TARP: A nuclear protein expressed in prostate and breast cancer cells derived from an alternate reading frame of the T cell receptor γ chain locus

Curt D. Wolfgang, Magnus Essand*, James J. Vincent, Byungkook Lee, and Ira Pastan†

Laboratory of Molecular Biology, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, Building 37, Room 4E16, Bethesda, MD 20892

Previously, we identified the expression of a prostate-specific form of T cell receptor γ chain (*TCRγ*) mRNA in the human prostate and demonstrated that it originates from epithelial cells and not from infiltrating T lymphocytes. Here, we show that this prostate-specific transcript is also expressed in three breast cancer cell lines and breast cancer tissues. Analysis of the cDNA sequence predicts that this transcript can encode two protein products of 7 and 13 kDa, and *in vitro* translation experiments showed that both proteins were made. The longer ORF encodes a 13-kDa truncated version of TCRγ, whereas the shorter alternative reading frame encodes a 7-kDa protein with five leucine residues in heptad repeats followed by a basic region. Studies with specific antibodies against each protein product revealed that both prostate and breast cancer cells contain only the 7-kDa protein, which is located in the nucleus. We have named this protein TCRγ alternate reading frame protein (TARP). These results demonstrate that an alternative protein product is encoded by the *TCRγ* locus in cells other than T lymphocytes.

expressed sequence tag | leucine zipper | nucleus

P rostate cancer is one of the leading causes of death for men in the United States. However, little is known about the mechanisms behind the oncogenic transformation of prostate cells. To better understand this process, it is important to determine genes that are specifically expressed in prostate cancer cells. In addition, identifying prostate-specific genes will aid in the early detection of prostate cancer as well as the development of the various types of immunotherapy being used to treat cancer. Many different methods are currently being used to identify genes that are tissue specific and/or up-regulated in certain cancers; they include subtractive hybridization (1), differential display (2), serial analysis of gene expression (3), and microarray analysis (4). Our lab has previously described a computer algorithm technique using the expressed sequence tag (EST) database to identify genes that are specifically and highly expressed in human prostate and prostate cancer (5). With this approach, we have identified several genes whose transcripts are specifically expressed in the prostate (5). One is *PAGE-4*, a human X-linked gene that shares homology with the GAGE/MAGE family of melanoma-associated cancer testis antigens (6). Another is the T cell receptor γ chain gene (*TCRγ*) that is derived from the nonrearranged *TCRγ* locus. It consists of a truncated transcript different from the transcript normally detected in lymphoid tissues (7). The expression of *TCRγ* in the prostate has also been detected in a subtraction and microarray analysis (8).

Expression of *TCRγ* in prostate was quite unexpected because expression of the *TCR* genes has been detected only in lymphoid tissues. However, the *TCRγ* transcript found in the prostate originates from epithelial cells of the prostate and not from infiltrating T lymphocytes. By RNA *in situ* hybridization, we showed that *TCRγ* mRNA is highly expressed in epithelial cells within the acinar ducts of the prostate, whereas the stromal cells and other cell types in the prostate are negative (7). Analysis of the prostate *TCRγ* mRNA lead to the discovery that the RNA is derived from a nonrearranged form of the *TCRγ* locus in prostate. The RNA starts within an intron directly upstream of the Jγ1.2 gene segment, contains three exons from the Cγ1 segment, and lacks a Vγ gene segment (Fig. 1*A*). In addition, the *TCRγ* transcripts found in the prostate have different sizes than the transcripts found in the thymus, spleen, and blood leukocytes (7). Two *TCRγ* transcripts are found in the prostate: 1,100 nucleotides (Fig. 1*B*) and 2,800 nucleotides, with the 1,100-nucleotide transcript being the predominant form (7). The 2,800-nucleotide transcript was determined to be a longer version of the 1,100-nucleotide transcript by using a downstream polyadenylation signal (M.E. and I.P., unpublished data). In this paper, we report that these transcripts encode a 7-kDa nuclear protein derived from an alternative reading frame different from the original TCRγ reading frame and are present in prostate and breast cancer cells. We name this protein TCRγ alternate reading frame protein (TARP).

Materials and Methods

Primers. Primers were as follows: TCRγ-upATGmut#1 (5′-TTACAGATAAACAACTTGATACAGATGTTTCCC-CCAAGCCC-3′); TCRγ-upATGmut#2 (5′-GGGCTT-GGGGGAAACATCTGTATCAAGTTGTTTATCTGTAA-3′); TCRγ-upATGmut#3 (5′-GATAAACAACTTGATG-CAGATATTTCCCCCCAAGCCC-3′); TCRγ-upATGmut#4 (5′-GGGCTTGGGGGAAATATCTGCATCAAGTTGTTT-ATC-3′); TCRγ-upATGmut#5 (5′-GATAAACAACTTGAT-ACAGATATTTCCCCCAAGCCC-3′); TCRγ-upATGmut#6 (5′-GGGCTTGGGGGAAATATCTGTATCAAGTTGTTTA-TC-3′); TCRγ-downATGmut#1 (5′-CCCAGGAGGGGAA-CACCATAAAGACTAACGACACATAC-3′); TCRγ-down-ATGmut#2 (5′-GTATGTGTCGTTAGTCTTTATGGTGT-TCCCCTCCTGGG-3′); TCR5.1 (5′-GATAAACAACTTG-ATGCAGATGTTTCC-3′); TCR3.1 (5′-TTATGATTTCTC-
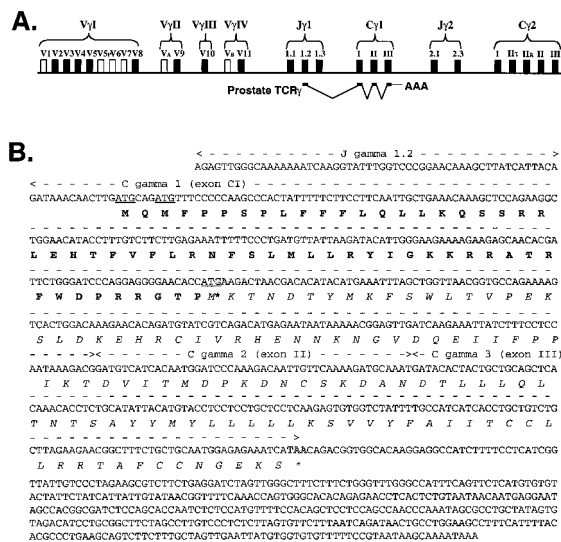

