

Elevated retinoic acid receptor β_4 protein in human breast tumor cells with nuclear and cytoplasmic localization

KAREN M. SOMMER*, LUCINDA I. CHEN*, PIPER M. TREUTING†, LYNNE T. SMITH‡, AND KAREN SWISSELM*§

Departments of *Pathology, †Comparative Medicine, and ‡Dermatology, University of Washington, Seattle, WA 98195

Communicated by Arthur B. Pardee, Dana–Farber Cancer Institute, Boston, MA, May 14, 1999 (received for review January 28, 1999)

ABSTRACT The transcription factor retinoic acid receptor β_2 (RAR β_2) is a potent inhibitor of breast cancer cells *in vitro*, and studies suggest that RAR β expression is lost in primary breast cancer. Although RAR β_2 is selectively down-regulated at the mRNA level in breast tumor cells, we show that expression of an RAR β protein is elevated in five of five breast tumor cell lines relative to normal human mammary epithelial cells. Subsequent analysis identified this protein as the translation product of the human RAR β_4 transcript. Unlike the previously characterized mouse RAR β_4 isoform, the human RAR β_4 retains only half of a DNA-binding domain and lacks a ligand-independent transactivation domain at its N terminus. The RAR β_4 protein localizes to the cytoplasm and to subnuclear compartments that resemble nuclear bodies. The structure and preliminary characterizations of human RAR β_4 , coupled with the observation that its expression is greatly elevated in breast tumor cell lines, support the hypothesis that RAR β_4 functions as a dominant-negative repressor of RAR-mediated growth suppression.

Retinoids are a class of lipid-soluble micronutrients that mediate cellular signals critical for embryonic morphogenesis, cell growth, and differentiation. Retinoid signals are transduced by two families of retinoid-activated transcription factors: retinoic acid receptors (RARs) and retinoid X receptors (RXRs). RARs and RXRs are members of the nuclear receptor superfamily. Both the RAR and RXR families are composed of three genes (α , β , and γ) that generate multiple RAR and RXR isoforms by using two promoters (P1 and P2) and alternative splicing (1, 2). For most retinoid-inducible genes, transcriptional activation occurs through RAR–RXR heterodimers acting at a retinoic acid- or retinoid X-response element (RAREs and RXREs, respectively) in the presence of biologically active metabolites (e.g., all-*trans*-retinoic acid or 9-*cis*-retinoic acid) of vitamin A (retinol).

Based on homology with other members of the steroid hormone receptor superfamily, six distinct domains (A–F) have been identified within RARs and RXRs (3). These domains contain functional units of the transcription factor (4), including a ligand-independent transcription activation function-1, AF-1 (regions A and B); a DNA-binding domain (region C); a hinge region with roles in nuclear translocation and corepressor binding (region D); and a domain involved in dimerization, ligand binding, and ligand-dependent transcriptional activation, AF-2 (region E). The role of region F is unknown. The greatest diversity of protein sequences between RAR isoforms is generated at the N terminus by the incorporation of unique A regions with the use of alternative splicing and different promoters. Isoforms of a given RAR gene generally contain identical protein sequences B–F.

There is strong evidence that of the three RAR genes, RAR β has the central role in the growth regulation of

mammary epithelial cells. RAR β maps to chromosome 3p24, a region that exhibits a high frequency (45%) of loss of heterozygosity in primary breast tumors (5). We and others find selective loss of RAR β in tumorigenesis (6–8). Retroviral transduction of breast tumor cell lines with RAR β_2 results in an inhibition of tumor cell proliferation (9). Other neoplasias also exhibit loss of RAR β expression, including lung neoplasias (10), premalignant oral lesions (11), and cervical (12), ovarian (13), and prostate neoplasias (14). These findings suggest that RAR β_2 plays a role in limiting the growth of many cell types and that the loss of this regulatory activity is associated with tumorigenesis.

The human RAR β gene [NR1B2, according to current nomenclature for the nuclear receptor superfamily (15)], consists of 12 exons and two promoters, P1 and P2. P1 is the 5′-most promoter and directs transcription of RAR β_1 , which includes exons 1–4 and 6–12 (16). P2 is located within intron 4 and promotes transcription of RAR β_2 (exons 5–12). Loss of RAR β_2 in breast tumorigenesis may occur by transcriptional down-regulation at P2 (7, 17). Transcriptional activation of P2 is induced by retinoic acid through an RARE at –55 to –39, relative to the start of transcription (β RARE), in close proximity to the TATA box. In electrophoretic gel mobility shift assays, unique complexes are observed in breast tumor cell lines compared with normal human mammary epithelial cells (HMECs) upon β RARE incubation with nuclear protein extracts (7). In this report, we describe efforts to ascertain whether aberrant RARs or RXRs are present in breast tumor cells that may suppress RAR β_2 transcription. This effort led to the identification of a human RAR β protein, RAR β_4 , that is up-regulated in breast tumor cell lines.

MATERIALS AND METHODS

Cell Lines and Reagents. The normal HMEC strain AG11132 was obtained from the Coriell Institute (National Institutes of Aging Cell Repository, Camden, NJ). Breast tumor cell lines Hs578T, MCF-7, MDA-MB-231, and ZR-75-1 were obtained from the American Type Tissue Collection (Manassas, VA). MDA-MB-435 tumor cells were a gift from J. Price (18). Cells were cultured by using media described elsewhere (19). Retinoids were not removed from sera. For each experiment, cells from a single T75 culture flask were expanded to 5 × 150 mm tissue culture dishes and harvested in parallel for subsequent RNA and protein extraction.

Identification of the Human RAR β_4 Transcript in Breast Tumor Cells. Total RNA was extracted from AG11132 and MDA-MB-435 cells as described (7). Reverse transcription (RT) of RNA (5 μ g) from MDA-MB-435 and AG11132 cells

Abbreviations: HMECs, human mammary epithelial cells; RAR, retinoic acid receptor; RARE, RAR-binding element; RT, reverse transcription; RXR, retinoid X receptor; UTR, untranslated region. Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF157483).

