Advances in Brief

WWOX, a Novel WW Domain-containing Protein Mapping to Human Chromosome 16q23.3–24.1, a Region Frequently Affected in Breast Cancer[1]

Andrzej K. Bednarek, Kendra J. Laflin, Rachael L. Daniel, Qiaoyin Liao, Kathleen A. Hawkins, and C. Marcelo Aldaz[2]

Department of Carcinogenesis, University of Texas M. D. Anderson Cancer Center, Science Park Research Division, Smithville, Texas 78957

Abstract

Studies were conducted with the final goal of identifying genes of interest mapping to the chromosome region 16q23.3–24.1, an area commonly affected by allelic losses in breast cancer. To this end we generated a detailed physical map of the genomic region spanning between sequence-tagged site markers D16S518 and D16S516. To identify candidate genes, we used shotgun genomic sequencing as well as isolation and analysis of transcripts mapping to the area of interest. We identified and cloned a novel gene, the genomic structure of which spans the whole region of interest. We named this gene *WWOX* because it contains two WW domains coupled to a region with high homology to the short-chain dehydrogenase/reductase family of enzymes. The ORF of *WWOX* is 1245 bp long, encoding a 414-amino acid protein. This gene is composed of nine exons. We performed a mutation screening of *WWOX* exons in a panel of breast cancer lines, most of which are hemizygous for the 16q genomic region indicated. We found no evidence of mutations, thus indicating that *WWOX* is probably not a tumor suppressor gene. However, we observed that one case of homozygous deletion as well as two previously described translocation breakpoints map to intronic regions of this gene. We speculate that *WWOX* may span the yet uncharacterized common fragile site FRA16D region. In expression studies we found overexpression of *WWOX* in breast cancer cell lines when compared with normal breast cells and tissues. The highest normal expression of *WWOX* was observed in hormonally regulated tissues such as testis, ovary, and prostate. This expression pattern and the presence of a short-chain dehydrogenase/reductase domain and specific amino acid features suggest a role for WWOX in steroid metabolism. Interestingly, the presence of WW domains in the structure of WWOX indicate the likelihood that this protein physically interacts with other proteins. The unique features of WWOX and its possible association with cancer processes make it an interesting target for further investigation.

Introduction

Chromosomal and genomic abnormalities affecting chromosome 16q have frequently been reported in cytogenetic and allelotypic studies of various epithelial tumors. We and others have demonstrated that LOH,[3] affecting the long arm of this autosome, is often observed in breast carcinomas and preinvasive breast lesions (1–5). Other tumor types, such as prostate and hepatic carcinomas, also exhibit similar abnormalities (6, 7). Due to these findings, considerable research

effort has been invested in an attempt to identify the putative tumor suppressor gene(s) that may reside in the distal portion of ch 16q (8–10). In a previous study we observed that one of the most commonly affected areas spanned the region between STS markers D16S515 and D16S504 with the most affected marker being D16S518 at 16q23.3–24.1 (11). The high incidence of LOH observed at preinvasive stages of tumor development led us to speculate that a candidate tumor suppressor gene or genes located at 16q23.3–24.1 may play an important role in early breast carcinogenesis (11).

In this report we describe the physical map of the region of interest between STS markers D16S518 and D16S516 and the cloning of a novel protein from within this area.

Materials and Methods

BAC Identification and Development of STSs and DNA Sequencing. YAC and BAC clones spanning the region of interest were identified by PCR screening of STSs. BACs were isolated from BAC library CITB-HSP-C (Research Genetics, Huntsville, AL). To generate novel STSs for contig building, BAC ends were sequenced using SP6 and T7 vector primers.

Genomic DNA shotgun sequencing was performed using DNaseI BAC DNA digestion and subsequent cloning into pZErO-1 vector (Invitrogen). Inserts were amplified with vector-specific primers. Cycle sequencing reactions were performed using ABI PRISM BigDye Terminator cycle sequencing chemistry (Perkin-Elmer/Applied Biosystems) and analyzed on an ABI 377 automated fluorescent sequencer (Perkin-Elmer/Applied Biosystems). When necessary, clones were sequenced manually with ^{32}P-labeled primers.

Isolation of Candidate cDNAs from the Region of Interest. cDNA clones were isolated following a modification of a solution hybrid capture method described by Futreal *et al.* (12), using BAC clones from the region of interest, as selector DNA, and isolating cDNA clones from a human mammary gland cDNA library (5′-STRECH; Clontech Laboratories, Inc.). All cloned cDNAs were sequenced and analyzed using the BLAST algorithm, searching all available GenBank human databases. The isolated cDNA clones were mapped back to the corresponding BAC (selector) DNAs and compared with the genomic DNA sequence.

WWOX cDNA Isolation and Exon-Intron Structure Determination. A consensus sequence was generated by alignment of the primary cloned cDNA sequence and matching ESTs. From this sequence primers were prepared to isolate the full-length cDNA. Two independent clones were isolated from a placenta cDNA library panel (Rapid Screen, OriGene Technologies, Inc.). Additionally, a second strategy was followed using the 5′- and 3′-rapid amplification of cDNA ends PCR method on a human mammary gland cDNA library (Marathon-Ready; Clontech) according to the manufacturer's protocol. The cDNAs isolated by this last method were cloned, sequenced, aligned, and compared with the clones isolated from the placenta library to determine the full-length cDNA. Primers for the 5′ and 3′ ends of *WWOX* cDNA were used as a first step to isolate additional BAC genomic clones. After the intron-exon junctions for a specific exon were determined, primers for the next exon were designed. The whole-length cDNA sequence was compared with the genomic sequence to determine the genomic structure of *WWOX*.

