Characterization of a retinoic acid responsive element isolated by whole genome PCR

M.Patricia Costa-Giomi, Marie-Pierre Gaub1, Pierre Chambon1 and Patricio Abarzua*
Department of Oncology, Roche Research Center, Hoffmann-LaRoche Inc., Nutley, NJ 07110, USA
and 1Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, U184 de Biologie Moléculaire
et de Génie Génétique de l'INSERM, 67085 Strasbourg, France

Received November 27, 1991; Revised and Accepted May 12, 1992

ABSTRACT
We have used whole genome PCR in an attempt to isolate novel retinoic acid (RA) responsive genes. We cloned several small genomic fragments from total human DNA containing putative retinoic acid responsive elements (RAREs) selected by direct binding to the retinoic acid receptor α (RARα). We report here that an oligonucleotide containing a sequence from one of the cloned human DNA fragments, and referred to as a1, functions as an authentic RARE. It is shown that both RARα and RARβ produced in Cos cells as well as in vitro translated RARα bind directly and sequence-specifically to the a1RARE. By mutational analysis it is demonstrated that the a1RARE consists of an imperfect direct repeat of the estrogen- and thyroid hormone-related AGGTCA half-site motif separated by a 5 bp spacer. The orientation and spacing of the half-site repeats are shown to play a critical role in RAR recognition. When cloned upstream of a TK-Luc reporter, the a1RARE is shown to confer responsiveness to RA in an orientation-independent fashion in F9 and CV-1 cells. The magnitude of the RA response mediated by the a1RARE differed in these cell lines.

INTRODUCTION
Retinoids, which include natural and synthetic derivatives of vitamin A (retinol), have been shown to have profound effects in a variety of biological processes including cell differentiation, embryogenesis and neoplasia. Retinoic acid (RA), one of the most potent natural retinoids, can dramatically change the biological and biochemical properties of normal and malignant cells in tissue culture, sometimes in opposite ways, depending on concentration and cell type (1). Retinoids are also potent inhibitors of chemical carcinogenesis in rodent models (2). During embryogenesis, RA has been found to be a very important endogenous signal in the development and regeneration of the limb and other structures in vertebrates (3). In rodents and human, RA is teratogenic causing a variety of cranio-facial and limb malformations as well as defects in brain development (4 – 7). Target tissues show unique temporal patterns of RA sensitivity that coincide with organogenesis (5).

The pleiotropic effects of retinoids have been thought to be mediated at the gene transcription level (8). Several cDNA clones representing transcripts that increase or decrease in abundance in cells treated with RA have been isolated (9 – 13), however a direct link between transcriptional regulation and RA was lacking until the recent identification of the nuclear RA receptors (RARs) (14,15). Molecular cloning and sequence analysis have led to the identification of three RAR subtypes (RARα, RARβ and RARγ) encoded by three independent genes (16 – 18). In addition, multiple mRNA isoforms have also been identified for each receptor subtype (19 – 21). Recently, a fourth retinoid nuclear receptor (RXR-α) has been cloned and shown to be distinct from the RARs (22). All these retinoid receptors belong to the steroid/thyroid hormone nuclear receptor superfamily (23 – 25). Members of this family function as hormone-inducible transcription factors and they are thought to bind as dimers to target sequences called hormone responsive elements (HRE). RARs possess a domain structure comprising six regions, designated A to F (24,25). The amino acid domains responsible for DNA and ligand binding are located in regions C and E, respectively. The C domain of the receptors includes two zinc-finger motifs believed to mediate DNA binding (26 – 28). The high level of amino acid homology among the RARs in the C region suggests that all three receptors are likely to recognize identical target sequences and, thus, may regulate identical sets of genes.

One of the ‘early’ response genes shown to be transcriptionally activated by RA is the RAR-β gene itself (29,30). The promoter region was cloned and the sequences responsible for RA inducibility identified as an imperfect direct repeat of the sequence (A/G)GTCCA (31,32). The RAREβ half-site repeat thus differs by a single non-conservative nucleotide change from the consensus TRE/ERE motif. However, the spacing and orientation of the half-sites appears to be unique. Recently, characterization of RAREs from the mouse laminin B1 (lamB1, refs. 33 and 34), mouse complement factor H (cfl, ref 35), human alcohol
dehydrogenase (adh3, ref 36) and phosphoenolpyruvate carboxykinase (pepck, ref 37) genes have all indicated that direct rather than inverted repeats are responsible for RA responsiveness. Moreover, in every case, the hexameric repeat is closely related to the AGGTCA motif. Despite these similarities, the spacing and number of half-sites vary among these natural RAREs. Whether these structural differences affect the affinity or the mechanism of binding of RARs to these elements is not known. It is also possible that receptor specificity may be determined by some of these structural variations. Identification and characterization of more RA-responsive genes should help to define what constitutes a consensus RARE and how this element differs from other HREs and it should also enhance our knowledge of how retinoids work.

Although cDNAs for genes up- or down-regulated by RA have been cloned (9–13), 'primary' response genes [defined as those whose transcription rates are increased or decreased in a protein synthesis-independent fashion by RA, examples are the RARβ and ERA-1 (Hox-1.6) genes] are the least represented. This paucity in known 'primary' RA-responsive genes is due to inherent limitations of screening large representiative cDNA libraries by conventional hybridization techniques (38,39). For example the RARE gene itself was never detected by differential screening of cDNA libraries from RA-treated cells (9–13). To overcome this problem, we have chosen a different approach to isolate RA-responsive genes based on cloning of discrete genomic fragments of DNA containing RAREs selected by direct binding to RARs. Then these fragments can be used as probes to isolate the RA-target genes. This technique, referred to as whole genome PCR (40), has several advantages compared to more conventional procedures. In principle, any cell or tissue could serve as a source of DNA, independent of whether or not RA-responsive genes are expressed in that particular cellular milieu. Furthermore, this approach should lead to the isolation of novel RA-regulated genes and also increase our understanding of RAREs. As a first step toward this goal, we report here the isolation and characterization of a novel human RARE from total genomic DNA. Our studies suggest that this approach may prove useful to identify genes regulated by retinoids.

