

The major transcription initiation site of the SV40 late promoter is a potent thyroid hormone response element

Béatrice Desvergne* and Tatiana Favez

Institut de Biologie Animale, Université de Lausanne, Bâtiment de Biologie, CH-1015 Lausanne, Switzerland

Received January 7, 1997; Revised and Accepted March 14, 1997

ABSTRACT

Thyroid hormone receptors (TRs) are members of the nuclear hormone receptor superfamily, which act as transcription factors upon binding to specific DNA sequences called thyroid hormone (T3) response elements (TREs). Such elements are found in the upstream regulatory region of promoters as well as in intragenic sequences of T3-responsive genes. In this report, we demonstrate that SV40 late (SVL) promoter activity is strongly down-regulated by TR in the absence of ligand. Addition of T3 releases this repression, but does not further induce SVL promoter activity. Electrophoretic mobility shift analyses reveal a TR binding element that overlaps with the SV40 major late transcription initiation site. This element closely fits the consensus TRE, formed of two hexanucleotides organized in a tandem repeat separated by 4 nt, and is able to confer T3 responsiveness on a heterologous promoter. We further show that, although the presence of TR leads to quantitatively modified expression of an SVL-driven reporter gene, neither displacement of the site of transcription initiation nor modification of the splicing pattern of the primary transcripts occur.

INTRODUCTION

Thyroid hormone receptors (TRs) belong to the superfamily of nuclear receptors, together with receptors such as those for steroid hormones, vitamin D and retinoic acid (reviewed in 1). They bind to specific DNA sequences, called thyroid hormone response elements (TREs), and act as transcription factors that modulate the promoter activity of TRE-containing genes. The TREs are composed of two half-sites derived from the consensus hexamer AGGTCA, arranged either as a palindrome, a direct repeat or as an inverted palindromic array, with a spacing between the two hexamers of 0, 4 and 6 nt respectively (reviewed in 2). TR preferentially binds as a heterodimer with RXR, another member of the nuclear receptor superfamily, but it can also bind as a homodimer or as a monomer. In *in vitro* binding assays as well as *in vivo* within the chromosomal chromatin context (3) TR binds to its element in the absence of thyroid hormone (T3). Binding of

unliganded TR results in transcriptional repression, as measured in functional assays, while the addition of T3 not only relieves this repression but allows a further activation of transcription above the derepressed basal level. However, with some specific response elements and promoter contexts T3 may lead to negative regulation of target gene expression (reviewed in 4).

For most DNA viruses, promoter activation depends on use of the transcriptional apparatus of the host cell. Analyses of transcriptional regulatory mechanisms in mammalian cells have largely benefitted from the use of these viruses as models (reviewed in 5). Indeed, the first described hormone response elements were glucocorticoid response elements found in the long terminal repeat (LTR) of mouse mammary tumor virus (MMTV) (reviewed in 6). Simian virus 40 (SV40), with its small circular double-stranded DNA genome, provided one of the first tools for characterizing enhancer elements (reviewed in 7,8). The SV40 late (SVL) promoter is a well-characterized TATA-less promoter and, as for most members of this class of promoters, transcription can initiate at multiple sites. However, the major start site maps at nt 325 with respect to the Buchman numbering of the SV40 genome (9) and accounts for 80–90% of total late RNA synthesis (10 and references therein). Upstream regulatory *cis*-elements modulating the level of late expression include the SV40 origin of replication, the three 21 bp repeats and the 72 bp direct repeat (11,12 and references therein). The core promoter of SVL consists of three elements which are sufficient for basal promoter activity and specify the start site of major late RNA synthesis. One of these elements is centred at nt 294 and functions as a binding site for TFIID (13), while the two other sites are centred at nt 325, i.e. the transcription initiation site itself, and at nt 352 (14). More recently, Wiley *et al.* (15) showed that some cellular transcriptional repressors, collectively called initiator binding proteins, bind to sequences located at and close to the transcription initiation site and might be involved in the SV40 early to late switch.

The strength of many viral promoters, together with the characteristics of specific viral sequences involved in messenger RNA maturation, also led to their development as tools for efficient expression of exogenous proteins in mammalian cells. Such expression vectors contain a viral promoter, which drives mRNA synthesis from inserted foreign cDNAs, as well as a sequence containing splicing acceptor/donor sites and a polyadenylation

*To whom correspondence should be addressed. Tel: +41 21 692 4140; Fax: +41 21 692 4115; Email: beatrice.desvergne@iba.unil.ch

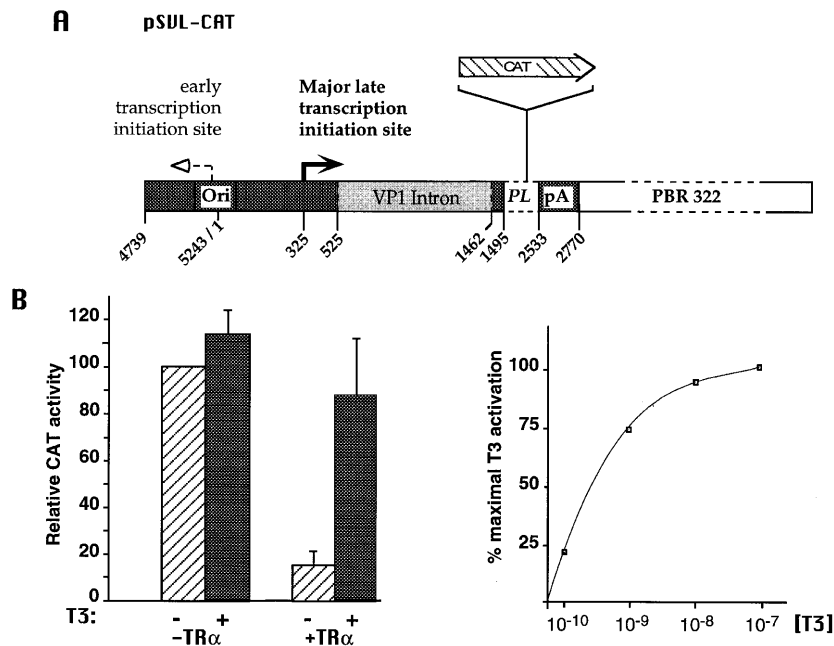


Figure 1. (A) Schematic representation of the vector pSVL with the site of insertion of the CAT coding sequence. The sequence coordinates of the SV40 genome fragments, boxed in grey, are given. PL, polylinker sequence. (B) Regulation of pSVL-CAT expression by TR α and T3. NIH 3T3 cells were transfected by electroporation with 5 μ g pSVL-CAT, 3.2 μ g pSG5-TR α (TR α +) or 3.2 μ g empty vector pSG5 (TR α -) and 5 μ g PUC19 as carrier DNA; 0.25 μ g pCMV β expression vector were added as an internal control for transfection efficiency. Cells were cultured for 72 h in the presence (+) or absence (-) of T3. (Left) CAT activity present in the corresponding cellular extracts normalized to β -galactosidase activity and expressed as a percentage of CAT reporter activity measured in the absence of TR and T3 at 10⁻⁷ M. Values are the mean \pm SD of three independent experiments, each performed in triplicate. (Right) T3 dose-response curve of pSVL-CAT activity in the presence of TR α . 100% corresponds to maximal T3 activation, obtained at 10⁻⁷ M.