Protein Sequence Analysis. The WWOX amino acid sequence was analyzed using the BLASTP and PSI-BLAST algorithms in search for matches or homologies in the GenBank protein databases. The identification of protein family domains was determined by using the Pfam domain models (PFAM:

multiple alignments and profile HMMs of protein domains, release 4.3, The Pfam Consortium, http://pfam.wustl.edu/).

Northern Analysis. Northern blots using 2 μg of poly(A) RNA from breast cancer cell lines and normal human breast epithelium were prepared using standard procedures. The multitissue Northern blot was purchased from Clontech. A 1553-bp *Eco*RI restriction fragment of *WWOX* clone, spanning the 5′ end and amino acid-coding region, was used as probe after labeling with [³²P]dCTP using random priming (Prime It II; Stratagene, La Jolla, CA). The membranes were hybridized in Rapid-hyb buffer (Amersham), followed by washing according to the manufacturer's protocol.

***In Vitro* Translation.** *In vitro* translation was performed using an *in vitro* transcription-translation reticulocyte lysate assay (TNT T7 Quick Coupled Transcription/Translation System; Promega) with full-length *WWOX* cDNA as template. [³⁵S]Methionine-labeled products were analyzed by SDS-PAGE followed by phosphorimager detection.

Mutation Screening. Genomic DNA isolated from a panel of 27 breast cancer cell lines was used to resequence each of the exons of *WWOX*. Primers for individual exon amplification and sequencing are specified in Table 1.

Results

General Characterization of the Chromosome 16q Region of Interest. We have focused on the 16q23.3–24.1 chromosome region as a consequence of our previous LOH analyses (11) and an additional high-resolution allelotypic study using a panel of 27 breast cancer cell lines. The latter study showed a very high incidence of hemizygosity within the area of interest affecting 70–80% of these cell lines and one primary tumor case with a homozygous deletion (13).

We followed two main approaches to characterize the chromosomal region of interest to isolate the putative tumor suppressor gene(s). After building a YAC and BAC contig spanning the D16S518–D16S516 region, we used conventional shotgun sequencing and cDNA isolation.

We isolated numerous cDNA clones from the area using a solution hybrid cDNA capture method (12). Thirty-five cDNAs were isolated and sequenced, of which 17 matched previously isolated ESTs, and 18 showed no matches in any of the GenBank databases. All of the isolated cDNA clones were mapped back to the corresponding BAC DNAs, and their sequences were compared with the genomic DNA sequence to identify evidence of exon-intron structure. Only one of these cDNAs showed such features and is described in the following section (*i.e., WWOX*).

We also sequenced approximately 400,000 bp of the region covered by the overlapping BACs 112B7, 249B4, 286F3, and 36O22 (Fig. 1) including a continuous sequence of 96,371 bp (accession no. AF179633). This genomic sequence was also analyzed for matching

EST clusters from GenBank databases. Of the numerous ESTs identified and analyzed, none showed evidence of ORF or exon-intron structure.

Isolation and Characterization of *WWOX*. After sequencing the mentioned cDNA we isolated two independent full-length cDNA clones from a placenta cDNA library using specific PCR primers that spanned the transcript. These full-length cDNAs showed a consensus sequence of 2264 bp long with a predicted ORF of 1245 bp, a 125-bp-long 5′ UTR, and a 3′ UTR 870 bp long with a polyadenylation signal AATAAA starting at position +2091 (cDNA reported to GenBank under accession no. AF211943). The putative start ATG codon is located within a strong Kozak sequence (TCAGCCatgG), which contains a highly conserved G residue (position +4) and purine (G residue at position −3; Ref. 14). An in-frame stop codon is present −30 bp from the predicted translation start site, indicating that the whole ORF was cloned. We named this gene *WWOX* for the reasons discussed in the following section.

Next we determined the exon structure and exon-intron boundaries of *WWOX* and confirmed the chromosomal location of the gene to the region of interest. To this end various combinations of PCR primers were designed based on the cDNA sequence and then mapped back to the corresponding BACs. Subsequent sequencing of the predicted exons permitted the exon-intron boundaries to be established at the genomic DNA level (Fig. 1).

WWOX is composed of nine exons, ranging in size from 58 to 1060 bp (Table 1). On the basis of analysis of the promoter region (reported to GenBank under accession no. AF212843), we determined that the first exon is localized within a CpG island starting at position −660 and extending into the first intron at 292 bp from the ATG translation start site. This area shows a 63% content of C + G and 8% CpG with the highest percentage within the area from −300 bp to ATG (68 and 11%, respectively).

We also observed that the 3′ end of *WWOX* had high sequence homology to a previous GenBank entry of a human putative oxidoreductase (U13395, locus *9621*).

