Activation of Prostate-Specific Antigen Precursor (pro-PSA) by Prostin, a Novel Human Prostatic Serine Protease Identified by Degenerate PCR[†,‡]

Thomas K. Takayama,* Corey A. Carter, and Ta Deng

Departments of Biochemistry and Urology, University of Washington, Box 357350, Seattle, Washington 98195-7350.

Received September 11, 2000; Revised Manuscript Received December 7, 2000

ABSTRACT: A novel serine protease was found in human prostate by degenerate oligonucleotide PCR amplification and cloned. The zymogen form of this enzyme, named prostinogen, is composed of 240 amino acid residues with an amino-terminal propiece of 5 residues and a 235-residue mature enzyme. The transcript has a signal peptide of 15 amino acid residues. The mature enzyme has 41% sequence identity with prostate specific antigen (PSA). Prostinogen was expressed in *Escherichia coli* and refolded from inclusion bodies. The zymogen, with a molecular mass of 28 kDa, was readily activated by agarose-immobilized trypsin to generate prostin, a serine protease, which cleaves the chromogenic substrate (*N*-benzoyl-L-Ile-L-Glu-L-Gly-L-Arg-*p*-nitroaniline hydrochloride) (S-2222). Recombinant prostin readily activates the precursor of PSA (pro-PSA) by cleavage of the amino terminal $Arg^{7}-Ile^{8}$ peptide bond. These results indicate that prostin may be a physiological activator of pro-PSA following its own proteolytic activation, as part of a cascade system involving a series of serine protease precursor proteins in the prostate.

Prostate specific antigen (PSA),[1] a serine protease produced by the prostate (*1*−*5*), has been extensively used as a marker for prostate cancer since the mid-1980s (*6*−*10*). It has chymotrypsin-type specificity toward its physiologic substrates, semenogelin I and II (*11*), which are the major gel-forming proteins in seminal plasma. PSA has also been implicated in growth regulation because it cleaves IGFBP3 (*12*) and PTHrp (*13*, *14*), which are proteins associated with cellular growth regulation and breast cancer metastasis, respectively. PSA is a member of the human tissue kallikrein family (*15*, *16*) and is synthesized as a zymogen (pro-PSA) (*15*). Pro-PSA is activated by human glandular kallikrein-1 (hGK1, hK2), another prostate-specific kallikrein (*17*, *18*), and possibly by other serine protease(s) produced within the prostate (*19*).

Several serine proteases in addition to PSA and hK2 have been isolated from the human prostate. Prostasin (*20*) was purified from seminal plasma using an aprotinin-agarose column. This protein is expressed in prostate, liver, salivary gland, kidney, lung, pancreas, colon, and bronchus. Its physiologic function has not yet been determined. More recently, another member of the tissue kallikrein family named prostase (KLK4 and KLK-L1) was reported, initially by Nelson et al. (*21*) and subsequently by other independent groups (*22*, *23*). This protein is produced primarily by the prostate but also has a high level of expression in the testis, mammary gland, adrenals, uterus, thyroid, and salivary glands (*23*). Although its function is unknown, it is homologous to porcine enamel matrix protease that is involved in tooth remodeling (*24*). TMPRSS2, a membrane-associated serine protease, was found by exon-trapping on chromosome 21 (*25*). It was subsequently shown to be expressed primarily in the prostate (*26*), as well as in the pancreas.[2] Immunohistochemistry studies have localized this protein to the basal cells of the prostate gland, but its physiological function is as yet unknown. All these serine proteases that have trypsin-type substrate specificities are potential physiologic activators of pro-PSA. Chymotrypsin-type enzymes other than PSA have also been detected in seminal plasma by several groups (*27*, *28*). These enzymes are not involved in pro-PSA activation because of differences in substrate specificity.

We previously demonstrated the activation of pro-PSA by several serine proteases, including hK2, and theorized that pro-PSA may require a cascade-like series of proteolytic activations (a PSA cascade) involving multiple serine proteases (*19*), analogous to the blood coagulation system (*29*). We proceeded to search for other serine proteases that are produced at significant levels within the prostate. A recently described, improved method of degenerate oligonucleotide PCR (*30*) was utilized to identify other potential activator-

[†] This work was supported in part by Grant DK02447 from the National Institutes of Health, the Richard M. Lucas Foundation, the American Foundation for Urologic Disease, and the American Cancer Society.

[‡] Nucleotide sequence reported in this paper has been submitted to the GenBank Data Bank with accession number AF303046.

* To whom correspondence should be addressed. E-mail: tomt@u.washigton.edu. FAX: (206) 543-5368.

[1] Abbreviations: PSA, prostate-specific antigen; hK2, human glandular kallikrein-1; pro-PSA, PSA precursor; rPSA, recombinant (active) PSA; pro-hK2, hK2 precursor; NES1, normal epithelial cell-specific gene 1; uPA, urokinase-type plasminogen activator; CODEHOP, consensus-degenerate hybrid oligonucleotide primers; RACE, rapid amplification of cDNA ends; S-2222, *N*-benzoyl-L-isoleucyl-L-glutamyl-glycyl-L-arginine-*p*-nitroaniline hydrochloride; S-2251, H-D-valyl-L-leucyl-L-lysine-*p*-nitroanilide dihydrochloride; S-2302, H-D-prolyl-L-phenylalany-L-arginine-*p*-nitroaniline hydrochloride; S-2586, 3-carbomethoxypropionyl-L-arginyl-L-prolyl-L-tyrosine-*p*-nitroaniline hydrochloride.

[2] Takayama et al., unpublished data.

10.1021/bi002129r CCC: $20.00
Published on Web 01/13/2001

Table 1: Serine Proteases Identified in Human Prostate cDNA Library by Degenerate Oligonucleotide PCR Amplification

1. prostinogen
2. prostate specific antigen (PSA, KLK3)
3. Human glandular kallikrein (hK2, KLK2)
4. prostase (KLK4, KLK-L1)
5. TMPRSS2
6. kidney kallikrein (KLK1)
7. testisin
8. NES1
9. protein C (coagulation factor)
10. trypsin
11. chymotrypsin
12. chymotrypsin B1
13. elastase IIA
14. granzyme K
15. unknown serine protease

(s) of the pro-PSA cascade. One product of this search was a novel serine protease precursor identified from a human prostate cDNA library (CLONTECH). This protein was then cloned, expressed, and refolded. It was activated by trypsin-agarose to generate the active form that was subsequently shown to activate pro-PSA. The proteolytic activity toward pro-PSA was significantly higher than that of hK2. We have named this protein's precursor form, prostinogen (GenBank Accession No. AF303046), and its active form, prostin. Prostin may well be a primary physiologic activator of pro-PSA.