Fig. 1. The prostate-specific *TCRγ* transcript. (*A*) Schematic of the *TCRγ* locus and how the prostate *TCRγ* is transcribed and spliced in prostate cells. The transcript consists of a Jγ1.2 segment, three Cγ1 exons, and an untranslated region. (*B*) Nucleotide and amino acid sequences of the prostate-specific *TCRγ* transcript. The full-length transcript is shown starting with the transcription start site and ending with the polyadenylation signal. Arrows above the corresponding nucleotides indicate exon boundaries. The predicted amino acid sequences for two potential ORFs are noted in bold or italics. Potential initiation methionines are underlined and stop codons are outlined.

TCCATTGCAGCAG-3′);TCRJγ1.2R(5′-AAGCTTTGTTC-CGGGACCAAATAC); B-Actin Forward (5′-ATCTGGCAC-CACACCTTCTACAATGAGCTGCG-3′); B-Actin Reverse (5′-CTTCATACTCCTGCTTGCTGATCCACATCTGC-3′). Primers were synthesized by Genosys (The Woodlands, TX) and Lofstrand Labs (Gaithersburg, MD).

Constructs. The *TARP* transcript cloned into pBluescript II SK(+) (Stratagene) was described (7). This plasmid is referred to as pBSSK-TCRγ in this manuscript. pBSSK-TCR-γmutATGup1, with the ATG at position 69 mutated to ATA, was constructed by using the Quickchange site-directed mutagenesis kit (Stratagene). The PCR used TCRγ-upATGmut#1 and TCRγ-upATGmut#2 as primers and pBSSK-TCRγ as template. pBSSK-TCRγmutATGup2, with the ATG at position 73 mutated to ATA, was constructed as above by using TCRγ-upATGmut#3 and TCRγ-upATGmut#4 as primers and pB-SSK-TCRγ as template. pBSSK-TCRγmutATGup-both, with the ATGs at positions 69 and 73 mutated to ATA, was constructed as above by using TCRγ-upATGmut#5 and TCRγ-upATGmut#6 as primers and pBSSK-TCRγmutATGup1 as template. pBSSK-TCRγmutATGdown, with the ATG at position 242 mutated to ATA, was constructed as above by using TCRγ-downATGmut#1 and TCRγ-downATGmut#2 as primers and pBSSK-TCRγ as template. pET-TCRγ contains nucleotides 242–469 of the *TARP* transcript (7) subcloned into the pET23a vector (Novagen). pET-TARP contains nucleotides 56–242 of the *TARP* transcript (7) subcloned into the pET23a vector. pVC4D-TARP contains nucleotides 69–242 of the *TARP* transcript (7) subcloned into the pVC4D vector (9).

Reverse Transcription–PCR (RT-PCR). Isolation of poly(A) RNA was performed by using the MicroFastTrack 2.0 kit (Invitrogen) according to the manufacturer's instructions. Poly(A) RNA (500 ng) or total RNA (5 μg) was denatured for 2 min at 70°C in the presence of 50 pmol of oligo(dT) primer (Invitrogen). Single-

stranded cDNAs were prepared in a 10-μl reaction mixture containing 250 μM dNTPs, 2 mM DTT, 8 units of RNasin (Roche Molecular Biochemicals, Indianapolis, IN), and 50 units of Superscript II RT (Life Technologies, Rockville, MD) and incubated for 90 min at 42°C. The samples were then diluted with 75 μl of 10 mM Tris·HCl, pH 7.5, and incubated at 72°C for 10 min. cDNA (3 μl) was used for PCR that contained 250 μM dNTPs, 25 pmol of each respective primer, and 1 unit of AmpliTaq DNA polymerase (Roche Molecular Biochemicals) and amplified for 35 cycles. Similar PCR conditions were used on the human breast RAPID-SCAN gene expression panel (OriGene Technologies, Rockville, MD). Primers TCRγJ1.2R, TCR5.1, and TCR3.1 were used to detect the TARP transcript, whereas primers B-Actin Forward and B-Actin Reverse were used to detect the actin transcript.

Northern Blot Hybridization. Northern blot hybridization with 2 μg of poly(A) RNA was performed as described (7).