**MATERIALS AND METHODS**

**Whole genome PCR**

High molecular weight DNA from human placenta (Clontech) was sonicated on ice to an average size of 150–250 bp and made blunt-ended by incubation with T4 DNA polymerase (4 units, New England Biolabs) and all four dNTPs. DNA was modified to a form appropriate for PCR amplification by ligation to 'catch-linkers' as previously described (40,41). To select for DNA sequences which bind to RAα, linker-modified DNA (0.5 μg) was incubated with nuclear extracts from Cos cells transfected with an RARα expression vector (see below) in 20 μl of DNA binding buffer (50 mM Hepes [N-2-Hydroxyethylpiperazine-N'-2-ethanesulfonic acid], pH 7.5, 50 mM KCl, 5 mM MgCl2, 10 μM ZnSO4, 1 mM dithiothreitol, 20% glycerol) containing 10 nM RA and 2 μg poly (dl-dC) (Pharmacia) for 30 min at room temperature. After DNA binding, 200 μl of immunoprecipitation buffer (40) and 5 μl of rabbit anti-RAα antisera SP171 (42) were added and incubation continued for 1 hr. Immunocomplexes were bound to protein-A-Sepharose (Sigma) for 1 hr at room temperature, recovered by centrifugation in a microfuge and washed three times with 1.5 ml of immunoprecipitation buffer. DNA was released from the complex, extracted with phenol:chloroform (1:1), and ethanol precipitated. Recovered DNA was amplified by 25 cycles of PCR with 'catch-linkers' as primers, and then used for additional rounds of selection.

After two and four rounds of binding and amplification, libraries were constructed as previously described (40,41). Fragments that were enriched by the selection procedure were identified by colony hybridization (43) with a probe prepared from the amplified and selected DNA (10–50 ng) by random primer labelling with [α-32P]dCTP. Highly repetitive sequences were removed by pre-annealing the probe to 1 mg/ml sheared, heat-denatured human placental DNA in 5×SSC (1×SSC is 0.15 M NaCl, 0.015 M Na Citrate) for 15 min at 68°C.

**Plasmid constructions**

All plasmid constructions were performed by standard cloning techniques (44). The RARα (hRARα0) and RARβ (hRARβ0) expression vectors were previously described (16,17). The RARγ and VDR expression vectors were a gift of Dr. P. LeMotte and Dr. W. Hunzinker (F. Hoffmann-LaRoche Ltd., Basle) respectively. The RXRα expression vector was a gift of Dr. J. Grippio (Hoffmann-LaRoche Inc., Nutley).

The reporter plasmid pTK-Luc was constructed by subcloning a BamHI fragment from plasmid ptkluc (45) into the BamHI site of the blueprints KS(+) plasmid. This fragment contains the HSV thymidine kinase promoter, positions −109 to +50, the firefly luciferase gene and SV40 splicing and polyadenylation signals. The α1 (113 bp) and α13 (167 bp) fragments were isolated by EcoRI digestion, made blunt by fill in with the Klenow fragment of *E. coli* DNA polymerase I, and cloned upstream of the TK promoter into a unique Smal site present in the multiple cloning site of blueprints KS (+) to generate the α1/TK-Luc and α13/TK-Luc reporters, respectively. To construct the α1RARE/TK-Luc reporter, a unique BglIII cloning site was created by inserting a BglII linker into the Smal site of pTK-Luc. A 73 bp synthetic fragment containing three tandem copies of the α1 sequence (Fig 1) and flanked by BamHI and BglII recessive ends was then ligated to BglII-digested pTK-Luc. The structure and orientation of the α1RARE/TK-Luc reporter was verified by restriction enzyme mapping and nucleotide sequencing. All plasmids were maintained in *E. coli* DH5α and prepared by banding twice in CsCl gradients.

**Synthetic oligonucleotides**

All oligonucleotides were synthesized on an Applied Biosystem DNA synthesizer, deblocked and purified by denaturing polyacrylamide gel electrophoresis. Two 33-base oligonucleotides were synthesized for each of the α1, 01 and TRE-GH probes. The sense and antisense strands of the α1 probe consisted of the sequences 5'-AAATGGATCCTCAAGGCGAGAAAGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAC-3' and 5'-AGGGAAGATCTTCTGTCCAGTCTTCTTCC-3', respectively. When these oligonucleotides are annealed, a 10 base 5' overhang is generated at each end. The oligonucleotides used to prepare the 01 and TRE-GH probes were synthesized with the same overhangs and the al sequence was replaced with information from the human RARβ gene (−59 to −33) and the rat growth hormone gene (−186 to −158), respectively (31,46). 32P-labelled probes were prepared by filling in the 5' overhangs with dATP, dTTP, dGTP, and [α-32P]dCTP using the Klenow fragment of *E. coli* DNA polymerase I. Oligonucleotides used as competitors were prepared in a similar manner except that unlabelled dCTP was added.
Oligonucleotides containing mutations in the α1 sequence were based on a 43-base oligonucleotide with the sequence 5'-GTGGATCCTTACAGGGCA-GGAGAAGTCAAGGAGA-GATCTTCTTAA-3' containing the same 23 bases of α1 information as above. A single oligonucleotide was synthesized for each mutant and annealed to a 14 base primer of the sequence 5'-TTAGGAGATCTTC-3'. Full-length mutant oligonucleotides used as competitors were made by extending the primer as described above.

