A constitutive nuclear localization signal from the second zinc-finger of orphan nuclear receptor TR2

Z Yu, C-H Lee, C Chinpaisal and L-N Wei

Department of Pharmacology, University of Minnesota Medical School, Minneapolis, Minnesota 55455, USA

(Requests for offprints should be addressed to L-N Wei, Department of Pharmacology, University of Minnesota Medical School, 3–249 Millard Hall, 435 Delaware St SE, Minneapolis, MN 55455, USA)

Abstract

The orphan nuclear receptor TR2 and its truncated isoform deleted in the ligand binding domain (LBD) were localized exclusively in the nuclei as revealed by two methods of detection. An anti-hemagglutinin (HA) antibody detected specific nuclear localization of HA-tagged receptors and the green fluorescent protein (GFP)-tagged receptors were found to be distributed in the nuclei of living cells. By deletion analyses, the sequence responsible for targeting this receptor into the nucleus was defined. A stretch of 20 amino acid residues (KDCVINKHRNRCQYCRILQR) within the second zinc-finger of this receptor is required for its nuclear localization and this signal is constitutively active. No nuclear localization signal was found in the N-terminus or the LBD. The GFP-tagged receptor remained biologically active, as evidenced by its repressive activity on the reporter that carried a binding site for this receptor, a direct repeat-5 (DR5). An electrophoretic mobility shift assay was performed to characterize the binding property of TR2 and its truncated isoform. TR2 bound to the DR5 as dimers whereas its truncated isoform bound as monomers.

Journal of Endocrinology (1998) 159, 53–60

Introduction

Nuclear receptors comprise a large family of transcription factors that regulate gene expression in a wide variety of biological processes such as growth, differentiation and development. These transcription regulators modulate the transcription efficiency of their target genes by binding to specific DNA sequences and recruiting co-repressors or co-activators to the transcription machinery for the target genes (Chen & Evans 1995, Horwitz et al. 1996, Seol et al. 1996, Glass et al. 1997). Within this superfamily a large number of receptors are cloned without the identification of known ligands, named orphan receptors (Giguere et al. 1988, Laudet et al. 1992).

We have previously isolated and characterized a mouse orphan receptor gene, TR2–11 (Lee et al. 1995). Two isoform transcripts have been identified (Lee et al. 1997), one encoding 590 amino acid residues translated from 13 exons, originally designated as TR2–11-f (abbreviated as TR2), and the other encoding 256 amino acid residues as a result of alternative splicing, designated as TR2–11-t (the truncated form). The TR2–11-t protein sequence is identical to the TR2 sequence from the N-terminus to the zinc-finger motif, and is truncated from the C-terminus to the hinge region (Lee et al. 1997). At both the RNA and protein levels, TR2 expression is most abundant in the advanced male germ cells, such as pachytene spermatocytes and round spermatids. In contrast, TR2–11-t is only weakly expressed in somatic cells and early germ cells, and is not expressed in advanced germ cells (Lee et al. 1996a). The biological activities of the two receptor isoforms have also been examined. By using a retinoic acid (RA)-responsive reporter that contains a direct repeat (DR)–5 (DR5) type RA response element (RARE) of the retinoic acid receptor β (RARβ) gene (Sucov et al. 1990), we have demonstrated that TR2, but not TR2–11-t, repressed RA induction of this reporter in a dose-dependent manner (Lee et al. 1996a, 1997). By using a reporter containing a DR4-type hormone response element of the mouse cellular retinoic acid binding protein I (CRABP-I) gene, we have also shown that TR2, but not TR2–11-t, strongly repressed this reporter (Chinpaisal et al. 1997). In other promoter systems such as the Simian virus (SV) 40 promoter (Lee & Chang 1995) and the erythropoietin gene promoter (Lee et al. 1996b), TR2 also functions primarily as a transcriptional repressor (Lin et al. 1995).

