Downstream Codons in the Retinoic Acid Receptor β-2 and β-4 mRNAs Initiate Translation of a Protein Isoform That Disrupts Retinoid-activated Transcription

Lucinda I. Chen‡§¶, Karen M. Sommer‡§, and Karen Swisshelm‡¶

From the ‡Department of Pathology, University of Washington, Seattle, Washington 98195

Retinoic acid receptors (RARs) are essential for the differentiation and maintenance of normal epithelium. In studies of RARs in breast cancer, there are striking differences in the expression of certain protein isoforms of the RARβ gene between cells derived from normal human mammary glands and those derived from breast tumors. While the protein isoforms RARβ2 and RARβ4 consist of the longest open reading frames of the RARβ2 and RARβ4 mRNAs, respectively, we find that a fraction of retinoic acid receptors bypass these upstream RARβ2 and RARβ4 protein start codons and initiate translation downstream. This downstream translation initiation site is identical in the RARβ2 and RARβ4 transcripts and generates a third RARβ protein isoform, here termed RARβ′ (formerly human RARβ4). RARβ′ lacks protein domains found in the N terminus of RARβ2 and RARβ4, including one of two zinc fingers required for DNA binding. However, RARβ′ retains the ability to heterodimerize with RXRs and interact with transcription cofactors. In reporter gene assays, RARβ′ repressed retinoic acid-activated transcription of co-transfected RARβ2, RARβ4, and RARα. This repression required the presence of acidic amino acids within the AF2 domain. These findings demonstrate an antagonistic role for RARβ′ in signaling by retinoic acid.

Vitamin A derivatives, called retinoids, are required for epithelial cell development and differentiation. Retinoid signals are transduced largely through ligand activation of two families of retinoid-activated transcription factors belonging to the nuclear receptor superfamily, retinoic acid receptors (RARs) and retinoid X receptors (RXRs). Receptors of both the RAR and RXR subfamilies share conserved protein regions designated A through F. The A and B regions include a ligand-independent transactivation domain (AF-1) whereas region C encompasses the DNA-binding domain. The D region acts as a conformational hinge between the DNA-binding domain and the remaining parts of the protein. The ligand binding, dimerization, and ligand-dependent transactivation (AF-2) domains are all found within regions E and F. RAR-RXR heterodimers bind to cis-acting DNA elements in gene promoter regions called retinoic acid response elements (RAREs) and transactivate target genes in the presence of a retinoid ligand. Both the RAR and RXR families are comprised of three genes (α, β, and γ) that generate multiple isoforms via the usage of two promoters, P1 and P2, and alternative splicing (1).

As the primary effectors of retinoid signaling, the RARs and RXRs themselves appear to be targets for disruption in tumorigenesis (2), including loss of heterozygosity (3), gene rearrangements (4, 5), mutations (6), and aberrant mRNA expression (7, 8). Retinoic acid receptor β (RARβ) in particular has been extensively studied in human carcinomas, and a body of evidence indicates that it may play a role in tumor suppression. Such findings include the loss of heterozygosity of its genetic locus (3p24) in primary tumors of the breast (3) and the loss of RARβ mRNA expression in primary breast (9, 10), lung (8, 11), and esophageal carcinomas (12). Loss of RARβ transcript expression has been linked to epigenetic silencing by methylation of the RARβ P2 promoter in primary breast (13–17) and lung tumors (18). Further evidence of RARβ suppression of tumor cell proliferation comes from cell culture studies, where ectopic expression of the RARβ2 isoform resulted in decreased proliferation of tumor cells lines derived from the breast (19, 20), lung (21, 22), cervix (23), pancreas (24), and squamous cell carcinomas (25).

In the mouse, the RARβ gene generates four distinct transcripts: splice variants RARβ1 and RARβ3 from transcription at promoter P1, and RARβ2 and RARβ4 from the RARE-containing P2 promoter (26, 27). In the human, only RARβ2 and RARβ4 transcripts have been identified in normal adult cells (28). Human RARβ1 is expressed in fetal tissues and some small cell lung carcinoma cell lines (29), whereas a human homologue of the RARβ3 isoform has not been detected. The RARβ2 and RARβ4 transcripts differ only in the content of their 5′-most exon (exon 5 relative to the entire RARβ gene), a result of alternative splicing (26). RARβ4 lacks 357 nucleotides encoded by this exon that are present in the RARβ2 mRNA, including the translation initiation site for the RARβ2 protein.

We have previously reported several unique findings regarding RARβ expression in cultured cells derived from normal and neoplastic mammary epithelium. Only two mRNA products of the RARβ gene (RARβ2 and RARβ4) had been detected in these cells by Northern blot analyses (30), and these transcripts were expressed at low levels in most breast tumor cell lines compared with human mammary epithelial cells (HMECs) isolated from normal breast tissue. However, the levels of RARβ protein in these cells could not be predicted on the basis of their transcript levels. Some cell lines with abundant RARβ2 and
Retinoic Acid Receptor Protein Isoforms

RARβ mRNAs expressed low levels of RARβ protein. Conversely, other cells, such as the breast carcinoma cell line MCF-7, expressed high levels of RARβ protein with very low transcript levels by Northern blot and RT-PCR analyses. We also noted the presence of three distinct RARβ protein species in these cell lines, although only two RARβ transcripts were present. Finally, we observed differential expression of these RARβ protein isoforms in cells derived from normal breast tissue compared with those from neoplastic mammary tissue (31).

In the present study, we introduced mutations into RARβ2 and RARβ4 5′ transcript leader sequences to determine which candidate codons operate to initiate protein synthesis within the cell. We performed a functional characterization of the protein generated from a downstream initiation codon of both the RARβ2 and RARβ4 transcripts, RARβ′, to determine its ability to bind a RARE, heterodimerize, and interact with transcription cofactors. We then tested the effect of this protein on retinoid-activated transcription in transient transfection assays. Our studies show distinct and opposing functions for RARβ2 and RARβ4 compared with RARβ′, and we propose that RARβ′ acts as an inhibitor of transcriptional activation mediated by RARs.

EXPERIMENTAL PROCEDURES

Plasmid Construction—To synthesize expression constructs of RARβ2 or RARβ4 transcripts including all nucleosides from +1 throughout the coding region, the promoter construct pβ2-747luc (32) was PCR-amplified between +1 and +155 (HindIII/RARβ forward primer, 5′-GAT AAG CTT GTG ACA GAA GTA GGA AGT-3′; HindIII site in bold; reverse primer, 5′-AGT GGA TCA TAC CCC GAC GGT GCC CAC AC-3′). Following digestion with HindIII and BamHI, the amplicon was ligated into the pBluescript II vector (Stratagene, La Jolla, CA), and the resulting plasmid was linearized with BamHI digestion. RARβ2 and RARβ4 cDNA fragments from +155 through +2107 or +155 through +1750, respectively, were obtained by the digestion of pCR(RARβ2) and pCR(RARβ4) (31) with BamHI. These RARβ fragments were ligated into the linearized plasmid described above to obtain plasmids pBS-RARβ2 and pBS-RARβ4.

