2068–2076 *Nucleic Acids Research, 1997, Vol. 25, No. 11*

Differential *in vivo* activation of the class II and class III snRNA genes by the POU-specific domain of Oct-1

Shona Murphy

Chemical Pathology Unit, Sir William Dunn School of Pathology, South Parks Road, Oxford, OX1 3RE, UK

Received February 24, 1997; Revised and Accepted April 14, 1997

ABSTRACT

Many snRNA genes contain binding sites for the ubiquitous transcription factor Oct-1. *In vitro* studies have shown that this factor potentiates binding of an essential transcription factor (PTF) to the proximal sequence element (PSE) of snRNA genes, and activates transcription. Using Gal4 fusion proteins, I show here that the POU-specific region of the DNA-binding domain of Oct-1 is sufficient both to potentiate PTF binding *in vitro* and to transactivate pol II- and pol III-dependent snRNA genes *in vivo*. A single amino acid change in this domain abrogates both activation and interaction with PTF. The N- and C-terminal regions of Oct-1 also activate transcription of both classes of snRNA genes. Wild-type levels of Pol II-dependent U2 expression require all activation domains, whereas efficient activation of the pol III-dependent 7SK and U6 genes is effected by the POU-specific domain alone. These results indicate that contacts between PTF and amino acids in the POU-specific domain of Oct-1 are critical for efficient transactivation of snRNA genes *in vivo*. The POU-specific domain of Oct-2A also activates these genes, but the N- and C-terminal domains are relatively inactive.

INTRODUCTION

The genes for the mammalian small nuclear RNAs (snRNAs) are transcribed by either pol II (e.g. U1–U5) or III (e.g. 7SK and U6) and have unusually simple and compact promoter structures (reviewed in 1,2). An essential proximal sequence element (PSE) is located at ~–50 in these genes and a second, enhancer-like element, often referred to as the distal sequence element or DSE, at ~–200. In the context of a PSE-containing promoter, an additional TATA box at –25 specifies transcription by pol III (3,4). Many of these genes are ubiquitously expressed and provide a good model to study the mechanics of the interactions between *trans*-acting factors required for basal and enhanced transcription by RNA polymerases II and III. The DSE enhances transcription of the snRNA genes from 10- to >100-fold and therefore fulfils a very important role in expression of these genes (reviewed in 1). In many cases the DSE contains binding sites for the octamer binding (Oct) factors that are also involved in the expression of diverse gene classes including the tissue-specific immunoglobulin genes (5) and the histone H2B gene (6 and refs therein). Octamer factor binding sites have been shown to be critical for efficient expression in both class II snRNA genes, for example the U2 gene (7), and class III snRNA genes, for example the U6 gene (8,9, reviewed in 5).

Oct-1 is probably responsible for transactivation of snRNA genes in most cell types since it is ubiquitously expressed, while expression of related factors is often tissue specific or developmentally regulated (reviewed in 10). For example, expression of Oct-2A is restricted mainly to B-cells. Oct factors contain a POU DNA binding domain that can be divided into a POU-specific domain, with a structure resembling the DNA binding domain of lambda repressor, and a POU-homeo domain similar to classic homeo domains (11–13). Although the DNA binding domains of Oct-1 and Oct-2A are highly homologous, there is little homology outside of this region (14) and this is reflected in differential activation of target genes. Tanaka *et al.* (15) showed that the N- and C-termini of Oct-1 can activate transcription of the U2 gene *in vivo* but are poor activators of a mRNA reporter. In these experiments the DNA binding domain of Oct-1 was swapped with the DNA binding domain of the related pituitary-specific factor, Pit-1, to direct the protein to a different binding site. An Oct-2A/Pit-1 hybrid activated transcription of the mRNA reporter much better than transcription of the U2 gene, suggesting that Oct-1 and Oct-2A have distinct abilities to activate different classes of gene. However, Yang *et al*. (16) have shown that, *in vivo*, Oct-2A with a Gal4 binding domain in place of the POU-homeo domain can transactivate a U2 gene containing Gal4 binding sites in place of the DSE. The only additional Oct-2A domain present in the study by Yang *et al*. (16) is the POU-specific domain and it is possible that this region is responsible for the transactivation of the U2 gene. The POU-specific domain of Pit-1 has 61% homology to that of Oct-2A (14), and the differences may affect the levels of activation detected in the study by Tanaka *et al*. (15). The DNA binding domain of either Oct-1 or Oct-2A can potentiate binding of the essential PSE-binding factor, PTF, to the PSE of snRNA genes, and the POU domain of Oct-1 can transactivate class III snRNA gene transcription *in vitro* (17). Thus the interaction between Oct-1 and PTF may play a critical role in the expression of snRNA genes. In which case, changes in the POU-domain that disrupt this interaction may well have a deleterious effect on transactivation of PSE-dependent genes.

Using Gal4 fusion proteins, Das *et al*. (18) found that multimerised short stretches of amino acids from Oct-1 activation domains were able to enhance snRNA transcription by both pol

Tel: +44 1865 275554; Fax: +44 1865 275556; Email: smurphy@worf.molbiol.ox.ac.uk

II and pol III better than those derived from Oct-2A. Also activation domains potent for enhancement of mRNA-type transcription were poor activators of snRNA gene expression, emphasising the differential responses of these promoter types to activators.

However, it is not yet clear what contribution the various domains of Oct-1 make to expression of class II and III snRNA genes *in vivo*, or if Oct-2A containing a POU-specific domain can activate either class of snRNA gene to the same level as Oct-1. Here I have used Gal4 binding domain fusion proteins to test the ability of regions of the DNA binding domain and the N- and C-termini of Oct-1 and Oct-2A to transactivate pol II- and pol III-dependent snRNA genes. Analysis of these regions, alone or in combination, gives a more complete picture of the requirements for efficient activation by DSE-bound Oct factors. In addition, the requirements for the different classes of gene have been compared.

MATERIALS AND METHODS

DNA templates and expression constructs

A U2 maxigene was made by inserting a linker containing an *Acc*65I site into the unique *Xba*I site of the marked U2 gene described by Mangin *et al*. (19), giving the sequence 5′-TCTAGTGGTACCAGCTCTAGA-3′. All reporters are in a pUC18 background. U2 DSE– was made by digesting the U2 maxigene with *Sma*I and religating to remove the region from –197 to –271 that contains the DSE. In U2 5G the *Sma*I fragment containing the DSE is replaced by five tandem repeats of a Gal4 binding site of sequence 5′-CGGAGTACTGTCCTCCG-3′ as described by Williams *et al*. (20). The 7SK maxigene has been described in (17) and the DSE– is identical to –210 described therein. In 7SK 5G, the DSE is replaced by five Gal4 binding sites of the sequence shown above, between the *Pst*I site at –243 and the *Bsm*I site at –210. In U6 maxigenes the promoter sequence of the 7SK maxigene from –243 to –8 or –200 to –8 is replaced by the U6 promoter sequence, between the *Pst*I site at –243 and the *Acc*65I site at –8. The resulting junction sequence at –8 is 5′-**AAAGGACG**GTACCCGAAG-3′. U6 sequences are shown in bold. In the U6 Gal4 reporter, five Gal4 sites were placed upstream of –200 in place of the DSE. In all Gal4 reporter genes the Gal4 binding sites span the position of the wild-type DSE.