Tissue culture, transfections, and luciferase assays
F9, CV1 or HeLa cells were plated at a density of 5×10⁴ cells per 100 mm dish in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal calf serum (FCS). F9 cells were cultured on gelatin-coated dishes as described previously (47). The following day, the medium was replaced with DMEM containing 10% charcoal-treated FCS and cells were transfected with 5 μg of luciferase reporter plasmid, 1 μg of receptor expression vector, 10 μg of the β-galactosidase expression vector pCH110 (Pharmacia) as an internal control for transfection efficiency and 14 μg of bluescript plasmid as carrier. Transfections were performed as described by Chen and Okayama (48) for 16–17 h at 37°C in 5% CO₂. After transfection, cells were washed with PBS and refed with DMEM plus 10% charcoal-treated FCS. All-trans-RA (Hoffmann-LaRoche Inc.), dissolved in 100% ethanol, was administered 4 h later to a final concentration of 1 μM. Untreated controls received only the vehicle. Cells were harvested 24 h after RA addition and luciferase (49) and β-galactosidase (50) activities were measured as previously described. Protein concentration was determined by the Bradford procedure (51). For each experiment, luciferase activity was normalized to that of β-galactosidase activity. All transfections were independently repeated at least three times unless otherwise noted.

Preparation of nuclear extracts
Cos cells were transfected with 7 μg of RARα or RARβ expression vectors by the DEAE-dextran procedure (44). After incubating the transfected Cos cells for another 48 h, the cells were harvested and nuclear extracts were prepared according to Lec et al (52). Briefly, cells were resuspended in one packed cell volume of buffer A (53) and allowed to swell on ice for 15 min. Cells were lysed by passing the suspension 13 times through a 1 ml syringe attached to a 26 1/2 gauge needle. The cell homogenate was centrifuged for 20 seconds in a microfuge and the nuclear pellet was resuspended in two thirds of the initial packed cell volume of buffer C (53) containing 0.4 M NaCl. Nuclei were extracted with stirring for 30 min on ice. Nuclear debris was removed by centrifugation for 5 min in a microfuge and the soluble nuclear extract was aliquoted, quick-frozen and stored at −80°C.

In vitro transcription and translation
RARα and RARβ expression vectors were linearized by BamHI restriction enzyme digestion and capped mRNA transcripts prepared using T7 RNA polymerase (Stratagene). These mRNAs were then used to program translation in a rabbit reticulocyte lysate system as recommended by the manufacturer (Promega). For gel retardation assays, 5 μl of the lysate was added to the binding reaction.

**Gel retardation assays**
Cos cells nuclear extracts (9–15μg) were preincubated in the presence of 1 μg of poly (dl-dC) in 20 μl of binding buffer (47 mM Hepes, pH 7.9, 50 mM KCl, 5 mM MgCl₂, 1 mM DTT, 0.05% NP40 and 10% glycerol) for 15 min at room temperature. Subsequently, 0.25 ng of the radiolabelled probe was added and the incubation was continued for 45 min. For the competition studies, a 50-fold molar excess of unlabeled competitor oligonucleotide was added to the reaction after the preincubation step and the mixture further incubated for 15 min before addition of the radiolabelled probe. When antibodies and peptides were used in the assay, they were preincubated with nuclear extracts in binding buffer for 1 h at 4°C prior to the addition of the probe. Protein/DNA complexes were analyzed by electrophoresis through a nondenaturating 6% polyacrylamide gel (80:1 acrylamide:bisacrylamide) prepared and electrophoresed in 1×TBE (0.1 M Tris, 0.1 M boric acid, and 1 mM EDTA, pH 8.3) for 2 h at 200 volts. Gels were pre-run for 30 min at room temperature.

**RESULTS**
Isolation of RARα binding sites by whole genome PCR
To isolate RARE sequences from total human DNA we followed the whole genome PCR procedure described by Kinzler and Vogelstein (40). In this technique, genomic DNA is first modified to a form that can be amplified by the polymerase chain reaction. Specific DNA sequences are then selected by binding to a regulatory protein followed by immunoprecipitation of the protein/DNA complex. Because DNA fragments recovered from the immunoprecipitation step are amplified by PCR, selection can be performed under stringent conditions with little concern about yield. Accordingly, sheared, linker-modified human DNA was incubated with nuclear extracts from RAR-α transfected Cos cells and receptor/DNA complexes were immunoprecipitated with rabbit polyclonal antibody SP171 (42). DNA was recovered, PCR amplified, and then used for further rounds of selection and amplification. After two and four of these cycles, DNA was digested with EcoRI and cloned. Colonies were screened for enriched sequences by hybridization to the same selected and amplified labelled DNA. Under these conditions, only colonies containing sequences whose abundance is 0.05% or higher of the total sequence complexity would be expected to score as positive (38,39).

After two cycles of selection and amplification less than 0.1% of the colonies were detectable following hybridization to the probe, and this required prolonged exposure to X-ray film. After four cycles, close to 100% of the colonies hybridized to the probe. Restriction enzyme analysis of plasmid DNA isolated from positive colonies selected at random showed about 30% to contain distinct fragment inserts. The remaining contained only linker sequences, suggesting a PCR artifact.

Ten independently-isolated positive clones containing inserts (average size 144 bp, range 76 to 291 bp) were sequenced by the dideoxy procedure and found to represent nine unique sequences. Extensive sequence analysis revealed most of the clones to contain only small stretches of homology among themselves and to previously described HREs (31,46,54). The most prevalent region of homology contained variants of the AGGTC motif found in EREs and TREs. Preliminary evidence suggesting that at least some of the fragments contained RAR
Figure 1. Sequence of a series of oligonucleotides containing putative RARα binding sites. Human DNA fragments isolated by whole genome PCR were sequenced and the regions showing significant homology were directly aligned. The sequences with the highest identity were then selected and double-stranded oligonucleotides synthesized. All oligonucleotides contained identical 10 bp overhangs for labelling and they are not shown. A 1 bp gap was introduced into α18 for optimal alignment. The AGGTCA-related motifs are shown within the boxes. Arrows indicate the orientation of the motifs.

