A frequent allele codes for a truncated variant of semenogelin I, the major protein component of human semen coagulum

Åke Lundwall1,3, Aleksander Giwercman2, Yasir Ruhayel2, Yvonne Giwercman2, Hans Lilja1, Christer Halfdén1 and Johan Malm1

1Department of Laboratory Medicine and 2Department of Urology, Lund University, University Hospital MAS, S-205 02 Malmö, Sweden
3To whom correspondence should be addressed. E-mail: ake.lundwall@klkemi.mas.lu.se

Human semen coagulum predominantly consists of high molecular mass complexes of the seminal vesicle secreted semenogelin I (SgI) and semenogelin II (SgII). Here we describe a previously unknown variant of the SgI gene that is present at an allele frequency of ~3% in the Swedish population. It gives rise to a protein with a molecular mass of 43 kDa, SgI43, which compared with the 50 kDa variant, SgI50, is lacking a tandem repeat of 60 amino acid residues that was probably deleted by homologous recombination. In spite of the size difference, SgI43 has many properties in common with SgI50, such as a very high isoelectric point and susceptibility to proteolytic degradation by prostate-specific antigen. Heterozygous carriers of the SgI43 allele neither show impaired fertility nor do they significantly differ from individuals homozygous for SgI50 with respect to sperm parameters such as semen volume, sperm count and fraction of motile spermatozoa.

Key words: allele/coagulum/semen semenogelin I/II/transglutaminase/truncated

Introduction

Semen consists of spermatozoa and seminal plasma provided by the accessory sex glands. In man, spermatozoa contribute only a few percent to the semen volume, while the major part is provided by secretions of the seminal vesicles and the prostate—around 2/3 and 1/3 of the volume respectively. At ejaculation, mixing of spermatozoa-rich epididymal fluid with fluids from accessory sex glands results in the formation of a loose, gel-like coagulum, which subsequently will liquefy. Normally the coagulum is completely dissolved within 20 min.

Major protein components of the human semen coagulum are the seminal vesicle-secreted semenogelin I (SgI) and semenogelin II (SgII), which are present as high molecular mass complexes in seminal vesicle fluid as a result of strong non-covalent forces and disulphide bridges (Lilja and Laurell, 1985). The complexes are disintegrated, by 4 mol/l urea in the presence of reducing agents at high pH, to yield three major molecular species (Malm et al., 1996). These are SgI of 50 kDa, SgII of 63 kDa, and a glycosylated form of SgII that is ~5 kDa larger than the parent molecule (Lilja, 1989; Lilja and Lundwall, 1992). The high molecular mass complexes are also disintegrated during semen liquefaction, due to proteolytic degradation of SgI and SgII by prostate-specific antigen (PSA) (Lilja, 1985). The semenogelin molecules are 78% similar in primary structure, which to a large extent consists of tandem repeats that show varying degrees of conservation. In the C-terminal half, the molecules carry very conserved tandem repeats of 60 amino acid residues (Lilja et al., 1989; Lilja and Lundwall, 1992). There are two such repeats in SgI and four in SgII; the difference of two repeats accounts for the size difference between the two polypeptide chains.

There are reports suggesting that SgI, and perhaps also SgII, is an inhibitor of sperm motility and capacitation (Robert and Gagnon, 1996; de Lamirande et al., 2001). However, it has also been reported that there is no correlation between semenogelin levels in seminal plasma and fertilization capacity of spermatozoa in vitro (Koistinen et al., 2002). The structure, physiochemical properties and site of synthesis, suggest that the semenogelin molecules might function to protect spermatozoa from harsh conditions at vaginal deposition and also facilitate transport in the genital tract of both the male and the female. Recent studies have shown expression of both SgI and SgII outside of the male genital tract, suggesting that the molecules might have functions of broader biological significance and not just confined to male fertility (Lundwall et al., 2002).

