



LUND UNIVERSITY  
Faculty of Medicine

---

# LU:*research*

*Institutional Repository of Lund University*

---

This is an author produced version of a paper published in The FEBS journal. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the published paper:

Malm, Johan and Jonsson, Magnus and Frohm, Birgitta and Linse, Sara.

"Structural properties of semenogelin I."  
FEBS J, 2007, Vol: 274, Issue: 17, pp. 4503-10.

<http://dx.doi.org/10.1111/j.1742-4658.2007.05979.x>

Access to the published version may  
require journal subscription.

Published with permission from: Blackwell

# Structural properties of semenogelin I

JOHAN MALM<sup>1</sup>, MAGNUS JONSSON<sup>1\*</sup>, BIRGITTA FROHM<sup>1</sup>, AND  
SARA LINSE<sup>2</sup>

<sup>1</sup>Department of Laboratory Medicine, Section for Clinical Chemistry, Lund University, Malmö University Hospital, SE-205 02 Malmö, Sweden, and <sup>2</sup>Department of Biophysical Chemistry, Lund University, SE-221 00 Lund, Sweden

Authors:

Johan Malm	johan.malm@med.lu.se
Magnus Jonsson*	magnus.jonsson@med.lu.se
Birgitta Frohm	birgitta.frohm@med.lu.se
Sara Linse	sara.linse@bpc.lu.se

\*To whom correspondence should be addressed: e-mail [magnus.jonsson@med.lu.se](mailto:magnus.jonsson@med.lu.se); telephone +46 40 33 14 37; fax +46 40 33 62 86.

Running title: Structural properties of semenogelin I

## Summary

The zinc-binding protein semenogelin I is the major structural component of the gelatinous coagulum that is formed in freshly ejaculated semen. Semenogelin I is a rapidly evolving protein with a primary structure that consists of six repetitive units, each comprising approximately 60 amino acid residues. We studied the secondary and tertiary structure of semenogelin I by circular dichroism (CD) spectroscopy and Trp fluorescence emission spectroscopy. Fitting to the far-UV CD data indicated that the molecule comprises 5–10%  $\alpha$ -helix and 20–30%  $\beta$ -sheet formations. The far-UV spectrum of semenogelin I is clearly temperature dependent in the studied range of 5°C to 90°C, and the signal at 222 nm increased with increasing temperature. The presence of  $\text{Zn}^{2+}$  did not change the secondary structure revealed by the far-UV CD spectrum, whereas it did alter the near-UV CD spectrum, which implies that rearrangements occurred on the tertiary structure level. The conformational change induced in semenogelin I by the binding of  $\text{Zn}^{2+}$  may contribute to the ability of this protein to form a gel.

Keywords: semenogelin; zinc; structure; fertility

## Introduction

In living systems, the interactions between proteins and metal ions control many central processes, such as memory, learning, blood clotting, muscle contraction, and vision.

Generally speaking, a metal ion can play a catalytic or a stabilizing role, it can induce a conformational change, or it can mediate protein-protein interplay. It was recently reported that the cooperation between  $Zn^{2+}$  and proteins controls both the formation and the breakdown of the loose gel in freshly ejaculated semen [1]. More specifically, it was found that these processes involve two classes of  $Zn^{2+}$ -binding proteins: the gel-forming semenogelins and a  $Zn^{2+}$ -regulated protease.

Semenogelins I and II (SgI and SgII)<sup>1</sup> are the predominant structural proteins in the loose gel formed in freshly ejaculated human semen. The concentration of SgI is 5–10 times higher than the level of SgII in semen, and these two molecules are the quantitatively dominating proteins in the fluid from the seminal vesicles, which contributes approximately 60% of the ejaculate volume [2, 3]. The secretion from the epididymis, which contains the spermatozoa, constitutes only a few percent of the ejaculate volume, and the remaining fraction of the semen (approx. 30%) comes mainly from the prostate and is rich in serine proteases and  $Zn^{2+}$  [4-6]. At ejaculation, the fluids are mixed to form a non-covalently linked gel-like structure that entraps the spermatozoa (Fig. 1). Within twenty minutes of ejaculation, the gel is almost completely liquefied by serine proteases, primarily prostate-specific antigen (PSA), which cleaves the SgI and SgII molecules to yield soluble fragments [4]. PSA is stored in the prostate in a  $Zn^{2+}$ -inhibited form, but it is activated upon mixing with SgI and SgII, both of which have a higher  $Zn^{2+}$ -binding capacity than PSA [1]. In parallel to this liquefaction, the spermatozoa become progressively more motile.

The concentration of  $Zn^{2+}$  is a hundred times higher in seminal plasma (i.e., semen without the spermatozoa) than in blood plasma [7]. The semenogelins are the major  $Zn^{2+}$ -binding proteins in seminal plasma [1], and there is indirect evidence that  $Zn^{2+}$  induces a conformational change in both intact semenogelin molecules and synthetic semenogelin peptides.