§To whom reprint requests should be addressed at: Box 357470, 1959 N.E. Pacific Street, Seattle, WA 98195. e-mail: kswiss@u.washington.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

was performed (SuperScript II, GIBCO/BRL) with a reverse primer located 2270 bp 3' of the RAR β_2 transcription start site (2270RP; 5'-ACCTAGGTAGACGTGTACAG-3'). Primer locations are diagrammed relative to the RAR β_2 transcript in Fig. 3. Nested PCR amplifications were performed (eLONGase, GIBCO/BRL). The initial PCR contained 2 μ l of the RT reaction, primers 26FP (5'-TGTTTCAGAGGCAGGAGGGTC-3') and 2270RP in 1.8 mM MgSO $_4$ eLONGase buffer, and were amplified 35 cycles at 94°C (30 sec), 58°C (30 sec), and 68°C (2 min 30 sec). A second PCR was performed as above with 4 μ l of the first reaction as the template. Primers used in the subsequent reaction were 110FP (5'-AACGC-GAGCGATCCGAGCAG-3') and 2107RP (5'-TGACT-GGGTCTTTGCCATGC-3'). DNA bands of \approx 2 kb and 1.65 kb were gel purified and ligated into pCR2.1 (Invitrogen) to create plasmids pCR(RAR β_2) and pCR(RAR β_4), respectively. Inserts were cycle sequenced (dRhodamine, Perkin-Elmer Applied Biosystems) and analyzed (ABI 377 Prism, Perkin-Elmer Applied Biosystems).

Plasmids. For *in vitro* translation assays, RAR β constructs were cloned into vector pSP64(PolyA) (Promega). Inserts were created by PCR amplification of pSG5(RAR β) (20) or pCR(RAR β_4), by using oligonucleotides with 5' chimeric restriction enzyme cleavage sites. (In the following five primers, the translation start sites, if any, of the respective isoforms are boldface and the enzyme cleavage sites are italicized.) All of the PCRs used the reverse primer *SacI*-RAR β RP (5'-TTGG-GAGCTCTTATTGCACGAGTGGTGAC-3'; the *SacI* site is italicized). PCRs were performed with the eLONGase enzyme mix. For the RAR β_2 coding sequence, PCR of pSG5(RAR β) was performed by using the forward primer *Bam*HI-RAR β_2 (5'-GGATCCAGGGAGATCATGTTTGACTGTATGG-3'). Two forward primers were used to generate a coding sequence downstream of the CUG at nucleotide 259 of the RAR β_4 mRNA; *Bam*HI-*wt*-RAR β_4 259 (5'-TAGAGGATCCGCCGAGCAAG-CCTGG-3') and *Bam*HI-*mut*-RAR β_4 259 (5'-TAGAGGATCCGCCGAGCAAGCATGG-3'). The latter oligonucleotide alters the CUG initiation site to AUG. The coding sequence downstream of the AUG at nucleotide 448 of the RAR β_4 transcript was amplified by PCR of pSG5(RAR β) with the forward primer *Bam*HI-RAR β_4 448 (5'-GGATCCAGAA-GAATATGATTTACACTTGTACC-3'). PCR products were gel purified, digested with *Bam*HI and *SacI*, and ligated into the *Bam*HI and *SacI* sites of pSP64(PolyA) to create pSP(RAR β_2), pSP(RAR β_4 259*wt*), pSP(RAR β_4 259*mut*), and pSP(RAR β_4 448), respectively.

The eukaryotic expression vector pSecTag (Invitrogen), containing a multiple cloning site 5' of a Myc epitope and six histidine residues [(His) $_6$], was modified to remove the secretory signal as follows: pSecTag was digested with *Kpn*I and the resulting 3' overhanging nucleotides were removed with exonuclease. After an *Nhe*I digestion and a Klenow fill-in reaction, the blunt-ended linearized vector (pTag) was religated. PCR amplification of pSG5(RAR β) used chimeric primers to construct restriction enzyme recognition sites adjacent to the amplified RAR β sequence. All PCR amplifications used the reverse primer RAR β pTag (5'-CCCTCTAGATTGCACGAGTGGTGACTG-3'; the *Xba*I restriction site is italicized). The forward primers *Bam*HI-RAR β_2 and *Bam*HI-RAR β_4 448 (above) were used to amplify inserts for pTag(RAR β_2) and pTag(RAR β_4), respectively. Amplified products were gel-purified, incubated with *Bam*HI and *Xba*I, repurified, and ligated into the *Bam*HI and *Xba*I sites of the pTag vector.

Northern and RT-PCR Analysis of RAR β_4 mRNA Expression. Total RNA was obtained from AG11132, Hs578T, MDA-MB-231, MDA-MB-435, MCF-7, and ZR-75-1 cells, and Northern blot analysis was performed on duplicate samples (15 μ g) (19). The resulting membrane was divided, and duplicate halves were hybridized with either an RAR β_2 - or an RAR β_2 / β_4 -specific probe. The RAR β_2 probe was the amplification

product of pCR(RAR β_2) with primers 295FP(5'-GGCTTT-TAGCTGGCTTGTCTG-3') and 623RP (5'-TGAGCAGT-GTGCCGATGCTG-3'). The RAR β_2 / β_4 probe was prepared by amplifying pCR(RAR β_2) with primers 110FP and 259RP (5'-GGCTTGCTCGGCCAATCCAG-3'). The location of isoform-specific probes is depicted in Fig. 3. Probes were gel-purified and labeled by random priming in the presence of [α - 32 P]dCTP. Probes were hybridized to the membranes overnight at 65°C with 4 \times 10 5 cpm/ml of the radiolabeled probe in a solution of 1% SDS, 1 M sodium chloride, 10% dextran sulfate and 66.7 μ g/ml sheared salmon sperm DNA. Membranes were washed twice in 0.1 \times SSC/0.5% SDS at 65°C for 30 min, after which autoradiographs were produced. The *Pst*I fragment of a plasmid containing a 36B4 insert (21) was used as a loading control probe.

