© 1994 Oxford University Press

Nucleic Acids Research, 1994, Vol. 22, No. 6 901–911

Regulation of the *Oct-4* gene by nuclear receptors

Ian Sylvester and Hans R.Schöler*
EMBL, Gene Expression Programme, Meyerhofstrasse 1, 69012 Heidelberg, Germany

Received January 24, 1994; Revised and Accepted February 22, 1994

ABSTRACT

To unravel the network of transcription factors established during development it is important to understand how genes specifically expressed during embryogenesis are regulated. Oct-4 is a transcription factor whose expression is associated with an undifferentiated cell phenotype in the early mouse embryo and is downregulated when such cells differentiate. An enhancer in the upstream region of *Oct-4* has previously been reported as being sufficient to mediate the cell-type specific expression and RA-dependent downregulation in EC cells, although the enhancer contains no retinoic acid receptor (RAR) binding sites. Here we report the identification of promoter elements important for the regulation of the *Oct-4* gene in EC cells. A region of the proximal *Oct-4* promoter contains an overlapping set of regulatory elements including a high affinity binding site for Sp1 and three direct repeats of an AGGTCA-like sequence with either +1 or 0 spacing. Binding and transient transfection assays reveal that *Oct-4* is subject to negative regulation by different members of the steroid-thyroid hormone receptor superfamily. Specifically, important roles for ARP-1 and RAR in *Oct-4* expression are indicated.

INTRODUCTION

Oct-4 (also termed Oct-3/NF-A3) was initially detected as a member of a family of proteins which are able to bind to the well characterised octamer motif (1), a *cis*-acting regulatory sequence found in many promoters and enhancers (for refs. 2, 3). Like other Oct factors Oct-4 is a member of the POU family of transcriptional activators which contain the DNA-binding POU domain (4). The POU domain is composed of the POU specific domain unique to the POU family, and the POU homeodomain distantly related to the classical Antennapedia homeodomain (5), both of which are required for sequence specific, high affinity DNA binding (6).

Oct-4 is a maternally expressed transcription factor (TF) present in unfertilised oocytes, undifferentiated embryonic stem cells and primordial germ cells. *Oct-4* expression is associated with an undifferentiated cell phenotype and is down-regulated upon differentiation (1, 7–10). *In vitro*, *Oct-4* expression in embryonal carcinoma (EC) cells is down-regulated when the cells are induced to differentiate by treatment with retinoic acid (RA). This is probably not an indirect effect of differentiation but rather specific repression of *Oct-4* by RA (8). The mechanism of this repression was recently investigated and an important cis-acting transcriptional control element defined which was termed the retinoic acid repressible enhancer. Located between nucleotides −1132 and −889 relative to the transcriptional start site the enhancer was reported as being sufficient for the cell-type specific expression of *Oct-4* and to mediate it's repression by RA (11). The cellular response to RA is mediated at the level of transcription by ligand inducible nuclear receptors belonging to the steroid-thyroid hormone receptor superfamily (reviewed in 12–16). Receptors for RA consist of three retinoic acid receptors (RARs), RARα, β and γ and three retinoid X receptors (RXRs), RXRα, β and γ (17–26). These bind as dimeric complexes to retinoic acid response elements (RAREs) in target genes mediating both positive and negative regulation in response to RA. Interestingly, RAR and RXR bind much more efficiently to both natural and synthetic RAREs as RAR-RXR heterodimers and the ability of RXR to interact with and enhance the binding of vitamin D_3 receptor (VD_3R), thyroid hormone receptor (T_3R) and RAR to their cognate response elements has suggested a central role for RXR in multiple hormonal response pathways (25, 27–31). Natural RAREs consist of direct repeats of the half-site consensus sequence AGG/TTCA spaced by a variable number of nucleotides (32–39) although work with synthetic hormone response elements (HREs) has shown the spacing to be of importance and that AGGTCA direct repeats separated by 5bp (DR+5) may constitute an optimum RARE (40, 41). The discovery of a natural retinoid X response element (RXRE) followed by subsequent work with synthetic RAREs has shown that an AGGTCA direct repeat with a spacing of 1 nucleotide (DR+1) is preferentially bound by RXR homodimers (42, 43). RXR is selectively bound and activated by it's ligand 9 cis RA which has been reported to induce RXR/RXR homodimer formation *in vitro* (15, 44, 46).

Among the members of the steroid-thyroid receptor superfamily, indeed comprising the majority, are a class of receptors which having no identified ligand have been termed orphan receptors (47 for review). One such family of orphan receptors are the chicken ovalbumin upstream promoter transcription factors (COUP-TFs), COUP-TF1/EAR-3, COUP-TF2/ARP-1 and EAR-2 (48–50). In addition to the ovalbumin gene, COUP-TFs recognise important regulatory elements in a number of other genes (50–58). Similar to RXR, COUP-TF has

*To whom correspondence should be addressed

been shown to bind preferentially to an AGGTCA DR+1 (43) which agrees with the structure of natural COUP-TF response elements. However it was recently shown that COUP-TFs are able to interact with both direct and palindromic repeats of the GGTCA motif separated by a variable number of nucleotides and that structual adaptation of the COUP-TF dimer underlies it's promiscuous DNA-binding ability (59).

Although initially characterised as positive regulators of ovalbumin gene transcription COUP-TFs have more recently been assigned a negative role in regulating hormonal response pathways. Specifically they have been shown to be able to repress VD_3R, T_3R, RAR and RXR mediated transactivation in the presence of their respective hormones (43, 51, 59–61). COUP-TFs were shown to exist as stable dimers in solution in HeLa cells although more recent conflicting evidence has suggested they may or may not be able to heterodimerise with RXR (43, 51, 59–61). Thus, the mechanism of COUP-TF repression is unclear but may involve either competitive binding between the COUP-TF dimer and the hormone receptor or the formation of repressing heterodimeric complexes with the receptors themselves or their co-regulator RXR.

In this study we have further investigated the regulation of the *Oct-4* gene. We identify regulatory elements in the *Oct-4* proximal promoter and show that *Oct-4* is subject to negative regulation by different members of the steroid-thyroid hormone receptor superfamily.

MATERIALS AND METHODS

Cell culture and transient transfection

Murine embryonal carcinoma cells, P19 (62) and RAC65 (63), were maintained on 0.1% gelatin-coated cell culture flasks in DMEM:EMEM medium (1:1) supplemented with 7.5% fetal calf serum, non-essential amino acids L-glutamine and penicillin-streptomycin. Twenty-four hours before transfection by the calcium-phosphate coprecipitation technique, P19 and RAC65 cells were seeded at a density of 1.3×10^5 (P19), 1.8×10^5 (RA-C65) per 6 cm dish. Fresh medium was added two hours prior to transfection. Each dish received a total of 10μg DNA including 5μg *Oct-4* luciferase reporter plasmid, 1μg RSV β-gal or βActin β-gal as internal standard, the indicated amount of effector plasmid, when used, and pBluescript KS to fill up to the required 10μg. Cells were exposed to the DNA-calcium coprecipitate for sixteen to twenty hours after which time fresh medium with or without 1×10^{-6} M RA was added. After an additional twenty-four hours the cells were harvested for determination of luciferase and β-gal activities. All luciferase activities given are normalized for β-galactosidase expression except in cotransfections involving ARP-1 where the activity of the internal standard was also seen to be affected. In these cases luciferase activity is normalized for protein concentration. The normalized luciferase activities represent the average of at least two independent experiments where each transfection was performed in duplicate.

Plasmid constructions

To construct the vector Oct-luc the *Oct-4* proximal promoter was first isolated as a *Xba*I (−413)−*Bam*HI (+150) 563 bp genomic fragment. It was shortened to 461bp by digestion at the *Ban*I site at +48 just upstream of the ATG. The 5′ overhangs were filled in using Klenow enzyme, *Hin*dIII linkers were added and subsequently cleaved to yield *Hin*dIII sticky ends. The promoter fragment was cloned into the *Hin*dIII site of the pBluescript polylinker upstream of the luciferase cDNA in pBluescript KS. Regeneration of the *Oct-4 Ban*I site yielded intact *Oct-4* promoter sequences from −413 to +53. Orientation of the promoter was checked by restriction digestion and sequencing. Next, *Oct-4* genomic sequence from *Bam*HI (−1222) to *Xba*I (−413) containing the *Oct-4* enhancer was subcloned into pBluescript SK, it was re-isolated as an *Eco*RI−-*Xba*I fragment and subcloned into pBS. The *Oct-4* enhancer-containing fragment was again re-isolated this time as a *Bam*HI−*Pst*I fragment and directionally cloned into the *Bam*HI and *Pst*I sites of the pBluescript polylinker upstream of the *Oct-4* promoter. Thus, Oct-luc contains 5′ *Oct-4* genomic sequences from −1222 to −413 separated by 12 bp of polylinker sequence from the remaining −413 to +53 promoter sequences. Subsequent mutations of the *Oct-4* proximal promoter to create the vectors Sp1mut.-luc and R2mut.-luc were introduced by PCR oligonucleotide-directed mutagenesis. In Sp1mut.-luc the Sp1 site −51 to −42 is changed to GGGtttGGGC and in R2mut.-luc the sequence −38 to −33 is changed to AacTCA. All introduced mutations were verified by sequencing.

Preparation of extracts

For preparation of the P19 and RAC65 EC cell extracts, cells were grown in the absence or presence of 1×10^{-6} M RA for the indicated time points. Cells were washed and harvested in ice cold PBS and collected by centrifugation (1000 rpm, 4′, 4°C). The cell pellet (up to 4×10^7 cells) was resuspended in 80−100μl of extraction buffer (150mM NaCl, 20mM Hepes pH 7.8, 0.2mM EDTA, 0.5mM DTT, 25% glycerol) containing the following, added just prior to extraction, protease inhibitors: PMSF (0.5mM), Leupeptin (0.5μg/ml), Pepstatin (0.7μg/ml), Aprotinin (2μg/ml), Bestatin (40μg/ml). Following sonnication (10 pulses, output control 3, 30% duty cycle) the extracts were cleared by centrifugation (13000 rpm, 15′, 4°C), protein concentration determined (64) and stored at −70°C. For the transiently expressed ARP-1 in P19 cells shown in Figure 4B, 10μg of pSG-ARP-1 was transfected into P19 cells and a whole cell extract prepared as above.

