

A Peptide from Human Semenogelin Self-Assembles into a pH Responsive Hydrogel

Birgitta Frohm, Jamie E. DeNizio, David S. M. Lee, Luigi Gentile, Ulf Olsson, Johan Malm, Karin S. Åkerfeldt, and Sara Linse

Supplementary Information

Semenogelin I

Semenogelin I, SgI, is the predominant structural protein in the loose gel formed in freshly ejaculated human semen. The protein displays high Zn^{2+} -binding capacity and a Zn^{2+} -dependent structure.¹ The protein binds to spermatozoa² and appears to play a role in optimizing conception.³ The 439 amino acid sequence is similar to intrinsically unfolded proteins,⁴ and contains six repeats with pair-wise high sequence similarity (Figure SI-1). The peptide studied here corresponds to residues 38-48 in SgI. The prostate-specific antigen (PSA), the major protease in seminal plasma, cleaves SgI at several specific sites (Figure SI-1, SI-2).⁵ One cleavage site, at residue 44, is located in the center of the gel forming peptide, indicating that the 38-48 region is important for gel formation and liquefaction in the natural processing of the protein.

Semenogelin I repeats

Similarity class I

261 NKISYQSSST EERRLHYGEN GVQKDVSQSS IYSQTEEKAQ GKSQKQITIP SQEQEHSQKA
321 NKISYQSSST EERRLHYGEN GVQKDVSQRS IYSQTEKLVA GKSQKQIQAP NPKQEPWHGE

Similarity class II

141 ERLWVHGLSK EQTSVSGAQK GRKQGGSS YVLQTEELVA NKQORETKNS HQNKGHYQNV
201 VEVREEHSSK VQTSLCPAHQ DKLQHGSKDI FSTQDELLVY NKNQHQTKNL NQDQQHGRKA

Similarity class III

381 KGESGQSTNR EQDLLSHEQK GRHQHGSHGG LDIVIEQED DSDRHLAQHL NNDRNPLFT
70 TTKSQRHLGG SQQLLHNKQE GRDHDKSKGH FHRVVIHHKG GKAHRGTQNP SQDQGNPSG

Figure SI-1. Sequence repeats of human semenogelin I, with the PSA cleavage sites indicated by arrows.

PSA cleavage map

QKGGSKGRLPSEFSQFPHGQKQGHYSGQKQTESK**GSFSIQYTYHV**DANDHDQSRKSQQYDLNALHKT-70
TKSQRHLGGSSQQLLHNKQEGRDHDKSKGHFHRVVIHHKGGKAHRGTQNPSSQDQGNPSGKGISSQYSNTE-140
ERLWVHGLSKEQTSVSGAQKGRKQGGSSYVLQTEELVANKQORETKNSHQNKGHYQNVVEVREEHSSK-210
VQTSLCPAHQDKLQHGSKDIFSTQDELLVYNKNQHQTKNLNQDQQHGRKANKISYQSSSTEERRLHYGEN-280
GVQKDVSQSSIYSQTEEKAQGKSQKQITIPSQEQEHSQKANKISYQSSSTEERRLHYGENGVQKDVSQRS-350
IYSQTEKLVAGKSQIQAPNPKQEPWHGENAKGESGQSTNREQDLLSHEQKGRHQHGSHGGLDIVIEQED-420
DSDRHLAQHLNNDRNPLFT-439

Figure SI-2. Amino acid sequence of human semenogelin I with PSA cleavage sites as arrows. The hydrogel peptide sequence (residues 38-48) is highlighted in red.

