Identification of a novel transcription factor, GAGATA binding protein, involved in androgen-mediated expression of prostate-specific antigen*

Chihuei Wang^‡, Fan Yeung¥, Po-Chun Liu^, Ricardo M. Attar†, Jieping Geng†, Leland W. K. ChungΦ, Marco Gottardis† and Chinghai Kao§‡

^ Orthopedic Research Center, and Department of Biology, Kaohsiung Medical University, Kaohsiung, 807, Taiwan.

†Department of Oncology, Bristol-Myers Squibb, Lawrenceville, NJ 08543

¥Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA 22908

§Department of Urology, University of Indiana, Indianapolis, IN 46202

ΦDepartment of Urology and Winship Cancer Institute, Emory University, Atlanta, Georgia, 30322.

Running Title: A novel cis-element, GAGATA.

Key Words: prostate-specific antigen, promoter, androgen receptor, prostate cancer, androgen-response elements, transcriptional factor.

* Supported by National Institutes of Health grant CA74042, Ed Thompson Scholar fund and NSC90-2320-B-037-054.

‡To whom correspondence should be addressed: Department of Urology, Indiana Cancer Pavilion, 535 N. Barnhill Drive, Suite 420, Indianapolis, IN 46202-5289. Tel: 317-278-3431; FAX: 317-278-3433; E-mail: chkao@iupui.edu. Orthopedic Research Center, Kaohsiung Medical University, Kaohsiung, 807, Taiwan. Tel: 886-7-3121101 ext 2553; FAX. 886-7-3219452; E-mail: chwang@kmu.edu.tw
Summary

Prostate-specific antigen (PSA) is the most valuable marker for the evaluation of prostate cancer progression. The expression of PSA is controlled by androgen receptor (AR) through its binding to androgen response elements (AREs). Several AREs have been identified within the 5.8 kb PSA promoter. The main activity of this 5.8 kb PSA promoter resides in a 455-bp enhancer core region located about 4 kb upstream of the TATA box. Our study suggests that in addition to the four androgen response elements (AREs) identified in the PSA enhancer core, another regulatory element (GAGATA), which is located at the region designated PSA3.1, also contributes to transcriptional regulation by androgens. Furthermore, electrophoretic mobility shift assay (EMSA) revealed that a putative transcriptional factor bound the GAGATA sequence in the PSA-producing prostate cancer cell. Further studies demonstrated that GAGATA factor preferentially bound the (G/C)(A/C/T)GATA sequence. The replacement of ATA with GGG in the GAGATA sequence completely eliminated the androgen-mediated transcriptional activity of the enhancer core. By using DNA-coupled magnetic beads and the Southwestern method, a 56-60 kDa protein was identified as the putative GAGATA binding factor. EMSA and Western blotting assay suggested that AR is not involved in androgen-mediated activation through PSA3.1. Therefore, we propose that binding of the GAGATA binding factor and AR to GAGATA and AREs of the PSA enhancer core are respectively required for the maximum transcriptional response to androgens.
Introduction

Prostate-specific antigen (PSA) is a member of the kallikrein-like serine protease family (1). PSA expression appears exclusively in prostate tissue and the increasing levels of this glycoprotein in serum are highly correlated with the progression of prostate cancer (2). PSA is routinely used for diagnosis, prognosis and monitoring of treatment responses and progression in patients with prostate cancer (3, 4). The production of PSA is tightly regulated by androgens. After androgen ablation therapy, prostate cancer patients may exhibit normal PSA levels (5, 6). However, once the tumor progresses and no longer responds to endocrine therapy, the expression of PSA rebounds and becomes androgen-independent (7, 8). The elucidation of the factors involved in the regulation of PSA expression could provide further understanding of the transition from the androgen-dependent to the androgen-independent phenotype in prostate cancer.

The PSA promoter and enhancer region have been intensively investigated. The DNA sequences responsible for the regulation of PSA expression have been mapped up to 6 kb from the start site of transcription (9, 10). There are several putative transcription factor binding sites in this region, such as several androgen response elements (AREs) in the region of -170 to –156 and –4148 to –4234, and an androgen response region (ARR) in the region of -395 to –376 (11, 12). Comparison of transgenic mouse lines carrying LacZ reporter genes, driven by either the 632-bp proximal PSA promoter or by the 6-kb PSA promoter, reveals that the latter contains most of the elements required for androgen regulation and prostate specificity of the PSA gene in vivo (9). Moreover, by comparing chromatin structures from the LNCaP prostate cancer cell line grown in the presence and absence of the synthetic androgen R1881, three R1881-induced Dnase1 hypersensitive
sites (DHS) have been identified in the region between 4.8 and 3.8 kb upstream of the PSA promoter (9). A 492 bp fragment (-4366 to -3874) containing a functional, high-affinity ARE has been identified as the enhancer core responsible for the expression of PSA (10, 13). In addition, three weak-affinity AREs and several GATA consensus sequences within the enhancer core are also essential for the function of androgen-mediated PSA expression (13, 14).

In the present study, we have identified a novel cis-element, (G/C)(A/C/T)GATA, which also contributes to androgen-regulated PSA expression. We have demonstrated that this sequence is a binding site for a high affinity and specific 56-60 kDa factor with a major role in the androgen-mediated function of the PSA promoter.
Materials and Methods

Plasmid Constructs. pGL3-PSAc/TATA reporter plasmid containing 440 bp PSA enhancer core and minimal TATA promoter was constructed as follows: the 440 bp PSA core was amplified by PCR using sequence-specific primers at both ends, followed by cloning into PCR vector (Invitrogen, Carlsbad, CA). A PSA core was then subcloned into a modified pGL3-basic that contained a simple TATA box.