Nuclear receptors must enter the nuclei to regulate gene expression. It has been demonstrated that the import is an energy-dependent process and is mediated by specific nuclear localization signals (NLSs) that interact with the nuclear pore complexes (Dingwall & Laskey 1991, Görlich 1997). Two types of NLS have been identified, the classical NLS and the M9 import signal. The classical NLS (Dingwall & Laskey 1991) contains positively charged residues and, for the nuclear receptors, is generally near, but not within, their DNA-binding domains (DBDs). The
SV40 large T antigen (Kalderon et al. 1984) contains many nuclear proteins, including the and does not contain characteristic basic residues. The NLSs identified in many nuclear proteins, including the estrogen receptor (Picard et al. 1990, Ylikomi et al. 1992), progesterone receptor (Guiochon-Mantel et al. 1989, 1991, Ylikomi et al. 1992), glucocorticoid receptor (Picard & Yamamoto 1987), and androgen receptor (Zhou et al. 1994) are characteristic of the classical NLSs. However, they differ in two respects. The SV40 large T antigen and HSP70 each contains only one cluster of basic amino acid residues, while most nuclear receptors contain several clusters. Some NLSs are constitutive while others are ligand-dependent, such as the NLSs in glucocorticoid and androgen receptors (Picard & Yamamoto 1987, Picard et al. 1990, Zhou et al. 1994). No orphan nuclear receptors have been characterized in terms of their cellular distribution or the identity of their NLSs.

In this study, we first investigated the cellular distribution of the orphan nuclear receptor TR2 and its truncated isoform TR2–11-t. Secondly, we determined the minimal requirement for its NLS and examined whether serum androgens or the identity of their NLSs. The cDNAs identified in many nuclear proteins, including the estrogen receptor (Picard et al. 1990, Ylikomi et al. 1992), progesterone receptor (Guiochon-Mantel et al. 1989, 1991, Ylikomi et al. 1992), glucocorticoid receptor (Picard & Yamamoto 1987), and androgen receptor (Zhou et al. 1994) are characteristic of the classical NLSs. However, they differ in two respects. The SV40 large T antigen and HSP70 each contains only one cluster of basic amino acid residues, while most nuclear receptors contain several clusters. Some NLSs are constitutive while others are ligand-dependent, such as the NLSs in glucocorticoid and androgen receptors (Picard & Yamamoto 1987, Picard et al. 1990, Zhou et al. 1994). No orphan nuclear receptors have been characterized in terms of their cellular distribution or the identity of their NLSs.

In this study, we first investigated the cellular distribution of the orphan nuclear receptor TR2 and its truncated isoform TR2–11-t. Secondly, we determined the minimal requirement for its NLS and examined whether serum contents affected its activity. Finally, we confirmed that the green fluorescent protein (GFP)-tagged receptor remained biologically active and that both TR2 and its truncated isoform were capable of binding to the same specific DNA target.

### Materials and Methods

#### Preparation of GFP-tagged fusions
The GFP expression vector (pEGFP-C1), purchased from Clontech (Palo Alto, CA, USA), was used as a molecular tag. The cDNAs encoding the full length (590 amino acids) and the truncated (256 amino acids) TR2 proteins (Lee et al. 1997) were each inserted into pEGFP-C1 at BglII and SmaI, and BglII and KpnI sites, and designated as TR2/GFP and TR2–11-t/GFP respectively. Constructs Δ1–219/GFP, Δ1–166/GFP, Δ140–154/GFP, Δ150–166/GFP and Δ168–219/GFP were generated by polymerase chain reaction (PCR). The primer sequences are listed in Table 1. The C-terminal deletion Δ1–138/GFP was constructed by digesting the full length TR2 cDNA with BglII and BamH1 and the N-terminal deletion Δ168–590/GFP was generated by PCR, each fragment was then inserted in frame into the pEGFP-C1 vector. The Δ140–166/GFP was constructed by digesting the construct Δ1–166/GFP with BamH1 and KpnI, deleting the sequence 1–139. All the sequences were confirmed by DNA sequencing and restriction enzyme digestion.