Total RNA (5 μ g) was analyzed by RT-PCR. After an initial RT from primer 1008RP (5'-GAGATCGTCCAACCTCAGCTG-3'), nested PCR using primers spanning the unique RAR β_4 splice junction [26FP and 1008RP, then 110FP and 933RP (5'-CCTGTCATTCTGACAGAT-3')] was performed. In the first round of PCR, 1 μ l of the RT reaction was amplified with *Taq* DNA polymerase (Roche Molecular Biochemicals, Indianapolis, IN), 35 cycles at 94°C (30 sec), 60°C (30 sec), and 72°C (1 min 30 sec). Two microliters of the first reaction was amplified with the second primer set, 40 cycles at 94°C (30 sec), 62°C (30 sec), and 72°C (1 min 15 sec). Duplicates of each sample were separated on agarose gels and then transferred onto a membrane. The duplicate membranes were probed with a radiolabeled oligonucleotide complementary either to the RAR β_2 sequence only (295FP) or to both the RAR β_2 and RAR β_4 isoforms (158FP; 5'-TCCGGAACGCATTCCGGAAGG-3'). Hybridization and wash conditions were as described (22).

Nuclear and Cytoplasmic Extracts. Nuclear extracts were obtained as described (7). The cytoplasmic fraction was also collected, i.e., the supernatant obtained after Dounce homogenization in buffer A and pelleting of the nuclear fraction. The cytoplasmic fraction was cleared by centrifugation at 24,000 \times g at 4°C for 2 hr and dialyzed against 100 vol of buffer D. Both nuclear and cytoplasmic protein extracts were frozen in ethanol/dry ice and stored at -80°C.

SDS/PAGE and Immunoblotting. *In vitro* translations were performed with the TnT SP6-coupled reticulocyte lysate (Promega). One microliter of *in vitro* translation product or 7 μ g of nuclear and cytoplasmic extracts were separated in duplicate on 12% SDS/PAGE and stained for total protein (SYPRO Orange, Bio-Rad) or blotted to a poly(vinylidene difluoride) membrane. The membrane was blocked in 1% nonfat dry milk (Bio-Rad) in PBS with 0.1% Tween-20 (PBS-T) for 1 hr at room temperature, followed by incubation with either an RAR β primary antibody (sc-552X; Santa Cruz Biotechnology) or a p170 topoisomerase II primary antibody (TopoGEN, Columbus, OH) at 1:1000 in PBS-T (1 hr). After four washes with PBS-T (5 min), the membrane was incubated with a horseradish peroxidase-conjugated anti-rabbit secondary antibody at 1:2000 in PBS-T (1 hr). The blot was visualized by chemiluminescent detection (Pierce).

Immunocytochemistry of Transiently Transfected Cells. Cells were plated at 2 \times 10 5 to 3 \times 10 5 cells per 9-cm 2 glass chamber slide and incubated for 1 to 2 days. For tumor cells, chamber slides were first coated with FBS 30 min before plating. Cells were transfected with 2.5 μ g of construct DNA using Superfect reagent (Qiagen, Valencia, CA). After fixation in cold 50% acetone/50% methanol (5 min), slides were incubated with 1% FBS in PBS (F-PBS) for 5 min. Antibody incubations were performed at room temperature for 1 hr. The primary anti-(His) $_6$ antibody (Invitrogen) was used at a 1:1000 dilution in F-PBS. A secondary antibody, FITC-conjugated goat anti-mouse antibody (Caltag, Burlingame, CA) was used at a 1:250 dilution in F-PBS. Slides were stained with 50 ng/ μ l 4',6-diamidino-2-phenylindole. Images analysis, including de-

convolution by epifluorescence with an air-cooled SenSys camera and IP LAB SPECTRUM software (Signal Analytics, Vienna, VA) was performed as described (23).

RESULTS

A Low Molecular Weight RARβ Protein in Tumor Cells. When Western blot analyses of known RARE-binding proteins were performed, we observed a striking difference in proteins with immunoaffinity for the RARβ antibody between the tumor and the normal cell protein extracts (Fig. 1). All breast tumor cells (five of five) contained two RARβ proteins with molecular masses of approximately 58.8 (53.0–64.7) kDa and 40.6 (36.5–44.7) kDa, compared with the single RARβ protein of 48.8 (44.0–53.7) kDa found in normal HMECs. The protein size range in parenthesis reflects a 10% margin of error for our molecular mass standards. The RARβ antibody was raised against amino acids at the C terminus, a region common among human RARβ isoforms (16). The molecular mass of RARβ₂ is predicted to be 50 kDa on the basis of amino acid sequence; this mass is close to the size seen in the normal HMEC nucleus. We have also observed the above-mentioned RARβ isoforms in breast tumor cell and normal HMEC proteins extracted from separate cell cultures. Preincubation of the RARβ antibody with a 5-fold excess (by weight) of the RARβ peptide used to generate the antibody completely abolished immunodetection of all three RARβ isoforms, confirming the specificity of our assay (data not shown). A p170 topoisomerase II immunoblot was used to confirm the presence and specificity of nuclear proteins in our tumor cell extracts. Because this protein is a marker of proliferating cells, its expression in nuclear proteins from tumor cells and normal cells is not expected to be equivalent. We can readily detect other nuclear proteins (e.g., RXRα, RARα) in the nuclear protein fraction of the normal AG11132 cells (data not shown).

The Putative Human RARβ₄ mRNA Is Identified in Mammary Cells. In previous experiments, two RARβ transcripts (≈3.1 kb and ≈2.8 kb) were observed in normal HMECs and breast tumor cell lines when probed with a pan-RARβ probe (7). RT-PCR was used to amplify exons 5 through 12 of the RARβ gene from the breast tumor cell line MDA-MB-435 and from normal HMEC AG11132. In both cells, two products were observed: one the expected size of ≈2 kb and the other ≈350 bp shorter. Sequence analysis of the amplicons revealed the deletion of nucleotides +269 through +625 from the

shorter product. Other than this deletion, there were no nucleotide alterations between +110 and +2107 (relative to the RARβ₂ transcript) when the smaller cDNA was compared with the RARβ₂ transcript. This region includes the sequence encompassing the entire RARβ₂ protein coding region. Nucleotides deleted from this isoform correspond exactly to those of a putative human RARβ₄ splice junction (24).

