PERGAMON

Journal of Steroid Biochemistry and Molecular Biology 68 (1999) 41–50

The Journal of Steroid Biochemistry & Molecular Biology

# Early gene expression during androgen-induced inhibition of proliferation of prostate cancer cells: a new suppressor candidate on chromosome 13, in the BRCA2-Rb1 locus

Peter Geck, Jozsef Szelei, Jesus Jimenez, Carlos Sonnenschein, Ana M. Soto *

*Department of Anatomy and Cell Biology, Tufts University School of Medicine, 136 Harrison Ave., Boston, MA 02111, USA*

Received 25 June 1998; received in revised form 20 September 1998; accepted 20 September 1998

**Abstract**

In the prostate gland cell numbers are regulated by androgens through three separate pathways: (a) inhibition of cell death (apoptosis), (b) induction of cell proliferation (step 1), and (c) inhibition of cell proliferation (step 2, proliferative shutoff). The precise regulation of these control pathways is still elusive. The human prostate carcinoma LNCaP cell line variants express a subset of proliferative pathways comparable to those present in normal prostate cells (LNCaP-FGC expresses both steps, LNCaP-LNO expresses step 2, LNCaP-TAC expresses step 1, LNCaP-TJA expresses neither). The purpose of the present work is to identify the genes involved in the androgen-induced proliferative arrest of these cells. Using a Wang–Brown subtracted library, a set of shutoff specific genes has been isolated. One of these new genes, AS3, shows high expression in the early regulatory phase of androgen-induced proliferative shutoff in the cell variants and in the prostates of castrated rats. The putative 1391-residue polypeptide has the molecular size of about 186 kDa. It has coiled-coil structures that usually participate in protein–protein interactions, a perfect leucine-zipper that suggests DNA binding, nuclear localization motifs, proline- and serine-rich domains, unique C-terminal acidic-basic repeats, and ATP- and DNA-binding motifs. The transcript has 34 exons in a 200,000 bp region on chromosome 13q12-q13, downstream of the breast cancer susceptibility gene BRCA2, and centromeric to the retinoblastoma (Rb1) locus. This area is subject to frequent allelic losses in cancers, and is believed to carry a number of cryptic suppressor genes. The AS3 gene seems to be a novel candidate in the regulation of androgen-induced proliferative arrest of human prostate cells. 

*Keywords:* LNCaP cells; Proliferation arrest; Subtractive library; Gene expression

## 1. Introduction

Androgens regulate prostate cell numbers and cell proliferation by three major mechanisms: (a) inhibition of cell death (apoptosis) [1], (b) induction of cell proliferation (step 1), and (c) inhibition of cell proliferation (proliferative shutoff, step 2) [2,3]. Androgens affect epithelial and stromal cell types which, in turn, interact in the prostate [4]. The complexity of the prostate gland precludes a clear understanding of the role of androgen on individual cell types in the whole animal. Thus, we chose for these experiments the human prostate LNCaP-FGC cell line that expresses hormone responsiveness and is used extensively for endocrine and molecular studies [3,5,6]. Proliferation is inhibited in these cells by sex steroid-stripped (charcoal–dextran treated) human serum (CDHuS) [3]. Low androgen concentrations cancel this inhibition (step 1) [3,6] and at higher levels androgens induce an irreversible proliferative shutoff (step 2) [3,6]. During the shutoff period, these cells remain in the $G_0/G_1$ phase of the cell cycle [6]. Prostate specific antigen (PSA) induction, however,

*Abbreviations:* CDHuS, charcoal-dextran treated human serum, R1881, a synthetic, nonmetabolizing androgen, methyltrienolone, R cDNA, cDNA of R1881-treated cells, CD cDNA, cDNA from CDHuS-treated cells, AS, androgen-induced shutoff related sequences, SA, serum-arrest related sequences.

* Corresponding author. Tel.: +1-617-636-6954; fax: +1-617-636-6536; e-mail: asoto@infonet.tufts.edu

0960-0760/99 $ - see front matter 
PII: S0960-0760(98)00165-4

is still dependent on androgens in these postmitotic cells [6]. The notion that the androgen-induced shutoff in LNCaP cells was mediated by TGF-β [7] has recently been refuted by reports on the lack of response to TGF-β by these cells [8].

In an attempt to reduce complexity, androgen target cell lines that express only one of the steps of the androgen regulated proliferative response were established. We reported the isolation of two LNCaP variants: the LNCaP-TAC variant which expresses step 1 only, and the LNCaP-TJA variant, which is resistant to the inhibitory effect of both CD serum and androgens [6]. LNCaP-LNO cells, established by Horoszewicz et al. [5], proliferate maximally in the presence of CDHuS, express an androgen-induced proliferative shutoff, and undergo $G_0/G_1$ arrest (step 2) at high androgen concentrations [6]. In addition to these human prostate cells, a new model to study the shutoff effect was developed by stable transfection of a wild type androgen-receptor construct into breast carcinoma MCF-7 cells. These MCF7-AR1 cells are also able to evoke a proliferative shutoff in response to androgens [9]. The established cell lines, together with recombinant DNA technology, made it possible to explore hypotheses regarding the control of cell proliferation, obviating confounding variables present in intact animals.

Using a differential subtractive amplification procedure [10], we identified a set of genes induced in the proliferative shutoff response (step 2). Since the proliferative shutoff is irreversible, we chose mRNA preparations from LNCaP-FGC cells exposed to 30 nM methyltrienolone (R1881) for 24 h, the minimal exposure time required for the expression of this phenotype [11]. Herein we report the characterization and mapping of a novel gene, AS3. Its expression pattern suggests, but does not prove that it maybe responsible for the induction of the androgen-mediated proliferative shutoff.

## 2. Materials and methods

### 2.1. Cell lines and tissue culture protocols

The LNCaP-FGC cell line (established from a metastatic lymph node of a patient with prostate adenocarcinoma) was generously supplied by Dr. Julius Horoszewicz from the State University of New York at Buffalo [5]. This cell line and the related cell lines mentioned in Section 1 were propagated in 5% FBS, as described earlier [6,12].

### 2.2. The Wang–Brown differential amplification

To isolate androgen-specific, low-abundance regulatory mRNA sequences expressed during the proliferative shutoff, we selected the Wang–Brown approach [10]. Briefly, short fragments of cDNAs were amplified first: then three cycles of subtractions and amplifications between the control and proliferation arrested cDNAs resulted in sequence pools that were differentially expressed [11]. LNCaP-FGC cells were treated with 30 nM R1881 to generate proliferative shutoff. R1881 (methyltrienolone) is a synthetic, nonmetabolized androgen (Roussell-UCLAF, Romainville, France). We reported earlier that exposure to androgen for 24 h was required to commit LNCaP-FGC cells to an irreversible proliferative shutoff [11]. Hence, at this point, the genes responsible for the shutoff must have been highly induced. LNCaP-FGC cells reversibly arrested by CDHuS were considered as the shutoff-negative control; they were harvested after 3 days of CDHuS treatment. Total RNA was prepared by the acidic guanidinium-thiocyanate method [13] and polyA$^+$RNA was purified by using the FastTrack kit (Invitrogen, San Diego, CA). Double-stranded cDNA pools from R1881-treated cells (R cDNA) and CDHuS-treated cells (CD cDNA) were synthesized using the Copy Kit (Invitrogen), with oligo-dT priming. After *Alu*I and *Rsa*I digestions and adaptor ligations, the constructs were PCR- amplified (GeneAmp Kit, Perkin-Elmer, Foster City, CA). The amplified CD cDNA were digested with *Eco*RI, photobiotinylated (driver cDNA) and hybridized at 20-fold molar excess to an aliquot of nonbiotinylated R cDNA. The hybridized nonspecific sequences were eliminated by subsequent Streptavidin chromatography. After 3 cycles of selection, the amplified expressed sequence tag (EST) pool of the androgen-induced shutoff AS (R cDNA pool minus CD cDNA pool) sequences was digested with *Eco*RI, cloned into the BlueScript SK vector (Stratagene, La Jolla, CA) and transformed into *E. coli* (OneShot strain, Invitrogen).

