tjian,R. (1997) Transcription properties of a cell type-specific TATA-binding protein, TRF. *Cell*, **91**, 71–83.

Helin,K., Lees,J.A., Vidal,M., Dyson,N., Harlow,E. and Fattaey,A. (1992) A cDNA encoding a pRB-binding protein with properties of the transcription factor E2F. *Cell*, **70**, 337–350.

Helin,K., Wu,C.-L., Fattaey,A.R., Lees,J.A., Dynlacht,B.D., Ngwu,C. and Harlow,E. (1993) Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperative trans-activation. *Genes Dev.*, **7**, 1850–1861.

Hen,R., Borrelli,E., Fromental,C., Sassone-Corsi,P. and Chambon,P. (1986) A mutated polyoma virus enhancer which is active in undifferentiated embryonal carcinoma cells is not repressed by adenovirus-2 E1A products. *Nature*, **321**, 249–251.

Herrmann,B.G., Labeit,S., Poustka,A., King,T.R. and Lehrach,H. (1990) Cloning of the T gene required in mesoderm formation in the mouse. *Nature*, **343**, 617–622.

Hoey,T., Warrior,R., Mank,J. and Levine,M. (1988) DNA-binding activities of the *Drosophila melanogaster* even-skipped protein are mediated by its homeodomain and influenced by protein context. *Mol. Cell. Biol.*, **8**, 4598–4607.

Hoffman,C.S. and Winston,F. (1987) A ten min DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *E.coli*. *Gene*, **57**, 267–272.

Hosler,B.A., LaRosa,G.P., Grippo,J.F. and Gudas,L.J. (1989) Expression of *REX-1*, a gene containing zinc finger motifs is rapidly reduced by retinoic acid in F9 teratocarcinoma cells. *Mol. Cell. Biol.*, **9**, 5623–5629.

Imperiale,M.J., Kao,H.T., Feldman,L., Nevins,J. and Strickland,S. (1984) Common control of the heat shock gene and early adenovirus genes: Evidence for a cellular E1A-like activity. *Mol. Cell. Biol.*, **4**, 867–874.

Ito,H., Fukuoka,Y., Murata,K. and Kimura,A. (1983) Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.*, **153**, 163–168.

