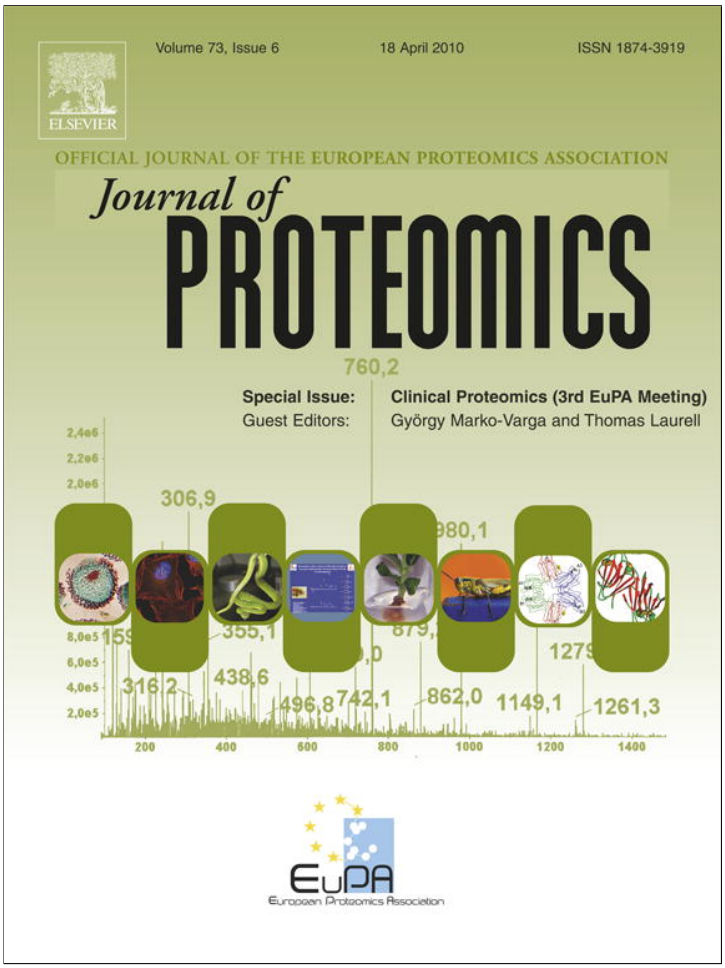
Identification of prostate specific antigen (PSA) isoforms in complex biological samples utilizing complementary platforms

Végvári, Ákos; Rezeli, Melinda; Welinder, Charlotte; Malm, Johan; Lilja, Hans; Marko-Varga, György; Laurell, Thomas

Published in:
Journal of Proteomics

DOI:
10.1016/j.jprot.2010.01.008

2010



JOURNAL OF PROTEOMICS 73 (2010) 1137–1147

available at www.sciencedirect.com

www.elsevier.com/locate/jprot

Identification of prostate-specific antigen (PSA) isoforms in complex biological samples utilizing complementary platforms

Ákos Végvári[a,*], Melinda Rezeli[a], Charlotte Welinder[b], Johan Malm[c], Hans Lilja[c,d,e,f], György Marko-Varga[a], Thomas Laurell[a]

[a]Div. Clinical Protein Science & Imaging, Dept. of Measurement Technology and Industrial Electrical Engineering, Lund University, BMC C13, SE-221 84 Lund, Sweden
[b]Dept. of Oncology, Clinical Sciences, Lund University, Barngatan 2B, SE-221 85 Lund, Sweden
[c]Dept. of Laboratory Medicine, Lund University, Malmö University Hospital, SE-205 02 Malmö, Sweden
[d]Dept. of Clinical Laboratories, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA
[e]Dept. of Surgery (Urology), Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA
[f]Dept. of Medicine (GU-Oncology), Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA

ARTICLE INFO

Article history:
Received 21 October 2009
Accepted 16 January 2010

Keywords:
Prostate-specific antigen isoforms
Seminal plasma
Mass spectrometry
MALDI LTQ Orbitrap XL

ABSTRACT

Measurements of the prostate-specific antigen (PSA) levels in blood are widely used as diagnostic, predictive and prognostic marker of prostate disease. The selective detection of molecular forms of PSA can contribute clinically to meaningful enhancements of the conventional PSA-test. As it is plausible that an in-depth search for structural variants of PSA gene products may increase our ability to discriminate distinct patho-biological basis and stages of prostate diseases, we have developed a multi-step protocol comprising gel-based methods followed by mass spectrometric identification.

Our current aim was to provide a comprehensive identification of PSA variants occurring in seminal fluid. We provide a proof-of-principle for this multiple step analytical approach to identify multiple PSA variants from complex biological samples that revealed distinct molecular characteristics. In addition, sequence-annotated protein bands in SDS–PAGE gels were compared to those detected by Western blots, and by monitoring the enzymatic activity in zymogram gels, using gelatin as a substrate. The high accuracy annotations were obtained by fast turnaround MALDI-Orbitrap analysis from excised and digested gel bands. Multiple PSA forms were identified utilizing a combination of MASCOT and SEQUEST search engines.

1. Introduction

Prostate-specific antigen (PSA) is a member of the kallikrein family of serine proteinases that is abundantly produced and released by epithelial cells of the prostate [1–3]. PSA expression remains exceedingly high also in advanced stages of prostate adenocarcinomas, and the amount of PSA released into blood can elevate up to 10^4-fold beyond normal levels in the blood (which is about six orders of magnitude lower than that in seminal fluid) due to increased tumor burden combined with histological changes in prostate architecture that accompany the malignant transformation [4,5].

PSA has restricted chymotrypsin-like endoproteolytic activity, which beyond the biological substrates, semenogelin 1 and 2 (SEMG1 and SEMG2) [1,6], also contributes cleavages of fibronectin [6], laminin and gelatin [7]. PSA is secreted as an

* *Corresponding author*. Tel.: +46 46 222 3721; fax: +46 46 222 4527.
E-mail address: akos.vegvari@elmat.lth.se (Á. Végvári).

1874-3919/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jprot.2010.01.008

inactive proenzyme (pro-PSA) into seminal fluid and can (at least *in vitro*) be activated by the kallikrein-related peptidase 2 (hK2) and other endopeptidases produced by the prostate [3,8–10]. PSA predominantly occurs in a free and active single-chain form, or as inactive (due to internal cleavages) in seminal fluid [1,2,10,11]. A minor proportion (≤5%) of the PSA in the seminal fluid is inactive, due to formation of covalent complexes with SERPINA5 (protein C inhibitor), which is abundantly released from the seminal vesicles [12–14]. In contrast, PSA in blood is non-catalytic and the predominant proportion is covalently linked to SERPINA3 (α-1-antichymotrypsin) [15,16].