### 2.3. Isolation of unique cDNAs from the differentially expressed sequence pool

Recombinants were collected randomly from the shutoff-positive AS pool of the Wang–Brown differential library and were plated. Using the labeled CD- and R-subtracted (CD cDNA pool minus R cDNA pool), PCR-amplified DNA master mixes as probes, double hybridizations revealed 11 and 14 clones that were present exclusively in the CD and R clone sets, respectively [11]. Multiple cross-hybridizations identified ten unique inserts.

### 2.4. DNA sequence analysis and GenBank search

To sequence the identified EST fragments, PCR sequencing reactions were performed using the dsDNA

(a)

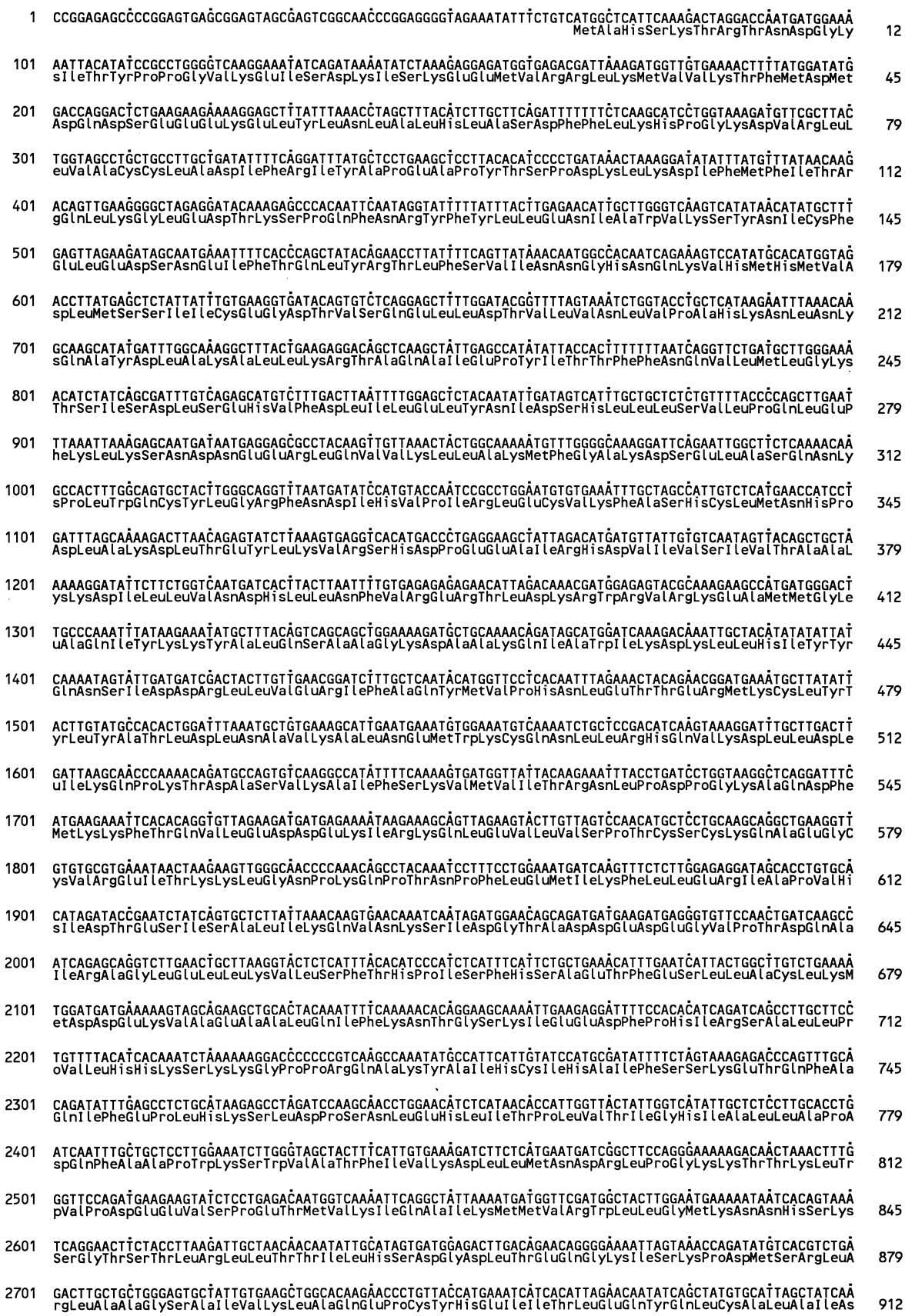


Fig. 1. AS3 cDNA sequence and open reading frame. Numbers on the left indicate positions in base pairs. The amino acid sequence of the open reading frame is depicted under the coding strand. Numbers on the right indicate amino acid positions. Destabilizing signals are underlined, and the polyadenylation signals are highlighted.