For mammalian expression, RARβ cDNAs were cloned into the first polynucleoty of vector pCS2+MT (33) to achieve read-through translation of 6 Myc epitopes at the protein C terminus. pCS2+MT was modified to contain a SmaI site by the insertion of a double-stranded oligonucleotide (top, 5′-GAT CCA GAT CTA GTG GAG AAT ATG TTA ATTGGACACCAAAACATTCC-3′; bottom, 5′-CGA TGA CCC GGG AGG ATG AGA TCT GTC ACC GGG CCA GAG TAT-3′; SmaI sites in bold) between the vector BamHI and ClaI sites. This modified vector was then digested with BamHI and SmaI for the following cloning steps. Plasmids pCS(RARβ2) and pCS(RARβ4), which contain nucleotides corresponding to +1 of the transcript through the entire RARβ2 or RARβ4 coding sequence (excluding the stop codon) were created by PCR amplification of pBS(RARβ2) or pBS(RARβ4) with primers containing either a HindIII site (HindIII/RARβ forward primer), or an XbaI site (reverse primer, 5′-CCCTCAGAAGCAGGATTGTGTA CCT-3′; XbaI site in bold). All PCR amplifications were performed using the Expand High Fidelity PCR System (Roche Molecular Biochemicals) and cycled using parameters detailed previously (31). Amplicons were digested with XbaI followed by a Klenow fill-in reaction and then subsequently digested with HindIII. The resulting fragments were ligated into the digested pCS2+MT vector.

A series of constructs derived from plasmids pCS(RARβ2) and pCS(RARβ4) were created with point mutations at putative translation initiation sites to either eliminate or to enhance ribosomal recognition as translation initiation sites. Site-directed mutagenesis was performed using the QuikChange kit (Stratagene, La Jolla, CA) according to directions. Site-directed mutagenesis plasmids and their targeted mutations are detailed in Table I. For mammalian expression, protein-coding sequences for the RARβ isoforms were cloned into the pTag vector (31) and included stop codons following the coding sequence so that the C-terminal (His)6 and Myc tags were not translated. A fragment containing the RARβ2 coding sequence was amplified from plasmid pCS(RARβ2 T2-) (see Table I) using forward primer BamHI-RARβ2 (33) and reverse primer 3′ BamHI RP (5′-TTGCGGAGATCCTATGCGACGAGTGTCAC-3′; BamHI site in bold). Following BamHI digestion, the insert was ligated into the BamHI site of pTag to create pTag/RARβ2/stop. Forward primer BamHI-mut-RARβ4 259 (31) and reverse primer 3′ NotI RP (5′-TTGGCGGCGCTTATTGCGAGTTGGTGA-3′, NotI site in bold) were used to amplify the RARβ4 coding sequence from pCS(RARβ4 T2-) with forward primer 5′-CCCTGGCATCCAGGAAAGATGTTGTTA CACTTGTC-3′, BamHI site in bold) and reverse primer 3′ NotI RP. Following digestion with BamHI and NotI, agrose gel-purified RARβ4 and RARβ′ coding sequences were ligated into the BamHI and NotI sites of pTag, generating plasmids pTag/RARβ4/stop (pTag/RARβ′/stop), respectively. pTag/RARβ4/stop (pTag/RARβ′/stop) was constructed by PCR amplification of pSP/RARRA2/AF2 (see below) with primers pSP-RARβ2/AF2 (5′-CCACTCGAGAAAGATGTTGTTA CACTTGTC-3′, PstI site in bold) and 3′ NotI RP, followed by BamHI digestion and ligation into the BamHI site of pTag, pTag/RARβ2/AF2 was created by PCR amplification of pSP/RARβ2/AF2 (see below) using primers PstRI-RARβ2′ FP (5′-CCACTCGAGAAAGATGTTGTTA CACTTGTC-3′, PstI site in bold) and 3′ NotI RP, followed by PstI and NotI digestion and ligation into the PstI and NotI sites of pTag.

For translation in vitro of a C-terminal-terminated RARβ2 isoform, plasmid pSP(RARβ2) was PCR-amplified with forward primer BamHI-RARβ2 (31) and the following reverse primer with an added SacI site (in bold): 5′-GAATTCCTAATTCTCTTGTAATTCTCTGC-3′. The gel-purified and ligated plasmid pSP(RARβ2) was constructed by PCR amplification of pSP/RARβ2/AF2 with primers pSP-RARβ2′ FP (5′-CCACTCGAGAAAGATGTTGTTA CACTTGTC-3′, PstI site in bold) and 3′ NotI RP (31) to replace codons for three glutamic acid residues in the AF2 domain with those encoding neutral amino acids alanine and valine (Fig. 8A). Site-directed mutagenesis was performed using the QuickChange kit as described by the manufacturer and resulted in plasmids pSP/RARβ2/AF2 and pSP/RARβ2/ΔAF2. For expresion of Myc-tagged proteins, 1.0 μg of plasmid pRL-VRx SO (34) from a Steven Collins, was amplified (forward primer, 5′-GAATTCCTGCGAGAATGTTGTTA CACTTGTC-3′) and reverse primer, 5′-CTTGGATCCGCAATGTTGTTA CACTTGTC-3′, EcoRI and XbaI digestion sites in bold, respectively), followed by EcoRI digestion, XbaI knock-in, XbaI digestion, and ligation into the HindIII (blunted by Klenow) and XbaI sites of pSP94.

For bacterial expression of N-terminal glutathione S-transferase (GST) fusion proteins, pl/RXSO1 (35) was digested with EcoRI. The fragment containing the RRα coding sequence was purified and then ligated into the EcoRI site of pGEX-4T-2 (Amersham Biosciences) to make pGEX(RRXSO1). Plasmids GSTC-SMRT and pGEX6P1AB1T1 were provided by光电 Jovanes (35) and Paul Meltzer (36), respectively. For expression of RARβ2 or RARβ4, amino- or carboxy-terminally tagged, synthetic oligonucleotides corresponding to nucleotides ~83~37 of the RARβ2 promoter (5′-ggg tca TTT GAG gaa gta TGT CAC GGG TAG GGT TCA CCG AAA GTG CA-3′; RARE in bold, non-consensus RARE in lowercase) were ligated into a BglII and NheI digested pG-PL-Promoter (Promega) to give pGβ/RARE, a retinoic acid-responsive reporter construct for use in the luciferase assays. All plasmid constructs described above were verified by automated sequence analysis.

Cell Culture and Transfection—The AG11132 human mammary epithelial cell (HEMC) strain was obtained from the Coriell Institute (National Institutes of Aging Repository, Camden, NJ). MCF-7 human breast cancer cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA). AG11132 and MCF-7 cells were cultured as previously described (37).

For expression of Myc-tagged proteins, 1.0 × 10^6 AG11132 or MCF-7 cells were plated in 60-mm^2 dishes. After 24 h, cells were transfected with 10 μg of one of the above pCS plasmids using 20 μl of SuperFect transfection reagent (Qiagen, Valencia, CA) following the manufacturer’s protocol. Forty-eight hours post-transfection, cells were scraped in lysis buffer (150 mm NaCl, 5 mm EDTA, 1/10 Trition X-100, 10 mm Tris-HCl, pH 7.4, 5 mm dithiothreitol, 0.1 mm phenylmethylsulfonyl fluoride). Following incubation on ice for 10 min, the cells were centrifuged at 13,000 × g for 10 min and supernatants were collected and stored at −80 °C for Western blot analysis. For luciferase assays, MCF-7 cells were plated in 24-well plates at 70,000 or 100,000 cells/well. The cells were transfected with a total of 0.24–0.27 μg DNA per well using Effectene transfection reagent (Qiagen) in medium containing 5% charcoal-dextran-treated fetal calf serum (Hyclone, Logan, UT). The amount of DNA per transfection was equalized using the pTag vector. The transfection control plasmid pRL (Renilla luciferase) was co-transfected at one-tenth the amount of pGL3RARE. Twenty-four hours following transfection, 0.1% ethanol or all-trans retinoic acid (atRA) (to final concentrations of 0.01 or 1 μM) was added. After another 24 h, the cells were lysed in 1× Passive Lysis Buffer (Promega) and
immediately assayed for luciferase activity or stored at −20 °C.