The CMV promoter driven expression system described in (16) was used to express proteins in HeLa cells. The parent vectors pGAL4/DBD and pOct-A/GAL4 were a kind gift from K. Seipel and W. Schaffner. All clones were made using the *Hin*dIII site downstream from the CMV promoter, and the *Bam*HI site in the polylinker downstream from the multiple translation terminator in pGAL4/DBD, to replace the Gal4 binding domain with the desired coding sequence. Linker sequences were added where necessary and several constructs were made using PCR. To prepare PS1 the *Hin*cII to *Eco*RI restriction fragment from +805 to +1103 of the Oct-1 gene (21) was cloned into a vector containing an upstream Kosaks and ATG followed by an *Eco*RV site flanked by *Hin*dIII and *Eco*RI/*Sfi*I/*Pst*I/*Bam*HI. The *Sfi* I/*Bam*HI fragment from Oct-2A Gal, containing the binding domain of Gal4, was then cloned downstream. The resultant unit was cloned *Hin*dIII–*Bam*HI into pGAL4/DBD. The Pit-1 cDNA (22) used to PCR out the POU-specific domain was a kind gift from the laboratory of M. G. Rosenfeld. All extra amino acids generated by cloning are noted on the diagrams in Figures 1, 3 and 4. To change E7 to R in PS1 and P+NC the Stratagene Quick Change mutagenesis system was used following the manufacturers' instructions. The sequence was changed from GACCTTGAGGAG to GACCTccgGGAG as described in (23) for mutation of the same region in the POU-domain of Oct-1.

Transfections and RNA and protein preparation

HeLa cells (90 mm plates) were lipofected with 5 μg of reporter gene and 0.25–0.5 μg of pUC plasmid containing the VAI gene using Lipofectamine (Boehringer Mannheim) according to the manufacturer's instructions. RNA was collected as described in (24). Although this method was developed to harvest cytoplasmic RNA, the RNA encoded by the U2, 7SK and U6 reporters was efficiently extracted by this method and the results were identical when a method for extracting total RNA was used (25) (data not shown). For simultaneous harvesting of RNA and protein, cells were scraped into 1 ml PBS and half taken for either RNA or protein preparation. For protein preparation the cells were washed with 1 ml TEN (10 mM Tris–HCl, pH 8.0, 1 mM EDTA, NaCl) and resuspended in 20 mM HEPES (pH 7.9), 20% glycerol; 0.2 mM EDTA, 400 mM KCl, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride, 2 μg/ml aprotinin. Following three freeze/thaw cycles using solid CO_2/ethanol the cells were pelleted for 5 min in an Eppendorf centrifuge and the supernatant was flash frozen in liquid nitrogen and stored at –80°C.

Gel retardation analysis

Gel retardation was performed essentially as described in (17) with the addition of $ZnSO_4$ to 45 μM and dI–dC was 50 μg/ml. The proteins extracted from transfected cells were diluted for use where appropriate (usually ~10 times). The amount of extract used for quantitative gel retardation analysis was calculated to ensure a large excess of probe and the transfection efficiency was determined by S1 analysis of the VAI using the matched RNA sample. PTF used for the gel retardation assay shown in Figure 2 was from the second step of the purification described in (26) and gave a high background of non-specific binding in this assay, possibly due to the presence of Zn ions. For Figure 4D the PTF used was a more highly purified preparation corresponding to the third step of the purification described in (26) and was kindly prepared by Diana Boyd. The probes contained a single Gal4 binding site in place of the Oct-1 binding site in O^+P^+ or O^+P^- (17).

S1 nuclease analysis

Oligonucleotides complementary to part of the encoded U2, 7SK and VAI RNAs were ^{32}P labelled using T4 polynucleotide kinase. The U2 probe was: 5′-**CTACACTTGATCCTCTAGAGCTG-GTACCACTAGAGGATCTTAGCCAAAAGGCCGAGA-AGCGAT**GCGCTCGCCTTCGCGCCC-3′, the 7SK probe was: 5′-**CCTGGCGATCAATGGGGTGACAGAACAAGCTT-AGTGTCGCAGCCAGATCGCCCTCACATCC**GAGGTAC-CCAAGCGGCGCAC-3′ and the VAI probe was: 5′-**ATGATACCCTTGCGAATTTATCCACCAGACCACGG-AAGAGTGCCC**GCTTACAGGC-3′. The RNA complementary nucleotides are shown in bold. Typically, 100 fmol of probe was annealed to one tenth of the RNA from a 90 mm plate of cells at 40°C overnight and S1 digestion was carried out at 30°C for 1 h. Products were fractionated on 6% urea/acrylamide gels. Gels