mRNA Expression of Human RARβ₄ in Mammary Cells. When probed with a portion of the sequence spliced from RARβ₂ in the novel isoform, a single ≈3.1-kb RARβ₂ transcript was detected in Northern blots of HMEC and tumor cell lines (Fig. 2A). RARβ₂ and the shorter (by approximately 350 bp) RARβ₄ species are observed when blots were hybridized with a probe complementary to the 5' untranslated region (UTR) that is specific for both RARβ₂ and RARβ₄ transcripts (Fig. 2A). RARβ₂ and RARβ₄ mRNAs were expressed at high levels in Hs578T and MDA-MB-435 breast tumor cells, and at lower levels in normal HMEC AG11132 and ZR-75-1 tumor cells. These expression results are highly repeatable with RNA derived from different cell cultures. RT-PCR, a more sensitive method of detecting RARβ₂ and RARβ₄ expression (Fig. 2B), confirmed the presence of RARβ₂ and RARβ₄ in the above cell lines and the lack of significant levels of these transcripts in the MCF-7 and MDA-MB-231 cell lines. The RARβ coding sequence can be detected reliably in these latter cells by using alternative primer sets that do not discriminate between RARβ₂ and RARβ₄ (e.g., 295FP and 1863RP, followed by 730FP and 1366RP, the primer names indicating the position in nucleotide 5' of the RARβ₂ transcription start site). Subsequent sequence analysis of genomic

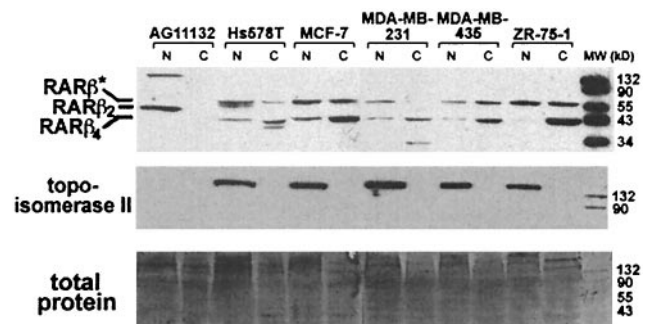


FIG. 1. A novel human RARβ protein is identified in breast cancer cell lines. Western blots of nuclear (N) and cytoplasmic (C) fractions of normal HMECs (AG11132) and breast cancer cell lines (Hs578T, MCF-7, MDA-MB-231, MDA-MB-435, and ZR-75-1) used a polyclonal antibody raised against a C-terminal RARβ epitope. The positions of molecular mass markers are indicated to the right. An ≈48.8-kDa protein (RARβ₂) is detected exclusively in the nucleus of normal HMECs. RARβ proteins with molecular masses of ≈58.8 kDa (RARβ*) and 40.6 kDa (RARβ₄) are apparent in the nucleus and cytoplasm of tumor cell lines only (five of five). The total protein stain and topoisomerase II (with a mass of 170 kDa) immunoblot from duplicate samples are included as a loading control and to confirm the specificity of nuclear protein extraction, respectively.

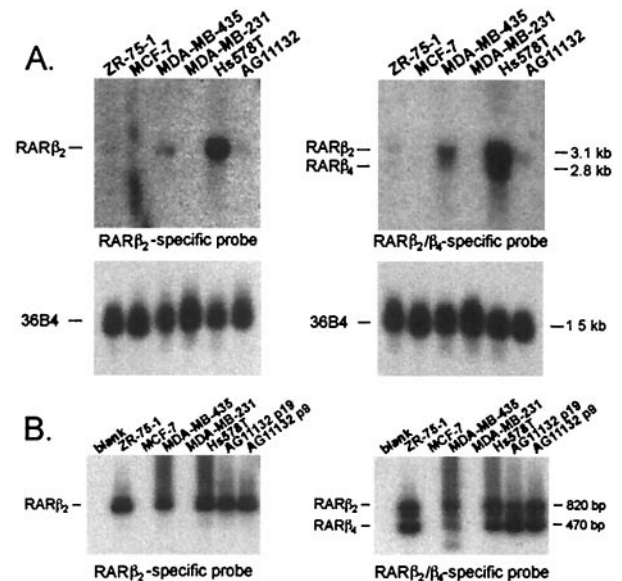


FIG. 2. Northern and RT-PCR analysis of RARβ₂ and RARβ₄ expression in mammary cells. (A) Northern blot of RNA from normal HMECs (AG11132) and breast tumor cell lines (Hs578T, MDA-MB-231, MDA-MB-435, MCF-7, and ZR-75-1) hybridized with either probe RARβ₂ or probe RARβ₂/β₄. The sizes of the resulting bands are indicated to the right. There were no RARβ transcripts visible in MCF-7 and MDA-MB-231 cells. Two RARβ transcripts are present at high levels in Hs578T and MDA-MB-435 cells, and at low levels in ZR-75-1 and normal AG11132 cells with probe RARβ₂/β₄. Only the higher molecular mass (3.1-kb) mRNA hybridized to probe RARβ₂. Membranes were subsequently stripped and rehybridized with 36B4, a loading control. (B) RARβ₂ and RARβ₄ expression was analyzed by using a sensitive RT-PCR assay. The region amplified spans the RARβ₄ splice junction. RNAs used in this experiment were separate aliquots of the same RNA used in the Northern blot assay. After transfer onto a membrane, blots were hybridized with an oligonucleotide specific for RARβ₂ or for RARβ₂/β₄, followed by autoradiography. The sizes of the resulting bands are indicated to the right. Passages 19 and 9 (p19, p9) were examined in the normal HMECs.

RAR β from these two cell lines, between -800 and +490, relative to the start of transcription (including the entire RAR β_2 5' UTR) showed no nucleotide alterations.

The Human RAR β_4 Protein Uses a Translation Start Site Within the DNA-Binding Domain. The RAR β_2 translation initiation site is within the sequence spliced from the RAR β_4 mRNA (Fig. 3). The 5'-most AUG of the RAR β_4 transcript that is in-frame with the RAR β coding sequence is located at nucleotide 448 of the RAR β_4 mRNA and corresponds to the methionine at amino acid 113 of the RAR β_2 protein. This AUG is within an appropriate nucleotide context for translation initiation (25) and would result in an ORF of 335 aa with an estimated molecular mass of 38 kDa. The mouse RAR β_4 protein initiates translation at a CUG that begins at nucleotide 259 of the mouse RAR β_4 transcript (24). There is an analogous CUG in the human RAR β_4 transcript within a nucleotide sequence that is similar to that of the mouse (86% identity within the surrounding 50 bp). If this CUG is used to initiate translation of the human transcript, the resulting protein would be 399 aa long and in-frame with the RAR β_2 coding sequence, with an estimated molecular mass of 45 kDa.