The pSEAP-PSA reporter construct containing 540 bp PSA promoter and 1500 bp enhancer region was constructed as follows: commercial human placental genomic DNA was used to generate the fragment containing the Bgl II site (+1) and the Hind III site (-539) of the human prostate specific antigen promoter by PCR (10). This fragment was subcloned into the pSEAP2/basic (Clontech, Palo Alto, CA) previously digested with Bgl II and Hind III to generate the pSEAP2/PSA540 construct. Subsequently a fragment bearing the fragment of human PSA upstream sequence between positions –5322 and –3873 was amplified by PCR from human placental genomic DNA. Xho I and Bgl II cloning sites were introduced and the resulting fragment was subcloned into pSEAP2/PSA540 digested with Xho I and Bgl II respectively, to generate the pSEAP2/PSA540/Enhancer construct (Fig. 6A). The mutation plasmid was generated with a QuickChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) by the replacement of GAGATA into GAGGGG of PSA3.1 in pSEAP/PSA540/Enhancer plasmid.

Flag-tagged AR expression vector was constructed as following. A primer encoding the FLAG peptide fused to the sequences encoding the amino-terminal 5 amino
acids of AR (ACC ACC ATG GAC GAC GAC GAC AAG GCC GAA GTG CAG TTA GGG C) was used in combination with an AR C-terminal primer (TCA CTG GGT GTG GAA ATA GAT GGG) to obtain a full length of AR cDNA with 8 amino acid FLAG tag. The PCR product of FLAG-tagged full length AR (f-AR) cDNA was cloned into pGEM-easyT (Promega, Madison, WI) to generate pGEM-easyT-f-AR. Then f-AR was excised from pGEM-easyT-fAR by Not I digestion, blunt-ended by T4 DNA polymerase (Invitrogen), and placed into the Stu I site of pIRES-EGFP to generate pIRES-EGFP-f-AR.

**Linker Scanning Mutagenesis.** By linker scanning analysis of the PSA core, 25 pGL3-PSAc/TATA derived mutations were generated. These plasmids carry a 17 bp replacement of the GAL4-binding site (CGG AGT ACT GTC CTC CG) within the 440 bp region. The forward primers and reverse primers comprised half of the GAL4-binding site, GTCCTCCG and AGTACTCCG at the 5'-end, followed by 20 bp of PSA core fragment corresponding to the sequence flanking the mutagenized bases. PCR reactions were carried out on a template of 100 pg pCL3-PSAc/TATA, by using Platinum Pfx high fidelity enzyme (Life Technologies, Inc., Rockville, MD). The PCR products were purified from 0.8% agarose gel and ligated at room temperature with a rapid ligation kit (Roche, Indianapolis, IN). Clones were screened by PCR with a GAL4-binding site and Glprimer2 from the pGL3-basic vector (Promega, Madison, WI).

**Cell transfection.** For the pGL3-PSAc/TATA reporter system, LNCaP cells were plated at 3.3x10^5 cells /well in 6-well plates for 2 days before transfection. Plasmid DNAs were
introduced into cells by complexing with DOTAP (Roche). Two and a half micrograms of the tested construct plasmids and 0.5 µg of the internal control CMV/β-Gal plasmids were mixed with 27 µl of 20 mM HEPES (pH 7.4), and then added into a tube containing 8 µl of DOTAP in 16 µl of HEPES with gentle mixing. DNA-lipid complexes were allowed to form for 15 min at room temperature prior to their addition to each well containing 1 ml of serum-free and phenol red-free RPMI 1640 medium. DNA-lipid-containing medium was replaced with fresh serum-free and phenol red-free RPMI 1640 medium. Cells were collected for luciferase assay after 36-48 hours additional incubation.

For the pSEAP2/PSA540/Enhancer reporter system, 1 x 10^7 LNCaP cells (ATCC, Rockville, MD) were collected in media containing 10% charcoal stripped FBS. The cell suspension was placed into a Gene Pulser Cuvetts (Bio-Rad, Hercules, CA) with 8 µg of the reporter construct and 1 µg of pGL3-control (Promega) as internal control, and electroporated using a Bio-Rad Gene Pulser at 210 volts and 960 µFaraday. Following the transfections, the cells were washed and incubated with media containing charcoal stripped fetal bovine serum in the absence (blank) or presence (control) of 1 nM dihydrotestosterone (DHT; Sigma Chemical, St. Louis, MO). Duplicates were used for each sample. After 48 hours, 25 µl of the supernatant was collected for SEAP activity and the remaining cells were used for luciferase assay.

To establish f-AR overexpressing cell lines, LNCaP cells were transfected with 3 µg of pIRES2-EGFP-f-AR by electroporation as described above. Following the transfection, the cells were washed and incubated with media containing 10% FBS in the
presence of 450 μg geneticin/ml. The stable transfectant, LNCaP-f-AR, was obtained through a 4-round selection.

**Luciferase Assay.** Cells were washed with PBS, lysed in lysis buffer (Promega) and collected. Cell debris was removed by microcentrifugation for 3 minutes. Twenty microliters of the supernatant were mixed with 100 μl of luciferase substrate (Promega) and measured by a luminometer (Monolight 2010, Analytical Luminescence Laboratory, San Diego, CA). Each luciferase reading was normalized by the OD405 value derived from the β-Gal assay.

**Secreted Alkaline Phosphatase (SEAP) Assay.** Twenty-five microliters of supernatant were incubated at 65 °C for 30 minutes and cooled down at 4 °C for 5 minutes. This supernatant was mixed with an equal volume of reaction buffer and chemiluminescent substrate (Tropix, Bedford, MA) for 20 minutes and measured by a TopCount (Packard Instrument Company, Inc., Torrance, CA) in the chemiluminescent mode. Each chemiluminescent reading was normalized by luciferase activity derived from the pGL3-control.

**Cell Culture and Preparation of Nuclear Extracts.** LNCaP cells were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum (Life Technologies, Inc.). Nuclear extracts were prepared by high salt extraction according to standard procedures (15). Briefly, cells grown at approximately 90% confluence were collected, washed with phosphate-buffered saline solution, and homogenized in a hypotonic buffer (10 mM...
HEPES, pH 7.9, 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM DTT and protease inhibitor cocktail (Roche)). After centrifugation at 2000x g, the nuclear pellets were incubated for 30 minutes on ice with continuous stirring in a high salt buffer (20 mM HEPES, pH 7.9, 25% Glycerol, 1.5 mM MgCl₂, 0.6 M KCl, 0.2 mM EDTA, 0.5 mM DTT and protease inhibitor cocktail). The extracts were centrifuged, and the supernatants were dialyzed against dialysis buffer (20 mM HEPES, pH 7.9, 20% Glycerol, 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT and protease inhibitor cocktail). The protein concentration of nuclear extracts was quantified by Brandford solution (Bio-Rad).

