

Nematode sperm maturation triggered by protease involves sperm-secreted serine protease inhibitor (Serpine)

Yanmei Zhao^a, Wei Sun^{a,b,1}, Pan Zhang^{c,1}, Hao Chi^{b,d,1}, Mei-Jun Zhang^{c,1}, Chun-Qing Song^c, Xuan Ma^a, Yunlong Shang^{a,b}, Bin Wang^a, Youqiao Hu^{a,b}, Zhiqi Hao^e, Andreas F. Hühmer^e, Fanxia Meng^a, Steven W. L'Hernault^f, Si-Min He^d, Meng-Qiu Dong^{c,2}, and Long Miao^{a,2}

^aLaboratory of Noncoding RNA, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; ^bGraduate School, Chinese Academy of Sciences, Beijing 100049, China; ^cNational Institute of Biological Sciences, Beijing, Beijing 102206, China; ^dKey Lab of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China; ^eThermo Fisher Scientific, San Jose, CA 94539; and ^fDepartment of Biology, Emory University, Atlanta, GA 30322

Edited by Timothy L. Karr, Arizona State University, Tempe, AZ, and accepted by the Editorial Board December 22, 2011 (received for review June 19, 2011)

Spermiogenesis is a series of poorly understood morphological, physiological and biochemical processes that occur during the transition of immotile spermatids into motile, fertilization-competent spermatozoa. Here, we identified a Serpin (serine protease inhibitor) family protein (As_SRP-1) that is secreted from spermatids during nematode *Ascaris suum* spermiogenesis (also called sperm activation) and we showed that As_SRP-1 has two major functions. First, As_SRP-1 functions *in cis* to support major sperm protein (MSP)-based cytoskeletal assembly in the spermatid that releases it, thereby facilitating sperm motility acquisition. Second, As_SRP-1 released from an activated sperm inhibits, *in trans*, the activation of surrounding spermatids by inhibiting vas deferens-derived As_TRY-5, a trypsin-like serine protease necessary for sperm activation. Because vesicular exocytosis is necessary to create fertilization-competent sperm in many animal species, components released during this process might be more important modulators of the physiology and behavior of surrounding sperm than was previously appreciated.

cell motility | regulated exocytosis | sperm competition | postcopulatory sexual selection | de novo sequencing

In most, if not all, animals, males produce sperm in their gonad that are fertilization-incompetent until they leave this organ and undergo further maturation. For instance, whereas mammalian spermatozoa form in the testes, they must undergo a maturation process called capacitation in the female reproductive tract before they become fertilization-competent (1). In nematodes, spermatids do not complete maturation into spermatozoa (spermiogenesis) until after they have left the testes. In the nematode *Caenorhabditis elegans*, sperm are made in both males and self-fertile hermaphrodites (there are no conventional females). Hermaphrodites have a testis that proliferates sperm and then it switches into an ovary and produces oocytes. The first ovulated oocyte pushes the stored spermatids from the gonad into the spermatheca, where they are rapidly activated into spermatozoa. Upon mating with hermaphrodite, male spermatids are ejaculated and activated within the uterus by exposure to unknown factor(s) in the seminal fluid (2). Male-derived sperm are preferentially used to promote outbreeding in a typical cross (3). This sperm precedence correlates with the larger size of male-derived sperm relative to hermaphrodite-derived sperm (4) and can even occur in certain fertilization-defective mutants (5). Although *in vitro* and *in vivo* studies have suggested that protease activity is involved in *C. elegans* sperm activation (6, 7), neither hermaphrodite nor male sperm activators have been identified.