Recombinant vaccinia virus expressing RARα, β, γ, RXRα and ARP-1 were constructed and amplified using standard protocols (65). For preparation of whole cell extracts HeLa cells infected with the recombinant virus were washed twice with PBS, harvested in TEN (40mM Tris pH 7.8, 10mM EDTA, 150mM NaCl) and pelleted by centrifugation (3000 rpm, 5′, 4°C). Cells were lysed in 50mM Hepes pH 7.9, 450mM NaCl, 0.5% triton X-100, and left on ice for 20′. Cell debris was removed (13000 rpm, 15′, 4°C) and following determination of protein concentration (64), clarified lysates were stored at −70°C.

Electrophoretic mobility shift assay (EMSA)

Oligonucleotides were synthesised with 5′ extensions and annealed double-stranded oligonucleotides were labelled by filling in with α^{32}P dCTP and Klenow enzyme. For the EMSAs performed with P19 and RAC65 EC cell extracts, the binding reaction consisted of 6μg whole cell extract, binding buffer (25mM Hepes pH 7.9, 0.5mM EDTA, 0.5mM DTT, 50mM NaCl, 10% glycerol), 1μg poly d(I-C) and 15000 cpm hot oligo (together with unlabelled competitor in the competition experiments), in a total volume of 15μl. For the EMSAs performed with overexpressed receptor-containing vaccinia extracts the binding reaction was: 3mg HeLa vaccinia extract, binding buffer (13.3mM Hepes pH 7.9, 0.07mM EDTA, 1mM

DTT, 33.3mM KCl, 3.3mM $MgCl_2$, 10% glycerol), 1μg poly d(I-C), 1μg/ml BSA and 15000 cpm hot oligo, again in a total volume of 15μl. Binding reactions were incubated for 20′, 4°C, followed by 10′, 37°C. For depletion of the Sp1 binding activities shown in Figure 7 preincubation of extract and unlabelled competitor was first carried out for 15′ at 4°C followed by addition of labelled oligonucleotide and incubation for a further 20′, 4°C, and 10′, 37°C. Following incubations the binding reactions were loaded onto pre-cooled, pre-run (1hr 300V), 5% polyacrylamide gels containing 0.25× TBE. The gels were run in the cold at 300V, dried and exposed (−70°C O/N).

Oligonucleotides used in this study

Bold letters indicate *Oct-4* sequences, lower case indicates introduced mutations.

OL1: 5′-AGCT**ACCCACCCAGGGGGCGGGGCCAGAGGTCAAGGCTAGAGG-GTG**-3′
OL2: 5′-CTGAACCTGGGTA**GGGGCGGGGC**TTCTAAAGTCGA-3′
OL3: 5′-CTGA**GGGCCAGAGGTCAAGGCTAGAGGGTG**GTCGA
OL4: 5′-CTGATGGGTA**GGGCCAGAGGTCA**TTCTAAAGTCGA-3′
OL5: 5′-CTGATGGGTA**aacCCAGAGGTCA**TTCTAAAGTCGA-3′
OL6: 5′-CTGATGGGTA**GGGCCAGaacTCA**TTCTAAAGTCGA-3[pri
OL7: 5′-CTGACTGGGTA**AGGTCAAGGCTA**TTCTAAAGTCGA-3′
OL8: 5′-CTGACTGGGTA**aacTCAAGGCTA**TTCTAAAGTCGA-3′
OL9: 5′-CTGACTGGGTA**AGGTCAaacCTA**TTCTAAAGTCGA-3′
OL10: 5′-CTGATGGGTA**GGGtCAGAGGTCA**TTCTAAAGTCGA-3′
OL11: 5′-CTGATGGGTAAGGTCAGAGGTCATTCTAAAGTCGA-3′
OL12: 5′-CTGA**CCACCCAGGGGGCGGGGC**TTCTAAT**CAAGGCTAGAG**TCGA-3′
β-RARE: 5′-TCGACGGGTAGGGTTCACCGAAAGTTCACTCGC-3′
ARP-1 bs: 5′-CTGACCACTGAACCCTTGACCCCTGCCCTT-3′

RESULTS

The proximal promoter of mouse *Oct-4* contains potential binding sites for nuclear receptors

In a recent study a *cis*-acting enhancer element was reported as being sufficient to confer the cell-type specific expression of *Oct-4* and also to mediate it's specific repression by RA, although this sequence contained no binding sites for RARs (11). As a further step towards locating regions important for RA-dependent down-regulation of *Oct-4* during differentiation we have investigated the promoter region of *Oct-4* for RAR elements and have found possible binding sites in the proximal promoter region (Figure 1A). R1, R2 and R3 denote three direct repeats of an AGGT-CA-like sequence, a sequence which has been shown to be the half site consensus of direct repeat response elements to VD_3R, T_3R, RAR, RXR and the orphan receptor COUP-TF, with R2 conforming perfectly to this consensus (41−43, 59). R1-R2 represents a DR+1 and as such is a potential binding site for RAR/RXR heterodimers, RXR/RXR homodimers, COUP-TF/COUP-TF homodimers and COUP-TF/RXR heterodimers (39, 42, 43). R2-R3 may also be discerned as a separate potential HRE, this time as a direct repeat with a spacing of 0 (DR+0), while R1-R3 with a spacing of seven is also a possible HRE. In addition, overlapping with the last three bp of R2 and encompassing the whole of R3 is a putative binding site for the orphan receptor ELP, differing from the reported consensus binding sequence of TCAAGGTCA in the reversed positions of the 3′ thymine and cytosine (66, 67). Like Oct-4, ELP is expressed in undifferentiated EC cells and down-regulated by RA. The involvement of ELP in the regulation of Oct-4 will be reported in detail elsewhere (I.Sylvester, Y.Yeom and H.Schöler in preparation).

As previously reported *Oct-4* is a TATA-less gene and contains multiple initiation sites. Investigation into mechanisms of TATA-less transcription has shown, in addition to the requirement of the general transcription factors, a requirement for Sp1 at promoters harbouring Sp1 binding sites (68). It has been postulated that Sp1 may recruit the general transcription factors to these TATA-less promoters (for references see 69). In this respect, the presence of a consensus decamer Sp1 binding site overlapping with the first repeat R1 in the *Oct-4* proximal promoter may be of importance. The position of the three direct repeats (R1-R3) approximately thirty bp upstream of the most 5′ of the initiation sites, where one may have expected a TATA box to be located, is also noteworthy.

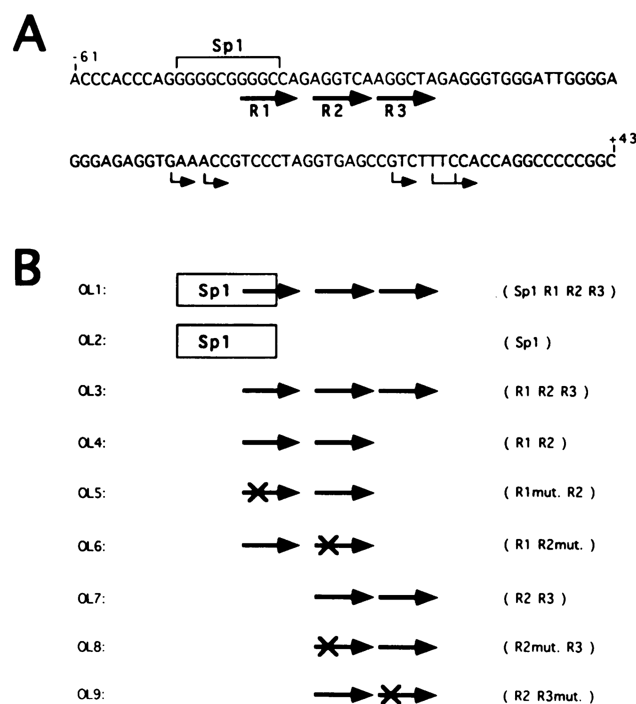

Figure 1. A: Genomic sequence of the mouse *Oct-4* proximal promoter. R1, R2 and R3 represent three direct repeats of an AGGTCA-like sequence, indicated by the arrows. A potential Sp1 binding site is highlighted. The right-angled arrows indicate the multiple transcription initiation sites of the TATA-less *Oct-4* gene (11, Y.I.Yeom unpublished data). **B**: Schematic representation of some of the oligonucleotides used as labelled probes in the EMSAs, relative to *Oct-4* proximal promoter sequences. The sequences of these oligonucleotides and those not related to the *Oct-4* promoter are given in MATERIALS AND METHODS. Numbering of nucleotides in (A) and (B) is according ref. 11.

Thus, within the proximal promoter, a region not previously shown to be important for *Oct-4* regulation, are an overlapping set of putative regulatory elements including potential DR+1 and DR+0 HREs. Subsequent *in vitro* binding and transient transfection studies were used to ascertain the contribution of these sequence elements towards the regulation of *Oct-4*.

RA-regulated complexes are formed on *Oct-4* proximal promoter sequences

As has already been mentioned *Oct-4* expression *in vitro* is specific for undifferentiated EC cells and is down-regulated by RA. Figure 2A shows an electrophoretic mobility shift assay (EMSA) using extracts from undifferentiated and RA-differentiated P19 and RAC65 EC cells. As expected Oct-4 binding activity can be detected on the octamer sequence-containing probe in extracts of undifferentiated P19 cells but is

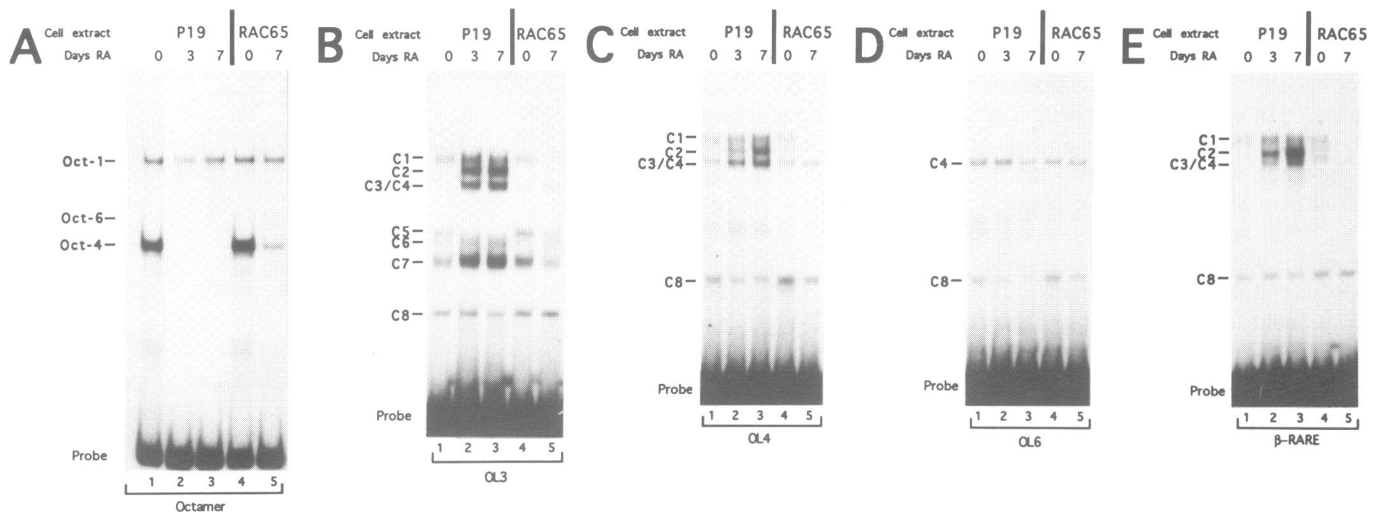

Figure 2. A comparison between the RA-repression of Oct-4 and formation of complexes on *Oct-4* proximal promoter sequences. **A**: Investigation of Oct-4 binding activity during EC cell differentiation. A radiolabelled octamer motif was incubated with untreated and RA-treated P19 (lanes 1 to 3) or RAC 65 (lanes 4 and 5) EC cell extracts. The positions of Oct-1, Oct-4 and Oct-6 are indicated. **B**: RA-independent and dependent binding activities on *Oct-4* proximal promoter sequences. The same set of five extracts used in (A) were incubated with a promoter fragment of *Oct-4* spanning the three direct repeats (OL3). This was repeated in **C** using a promoter fragment containing only the first two repeats (OL4) and in **D** with the same oligonucleotide but with a mutation in the second repeat (OL6). **E**: Binding activities induced by RA on the β-RARE. Radiolabelled β-RARE was incubated with the same set of extracts to compare the complexes with those on the *Oct-4* promoter. The complexes bound in (B) to (E) were named C1 to C8. For a schematic representation and sequences of the oligonucleotides see Figure 1B and MATERIALS AND METHODS. The cells were treated with RA for the time points indicated above the figures as described in MATERIALS AND METHODS.

absent after three days of treatment with RA (Figure 2A, compare lanes 1 and 2). This is in accordance with previous results showing the level of Oct-4 mRNA to be reduced by 90% 12 hours after adding RA to P19 EC cells and undetectable after one day (8). RAC65 cells are a mutant derivative of P19 cells which cannot be differentiated in the presence of RA (63, 70). The RA resistance of RAC65 cells is due to expression of a truncated RARα which is able to bind DNA but unable to bind ligand and acts as a dominant negative repressor of RA-responsive target genes (71, 72). Because RAC65 cells are defective in RA transactivation pathways treatment under the same conditions as P19 cells allows RA-independent aspects of *Oct-4* downregulation to be addressed. In this respect it is interesting that in RA-treated RAC65 cells Oct-4 binding activity is also decreased (Figure 2A, compare lanes 4 and 5).

The same extracts were used in an attempt to detect RA-dependent binding activities on the putative regulatory elements of the *Oct-4* proximal promoter. With the promoter probe comprising the three direct repeats eight binding activities (termed C1 to C8) could be detected (Figure 2B, see Figure 1B for a schematic representation of oligonucleotides used in the EMSAs). These eight complexes can be divided into three groups according to their response to RA: (1) unaffected, (2) induced or (3) down-regulated (compare lanes 1, 2 and 3, Figure 2B; see also C and D). Complexes C4, and C8 are unaffected by RA-treatment. In contrast to this are binding activities C2 and C3 which are clearly induced by RA-treatment of P19 EC cells. The RA-induced C3 complex runs at the same position as the uninduced C4 complex (see Figure 2D for C4). The C1, C6 and C7 binding activities are also increased by RA treatment but to a lesser extent (compare lanes 1 and 2, Figure 2B). Finally, the C5 binding complex, present in undifferentiated P19 cells, is down-regulated upon RA-treatment (compare lanes 1, 2 and 3, Figure 2B). In contrast to differentiated P19 cells, in RA treated RAC65 cells the C2 and C3 binding activities are not correspondingly induced (Figure 2B, lanes 4 and 5).

The Sp1 site in the Oct-4 promoter was used to investigate if Sp1 is regulated during EC cell differentiation. With both, P19 and RAC65 cells, only a slight decrease to about 80% of the initial Sp1 binding activity was observed after 7 days of RA treatment (data not shown).

Thus, factors present in undifferentiated cells in which *Oct-4* is expressed and factors induced by RA, and therefore present under differentiated conditions when *Oct-4* is no longer expressed, are able to bind to the putative regulatory sequences contained in the *Oct-4* proximal promoter.

The RA-induced complexes C2 and C3 bind to the first repeat R1-R2

As mentioned above the *Oct-4* proximal promoter contains a putative DR+1 HRE (R1-R2). Given that a natural RARE comprising a DR+1 element has recently been described (39) we next addressed the question of whether the *Oct-4* R1-R2 DR+1 element was important for the RA-regulated EC cell binding activities previously detected on the proximal promoter. An EMSA performed with undifferentiated and RA-differentiated EC cell extracts shows the R1-R2 sequence alone to be sufficient for binding of the P19 cell RA-induced C2 and C3 complexes (Figure 2C, compare lanes 1, 2 and 3). As before, these complexes are not correspondingly induced in RA-treated RA-C65 cells (Figure 2C, lanes 4 and 5). Interestingly, the complexes C5 to C7 did not bind to R1-R2 suggesting that the factors require sequences outside of R1-R2 (Figure 2C, lane 1). A mutation of the R2 half site, the half site which conforms perfectly to the consensus half site for HREs, prevents the RA-induced EC cell complexes from binding (Figure 2D). The EC cell C1 complex is also abolished upon mutation of the R2 half site (compare lanes 1 of Figure 2C and D).

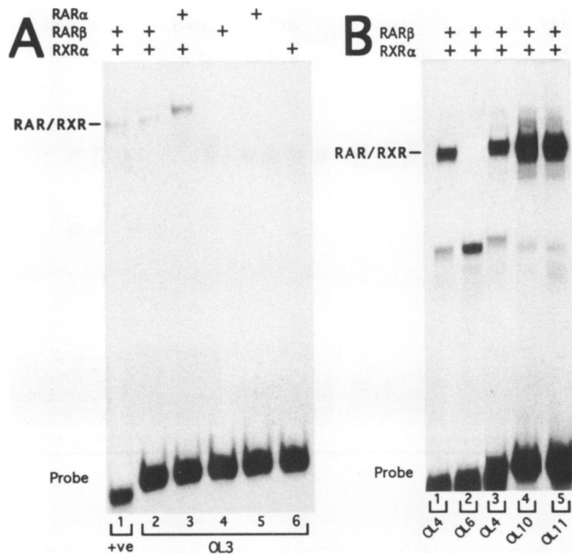

Figure 3. Investigation of RAR/RXR binding to the *Oct-4* proximal promoter. **A**: RAR/RXR heterodimers bind to *Oct-4* proximal promoter sequences. A radiolabelled promoter fragment spanning the three direct repeats (OL3) was incubated with vaccinia expressed RAR/RXR in combination or alone. RARβ/RXRα (lane 2), RARα/RXRα (lane 3), RARβ (lane 4), RARα (lane 5) and RXRα (lane 6). The position of the RAR/RXR complex is indicated by the position of RARβ/RXRα bound to the radiolabelled β-RARE (lane 1, labelled +ve). **B:** Determination of R1-R2 as being a direct repeat with R2 critical for binding. Vaccinia expressed RARβ/RXRα was incubated with radiolabelled R1-R2 (OL4, lanes 1 and 3), R1-R2mut. (OL6, lane 2), a probe where R1-R2 is made a better direct repeat (OL10, lane 4), and a perfect AGGTCA direct repeat (OL11, lane 5). See Figure 1B and MATERIALS AND METHODS for oligonucleotides.

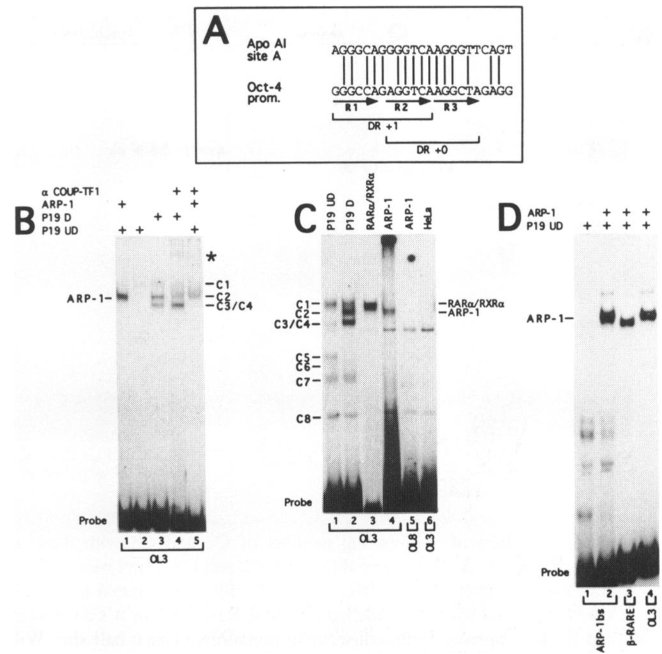

Figure 4. Investigation of the C2 complex. **A**: Sequence comparison between site A of the APO AI gene (Ladias and Karathanasis, 1991) and the three direct repeats of the *Oct-4* proximal promoter. **B**: The RA-induced C2 complex contains COUP-TFs. A radiolabelled promoter fragment containing the three direct repeats (OL3) was incubated with undifferentiated P19 EC cell extract containing transiently overexpressed ARP-1 (lane 1), undifferentiated P19 EC cell extract (lane 2), differentiated P19 cell extract (lane 3), differentiated P19 cell extract plus 1μl αCOUP-TF1 Ab (lane 4) and ARP-1-containing P19 undifferentiated extract used in lane 1 but with the addition of 1μl αCOUP-TF1 Ab (lane 5). (*) indicates 'supershifted' complex in the presence of Ab. **C**: The running position of ARP-1 but not RAR/RXR corresponds to a RA-induced complex on the *Oct-4* proximal promoter. A radiolabelled promoter fragment containing the three direct repeats (OL3) was incubated with extracts of undifferentiated (lane 1) or differentiated (lane 2) P19 cells, vaccinia expressed RARα/RXRα (lane 3) and vaccinia expressed ARP-1 (lane 4). The respective positions of the RAR/RXR and ARP-1 complexes are indicated. A radiolabelled oligonucleotide to which ARP-1 cannot bind (OL8) is incubated with vaccinia expressed ARP-1 (lane 5) to control for protein integrity relative to the precipitation effect seen in lane 4 when ARP-1 interacts with a specific sequence. Incubation with HeLa cell extract is shown in lane 6 as a control. **D:** Comparison of ARP-1 binding to three different recognition sequences. Radiolabelled consensus ARP-1 binding site from the APO AI gene is incubated with undifferentiated P19 EC cell extract (lane 1) and undifferentiated P19 EC cell extract containing transiently overexpressed ARP-1 (lane 2). Radiolabelled β-RARE is incubated with the same ARP-1 containing P19 cell extract (lane 3), as is a radiolabelled *Oct-4* promoter fragment containing the three direct repeats (OL3, lane 4). See Figure 1B and MATERIALS AND METHODS for oligonucleotides.

For comparison we have examined the RA-induction of EC cell binding activities using the well-characterized RARE (β-RARE) of the RARβ2 promoter as a probe (33, 34, 38). After treatment of P19 cells with RA for three days an upper triplet of binding activities are seen, the lower two of which have clearly been induced (Figure 2E, compare lanes 1 and 2). The middle complex of this upper triplet continues to be induced upon further RA-treatment (Figure 2E, compare lanes 2 and 3). It has been reported that RA treatment of P19 cells leads to an increase in the expression of RARα transcripts and an induction of RARβ transcripts (71, 73). RARβ transcripts are induced by transactivation through the autoregulatory β-RARE by RARα and RARβ in the presence of RA (33, 38). Thus, the RA-induced complexes formed on the β-RARE, shown in Figure 2E, might contain RARα/RXR and RARβ/RXR heterodimers. Although these complexes have been given the same nomenclature as those formed on the R1-R2 DR+1 of the *Oct-4* proximal promoter we cannot say with certainty that the complexes are the same in both cases. However, there is a similarity in the binding patterns of both uninduced and RA-induced complexes formed on the β-RARE and the R1-R2 sequence, suggesting RAR/RXR heterodimers may be present in the C2 and C3 RA-induced complexes formed on the *Oct-4* proximal promoter. Therefore, the R1-R2 DR+1 in the *Oct-4* proximal promoter is sufficient for binding of the C2 and C3 RA-induced EC cell binding activities, the binding of which is dependent on the integrity of the R2 half site.

RARs bind to the first repeat R1-R2 *in vitro*

Given the possibility that C2 or C3 contain endogenous RARs we next investigated whether RARs would bind to this region. RARs were overexpressed using recombinant vaccinia virus in HeLa cells and extracts used in EMSAs. All three RARs α, β and γ, in combination with RXR, bind to a probe spanning the three direct repeats R1 to R3 with equal affinity and competition shows the binding to be specific (Figure 3A, lanes 2 and 3 and data not shown). RAR/RXR heterodimerization is required for binding and neither RAR nor RXR alone will bind to this region of the proximal promoter (Figure 3A, lanes 4 to 6). This is in agreement with reported data which has shown RAR/RXR heterodimerization is required for efficient binding to both natural

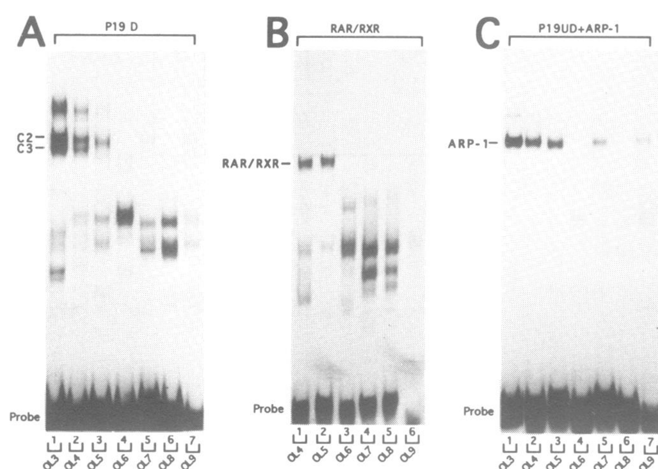

Figure 5. Comparison of the binding profiles of C2 and C3 with those of RAR/RXR and ARP-1. **A**: Binding profile of the C2 and C3 complexes on *Oct-4* proximal promoter sequences. An extract of P19 cells differentiated for 3 days with RA was incubated with radiolabelled complete R1-R2-R3 or it's constituent R1-R2 and R2-R3 sequences with consecutive mutations in each half site. Wild type R1-R2-R3 (OL3, lane 1), R1-R2 (OL4, lane 2), R1mut.-R2 (OL5, lane 3), R1-R2mut. (OL6, lane 4), wild type R2-R3 (OL7, lane 5), R2mut.-R3 (OL8, lane 6) and R2-R3mut. (OL9, lane 7). **B**: Binding profile of RAR/RXR on *Oct-4* proximal promoter sequences. Vaccinia expressed RARβ/RXRα was incubated with radiolabelled R1-R2 or R2-R3 sequences with consecutive mutations in each half site. Wild type R1-R2 (OL4, lane 1), R1mut.-R2 (OL5, lane 2), R1-R2mut. (OL6, lane 3), wild type R2-R3 (OL7, lane 4), R2mut.-R3 (OL8, lane 5) and R2-R3mut. (OL9, lane 6). **C**: Binding profile of ARP-1 on *Oct-4* proximal promoter sequences. An undifferentiated P19 EC cell extract containing transiently overexpressed ARP-1 was incubated with radiolabelled R1-R2-R3, R1-R2 or R2-R3 sequences with consecutive mutations in each half site. Wild type R1-R2-R3 (OL3, lane 1), R1-R2 (OL4, lane 2), R1mut.-R2 (OL5, lane 3), R1-R2mut. (OL6, lane 4), wild type R2-R3 (OL7, lane 5), R2mut.-R3 (OL8, lane 6) and R2-R3mut. (OL9, lane 7). See Figure 1B and MATERIALS AND METHODS for oligonucleotides.

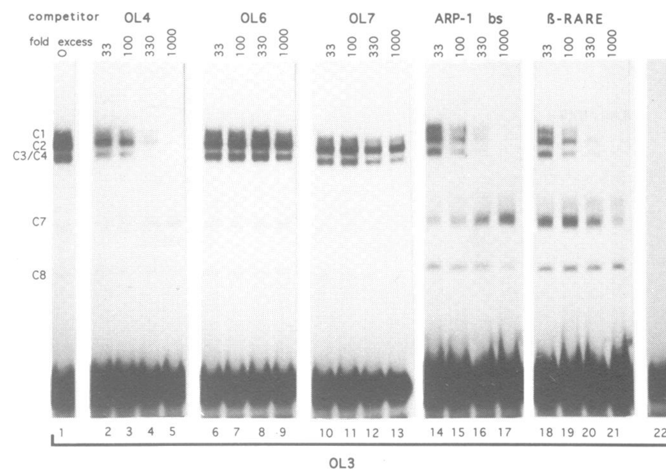

Figure 6. Investigation of the specificity of complex formation on Oct-4 proximal promoter sequences by competition EMSAs. An extract of P19 EC cells differentiated for three days with RA was incubated with a radiolabelled promoter fragment containing the three direct repeats (OL3) in the presence of increasing amounts of unlabelled competing oligonucleotides. The unlabelled competitors were: Wild type R1-R2 (OL4, lanes 2 to 5), R1-R2mut. (OL6, lanes 6 to 9), R2-R3 (OL7, lanes 10 to 13), the consensus ARP-1 binding site from the APO AI gene (ARP-1 bs, lanes 14 to 17) and the β-RARE (lanes 18 to 21). The fold excess of each competitor is shown above the lanes. Free probe alone is shown in lane 22. See Figure 1B and MATERIALS AND METHODS for oligonucleotides.

and synthetic RAREs (25, 27–31). The position of an RAR/RXR complex bound to the β-RARE is indicated (Figure 3A, lane 1).

To locate the RAR/RXR binding site in the *Oct-4* promoter the same oligonucleotides as for the detection of C2 and C3 in the previous experiment were used (Figure 3B, lanes 1 and 2). Similar to C2 and C3, RAR/RXR was found to bind to R1-R2 and could be prevented from binding by the same mutation in the R2 half site. In the same experiment a single bp alteration in the R1 sequence (GGGCCA to GGGtCA), which in the context of R1-R2 improves it's sequence requirements as a direct repeat, improves binding to an extent similar to the binding of RAR/RXR to a perfect DR+1 (AGGTCA(1)AGGTCA) (Figure 3B, compare lanes 3, 4 and 5). However, a mutation which improves R1-R2 sequence requirements as an inverted palindrome has no effect on receptor binding (data not shown) suggesting the nature of R1-R2 to indeed be a direct, although imperfect, repeat. Therefore, RARs are able to bind, in combination with RXR, to this region of the *Oct-4* proximal promoter and as with the RA-induced C2 and C3 complexes, the R2 half site is critical for binding.

One of the RA-induced EC cell complexes contains COUP-TFs

We have shown binding of RAR/RXR heterodimers to the R1-R2 DR+1 sequence *in vitro*. As mentioned, a natural RARE comprising a DR+1 element has recently been described (39). However, DR+1 elements have also been proposed to be selective binding sites for RXR alone and COUP-TFs, and many natural COUP-TF response elements have a DR+1 structure (42, 43, 59). A comparison between site A of the APO AI gene bound by the orphan receptor COUP-TF2 (ARP-1) (50) and the three direct repeats (R1 to R3) in the *Oct-4* proximal promoter shows a high degree of conservarion (Figure 4A). Although the similarity is strongest over the DR+0 of R2-R3 it clearly extends to R1. We were therefore prompted to investigate if members of the COUP-TF family, specifically ARP-1 (COUP-TF2), might also play a role in the regulation of *Oct-4*.

ARP-1 was transiently overexpressed in undifferentiated P19 EC cells and an extract used in EMSAs. ARP-1 was found to bind to the R1-R3 sequence of the *Oct-4* proximal promoter and to run at the same position as the RA-induced C2 complex in differentiated P19 cells (Figure 4B, compare lanes 1 and 3). Moreover, inclusion of a polyclonal αCOUP-TF1 Ab (kindly provided by Dr. M-J. Tsai) in the binding reaction with the differentiated P19 cell extract specifically decreases the intensity of the RA-induced C2 complex and leads to the appearance of a 'supershifted' complex (Figure 4B, compare lanes 3 and 4). αCOUP-TF1 Ab was used due to the high degree of homology between COUP-TF1 and COUP-TF2 (ARP-1). The overall amino acid idenity between COUP-TF1 and COUP-TF2 is 87% with the DNA and putative ligand-binding domains being 98 and 97% conserved respectively (74). The αCOUP-TF1 Ab does not completely remove the C2 binding activity although it decreases it to the same extent as it does the transiently overexpressed ARP-1 (COUP-TF2) in undifferentiated cells, with which it is cross-reacting, suggesting that C2 represents ARP-1 and not COUP-TF1 (Figure 4B, compare lanes 3 and 4 with lanes 1 and 5).

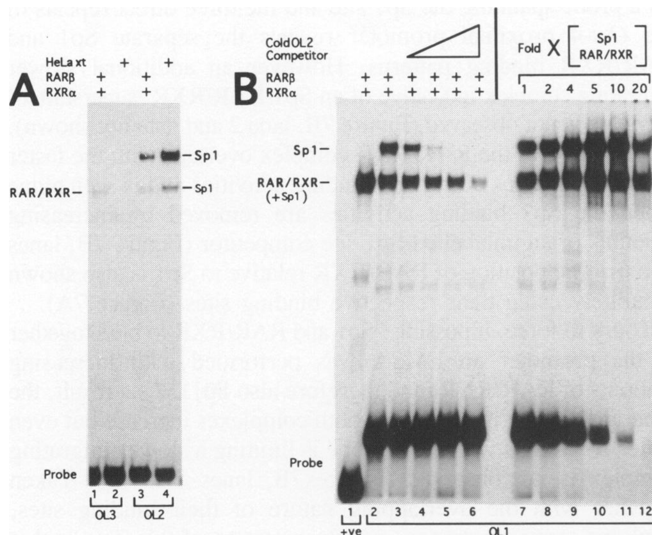

Figure 7. Non-cooperativity between Sp1 and RAR/RXR on the *Oct-4* proximal promoter. **A**: Comparison of RAR/RXR and Sp1 running positions. Vaccinia expressed RARβ/RXRα was incubated with a radiolabelled promoter fragment containing only the three direct repeats (OL3, lane 1) or one containing only the Sp1 site (OL2, lane 3). A comparison between lanes 1 and 3 shows the presence of Sp1 in the RAR/RXR extract. Wild type HeLa extract (lanes 2 and 4) is included as a control. **B**: Titration of Sp1 and RAR/RXR at the *Oct-4* proximal promoter. Vaccinia expressed RARβ/RXRα was incubated with a radiolabelled promoter fragment spanning both the Sp1 site and the three direct repeats (OL1, lanes 2 to 12). To compete for Sp1 binding, increasing amounts, compared to the labelled probe, of unlabelled Sp1 site competitor (OL2) were additionally incubated in lanes 3 to 6. The molar ratios were: 10× (lane 3), 100× (lane 4), 1000× (lane 5) and 10,000× (lane 6). In an attempt to observe mutual binding between Sp1 and RAR/RXR increasing amounts of extract containing RAR/RXR and Sp1 were incubated (lanes 7 to 12). Even under limiting probe conditions (lanes 11 and 12), no slower migrating complexes were formed. The increase of extract used is given above each lane. In lane 1, RARβ/RXRα was incubated with radiolabelled β-RARE to show the position of the RAR/RXR complex. See Figure 1B and MATERIALS AND METHODS for oligonucleotides.

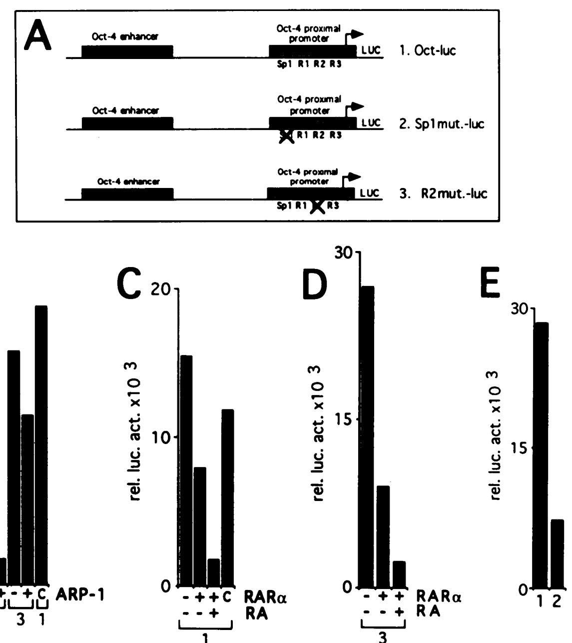

Figure 8. Negative regulation of the *Oct-4* promoter by transiently expressed nuclear receptors. **A**: Schematic representation of wild type and mutated *Oct-4* reporter constructs. The crosses indicate which element had been mutated. The numbers and the construct names correspond to the numbers below (B) to (E). **B**: Repression of transient *Oct-4* expression by ARP-1. 5μg of wild type Oct-luc and R2mut.-luc reporter constructs were transiently transfected into P19 EC cells in the absence (−) or presence (+) of 0.1μg coexpressed pSG-ARP-1. 'C' indicates coexpression of 0.1μg parental pSG vector. **C**: Ligand dependent repression of *Oct-4* transient expression by RARα. 5μg of wild type Oct-luc reporter construct was transiently transfected into RAC 65 EC cells in the absence (−) or presence (+) of 1μg coexpressed pSG-RARα. Absence or presence of ligand (RA) is indicated in the same way. 'C' indicates coexpression of 1μg parental pSG vector. **D**: Repression of the *Oct-4* proximal promoter with a mutated R2 repeat by RARα in RAC 65 cells. 5μg of the R2mut.-luc reporter construct was transiently transfected into RAC 65 EC cells under the same conditions as in (C). **E**: Mutation of the proximal promoter Sp1 site decreases *Oct-4* promoter activity. 5μg of wild type Oct-luc and Sp1mut.-luc reporter constructs were transiently transfected into P19 EC cells. All activities in (B) to (E) are presented in bar graph form. The conditions for the transient transfections are given in MATERIALS AND METHODS.

Given also the possibility of the presence of RAR/RXR heterodimers in the RA-induced EC cell complexes on the *Oct-4* proximal promoter (see above) an EMSA was performed using vaccinia expressed ARP-1 and RAR/RXR to directly compare their electrophoretic mobilities when bound to the R1-R3 sequence. Again ARP-1 runs at the position of the RA-induced C2 complex and thus this result correlates with the Ab result described above, indicating that the C2 complex may represent ARP-1 (Figure 4C, compare lanes 2 and 4). However, the RAR/RXR complex runs between C1 and C2 and thus at a different position to both the C2 and C3 RA-induced complexes (Figure 4C, compare lanes 2, 3 and 4).

Since the mobility of the most prominent RA-induced complex on the β-RARE (Figure 2E) is identical to that of C2 on R1-R2 (data not shown) we considered it to be possible that ARP-1 also binds to the β-RARE. To test this we compared the binding of ARP-1 to it's binding site in the ApoA1 gene (oligo A), R1-R3 and β-RARE (Figure 4D). As expected, ARP-1 complexes were detected with each oligonucleotide, showing that ARP-1 also binds to the β-RARE (lanes 3 to 5). Therefore, the RA-induced complexes detected on the β-RARE in Figure 2E may represent members of the COUP-TF family and endogenous RAR/RXR complexes are undetected in the EMSA even on a natural RARE.

Thus, although RARs are able to bind to the R1-R2 DR+1 *in vitro* they seem to have a different mobility compared to the RA-induced complexes. However, endogenous RARs may not be induced to the same extent or with the same kinetics as the detected complexes and therefore may have been undetected on *Oct-4* promoter sequences, indeed they seem to be undetectable by EMSA on even the β-RARE. Alternatively, artifacts of the vaccinia expression system may be responsible for the different mobilities. On the other hand, the binding and Ab data suggest strongly that one of the RA-induced complexes contains COUP-TFs. Specifically, a direct role for ARP-1 (COUP-TF2) in the regulation of *Oct-4* is possible.

RAR/RXR and ARP-1 show similar binding profiles to C2 on *Oct-4* proximal promoter sequences

As mentioned, the different mobility of RAR/RXR in the EMSA to C1, C2 or C3 might be due to modifications caused by the vaccinia expression system. To further address the identity of

the RA-induced C2 and C3 complexes, a direct binding profile comparison on *Oct-4* promoter sequences was carried out between RA-treated P19 cell extract, RAR/RXR and ARP-1. The three direct repeats were dissected with R1-R2 and R2-R3 sequences being used separately in EMSAs. Mutations were also introduced to assess the half site contribution of each direct repeat.

The RA-induced C2 and C3 complexes were found to bind more strongly to an oligonucleotide spanning all three direct repeats than to either R1-R2 or R2-R3 alone with R2-R3 constituting the weaker binding site (Figure 5A, compare lanes 1, 2 and 5). Mutation of the first three bases of R1 from GGG to AAC has no effect on C2 binding to R1-R2 whereas C3 binding is decreased (compare lanes 2 and 3). However the same introduced mutation at R2 completely abolishes binding of both C2 and C3 (lanes 4 and 6). A slight decrease of C2 and C3 binding to R2-R3 was observed when the mutation was introduced into R3 (compare lanes 5 and 7). Thus, C2 and C3 show a similar binding profile on the oligonucleotides used with the exception of the R1 mutation which selectively affects C3 binding to R1-R2.

Similarly to C2 and C3 RAR/RXR was found to bind efficiently to R1-R2 and to be dependent on the R2 half site and as with C2, but not C3, mutation of R1 had no effect on binding (Figure 5B, lanes 1, 2 and 3). RAR/RXR was found to bind weakly to R2-R3 compared to R1-R2 (Figure 5B, compare lanes 1 and 4). As with RAR/RXR and C2, ARP-1 was found to bind efficiently to R1-R2 and to be dependent on the R2, but not R1, half site for binding (Figure 5C, lanes 2, 3 and 4). However, ARP-1 was able to bind much more efficiently to R2-R3 than had RAR/RXR (lane 5). Interestingly, as with the RA-induced complexes, the binding of ARP-1 to the complete three direct repeats is stronger than to the separate R1-R2 or R2-R3 sequences (lanes 1, 2 and 5). Thus this comparison shows clearly a similarity in binding profile between the RA-induced C2 complex, RAR/RXR and ARP-1 with the possible exception that RAR/RXR binds more weakly to R2-R3 than either C2 or ARP-1. It is therefore possible that endogenous RARs do bind to the Oct4 proximal promoter and that either they were not detected amongst the induced complexes or that modifications due to the vaccinia expression system cause them to run at a different position to the induced complexes as shown in Figure 4C.

Finally, the specificity of the RA-induced complexes was additionally confirmed by competition EMSAs (Figure 6). A probe spanning the three direct repeats was incubated with RA-treated P19 cell extract in the presence of increasing amounts of unlabelled R1-R2, R1-R2mut., R2-R3, consensus ARP-1 binding site or β-RARE sequences. The RA-induced C2 and C3 complexes are efficiently competed by R1-R2 (lanes 2 to 5), Arp-1 bs (lanes 14 to 17) and β-RARE (lanes 18 to 21) sequences. As expected, competition is less efficient with R2-R3 sequences (lanes 10 to 13) and absent with R1-R2mut. sequences (lanes 6 to 9). The competition data are thus in full agreement with the binding data presented in Figures 2 and 5.

Non-cooperativity on the *Oct-4* proximal promoter

As outlined above *Oct-4* is a TATA-less gene containing an Sp1 site in the proximal promoter region. The Sp1 binding site in the *Oct-4* proximal promoter overlaps with the R1 half site suggesting that *in vitro* binding of Sp1 and an RAR/RXR complex at R1-R2 may be mutually exclusive. The vaccinia expressed RAR and RXR used in this study were made in HeLa cells and since the receptors were not purified the extracts prepared contain Sp1 (Figure 7A, lane 3). An EMSA performed with RAR/RXR on a probe spanning the Sp1 site and the three direct repeats of the *Oct-4* proximal promoter reveals the separate Sp1 and RAR/RXR binding patterns. However an additional slower migrating complex indicative of an Sp1/RAR/RXR 'supershifted' complex is not observed (Figure 7B, lane 2 and data not shown). The position of the RAR/RXR complex overlaps with the faster migrating species of the Sp1 binding activities and is seen alone when the Sp1 binding activities are removed by increasing amounts of an unlabelled Sp1 site competitor (Figure 7B, lanes 3 to 6). The position of RAR/RXR relative to Sp1 is also shown separately using their respective binding sites (Figure 7A).

To try to force, if possible, Sp1 and RAR/RXR to bind together to the promoter an EMSA was performed with increasing amounts of RAR/RXR (and therefore also Sp1). As a result, the separate binding intensities of both complexes increase but even under conditions where the probe is limiting a slower migrating complex is not observed (Figure 7B, lanes 7 to 12). Taken together with the overlapping nature of their binding sites, implying steric hindrence, an interpretation of this data is that Sp1 and RAR/RXR binding *in vitro* to the *Oct-4* proximal promoter is non-cooperative.

Negative regulation of *Oct-4* transient expression by members of the nuclear receptor superfamily

The binding studies have suggested that members of the COUP-TF and RAR family may be involved directly in the RA-dependent down-regulation of *Oct-4*. The possible involvement of ARP-1 and RAR in the regulation of the *Oct-4* gene was further addressed in transient transfection studies. Transient transfection into P19 cells reveals repression of the wild type *Oct-4* reporter construct in the additional presence of co-expressed ARP-1 but not control vector DNA (Figure 8B, see also Figure 8A for a schematic representation of reporter constructs). Repression by ARP-1 is highly effective since an approximate four fold repression is seen with as little as 10ng ARP-1 DNA (data not shown) rising to a maximum eleven fold repression at 100ng. Mutation of the R2 half site, shown in the binding studies to prevent RA-induced C2 and C3 complex formation, relieves the repression effect (Figure 8B). The upstream enhancer had been included in these constructs because the proximal promoter by itself was less active and thus it was difficult to evaluate repression by ARP-1 (data not shown).

Transient co-transfection studies in RAC65 cells, known to contain low amounts of endogenous RARs (71), showed ligand-independent repression of the wild type *Oct-4* reporter by all three receptors α , β and γ. The effect was more marked for RARβ and RARγ and was dose dependent (data not shown). Ligand-dependent repression of the *Oct-4* reporter was only seen with co-transfection of RARα, and a representative experiment is shown in Figure 8C. However, repression by RARα in the presence of RA is still observed when the R2 half site is mutated, the same mutation which prevents RARs from binding to the R1-R2 sequence *in vitro* (Figure 8D).

The transient expression studies indicate that RAR/RXR and ARP-1 act as negative regulators of the *Oct-4* gene and especially in the case of ARP-1 this interaction appears to be mediated via *Oct-4* promoter sequences. If steric hindrance of Sp1 by receptors plays a role in *Oct-4* regulation, as might be suggested by the non-cooperativity of binding between Sp1 and a complex at R1-R2, mutating the Sp1 site in the promoter may have a negative effect on *Oct-4* expression. The importance of Sp1 for *Oct-4* promoter activity was tested by introduction of a mutation into

the Sp1 site known to abolish *in vitro* binding of Sp1 (data not shown). As a result, transient Oct-4 expression is decreased despite the presence of the upstream enhancer (Figure 8E). This demonstrates the importance of Sp1 bound at the proximal promoter to *Oct-4* expression *in vitro* and that it is an ideal target for negative regulation by nuclear receptors.

DISCUSSION

Oct-4 is the earliest known gene in the mammalian embryo to be differentially regulated. It's expression is associated with the undifferentiated cells of the early mouse embryo, namely with those cells which give rise to the embryonic ectoderm, and is strictly down-regulated upon differentiation. Assuming that Oct-4 plays a developmental role during these stages, the understanding of how the *Oct-4* gene itself is regulated is of importance. Because the genetics of the early mouse is only poorly established we are using EC cells, believed to mimick the stem cells of the early embryo, as a model system to study the regulation of the Oct-4 gene *in vitro*.

Investigation of the Oct-4 proximal promoter revealed an overlapping set of putative regulatory elements. A high affinity decamer consensus Sp1 binding sequence overlapping with three direct repeats of an AGGTCA-like sequence containing two potential HREs, a DR+1 (R1-R2) and a DR+0 (R2-R3). Using extracts prepared from both undifferentiated and RA-differetiated EC cells we could detect multiple binding activities on the Oct-4 proximal promoter. Most interestingly, two complexes (C2 and C3) were strongly induced upon RA treatment and are therefore candidates for negative regulators of the Oct-4 gene in response to RA. The C2 and C3 complexes were found to bind to both potential HREs, although more efficiently to the DR+1 of R1-R2, and to be absolutely dependent on the integrity of the R2 half site.

Given the recent discovery of a natural DR+1 RARE, RARs were considered candidates for the RA-induced complexes. Binding studies revealed RAR/RXR heterodimers to be able to interact strongly with the R1-R2 sequence and again this binding was dependent on the R2 half site. However, the mobility in an EMSA of an RAR/RXR complex on the Oct-4 promoter did not seem to correlate with any of the RA-induced complexes suggesting endogenous receptors may not be components of the RA-induced complexes, although several reasons make this unlikely. Within one hour of treating P19 cells with RA the levels of the already present RARα are increased while expression of RARβ is induced (71, 73). The levels of RARβ transcripts reach a maximum at two days after which time they substantially decrease. The RA-induced complexes are detected after three days of RA treatment at which time endogenous levels of RAR/RXR may not be high enough to be detected by EMSA. Alternatively, modifications of the RARs overexpressed using the vaccinia system could have lead to the discrepency in mobilities. Finally, a comparison between the binding requirements of the RA-induced complexes and RAR/RXR on Oct-4 promoter sequences showed a strong correlation between C2 and RAR/RXR.

Other potential candidates for components of the RA-induced complexes, and therefore as negative regulators of Oct-4 expression, are members of the COUP-TF family of orphan receptors. Many natural COUP-TF response elements have a DR+1 structure and, in addition to the presence of a DR+1 element in the proximal promoter of *Oct-4*, the three direct repeats show homology to the binding site of ARP-1 (COUP-TF2) within the APO AI gene (50). Overexpressed ARP-1 in undifferentiated P19 cells runs at the same position in an EMSA as the C2 complex in RA-differentiated cells, and a direct comparison revealed them to have the same binding requirements for *Oct-4* promoter sequences. Moreover, using Abs, COUP-TFs were shown to be present in the RA-induced C2 binding activity, suggesting a direct role for COUP-TFs, specifically ARP-1, in the RA-dependent down-regulation of *Oct-4*.

Thus, the binding studies revealed putative candidate factors involved in negative regulation of the *Oct-4* gene. Subsequent transient transfections in EC cells using *Oct-4* reporter constructs containing the identified proximal promoter elements were used to address the functional relevence of the candidate factors in the regulation of *Oct-4*. Co-expressed ARP-1 was shown to negatively regulate transient *Oct-4* expression and in agreement with the binding studies this was dependent on an intact R2 half site. *Oct-4* was also shown to be repressed in a ligand-dependent manner by RARα although the same mutation in the R2 half site which abolishes *in vitro* binding of the RARs to the *Oct-4* promoter still allowed this repression. This may indicate that the repression of *Oct-4* by RARα, in contradiction to the binding data, is more indirect. Expression of RARα in RAC65 cells has been shown to compensate for the altered endogenous RARα thus sensitizing the cells to the differentiating-inducing effects of RA (72, 73). It is therefore possible that transient transfection of RARα into RAC65 cells in the presence of RA shown in this study represses *Oct-4* because normally inactive RA-response pathways are activated. This could result for example in the turn-down of positive regulatory factor(s) also subject to regulation by RA. It is interesting that one of the complexes formed on the *Oct-4* promoter was specific for undifferentiated cells and was down-regulated by RA. Thus, the turn-down of *Oct-4* resulting from introducing RARα in the presence of RA into RAC65 cells may reflect the different pathways which all act normally in EC cells to facilitate the rapid turn-down of *Oct-4* expression in response to RA.

However, a more direct role for the RARs, as suggested by the binding data, is also possible. RARα may be able to bind and repress directly *Oct-4* expression at a mutated element due to stabilizing protein:protein interactions which are absent in a gel retardation assay. Such an effect has been shown for the human interferon-β promoter where HMG I(Y) increases the affinity of NF-κB and ATF-2 for their respective binding sites (75). Alternatively, the DR+7 of R1-R3 may be utilized in the absence of R2, or that there may be a redundancy of RAR/RXR binding sites in the upstream *Oct-4* sequences. The initial signal for *Oct-4* turn-down may not be mediated by the COUP-TFs but rather they might act subsequently. This initial signal could be provided by the already present RARα bound with RXR to the DR+1 of R1-R2. Thus, it is possible that the activity of the *Oct-4* promoter could be decreased quickly without *de novo* protein synthesis. The presence of an RAR/RXR complex at R1-R2 might antagonize the interaction of Sp1 with it's binding site which we have shown to be a prerequisite for promoter activity. In this respect, the observed non-cooperativity of binding between Sp1 and RAR/RXR on the *Oct-4* proximal promoter *in vitro*, which may suggest mutual exclusivity, is of importance. The binding of RAR/RXR need not be one of high affinity, only that it antagonizes Sp1 binding. A more potent repressor induced with slower kinetics might then ensure more stable repression of *Oct-4*. ARP-1 is a candidate for such a late repressor of *Oct-4* expression.

Of course this does not take into account the additional points by which *Oct-4* is repressed that lie outside of the RA-transactivation pathway. For example as shown by the decrease in Oct-4 binding activity in the RA-resistant RAC65 cells upon treatment with RA and yet non-induction of the C2 and C3 complexes as shown in Figure 2. Finally, the last stage in repression of the *Oct-4* gene, ensuring complete transcriptional silence in differentiated cells, may be brought about by changes in chromatin, such as methylation of *Oct-4* upstream sequences as has recently been shown (76). In conclusion, studies by others and ourselves have led to the identification of regulatory elements important for the expression of the *Oct-4* gene *in vitro*. The subsequent use of *in vivo* techniques, for example the generation of transgenic mice, should demonstrate the contribution of the defined regulatory sequences to the temporal and spatial expression pattern of *Oct-4* during murine development.

Note added

While this work was under review two different groups of investigators have also reported the exsistence of the same negatively acting RARE in the Oct-4 proximal promoter shown in this study (77, 78). In addition to repression by ARP-1 (COUP-TF2), the other members of the COUP-TF family COUP-TF1/EAR-3 and EAR-2 were also shown to be able to repress transient Oct-4 expression and that this was dependent on an intact R2 half site (78), although, in contradiction to the results presented here, a repression of Oct-4 by RARs could not be demonstrated.

ACKNOWLEDGEMENTS

We are very grateful to Henk Stunnenberg for his continued help and support and to Jon Schoorlemmer for discussing the possibility of COUP-TF involvement in *Oct-4* regulation. We thank Dr. M-J. Tsai for the kind gift of the polyclonal αCOUP-TF1 antibody. Thanks to Domingo Barettino for the pSG-ARP-1 expression vector, Thomas H. Bugge for initial preparations of vaccinia extracts and Susan Weston, Samantha O'Loughlin and Viviane Adam for oligonucleotide synthesis. Finally we thank members of the gene expression programme for helpful discussions and critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (DFG grant Scho 340/2-1).

REFERENCES

1. Schöler,H.R., Hatzopoulos,A.K., Balling,R., Suzuki,N. and Gruss,P. (1989) *EMBO J.*, **8**, 2543–2550.
2. Hatzopoulos,A.K., Schlokat,U. and Gruss,P. (1988) in Hames,B.D. and Glover,D.M. (eds), Transcription and splicing, frontiers in molecular biology. IRL Press, Oxford, pp. 43–96.
3. Verrijzer,C.P. and van der Vliet,P.C. (1993) *Biochim. et. Biophys. Acta*, **1173**, 1–21.
4. 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. and Horvitz,H.R. (1988) *Genes Dev.*, **2**, 1513–1516.
5. Scott,M.P., Tamkun,J.W. and Hartzell,G.W. (1989) *Biochim. Biophys. Acta.*, **989**, 25–48.
6. Schöler,H.R. (1991) *Trends Genet.* **7**, 323–328.
7. Schöler,H.R., Balling,R., Hatzopoulos,A.K., Suzuki,N and Gruss,P (1989) *EMBO J.*, **8**, 2551–2557.
8. Okamoto,K., Okazawa,H., Okuda,A., Sakai,M., Muramatsu,M. and Hamada,H. (1990) *Cell*, **60**, 461–472.
9. Rosner,M.H., Vigano,M.A., Ozato,K., Timmons,P.M., Poirier,F., Rigby,P.W.J. and Staudt,L.M. (1990) *Nature*, **345**, 686–692.
10. Schöler,H.R., Dressler,G.R., Balling,R., Rhodewohld,H. and Gruss,P. (1990) *EMBO J.*, **9**, 2185–2195.
11. Okazawa,H., Okamoto,K., Ishino,F., Ishino-Kaneko,T., Takeda,S., Toyoda,Y., Muramatsu,M. and Hamada,H. (1991) *EMBO J.*, **10**, 2997–3005.
12. Evans,R.M. (1988) *Science*, **240**, 889–895.
13. Green,S. and Chambon,P. (1988) *Trends Genet.*, **4**, 309–314.
14. Beato,M. (1989) *Cell*, **56**, 335–344.
15. Leid,M., Kastner,P. and Chambon,P. (1992) *Trends Biochem. Sci.*, **17**, 427–433.
16. Stunnenberg,H.G. (1993) *BioEssays*, **15**, 309–315.
17. Giguere,V., Ong,E.S., Segui,P. and Evans,R.M. (1987) *Nature*, **330**, 624–629.
18. Petkovich,M., Brand,N.J., Krust,A. and Chambon,P. (1987) *Nature*, **330**, 444–450.
19. Benbrook,D., Lernhardt,E. and Pfahl,M. (1988) *Nature*, **333**, 669–672.
20. Brand,N., Petkovich,M., Krust,A., Chambon,P., de Thé,H., Marchio,A., Tiollais,P. and Dejean,A. (1988) *Nature*, **332**, 850–853.
21. Hamada,K., Gleason,S.L., Levi,B.Z., Hirschfeld,S., Apella,E. and Ozato,K. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 8289–8293.
22. Krust,A., Kastner,P., Petkovich,M., Zelent,A. and Chambon,P. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5310–5314.
23. Zelent,A., Krust,A., Petkovich,M., Kastner,P. and Chambon,P. (1989) *Nature*, **339**, 714–717.
24. Mangelsdorf,D.J., Ong,E.S., Dyck,J.A. and Evans,R.M. (1990) *Nature*, **345**, 224–229.
25. Leid,M., Kastner,P., Lyons,R., Nakshatri,H., Saunders,M., Zacharewski,T., Chen,J.Y., Staub,A., Garnier,J.M., Mader,S. and Chambon,P. (1992) *Cell*, **68**, 377–395.
26. Mangelsdorf,D.J., Borgmeyer,U., Heyman,R.A., Zhou,J.Y., Ong,E.S., Oro,A.E., Kakizuka,A. and Evans,R.M. (1992) *Genes Dev.*, **6**, 329–344.
27. Yu,V.C., Delsert,C., Andersen,B., Holloway,J.M., Devary,O., Näär,A.M., Kim,S.Y., Boutin,J.M., Glass,C.K. and Rosenfeld,M.G. (1991) *Cell*, **67**, 1251–1266.
28. Bugge,T.H., Pohl,J., Lonnoy,O. and Stunnenberg,H.G. (1992) *EMBO J.*, **11**, 1409–1418.
29. Kliewer,S.A., Umesono,K., Mangelsdorf,D.J. and Evans,R.M. (1992) *Nature*, **355**, 446–449.
30. Marks,M.S., Hallenbeck,P.L., Nagata,T., Segars,J.H., Apella,E., Nikodem,V.M and Ozato,K. (1992) *EMBO J.*, **11**, 1419–1435.
31. Zhang,X.K., Hoffman,B., Tran,P.B.V., Graupner,G. and Pfahl,M. (1992) *Nature*, **355**, 441–446.
32. Vasios,G.W., Gold,J.D., Petkovich,M., Chambon,P. and Gudas,L.J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9099–9103.
33. de Thé,H., Vivanco-Ruiz,M., Tiollais,P., Stunnenberg,H. and Dejean,A. (1990) *Nature*, **343**, 177–180.
34. Sucov,H.M., Murakami,K.K. and Evans,R.M. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 5392–5396.
35. Leroy,P., Nakshatri,H. and Chambon,P. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 10138–10142.
36. Smith,W.C., Nakshatri,H., Leroy,P., Rees,J. and Chambon,P. (1991) *EMBO J.*, **10**, 2223–2230.
37. Vasios,G., Nader,S., Gold,J.D., Leid,M., Lutz,Y., Gaub,M.P., Chambon,P. and Gudas,L. (1991) *EMBO J.*, **10**, 1149–1158.
38. Vivanco Ruiz,M., Bugge,T.H., Hirshmann,P. and Stunnenberg,H.G. (1991) *EMBO J.*, **10**, 3829–3838.
39. Durand,B., Saunders,M., Leroy,P., Leid,M. and Chambon,P. (1992) *Cell*, **71**, 73–85.
40. Näär,A.M., Boutin,J-M., Lipkin,S.M., Yu,V.C., Holloway,J.M., Glass,C.K. and Rosenfeld,M.G. (1991) *Cell*, **65**, 1267–1279.
41. Umesono,K., Murakami,K.K., Thompson,C.C. and Evans,R.M. (1991) Cell, **65**, 1255–1266.
42. Mangelsdorf,D.J., Umesono,K., Kliewer,S.A., Borgmeyer,U., Ong,E.S. and Evans,R.M. (1991) *Cell*, **66**, 555–561.
43. Kliewer,S.A., Umesono,K., Heyman,R.A., Mangelsdorf,D.J., Dyck,J.A. and Evans,R.M. (1992) *Proc. Natl. Acad. Sci.* USA, **89**, 1448–1452.
44. Heyman,R.A., Mangelsdorf,D.J., Dyck,J.A., Stein,R.B., Eichele,G., Evans,R.M. and Thaller,C. (1992) *Cell*, **68**, 397–406.
45. Levin,A.A., Sturzenbecker,L.J., Kazmer,S., Bosakowski,T., Huselton,C., Allenby,G., Speck,J., Kratzeisen,C., Rosenberger,M., Lovey,A. and Grippo,J.F. (1992) *Nature*, **355**, 359–361.
46. Zhang,X.K., Lehmann,J., Hoffman,B., Dawson,M.I., Cameron,J., Graupner,G., Hermann,T., Tran,P. and Pfahl,M. (1992) *Nature*,**358**, 587–591.
47. O'Malley,B.W. and Conneely,O.M. (1992) *Mol. Endocrinol.*, **6**, 1359–1361.

48. Miyajima,N., Kadowaki,Y., Fukushige,S.I., Shimizu,S.I., Semba,K., Yamanoshi,Y., Matsubara,K.I., Toyoshima,K. and Yamamoto,T. (1988) *Nucleic Acids Res.*, **16**, 11057–11074.
49. Wang,L.H., Tsai,S.Y., Cook,R.G., Beattie,W.G., Tsai,M-J. and O'Malley,B.W. (1989) *Nature,* **340**, 163–166.
50. Ladias,J.A.A. and Karathanasis,S.K. (1991) *Science*, **251**, 561–565.
51. Sagami,I., Tsai,Y., Wang,H., Tsai,M-J. and O'Malley,B.W. (1986) *Mol. Cell. Biol.*, **6**, 4259–4267.
52. Hwung,Y.P., Crowe,D., Wang,L.H., Tsai,S.Y. and Tsai,M.J. (1988) Mol. Cell. Biol., **8**, 2070–2077.
53. Hwung,Y.P.,Wang,L.H., Tsai,S.Y. and Tsai,M.J. (1988) J. Biol. Chem., **263**, 13470–13474.
54. Wijnolds,J., Philipsen,N.J. and Ab,G. (1988) *EMBO J.*, **7**, 2757–2763.
55. Orchard,K., Perkins,N., Chapman,C., Harris,J., Emery,V., Goodwin,G.,Latchman,D. and Collins,M. (1990) *J. Virol.*,**64**, 3234–3289.
56. Cooney,A.J., Tsai,S.Y., O'Malley,B.W. and Tsai,M.J. (1991) *J. Virol.*, **65**, 2853–2860.
57. Liu,Y. and Teng,C.T. (1991) *J. Biol. Chem.*, **266**, 21880–21885.
58. Wijnolds,J., Muller,E. and Ab,G. (1991) *Nucleic Acids Res.*, **19**, 33–41.
59. Cooney,A.J., Tsai,S.Y., O'Malley,B.W. and Tsai,M.J. (1992) *Mol. Cell. Biol.*, **12**, 4153–4163.
60. Tran,P., Zhang,X.K., Salbert,G., Herrman,T., Lehmann,J.M. and Pfahl,M. (1992) *Mol. Cell. Biol.*, **12**, 4666–4676.
61. Widom,R.L., Rhee,M. and Karathanasis,S.K. (1992) *Mol. Cell. Biol.*, **12**, 3380–3389.
62. Rudniki,M.A. and McBurney,M.W. (1987) in Robertson,E.J. (ed), Teratocarcinomas and embryonic stem cells a practical approach. IRL Press, Oxford, pp 19–49.
63. Jones-Villeneuve,E.M.V., Rudniki,M.A., Harris,J.F. and McBurney, M.W. (1983) *Mol. Cell. Biol.*, **3**, 2271–2279.
64. Kalb,JR.,V.F. and Bernlohr,R.W. (1977) *Anal. Biochem.*, **82**, 362–371.
65. Stunnenberg,H., Lisage,H., Philipson,L., van Miltenburg,R.T. and van der Vliet,P.C. (1988) *Nucleic Acids Res.*, **16**, 2431–2444.
66. Tsukiyama,T., Niwa,O. and Yokoro,K. (1989) *Mol. Cell. Biol.*, **9**, 4670–4676.
67. Tsukiyama,T., Ueda,H., Hirose,S. and Niwa,O. (1992) *Mol. Cell. Biol.*, **12**, 1286–1291.
68. Pugh,B.F. and Tjian,R. (1991) *Genes Dev.*, **5**, 1935–1945
69. Weis, L. and Reinberg, D. (1992) FASEB J. 6, 3300–3309.
70. Campione-Piccardo,J., Sun,J.J., Craig,J. and McBurney,M.W. (1985) *Dev. Biol.*, **109**, 25–31.
71. Pratt,M.A.C., Kralova,J. and McBurney,M.W. (1990) *Mol. Cell. Biol.*, **10**, 6445–6453.
72. Kruyt,F.A.E., van der Veer,L.J., Mader,S., van den Brink,C.E., Feijen,A., Jonk,L.J.C., Kruijer,W. and van der Saag,P.T. (1992) *Differentiation*, **49**, 27–37.
73. Kruyt,F.A.E., van den Brink,C.E., Defize,L.H.K., Donath,M.J., Kastner,P., Kruijer,W., Chambon,P. and van der Saag,P.T. (1991) *M.O.D.*, **33**, 171–178.
74. Wang,L.H., Ing, N.H., Tsai,S.Y., O'Malley,B.W. and Tsai,M-J. (1991) *Gene Express*. **1**, 207–216.
75. Du, W., Thanos,D., and Maniatis,T. (1993) *Cell*, **74**, 887–898.
76. Ben-Shushan,E., Pikarsky,E., Klar,A. and Bergman,Y. (1993) *Mol. Cell. Biol.*, **13**, 891–901.
77. Pikarsky,E., Sharir,H., Ben-Shushan,E. and Bergmann,Y. (1994) *Mol. Cell. Biol.*, **14**, 1026–1038.
78. Schoorlemmer,J., van Puijenbroek,A., van den Eijnden,M., Jonk,L., Pals,C. and Kruijer,W. (1994) *Mol. Cell. Biol.*, **14**, 1122–1136.