The genes of SgI, SEMG1, and SgII, SEMG2, have been assigned to the cytogenetic region q12-13.1 on the long arm of human chromosome 20 (Ulsväck et al., 1992). The two genes were formed by duplication of an 8 kb region and are now separated by 11.6 kb intergenic DNA (Lundwall, 1996b). Both genes consist of three exons. The first exon encodes the signal peptide and the two N-terminal residues of the mature protein (Ulsväck et al., 1992). A large second exon contains the remaining coding nucleotides, the stop-codon and ~20% of the 3’ non-translated nucleotides. The last exon, encompassing the rest of the 3’ non-translated nucleotides, carries the poly-adenylation signal. Phylogenetic studies show that the second exon of the semenogelin genes have evolved in an unusual way. This was first seen in a comparison between the human semenogelin genes and genes of the major coagulating protein in mouse and rat semen, semenoclotin and SVS II (Lundwall and Lazure, 1995; Lundwall, 1996a). Despite major differences in primary structure of their protein products, these genes...
are homologous. The explanation of this unusual situation is that the second exon has expanded by way of repeated tandem duplications and replication slippage after the separation of the primate and murine lineage. This subsequently translated into highly differing protein structures as coding nucleotides were affected. In contrast, the first and the last exon of the genes show normal conservation, as do introns and the very ends of the second exon. Other studies have later shown that the expansion of the second exon is also present within the primate lineage. For instance, SgI of the cotton-top tamarin is 32% larger than the human orthologue due to duplications of 174 bp in the second exon that give rise to five tandem repeats that are unique to the tamarins (Lundwall, 1998). Similarly, a duplication of 360 bp in the rhesus monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Ulvsbäck and Lundwall, 1997). Thus, the monkey SgII gene yielded a molecule that was substantially larger (Lundwall, 1998). Similarly, a duplication of 360 bp in the rhesus monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Ulvsbäck and Lundwall, 1997). Thus, the monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Lundwall, 1998). Similarly, a duplication of 360 bp in the rhesus monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Lundwall, 1998). Similarly, a duplication of 360 bp in the rhesus monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Lundwall, 1998). Similarly, a duplication of 360 bp in the rhesus monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Lundwall, 1998). Similarly, a duplication of 360 bp in the rhesus monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Lundwall, 1998). Similarly, a duplication of 360 bp in the rhesus monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Lundwall, 1998). Similarly, a duplication of 360 bp in the rhesus monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Lundwall, 1998). Similarly, a duplication of 360 bp in the rhesus monkey SgII gene yielded a molecule that was substantially larger than its human counterpart (Lundwall, 1998).

**Materials and methods**

**Ethics**

Written consents were given by voluntary donors who provided blood and semen samples for these studies. The investigation was approved by the Research Ethics Committee at Lund University (LU 385-99).