Figure 2. In vitro binding of RARα and RARβ to oligonucleotide α1. 10 μg of nuclear extracts from mock-transfected Cos cells (M; lanes 1, 4, 7 and 10), or Cos cells transiently transfected with RARα (α; lanes 2, 5 and 8) or RARβ (β; lanes 3, 6, 9 and 12) were incubated with [32P]-labelled α1 oligonucleotide. For competition experiments, a 50-fold molar excess of unlabelled α1 oligonucleotide or RARE/3 (RARE/3; ref 31) almost completely abolished the formation of the shifted complex (lanes 4—6, Fig 2), suggesting that RARs are not involved in the formation of this complex. Identical results were obtained in competition assays (Fig 2, lane 2 and 3). In contrast, complex 2 (C2, Fig 2) was detected when both mock- and RAR-transfected nuclear extracts were incubated with the α1 probe. Complex 1 (C1, Fig 2) was only present when nuclear extracts from either RARα- or RARβ-transfected Cos cells were used in the binding reaction (Fig 2, lane 2 and 3). In contrast, complex 2 (C2, Fig 2) was detected when both mock- and RAR-transfected nuclear extracts were incubated with the α1 probe. Sequence-specific binding of the RARs to the α1 probe was performed in the absence of antibody (—: panel A, lanes 1 and 2; panel B, lanes 1 and 2) or in the presence of non-immune rabbit serum (C: panel A, lanes 3 and 4; panel B, lanes 3 and 4), anti-RARα antibody SP171 (α: panel A, lanes 5 to 8) or anti-RARβ antibody SP172 (β: panel B, lanes 5 to 10). Nuclear extracts from mock-transfected Cos cells (panel A; lanes 1, 3, 5 and 7, panel B; lanes 1, 3, 5, 7 and 9) or Cos cells transfected with RARα (α: panel A, lanes 2, 4, 6 and 8) or RARβ (β: panel B, lanes 2, 4, 6, 8 and 10) were pre-incubated with antibodies for 60 min at 4°C prior to the addition of the radiolabelled probe. To show the specificity of the antibodies, a RARα peptide (panel A; lanes 7 and 8; panel B; lanes 9 and 10) or a RARβ peptide (panel B; lanes 7 and 8) containing the SP171 and SP172 epitopes respectively were incubated for 60 min at 4°C with the respective anti-serum before addition to the binding reaction. The position of the specific complexes C1 and C2, and that of the shifted complex are indicated by arrows.

Figure 3. Characterization of the protein/DNA complexes formed with the α1 probe with anti-RAR polyclonal antibodies. Binding of nuclear extracts to the α1 probe was performed in the absence of antibody (—: panel A, lanes 1 and 2; panel B, lanes 1 and 2) or in the presence of non-immune rabbit serum (C: panel A, lanes 3 and 4; panel B, lanes 3 and 4), anti-RARα antibody SP171 (α: panel A, lanes 5 to 8) or anti-RARβ antibody SP172 (β: panel B, lanes 5 to 10). Nuclear extracts from mock-transfected Cos cells (panel A; lanes 1, 3, 5 and 7, panel B; lanes 1, 3, 5, 7 and 9) or Cos cells transfected with RARα (α: panel A, lanes 2, 4, 6 and 8) or RARβ (β: panel B, lanes 2, 4, 6, 8 and 10) were pre-incubated with antibodies for 60 min at 4°C prior to the addition of the radiolabelled probe. To show the specificity of the antibodies, a RARα peptide (panel A; lanes 7 and 8; panel B; lanes 9 and 10) or a RARβ peptide (panel B; lanes 7 and 8) containing the SP171 and SP172 epitopes respectively were incubated for 60 min at 4°C with the respective anti-serum before addition to the binding reaction. The position of the specific complexes C1 and C2, and that of the shifted complex are indicated by arrows.

putative sites as found in five of the cloned fragments. The spatial organization and sequence composition showed a degree of resemblance to the RARE found in the RAR-β gene (31,32). We show below by direct binding and functional assays that the α1 oligonucleotide functions as an authentic RARE.

Binding of RAR-α and RAR-β to α1 oligonucleotide
To test if RARs would bind to any of the sequences shown in Fig 1, double stranded oligonucleotides were synthesized, end-labelled with [α-32P]dCTP, and incubated with nuclear extracts prepared from Cos cells transiently transfected with RAR expression vectors. Protein/DNA complexes were resolved by gel retardation assays and detected by autoradiography. Under these experimental conditions only the α1 oligonucleotide yielded specific bands. Two major protein/DNA complexes could be detected by this assay with the α1 probe. Complex 1 (C1, Fig 2) was only present when nuclear extracts from either RARα- or RARβ-transfected Cos cells were used in the binding reaction (Fig 2, lane 2 and 3). In contrast, complex 2 (C2, Fig 2) was detected when both mock- and RAR-transfected nuclear extracts were incubated with the α1 probe. Sequence-specific binding of the RARs to the α1 probe was demonstrated in competition assays (Fig 2 lanes 4—12). Addition of a 50-fold molar excess of unlabelled α1 or 01 oligonucleotide (RARE/3, ref 31) almost completely abolished the formation of C1, whereas an oligonucleotide containing a TRE from the rat growth hormone gene failed to compete under these conditions. On the other hand, formation of C2 was only competed by α1 (lanes 4—6, Fig 2), suggesting that RARs are not involved in the formation of this complex. Identical results were obtained...
when the $\alpha 1$ fragment was used as the probe. However, the intensity of the RAR-specific complex was reduced (data not shown).