Interestingly, the semenogelins were recently identified in another  $Zn^{2+}$ -rich environment—the retina [8].

The primary structures of SgI (439 amino acid residues [2]) and SgII (559 amino acid residues) are very similar (78% amino acid identity), and there are comparable 60-amino-acid residue repeats in the proteins, six in SgI and eight in SgII [9]. The two different genes encoding these proteins are located 11.5 kbp apart on the long arm of chromosome 20 [10]. The semenogelins are rapidly evolving proteins that are coded by three exons: the first gives rise to the signal peptide, the second encodes the secreted protein, and the third expresses the untranslated 3' part of the mRNA [11]. Neither the primary structure nor the repetitive elements of the intact semenogelins are similar to motifs seen in other proteins, and thus the structure cannot be predicted from class neighbors.

The capacity of the semenogelins to form a gelatinous mass may influence the ability of the spermatozoa to reach and fuse with an ovum [12]. However, studies have not yet elucidated the molecular mechanisms responsible for creation of the gelatinous coagulum upon ejaculation. The gel is liquefied under denaturing conditions, which indicates that the conformation of the proteins is important for the integrity of this semi-solid mass [13]. The repetitive nature of the semenogelins, as well as their susceptibility to protease degradation, suggests that these molecules have a non-globular structure [14]. To gain a better understanding of the biophysical mechanisms of gel formation in seminal plasma, we studied SgI with regard to its structural properties and the influence of  $Zn^{2+}$  on those characteristics. The degree and stability of secondary structures was estimated by far-UV circular dichroism (CD) spectroscopy, and the tertiary structure was studied by both near-UV CD spectroscopy and tryptophan fluorescence emission spectroscopy.

## **Results**

SgI shows low solubility in buffers that are not supplemented with urea, whereas it is fully or partially denatured when exposed to high concentrations of urea. Therefore, the first step in the CD experiments was to find the optimal concentration of urea to use in structural investigations. We treated SgI with different levels of urea (0.2–2 M) in the presence of 0.5 M NaCl, and recorded CD spectra in the range 250–200 nm (Fig. 2). The signal at 222 nm originates chiefly from the peptide bonds in the backbone of the protein, and hence it correlates with the degree of secondary structure. At this wavelength, there is only minor interference from urea, although the aromatic side chains may have some effect. At urea concentrations above 0.8 M, the absolute value of the signal is slightly decreased, which indicates a lower degree of structure and an increasing tendency towards random coil.

SgI was exposed to 5 mM Tris/HCl (pH 9.7) supplemented with 0.5 M NaCl and 0.5 M urea, and the CD signal at 222 nm was measured as a function of temperatures gradually increasing from 25°C to 90°C, as illustrated in Figure 3A. The graph in the figure shows a linear increase in negative ellipticity, which reflects increasing secondary structure with rising temperature. The sensitivity to temperature was further evaluated by recording far-UV CD spectra at temperatures in the range 5–90°C, in both a 0.1-cm and a 0.01-cm cuvette (Fig. 3, B and C). When using the 0.01 cm cuvette, the protein concentration was raised to compensate for the shorter path length, which preserved the amplitude of the signal and resulted in less interference from urea.

SgI has an isoelectric point above 9.5, and it appears to be more soluble at pH values greater than 8. Therefore, we recorded far-UV CD spectra at pH values of 2.5, 6.3, 8.1, and 9.7 in a buffer containing 0.5 M urea. The spectra were not changed by low pH values or by addition of 2%, 10%, or 15% trifluoroethanol (data not shown). Furthermore, since SgI is a Zn<sup>2+</sup>-binding protein, we performed titration with 0–200 μM ZnAc and found that the presence of Zn<sup>2+</sup> did not alter the far-UV CD-spectra (data not shown).

The CD spectra obtained for SgI in a 0.01-cm cuvette at 5–90°C in a buffer supplemented with 0.5 M urea were used as input data in CDPro (which includes the programs CONTINLL, Selcon 3, and CDSSTR) to estimate the extent to which the different types of secondary structure were present. The algorithms in the programs compares the CD spectra for SgI with those recorded for a set of reference proteins. Only results corresponding to a voltage below 600 mV were used in the prediction, which gave a lower limit of 208 nm for the spectra. The findings are summarized in Table 1. Regardless of the temperature, all three methods provided roughly the same results with ~5–10% α-helical structure and ~20–30% of the residues in β-sheets. The smallest variation was shown by CDSSTR, which indicated that the sums of α-helix and β-sheet structure were 4–7% and 23–26%, respectively. Considering the set of reference proteins, green fluorescent protein (11% α-helix and 37% β-sheet conformation) yielded the far-UV CD spectrum that was most similar to that of SgI. The fit between the far-UV CD spectra of SgI and green fluorescent protein is illustrated in Figure 4. The ellipticity at 222 nm is often used to roughly estimate the degree of secondary structure in a protein. Therefore, we plotted the percentages of α-helical structure and the sum of α-helix