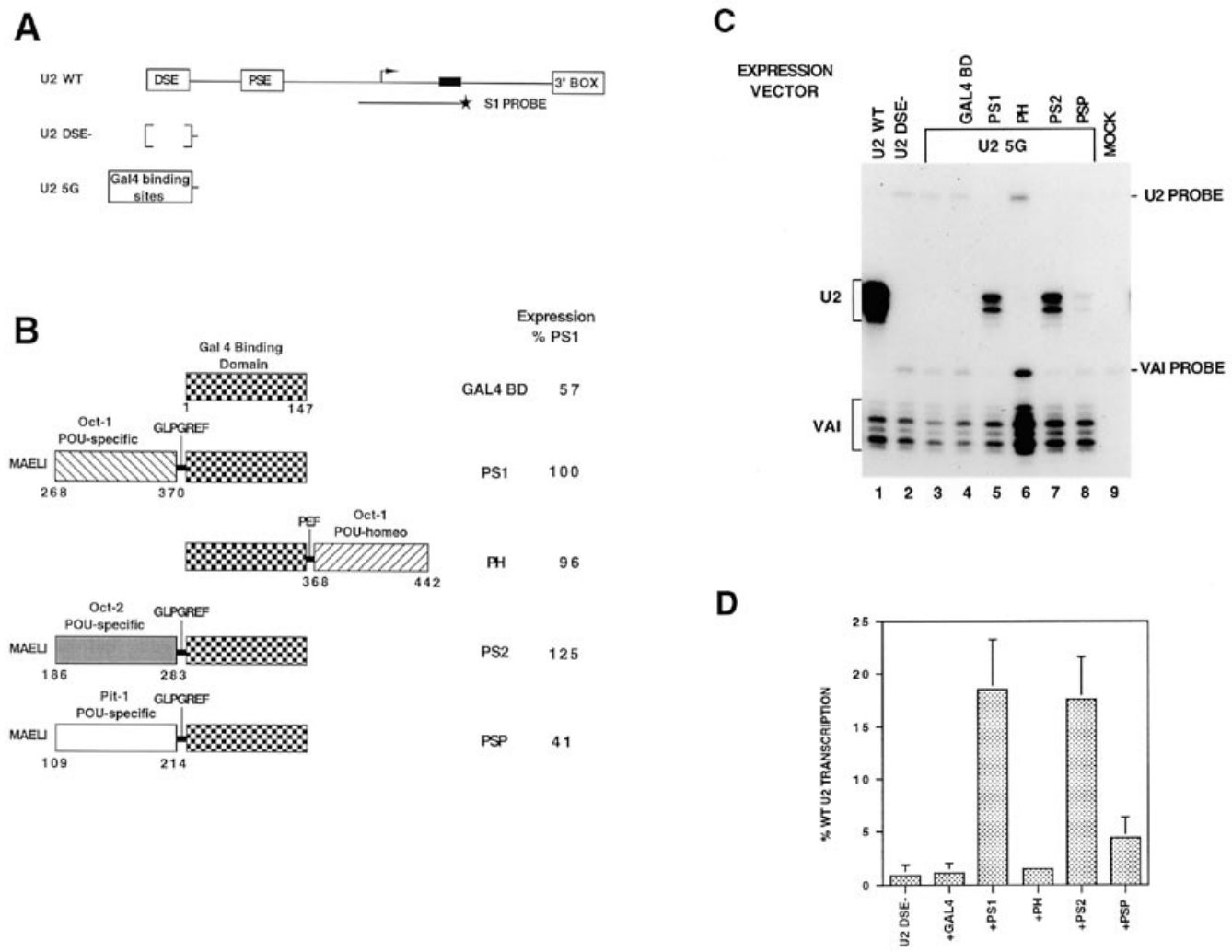

Figure 1. The POU-specific domain of Oct-1 activates pol-II transcription of the U2 snRNA gene. **(A)** The structure of the U2 maxigenes used for analysis. All constructs are identical apart from the removal of the DSE (U2 DSE–) or the replacement of the DSE with five Gal4 binding sites (U2 5G). DSE = distal sequence element, PSE = proximal sequence element, 3′ box = conserved downstream transcription termination signal. The arrow indicates the start site and direction of transcription. The black box indicates the position of an insert to distinguish the transfected U2 gene from the endogenous U2 gene. The star indicates the position of labelling of the S1 probe. **(B)** The structure of the chimeric proteins encoded by expression vectors co-transfected with U2 maxigenes for activation studies. In this and subsequent figures the numbers noted at the borders of the domains correspond to the position of the respective amino acid in the native protein. Extra amino acids encoded by linker sequences are noted where present. The % expression level as determined by gel retardation analysis (DNA binding activity) relative to PS1 is noted at the side of the figure. **(C)** S1 analysis of RNA from cells transfected with U2 reporter genes and constructs expressing the proteins described in (B). In this and subsequent experiments the positions of the probes and the protected DNA on the polyacrylamide gel are indicated on the diagram. The U2 maxigene and the expression plasmid co-transfected with the VAI and U2 genes are indicated above the lanes. MOCK indicates that no DNA was transfected. **(D)** Graphic representation of the results from a series of experiments testing U2 gene activation by the proteins shown in (A). GAL4 = Gal4 DNA binding domain here and in subsequent figures. Positive standard deviations only are shown.

were dried and used to expose either autoradiographic film or a phosphorimager screen. All quantitations were performed using a Molecular Dynamics PhosphorImager.

RESULTS

The POU-specific domain of Oct-1 activates pol II-dependent transcription of the U2 snRNA gene

The DNA binding POU domain of Oct-1 can activate pol III-dependent snRNA gene transcription *in vitro* and the POU domain of either Oct-1 or Oct-2A is sufficient to potentiate binding of PTF to the PSE of snRNA genes (17). However, further characterisation of the region required for PTF interaction is complicated by the fact that the interaction assay requires the DNA binding function of the Oct factor. In addition, *in vivo* analysis of the native POU domain requires that the problem of interference by endogenous Oct-1 is overcome. I have therefore used hybrid proteins containing the DNA binding domain of Gal4 to determine the ability of regions of the POU domain proteins Oct-1, Oct-2 and Pit-1 to activate snRNA gene transcription *in vivo*.

A U2 maxigene where the DSE was replaced by five Gal4 binding sites (Fig. 1A, U2 5G) was used to analyse the activation potential of various fusion proteins (Fig. 1B). VAI was used as a co-transfection control and transcripts were analysed by S1 nuclease mapping (see Materials and Methods). The VAI protected DNA is 45 nucleotides (nt) and the U2 protected DNA

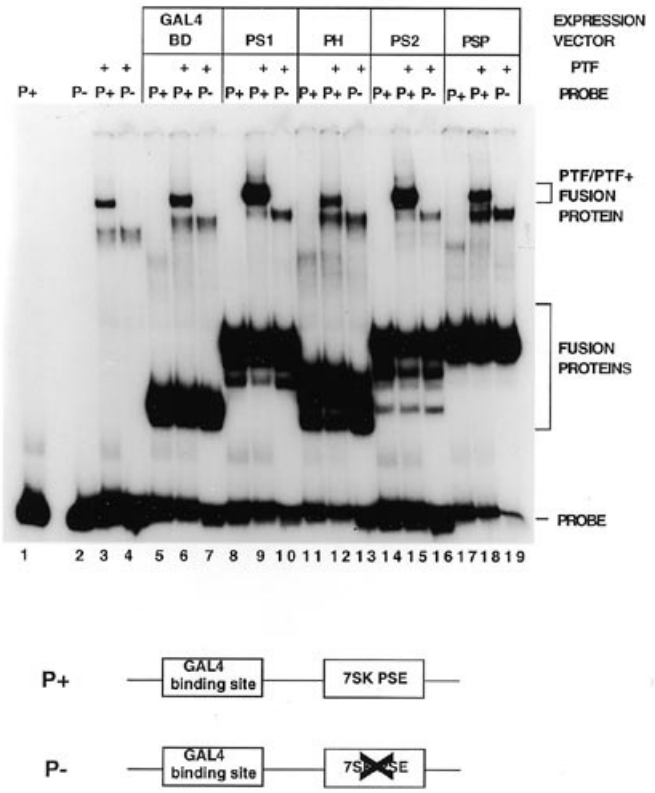

Figure 2. The POU-specific domains of Oct-1 and Oct-2A potentiate binding of PTF to its binding site. Gel retardation analysis of PTF in the absence and presence of the Gal4 fusion proteins described in Figure 1A. The position of probes, fusion proteins, PTF and PTF/fusion protein containing complexes is noted at the side of the figure. The structure of the probes is shown at the bottom and the probe used in each reaction noted above the figure. The inclusion of PTF and the vector encoding each fusion protein is noted above the lanes.

is 63 nt. The results of these experiments are shown in Figure 1C and expressed as percentages of wild-type U2 levels in a graphic representation of several experiments in Figure 1D. The relative level of functional expression, i.e. binding activity, of each protein was determined by gel retardation analysis of extracts from transfected cells. This is noted as a percentage of expression of the Oct-1 POU-specific/Gal4 binding domain protein, PS1 in Figure 1B and subsequent figures.