Western Blot Analysis—12 μl of each cell lysate from MCF-7 and AG11132 cells transfected with pCS-derived plasmids (above) were separated on 12% SDS-PAGE gels and transferred to polyvinylidene difluoride membrane as described previously (31). Membranes were incubated in 1% milk in phosphate-buffered saline with 0.1% Tween-20 (PBS-T) for 1 h at room temperature and then incubated with a monoclonal anti-Myc antibody (1:200 in PBS-T) produced as a medium supernatant from the Myc1-9E10.2 hybridoma line (38) for 1 h at room temperature. After extensive washing in PBS-T, membranes were incubated 1 h at room temperature with anti-mouse IgG-horseradish peroxidase secondary antibody (1:20,000 in PBS-T; Pierce, Rockford, IL), re-washed in PBS-T, then incubated 5 min at room temperature in SuperSignal West Pico chemiluminescent substrate (Pierce), followed by autoradiography.

In Vitro Translation—1 μg of plasmid pSP(RARβ2), pSP(RARβ4 259mut) (coding sequence corresponds to RARβ4), pSP(RARβ4 448) (coding sequence corresponds to RARβ4 500) and pSP(RARβ4 ΔAF2), pSP(RARβ4 ΔAF3), or pSP(RARβ4 ΔAFα) was incubated with 50 μl of TNT Quick Master Mix (Promega) and either 20 μl of [-35S]methionine (PerkinElmer Life Science Products) or 20 μl methionine according to the manufacturer’s instructions. Translated products were separated by SDS-PAGE followed by autoradiography.

GST Pull-down Assay—BL21 cells (Amersham Biosciences) were transformed with pGEX constructs described above. Purification of GST fusion proteins was performed as recommended by the manufacturer (Amersham Biosciences). Glutathione-agarose beads (Sigma) conjugated with GST(A1B1.T1), GST(C-SMRT), or GST-RXRα were incubated with 0.1% ethanal or 1 μl atRa and 3 μl of in vitro translated [-35S]methionine-labeled RARβ proteins (pre-incubated with 0.1% ethanal or 50 μl atRa for 1 h on ice) in binding buffer (60 mM NaCl, 1 mM EDTA, 20 mM Tris-HCl, pH 8.0, 0.05% IGEPA, 1 mM dithiothreitol, 6 mM MgCl2, 8% glycerol). Mixtures were incubated on ice for 2 h and then washed with NENT buffer (same as the binding buffer, except with 500 mM NaCl). After elution in 2× sample buffer (125 mM Tris-HCl, pH 6.8, 4% SDS, 10% glycerol, 0.1% bromphenol blue), complexes were separated on 12% SDS-PAGE and visualized by autoradiography.

Electrophoretic Mobility Shift Assay (EMSA)—Approximately 0.1–0.5 μg of in vitro translated RARβ4, RARβ2, RARβ4 AFα, RARβ4 AFβ, RARβ4 AFβ α2F2, or RXRs were incubated in binding buffer (100 mM KCl, 25 mM Tris-HCl, pH 8.0, 1 mM dithiothreitol, 1 mM EDTA, 20% glycerol) overnight on ice at 4 °C. The following day, 50,000–80,000 cpm of [-3P]5'-end-labeled, double-stranded RARE (top strand, 5'-GGG TAG GGT-3') was added, and the resulting mixture was incubated on ice for 15–20 min. Complexes were separated on 5% non-denaturing acrylamide gels. For supershift assays, 1 μg of RARβ or RXRα antibody (Santa Cruz Biotechnology, Santa Cruz, CA) was included in the overnight incubation at 4 °C.

Luciferase Assay—Firefly and Renilla luciferase activities were measured on a Zylux FB15 luminometer (Zylux Corp., Maryville, TN) using the Dual-Luciferase Reporter Assay System (Promega). For each experiment, duplicate transfections were assayed. Firefly relative light units were divided by Renilla relative light units and normalized to vector control (without addition of AtRa). The average of the 2–3 experiments was calculated, and statistical significance was determined using the two-tailed Student’s t test for samples with equal variance. Statistical analyses compared the relative luciferase activity obtained for samples with an added component, such as a cofactor, with those obtained for identical samples without the added component.

RESULTS

The RARβ2 and RARβ4 Transcripts Each Generate Two Protein Isoforms—We have previously reported the expression of three discrete RARβ protein isoforms in HMECs and breast tumor cells that express only two RARβ mRNAs (RARβ2 and RARβ4) (Fig. 1) (31). We hypothesized that at least one of the RARβ transcripts directs the synthesis of multiple proteins by the mechanism of leaky scanning (39, 40). Expression plasmids pCS(RARβ2) and pCS(RARβ4) were created containing the entire transcript leader region and coding sequence of each transcript cloned in-frame to six C-terminal Myc epitopes (Fig. 2A). These plasmids were transiently transfected into both HMECs (AG11132) and MCF-7 cells, and soluble protein extracts were analyzed by Western blots using an antibody raised against the Myc epitope. We found that the pCS(RARβ2) and pCS(RARβ4) transcripts each generated two proteins in both cell lines assayed (Fig. 2B). The apparent masses of these proteins were comparable to the endogenous RARβ proteins when the ~14-kDa tag at the C terminus of the ectopically expressed proteins is accounted for. Note that in vitro translated proteins generated from each of these putative translation start sites migrated in SDS-PAGE with one of the three endogenous RARβ proteins found in HMECs or MCF-7 cells (31). The lowest molecular weight protein appeared to be common to both transcripts, whereas the higher molecular weight protein was more specific to the transcript from which it was derived. The lower molecular weight protein was expressed at higher levels than that of the larger protein in MCF-7 cells, and at approximately equivalent amounts in the HMEC lysates (Fig. 2, B and C). The higher molecular weight protein was the predominant protein produced from the RARβ2-derived transcript in both MCF-7 cells and HMECs (Fig. 2, B and C). In identical experiments, these patterns of protein expression directed by transiently transfected pCS(RARβ2) and pCS(RARβ4) plasmids into MCF-7 cells were replicated in two other breast tumor cell lines, MDA-MB-231 and ZR-75-1 (data not shown).

To identify the translation start sites utilized by each transcript, we introduced single mutations into pCS(RARβ2) and pCS(RARβ4) to abolish ribosome recognition of putative translation start sites (Table 1 and Fig. 2A). Mutations were targeted to eradicate translation of either the putative 5′-most translation initiation codon (termed T1 for the first translation initiation site) or the putative downstream initiation codon (T2) in both the RARβ2- and RARβ4-derived transcripts. In both HMECs and MCF-7 cells, mutation of T1, a previously identified translation initiation codon at nucleoside +469 of the RARβ2 transcript, from AUG to AUA altered the ratio of expression of the two proteins. The proportion of the larger protein decreased, whereas that of the smaller protein increased (Fig. 2C). Conversely, mutation of an AUG at +805 of the RARβ2 transcript (T2) resulted in a loss of expression of the lower molecular weight pCS(RARβ2) protein, whereas expression of the higher molecular weight protein was maintained (Fig. 2C). T2 of the RARβ2 mRNA is found within sequence that is identical to a translation start site previously identified in the human RARβ4 transcript (31).