To ascertain the position of the translation start site generating the endogenous 40.6-kDa RAR β protein observed in breast tumor cell lines, the entire coding sequence of these two candidate translation start sites and 9 bp of the adjacent 5' untranslated region were cloned downstream of an SP6 promoter [pSP-(RAR β_4 448) and pSP(RAR β_4 259wt), respectively]. We also constructed clones with the RAR β_2 coding sequence [pSP-(RAR β_2)] for comparison, and constructed a version of pSP(RAR β_4 259wt), with the initial CUG mutated to AUG to ensure *in vitro* initiation of translation [pSP(RAR β_4 259mut)]. *In vitro* translation was performed and the sizes of the protein products generated *in vitro* were compared with the endogenous proteins by using Western blot analysis. Fig. 4 shows that the *in vitro* translated RAR β_2 protein is comparable in size to the higher molecular mass RAR β isoform found in the tumor cell extracts (58.8 kDa). The protein generated from pSP(RAR β_4 448), containing the putative coding region initiated by the AUG starting at position 448, has an apparent molecular mass identical to the 40.6-kDa protein found endogenously in tumor cell protein extracts. In tumor cells, both RAR β_2 and RAR β_4 have apparent molecular masses higher than their theoretical masses based on

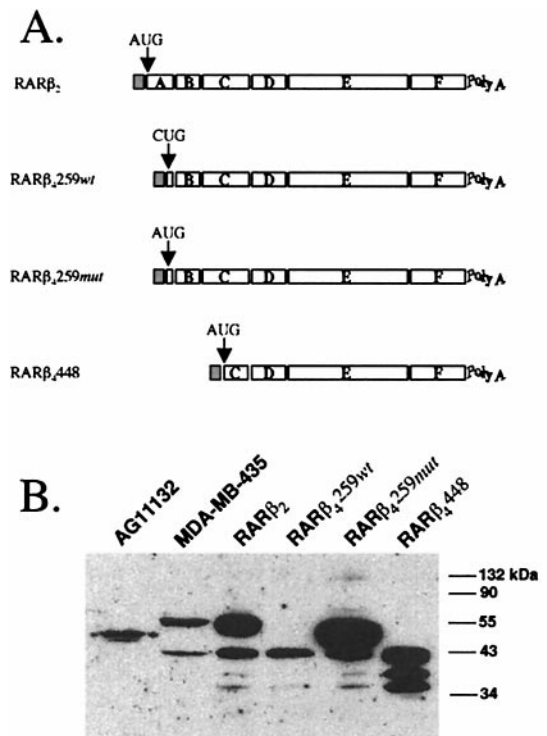


Fig. 4. *In vitro* translation of RAR β_4 candidate ORFs. (A) Transcripts used for *in vitro* translation assays are diagrammed. The translation start site for each transcript is indicated by an arrow, and the resulting protein domains (A-F) are denoted by open boxes. (B) Comparison of *in vitro* translation products RAR β_2 , RAR β_4 259wt, RAR β_4 259mut, and RAR β_4 448 with nuclear protein extracts from normal HMECs (AG1132) and breast cancer cells (MDA-MB-435). After SDS/PAGE, proteins were immunoblotted using the anti-RAR β antibody. Migration of molecular mass markers is indicated in kDa to the right of the figure. The 40.6-kDa protein observed in breast tumor cells (MDA-MB-435) corresponds to the product of translation initiated at nucleotide 448 (RAR β_4 448). The 58.8-kDa RAR β isoform in tumor cells is the size obtained from *in vitro* translation of the full RAR β_2 coding sequence. The 40.6-kDa RAR β protein was a product of all of the *in vitro* translation reactions, probably caused by leaky translation by the reticulocyte lysate.

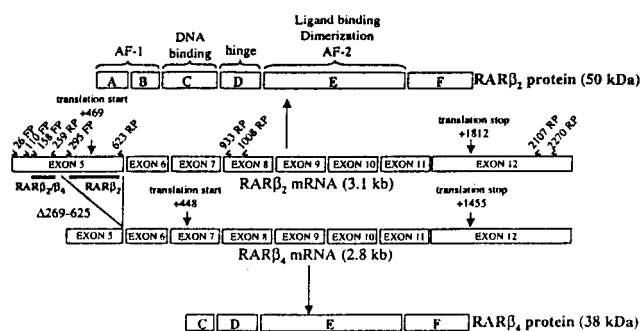


Fig. 3. Comparison of RAR β_2 and RAR β_4 mRNA and protein structures. Protein domains (A-F) of the RAR β isoforms are depicted to scale above (RAR β_2) or below (RAR β_4) diagrams representing their respective mRNA coding sequences. The molecular masses (in kDa) are the theoretical values predicted on the basis of amino acid sequence. RAR β_2 nucleotide sequence from 269 to 625 is deleted from RAR β_4 by alternative splicing in exon 5. The nucleotide positions of the translation start and stop sites are indicated with short arrows along the mRNA diagrams. Note that the RAR β_2 translation start site is included within the sequence deleted from RAR β_4 . The RAR β_4 translation begins within the C region, resulting in loss of the AF-1 and half of the DNA-binding domain. Oligonucleotide primers used for cloning or detection of RAR β_2 and RAR β_4 transcripts are denoted by > and <. The positions of sequence complementary to RAR β_2 - or RAR β_2/β_4 -specific probes are indicated by bars at the exon-5 end.

amino acid sequence. The *in vitro* translation product of pSP(RAR β_4 259mut), with a mutant AUG beginning at nucleotide 259 to initiate translation, is approximately 44.0 kDa and does not correspond with the size of any of the endogenous RAR β proteins. Translation of pSP(RAR β_4 259wt), having the wild-type CUG at nucleotide 259 to initiate translation, resulted primarily in the 40.6-kDa size produced by pSP(RAR β_4 448) (Fig. 4), thereby indicating that the wild type CUG at position 259 does not strongly direct translation initiation. Translation then initiates at the downstream AUG. These data are consistent with translation initiation of the human RAR β_4 transcript at the AUG, beginning at nucleotide 448. This AUG is maintained in the cDNA sequence of the two tumor cell lines sequenced (MDA-MB-435 and ZR-75-1) and in the normal AG1132 cDNA.