Unlike *C. elegans*, the intestinal parasitic nematode *Ascaris suum* (or *Ascaris* hereafter) has males and true females, but no hermaphrodites. However, sperm of both species do share several similarities (8). First, they are unusual in that their activated spermatozoa lack flagella. Rather, nematode spermatozoa move by

using pseudopods generated during spermiogenesis, also called sperm activation (2). Second, unlike other types of amoeboid motility that are based on actin, the motility of nematode spermatozoa is based on controlled assembly/disassembly of a major sperm protein (MSP) cytoskeleton (8). Third, like male-derived sperm in *C. elegans*, *Ascaris* sperm activation occurs postinsemination. Fourth, sperm of both *C. elegans* and *Ascaris* contain structurally similar membranous organelles (MOs) (2), which is a type of intracellular vesicle with similarity to lysosomes (9). During sperm activation, fusion of MOs with the plasma membrane (PM) of spermatids is necessary for spermatozoan motility and male fertility (10, 11). However, the exact function of MOs and their components that are released into the extracellular space during fusion are not well understood.

Ascaris sperm are highly suitable for answering questions about how sperm prepare for fertilization because: sperm activation can be studied *ex vivo* (12), sperm motility has been reconstituted in cell-free sperm extracts (13, 14), and all relevant components can be obtained in the large quantities required for biochemical analysis (12, 15). In this study, we identified two *Ascaris* proteins, As_SRP-1 [a member of the Serpin (serine protease inhibitor) superfamily] and As_TRY-5 (a trypsin-like serine protease). We showed that nematode sperm maturation triggered by vas deferens-derived As_TRY-5 involves sperm-secreted As_SRP-1 and that secreted As_SRP-1 in the medium inhibits activation of surrounding spermatids. This dual function of sperm-secreted As_SRP-1 might play a significant role during postcopulatory sexual selection.

Results

As_SRP-1 (1CB4 antigen) Is Translocated During *Ascaris* Sperm Activation. We found that the 1CB4 monoclonal antibody that recognizes *C. elegans* MOs (11, 16–18) also recognized *Ascaris* sperm MOs (Fig. 1 *A* and *B*). Immunofluorescence staining of permeabilized *Ascaris* spermatids or spermatozoa with 1CB4 revealed

Author contributions: M.-Q.D. and L.M. designed research; Y.Z., W.S., P.Z., H.C., M.-J.Z., C.-Q.S., X.M., Y.S., B.W., Y.H., S.-M.H., M.-Q.D., and L.M. performed research; Z.H., A.F.H., S.W.L., and S.-M.H. contributed new reagents/analytic tools; Y.Z., W.S., P.Z., H.C., M.-J.Z., C.-Q.S., X.M., Y.S., B.W., Y.H., F.M., S.W.L., S.-M.H., M.-Q.D., and L.M. analyzed data; and Y.Z., F.M., S.W.L., M.-Q.D., and L.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. T.L.K. is a guest editor invited by the Editorial Board.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession nos. JF894302 (*As_srp-1*) and JF894303 (*As_srp-1*)].

¹W.S., P.Z., H.C., and M.-J.Z. contributed equally to this work.

²To whom correspondence may be addressed. E-mail: lmiao@moon.ibp.ac.cn or dongmengqiu@nibs.ac.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1109912109/-DCSupplemental.