Interestingly, we found that the physical map of *WWOX* spans the whole region of interest: we mapped exon 1 of *WWOX* to the BAC containing D16S518 and exon 9 and the 3′ UTR to the BAC containing D16S516 (Fig. 1). We estimate that *WWOX* spans a large genomic region of ~1 Mb in size. Although the exact lengths of the intronic portions of this gene were not determined, we based this estimate on the known size of the YAC clones, the average size of BAC clones (~150 kb), and our shotgun sequencing (described above) wherein the

Table 1 *WWOX intron-exon junctions and PCR and sequencing primers*

Exon/intron	Starting position in cDNA	Exon length (bp)	Acceptor splice site[a]	Donor splice site[a]	Intronic primers for mutation screening (5′-3′)	PCR product size (bp)
1	1	232		ACGCCAA/gtaaggggggc	1F-ACTGGATTTCAGCTTCGTGGTCG 1R-TCCGTGGGCTGTGCAGGGTC	334
2	233	65	tttttaacag/TCACACC	GCAGGAG/gtttgtatgt	2F-TTCCCCCTACTTCCTTCTTATATCTGGC 2R-ATCCTCACTCCACCCTATGATCTCATC	373
3	298	58	tgtgtttcag/ATTTGCC	TTGTTGA/gtaagtgtct	3F-ATGGTCTTTACTTCTCCCTGGCAC 3R-ACTTCTGCTAAGATTACAGATACACACTG	244
4	356	179	tttttgggcag/CCATATA	GGAATAG/gtaggctctt	4F-AGTTCTTTCAGGTTTAAGGAATAAGC 4R-TAGATCTAAGTGGATCTCATTATAGCAG	370
5	535	107	taaaccatag/GGTTCGA	AGAATGG/gtaagcgctt	5F-ACTTGGGGTAATTTAAGTGGTGCTC 5R-AACTTTACACACTCCACTGAAATCTCC	296
6	642	89	ctcattgcag/CATAAAG	GAATGT/gtgagtgttcc	6F-ATTAAACAGGGGAATTCCGAC 6R-TCTCCCAATTGTGTTCATCTG	278
7	731	186	tttttttcag/GCCTCTTC	CCCATCG/gtgggtttga	7F-ACATCCATGGATCCCGAAG 7R-TGGTATGAGAAAGGGGATAAGTG	318
8	917	265	tattttttaag/ATTTACA	GTCCATG/gtaagagaac	8F-TGCACCCAGCATTCCTTAGATTTCC 8R-ACCAGACTCATGCCCGCAAG	459
9	1182	1060	ggatttccag/CAACAGG		8F-AAATGACGCCATCTCATCACTCC 8R-TGTTTTCCTGGCATCTACGAGAAG	1259

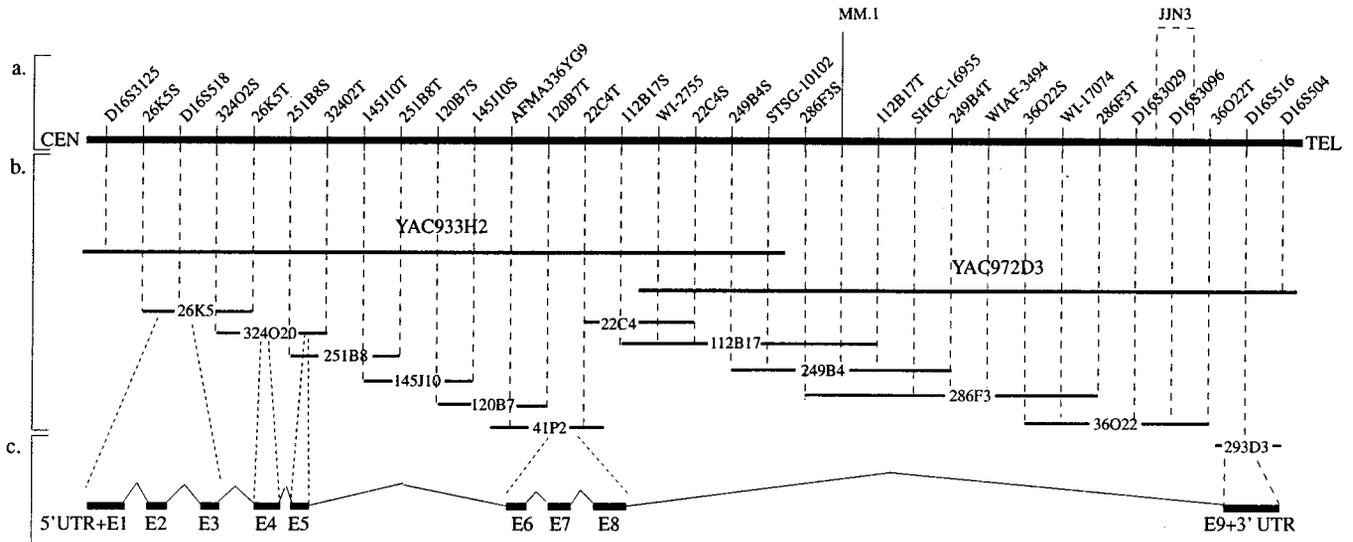[a] Exonic sequences are indicated in uppercase; intronic sequences are indicated in lowercase.

2141

Fig. 1. High-resolution physical map of the 16q23.3–24.1 chromosomal region containing WWOX. *a*, STSs are ordered from centromere (*left*) to telomere (*right*). The relative distance between STS markers is arbitrary. The nomenclature used for the STSs derived from the ends of specific BACs indicates first the BAC clone address followed by *S* or *T*, representing the SP6 or T7 (vector sequencing primers) end of the insert, respectively. *MM.1* and *JJN3* indicate the relative positions of translocation breakpoints observed in myelomas, t(14;16)(q32;q23) (15). Based on the length of unordered contigs obtained from shotgun sequencing, the distance between markers 112B17S and 36O22T is approximately 400 kbp. *b*, The isolated BACs and YACs are represented as *horizontal lines*; *dashed vertical lines* represent the positions of STS markers in corresponding BACs and YACs. The sizes of clones are not to scale. There is only one gap in this BAC contig (between BACs 36O22 and 293D3). The case of homozygous deletion mentioned in "Results" (13) spanned from STS 249B4S to D16S3029, inclusive. *c*, *Black rectangular blocks* represent *WWOX* exons as numbered. The approximate position of each exon in each corresponding BAC clone is illustrated.