In Vitro Transcription-Coupled Translation. *In vitro* transcription-coupled translation reactions were described (7). pBSSK-TCRγ, pBSSK-TCRγmutATGdown, pBSSK-TCRγmutATGup1, pBSSK-TCRγmutATGup2, and pBSSK-TCRγmutATGup-both were used as templates.

Cell Culture. LNCaP, PC3, MCF7, BT-474, and SK-BR-3 cells were maintained in RPMI medium 1640 (Quality Biologicals, Gaithersburg, MD) at 37°C with 5% CO_2. The medium contained 10% FBS (Quality Biologicals), 2 mM L-glutamine, 1 mM sodium pyruvate, and penicillin/streptomycin. Hs57Bst cells were maintained in RPMI medium 1640 at 37°C with 5% CO_2. The medium contained 10% FBS, 30 ng/ml epidermal growth factor (Harlan Laboratories, Cincinnati, OH), 2 mM L-glutamine, 1 mM sodium pyruvate and penicillin/streptomycin.

Antibody Production. Polyclonal ΔPE-TARP antibodies were made as follows. pVC4D-TARP, which contains the entire TARP ORF fused to the C′ terminus of a catalytically inactive form of the *Pseudomonas* exotoxin (ΔPE) (9), was expressed in Epicurian Coli BL21-CodonPlus (DE3)-RIL cells (Stratagene). Preparation of inclusion bodies and rabbit immunization were described (10). The antiserum was purified by using the ImmunoPure IgG (protein A) Purification Kit according to the manufacturer's instructions (Pierce).

TCRγ antibodies were made as described above by using pET-TCRγ, an expression plasmid containing the extracellular domain of TCRγ fused to a C′-terminal six-His tag. Before immunization, the His-tagged TCRγ protein was purified by using a Ni-NTA agarose column according to the manufacturer's instructions (Qiagen, Chatsworth, CA).

Preparation of Cell Extracts. Whole-cell protein extracts were prepared as follows. Growing cells (5×10^6) from each cell line were harvested and resuspended in 1× RIPA buffer containing proteinase inhibitors (50 mM Tris·HCl, pH 7.5/150 mM NaCl/1 mM EDTA/0.1% Triton X-100/1 mM PMSF/1 μg/ml aprotinin/1 μg/ml leupeptin). The extracts were sonicated briefly and clarified by centrifugation. Protein concentrations were determined by using the Coomassie Plus Protein Assay reagent according to the manufacturer's instructions (Pierce). Protein extracts from prostate tissue were prepared by grinding 0.5 g of prostate cancer tissue frozen at −80°C into a fine powder by using a cold mortar and pestle. The powdered tissue was collected, resuspended in 1× RIPA, and processed as described above.

Nuclear, membrane, and cytoplasmic extracts from prostate and breast cell lines were prepared based on protocols published (11, 12).
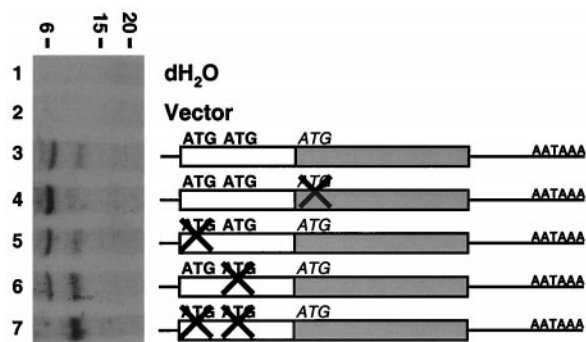
Fig. 2. *In vitro* translation analysis of the prostate-specific *TCRγ* transcript. The prostate-specific *TCRγ* transcript encodes two proteins *in vitro*. [^{35}S]Met-labeled *in vitro*-translated proteins were run on a Tris/Tricine-buffered 16.5% polyacrylamide gel and analyzed by autoradiography. A schematic representation of the mutant constructs used is shown on the right. An open box represents the first reading frame with potential initiation codons in bold, whereas the second reading frame is represented by a shaded box with the potential initiation codon in italics. "X" indicates an ATG codon mutated to ATA. Size markers in kDa are indicated on the top.

Western Blot Analysis. Protein extract (20 or 40 μg) and 1 μg of recombinant His-TARP or 100 ng of recombinant His-TCRγ were run on a 16.5% polyacrylamide gel buffered with Tris/*N*-tris(hydroxymethyl)methylglycine (Tricine) (Bio-Rad) and transferred to a 0.2-μm Immun-Blot poly(vinylidene difluoride) membrane (Bio-Rad) in transfer buffer [25 mM Tris/192 mM glycine/20% (vol/vol) methanol, pH 8.3] at 4°C for 4 h at 30 V. Filters were probed with either 10 μg/ml ΔPE-TARP antiserum or 1 μg/ml TCRγ antiserum, and their respective signals were detected by using a chemiluminescence Western blotting kit according to the manufacturer's instructions (Roche Molecular Biochemicals).

Results

In Vitro Translation of the Prostate-Specific *TCRγ* Transcript. Previously, we identified the expression of *TCRγ* mRNA in the human prostate and demonstrated that it originates from epithelial cells of the prostate and not from infiltrating γδ T lymphocytes (7). Analysis of this transcript revealed two potential reading frames. One encodes a protein with a calculated molecular mass of 7 kDa (Fig. 1*B*; bold amino acids), and one encodes a protein with a calculated molecular mass of 13 kDa (Fig. 1*B*; italicized amino acids). To determine the protein product encoded by this transcript, we analyzed its translational activity by *in vitro* transcription-coupled translation with the full-length prostate *TCRγ* cDNA as template. As demonstrated (7), two protein products of about 7 and 13 kDa were obtained (Fig. 2; lane 3). No specific protein products were obtained when the reactions were done without the *TCRγ* cDNA template or with vector only (Fig. 2; lanes 1 and 2). Therefore, the *in vitro* translation results correlate well with the predicted protein sizes.