Furthermore, several studies have implicated that considerable structural heterogeneity of the free forms of PSA occurs in serum, seminal plasma, and hyperplastic-, or cancerous-tissue [2,5,11]. It is suggested that PSA isoforms may appear upon the translation of alternative hKLK3 transcripts [17]. In contrast to the large number of KLK transcript variants recently identified, very few of them have been confirmed at the protein level (Table 1 in Ref. [18]). Previous studies presented nicked forms of PSA that were associated with prostatic diseases, linking the Lys145/Lys182 forms as over-represented in benign prostatic hyperplasia (BPH), while the nicked pro-PSA forms to be more abundant in cancer [19–21].

Table 1 – The concentration of free PSA determined by DELFIA assay and the molecular weight of protein bands observed by gel electrophoretic measurements.

Sample nr.	fPSA (mg/mL)	Mw of bands detected by		
		MS (SDS–PAGE)	Western blot	Zymogram
1	0.800	33.98		34.14
		32.78		
		31.23		31.17
		30.35		
		25.66	25.12	24.75
		23.73		
		22.83		23.06
2	1.855			35.49
				33.99
		30.95		31.44
		30.08		
		28.94	25.04	
		27.63		
		24.75		
		23.54		23.95
				22.64
3	1.265			33.79
		31.87		
		29.88		
		27.74		
		24.24	25.09	
		23.23		23.77
				22.49
4	0.505	33.98		33.72
		32.78		
		31.23		31.46
			28.95	
			27.10	
			24.92	
				23.98
				22.54

As it is plausible that an in-depth search for structural variants of PSA gene products may increase our ability to discriminate distinct patho-biological basis and stages of prostate disease, we have therefore developed a multi-step protocol comprising gel-based methods followed by mass spectrometric identification. We hereby present a proteomic platform, including gel electrophoretic separations and mass spectrometric analyses in order to verify and outline the multitude forms and enzymatic function of PSA. Our goal was to demonstrate the expression of PSA forms with annotated sequences and quantities in seminal plasma samples from a healthy volunteer and from young men undergoing investigation for infertility. We present data that identify various molecular forms of PSA.

2. Material and methods

2.1. *Seminal plasma*

Seminal plasma was obtained from a healthy volunteer and from men undergoing investigation for infertility prior to final diagnosis of any dysfunction. The samples were provided by the Center for Reproductive Medicine at Malmö University Hospital, following the guidelines of the Helsinki Declaration. The collection of seminal plasma was approved by the ethical board at Lund University (approval number: LU 532-03), and the samples were processed according to the WHO guidelines (WHO, 1999). Seminal plasma was obtained by centrifugation at 10000 *g* for 10 min. Free PSA (fPSA) ranged between 0.505 and 1.855 mg/mL (see Table 1) as determined by the DELFIA assay [22] (Perkin Elmer, Turku, Finland). Samples were stored at −20 °C until use.

2.2. *Gel electrophoretic techniques*

PSA expression was detected by SDS–PAGE gel electrophoresis in 0.75 mm thick, 12%T, 2.67%C polyacrylamide gels under reducing conditions. The samples were reduced with DTT (at a final concentration of 4 mM) while boiled for 5 min and alkylated with addition of 2 μL of 0.5 M iodoacetamide at room temperature. Following preparation, 10 μL of samples and 3 μL of protein molecular weight standards (PageRuler™ Prestained Protein Ladder Plus from Fermentas, St. Leon-Rot, Germany) were applied. The gels were run at a constant voltage of 125 V for 1 h 10 min in a MiniProtean III electrophoresis unit (Bio-Rad, Hercules, CA). Staining was performed with Coomassie blue R350 (GE Healthcare, Uppsala, Sweden) for 1 h and destaining according to the manufacturer's recommendations.

The Western blot analysis was performed using 0.75 mm thick, 12%T, 2.67%C polyacrylamide gels run at non-reducing conditions. The separated protein bands were then transferred onto PVDF membranes (0.45 μm pore size Invitrolon™ PVDF, Invitrogen, Carlsbad, CA) using a tank blotting (Bio-Rad, Hercules, CA) at constant current of 350 mA for 1.5 h. The membranes were blocked with 1% of BSA in 50 mM Tris–HCl/150 mM NaCl/0.1% Tween-20, pH 7.4. The primary antibody was 2E9 diluted to 1:5000 [23], whereas the secondary antibody (ECL anti-mouse IgG), used in dilution 1:10000, and the

horseradish peroxidase reagent system (ECL) were from GE Healthcare (Uppsala, Sweden).

Enzymatic activity of PSA forms was monitored by zymogram gel electrophoresis using non-boiled sample aliquots in 0.75 mm thick, 10%T, 2.67%C polyacrylamide gels containing 0.1% gelatin (from porcine skin; Sigma, Steinheim, Germany) under non-reducing conditions in the presence of SDS. In order to the proteins refold and their regain enzymatic activity, SDS was removed by washing it off with Triton X-100 and then the zymographic gels were incubated overnight at 37 °C in Zymogram Developing Buffer (Invitrogen, Carlsbad, CA). Staining was performed with Coomassie blue R350 for 1 h and de-stained according to the manufacturer's recommendations.

2.3. Quantification

Densitometric quantification of Coomassie blue stained protein bands were performed on diluted seminal plasma proteins separated in SDS–PAGE gels under reducing conditions. The molecular weights of proteins were determined by calculation from the protein standards using Quantity One software (Bio-Rad, Hercules, CA).

2.4. Mass spectrometry

Following in-gel digestion of the major protein bands, observed within the size region of 20–36 kDa on SDS–PAGE gels, samples were analyzed by MALDI-MS providing both peptide mass fingerprint (PMF) and sequencing data. The mass spectrometric analysis was performed on a MALDI LTQ Orbitrap XL instrument (Thermo Scientific, Waltham, MA). Mass spectra were obtained in positive mode within a mass range of 600–4000 Da, using 60,000 resolution (determined at 400 *m/z*). Twenty full mass scans (2 microscans/scan) were collected from each spot. The laser, operated at 10 μJ, was moved automatically after 2 microscans to avoid overburning of the sample. MS/MS data collection was performed on the top 25 peaks that matched with the list of masses derived by theoretical (*in silico*) digestion of 35 PSA variants by PeptideMass (available at the ExPASy Proteomics Server site, http://expasy.org/sprot/), choosing iodoacetamide as alkylation agent, oxidation on methionine and allowing 1 miss-cleavage. The minimal signal required for MS/MS data acquisition was 500 counts. Normalized collision energy was 50% during an activation time of 30 ms and activation Q of 0.250 was applied. All experiments were performed in triplicates.