and  $\beta$  sheets for the individual proteins in the reference set versus their  $\Delta\epsilon$  values at 222 nm (Fig. 5). The correlation between the secondary structure and  $\Delta\epsilon$  at 222 nm was calculated by linear regression. Using the correlation between  $\alpha$ -helix and  $\Delta\epsilon$  at 222 nm as the reference proteins to approximate the degree of  $\alpha$ -helical structure in SgI at the temperatures 5°C, 20°C, 37°C, 45°C, 65°C, and 90°C resulted in values of 4.5%, 6.4%, 8.5%, 8.6%, 10%, and 12%, respectively. The sum of  $\alpha$ -helix and  $\beta$ -sheet conformation in SgI approximated by the same method (Fig. 5B) gave values of 31%, 32%, 34%, 34%, 35%, and 36% at the corresponding temperatures. These figures are in the same range as the predictions based on the spectra. According to  $\Delta\epsilon$  at 222 nm, the degree of secondary structure in SgI seems to increase with increasing temperature.

SgI contains six Phe, 14 Tyr, and two Trp residues, which we used to monitor the tertiary structure. Fluorescence intensity was recorded between 320 and 450 nm, using the excitation wavelengths 295 nm (affecting mainly Trp residues) and 280 nm (exciting both Trp and Tyr residues) in the presence of 0.5 M urea at 25°C. A broad peak at ~350 nm was noted at both excitation wavelengths, albeit with slightly lower intensity at 295 nm (Fig. 6). Raising the urea concentration to 7.4 M resulted in a sharp and higher peak in the spectra at both excitation wavelengths. The emission in the 295-nm and 280-nm spectra increased by ~75% and ~20%, respectively. The rise in fluorescence intensity in the presence of 7.4 M urea indicates that the fluorescence of the Trp residue was quenched (e.g., by a charged residue) at 0.5 M concentration of urea. That finding suggests that native SgI probably has some degree of tertiary structure that is sensitive to denaturation.

Near-UV CD spectra were recorded for SgI at different temperatures in the range 5–45°C in the presence and absence of 20  $\mu\text{M}$   $\text{Zn}^{2+}$  (Fig. 7). No gel or precipitation was observed under these conditions (0.5 M urea, 0.5 M NaCl, 5 mM Tris (pH9.7), 20  $\mu\text{M}$  SgI and 20  $\mu\text{M}$   $\text{Zn}^{2+}$ ). For proteins, such measurements reveal the structural confinement of the side chains of aromatic residues. In a folded protein, these residues may be situated in an asymmetric environment, with reduced rotational mobility, and therefore the near-UV CD signal depends on the tertiary structure of the protein. We observed a distinct increase in negative ellipticity between 260 and 285 nm in the presence of  $\text{Zn}^{2+}$ , which indicates that binding of the ion induces either a change in the tertiary structure or decreased rotational freedom of aromatic side chains.

## Discussion

Considering our results, the far-UV CD spectra of semenogelin I indicate that the protein contains secondary structure, and predictions made using computer-based models suggest that 4–8% and 20–30% of the molecule consist of  $\alpha$ -helix and  $\beta$ -sheet structure, respectively. The degree of secondary structure increases at higher temperatures, which implies that the protein is heat stable. Also, the SgI molecule has tertiary structure that changes in the presence of  $\text{Zn}^{2+}$ .

SgI and green fluorescent protein (which is an energy transfer acceptor in jelly fish) are similar with regard to predicted secondary structure content but not with respect to their primary structure (compared by use of BLASTP 2.2.13, matrix BLOSUM62 with default settings available at <http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>). The benefit of a heat stable secondary structure for SgI is not obvious from a physiological perspective. The ability to build up and maintain a structure within a particular temperature range is mainly an intrinsic property that is determined by the amino acid sequences [15]. As mentioned in the introduction, due to the rapid evolution of the semenogelins, SgI has a primary structure that differs greatly from motifs seen in other structurally well-characterized proteins. Thus, the amino acid sequence cannot be used to predict the structure of SgI or to ascertain whether this protein has secondary or tertiary structural similarities to other thermophilic proteins.

The SgI molecule has two Trp residues. At high concentrations of urea, we found that the Trp fluorescence emission spectra for SgI exhibited increased signal intensity compared to the spectra recorded under non-denaturing conditions. Many globular proteins show decreased Trp fluorescence intensity upon denaturation as the result of quenching due to collisions with solvent water. However, the opposite can be seen when the Trp fluorescence is quenched in the folded state, for example by a nearby disulfide or prosthetic group. Consequently, there is no experimental evidence that the SgI molecule has globular properties, although it clearly possesses tertiary structure that is sensitive to denaturation by urea.