Further evidence for the involvement of RARs in the formation of $C1$, but not $C2$, was obtained from retardation assays in the presence of RAR-specific antibodies (Fig 3A and 3B). Incubation of binding reactions with anti-serum SP171, but not non-immune serum, led to the appearance of a higher molecular weight complex with a concomitant decrease in the intensity of $C1$ when RAR$\alpha$-transfected Cos cell nuclear extracts were the source of receptor. Moreover, an RAR$\alpha$-specific peptide containing the epitope recognized by the SP171 anti-serum blocked the formation of this large complex (compare lanes 6 and 8 in Fig 3A). Similarly, incubation of RAR$\beta$-transfected Cos cell nuclear extracts with rabbit polyclonal antibody SP172 raised against a RAR$\beta$ peptide (42) also led to the formation of a slowly migrating complex with an accompanying decrease in the intensity of $C1$ (Fig 3B lane 6). Incubation with the RAR$\beta$ peptide, but not an RAR$\alpha$ peptide, blocked the shift to the high molecular weight form of the complex (compare lanes 8 and 10 with lane 6 in Fig 3B). In contrast, neither antibody affected the intensity or migration of $C2$. These results demonstrate that both RAR$\alpha$ and RAR$\beta$ bind with high affinity to the $\alpha 1$ oligonucleotide. Furthermore, the presence of a second specific protein/DNA complex also generated by the mock-transfected Cos cell nuclear extracts suggests the existence of other nuclear factor(s) that bind to the same element as the RARs. It is unlikely that endogenous monkey (Cos cells) RARs are responsible for $C2$ since the 01 oligonucleotide failed to compete, even though it is a highly efficient RARE activated by endogenous RARs in human, monkey and mouse cells (31,32).

To investigate whether the formation of the complex between RARs and the $\alpha 1$ probe requires additional factor(s) present in nuclear extracts, in vitro transcribed RAR$\alpha$ mRNA was used to program a rabbit reticulocyte lysate and the in vitro translated receptor was then tested in gel retardation assays for binding to the $\alpha 1$ oligonucleotide. As shown in Fig 4, a protein/DNA complex could be detected only when the $\alpha 1$ probe was incubated with reticulocyte lysate programmed with the RARE$\alpha$ mRNA. Addition of polyclonal antibody SP171 to the binding reaction induced a shift in this band toward a higher molecular weight form, demonstrating the presence of the receptor in this complex (Fig 4, lane 6). This antibody-associated shift was not observed when non-immune serum was added to the binding reaction (Fig 4, lane 4). Similar results were obtained for RAR$\beta$ (data not shown). These data demonstrate that RARs can bind directly to the $\alpha 1$ probe. Note that the addition of the SP171 antiserum appears to increase the intensity of the complex formed between in vitro translated RAR$\alpha$ and the $\alpha 1$ probe (compare lane 6 with lanes 2 and 4 in Fig 4). Likewise, we have observed that addition of nuclear extracts from untransfected Cos cells, but not bovine serum albumin, increases the amount of specific complex at least 5-fold, suggesting that accessory nuclear factor(s), although not required for binding, may either increase the stability of the RAR/DNA complex or facilitate dimer formation. These nuclear factor(s) may be limiting since a 5-fold dilution of the nuclear extracts leads to a similar decrease in the amount of specific complex (55). This is in agreement with published reports demonstrating the existence of multiple cell type-specific factors that increase the binding affinity of the RARs for response elements (56). Based on the high affinity binding of the RARs to the $\alpha 1$ oligonucleotide we conditionally designated it as $\alpha 1$RARE.

We failed to observe specific RAR/DNA complexes with the other oligonucleotides shown in Fig 1, however, such a result can occur if the gel retardation assay is not sensitive enough to detect weaker, but significant interactions between RARs and some DNA sequences. In particular, the $\alpha 24$ oligonucleotide shares 62% overall sequence homology with the RARE from the RAR$\beta$ gene (see Fig 9A). Moreover, its 3'-end matches perfectly at 9 out of 12 positions to a shorter but still active version of the RARE$\beta$ (32), yet we could not detect a RAR-specific complex. Consequently, we tested these oligonucleotides for RAR
would have a more dramatic effect upon the ability of the mutant oligonucleotide to compete for the formation of the specific complex, and indeed that was the case. Any mutation affecting the motif AGGTCA (e.g. m2, m3, m4 and m5) had profound effects upon the ability of the mutant DNAs to compete against the α1 probe for RAR binding (Fig 6A). By contrast, either a single (m8) or a double (m7) base change immediately following this hexameric element did not appreciably affect competition, suggesting that these nucleotide positions are not crucial for sequence-specific binding by the RARs. Similarly, m5 and m6 were equally ineffective in competing for complex 1, suggesting that the adenosine residue immediately preceding the hexameric motif is not involved in recognition. Mutations within the region separating the hexameric half-sites either have no effect (m13) or slightly decreased the ability of the oligonucleotide to compete (m9). Altogether, these results delineate the two AGGTCA-related motifs as critical for sequence-specific RAR binding. Note that mutant competition data for complex 2 follow the same general pattern as for the RAR specific complex except for mutant 2. In m2, the T to G transversion at DR2 only marginally affected competition for complex 2 while still impairing its ability to compete for complex 1. This observation suggests that the protein component(s) involved in formation of complex 2 may recognize related but not identical sequences.