**Genetic analyses**

DNA was isolated from blood using the Qiagen Genomic-tip system or the QIAamp 96 DNA Blood kit (VWR International, Lund, Sweden). Southern blot analysis was performed essentially as previously described (Lundwall, 1998). A BP1025-probe, which carries the central part of the SgII transcript and recognizes both the SgI gene and the SgII gene, was labelled with [32P]dCTP to a specific activity exceeding 109 d.p.m./μg using the Megaprime DNA labelling system (Amersham Biosciences, Uppsala, Sweden). DNA fragments, corresponding to the EcoRI fragment holding most of the SgI gene, were generated by PCR in a MJ Research PTC-200 peltier thermal cycler (MJ Research, Watertown, MA) using Advantage 2 DNA polymerase mix (Clontech, BD Biosciences, Stockholm, Sweden) according to the supplier’s recommended standard protocol for a volume of 50 μl. The template was 100 ng of DNA and the primers were 5’-CCCTAGCTTCATCCACCAGTGAATCAGCTT-3’ overlapping the EcoRI site located 0.7 kb upstream of the SgI gene and 5’-AAGATAGGAATTCTAGGACCTCATGATAGGCCA-3’ overlapping the EcoRI site 0.2 kb downstream of exon 2. The PCR program consisted of an initial denaturation step at 95°C for 1 min followed by 35 cycles of 95°C denaturation for 30 s and 68°C annealing and extension for 3 min. The conserved repeat region of SgII was amplified using the same protocol with the annealing and extension time shortened to 1 min. The two oligonucleotides 5’-ACTACAAACTTCACTCCATGACC-3’ and 5’-GGGCTATGTAGATATTGGATTTCTGTCT-3’ served as primers. Another PCR protocol was devised for screening of SEMG1 alleles. In this case, the PCR was performed with 0.5 IU of AmpliTaq Gold (Applied Biosystems, Stockholm Sweden) on 10 ng of DNA in 10 mmol/l Tris–HCl pH 8.3, 50 mmol/l KCl, 3 mmol/l MgCl2, 0.2 mmol/l dNTP and 0.4 μmol/l of each primers with the following sequences 5’-TGCGCACCAAGACAAACTC-CAACATGGAT-3’ and 5’-TGACGATCACTGTCATCTTCCTGCTCTAT-3’. The PCR amplification had an initial activation step at 95°C for 5 min, which was followed by 30 cycles consisting of denaturation at 95°C for 30 s, annealing at 68°C for 30 s and synthesis at 72°C for 45 s. A final step consisted of an incubation at 72°C for 10 min. PCR products were analysed by electrophoresis in agarose gels and DNA fragments were visualized by staining with ethidium bromide at a concentration of 1 μg/ml. DNA sequencing was performed directly on PCR products using Big Dye DNA Sequencing kit (Applied Biosystems) and an Applied Biosystems 373 DNA sequencer upgraded to Big Dye Chemistry. Sequencing primers were synthetic oligonucleotides of 20 residues that were made in accordance with the structure of the SgI gene. The PCR primers used for amplification of the SgI gene were also used as sequencing primers. Nucleotide sequences were assembled and analysed using the GCG Wisconsin computer package (Accelrys Inc., San Diego, CA).

**Protein analyses**

Semen samples were obtained from voluntary donors by masturbation. Freshly delivered ejaculates were either allowed to liquefy at room temperature for 90 min, or directly collected in 40 mmol/l Tris–HCl, pH 9.7, 4 mmol/l urea, 25 mmol/l EDTA, 30 mmol/l diithiothreitol, 3 mmol/l benzamidine and 0.5 mmol/l Pefablock (Roche Applied Science) in order to prevent the liquefaction process as previously described (Malm *et al.*, 1996). SgI was purified by chromatography on heparin–Sepharose (Amersham Biosciences, Uppsala, Sweden) (Malm *et al.*, 1996). SDS–PAGE was run with reduced and alkylated samples in 12% polyacrylamide gels using the Mini PROTEAN II system (Bio-Rad Life Science, Sundbyberg, Sweden) (Laemmli, 1970). Western blot analysis was performed on Immobilon membranes (Millipore AB, Sundbyberg, Sweden) after transfer with a semi-dry electroblotting apparatus (Ancos, Højby, Denmark) as described previously (Malm *et al.*, 1996). The monoclonal antibody, Mab9, raised against SgI, that also recognizes SgII, has been described elsewhere (Bjartell *et al.*, 1996). Conditions for digestion of SgI with PSA have also been described in an earlier work (Malm *et al.*, 2000).

**Semen analysis**

Semen analysis was performed according to the World Health Organization’s most recent guidelines (World Health Organization, 1999). Semen samples were collected from 305 young Swedish military conscripts. They were asked to be abstinent for a period of 2–3 days prior to collecting the semen sample, but

![Figure 1. Identification of a novel semenogelin allele by Southern blot analysis. An EcoRI digest of 15 μg of human DNA was separated by electrophoresis in 0.7% agarose gel, transferred to nylon membrane, and probed with a DNA-fragment recognizing both SgI and SgII. The resulting autoradiogram is shown with sizes of molecular standards, given in kb, to the left. The lettering to the right indicates the source of hybridizing fragments. New = the novel semenogelin allele.](image-url)
in each case the time span of the actual abstinence period was recorded. Ejaculates were obtained by masturbation into a wide mouthed plastic cup. Semen volume was determined by weighing the plastic cup before and after ejaculation. One gram of semen was considered to be equivalent to 1 ml. The concentration of spermatozoa was assessed using an improved Neubauer haemocytometer. Samples were drawn using positive displacement pipettes. Sperm motility was analysed using CRISMAS Computer Aided Sperm Analysis system, as recently described (Elzanaty et al., 2002). Motile sperm were defined as those having curvilinear velocity above 25 \( \mu \text{m/s}. \) Statistical analysis was computed with the SPSS 11.0 software (SPSS Inc., Chicago, USA). Using a Mann–Whitney test for unpaired data, semen volume, sperm concentration and percentage of motile sperm were compared between the group of males who were heterozygous for SgI 43 and those homozygous for SgI 50. P-values below 0.05 were considered statistically significant.

### Results

#### A truncated SgI gene

Human genomic DNA was analysed by Southern blot hybridization using a probe that recognizes both the SgI gene and the SgII gene. In one sample, the probe detected an EcoRI fragment of 2.4 kb that had not previously been observed (Figure 1). As the novel EcoRI fragment was close in size to the fragment encompassing most of the SgI gene, it was hypothesized that the 2.4 kb fragment might represent a new SgI allele. To test this hypothesis, a PCR was devised using primers that overlap the EcoRI sites flanking the SgI gene. The PCR yielded two products; one of 2.6 kb that agrees with the expected size of the SgI gene, and a second that is ~0.2 kb smaller and agrees in size with the novel EcoRI fragment. The mixture of the two fragments was subjected to DNA sequencing. A single nucleotide sequence was obtained that extended from the 5’ end of the fragments to nucleotide 1199 of the SgI gene, where the sequencing produced equal yields of A and C nucleotide. Further downstream there were additional positions with double base calls, which suggested that there were two sequences—one from the previously described allele and one from an allele that is lacking a tandem repeat of 180 bp, present in the former (Figure 2). The same conclusion was drawn by sequencing from the 3’ end of the fragments. Also in this case, a single sequence was initially produced, but from position 1324 and upstream, certain positions gave rise to double base calls that were consistent with an allele truncated by 180 bp. Furthermore, based on positions where the aligned repeat sequences differ, but where the sequencing of the PCR products yielded single base calls, it was concluded that the new SgI variant was lacking 180 bp in the region between position 1145 and 1378 of the ‘normal’ SgI gene. This was also confirmed by sequencing of cloned DNA of the shorter allele generated by PCR. Presumably, this was caused by recombination between nucleotides 1145–1198 and 1325–1378 that are identical in sequence. At the protein level, this translates into a deletion of 60 amino acid residues between position 271 and 348 of the ‘normal’ SgI molecule (Figure 3). The molecular mass of the new allele product was calculated to be 42.8 kDa as compared with 49.6 kDa for the ‘normal’ protein: from here on the former product is called SgI 43 while the latter product is denoted SgI 50 in order to distinguish the two forms. The high isoelectric point, characteristic to the semenogelins, is not substantially affected by the deletion; there is a slight increase in pI from 10.07 for SgI 50 to 10.16 for SgI 43.

#### The novel allele is relatively common

In order to investigate whether the SgI 43 variant is unique or common in the population, DNA of 100 randomly selected individuals were analysed by PCR. The majority of samples yielded a single PCR product of 2.6 kb, but in five cases there was also an additional PCR product of 2.4 kb, that is indicative of the truncated SgI 43 allele. DNA sequencing of these PCR products revealed the same sequence pattern as in the original sample. Thus, all analysed subjects carry the same truncated allele. In order to facilitate the analyses of a larger number of samples, a new PCR procedure, with focus on the truncated region, was devised for allele detection (Figure 4). This method was subsequently used for analysis of DNA samples from another 305 voluntary donors. In this population, which consisted of young Swedish military conscripts, another 20 heterozygous carriers of SgI 43 were found. Therefore, 25 heterozygous carriers in all were found in a population of 405 individuals, equating to an allele frequency of 3.1% (95% CI, 2.0–4.5%).

An analysis was also made of the homologous DNA segment of the SgII gene; i.e. the region encompassing four highly conserved repeats of 180 bp. PCR amplification yielded a single product of ~1.0 kb in the 305 samples analysed. The size agrees with the predicted size of 1054 bp that was calculated from the nucleotide sequence of SgII cDNA. This means that the frequency of an allele with a variant SgII size is close to zero (95% CI, 0.0–0.6%).

#### A truncated SgI molecule in seminal plasma

Semen samples from an individual that is homozygous for the SgI 50 allele and a heterozygous carrier of the novel SgI 43 allele were

---

**Figure 2.** Extension of the deletion. An alignment of two conserved repeats of the gene encoding SgI 50 is shown with nucleotide positions indicated. Stars highlight positions where the repeat sequences disagree. Lines started by seq denote sequences obtained by subjecting a mixture, with PCR products of 2.6 and 2.4 kb, to DNA sequencing. The line indicated by + STRAND, denotes DNA sequencing of the coding strand and the line indicated by – strand, denotes DNA sequencing of the non-coding strand; small letters are used to indicate that the given sequence is the complement to the one obtained by DNA sequencing. A dash indicates positions where the generated sequence agrees with that of the aligned sequences. Single base calls at sites where the repeat sequences differ are observed before the deletion while double base calls at differing sites are recovered after the deletion.
compared by SDS±PAGE and Western blot analysis (Figure 5). In samples collected under conditions where liquification is inhibited, both the homozygous and the heterozygous sample display protein bands representing SgI of 50 kDa and the two forms of SgII that migrate similar to the 67 kDa standard. However, in the sample from the heterozygous carrier there is also an extra molecular species of ~43 kDa that is immuno-stained by an antibody that recognizes a shared epitope of SgI and SgII. This demonstrates that the new allele is expressed in seminal vesicles and, judging from the staining intensity on the polyacrylamide gel, the expression level is as high as for SgI50. The new allele does not seem to grossly affect the stability of the semen coagulum as SgI43 is degraded to small molecular mass fragments during semen liquefaction, similar to SgI50.

On purification, by chromatography on heparin±Sepharose, SgI50 and SgI43 co-elute in fractions that are free of SgII. A sample of the purified mixture of truncated and normally sized SgI was incubated with PSA in order to study the former component’s susceptibility to PSA digestion. For comparison, a sample with SgI50 was digested

compared by SDS–PAGE and Western blot analysis (Figure 5). In samples collected under conditions where liquification is inhibited, both the homozygous and the heterozygous sample display protein bands representing SgI of 50 kDa and the two forms of SgII that migrate similar to the 67 kDa standard. However, in the sample from the heterozygous carrier there is also an extra molecular species of ~43 kDa that is immuno-stained by an antibody that recognizes a shared epitope of SgI and SgII. This demonstrates that the new allele is expressed in seminal vesicles and, judging from the staining intensity on the polyacrylamide gel, the expression level is as high as for SgI50. The new allele does not seem to grossly affect the stability of the semen coagulum as SgI43 is degraded to small molecular mass fragments during semen liquefaction, similar to SgI50.

On purification, by chromatography on heparin–Sepharose, SgI50 and SgI43 co-elute in fractions that are free of SgII. A sample of the purified mixture of truncated and normally sized SgI was incubated with PSA in order to study the former component’s susceptibility to PSA digestion. For comparison, a sample with SgI50 was digested

compared by SDS±PAGE and Western blot analysis (Figure 5). In samples collected under conditions where liquification is inhibited, both the homozygous and the heterozygous sample display protein bands representing SgI of 50 kDa and the two forms of SgII that migrate similar to the 67 kDa standard. However, in the sample from the heterozygous carrier there is also an extra molecular species of ~43 kDa that is immuno-stained by an antibody that recognizes a shared epitope of SgI and SgII. This demonstrates that the new allele is expressed in seminal vesicles and, judging from the staining intensity on the polyacrylamide gel, the expression level is as high as for SgI50. The new allele does not seem to grossly affect the stability of the semen coagulum as SgI43 is degraded to small molecular mass fragments during semen liquefaction, similar to SgI50.

On purification, by chromatography on heparin–Sepharose, SgI50 and SgI43 co-elute in fractions that are free of SgII. A sample of the purified mixture of truncated and normally sized SgI was incubated with PSA in order to study the former component’s susceptibility to PSA digestion. For comparison, a sample with SgI50 was digested

compared by SDS±PAGE and Western blot analysis (Figure 5). In samples collected under conditions where liquification is inhibited, both the homozygous and the heterozygous sample display protein bands representing SgI of 50 kDa and the two forms of SgII that migrate similar to the 67 kDa standard. However, in the sample from the heterozygous carrier there is also an extra molecular species of ~43 kDa that is immuno-stained by an antibody that recognizes a shared epitope of SgI and SgII. This demonstrates that the new allele is expressed in seminal vesicles and, judging from the staining intensity on the polyacrylamide gel, the expression level is as high as for SgI50. The new allele does not seem to grossly affect the stability of the semen coagulum as SgI43 is degraded to small molecular mass fragments during semen liquefaction, similar to SgI50.

On purification, by chromatography on heparin–Sepharose, SgI50 and SgI43 co-elute in fractions that are free of SgII. A sample of the purified mixture of truncated and normally sized SgI was incubated with PSA in order to study the former component’s susceptibility to PSA digestion. For comparison, a sample with SgI50 was digested

compared by SDS±PAGE and Western blot analysis (Figure 5). In samples collected under conditions where liquification is inhibited, both the homozygous and the heterozygous sample display protein bands representing SgI of 50 kDa and the two forms of SgII that migrate similar to the 67 kDa standard. However, in the sample from the heterozygous carrier there is also an extra molecular species of ~43 kDa that is immuno-stained by an antibody that recognizes a shared epitope of SgI and SgII. This demonstrates that the new allele is expressed in seminal vesicles and, judging from the staining intensity on the polyacrylamide gel, the expression level is as high as for SgI50. The new allele does not seem to grossly affect the stability of the semen coagulum as SgI43 is degraded to small molecular mass fragments during semen liquefaction, similar to SgI50.

On purification, by chromatography on heparin–Sepharose, SgI50 and SgI43 co-elute in fractions that are free of SgII. A sample of the purified mixture of truncated and normally sized SgI was incubated with PSA in order to study the former component’s susceptibility to PSA digestion. For comparison, a sample with SgI50 was digested

compared by SDS±PAGE and Western blot analysis (Figure 5). In samples collected under conditions where liquification is inhibited, both the homozygous and the heterozygous sample display protein bands representing SgI of 50 kDa and the two forms of SgII that migrate similar to the 67 kDa standard. However, in the sample from the heterozygous carrier there is also an extra molecular species of ~43 kDa that is immuno-stained by an antibody that recognizes a shared epitope of SgI and SgII. This demonstrates that the new allele is expressed in seminal vesicles and, judging from the staining intensity on the polyacrylamide gel, the expression level is as high as for SgI50. The new allele does not seem to grossly affect the stability of the semen coagulum as SgI43 is degraded to small molecular mass fragments during semen liquefaction, similar to SgI50.

On purification, by chromatography on heparin–Sepharose, SgI50 and SgI43 co-elute in fractions that are free of SgII. A sample of the purified mixture of truncated and normally sized SgI was incubated with PSA in order to study the former component’s susceptibility to PSA digestion. For comparison, a sample with SgI50 was digested

compared by SDS±PAGE and Western blot analysis (Figure 5). In samples collected under conditions where liquification is inhibited, both the homozygous and the heterozygous sample display protein bands representing SgI of 50 kDa and the two forms of SgII that migrate similar to the 67 kDa standard. However, in the sample from the heterozygous carrier there is also an extra molecular species of ~43 kDa that is immuno-stained by an antibody that recognizes a shared epitope of SgI and SgII. This demonstrates that the new allele is expressed in seminal vesicles and, judging from the staining intensity on the polyacrylamide gel, the expression level is as high as for SgI50. The new allele does not seem to grossly affect the stability of the semen coagulum as SgI43 is degraded to small molecular mass fragments during semen liquefaction, similar to SgI50.

On purification, by chromatography on heparin–Sepharose, SgI50 and SgI43 co-elute in fractions that are free of SgII. A sample of the purified mixture of truncated and normally sized SgI was incubated with PSA in order to study the former component’s susceptibility to PSA digestion. For comparison, a sample with SgI50 was digested

compared by SDS±PAGE and Western blot analysis (Figure 5). In samples collected under conditions where liquification is inhibited, both the homozygous and the heterozygous sample display protein bands representing SgI of 50 kDa and the two forms of SgII that migrate similar to the 67 kDa standard. However, in the sample from the heterozygous carrier there is also an extra molecular species of ~43 kDa that is immuno-stained by an antibody that recognizes a shared epitope of SgI and SgII. This demonstrates that the new allele is expressed in seminal vesicles and, judging from the staining intensity on the polyacrylamide gel, the expression level is as high as for SgI50. The new allele does not seem to grossly affect the stability of the semen coagulum as SgI43 is degraded to small molecular mass fragments during semen liquefaction, similar to SgI50.

On purification, by chromatography on heparin–Sepharose, SgI50 and SgI43 co-elute in fractions that are free of SgII. A sample of the purified mixture of truncated and normally sized SgI was incubated with PSA in order to study the former component’s susceptibility to PSA digestion. For comparison, a sample with SgI50 was digested
The seminal vesicles concluded that both SgI and SgII have undergone
decomposition (Lundwall, 1998). The most obvious mechanism to explain this
effect is small duplications, affecting the C-terminal half of the
molecules, creating varying numbers of tandem repeats with sizes of
~60 amino acid residues. However, the molecular variation among
primates is not merely a matter of repeat length, but is related to the
sequence that is absent from SgI43 and thus can be related to the
structural alterations relatively recently during evolution
(Lundwall and Lazure, 1995; Ulvsbäck and Lundwall, 1997;
Lundwall, 1998). The most obvious mechanism to explain this
phenomenon is small duplications, affecting the C-terminal half of the
molecules, creating varying numbers of tandem repeats with sizes of
~60 amino acid residues. However, the molecular variation among
primates is not merely a matter of repeat number, as the residues
affected by duplication differ among species. In this report, we
describe a previously unknown allele encoding a truncated form of
human SgI. As the size difference between the products of the two
human SgI alleles is 60 amino acid residues, it appears as if size
alteration of semenogelin molecules is still proceeding. It follows that
SgII might also be affected by size variation. In the present study no
heterogeneity was detected in the SgII gene. However, due to the
limited number of subjects under study, the existence of a size variant
of SgII occurring at a low frequency cannot be excluded. It is also
possible that the outcome of the SgII allele screening would have been
different with a study population from a different ethnic background.

The latter is also an interesting aspect that pertains to the novel SgI
allele. Although the allele distribution in different ethnic populations
has not been the subject of the present investigation, it appears that the
SgI43 allele is widely spread as the screening also revealed carriers of
non-European descent.

The novel SgI43 allele could either be parental to, or a deletion
product of, the SgI50 allele. Results showing that semenogelin
molecules became larger during evolution because of repeated
duplications are in favour of the former alternative (Ulvsbäck and
Lundwall, 1997; Lundwall, 1998). However, these studies on
semenogelin genes in monkeys also show that recently added tandem
repeats are almost perfect copies of each other. The most recent
duplication in SgI50 probably gave rise to the tandem repeats located
between residues 259–318 and 319–378, as these are the most
conserved repeated sequences in the molecule. Both regions are
affected by the amino acid sequence that is absent from SgI43 and thus,
the shorter allele is probably the result of a deletion.

There seems to be no major effect on reproductive function
associated with heterozygosity for the SgI43 allele. Both fertilizing
ability and sperm parameters appear to be within the normal range.
However, a limited effect cannot be excluded and might be revealed
by association analyses of the SgI43 allele and a larger number of
parameters of importance for male reproductive function. In particu-
lar, studies of individuals that are homozygous for the novel allele will
be important and might have a strong impact on the functional
evaluation of the molecule. Such individuals will probably not be hard
to find, given the allele frequency, which suggest that there may be one
homoyzogous individual for SgI43 in a population of around 1100
individuals.

At or very shortly after ejaculation, the semen of many mammals
forms a coagulum. There is a tremendous species variation in

---

**Table I.** Comparison of the time of abstinence and seminal parameters in Swedish young males
homozygous (n = 20) for the 43 kDa variant (Sgl43) or homozygous (n = 285) for the 50 kDa variant
(Sgl50) of Sgl

| Reproductive parameter | Median (range) | Mean (SD) | P*
|------------------------|---------------|-----------|---
| Time of abstinence (h)  | Sgl50/Sgl50   | Sgl43/Sgl50 | Sgl100/Sgl50  | Sgl143/Sgl100 |
| Semen volume (ml)      | 68 (12–504)   | 63 (43–168) | 86 (58)       | 76 (35)       | 0.48       |
| Sperm concentration (10^6/ml) | 3.2 (0.3–8.4) | 2.8 (1.6–7.6) | 3.2 (1.3) | 3.2 (1.5) | 0.52 |
| Total sperm count (10^6) | 54 (0.1–390)  | 62 (11–205) | 73 (68)       | 64 (41)       | 0.82       |
| Motile sperm (%)       | 167 (0.5–1180)| 158 (31–679) | 209 (187)     | 197 (147)     | 0.88 |

*a* Determined by Mann–Whitney test for unpaired data.

---

**Discussion**

Previous studies of genes encoding predominant proteins secreted by
the seminal vesicles concluded that both SgI and SgII have undergone
simultaneously. The result was analysed by SDS–PAGE (Figure 6).
As can be seen, SgI43 and SgI50 are equally susceptible to degradation
by PSA. Thus, despite the size difference, the novel SgI variant
appears to have many shared properties with the more common SgI
variant of 50 kDa.

**Semen parameters**

Semen samples from individuals heterozygous for SgI43 did not show
abnormal liquefication or any other deviation from normal samples
that can be observed by macroscopic inspection. A comparison of
sperm parameters between those 20 individuals that are heterozygous
for the SgI43 allele to those homozygous for SgI50, did not reveal any
statistically significant difference with respect to semen volume,
sperm concentration, total sperm count or the percentage of motile
sperm (Table I). The values were independent of sexual abstinence, as
the recorded period of abstinence did not significantly differ between
the two groups. In a questionnaire, the 305 sperm donors were asked if
they had caused a pregnancy. A positive response was given by four
(20%) in the group heterozygous for SgI43 and 24 (8%) among the
remaining 285 individuals homozygous for SgI50.

---

**Figure 6.** Digestion by PSA. Purified Sgl from an individual homozygous
for the normal SgI50 allele (N) and a heterozygous carrier of the SgI43 allele
(V) were digested with purified PSA for varying times. Protein components
were separated by SDS–PAGE and stained by Coomassie Brilliant Blue. A
molecular standard with sizes given in kDa is shown in the left part of the
figure. The position of PSA is indicated by an arrow.
physicochemical properties of the coagulum, ranging from the loose, gel-like, human coagulum to the transglutaminase-stabilized coagulum of rodents that is reminiscent of a blood clot. Perhaps this macroscopic variation is a reflection of the rapid evolution of the coagulum proteins secreted by the seminal vesicles. One might, therefore, expect that SgI14 could affect the properties of human coagulum. At least in relation to liquefaction and susceptibility to proteolytic degradation by PSA, there is no apparent difference compared with SgI10, suggesting that the truncated molecule does not affect the stability of the coagulum. Whether it affects semen viscosity or coagulum strength remains to be studied.

Acknowledgements
The technical assistance by Birgitta Frohm, Stefan Strömberg and Agneta Östensson is acknowledged. This work was supported by grants from the Swedish Cancer Society (project nos 4564, 4423 and 3555), the Swedish Research Council (project nos 7903 and 14199) and Fundacion Federico SA.

References


Submitted on January 15, 2003; accepted on March 6, 2003