signal. While attempting to take advantage of pSVL, a SVL-derived vector, to drive expression of TRs, preliminary experiments suggested to us that TR itself was affecting the efficiency of the pSVL vector. In this work we demonstrate that a functional and potent TRE overlaps with the SV40 major late transcription initiation site and we further characterize the regulatory role of TR in this unusual context.

MATERIALS AND METHODS

Plasmid constructions

pSVL-CAT was obtained by inserting the chloramphenicol acetyltransferase (CAT) coding sequence into the *Bam*HI site of pSVL (Pharmacia LKB Biotechnology). For the pTRE_{SVL}-TKCAT construct, a synthetic double-stranded oligonucleotide encompassing the SVL sequence from nt 308 to 339 and flanked on its 5'- and 3'-side by a *Hind*III and *Bam*HI restriction site respectively was inserted into the pBLCAT2 reporter plasmid (16) upstream of the -105 base pair of the TK promoter. In pTRE_{SVL}-CAT, the same oligonucleotide was introduced into the promoterless vector pBLCAT3 (16). For the vector pSVL-neo, the SVL sequence from nt 39 to 417 was PCR amplified using corresponding 5' and 3' primers with an extended sequence encompassing a *Nde*I and *Hind*III restriction site respectively. The MMTV LTR was removed from pMAM-neo (Clontech) by *Nde*I/*Hind*III digestion and replaced by the amplified fragment. For *in vitro* transcription/translation and for transfections, the rat TR α cDNA was inserted into a modified pSG5 (pSG5PL, a gift from Dr H.Richard-Foy).

Cell culture and transfections

NIH 3T3 cell maintenance and transfection were performed as described previously (17). pCMV β (Clontech), which expresses β -galactosidase, was used as an internal control of transfection efficiency. Following transfection, cells were harvested after a 72 h incubation in hormone-free medium (18) with or without 10⁻⁷ M T3 (Sigma). CAT and β -galactosidase activities were assayed as described (17,19). Total RNA was prepared from transfected cells using the TRIZOL reagent (Gibco), according to the manufacturer's instructions. To eliminate possible contamination with co-purified small DNA molecules such as plasmids, RNA preparations were further digested with a RNase-free DNase (RQ1 DNase) for 30 min at 37°C.

Electrophoretic mobility shift assay

The probes, DNA restriction fragments or double-stranded oligonucleotides were ³²P-end-labelled by a Klenow fragment filling-in reaction. *In vitro* transcription was performed on linearized pSG5-TR α and followed by *in vitro* translation according to the manufacturer's protocols (Promega). Binding reactions were carried out by incubating 3 μ l unprogrammed or programmed reticulocyte lysate and 0.5 μ l nuclear extract from Sf9 cells infected with a recombinant baculovirus overexpressing mouse RXR β (prepared as described in 20,21), in the presence of 22 mM HEPES, 1.25 mM dithiothreitol, 2.5 mM MgCl₂, 40 mM KCl, 12% glycerol and 250 μ g/ml poly(dA-dT) in a 20 μ l reaction to which 20 000 c.p.m. labeled probe were added. For competition experiments, the indicated molar excess of unlabelled double-stranded oligonucleotide was added to the mixture. When used,

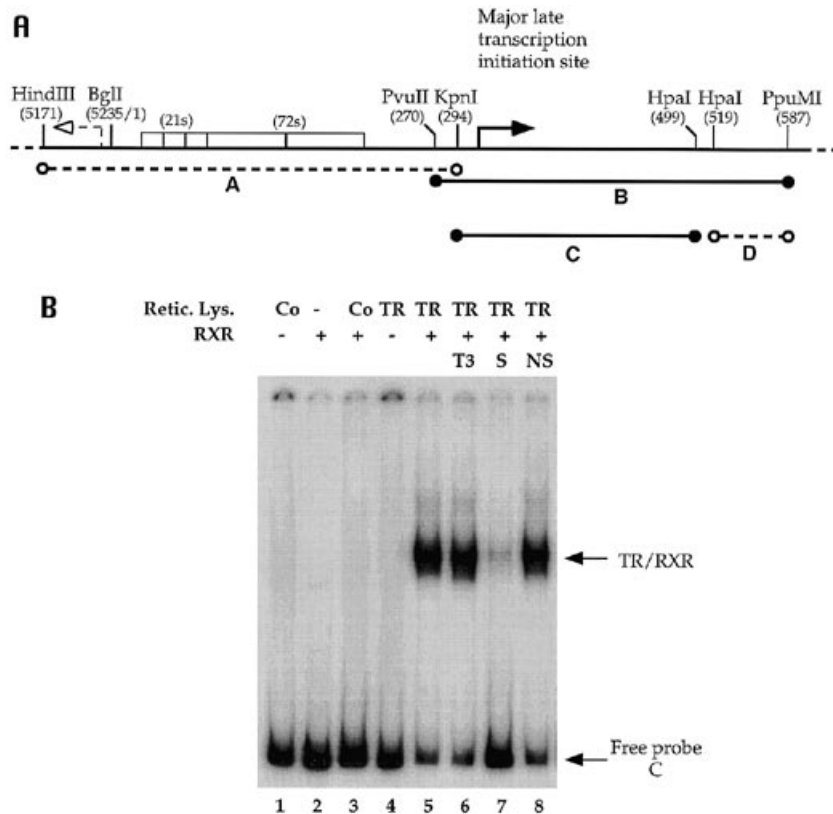


Figure 2. The SVL promoter fragment from the *KpnI* (294) to the *HpaI* (498) site contains a binding site for TR α . (A) Schematic diagram of the probes tested in binding assays with TR α and RXR. Probes A and D (dotted lines) did not allow formation of TR–RXR complexes, while probes B and C (plain lines) were positive. Boxes (21s) and (72s) correspond to enhancer elements characterized for the SV40 early and late promoters. (B) Binding reactions with probe C. *In vitro* translated TR α (TR) or unprogrammed lysate (Co) in the presence or absence of nuclear extracts from Sf9 cells infected with a recombinant baculovirus expressing RXR β (+ or – as indicated in the figure) were mixed with 20 000 c.p.m. radiolabelled probe C and the reaction performed as described in Materials and Methods. In lane 6, T3 has been added at a final concentration of 10^{-7} M. Competition for binding to the probe was achieved using a 100-fold molar excess of an unlabelled double-stranded oligonucleotide either specific, encompassing TRE_{ME} (S; lane 7), or non-specific, bearing an unrelated sequence (NS; lane 8).