To verify that the two protein products observed in Fig. 2 were caused by the use of two independent ORFs, site-directed mutagenesis was performed to mutate individual initiation ATG codons to ATA. Mutation of the third ATG, corresponding to the predicted start of the second ORF, resulted in the loss of the 13-kDa protein product (Fig. 2; lane 4). However, mutation of the first ATG, corresponding to the predicted start of the first ORF, did not result in the loss of the 7-kDa protein product (Fig. 2; lane 5). Analysis of the first ORF revealed a second in-frame ATG codon 2 aa downstream of the first ATG. Surprisingly, mutation of the second ATG did not result in the loss of the 7-kDa protein product (Fig. 2; lane 6). However, mutation of both the first and second ATGs in the first reading frame
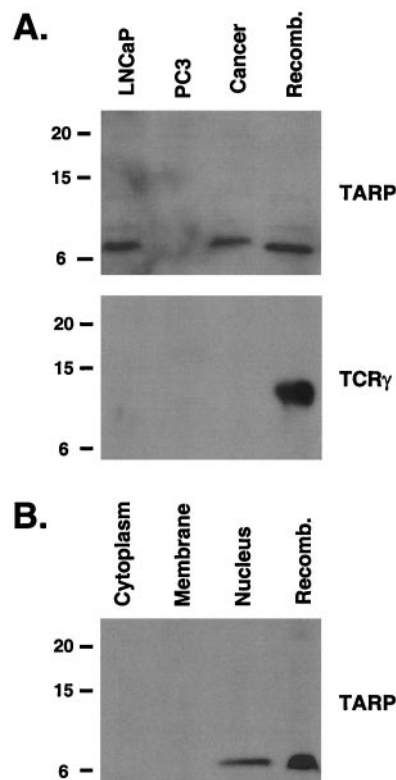

Fig. 3. TARP is nuclear protein expressed in prostate extracts. (*A*) Western blot of protein extracts derived from LNCaP cells, PC3 cells, or a prostate tumor sample (Cancer). Each protein extract (20 μg) was run on a Tris/Tricine 16.5% polyacrylamide gel and probed with an antibody against TARP (*Upper*) or TCRγ (*Lower*). As a positive control, 1 μg of His-tagged TARP (*Upper*) or 100 ng of His-tagged TCRγ (*Lower*) was run on the gel (Recomb.). Size markers in kDa are indicated on the left. (*B*) Western blot of the cytoplasmic fraction (Cytoplasm), membrane fraction (Membrane), and nuclear fraction (Nucleus) of LNCaP cells. Each fraction (40 μg) was run on a Tris/Tricine 16.5% polyacrylamide gel and probed with an antibody against TARP. As a positive control, 1 μg of His-tagged TARP was run on the gel (Recomb.). Size markers in kDa are indicated on the left.

resulted in the loss of the 7-kDa protein. These results indicate that the first reading frame can start at either ATG.

Inspection of the amino acid sequence encoded by the second ORF indicates that it is derived from one of the ATG codons in the original TCRγ reading frame. Therefore, this protein will be referred to as TCRγ in the rest of this paper. On the other hand, the protein product encoded by the first ORF shows no resemblance to any published protein sequence in GenBank. Therefore, this protein will be referred to as <u>T</u>CRγ <u>a</u>lternate <u>r</u>eading frame <u>p</u>rotein (TARP).

TARP Is a Nuclear Protein Expressed in Prostate Cancer Cells. To determine whether TARP or TCRγ exists in prostate cancer cells, we generated antibodies against both proteins and performed Western blot analyses on different prostate cancer cell extracts. As shown in Fig. 3*A Upper*, TARP was detected in the prostate cancer LNCaP cell line and a prostate cancer tumor extract. The 7-kDa band comigrates with the recombinant His-TARP, suggesting that the product detected in the LNCaP and cancer extracts is TARP. Previously, we demonstrated that the prostate-specific *TCRγ* transcript is not expressed in the prostate cancer PC3 cell line (7). Therefore, we used PC3 cell extracts as a negative control and demonstrated that the 7-kDa band was absent in these extracts (Fig. 3*A Upper*). Importantly,
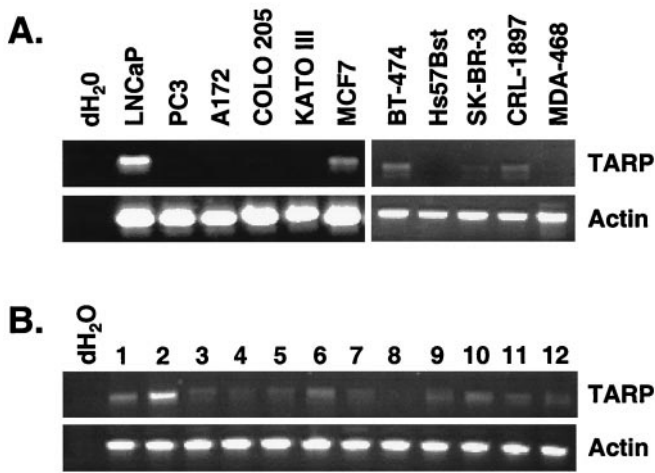
Fig. 4. *TARP* mRNA is expressed in breast cancer cells. (*A*) RT-PCR was performed with primers specific for *TARP* (*Upper*) or *actin* (*Lower*) by using RNA derived from the following cell lines: prostate (LNCaP and PC3), neuroblastoma (A172), colon (COLO 205), gastric (KATO III), and breast (MCF7, BT-474, Hs57Bst, SK-BR-3, CRL-1897, and MDA-468). RT-PCRs performed without template are indicated as dH$_2$O. (*B*) PCR was performed by using cDNAs derived from 12 human breast cancer tissue samples (lanes 1–12) by using primers specific for *TARP* (*Upper*) or *actin* (*Lower*). PCR reactions performed without template are indicated as dH$_2$O. For both panels, 20% of the PCR products were run on a 1% agarose gel and visualized by ethidium bromide staining.