Peptide purification

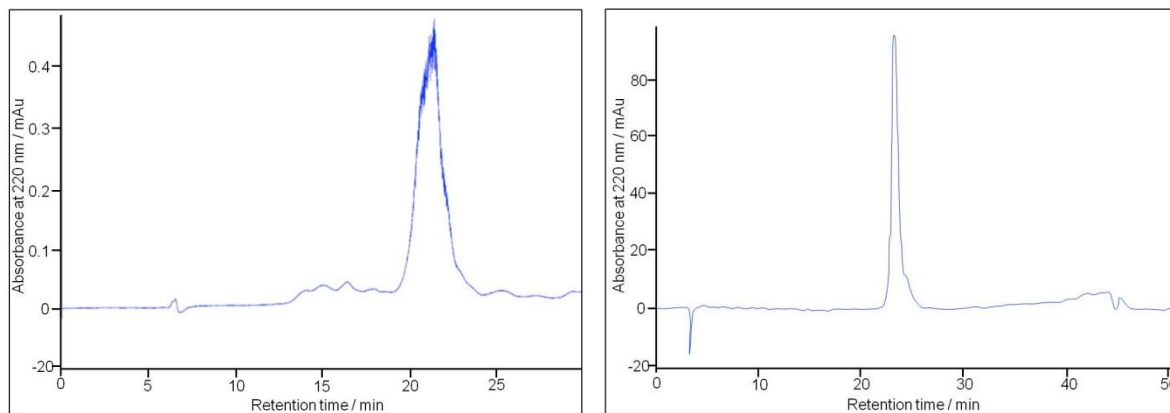


Figure SI-3. Left: Purification of crude SgI:38-48 peptide by reversed phase HPLC using a preparative column (Vydac C4, 10-15 μm , 22 x 250 mm). **Right:** Analytical HPLC of purified SgI:38-48 peptide using an analytical column (Vydac C4, 5 μm , 4.6 x 250 mm).

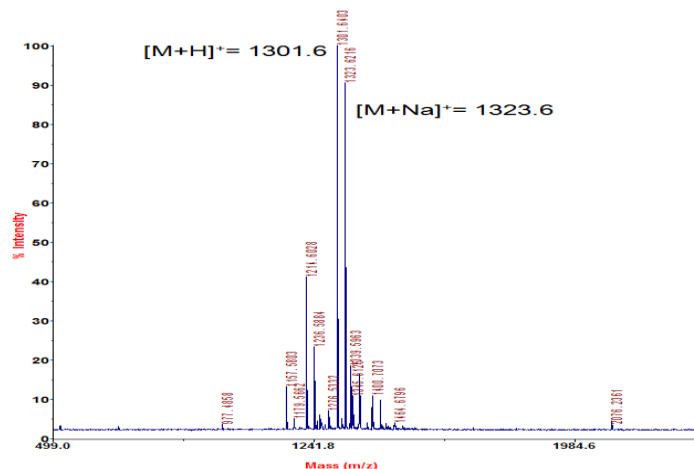


Figure SI-4: Maldi-TOF mass spectrometry of purified SgI:38-48 peptide.