T3 was at a final concentration of 10^{-7} M. Incubation, electrophoresis and exposure were performed as previously described (17).

S1 nuclease

Probe preparation. The vector pSVL-neo was digested with *SphI* and used as template in asymmetric PCR. A 32 P-end-labelled oligonucleotide complementary to 25 bases within the neomycin resistance gene cDNA was used as unique primer (see Fig. 7A). After 30 cycles at an annealing temperature of 55°C, the PCR reaction was subjected to electrophoresis on a 6% polyacrylamide–8 M urea gel. The band corresponding to the specific labelled single-stranded DNA was identified after a 2 min exposure of the wet gel, eluted in 0.5 M ammonium acetate, 1 mM EDTA, precipitated and resuspended in TE.

S1 nuclease assay. Aliquots of 20 μ g total RNA from NIH 3T3 cells, transfected and treated as indicated, were hybridized overnight at 30°C with 50 000 c.p.m. probe in the presence of 80% formamide, 0.4 M NaCl, 40 mM PIPES, pH 6.4, 1 mM EDTA, pH 8.0. After a 10-fold dilution of the hybridization mixture with 0.28 M NaCl, 4.5 mM ZnSO₄, 50 mM sodium acetate, pH 4.5, 300 U S1 nuclease (Pharmacia) were added for a 1 h incubation at 30°C in the presence of 2 μ g single-stranded

calf thymus DNA. The reaction was stopped by adding 80 μ l stop solution (4 M ammonium acetate, 20 mM EDTA, pH 8.0, 40 μ g/ml tRNA), precipitated and resuspended in 3 μ l TE and 4 μ l RNA loading buffer (80% formamide, 1 mM EDTA, pH 8.0, 0.1% bromophenol blue, 0.1% xylene cyanol). After denaturation for 5 min at 95°C, the samples were electrophoresed on 6% polyacrylamide–8 M urea gels in 1 \times TBE. Autoradiography of the gel was performed with a 10 day exposure.

RT-PCR

Aliquots of 20 μ g total RNA from NIH 3T3 cells, transfected and treated as indicated, were co-precipitated with 3 pmol CAT primer or 3 pmol act2 primer, resuspended in 80 mM Tris–HCl, pH 8.3, 80 mM KCl and hybridized overnight at 52°C. The reverse transcriptase buffer (33 mM Tris–HCl, pH 8.3, 33 mM KCl, 10 mM MgCl₂, 16 mM DTT, 1 mM 4 mix dNTP, 10 U RNasin) and 16 U AMV reverse transcriptase were added for an incubation of 1 h at 42°C. The reaction was stopped with phenol/chloroform extraction of the cDNA, which was then ethanol precipitated and resuspended in water. The cDNA was then used as template for the PCR amplification using four different pairs of primers: p1/CAT, p2/CAT, p1/p3 and act1/act2. Aliquots of 2.5 μ g cDNA were mixed with 2 μ M each primer, 0.2 mM each dNTP, 2 mM MgCl₂, 20 mM Tris–HCl, pH 8.4,

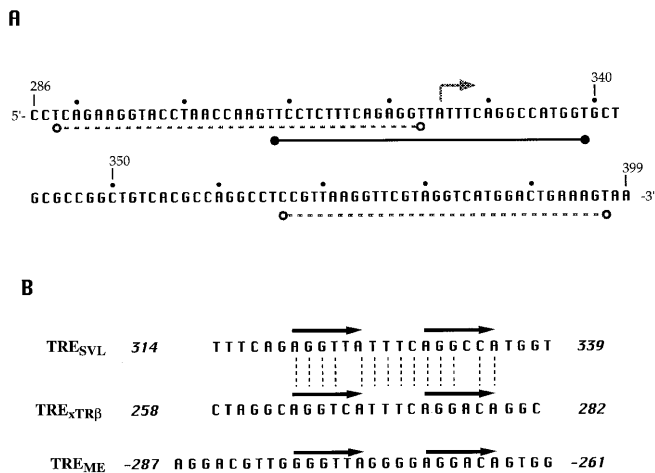


Figure 3. (A) Sequence surrounding the major transcription initiation site of the SVL promoter. The sequences of the oligonucleotides used in competition assays are underlined; plain line, efficient competition for TR-RXR binding to probe C (described in Fig. 2); dotted lines, no competition. (B) Sequence and coordinates of the TR-RXR binding site contained in the SVL promoter (TRE_{SVL}) aligned with those of the *Xenopus* TR β A gene ($TRE_{xTR\beta}$) and those of the malic enzyme gene (TRE_{ME}). The double arrows indicate the two hexamers corresponding to consensus half-sites, organized for each of these TREs as a direct repeat with a four base pair spacing.

50 mM KCl and 1 U Taq polymerase. After 39 cycles, the PCR reaction products were loaded on a 1.2% agarose gel in 1 \times TBE. The respective position of the primers p1 (AACGCCTTTTTGTGTTTGT), p2 (TATCATTTGGGCACACCTAT), p3 (CTACAGCCATTCCTGGTTGT) and CAT (GCAACTGACTGAAATGCC TC) are shown in Figure 6A. The primers act1 (ATATCGCTGCGCTGGTCGTCGACAA) and act2 (AGCACAGCCTGGATGGCTACGTACA) correspond to the mouse β -actin cDNA from nt 91 to 115 and from nt 499 to 475 respectively.