The RAR β_4 Protein Localizes to the Nucleus and Cytoplasm. To verify that the protein product of the RAR β_4 transcript is capable of nuclear translocation, we performed fluorescent immunocytochemistry on cells that were transiently transfected with (His) $_6$ -tagged RAR β proteins. The human RAR β_2 and RAR β_4 coding sequences were cloned into the mammalian expression vector pTag. These constructs [pTag(RAR β_2) and pTag(RAR β_4)] were transiently transfected into normal HMECs and MDA-MB-231 breast tumor cells. Subsequent immunocytochemistry with a monoclonal anti-(His) $_6$ antibody confirmed the presence of RAR β_4 within the nuclear fraction that was observed by Western blot anal-

ysis. The majority of the (His)₆-RARβ₄ protein in both normal and tumor cells (Fig. 5*B* and *E*, respectively) is cytoplasmic, in patterns resembling cytoskeletal staining. The nuclear signal of the (His)₆-RARβ₄ transfected cells was punctate in both cell lines and ranged from multiple small dots to larger, coalescing clusters. Deconvolution of the AG11132 cells that were transiently transfected with pTag(RARβ₄) confirmed the localization of the FITC punctate signal to the nucleus (Fig. 5*C*). As expected, the exogenous (His)₆-RARβ₂ protein localized primarily to the nucleus of both normal (Fig. 5*A*) and tumor (Fig. 5*D*) cells. The nuclear staining pattern of (His)₆-RARβ₂ appeared to be uniform throughout the nucleoplasm.

DISCUSSION

Two RARβ proteins were detected in the nuclei of breast tumor cell lines; these proteins had molecular masses distinctly different from the single RARβ protein found in the normal HMEC nucleus. It is probable that the 48.8-kDa species in normal cells is RARβ₂, on the basis of its size and the presence of the RARβ₂ mRNA. This is also the size observed in Western blot analyses of human RARβ₂ (26) and of the mouse RARβ₂ (27). We have determined that the 40.6-kDa isoform seen in breast tumor cells is human RARβ₄ (see discussion below). The 58.8-kDa RARβ species found exclusively in breast tumor cells has an electrophoretic mobility identical to RARβ₂ generated by *in vitro* translation. It is possible that this 58.8-kDa species represents posttranslationally modified RARβ₂. Other investigators have described phosphorylation of the mouse RARβ₂ during *in vitro* translation with rabbit reticulocyte lysates (27). RARβ₂ transcripts were detected in three of the five breast tumor cell lines examined, and these cell lines contained only the 58.8-kDa RARβ in addition to RARβ₄. Only two RARβ transcripts can be observed in these cells with a pan-RARβ probe (7). By hybridizing Northern blots with specific RARβ₂ or RARβ₂/RARβ₄ probes, these two transcripts were identified as RARβ₂ and RARβ₄ (Fig. 2*A*).

Although Nagpal *et al.* (24) described the exon-5/exon-6 junction of a putative human RARβ₄ cDNA over 6 years ago, neither a complete analysis of the RARβ₄ mRNA nor a reported human RARβ₄ protein product has been reported. Other investigators have detected a portion of the putative RARβ₄ transcript in lung tissue and in lung carcinoma cell lines (28). Here we

describe the human RARβ₄ transcript more extensively; we identify its protein product, which is expressed endogenously in human breast tumor cell lines. The RARβ₄ mRNA is probably generated by splicing of the nascent RARβ P2 transcript by means of an alternative exon-5 splice donor to that used to create RARβ₂. Both RARβ₂ and RARβ₄ contain exon-5 sequence that is specific for P2 transcription.

There is compelling evidence that RARβ₂ and RARβ₄ are posttranscriptionally regulated in breast epithelial cells. This appears dramatically when the protein and RNA levels in the MCF-7 cell line are compared with those in the MDA-MB-231 line (Figs. 1 and 2). Expression of RARβ₂ and RARβ₄ mRNA is undetectable by Northern analysis (Fig. 2*A*) and can be observed only by RT-PCR. Surprisingly, RARβ₂ and RARβ₄ proteins are present in these same cells at levels comparable to those cells with relatively high levels of the transcripts (e.g., MDA-MB-435) (Fig. 1). Conversely, although normal HMECs transcriptionally express RARβ₄, the RARβ₄ protein is not detected in these cells by Western blot assays. These two pieces of data suggest differential protein synthesis or stability for RARβ₄ between normal HMECs and breast cancer cells.

Sequence analysis of the RARβ₄ 5' untranslated region revealed the presence of 5 uORFs (Fig. 6). uORFs occur when AUGs are present 5' of the coding sequence initiation site. According to the scanning model of translation, a 40S preinitiation complex binds the mRNA cap site and proceeds in the 3' direction until reaching the first AUG that is within an adequate nucleotide context for translation initiation (29). Translation then proceeds until an in-frame stop codon is reached. Depending on uORF length and intercistronic distance, reinitiation of translation at a downstream AUG may occur (29, 30). Because reinitiation is inefficient, the presence of AUG sites in the 5' leader sequence is inhibitory to translation at downstream coding regions. Whereas upstream AUG sequences occur only rarely in mRNA transcripts, those having this mRNA structure are important regulators of cellular growth and differentiation, i.e., tumor suppressors, oncogenes, homeobox genes, and cell cycle control genes (31). uORFs, then, may provide an additional level of regulation for genes that are critical to the maintenance of cellular homeostasis. In human RARβ₄, uORFs 3 and 4 begin with AUGs within a strong context for translation initiation; translation of either

