Proteomic analysis of heparin-binding proteins from human seminal plasma: a step towards identification of molecular markers of male fertility

Vijay Kumar[1], Md Imtaiyaz Hassan[1], Anil Kumar Tomar[1], Tara kashav[2], Jaya Nautiyal[1], Sarman Singh[3], Tej P Singh[1] and Savita Yadav[1,*]

[1]*Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110 029, India*
[2]*School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India*
[3]*Department of Lab Medicine, All India Institute of Medical Sciences, New Delhi 110 029, India*

**Corresponding authors (Email, savita11@gmail.com)*

Glycosaminoglycans, especially heparin, are involved in various cell processes such as apoptosis, cell cycle control, platelet activation, capacitation, acrosome reaction and sperm decondensation. Heparin-binding proteins (HBPs) are essential constituents of human seminal fluid, which bind to sperm lipids containing the phosphorylcholine group and mediate the fertilization process. We utilized a proteomic set-up consisting of affinity chromatography, isoelectric focusing (IEF) coupled with matrix-assisted laser desorption/ionization time-of-flight tandem mass spectrometry (MALDI TOF/MS) for protein analysis of human HBPs. We resolved 70 different spots on two-dimensional (2-D) gel and subsequently identified these proteins. Forty different types of proteins were identified. Functional analysis revealed that 38% of the proteins belonged to the enzyme category, 20% were involved in RNA processing and transcription, 18% in structure and transport function, and 16% in cell recognition and signal transduction. We also identified 8% of proteins with unknown functions, although their expression in seminal fluid has been documented. Proteins of seminal fluid that bind heparin may be directly involved in sperm capacitation and acrosome reaction (AR), which are the two critical steps for fertilization. This information on HBPs would be useful for identifying potential biomarkers of fertility in the near future.

[Kumar V, Hassan M I, Tomar A K, Kashav T, Nautiyal J, Singh S, Singh T P and Yadav S 2009 Proteomic analysis of heparin-binding proteins from human seminal plasma: a step towards identification of molecular markers of male fertility; *J. Biosci.* **34** 000–000]

1. Introduction

Males of many species with either external or internal fertilization release their spermatozoa along with seminal fluid. Seminal fluid is a complex mixture of secretions originating from the testis, epididymis and male accessory glands such as the seminal vesicles, ampulla, prostate and bulbourethral glands (Pilch and Mann 2006). It serves as a vehicle that carries ejaculated sperm to the female genital tract. A common feature of most seminal plasma proteins is their ability to interact with different types of inorganic and organic materials present in seminal fluid (Russell *et al.* 1984; Cameron *et al.* 2007). Some of these proteins are bound to the sperm surface during ejaculation and thus protein-coating layers are formed (Varilova *et al.* 2006). In the female reproductive tract, seminal plasma proteins

Keywords. 2D gel electrophoresis; human seminal plasma; heparin-binding proteins; MALDI-TOF; sperm capacitation

Abbreviations used: 2-DE, two-dimensional gel electrophoresis; ACN, acetonitrile; ACTH, adrenocortocotrophic hormone; AR, acrosome reaction, DTT, dithiothreitol; GAG, glycosaminoglycan; HBP, heparin-binding protein, IEF, isoelectric focusing; IPG, immobilized pH gradient; KRAB, Kruppel-associated box; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight, MS, mass spectrometry, NCBI, National Center for Biotechnology Information; PMF, peptide mass fingerprinting; PSA, prostate-specific antigen; TFA, trifluoroacetic acid; ZFP, zinc-finger protein

bound to the sperm surface most probably participate first in the formation of the oviductal sperm reservoir (Evans and Kopf 1998; Jansen *et al.* 2001); second, in the control of sperm capacitation by the intensive action of negative (decapacitation factors) and positive regulatory (capacitation-stimulating) factors, and finally in central fertilization events such as sperm–zona pellucida interaction and sperm–egg fusion (Primakoff and Myles 2002; Yi *et al.* 2007). Seminal plasma proteins have also been shown to modulate the immune responses in the uterine environment. Thus, the combined effects of seminal plasma components support the survival of sperm within the female reproductive tract and ensure that functionally competent sperm meets the ovulated egg at the site of fertilization (Kanwar *et al.* 1979).

The basic requirement for fertilization to occur is the presence of spermatozoa at the site of fertilization and their capacitation (Nixon *et al.* 2007). Another subsequent critical step is the acrosome reaction (AR), which is normally induced by the zona pellucida as spermatozoa bind to the egg (Wassarman 1999). A prerequisite for acrosome reaction is capacitation, which is enhanced by glycosaminoglycans (GAGs) (Handrow *et al.* 1982; Lenz *et al.* 1983; Parrish *et al.* 1988). GAGs are disaccharide polymeric components of proteoglycans found in many tissues. Among the several GAGs tested, heparin is the most potent enhancer of capacitation in bovine and rabbit spermatozoa (Parrish *et al.* 1988). It has been previously shown that the heparin in seminal fluid stimulates sperm capacitation in bulls. Interestingly, induction of sperm capacitation in the female reproductive tract is aided by HBPs, secreted by the male accessory sex glands (Miller *et al.* 1990). Seminal fluid HBPs are supposed to attach themselves to the sperm surface, especially lipids containing the phosphoryl-choline group, thus allowing heparin-like GAGs in the female reproductive tract to activate sperm capacitation (Miller *et al.* 1987). Thus, seminal fluid HBPs play a vital role in spermatozoon survival and the overall fertilization process, and any alteration in these proteins can be directly related to infertility. HBPs have been identified in various species, e.g. bovine (Chandonnet *et al.* 1990), equine (Calvete *et al.* 1995, 1997) and canine (de Souza *et al.* 2006). However, very little attention has been paid to the identification and characterization of HBPs from human seminal fluid. We have earlier reported the purification of seven HBPs from human seminal fluid (Kumar *et al.* 2008). In this study, we report the identification of HBPs from human seminal fluid through two-dimensional gel electrophoresis (2-DE) and matrix-assisted laser desorption/ionization time-of-flight/mass spectrometry (MALDI-TOF/MS). We believe that a thorough analysis of the HBPs in seminal plasma in particular will serve as a reference for future studies in basic physiology, as well as the discovery of biomarkers of male infertility.