Spectra were processed by Xcalibur software v2.0.7 (Thermo Scientific, Waltham, MA), averaging 20 full mass scans followed by quantitative analysis summing the isotope intensities by using the Xtract function in Xcalibur. The peptide mass fingerprint search was performed on the MASCOT server (http://www.matrixscience.com), using the NCBInr database, narrowed down to *Homo sapiens* (release 20091024, 224815 *H. sapiens* sequences out of 9,937,670 sequences). Side chain modifications were set to fixed carbamidomethylation at cystein and variable oxidation at methionine residues, respectively. Peptide tolerance was kept at 10 ppm. The samples were considered positive PSA identities, if at least 2 specific peptides were identified in at least two of the triplicates.

The MS/MS search was also performed on the NCBInr database, using the MASCOT search engine and allowing 1 missed cleavage. Peptide tolerance was 10 ppm, whereas MS/MS tolerance was chosen to 0.4 Da. The MS/MS search was repeated with the SEQUEST search engine, using the UniProt_human database (224,015 sequences), allowing 2 miss-cleavages. The peptide and fragment ion tolerances were set to 2 and 1 amu, respectively. The following filters were used on the search results: Delta CN≥1; R_{sp}≤4; X_{corr} vs charge state≥1.5 and protein probability≤0.5. PSA identification was accepted if it contained at least one identified peptide.

3. Results

3.1. Concept of the multiple analytical approach

We have developed a methodological strategy for successful identification of PSA forms in seminal plasma, which includes various gel-based separation methods combined with a mass spectrometric read-out. The protein expression was monitored by reducing SDS–PAGE and Western blot analyses, whereas the enzymatic activity was detected by zymographic gel electrophoresis (Fig. 1). Following densitometric quantification from gels, relevant protein bands were excised and trypsinated prior to mass spectrometric analysis. The identification of PSA from gel bands was performed by PMF (inclusion criterion: at least two PSA specific peptides) and also by MALDI-MS/MS analysis, sequencing the tryptic peptides of PSA (Fig. 1). This toolbox concept was applied and exemplified using individual samples collected from men undergoing investigation for infertility and a healthy volunteer.

3.2. Gel electrophoretic read-outs

The Western blot was employed as a first analysis step for an early verification of PSA in native polyacrylamide gels. The 2E9 monoclonal antibody, used for detection, is known to recognize PSA as both free and as complexed with α-1-antichymotrypsin [23]. The single band observed within the range of 23–28 kDa of Samples 1–3 (see Fig. 2 and Table 1) agreed well with the previous finding that PSA predominantly exists in free form in seminal plasma. However, in case of Sample 4, three PSA variants were clearly detected.

By immobilizing gelatin as the substrate in the zymographic assay, enzymatic activity was measured as the read-out. We were able to identify a number of enzymatically active protein bands. Interestingly, the zymogram gels revealed 3–5 activity bands within the Mw-range 23–35 kDa, of which some migrated at positions corresponding to that Coomassie-stained bands in the SDS–PAGE that were identified by the specific monoclonal antibody (2E9) against PSA in Western blots (see Fig. 2 and Table 1).

In the SDS–PAGE experiments, we were able to identify 10–14 gel bands, respectively, Coomassie-stained within the 20–35 kDa region (see Fig. 2).

The zymograms revealed strikingly similar patterns of 3–5 enzymatically active protein bands in these samples. Nine out of sixteen zymogram bands could be associated with corresponding PSA bands in SDS–PAGE gels at similar molecular

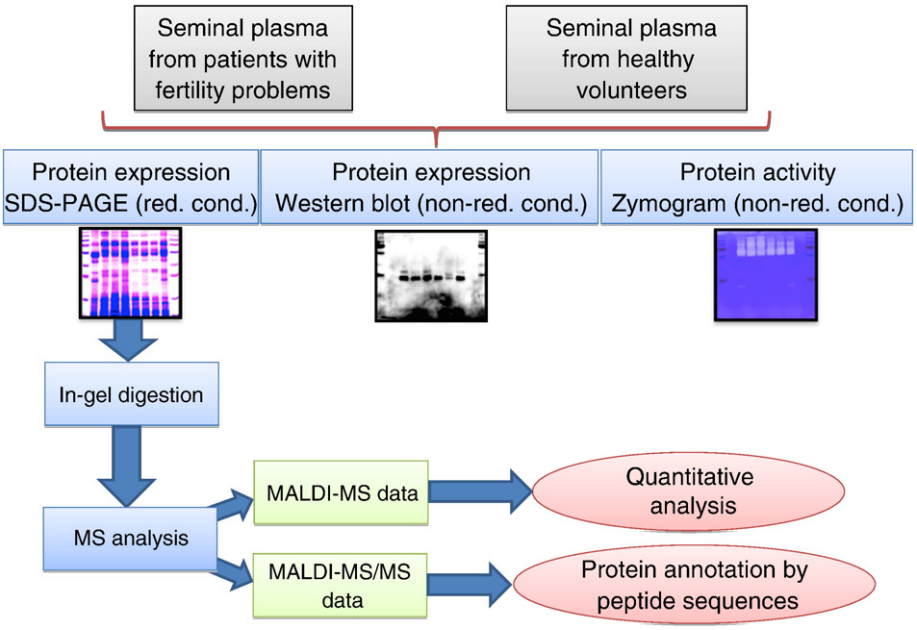


Fig. 1 – Depiction of the analytical strategy and the overall workflow employed to identify PSA forms in seminal plasma samples.

weights. By contrast, only one, enzymatically active PSA form at position 24.75 kDa in the zymogram of Sample 1 was confirmed by Western blot.

Based on the strict inclusion criteria defined in Section 2.4 for the MS read-out, 3–7 Coomassie-stained protein bands contained PSA in the SDS–PAGE experiments (Table 2). We observed a close to perfect match of positions (Mw) at 34, 33 and 31 kDa in Samples 1 and 4; at 30 and 23 kDa in Samples 2 and 3.

Only Coomassie-stained SDS–PAGE gel bands were subjected to quantification. The sum of the two most abundant peptide intensities in the bands, where PSA was clearly identified by MALDI-MS (see below), corresponded closely to the free PSA levels determined by a commercial PSA Standard Kit (time-resolved fluorescence immunoassay, Prostatus, Perkin Elmer, Turku, Finland).

3.3. *MALDI-Orbitrap MS analysis*

The bands at 20–36 kDa, containing reduced and alkylated proteins, were cut and subjected to proteolytic digestion with trypsin prior to MS analysis. In our mass spectrometric method we utilized the FT analyzer (*i.e.*, the Orbitrap) for full mass scans at high mass resolution (60,000) in order to produce high quality spectra (see Fig. 3). The fact that the Orbitrap provides high mass accuracy, peptide mass fingerprinting is an excellent means for protein identification, as a complement to MS/MS sequencing.