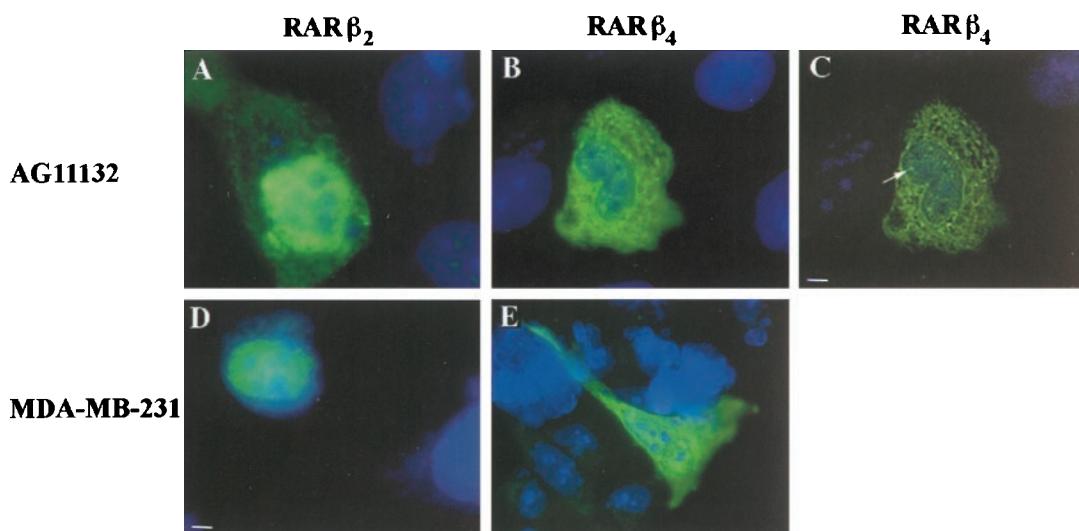


FIG. 5. Protein localization of (His)₆-RARβ₂ and (His)₆-RARβ₄ in normal HMECs and breast tumor cells. Immunocytochemistry was performed on normal HMECs (AG11132) and breast tumor cells (MDA-MB-231) transiently transfected with pTag(RARβ₂) and pTag(RARβ₄). (His)₆-tagged proteins were detected with a FITC-conjugated secondary antibody (green) with a 4',6-diamidino-2-phenylindole (DAPI) nuclear counterstain (blue). The majority of the (His)₆-RARβ₂ protein is nuclear in the normal (*A*) and the breast tumor cells (*D*). (His)₆-RARβ₄ protein in normal (*B*) and tumor (*E*) cells was detected both in the nucleus and in filamentous structures in the cell cytoplasm. The image in *B* was deconvolved to confirm nuclear localization of the FITC signal (*C*). Nuclear (His)₆-RARβ₄ localized to structures resembling nuclear bodies (an example is indicated with an arrow). (Bars = 5 μm.)

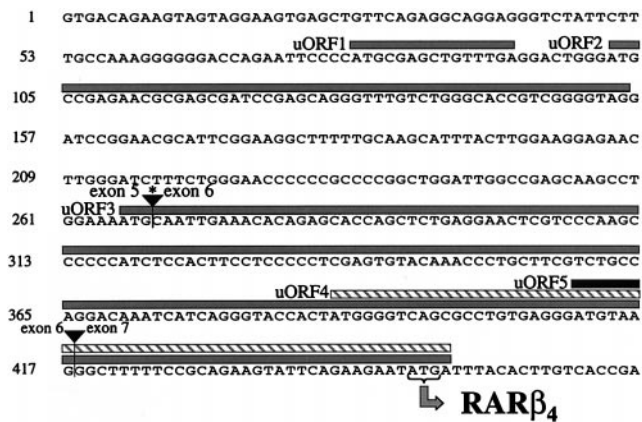


FIG. 6. uORFs in the 5' untranslated region of the RAR β_4 transcript. Nucleotides are numbered relative to the transcription start site. Nucleotides 269–625 are removed from exon 5 in the RAR β_4 transcript, resulting in a novel exon 5 and exon 6 junction (*). Translation of RAR β_4 begins at +448 of the RAR β_4 transcript, indicated above right-angled arrow. uORFs in the 5' (UTR) are indicated above their coding sequence (filled arrowheads = exon junctions).

of these uORFs would proceed out-of-frame and past the RAR β_4 translation start site. For RAR β_4 translation to occur, initiation at uORFs 3 and 4 must be bypassed. We hypothesize that the lack of RAR β_4 protein in normal HMECs results from tight inhibition of translation through these uORFs. Conversely, the presence of RAR β_4 protein in breast tumor cells with low transcript levels might indicate that these cells have evolved mechanisms of escaping this translational control. It has been shown that RAR β_2 uORFs translationally regulate the mouse RAR β_2 gene in a tissue-specific manner (32). Underscoring the presence of aberrant translational controls in certain breast tumor cell lines, a 33-kDa RAR β protein present in MDA-MB-231 cytoplasmic extracts (Fig. 1) corresponds to translation initiation at an internal AUG, beginning at nucleotide 545 of the RAR β_4 mRNA.

The CUG sequence starting at nucleotide 259 of the mouse RAR β_4 transcript directs translation initiation in COS-1 cells (24). We have determined that the analogous CUG in the human isoform does not appear to be recognized as an initiation site by the human translation machinery *in vivo* or by rabbit reticulocyte lysates *in vitro*, despite an 86% sequence identity. Instead, it is likely that translation of the RAR β_4 protein in breast tumor cells initiates at nucleotide 448 of the RAR β_4 mRNA. The resulting RAR β_4 protein consists of the C-terminal 335 aa of RAR β_2 . While maintaining the hinge region and its putative nuclear localization signal, the ligand-binding/dimerization domains, and the C-terminal transcription activation region (AF-2), the resulting human RAR β_4 protein lacks half of the DNA-binding domain, including one of two zinc fingers, and the entire N-terminal transcription activation domain (AF-1) of RAR β_2 (Fig. 3). The protein product of the human RAR β_4 mRNA is therefore quite different from the protein product generated in the mouse; in the mouse, it is identical to RAR β_2 , with the exception of a unique A region of 4 aa (24).

Although the human RAR β_4 lacks a complete DNA-binding domain and an AF-1 region, it retains domains required for ligand binding, dimerization, and interactions with nuclear receptor cofactors. Additionally, we show that RAR β_4 is selectively expressed at the protein level in breast tumor cells and that it is capable of nuclear localization. We hypothesize that RAR β_4 in the tumor cell nucleus maintains the ability to interact with RXR heterodimers and essential coactivators of the transcription ma-

chinery, sequestering these factors from transcriptionally productive interactions with RARs binding at target RAREs.