2. Materials and methods

2.1 *Chemicals and reagents*

Acetonitrile (ACN), methanol, acetic acid, ammonium bicarbonate, ammonium dihydrogen phosphate, phosphoric acid and sodium hydroxide were obtained from Sisco research laboratories (India). Iodoacetamide, urea and trifluoroacetic acid (TFA) were purchased from Sigma Aldrich (St Louis, MO, USA). Dithiothreitol (DTT) was obtained from GE Healthcare (Uppsala, Sweden). Trypsin mass grade was from Sigma Aldrich (St Louis, MO, USA).

2.2 *Sample collection and affinity chromatography on heparin sepharose*

Human semen samples were obtained from the Department of Laboratory Medicine, All India Institute of Medical Sciences, New Delhi. Sample collection and affinity chromatography were performed as described previously (Kumar *et al.* 2008). The absorbance of eluted fractions was measured at 280 nm. The fractions containing proteins were pooled, concentrated and desalted using ultrafiltration (Amnicon, USA), lyophilized and used for the subsequent analysis.

2.3 *2-DE*

The sample was prepared according to the method described previously for crude human seminal plasma proteome (Hassan *et al.* 2007). The eluted fractions obtained from a heparin–sepharose CL-6B column were precipitated at 4°C for 60 min using a precipitation kit according to the manufacturer's protocol (GE-Healthcare Upsala, Sweden). Following further centrifugation at 10 000 rpm for 30 min, the supernatant was removed and the pellets were washed with cold acetone (three times). Protein concentration was assessed by the 2D quant kit (GE healthcare, Upsala Sweden), and 100 μg protein was resuspended in 2-DE rehydration buffer (350 μl, 8 M urea, 2% CHAPS, 18 mM DTT, 0.5% immobilized pH gradient [IPG] buffer, bromophenol blue) and passively rehydrated into 13 cm immobilized pH gradient strips (IPG 3–10, GE- Healthcare Upsala, Sweden) for 16 h First-dimension IEF was performed by the IPG phor™ system according to the manufacturer's protocol (GE-Healthcare Upsala, Sweden). The instrument was programmed for 8 h rehydration, 4 h active rehydration at 30 V, 1 h at 300 V, 1 h at 8000 V (gradient), and 10 h at 8000 V (steady-state level). During the actual run, the final step was shortened to only 4–5 h (at 40 000 V h). Upon completion of the first dimension, IPG strips were incubated with gentle shaking in the equilibration buffer (50 mM Tris-HCl, pH 8.8, 6 M urea, 30% glycerol, 2% SDS, a trace of bromophenol

blue containing 1% DTT) for 15 min. Second-dimension IEF was performed on 12% polyacrylamide gels. Gels were run for 12–16 h, fixed, stained with silver stain (16 h) and then destained overnight. Destained gels were then scanned using an Image Scanner™ flatbed scanner with Magic Scan™ 32 software. Gels were analysed using the Image Master™ 2-D Elite Software (GE-Healthcare Upsala, Sweden). For comparative analysis, the relative intensity of each pair of matching spots was measured and the relative abundance and differential abundance values were displayed.

2.4 *In-gel digestion of protein spots*

The spots from 2-DE were manually excised, destained with 250 μl of 50 mM sodium thiosulphate and 15 mM potassium ferricyanide, washed twice with 500 μl of 50 mM ammonium carbonate/50% acetonitrile, dehydrated with 500 μl of 100% acetonitrile for 20 min at room temperature, reduced for 30 min at 56°C with 150 μl of 10 mM DTT in 100 mM ammonium bicarbonate, alkylated with iodoacetamide (100 μl of 50 mM iodoacetamide in 100 mM ammonium bicarbonate) in the dark at room temperature for 30 min. After extensive washing with ammonium bicarbonate/100% acetonitrile and dehydration, the protein spots were digested at 4°C with 30 μl of 0.02 μg/μl trypsin (modified porcine TPCK-treated, sequencing grade, SIGMA, USA) in 50 mM ammonium bicarbonate and incubated at 37°C for 16–18 h (overnight). The tryptic peptides were then extracted with 60% acetonitrile, 1% trifluoroacetic acid (TFA) and the organic solvent was evaporated in a vacuum centrifuge (12 000 *g* for 30 s) to near dryness (10 μl). Five microlitres of the in-gel digest sample was then mixed with 5 μl of resuspension solution (50% acetonitrile, 0.1% TFA) and gently agitated on a vortex. Of the reconstituted in-gel digested sample, 2 μl was then mixed with 1 μl of alpha-cyano-4-hydroxycinnamic acid matrix (10 mg/ml in 50% ACN, 0.1%TFA; Bruker, Daltonics) and spotted onto a MALDI-target plate. Analysis was performed directly after spotting.

2.5 *MALDI-TOF MS and data analysis*

Mass spectra were collected on a Bruker Autoflex MALDI-TOF mass spectrometer (Bruker Daltonics) equipped with a pulsed 337 nm N2 laser. Operating conditions were: ion source 1 = 19.00 kV, ion source 2 = 16.50 kV, lens voltage = 8.80 kV, reflector voltage = 20.0 kV, optimized pulsed ion extraction time = 80 ns, matrix suppression = 400 Da, and positive reflection mode. Spectra were calibrated externally using a standard peptide mixture (angiotensin II [1046.5 Da], angiotensin I [1296.7 Da], substance P amide [1347.7 Da], bombesin [1619.8 Da), adrenocorticotrophic hormone (ACTH) fragment 1–17 [2093.1 Da], ACTH fragment 18–39 [2465.2 Da], and somatostatin 28 [3147.5 Da], Bruker). Peptide monoisotopic signals were analysed using the SNAP algorithm, implemented in the FlexAnalysis software (Bruker Daltonics). Real data were acquired with the help of the WarpLC Software (Bruker Daltonics) for automatic TOF-MS spectra acquisition, background signal filtering, and grouping of signals into a peptide profile with respect to their distribution and intensity. The acquired data were searched against non-redundant databases from the National Center for Biotechnology Information (NCBI) and Swiss-Prot with the automated database-searching program using the Mascot search engine version 2.2 (Matrix Science, Boston, MA, USA). The following specified parameters were applied for database search: database (Swiss-Prot v. 51.6); taxonomy (*Homo sapiens*); proteolytic enzyme (trypsin); peptide mass tolerance (25 ppm); fragment mass tolerance (0.25 Da); global modification (carbamidomethyl [Cys]); variable modification (oxidation [Met]); peptide charge state (+1); and maximum missed cleavage (1). The criteria used to accept identifications included the extent of sequence coverage, the number of peptides matched, and the probabilistic score (MOWSE score), as indicated in table 1.

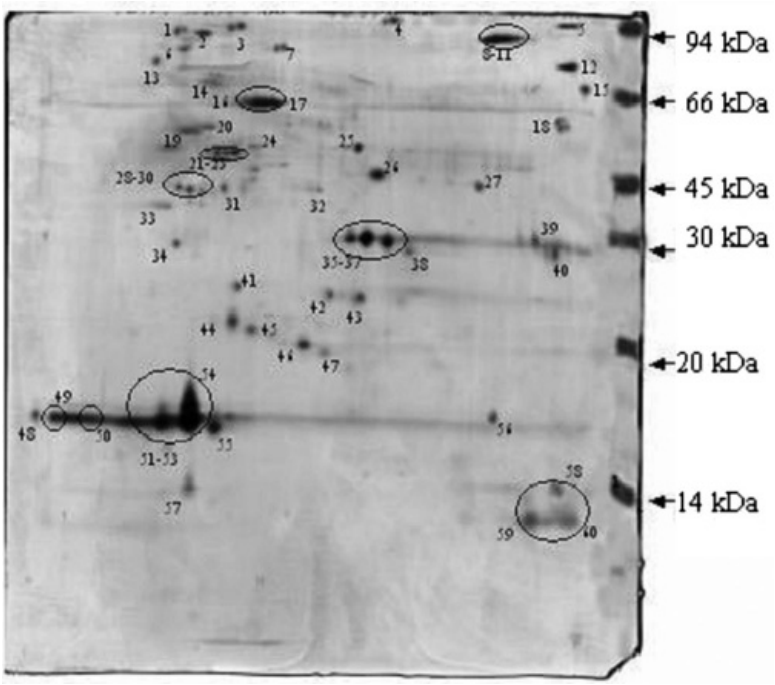
3. Results

3.1 *2-D gel analysis and identification*

Seventy spots (ranging from 10 to 90 kDa and an isoelectric point (pI) of 4–10) corresponding to HBPs were resolved by 2-DE from normal human seminal fluid (figure 1). These spots were excised from the gel, trypsinized and the resulting peptides were analysed by MALDI-TOF MS, as described in Materials and methods. The masses of monoisotropic peaks obtained were used to identify each protein by comparing theoretical digestions of proteins by trypsin. Out of the 70 spots resolved, 60 spots were identified by peptide mass fingerprinting (PMF) analysis. The list of identified proteins is shown in table 1. Ten of the spots failed to give any identification, giving an overall identification success rate of 86% for this study. A number of the proteins appeared as multiple spots on 2-D gel, possibly due to post-translational modifications, proteolytic breakdown or due to expression of highly related gene sequences. Thus, the 60 spots identified corresponded to 40 different proteins. The experiments were repeated three times from five different pooled fractions; between each of the samples the variation was less than 5%.

3.2 *Classification of HBPs*

We subsequently grouped all the identified proteins into functional clusters using Blast2GO functional annotations (Conesa *et al.* 2005) to define the cellular processes in which



Figures 1. 2-D gel analysis of heparin-binding proteins (HBPs) from human seminal plasma. 100 µg of the HBP sample was subjected to isoelectric focusing (IEF) using 13 cm pI 3–10 immobilized pH gradient (IPG) strips, followed by second- dimension fractionation on 12% polyacrylamide gels and silver stained. Spots from the gels were excised and subjected to in-gel tryptic digestion followed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis.

Table 1. List of human seminal plasma heparin-binding proteins (HBPs)

S Spot number (s)	Protein name	Accession number	Molecular mass/p*I*		Score	% Seq. cov/ no. of peptides matched
			Determined	Expected		
1	Catenin alpha-3	Q9UI47	90/5.3	99.8/5.8	39.20	30.6/3
2	SAPS domain family member 3	Q5H9R7	90/5.4	97.6/4.5	74.50	56.8/9
3	DNA topoisomerase 1	P11387	90/5.9	90.7/9.3	89.3	24.6/22
4	Putative uncharacterized protein DKFZp686K139 [fragment]	Q6MZF4	94/7.3	122/6.2	101.0	17.2/10
5	Zinc finger protein 225	Q9UK10	82/9.4	82.4/9.1	61.20	20.8/14
6	Matrix metalloproteinase-9 [precursor]	P14780	78/5.3	78.4/5.7	61.20	20.8/14
7	Chromodomain-helicase DNA-binding protein 3 fragment	Q2TAZ1	78/6.2	73.5/5.7	41.0	18.9/5
8	Lactoferrin	P02788	76/9.4	78/8.5	334.0	45.1/26
9	Lactoferrin	P02788	76/9.4	78/8.5	244.0	37.2/20
10	Lactoferrin	P02788	76/9.4	78/8.5	248.0	42.0/22
11	Lactoferrin	P02788	76/9.4	78/8.5	248.0	42.0/22
12	Human chorionic gonadotrophin hormone (hCG)	P22888	80/9.2	78.6/8.8	55.00	45.5/5
13	Putative uncharacterized protein DKFZp686C12204	Q7Z2Z3	74/5.4	71.0/7.4	89.90	20.8/13
14	Matrix metalloproteinase-2	P08253	72/5.4	73.8/5.26	84.50	30.8/10
15	ATP-dependent DNA helicase PIF1	Q9H611	70/9.4	69.7/9.8	42.70	23.3/5
16	Human serum albumin	P02768	66/6.0	69.3/5.9	89.8	18.8/10

Table 1. (*Continued*)

17	Hman serum albumin	P02768	66/6.0	69.3/5.9	100.0	17.2/9
18	Unnamed protein product		65/9.4		32.80	18.8/3
19	Semenogelin II	Q02383	64/5.4	65.4/9.0	64.10	24.6/16
20	Semenogelin II	Q02383	65/5.6	65.4/9.0	58.0	32.8/14
21	Semenogelin I	P04279	52/5.8	52.1/9.3	54.50	31.3/6
22	Semenogelin I	P04279	52/5.9	52.1/9.3	55.9	18.8/5
23	Semenogelin I	P04279	52/5.9	52.1/9.3	64.1	17.2/9
24	ATP/GTP-binding protein-like 4	Q5VU57	55/6.0	62.3/8.8	53.10	21.3/12
25	Keratin, type II cytoskeletal 1	P04264	55/6.8	66.0/8.1	81.4	20.2/7
26	Alkaline phosphatase placental-like [precursor]	P10696	54/7.1	57.3/5.9	17.2	11.3/2
27	Zinc-finger protein 101	Q8IZC7	50/8.6	50.3/9.6	83.7	26.3/7
28	Fibronectin type 3	Q8TC84	38/5.2	38.3/8.9	74.5	56.8/9
29	Fibronectin type 3	Q8TC84	38/5.3	41.0/9.0	68.2	56.7/4
30	Fibronectin type 3	Q8TC84	37/5.4	37.2/8.4	95.70	54.4/11
31	Zinc-finger protein 169	Q5SR55	38/5.8	46.9/9.6	79.10	23.4/ 11
32	Protein-tyrosine phosphatase-like member A	Q9HB93	36/6.2	32.7/9.3	42.0	14.5/4
33	Steroid 21-hydroxylase	P08686	36/5.2	55.8/7.7	36.30	10.1/4
34	Elongation factor 1-delta	P29692	30/5.2	31.1/4.9	32.60	18.9/3
35	Prostate-specific antigen precursor	Q8IWY3	29/7.2	29.2/9.0	122.0	52.0/6
36	Prostate-specific antigen precursor	Q8IWY3	29/7.4	29.2/9.0	180.0	59.9/14
37	Prostate-specific antigen precursor	Q8IWY3	29/7.6	29.2/9.0	129.0	46.8/11
38	Prostate-specific antigen	P07288	28/7.8	28.7/7.6	109.0	44.3/14
39	Apolipoprotein-I	P02647	29/9.3	30.7/5.5	62.0	12.9/6
40	Similar to TIM isoform 1	P60174	28/9.4	26.6/6.4	34.50	14.3/3
41	Ras-related protein Rab-3D	O95716	25/5.8	24.2/4.7	65.0	32.9/6
42	Putative uncharacterized protein MGC3196	A0PJW6	24/7.0	22.0/11	44.30	36.0/6
43	Metalloproteinase inhibitor 1 [precursor]	P01033	24/7.2	23.1/8.4	38.80	16.3/6
44	Lactoferrin fragment	Q8IX02	25/5.8	25/9.4	58.20	18.8/5
45	Lactoferrin fragment	Q8IX02	23/6.0	25/9.4	58.2	18.8/5
46	Prostaglandin-H2 D-isomerase	P41222	22/6.8	21/7.6	89.80	18.8/10
47	Cystatin-C-precursor	P01034	20/7.0	15.7/9.0	113.0	49.1/8
48	Angiogenin [precursor]	P03950	17/3.6	16.5/9.7	55.90	18.8/ 5
49	Putative uncharacterized protein P14	Q53TM1	17/4.2	14.5/9.4	34.30	16.5/4
50	Pregnancy-associated endometrial alpha-2-globulin	Q5T6T0	17/4.6	14.9/7.8	44.30	36.0/6
51	Fibronectin [fragment]	O95608	17/5.2	13.3/9.8	140.0	87.1/8
52	Fibronectin [fragment]	O95608	17/5.4	13.3/9.8	81.4	72.6/7
53–54	Fibronectin type III domain containing 5	Q7Z676	18/5.5	18.2/5.0	74.30	78.5/7
55	Seminal plasma protein –like-1	Q075Z2	16.5/5.6	15.6/6.8	100.0	17.2/8
56	Serine protease inhibitor Kazal type-2	P20155	17/8.8	9.2/9.3	173.0	46.0/22
57	Kunitz-type protease inhibitor 3	P49223	12/5.4	10.2/6.2	47.2	14.4/6
58	Semenogelin I [fragment]	Q6Y806	12/9.5	10.4/9.74	55.0	31.3/8
59	Semenogelin I [fragment]	Q6Y806	10/9.4	10.4/9.74	50.6	14.5/6
60	Semenogelin I [Fragment]	Q6Y808	10/9.4	10.3/9.74	80.2	34.0/7

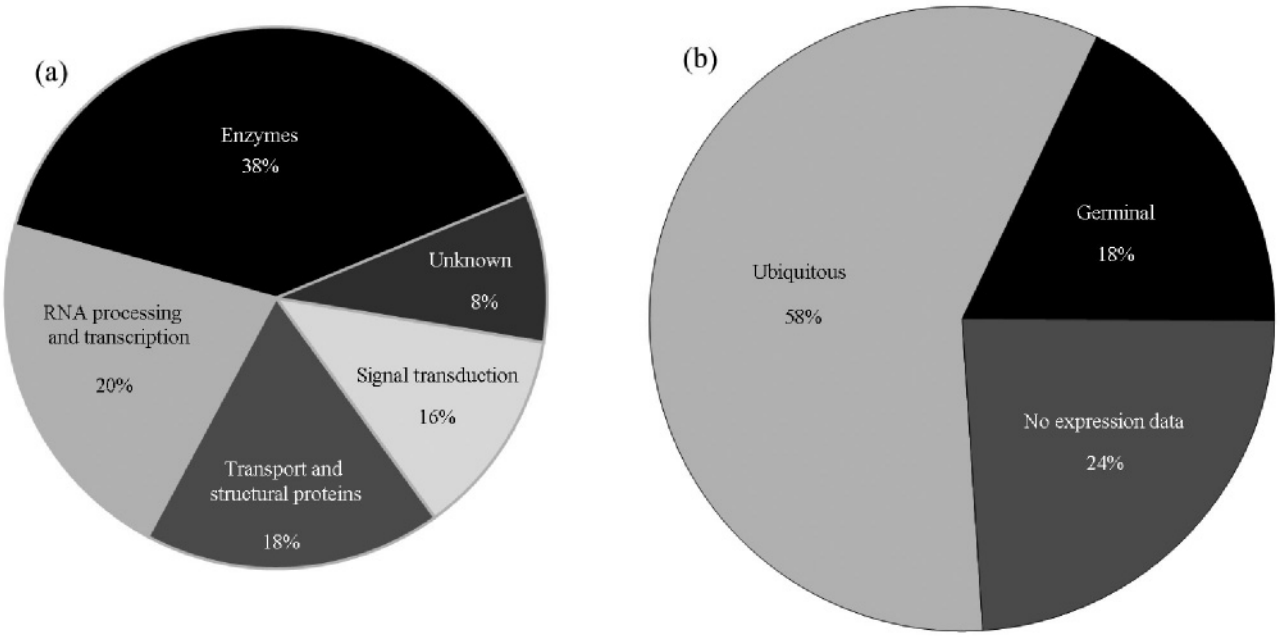


Figure 2. **(a)** Classification of human seminal plasma heparin-binding proteins (HBPs) into functional groups as a percentage of the total number of proteins identified. **(b)** Expression profile of human seminal plasma HBPs

they participate (figure 2A). The most abundant group corresponded to the proteins involved in protein metabolism (38%). Twenty per cent of the proteins were involved in RNA processing and transcription, 18% in transport and structural proteins, and 16% in cell recognition and signal transduction. Considering the expression pattern of these identified proteins, we found that 18% of the proteins were specifically expressed in different origins of seminal fluid. Of the proteins, 58% are ubiquitously expressed and no expression data in seminal fluid have so far been described for 24% of the proteins identified (figure 2B). Therefore, the identification of these 24% of proteins in human seminal fluid now provides the first experimental protein expression data available.

4. Discussion

In this study, we identified 60 spots out of 70 resolved spots corresponding to the HBPs extracted from normal human seminal fluid. These 60 spots corresponded to 40 different proteins. Based on their probable functions, these proteins were grouped into four major functional categories: (i) enzymes (22 proteins), (ii) RNA processing and transcription (12 proteins), (iii) transport and structural proteins (10 proteins), and (iv) cell recognition and signal transduction (7 proteins).

To validate their heparin-binding activity, we conducted kinetic studies using surface plasmon resonance methods. HBPs such as semenogelin I, semenogelin II, fibronectin fragment, lactoferrin and prostate-specific antigen (PSA) were passed through immobilized biotinylated heparin. The binding constants (K_A and K_D) were in the range of 10^4–10^9 1/M and 10^{-6}–10^{-10} M (results not shown). A few of the identified proteins have also been previously reported as HBPs, such as lactoferrin (Mann *et al.* 1994), fibronectin (Walker and Gallagher 1996) and PSA (*Hessick et al. 2005*). Various studies have reported that heparin-binding sites consist of clusters of basic amino acid residues such as XBBXBX, XBBBXXBX (Cardin and Weintraub 1989), BBXB (Mann *et al.* 1994; Proudfoot *et al.* 2001), XBBX, XBBBX and BXXB (Fromm *et al.* 1997), where B is a basic amino acid and X can be any amino acid. Sequence analysis of all the HBPs identified in our study revealed that most of them have such clusters of basic residues. Though we have shown the clusters of BBXB (table 2), some of these sequences also contain other reported consensus sequences rich in basic residues.

The first major group of HBPs includes a number of enzymes engaged in catalytic activities (38%). This high number of enzymes, especially proteases present in seminal plasma as HBPs, is not surprising as it has been described that almost 60% of the seminal fluid proteome is involved in enzymatic activity (Povoa 1962; Pilch and Mann 2006). Much evidence has been provided to show that proteases may be active partners in establishing and maintaining testicular architecture, and in facilitating germ cell migration throughout the spermatogenic developmental processes. To prevent unwanted proteolysis and keep these sperm proteases in an inactive state, these proteolytic enzymes are strongly regulated by protease inhibitors present in seminal

Table 2. Sequence analysis of human seminal plasma heparin-binding proteins (HBPs)

Identified HBP	Clusters of BBXB (B=basic amino acid, X=any other amino acid)
Catenin alpha 3	$_{48}$KKGR, $_{324}$HRER, $_{352}$KKER, $_{368}$KKTR, $_{681}$KKVK, $_{878}$KKIH
SAPS domain family member 3	$_{144}$KKKH, $_{257}$HKEK
DNA topoisomerase 1	$_{39}$KKEK, $_{103}$KKEK, $_{150}$KKIK, $_{158}$KKEK, $_{174}$KKPK, $_{448}$RRLK
Putative uncharacterized protein DKFZp686K139[fragment]	$_{65}$RRLR, $_{244}$RRPH, $_{621}$HKRH
Zinc-finger protein 225	$_{251}$RKLH, $_{342}$KRYK, $_{389}$RHVR, $_{503}$RRVH, $_{531}$RRVH, $_{587}$KRLH, $_{687}$KRYK
Chromodomain-helicase-DNA-binding protein 3 fragment	$_{2}$KRVR, $_{322}$HRER, $_{402}$HRRH, $_{456}$RRFK, $_{632}$KKEK
Lactoferrin	$_{2}$RRRR, $_{28}$RKVR, $_{342}$RRAR
Matrix metalloproteinase-2	$_{98}$RKPR, $_{115}$RKPK
ATP-dependent DNA helicase PIFI	$_{158}$RRLR, $_{615}$RRGR
Human serum albumin	$_{33}$HRFK, $_{183}$KRYK, $_{559}$HKPK
Semenogelin I	$_{296}$RRLH, $_{356}$RRLH, $_{425}$RHQH
Semenogelin II	$_{281}$RKAH
ATP/GTP-binding protein-like 4	$_{464}$RKEK
Alkaline phosphtase placental-like [precursor]	$_{78}$KKDK
Zinc-finger protein 101	$_{119}$RHMR, $_{141}$RKQK, $_{214}$KHGK, $_{299}$RHER, $_{357}$KKTH
Fibronectin type 3	$_{43}$KKAK, $_{335}$KKQR
Zinc-finger protein 169	$_{106}$RHFR, $_{118}$KRIH, $_{228}$RHQR, $_{284}$RHQR, $_{311}$RHQR, $_{368}$RHQR, $_{398}$RRTK
Steroid 21-hydroxylase	$_{224}$RRLK, $_{232}$KRDH
Elongation factor 1-delta	$_{185}$KKAK
PSA precursor	$_{196}$RKYR
RAS-related protein RAB-3D	$_{66}$RHDK
Pregnancy-associated endometrial alpha-2-globulin	$_{61}$KKFK
Fibronectin type III domain containing 5	$_{27}$RHLK

plasma. Protease inhibitors are present in the seminal fluid of many mammalian species, for example, human (Schiessler *et al.* 1976; Fink *et al.* 1990), bull (Cechova and Fritz 1976; Fritz *et al.* 1976) and boar (Veselsky *et al.* 1985), and they have been shown to play an important role in male fertility. They are also involved in regulation, processing or degradation of seminal fluid proteins and coagulation of semen (Zheng *et al.* 1994; He *et al.* 1999). A challenge for the future will be to identify the full complement of proteases and their regulatory mechanisms. This will enable the design of additional studies to precisely define the role and relative importance of each in the complex steps of testis development and spermatogenesis.

The next most abundant group of proteins identified corresponds to those involved in RNA processing and transcription (20%). This finding was not surprising as the majority of these proteins are nuclear or intracellular in origin. Within this group, DNA-dependent Cys2 His2 zinc-finger proteins (ZFPs) accounted for the most. In mammals, ZFPs probably constitute the largest families of transcription factor. The identified ZFPs contain Cys2 His2 zinc-finger motifs and at least one Kruppel-associated box (KRAB) domain. This Kruppel-related zinc-finger motif is present in approximately 3% of the proteins encoded by mouse and human genomes. This family effectively represses transcription through interaction with transcriptional intermediary factor 1β (TIF1β), a transcriptional co-repressor involved in gene silencing through heterochromatin formation (Looman *et al.* 2002; Urrutia 2003). The potential explanation for the proteins involved in transcription is that these proteins may be leftovers of the spermatogenesis process or may have important, but unknown, functions related to the fertilization and development of spermatozoa.

The third category of proteins consists of structural and transport proteins (18%). Among all structural proteins, the most abundant are gel-forming proteins secreted by the seminal vesicles: semenogelin I, semenogelin II and fibronectin, which are cleaved by kallikrein-like proteases. They form a viscous gel, entrapping spermatozoa immediately after ejaculation to provide protection from physical damage. The most abundant transport protein is lactoferrin, which is present in solution and has an antimicrobial role to play in seminal fluid. Lactoferrin binds

to the sperm surface membrane during ejaculation, and seems to be a major sperm-coating antigen (Goodman *et al.* 1981). It has been suggested that the affinity of lactoferrin for sperm and its potential role in protecting sperm from the immunocompetent environment of the female reproductive tract might be regulated by its clustering with some of the other proteins of seminal plasma (Thaler *et al.* 1990). Serum albumin, an important constituent of seminal plasma that has a role as a sink for cholesterol, is removed from the sperm membrane during capacitation (Mortimer *et al.* 1998; Bedu-Addo *et al.* 2005).

Among the identified proteins, seven of these are involved in cell recognition and signal transduction (figure 2a). These proteins are consistent with the functions required for sperm capacitation, interaction with the zona pellucida, acrosomal reaction and oocyte penetration. This group includes cadherin-associated protein, Ras-related Rab 3D, keratinocyte growth factor, mitogen-activated protein kinase kinase kinase (MLT) and elongation factor 1delta. We also identified five proteins (8%) with unknown function. However, the expression of some of these proteins in the testes and prostate is well known.

Proteins that have been previously described as HBPs, for example, protein C inhibitor (PCI) and its complex with PSA (Suzuki *et al.* 2007; Espana *et al.* 1991), hepatocyte growth factor (HGF) (Schwall *et al.* 1996; Sakata *et al.* 1997), phospholipase A2 (PLA2) (Bugs *et al.* 2005) and human spermadhesins-like (hsa) proteins (Kraus *et al.* 2001) were not detected in our study. Thus, we conclude that our experimental approach has permitted the identification of a significant number of new HBPs but some of them still escape detection by this approach, possibly due to the type of sample used for the experiments or the amount of these proteins in the sample. Thus, in the near future, further investigations using samples from different species will probably add to the list of new HBPs.

In summary, proteomic analysis of body fluids, especially human seminal plasma, is of great interest and has potential clinical implications. We have offered a proteomic strategy based on affinity chromatography followed by 2-DE, in conjunction with subsequent enzymatic digestion followed by MALDI-TOF/MS. This has been demonstrated to be a promising alternative approach for protein identification in complex body fluids. Our strategy proved to be robust, easy to use, and less work intensive than traditional proteomic approaches. The combination of proteomic analysis and functional studies will be a successful tool to discover new cellular pathways in which HBPs may be involved. Moreover, the identification of such proteins in seminal plasma could provide better insights into the physiological processes and the nature of subfertility or infertility.

Acknowledgements

This work was supported by grants from the Department of Biotechnology (DBT), New Delhi. VK and AKT are thankful to the Council of Scientific and Industrial Research (CSIR) and Department of Science and Technology (DST), respectively, for fellowships.

References

Bedu-Addo K, Lefievre L, Moseley F L, Barratt C L and Publicover S J 2005 Bicarbonate and bovine serum albumin reversibly 'switch' capacitation-induced events in human spermatozoa; *Mol. Hum. Reprod*. **11** 683–691

Bugs M R, Bortoleto-Bugs R K and Cornelio M L 2005 The interaction between heparin and Lys49 phospholipase A2 reveals the natural binding of heparin on the enzyme; *Int. J. Biol. Macromol*. **37** 21–27

Calvete J J, Raida M, Gentzel M, Urbanke C, Sanz L and Topfer-Petersen E 1997 Isolation and characterization of heparin- and phosphorylcholine-binding proteins of boar and stallion seminal plasma. Primary structure of porcine pB1; *FEBS Lett*. **407** 201–206

Calvete J J, Reinert M, Sanz L and Topfer-Petersen E 1995 Effect of glycosylation on the heparin-binding capability of boar and stallion seminal plasma proteins; *J. Chromatogr.* **A711** 167–173

Cameron E, Day T and Rowe L 2007 Sperm competition and the evolution of ejaculate composition; *Am. Nat*. **169** E158–E172

Cardin A D and Weintraub H J 1989 Molecular modeling of protein–glycosaminoglycan interactions; *Arteriosclerosis* **9** 21–32

Cechova D and Fritz H 1976 Characterization of the proteinase inhibitors from bull seminal plasma and spermatozoa; *Hoppe Seylers Z. Physiol. Chem*. **357** 401–408

Chandonnet L, Roberts K D, Chapdelaine A and Manjunath P 1990 Identification of heparin-binding proteins in bovine seminal plasma; *Mol. Reprod. Dev.* **26** 313–318

Conesa A, Gotz S, Garcia-Gomez J M, Terol J, Talon M and Robles M 2005 Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research; *Bioinformatics* **21** 3674–3676

de Souza F F, Martins M I, dos Santos Fernandes C E, Ribolla P E and Lopes M D 2006 Herapin-binding proteins of canine seminal plasma; *Theriogenology* **66** 1606–1609

Espana F, Gilabert J, Estelles A, Romeu A, Aznar J and Cabo A 1991 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*. **64** 309–320

Evans J P and Kopf G S 1998 Molecular mechanisms of sperm–egg interactions and egg activation; *Andrologia* **30** 297–307

Fink E, Hehlein-Fink C and Eulitz M 1990 Amino acid sequence elucidation of human acrosin-trypsin inhibitor (HUSI-II)

reveals that Kazal-type proteinase inhibitors are structurally related to beta-subunits of glycoprotein hormones; *FEBS Lett*. **270** 222–224

Fritz H, Tschesche H and Fink E 1976 Proteinase inhibitors from boar seminal plasma; *Methods Enzymol*. **45** 834–847

Fromm J R, Hileman R E, Caldwell E E O, Weiler J M and Linhardt R J 1997 Pattern and spacing of basic amino acids in heparin binding sites; *Arch. Biochem. Biophys*. **343** 92–100

Goodman S A and Young L G 1981 Immunological identification of lactoferrin as a shared antigen on radioiodinated human sperm surface and in radioiodinated human seminal plasma; *J. Reprod. Immunol.* **3** 99–108

Handrow R R, Lenz R W and Ax R L 1982 Structural comparisons among glycosaminoglycans to promote an acrosome reaction in bovine spermatozoa; *Biochem. Biophys. Res. Commun*. **107** 1326–1332

Hassan M I, Kumar V, Kashav T, Alam N, Singh T P and Yadav S 2007 Proteomic approach for purification of seminal plasma proteins involved in tumor proliferation; *J. Sep. Sci*. **30** 1979–1988

He S, Gaca M D, McEuen A R and Walls A F 1999 Inhibitors of chymase as mast cell-stabilizing agents: contribution of chymase in the activation of human mast cells; *J. Pharmacol. Exp. Ther*. **291** 517–523

Hessick C M, Greene B, Abbey B, Sulkowski E and Srikrishnan T 2005 Chromatography of prostate-specific antigen (PSA) on heparin-sepharose and IDA-Me (II)-agarose; *Proc. Am. Assoc. Cancer Res*. **46** 1352

Jansen S, Ekhlasi-Hundrieser M and Topfer-Petersen E 2001 Sperm adhesion molecules: structure and function; *Cells Tissues Organs* **168** 82–92

Kanwar K C, Yanagimachi R and Lopata A 1979 Effects of human seminal plasma on fertilizing capacity of human spermatozoa; *Fertil. Steril*. **31** 321–327