There are reasons to believe that SgI is stabilized by binding of  $\text{Zn}^{2+}$ . Previous studies have demonstrated that both SgI and SgII have a high  $\text{Zn}^{2+}$ -binding capacity, with  $K_D$  values in the micromolar range and a stoichiometry of at least 10 zinc ions per molecule. In the body, the semenogelins and  $\text{Zn}^{2+}$  are stored separately, and they are not exposed to each other until

ejaculation leads to mixing of the semenogelin-rich secretion from the seminal vesicles and the  $Zn^{2+}$ -rich secretion from the prostate to form a coagulum. Hypothetically, this coagulation phenomenon might occur because binding of  $Zn^{2+}$  alters the tertiary structure of the semenogelins to a more stable form, and that particular conformation can participate in stable non-covalent interactions with the surrounding structural proteins (mainly other semenogelin molecules, but possibly also fibronectin). Another plausible explanation is that  $Zn^{2+}$  simply bridges the semenogelins. The semi-solid consistency of the gel suggests that the semenogelins have a more rigid tertiary/quaternary structure when acting as components of the coagulum than when they appear in solution. The importance of the protein structure and stability in this context is further emphasized by the fact that it takes a high concentration of urea to dissolve the coagulum. Our results imply that not only does the SgI molecule display secondary structure, but it also harbours tertiary structure that is changed by exposure to  $Zn^{2+}$ . The observation that high concentrations of urea dissolve the gel strengthens the assumption that the structure of SgI (as the dominating protein in the coagulum) is important for its ability to induce formation of a gelatinous mass.

## **Experimental procedures**

*Human SgI.* Human semen specimens were collected from healthy volunteer sperm donors (through masturbation) at the fertility laboratory (Malmö University Hospital, Malmö, Sweden). SgI was purified essentially as described by Jonsson et al. [1]. The concentration of SgI was determined by assessment performed after acid hydrolysis (24 h in 6 M HCl at 110°C *in vacuo*) on a Beckman 6300 amino acid analyzer. The protein was diluted to appropriate concentrations for each experiment.

*CD spectroscopy.* To investigate the conformation of SgI under different conditions, far-UV and near-UV spectra were recorded using a Jasco J720 spectropolarimeter equipped with a Peltier heating element temperature controller Jasco PT343 (Jasco Inc., Easton, MD, USA). Secondary structure parameters were estimated using the computer software package CDPro (<http://lamar.colostate.edu/~sreerama/CDPro>; [16, 17] to compare CD spectra recorded for SgI at different temperatures and a cell path length of 0.01 cm with the spectra of reference proteins (basis set 5).

Far-UV CD spectra (250–200 nm) of SgI were recorded at 25°C using different concentrations of urea. The concentration of SgI was 7.6  $\mu$ M in 5 mM Tris buffer (pH 9.7)

containing 0.5 M NaCl, and a cell path length of 0.1 cm used. The concentration of urea was varied between 0.2 M and 2 M. The spectra illustrated represent an average of two scans (scan rate 10 nm/min, response 16 s, resolution 1 nm, step 1 nm) from which a background spectrum recorded for the buffer without protein was subtracted.

Melting curves were measured at 222 nm at a cell path length of 0.1 cm by slowly increasing the temperature from 25°C to 90°C (1°C per minute), using samples containing 7.6  $\mu$ M SgI in 5 mM Tris buffer (pH 9.7) supplemented with 0.5 M NaCl and 0.5 M urea.

SgI concentrations of 7.6  $\mu$ M and 63  $\mu$ M in 5 mM Tris buffer (pH 9.7) containing 0.5 M NaCl and 0.5 M urea were used to record far-UV CD spectra at different temperatures in cells with path lengths of 0.1 cm (250–200 nm) and 0.01 cm (250–180 nm), respectively. The spectra reported were run at 5°C, 25°C, 37°C, 45°C, 65°C, and 90°C, and they represent an average of two (0.1-cm cuvette) or eight (0.01-cm cuvette) scans corrected for background.

Near-UV CD spectra (320–250 nm) of Sg I were recorded in the presence and absence of 20  $\mu$ M  $Zn^{2+}$  at 5°C, 20°C, 37°C, and 45°C. The protein concentration was 20  $\mu$ M in 5 mM Tris buffer (pH 9.7) containing 0.5 M NaCl and 0.5 M urea, and a cell path length of 1 cm was used. The spectra reported were recorded at 5°C, 20°C, 37°C, and 45°C and each represents an average of ten scans. Due to the low ellipticity signal, background correction was performed by subtracting the mean value of the data points obtained between 320 nm and 311 nm. The  $Zn^{2+}$  concentration (20  $\mu$ M) was chosen so as to avoid precipitation which interferes with the CD measurements. When a higher  $Zn^{2+}$  concentration (100  $\mu$ M) was used, no reliable signal was obtained due to high background absorbance.

*Fluorescence measurements.* Fluorescence spectra of SgI at different concentrations of urea were recorded using an LS 50B spectrofluorometer (Perkin Elmer, Inc., Wellesley, MA, USA) with excitation and emission band passes set at 3 nm and 6 nm, respectively. Trp spectra were obtained with excitation wavelengths of 280 nm and 295 nm, and the emission was scanned in the range 320–450 nm. The protein was used at a concentration of 7.6  $\mu$ M in 5 mM Tris buffer (pH 9.7) containing 0.5 M NaCl and 0.5 M or 7.4 M urea.