~400,000 bp corresponded to the intron 8 area of *WWOX* (Fig. 1). Interestingly, analysis of sequence contigs from this large intronic region allowed us to identify two previously described translocation breakpoints mapping to this same area. These translocation breakpoints, MM.1 and JJN3, have been described in multiple myeloma involving chromosomes 14 and 16, *i.e.,* t(14;16)(q32;q23) (15). Further sequence analysis of this area also identified the existence of a pseudogene for ribosomal protein S3 matching to sequence within this intronic region (AF179633). To our knowledge no other genes have been reported in this area (*i.e.,* D165518–D165516), and we have not identified any other candidates.

WWOX Protein Structure. The 1245-bp *WWOX* ORF encodes for a 414-amino acid protein (Fig. 2). The BLAST and PSI-BLAST algorithms were used to search for matches in GenBank databases. Interestingly, the NH_2 terminus of the putative WWOX protein showed homology to ubiquitin ligases such as NEDD4, YES-associated protein YAP65, and other WW domain-containing proteins (data not shown). Further amino acid sequence analysis using the PROSITE database identified two regions within the putative WWOX protein (amino acids 18–47 and 59–88), which have high homology to WW domain sequences. The first motif exhibits typical features of a WW domain; it is 26 bp long with the two highly conserved tryptophan and one proline residues. In the second WW domain one tryptophan is replaced by a tyrosine residue; this is an alternative functional replacement, which is also found in other WW domain proteins (Fig. 2).

The amino acid sequence also revealed homology to numerous proteins known as members of the SDR family. The SDR family encompasses a wide variety of enzymes, which act on diverse hydroxy and keto substrates. The most conserved features of SDR proteins are two domains constituting the cofactor, G*XXX*G*X*G, and substrate, Y*XXX*K, binding sites (16). Further WWOX amino acid sequence analysis identified both the coenzyme, NAD(H) or NADP(H), binding site **G**ANS**G**IG at position 131–137 and the potential substrate binding site **Y**NRS**K** at positions 293–297 (Fig. 2).

Due to the presence of the WW domains and the homology to SDR, we named this novel protein WWOX. Analysis using the PSORT algorithm predicts that WWOX has no NH_2-terminal signal peptide, and its localization is probably in the cytoplasm (17).

WWOX Expression Analysis. Northern blot analysis with a probe derived from 5′ end of *WWOX* revealed a 2.2-kb mRNA (Fig. 3), which is in agreement with the length of the cDNA clone *i.e.,* 2264 bp.

Analysis of the *WWOX* expression pattern in normal human tissues showed that expression was highest in testis, prostate, and ovary and significantly lower in the other examined tissues, including bulk breast tissues, which showed very low levels of expression (Fig. 3*A*).

We also analyzed the expression of *WWOX* in normal mammary epithelial cells in culture and in breast cancer cell lines. All of the breast cancer lines analyzed showed higher *WWOX* expression than normal breast tissue and normal mammary epithelial cells (HME-87; Fig. 3, *A* and *B*). However, among the breast cancer lines analyzed, *WWOX* expression levels varied from relatively low in T47D and MDA-MB435 to high overexpression in ZR75–1 and MCF-7 cells (Fig. 3*B*). Of these cell lines, previous extensive allelotypic analysis using highly polymorphic STS markers allowed us to determine that T47D, ZR-75, MDA-MB435, MDA-MB453, SKBR3, and UACC812 were among the group of breast cancer lines with no evidence of STS heterozygosity along most or all of 16q (13). This indicated the high likelihood that they had lost all or a large portion of one of the original parental 16q arms including the region spanned by *WWOX*. However, the putative hemizygous status of *WWOX* did not impede expression (Fig. 3*B*). We also found no correlation between the estrogen receptor status of the breast cancer lines and the levels of *WWOX* expression.

To examine whether the translation of the ORF of *WWOX* produced the predicted protein, we used an *in vitro* transcription-translation system. SDS-PAGE analysis of the translated product revealed a single protein product of ~46 kDa (Fig. 3*C*). This agrees with the predicted molecular mass of WWOX based on its amino acid sequence (*i.e.,* 46,676.8 Da).