**Electrophoretic mobility shift assay (EMSA).** EMSA was performed as described previously (16). Briefly, 8 µg of nuclear extracts were incubated for 20 min. at room temperature with 0.014 pmole ³²P end-labeled oligonucleotide-probes with or without unlabeled DNA competitors and 2 µg of poly (dI/dC) in binding buffer (20 mM HEPES, pH 7.9, 1 mM EDTA, 100 mM KCl, 1 mM dithiothreitol and 10% glycerol). Another EMSA was carried out using the same binding conditions as the ones described above, except that 10 ng of purified AR DNA binding domain-ligand binding domain (DBD-LBD)¹ instead of nuclear extracts, and 500 ng of BSA were included in the reactions. The samples were electrophoresed on a pre-run 6 % native acrylamide-bisacrylamide (28:2) gel in 1 X Tris-borate-EDTA at 4° C. Gel was dried and autoradiographed. The oligonucleotides used are described in figure 2.

**Preparation of DNA-coupled magnetic beads.** Biotin-5’-GAT CAT CCT TGC AAG ATG ATA TCT CTC T-3’ was used as the oligonucleotide. The double-stranded
oligonucleotide was created by annealing this oligomer with its antisense oligomer. Twenty microliters of the double-stranded oligomer (10 µM) were incubated with 0.6 mg of Strepavidin MagneSphere Paramagnetic particles (Promega) in 10 mM Tris-HCl buffer (pH 7.9) containing 1M NaCl for 20 minutes to prepare DNA-coupled magnetic beads. Then, DNA-coupled magnetic beads were washed by 500 µl of the same buffer three times. Coupling efficiency was evaluated by comparing the DNA content before and after incubation by using 1.5% agarose gel and EtBr staining.

DNA binding assay. Before use, DNA-coupled magnetic beads were washed three times with 500 µl of buffer (20 mM Hepes, pH7.9, 1 mM EDTA, 100 mM KCl, 1 mM dithiothreitol and 10% glycerol). After collecting the beads with a MagneSphere Magnetic Separation Stand (Promega), supernatant was carefully removed. Binding reactions were carried out with 3 mg of LNCaP nuclear extract, and incubated at room temperature for 20 minutes. Then beads were collected with a magnet stand, unbound protein in supernatant was removed and subjected to further analysis, and beads were washed three times in 500 µl of buffer (20 mM Hepes, pH7.9, 1 mM EDTA, 100 mM KCl, 1 mM dithiothreitol, 10% glycerol, and protease inhibitor cocktail). Afterward, beads were collected with a magnet stand, resuspended in 50 mM Tris-Cl (pH 6.8), 100 mM dithiothreitol, 2% SDS, and 10% glycerol to extract the proteins that bound to DNA. These bound proteins were subjected to SDS-polyacrylamide gel electrophoresis (PAGE), followed by Southwestern blotting assay. 

SDS-PAGE analysis and Western blotting. Nuclear extracts (30 µg), and unbound supernatants were heated at 100 °C for 3 min in 60 mM Tris-Cl (pH 6.8), 1% 2-mercaptoethanol, 2% SDS, and 10% glycerol, and then were separated on 10% SDS-PAGE, followed by Western blotting. In the Western blotting assay, the proteins on the SDS-PAGE were transferred onto nitrocellulose membrane (Millipore). Then the membrane was treated by ANTI-FLAG M2 monoclonal antibody (mAb) (Sigma), followed by horseradish peroxidase-conjugated secondary antibody (PIERCE). The
proteins that interacted with ANTI-FLAG mAb were detected by using SuperSignal West Pico Chemiluminescent Substrate (PIERCE), and exposed to X-ray film.

**Southwestern analysis.** Nuclear extracts (30 µg), the unbound proteins (30 µg) in supernatant, and the bound proteins extracted from DNA-coupling magnetic beads were heated 90 °C for 3 min in 50 mM Tris-Cl (pH 6.8), 100 mM dithiothreitol, 2% SDS, and 10% glycerol, resolved on 8% SDS-PAGE and electrophoretically transferred to nitrocellulose membranes. The membranes were incubated with phosphate-buffered saline for 30 min, blocked 4 h at room temperature with 20 mM HEPES (pH 7.9), 1 mM EDTA, 100 mM KCl, 1 mM dithiothreitol, 10% glycerol and 3% nonfat milk, and then incubated overnight at room temperature in 2 ml of 20 mM HEPES (pH 7.9), 1 mM EDTA, 100 mM KCl, 1 mM dithiothreitol, 10% glycerol plus 10^7 cpm of radiolabeled PSA3.1 or PSA3.1 mutant probes. Afterward, membranes were subjected to three washes of 15 min each at room temperature in a buffer of 20 mM HEPES (pH 7.9), 1 mM EDTA, 100 mM KCl, 1 mM dithiothreitol, 10% glycerol and 0.025% Nonidet P-40, prior to drying and autoradiography.
Results

Identification of a cluster of cis-elements involved in the regulation and expression of PSA gene. As previously reported by Huang et al. (13), the enhancer core of the PSA promoter contains four AR binding sites responsible for androgen-mediated transcriptional activity. In order to identify new cis-elements other than the AREs that are involved in the regulation of PSA expression, linker scanning mutagenesis was performed using 25 sets of oligonucleotides in the enhancer core region (Figure 1A). The resulting mutation enhancers were tested for activity in the transactivation assays in the absence or presence of DHT. All the mutations did not affect the basal transcriptional activity (Fig. 1B). Several regions were identified as essential components for the androgen-mediated transcription activity of the PSA enhancer core (Fig. 1C). L6, LO and R5 were previously reported as weak AR binding sites, L4 as a strong AR binding site (13), and L1 and L7 as GATA transcription factor binding sites (14). In addition, the region that extends from R1 to R3 was also identified as the essential region for androgen-mediated PSA expression (Figure 1B).