no 7-kDa bands were detected when the prebleed antiserum or an antiserum against the *Pseudomonas* exotoxin (PE; see *Materials and Methods*) was used (data not shown). TCRγ was not detected in any of these extracts even though the recombinant protein showed a very strong signal with the antibody used (Fig. 3*A Lower*). These data indicate that the prostate-specific *TCRγ* transcript encodes TARP.

To determine the cellular localization of TARP, we prepared nuclear, cytoplasmic, and membrane fractions from LNCaP cells. As shown in Fig. 3*B*, TARP was detected in the nucleus and not in the cytoplasm or membrane fraction. Similar results were obtained by using nuclei purified by fractionating the cell extracts through a sucrose cushion (12) (data not shown).

The *TARP* Transcript Is Expressed in Breast Cells. We reported that the *TCRγ* EST cluster also contains some ESTs from brain libraries (5). After this initial report, additional ESTs have been deposited into the database and the cluster now contains ESTs from breast, colon, kidney, and gastric libraries as well. To determine whether the existence of these ESTs indicates the expression of the *TARP* transcript in these cells or it was caused by the presence of infiltrating $\gamma\delta$ T lymphocytes when these libraries were made, we performed RT-PCR on various cell lines to test for the presence of the *TARP* transcript. As shown in Fig. 4*A*, expression of the *TARP* transcript was detected in the breast cell lines MCF7, BT-474, SK-BR-3, and CRL-1897. No signals were detected in the neuroblastoma cell line A172, glioblastoma cell line IMR32, colon cell line COLO 205, gastric cell line KATO III, or kidney cell lines COS7 and 293 (Fig. 4*A* and data not shown). To determine whether the *TARP* transcript is expressed in human breast tissues in addition to cell lines, we tested 12 different normal breast and 12 different breast cancer cDNAs by using a RAPID-SCAN panel (OriGene Technologies, Rockville, MD). *TARP* mRNA was shown to be abundant in some of the breast cancer samples (Fig. 4*B Upper*), whereas it was barely detectable in the normal breast samples after 35 rounds of PCR (data not shown). Significantly, no signals were detected
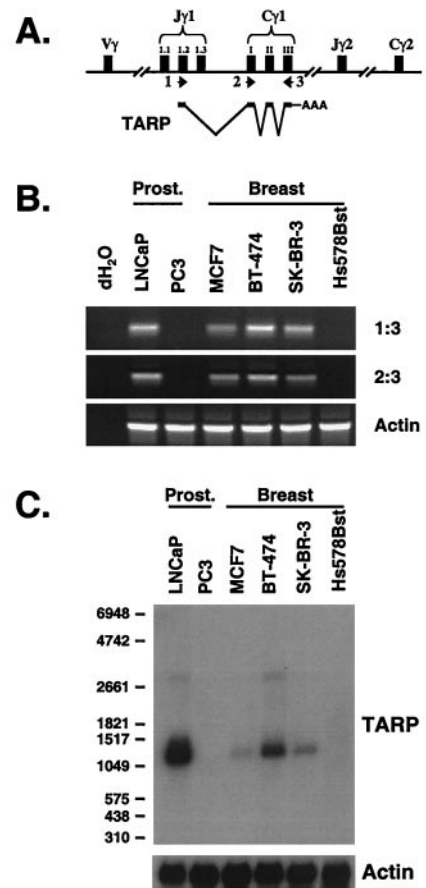

Fig. 5. The *TARP* transcript found in the breast cell line is the same as the prostate-specific form. (*A*) Schematic of the *TCRγ* locus and how *TARP* is transcribed and spliced in prostate cells. Primers used for RT-PCR analysis in *B* are indicated. (*B*) RT-PCR analysis of *TARP* mRNA expression. PCRs with *TARP* primers 1 and 3 (*Top*), *TARP* primers 2 and 3 (*Middle*), or *actin* primers (*Bottom*) were performed with cDNAs derived from prostate cell lines (LNCaP and PC3) and breast cell lines (MCF7, BT-474, SK-BR-3, and Hs578Bst). RT-PCRs performed without template are indicated as dH$_2$O. PCR products (20%) were run on a 1% agarose gel and visualized by ethidium bromide staining. (*C*) Northern blot analysis of *TARP* transcripts. Poly(A) mRNA (2 μg) from prostate cell lines (LNCaP and PC3) and breast cell lines (MCF7, BT-474, SK-BR-3, and Hs578Bst) were analyzed by using a constant domain fragment as probe. The autoradiograph was generated after a 24-h exposure (*Upper*). The same filter was stripped and analyzed with a human β-actin RNA probe to verify equal loading. The autoradiograph was generated after a 1-h exposure (*Lower*). RNA size markers in the nucleotides are indicated on the left.