Mutation Screening in Breast Cancer Cell Lines. As mentioned above, *WWOX* spans the whole chromosomal area of interest between STS markers D16S518 and D16S516. The high incidence of LOH affecting this region led us previously to speculate on the existence of

2142

a)

```
               10         20         30         40
    MAALRYAGLDDTDSEDE LPPGWEERTTKDGWVYYANHTEEKTQWEHP KTG

               60         70         80         90
    KRKRVAGD LPYGWEQETDENGQVFFVDHINKRTTYLDP RLAFTVDDNPTK

              110        120        130        140
    PTTRQRYDGSTTAMEILQGRDFTGKVVVVT GANSGIG FETAKSFALHGAH

              160        170        180        190
    VILACRNMARASEAVSRILEEWHKAKVEAMTLDLALLRSVQHFAEAFKAK

              210        220        230        240
    NVPLHVLVCNAATFALPWSLTKDGLETTFQVNHLGHFYLVQLLQDVLCRS

              260        270        280        290
    APARVIVVSSESHRFTDINDSLGKLDFSRL S PTKNDYWAMLA YNRSK LCN

              310        320        330        340
    ILFSNELHRRLSPRGVTSNAVHPGNMMYSNIHRSWWVYTLLFTLARPFTK

              360        370        380        390
    SMQQGAATTVYCAAVPELEGLGGMYFNNCCRCMPSPEAQSEETARTLWAL

              410
    SERLIQERLGSQSG
```

Fig. 2. *a*, Predicted amino acid sequence of WWOX reveals two WW domains and a short-chain dehydrogenase domain. WW domains are *boxed*, and conserved tryptophans and prolines are shown in *bold*; note that in the second WW domain one tryptophan is replaced by a tyrosine (conservative change). The short-chain dehydrogenase domain is *underlined*, and the conserved residues Y*XXX*K and S, characteristic of a substrate binding site, are highlighted in *black boxes*; bases G*XXX*G*X*G, typical of a coenzyme binding site, are shown in *bold italic*. *b*, Comparative alignment of the WW domains from WWOX *versus* the Pfam WW domain prototypic amino acid sequence. In the consensus sequence, highly conserved amino acids are shown in *uppercase*. The *middle line* illustrates how well the WWOX sequence compares with the WW consensus; satisfactory replacements are represented by +. Numbering corresponds to WWOX amino acid sequence.

b)

```
WW consensus       lppgWeerkdpdGrpYYyNhnTkeTqWekP
                   lppgWeer++ dG +YY Nh    +TqWe+P
WWOX1        18 LPPGWEERTTKDGWVYYANHTEEKTQWEHP  47


WW consensus       lppgWeerkdpdGrpYYyNhnTkeTqWekP
                   lp gWe+ +d++G++++++h++k+T++ +P
WWOX1        59 LPYGWEQETDENGQVFFVDHINKRTTYLDP  88
```

a putative tumor suppressor gene within this area. Thus, to investigate whether *WWOX* is a target for mutations in breast cancer, we performed a mutation screening on a panel of 27 breast cancer cell lines. This panel of breast cancer cell lines was of particular interest, because we have previously observed they exhibit a high incidence of hemizygosity within the chromosome 16q area of interest (13), *i.e.*, cell lines in which the majority had already lost one *WWOX* allele via chromosomal rearrangements. Only one breast cancer case displayed a homozygous deletion in this region (13). We now determined that this deletion spanned part of the intron 8 region of *WWOX* from STS 249B4S to D16S3029 inclusive (see Fig. 1).

Each of the *WWOX* exons was amplified from the genomic DNA of each of the 27 breast cancer lines, and the products were sequenced. The intronic primers used for amplification and sequencing are detailed in Table 1.

We found no evidence of deletions or insertions in the examined DNAs. Two substitutions found appear to have a polymorphic rather than a mutational nature. The first, a C/T found at position −5 in the Kozak sequence, was observed in 50% of the tumor samples, but it was also observed that DNA isolated from normal mammary gland

was polymorphic at this position. The second substitution, a G/A at position +534, results in an amino acid change, alanine to threonine, but because of the approximate frequency of 50% for either allele in the samples examined and because heterozygosity at this position was also found in normal DNA, we concluded that it also represents a polymorphism.

Discussion

In these studies we have outlined the physical map of a 16q chromosomal region commonly affected by abnormalities in breast and other cancers. The area spans from STS marker D16S518 to D16S516 in the 16q23.3–24.1 region. Additionally, we have cloned a novel gene mapping to this area, with unique characteristics. This gene, *WWOX*, contains two WW domains on the NH$_2$ terminus of the protein and a SDR central domain. By analogy to other WW domain-containing proteins, the WW motifs should play a role in protein-protein interactions. The SDR domain is expected to bind low molecular weight ligands and cofactors, and the corresponding putative binding motifs have also been identified in WWOX.
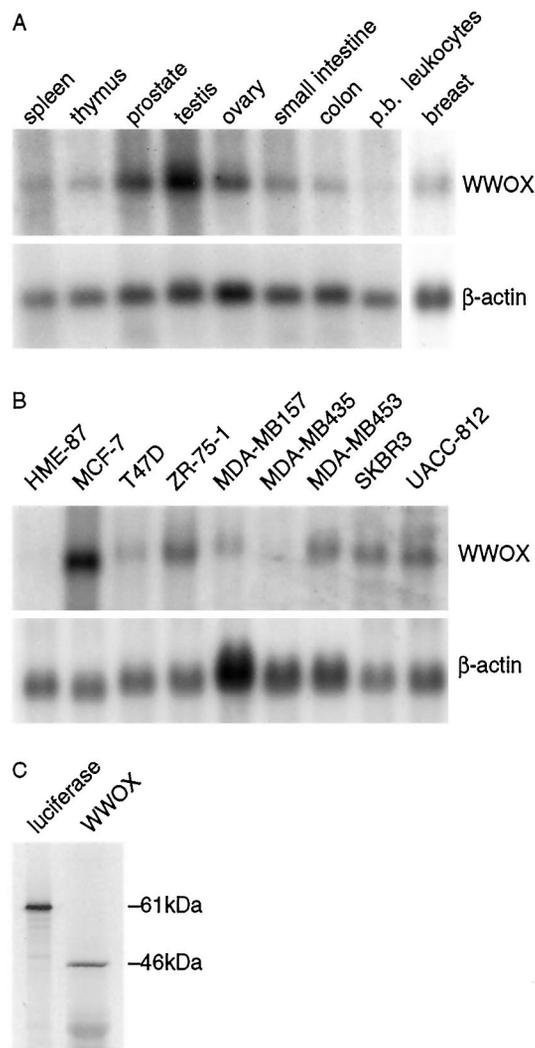
2143