## **Acknowledgements**

This study was supported by grants from the Swedish Research Council (project no. 14199), the Alfred Österlund Foundation, the Malmö University Hospital Cancer Foundation, Scania County Council Research and Development Foundation, the Foundation of Malmö University Hospital, and Fundacion Federico S.A.

## **Footnotes**

<sup>1</sup>Abbreviations: SgI, semenogelin I; SgII, semenogelin II; PSA, prostate-specific antigen; CD, circular dichroism.

## References

1. Jonsson M, Linse S, Frohm B, Lundwall A & Malm J (2005) Semenogelins I and II bind zinc and regulate the activity of prostate-specific antigen. *Biochem J* **387**, 447-453.
2. Lilja H, Abrahamsson PA & 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.
3. Malm J, Hellman J, Magnusson H, Laurell CB & 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.
4. Lilja H (1985) A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. *J Clin Invest* **76**, 1899-1903.
5. Chapdelaine P, Paradis G, Tremblay RR & Dube JY (1988) High level of expression in the prostate of a human glandular kallikrein mRNA related to prostate-specific antigen. *FEBS Lett* **236**, 205-208.
6. Arver S (1982) Zinc and zinc ligands in human seminal plasma. III. The principal low molecular weight zinc ligand in prostatic secretion and seminal plasma. *Acta Physiol Scand* **116**, 67-73.
7. Burtis CA & Aschwood ER (1999) Tietz textbook of clinical chemistry, 3<sup>rd</sup> ed. W. B. Saunders company, Philadelphia.
8. Bonilha VL, Rayborn ME, Shadrach K, Ake L, Malm J, Bhattacharya SK, Crabb JW & Hollyfield JG (2006) Characterization of semenogelin proteins in the human retina. *Exp Eye Res* **83**, 120-127.
9. Lilja H & Lundwall A (1992) Molecular cloning of epididymal and seminal vesicular transcripts encoding a semenogelin-related protein. *Proc Natl Acad Sci U S A* **89**, 4559-4563.
10. Ulvsback M, Lazure C, Lilja H, Spurr NK, Rao VV, Loffler C, Hansmann I & Lundwall A (1992) Gene structure of semenogelin I and II. The predominant proteins in human semen are encoded by two homologous genes on chromosome 20. *J Biol Chem* **267**, 18080-18084.
11. Lundwall A & Lazure C (1995) A novel gene family encoding proteins with highly differing structure because of a rapidly evolving exon. *FEBS Lett* **374**, 53-56.

12. Robert M & Gagnon C (1999) Semenogelin I: a coagulum forming, multifunctional seminal vesicle protein. *Cell Mol Life Sci* **55**, 944-960.
13. Lilja H & Laurell CB (1985) The predominant protein in human seminal coagulate. *Scand J Clin Lab Invest* **45**, 635-641.
14. Wright PE & Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* **293**, 321-331.
15. Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230.
16. Sreerama N & Woody RW (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal Biochem* **209**, 32-44.
17. Johnson WC (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins* **35**, 307-312.

**Table Legend**

*Table 1.* Prediction of the secondary structure of semenogelin I by computer-based analysis of far-UV CD measurements

## Figure legends

*Figure 1.* Schematic flow chart illustrating the coagulation and liquefaction of human semen. (A) Components of the semen are stored separately and mixed upon ejaculation. (B) Mixing the prostate secretion rich in  $Zn^{2+}$  and zinc-inhibited PSA with the seminal fluid that contains large amounts of semenogelins results in that the semenogelins bind the major fraction of  $Zn^{2+}$ . This induces a conformational change of SgI that enables gel-formation and diminishes the concentration of free  $Zn^{2+}$ . As a consequence of the diminished free  $Zn^{2+}$  concentration PSA is activated (C) PSA cleaves the semenogelins, which results in liquefaction of the gel and motile spermatozoa are released.

*Figure 2.* Far-UV CD spectra of SgI in buffer containing urea at concentrations of 0.2 (■), 0.5 (▲), 0.8 (□) and 2 M (△). Only results corresponding to a voltage below 600 mV are shown.

*Figure 3.* CD measurements of SgI at different temperatures in a buffer containing 0.5 M urea. (A) Mean residue ellipticity recorded at 222 nm plotted versus temperature. (B) CD spectra of SgI recorded at different temperatures using a 0.1-cm cuvette. Spectra were collected at temperatures of 5°C, 25°C, 45°C, 65°C, and 90°C. (C) As in (B) except using a 0.01-cm cuvette and a temperature of 37°C. Only results corresponding to a voltage below 600 mV are illustrated.

*Figure 4.* (A) Examples of fitting to the experimental data presented in Figure 3C performed using CDPro. The curve for 45°C was excluded because it gave essentially the same results as the curve for 37°C. (B) CD spectrum of green fluorescent protein (GFP), which, according to CDPro, was most similar to the spectra of SgI (considering all the proteins in the reference set).