In a similar fashion, mutation of a CUG to a CUU (T1) of the RARβ4 transcript, a CUG analogous to one that
initiates translation of the mouse RARβ4 (26), eradicated the expression of the larger pCS(RARβ4) protein in both HMECs and MCF-7 cells (Fig. 2C). It is interesting to note that mutation of the AUG translation start site to a codon (AUU) that should not initiate translation decreased, but did not completely abolish, the presence of the larger protein. Mutation of the previously identified translation start site at +448 relative to the RARβ4 transcript (T2; equivalent to the AUG located at +805 in the RARβ2 transcript) to AUU abolished the expression of the lower molecular weight protein in both cell lines tested. These results clearly indicated that the RARβ2 and RARβ4 transcripts each produce two RARβ proteins. RARβ2 initiated translation at AUGs located at +469 (T1) and +805 (T2) of the RARβ2 transcript, whereas RARβ4 used a CUG at +259 (T1) and an AUG at +448 (T2) of the RARβ4 transcript for translation initiation (Fig. 2C). The lower molecular weight protein generated at T2 in both RARβ2- and RARβ4-derived transcripts, previously called human RARβ4, is identical on the basis of the coding sequence downstream of T2 in both transcripts; we therefore re-designated this protein as RARβ'. We refer to the protein product of translation initiation at the CUG codon at T1 of the RARβ4 transcript as RARβ4, the analogous isoform to the previously reported mouse RARβ4 protein (26). RARβ' was the primary protein product of the RARβ4-derived mRNA in the MCF-7 breast cancer cell line, whereas RARβ4 and RARβ' were generated at approximately equivalent amounts in HMECs. In contrast, RARβ' was produced at significantly lower levels relative to the RARβ2 protein in both cell types when the RARβ2-derived transcript was expressed (Fig. 2C).

The CUG T1 translation initiation site of RARβ4 does not conform to the Kozak consensus sequence predicting robust recognition by the scanning ribosome (GCC/A/G/C/C/CAUGG, translation initiation codon in bold; Ref. 41). The most important bases for ribosome recognition in this sequence are the initiation codon and the nucleosides in the −3 and +4 positions relative to the start of translation. Since the RARβ4 start site is not an ideal initiation codon, we hypothesized that two proteins were generated from the human RARβ4 transcript by a mechanism known as leaky scanning. To determine whether translation of RARβ' at T2 in the RARβ4 mRNA was a result of leaky scanning past the non-canonical CUG initiation codon at T1 (42), we performed site-directed mutagenesis to create RARβ4 translation initiation sites with a greater identity to the Kozak consensus sequence (Table I; Fig. 3A). As shown in Fig. 3B, mutation of T1 of the RARβ4-derived transcript from a CUG to an AUG resulted in an increased usage of the 5' most translation initiation site along with a decrease in RARβ' translation, consistent with a leaky scanning model. This effect was observed identically in MCF-7 cells and HMECs.

Although the Kozak sequence of the T1 translation initiation site of RARβ2 should be adequate for translation initiation (Table I), we created two mutants of the pCS(RARβ2) construct to determine if improving the Kozak sequence at T1 would abolish translation at T2. The first mutation (T1+) exchanged a guanosine for the uracil at +4 relative to the start of translation, a key nucleoside for translation initiation. The second construct (T1+++) contained mutations altering the entire context of the T1 site to conform exactly to the Kozak sequence (Table I). As shown in Fig. 3B, mutation of the +4 nucleoside from uracil to guanosine (T1+) did not alter the expression of the RARβ' isoform at T2 in either HMECs or MCF-7 cells. However, the RARβ2-derived transcript having the T1 initiation site in a perfect Kozak context (T1++++) displayed no trans-
fore concluded that RAR\(_2\) and RAR\(_4\) T1 abolishes translation initiation downstream at T2. A, diagram of RAR\(_2\)- and RAR\(_4\)-derived transcripts generated from pCS vectors harboring point mutations to enhance ribosome recognition of the putative 5’-most (T1) translation initiation sites. The 5’ transcript leader sequence is represented by a line and protein coding sequence by a box. The positions of the T1 and T2 sites found on each transcript are indicated by arrows. The six Myc tags in-frame with RAR\(_2\) coding sequence at the C terminus is represented by stippled boxes. Plus signs denote constructs having mutations at the T1 site to increase the identity of the surrounding nucleotides to that of the optimal Kozak sequence. Two plus signs designate an RAR\(_2\) T1 mutation that is exactly homologous to the optimal Kozak sequence. The size of the arrowhead, T1 or T2, indicates relative optimal Kozak consensus sequence. B and C, anti-Myc immunoblots derived from whole cell lysates of AG11132 HMEC and MCF-7 breast carcinoma cells transfected with the above pCS plasmids. The plasmid transfected is identified above each sample. The migration of molecular weight markers is indicated to the left of each gel.

**Fig. 3.** Improving the Kozak context of the RAR\(_2\) and RAR\(_4\) T1 abolishes translation initiation downstream at T2. A, diagram of RAR\(_2\)- and RAR\(_4\)-derived transcripts generated from pCS vectors harboring point mutations to enhance ribosome recognition of the putative 5’-most (T1) translation initiation sites. The 5’ transcript leader sequence is represented by a line and protein coding sequence by a box. The positions of the T1 and T2 sites found on each transcript are indicated by arrows. The six Myc tags in-frame with RAR\(_2\) coding sequence at the C terminus is represented by stippled boxes. Plus signs denote constructs having mutations at the T1 site to increase the identity of the surrounding nucleotides to that of the optimal Kozak sequence. Two plus signs designate an RAR\(_2\) T1 mutation that is exactly homologous to the optimal Kozak sequence. The size of the arrowhead, T1 or T2, indicates relative optimal Kozak consensus sequence. B and C, anti-Myc immunoblots derived from whole cell lysates of AG11132 HMEC and MCF-7 breast carcinoma cells transfected with the above pCS plasmids. The plasmid transfected is identified above each sample. The migration of molecular weight markers is indicated to the left of each gel.

**Fig. 3.** Improving the Kozak context of the RAR\(_2\) and RAR\(_4\) T1 abolishes translation initiation downstream at T2. A, diagram of RAR\(_2\)- and RAR\(_4\)-derived transcripts generated from pCS vectors harboring point mutations to enhance ribosome recognition of the putative 5’-most (T1) translation initiation sites. The 5’ transcript leader sequence is represented by a line and protein coding sequence by a box. The positions of the T1 and T2 sites found on each transcript are indicated by arrows. The six Myc tags in-frame with RAR\(_2\) coding sequence at the C terminus is represented by stippled boxes. Plus signs denote constructs having mutations at the T1 site to increase the identity of the surrounding nucleotides to that of the optimal Kozak sequence. Two plus signs designate an RAR\(_2\) T1 mutation that is exactly homologous to the optimal Kozak sequence. The size of the arrowhead, T1 or T2, indicates relative optimal Kozak consensus sequence. B and C, anti-Myc immunoblots derived from whole cell lysates of AG11132 HMEC and MCF-7 breast carcinoma cells transfected with the above pCS plasmids. The plasmid transfected is identified above each sample. The migration of molecular weight markers is indicated to the left of each gel.