in reactions lacking cDNA. *actin* was used to show that similar amounts of cDNA were present in each lane (*Lower*). The weak signals in the normal breast samples correlate well with the lack of *TARP* signal shown in Figs. 4A and 5 for the Hs57Bst cell line, a breast cell line derived from normal breast tissue. These results suggest that expression of the *TARP* transcript in the breast is increased after oncogenic transformation. However, more studies are needed before any definitive conclusions can be made.

To determine whether the *TARP* transcript observed in the breast cell lines is the same as the transcript found in the prostate cell line, we performed RT-PCR using primers against different regions of the *TARP* transcript. As shown in Fig. 5*A*, the *TARP* transcript in prostate contains a portion of the Jγ1.2 gene segment, three Cγ1 exons, and some untranslated sequence followed by a poly(A) tail (7). Primer set 1 and 3 amplifies the entire *TARP* transcript (Fig. 5*B Top*), whereas primer set 2 and 3 amplifies the Cγ1 region only (Fig. 5*B Middle*). As shown in Fig.
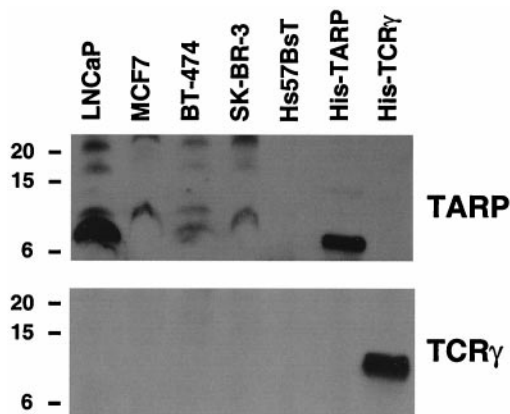
Wolfgang *et al.*

Fig. 6. TARP exists in the nuclei of breast cancer cells. Western blot of nuclear extracts derived from LNCaP, MCF7, BT-474, SK-BR-3, and Hs57BsT cells. Each nuclear extract (40 μg) was run on a Tris/Tricine 16.5% polyacrylamide gel and probed with an antibody against TARP (*Upper*) or TCRγ (*Lower*). As a positive control, 1 μg of His-tagged TARP (His-TARP) and 100 ng of His-tagged TCRγ (His-TCRγ) were run on the gels. Size markers in kDa are indicated on the left.

Fig. 7. Potential functional domains of TARP. (*A*) TARP contains a potential leucine zipper motif and phosphorylation sites. A potential leucine zipper motif is indicated with boxed leucine followed by a basic region that is underlined. cAMP- and cGMP-dependent protein kinase phosphorylation sites (amino acids 46–49 and 55–58) and protein kinase C phosphorylation sites (amino acids 19–21 and 20–22) are outlined. (*B*) Protein sequence comparison of TARP with Tup1. Amino acid sequences for TARP (42–57), *D. discoideum* Tup1 (dTup1, 521–536), and *S. cerevisiae* Tup1 (yTup1, 626–660) are shown. Conserved residues are boxed.

5*B*, similar-sized bands were detected in three breast cell lines (MCR7, BT-474, and SK-BR-3) as compared with the prostate cell line (LNCaP) by using either primer set. Importantly, no signals were detected in the reactions lacking cDNA (dH$_2$O) and similar amounts of cDNA were used as demonstrated by the actin control (Fig. 5*B Bottom*). These data indicate that the *TARP* transcript found in the breast cell lines is the same as the transcript found in the prostate cell line. To further support this conclusion, we analyzed the *TARP* transcript sizes from each cell line by a Northern blot. Previously, we showed that 1,100- and 2,800-nucleotide transcripts exist in LNCaP cells, with the 1,100-nucleotide transcript being the predominant form (7). As shown in Fig. 5*C*, similar-sized *TARP* transcripts were found in three breast cell lines (MCF7, BT-474, and SK-BR-3) when compared with the prostate cell line (LNCaP), although at a weaker intensity. Therefore, we conclude that *TARP* mRNA is expressed in prostate and breast cancer cells.

To determine whether TARP protein exists in the breast cancer cell lines, we performed Western blotting with breast cancer nuclear extracts by using an antibody against TARP. As shown in Fig. 6 *Upper*, TARP-reactive bands were detected in MCF7, BT-474, and SK-BR-3 cells. TARP was not detected in the membrane or cytoplasmic fractions in these breast cancer lines (data not shown). Importantly, TARP is the protein product encoded by the *TARP* transcript in the breast cell lines because TCRγ was not detected in any of these nuclear extracts even though the recombinant protein showed a very strong signal with the antibody used (Fig. 6 *Lower*). These data indicate that TARP also exists in breast cancer cells.