*Figure 5.* The percentage of  $\alpha$ -helical conformation (A) and the sum of the proportions of  $\alpha$ -helix and  $\beta$ -sheet structure (B) for each protein in the reference set plotted versus their  $\Delta\epsilon$  at 222 nm. The line in each graph represents the correlation (calculated by linear regression) between the secondary structure and  $\Delta\epsilon$  at 222 nm of the proteins in the reference set.

*Figure 6.* Trp fluorescence emission spectrometry of SgI. The excitation wavelengths 280 nm (A) and 295 nm (B) were used to analyze SgI in buffer containing urea at a concentration of 7.4 M (1) or 0.5 M (2).

*Figure 7.* Near-UV CD spectra of SgI at different temperatures in the presence (A) and the absence (B) of  $Zn^{2+}$ . The analysis was performed at the following temperatures (from top to bottom): (—) 5°C, (---) 20°C, (—) 37°C, and (- - -) 45°C.

Table 1

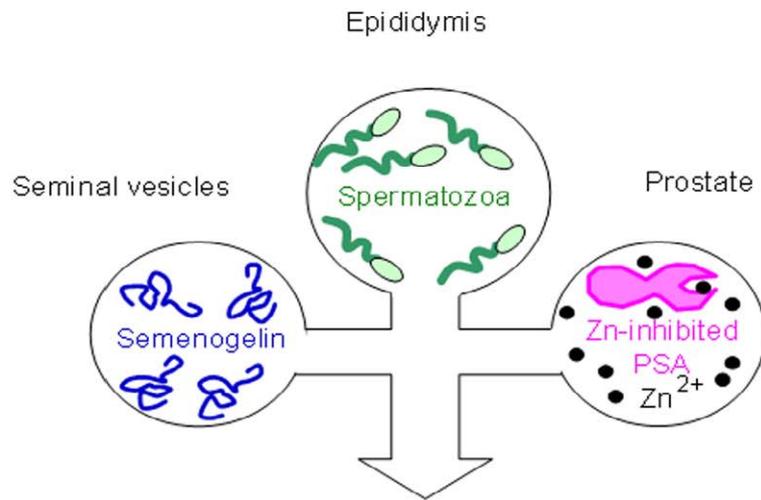
Program	Temperature °C	240—208 nm		
		$\alpha$ *	$\beta$ *	Related protein
CONTINLL	5	6	29	GFP
	20	7	26	GFP
	37	8	26	GFP
	45	9	24	GFP
	65	9	23	GFP
	90	9	22	GFP
Selcon 3	5	10	35	GFP
	20	9	23	GFP
	37	9	23	GFP
	45	10	24	GFP
	65	13	19	GFP
	90	11	24	GFP
CDSSTR	5	4	26	
	20	5	24	
	37	6	25	
	45	5	24	
	65	6	25	
	90	7	23	