EXPERIMENTAL PROCEDURES

Degenerate Oligonucleotide PCR Amplification. Degenerate primers were designed using the consensus-degenerate hybrid oligonucleotide primers (CODEHOP) algorithm (*30*) (http://blocks.fhcrc.org/codehop.html/) by entering the consensus sequences flanking the catalytic histidine residue (for 5′ end primer) of 10 serine proteases (PSA, chymotrypsin, factor X, factor XII, hK2, plasma kallikrein, plasmin, prothrombin, tissue plasminogen activator, and trypsin) and the catalytic serine residue (for the 3′ end primer) of eight of the same serine proteases. The resultant 5′ primer was 5′-GCT GGG TGC TGA CCG CNG CNC AYT G-3′, and the 3′ primer was 5′-CCA GGG GGC CGC CNS WRT CNC C-3′. Degenerate alphabet was from CODEHOP (*30*) (http://blocks.fhcrc.org/blocks/help/CODEHOP/degen.html). Marathon cDNA library from the human prostate (CLONTECH) 2 μL, degenerate primers at 1.2 μM (final) each, Advantage PCR mix and buffer (CLONTECH), and 0.2 mM (final) dNTP were mixed in a total volume of 20 μL for PCR amplification according to the manufacturer's instructions. Temperature cycles were 94 °C for 5 min; 35 cycles of 94 °C for 1 min, 65 °C for 1 min, 68 °C for 1 min; followed by 68 °C for 5 min. The PCR products were subcloned into 2.1-TOPO vector using a TA cloning kit (Invitrogen). TA flanking primers were used for PCR amplification of the expected ~450-bp inserts. By restriction digestion of the inserts, the redundant and known transcripts were identified. For example, *Nae*I was used for prostase, elastase, protein C, and TMPRSS2; *Bsu36*I was for PSA and hK2; *Sty*I was for TMPRSS2, chymotrypsin, chymotrypsin B1, NES1 (*31*), kidney kallikrein, and prostin; *Apo*I was for granzyme K and kidney kallikrein (see Table 1) (*32*). Unique inserts were sequenced using M13 forward or reverse primers employing the ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction kit (Perkin-Elmer), G60 spin columns (BioRad), and the Applied Bioscience Automated sequencer (Perkin-Elmer).

Cloning of Prostinogen. The 5′ and 3′ cDNA sequences were obtained from a prostate Marathon-ready cDNA library (a generous gift from Wen-Feng Xu of Zymogenetics, Inc.) by rapid amplification of cDNA ends (RACE) (*33*). The RACE PCR amplification methods were optimized using a "touchdown" thermal cycling program (*34*) and sequential nested-primer reactions. Oligonucleotide primers were designed from the 450-bp sequence from the degenerate PCR above. PCR (*33*) was performed with a Perkin-Elmer/Cetus DNA thermal cycler and thermostable DNA polymerase KlenTaq (CLONTECH) according to the manufacturer's instructions. PCR products were cloned (TOPO TA cloning kit, Invitrogen) and sequenced by the methods described above.

Northern Blot Analysis and Multiple Tissue PCR. A region on the 3′ untranslated cDNA (467 bp) that does not have significant sequence identity with other serine proteases was PCR amplified using two primers, UNT F 5′-TTC TAG CCT ATC TCC TGT GCC CCT GAC TGA-3′ and UNT R 5′-TTC AGA CGA ATC ATC TCA TTT AAT CCT CAC-3.′ The subsequently purified (Qiagen) PCR product and Beta-actin cDNA (CLONTECH) were radiolabeled with [α-^{32}P]-dCTP (6000 Ci/mmol; 1 Ci = 37 Gbq; Amersham) by random priming (Prime-It II kit, Stratagene) to a specific activity of 10^9 cpm/μg. Human multiple-tissue Northern blots (CLONTECH) representing 16 different human tissues were prehybridized with ExpressHyb solution (CLONTECH) according to the manufacturer's instructions with a final probe concentration of 10^6 cpm/mL. The blots were washed for 1 h at room temperature in 2× standard saline citrate/0.05% SDS and for 1 h at 65 °C in 0.1× standard saline citrate/0.1% SDS and visualized by using a Molecular Dynamics Phosphorimager SF after overnight exposure. For multiple tissue PCR (CLONTECH), PCR amplification of cDNA from various human tissues was performed according to the manufacturer's instructions for 32 cycles using two PCR primers: UP2 5′-AGG GAA ACC ATG AAG AGG AAC TGA CT-3′ and LP2 5′-AAC AGC CCT TTG AAG TCA GCA ATA AG-3′ and produced the expected 476-bp band. Control primers were used for 30 cycles.

Expression and Refolding of Prostinogen. The majority of procedures were described previously (*19*). The cDNA of prostinogen (including a 9-amino acid sequence upstream to the activation site) was inserted immediately behind the initiatior methionine codon of a pET12 vector to construct pET12-prostinogen according to the published methods (*19*). Primers used were PET F 5′-CAT ATG GCA TCC ACA GCC CAG GAT GGT GAC-3′ and PET R 5′-GGA TCC TCA GTT CCT CTT CAT GGT TTC CCT-3′. *Escherichia coli* strain BL21(DE3) was transformed, the protein was expressed, and the inclusion bodies were purified and solubilized by the previously published methods (*19*) with some modifications. Solubilization of the inclusion body required a pH of 3.8. The solubilized inclusion bodies (~40 mg in 20 mL) were refolded by diluting into 8 L of 2 M urea, 50 mM Tris-HCl, pH 8.8, 0.5 M NaCl, 0.1 M NH_4Cl, 1 mM EDTA, 10 mM benzamidine, 1.25 mM reduced glutathione, and 0.5 mM oxidized glutathione followed by slowly stirring for 24 h at 4 °C. The final concentration of

the protein was ~5 μg/mL. The refolding buffer was degassed and flushed under nitrogen prior to use. The refolded sample was then dialyzed against 80 L of 10 mM Tris-HCl, pH 8.8, with three buffer changes. NaCl at 0.1 M and Tween 20 at 0.01% were added, and the sample was concentrated to 4 mL by ultrafiltration using a PM10 membrane (Amicon). The concentrated protein was stored at −20 °C until use.