We next examined if the region extending from R1 to R3, (referred to as PSA3 in this study) could interact with AR. PSA3 was used as a probe in EMSA with E. coli purified AR DBD-LBD (Figure 2A). We then compared this site to the C3 (I) fragment, a strong ARE (17, 18), the L4 fragment (referred to as PSA1 ARE in this study) and L1-LO (referred to as PSA2 ARE in this study). As Figure 2B shows, only C3 (I) ARE and PSA1 ARE interacted with AR DBD-LBD. To further elucidate the binding affinity of proteins to these sites, we used LNCaP nuclear extracts to perform another set of EMSA in the
absence or presence of competitor oligonucleotides. The results revealed the same binding pattern for C3 (I) ARE, PSA1 and PSA2 (Figure 2C). In addition, PSA1 and PSA2 AREs can be competed by 200-fold of C3 (I) ARE and by their own sequences (Figure 2C). However, PSA3 not only demonstrated a different binding pattern when compared to the other three oligonucleotides, it could only be competed by its own sequence (Figure 2C). These results confirmed that both PSA1 and PSA2 could interact with AR, and PSA1 ARE had a stronger binding affinity for AR than PSA2. In addition, these results suggest that a different factor could interact with PSA3.

The first 24 nucleotides of PSA3 are essential for protein-DNA interaction. To locate the interaction motif of PSA3 with an unknown binding factor(s), three oligonucleotides, PSA3.1, PSA3.2 and PSA3.3, were synthesized. PSA3.1 represented the first 24 nucleotides, PSA3.2 represented the last 24 nucleotides, and PSA3.3 represented the 13th to the 36th nucleotide (data not shown). Only PSA3.1 exhibited specific interaction in EMSAs using LNCaP nuclear extracts (data not shown). This result suggested that the interaction motif of PSA3 was located in the PSA3.1 oligonucleotide. The PSA3.1 fragment contained two putative transcription factor-binding sites. The first binding sequence, GAGAGA, a potential binding site of the GAGA factor, has been reported to be a chromatin-remodeling factor isolated from Drosophila (19). The second binding sequence, AGATAT, is a putative binding site of the GATA transcriptional factor family (20, 21, 22).
The identification of the GAGATA binding site. To exactly locate the binding site of this unknown factor, we sequentially replaced the first 12 nucleotides of PSA3.1 with mutated sequences. Every 3 nucleotides were used as a unit to generate four PSA3.1 mutants, designated as follows: m1 (AGA to CCC), m2 (GAG to TTT), m3 (ATA to GGG) and m4 (TCA to AAA) (Figure 3A). The interaction of WT P32-labelled PSA3.1 with unknown factors cannot be abolished by mutating GAG to TTT (m2) and ATA to GGG (m3) in the EMSA (Fig. 3B). The conclusion is that GAGATA is the essential motif for this unknown factor to bind. We excluded the possibility that the PSA3.1 binding factor is a GAGA binding protein, since the m1 mutant (AGA to CCC) with mutation at the center binding sequence of GAGA binding protein, was able to compete with wt PSA3.1 oligonucleotide in the EMSA (Figure 3B).

GAGATA is the novel cis-element responsible for binding the transcription factors. Comparing the PSA3.1 sequence with the consensus sequence (T/A)GATA(A/G) of GATA transcription factors (20, 23, 24) revealed that the PSA3.1 sequence lacks the A or G in the base position immediately 3' to the GATA core. PSA3.1 contains a T at this position. Moreover, the first G nucleotide of the GAGATA sequence in PSA3.1 is not included in the consensus sequence (T/A)GATA(A/G) of GATA transcription factors. Thus, we examined if the first G nucleotide of the GAGATA sequence was required for the binding of GAGATA binding factor. The result showed that replacing the first nucleotide of the GAGATA sequence with A or T decreased the binding capacity of GAGATA binding factor to about 2% and 20%, respectively (Fig. 4A). Substituting the first nucleotide with C had a relatively low effect, losing 50% of its original binding
capacity (Fig. 4A). Together the results indicate that the first nucleotide of the GAGATA sequence is essential for the binding of GAGATA binding factor. Furthermore, we analyzed the preference of the GAGATA binding factor for the second nucleotide of the GAGATA sequence. Replacing the second nucleotide of the GAGATA with G abolished the interaction between GAGATA binding factor and PSA3.1 (Fig. 4A). In contrast, changing the second nucleotide of GAGATA sequence to C and T decreased the GAGATA binding factor-PSA3.1 interaction by about 30% and 37%, respectively (Fig. 4A). This result indicated that GAGATA binding factor has stronger interaction with the second nucleotide being replaced by A, C or T, G(A/C/T)GATA, whereas GATA transcription factors have stronger interaction with the first nucleotide being replaced by A or T, (T/A)GATA(A/G). These binding patterns were consistent in relative cold competition in EMSA (Fig. 4B). Together, the results suggested that GAGATA is a novel cis-element essential for the binding of the GAGATA binding factor.

EMSA was performed to determine whether the GAGATA sequence is the recognition site of a known GATA transcription factor. After examining whether the consensus sequence of GATA transcription factors could compete with the specific binding site of PSA3.1, and also if the binding pattern of the consensus sequence in GATA was identical with that of PSA3.1, our results revealed that the consensus sequence of GATA could not compete with PSA3.1 for the binding of GAGATA factor (Figure 5B). Moreover, the binding patterns of PSA3.1 and the binding patterns of the consensus sequence of GATA were different, and as a probe, the GATA consensus sequence failed to have any specific binding in EMSA (Figure 5B). Therefore, the results
indicated that GAGATA binding factor is a novel specific DNA binding protein and known GATA transcription factors are undetectable in LNCaP nuclear extracts by EMSA.