Kamei,Y. *et al*. (1996) A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell*, **85**, 403–414.
Keaveney,M., Berkenstam,A., Feigenbutz,M., Vriend,G. and Stunnenberg,H.G. (1993) Residues in the TATA-binding protein required to mediate a transcriptional response to retinoic acid in EC cells. *Nature*, **365**, 562–566.
Kelly,F. and Condamine,H. (1982) Tumor viruses and early mouse embryos. *Biochim. Biophys. Acta*, **651**, 105–141.
LaRosa,G.L. and Gudas,L.J. (1988) An early effect of retinoic acid: cloning of an mRNA (Era-1) exhibiting rapid and protein synthesis-independent induction during teratocarcinoma stem cell differentiation of murine embryonic carcinoma stem cells. *Cell*, **49**, 507–513.
La Thangue,N.B. and Rigby,P.W.J. (1987) An adenovirus E1A-like transcription factor is regulated during differentiation of murine embryonic carcinoma stem cells. *Cell*, **49**, 507–513.
Lee,K.A.W., Hai,T.-Y., Sivaraman,L., Thimmappaya,B., Hurst,H.C., Jones,N.C. and Green,M.R. (1987) A cellular protein, activating transcription factor, activates transcription of multiple E1A-inducible adenovirus early promoters. *Proc. Natl Acad. Sci. USA*, **84**, 8355–8359.
Lee,W.S., Kao,C.C., Bryant,G.O., Liu,X. and Berk,A.J. (1991) Adenovirus E1A activation domain binds the basic repeat in the TATA box transcription factor. *Cell*, **67**, 365–376.
Li, X.-Y. and Green,M.R. (1996) Intramolecular inhibition of activating transcription factor-2 function by its DNA-binding domain. *Genes Dev.*, **10**, 517–527.
Lill,N.L., Grossman,S.R., Ginsberg,D., DeCaprio,J. and Livingston,D.M. (1997) Binding and modulation of p53 by p300/CBP coactivators. *Nature*, **387**, 823–827.
Liu,F. and Green,M.R. (1994) Promoter targeting by adenovirus E1a through interaction with different cellular DNA-binding domains. *Nature*, **368**, 520–525.
Livingstone,C., Patel,G. and Jones,N. (1995) ATF-2 contains a phosphorylation-dependent transcriptional activation domain. *EMBO J.*, **14**, 1785–1797.
Luckow,B. and Schütz,G. (1987) CAT constructs with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. *Nucleic Acids Res.*, **15**, 5490.
Luo,Y. and Roeder,R.G. (1995) Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. *Mol. Cell. Biol.*, **15**, 4115–4124.
Maehara,Y., Fujiyoshi,T., Takahashi,T., Yamamoto,M. and Endo,H. (1985) 1.5 kb mRNA abundantly expressed in rat tumors encodes 37 kilodalton protein *in vitro*. *Biochem. Biophys. Res. Commun.*, **131**, 800–805.
Maekawa,T., Matsuda,S., Fujisawa,J., Yoshida,M. and Ishii,S. (1991) Cyclic AMP response element-binding protein, CRE-BP1, mediates the E1A-induced but not the Tax-induced *trans*-activation. *Oncogene*, **6**, 627–632.
Martin,G.R. (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA*, **78**, 7634–7638.
Miyajima,A., Miyajima,I., Arai,K. and Arai,N. (1984) Expression of plasmid R388- encoded type II dihydrofolate reductase as a dominant selective marker in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **4**, 407–414.
Murphy,S.P., Garbern,J., Odenwald,W.F., Lazzarini,R.A. and Linney,E. (1988) Differential expression of the homeobox gene *HOX-1-3* in F9 embryonic carcinoma cells. *Proc. Natl Acad. Sci. USA*, **85**, 5587–5591.
Murray,E.J., Stott,D. and Rigby,P.W.J. (1991) Sequences and factors required for the F9 embryonic stem cell E1a-like activity. *Mol. Cell. Biol.*, **11**, 5534–5540.
Okamoto,K., Okazawa,H., Okuda,A., Sakai,M., Muramatsu,M. and Hamada,H. (1990) A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell*, **60**, 461–472.
Okayama,H., Kawaichi,M., Brownstein,M., Lee,F., Yokota,T. and Arai,K. (1987) High efficiency cloning of full-length cDNA. *Methods Enzymol.*, **154**, 3–28.
Okuda,A., Imagawa,M., Maeda,Y., Sakai,M. and Muramatsu,M. (1989) Structural and functional analysis of an enhancer GPEI having a phorbol 12-*O*-tetradecanoate 13-acetate responsive element-like sequence found in the rat glutathione transferase P gene. *J. Biol. Chem.*, **264**, 16919–16926.
Okuda,A., Imagawa,M., Sakai,M. and Muramatsu,M. (1990) Functional cooperativity between two TPA responsive elements in undifferentiated F9 cells. *EMBO J.*, **9**, 1131–1135.
O'Reilly,D.R., Miller,L.K. and Lukow,V.A. (1992) BACULOVIRUS EXPRESSION VECTORS. *A Laboratory Manual.* Oxford University Press, New York.
Rosner,M.H., Vigano,M.A., Ozato,K., Timmons,P.M., Poirier,F., Rigby,P.W.J. and Staudt,L.M. (1990) A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature*, **345**, 686–692.
Schöler,H.R., Ruppert,S., Suzuki,N., Chowdhury,K. and Gruss,P. (1990) New type of POU domain in germ line-specific protein Oct-4. *Nature*, **344**, 435–439.
Schöler,H.R., Ciesiolka,T. and Gruss,P. (1991) A nexus between Oct-4 and E1A: Implications for gene regulation in embryonic stem cells. *Cell*, **66**, 291–304.
Shew,J.-Y., Lin,B.T.-Y., Chen,P.-L., Tseng,B.Y., Yang-Feng,T. and Lee,W.-H. (1990) C-terminal truncation of the retinoblastoma gene product leads to functional inactivation. *Proc. Natl Acad. Sci. USA*, **87**, 6–10.
Sikorski,R.S. and Hieter,P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*, **122**, 19–27.
Strubin,M., Newell,J.W. and Matthias,P. (1995) OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell*, **80**, 497–506.
Tanaka,K., Nakafuku,M., Tamanoi,F., Kaziro,Y., Matsumoto,K. and Toh-e, A. (1990) *IRA2*, a second gene of *Saccharomyces cerevisiae* that encodes a protein with a domain homologous to mammalian ras GTPase-activating protein. *Mol. Cell. Biol.*, **10**, 4303–4313.
Verrijzer,C.P. and Tjian,R. (1996) TAFs mediate transcriptional activation and promoter selectivity. *Trends Biochem. Sci.*, **21**, 338–342.
Wagner,S. and Green,M.R. (1993) HTLV-1 Tax protein stimulation of DNA binding of bZIP proteins by enhancing dimerization. *Science*, **262**, 395–399.
Wilkinson, D.G. (1992) Whole mount *in situ* hybridization of vertebrate embryos. In Wilkinson,D.G. (ed.), *In situ hybridization. A practical approach*. IRL Press, Oxford, pp. 75–83.
Xin,J.-H., Cowie,A., Lachance,P. and Hassell,J.A. (1992) Molecular cloning and characterization of PEA3, a new member of the Ets oncoprotein family that is differentially expressed in mouse embryonic cells. *Genes Dev.* **6**, 481–496.
Yang-Yen,H.-F., Chambard,J.-C., Sun,Y.-L., Smeal,T., Schmidt,T.J., Drouin,J. and Karin, M. (1990) Transcriptional interference between c-jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein–protein interaction. *Cell*, **62**, 1205–1215.
Yao,Y., Yamamoto,K., Nishi,Y., Nogi,Y. and Muramatsu,M. (1996) Mouse RNA polymerase I 16-kDa subunit able to associate with 40-kDa subunit is a homolog of yeast AC19 subunit of RNA polymerase I and III. *J. Biol. Chem.*, **271**, 32881–32885.
Yuan,H., Corbi,N., Basilico,C. and Dailey,L. (1995) Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev.*, **9**, 2635–2645.
Zhou,Q., Boyer,T.G. and Berg,A.J. (1993) Factors (TAFs) required for activated transcription interact with TATA box-binding protein conserved core domain. *Genes Dev.*, **7**, 180–187.

Received September 26, 1997; revised February 9, 1998; accepted February 10, 1998