\* % of total sum of secondary structure

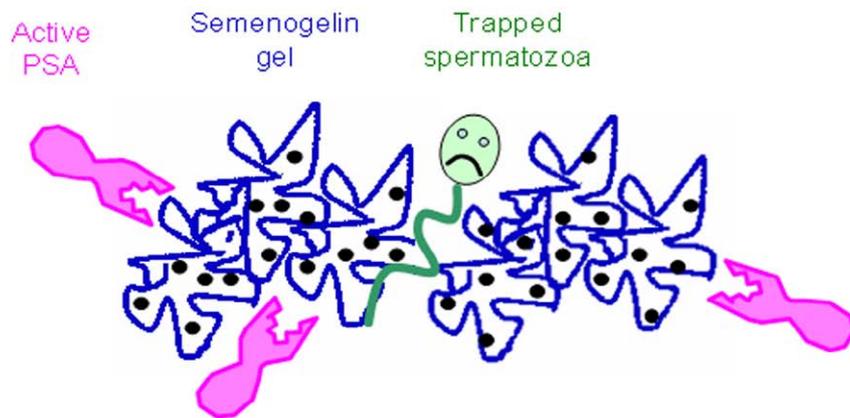
---

Figure 1

**A**



**B**



**C**

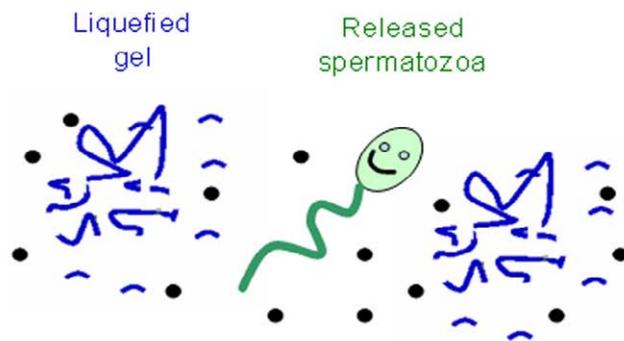


Fig. 2

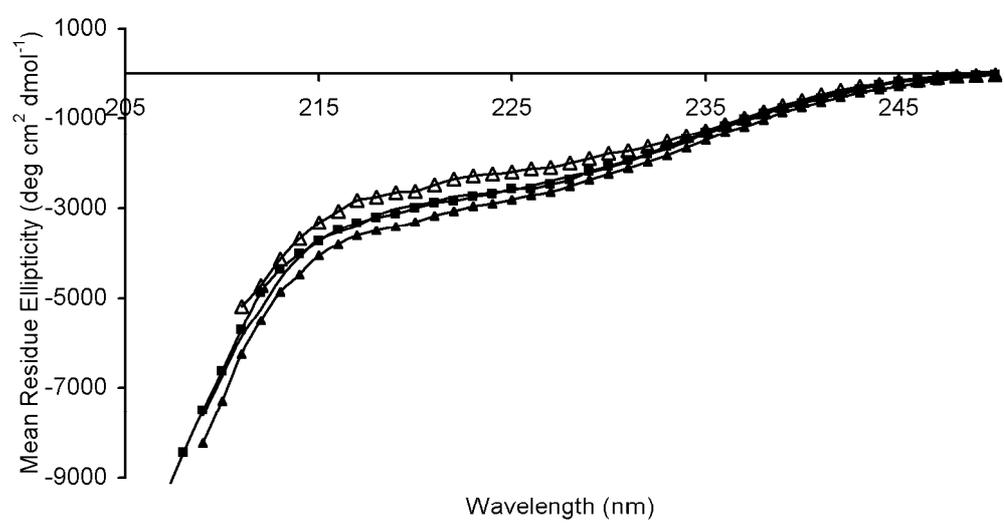