(b)

```
2801 CGATGAATGCTATCAAGTAAGACAAGTGTTTGCCCAGAAACTTCACAAAGGCCTTTCCCGTTTACGGCTTCCACTTGAGTATATGGCAATCTGTGCCCTT
     nAspGluCysTyrGlnValArgGlnValPheAlaGlnLysLeuHisLysGlyLeuSerArgLeuArgLeuProLeuGluTyrMetAlaIleCysAlaLeu  945

2901 TGTGCAAAAGATCCTGTAAAGGAGAGAAGAGCTCATGCTAGGCAATGTTTGGTGAAAAATATAAATGTAAGGCGGGAGTATCTGAAGCAGCATGCAGCTG
     CysAlaLysAspProValLysGluArgArgAlaHisAlaArgGlnCysLeuValLysAsnIleAsnValArgArgGluTyrLeuLysGlnHisAlaAlaV  979

3001 TTAGTGAAAAATTATTGTCTCTTCTACCAGAGTATGTTGTTCCATATACAATTCACCTTTTGGCACATGACCCAGATTATGTCAAAGTACAGGATATTGA
     alSerGluLysLeuLeuSerLeuLeuProGluTyrValValProTyrThrIleHisLeuLeuAlaHisAspProAspTyrValLysValGlnAspIleGl 1012

3101 ACAACTTAAAGATGTTAAAGAATGTCTTTGGTTTGTTCTGGAAATATTAATGGCTAAAAATGAAAATAACAGTCACGCTTTTATCAGAAAGATGGTAGAA
     uGlnLeuLysAspValLysGluCysLeuTrpPheValLeuGluIleLeuMetAlaLysAsnGluAsnAsnSerHisAlaPheIleArgLysMetValGlu 1045

3201 AATATTAAACAAACAAAAGATGCCCAAGGACCAGATGATGCAAAAATGAATGAAAAACTGTACACTGTGTGTGATGTTGCCATGAATATCATCATGTCAA
     AsnIleLysGlnThrLysAspAlaGlnGlyProAspAspAlaLysMetAsnGluLysLeuTyrThrValCysAspValAlaMetAsnIleIleMetSerL 1079

3301 AGAGTACTACATACAGTTTGGAATCTCCTAAAGACCCGGTACTACCAGCTCGTTTCTTCACTCAACCTGACAAGAATTTCAGTAACACCAAAAATTATCT
     ysSerThrThrTyrSerLeuGluSerProLysAspProValLeuProAlaArgPhePheThrGlnProAspLysAsnPheSerAsnThrLysAsnTyrLe 1112

3401 GCCTCCTGAAATGAAATCATTTTTCACTCCTGGAAAACCTAAAACAACCAATGTTCTAGGAGCTGTTAACAAGCCACTTTCATCAGCAGGCAAGCAATCT
     uProProGluMetLysSerPhePheThrProGlyLysProLysThrThrAsnValLeuGlyAlaValAsnLysProLeuSerSerAlaGlyLysGlnSer 1145

3501 CAGACCAAATCATCACGAATGGAAACTGTAAGCAATGCAAGCAGCAGCTCAAATCCAAGCTCTCCTGGAAGAATAAAGGGGAGGCTTGATAGTTCTGAAA
     GlnThrLysSerSerArgMetGluThrValSerAsnAlaSerSerSerSerAsnProSerSerProGlyArgIleLysGlyArgLeuAspSerSerGluM 1179

3601 TGGATCACAGTGAAAATGAAGATTACACAATGTCTTCACCTTTGCCGGGGAAAAAAAGTGACAAGAGAGACGACTCTGATCTTGTAAGGTCTGAATTGGA
     etAspHisSerGluAsnGluAspTyrThrMetSerSerProLeuProGlyLysLysSerAspLysArgAspAspSerAspLeuValArgSerGluLeuGl 1212

3701 GAAGCCTAGAGGCAGGAAAAAAACGCCCGTCACAGAACAGGAGGAGAAATTAGGTATGGATGACTTGACTAAGTTGGTACAGGAACAGAAACCTAAAGGC
     uLysProArgGlyArgLysLysThrProValThrGluGlnGluGluLysLeuGlyMetAspAspLeuThrLysLeuValGlnGluGlnLysProLysGly 1245

3801 AGTCAGCGAAGTCGGAAAAGAGGCCATACGGCTTCAGAATCTGATGAACAGCAGTGGCCTGAGGAAAAGAGGCTCAAAGAAGATATATTAGAAAATGAAG
     SerGlnArgSerArgLysArgGlyHisThrAlaSerGluSerAspGluGlnGlnTrpProGluGluLysArgLeuLysGluAspIleLeuGluAsnGluA 1279

3901 ATGAACAGAATAGTCCGCCAAAAAAGGGTAAAAGAGGCCGACCACCAAAACCTCTTGGTGGAGGTACACCAAAAGAAGAGCCAACAATGAAAACTTCTAA
     spGluGlnAsnSerProProLysLysGlyLysArgGlyArgProProLysProLeuGlyGlyGlyThrProLysGluGluProThrMetLysThrSerLy 1312

4001 AAAAGGAAGCAAAAAAAAAATCTGGACCTCCAGCACCAGAGGAGGAGGAAGAAGAAAGACAAAGTGGAAATACGGAACAGAAGTCCAAAAGCAAACAG
     sLysGlySerLysLysLysSerGlyProProAlaProGluGluGluGluGluGluArgGlnSerGlyAsnThrGluGlnLysSerLysSerLysGln 1345

4101 CACCGAGTGTCAAGGAGAGCACAGCAGAGAGCAGAATCTCCTGAATCTAGTGCAATTGAATCCACACAGTCCACACCACAGAAAGGACGAGGAAGACCAT
     HisArgValSerArgArgAlaGlnGlnArgAlaGluSerProGluSerSerAlaIleGluSerThrGlnSerThrProGlnLysGlyArgGlyArgProS 1379

4201 CAAAAACGCCATCACCATCACAACCAAAAAAAAAATGTGTAAGTTGTAAATATTACATTTCAAACCAATTTCAAATTATTTTGCAAAAGTTCCTAAATTTG
     erLysThrProSerProSerGlnProLysLysAsnValEnd                                                          1391

4301 TAAACATACATATTGCTGTATTTAAATTCCATATATTTAGCCCCATTACACTAGGTACGGCGGCGAAGTGCTAAAAGGGAACGGCGATGAACAAATGTAA

4401 TTAATAACTTTCTCTGTGAAAGCTTTGGAAAAATCTTTTTTTTTTTTTTTTTTTTTTGGTCAAGCTTGAGGCTGAATAAAGCCTTTGATGCACAAAATGG

4501 GACTGCTGAAGAGTGGACAGTTGGACCTTACTTTGGTGACCCCATACATTTGTGGTCACATGCTTTAGCCATACACATGGTAACATTGACTATGGAGTCT

4601 TGTGAAAGTGTAATGTGCGATGGCTATGTAGACATAAAGAAGAAACTTGTAAATATCTTTTTTCTTTTTTTTAATGTTTCTGATTTCTGAAGTGCTTGTA

4701 TAGCTTTTATCTGCGGCTTTAAACTGACAGTACCCGACTGTTTATTGGATCTATTGATTTGAAAAGAATTTGTTAGGATAGATCTTAAGCAGTAATCTGT

4801 CAGTGTTTGTATTTGTATTTTCTGCAATTTTACTGTGAAAAAAAATTTGTTTTCAACAATTGGTGTCATTTTCTTGATGTCACTATTTGTTGGAGAGTTA

4901 AATGGTCTCTTCCCTTTGTGTATCTTACCTAGTGTTTACTCCTGGGCACCCTTAATCTTCAGAGGTGCTAAATTGTCTGCCATTACACCAGAAGGATGCC

5001 TCTGATAGGAGGACAACCATGCAAATTGTGAAATAGTCCTGAAGTTCTTGGATTACTTTACACCTCAGTATTGATTTGTCCCAGAATTTTCTGGCCTTTC

5101 ATGGCAATGAAAATTTTAAGAAGAAAGATTTAAAGTATTTAATTTTAAAGAGTGTGTTATAAAATAATGTACTGAATTCTTTATCCCATTTTATCATCC

5201 TTTCAGTTTTTATTAATCTACTGTATCAATAAAATTCTGTAATTTGAATGAGTAAAAAAAAAAAAAAAAAA (5253)
```

Fig 1 (*continued*)