#### Cell culture and transfection
The monkey kidney cell line, COS-1, was used in this study. Cells were grown in Dulbecco’s modified Eagle’s medium (DMEM; Gibco, Gaithersburg, MD, USA) supplemented with 10% fetal bovine serum (FBS). For fluorescence microscopy, cells were plated onto coverglass in 35 mm dishes at a density of 5 × 10^4 cells/ml. Six hours after plating, the cultures were changed to medium containing either dextran–charcoal stripped (DCC) or regular serum. Cells were transfected with 5 µg plasmid DNA using the calcium phosphate precipitate method (Graham & van der Eb 1973). For reporter gene activity, COS-1 cells were plated on 24-well plates in DCC medium and co-transfected with 0.3 µg of the reporter, RAREβ-tk-luciferase, 0.1 µg TR2/GFP and 0.05 µg of the internal control plasmid SV-lacZ (a SV40 promoter-driven β-galactosidase reporter). All the reporter plasmids have been described previously (Lee et al. 1997). Retinoic acid (RA, 5 × 10^{-7} M) was added 24 h after transfection and cells were harvested 24 h after the RA induction. The luciferase activity was normalized to the lacZ activity to

### Table 1 Sequences of primers used for generation of fusion constructs

<table>
<thead>
<tr>
<th>Construct</th>
<th>Primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ1–219/GFP</td>
<td>Sense: 5’-ACATATGGTGAGCAAGGG-3’</td>
</tr>
<tr>
<td></td>
<td>Antisense: 5’-CCTCGAGTCATCCCGAAGGCAC-3’</td>
</tr>
<tr>
<td>Δ1–166/GFP</td>
<td>Sense: 5’-ACATATGGTGAGGAAAGG-3’</td>
</tr>
<tr>
<td></td>
<td>Antisense: 5’-CCTCGAGTCATCCCGAAGGCAC-3’</td>
</tr>
<tr>
<td>Δ140–154/GFP</td>
<td>Sense: 5’-TGGAGACAAAGCATGACAAAGG-3’</td>
</tr>
<tr>
<td></td>
<td>Antisense: 5’-CCTCGAGTCGTCATCCCGAAGGCAC-3’</td>
</tr>
<tr>
<td>Δ150–166/GFP</td>
<td>Sense: 5’-GGGATCCGAGAAACCGCTGCCA-3’</td>
</tr>
<tr>
<td></td>
<td>Antisense: 5’-CCTCGAGTCGTCATCCCGAAGGCAC-3’</td>
</tr>
<tr>
<td>Δ168–219/GFP</td>
<td>Sense: 5’-GGCTTCGAGTCGATCCACCGTCGTCG-3’</td>
</tr>
<tr>
<td></td>
<td>Antisense: 5’-CCTCGAGTCGTCGTCGTCGTCG-3’</td>
</tr>
</tbody>
</table>

Journal of Endocrinology (1998) 159, 53–60
obtain the relative luciferase units (RUL) as described previously (Lee et al. 1996a, 1997). Three independent experiments were performed, each in duplicate cultures.

Fluorescence microscopy and imaging

The transfected cells on coverglass were washed briefly with 1 × PBS 24 h after transfection, and fixed in 2% formaldehyde and 0.2% glutaraldehyde in PBS at room temperature for 15 min. The coverslips were mounted on slides with a mounting fluid containing 0.1% (w/v) p-phenylenediamine and 40% (v/v) glycerol in PBS (pH 8.0). Fluorescence signals, illuminated by a 150-W xenon arc lamp (Opti-Quip, Highlan Mills, NY, USA), were visualized with a Nikon CF fluor × 20 objective on a Diaphot 300 inverted microscope (Nikon, Melville, NY, USA). The image was captured using a cooled charged-coupled device camera (PXL KAF 1400, Photometrics, Tucson, AZ, USA) and IL Lab Spectrum (Signal analytics, Vienna, VA, USA) on an Apple Power Macintosh computer.

Antibody staining

A hemagglutinin (HA) epitope (YPYDVPDYA) derived from the human influenza HA was fused to a TR2 expression vector at the N-terminus under the control of the human Harvey ras promoter (Ishii et al. 1985), designated as HA-TR2. COS-1 cells were transiently transfected with HA-TR2. Thirty-six hours later, cells were fixed with a fixative as described above, permeabilized with a solution containing 0.5% Triton X-100 and stained with a fluorescent-conjugated anti-HA antibody (Boehringer Mannheim, Indianapolis, IN, USA). Fluorescent signal was visualized and images were captured as described above.