Fig.3

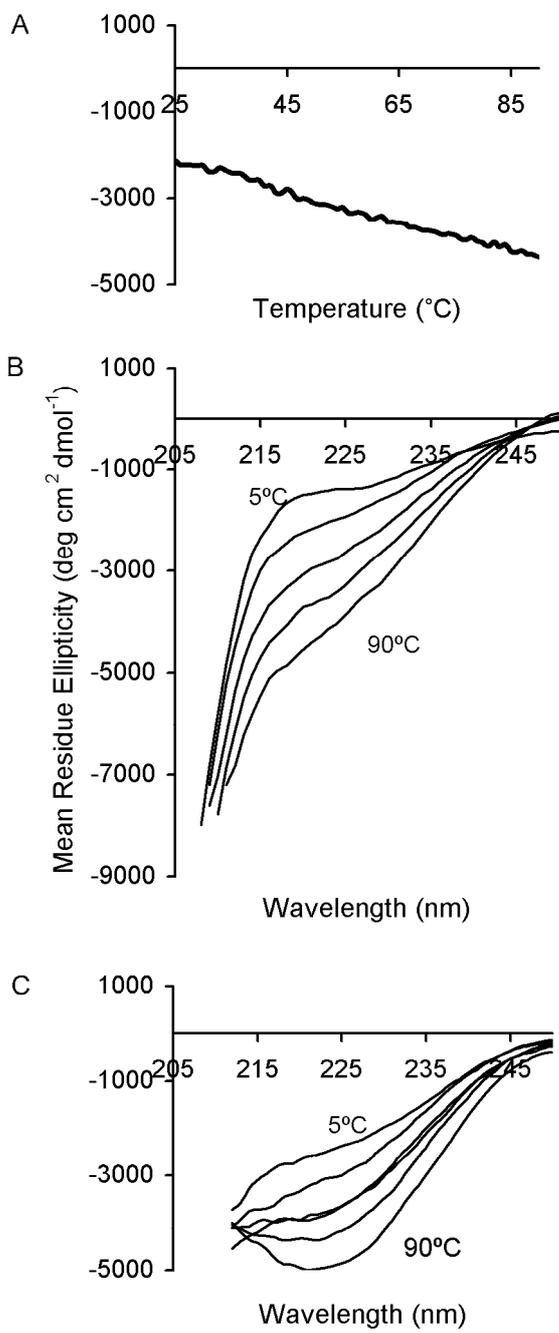


Fig. 4

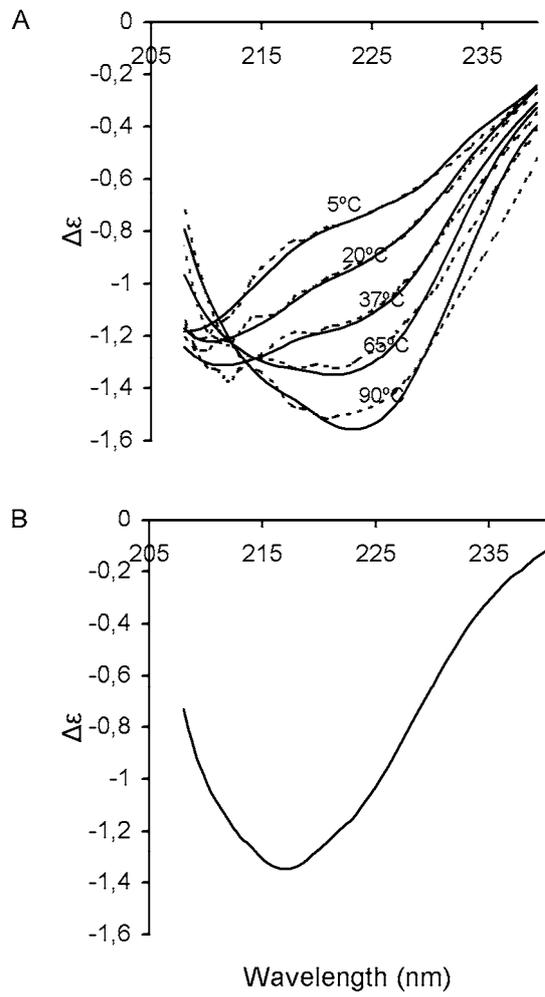


Fig. 5

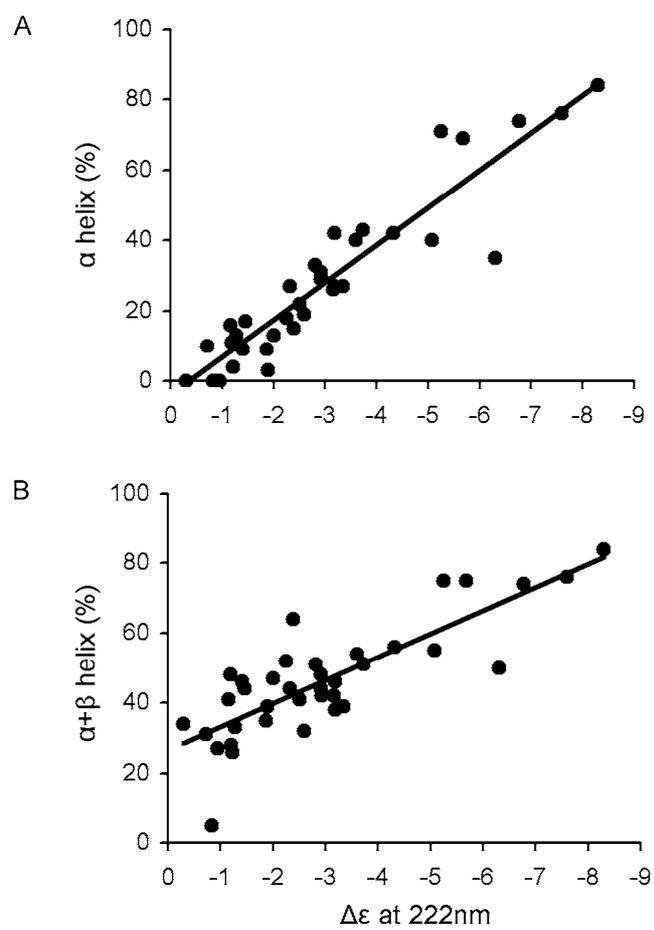


Fig. 6

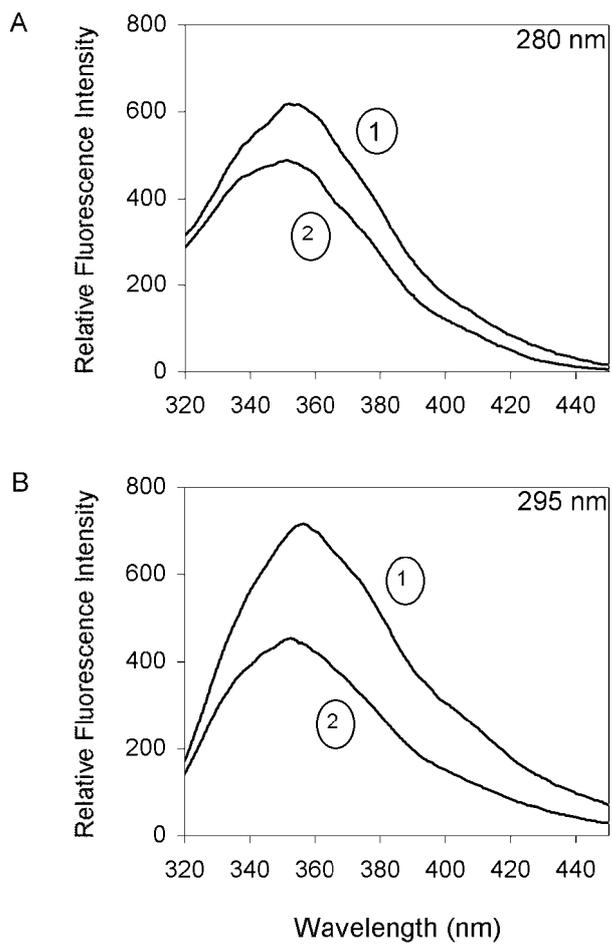


Fig. 7