Activation of Prostinogen by Immobilized Trypsin. Trypsin-agarose (Sigma), 200 μL was gently washed four times by mixing with 1.0 mL of 1 mM HCl and centrifuged at low speed (Picofuge, Stratagene). Trypsin activity of the wash was measured as follows: the trypsin-agarose pellet was suspended in 100 μL of wash solution, and the supernatant (45 μL) obtained by centrifugation was taken for assay. The sample was incubated at room temperature with 5 μL of 10× buffer (1 M NaCl, 1 M Tris-HCl, pH 7.5) and 50 μL of 1 mM S-2222, and the absorbance change (405 nm) was measured using a plate reader as described (*19*). After confirming that there was no residual activity in the wash solution, the trypsin-agarose pellet was resuspended with 100 μL of water and used for the activation of prostinogen. Crude refolded prostinogen (100 μL, 0.19 mg) was mixed with 5 μL of trypsin-agarose and 5 μL of buffer containing 1 M NaCl, 1 M Tris-HCl, pH 7.5, in a total volume of 110 μL. The sample was incubated at 37 °C; gentle mixing was required every minute to keep the trypsin-agarose in suspension. The reaction was terminated by 20 s centrifugation at low speed (Picofuge) of the 25 μL samples taken at various time points (0, 15, 30, and 60 min). After centrifugation, 15 μL of the 25 μL aliquots were removed for measurement of amidolytic activity. Care was taken not to pipet any trypsin-agarose at the bottom of the tube. The sample was mixed with 30 μL of water and 5 μL of 10× buffer, and then transferred to an ELISA plate; 50 μL of 1 mM S-2222 was added to each sample (0.5 mM final). The absorbance increase at 405 nm was monitored at room temperature at 1-min intervals for 15 min. Control reactions were carried out in the absence of prostinogen.

Activation of pro-PSA by Prostin and by hK2. Recombinant pro-PSA (*19*) (5 μg) was incubated at 37 °C with 33.7 μg of crude prostin (prepared by trypsin activation of crude prostinogen as described above) in a final 50 μL of 0.1 M Tris-HCl, pH 7.5, and 0.1 M NaCl. At various time intervals, the reaction was stopped by the addition of 5 μL of aprotinin (0.02 mg/mL). To measure the amidolytic activity of the newly generated active rPSA, 50 μL of 1 mM S-2586 (0.5 mM final) was added to the above reaction, and the absorbance was monitored as above. As a control, 33.7 μg of crude prostinogen (supernatant sample from time zero activation with trypsin-agarose above) was used instead of the activated prostin. Recombinant hK2 (*19*) (0.75 μg) was incubated with recombinant pro-PSA (5 μg) and the enzyme activity of rPSA (activated PSA) was measured in the same fashion as by prostin above. The amount of hK2 was the same as prostin, estimated to be 0.75 μg by SDS−PAGE [calculation was based on the intensity of 28 kDa band contained in a crude prostin sample compared to the known amount of PSA (2 μg) in an adjacent lane, Figure 5, lanes 1 and 2].

Measurement of the Amidolytic Activity of Prostin Against other Chromogenic Substrates. Prostin (20.5 μg of crude preparation containing an estimated 0.45 μg of prostin) in the same buffer condition as for S-2222 above, was incubated at room temperature with 50 μL of S-2586 (Arg−Pro−Tyr-*p*NA), S-2251 (Val−Leu−Lys-*p*NA), or S-2302 (Pro−Phe−Arg-*p*NA) 0.5 mM final, and the increase in absorbance (405 nm) was measured by a plate reader as above. The amidolytic activities were expressed as the amounts of *p*NA generated (mg of protein)$^{-1}$ min^{-1}.

Other Procedures. Protein concentrations were determined by micro BCA protein assay (Pierce) using bovine serum albumin as a standard. Protein sequence analysis and SDS−PAGE were performed according to the previously published methods (*19*). Protein bands were visualized by staining with Coomassie Brilliant Blue. Trypsin-agarose (TPCK-treated bovine pancreatic trypsin attached to beaded agarose) was from Sigma (T-4019). The remaining materials were as described previously (*19*).

RESULTS

Degenerate Oligonucleotide PCR Amplification and Cloning of Prostinogen. Degenerate oligonucleotide PCR primers were designed using short stretches of conserved peptide sequences from 8 to 10 different serine proteases (see Experimental Procedures). These proteases are primarily trypsin-type, except for PSA and chymotrypsin. This is because we wished to focus our initial search to trypsin-type proteases, which in general are activators of pro-enzymes in various proteolytic cascades. Other conserved chymotrypsin-type sequences may be used in the future design to search for serine proteases involved in relatively nonspecific degradative processes. PCR amplification of a human prostate cDNA library produced the expected 450 bp-sized fragments. The fragments were subcloned into a TA vector and transformed, and the bacterial colonies were isolated. Of ~500 individual colonies (inserts), over 90% were identified as containing redundant and/or known serine protease sequences by restriction digest analysis (see Experimental Procedures). The unique inserts were sequenced and compared to known serine proteases (BLAST search). The presence of 15 different serine proteases was thus demonstrated (Table 1). This plethora of serine proteases identified in a single cDNA library using a single pair of degenerate primers was a highly unusual result. At the time of the initial sequence analyses, four of the 15 serine proteases had not been reported. Subsequently, two were described in the literature, prostase (*21*) and testisin (*35*). The 450-bp novel sequences from the remaining two clones were used to design oligonucleotide primers for subsequent 5′ and 3′ RACE from human prostate cDNA library (CLONTECH). The sequence of the 5′ end including the start codon could not be obtained from the Clontech cDNA library. Therefore, 5′ RACE was performed on a human prostate cDNA library from Zymogenetics, Inc. One of the two RACE reactions was successful. A composite sequence (Figure 1) was generated from the overlapping sequences of the 450-bp middle segment and the fragments from the 5′ and 3′ RACE reactions. Subsequent PCR amplification using primers complementary to the start and stop codon regions (1 to 768) generated the expected sequence of prostinogen, and the sequence was confirmed for both strands.