The half-site repeats in the α1RARE as defined above are spaced by a five nucleotide long ‘linker’ sequence. In this arrangement, each position in one half-site is separated by 11 bp from its homolog. This locates both half-sites on the same side of the DNA helix as it has also been suggested for the βRARE and the lamB1 RARE (32–34). By analogy to other members of the nuclear receptor family, RARs are believed to bind as dimers to RAREs and the stereo-alignment of the half-sites may serve to orient the two monomers along the DNA. Therefore, alterations of the spacing between the repeats might be expected to be deleterious for RAR binding. Indeed, insertion or deletion of a single nucleotide into the spacer of the α1RARE did abolish the ability of these mutant oligonucleotides to compete for binding to the RAR-specific complex (m15 and m10, Fig 6B). Unexpectedly however, deletion of two or three bases from the spacer region was not detrimental for binding, since m11 and m12 competed as well as the wild-type oligonucleotide for binding to RARα (Fig 6B, lanes 6 and 7). Our results suggest that at least for sequence-specific DNA binding, positioning of the half-sites on the same side of the DNA helix might not be a critical parameter. Most likely, deletion or insertion of a single bp in the spacer region as in m10 and m15 disrupts the spatial geometry between the receptor dimer and the half-sites, since each nucleotide added or removed would shift the half-sites positions by 3.4 Å and impose a 36° rotation around the DNA helix (57). This detrimental effect on receptor binding may be partially (e.g. m11) or totally (e.g. m12) compensated by rotating and moving the half-sites closer together to within a distance where functional interaction with the receptor dimer may take place. Note that the size of the spacer does not impair the ability of the mutations to compete for complex 2, demonstrating that formation of complex 1 and 2 have different structural requirements.

RARs have been shown to activate both direct and inverted repeat responsive elements (31,32,34,35,37,58–60). This promiscuity displayed by the RARs suggested that the orientation of the half-sites might be functionally equivalent between the direct and inverted form. To test this possibility, we synthesized a mutant α1 oligonucleotide (Fig 5C, m14) containing the second

---

**Figure 6.** Sequence and structural requirements for RAR specific binding to the α1 RARE. Double-stranded mutant oligonucleotides shown in Fig 6 were used as competitors at a 50-fold molar excess in binding reactions containing nuclear extracts from RARα-transfected Cos cells and labelled α1 probe. (A) Importance of the nucleotide sequence and involvement of the half-site repeats in RARα recognition of the α1RARE. (B) Effects of spacer sequence (lanes 4 and 5), spacer length (lanes 2, 3, 6, 7 and 8), and orientation of the half-sites (lane 9) for binding of RARα to the α1RARE. Arrows indicate specific complex C1 and C2.

Mutational analysis of the α1RARE

To help define the sequences and structural determinants critical for RAR binding to the α1RARE, we synthesized several mutant oligonucleotides (m1 to m15, shown in Fig 5) and assayed them for their ability to compete for binding to the RARs present in nuclear extracts from RARα-transfected Cos cells (Fig 6). Mutants 1 and 2 both contain a single nucleotide change converting them to a perfect direct repeat of motif 2 (DR2) or 1 (DR1) respectively. In gel retardation assays, m1 competed even more effectively than the wild-type α1 oligonucleotide for binding to the RARα-specific complex (Fig 6A, compare lane 2 and 3). In contrast, m2 was a very poor competitor even at the 50-fold molar excess used in these experiments (Fig 6A, lane 4). Similar quantitative results were obtained when we used a 25- or 100-fold molar excess of competitor or when RARβ extracts were used (data not shown). These dramatic changes induced by single base substitutions suggest that high affinity binding of RARs to the half-sites closer together to within a distance where functional interaction with the receptor dimer may take place. Note that the size of the spacer does not impair the ability of the mutants to compete for complex 2, demonstrating that formation of complex 1 and 2 have different structural requirements.

The α1RARE contains two imperfect direct repeats, each consisting of 9 nucleotides (Fig 1). HRE for other members of the steroid/thyroid receptor superfamily have been shown to contain hexameric half-sites arranged as inverted repeats (26–28,46,54). To investigate the possible hexameric nature of the half-sites in the α1RARE, mutant oligonucleotides were synthesized containing either single or double base substitutions within DR2 (Fig 5). Since we have already shown above that the DR1 sequence AGGGCA is not optimal for binding, we expected that any mutation critical for RAR recognition in DR2 binding in competition assays against the α1 probe to increase the probability of detecting weak but sequence-specific interactions (data not shown). As would be expected from sequence similarity, under these conditions the α2α oligonucleotide was able to compete for binding to the RARα almost as efficiently as the α1 sequence itself, demonstrating that at least two of the nine immunoselected human DNA fragments contain authentic RAR binding sites.

---

**Probe**

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

**Spacers**

- m1
- m2
- m3
- m4
- m5
- m6
- m7
- m8
- m9

**Oligonucleotides**

- ml
- ml
- ml
- ml
- ml
- ml
- ml
- ml

**Figure 6**

Figure 6. Sequence and structural requirements for RAR specific binding to the α1 RARE. Double-stranded mutant oligonucleotides shown in Fig 6 were used as competitors at a 50-fold molar excess in binding reactions containing nuclear extracts from RARα-transfected Cos cells and labelled α1 probe. (A) Importance of the nucleotide sequence and involvement of the half-site repeats in RARα recognition of the α1RARE. (B) Effects of spacer sequence (lanes 4 and 5), spacer length (lanes 2, 3, 6, 7 and 8), and orientation of the half-sites (lane 9) for binding of RARα to the α1RARE. Arrows indicate specific complex C1 and C2.
Simplex Virus Thymidine Kinase (TK) promoter driving the containing three tandem copies of the al sequence (alRARE/TK-Luc) were cloned upstream of the Herpes (a1/TK-Luc) as well as a double stranded oligonucleotide To investigate whether the putative a1 RARE could confer RA responsiveness to a heterologous promoter, the al fragment (a1/TK-Luc) was cloned upstream of the TK promoter. The reporter plasmid a13/TK-Luc contains a human genomic fragment that is not bound by the RARs and it was used as a negative control. The reporter plasmid, TK-Luc, contains only the TK promoter driving the luciferase gene. (B) Reporter plasmids a1RAREy/TK-Luc (F) and a1RAREy/TK-Luc (R) each contain three copies of the a1 RARE (Fig 1) in the forward and reverse orientation, respectively, cloned upstream of the TK promoter. Luciferase reporter plasmids (5 μg) and β-galactosidase expression vector pCH 110 (10 μg) were cotransfected into F9 cells with (+) or without (−) an RARα expression plasmid (1 μg). After transfection, 1 μM RA (+) or vehicle alone (−) were added to the culture media. Cell extracts were prepared 24 h later and assayed for luciferase and β-galactosidase activity. Relative luciferase activity has been corrected for protein concentration and normalized for β-galactosidase activity.