*Mutation of GAGATA and PSA1 ARE in the PSA enhancer core completely inhibited DHT-mediated expression of reporter gene.* To examine whether the binding of GAGATA factor to PSA3.1 is essential for PSA enhancer core function, we replaced the ATA sequence in PSA3.1 with GGG on the pSEAP2/PSA540/Enhancer plasmid (Fig. 6A). Furthermore, this mutation and PSA1 ARE mutation were compared in a transient transfection assay. The PSA1 ARE mutation has been shown to lose its androgen-mediated transcription activity (13). As with the PSA1 ARE mutation, the substitution of ATA with GGG has completely inhibited the DHT-mediated expression of reporter gene in the presence of 1 nM DHT (Fig. 6B). This result indicated that GAGATA binding factor not only interacts physically with PSA enhancer core but also plays an essential role in DHT-mediated enhancer core function, suggesting that the AR-mediated enhancer function of the PSA promoter is not dependent on a single factor, but requires the cooperation of different transcription factors.

*The putative GAGATA binding factor is an approximately 56-60 kDa protein.* To further characterize this GAGATA binding factor, we incubated nuclear extracts of LNCaP-f-AR stable cell line with magnetic beads coupled with double-stranded biotin-PSA3.1. This immobilized DNA magnetic bead demonstrated nearly 100 percent binding capacity for the GAGATA binding factor (Fig. 7A). Southwestern was performed to identify the proteins that specifically interact with GAGATA cis-element. Significantly, an
approximate 56-60 kDa protein that bound to PSA3.1 sequence was detected in the DNA-coupling magnetic bead bound fraction, and mutations that alternated GAGATA motif eliminated binding in Southwestern blotting analysis (Fig. 7B). The concentration of this protein in crude nuclear extracts might be too low to be detected by Southwestern blotting assay (Fig. 8B). Taken together with reporter gene assay (Fig. 6), the evidence suggested that this 56-60 kDa protein might be the GAGATA binding factor and is a positive transcription regulator of PSA expression. Moreover, the molecule mass of this GAGATA binding factor is inconsistent with known GATA transcription factors. The known molecular masses of GATA-1, GATA-2, GATA-3, GATA-4, GATA-5, and GATA-6 are 46 kDa, 46 kDa, 50 kDa, 50 kDa, < 43 kDa and < 50 kDa respectively. This confirms that GAGATA binding factor is a novel transcription factor.

*AR does not interact with PSA3.1.* To investigate the involvement of AR interaction with PSA3.1, we performed EMSA to examine whether the antibodies against f-AR could super-shift PSA3.1-protein complex into higher position. Antibodies against f-AR efficiently abolished interaction of f-AR with ARE, suggesting that these antibodies can interact with the native form of f-AR (Fig. 8B). However, they did not super-shift and did not abolish PSA3.1-DNA interaction (Fig. 8A). These results suggested that AR might not interact with PSA3.1 either directly or indirectly. To further validate this conclusion, we performed Western blotting to check whether f-AR could be pulled down by PSA3.1-coupling magnetic beads efficiently. Comparison of crude nuclear extract with the unbound supernatant indicated that most f-AR did not bind to DNA-coupled magnetic beads with GAGATA binding factor (Fig. 8C). This result was consistent with EMSA
observation, suggesting that AR is not involved in androgen-mediated activation through PSA3.1.
Discussion

The cooperation of specific sets of transcription factors unique to each cell lineage at different stages of differentiation can achieve tissue-specific gene expression. Multiple transcription factors, some prostate-specific, might be involved in the regulation of the PSA gene. Previous studies have indicated that the 6-kb promoter region, but not the 632-bp proximal PSA promoter, is responsible for prostate-specific reporter gene expression in transgenic mice (9). Within this 6-kb region, a 492 bp enhancer core, -4366 to -3874, and a promoter region, -564 to +1, contribute to most of the PSA expression (9, 12). A recent report suggested that multiple low-affinity AREs and a strong ARE in the enhancer core integrate with the high expression of AR in prostate, contributing to the specificity and activity of PSA gene expression (13).

In addition, six respective consensus GATA sites within this enhancer core, located on L7, L1 and R1 of the PSA enhancer core, have been proven to be essential for PSA gene expression (14). The cis-element responsible for the transcription activity of PSA enhancer core in the R1 region was GAGATA, instead of AGATAT, and the L7 region did not seem to contribute much to the PSA enhancer core activity (Fig. 1C). The L1 region consists of two perfect GATA consensus sequences (14), like the L4 and R1, essential for the androgen-mediated transcriptional activity of PSA enhancer core (Fig. 1C). However, we questioned whether the known GATA transcription factors contribute to the transcriptional activity of the L1 region, since none of the known GATA transcriptional factors were detected using the known GATA consensus sequence in the nuclear extracts of LNCaP cells in EMSA (Fig 6A and B). More likely, the cis-elements
responsible for the transcriptional activity of the L1 region might be GAAATA (antisense), rather than the initially proposed GATA consensus sequence (14). We speculated that the GAGATA binding factors might bind to the GAAATA sequence, which is only one nucleotide different from GAGATA. These findings suggest that both GAGATA binding factor, which binds to GAGATA or putative GAGATA binding sequence, and AR, which interacts with the high-affinity ARE and low-affinity ARE in PSA enhancer core, are the determining factors in androgen-regulated PSA expression in prostate tissue.

Several androgen-regulated target genes have been demonstrated to undergo complicated regulation mechanisms mediated by AR alone or in combination with other transcription factors. Kasper et al. (25) have indicated that the androgen-specific regulation of probasin gene transcription requires two AR binding sites, where the binding of AR to both sites occurs in a cooperative, mutually dependent manner. Lin et al. (26) demonstrated that the three L-plastin AREs could cooperate with each other to become a potent testosterone/AR-responsive unit, which was likely responsible for the inducibility of the L-plastin gene by testosterone.

In addition, two genes have been shown to require additional DNA-binding proteins for complete regulation by androgens, the mouse sex-limited protein (Slp) (27, 28) and p21 (29). Two AML/CBFα binding sites within the enhancer of the mouse Slp gene are necessary for the androgen-mediated activation of Slp gene. Both AML/CBFα binding sites are proximal to the ARE in the Slp gene and AML/CBFα can interact with AR. The binding of androgens to AR may induce a conformation change that allows AR to interact with AML/CBFα resulting in the activation of Slp gene. The study of androgen
induction of p21 gene suggested that AR and Sp1 not only bind to their respective consensus sites within the p21 promoter, but also complex with one another, thereby recruiting coactivators and general transcription factors, thus inducing p21 transcription. In contrast to previous studies, our results do not seem to support an interaction of AR with GAGATA binding factor, since the GAGATA binding factor-DNA complex did not shift position in the presence of a high concentration of C3 (I) ARE. However, we cannot rule out the possibility that AR interacts directly with GAGATA binding factors in vivo.