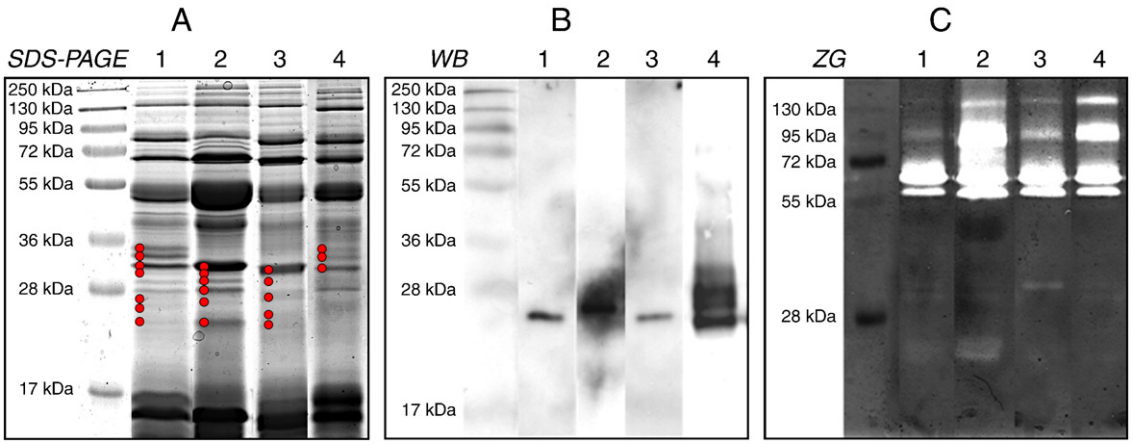


Fig. 2 – The summarized comparison of the gel bands observed in (A) SDS–PAGE, (B) Western blot (WB) and (C) zymogram (ZG) experiments. In the lanes of SDS–PAGE the bands with PSA (as identified by MS) are marked with red dots.

Table 2 – The annotated PSA sequence data from database search by both PMF and MS/MS using MASCOT and SEQUEST.

	PMF MASCOT	MS/MS MASCOT	MS/MS SEQUEST
Sample 1			
Band 5	673.3764	*VVHYR*[a]	
	757.4915	SVILLGR	
	854.3997		
	870.3945		
	1077.5023	IVGGWECEK	IVGGWECEK
	1272.6680		
	1407.7491	HSQPWQVLVASR	HSQPWQVLVASR
		FLRPGDDSSHDLMLLR	FLRPGDDSSHDLMLLR
	1887.9388	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
		LQCVDLHVISNDVCAQVHPQK	LQCVDLHVISNDVCAQVHPQK
			KLQCVDLHVISNDVCAQVHPQK
Band 6	673.3760	VVHYR	
	757.4913		
	854.3997	*FMLCARG*	
	870.3946		
	1077.5023	IVGGWECEK	IVGGWECEK
	1272.6676	LSEPAELTDAVK	LSEPAELTDAVK
	1407.7492	HSQPWQVLVASR	HSQPWQVLVASR
		FLRPGDDSSHDLMLLR	FLRPGDDSSHDLMLLR
	1887.9387	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
	2460.2116	LQCVDLHVISNDVCAQVHPQK	LQCVDLHVISNDVCAQVHPQK
		KLQCVDLHVISNDVCAQVHPQK	KLQCVDLHVISNDVCAQVHPQK
Band 7	673.3761	VVHYR	
	757.4915		
	854.3997	*FMLCARG*	
	870.3945		
	1077.5025	IVGGWECEK	IVGGWECEK
	1407.7499	HSQPWQVLVASR	HSQPWQVLVASR
		FLRPGDDSSHDLMLLR	FLRPGDDSSHDLMLLR
	1887.9393	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
	2460.2132		LQCVDLHVISNDVCAQVHPQK
			KLQCVDLHVISNDVCAQVHPQK
Band 8	673.3764		
	757.4919		
	870.3949		
	1077.5031	*IVGGWECEK*	*IVGGWECEK*
	1272.6686		
	1407.7503	HSQPWQVLVASR	*HSQPWQVLVASR*
	1887.9400	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
	2460.2123		
Band 11	757.4918		
	1077.5028	IVGGWECEK	IVGGWECEK
	1272.6685		LSEPAELTDAVK
	1407.7498	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9397	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
Band 12	757.4920		
	1077.5023	IVGGWECEK	IVGGWECEK
	1272.6687		LSEPAELTDAVK
	1407.7499	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9404	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
Band 13	757.4919		
	1077.5035		*IVGGWECEK*
	1272.6686		
	1407.7502	*HSQPWQVLVASR*	*HSQPWQVLVASR*
	1887.9397		
Sample 2			
Band 7	673.3768	VVHYR	
	757.4924	*SVILLCG*	
	854.4007	FMLCARG	FMLCARG
	870.3959		
	1077.5039	IVGGWECEK	IVGGWECEK

(continued on next page)

Table 2 (*continued*)

	PMF MASCOT	MS/MS MASCOT	MS/MS SEQUEST
Sample 2			
Band 7	1272.6697		
	1407.7518	HSQPWQVLVASR	HSQPWQVLVASR
	1871.9477	FLRPGDDSSHDLMLLR	FLRPGDDSSHDLMLLR
	1887.9430	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
	2344.2238	AVCGGVLVHPQWVLTAAHCIR	AVCGGVLVHPQWVLTAAHCIR
	2460.2184	LQCVDLHVISNDVCAQVHPQK	LQCVDLHVISNDVCAQVHPQK
	2588.3118	KLQCVDLHVISNDVCAQVHPQK	KLQCVDLHVISNDVCAQVHPQK
	3493.6930		
	3509.6916	*HSLFHPEDTGQVFQVSHSFPHPLYDMSLLK*	HSLFHPEDTGQVFQVSHSFPHPLYD**M**SLLK
Band 8	673.3771	VVHYR	
	757.4928		
	870.3961		
	1077.5040	IVGGWECEK	IVGGWECEK
	1272.6694		
	1407.7516	HSQPWQVLVASR	HSQPWQVLVASR
		FLRPGDDSSHDLMLLR	*FLRPGDDSSHDLMLLR*
	1887.9418	FLRPGDDSSHDL**M**LLR	*FLRPGDDSSHDLMLLR*
			AVCGGVLVHPQWVLTAAHCIR
		LQCVDLHVISNDVCAQVHPQK	*LQCVDLHVJSNDVCAQVHPQK*
		KLQCVDLHVISNDVCAQVHPQK	*KLQCVDLHVISNDVCAQVHPQK*
Band 9/10[b]	757.4920		
	870.3956		
	1077.5030		
	1272.6692		
	1407.7510		
	1887.9409	FLRPGDDSSHDL**M**LLR	
	2588.3101		
Band 12		IVGGWECEK	IVGGWECEK
	1272.6719		LSEPAELTDAVK
	1407.7533	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9444	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
Band 13	757.4927	*SVILLCG*	
	1077.5040	IVGGWECEK	IVGGWECEK
	1272.6694	LSEPAELTDAVK	LSEPAELTDAVK
	1407.7524	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9430	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
	2344.2264		
	3509.6918		
Sample 3			
Band 5[c]	673.3764	VVHYR	
	757.4920		
	854.4000		
	870.3951		
	1077.5033	IVGGWECEK	IVGGWECEK
	1272.6693		LSEPAELTDAVK
	1407.7507	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9403	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
	2460.2145		LQCVDLHVISNDVCAQVHPQK
	2588.3098	*KLQCVDLHVJSNDVCAQVHPQK*	KLQCVDLHVISNDVCAQVHPQK
Band 6[c]	673.3768	VVHYR	
	757.4923	*SVILLGR*	
	854.4005		
	870.3956		
	1077.5034	IVGGWECEK	IVGGWECEK
	1272.6690	*LSEPAELTDAVK*	LSEPAELTDAVK
	1407.7513	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9413	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
		AVCGGVLVHPQWVLTAAHCIR	*AVCGGVLVHPQWVLTAAHCIR*
	2460.2154	LQCVDLHVISNDVCAQVHPQK	LQCVDLHVISNDVCAQVHPQK
	2588.3091	KLQCVDLHVISNDVCAQVHPQK	KLQCVDLHVISNDVCAQVHPQK
	3509.6976		
Band 7	673.3768	VVHYR	
	757.4922		

Table 2 (*continued*)

	PMF MASCOT	MS/MS MASCOT	MS/MS SEQUEST
Sample 3			
Band 7	870.3953		
	1077.5031	IVGGWECEK	IVGGWECEK
	1272.6689		
	1383.6693		
	1407.7503	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9400	*FLRPGDDSSHDLMLLR*	
Band 9	757.4919		
	870.3950		
	1077.5025	IVGGWECEK	IVGGWECEK
	1272.6685	*LSEPAELTDAVK*	
	1383.6689		
	1407.7497	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9398		
Band 10	757.4928		
	1077.5027	IVGGWECEK	IVGGWECEK
	1272.6699	LSEPAELTDAVK	
	1383.6707		
	1407.7517	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9412	FLRPGDDSSHDL**M**LLR	
Band 11	757.4930		
	1077.5046	IVGGWECEK	IVGGWECEK
	1272.6704	LSEPAELTDAVK	
	1407.7520	HSQPWQVLVASR	HSQPWQVLVASR
	1887.9430	FLRPGDDSSHDL**M**LLR	FLRPGDDSSHDL**M**LLR
	3509.6958		
Sample 4			
Band 5	673.3770		
	757.4919		
	854.4009		
	870.3961		
	1077.5030	*IVGGWECEK*	IVGGWECEK
	1272.6692		
	1407.7512	HSQPWQVLVASR	
	1871.9443	*FLRPGDDSSHDLMLLR*	FLRPGDDSSHDLMLLR
Band 6	673.3769		
	757.4925		
	854.4008	*FMLCARG*	
	870.3954		
	1077.5037	IVGGWECEK	IVGGWECEK
	1272.6693	*LSEPAELTDAVK*	LSEPAELTDAVK
	1407.7510	HSQPWQVLVASR	HSQPWQVLVASR
	1871.9460	*FLRPGDDSSHDLMLLR*	
	1887.9445		FLRPGDDSSHDL**M**LLR
Band 7	757.4923		
	854.4007	*FMLCARG*	
	870.3960		
	1077.5034	IVGGWECEK	IVGGWECEK
	1272.6694	LSEPAELTDAVK	LSEPAELTDAVK
	1407.7512	HSQPWQVLVASR	HSQPWQVLVASR
	1871.9465	FLRPGDDSSHDLMLLR	
	1887.9427		

[a] Sequences in italics indicate peptides observed only in one of the experiments performed in triplicates.
[b] Bands were cut in a single gel piece.
[c] Bands 5 and 6 appeared as single protein band in gel but cut into half and analyzed separately.

The PMF search, using MASCOT search engine, was performed on the monoisotopic mass lists obtained by extracting 20 averaged full mass scans, resulting in unambiguous PSA identification with amino acid sequence coverage up to 73%. The delta mass values (accuracy) were in most cases close to zero, but never higher than 3 ppm.

3.4. *MS/MS strategy based on "in silico" PSA sequences*

A theoretical monoisotopic mass list of possible PSA specific peptides was generated *in silico*. The MS/MS fragments expected from all of the available PSA forms were copied into the MS data acquisition method and used for the selection

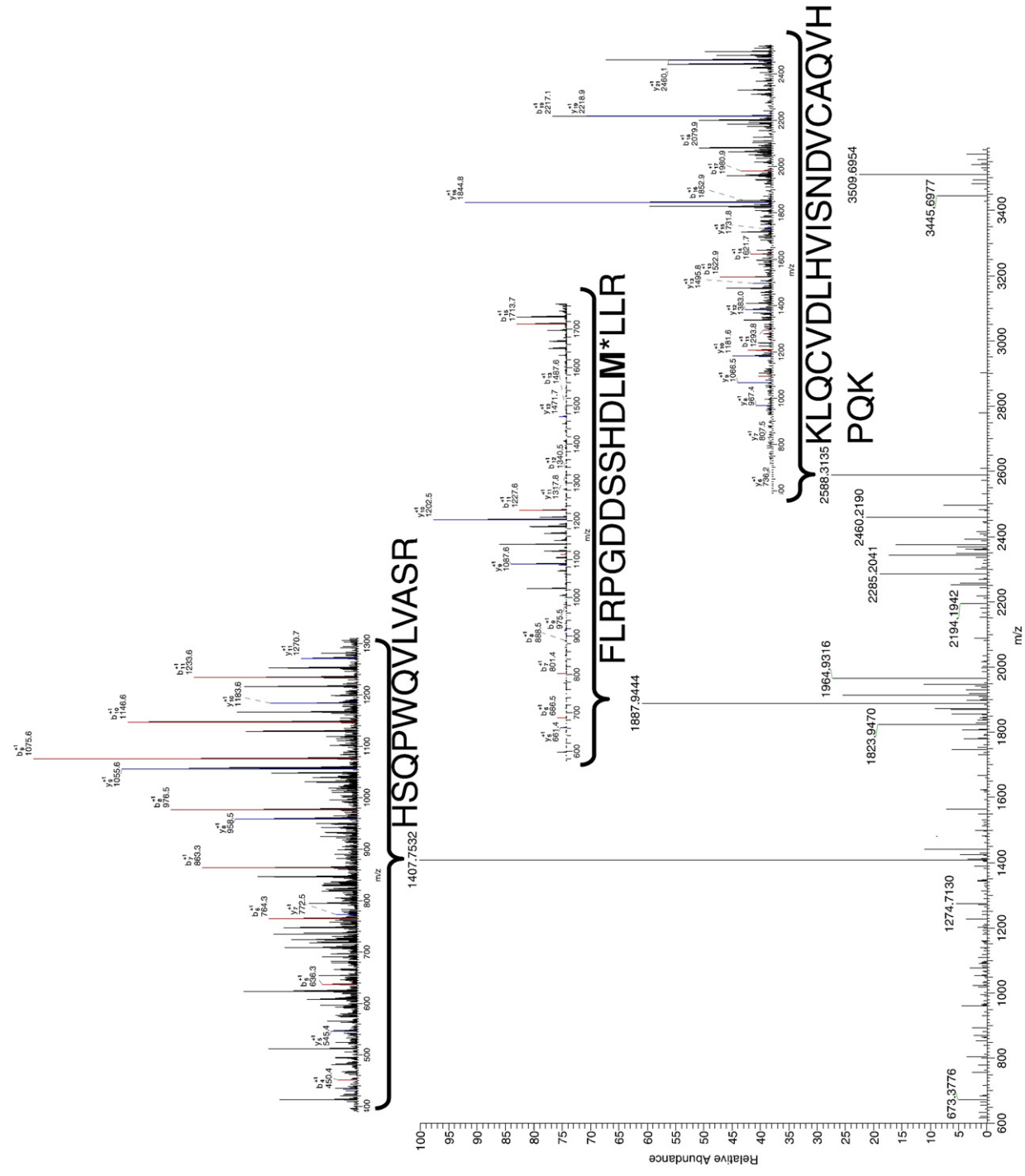


Fig. 3 – A typical example of the MS and MS/MS spectra obtained with the MALDI LTQ Orbitrap XL instrument showing three fragmentation spectra and their corresponding sequences in the inserts.

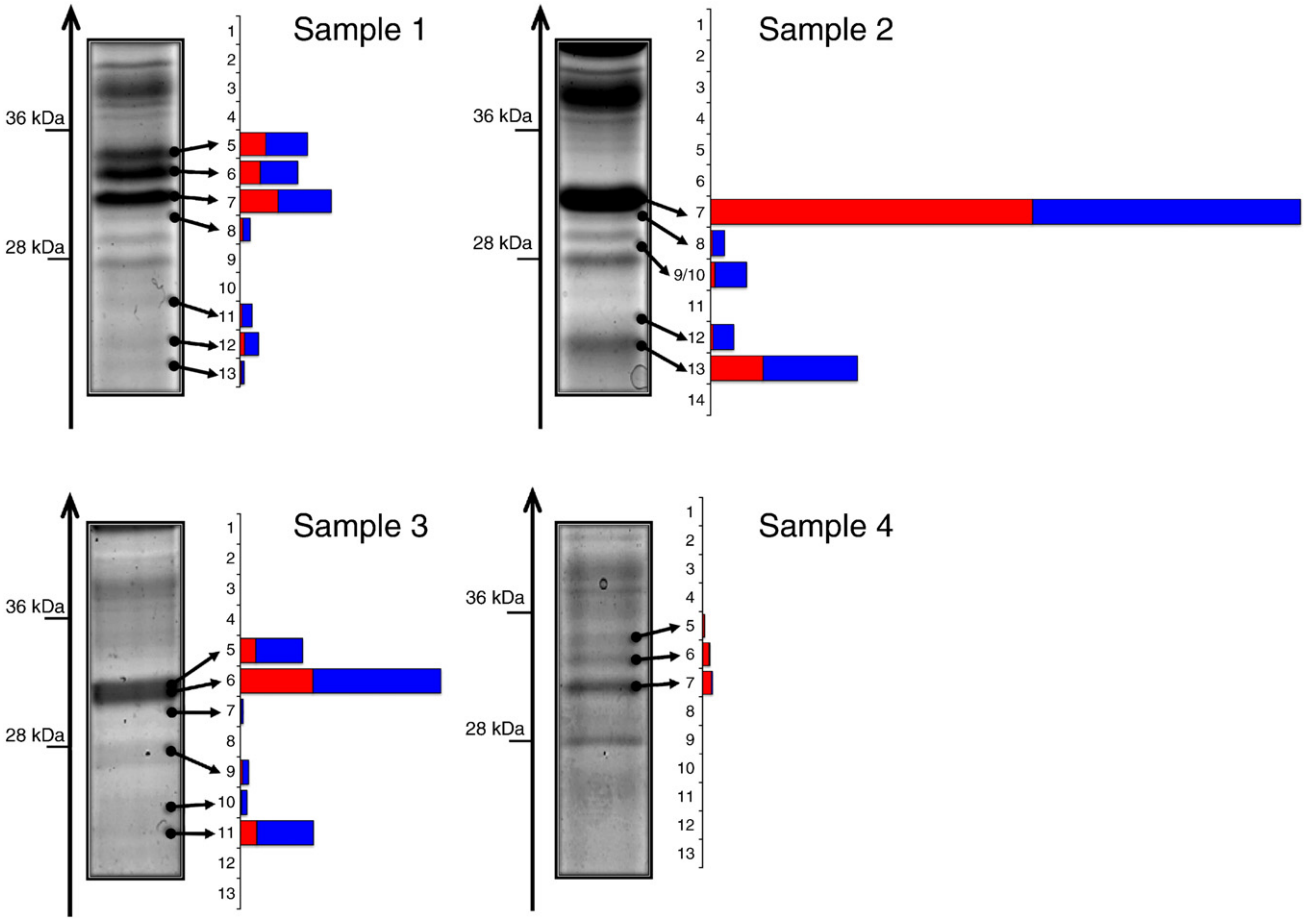


Fig. 4 – MS identification of PSA bands in gels depicted as it was determined by MALDI-MS data in peptide mass fingerprint search in the NCBInr database. The bars show the average of cumulative intensities of the two most abundant peptides in corresponding bands in SDS–PAGE gels. The volumes of the boxes refer to the contribution to the cumulative intensity of these peptides HSQPWQVLVASR (*m/z* = 1407.753) and FLRPGDDSSHDLM*LLR (*m/z* = 1887.945), which are shown in red and blue colors, respectively.

of parent ions. As part of our MS strategy, we selected twenty-two different sequences of PSA variants, collected by searching through the UniProtKB/SwissProt, UniProtKB/TrEMBL and NCBI protein databases, in which we identified 35 PSA forms. This indicates the existence of several duplicates assigned under different names. Due to the fact that many of the resulting peptides were identical, the total number of unique sequences was determined to be 138. We chose to exclude those with *m/z* less than 500 Da, but peptides with both modified and unmodified cystein and methionine side-chains were included.