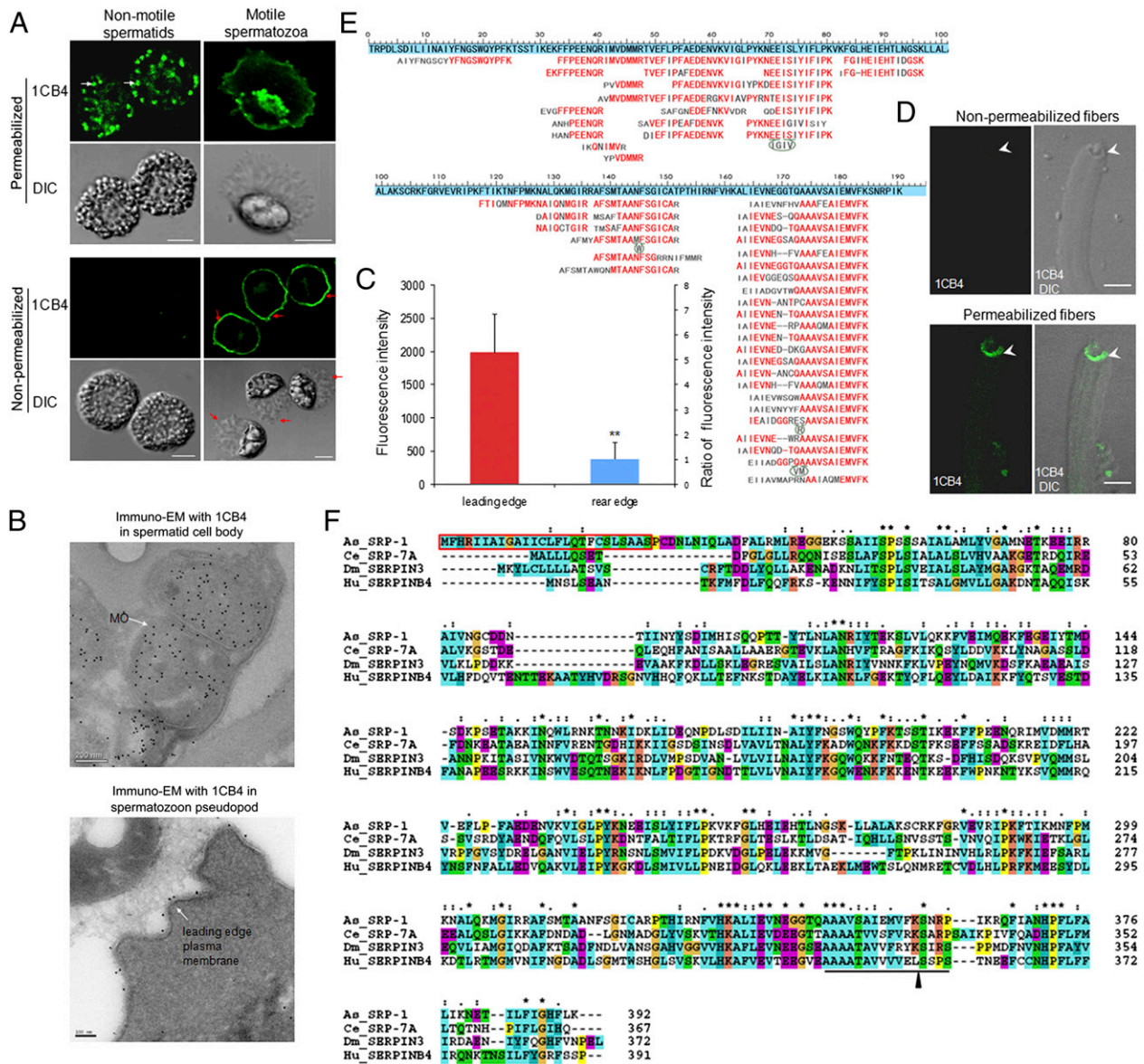


Fig. 1. As_SRP-1 (protein recognized by the 1CB4 antibody) is translocated during *Ascaris* sperm activation. (A) The 1CB4 monoclonal antibody labeled MOs and the leading edge of spermatozoon PM in *Ascaris*. White arrows, MOs; red arrows, the leading edge of the pseudopod. (Scale bars, 5 μ m.) (B) 1CB4 immunogold was detected in the MO of spermatid (Upper) and the outer PM of spermatozoon leading edge (Lower) by Cryo-immuno-EM. [Scale bars, 200 nm (Upper) and 100 nm (Lower).] (C) 1CB4 fluorescence intensity at the leading edge and rear edge was measured in nonpermeabilized spermatozoa (A, Lower) by MetaMorph. Results are means \pm SD ($n = 50$ spermatozoa). ** $P < 0.01$ (Student t test). (D) In an in vitro MSP fiber assembly assay, the 1CB4 immunostaining (green) was detected only in vesicles that were permeabilized. (Scale bars, 5 μ m.) (E) The protein recognized by the 1CB4 antibody was identified as a Serpin family protein by de novo sequencing (*SI Materials and Methods*). Set against a blue background is the partial As_SRP-1 protein sequence translated from *Ascaris* expressed sequence tag (EST) sequences. Sequences deduced from mass spectra of As_SRP-1 peptides generated by trypsin. Red, matched residues; gray, unmatched residues; dash, a gap; bulge, extra residues found in the de novo peptide sequences. (F) Amino acid sequence of As_SRP-1 was aligned with three other Serpins, including Ce_SRP-7A (*C. elegans*, NP_001023823), Dm_SERPIN3 (*D. melanogaster*, NP_524956), and Hu_SERPINB4 (human, NP_002965). The reactive site loop (RSL) is underlined; arrowhead, the putative scissile bond; red box, a predicted signal peptide; asterisk, identical amino acid; colon, amino acid with high similarity; dot, amino acid with less similarity.

punctuate, peripherally located structures, similar to what is seen in *C. elegans* (11, 17). Cryo-immuno-EM with 1CB4 confirmed that immunogold labeled tightly-packed stacks of membranes inside sperm (Fig. 1B, Upper), characteristics of MOs in *C. elegans* (2). Different from previous immunofluorescence studies in *C. elegans*, 1CB4 also stained the leading edge PM of *Ascaris* spermatozoon (Fig. 1A). The 1CB4 staining in the leading edge spermatozoon is further shown to be on the outer PM by three lines of our evidence. First, in nonpermeabilized spermatozoa, 1CB4 immunofluorescence was readily observed on the cell surface (Fig. 1A). Second,