The detailed molecular mechanisms by which AR and GAGATA binding factor cooperate in the androgen-mediated regulation of PSA gene expression still require further elucidation. But this cooperation seems to have neither a synergistic nor an additive effect, since the mutation of either binding site completely abolishes the transcriptional activity induced by androgens. Therefore, it is more likely that the simultaneous binding of GAGATA binding factor and AR to their respective DNA-binding sites results in the maximum effect for the recruitment of general transcriptional machinery ensuring the transcriptional response induced by androgens.

The combination of different transcription factors to regulate gene expression has been widely observed in nature. In most cases, the regulatory region of genes containing different response elements requires a combination of various transcription factors to determine the expression of the genes in an efficient and specific manner (30). In the PSA 455-bp enhancer core region, integration of different transcription factors such as GAGATA binding factors and AR might ensure the expression of PSA gene exclusively in prostate tissue. Different signal transduction pathways may also contribute to the activation of different transcriptional factors in the regulation of a specific gene expression.
Whether other signal transduction pathways are involved in the activation of GAGATA binding factors still remains to be investigated. However, a previous study has indicated that the cross-talk between AR and protein kinase A signal transduction pathways contributes to the androgen-independent induction of PSA gene expression (31). Further cloning and study of GAGATA binding factors may allow us to explore the regulating mechanism of androgen-dependent and androgen-independent PSA gene expression.

In sum, our experiments show that AR and GAGATA binding factor are the most important regulators of PSA expression. Mutation of either AR or GAGATA factor binding sites almost completely abolishes androgen-mediated PSA enhancer activity. We proposed that the binding of the GAGATA binding factor and AR to their respective regulatory DNA sequences on the PSA gene is required for the maximum transcription response to androgens. Further investigation of the possible involvement of the GAGATA binding factor to the androgen-independent induction of PSA gene expression is underway.

**Acknowledgments.** We thank Wayne Hsu and Gary Mawyer for editing the manuscript.
References

1. Henttu, P. and Vihko, P. (1994) *Ann. Med.* **26**, 157-164

2. Moul, J. W. (2000) *J. Urol.* **163**, 1632-1642

3. Castaldo, G., Cecere, G., di, F., V, Prezioso, D., d'Armiento, M., and Salvatore, F. (1997) *Clin. Chim. Acta* **265**, 65-76

4. Kim, J. and Logothetis, C. J. (1999) *Urol. Clin. North Am.* **26**, 281-290

5. Ruckle, H. C. and Oesterling, J. E. (1993) *World J. Urol.* **11**, 227-232

6. Richie, J. P. (1999) *Urology* **54**, 15-18

7. Sadar, M. D., Hussain, M., and Bruchovsky, N. (1999) *Endocr. Relat. Cancer* **6**, 487-502

8. Bhayani, S. B. and Andriole, G. L. (1999) *Semin. Urol. Oncol.* **17**, 148-153

9. Cleutjens, K. B., van der Korput, H. A., Ehren-van Eekelen, C. C., Sikes, R. A., Fasciana, C., Chung, L. W., and Trapman, J. (1997) *Mol. Endocrinol.* **11**, 1256-1265

10. Schuur, E. R., Henderson, G. A., Kmetec, L. A., Miller, J. D., Lamparski, H. G., and Henderson, D. R. (1996) *J. Biol. Chem.* **271**, 7043-7051

11. Riegman, P. H., Vlietstra, R. J., van der Korput, J. A., Brinkmann, A. O., and Trapman, J. (1991) *Mol. Endocrinol.* **5**, 1921-1930

12. Cleutjens, K. B., van Eekelen, C. C., van der Korput, H. A., Brinkmann, A. O., and Trapman, J. (1996) *J. Biol. Chem.* **271**, 6379-6388

13. Huang, W., Shostak, Y., Tarr, P., Sawyers, C., and Carey, M. (1999) *J. Biol. Chem.* **274**, 25756-25768

14. Perez-Stable, C. M., Pozas, A., and Roos, B. A. (2000) *Mol. Cell Endocrinol.* **167**, 43-53

15. Abmayr, S. and Workman, J. (1993) DNA-Protein interaction. In Ausubel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J., and Struhl, K., editors. *Current Protocols in Molecular Biology*, WILEY INTERSCIENCE,

16. Wang, C., Young, W. J., and Chang, C. (2000) *Endocrine.* **12**, 69-76
17. Tan, J. A., Marschke, K. B., Ho, K. C., Perry, S. T., Wilson, E. M., and French, F. S. (1992) *J.Biol.Chem.* **267**, 4456-4466

18. De Vos, P., Claessens, F., Winderickx, J., Van Dijck, P., Celis, L., Peeters, B., Rombauts, W., Heyns, W., and Verhoeven, G. (1991) *J.Biol.Chem.* **266**, 3439-3443

19. Tsukiyama, T. and Wu, C. (1995) *Cell* **83**, 1011-1020

20. Charron, F. and Nemer, M. (1999) *Semin.Cell Dev.Biol.* **10**, 85-91

21. Merika, M. and Orkin, S. H. (1993) *Mol.Cell Biol.* **13**, 3999-4010

22. Laverriere, A. C., MacNeill, C., Mueller, C., Poelmann, R. E., Burch, J. B., and Evans, T. (1994) *J.Biol.Chem.* **269**, 23177-23184

23. Weiss, M. J. and Orkin, S. H. (1995) *Exp.Hematol.* **23**, 99-107

24. Simon, M. C. (1995) *Nat.Genet.* **11**, 9-11

25. Kasper, S., Rennie, P. S., Bruchovsky, N., Sheppard, P. C., Cheng, H., Lin, L., Shiu, R. P., Snoek, R., and Matusik, R. J. (1994) *J.Biol.Chem.* **269**, 31763-31769