**Fig. 3.** Improving the Kozak context of the RAR\(_2\) and RAR\(_4\) T1 abolishes translation initiation downstream at T2. A, diagram of RAR\(_2\)- and RAR\(_4\)-derived transcripts generated from pCS vectors harboring point mutations to enhance ribosome recognition of the putative 5’-most (T1) translation initiation sites. The 5’ transcript leader sequence is represented by a line and protein coding sequence by a box. The positions of the T1 and T2 sites found on each transcript are indicated by arrows. The six Myc tags in-frame with RAR\(_2\) coding sequence at the C terminus is represented by stippled boxes. Plus signs denote constructs having mutations at the T1 site to increase the identity of the surrounding nucleotides to that of the optimal Kozak sequence. Two plus signs designate an RAR\(_2\) T1 mutation that is exactly homologous to the optimal Kozak sequence. The size of the arrowhead, T1 or T2, indicates relative optimal Kozak consensus sequence. B and C, anti-Myc immunoblots derived from whole cell lysates of AG11132 HMEC and MCF-7 breast carcinoma cells transfected with the above pCS plasmids. The plasmid transfected is identified above each sample. The migration of molecular weight markers is indicated to the left of each gel.

**Fig. 3.** Improving the Kozak context of the RAR\(_2\) and RAR\(_4\) T1 abolishes translation initiation downstream at T2. A, diagram of RAR\(_2\)- and RAR\(_4\)-derived transcripts generated from pCS vectors harboring point mutations to enhance ribosome recognition of the putative 5’-most (T1) translation initiation sites. The 5’ transcript leader sequence is represented by a line and protein coding sequence by a box. The positions of the T1 and T2 sites found on each transcript are indicated by arrows. The six Myc tags in-frame with RAR\(_2\) coding sequence at the C terminus is represented by stippled boxes. Plus signs denote constructs having mutations at the T1 site to increase the identity of the surrounding nucleotides to that of the optimal Kozak sequence. Two plus signs designate an RAR\(_2\) T1 mutation that is exactly homologous to the optimal Kozak sequence. The size of the arrowhead, T1 or T2, indicates relative optimal Kozak consensus sequence. B and C, anti-Myc immunoblots derived from whole cell lysates of AG11132 HMEC and MCF-7 breast carcinoma cells transfected with the above pCS plasmids. The plasmid transfected is identified above each sample. The migration of molecular weight markers is indicated to the left of each gel.

**Fig. 3.** Improving the Kozak context of the RAR\(_2\) and RAR\(_4\) T1 abolishes translation initiation downstream at T2. A, diagram of RAR\(_2\)- and RAR\(_4\)-derived transcripts generated from pCS vectors harboring point mutations to enhance ribosome recognition of the putative 5’-most (T1) translation initiation sites. The 5’ transcript leader sequence is represented by a line and protein coding sequence by a box. The positions of the T1 and T2 sites found on each transcript are indicated by arrows. The six Myc tags in-frame with RAR\(_2\) coding sequence at the C terminus is represented by stippled boxes. Plus signs denote constructs having mutations at the T1 site to increase the identity of the surrounding nucleotides to that of the optimal Kozak sequence. Two plus signs designate an RAR\(_2\) T1 mutation that is exactly homologous to the optimal Kozak sequence. The size of the arrowhead, T1 or T2, indicates relative optimal Kozak consensus sequence. B and C, anti-Myc immunoblots derived from whole cell lysates of AG11132 HMEC and MCF-7 breast carcinoma cells transfected with the above pCS plasmids. The plasmid transfected is identified above each sample. The migration of molecular weight markers is indicated to the left of each gel.

**Fig. 3.** Improving the Kozak context of the RAR\(_2\) and RAR\(_4\) T1 abolishes translation initiation downstream at T2. A, diagram of RAR\(_2\)- and RAR\(_4\)-derived transcripts generated from pCS vectors harboring point mutations to enhance ribosome recognition of the putative 5’-most (T1) translation initiation sites. The 5’ transcript leader sequence is represented by a line and protein coding sequence by a box. The positions of the T1 and T2 sites found on each transcript are indicated by arrows. The six Myc tags in-frame with RAR\(_2\) coding sequence at the C terminus is represented by stippled boxes. Plus signs denote constructs having mutations at the T1 site to increase the identity of the surrounding nucleotides to that of the optimal Kozak sequence. Two plus signs designate an RAR\(_2\) T1 mutation that is exactly homologous to the optimal Kozak sequence. The size of the arrowhead, T1 or T2, indicates relative optimal Kozak consensus sequence. B and C, anti-Myc immunoblots derived from whole cell lysates of AG11132 HMEC and MCF-7 breast carcinoma cells transfected with the above pCS plasmids. The plasmid transfected is identified above each sample. The migration of molecular weight markers is indicated to the left of each gel.

**Fig. 3.** Improving the Kozak context of the RAR\(_2\) and RAR\(_4\) T1 abolishes translation initiation downstream at T2. A, diagram of RAR\(_2\)- and RAR\(_4\)-derived transcripts generated from pCS vectors harboring point mutations to enhance ribosome recognition of the putative 5’-most (T1) translation initiation sites. The 5’ transcript leader sequence is represented by a line and protein coding sequence by a box. The positions of the T1 and T2 sites found on each transcript are indicated by arrows. The six Myc tags in-frame with RAR\(_2\) coding sequence at the C terminus is represented by stippled boxes. Plus signs denote constructs having mutations at the T1 site to increase the identity of the surrounding nucleotides to that of the optimal Kozak sequence. Two plus signs designate an RAR\(_2\) T1 mutation that is exactly homologous to the optimal Kozak sequence. The size of the arrowhead, T1 or T2, indicates relative optimal Kozak consensus sequence. B and C, anti-Myc immunoblots derived from whole cell lysates of AG11132 HMEC and MCF-7 breast carcinoma cells transfected with the above pCS plasmids. The plasmid transfected is identified above each sample. The migration of molecular weight markers is indicated to the left of each gel.
experiments were performed using two different concentrations of atRA for induction: a pharmacological dosage of 1 μM and one approximating physiological levels at 0.01 μM. Interestingly, both dosages of atRA produced equivalent results, including degree of atRA induction (2-3-fold), for all samples transfected with RARβ2 or RARβ4 (Fig. 6). In the following experiments, we report results of transfections in the MCF-7 cell line, although equivalent results were obtained in the other two cell lines tested. We also induced reporter gene expression using 1 μM of atRA.

When RARβ2, RARβ4, and RARβ2’ expression vectors were transfected individually, cells with the highest amounts (10×) of RARβ2 and RARβ4 did not display significantly greater levels of reporter gene activity than those with the lowest amount, either in the absence or presence of atRA (Fig. 6). However, the reporter gene activity of the cells transfected with the highest amount of RARβ2’ was slightly lower than that with the lowest amount of RARβ2’ with atRA (Fig. 6, p < 0.05). RARβ2- and RARβ4-transfected cells showed a 1.2–1.5-fold elevated level of reporter gene activity compared with in the absence of atRA (Fig. 6, p < 0.05). With the addition of 1 μM atRA, RARβ2- and RARβ4-transfected cells displayed a 2-fold greater reporter gene activity than vector-transfected cells (p < 0.01). There was no significant difference between RARβ2- and RARβ4-transfected cells in the amount of reporter gene activity, with or without atRA addition. AtRA-treated cells transfected with RARβ2’, however, showed similar levels of reporter gene activity as those transfected with the empty expression vector (Fig. 6).

Co-transfection of increasing amounts of RARβ2’ with either RARβ2 or RARβ4 resulted in a decrease of reporter gene activity, with the highest amounts of RARβ2’ (10×) diminishing the expression of the reporter to the level of the empty vector (Fig. 7, A and B; p < 0.05 and p < 0.01, respectively, in the presence of atRA). The inhibitory effect of RARβ2’ on reporter gene expression was opposite to what was observed with co-transfection of RARβ4. Increasing the amount of RARβ2 co-transfected with constant amounts of RARβ2 resulted in an increase in reporter gene activity (Fig. 7A, p < 0.01 with the highest amounts of RARβ4). Because the AUG to AUA at T2 resulted in an increase in reporter gene activity as those transfected with the empty expression vector (Fig. 6).