Electrophoretic mobility shift assay (EMSA)

The double stranded DNA fragments containing the DR5 of RARβ promoter (5'-AGCTTAAGGGTTCACCGA AAGTTCACTCGATATATGCT-3') were labeled with [α-32P]dCTP using Klenow fragment (Boehringer Mannheim). The TR2 proteins were synthesized using a T7-coupled transcription/translation reticulocyte lysate system (TNT system, Promega, Madison, WI, USA). The in vitro translated products were incubated with 32P-labeled probes at room temperature in a solution containing 20 mM Tris–HCl pH 7.5, 1 mM dithiothreitol, 1 mM EDTA, 100 mM KCl, 1 µg poly dI–dC, and 10% glycerol. Protein–DNA complexes were analyzed by electrophoresis in 5% non-denaturing polyacrylamide gels, followed by autoradiography.

Results

Intracellular localization of TR2 and TR2–11-t

In order to determine whether the previously observed difference in the biological activities of the full length and truncated mouse TR2 was due to any difference in their cellular distribution, we fused the full length and truncated receptor cDNA, individually, into the C-terminus of the TR2–11-t molecule including the N-terminal domain and the DBD, encoding sufficient information for importing the receptors into the nucleus.

The NLS of TR2

To identify the sequence that is responsible for TR2 receptor nuclear localization, we generated GFP fusion proteins of TR2, deleted in various portions of the
receptor, and expressed these fusion proteins in COS-1 cells. The maps of these fusion constructs are shown in Fig. 2A and the corresponding fluorescence images of these constructs (numbered 1–8) are shown in panels 1–8 respectively in Fig. 2B. For the C-terminal deletions (fusions 1 and 2 in Fig. 2A), the fusion proteins remained exclusively in the nucleus as long as the zinc finger motif was retained (Fig. 2B, panels 1 and 2). However, when the receptor was truncated up to residue 139 (fusion 3 in Fig. 2A, deleting the second zinc finger), the fusion protein became evenly distributed in both the nucleus and the cytoplasm, a pattern resembling that of the GFP control (Fig. 2B, panel 3). This result suggested that amino acid residues 140 to 166 (constituting the second zinc finger) were important for the nuclear localization. To determine if this sequence alone was sufficient for nuclear localization, the segment from residue 140 to 166 was fused to the GFP vector as fusion 4. As expected, this sequence of amino acid residues 140 to 166 was sufficient for nuclear localization as all the fluorescent signals remained in the nucleus. However, the intranuclear pattern was slightly different from that of fusions 1 and 2, indicating some other information for intranuclear distribution might be present in the A/B domains of the receptor molecules. The segment from residues 140 to 166 was further dissected into the N-terminal half and the C-terminal half as fusions 5 and 6 respectively. As shown in panels 5 and 6 of Fig. 2B, both fusion proteins were distributed homogeneously inside the cells, suggesting that the sequence 140–160 was required as a whole for efficient nuclear import.

To determine if additional NLS might be present in the hinge region or the ligand binding domain (LBD) of the receptor, we generated two more fusions by tagging the GFP to the hinge region linked to its LBD (fusion 7) or the hinge region alone (fusion 8) and their intracellular distribution patterns were examined. As shown in panels 7 and 8 in Fig. 2B, the fusion containing both the hinge and the LBD (panel 7) was localized in the cytoplasm whereas the fusion containing only the hinge region was distributed

![Figure 2](https://example.com/figure2)

Figure 2 Dissection of the NLS of TR2. (A) Serial deletions of TR2 were fused to the GFP and the TR2 portion in each fusion is shown for each construct. Dots indicate sequences that are identical to that shown for construct 1. Amino acid numbers from TR2 are indicated for each construct. A summary of their nuclear (N) and/or cytoplasmic (C) localization is provided on the right. (B) Fluorescent signals of COS-1 cells transfected with each fusion. The number of each panel corresponds to the number of each construct shown in panel A. The images were all taken 24 h following transfection.
in the whole cells (panel 8). Therefore, no other NLS is present in the hinge or the LBD regions of TR2. A summary of nuclear/cytoplasmic localization of these constructs is provided in the right-hand column of Fig. 2A.

Based upon its amino acid sequence, the TR2 full length receptor contains all the functional domains characteristic of the type II zinc-finger nuclear receptor, including the A/B domains (amino acid residues 1 to 98), the C domain (the DBD, from residues 99 to 166), followed by the D domain (hinge region) and the E domain (LBD). It is concluded that TR2 receptor contains only one NLS which is present between residues 139 and 166 and constitutes the second zinc finger motif within the DBD. No other NLS is encoded in other parts of this receptor, such as the N-terminus, the hinge or the LBD.