There is no signal corresponding to VAI or U2 transcripts when the cells are not transfected (Fig. 1C, Mock, lane 9). The wild-type U2 promoter (U2 WT, lane 1) gives a high level of expression of the marked U2 transcript and this has been designated as 100%. Removal of the DSE from the U2 gene, or replacement of the DSE by five Gal4 sites, reduces expression to less than 1% of the wild-type level (U2 DSE–, lane 2, 0.85%, U2 5G, lane 3, 0.86%). Co-expression of the Gal4 DNA binding domain alone with U2 5G has little effect on U2 expression (GAL4 BD, lane 4, 1.1%). However, the POU-specific domain of Oct-1 linked to the Gal4 binding domain clearly activates transcription of the U2 gene (PS1, lane 5, 18.43%), but not to the level of the wild-type gene. In contrast, the homeodomain of Oct-1 (PH, lane 6, 1.5%) has little effect on transcription of the U2 gene. The POU-specific domain of Oct-2A is as effective as the POU-specific domain of Oct-1 at activating U2 transcription (PS2, lane 7, 17.6%), whereas the POU-specific domain of the related transcription factor Pit-1 (PSP, lane 8) is much less effective than either PS1 or PS2, effecting only 4.3% of the wild-type level of transcription.

These results indicate that the POU-specific domains of Oct-1 and Oct-2A and, to a lesser extent, Pit-1, can activate transcription of the U2 gene in the absence of the POU-homeodomain and other parts of these proteins.

The expression levels of these proteins do differ slightly and PSP is expressed at just less than half the level of PS1 (Fig. 1B). However, the effect, if any, of these differences in expression is unclear since the levels may be saturating.

The POU-specific domains of Oct-1 and Oct-2A potentiate binding of PTF to its binding site

To determine whether the Gal4 fusion proteins functionally interact with PTF, I have used a gel retardation assay (Fig. 2). A probe containing a Gal4 binding site adjacent to the PSE of the 7SK gene was prepared and a matched probe containing a mutated PSE was used as a control (see Fig. 2). The Gal4 binding site is more closely juxtaposed to the 7SK PSE in the probe than the DSE is in the wild-type promoter. However, 7SK gene constructs where Oct-1 binding sites are placed close to the PSE behave essentially the same as the wild-type promoter *in vivo* (17). Extracts from transfected HeLa cells were used as a source of the Gal4 fusion proteins and PTF was assayed for binding in the presence and absence of fusion proteins bound to the Gal4 site. PTF alone binds poorly to the probe (lane 3) and Gal4 has little effect on PTF binding (lane 6, 1.3-fold). However, both PS1 and PS2 increase PTF binding by ~5-fold (lanes 9 and 15), equivalent to that effected by the entire POU domain of Oct-1 using the same preparation of PTF (data not shown). PSP slightly enhances PTF binding (lane 18, 1.8-fold), while PH gives the same amount of potentiation as the Gal4 binding domain alone (lane 12, 1.2-fold).

Thus, there is a correlation between the ability of the fusion proteins to potentiate PTF binding and activate U2 transcription.

The N- and C-terminal domains of Oct-1 further enhance U2 transcription

Tanaka *et al.* (15) have used a reporter system where the POU-domain of Oct-1 is replaced by the POU domain of the pituitary-specific transcription factor Pit-1 to show that the N- and C-terminal domains of Oct-1 can specifically enhance transcription of the U2 gene. I have therefore tested the effect of addition of these domains to the minimal PS1 construct. The relative activation potential of the N- and C-termini of Oct-2A was also assessed by using the Oct-2A construct pOct-2AGAL4 described by Yang *et al.* (16), that encodes Oct-2A with the POU-homeo domain replaced by the Gal4 DNA-binding domain. Figure 3A shows the structure of the chimeric proteins used for this analysis. The results of S1 analysis are shown in Figure 3B and represented graphically in Figure 3C.

The addition of either the N- or the C-terminus of Oct-1 results in a marked increase of U2 transcription (P+N, 64.2% and P+C, 74%, Fig. 3B, lanes 6 and 7) compared with the POU-specific domain alone (PS1, lane 5). A protein including both termini gives slightly higher activation than proteins with either region alone (P+NC, lane 8, 87.8%). The 'full length' Oct-2A (O2G,

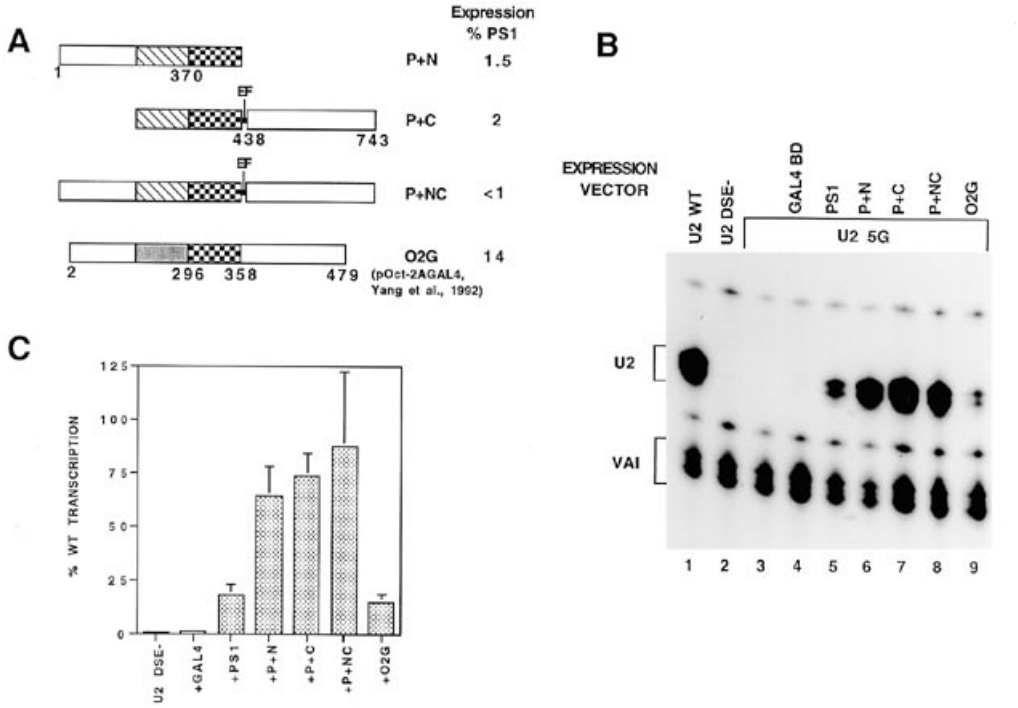

Figure 3. The N- and C-terminal domains of Oct-1 further enhance U2 expression. **(A)** The structure of the chimeric proteins encoded by expression vectors co-transfected with U2 maxigenes. Linker amino acids already noted in Figure 1(A) have been omitted. **(B)** S1 analysis of RNA from cells transfected with U2 reporter genes and constructs expressing the proteins described in (A). **(C)** Graphic representation of the results of a series of experiments testing activation of U2 expression by the proteins described in (A).

lane 9) gives approximately the same activation as the Oct-2A POU-specific domain alone (17.6%).

These results indicate that the N- and C-domains of Oct-1 also activate U2 transcription in this assay system. In contrast, the mapped activation domains of Oct-2A appear to have relatively little effect on U2 transcription.

The levels of expression of P+N, P+C and P+NC (noted in Fig. 3A) are extremely low relative to PS1. Therefore, the activation levels measured may be an underestimate if the binding sites on the reporter gene are not saturated. O2G, however, is expressed at a much higher level than P+NC, indicating that the relatively low activation effected by this fusion protein reflects the reduced ability of the N- and C-terminal domains of Oct-2A to activate U2 transcription.