RESULTS

Activity of the reporter gene pSVL-CAT is repressed by unliganded TR α

The pSVL vector contains a SV40 genome fragment from nt 4739 to 1495 comprising the origin of replication, the SVL promoter, the leader sequences and the VP1 intron, as well as the SV40 polyadenylation signal from nt 2533 to 2770 (Fig. 1A). Because of the removal of the first ATG of VP1, this vector allows expression of protein starting with the first ATG of the cDNA inserted in the polylinker (PL) sequence. To monitor activity of the SVL promoter and to analyse its regulation, we inserted the CAT coding sequence into the *Bam*HI site and used this reporter gene in transfection assays performed in NIH 3T3 cells, which do not express detectable levels of endogenous TR. Expression of the α subtype of TR (TR α) reduces basal activity of the pSVL-CAT construct by 85%. Addition of T3 to the cell culture medium relieves this repression in a dose-dependent manner, restoring CAT activity close to the basal level at 10^{-7} M (Fig. 1B). The same results were obtained when using TR β 1, the other TR subtype, though derepression in the presence of T3 was less efficient than that observed with TR α . In contrast, overexpression of c-erbA α 2, which is an alternative product of the TR α gene and which does not bind T3 (22), does not modify SVL promoter

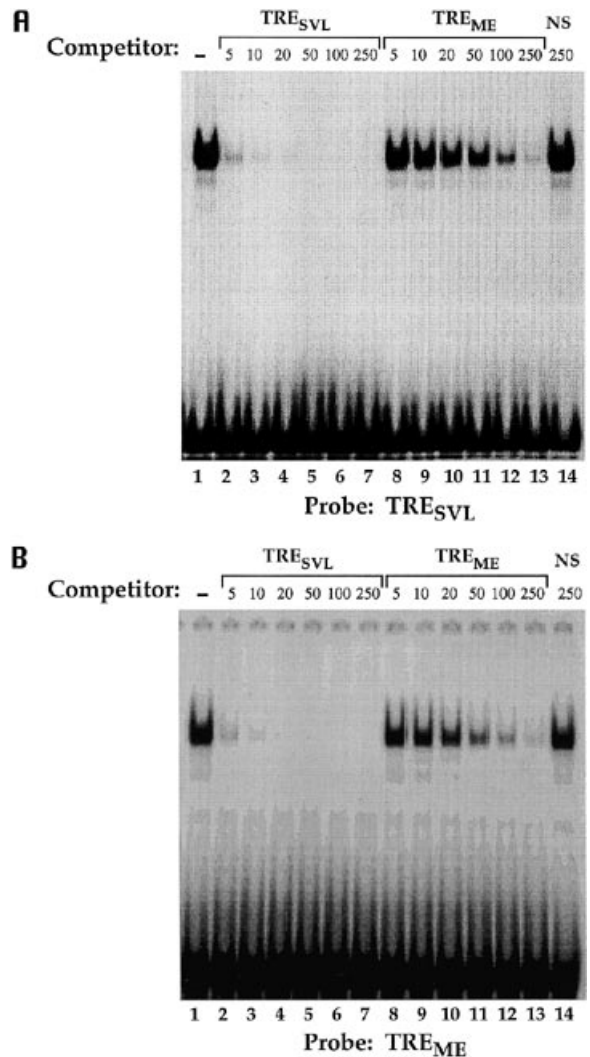


Figure 4. TRE_{SVL} is a high affinity binding site for the TR α -RXR heterodimer. Competition analyses were performed using EMSA as described in Materials and Methods. (A) EMSA with the oligonucleotide containing the TRE_{SVL} sequence as probe. (B) EMSA with the oligonucleotide containing the TRE_{ME} sequence as probe. Competitors correspond to 33 bp double-stranded oligonucleotides encompassing either TRE_{SVL} (lanes 2–7) or TRE_{ME} (lanes 8–13) or an unrelated sequence (NS, lane 14) at the indicated fold molar excess.

activity in the absence or presence of T3 (data not shown). Thus, these results provide evidence for regulation of SVL promoter activity by TR and T3.

The heterodimer TR α -RXR binds to an element within the SVL promoter

We thus searched for the putative TRE mediating TR action in pSVL-CAT by performing electrophoretic mobility shift analyses (EMSA) with TR and RXR. We first tested two probes, A and B, covering a total of 651 bp of the promoter sequence from nt 5171 to 587 (Fig. 2A). Since only probe B exhibited strong binding of the TR-RXR complex (data not shown), probes C and D derived from B were used and the TR-RXR binding region was mapped to the *Kpn*I-*Hpa*I DNA fragment C (nt 294–499). As seen in Figure 2B, neither TR α alone nor RXR β alone is able to bind to this probe, while a strong complex, migrating as a single band,

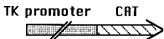



Vector :		T3 mediated fold induction	
		-TR α	+TR α
pBLCAT2		1.04±0.02	1.2±0.6
pTRE _{SVL} -TK-CAT		7.3±2.2	144±65
pTK1AM		1.2±0.4	558±76
pTRE _{SVL} -CAT		No Detect.	No Detect.

Figure 5. The SVL sequence which binds TR α is a functional TRE. Five micrograms of various CAT reporter genes as indicated were transfected into NIH 3T3 cells together with 12 μ g PUC19 (carrier DNA) and 0.25 μ g pCMV β expression vector (internal control) and with 3.2 μ g expression vector pSG5-TR α (TR+) or the empty vector PSG5PL (TR-). The fold induction corresponds to the ratio of the normalized CAT activity present in T3-treated cells over that present in non-treated cells and is the mean \pm SD of three independent experiments, each performed in duplicate. No transcriptional activity was detected (No Detect.) with the pTRE_{SVL}-CAT construct.

appears when both TR α and RXR are present in the binding reaction. As expected for a TR α -RXR heterodimer, addition of T3 does not alter the amount of this complex but provokes a very small but reproducible down-shift (Fig. 2B, compare lanes 5 and 6). The retarded band is specifically competed out by a 100-fold excess of an oligonucleotide corresponding to the TRE of the malic enzyme gene (TRE_{ME}) (17) but not by the same excess of an unrelated competitor (Fig. 2B, lanes 7 and 8).