RARβ2’ from these plasmids. Additionally, the CUG translation initiation codon of RARβ4 T1 was altered to an AUG to ensure high levels of translation. A Renilla luciferase reporter construct was co-transfected as a measure of transfection efficiency, and total amounts of DNA transfected per sample were equalized using the empty expression vector. The following
noid responsive promoters: pGL[H9252]RARE Luc or p[H9252]2-747luc (consisting of the RAR\[H9252] gene P2 promoter, from \[H11002]747 to \[H11001]155 bp, upstream of a luciferase gene, Ref. 32). Mass cultures of cells stably transfected with either of the two vectors with luciferase gene expression driven by an atRA-inducible promoter displayed significant reporter gene expression without atRA addition and increased reporter gene expression between 2.6- and 6.6-fold after exposure to 1\[H9262]M atRA for 24 h, similar to what is observed in transient transfection assays. Cells transfected with the control vector exhibited a slight decline in luciferase expression with atRA induction (data not shown).

**Negatively Charged Amino Acids within the AF2 Domain Are Required for RAR\[H9252] Inhibition of atRA-activated Transcription**—Evidence that some transcription co-activators are present at limiting concentrations within the cell (47–49) led us to hypothesize that RAR\[H9252] represses transcription by competing with full-length RARs for essential cofactors for transcriptional activation. For RARs, protein interactions with transcription co-activators are dependent upon negatively charged amino acids within the AF2 domain (50). We performed site-directed mutagenesis of the expression plasmids pTag(RAR\[H9252]2) and pTag(RAR\[H9252]/H9252/H11032) to alter codons for the three glutamic acid residues within the AF2 domain to encode the neutral amino acids valine and alanine (see Fig. 8A). These three amino acid substitutions had no effect on RAR\[H9252] binding as a heterodimer with RXR\[H9252] to a RARE in vitro (Fig. 4B, lanes 6–8) when compared with EMSAs with the wild type protein (Fig. 4B, lane 5).

**Fig. 5.** Wild-type RAR\[H9252] isoforms interact with RXR\[H9252] and nuclear receptor cofactors AIB1 and SMRT, whereas mutations within the AF2 domain abolish AIB1 interactions. Presented are autoradiographs of proteins eluted from GST pull-down assays separated by SDS-PAGE. GST proteins were incubated with in vitro translated, \[^{35}S\]methionine-labeled RAR\[H9252]2 (A), RAR\[H9252]4 (B), RAR\[H9252] (C), RAR\[H9252]/H9252 (D), RAR\[H9252]/H9252AF2 (E), or RAR\[H9252]/H9252AF2 (F). The GST proteins used in each assay are indicated at the top of each gel and include a control GST protein (GST), or GST fusion proteins including full-length RXR\[H9252], C-terminal AIB1, or C-terminal SMRT. GST pull-downs were performed either in the absence (−) or presence (+) of atRA. A control sample C, consisting of 15% of the in vitro translated RAR\[H9252] protein utilized in the pull-down assays, was loaded onto the first lane of each gel.

**Fig. 6.** Expression of RAR\[H9252] or RAR\[H9252] but not RAR\[H9252] increases the atRA-induced expression of a luciferase reporter gene in MCF-7 cells. A vector-only control (v) or increasing amounts of expression plasmids of RAR\[H9252] (b2), RAR\[H9252] (b4), or RAR\[H9252] (b') were transiently transfected into MCF-7 cells along with a pGL3-RARE-firefly luciferase reporter plasmid and a Renilla luciferase reporter construct as a transfection control. Transfected cells were treated with the ethanol vehicle (white bars) or atRA (shaded bars). For each sample, the value of the firefly luciferase activity was normalized to that of the Renilla luciferase transfection control. Each bar represents the average of three experiments performed in duplicate, normalized to values obtained by samples transfected with vector alone (without atRA). Error bars represent S.D. *p < 0.05 versus the empty vector and 1 \[H9262]M atRA incubation.
that of the sample, the value of the firefly luciferase activity was normalized to RAR transfections performed for each sample were treated with either the 5-, or 10-fold excess of the RAR, respectively, and are represented here as shorter boxes in the diagram. Two unique amino acids found in RAR are indicated by dark shading at the N terminus of the protein. Included is the amino acid sequence of the AF2 domain (common to all three isoforms) as well as the amino acid changes generated by site-directed mutagenesis to create RAR2ΔAF2 and RAR2ΔAF2. B, MCF-7 cells were transiently transfected with a βRARE-firefly luciferase reporter plasmid, Renilla luciferase reporter, and expression plasmids of RAR2 (b2), RAR2 (b4), RAR2 (b2m), RAR2 (b M), or empty vector (v) as a control. In samples 6–8, 1–5, or 10-fold excess RAR2ΔAF2 plasmid over RAR2 plasmid were co-transfected. Transfected cells were treated with ethanol vehicle (white bars) or 1 μM atRA (shaded bars). For each sample, the value of the firefly luciferase activity was normalized to that of the Renilla luciferase transfection control. Each bar represents the average of three experiments performed in duplicate, normalized to values obtained by samples transfected with vector alone (without atRA). Error bars represent S.D. *, p < 0.05 versus RAR2 without RAR2ΔAF2 co-transfection and induced with 1 μM atRA (sample 2).

5). RAR2 proteins with AF2 domain mutations (RAR2ΔAF2) also behaved similarly to the wild type RAR2 in EMSAs, remaining unable to bind the RARE in vitro, with or without RXRa (Fig. 4B, lanes 4 and 10). GST pull-down assays were performed to examine the effect of the AF2 mutations on heterodimer, co-repressor, and co-activator interactions. As expected, binding of RAR2ΔAF2 and RAR2ΔAF2 proteins to GST-RXRa and GST-SMRT was equivalent between the wild type proteins in the presence and absence of atRA (compare 10-fold excess of the RAR plasmid. C, RXRa (a) expression construct was co-transfected with a 1-, 5-, or 10-fold excess of RAR plasmid. In all panels, statistics were performed comparing samples co-transfected with RAR2 to equivalent samples without RAR2 addition, *, p < 0.05 and ***, p < 0.01.
Fig. 5, panels A and E, and C and F). Unlike their wild type counterparts, RARβΔAF2 and RARβ′ΔAF2 did not bind GST-AIB1 either before or after atRA addition (Fig. 5, panels E and F) confirming that these mutations indeed disrupt co-activator interactions. This conclusion was also suggested by the results of transient transfection assays. An equivalent amount of atRA-induced (βRARE)Luc reporter gene expression was obtained following the introduction of RARβΔAF2 into MCF-7 cells as a vector only control (Fig. 8B).

To determine whether the ability to bind co-activators is a requirement for the repression of retinoic acid-induced transcription, we compared the effect of RARβ′ and RARβ′ΔAF2 expression in transient transfection assays. As seen with RARβΔAF2, transfection of MCF-7 cells with a RARβ′ΔAF2 expression vector did not affect atRA-induced reporter gene expression beyond that observed with the empty expression vector (Fig. 8B). We then co-transfected vectors expressing RARβ2 and 1-, 5-, and 10-fold excess RARβ′ΔAF2 into MCF-7 cells. Whereas increasing RARβ′ represses the retinoic acid induction of the (βRARE)Luc reporter gene, increasing the amount of RARβ′ΔAF2 transfected with RARβ2 had a positive effect on the atRA-induction of luciferase activity (Fig. 8B, p < 0.05 with the highest amounts of RARβ′ΔAF2 in the presence of atRA). This result demonstrated that negatively charged residues within the AF2 domain, which are involved in co-activator binding, are essential for RARβ′ inhibition of RARβ2-activated transcription.