A single amino acid change in PS1 reduces U2 activation and the potentiation of PTF binding

As noted above, the POU-specific domain of Pit-1 does not efficiently potentiate PTF binding *in vitro*. Mittal *et al.* (23) also found that the POU domain of Pit-1 interacts poorly with the PTF-related complex $SNAP_c$ (see Discussion). By comparing the amino acid sequences of the POU domains of Oct-1 and Pit-1, and exchanging amino acids between the two, these authors have identified the glutamic acid at +7 in the POU-specific region of Oct-1 as a critical determinant of $SNAP_c$ interaction. At the same position in the POU-specific domain of Pit-1 there is an arginine. To test the effect of mutating the glutamic acid at +7 of the Oct-1 POU-specific region on U2 activation *in vivo* and on PTF potentiation *in vitro*, I have changed it to arginine in both PS1 and P+NC constructs. I have also attached the N- and C-terminal domains of Oct-1 directly to the binding domain of Gal4 to test the activation potential of these regions in the complete absence of the POU-domain. Diagrams of these constructs are shown in Figure 4A and the results of analysis in Figures 4B and C.

The amino acid change from glutamic acid to arginine (E to R) at position +7 of the POU-specific domain has a dramatic effect on the level of activation of the U2 gene in the context of PS1 (PS1M, Figure 4B, lane 6), reducing the transcription level from 18.43 to 5.3% and in the context of P+NC (P+NCM, lane 8) reducing the level from 87.8 to 37%. It is notable that the level of activation by PS1M is very close to that effected by the POU-specific domain of Pit-1 (PSP, Fig. 1D, 4.3%).

The N- and C-terminal domains of Oct-1 alone give activation to 18 and 11.2% of the wild-type U2 level respectively (lanes 9 and 10), indicating that these domains can activate independently of the POU-specific domain. Interestingly, the level of activation affected by a combination of the POU-specific domain and either the N- or C-terminus (P+N and P+C, see Fig. 3D) is more than the sum of activation by each domain alone.

Gel retardation analysis to test the effect of this E to R change on PTF interaction is shown in Figure 4D. Compared with PS1, PS1M has a much reduced ability to potentiate PTF binding (compare lanes 8 and 11).

Since PS1M is expressed at the same level as PS1 (see Fig. 4A), reduced activation can be directly attributed to the one amino acid change and is likely to be due to the reduced interaction with PTF. However, PS1M still potentiates PTF binding significantly better than the Gal4 DNA binding domain alone (compare lanes 5 and

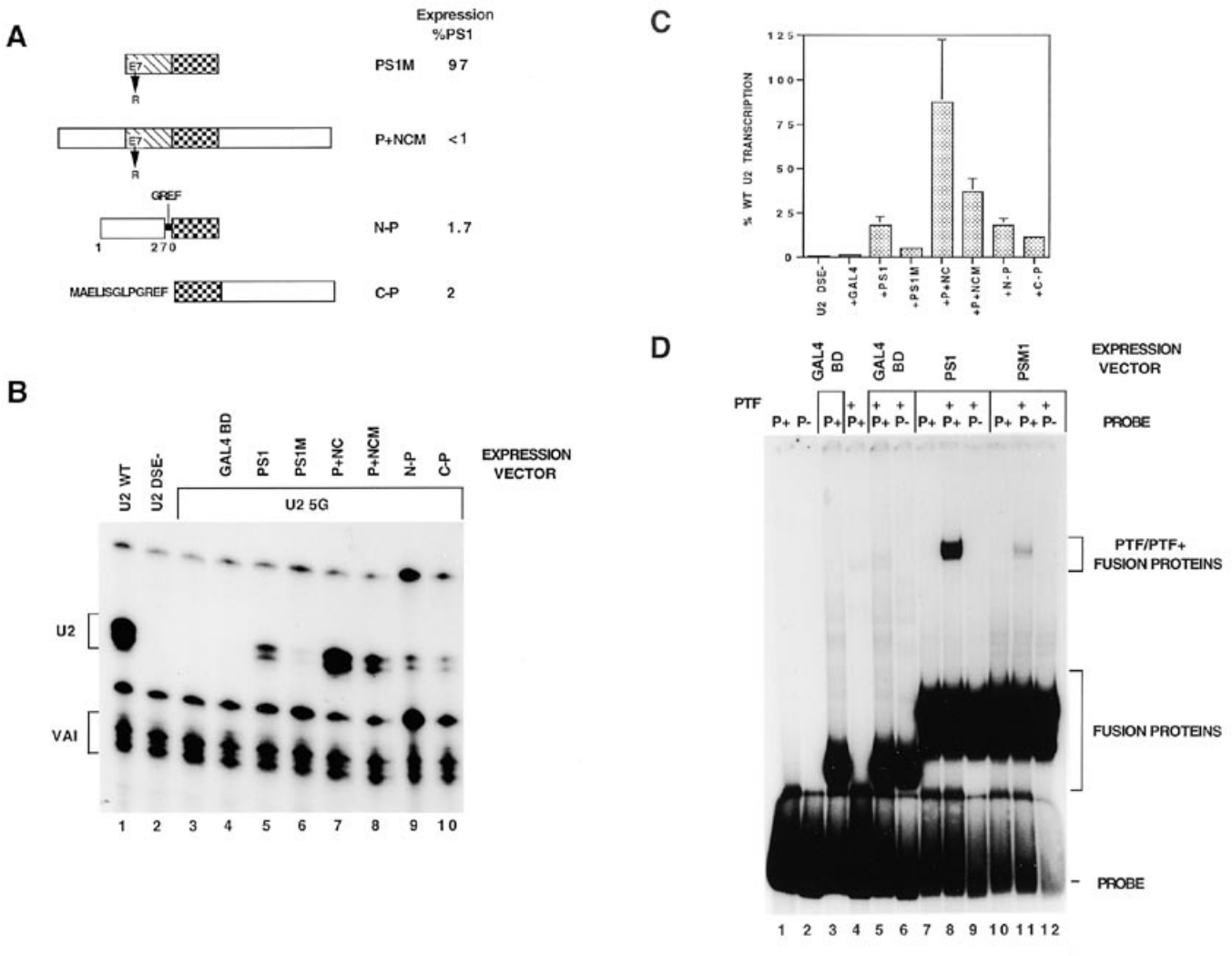

Figure 4. One amino acid change in PS1 reduces U2 activation and the potentiation of PTF binding. **(A)** The structure of the chimeric proteins encoded by expression vectors transfected with U2 maxigenes. **(B)** S1 analysis of RNA from cells transfected with U2 reporter genes and constructs expressing the proteins described in (A). **(C)** Graphic representation of the results of a series of experiments testing activation of U2 expression by the proteins shown in (A). **(D)** Gel retardation analysis of PTF in the absence and presence of the Gal4 fusion proteins described in (A). Details are as for Figure 2.

11). This may explain why this protein can still activate to some extent *in vivo.*

P+NC and P+NCM are expressed at very low but equivalent levels allowing a direct comparison of activation levels. Thus, a one amino acid change can also significantly affect activation in this context.