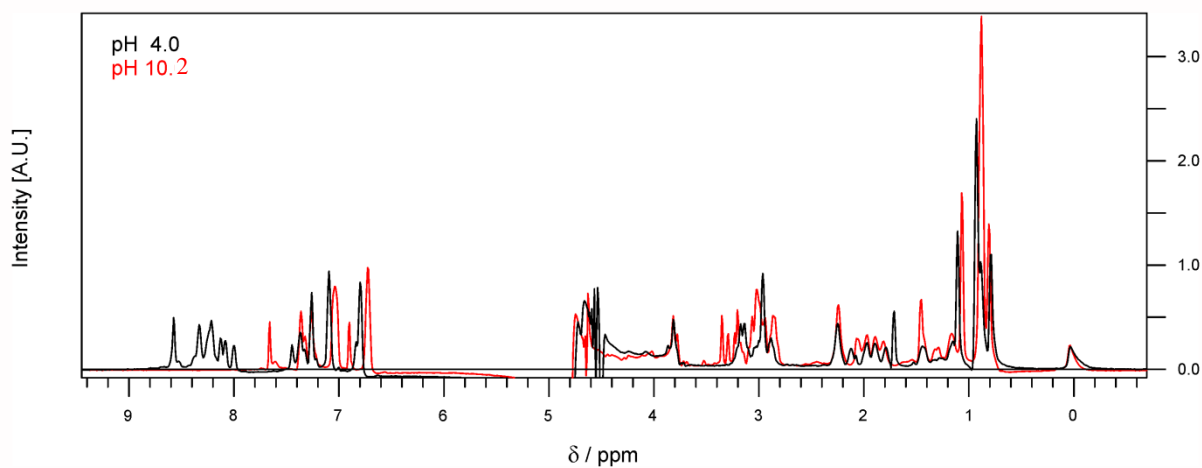


Figure SI-5: Overlay of ^1H NMR spectra of purified SgI:38-48 peptide at pH 4.0 (red) and 10.2 (black) in H_2O with 10% D_2O .

Parallel versus anti-parallel arrangement

To evaluate the likelihood for anti-parallel versus parallel β -sheet arrangement of Sg38-48, a potential energy was calculated for each orientation using the empirical parameters by Keskin *et al.*⁶ for solvent-mediated and residue-residue contact potential (Figure SI-5). The calculations were made in register and for each possible out-of-register arrangement. This analysis suggests that the in-register arrangements are most stable, with a slight preference for anti-parallel over parallel arrangement. This analysis does not take into account the electrostatic repulsion between end charges, which would further disfavor the parallel orientation. In summary, the in-register anti-parallel arrangement is most likely. The second best arrangement seems to be a two-residue-out-of register mode +2 or -2 (=+9).

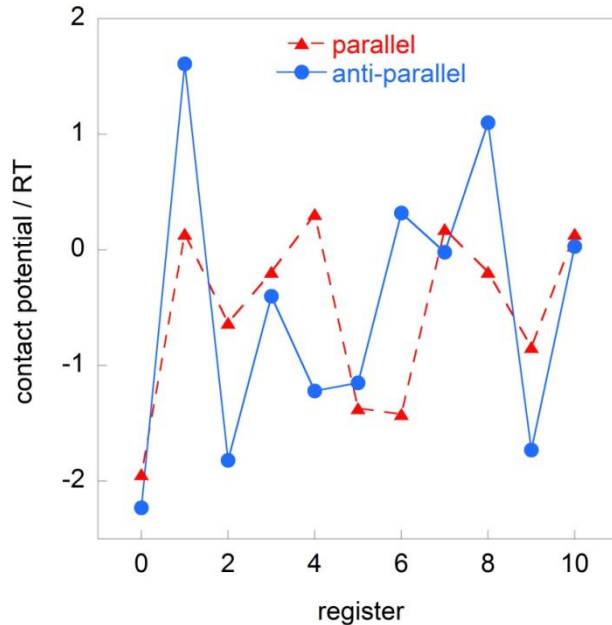


Figure SI-6. Contact potential summed over all pair-wise interactions in each possible in-register and out-of-register parallel (red triangles) and anti-parallel (round blue dots) arrangements, based on Table 1 in Keskin et al 1998.

References

1. M Jonsson, S. Linse, B. Frohm, Å. Lundwall, J. Malm. *Biochem. J.* 2005, **387**, 477-453
2. M. Jonsson, B. Frohm, J. Malm, J. Androl. 2010, **31**, 560-565.
3. K.P. Tremellen, D. Valbuera, J. Landeras, A. Ballesteros, J. Martinez, S. Mendoza, R.J. Norman, S. A. Robertson, C. Simón. *Hum. Reprod.* 2000, **15**, 2653-2658.
4. V.N. Uversky. *J Biomed. Biotechnol.* 2010, 2010, 568068.
5. J. Malm, J. Hellman, P. Hogg, H. Lilja. *Prostate* 2000, **45**, 132-139.
6. O. Keskin, I. Bahar, A.Y. Badretdinov, O.B. Ptitsyn, R.L. Jernigan. *Protein Science* 1998 **7**, 2578-2586