We thank colleagues at the University of Washington, Drs. D. Morris (Biochemistry), K. Stephens (Medical Genetics), S. Hosier and R. Robertson (Pathology), and B. Underwood (Dermatology) for technical support and helpful discussions. The pTag vector was provided by Dr. B. Sopher (Pathology). Plasmids pG5(RAR β_2) and pGEM(36B4) were provided by Dr. P. Chambon, Institut de Génétique et de la Biologie Moléculaire et Cellulaire, Strasbourg, France. This work was supported by a Genetic Approaches to Aging Research Training Grant (2 T32 AG00057) to K.M.S., by grants from the Molecular Training Program in Cancer Research (National Institutes of Health T32 CA 09437) to L.I.C., from the Biomedical Research Training for Veterinary Scientists (T32 RR07019) to P.M.T., from the George F. Odland Endowed Funds in Dermatology (AR2 1557) to L.T.S., from the U.S. Army (DMAD 17-94-J-40288), and from the American Cancer Society (JFRA-534) to K.S., and by a Clinical Nutrition Research Unit Pilot Award (P30-DK 35816) to K.S.

1. Chambon, P. (1996) *FASEB J.* **10**, 940–954.
2. Giguere, V. (1994) *Endocr. Rev.* **15**, 61–79.
3. Krust, A., Green, S., Argos, P., Kumar, V., Walter, P., Bornert, J.-M. & Chambon, P. (1986) *EMBO J.* **5**, 891–897.
4. Mangelsdorf, D. J., Umesono, K. & Evans, R. M. (1994) in *The Retinoids: Biology, Chemistry, and Medicine*, eds. Sporn, M. B., Roberts, A. B. & Goodman, D. S. (Raven, New York), pp. 319–349.
5. Deng, G., Lu, Y., Zlotnikov, G., Thor, A. D. & Smith, H. S. (1996) *Science* **274**, 2057–2059.
6. Roman, S. D., Clarke, C. L., Hall, R. E., Alexander, I. E. & Sutherland, R. L. (1992) *Cancer Res.* **52**, 2236–2242.
7. Swisshelm, K., Ryan, K., Lee, X., Tsou, H. C., Peacocke, M. & Sager, R. (1994) *Cell Growth Differ.* **5**, 133–141.
8. Jing, Y., Zhang, J., Bleiweiss, I. J., Waxman, S., Zelent, A. & Mira, Y. L. R. (1996) *FASEB J.* **10**, 1064–1070.
9. Seewaldt, V. L., Johnson, B. S., Parker, M. B., Collins, S. J. & Swisshelm, K. (1995) *Cell Growth Differ.* **6**, 1077–1088.
10. Houle, B., Leduc, F. & Bradley, W. E. (1991) *Genes Chromosomes Cancer* **3**, 358–366.
11. Lotan, R., Xu, X. C., Lippman, S. M., Ro, J. Y., Lee, J. S., Lee, J. J. & Hong, W. K. (1995) *N. Engl. J. Med.* **332**, 1405–1410.
12. Geisen, C., Denk, C., Gremm, B., Baust, C., Karger, A., Bollag, W. & Schwarz, E. (1997) *Cancer Res.* **57**, 1460–1467.
13. Sabichi, A. L., Hendricks, D. T., Bober, M. A. & Birrer, M. J. (1998) *J. Natl. Cancer Inst.* **90**, 597–605.
14. Campbell, M. J., Park, S., Uskokovic, M. R., Dawson, M. I. & Koeffler, H. P. (1998) *Endocrinology* **139**, 1972–1980.
15. Nuclear Receptors Nomenclature Committee (1999) *Cell* **97**, 161–163.
16. Toulouse, A., Morin, J., Pelletier, M. & Bradley, W. E. C. (1996) *Biochim. Biophys. Acta* **1309**, 1–4.
17. Liu, Y., Lee, M. O., Wang, H. G., Li, Y., Hashimoto, Y., Klaus, M., Reed, J. C. & Zhang, X. (1996) *Mol. Cell Biol.* **16**, 1138–1149.
18. Price, J. E., Polyzos, A., Zhang, R. D. & Daniels, L. M. (1990) *Cancer Res.* **50**, 717–721.
19. Swisshelm, K., Machl, A., Planitzer, S., Robertson, R., Kubbies, M. & Hosier, S. (1999) *Gene* **226**, 285–295.
20. Brand, N., Petkovich, M., Krust, A., Chambon, P., de The, H., Marchio, A., Tiollais, P. & Dejean, A. (1988) *Nature (London)* **332**, 850–853.
21. Laborda, J. (1991) *Nucleic Acids Res.* **19**, 3998.
22. Omiecinski, C. J., Walz, F. G. & Vlasuk, G. P. (1985) *J. Biol. Chem.* **260**, 3247–3250.
23. Smith, L. T., Underwood, R. A. & McLean, W. H. I. (1999) *Brit. J. Dermatol.* **140**, 582–591.
24. Nagpal, S., Zelent, A. & Chambon, P. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2718–2722.
25. Kozak, M. (1989) *J. Cell Biol.* **108**, 229–241.
26. Houle, B., Rochette-Egly, C. & Bradley, W. E. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 985–989.
27. Rochette-Egly, C., Gaub, M.-P., Lutz, Y., Ali, S., Scheuer, I. & Chambon, P. (1992) *Mol. Endocrinol.* **6**, 2197–2209.
28. Berard, J., Gaboury, L., Landers, M., De Repentigny, Y., Houle, B., Kothary, R. & Bradley, W. E. (1994) *EMBO J.* **13**, 5570–5580.
29. Kozak, M. (1989) *Mol. Cell Biol.* **9**, 5134–5142.
30. Geballe, A. P. & Morris, D. R. (1994) *Trends Biochem. Sci.* **19**, 159–164.
31. Kozak, M. (1991) *J. Cell Biol.* **115**, 887–903.
32. Zimmer, A., Zimmer, A. M. & Reynolds, K. (1994) *J. Cell Biol.* **127**, 1111–1119.