Cryo-immuno-EM with 1CB4 revealed the clear immunogold labeling along the outer PM of spermatozoa (Fig. 1B, Lower). Third, from an in vitro MSP motility assay (Fig. 1D), in which the leading edge PM-derived vesicles from spermatozoa extracts recruit cytosolic components to trigger MSP fiber assembly (13), we found that 1CB4 immunofluorescence could be detected only in permeabilized fiber-growing vesicles. Given that these vesicles acquire an inside-out configuration during cell lysis (13), this result is consistent with the outer PM-localization of the 1CB4 target in *Ascaris* spermatozoa. Moreover, immunofluorescence quanti-

fication of the 1CB4 staining in nonpermeabilized spermatozoa demonstrates that the 1CB4 on the outer PM of spermatozoa was distinctly asymmetrical, i.e., the fluorescence intensity along the leading edge PM was 5.3-fold higher than that in the rear edge PM (Fig. 1C), in agreement with previous observations by quantitative immuno-EM in *C. elegans* spermatozoa (11).

1CB4 is a monoclonal antibody generated using homogenates of whole *C. elegans* (16). Although it has been extensively used for labeling MOs in *C. elegans* (11, 17), the molecular identity of the antigen recognized by 1CB4 has not been determined. By using Western blotting, we found that a single polypeptide (~46 kDa) is recognized by 1CB4 in *Ascaris* sperm extract, and it was mostly in a soluble, cytosolic fraction (Fig. S1A). Isolation and purification of the 46 kDa protein were achieved by following the 1CB4 signal in Western blots from different cellular fractions (Fig. S1B). Initial MS analysis of the purified protein using a conventional database search strategy was ineffective because this protein was not in the database. We resorted to de novo sequencing analysis using the pNovo program (19) and extracted sequences directly from the tandem mass spectra of peptides derived from this protein (Fig. 1E and Fig. S2). We synthesized two peptides according to the pNovo result and found that the identification of these two sequences was fully supported by the parent masses and high-resolution MS/MS spectra of the synthetic peptides (Fig. S2). BLAST searches of these peptides against predicted *Ascaris* protein sequences in NEMBASE3 (20) revealed that the most abundant protein in the sample was a Serpin (Fig. 1E), belonging to the Serpin superfamily (we named it As_SRP-1). Using rapid amplification of cDNA ends by PCR (RACE PCR), we cloned the full-length cDNA of *As_srp-1* and deduced its amino acid sequence (Fig. 1F). When the original MS