26. Lin, C. S., Lau, A., Yeh, C. C., Chang, C. H., and Lue, T. F. (2000) *DNA Cell Biol.* **19**, 1-7

27. Adler, A. J., Scheller, A., and Robins, D. M. (1993) *Mol.Cell Biol.* **13**, 6326-6335

28. Ning, Y. M. and Robins, D. M. (1999) *J.Biol.Chem.* **274**, 30624-30630

29. Lu, S., Jenster, G., and Epner, D. E. (2000) *Mol.Endocrinol.* **14**, 753-760

30. Carey, M. (1998) *Cell* **92**, 5-8

31. Sadar, M. D. (1999) *J.Biol.Chem.* **274**, 7777-7783
Footnotes

1Wang, C., Jiang, P., Geng, J., Weinmann, R., Attar, M. R., Gottardis, M and Wang, G. (Manuscript submitted).
Figure Legend

FIG. 1. Identification of a cluster of cis-elements in the PSA enhancer core. A, sequence of the PSA enhancer core. Regions that are replaced by linker scanning mutagenesis are underlined and indicated as L14-1, LO and R1-10. The number of the PSA enhancer core sequences relative to the transcriptional start site was shown as indicated, and the previous identified ARE sites (13) were marked by boxes. B, effect of linker replacement mutations on enhancer activity in LNCaP cells in the absence of 1 nM DHT. C, effect of linker replacement mutations on enhancer activity in LNCaP cells in the presence of 1 nM DHT. Linker scanning mutants were indicated as described in A above. Cs represents wild-type control. Relative luciferase activities obtained from three independent experiments were represented as a bar graph.

FIG. 2. Formation of PSA3-protein complexes in EMSA. A, the oligonucleotide sequences used in EMSA. Sequences corresponding to the linker replacement mutations are indicated as shadow boxes. C3 (I) ARE was used as positive control and its sequence is indicated. B, EMSA of purified E. coli expressed AR DBD-LBD-ARE complexes. 0.014 pmoles of about $10 \times 10^4$ cpm of C3 (I) ARE (lane 1), PSA1 ARE (lane 2), PSA2 ARE (lane 3) and PSA3 (lane 4) were incubated with AR DBD-LBD and analyzed by EMSA as described under “Materials and Methods.” C, EMSA of AR-ARE complexes and PSA3-proteins complexes. 0.014 pmoles of about $10 \times 10^4$ cpm of C3 (I) ARE, PSA1 ARE, PSA2 ARE and PSA3 were incubated with LNCaP nuclear extract, and complexes were analyzed by EMSA as described in “Materials and Methods.” Incubation was in the absence or presence of 200 x cold oligonucleotide as indicated.
FIG. 3. Effects of mutations of PSA3.1 in EMSA. A, sequences of 4 different mutants of PSA3.1. The mutated sequence was underlined. B, competition of PSA3.1 by 4 different PSA3.1 mutants in EMSA. 0.014 pmoles of about 10 x 10^4 cpm of WT PSA3.1 were incubated with LNCaP nuclear extracts. The analysis method was the same as described in FIG. 2C. Incubations were in the absence or presence of 10x and 100x cold WT and various mutated cold oligonucleotides as indicated. S represents specific binding and NS represents non-specific binding.

Fig. 4. Effects of point mutations of GAGATA in EMSA. A, the substitution of the first nucleotide of GAGATA with A, C, or T, and the substitution of the second nucleotide of GAGATA with G, C, or T, respectively, affecting PSA3.1-GAGATA binding factor interaction. The density of the GAGATA binding factor-DNA complex was quantified by image analysis software and presented as a bar graph. 0.014 pmoles of about 10 x 10^4 cpm of WT and the respective point mutations of the PSA3.1 were incubated with LNCaP nuclear extracts. WT and mutated forms were shown as indicated. The analysis method was the same as described in FIG. 2C. B, competition of PSA3.1 by 6 different PSA3.1 mutated oligonucleotides in EMSA. The density of the GAGATA binding factor-DNA complex was quantified by image analysis software and presented as a bar graph. 0.014 pmoles of about 10x10^4 cpm of PSA3.1 were incubated with LNCaP nuclear extracts. Incubations were in the absence or presence of 2x, 10x and 100x cold WT and mutation.
cold oligonucleotide sequences as indicated. The analysis method was the same as the one described in FIG. 2C.

FIG. 5. GAGATA binding factor is not a known GATA transcription factor. A, sequences of PSA3.1, PSA3.1 mutant, consensus GATA and GATA mutant, and the binding sequence as indicated by underline. B, competition of PSA3.1 by PSA3.1, PSA3.1 mutant, GATA and GATA mutant GATA in EMSA and competition of GATA by GATA, GATA mutant, PSA3.1 and PSA3.1 mutant in EMSA. 0.014 pmoles of about 10x10^4 cpm of PSA3.1 and GATA, respectively, were incubated with LNCaP nuclear extracts. The analysis method was the same as described in FIG. 2C. Incubations were in the absence or presence of 10x cold WT and mutation cold oligonucleotides as indicated.

FIG. 6. Effects of mutation of the GAGATA binding site for PSA enhancer core activity in LNCaP cells. A, the position of PSA promoter and PSA enhancer core in pSEAP2/PSA540/Enhancer construct. The substitution of ATA by GGG in PSA3.1 is indicated below the figure. B, comparison of SEAP activity between WT and mutants in the absence or presence of 1 nM DHT. Experimental details of the transfection and SEAP reporter gene activity are described in "Materials and Methods." The SEAP activities of various constructs were normalized by luciferase activity with pGL3-control and presented as a bar graph. AREm and GAGATAm represent the ARE mutation in PSA1 and the GAGATA mutation in PSA3.1, respectively.
FIG. 7. An about 56-60 kDa protein is the putative GAGATA binding factor. A, EMSA showing that DNA-coupled magnetic beads had nearly 100 percent binding capacity for the GAGATA binding proteins. The complex of GAGATA binding protein and PSA3.1 DNA formed from nuclear extract before incubating with DNA-coupled magnetic beads is shown in lane 1. In contrast, no complex formed from the same nuclear extract after incubating with DNA-coupled magnetic beads is shown in lane 2. EMSA experimental details were as described in FIG. 2C. B, Southwestern blotting analysis for SDS-PAGE of proteins from crude nuclear extracts, unbound supernatant and DNA-coupled magnetic bead binding fraction demonstrated that an about 56-60 kDa protein can specifically interact with PSA3.1 but not PSA3.1 mutation probe. Thirty microgram of nuclear extract, 30 µg of unbound supernatant, and the proteins extracted from DNA-coupling magnetic beads were separated on 10% SDS-polyacrylamide gels, followed by Southwestern blot assay. Experimental details of SDS-PAGE and Southwestern blotting assay were described in “Materials and Methods.” The relative molecular size was indicated on the left. N; crude nuclear extract. S; supernatant; B; DNA-coupling magnetic bead bound fraction.