Sequence analysis of DNA fragment C reveals three sites bearing some resemblance to a TRE, in the region shown in Figure 3A. The first putative binding site corresponds to the DNA sequence from nt 289 to 323, previously shown to be an element of the SVL core promoter. The second site, from nt 310 to 339, contains the major late initiation site, while the third site, from nt 366 to 394, has recently been described as containing a binding site for human testicular receptor 2 (TR2) and testicular receptor 4 (TR4), which are orphan receptors that also belong to the nuclear receptor superfamily (23). The corresponding double-stranded oligonucleotides were synthesized and used as competitors in an EMSA. Only oligonucleotide 308/339 is able to efficiently compete out formation of the TR-RXR retarded complex bound to the *KpnI*-*HpaI* DNA fragment, while oligonucleotides 288/323 and 366/397 give only poor or no competition respectively (data not shown). In Figure 3B, the newly identified TR-RXR binding site is compared with the well-characterized TRE_{ME}. A good level of similarity is found when considering the sequence of the two hexanucleotides forming each half-site (10 nt out of 12). However, the resemblance of the SVL element with a TRE recently characterized in the promoter region of the *Xenopus* TR β A gene (3) is even more striking, since the identity also encompasses the four spacing base pairs.

To assess the relative strength in terms of TR-RXR binding of this newly discovered TRE, hereafter referred to as TRE_{SVL}, we used it in reciprocal competition with TRE_{ME}. As seen in Figure 4A, binding of the TR α -RXR complex to the SVL probe is more efficiently competed out by TRE_{SVL} than by TRE_{ME}. In agreement with this result, TR α -RXR complexes bound to the ME probe are also more efficiently displaced by TRE_{SVL} than by

TRE_{ME} (Fig. 4B), confirming that TRE_{SVL} is a potent TR binding site.

The TR α -RXR binding site of the SVL promoter acts as a positive TRE in a heterologous promoter

To test if TRE_{SVL} has the capability of mediating an *in vivo* T3 response, we inserted it upstream of the thymidine kinase (TK) promoter in pBLCAT2 and tested this reporter gene activity by transfection into NIH 3T3 cells. For comparison we used the same reporter construct containing the previously described TRE_{ME} (pTK1AM; 17). Co-transfection of the TR α -expressing vector leads to repression of the basal activity of TRE-containing promoters while it does not affect that of pBLCAT2 (data not shown). Under these conditions, addition of T3 does not simply relieve this repression but also provokes a high level of induction above the basal level, resulting in a final 150-fold induction of pTRE_{SVL}-TKCAT activity and 550-fold for pTK1AM activity (Fig. 5). Interestingly, in the absence of overexpressed TR α , the basal activity of pTRE_{SVL}-TKCAT is significantly higher than that of the parent vector pBLCAT2. Moreover, T3 does not significantly affect expression of the control vector or of pTK1AM while it provokes a significant 7-fold induction of pTRE_{SVL}-TKCAT activity, for reasons which are not clear at this time.

Since TRE_{SVL} corresponds to the main transcription initiation site of the SVL promoter, we inserted it into the promoterless vector pBLCAT3 and tested it in transfections. However, neither basal nor induced activity of this new pTRE_{SVL}-CAT construct could be detected in the presence or absence of TR α and/or T3 (Fig. 5).

In summary, these data demonstrate that TRE_{SVL} can confer T3 responsiveness to a heterologous promoter and has the property of a regular positive TRE.

TR does not affect splicing of RNAs expressed from the pSVL vector

One particularity of the vector pSVL used in this study is the presence of the VP1 intron sequence, from nt 525 to 1462. Because T3 has been shown to alter alternative splicing of the tau gene primary transcript during development (24), we verified that such a mechanism was not involved in regulation by T3 of gene expression driven by the SVL promoter. For this purpose we designed a RT-PCR-based assay to detect the presence of spliced and unspliced CAT mRNA. After transfection into NIH 3T3 cells of pSVL-CAT with or without the TR expression vector and in the presence or absence of hormone, followed by RNA extraction, we used the strategy summarized in Figure 6A to specifically amplify the CAT mRNAs by RT-PCR. The first pair of primers, p1/CAT, specifically amplifies a product of 320 bp, corresponding to spliced CAT mRNA, as well as a product of 1260 bp, corresponding to unspliced CAT mRNA, from transfected cells but not from control cells (Fig. 6B, left, lanes 2-5). To further assess the presence of the unspliced form, we used two different pairs of primers designed for specific amplification of unspliced CAT mRNA: p2/CAT and p1/p3 (Fig. 6A). The corresponding RT-PCR assay indeed revealed single specific products of 400 and 480 bp respectively, whose presence was observed in the presence as well as in the absence of TR and T3 (Fig. 6B, middle, and data not shown respectively). As a control we performed a parallel RT-PCR assay using actin-specific primers (Fig. 6B, right). Thus, while our experimental design did not allow for