data were searched against a database containing the newly cloned As_SRP-1 sequence using either Mascot or pFind, As_SRP-1 was identified as the top hit, and the overlap between the database search result and the de novo sequencing result was extensive (Figs. S3 and S4). Amino acid alignment showed that As_SRP-1 shares strong sequence homology with members of the clade B Serpin family. A highly conserved reactive site loop (RSL) containing a putative scissile bond (21) was detected in the sequence of As_SRP-1 (Fig. 1F). When expressed in *E. coli*, the recombinant As_SRP-1 displayed the same molecular mass as that of native As_SRP-1 and was recognized by both 1CB4 and the polyclonal antibody we raised against purified native As_SRP-1 (Fig. S1C). These data demonstrate that the target of the 1CB4 monoclonal antibody in *Ascaris* is As_SRP-1.

As_SRP-1 Is Essential for MSP-Based Sperm Motility in *Ascaris*. The localization of As_SRP-1 on the outer PM of spermatozoon and its asymmetrical distribution at the leading edge (Fig. 1A–D) suggest that this protein probably plays a role in MSP cytoskeleton dynamics and sperm motility. To examine this possibility, we performed both ex vivo and in vitro experiments. When spermatozoa were perfused with the As_SRP-1 antiserum (1:100 or 1:50 dilution), spermatozoa stopped crawling, their MSP cytoskeleton disappeared (66% or 98%, respectively), and cells rounded up (Fig. 2A and C). These defects in cytoskeleton dynamics and sperm morphology were almost completely reversed when the antiserum was first neutralized by adding purified native As_SRP-1, with < 5% of spermatozoa exhibiting defects (Fig. 2A and C). Furthermore, As_SRP-1 localization on the inner leaflet of the vesicle membrane (equivalent to outer PM) (Fig. 1D) is important for MSP fiber assembly in vitro