DISCUSSION

We previously noted the presence of three RARβ proteins in cell lines that express only the RARβ2 and RARβ4 transcripts of the RARβ gene (31). We had hypothesized that a post-translational modification might account for the presence of three discrete RARβ proteins where only two RARβ mRNA isoforms were expressed. However, we were unable to detect any differential phosphorylation or glycosylation of the RARβ proteins found in HMECs and MCF-7 cells (data not shown). Here we show that three RARβ proteins are created from the two individual RARβ transcripts found in human breast epithelial cells. Two of these proteins, RARβ2 and RARβ4, are unique to their respective transcripts and represent translation initiation products from their 5′-most translation initiation sites (T1). We also show that the third RARβ protein, previously termed RARβ4 but which we now call RARβ′, is generated from both transcripts by a downstream AUG (T2). As this AUG is located 3′ of exon 5, the only variably spliced exon in RARβ2 and RARβ4, this AUG is present identically in both mRNAs. It is unlikely that the RARβ′ isoform present in our transient transfection assays (Figs. 2 and 3) is a protease cleavage artifact, as the amount of RARβ′ dramatically increases in RARβ2- and RARβ4-derived transcripts having mutations that abrogate translation initiation upstream (Fig. 2C). Additionally, strengthening the Kozak consensus sequence of the RARβ2 and RARβ4 initiation codons (T1) abolishes expression of RARβ′ (T2) from the RARβ2- and RARβ4-derived transcripts (Fig. 3, B and C). These results are contradictory to what one would expect if RARβ′ were a cleavage product of the full-length RARβ2 or RARβ4.

RARβ′ was the primary protein product of the RARβ4-derived transcript found in the MCF-7 breast cancer cell line in our transient transfection assays (Fig. 2, B and C), a result observed in MDA-MB-231 and ZR-75-1 breast cancer cell lines (data not shown). We observed equivalent translation initiation at the RARβ4 mRNA T1 and T2 in HMECs. While RARβ′ was also generated from the RARβ2-derived transcript, it was produced at much lower levels than that of the RARβ2 protein (Fig. 2B). We tested whether leaky scanning played a role in the expression of the downstream translation initiation site. According to the scanning model of translation, after recognition and binding to an RNA cap, the 40 S ribosomal subunit will scan in the 3′ direction along the transcript until it reaches an AUG that is in an appropriate sequence context to act as a signal for translation initiation. Deviations from the Kozak sequence at key positions (the initiating AUG or –3 and +4 relative to the start of translation) may decrease or abolish ribosome recognition of the translation start site. In this event, a proportion of scanning ribosomes will bypass such a translation start site and continue to scan down the mRNA until reaching an AUG within an appropriate Kozak context, a phenomenon termed leaky scanning (reviewed in Ref. 39). In the case of the RARβ4 mRNA, we found that leaky scanning is likely a consequence of the weak recognition by ribosomes of the CUG utilized at T1 as a translation start site. There is a markedly increased usage of this T1 upon mutation to AUG, and a resulting loss of translation downstream at T2 (Fig. 3B). Non-AUG codons only rarely initiate translation, and are inefficiently recognized by the scanning ribosome. When non-AUG codons are utilized to initiate translation, translation at the first AUG codon is usually also observed as well. It has been proposed that translation initiation at non-AUG codons may be a mechanism of creating an N-terminally extended version of the encoded protein (42).

Leaky scanning may also have a role in the production of RARβ′ from the RARβ2 mRNA, as mutation of the nucleosides surrounding the RARβ2 translation start site to exactly conform to the Kozak consensus sequence abolished translation initiation downstream. It is interesting to note that the RARβ2-derived transcript having an AUG to AU4 mutation at the T1 initiation codon did not abolish translation at T1 as expected, although it did result in greater usage of the downstream T2 as expected by the leaky scanning model (Fig. 2C). We are currently investigating the possibility that structural or sequence elements in the transcript leader sequence are also involved in the RARβ2 mRNA T1 site recognition. We do not yet understand the mechanism by which the endogenous RARβ′ isoform is expressed in breast tumor cells at higher levels than in normal HMECs. It has been demonstrated previously that the ratio of C/EBPα and C/EBPβ protein isoforms that are generated by leaky scanning can be dramatically shifted during mammary gland lactation (51) and breast epithelial cell tumorigenesis (52). This shift in translation start site usage is regulated in adipocytes by the eukaryotic initiation factors eIF4F, eIF2α and eIF4E (53). eIF4E is up-regulated in a high percentage of primary human breast tumors (54), and the levels of eIF4E expression in breast tumors are thought to be of prognostic significance in determining the risk of tumor recurrence (55). It is interesting to note that the presence of evolutionarily conserved uORFs in the C/EBPα and C/EBPβ transcripts are required for proper regulation of isoform expression by leaky scanning (56). As upstream open reading frames of the RARβ2 transcript leader region have been shown to regulate tissue-specific expression of the mouse RARβ2 protein at the level of translation initiation (56, 57), future experiments will also be performed to identify whether the levels of RARβ isoforms are regulated by the cell at the level of translation initiation.

Structurally, the RARβ4 protein is identical to RARβ2 except that it lacks the N-terminal 49 amino acids that make up the AF-1 domain of RARβ2. Also identical to RARβ2 at the C terminus, RARβ′ lacks the first 113 amino acids of RARβ2. This 113-amino acid region contains the AF-1 domain and one of two zinc finger motifs of the DNA-binding site. Other naturally occurring transcription factor isoforms have been described that also lack a DNA-binding domain. Most analogous
to RAR\(\beta\) is the 11.4 kb progesterone receptor C mRNA that encodes an N-terminally truncated progesterone receptor. Like RAR\(\beta\), progesterone receptor C is made up of amino acids that comprise the dimerization, ligand-binding, and ligand-activated transactivation domains of the full-length progesterone receptor but lacks all peptides N-terminal to the second zinc finger of the DNA-binding domain (58). Another example is the Id family of proteins. Id (or “inhibitor of DNA binding”) proteins are closely related to the basic helix-loop-helix family of transcription factors that play key roles in cell type determination and differentiation. Id proteins contain a conserved helix-loop-helix dimerization motif but lack an upstream basic domain that mediates DNA binding (59). Id proteins inhibit DNA binding of other basic helix-loop-helix transcription factors to their DNA elements and inhibit transactivation mediated by basic helix-loop-helix transcription factors in reporter gene assays (59).