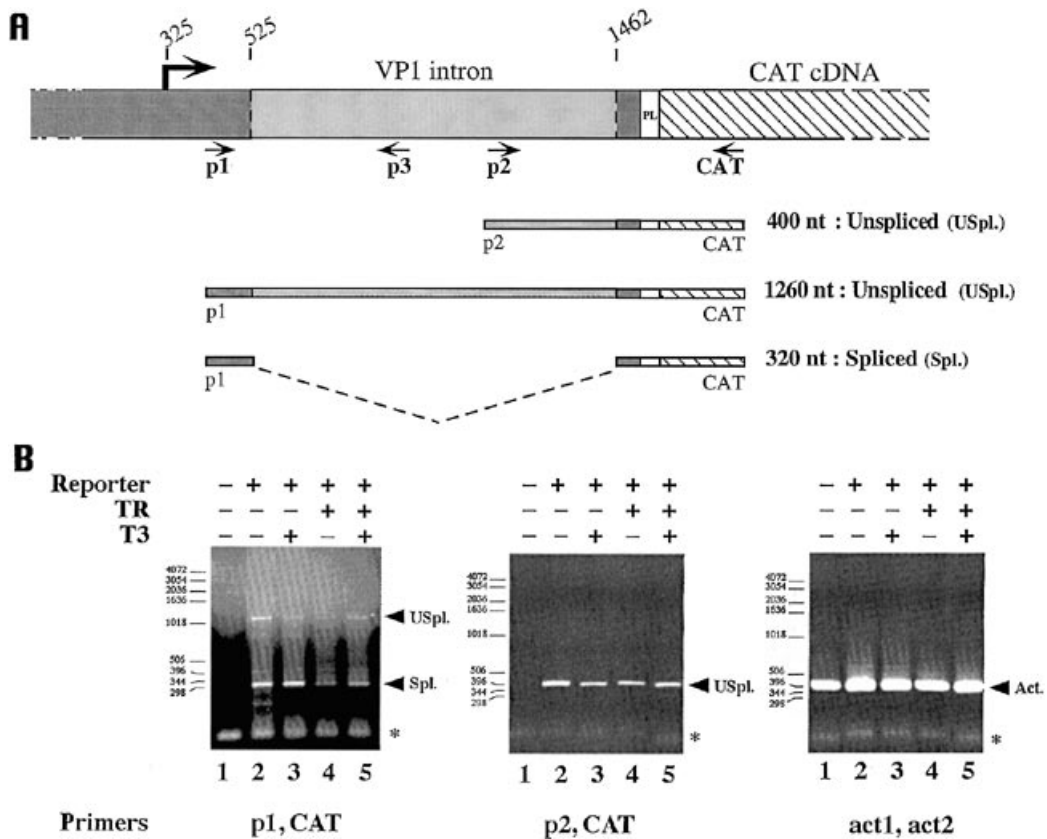


Figure 6. Splicing of CAT mRNAs in the presence and absence of TR. (A) Localization of the primers (p1, p2, p3 and CAT) used to test the presence of spliced and unspliced mRNA produced by pSVL-CAT transfected into NIH 3T3 cells and representation of the amplified product. The primers act1 and act2, used in a control reaction, amplified a 410 bp fragment and are described in Materials and Methods. (B) RT-PCR assays were performed on 20 μ g total RNA extracted from NIH 3T3 cells transfected with pSVL-CAT, together with pSG5-TR α or alone (TR+ or -) and cultured in the presence or absence of T3 (T3+ or -). Reactions were loaded on a 1.2% agarose gel, stained with ethidium bromide and recorded as digital images. The free primers are marked by a star. The closed arrowheads indicate the positions of the specific bands. Lane 1 in each panel corresponds to a reaction performed with total RNA prepared from non-transfected NIH 3T3 cells.

quantitative evaluation, we demonstrate that both spliced and unspliced CAT mRNAs are detected, regardless of the presence or absence of TR and T3, and thus that TR-dependent repression of CAT activity is not due to an alteration of messenger RNA processing.

Binding of TR at the major late transcription initiation site does not shift transcription initiation

Because TRE_{SVL} comprises the major transcription initiation site of the SVL promoter, we analysed the possibility of a change of position of this site in the presence of TR. Preliminary experiments using primer extension as well as S1 nuclease and RNase protection assays suggested that the presence of both spliced and unspliced messages was impeding these experiments. We thus constructed a pSVL-neo vector in which the VP1 intron sequence was removed together with the splicing junctions (Fig. 7A). Total RNA was then prepared from NIH 3T3 cells transfected with the pSVL-neo vector with or without the TR expression vector. To determine the transcription initiation site, we performed an S1 nuclease assay using a single-stranded DNA probe encompassing the TRE/transcription initiation site at nt 325 (Fig. 7A). The experimental design did not allow for quantification of the transcripts and the signal strength differences in the experiment shown in Figure 7B were not reproducible. However, all three

independent transfections and S1 nuclease assays performed showed that within the 400 bases covered by the probe, only one transcription initiation site is detectable, regardless of the culture and transfection protocol (Fig. 7B). A sequencing reaction performed with a primer corresponding to the neomycin end of the probe allowed us to precisely position this initiation site at nt 325, which is indeed the usual major late transcription initiation site. Thus, while binding of TR to the transcription initiation site strikingly reduces SVL-driven expression, it does not lead to a change of position of transcription initiation.

DISCUSSION

In this paper we have identified and functionally characterized a TRE present at the major late transcription initiation site of the SVL promoter. Many binding sites for transcription factors were first described in viral promoters (reviewed in 5) and hormone response elements (HREs) such as the one described herein are no exception. The role of such response elements in the viral cycle remains to be elucidated. However, one very important fact is that viral promoters continue to be important tools to further our understanding of the mechanisms of transcription regulation of mammalian genes. In that respect, our discovery of a TRE at an unusual location in the SVL promoter has two consequences. First, it gives us a new tool to extend our understanding of the T3

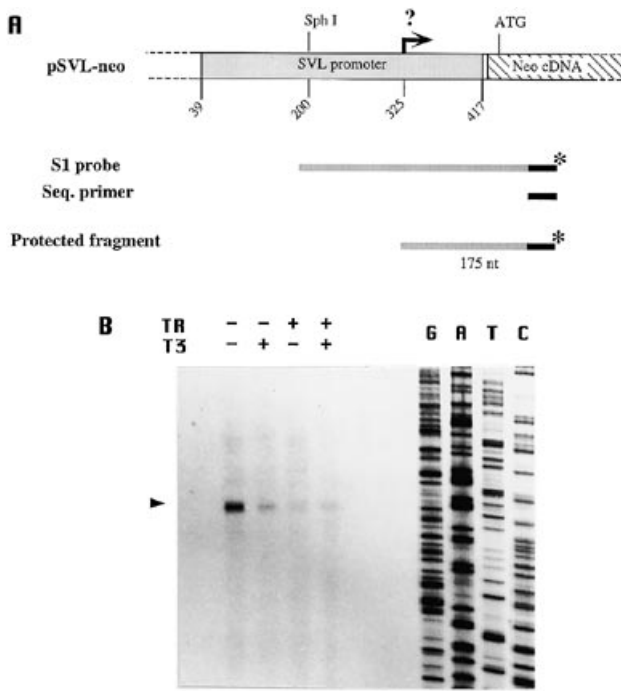


Figure 7. S1 nuclease analysis of the pSVL-neo transcripts. **(A)** Schematic diagram of the SVL-neo vector which does not contain the SVL intron sequence. The single-stranded S1 probe is drawn with its labelled end indicated by a star. The same primer (black line) was used for generating the S1 probe and for the sequencing reaction loaded in parallel with the S1 nuclease reactions (see Materials and Methods). **(B)** S1 nuclease analyses were performed on 20 μ g RNA extracted from NIH 3T3 cells transfected with pSVL-neo, alone or together with pSG5-TR α (TR- or +) and cultured in the presence or absence of T3 (T3+ or -). A start site is observed at nt 325, corresponding to the major initiation site used by the wild-type SV40 virus for its late transcription. No other initiation sites appear upon addition of TR α and/or T3.

mechanism of action. Second, it implies that one should be cautious when using a SVL-derived expression vector in studies on thyroid hormone responsiveness. A similar warning has also been propounded for the use of the luciferase reporter gene, which bears, within the structural gene sequence, a potent T3-dependent negative regulatory element (25).