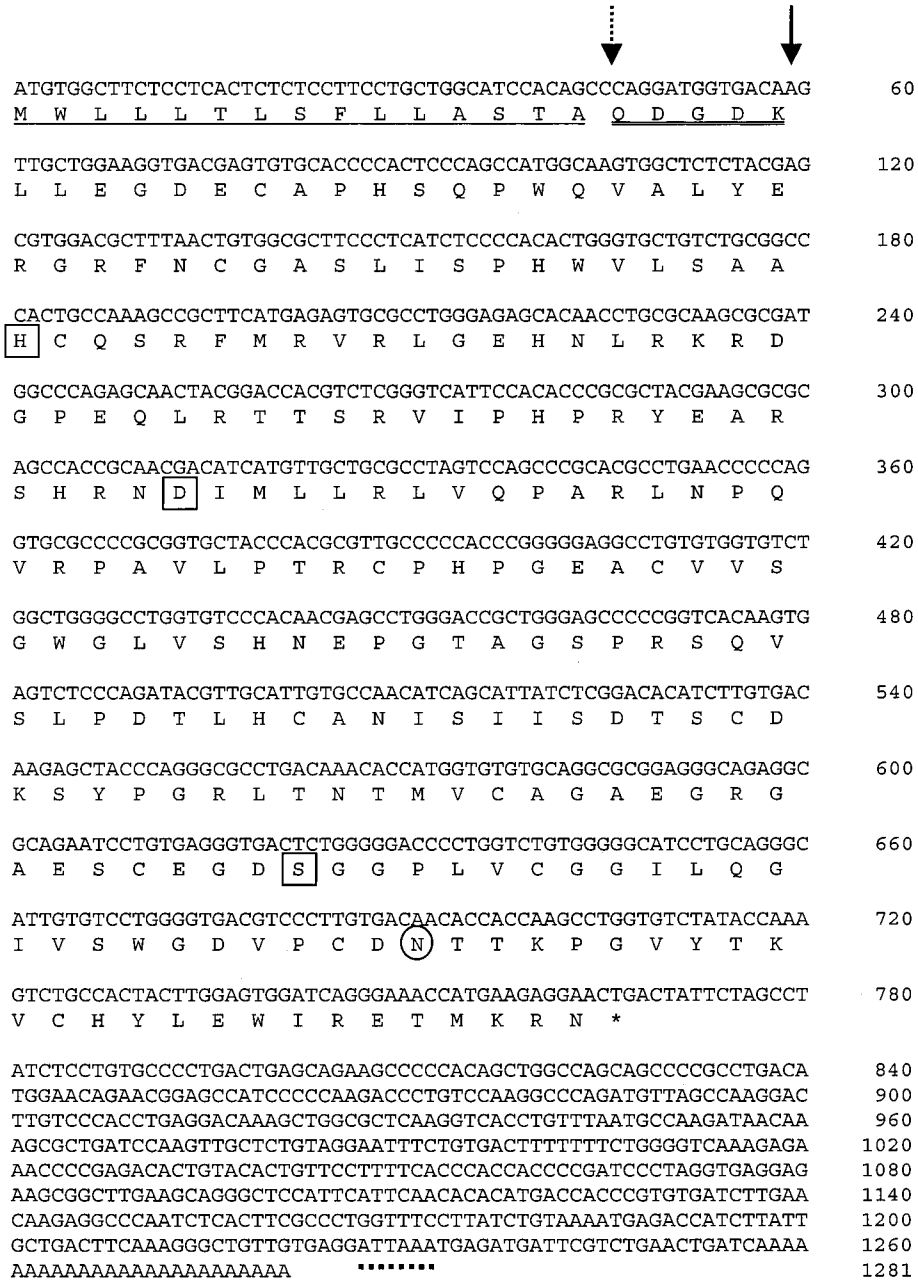

FIGURE 1: Nucleotide sequence of prostinogen cDNA and deduced amino acid sequence. The proposed signal peptide sequence is underlined and its cleavage site is indicated with a dotted downward arrow. The pro-piece sequence is underlined twice and the activation cleavage site is indicated with a solid downward arrow. The catalytic triad residues are boxed. A variant polyadenylation signal, ATTAAA, is underlined with a dotted line. The potential *N*-glycosylation site is circled.

The human multiple tissue Northern blot analysis (see below for details) using the radiolabeled probes based on the 3′ prime RACE demonstrated the presence of a 1.3-kb transcript in testis and also a very low level in the prostate, an approximately 1.5 kb transcript was also seen at low levels in the human colon.

Sequence Analysis and Tissue Distribution of Prostinogen. Figure 1 shows the amino acid and cDNA sequences of human prostinogen. It is the precursor of a serine protease composed of a signal peptide of 15 residues, a propiece of 5 residues, and a serine protease of 235 residues. The zymogen has a calculated molecular mass of 28.0 kDa and a pI of 8.27 in the absence of glycosylation. The expected cleavage between Lys^{5}−Leu^{6} results in the mature enzyme having the catalytic triad His^{46}, Asp^{90}, and $Ser.^{193}$ When compared to other serine proteases identified in the prostate gland, prostin has 41% amino acid sequence identity with PSA, 43% with hK2, 38% with prostase, 31% with TMPRSS2, and 35% with prostasin (Figure 2). A BLAST search with the prostinogen sequence also identified the partial sequence of a serine protease-like protein named ACO (Genbank Accession Number S45356) (*36*), found in the human brain and having 69% sequence identity with prostin. This sequence was reported as a serine protease-like protein although the investigators were unable to obtain the 5′ end sequence including the activation site.

The presence of the acidic residue Glu^{187} in prostinogen is similar to most trypsin-type enzymes that have an Asp at this position. The active enzyme has an *N*-terminal sequence of Leu−Leu−Glu−Gly instead of the more common Ile−Val−Gly−Gly sequence. The Lys^{5} prior to the Leu−Leu−Glu−Gly indicates that the activation of prostinogen requires

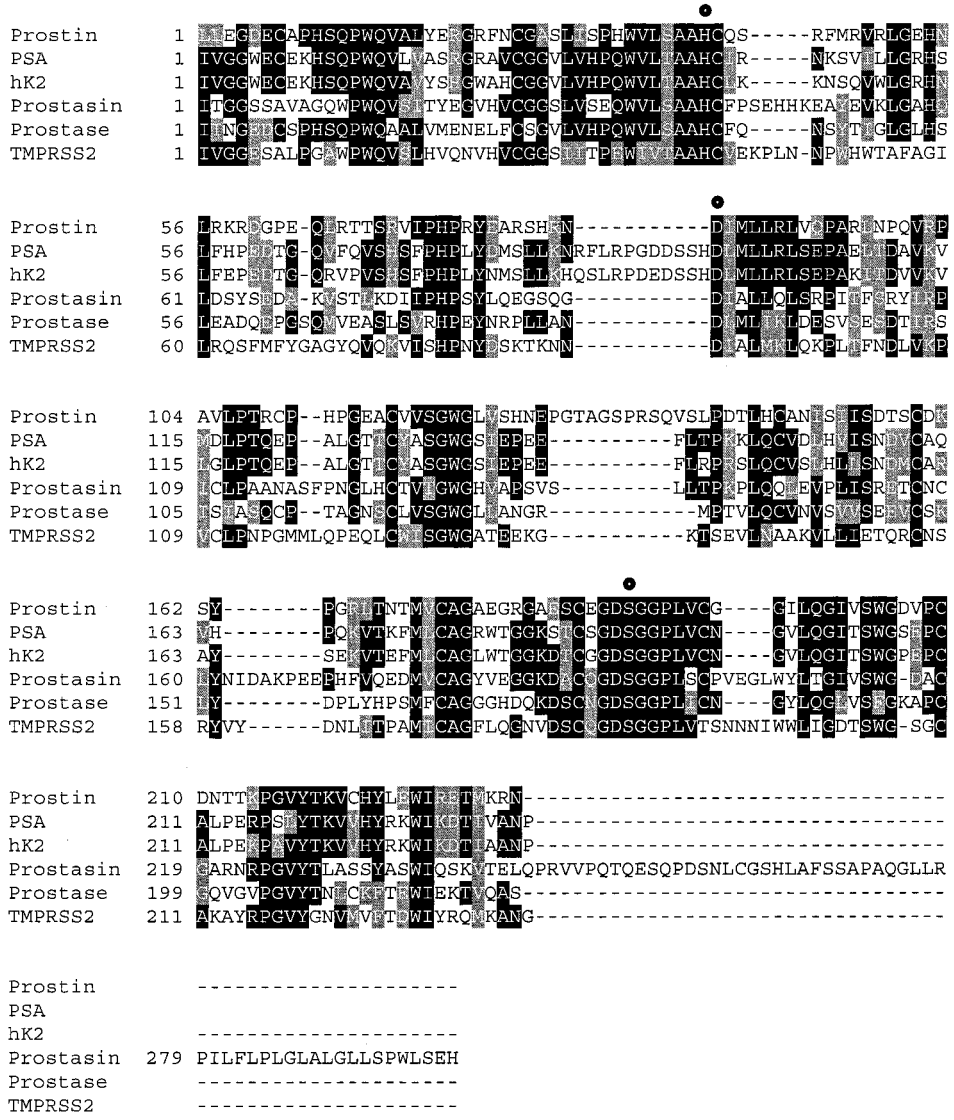

FIGURE 2: Comparison of prostin with other serine proteases identified in the prostate by amino acid sequence alignment. The identical sequences are shaded black; the conserved sequences are shaded gray. Dashes represent gaps to bring the sequences into better alignment. The catalytic triad, His, Asp, Ser, is circled.

another trypsin-type serine protease. A potential *N*-glycosylation site occurs at Asn^{216} followed by an *N*-glycosylation signal sequence of Asn−Thr−Thr−Lys.

Northern blots with multiple tissue Northern membranes (CLONTECH) showed that prostinogen is produced primarily in the human testis (Figure 3). A low level transcript of slightly larger size was also identified in the colon (Figure 3). Results of the multiple tissue (MTC panel I and II) PCR (CLONTECH) demonstrated that there was also expression of prostinogen by the prostate, kidney, and pancreas (Figure 4).

Expression and Refolding of Prostinogen. E. coli BL21-(DE3) was transformed with the expression plasmid, pET12-prostinogen. The cells were induced by IPTG and harvested, and inclusion bodies were prepared as described previously (*19*). The inclusion bodies were solubilized, and prostinogen was refolded as described in Experimental Procedures. As in the refolding of pro-PSA (*19*), a significant proportion of the soluble prostinogen was incorrectly folded. This was illustrated by the multiple bands in the unreduced sample (Figure 5, lane 1). When the samples were reduced, the predominant band of 32 kDa was seen for both prostinogen and pro-PSA (Figure 5, lanes 3 and 4). These results showed that the multiple bands seen in the nonreduced sample (Figure 5, lane 1) were essentially all incorrectly refolded proteins. Under the above conditions, 3 mg of soluble prostinogen was routinely obtained from 20 mg of inclusion bodies. On the basis of the intensity of the 28-kDa band of prostinogen (Figure 5, lanes 1) as compared to the 28-kDa band of purified pro-PSA (Figure 5, lane 2), approximately 33 μg of correctly folded prostinogen was obtained with an overall yield of ~0.2%. Pro-PSA was refolded with an overall yield of 24% (*19*). The homogeneous preparation of correctly refolded prostinogen will require improvement of refolding conditions.

Activation of Prostinogen with Trypsin-Agarose. We utilized trypsin-agarose for the prostinogen activation experiments, in which prostinogen was incubated with trypsin-agarose, the trypsin-agarose was removed by centrifugation, and the generated prostin activity was measured using S-2222. Extensive wash of trypsin-agarose was required to remove any leached trypsin activity, because trypsin also cleaves S-2222. The crude refolded prostinogen was incubated with thoroughly washed trypsin-agarose at an enzyme-

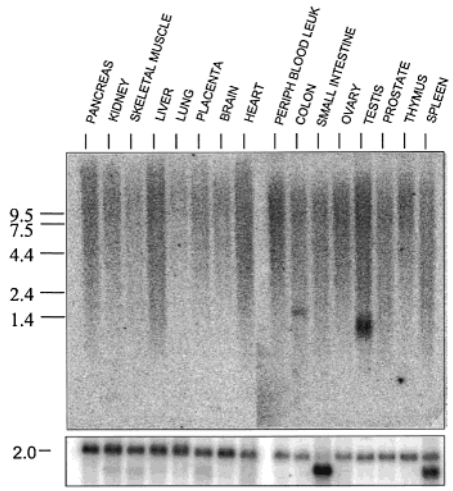

FIGURE 3: Multiple tissue Northern analysis of prostinogen. Prostinogen is expressed primarily in the testis as an approximately 1.3 kb transcript. A faint 1.5-kb band is seen in the colon. The human glyceraldehyde-3-phosphate dehydrogenase control is shown on the lower panel.

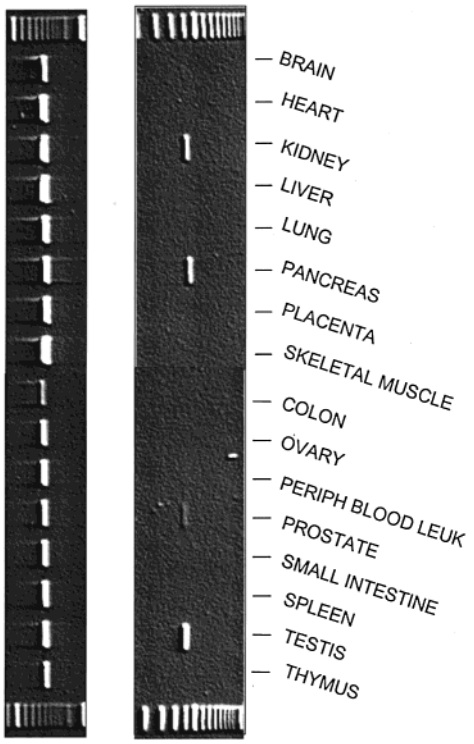

FIGURE 4: Multiple tissue cDNA PCR analysis of prostinogen. The expected 476-bp band is seen in the kidney, pancreas, prostate, and testis. The human glyceraldehyde-3-phosphate dehydrogenase control is shown on the lower panel.

to-substrate weight ratio of 1:7600, assuming the specific activity of trypsin-agarose is equal to free trypsin. The concentration of the refolded prostinogen was estimated to be 22 μg/mg of crude prostinogen from the intensity of the 28-kDa band. After the incubation and removal of trypsin-agarose, the prostin activity was measured using S-2222. As shown in Figure 6, the enzyme activity of prostin gradually increased over a 1-h incubation, whereas the incubation of trypsin-agarose or prostinogen alone showed no activity. At the 1 h incubation, the calculated amidolytic activity of prostin was 2.1 μmol of *p*NA (mg of prostin)$^{-1}$ min^{-1} toward

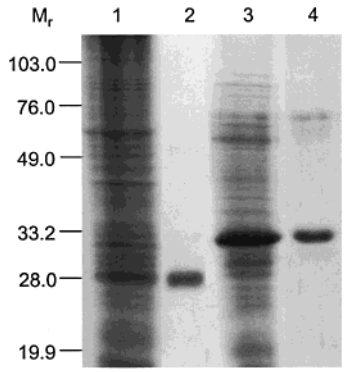

FIGURE 5: SDS−PAGE analysis of recombinant prostinogen and recombinant pro-PSA. Lane 1, prostinogen (unpurified, 15 μg); lane 2, pro-PSA (2 μg); lanes 3 and 4, reduced prostinogen and pro-PSA, respectively.

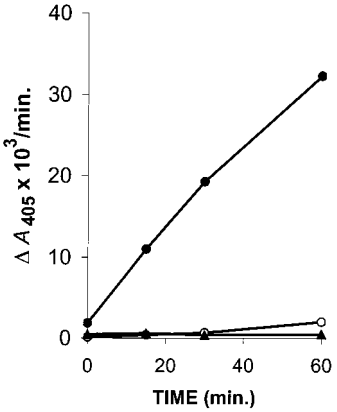

FIGURE 6: Time course of prostinogen activation with trypsin-agarose. Unpurified prostinogen (190 μg) was activated with trypsin-agarose (5 μL) in 110 μL total volume. The reaction was terminated at various times by centrifugation of 25 μL aliquots. Enzyme activity of 15 μL of the samples was measured using the substrate S-2222 (0.5 mM final). Activated prostin (black circles) showed substantial amidolytic activity, while prostinogen (triangles) and the control (sample without prostinogen; open circles) did not show any activity.

S-2222 (0.5 mM final). Prostin was completely inhibited by aprotinin.

Activation of pro-PSA by Prostin. Although hK2 was previously shown to activate pro-PSA, other serine proteases in seminal plasma are also potential physiologic activators of pro-PSA (*19*). It was previously demonstrated that another yet unidentified partially purified serine protease could also activate pro-PSA (*19*). In the previous study, the degree of activation of pro-PSA by hK2 was shown to be significantly lower than that by trypsin (*19*). To test for serine proteases that might have higher rates of pro-PSA activation, we therefore examined pro-PSA activation by prostin. Prostin was generated from prostinogen as described above. Pro-PSA was incubated with prostin or prostinogen (supernatant sample from time zero activation with trypsin-agarose above) at an approximate molar ratio of 7:1 (based on SDS−PAGE estimate as above), and the reaction was stopped with aprotinin as describe above. Aliquots were analyzed for PSA activity employing S-2586 as substrate (Figure 7). Prostin at this concentration had no activity toward S-2586. A rapid activation of the PSA zymogen by prostin was observed within 10 min, whereas no activity was observed with the reactions containing prostinogen. The activation level achieved by prostin was equal to the maximum (100%) activation of

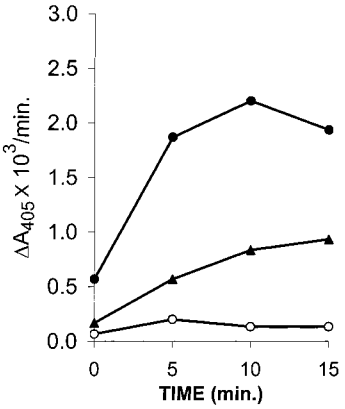

FIGURE 7: Time course of pro-PSA activation by prostin. Pro-PSA (5 μg) was activated with unpurified prostin (34 μg, black circles), unpurified prostinogen (34 μg; open circles), or hK2 (0.75 μg; triangles), and the reaction was terminated with aprotinin (0.1 μg). The estimated concentrations of prostin and prostinogen were equivalent to the level of hK2. Enzyme activity was measured using S-2586 (0.5 mM final).

pro-PSA by trypsin (*19*). The level of activation by hK2 was significantly lower than that of prostin. Therefore, prostin is the most efficient activator of pro-PSA identified to date.

Measurements of Amidolytic Activity Against other Chromogenic Substrates. Prostin was incubated with S-2586, S-2251, and S-2302, and the increase in absorbance (405 nm) was measured according to Experimental Procedures. On the basis of the estimated level of pure prostin in the crude preparation (see above), the calculated specific activity was 0.12 μmoles of *p*NA (mg of prostin)$^{-1}$ min^{-1} against S-2302 (0.5 mM final). Prostin had much higher activity toward S-2222 [2.10 μmol of *p*NA (mg of prostin)$^{-1}$ min^{-1} at 0.5 mM final]. Prostin showed low level activity against S-2251 and had no amidolytic activity against S-2586. Thus, prostin had significant activity against both S-2222 and S-2302, which have an Arg-*p*NA cleavage site. These data and the presence of the acidic residue Glu^{187} indicate clearly that prostin is a trypsin-type protease.

DISCUSSION

In our search for potential activator(s) of the PSA cascade, we identified a novel serine protease, prostinogen, from a human prostate cDNA library. It contains the catalytic triad, His, Asp, and Ser, of serine proteases in the conserved positions, and it has a substantial amino acid sequence identity with PSA, hK2, prostase, TMPRSS2, and prostasin. Recombinant prostinogen was produced in *E. coli* and refolded. The activation of prostinogen with trypsin-agarose generated the active form, prostin. Prostin exhibited amidolytic activity toward S-2222. Furthermore, prostin activated pro-PSA, much better than hK2, which was previously shown to activate pro-PSA (*19*). These findings indicate that prostin may be a primary physiologic activator of pro-PSA.

Serine proteases are involved in numerous biological processes. Some enzymes, such as trypsin and chymotrypsin, rapidly digest a wide variety of proteins. Others, mostly trypsin-type proteases, activate only a restricted number of pro-enzymes through limited proteolysis. For example, in the blood coagulation system, the clotting factors activate specific downstream pro-enzymes in a "cascade-like" manner, such that a minute amount of upstream activation can result in an amplification that leads to rapid clot formation (*29*). In the prostate, a similar cascade-like process may well be operational (*19*). During fertilization, a seminal clot must form almost instantaneously, yet it must also dissolve within minutes to allow the release of motile sperm. Every zymogen of prostate-associated serine proteases discovered to date, including prostin, hK2, prostasin, prostase, and TMPRSS2, requires limited proteolysis for its activation. Because these proteases are trypsin-type in specificity, they are all potential activators of PSA and of each other [except prostase which has an unusual cleavage site, Gln^4−Ile^5 (*21*) for activation]. Of these proteases, only hK2 and prostin have been tested so far for their ability to activate pro-PSA. Prostin has stronger activity toward pro-PSA than does hK2; therefore, it is a potential primary physiologic activator of pro-PSA. Furthermore, because pro-hK2 has a very similar pro-piece to that of pro-PSA, prostin may also activate pro-hK2. The interactions between these serine proteases are greatly clarified by the production and characterization of recombinant proteins.

The likely role of serine proteases in prostatic diseases is perhaps somewhat subtle and more complex. This is because there are no known protein deficiencies that result in particular diseases, analogous to hemophilias that result from defects in the blood coagulation cascade. It is therefore difficult to accurately attribute a specific pathological function to a given enzyme. Nonetheless, PSA has been shown to cleave IGFBP3 (*12*) and PTHrp (*12*), which may lead to abnormal growth of prostate cells. The ultimate consequence may be the development of cancer or benign prostatic enlargement. Perhaps the strongest evidence thus far for the role of serine proteases in prostate cancer is the fact that the cancer cells overexpress urokinase-type plasminogen activator (uPA) (*37*, *38*). uPA activates plasminogen to generate plasmin (*39*). Plasmin, in turn activates metalloproteases, including procollagenase (*40*, *41*). Prostate cancer cells thus may overproduce proteolytic enzymes that aid their invasion through adjacent connective tissue. We have shown that recombinant hK2 can activate the precursor form of uPA to urokinase (*19*). This reaction points to the possibility that the interaction between hK2 and uPA may be involved in malignant processes, and as a result, hK2 may become a future target for anticancer drugs. hK2 is produced only by the prostate, while uPA is ubiquitous. Future studies will help determine if overexpression of specific serine proteases is involved in the development and progression of certain cancers.

There are several published procedures for the identification of novel serine proteases. Blood clotting factors have been identified by studies of patients having bleeding disorders caused by deficiencies of specific proteins. More recently, molecular biology techniques have been used to identify members of protein families that are not clearly associated with a particular pathological state. For example, because serine proteases have very distinctive conserved stretches of amino acid sequences, similar genes have been identified in cDNA databases. Other approaches include DNA subtraction, exon trapping, and degenerate oligonucleotide PCR amplification. In this study, we were able to improve the conventional degenerate oligonucleotide technique by applying the method of Rose et al. in designing more efficient primers based on the conserved regions of

serine proteases (*30*). As a result, a total of 15 different serine proteases were identified in a human prostate cDNA library (CLONTECH). It is likely that these PCR primers could be used to identify additional serine proteases in other tissue (normal or pathological) cDNA libraries as well.[3] Furthermore, by designing new primers based on other serine proteases, e.g., chymotrypsin-type proteases, we may discover other novel proteases having completely different physiologic functions.

Of the 15 different prostate serine proteases identified here, we found one novel protease that can activate pro-PSA. The zymogen of this protease, named prostinogen, was expressed in *E. coli* and refolded. The refolding procedure was, however, challenging on several fronts. First, solubilization of the inclusion bodies required a much lower pH (3.8) than that used for recombinant pro-PSA (8.5). When testing the pH condition of the refolding buffer, we found that a pH of 8.8 or higher was required for keeping prostinogen in solution. During the shift of pH from 3.8 to 8.8, the overall yield decreased, probably due to formation of protein aggregates in the refolding buffer. Another reason for the lower yield may have been caused by two additional Cys residues in prostinogen, which could form improper disulfide bonds and thus result in protein aggregates. SDS−PAGE analysis of refolded pro-PSA prior to purification had fewer high molecular weight bands than did prostinogen (data not shown), indicating that pro-PSA has less tendency to form protein aggregates. Finally, several other conditions were tested to optimize refolding but failed to produce any improvement. These procedures included addition of polyethylene glycol or arginine, or altering salt concentrations and the levels of oxidized/reduced glutathione. In the future, the refolding protocol may be improved substantially by testing additional conditions and utilizing more sensitive methods to quantitate properly refolded protein, such as by enzyme assays described above.

Northern blot analysis and multiple tissue PCR amplification demonstrated that prostinogen is expressed at high levels in the human testis and at somewhat lower levels in the prostate, kidney, and pancreas. These findings may be a result of the biological characteristics that are salient to prostinogen. For example, its production by the testis suggests possible androgen regulation of its expression. Its expression by the kidney may indicate a common embryological origin, as the fetal development of the kidneys, ureters, bladder, and prostate are closely associated. Pancreatic expression, on the other hand, suggests that prostin may also have a role in digestion. In pathological states such as cancer metastasis, prostin may participate in the degradation of tissue.

Future studies of prostinogen will better elucidate the potential biological role of this protein in prostatic diseases. The development of antibodies directed against peptide epitopes or purified recombinant protein will improve the estimation of its expression levels in prostate cancer as well as in benign prostatic hyperplasia and prostatitis. Improved refolding and purification procedures for recombinant prostinogen may facilitate the discovery of other physiologic substrates and lead to the identification of its own physiological activator(s). The ultimate goal of these studies is the development of better tools for the diagnosis and management of prostatic diseases.

[3] Takayama et al., unpublished data.

ACKNOWLEDGMENT

We are grateful to Drs. Earl W. Davie, Kazuo Fujikawa, and Dominic W. Chung for their support and critical review of the manuscript. We thank Drs. Paul H. Lange and Robert Vessella for their support and encouragement. We also thank Dr. Kathleen P. Pratt for her helpful suggestions. We are grateful to Emily Chen and Sharie Chen for their technical assistance.