The POU-specific domain of Oct-1 alone efficiently activates 7SK and U6 transcription *in vivo*

Since the POU-specific domain of Oct-1 is sufficient to activate class II-snRNA gene transcription *in vivo*, I have also investigated the effect of the PS1 and other fusion proteins on transcription of the pol III-dependent 7SK and U6 snRNA genes in transfected HeLa cells. The enhancer region of marked 7SK and U6 genes was replaced by five Gal4 binding sites. The structure of the 7SK maxigenes used is shown in Figure 5A. U6 maxigene constructs are very similar (see Materials and Methods). The expression constructs described in Figures 1–4 were co-transfected with the reporters, RNA was collected and assayed by S1 analysis. VAI was used as a transfection control as before and the protected DNA for the 7SK and U6 reporters was 61 nt.

The results of analysis of transactivation of the 7SK reporter gene (7SK 5G) by the various fusion proteins are shown in Figure 5B and the results of several experiments using 7SK or U6 reporters are illustrated graphically in Figure 5C.

7SK WT expresses a high level of marked 7SK RNA (Fig. 5B, lane 1). Removal of the DSE (7SK DSE–, lane 2) or replacement of the DSE with five Gal4 binding sites (7SK 5G, lane 3) reduces transcription to 6–8% of 7SK WT levels. Co-transfection of the Gal4 binding domain with 7SK 5G has little effect (lane 4). Surprisingly, the POU-specific domains of both Oct-1 and Oct-2A activate 7SK transcription to essentially wild-type levels (Fig. 5B, PS1, lane 5, 108.3% and PS2, lane 7, 95%). As in the case of U2 gene activation, the POU-homeo domain has little effect (PH, lane 6, 9.1%) and the POU-specific domain of Pit-1 (PSP, lane 8, 28%) is much less effective than either PS1 or PS2. Thus, activation by these fusion proteins appears to correlate with PTF interaction for both the class II U2 gene and the class III 7SK gene.

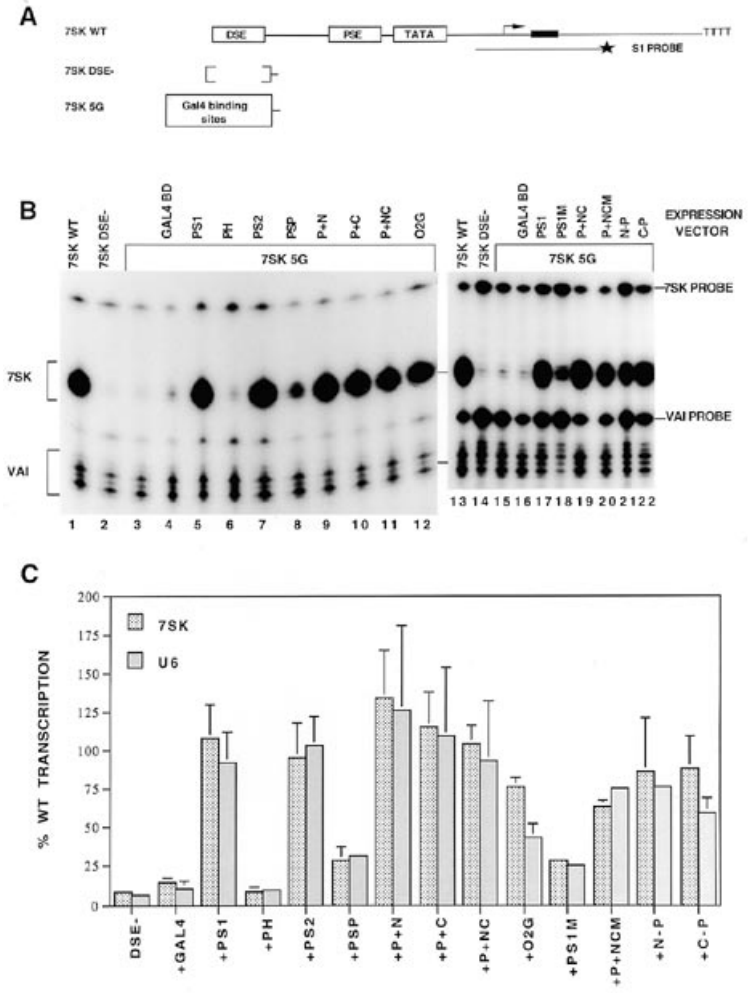

Figure 5. The POU-specific domain of Oct-1 alone efficiently activates 7SK and U6 expression *in vivo*. **(A)** The structure of the 7SK maxigenes used for transactivation analysis. All constructs are identical apart from the removal of the DSE (7SK DSE–) or the replacement of the DSE with five Gal4 binding sites (7SK 5G). TATA = the T/A rich sequence at –25. TTTT = the transcription termination signal. All other details are as for Figure 1A. **(B)** S1 analysis of RNA from cells transfected with 7SK maxigenes and constructs expressing the proteins described in Figures 1B, 3A and 4A. **(C)** Graphic representation of the results of a series of experiments testing activation of 7SK and U6 expression by the proteins described in Figures 1B, 3A and 4A.

Addition of either the N- or C-terminal domain of Oct-1 increases activation slightly (P+N, lane 9, 142% and P+C, lane 10, 118%) whereas addition of both regions gives the same level of activation as PS1 (P+NC, lane 11, 103.75%). The 'full length' Oct-2A fusion protein activates less than PS2 (O2G, lane 12, 68%).

Mutation of the glutamic acid at +7 of the POU-specific domain to arginine in the context of PS1 (PS1M, see Fig. 4A) reduces activation to <30% of the level effected by PS1 (lane 18, 28.75%), suggesting that, also in the case of the class III snRNA genes, activation is dependent on PTF interaction. However, mutation of the POU-specific domain in the presence of the N- and C-terminal domains has a less dramatic effect (compare P+NC, lane 19, 103.75% and P+NCM, lane 20, 63%). Both the N- and C-terminal domains significantly activate 7SK and U6 transcription in the absence of the POU-specific domain (N-P, lane 21, 86.5%, C-P, lane 22, 88.5%) although not quite as well as the POU-specific domain alone.

The results are very similar for the 7SK and U6 reporters (see Fig. 5C), except that the O2G protein activates the U6 gene to <50% of the wild-type and the 7SK gene to 68% of the level of 7SK WT.

Clearly, the minimal POU-specific domain of Oct-1 is able to activate pol III-dependent transcription of snRNA genes to wild-type levels. Indeed, the addition of domains capable of significant activation of the U2 gene appears to have little additional effect. However, it is not clear if higher levels of activation would be affected if P+N, P+C and P+NC and P+NCM were expressed at higher levels. Interestingly, the other domains outside the DNA binding domain can also independently efficiently activate pol III-dependent transcription of snRNA genes.

As I found for the U2 gene analysis, O2G is a less effective activator than the Oct-1 equivalent P+NC, suggesting that the N- and C-terminal domains of Oct-2A also have a reduced ability to activate class III snRNA genes.