Although, fragmentation in the ion trap can be further analyzed in the FT analyzer, we decided to use the ion trap for MS/MS data generation, trading the high mass resolution for high sensitivity. In evidence, comparing the MALDI Obritrap results with those obtained by nanoLC-ESI (on a LTQ XL ETD mass spectrometer, Thermo Scientific) the tryptic peptide patterns of PSA were found to overlap, whereby these findings could be confirmed (data not shown). PSA was identified based on fragment (MS/MS) data, using both the MASCOT, as well as the SEQUEST search engines (Table 2). We were able to successfully identify 12 PSA peptides using these stringent inclusion and exclusion criteria. All of the MS/MS spectra showed at least 50% complete b- and y-ion series (insets in Fig. 3). We identified 14 PSA peptides by PMF, which were in close agreement with our MS/MS hits.

The semi-quantitative analysis of MALDI-MS data, (based on the sum of isotopic patterns in 20 averaged full mass scans), showed good correlation by using HSQPWQVLVASR and FLRPGDDSSHDLM*LLR, as PSA-target sequences (as illustrated in Fig. 4) when compared with the fPSA values measured by an ELISA standardized against the WHO-endorsed calibration. In addition, we were able to prove that the averaged *m/z* intensities of the two selected target peptides correlate with the densitometric quantification.

4. Discussions

Several PSA isoforms have been described in the literature throughout the years, such as the alternative splice variants, and different glycosylations [24], of which some could be associated with diseases [20]. However, no systematic study has ever been conducted at the protein expression level. One of the reasons is probably the lack of robust and rigorously validated

analytical protocols that are capable of demonstrating an association between expressed PSA forms and disease. Using only a minute amount of seminal plasma (3.6 μl), our current experience and validation of performance of the presented toolbox concept indicates that it will be suited to run large series of samples from patients vs controls, *e.g.*, using collection of bio-repositories from large cohorts.

Interestingly, the highly selective and sensitive Western blot analysis could only determine one protein band with a molecular weight of 25 kDa in three samples. MS data did provide evidence that the corresponding bands in SDS–PAGE gel at 25 kDa in Samples 1 and 2 contained PSA. However, the PSA level in these bands appeared to be much lower in comparison to other PSA forms (as indicated from the gel bands) identified by the MS. This important observation may open up further in-depth studies of additional PSA forms that could be related to a given biological function. The zymographic separations could verify an enzymatically active protein within the 22.5–35.5 Mw region, but only one band in Sample 2 correlated well with the WB findings. Nine corresponding protein bands could also be localized in the SDS–PAGE gel positions, where PSA was identified by MS in all cases. Consequently, the mass spectrometric analysis plays a central role in the identification of the protein in semen samples. Hence, MS data appear critical to further understanding of PSA expression and disease linkage. Furthermore, we expect a good agreement in-between this mass spectrometry-based assay and the clinical diagnostic ELISA used as an international standard, which measures the affinity of antigen-, and antigen-like proteins in the final read-out.

Rapid screening is possible with the new generation of MALDI instrumentation, providing qualitative MS/MS sequencing by the MALDI LTQ Orbitrap XL, wherein the number of samples no longer remain to be a major limitation. This platform also has the advantage that it allows soft laser energy to be applied to the sample within the crystal spot, resulting in high sequence coverage and signal intensity in MS/MS peptide sequence peaks. The ion trap principle can provide high signal yields that, from a practical aspect, make it possible to analyze repeatedly with minimal sample loss. This feature has a particular advantage in running MS/MS sequencing that is unusual compared to traditional MALDI instrumentation, where high frequency lasers will deplete the sample rather quickly.

On an experimental basis, MS/MS searches resulted in hits on several PSA isoforms by blast searching through the NCBI and the UniProt databases. Additional PSA isoforms could be recognized in comprehensive searches through additional databases, such as the US Patent database. Our findings verify that unique iso-specific PSA forms do exist. However, the iso-annotation with the entire sequence coverage, as well as the possible modifications is still not sufficiently verified and presently remains unclear.

We found identical PSA variants that are listed under other names and/or annotations. One example is the isoform CRA_h on locus EAW71929 (gi:119592335) and the splice variant RP5 on locus CAD54617 (gi:24370944) that have identical amino acid sequences in the NCBI database. Furthermore, many of these proposed isoforms were splice variants; some of them were artificial sequences, which were not yet confirmed at the protein expression level.

In conclusion, we have shown that our platform enables the identification PSA molecular forms by means of peptide mass fingerprinting and MS sequencing. The application of our platform on seminal plasma samples clearly indicates that MS annotation can identify more molecular forms of PSA compared with Western and zymographic analyses. We propose that by using our strategy to identify and verify some of these novel variants at the protein level, which can open up a better understanding of the disease mechanisms. This may also include whether PSA has a causal influence in the development or progression of prostate disease or infertility. We also envision that the proposed analytical platform can be applicable to PSA determination in prostate cancer as preceding sample enrichment may overcome prior limitations due the low abundance of the protein target in the blood.

5. Conclusions

In this study we present a useful combination of techniques, which can be employed for collection of refined information regarding patterns of PSA isoforms in clinical samples. A profound difference was observed at the protein expression level comparing four samples from patients undergoing investigation for infertility and a healthy volunteer. The Western blot analysis proved that the monoclonal antibody (2E9) used for detection could clearly interact with a single or three molecular variants present in samples. However, the MS analysis proved the existence of 15 PSA forms with different molecular weights at various concentration ratios.

We believe that our observation of various isoforms of PSA in patients reported in this communication may contribute to the further identification of disease-relevant heterogeneity of PSA, including transcriptional and post-translational modifications present due to various stages and causes of prostate disease. This may also allow comprehensive structural characterization of this important target protein in samples with low concentration of PSA, such as blood, in the near future. This is particularly interesting since circulating PSA exists predominantly in complexed form and the structurally and functionally detailed relationship between isoform patterns in prostate tissue and blood is presently unclear.

Conflict of interest

Dr. Hans Lilja holds patents for free PSA and hK2 assays.

Acknowledgments

We would like to thank Gun-Britt Eriksson at the Dept. of Clinical Chemistry, University Hospital, Malmö, Sweden, for technical assistance; Martin Hornshaw and Egon Rosén at Thermo Fisher Scientific, for mass spectrometry support. The authors are grateful for funding support from the Swedish Research Council, Vinnova and Foundation for Strategic Research – The Programme: Biomedical Engineering for Better

Health – grant no.: 2006-7600, Swedish Research Council grant no: 2006-6020 and grant no: K2009-54X-20095-04-3, Swedish Cancer Society (08-0345), Knut and Alice Wallenberg Foundation, Crafoord Foundation, Carl Trygger Foundation, Fundación Federico SA, Royal Physiographic Society, Sten Lexner Foundation and Hecht Foundation, Sidney Kimmel Center for Prostate and Urologic Cancers, National Cancer Institute Specialized Programs of Research Excellence (P50-CA92629) and David H. Koch through the Prostate Cancer Foundation.

REFERENCES

[1] Lilja H. A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. J Clin Invest 1985;76:1899–903.
[2] Watt KWK, Lee P-J, M'Timkulu T, Chan W-P, Loor R. Human prostate-specific antigen: structural and functional similarity with serine proteases. Proc Natl Acad Sci USA 1986;83:3166–70.
[3] Lundwall Å, Lilja H. Molecular cloning of human prostate specific antigen cDNA. FEBS Lett 1987;214:317–22.
[4] Stenman U-H, Leinonen J, Zhang W-M, Finne P. Prostate-specific antigen. Semin Cancer Biol 1999;9:83–93.
[5] Lilja H, Ulmert D, Vickers AJ. Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. Nat Rev Cancer 2008;8:268–78.
[6] Lilja H, Oldbring J, Rannevik G, Laurell CB. Seminal vesicle-secreted proteins and their reactions during gelation and liquefaction of human semen. J Clin Invest 1987;80:281–5.
[7] Pezzato E, Sartor L, Dell Aica I, Dittadi R, Gion M, Belluco C, et al. Prostate carcinoma and green tea: PSA-triggered basement membrane degradation and MMP-2 activation are inhibited by (-)epigallocatechin-3-gallate. Int J Cancer 2004;112:787–92.
[8] Lövgren J, Rajakoski K, Karp M, Lundwall Å, Lilja H. Activation of the zymogen form of prostate-specific antigen by human glandular kallikrein 2. Biochem Biophys Res Commun 1997;238:549–55.
[9] Takayama TK, Fujikawa K, Davie EW. Characterization of the precursor of prostate-specific antigen. J Biol Chem 1997;272:21582–8.
[10] Emami N, Diamandis EP. Human kallikrein-related peptidase 14 (KLK14) is a new activator component of the KLK proteolytic cascade. J Biol Chem 2008;283:3031–41.
[11] Christensson A, Laurell CB, Lilja H. Enzymatic activity of prostate-specific antigen and its reactions with extracellular serine proteinase inhibitors. Eur J Biochem 1990;194:755–63.
[12] Christensson A, Lilja H. Complex-formation between protein-C inhibitor and prostate-specific antigen in-vitro and in human semen. Eur J Biochem 1994;220:45–53.
[13] Espana F, Gilabert J, Estelles A, Romeu A, Aznar J, Cabo A. Functionally active protein-C inhibitor plasminogen-activator inhibitor-3 (PCI/PAI-3) is secreted in seminal-vesicles, occurs at high-concentrations in human seminal plasma and complexes with prostate-specific antigen. Thromb Res 1991;64:309–20.
[14] Laurell M, Christensson A, Abrahamsson PA, Stenflo J, Lilja H. Protein-C inhibitor in human-body fluids — seminal plasma is rich in inhibitor antigen deriving from cells throughout the male reproductive-system. J Clin Invest 1992;89:1094–101.
[15] Stenman UH, Leinonen J, Alfthan H, Ranniko S, Tuhkanen K, Alfthan O. A complex between prostate-specific antigen and a1-antichymotrypsin is the major form of prostatic-specific antigen in serum of patients with prostatic cancer: assay of the complex improves clinical sensitivity for cancer. Cancer Res 1991;51:222–6.
[16] Lilja H, Christensson A, Dahlen U, Matikainen MT, Nilsson O, Pettersson K, et al. Prostate-specific antigen in serum occurs predominantly in complex with alpha 1-antichymotrypsin. Clin Chem 1991;37:1618–25.
[17] Heuzé N, Olayat S, Gutman N, Zani M-L, Courty Y. Molecular cloning and expression of an alternative hKLK3 transcript coding for a variant protein of prostate-specific antigen. Cancer Res 1999;59:2820–4.
[18] Tan OL, Whitbread AK, Clements JA, Dong Y. Kallikrein-related peptidase (KLK) family mRNA variants and protein isoforms in hormone-related cancers: do they have a function? Biol Chem 2006;387:697–705.
[19] Mikolajczyk SD, Millar LS, Marker KM, Wang TJ, Rittenhouse HG, Marks LS, et al. Seminal plasma contains BPSA, a molecular form of prostate-specific antigen that is associated with benign prostatic hyperplasia. Prostate 2000;45:271–6.
[20] Hilz H, Noldus J, Hammerer P, Buck F, Lück M, Huland H. Molecular heterogeneity of free PSA in sera of patients with benign and malignant prostate tumors. Eur Urol 1999;36:286–92.
[21] Nurmikko P, Pettersson K, Piironen T, Hugosson J, Lilja H. Discrimination of prostate cancer from benign disease by plasma measurement of intact, free prostate-specific antigen lacking an internal cleavage site at Lys145–Lys146. Clin Chem 2001;47:1415–23.
[22] Dickson EF, Pollak A, Diamandis EP. Ultrasensitive bioanalytical assays using time-resolved fluorescence detection. Pharmacol Ther 1995;66:207–35.
[23] Pettersson K, Piironen T, Seppala M, Liukkonen L, Christensson A, Matikainen MT, et al. Free and complexed prostate-specific antigen (PSA) — in-vitro stability, epitope map, and development of immunofluorimetric assays for specific and sensitive detection of free PSA and PSA-alpha(1)-antichymotrypsin complex. Clin Chem 1995;41:1480–8.
[24] White KY, Rodemich L, Nyalwidhe JO, Comunale MA, Clements MA, Lance RS, et al. Glycomic characterization of prostate-specific antigen and prostatic acid phosphatase in prostate cancer and benign disease seminal plasma fluids. J Proteome Res 2009;8:620–30.