FIG. 8. AR does not interact with PSA3.1. A, ANTI-FLAG M2 mAb could not supershift and not deplete GAGATA binding protein-PSA3.1 complex. 0.014 pmoles of about 10x10^4 cpm of PSA3.1 were incubated with LNCaP-f-AR nuclear extracts and various amounts of ANTI-FLAG M2 mAb. B, ANTI-FLAG M2 mAb could abolish f-AR to interact with PSA2 ARE. 0.014 pmoles of about 10x10^4 cpm of PSA2 ARE were incubated with LNCaP-f-AR nuclear extracts and 4 µg of ANTI-FLAG M2 mAb.
Experimental details for EMSA are as described in FIG. 2C. C, Western blotting indicated that PSA3.1-coupled magnetic beads could not pull down AR. Nuclear extracts (30 µg), and unbound supernatants (30 µg) were separated on 10% SDS-PAGE, and followed by Western blotting as described in “Materials and Methods.” The density of f-AR was quantified by image analysis software and presented as a bar graph. N; crude nuclear extract. S: supernatant.
### A

**PSA Enhancer Core**

|         | L          | R          |
|---------|------------|------------|
| 14      |            |            |
| 13      |            |            |
| 12      |            |            |
| 11      |            |            |
| 10      |            |            |
| 9       |            |            |
| 8       |            |            |
| 7       |            |            |
| 6       |            |            |
| 5       |            |            |
| 4       |            |            |
| 3       |            |            |
| 2       |            |            |
| 1       |            |            |
| 0       |            |            |
| 1       |            |            |
| 2       |            |            |
| 3       |            |            |
| 4       |            |            |
| 5       |            |            |
| 6       |            |            |
| 7       |            |            |
| 8       |            |            |
| 9       |            |            |
| 10      |            |            |

- **L4 PSA1 ARE**
  - TGGAGGAAACATATTGTATCGATTGT

- **L1-L0 PSA2 ARE**
  - AGGACAGTAAGCAAGCCTGGATCT

- **R1-R2-R3 PSA3**
  - AGAGAGATATCATCTTGCAAGGATGCTGCTTTACAAACATCCTG
  - GAA

- **C3(1) ARE**
  - ACATAGTACGTGATGTTCCTCAAGA

### B

### C

|          | C3(1) | PSA1 | PSA2 | PSA3 |
|----------|-------|------|------|------|
| 200X C3(1) |       |      |      |      |
| 200X PSA1   |       |      |      |      |
| 200X C3(1)   |       |      |      |      |
| 200X PSA2   |       |      |      |      |
| 200X PSA3 |       |      |      |      |
| 200X C3(1) |       |      |      |      |
Figure 3

A

PSA 3.1

wt     AGAGAGATATCATCTTGCAGGAT
m1     CCGAGATATCATCTTGCAGGAT
m2     AGATTTATATCATCTTGCAGGAT
m3     AGAGAGGGTATCATCTTGCAGGAT
m4     AGAGAGATAAAATCATCTTGCAGGAT

B

100X 100X 100X 100X 100X
10X   10X   10X   10X   10X

wt     m1     m2     m3     m4

S     NS
Figure 4

A

| wt  | mutant |
|-----|--------|
| GMGATA | GGGATA |
| GGGATA | GTGATA |
| AAGATA | CAGATA |
| TAGATA |        |
Figure 5

A

|          | PSA3.1       | PSA3.1m        |
|----------|--------------|----------------|
|          | AGAGAGATA    | AGAGAGGGG      |
|          | TCATCTTGCAAGGAT | CATCTTGCAAGGAT |
| GATA     | CACTTGATAA   | CACTCTTTAA     |
| GATAm    | CAGAAAGTCTTAACCTCT | CAGAAAGTCTTAACCTCT |

B

|          | PSA3.1       | GATA           |
|----------|--------------|----------------|
|          | PSA3.1       | PSA3.1m        |
|          | GATA         | GATAm          |
|          | 10x10x10x10x | 10x10x10x10x   |

S

NS
Figure 6

A

-5322 -4366 -4140-4050 -3873-539 -400 -170 -1

Enhancer core Promoter SEAP

ARE PSA3.1 ARR ARE

PSA3.1
GAGATAm

AGAGAGATATCATCTTGCAAGGAT
AGAGAGGGGTCATCTTGCAAGGAT
Figure 7

A

GAGATA binding protein-DNA complex

B

| Probe | PSA3.1 | PSA3.1m |
|-------|--------|---------|
| kDa   | N      | S       | B      |
| 220   |        |         |        |
| 98    |        |         |        |
| 64    |        |         |        |
| 50    |        |         |        |

GAGATA binding protein
|       | A                  | B                  | C                  |
|-------|-------------------|-------------------|-------------------|
| mAb (µg) | 0  2  4  6       | 0  4              |                   |
|        | GAGATA binding    | f-AR-DNA complex  | N S               |
|        | protein-DNA       |                   |                   |
|        | complex           |                   |                   |

Figure 8
Identification of a novel transcription factor, GAGATA binding protein, involved in androgen-mediated expression of prostate-specific antigen
Chihuei Wang, Fan Yeung, Po-Chun Liu, Ricardo M Attar, Jieping Geng, Leland W.K Chung, Marco Gottardis and Chinghai Kao

*J. Biol. Chem.* published online June 2, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M207862200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts