Semenogelin I and II, the predominant human seminal plasma proteins, are also expressed in non-genital tissues

Åke Lundwall[1,3], Anders Bjartell[2], A.Yvonne Olsson[1] and Johan Malm[1]

Departments of [1]Laboratory Medicine and [2]Urology, Lund University, University Hospital MAS, S-205 02 Malmö, Sweden

[3]To whom correspondence should be addressed. E-mail: ake.lundwall@klkemi.mas.lu.se

Semenogelin I (SgI) and semenogelin II (SgII) are the dominating protein components of the coagulum formed by freshly ejaculated human semen. The primary source of these proteins is the seminal vesicles and, apart from a small production of SgII in epididymis, they have not been detected in other tissues. In this report, we have re-examined the distribution of SgI and SgII transcripts and protein by RT–PCR and immunohistochemistry. Both SgI and SgII transcripts were demonstrated in several tissues, with the strongest signals coming from seminal vesicles, vas deferens, prostate, epididymis and trachea. Transcripts in the gastro-intestinal tract and skeletal muscle almost exclusively encoded SgI, whereas in kidney and testis, SgII transcripts were predominant. By immunohistochemistry, the basal cell layer of the secretory epithelium in prostate, trachea and bronchi was stained by antibodies recognizing both SgI and SgII. This is in contrast to the seminal vesicle and vas deferens, where the luminal cells were stained. The staining of skeletal muscle cells and a few scattered cells in the central nervous system suggests that semenogelin expression is not restricted to epithelial cells.

Key words: coagulation/semen/semenogelin/seminal plasma/transglutaminase

Introduction

Semenogelin I (SgI) and semenogelin II (SgII) are dominating proteins of human seminal plasma that together with fibronectin give rise to the gel-like coagulum of newly ejaculated semen (Lilja, 1985; Malm *et al.*, 1996). They both originate from the glandular epithelium of the seminal vesicles, which secrete them at very high concentrations. At lower concentrations, SgII is also secreted by the epithelium of the epididymis (Bjartell *et al.*, 1996).

SgI is a non-glycosylated protein of 439 aa residues with a molecular mass of 50 kDa (Lilja *et al.*, 1989). The SgII molecule of 63 kDa consists of 559 aa residues with a primary structure that is 78% similar to that of SgI (Lilja and Lundwall, 1992). It has one potential site for N-linked glycosylation and around half of the molecules in seminal plasma are glycosylated, yielding two molecular species with an apparent mass difference of 5 kDa (Lilja and Laurell, 1985). The repeated structure with motifs abundant in Gln, Ser, Gly and Lys residues is characteristic to both SgI and SgII and most conspicuous are highly conserved repeats of 60 aa residues in the central to C-terminal part of the molecules (Lilja and Lundwall, 1992). In the SgI molecule, there are two such repeats and in the SgII molecule there are four. The two extra repeats in SgII account for the size difference between the two semenogelin molecules. By way of the single Cys residue present in SgI and the two Cys residues in SgII, the molecules are present as covalent homo- and hetero-multimers (Lilja and Laurell, 1985). In ejaculated sperm, the molecules are also held together by non-covalent forces to yield a gel-like structure that can be dissolved by 4 mol/l urea and high pH in the presence of reducing agents (Malm *et al.*, 1996). The gel structure also dissolves spontaneously within minutes after ejaculation as the result of proteolytic degradation of the semenogelin molecules by

prostate-specific antigen (PSA) (Lilja, 1985). Both SgI and SgII have been shown to function as excellent substrates of transglutaminase and the repeated structure with an abundance of Gln and Lys is probably a major explanation for this property (Peter *et al.*, 1998).

Phylogenetic studies show that the genes encoding semenogelins and related seminal vesicle-secreted proteins have undergone dramatic changes during evolution. This is illustrated by the structural differences between the semenogelins and the predominant proteins expressed by rodent seminal vesicles (Lundwall and Lazure, 1995). There is a very limited similarity in primary structure between the semenogelins and the major proteins secreted by the rat and mouse seminal vesicles, even though they originate from homologous genes. This is because of the unusual evolution of otherwise ordinarily conserved genes. The major part of the secreted protein is encoded by a single exon and the structural heterogeneity is caused by either a differing selection of splice site or by an internal expansion of the exon during evolution (Lundwall and Lazure, 1995). The gene coding for the major clot protein of guinea-pig semen is an interesting example of the former (Hagstrom *et al.*, 1996). The upstream promoter region as well as the first exon and first intron is similar in sequence to the two human semenogelin genes. Furthermore, the first intron of the guinea-pig gene also carries ~0.5 kb that are homologous to the protein coding nucleotides in the beginning of the second exon of the semenogelin genes. It is therefore very likely that a relatively recent ancestor of the guinea-pig expressed a primitive semenogelin-like molecule that during evolution was omitted through the selection of new splice sites. The high sequence similarity in the 5′ end of the genes, including the upstream promoter region, and a mutual high expression level in seminal vesicles suggests that the guinea-pig gene might be regulated in a similar manner to the two semenogelin genes. Because the transcription of the guinea-pig gene has been reported

Table I. PCR primers. Nucleotides sequence of primers used for RT–PCR on transcripts of semenogelin I (SgI), semenogelin II (SgII) and adenine phosphoribosyltransferase (APRT).

Primer name	Primer sequence	Product size (bp)
5′ SgI primer	5′-GCAGACACCAACATGGATCTCA-3′	180
3′ SgI primer	5′-CTGAGGTCAACTGACACCTTGA-3′	
5′ SgII primer	5′-AGCATGAGGTTGCCCAAGATGA-3′	130
3′ SgII primer	5′-GAGGTCGGGTGACACCTTGC-3′	
5′ APRT primer	5′-GACTACATCGCAGGCCTAGAC-3′	257
3′ APRT primer	5′-GCAGCGTTCATGGTTCCACCA-3′	

to occur not only in seminal vesicles, but also in testis, kidney, liver and lung (Hagstrom *et al.*, 1996), it is not unlikely that the human semenogelin genes are expressed in more tissues than previously anticipated. Studies have indicated a role of semenogelin molecules in relation to capacitation and motility of sperm (Robert and Gagnon, 1996; de Lamirande *et al.*, 2001). A non-genital expression of SgI and SgII would suggest that the molecule also has functions that are unrelated to sperm. To investigate this, semenogelin transcripts and immunoreactivity were analysed in a collection of tissue samples.

Materials and methods

Tissue and semen samples

The Helsinki Declaration regarding the use of human tissues was followed. For immunohistochemistry, archival tissues obtained at operations were used. All tissues were fixed within 30 min after removal in Bouin's fixative (for 4–18 h) or 4% buffered paraformaldehyde (overnight). All tissues were histopathologically normal according to haematoxylin–eosin staining.

RNA samples for RT–PCR were isolated from tissue specimens homogenized in 4 mol/l guanidiniumthiocyanate as previously described (Chomczynski and Sacchi, 1987). RNA preparations were made from single tissue specimens, except from epididymis and kidney where respectively five and four preparations were made. Tissue samples from the urogenital tract and mammary glands were from patients undergoing surgical treatment for neoplastic disease. The other specimens were taken at autopsy ~20 h post-mortem. RNA from human lung, pancreas, salivary gland, skeletal muscle and trachea were purchased from Clontech (Palo Alto, CA, USA). RNA samples were stored at –70°C.

Semen samples were collected from voluntary donors by masturbation. Liquefaction was achieved by incubation for 60–90 min at room temperature. Cells were removed by centrifugation at 800 *g* for 15 min. Unliquefied samples were collected in 40 mmol/l Tris–HCl, pH 9.7, 4 mol/l urea, 25 mmol/l EDTA and 30 mmol/l dithiothreitol as previously described (Malm *et al.*, 1996). All semen samples were stored at –20°C.

RT–PCR

Oligo-dT-primed cDNA synthesis, using 3 µg of total RNA, was performed in a volume of 15 µl with the First-Strand cDNA synthesis kit (Amersham Pharmacia Biotech, Uppsala, Sweden). The subsequent PCR was performed with 2 µl cDNA, equivalent to ~0.4 µg of RNA, in a volume of 10 µl using the Advantage 2 PCR kit (Clontech) and 0.2 µmol/l gene-specific primers (Table I). The PCR was run in a DNA Engine, PTC-200 (MJ Research, Watertown, MA, USA) with a program that had an initial 1 min denaturation at 95°C followed by 35 cycles consisting of 30 s at 95°C and 1 min at 68°C. At the end of the program, there was also an extra 1 min incubation at 68°C. Reaction products were analysed by electrophoresis in 2.5% (w/v) agarose gels and visualized by exposure to UV light following staining with ethidium bromide. Each sample was analysed at least three times for both SgI and SgII: once with primers for a single gene and twice with primers for both genes.

Immunological methods

Antiserum against SgII was raised in rabbits by standard procedures (Harlow and Lane, 1988) using protein purified from unliquefied semen (Malm *et al.*,

1996). Immunohistochemistry was performed using a detection kit, Dako ChemMate™ Detection Kit peroxidase/carbazole, rabbit/mouse and a staining machine, Dako TechMate™ 500/1000 Instrument (Dako A/S, Denmark). Briefly, the sections were deparaffinized in xylene, rehydrated and treated with 0.3% H_2O_2 in methanol for 30 min at room temperature to quench endogenous peroxidase activity. For antigen retrieval, tissue sections were first incubated with sodium citrate (10 mmol/l, pH 6.0) and boiled in a microwave oven at 750 W for 2×3 min, and then digested with proteinase K (20 µg/ml in 20 mmol/l Tris–HCl, 2 mmol/l $CaCl_2$, pH 7.5) for 25 min at 37°C. The sections were incubated with the antiserum against SgII (diluted 1:4000) for 60 min at room temperature, after which they were incubated with biotinylated secondary antibodies against rabbit IgG, which were included in the ChemMate kit, for 60 min at room temperature. The immunoreactivity was visualized using the manufacturer's protocol for the peroxidase/AEC (3-amino-9-ethyl-carbazole) or DAB (3,3′diaminobenzidine tetrahydrochloride) reagents in the ChemMate kit. The sections were counterstained with Mayer's haematoxylin solution. As a negative control, adjacent tissue sections were processed by replacing the primary antibody with non-immune rabbit IgG diluted 1:8000 (Dako). At least three samples were analysed from each tissue.

Results

Detection of SgI and SgII transcripts

An RT–PCR assay was developed for easy and swift detection of semenogelin transcripts in different tissues. PCR primers were selected from cDNA sequences that showed low conservation between SgI and SgII. In order to reduce false signals by priming on genomic DNA present in small quantities in RNA preparations, the 3′ primers were designed to span the second intron of the genes. The transcript of the housekeeping gene encoding adenine phosphoribosyltransferase (APRT) served as control of cDNA synthesis and PCR (Hidaka *et al.*, 1987). Also with this transcript, false priming was avoided by selecting intron-spanning primers; in this case the 5′ and 3′ PCR primers were designed to span introns 2 and 4 of the ARPT gene. The assay was optimized using RNA from seminal vesicles and epididymis, and distinct APRT, SgI and SgII signals were obtained using the protocol with 35 PCR cycles. By running the assay with samples of cloned cDNA of known concentration, the detection limit was estimated to be ~1 fg of transcript.

Figure 1 shows the PCR products run on agarose gel stained with ethidium bromide. The 257 bp APRT product is seen with samples from all tissues as expected. Several samples also give rise to strong signals for both the 180 bp SgI product and the 130 bp SgII product. In the gastro-intestinal tract there are mainly SgI transcripts, whereas in kidney, the SgII transcripts dominate. However, the relative abundance of SgI to SgII transcripts varied in different kidney samples and in some the SgI signal was as equally strong as the SgII signal. This did not depend on the anatomical location as the relative signal intensity was the same whether samples were taken from the cortex or the medulla, although the intensity was consistently stronger in medulla samples. Several epididymis specimens were analysed and invariably both SgI and SgII signals were found. However, as with the kidney samples, the staining intensity of epididymis samples varied so that the signal from some preparations yielded a stronger SgII than SgI signal and vice versa. The SgI signal in epididymis was a surprise as results from previous investigations (Bjartell *et al.*, 1996) suggested that only the SgII gene is transcribed in this tissue, with a restricted distribution to the caudal part. Because cauda epididymis empties into vas deferens, there is a possibility that epididymal RNA preparations were contaminated with RNA from vas deferens. From the close relationship of the epithelium of vas deferens and seminal vesicles it is reasonable to assume
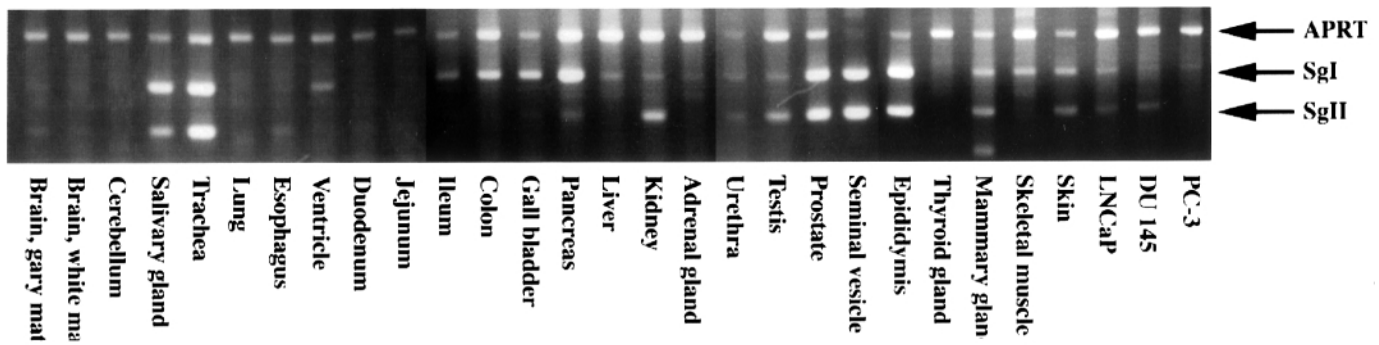
Figure 1. Expression of semenogelin (Sg)I and SgII mRNA. Transcripts were identified by RT–PCR and electrophoresis. The arrows to the right indicate the location of PCR products of adenine phosphoribosyltransferase (ARPT), SgI and SgII transcripts. The source of RNA is indicated below each lane.
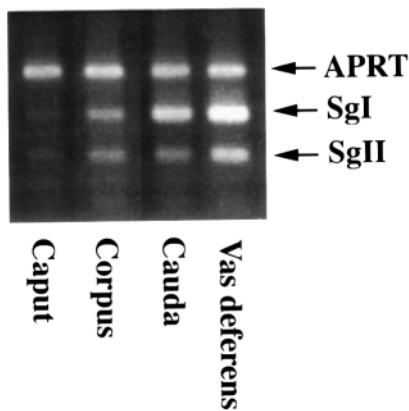

Figure 2. Transcripts in epididymis. Epididymis was dissected into its constituent parts from which RNA was extracted. Transcripts were amplified by RT–PCR and analysed by electrophoresis. Transcripts of adenine phosphoribosyltransferase (ARPT), semenogelin I (SgI) and semenogelin II (SgII) are indicated by the arrows, with the RNA source is indicated below each lane.

Figure 3. Western blotting. Samples of 10 µl separated by SDS–PAGE and analysed by Western blotting using an antiserum raised against human semenogelin (Sg)II, but which also recognizes SgI. Samples are unliquefied human seminal plasma diluted 1:70 (lane1), liquefied human seminal plasma diluted 1:10 (lane 2) and human blood plasma diluted 1:10 (lane 3). Lane 4 show size markers stained by Coomassie Brilliant Blue, with sizes in kDa indicated to the right. The arrows indicate the location of SgI and SgII.

that vas deferens express SgI and SgII. Thus, the SgI signal in epididymis might originate from vas deferens contamination. This was investigated by dissection of epididymis into its constituent parts which were subsequently analysed by RT–PCR. As can be seen in Figure 2, both vas deferens and cauda epididymis show strong SgI and SgII signals, but weaker and very faint signals for both transcripts are also seen in corpus and caput epididymis respectively. As it is unlikely that the signals from corpus and caput epididymis are the result of vas deferens contamination, both the SgI and SgII genes are probably transcribed in epididymis.

Because of the strong signal from prostate RNA, transcripts were also analysed in three prostate cancer cell lines. No, or very few, transcripts were detected in RNA from the hormone-insensitive cell line PC-3, but from the likewise hormone-insensitive DU 145 cell line, a faint SgII signal was detected. In the hormone-sensitive cell line LNCaP both SgI and SgII transcripts were detected, but compared with prostate tissue, the signal from cell lines was weak. This might indicate that the semenogelin genes are down-regulated during malignant transformation or that stromal cells instead of epithelial cells are the main source of prostate SgI and SgII.

Immunohistochemical localization of semenogelin antigenicity

The specificity of the antiserum against SgII was analysed by Western blotting (Figure 3). As can be seen, it could recognize both SgI and SgII as well as their degradation products. The equally strong immunostaining of SgI and SgII suggests that the
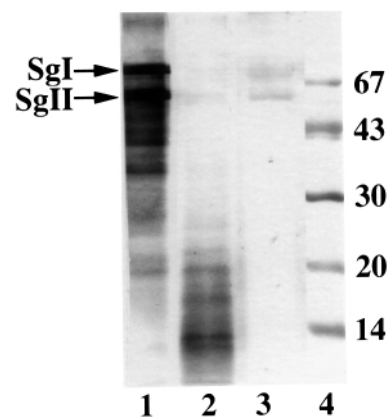
antiserum is unable to distinguish between SgI and SgII. Semenogelin immunoreactivity was thereafter analysed in tissue sections from a variety of organs. Tissue specimens from the gastrointestinal tract, with moderate amounts of SgI transcripts according to RT–PCR, showed no specific semenogelin signal, as the weak staining generated by immune sera did not significantly exceed background staining with non-immune sera. In contrast to this are the discrete semenogelin signals seen in neural tissues, despite negative RT–PCR results. The immunostaining appeared to be specific for a discrete cell type that is present both in medulla oblongata and cerebellum (Figure 4A–C). The seminal vesicle secretion contains very high concentrations of both SgI and SgII. In tissue sections of the seminal vesicles this appeared as an intense immunostaining of the secretory epithelium. An equally intense staining was seen in vas deferens and ampulla vas, indicating high semenogelin concentration also in secretion from the epithelium of the spermatic ducts (Figure 4D,E). A much weaker immunostaining was observed in breast tissue, prostate, trachea, bronchi, kidney and skeletal muscle (Figure 4F–L). In skeletal muscle the immunostaining appeared to be specific for myocytes, whereas in breast tissue the epithelium lining the mammary ducts was stained and in the kidneys the proximal tubuli were stained. The experiments with prostate, trachea and bronchi were very interesting as they all appeared to show immunostaining in the basal cell layer and not in the luminal layer, as in the seminal vesicles. Because of the result
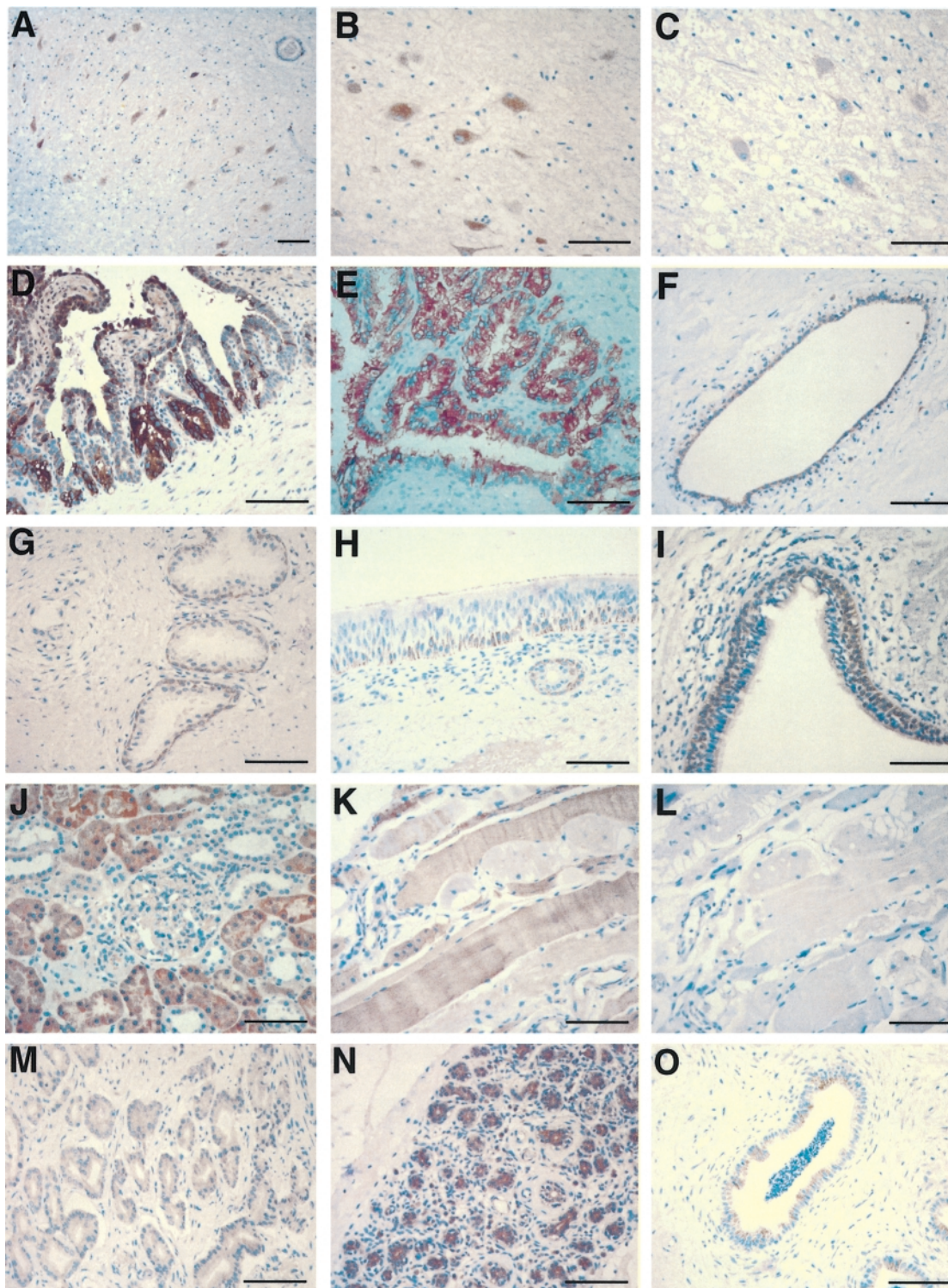
Figure 4. Immunohistochemistry. Demonstration of semenogelin (Sg) immunoreactivity in tissue sections using rabbit antiserum which was raised against SgII, but which also recognizes SgI, and secondary horse radish peroxidase-labelled goat anti-rabbit IgG antibodies. Immunostaining was developed with 3-amino-9-ethyl-carbazole (red) or 3,3′diaminobenzidine tetrahydrochloride (brown) and the tissue sections were counterstained with Mayer's haematoxylin solution. The scale bars are 40 μm. Perinuclear immunostaining of neurons in the central nervous system in (**A**) cerebellum and (**B**) medulla oblongata. (**C**) Negative control of immunostaining in medulla oblongata is demonstrated by substitution of immune serum with normal rabbit serum. Intense immunostaining in (**D**) seminal vesicle and (**E**) ampulla vas deferens. Moderate immunostaining in epithelium of (**F**) mammary duct and in the basal cell layer of (**G**) prostate, (**H**) trachea and (**I**) bronchi. Immunostaining of (**J**) proximal tubuli in kidney, (**K**) myocytes in skeletal muscle and (**L**) negative control of the latter using non-immune serum. Expression of semenogelin in cancer as demonstrated by immunostaining of (**M**) prostate cancer and (**N**) mammary cancer. (**O**) Immunostaining of the secretory epithelium in epididymis.

with normal tissues, expression was also analysed in prostate and breast cancer. As can be seen, semenogelin immunoreactivity persists in both prostate and breast cancer (Figure 4M,N). Immunostaining in epididymis is displayed as a positive control of the antiserum's performance on a tissue expressing moderate levels of semenogelin molecules (Figure 4O).

Discussion

In this study, we showed that semenogelin transcripts and antigenicity are present in several tissues and not confined to the seminal vesicles and epididymis as was previously assumed. Following ejaculation, SgI and SgII are degraded within minutes by PSA during the process known as semen liquefaction (Lilja, 1985). Prolonged incubation leads to the formation of very small fragments and a progressive loss of semenogelin immunoreactivity. Also, other proteases readily degrade the semenogelin molecules and therefore it might be assumed that their general structure renders them very sensitive to proteolytic digestion (Deperthes *et al.*, 1996). Because of this, proteolytic degradation could be one explanation for the failure in previous reports to detect semenogelin antigenicity in tissues other than seminal vesicles and epididymis. At least in the seminal vesicles, the expression levels of the semenogelins are extremely high so even if most of the semenogelin molecules are degraded, there will still be substantial quantities of antigen left over to be detected by the anti-semenogelin antibodies. The improved fixation procedure in combination with microwave radiation to expose antigen in this study has probably also increased the sensitivity.

It has previously been shown by in-situ hybridization that only SgII is synthesized in the epididymis (Bjartell *et al.*, 1996). This is in contrast to this report where both SgI and SgII transcripts were detected. The discrepancy can probably be explained, at least partly, by sample heterogeneity. As the epididymal ratio of SgI to SgII transcripts appears to vary significantly between preparations, the previous investigation could have been conducted on samples with high levels of SgII transcripts and low levels of SgI transcripts, whereby the weak signal from the latter could have been ignored. It is not clear what governs the variation in SgI to SgII transcript ratios, but it appears to predominantly reflect variations in SgI transcript levels. Thus, SgI transcript levels might be prone to fluctuations, something that could also be a factor behind the failure to detect semenogelin immunoreactivity in the gastro-intestinal tract. The expression of the SgI gene could, for instance, have been different in samples taken for RT–PCR compared to those taken for immuno-histochemistry. This can probably be addressed by simultaneous analyses of transcripts and immunoreactivity in larger samples. Another indication that the SgI gene is differently regulated to the SgII gene was the discrepant expression in the hormone-sensitive cell line LNCaP and the hormone-insensitive cell line DU 145. The transcription of both the SgI gene and the SgII gene in LNCaP cells versus almost sole transcription of the SgII gene in DU 145 cells suggests that the SgI gene is more sensitive to hormone stimulation.

UniGene (http://www.ncbi.nlm.nih.gov/UniGene/) is a database which contain clusters of expressed sequence tags (EST) that enables identification of tissues in which a specific gene is transcribed. In November 2001, the database contained 241 EST sequences encoding SgI and 84 EST sequences encoding SgII. Of the SgI transcripts, 226 originated from prostate, 11 from skeletal muscle, two from breast tissue and two from kidney, and of the SgII transcripts, 78 originated from prostate, six from skeletal muscle and one each from colon and pooled colon, kidney and stomach RNA. The EST sequences confirm the here reported expression of both SgI and SgII in the prostate. The rest of the EST sequences also agree with our data, except for those suggesting SgII transcription in skeletal muscle and colon. The number of identified EST sequences in skeletal muscle exclude the possibility that they represent cryptic transcripts. More likely they represent transcripts from discrete subsets of cells that might differ in number between different parts of the tissue. A similar reasoning might also explain the negative RT–PCR in neural tissue in spite of the immunostaining of certain cells in medulla oblongata and cerebellum.

SgI has been reported to be an inhibitor of sperm motility and capacitation (Robert and Gagnon, 1996; de Lamirande *et al.*, 2001). These activities have not been independently verified, so the biological role of the semenogelins remains somewhat unclear. The here reported finding of both SgI and SgII transcripts and antigen in a variety of tissues suggest that the molecules have a function that is more far reaching than previously believed and not confined to a role connected with male fertility. In fact, it might very well be that the seminal vesicle synthesis should be regarded as a special case, where the protein is over expressed so as to exert a specialized function in the male genital tract, while the function in other tissues is of more general importance. In this context it is interesting to note that both SgI and SgII were recently found in cell surface adhesion complexes of small cell lung carcinoma cell lines (Rodrigues *et al.*, 2001). The molecules could therefore very well turn out to be adhesion molecules.