The constitutive NLS of TR2
Since several fusion proteins deleted in the LBD were intranuclear, the entry of TR2 into the nucleus appeared to be ligand-independent. To confirm this point further, the fusion construct containing the full length TR2 (TR2/GFP) or the truncated receptor without LBD (Δ1–219/GFP) was expressed in COS-1 cells in the presence of regular medium or DCC medium and their fluorescence distribution patterns were compared. As shown in Fig. 3, both TR2/GFP and Δ1–219/GFP were exclusively intranuclear, regardless of the medium contents. This result confirms that the identified NLS of TR2 can target the protein into the nucleus constitutively and its LBD encodes no additional signal for its import into the nucleus.

Biological activity of GFP-tagged TR2
Our previous studies demonstrated that expression of the full length TR2, but not TR2–11-t, repressed RA induction of an RA-responsive reporter, RAREβ-tk-luciferase. In order to determine whether the GFP-tagged receptors retained their biological properties, we examined the repressive activity of the GFP-tagged TR2 in transient transfection assays. COS-1 cells were transfected with the RAREβ-tk-luciferase reporter together with TR2/GFP and an internal control SV-lacZ vector. The specific luciferase activity (by normalizing to the internal control lacZ activity) in the presence of RA was compared with that of the control experiments, and the fold reporter induction by RA was determined. As shown in Fig. 4, TR2/GFP remained biologically active as it repressed RA induction of the reporter by approximately 60%, a level comparable to that achieved by the expression of the wild type TR2 (Lee et al. 1997). As observed in our previous studies (Lee et al. 1997), the GFP-TR2–11-t lacked the repressive activity on this reporter. Therefore, it is concluded that the GFP-tagged TR2 receptors remain biologically functional.

DNA binding property of TR2 and TR2–11-t
Only TR2, but not TR2–11-t, functioned as a repressor for all the response elements tested (Chinpaisal et al. 1997,
Lee et al. (1997). However, both the GFP-tagged full length and truncated receptors displayed very similar patterns of nuclear localization (Fig. 1). We then examined whether these two receptor isoforms had different DNA binding properties. An EMSA was conducted using receptor proteins expressed in TNT reactions to bind \( ^{32} \text{P} \)-labeled DR5 DNA fragments as shown in Fig. 5. It appeared that both receptor isoforms were able to bind to this DNA fragment specifically, since the bands shifted by either the full length or the truncated receptors (lanes 4 and 8 respectively) were competed out by a 100-fold excess of unlabeled DNA fragments containing only a half site of DR5 used for reactions loaded in lanes 1 and 5. The arrow on the left hand side of the gel indicates the bands shifted by specific dimeric TR2 binding. The arrow on the right hand side of the gel indicates the bands shifted by monomeric TR2–11-t binding.

**Discussion**

We have utilized the GFP as an in vivo tag to detect the intracellular localization of orphan receptor TR2. We have found that both the full length and the truncated isoforms are exclusively localized in the nucleus. The nuclear localization is further confirmed by staining cells expressing HA epitope-tagged receptors with an anti-HA antibody. In addition, the GFP-tagged TR2 remains biologically functional, suggesting that the addition of this GFP domain has not altered the biological function of TR2 in this reporter assay. The EMSA results add further support by showing that both the full length and the truncated receptors are able to bind to the DR5 element. Interestingly, only TR2–11-t, but not TR2, can bind a single site of this direct repeat. It is possible that the large LBD domain of TR2 may hinder its binding to DNA without a cooperative interaction with a partner binding to the other site of the repeat. The physiological relevance of the binding of the truncated receptor remains elusive.