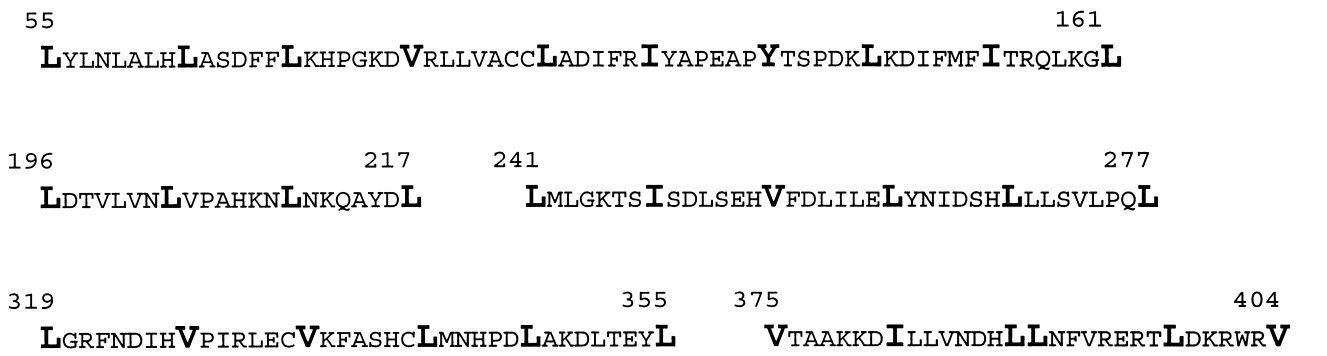


Fig. 2. The N-terminal leucine repeat structure of the AS3 polypeptide. Numbers above the AS3 sequence indicate the positions of the blocks where uninterrupted leucine (or isoleucine, valine) heptades occur.

Sequencing System (Life Technologies, Gaithersburg, MD). The EST DNA sequences were tested for homology to known DNA sequences using the FASTA and BLAST (National Center for Biotechnology Information, Bethesda, MD) programs. Five inserts were found with no match in GenBank [11]. For further analysis, the mRNA with the highest induction in shutoff positive LNCaP-FGC cells (AS3, >5–6-fold of the 5.3 kb mRNA, and >3–4-fold of the 8 kb isoform) was selected.

### 2.5. Isolation of the full length AS3 cDNA sequence

Based on the 262 bp known AS3 tag sequence, nested primer pairs were designed to amplify the full length cDNA sequence from a cDNA library. The cDNA libraries were generated by Human Genome Sciences (Rockville, MD), using polyA$^+$mRNA preparations from androgen-treated or CDHuS-treated proliferation-arrested LNCaP-FGC cells. The Lambda ZAPII (UniZAP) phage was used as a vector carrying *Eco*RI and *Xho*I cloning sites. The PCR reaction was designed to amplify the cloned unknown cDNA segments between the known tag sequence and the flanking vector sequences. Since the orientation of the tag sequence was not known, both ends of the insert were amplified in both directions. The vector primers were commercially available sequencing primers: M13 Reverse and T3 primers at the *Eco*RI site, and M13–20 and T7 primers at the *Xho*I site. For the PCR reaction, the Expand High Fidelity kit was used with 1 μl phage suspension as template (Boehringer-Mannheim). A 40 cycle amplification in a Perkin-Elmer 9600 thermocycler resulted in a 1370 bp fragment at the 5′ end, and a 3250 bp fragment at the 3′ end. The PCR products were purified by using Qiagen columns, and sequenced by automatic sequencing using the primer walking strategy. The sequencing data showed that the open reading frame in the 5′ end fragment did not have an authentic AUG codon. To search for the missing 5′ end of the transcript, we used the Prostate Specific Marathon Ready cDNA preparation from Clontech. Amplifications with the Clontech anchored primer and a set of AS3 specific primers resulted in a 419 bp fragment. The DNA was cloned and sequencing data showed that it carried the N-terminal 118 amino acids of the open reading frame. The nucleotide sequence reported in this paper has been submitted to the GenBank with accession number U95825.

### 2.6. Protein sequence analysis, motif and profile search

Computer translation of the open reading frame was performed using the Translate program of the Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI. β-strand and α-helix structures were calculated by the Chou–Fasman method using PepStructure and PepPlot programs. Motif and profile predictions were calculated using various programs of the Wisconsin Package, or by using remote servers offering sequence analyses of protein functional domains through the Internet. The following remote servers were used: PROWEB (http://www.proweb.org); BLOCKS (http://www.blocks.fhcrc.org); PRODOM (http://www.toulouse.inra.fr/prodom/); PRINTS (http://www.biochem.ucl.ac.uk/cgi-bin/attwood/) and the Protein Kinase Resource (http://www.sdsc.edu/Kinases/).

## 3. Results

### 3.1. Characterization of the AS3 cDNA sequence

Computer analysis of the sequenced 5253 bp AS3 cDNA identified a long open reading frame (Fig. 1). The initiator methionine is at position 66, the stop codon was found at position 4239, and the region codes for a polypeptide of 1391 residues. The initiator is the first AUG codon downstream from the 5′ end of the sequence, and it is in strong Kozak-context [14,15]. The Northern blot size of the transcript is between 5.3