direct repeat in the opposite orientation relative to the wild-type sequence while preserving the half-site spacing, thus converting it to the inverted repeat form. This oligonucleotide was tested for its ability to compete against the a1 probe for binding to the RARs. As shown in Fig 6B, lane 9, m14 was completely inactive as a competitor for the specific RAR complex, demonstrating that the direct repeat arrangement of the half-sites in the a1RARE is highly specific for RAR recognition. Similarly, m14 also failed to compete for complex 2, suggesting that the protein(s) involved in formation of this complex also recognize direct rather than inverted repeats.

Functional analysis of the a1RARE
To investigate whether the putative a1RARE could confer RA responsiveness to a heterologous promoter, the a1 fragment (a1/TK-Luc) as well as a double stranded oligonucleotide containing three tandem copies of the a1 sequence (a1RAREy/TK-Luc) were cloned upstream of the Herpes Simplex Virus Thymidine Kinase (TK) promoter driving the firefly luciferase reporter gene. These constructs were then transiently cotransfected with or without an RARα expression vector into F9 embryonal carcinoma cells and tested for RA inducibility.

Under the experimental conditions used, the TK promoter itself was unresponsive to RA (Fig 7A and B). In the absence of cotransfected RARα, the a1/TK-Luc reporter showed basal luciferase activity comparable to the TK-Luc reporter, with a slight increase upon addition of RA. In contrast, when RARα was cotransfected, a modest but reproducible 2- to 3-fold increase in luciferase activity was observed following addition of RA (Fig 7A). As a control we constructed a reporter (a13/TK-Luc) containing a genomic human fragment (a13) isolated by the same procedure. The a13 fragment neither binds nor competes for binding to the RARs by gel retardation assays (data not shown). As expected, the a13/TK-Luc reporter failed to show any increase in luciferase activity upon addition of RA with or without cotransfected RARα (Fig 7A).

Fig 7B summarizes the results obtained with the a1RAREy/TK-Luc reporter. The reporter with the response element cloned in either orientation showed basal luciferase activity comparable to the TK-Luc reporter, with slight increase upon addition of RA. In contrast, when RARα was cotransfected, a significant and reproducible 6- to 8-fold increase in luciferase activity was observed following addition of RA (Fig 7A). As a control we constructed a reporter (a13/TK-Luc) containing a genomic human fragment (a13) isolated by the same procedure. The a13 fragment neither binds nor competes for binding to the RARs by gel retardation assays (data not shown). As expected, the a13/TK-Luc reporter failed to show any increase in luciferase activity upon addition of RA with or without cotransfected RARα (Fig 7A).

Table I. Activation of various TK-Luc reporter plasmids by RA and RARα in CV-1 cells.

<table>
<thead>
<tr>
<th>Reporter plasmid</th>
<th>Transfected Receptor (−)</th>
<th>Transfected Receptor (+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TK-Luc</td>
<td>−RA</td>
<td>+RA</td>
</tr>
<tr>
<td>a1RAREy/TK-Luc(F)</td>
<td>1.0</td>
<td>1.4</td>
</tr>
<tr>
<td>a1RAREy/TK-Luc(R)</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>α26RAREy/TK-Luc(F)</td>
<td>1.0</td>
<td>1.2</td>
</tr>
</tbody>
</table>

(a) Relative luciferase activity was determined by dividing the normalized luciferase activity from each experiment by the normalized luciferase activity for untreated, non-receptor transfected cells.
(b) Forward orientation of the a1RARE (as shown in Fig 1) relative to the TK promoter.
(c) Reverse orientation of the a1RARE relative to the TK promoter.
(d) Plasmid α26RAREy/TK-Luc contains three copies of oligonucleotide α26 (Fig 1).
negative control, we cotransfected an antisense RARα expression vector. As shown in Fig 8, addition of RA induced an increase in luciferase with all four receptors tested, but not with the antisense control. The strongest response was obtained with the RARβ expression vector. We do not know the reasons for this preferential transactivation by RARβ. Our binding studies with RARα and RARβ did not reveal any significant differences in binding affinities to the α1 probe. The only reproducible difference so far observed between nuclear extracts from RARα and RARβ transfected Cells is the greater abundance of complex 2 with RARβ (compare lanes 1 and 3 in Fig 2). A complex with similar properties has also been detected in nuclear extracts from F9 and Hela cells (data not shown), suggesting that complex 2 is not a particular attribute of Cos cells. Whether this complex is physiologically relevant in transactivation of the α1RARE remains to be investigated.

RXRα was identified by molecular cloning and sequencing analysis as a new member of the steroid/thyroid receptor superfamily. Transcriptional activation studies revealed that RXRα responds specifically to Vitamin A metabolites, including RA, but it was considered unlikely that RA itself was the ligand since RXRα only shares 27% amino acid homology with the ligand binding domain of RARα (22). Additionally, recent studies have shown that the natural ligand that activates RXRα is 9-cis RA, a new stereoisomer of RA (61,62). As shown in Fig 8, RXRα was able to induce luciferase activity from the α1RARE reporter upon addition of RA. The levels of induction were slightly lower than for RARα. One possible explanation of these results is that F9 cells can isomerize some all-trans RA to 9-cis RA, which it acts as a ligand for RXRα.