Discussion

We report the identification of a 7-kDa nuclear protein encoded by a specific transcript derived from the *TCRγ* locus expressed in prostate and breast cancer cells. Because the protein is encoded from a reading frame different from TCRγ, we name it TARP for <u>T</u>CRγ <u>a</u>lternate <u>r</u>eading frame <u>p</u>rotein. Besides being translated from an alternate reading frame of a transcript originating within an intron of the *TCRγ* locus, TARP has two other unusual features. First, it is surprising to find such a small peptide in the cell because most are usually secreted. Second, TARP lacks a good Kozak sequence (13). Because the TCRγ reading frame contains a good Kozak sequence, we initially hypothesized that a truncated TCRγ protein was encoded. However, as shown in Fig. 3, our initial hypothesis was incorrect.

It is of interest that the *in vitro* translation results indicate a preference for TARP and that either ATG in the TARP reading frame can be used to initiate protein synthesis. Protein sequencing will be needed to determine which ATG is used to initiate TARP synthesis.

A very interesting feature of the TARP sequence is that it contains five leucines in heptad repeats, suggesting that TARP may contain a leucine zipper dimerization motif (Fig. 7*A*). For this to be true, TARP must contain an amphipathic helix. One indication that TARP may contain an amphipathic helix is that serine and proline residues, which are believed to serve as a helix initiator, are found immediately before the first leucine repeat. Second, many charged amino acids are found within the heptad repeats, thereby giving the helix an amphipathic nature and potentially serving as salt bridges with other helices. Even though the presence of leucine in heptad repeats is a good indication of a leucine zipper motif, there are proteins identified containing five leucines in heptad repeats that are not considered leucine zipper proteins. For example, the crystal structures for karyopherin (14), *Bacillus stearothermophilus* pyrimidine nucleoside phosphorylase (15), and *Thermus thermophilus* phenylalanyl-tRNA synthetase (16) have shown that these proteins do not contain α-helical structures in the region where the sequence contains five leucines in heptad repeats. Interaction and structure studies are needed to determine the significance of the leucine repeats found in TARP.

Another unusual feature of the TARP amino acid sequence is that a region of basic amino acids follows the potential leucine zipper motif (Fig. 7*A*), suggesting a possible DNA-binding motif. However, the orientation of the basic region is unique in that it follows the leucine repeats rather than precedes them. Most leucine zipper proteins that bind DNA have the basic region before the leucine repeats (17). The basic region in TARP may only be functioning as a nuclear localization signal, but the fact that TARP is a nuclear protein strengthens the hypothesis that TARP may bind DNA. Functional studies are needed before any definitive conclusions can be made.

To determine if TARP shares homology with any known proteins, we performed a protein BLAST search against GenBank. This search indicated that the amino acid sequence of TARP shares some homology to *Dictyostelium discoideum* thymidine uptake 1 (Tup1) (GenBank accession no. AAC29438) and *Saccharomyces cerevisiae* Tup1 (18) (Fig. 7*B*). Yeast Tup1 is

normally found in a complex with Cyc8 (Ssn6) and is required for transcriptional repression of genes that are regulated by glucose, oxygen, and DNA damage (19). Neither Cyc8 (Ssn6) nor Tup1 binds DNA, but each acts as a part of a corepressor complex through interactions with specific DNA-binding proteins such as α2, Mig1, Rox1, and **a**1 (19). The C′-terminal half of Tup1 contains six repeats of a 43-aa sequence rich in Asp and Trp, known as WD-40 or β-transducin repeats (18, 20). WD-40 repeats have been identified in many proteins and play a role in protein–protein interactions. Importantly, Tup1 has been shown to interact with α2 through two of its WD-40 repeats (21). It is interesting to note that TARP shares homology with the fifth WD-40 repeat of Tup1 (Fig. 7B). Because TARP is a nuclear protein, its homology with Tup1 suggests that TARP may be a member of a functional nuclear protein complex involved in transcriptional regulation. Therefore, it is necessary to identify TARP-interacting proteins to determine its function.

The TARP antibody recognizes a doublet in prostate and breast nuclear extracts (Fig. 6 *Upper*). The faster 7-kDa band comigrates with the His-TARP recombinant protein, whereas the weaker band runs at a larger molecular mass. One possible explanation for the 9-kDa band is posttranslational modifications. To determine if TARP contains any known posttranslational modification sites, we analyzed the TARP amino acid sequence by using the PROSITE program of the Swiss Institute of Bioinformatics ExPASy proteomics server (http://www.expasy.ch) (22, 23). As shown in Fig. 7A, many potential phosphorylation sites were found, including cAMP- and cGMP-dependent protein kinase phosphorylation sites (RRAT and RRGT) and protein kinase C phosphorylation sites (SSR and SRR). Phosphorylation has been shown in many cases to cause a protein to run at a larger apparent molecular weight on an SDS/PAGE gel. If this is the case, the results from Fig. 6 may indicate that the unmodified form is prevalent in LNCaP cells and that only the phosphorylated form is present in MCF7 and SK-BR-3 cells. Additional experiments are clearly needed to determine the true nature of the 9-kDa band and whether TARP is posttranslationally modified when expressed in prostate and breast cancer cells.