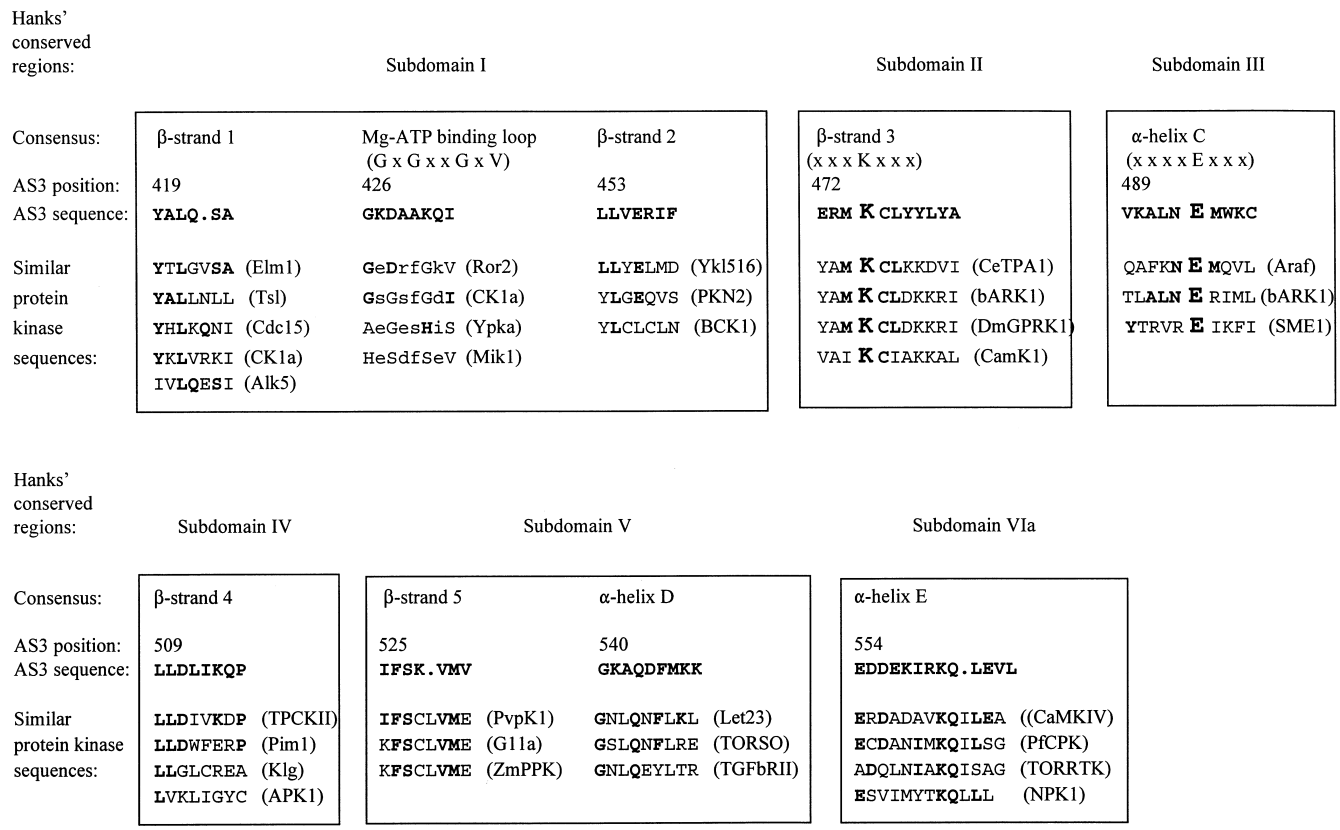


Fig. 3. Sequence comparisons of the putative Mg-nucleotide triphosphate binding subdomains of AS3 with corresponding subdomains of various protein kinases. The boxes represent the Hanks' conserved subdomains, as indicated above each box. The top lines within the boxes show the consensus β-strand, loop, and α-helical secondary structure elements. The numbers in the second lines indicate the positions of the corresponding conformations in the AS3 sequence. The actual AS3 motifs are shown in the third line. Hanks' subdomains from protein kinases of close similarity are represented in the lines below the AS3 sequence. The names of the kinases are indicated in parentheses. Identical residues are highlighted. In the Mg-ATP binding loop, the x and lower case letters indicate nonconserved amino acids.

and 5.5 kb [11] and since the sequence we report here has 5253 nucleotides plus the poly-A tail, the 5′ end of our sequence is at or within a few nucleotides of the 5′ physical end, further suggesting that the initiator is authentic. The 5′ noncoding region is high in GC nucleotides (63.3%), but it has no recognizable secondary structure elements or other sequence features. The 3′ noncoding region has several destabilizing AT-rich elements (underlined in Fig. 1), typical of transcripts claimed to play a role in proliferation regulation [16,17]. The polyadenylation signal of the transcript is 25 bp upstream of the consensus GT-rich cleavage site (highlighted in Fig. 1).

### *3.2. Characterization of the putative AS3 polypeptide sequence*

The expected molecular weight of the putative polypeptide is 186 kDa. The N-terminal 400 amino acid domain is characterized by a unique arrangement of 31 aliphatic residues (21 of them are leucines). Every seventh position (with minor variations) is occupied by a leucine or similar hydrophobic residues and in the five subdomains shown in Fig. 2 the pattern is uninterrupted. The arrangement is typical for coiled-coil structures where one side of the long α-helixes is hydrophobic and usually participates in protein–protein interactions [18,19]. The leucine-zipper motif of DNA binding proteins is a specific subclass of this general pattern and in fact, the subdomain between positions 196 and 217 in the AS3 sequence is a perfect leucine-zipper.

The AS3 polypeptide sequence between positions 400 and 600 has elements of a conserved Mg-ATP binding domain of various nucleotide triphosphate binding proteins including protein kinases. In Fig. 3, the AS3 sequence is shown in the conserved subdomain arrangements established by Hanks [20]. The conserved β-strand, α-helix structures and highly conserved critical residues are also indicated [21], together with the corresponding sequences of various protein kinases. Although the complete AS3 sequence did not appear to be related to any particular protein kinase or ATP binding proteins, partial homology within the

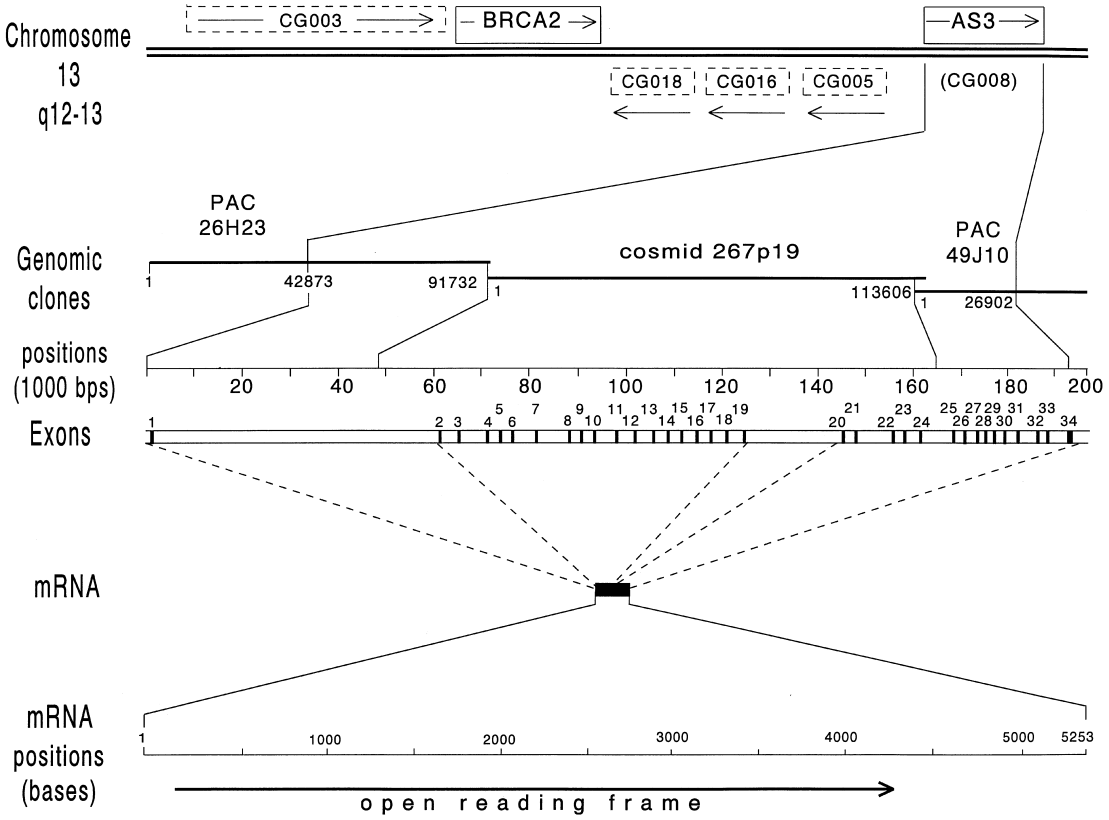


Fig. 4. The genomic, cosmid, and exon maps of AS3 cDNA. The chromosomal panel represents a 1Mb genomic region around BRCA2. Boxes with CG numbers are genomic areas where expression of transcripts were detected. The centromer is at the left. PAC26H23 (Acc. : Z84467) overlaps with cosmid 267p19 (Acc#: Z75889), which, in turn, overlaps with PAC49J10 (Acc#: Z84572). Numbers below the PAC and cosmid lines indicate positions within the genomic clone. The scale above the exon map indicate the genomic distance (×thousand). In the exon panel, black boxes represent the exons, while the numbers above them indicate exon numbers. In the mRNA panel, the numbers indicate nucleotide positions.

subdomains is maintained, and probably indicates that the domain is functional. Sequence comparisons with other domains of protein kinases did not show high levels of conservation.

A putative nuclear localization sequence (NLS) (**KK**FTQVLEDDE**KIRK**) resembling that of the androgen receptor [22] and DNA polymerase-α [23] was localized at position 547. The C-terminal region of the putative AS3 polypeptide contains several sequence elements that show similarities to DNA binding proteins. Motifs and ProfileScan searches in the Wisconsin Package indicated helix-loop-helix and Homeo-box signature sequences in the area, and a remote search on the BLOCKS server also identified DNA binding block elements in the C-terminal sequences.

A serine-rich domain at position 1139, and a proline/glycine-rich domain at the 1284 position were also found. The C-terminal domain (about 200 amino acids) is highly charged and arranged in unique repeats of seven alternating acidic and basic domains.

### 3.3. Characterization of the genomic coding region

A homology search in GenBank identified the AS3 genomic region on chromosome 13q12-q13. The area is represented in cosmid 267p19 and on P1 artificial chromosomes PAC26H23 and PAC49J10. We identified consensus splicing donor and acceptor sites [24]; the entire exon–intron structure of the AS3 gene was resolved comparing the cDNA sequence and the genomic sequence by the BLAST program (Fig. 4). The actual cosmid and cDNA positions are listed in Fig. 5, and the arrangement of exons is depicted in Fig. 4. The area covers nearly 200,000 bp; the average size of the exons is 100–150 bp.

The AS3 genomic area is centromeric to the RB1 locus, and telomeric to BRCA2. The AS3 gene is transcribed in the same direction as BRCA2, and in the coding strand it is downstream from the breast cancer gene [25]. On the opposite strand upstream of AS3, three regions were assigned to cDNAs of unknown functions. An expressed sequence (CG008) has been assigned to this area [25], and represents a portion of

```
            {42873}  1                                46  {42919}
                     CCGGAGAG.... Exon 1 ....ACCCGGAG * gtaggaa...

             (13347)  47                             173  (13475)
....ttttcttgtttcag * GGGTAGAA.... Exon 2 ....GATTAAAG * gtgagta...

             (16397)  174                            377  (16602)
..ttttatttttgtatag * ATGGTTGT.... Exon 3 ....AACTAAAG * gcaagta...

             (22832)  378                            464  (22920)
..tctttttttatttaag * GATATATT.... Exon 4 ....TACTTGAG * gtaagca...

             (23028)  465                            562  (23125)
......ccttatttttag * AACATTGC.... Exon 5 ....GTTATAAA * gtaagtt...

             (23747)  563                            689  (23873)
.....ttttgaattgcag * CAATGGCC.... Exon 6 ....CTCATAAG * gtgagta...

             (32357)  690                            854  (32439)
....tttatgtttttcag * AATTTAAA.... Exon 7 ....TTACCACT * gtaagtc...

             (37809)  855                            911  (37951)
...ctttctcctcaaaag * TTTTTTAA.... Exon 8 ....AATTAAAG * gtaactt...

             (40437)  912                           1027  (40554)
.......ttttattttag * AGCAATGA.... Exon 9 ....TTGGGCAG * gtatatg...

             (43428)  1028                          1122  (43524)
...tttatattttatcag * GTTTAATG.... Exon 10....CTTAACAG * gtactat...

             (48471)  1123                          1268  (48617)
.....tgttatctttcag * AGTATCTT.... Exon 11....ACAAACGA * gtaagta...

             (51727)  1269                          1420  (51880)
....tttttgtttttaag * TGGAGAGT.... Exon 12....GATGATCG * gtaagtt...

             (53049)  1421                          1534  (53164)
...tctgctttttgtag * ACTACTTG.... Exon 13....GCTGTGAA * gtatgtt...

             (58816)  1535                          1616  (58898)
....tttgtgtttttcag * AGCATTGA.... Exon 14....AACCCAAA * gtaagta...

             (61447)  1617                          1665  (61497)
...ttgtgtgatttacag * ACAGATGC.... Exon 15....TATTACAA * gtaagtt...

             (64323)  1666                          1805  (64464)
.......tttattttaag * GAAATTTA.... Exon 16....GTTGTGTG * gtaagga...

             (65916)  1806                          1921  (66033)
...taatctgtattacag * CGTGAAAT.... Exon 17....TCTATCAG * gtatttg...

             (71527)  1922                          2027  (71633)
...ttggtcatattttag * TGCTCTTA.... Exon 18....TGCTTAAG * gtaagta...

             (74539)  2028                          2188  (74700)
...tgattcattttatag * GTACTCTC.... Exon 19....ATCAGATC * gtgagtt...

             (96694)  2189                          2312  (96818)
....ttttttttttaatag * AGCCTTGC.... Exon 20....TATTTGAG * gtaatga...

             (99765)  2313                          2471  (99925)
...tcccctcatttttcag * CCTCTGCA.... Exon 21....ATGATCGG * gtaattt...

            (105674)  2472                          2540  (105744)
...ctcgtttatttttag * CTTCCAGG.... Exon 22....TGGTCAAA * gtgagta...

            (107185)  2541                          2677  (107322)
...ttgtctcttaaatag * ATTCAGGC.... Exon 23....AAAATTAG * gtatgca...

            (110571)  2678                          2801  (110696)
...ctactcatttttcag * TAAACCAG.... Exon 24....CTATCAAC * gtaagga...

              [4319]  2802                          3006  [4524]
....ttgtgtctttacag * GATGAATG.... Exon 25....TGTTAGTG * gtaagca...

              [6829]  3007                          3121  [6945]
.....ttttctttttcag * AAAAATTA.... Exon 26....GTTAAAGA * gtaagac...

              [9074]  3122                          3254  [9208]
....ttttttttttttag * ATGTCTTT.... Exon 27....TGAATGAA * gtatgta...

              [9522]  3255                          3374  [9642]
.....tatactattgcag * AAACTGTA.... Exon 28....CTGACAAG * gtagtta...

             [10614]  3375                          3437  [10679
...ttctcttggttgtag * AATTTCAG.... Exon 29....CTGGAAAA * gtatgtt...

             [11561]  3438                          3583  [11709]
...catttctcatttcag * CCTAAAAC.... Exon 30....AAGGGGAG * gtaagtg...

             [15476]  3584                          3689  [15583]
...tgtctgtattaaaag * GCTTGATA.... Exon 31....TTGTAAGG * gtgagat...

             [21107]  3690                          4129      [21548]
...ttttttttccccctag * TCTGAATT.... Exon 32....CAGCAGAG * gtaagca...

             [21640]  4130                          4354  [21866]
...tcttccccaaagcag * AGCAGAAT.... Exon 33....TACACTAG * gtaagat...

             [26002]  4355                          5253  [26902]
.....ctttccttttaag * GTACGGCG.... Exon 34....GAATGAGT * (poly-A)
```

Fig. 5. Genomic and cDNA positions of exons in the AS3 transcript. Asterisks represent the exon–intron boundaries. The area between asterisks represents the exons. Exon sequences are in upper case, the numbers represent cDNA positions. Lower case letters are intron sequences. Numbers of the first exon indicate positions in PAC26H23. Numbers in parenthesis refer to positions on cosmid 267p19, while numbers in brackets refer to PAC49J10 positions.

the AS3 transcript. The N terminal 354 amino acids of the open reading frame are missing in the CG008 sequence in GenBank. The CG008 open reading frame terminates at amino acid 738 of the AS3 sequence. The sequencing data reported herein and the published genomic sequence are identical, confirming the correct sequence of AS3. The extra C at nucleotide position 1109 in the CG008 sequence suggests a possible sequencing error that results in a frame shift and a stop codon at position 1152 of the CG008 sequence.

## 4. Discussion

The mechanisms whereby androgens affect the proliferation of their target cells, and specifically those in the human prostate, are incompletely understood [2–4]. We have proposed that androgens regulate prostate cell proliferation through a two-step mechanism [3]. In step 1, androgens would neutralize a putative serum-borne proliferation inhibitor. In step 2, androgen-induced gene products would directly arrest prostate cell proliferation (proliferative shutoff). The androgen-induced proliferative shutoff appears to be mediated by the androgen receptor [9]. Antiandrogen inhibition of the shutoff phenomenon is also consistent with androgen receptor involvement [26]. Here, we report the complete cDNA sequence, the genomic localization, the exon–intron map, and the putative protein structure of one of the candidate genes, AS3.

To identify cDNAs coding for inhibitors that are induced by androgens, a subtractive strategy was used, whereby the proliferation of LNCaP-FGC cells was arrested using two different treatments, namely, CD serum and high androgen concentrations. This selective approach takes advantage of the fact that the cells were equally arrested at the G1 stage of the cell cycle by different mechanisms [6]. Regulatory mRNAs are likely to be expressed at low copy numbers; we adopted a protocol where repeated PCR cycles can selectively amplify these sequences. The final subtracted pool, therefore, was expected to represent high ranking regulatory elements in the androgen-induced proliferative shutoff (step 2).

Expression analysis of the AS3 transcript demonstrated proliferation arrest-specific induction patterns, starting early (4–6 h) after androgen exposure [11]. It peaked at 18–20 h, about 4 h before the commitment for proliferative shutoff was detected, suggesting that this gene is a candidate for a shutoff mediator. Furthermore, expression of the AS3 transcript positively correlated with proliferation arrest; this gene was expressed only in shutoff-positive cell lines and variants [11]. LNCaP-FGC cells proliferated maximally in CDHuS supplemented with 30 pM R1881; under these conditions AS3 was not expressed. To the contrary,

AS3 was strongly induced at 0.3–30 nM R1881 and the cells were inhibited from proliferating [11]. Additional arguments favoring the notion that the AS3 gene codes for an inhibitor of the proliferation of prostate cells, are the increase of AS3 mRNA levels while proliferation in the rat prostate was arrested by prolonged androgen administration [11], and a comparable effect was recorded on MCF7-AR1 cells [11].

The AS3 coding region was localized at the BRCA2-Rb1 suppressor area on chromosome 13, immediately downstream of the breast cancer susceptibility gene BRCA2. Epidemiological studies have suggested a link between breast and prostate cancers [27,28] implying shared genetic suppressor elements in both. Studies of breast cancer families with high loss of heterogenicity (LOH) in the BRCA2 area showed that high prostate cancer incidence occurred in 4 out of 5 families investigated by Gudmundson et al. [29]. In the majority of the male relatives with prostate cancer in these families (86%), allelic losses in the BRCA2 area were also detected [29], and some of them hit the region immediately downstream of the BRCA2 gene [29–31].

Putative suppressors in the immediate vicinity of BRCA2 are not limited to sex hormone-related cancers. Recent studies on chronic lymphoid leukemia detected a 1Mb allelic loss in this region, with no mutations in the BRCA2 gene, pointing to a cryptic suppressor next to this gene [32,33]. We show here that the coding sequence of AS3 lies within this area.

Computer analysis of the AS3 open reading frame identified putative functional domains typically found in regulatory proteins. The coiled-coil domain mediates protein–protein interactions, such as those involved in dimerization or oligomerization. Leucine-zippers are found in many DNA binding proteins. Homology and profile search also identified various other DNA binding motifs in the AS3 sequence, strongly suggesting that the protein may act as a DNA binding factor.

A BLASTP search performed on the GenBank database resulted in a single high score similarity with the bimD gene of the eukaryotic *Aspergillus nidulans*, where 50% of the amino acid sequence was functionally similar in portions of the coiled–coil domain and the putative DNA binding domain at the C-terminus. The bimD protein has a basic leucine-zipper and a C-terminal charged (acidic) domain, similar to AS3, and appears to function as a DNA binding protein [34]. Both the AS3 and the bimD proteins also have nuclear localization consensus sequences. Significantly, overexpression of bimD in *A. nidulans* results in cell cycle arrest in G1/S phase [34]. The peak expression of AS3 also coincides with the arrest of the cell cycle in G1 phase in the androgen-induced proliferative shutoff [11].

A unique feature of the AS3 polypeptide is the presence of a relatively well conserved Mg-nucleotide triphosphate binding pocket, as well as elements of the catalytic domain of protein kinases. Functional analysis of this domain is in progress using GST-fusion constructs. The fusion construct can form a complex and phosphorylate two substrate proteins in LNCaP-FGC extracts (data not shown). It is not yet clear whether this effect is catalyzed directly by the AS3 fusion protein, or the AS3 fusion protein is involved indirectly, through binding a protein kinase and its substrates (docking function).

In summary, this work represents an analysis of AS3, an androgen-induced, proliferative shutoff candidate gene. AS3 has been mapped to a genomic region with frequent allelic losses in various cancers.

## Acknowledgements

This work was supported in part by PHS NIH grants CA-55574, CA-13410, AG-13807 ES 08314 and NSF DCB-9603779. JJ was a recipient of a postdoctoral fellowship from Fundacion Ramon Areces (Madrid, Spain). The able technical help of Marcelle Desronvil, Anita Oles, Mariana Fernandez and Maria Luizzi is acknowledged. We are grateful to Dr. Marion Gordon, Nancy Prechtl and Cheryl Michaelson for helpful suggestions in the preparation of the manuscript. We also gratefully acknowledge Dr. Wang and Dr. Brown for providing their detailed unpublished protocol. The assistance of The Center for Reproductive Research at Tufts University (P30 HD 28897) is acknowledged.

## References

[1] J.T. Isaacs, Antagonistic effect of androgen on prostatic cell death, Prostate 5 (1985) 545–557.

[2] N. Bruchovsky, B. Lesser, E. Van Doorn, S. Craven, Hormonal effects on cell proliferation in rat prostate, Vit. & Horm. 33 (1975) 61–102.

[3] C. Sonnenschein, N. Olea, M.E. Pasanen, A.M. Soto, Negative controls of cell proliferation: human prostate cancer cells and androgens, Cancer Res. 49 (1989) 3474–3481.

[4] S.W. Hayward, M.A Rosen, G.R. Cunha, Stromal–epithelial interactions in the normal and neoplastic prostate, Br. J. Urol. Suppl. 2 (1997) 18–26.

[5] J.S. Horoszewicz, S.S. Leong, E. Kawinski, J.P. Karr, H. Rosenthal, T.M. Chu, E.A. Mirand, G.P. Murphy, LNCaP model of human prostatic carcinoma, Cancer Res. 43 (1983) 1809–1818.

[6] A.M. Soto, T.M. Lin, K. Sakabe, N. Olea, D.A. Damassa, C. Sonnenschein, Variants of the human prostate LNCaP cell line as a tool to study discrete components of the androgen-mediated proliferative response, Oncology Res. 7 (1995) 545–558.

[7] I.Y. Kim, J.H. Kim, D.J. Zelner, H.J. Ahn, J.A. Sensibar, C. Lee, Transforming growth factor-β1 is a mediator of androgen-regulated growth arrest in an androgen-responsive

prostatic cancer cell line LNCaP, Endocrinology 137 (1996) 991–999.
[8] I.Y. Kim, H.J. Ahn, D.J. Zelner, J.W. Shaw, J.S. Sensibar, J.H. Kim, M. Kato, C. Lee, Genetic change in transforming growth factor-β receptor type I correlates with insensitivity to transforming growth factor-β1 in human prostate cancer cells, Proc. Am Assoc. Cancer Res. 37 (1996) 124.
[9] J. Szelei, J. Jimenez, A.M. Soto, M.F. Luizzi, C. Sonnenschein, Androgen-induced inhibition of proliferation in human breast cancer MCF7 cells transfected with androgen receptor, Endocrinology 138 (1997) 1406–1412.
[10] Z. Wang, D.D. Brown, A gene expression screen, Proc. Natl. Acad. Sci. USA 88 (1991) 11505–11509.
[11] P. Geck, J. Szelei, T- M Lin, C. Sonnenschein, A.M. Soto, Expression of novel genes linked to the androgen induced proliferative shutoff in prostate cancer cells, J. Steroid Biochem. Mol. Biol. 63 (1997) 211–218.
[12] A.M. Soto, C. Sonnenschein, The role of estrogens on the proliferation of human breast tumor cells (MCF-7), J. Steroid Biochem. 23 (1991) 87–94.
[13] P. Chomczinsky, N. Sacchi, Single step method of RNA isolation by acid guanidinium–phenol–chloroform extraction, Anal. Biochem. 162 (1987) 156–159.
[14] M. Kozak, Structural features in eukaryotic mRNAs that modulate the initiation of translation, J. Biol. Chem. 266 (1991) 19867–19870.
[15] M. Kozak, An analysis of vertebrate mRNA sequences: intimations of translational control, J. Cell Biol. 115 (1991) 887–903.
[16] G. Shaw, R. Kamen, A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation, Cell 46 (1986) 659–667.
[17] C.-Y. Chen, A.-B. Shyu, AU-rich elements: characterization and importance in mRNA degradation, Trends Biochem. Sci. 20 (1995) 465–470.
[18] A. Lupas, Coiled-coils: new structures and new functions, Trends Biochem. Sci. 22 (1996) 375–382.
[19] A.J. Beavil, R.L. Edmeades, H.J. Gould, B.J. Sutton, α-helical coiled–coil stalks in the low affinity receptor for IgE (FceRII/CD23) and related C-type lectins, Proc. Natl. Acad. Sci. USA 89 (1992) 753–757.
[20] S.K. Hanks, A.M. Quinn, Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members, Meth. Enzymol. 200 (1991) 38–62.
[21] S.S. Taylor, D.R. Knighton, J. Zheng, L.F. Ten Eyck, J.M. Sowadsky, Structural framework for the protein kinase family, Annu. Rev. Cell Biol. 8 (1992) 429–462.
[22] Z.-X. Zhou, M. Sar, J.A. Simental, M.V. Lane, E.M. Wilson, A ligand dependent bipartite nuclear targeting signal in the human androgen receptor, J. Biol. Chem. 269 (1994) 13115–13123.
[23] D. Bouvier, G. Baldacci, The N-terminus of fission yeast DNA polymerase α contains a basic pentapeptide that acts in vivo as a nuclear localization signal, Mol. Biol. Cell 6 (1995) 1697–1705.
[24] M.B. Shapiro, P. Senapathy, RNA splice junctions in different classes of eukaryotes: sequence statistics and functional implications in gene expression, Nucl. Acids Res. 15 (1987) 7155–7174.
[25] F.J. Couch, J.M. Rommens, S.L. Neuhausen, C. Belanger, M. Dumont, K. Abel, R. Bell, S. Berry, Generation of an integrated transcription map of the BRCA2 region on chromosome 13q12-q13, Genomics 36 (1997) 86–99.
[26] N. Olea, K. Sakabe, A.M. Soto, C. Sonnenschein, The proliferative effect of 'anti-androgens' on the androgen-sensitive human prostate tumor cell line LNCaP, Endocrinology 126 (1990) 1457–1463.
[27] E.U. Thiessen, Concerning a familial association between breast cancer and both prostatic and uterine malignancies, Cancer 34 (1974) 1102–1107.
[28] H. Tulinius, V. Egilsson, G. Olafsdottir, H. Sigvaldason, Risk of prostate ovarian and endometrial cancer among relatives of women with breast cancer, Br. Med. J. 305 (1992) 855–857.
[29] J. Gudmundsson, G. Johannesdottir, J.T. Bergthorsson, A. Arason, S. Ingvarsson, V. Egilsson, R.B. Barkardottir, Different tumor types from BRCA2 carriers show wild type chromosome deletions on 13q12-q13, Cancer Res. 55 (1995) 4830–4832.
[30] J. Van den Berg, O. Johansson, S. Hakansson, H. Olsson, A. Borg, Allelic loss at chromosome 13q12-q13 is associated with poor prognosis in familial and sporadic breast cancer, Br. J. Cancer 74 (1996) 1615–1619.
[31] A.-M. Cleton-Jansen, N. Collins, S.R. Lakhani, J. Weissenbach, J.P. Devilee, C.J. Cornelisse, M.R. Stratton, Loss of heterozygosity in sporadic breast tumors at the BRCA2 locus on chromosome 13q12-q13, Br. J. Cancer 72 (1995) 1241–1244.
[32] J.A. Garcia-Marco, C. Caldas, C.M. Price, L.M. Wiedemann, A. Ashworth, D. Catovsky, Frequent somatic deletion of the 13q12.3 locus encompassing BRCA2 in chronic lymphocytic leukemia, Blood 88 (1996) 1568–1575.
[33] C. Caldas, A. McGregor, L. Wiedemann, D. Catovsky, A. Ashworth, J. Garcia-Marco, Evidence for the existence of a new tumor suppressor gene near the BRCA2 locus, Proc. Am. Assoc. Cancer Res. 38 (1997) 191.
[34] S.H. Denison, E. Kafer, G.S. May, Mutation in the bimD gene of Aspergillus nidulans confers a conditional mitotic block and sensitivity to DNA damaging agents, Genetics 134 (1992) 1085–1096.