Fig. 3. Expression analysis of WWOX in breast cancer cell lines and normal human tissues and *in vitro* translation of WWOX protein. *A,* Northern blot analysis of poly(A) RNA obtained from normal human tissues as indicated. The WWOX probe detects a mRNA product 2.2 kb in size. Note that the highest WWOX expression is observed in hormonally regulated tissues such as testis, ovary, and prostate. *B,* Northern blot analysis of a poly(A) RNA obtained from human breast cancer cell lines as indicated and normal human mammary epithelial cells (*HME-87*). Both blots were probed with a 1553-bp *Eco*RI fragment of the WWOX cDNA clone containing the 5'UTR and most of its ORF; a *β*-actin probe was used as a control for RNA loading. *C,* The *in vitro* translation of the WWOX cDNA generates a single protein product of 46 kDa in size.

The protein motif called WW or WWP was identified in very different types of proteins, including peptidyl-prolyl isomerases involved in mitosis regulation (18, 19), the cytoskeletal protein dystrophin (20), spliceosome-associated proteins (21), the ubiquitin-protein ligase NEDD4 (22), and signal-transducing protein YAP65 (23). This domain is characterized by the presence of very conserved proline and tryptophan residues (20, 24, 25). WW domains are known to interact with the proline-rich motifs of other proteins. Thus far, four different WW binding proline-rich motifs have been identified: PP*XY* (23), PPLP (26), PGM/PPR (21), and phosphoserine/phosphothreonine (27). At this point it is not possible to predict which type of proline motif the WWOX protein would have affinity for.

SDRs represent a wide spectrum of enzymes. The protein domain database (PROSITE) identifies >60 different proteins from bacteria, fungi, plants, and animals that belong to this family. These are typically enzymes that metabolize different alcohols, sugars, keto-acyls, retinoids, steroids, and other hydroxy and keto substrates. One

important group among the SDR proteins is the family of hydroxysteroid dehydrogenases. The average size of SDR enzymes is 250–300 amino acids. Although overall similarity between the different SDRs can be as low as 15–30%, a small conserved substrate binding motif, Y*XXX*K, and a coenzyme binding motif, G*XXX*G*X*G, are characteristic for these proteins (16). Although the 414 amino acids of WWOX make it larger than the average SDR enzyme, the WWOX dehydrogenase domain exhibits the typical sequence features and distances between conserved motifs that are characteristic of SDR enzymes (28). To our knowledge WWOX is, as yet, the only protein described that contains both binding motifs for low molecular weight ligands and substrates and WW domains.

WWOX has one additional putative signature, which is a serine residue 12 amino acids upstream of the **Y**NRS**K** substrate binding motif. This serine is at a nearly identical location to that observed in steroid dehydrogenases (usually position −13 from Tyr), which is suggested to play an important role in the catalysis of steroid substrates (28).

At this stage it is difficult to predict the role that WWOX may play in the metabolism of the cell. Nevertheless, its unique features make it an interesting target for further investigation. Hydroxysteroid dehydrogenases and reductases usually show a wide tissue expression profile, although some enzymes of this family are tissue specific. Northern analysis has shown that *WWOX* transcripts are highly represented in hormonally active tissues, with testis showing the highest expression. This tissue specificity, in addition to the SDR domain features, leads us to speculate that the WWOX substrate for dehydrogenase and reductase activity is probably a steroid. Because WWOX has the ability to interact with other proteins via the WW domains, the possibility exists that this protein may play a role in steroid-receptor interaction regulation. It is also possible to predict, on the basis of amino acid sequence analysis, that WWOX may localize to the cytoplasm of cells (17).

We found no evidence of *WWOX* mutations in a variety of breast cancer lines suggesting that *WWOX* is not a tumor suppressor gene. However, we found that *WWOX* is overexpressed in breast cancer cells when compared with normal tissues. It is interesting to note that the cell line with the highest *WWOX* mRNA expression is the MCF7 line, which is characterized by its high dependence on estradiol for growth. This invites the speculation that perhaps WWOX plays a role in estradiol-estrogen receptor interaction regulation.

It is also intriguing that *WWOX* spans a chromosomal area characterized by a very high incidence of allelic loss and chromosomal rearrangements. Furthermore, we have mapped two previously described chromosomal breakpoints, MM.1 and JJN3, to the last intron of *WWOX* (Fig. 1). These specific 16q translocation breakpoints, t(14;16)(q32;q23), were previously described in multiple myeloma (15). Consequently, at least one of the alleles of *WWOX* should be truncated in some multiple myeloma lines. Hence, the potential role for WWOX inactivation in multiple myeloma needs to be investigated. In those myeloma studies it was also observed that other additional translocation breakpoints such as KMS11 and ANBL6 (see Ref. 15) also map to the same region covered by YACs 933h2 and 972d3, hence, in very close vicinity of *WWOX*. Nevertheless, the putative oncogenic target for transcriptional dysregulation in the myeloma translocations was proposed to be the c-*maf* oncogene, which is located telomeric of *WWOX* and in the opposite 5'-3' orientation (15). Due to those observations, we investigated whether the c-*maf* oncogene showed any expression alterations in the breast cancer lines shown in Fig. 3. In short, we found no abnormalities of c-*maf* expression when comparing breast cancer lines with normal breast cells and tissues (data not shown). It is also worth mentioning that translocations and deletions affecting ch16q23 have been described as