REFERENCES

1. Hara, M., Inoue, T., Koyanagi, Y., Fukuyama, T., and Iki, H. (1969) *Nippon Hoigaku Zasshi 23***,** 333.
2. Li, T. S., and Beling, C. G. (1973) *Fertil. Steril. 24*, 134−44.
3. Papsidero, L. D., Kuriyama, M., Wang, M. C., Horoszewicz, J., Leong, S. S., Valenzuela, L., Murphy, G. P., and Chu, T. M. (1981) *J. Natl. Cancer Inst. 66*, 37−42.
4. Sensabaugh, G. F. (1978) *J. Forensic Sci. 23*, 106−15.
5. Wang, M. C., Valenzuela, L. A., Murphy, G. P., and Chu, T. M. (1979) *Invest. Urol. 17*, 159−63.
6. Wang, M. C., Papsidero, L. D., Kuriyama, M., Valenzuela, L. A., Murphy, G. P., and Chu, T. M. (1981) *Prostate 2*, 89−96.
7. Stamey, T. A., Yang, N., Hay, A. R., McNeal, J. E., Freiha, F. S., and Redwine, E. (1987) *N. Engl. J. Med. 317*, 909−16.
8. Oesterling, J. E., Chan, D. W., Epstein, J. I., Kimball, A. W. J., Bruzek, D. J., Rock, R. C., Brendler, C. B., and Walsh, P. C. (1988) *J. Urol. 139*, 766−72.
9. Hudson, M. A., Bahnson, R. R., and Catalona, W. J. (1989) *J. Urol. 142*, 1011−7.
10. Lange, P. H., Ercole, C. J., Lightner, D. J., Fraley, E. E., and Vessella, R. (1989) *J. Urol. 141*, 873−9.
11. Lilja, H. (1985) *J. Clin. Invest. 76*, 1899−903.
12. Cohen, P., Graves, H. C., Peehl, D. M., Kamarei, M., Giudice, L. C., and Rosenfeld, R. G. (1992) *J. Clin. Endocrinol. Metab. 75*, 1046−53.
13. Cramer, S., Chen, Z., and Peehl, D. (1996) *J. Urol. 156*, 526−531.
14. Iwamura, M., Hellman, J., Cockett, A., Lilja, H., and Gershagen, S. (1996) *Urology 48*, 317−325.
15. Lundwall, A., and Lilja, H. (1987) *FEBS Lett. 214*, 317−22.
16. Riegman, P. H., Vlietstra, R. J., Klaassen, P., van der Korput, J. A., van Kessel Geurts, A., Romijn, J. C., and Trapman, J. (1989) *FEBS Lett. 247*, 123−6.
17. Kumar A., Mikelajczyk, S. D., Goel, A. S., Millar, L. S., Saedi, M. S. (1997) *Cancer Res. 57*, 3111−4.
18. Lovgren, J., Rajakoski, K., Karp, M., Lundwall, A., and Lilja, H. (1997) *Biochem. Biophys. Res. Commun. 238*, 549−55.
19. Takayama, T., Fujikawa, K., and Davie, E. (1997) *J Biol Chem 272***,** 21582−21588.
20. Yu, J., Chao, L., and Chao, J. (1994) *J. Biol. Chem. 269*, 18843−18848.
21. Nelson, P., Gan, L., Ferguson, C., Moss, P., Gelinas, R., Hood, L., and Wang, K. (1999) *Proc. Natl. Acad. Sci. 96*, 3114−3119.
22. Stephenson, S., Verity, K., Ashworth, L., and Clements, J. (1999) *J. Biol. Chem. 274*, 23210−23214.
23. Yousef, G., Obiezu, C., Luo, L., Black, M., and Diamandis, E. (1999) *Cancer Res. 59*, 4252−4256.
24. Simmer, J., Fukae, M., Tanabe, T., Yamakoshi, Y., Uchida, T., Xue, J., Margolis, H., Shimizu, M., DeHart, B., and Bartlett, J. (1998) *J. Dent. Res. 77*, 377−86.
25. Paoloni-Giacobino, A., Chen, H., Peitsch, M., Rossier, C., and Antonarakis, S. (1997) *Genomics 44*, 309−320.
26. Lin, B., Ferguson, C., White, J., Wang, S., Vessella, R., True, L., Hood, L., and Nelson, P. (1999) *Cancer Res. 59*, 4180−4184.
27. Akiyama, K., Nakamura, T., Iwanaga, S., and Hara, M. (1987) *FEBS Lett. 225*, 168−72.

28. Takayama, T., Fujikawa, K., and Davie, E. (1997) *J. Biol. Chem. 272*, 21582−21588.
29. Davie, E. W., Fujikawa, K., and Kisiel, W. (1991) *Biochemistry 30*, 10363−70.
30. Rose, T., Schultz, E., Henikoff, J., Pietrokovski, S., McCallum, C., and Henikoff, S. (1998) *Nucleic Acids Res. 26*, 1628−1635.
31. Liu, X. L., Wazer, D. E., Watanabe, K., and Band, V. (1996) *Cancer Res. 56*, 3371−3379.
32. Barrett, A. J., Rawlings, N. D., and Woessner, J. F., Eds. (1998) *Handbook of Proteolytic Enzymes*, Academic Press, San Diego, CA.
33. Frohman, M., Dush, M. K., and Martin, G. R. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8998−9002.
34. Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Mattick, J. S. (1991) *Nucleic Acids Res. 19*, 4008.
35. Hooper, J., Nicol, D., Dickinson, J., Eyre, H., Scarman, A., Normyle, J., Stuttgen, M., Douglas, M., Loveland, K., Sutherland, G., and Antalis, T. (1999) *Cancer Res. 59*, 3199−205.
36. Dihanich, M., and Spiess, M. (1994) *Biochim. Biophys. Acta 1218*, 225−8.
37. Achbarou, A., Kaiser, S., Tremblay, G., SteMarie, L. G., Brodt, P., Goltzman, D., and Rabbani, S. A. (1994) *Cancer Res. 54*, 2372−7.
38. VanVeldhuizen, P., Sadasivan, R., Cherian, R., and Wyatt, A. (1996) *Am. J. Med. Sci. 312*, 8−11.
39. Lijnen, H. R., Zamarron, C., Blaber, M., Winkler, M. E., and Collen, D. (1986) *J. Biol. Chem. 261*, 1253−8.
40. Stricklin, G., Bauer, E., Jeffrey, J., and Eisen, A. (1977) *Biochem. 16*, 1607−1615.
41. He, C., Wilhelm, S., Pentland, A., Marmer, B., Grant, G., Eisen, A., and Goldberg, G. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2632−2636.

BI002129R