Interestingly, two non-classical TREs have been previously described in viral promoters. One of these is present in the thymidine kinase gene of Herpes sarcoma virus (HSV TK). It requires a functional interaction with the CTF/NF1 binding sequence located 80 bp upstream of it to be active (26). This interaction is likely to be tissue specific, since T3 regulation of the TK promoter has been described in GH4C1 cells but is not found in other cell lines (17,27-31 and this study). The second TRE, in the Rous sarcoma virus LTR, mediates ligand-independent activation by TR, activation that is inhibited by T3 (32). In contrast to these two atypical TREs, the sequence and function of TRE_{SVL} described herein better resembles that of mammalian gene TREs, as will be discussed below.

The positive regulation by T3 of gene expression mediated by TRE-bound TR can be described as a stepwise process. In the absence of T3, TR binding to a regular positive TRE represses basal promoter activity. Addition of T3 leads first to promoter derepression, followed by a further stimulation well over basal levels of target gene activity. At the molecular level, the

mechanisms involved are thought to operate as follows. Repression by TR in the absence of hormone implies that TR heterodimerizes with its RXR partner and binds to the TRE, as demonstrated both *in vitro* and *in vivo* (3). Upon binding to the TRE, the TR may directly interact with TFIIB (33) or with TBP (34) to repress RNA polymerase II activity. Alternatively, a newly discovered mechanism involves the interaction of TR with silencing factors such as N-CoR and SMRT (35-37). Addition of T3 first releases the interaction of these silencing factors with TR. Additionally, T3 allows interaction of TR with co-activators such as SRC-1 and CBP/P300 (38,39). It is then tempting to match derepression with T3-induced release of repressors and further induction of target gene expression with T3-induced interaction of TR with co-activators.

With respect to the patterns of T3-mediated regulation of gene expression, the TRE_{SVL} that we have characterized is located at the transcription initiation site but, nonetheless, does not drive T3-dependent inhibition of SVL promoter activity, as has been described for some TREs found in close proximity to the transcription initiation sites they control (reviewed in 4). The observed pattern rather corresponds to an altered positive regulation in which the addition of T3 would lead to relief of repression mediated by TR, possibly by releasing interaction with a co-repressor, while the second step of the activation process, i.e. interaction with a co-activator, would not take place or not lead to an efficient TR-co-activator complex. When removed from its native context and inserted into the heterologous TK promoter, TRE_{SVL}, which is arranged as a DR4 element, behaves strictly as a classical positive TRE, implying that the TRE_{SVL} structure *per se* allows the molecular steps described above, i.e. repression, derepression and stimulation. Therefore, the position of the TRE at the transcription initiation site of the SVL promoter may account for the lack of stimulation after derepression.

Our demonstration that transcription starts at the same position in the SVL promoter regardless of the presence or absence of TR and T3 raised the question of how TR-RXR and the transcription initiation machinery co-localize to the same sequence. In contrast to most of the transcription initiation sites of promoters transcribed by RNA polymerase II, the SVL cap site is required for basal promoter activity (14), suggesting that this site may bind both a transcription factor and the RNA polymerase II complex at the same time. Interestingly, this sequence also corresponds to one of the sites that bind the cellular transcriptional repressors which Wiley *et al.* (15) purified from HeLa cell nuclear extracts. The active fraction is mainly composed of three proteins, called initiator binding proteins (IBPs), which bind with high affinity to discrete sites, surrounding and located at the major transcription initiation site of the SVL (see Fig. 8). The human estrogen-related receptor (hERR1) is one major IBP component (15), while testicular receptor 2 and testicular receptor 4, as well as COUP-TFI and COUP-TFII, can interact with the sequence centred at nt 380 (23,40) and for the two latter factors, to a lesser extent to the binding site centred at nt 325 (41). In infected cells, titration of these cellular factors may participate in the SV40 early to late switch (15). Interestingly, all these factors are orphan nuclear receptors and members of the steroid-thyroid hormone receptor family. The importance of TR-mediated repression in this context of multiple interactions still remains to be elucidated.

In summary, we have demonstrated that the major late transcription initiation site of the SV40 genome, previously characterized as a binding site for transcriptional repressors, is a potent TRE. The presence of TR leads to strong repression of

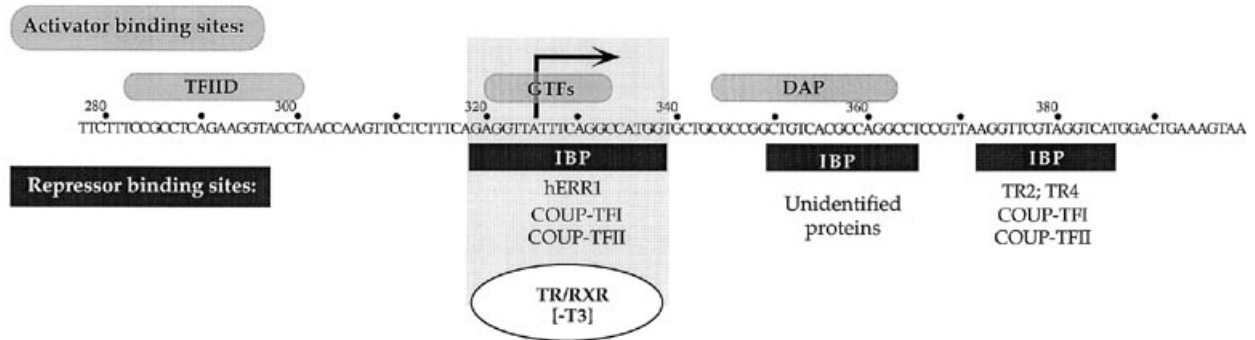


Figure 8. Compilation of the factors binding to the region surrounding the SVL promoter transcription initiation site. Light grey boxes, core promoter elements that bind factors stimulating transcription as identified by Ayer and Dynan (14,42); GTFs, general transcription factors; DAP, downstream activator protein; dark grey boxes, repressor binding sites interacting with IBPs. hERR1 has been identified as one of the IBPs, while COUP-TFI, COUP-TFII, TR2 and TR4 also bind to the indicated sites (15,23,40,41) and repress transcription.

SVL promoter activity, but this repression can be relieved by addition of T3. Both TR-mediated repression and derepression do not affect the site of transcription initiation nor splicing of the VP1 intron. These results raise the question, not yet resolved, of how does the transcription machinery overcome cluttering at the initiation site. Finally, our results demonstrate that one should be cautious when using SVL in expression vectors designed for the study of hormone responsiveness.

ACKNOWLEDGEMENTS

We wish to thank Daniel Bachman for technical help and Dr Richard-Foy and Abdelmadjid Hiji for the gift of pSVL-PL and RXR-containing Sf9 cellular extracts respectively. We also thank Dr W.Wahli for critical suggestions during the preparation of the manuscript and Dr E.Beale for careful reading of the manuscript.

REFERENCES

- Gronemeyer,H. and Laudet,V. (1995) *Protein Profile*, **2**, 1173–1308.
- Desvergne,B. (1994) *Mol. Cell. Endocrinol.*, **100**, 125–131.
- Wong,J., Shi,Y.-B. and Wolffe,A.P. (1995) *Genes Dev.*, **9**, 2696–2711.
- Williams,G.R. and Brent,G.A. (1995) In Weintraub,B.D. (ed.), *Molecular Endocrinology: Basic Concepts and Clinical Correlations*. Raven Press, New York, NY, pp. 217–239.
- Jones,N.C., Rigby,P.W.J. and Ziff,E.B. (1988) *Genes Dev.*, **2**, 267–281.
- Martinez,E. and Wahli,W. (1991) In Parker,M.G. (ed.), *Nuclear Hormone Receptors*. Harcourt Brace Jovanovitch, London, UK, pp. 125–153.
- Serfling,E., Jasin,M. and Schaffner,W. (1985) *Trends Genet.*, **1**, 224–230.
- Maniatis,T., Goodbourn,S. and Fischer,J.A. (1987) *Science*, **236**, 1237–1244.
- Buchman,A.R., Burnett,L. and Berg,P. (1981) In Tooze,J. (ed.), *DNA Tumor Viruses*. Cold Spring Harbor Laboratory, USA, pp. 799–841.
- Somasekhar,M.B. and Mertz,J.E. (1985) *J. Virol.*, **56**, 1002–1013.
- Brady,J. and Khoury,G. (1985) *Mol. Cell. Biol.*, **5**, 1391–1399.
- Keller,J.M. and Alwine,J.C. (1985) *Mol. Cell. Biol.*, **5**, 1859–1869.
- Wiley,S.R., Kraus,R.J. and Mertz,J.E. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 5814–5818.
- Ayer,D.E. and Dynan,W. (1988) *Mol. Cell. Biol.*, **8**, 2021–2033.
- Wiley,S.R., Kraus,R.J., Zuo,F.G., Murray,E.E., Loritz,K. and Mertz,J.E. (1993) *Genes Dev.*, **7**, 2206–2219.
- Luckow,B. and Schütz,G. (1987) *Nucleic Acids Res.*, **15**, 5490.
- Desvergne,B., Petty,K.J. and Nikodem,V.M. (1991) *J. Biol. Chem.*, **266**, 1008–1013.
- Samuels,H.H., Stanley,F. and Casanova,J. (1979) *Endocrinology*, **105**, 80–85.
- Seed,B. and Sheen,J. (1988) *Gene (Amst.)*, **67**, 271–277.
- Schreiber,E., Matthias,P., Müller,M.M. and Schaffner,W. (1989) *Nucleic Acids Res.*, **17**, 6419.
- Keller,H., Dreyer,C., Medin,J., Mahfoudi,A., Ozato,K. and Wahli,W. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 2160–2164.
- Lazar,M.A. (1993) *Endocrinol. Rev.*, **14**, 184–193.
- Lee,H.-J. and Chang,C. (1995) *J. Biol. Chem.*, **270**, 5434–5440.
- Aniello,F., Couchie,D., Bridoux,A.-M., Gripois,D. and Nunez,J. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4035–4039.
- Tillman,J.B., Crone,D.E., Kim,H.-S., Sprung,C.N. and Spindler,S.R. (1993) *Mol. Cell. Endocrinol.*, **95**, 101–109.
- Park,H.-Y., Davidson,D., Raaka,B.M. and Samuels,H.H. (1993) *Mol. Endocrinol.*, **7**, 319–330.
- Damm,K., Thompson,C.C. and Evans,R.M. (1989) *Nature*, **339**, 593–597.
- Farsetti,A., Desvergne,B., Hallenbeck,P., Robbins,J. and Nikodem,V.M. (1992) *J. Biol. Chem.*, **267**, 15784–15788.
- Glass,C.K., Holloway,J.M., Devary,O.V. and Rosenfeld,M.G. (1988) *Cell*, **54**, 313–323.
- Glass,C.K., Lipkin,S.M., Devary,O.V. and Rosenfeld,M.G. (1989) *Cell*, **59**, 698–708.
- 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.
- Saatcioglu,F., Deng,T.L. and Karin,M. (1993) *Cell*, **75**, 1095–1105.
- Baniahmad,A., Ha,I., Reinberg,D., Tsai,S., Tsai,M.-J. and O'Malley,B.W. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 8832–8836.
- Fondell,J.D., Roy,A.L. and Roeder,R.G. (1993) *Genes Dev.*, **7**, 1400–1410.
- Horlein,A.J., Näär,A.M., Heinzel,T., Torchia,J., Gloss,B., Kurokawa,R., Ryan,A., Kamel,Y., Soderstrom,M., Glass,C.K. and Rosenfeld,M.G. (1995) *Nature*, **377**, 397–404.
- Kurokawa,R., Soderstrom,M., Horlein,A., Halachmi,S., Brown,M., Rosenfeld,M.G. and Glass,C.K. (1995) *Nature*, **377**, 451–454.
- Chen,J.D. and Evans,R.M. (1995) *Nature*, **377**, 454–457.
- Onate,S.A., Tsai,S.Y., Tsai,M.-J. and O'Malley,B.W. (1995) *Science*, **270**, 1354–1357.
- Kamei,Y., Xu,L., Heinzel,T., Torchia,J., Kurokawa,R., Gloss,B., Lin,S.-C., Heyman,R.A., Rose,D.W., Glass,C.K. and Rosenfeld,M.G. (1996) *Cell*, **85**, 403–414.
- Lee,H.-J., Lee,Y., Burbach,J.P.H. and Chang,C. (1995) *J. Biol. Chem.*, **270**, 30129–30133.
- Zuo,F. and Mertz,J.E. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 8586–8590.
- Ayer,D.E. and Dynan,W.S. (1990) *Mol. Cell. Biol.*, **10**, 3635–3645.