**DISCUSSION**

We have shown above that both RARα and RARβ produced in Cos cells bind in a sequence-specific manner to the α1 oligonucleotide in gel retardation assays. Moreover, RARs translated in vitro also bound the α1 probe, demonstrating that no additional nuclear protein is required for this specific interaction. A second retarded complex detected by the α1 probe was also present in nuclear extracts from mock-transfected Cos cells. The strongest evidence against a role for RARs in formation of complex 2 is the failure of the βRARE oligonucleotide (01, Fig 2) to compete for this complex and the inability of both anti-RAR antibodies to induce a shift in the mobility of complex 2. In addition, the structural requirements and, to lesser degree, the sequence requirements for binding displayed by complex 2 were different from those exhibited by the RARs. We presently do not know the identity of the protein(s) involved in formation of complex 2. Interestingly, complex 2 is always more abundant in nuclear extracts prepared from RARβ-transfected Cos cells (compare lane 3 with lanes 1 and 2 in Fig 2). Since we culture Cos cells in the absence of RA, induction of this activity by retinoids is unlikely. Further studies will be required to identify the factor(s) involved in this complex.

The α1RARE contains nine-bp-long imperfect direct repeats, however, our mutational analysis indicates that the functional half-site unit is the motif AGGTCA, also shown to be important for ERE and TRE function. Nucleotide positions outside the hexameric half-sites seem to contribute little to sequence-specific binding. We have also shown that the spacing and orientation of the hexameric repeats play a critical role in RAR binding. Our binding studies with RARα and RARβ did not reveal any significant differences in specific footprint of RARE0 and the lamBl RARE (32,34). Instead, certain motifs are shown within the boxes. Arrows indicate the spatial orientation of the half-sites.
The existence of a natural 2 bp spaced RARE is consistent with our previous results. Similarly, for maximal activation of the pepck-RARE, sequences with the spacing of the half-sites (71) and is similar to ml2. The binding of the RARs to the lamBl-RARE was weak (34). These results suggest a consensus direct repeat RARE consisting of half-sites spaced by 5 nucleotides, as well as nucleotides located between the half-site repeats, are less conserved. Linear alignment of the sequences suggests a consensus direct repeat RARE consisting of half-sites of the sequence AG(G/T)TCA spaced by 5 bp. Three deviations from this consensus are the crbpI-, lamB1-, and pepck-RAREs (Fig 9C). The crbpI-RARE solely differs from the consensus in the spacing of the half-sites (71) and is similar to m12. The existence of a natural 2 bp spaced RARE is consistent with our mutational analysis. By contrast, the lamB1-RARE contains three non-uniformly spaced AGGTCA-related motifs and all three have been shown to be required for maximal activation. In addition, binding of the RARs to the lamB1-RARE was weak (34). Similarly, for maximal activation of the pepck-RARE, sequences directly upstream of the 1 bp spaced imperfect direct repeats have been shown to be required (37). These sequences are not related to the AGGTCA motif and they are not essential for RAR binding, suggesting that additional protein(s) may be involved in the receptor overexpression observed in F9 cells. Our studies have also demonstrated that binding of RARs produced in Cos cells to the α1RARE probe is strong. Still, the induction levels of the α1RARE, were not only dependent upon transfected receptor in F9 cells, but also unexpectedly low in CV-1 and Hela cells. Thus, this lack of correlation between binding and transcriptional activation makes it unlikely that DNA binding alone can explain these disparate observations. One interpretation of these results is that cell-type specific factor(s) may exist that modulate the receptor response by binding directly to the RAR-RA complex or by binding to factors required for transcription. Alternatively, overlapping binding sites for the RARs and complex 2 in the α1RARE may antagonize receptor binding and repress activation. Overlapping binding sites at the α-fetoprotein and human osteocalcin gene gene promoters have been implicated in mutual repression by c-jun/c-fos and members of the nuclear receptor superfamily (60,72). Whether complex 2 plays any role in repression and/or activation of the α1RARE remains to be shown. However, we find it unlikely that binding antagonism occurs at the α1RARE between RARs and the factors involved in formation of complex 2, because we have observed that complex 2 or a complex 2-like activity seems more abundant in untransfected F9 cells than in Cos and Hela cells (data not shown).

Regardless of the mechanism that may be used to achieve the differential activation of the α1RARE in a cell-type specific manner, our results show that F9 cells seem more permissible for RA-induced transcriptional activation mediated by this enhancer element. The embryonic nature of these cells and the fact that RA is potent inducer of their differentiation suggest that the α1RARE may regulate a gene involved in development. A search in the sequence data bank (GenBank release 65) did not uncover significant homology between the α1 fragment and known gene sequences. Southern blot analysis of human genomic DNA revealed discrete bands consistent with a low copy number sequence (data not shown). We are currently screening a genomic library to isolate the gene(s) regulated by the α1RARE.

In conclusion, our work has shown that it is possible to isolate authentic RAREs from a highly complex DNA mixture by using whole genome PCR. This approach may not only be useful to isolate RA-responsive genes, but it may also serve to identify novel genes regulated by other sequence-specific DNA binding proteins, provided that a stringent selection procedure can be developed. One possible drawback of our selection procedure based on coimmunoprecipitation of RAR binding sites seems to be that the polyclonal antibody we used may stabilize weak interactions between RARs and DNA. Since low-affinity sites should be more abundant in a random sequence, for this approach to be more efficient a different or modified selection strategy may be needed (41). When the promoter region of a gene has been cloned, whole genome PCR could provide a quick alternative to end-point deletion analysis to identify a regulatory sequence bound by a known protein factor.
ACKNOWLEDGMENTS

We thank J. Grippo for the RXRα expression vector and for critical reading of the manuscript; P. LeMotte for RARγ expression vector; M.L. DePamphilis for plasmid ptkluc; R. Speth for help with nucleotide sequencing; W. McComas and R. Motyka for oligonucleotide synthesis; W. Danho for peptide synthesis; and D. Schoenhaut for critical comments during preparation of the manuscript. We are especially grateful to M. I. Sherman for support and encouragement during the early stages of this work.

REFERENCES