In this paper, we report the expression of TARP mRNA and protein in breast cancer cells. Our initial studies of the *TARP* transcript did not reveal *TARP* expression in the breast (7). One possible explanation is that *TARP* is expressed at low levels in the normal breast and is difficult to detect. As described in *Results*, very weak signals were detected in a PCR analysis of normal breast samples compared with the strong signals detected in the cancer samples. Therefore, the presence of *TARP* in breast cancer cells may indicate that *TARP* expression is induced after the oncogenic transformation of breast cells. In addition, the existence of *TARP* in breast cancer cells may indicate that *TARP* is regulated by estrogen. This hypothesis is strengthened by the identification of an element within the intronic promoter of *TARP* that combines an androgen response element (ARE) with an estrogen response element. This hybrid element consists of two half-sites specific to the ARE at the 5′ end and the estrogen response element at the 3′ end (ref. 24 and unpublished data). Additional experiments are needed to determine if estrogen regulates *TARP*. There are instances, however, where mutant AREs cause the expression of certain prostate-specific genes in breast tumors. For example, prostate-specific antigen has been shown to be expressed in breast tumors (25). Molecular analysis of the aberrant expression of prostate-specific antigen lead to the discovery of a single point mutation in one of the AREs found within the prostate-specific antigen promoter. It is believed that this mutation leads to the loss of androgen-regulated prostate-specific antigen expression in breast tumors (25). It is unclear at this time whether a similar mutation in the *TARP* promoter occurs in the three breast cell lines tested.

The prostate is dependent on androgens for maintenance of its structure and function. When prostate cells become malignant, they often lose their androgen dependence. In this study, we used two prostate cell lines that differ in their dependence on androgen for growth: LNCaP and PC3 cells. The androgen receptor is present in the androgen-dependent LNCaP cell line, but is absent in the androgen-independent PC3 cell line (26). As shown in Fig. 3, *TARP* is expressed in LNCaP cells but not in PC3 cells. This result suggests that *TARP* expression may be regulated by androgen stimulation. The identification of an ARE-like element within the *TARP* promoter strengthens the idea that *TARP* is induced by androgens. Further experiments are needed to determine whether androgens induce *TARP* mRNA expression. Expression in LNCaP cell but not in PC3 cells may indicate that TARP is important in regulating androgen-dependent responses.

1. Hara, T., Harada, N., Mitsui, H., Miura, T., Ishizaka, T. & Miyajima, A. (1994) *Blood* **84,** 189–199.
2. Liang, P. & Pardee, A. B. (1992) *Science* **257,** 967–971.
3. Velculescu, V. E., Zhang, L., Vogelstein, B. & Kinzler, K. W. (1995) *Science* **276,** 1268–1272.
4. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. (1995) *Science* **270,** 467–470.
5. Vasmatzis, G., Essand, M., Brinkmann, U., Lee, B. & Pastan, I. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 300–304.
6. Brinkmann, U., Vasmatzis, G., Lee, B., Yerushalmi, N., Essand, M. & Pastan, I. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 10757–10762.
7. Essand, M., Vasmatzis, G., Brinkmann, U., Duray, P., Lee, B. & Pastan, I. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9287–9292.
8. Xu, J., Stolk, J. A., Zhang, X., Silva, S. J., Houghton, R. L., Matsumura, M., Vedvick, T. S., Leslie, K. B., Badaro, R. & Reed, S. G. (2000) *Cancer Res.* **60,** 1677–1682.
9. Bruggemann, E. P., Chaudhary, V., Gottesman, M. M. & Pastan, I. (1991) *BioTechniques* **10,** 202–209.
10. Brinkmann, U., Pai, L. H., FitzGerald, D. J., Willingham, M. & Pastan, I. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 8616–8620.
11. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. (1983) *Nucleic Acids Res.* **11,** 1475–1489.
12. Sladek, F. M., Zhong, M., Lai, E. & Darnell, J., J. E. (1990) *Genes Dev.* **4,** 2353–2365.
13. Kozak, M. (1986) *Cell* **44,** 283–292.
14. Chook, Y. M. & Blobel, G. (1999) *Nature (London)* **399,** 230–237.
15. Pugmire, M. J. & Ealick, S. E. (1998) *Structure (London)* **6,** 1467–1479.
16. Mosyak, L., Reshetnikova, L., Goldgur, Y., Delarue, M. & Safro, M. G. (1995) *Nat. Struct. Biol.* **2,** 537–547.
17. Hurst, H. C. (1994) *Protein Profile* **1,** 123–168.
18. Williams, F. E. & Trumbly, R. J. (1990) *Mol. Cell. Biol.* **10,** 6500–6511.
19. Tzamarias, D. & Struhl, K. (1995) *Genes Dev.* **9,** 821–831.
20. Fong, H. K. W., Hurley, J. B., Hopkins, R. S., Miake-Lye, R., Johnson, M. S., Doolittle, R. F. & Simon, M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2162–2166.
21. Komachi, K., Redd, M. J. & Johnson, A. D. (1994) *Genes Dev.* **8,** 2857–2867.
22. Appel, R. D., Bairoch, A. & Hochstrasser, D. F. (1994) *Trends Biochem. Sci.* **19,** 248–260.
23. Hofmann, K., Bucher, P., Falquet, L. & Bairoch, A. (1999) *Nucleic Acids Res.* **27,** 215–219.
24. Zilliacus, J., Wright, A. P. H., Carlstedt-Duke, J. & Gustafsson, J.-Å. (1995) *Mol. Endocrinol.* **9,** 389–400.
25. Majumdar, S. & Diamandis, E. (1999) *Br. J. Cancer* **79,** 1594–1602.
26. Tilley, W. D., Wilson, C. M., Marcelli, M. & McPhaul, M. J. (1990) *Cancer Res.* **50,** 5382–5386.