DISCUSSION

Several of the factors that bind directly to the promoters of snRNA genes have been identified. TBP binds directly to the TATA box of the class III snRNA genes (27,28). This factor is also required for transcription of the class II snRNA genes but is recruited indirectly in this case (see below). Three activities

binding specifically to the PSE have been described: factors PTF (17), and PBP (29), and the complex $SNAP_c$ (30), but these are likely to be very similar in composition (see 17,23,28,31,32 for discussions of this point). $SNAP_c$ contains TBP in association with at least some of the same subunits present in PTF, and interaction with the PSE-binding factor may deliver TBP to the TATA-less, class II snRNA genes (30,33). While the TATA box and PSE elements are reasonably well conserved in size, position and sequence in mammalian snRNA genes, DSEs are more heterogeneous and may be recognised by a variety of factors including Oct-1 and Sp1 (see 1 for a review). Octamer binding sites are present in most cases and usually play a critical role in DSE function. Thus, dissection of the mechanism of activation through these sites will help to understand DSE function. Since the ubiquitous Oct-1 and the B-cell restricted Oct-2 bind equally well to the octamer site there is the additional question of the relative function of these proteins in tissues where they are both expressed.

The POU-specific domain of Oct-1 activates snRNA gene transcription through interaction with PTF

The POU-specific domain of Oct-1 activates U2, 7SK and U6 gene transcription in the absence of any other part of the parent protein. This is perhaps surprising since this region is an integral part of the DNA binding domain and serves to impart greater sequence recognition and specificity to DNA binding (34), whereas activation and DNA binding are usually separate functions of modular transcription factors. However, the involvement of the Oct-1 DNA binding domain in activation is not restricted to this example. Oct-1 is required for the recruitment of cellular co-factors and the VP16 activator protein to binding sites in the promoters of *Herpes simplex* genes. The POU domain is sufficient for this process and amino acids within the homeo domain are critical for this interaction (35–37). The POU domain of Oct-1 is also sufficient for activation of H2B gene transcription, possibly through interaction with an H2B gene specific co-activator (6). In addition, the B-cell co-activator, variously known as OCA-B, Bob1 and OBF-1 (6,38,39), is recruited to immunoglobulin gene promoters through interactions with the POU domain of Oct-1 and Oct-2, and amino acids in the POU-specific domain are required for this interaction (40). However, the results presented here represent the first demonstration that the short segment of the POU-specific domain by itself can function as an activator *in vivo*.

Conversion of the glutamic acid at +7 to arginine within the isolated POU-specific domain decreases interaction with PTF and activation *in vivo* confirming that interaction is required for activation in the cell. The POU-homeodomain is neither sufficient nor necessary for activation or PTF interaction. A surface for PTF interaction, therefore, lies entirely within the POU specific domain.

Efficient activation of transcription of the U2 gene by Oct-1 requires N- and C-terminal regions in conjunction with the POU-specific domain

Although the POU-specific domain of Oct-1 activates transcription of both classes of snRNA gene, the extent of activation, compared with wild-type levels, is not the same. The level of activation of the U2 reporter is relatively low and this is significantly increased by adding regions N- or C-terminal to the DNA binding domain. These regions improve activation when added singly and even better when added together. In the absence of a functional POU-specific domain the N- and C-terminal regions activate poorly, emphasising that efficient activation requires more than one activation function.

Transcription of both the 7SK and U6 genes, on the other hand, is activated to wild-type levels by the POU-specific domain of Oct-1 alone. The regions located N- and C-terminal do not greatly increase activation in the presence of the POU-specific domain. However, the level of 7SK and U6 transcription is relatively high in the absence of the enhancer region. This may be because these class III genes have an additional TATA box that recruits TBP directly, making preinitiation complex formation more efficient in the absence of the DSE. In fact, the POU-specific domain of Oct-1 increases transcription of the U2 reporter by ~18-fold and 7SK and U6 transcription by roughly the same amount. The involvement of all activation domains may be crucial for high-level transcription of genes with an otherwise weak promoter. Indeed, loss of interaction between PTF and the POU-specific domain of Oct-1 unmasks the potential of the N- and C-termini to activate the class III genes. Also, the isolated N- and C-terminal domains almost restore wild-type levels of transcription, although neither is as active as the POU-specific domain. Each domain, then, has the ability to elicit high levels of pol III-dependent transcription. This redundancy of activation domains in one of the key activators of snRNA gene transcription parallels the redundancy of transcription factor binding sites within the promoter of the 7SK gene (28).

What are the targets for the N- and C-terminal domains of Oct-1?

Oct-1 has stretches of glutamines N-terminal to the DNA binding domain and a serine/threonine rich region C-terminal to the DNA binding domain. Tanaka *et al.* (15) have shown that the N-terminus of Oct-1 activates U2 transcription *in vivo* but is unable to activate a β-globin reporter. A multimer of a short glutamine-rich stretch from this domain can also selectively activate snRNA genes (18). Tanaka *et al.* (15) also found that the C-terminus of Oct-1 was a better activator of U2 transcription than the N-terminus and mapped activation to an 133 amino acid region close to the DNA binding domain containing the serine/threonine-rich stretch. In this study the N-terminus is a better activator of snRNA gene transcription which may be due to differences in the assay system. Although Oct-2A also has a serine/threonine stretch in the C-terminal region, this is not active in transcription of the U2 gene (here and ref. 15), indicating that the Oct-1 activation domain consists of more than just a high concentration of these amino acids.

The extreme C-terminus of Oct-1 has a region of homology to the C-terminus of Oct-2B, an alternatively spliced variant from the Oct-2A gene, and Tanaka *et al.* (15) have shown that the C-terminus of Oct-2B can activate U2 transcription. Thus, there may be two distinct activator regions located in the C-terminus of Oct-1. The targets of these activation domains are not known. The POU-specific domain works synergistically with either of the terminal domains for activation of the U2 gene, suggesting that the targets are distinct but that the pathway affected is the same. Either the targets are shared between the pol II- and pol III-specific initiation complexes, or different parts of the domains

must contact distinct targets in each system, which seems unlikely. Shared components include PTF and TBP, and there is homology between one of the components of the pol III-specific factor TFIIIB and the pol II-specific factor TFIIB (41), that are required for transcription of class II and class III snRNA genes respectively (26,41,42). However, TBP and TFIIB are also required for transcription of mRNA genes and the targets appear to be specific for snRNA genes. PTF is therefore a likely candidate. Alternatively, targets may be co-factors or basic factors that are as yet unidentified components of snRNA gene preinitiation complexes. Subunits of PTF/SNAP$_c$ interact directly with TBP (33,43,44) and PTF/SNAP$_c$ potentiates binding of TBP to the TATA box of class III snRNA genes (45; D.C.Boyd and S.Murphy, unpublished results). In the absence of a strong interaction of PTF with the PSE, or in the absence of a TATA box, TBP may respond directly or indirectly to the N- and C-terminal domains of Oct-1. In any case, it is likely that these regions, like the POU-specific domain, have an effect on the efficiency of initiation through stabilisation of the pre-initiation complex.

Since Oct-1 interacts with at least two distinct targets, its association with the transcription complex is likely to be very stable. In contrast, Oct-2 interacts only through the POU-specific domain and may be more easily displaced. Thus, Oct-1 is probably responsible for activation even when Oct-2A is also present in the cell.

In conclusion, this study demonstrates that the Oct-1/PTF interaction detected *in vitro* is crucial for efficient expression of snRNAs *in vivo* and supports our earlier suggestion that other activation domains work in conjunction with the POU-specific region of the DNA binding domain (17).