By deletion analysis, a single NLS is located within the second zinc-finger of TR2, which is retained in both receptor isoforms. This NLS is constitutively active and is characteristic of the classical type NLS based upon its amino acid sequence. The dissected NLS as shown in fusion 4 in Fig. 2A, although capable of directing proteins into the nucleus (panel 4 of Fig. 2B), has an intranuclear distribution pattern which is slightly different from that of the fusions retaining their N-terminal sequences (fusions 1 and 2 of Fig. 2A). It is possible that other signals located in this N-terminal segment (domain A/B) may play a role in receptor targeting to specific nuclear locations. Since fusion 3, which retains only the A/B domain and the first zinc-finger, is distributed in the whole cells, this N-terminal sequence certainly lacks the information for nuclear localization. Additionally, the fusion which lacks the NLS and contains the hinge region linked to the LBD (\( \tilde{\text{A168–}590/\text{GFP}} \)) is present only in the cytoplasm, probably due to its relatively large molecular weight which prevents its diffusion into the nucleus (greater than 70 kDa). All the other fusions that have lost the NLS are relatively small in size; therefore, they are all distributed in both the nucleus and the cytoplasm.

The minimal sequence (KDCVINKHHRNRCQY CRLLQR) responsible for this NLS contains several lysine and arginine residues (bold letters). All these basic amino acids are required for efficient nuclear localization, since deletion of the two lysine residues at the N-terminus (fusion 6) or the last two arginine residues from the

**Figure 5** EMSA for TR2 and TR2–11-t. The EMSA assay was conducted as described in Materials and Methods. Lanes 1–4 contained TR2 proteins prepared in TNT, lanes 5–8 contained TR2–11-t proteins also made in TNT. For competition experiments, unlabeled DR5 fragment was included in reactions loaded in lanes 2 and 6 at a one hundred times molecular excess, and in lanes 3 and 7 at a ten times molecular excess. For competition with the half site, a one hundred times molecular excess of unlabeled DNA fragments containing only a half site of DR5 was used for reactions loaded in lanes 1 and 5. The arrow on the left hand side of the gel indicates the bands shifted by specific dimeric TR2 binding. The arrow on the right hand side of the gel indicates the bands shifted by monomeric TR2–11-t binding.
C-terminus (fusion 5) rendered the GFP fusion proteins distributed in the whole cells. Based on the characteristic of this sequence, this NLS can be classified into the classical NLS category. However, it is different from other classical NLSs in several respects. First, the entire NLS is localized within the second zinc finger of TR2. In contrast, all the identified NLSs of nuclear receptors are localized several amino acids downstream from the DBD, i.e. 10 for progesterone, glucocorticoid and androgen receptors (Guiochon-Mantel et al. 1989), 11 for estrogen receptor (Guiochon-Mantel et al. 1989), 14 for retinoid X receptor α and RARβ (Giguere et al. 1987, de The et al. 1987), 15 for thyroid hormone receptor (Weinberger et al. 1986), and 18 for vitamin D receptor (Baker et al. 1988).

Secondly, the NLS in the TR2 is constitutive and its function is independent of the serum contents. This is different from glucocorticoid and androgen receptors where hormone-inducible NLSs are found in each LBD (Picard & Yamamoto 1987, Picard et al. 1990, Zhou et al. 1994). In these cases, the hormone-inducible and constitutive NLSs cooperate to direct the receptors to the nucleus. Finally, most nuclear receptors are very dynamic in their intracellular localization, shuttling continuously between the cytoplasm and the nucleoplasm. However, TR2 seems less dynamic. We have followed the live images of the TR2–11-t/GFP and TR2–11-t/GFP fluorescence signals in the transfected cells for 24 h and found that the fluorescence signal remained in the nucleus during the entire period of observation. No signals could be captured in the cytosol (data not shown).

In summary, we have identified a constitutive NLS of the mouse TR2. This NLS is located within the second zinc-finger of the DBD and this motif alone is sufficient to target the TR2 receptor protein into the nucleus. Based upon the characteristic of this signal, it may be classified as a classical NLS. This is the first NLS characterized for an orphan nuclear receptor.

Acknowledgements

This work was supported by NIH grant DK46866, DA11190, a Grant-in-Aid from the Graduate School of University of Minnesota, and a grant SMF2005–98 from the Minnesota Medical Foundation to L N W. C C was supported by a fellowship from the Royal Thai Government. We thank the Core B of a PPG (DA08131) for help in oligonucleotide synthesis. We thank Dr Paul Sammak and the Biological Image Preparation Laboratory for the use of their fluorescence microscopy and image-processing equipment.

References


Received 5 February 1998
Accepted 19 May 1998