primary cytogenetic anomalies in several breast cancers (29, 30). In our studies, we found one breast cancer case with a homozygous deletion also mapping within the same intron 8 region in which the translocation breakpoints were mapped. Interestingly, the locus for the yet uncharacterized common fragile site FRA16D has been cytogenetically mapped to this very same chromosome region, 16q23.2–23.3 (31, 32).

All the evidence suggests that the whole genomic region spanned by *WWOX* and, in particular, the intron 8 region appear to be an area prone to chromosomal fragility. Although highly speculative, this area could be the same as the mentioned common fragile site FRA16D, because it displays features of genomic fragility similar to those observed in other better-characterized common fragile site loci (*e.g.,* FRA3B; Refs. 31, 32).

Our identification of *WWOX* and its possible association with cancer processes make it an interesting gene that deserves further investigation.

Acknowledgments

Note Added in Proof

While this manuscript was in press, additional chromosome 16q genomic sequence was reported by the DOE Joint Genome Institute under GenBank Accession number AC009044. This information allowed us to confirm WWOX's genomic structure for exons 1–5 and estimate the length of the corresponding introns as follows: intron 1 (8,538 bp); intron 2 (1,295 bp); intron 3 (5,141 bp); intron 4 (49,019 bp); and intron 5 (>80,000 bp).

References

1. Sato, T., Tanigami, A., Yamakawa, K., Akiyama, F., Kasumi, F., Sakamoto, G., and Nakamura, Y. Allelotype of breast cancer: cumulative allele losses promote tumor progression in primary breast cancer. Cancer Res., *50:* 7184–7189, 1990.

2. Tsuda, H., Callen, D., Fukutomi, T., Nakamura, Y., and Hirohashi, S. Allele loss on chromosome 16q24..2-qter occurs frequently in breast cancer irrespectively of differences in phenotype and extent of spread. Cancer Res., *54:* 513–517, 1994.

3. Cleton-Jansen, A., Moerland, E., Kuipers-Dijkshoorn, N., Callen, D., Sutherland, G., Hansen, B., Devilee, P., and Cornelisse, C. At least two different regions are involved in allelic imbalance on chromosome arm 16q in breast cancer. Genes Chromosomes Cancer, *9:* 101–107, 1994.

4. Aldaz, C. M., Chen, T., Sahin, A., Cunningham, J., and Bondy, M. Comparative allelotype of in situ and invasive human breast cancer: high frequency of microsatellite instability in lobular breast carcinomas. Cancer Res., *55:* 3976–3981, 1995.

5. Chen, T., Sahin, A., and Aldaz, C. Deletion map of chromosome 16q in ductal carcinoma *in situ* of the breast: refining a putative tumor suppressor gene region. Cancer Res., *56:* 5605–5609, 1996.

6. Carter, B. S., Ewing, C. M., Ward, W. S., Treiger, B. F., Aalders, T. W., Schalken, J. A., Epstein, J. I., and Isaacs, W. B. Allelic loss of chromosomes 16q and 10q in human prostate cancer. Proc. Natl. Acad. Sci. USA, *87:* 8751–8755, 1990.

7. Nishida, N., Fukuda, Y., Kokuryu, H., Sadamoto, T., Isowa, G., Honda, K., Yamaoka, Y., Ikenaga, M., Imura, H., and Ishizaki, K. Accumulation of allelic loss on arms of chromosomes 13q, 16q and 17p in the advanced stages of human hepatocellular carcinoma. Int. J. Cancer, *51:* 862–868, 1992.

8. Whitmore, S. A., Crawford, J., Apostolou, S., Eyre, H., Baker, E., Lower, K. M., Settasatian, C., Goldup, S., Seshadri, R., Gibson, R. A., Mathew, C. G., Cleton-Jansen, A. M., Savoia, A., Pronk, J. C., Auerbach, A. D., Doggett, N. A., Sutherland, G. R., and Callen, D. F. Construction of a high-resolution physical and transcription map of chromosome 16q24.3: a region of frequent loss of heterozygosity in sporadic breast cancer. Genomics, *50:* 1–8, 1998.

9. Crawford, J., Ianzano, L., Savino, M., Whitmore, S., Cleton-Jansen, A. M., Settasatian, C., d'apolito, M., Seshadri, R., Pronk, J. C., Auerbach, A. D., Verlander, P. C., Mathew, C. G., Tipping, A. J., Doggett, N. A., Zelante, L., Callen, D. F., and Savoia, A. The PISSLRE gene: structure, exon skipping, and exclusion as tumor suppressor in breast cancer. Genomics, *56:* 90–97, 1999.

10. Savino, M., d'Apolito, M., Centra, M., van Beerendonk, H. M., Cleton-Jansen, A. M., Whitmore, S. A., Crawford, J., Callen, D. F., Zelante, L., and Savoia, A. Characterization of copine VII, a new member of the copine family, and its exclusion as a candidate in sporadic breast cancers with loss of heterozygosity at 16q24.3. Genomics, *61:* 219–226, 1999.

11. Chen, T., Sahin, A., and Aldaz, C. M. Deletion map of chromosome 16q in ductal carcinoma in situ of the breast: refining a putative tumor suppressor gene region. Cancer Res., *56:* 5605–5609, 1996.

12. Futreal, P. A., Cochran, C., Rosenthal, J., Miki, Y., Swenson, J., Hobbs, M., Bennett, L. M., Haugen-Strano, A., Marks, J., Barrett, J. C., Tavtigian, S. V., Shattuck-Eidens, D., Kamb, A., Skolnick, M., and Wiseman, R. W. Isolation of a diverged homeobox gene, MOX1, from the BRCA1 region on 17q21 by solution hybrid capture. Hum. Mol. Genet., *3:* 1359–1364, 1994.

13. Bednarek, A. K., and Aldaz, C. M. Characterization of transcripts from a commonly deleted area of chromosome 16 (q23.3-q24.1) in human breast cancer (Abstract). Proc. Am. Assoc. Cancer Res., *39:* 872, 1998.

14. Kozak, M. Interpreting cDNA sequences: some insights from studies on translation. Mamm. Genome, *7:* 563–574, 1996.

15. Chesi, M., Bergsagel, P. L., Shonukan, O. O., Martelli, M. L., Brents, L. A., Chen, T., Schrock, E., Ried, T., and Kuehl, W. M. Frequent dysregulation of the c-maf proto-oncogene at 16q23 by translocation to an Ig locus in multiple myeloma. Blood, *91:* 4457–4463, 1998.

16. Jornvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J., and Ghosh, D. Short-chain dehydrogenases/reductases (SDR). Biochemistry, *34:* 6003–6013, 1995.

17. Nakai, K., and Kanehisa, M. A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics, *14:* 897–911, 1992.

18. Lu, K. P., Hanes, S. D., and Hunter, T. A human peptidyl-prolyl isomerase essential for regulation of mitosis. Nature (Lond.), *380:* 544–547, 1996.

19. Maleszka, R., Hanes, S. D., Hackett, R. L., de Couet, H. G., and Miklos, G. L. The *Drosophila melanogaster* dodo (dod) gene, conserved in humans, is functionally interchangeable with the ESS1 cell division gene of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA, *93:* 447–451, 1996.

20. Bork, P., and Sudol, M. The WW domain: a signalling site in dystrophin? Trends Biochem. Sci., *19:* 531–533, 1994.

21. Bedford, M. T., Reed, R., and Leder, P. WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: the proline glycine and methionine-rich motif. Proc. Natl. Acad. Sci. USA, *95:* 10602–10607, 1998.

22. Staub, O., Dho, S., Henry, P., Correa, J., Ishikawa, T., McGlade, J., and Rotin, D. WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na$^+$ channel deleted in Liddle's syndrome. EMBO J., *15:* 2371–2380, 1996.

23. Chen, H. I., and Sudol, M. The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. Proc. Natl. Acad. Sci. USA, *92:* 7819–7823, 1995.

24. Andre, B., and Springael, J. Y. WWP, a new amino acid motif present in single or multiple copies in various proteins including dystrophin and the SH3-binding Yes-associated protein YAP65. Biochem. Biophys. Res. Commun., *205:* 1201–1205, 1994.

25. Hofmann, K., and Bucher, P. The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors (Letter). Trends Biochem. Sci., *20:* 347–349, 1995.

26. Chan, D. C., Bedford, M. T., and Leder, P. Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. EMBO J., *15:* 1045–1054, 1996.

27. Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. Function of WW domains as phosphoserine- or phosphothreonine-binding modules. Science (Washington, DC)., *283:* 1325–1328, 1999.

28. Duax, W. L., and Ghosh, D. Structure and function of steroid dehydrogenases involved in hypertension, fertility, and cancer. Steroids, *62:* 95–100, 1997.

29. Dutrillaux, B., Gerbault-Seureau, M., and Zafrani, B. Characterization of chromosomal anomalies in human breast cancer. A comparison of 30 paradiploid cases with few chromosome changes. Cancer Genet. Cytogenet., *49:* 203–217, 1990.

30. Pandis, N., Heim, S., Bardi, G., Idvall, I., Mandahl, N., and Mitelman, F. Whole-arm t(1;16) and i(1q) as sole anomalies identify gain of 1q as a primary chromosomal abnormality in breast cancer. Genes Chromosomes Cancer, *5:* 235–238, 1992.

31. Sutherland, G. R., Baker, E., and Richards, R. I. Fragile sites still breaking. Trends Genet., *14:* 501–506, 1998.

32. Smith, D. I., Huang, H., and Wang, L. Common fragile sites and cancer (Review). Int. J. Oncol., *12:* 187–196, 1998.

WWOX, a Novel WW Domain-containing Protein Mapping to Human Chromosome 16q23.3–24.1, a Region Frequently Affected in Breast Cancer

Andrzej K. Bednarek, Kendra J. Laflin, Rachael L. Daniel, et al.

Updated version	Access the most recent version of this article at: http://cancerres.aacrjournals.org/content/60/8/2140

Cited articles	This article cites 31 articles, 10 of which you can access for free at: http://cancerres.aacrjournals.org/content/60/8/2140.full#ref-list-1
Citing articles	This article has been cited by 51 HighWire-hosted articles. Access the articles at: http://cancerres.aacrjournals.org/content/60/8/2140.full#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/60/8/2140. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.