The expression in lung cancers indicates that the semenogelin molecules might serve as diagnostic and/or prognostic tumour markers (Rodrigues *et al.*, 2001). In this report we also show semenogelin expression in prostate and mammary cancer. Presently we are assessing the utility of semenogelin immunostaining as a prognostic tool in these cancers.

Acknowledgements

References

Bjartell, A., Malm, J., Möller, C., Gunnarsson, M., Lundwall, Å. and Lilja, H. (1996) Distribution and tissue expression of semenogelin I and II in man as demonstrated by *in situ* hybridization and immunocytochemistry. *J. Androl.*, **17**, 17–26.

Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*, **162**, 156–159.

de Lamirande, E., Yoshida, K., Yoshiike, T.M., Iwamoto, T. and Gagnon, C. (2001) Semenogelin, the main protein of semen coagulum, inhibits human sperm capacitation by interfering with the superoxide anion generated during this process. *J. Androl.*, **22**, 672–679.

Deperthes, D., Frenette, G., Brillard-Bourdet, M., Bourgeois, L., Gauthier, F., Tremblay, R.R. and Dube, J.Y. (1996) Potential involvement of kallikrein hK2 in the hydrolysis of the human seminal vesicle proteins after ejaculation. *J. Androl.*, **17**, 659–665.

Hagstrom, J.E., Fautsch, M.P., Perdok, M., Vrabel, A. and Wieben, E.D. (1996) Exons lost and found. Unusual evolution of a seminal vesicle transglutaminase substrate. *J. Biol. Chem.*, **271**, 21114–21119.

Harlow, E. and Lane, D. (eds) (1988) *Antibodies. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, p. 92.

Hidaka, Y., Tarle, S.A., O'Toole, T.E., Kelley, W.N. and Palella, T.D. (1987) Nucleotide sequence of the human APRT gene. *Nucleic Acids Res.*, **15**, 9086.

Lilja, H. (1985) A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. *J. Clin. Invest.*, **76**, 1899–1903.

Lilja, H. and Laurell, C.-B. (1985) The predominant protein in human seminal coagulate. *Scand. J. Clin. Lab. Invest.*, **45**, 635–641.

Lilja, H. and Lundwall, A. (1992) Molecular cloning of epididymal and seminal vesicular transcripts encoding a semenogelin-related protein. *Proc. Natl Acad. Sci. USA*, **89**, 4559–4563.

Lilja, H., Abrahamsson, P.A. and Lundwall, A. (1989) Semenogelin, the predominant protein in human semen. Primary structure and identification of closely related proteins in the male accessory sex glands and on the spermatozoa. *J. Biol. Chem.*, **264**, 1894–1900.

Lundwall, A. and Lazure, C. (1995) A novel gene family encoding proteins with highly differing structure because of a rapidly evolving exon. *FEBS Lett.*, **374**, 53–56.

Malm, J., Hellman, J., Magnusson, H., Laurell, C.B. and Lilja, H. (1996) Isolation and characterization of the major gel proteins in human semen, semenogelin I and semenogelin II. *Eur. J. Biochem.*, **238**, 48–53.

Peter, A., Lilja, H., Lundwall, A. and Malm, J. (1998) Semenogelin I and semenogelin II, the major gel-forming proteins in human semen, are substrates for transglutaminase. *Eur. J. Biochem.*, **252**, 216–221.

Robert, M. and Gagnon, C. (1996) Purification and characterization of the active precursor of a human sperm motility inhibitor secreted by the seminal vesicles: identity with semenogelin. *Biol. Reprod.*, **55**, 813–821.

Rodrigues, R.G., Panizo-Santos, A., Cashel, J.A., Krutzsch, H.C., Merino, M.J. and Roberts, D.D. (2001) Semenogelins are ectopically expressed in small cell lung carcinoma. *Clin. Cancer Res.*, **7**, 854–860.