Kraus M, Ticha M and Jonakova V 2001 Heparin-binding proteins of human seminal plasma homologous with boar spermadhesins; *J. Reprod. Immunol*. **51** 131–144

Kumar V, Hassan M I, Kashav T, Singh T P and Yadav S 2008 Heparin-binding proteins of human seminal plasma: purification and characterization; *Mol. Reprod. Dev.* **75** 1767–1774

Lenz R W, Ball G D, Lohse J K, First N L and Ax R L 1983 Chondroitin sulfate facilitates an acrosome reaction in bovine spermatozoa as evidenced by light microscopy, electron microscopy and in vitro fertilization; *Biol. Reprod*. **28** 683–690

Looman C, Abrink M, Mark C and Hellman L 2002 KRAB zinc finger proteins: an analysis of the molecular mechanisms governing their increase in numbers and complexity during evolution; *Mol. Biol. Evol*. **19** 2118–2130

Mann D M, Romm E and Migliorini M 1994 Delineation of the glycosaminoglycan-binding site in human inflammatory response protein lactoferrin; *J. Biol. Chem*. **269** 23661–23667

Miller D J , Winer M A and Ax R L 1990 Heparin-binding proteins from seminal plasma bind to bovine spermatozoa and modulate capacitation by heparin; *Biol. Reprod*. **42** 899–915

Miller D J, First N L and Ax R L 1987 Isolation and characterization of seminal fluid proteins that bind heparin; *Adv. Exp. Med. Biol*. **219** 597–601

Mortimer S T, Swan M A and Mortimer D 1998 Effect of seminal plasma on capacitation and hyperactivation in human spermatozoa; *Hum. Reprod*. **13** 2139–2146

Nixon B, Aitken R J and McLaughlin E A 2007 New insights into the molecular mechanisms of sperm–egg interaction; *Cell Mol. Life Sci*. **64** 1805–1823

Parrish J J, Susko-Parrish J, Winer M A and First N L 1988 Capacitation of bovine sperm by heparin; *Biol. Reprod*. **38** 1171–1180

Pilch B and Mann M 2006 Large-scale and high-confidence proteomic analysis of human seminal plasma; *Genome Biol*. **7** R40

Povoa H Jr 1962 Enzyme localization in proteins separated by paper electrophoresis in human seminal plasma; *Experientia* **18** 552

Primakoff P and Myles D G 2002 Penetration, adhesion, and fusion in mammalian sperm-egg interaction; *Science* **296** 2183–2185

Proudfoot A E I, Fritchley S, Borlat F, Shaw J P, Vilbois F, Zwahlen C, Trkola A, Marchant D, Clapham P R and Wells T N C 2001 The BB*X*B motif of RANTES is the principal site for heparin binding and controls receptor selectivity; *J. Biol. Chem*. **276** 10620–10626

Russell L D, Peterson R N, Hunt W and Strack L E 1984 Posttesticular surface modifications and contributions of reproductive tract fluids to the surface polypeptide composition of boar spermatozoa; *Biol. Reprod.* **30** 959–978

Sakata H, Stahl S J, Taylor W G, Rosenberg J M, Sakaguchi K, Wingfield P T and Rubin J S 1997 Heparin binding and oligomerization of hepatocyte growth factor/scatter factor isoforms. Heparan sulfate glycosaminoglycan requirement for Met binding and signaling; *J. Biol. Chem*. **272** 9457–9463

Schiessler H, Fink E and Fritz H 1976 Acid-stable proteinase inhibitors from human seminal plasma; *Methods Enzymol*. **45** 847–859

Schwall R H, Chang L Y, Godowski P J, Kahn D W, Hillan K J, Bauer K D and Zioncheck T F 1996 Heparin induces dimerization and confers proliferative activity onto the hepatocyte growth factor antagonists NK1 and NK2; *J. Cell Biol.* **133** 709–718

Suzuki K, Kise H, Nishioka J and Hayashi 2007 The interaction among protein C inhibitor, prostate-specific antigen, and the semenogelin system; *Semin. Thromb. Hemost.* **33** 46–52

Thaler C J, Vanderpuye O A, McIntyre J A and Faulk W P 1990 Lactoferrin binding molecules in human seminal plasma; *Biol. Reprod.* **43** 712–717

Urrutia R 2003 KRAB-containing zinc-finger repressor proteins; *Genome Biol*. **4** 231

Varilova T, Semenkova H, Horak P, Madera M, Pacakova V, Ticha M and Stulik K 2006 Affinity liquid chromatography and capillary electrophoresis of seminal plasma proteins; *J. Sep. Sci*. **29** 1110–1115

Veselsky L, Jonakova V and Cechova D 1985 A Kunitz type of proteinase inhibitor isolated from boar seminal vesicle fluid; *Andrologia* **17** 352–358

Walker A and Gallagher J T 1996 Structural domains of heparan sulphate for specific recognition of the C-terminal heparin-binding domain of human plasma fibronectin (HEPII); *Biochem. J.* **317** 871–877

Wassarman P M 1999 Mammalian fertilization: molecular aspects of gamete adhesion, exocytosis, and fusion; *Cell* **96** 175–183

Yi Y J, Manandhar G, Oko R J, Breed W G and Sutovsky P 2007 Mechanism of sperm–zona pellucida penetration during mammalian fertilization: 26S proteasome as a candidate egg coat lysine; *Soc. Reprod. Fertil. Suppl*. **63** 385–408

Zheng X, Geiger M, Ecke S, Bielek E, Donner P, Eberspacher U, Schleuning W D and Binder B R 1994 Inhibition of acrosin by protein C inhibitor and localization of protein C inhibitor to spermatozoa; *Am. J. Physiol*. **267** C466–472

MS received 8 July 2009; accepted 15 October 2009

*e*Publication: 27 November 2009

Corresponding editor: SUBHASH C LAKHOTIA