ACKNOWLEDGEMENTS

I would like to thank Diana C. Boyd and Paula Cuello for all their help and support during this work and for their comments on the manuscript and Paris Margaritis for helping so much with the figures. I am very grateful to Robert G. Roeder, Jong-Bok Yoon, Neil Davies, Dave Pritlove, Charlie Birse, Ana Pombo and Gunnar Westin for many helpful discussions. I am especially indebted to Nick Proudfoot and George Brownlee for space, support, patience and kindness. I am an MRC Senior Fellow supported by MRC grant No. G9228550MB.

REFERENCES

1 Hernandez, N. (1992) In McKnight, S. L., and Yamamoto, K. R. (eds.), *Transcriptional Regulation.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, Vol. 1, pp. 281–313.
2 Lobo, S. M. and Hernandez, N. (1994) In Conaway, R. C. and Conaway, J. W. (eds.), *Transcription, mechanisms and regulation.* Raven Press, New York, pp. 127–159.
3 Mattaj, I. W., Dathan, N. A., Parry, H. D., Carbon, P. and Krol, A. (1988) *Cell*, **55,** 435–442.
4 Lobo, S. M. and Hernandez, N. (1989) *Cell*, **58,** 55–67.
5 Staudt, L. M. and Lenardo, M. J. (1991) *Annu. Rev. Immunol.*, **9,** 373–398.
6 Luo, Y. and Roeder, R. G. (1995) *Mol. Cell. Biol.*, **15,** 4115–4124.
7 Janson, L., Weller, P. and Pettersson, U. (1989) *J. Mol. Biol.*, **205,** 387–396.
8 Kunkel, G. R. and Pederson, T. (1988) *Genes Dev.*, **2,** 16–204.
9 Das, G., Henning, D., Wright, D. and Reddy, R. (1988) *EMBO J.*, **7,** 503–512.
10 Scholer, H. R. (1991) *Trends Genet.*, **7,** 323–328.
11 Assa-munt, N., Mortishire-Smith, R. J., Aurora, R., Herr, W. and Wright, P. E. (1993) *Cell*, **73,** 193–205.
12 Dekker, N., Cox, M., Boelens, R., Verrijzer, C. van der Vliet, P. C. and Kapstein, R. (1993) *Nature*, **362,** 852–855.
13 Klemm, J. D., Rould, M. A., Aurora, R, Herr, W. and Pabo, C. O. (1994) *Cell*, **77,** 21–32.
14 Herr, W., Sturm, R. A., Clerc, R. G., Corcoran, L. M., Baltimore, D., Sharp, P. A., Ingraham, H. A., Rosenfeld, M. G., Finney, M., Ruvkun, G. *et al.* (1988) *Genes Dev.*, **2,** 1513–1516.
15 Tanaka, M., Lai, J.-S. and Herr, W. (1992) *Cell*, **68,** 755–767.
16 Yang, J., Muller-Immergluck, M. M., Seipel, K., Janson, L., Westin, G., Schaffner, W. and Pettersson, U. (1991) *EMBO J.*, **10,** 2291–2296.
17 Murphy, S., Yoon, J. B., Gerster, T. and Roeder, R. G. (1992) *Mol. Cell. Biol.*, **12,** 3247–3261.
18 Das, G., Hinkley, C. S. and Herr W. (1995) *Nature*, **374,** 657–660.
19 Mangin, M., Ares, M. and Weiner, A. M. (1986) *EMBO J.*, **5,** 987–995.
20 Williams, R. D., Lee, B. A., Jackson, S. P. and Proudfoot, N. J. (1996) *Nucleic Acids Res.* **24,** 549–557.
21 Sturm, R. A., Das, G. and Herr, W. (1988) *Genes Dev.* **2,** 1582–1599.
22 Ingraham, H. A,, Chen, R., Mangalam, H. J., Elsholtz, H. P., Flynn, S. E., Lin, C. R., Simmons, D. M., Swanson, L. and Rosenfeld, M. G. (1988) *Cell*, **55**, 519–529.
23 Mittal, V., Cleary, M. A., Herr, W. and Hernandez, N. (1996) *Mol. Cell. Biol.*, **16,** 1955–1965.
24 Whitelaw, E., Hogben, P., Hanscombe, O. and Proudfoot, N. J. (1989) *Mol. Cell. Biol.*, **9,** 241–251.
25 Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.*, **162,** 156–159.
26 Yoon, J.-B., Murphy, S., Bai, L., Wang, Z. and Roeder, R. G. (1995) *Mol. Cell. Biol.*, ***15,*** 2019–2027.
27 Lobo, S. M., Lister, J., Sullivan, M. L. and Hernandez, N. (1991) *Genes Dev.*, ***5,*** 1477–1489.
28 Boyd, D. C., Turner, P. C., Watkins, N. J., Gerster, T and Murphy, S. (1995) *J. Mol. Biol.*, **253,** 677–690.
29 Waldschmidt, R., Wanadi, I. and Seifart, K. H. (1991) *EMBO J*, **8,** 2595–2603.
30 Sadowski, C. L., Henry, R. W., Lobo, S. M. and Hernandez, N. (1993) *Genes Dev.*, **7,** 1535–1548.
31 Henry, R. W., Ma, B., Sadowski, C. L., Kobayashi, R. and Hernandez, N. (1996) *EMBO J.*,**15,** 7129–7136.
32 Bai, L., Wang, Z., Yoon, J-B. and Roeder, R. G. (1996) *Mol. Cell. Biol.* **16,** 5419–5426.
33 Henry, R. W., Sadowski, C. L., Kobayashi, R. and Hernandez, N. (1995) *Nature*, **374,** 653–656.
34 Klemm, J. D. and Pabo, C. O. (1996) *Genes Dev.* **10,** 27–36.
35 Pomerantz, J. L., Kristie, T. M. and Sharp, P. A. (1992) *Genes Dev.*, **6,** 2047–2057.
36 Walker, S., Hayes, S. and O'Hare, P. (1994) *Cell*, **79,** 841–852.
37 Cleary, M. A., Stern, S., Tanaka, M. and Herr, W. (1993) *Genes Dev.*, **7,** 72–83.
38 Gstaiger, M., Knoepfel, L., Georgiev, O., Schaffner, W. and Hovens, C. M. (1995) *Nature*, **373,** 360–362.
39 Strubin, M., Newell, J. W. and Matthias, P. (1995) *Cell*, **80,** 497–506.
40 Gstaiger, M., Georgiev, O., van Leeuwen, H., van der Vliet, P. and Schaffner, W. (1996) *EMBO J.*, **15,** 2781–2790.
41 Wang, Z. and Roeder, R. G. (1995) *Proc. Natl. Acad. Sci. USA*, **92,** 7026–7030.
42 Bernues, J., Simmen, K. A., Lewis, J. D., Gunderson, S. I., Moncollin, M., Egly, J.-M. and Mattaj, I. W. (1993) *EMBO J.*, **12,** 3575–3585.
43 Sadowski, C. L., Henry, R. W., Kobayashi, R. and Hernandez, N. (1996) *Proc. Natl. Acad. Sci.* USA, **93,** 4289–4293.
44 Yoon, J.-B. and Roeder, R. G. (1996) *Mol. Cell. Biol.*, **16,** 1–9.
45 Mittal, V. and Hernandez, N. (1997) *Science*, **275,** 1136–1139.