In transient transfection assays, atRA induced expression directed by the retinoid-inducible promoter −2-fold in cells having only endogenous RARs and −3.4-fold in cells that were transfected with full-length RARs (e.g. Fig. 6). In comparison, and representative of the literature, Durand et al. (50) report an increase in atRA activation of a DR-5 RAR\(\alpha\)-activated promoter from 2-fold to 15–30-fold after ectopic expression of RAR\(\alpha\)s into COS-1 cells. There are several lines of evidence to suggest that the amplitude of atRA-induced gene expression that we find in our transient transfection assays is likely an accurate reflection of the degree of atRA-induction of RAR\(\beta\) in the MCF-7 cells and is not a result of transfection or chromosome packaging artifacts. Quantification of endogenous RAR\(\beta\) expression in MCF-7 cells by RNase protection assays revealed a 6-fold increase in the transcript after 3 days of exposure to 1 \(\mu M\) atRA (60). The level of atRA-induced activation that we observe in our cells is consistent between reporter genes that are stable transiently transfected plasmids. Additionally, MCF-7 cells express significant quantities of RAR\(\alpha\), RAR\(\beta\), RAR\(\gamma\), and RXR\(\alpha\) proteins by Western blot analyses (data not shown), and would be expected to activate reporter gene expression without further addition of receptors upon retinoid activation. Reporter gene expression without atRA addition to the cell culture media likely results from activation effected by the presence of endogenous retinoic acids in serum, which are −10 nM for atRA alone in fasting human and rat serum (61). Although treatment of serum with charcoal-dextran reduces the steroid hormones in serum, it only poorly removes the hydrophobic retinoic acids. Note that the ED\(_{50}\) of atRA activation of RAR\(\alpha\)-mediated transcription is 0.6 nM (62) and that addition of atRA to a concentration of only 10 nM induced similar amounts of reporter gene expression as that of the pharmacological dose of 1 \(\mu M\) routinely used in transient transfection assays (Fig. 6). In functional characterizations of the three RAR\(\beta\) isoforms that are expressed in normal and breast tumor-derived mammary cell lines, we find that the RAR\(\beta\)4 protein behaves similarly to RAR\(\beta\)2 in binding the \(\beta\)RARE as a heterodimer with RXR\(\alpha\) (Fig. 4A). Additionally, both RAR\(\beta\)2 and RAR\(\beta\)4 show equivalent transactivation of a reporter gene containing a DR-5 RARE in the presence of physiological (0.01 \(\mu M\)) and pharmacological (1 \(\mu M\)) levels of atRA (Fig. 6). In contrast, RAR\(\beta\)\(\beta\) does not bind the DR-5 RARE in vitro (Fig. 4). Importantly, RAR\(\beta\) expression antagonizes retinoic acid-activated transcription mediated by RAR\(\beta\)2, RAR\(\beta\)4, and another RAR family member, RXR\(\alpha\) (Fig. 7). The inhibition by RAR\(\beta\)\(\beta\) of transcription mediated by other RAR subtypes (Fig. 7) demonstrates that RAR\(\beta\)\(\beta\) may exert a general repression of transcription by other nuclear clear receptors. Because RAR\(\beta\) with mutations in the AF2 cofactor interacting domain failed to repress atRA-induced transcription of RAR\(\beta\)2 (Fig. 8F), it is likely that the ability to compete for transcription cofactors is a requirement for RAR\(\beta\)\(\beta\) inhibition of retinoid-activated transcription.

We hypothesize that the RAR\(\beta\)2 and RAR\(\beta\)4 transcripts generate both positive-acting (RAR\(\beta\)2 and RAR\(\beta\)4 proteins) and negative-acting (RAR\(\beta\)\(\beta\) protein) factors for transactivation at RAREs as a mechanism of fine-tuning atRA-induced transcription. Unlike some other reported dominant negative nuclear receptors (6, 50, 63), RAR\(\beta\)\(\beta\) does not bind cis-acting DNA elements and therefore cannot directly inactivate gene transcription. RAR\(\beta\)\(\beta\) likely represses by stoichiometric competition, away from the RARE, against other transcription factors within the cell (e.g. RAR\(\alpha\), RAR\(\beta\), and RAR\(\gamma\)) for transcription cofactors. Additionally, although the nuclear localization signal is retained in the RAR\(\beta\)\(\beta\) protein, a significant fraction of RAR\(\beta\)\(\beta\) is located in the cytoplasm (31), decreasing the amount of RAR\(\beta\)\(\beta\) available to compete for transcription cofactors.

Although previous reports suggest that RAR\(\beta\)4 mRNA expression may be tumorigenic (64), we find that the RAR\(\beta\)4 protein activated transcription from a RARE at similar levels as RAR\(\beta\)2 in our reporter assays (Figs. 6 and 7). We also find that the RAR\(\beta\)4 transcript is present at higher levels in normal mammary epithelial cells than most breast tumor cell lines (31). It is possible that tumorigenic effects reported for RAR\(\beta\)4 expression may in fact be due to overexpression of the RAR\(\beta\)2 isoform at T2 in the RAR\(\beta\)4 transcript. Whereas the mechanism regulating expression of these proteins in HMECs and breast cancer cells awaits further investigation, it is clear that future expression studies of the RAR\(\beta\) gene products must include an analysis of the RAR\(\beta\) isoforms at the protein level.

Our finding of RAR\(\beta\)\(\beta\) as an inhibitor of retinoic acid-mediated transcriptional activation reveals an additional layer of complexity to retinoid signaling. As the evidence presented here suggests that RAR\(\beta\)\(\beta\) competes with other RARs for transcription co-activators, it would likely inhibit other nuclear receptors that partner with these co-activators, implicating RAR\(\beta\)\(\beta\) as a potential regulator of signaling by other pathways in addition to the RAR family.

Acknowledgments—We thank David Morris and Jonathan Gallant for valuable discussions. We thank Steven Bressler and Jia-ju Zhang for reagents and for technical advice in purifying the GST proteins. We thank Steven Collins for the plasmid L/RXR\(\alpha\)SN, Ronald Evans for plasmid GST(C-SMRT), Paul Meltzer for plasmid pGEX6P(AIB1.T1), and Monica Peacocke and Hui Tsou for the plasmid pEG2-747uc.

REFERENCES

1. Chambon, P. (1996) FASEB J. 10, 940–954
2. Xu, X. C., and Lotan, R. (1999) in Handbook of Experimental Pharmacology (Nau, H., and Blamer, W. S., eds), p. 139, Springer-Verlag, Berlin-Heidelberg
3. Deng, G., Lu, Y., Zlotnikov, G., Thor, A. D., and Smith, H. S. (1996) Science 274, 2057–2059
4. Borrow, J., Geddaard, A. D., Sheer, D., and Solomon, E. (1999) Science 289, 1577–1580
5. de The, H., Chomienne, C., Lanotte, M., Degos, L., and Dejean, A. (1990) Nature 347, 558–561
6. Pratt, M. A., Kralova, J., and McBurney, M. W. (1990) Mol. Cell. Biol. 10, 6445–6453
7. Lawrence, J. A., Merino, M. J., Simpson, J. F., Mannow, R. E., Page, D. L., and Steeg, P. S. (1996) Cancer Epidemiol. Biomarkers Prev. 7, 29–35
8. Picard, E., Seguin, C., Monhoven, N., Rochette-Egly, C., Siat, J., Borrelly, J., Martinet, Y., Martinet, N., and Vignaud, J. M. (1999) J. Natl. Cancer Inst. 91, 1099–1106
9. Widschwendter, M., Berger, J., Daxenbichler, G., Muller-Holzner, E., Widschwendter, A., Mayr, A., Marth, C., and Zeimet, A. G. (1997) J. Surg. Oncol. 65, 4158–4161
10. Xu, X. C., Sneige, N., Liu, X., Nandagiri, R., Lee, J. J., Lukmanji, F., Gebert, J. F., Moghal, N., Frangi, J. V., Sugarbaker, D. J., and Neel, B. G. (1991) Oncogene 6, 1859–1868
11. Qi, H., Zhang, W., El-Naggar, A. K., Lippman, S. M., Lin, P., Lotan, R., and Xu, X. C. (1999) Am. J. Pathol. 155, 1519–1523