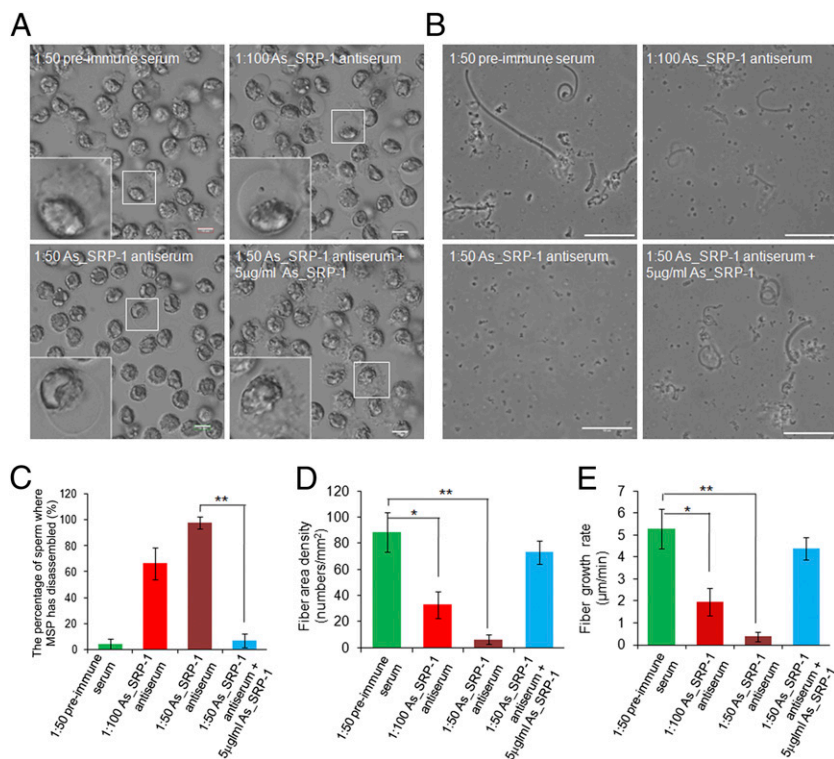
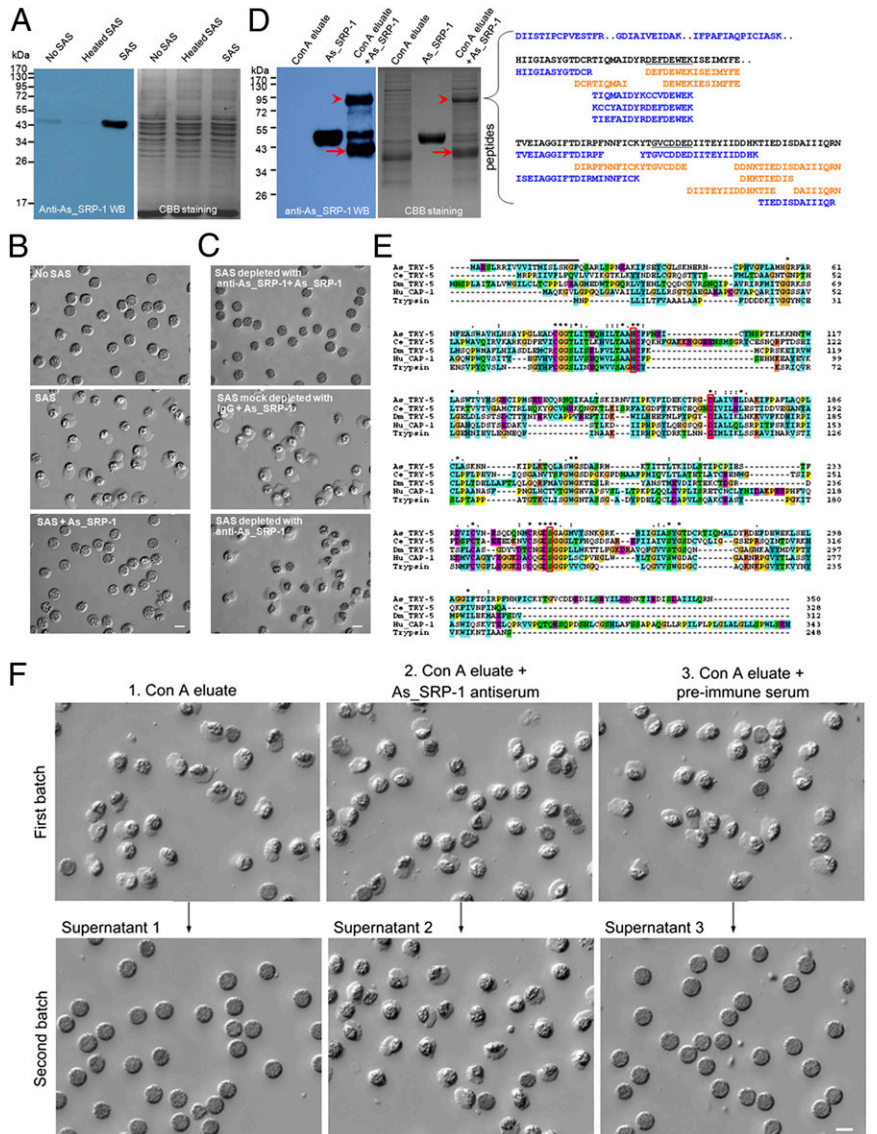


Fig. 2. As_SRP-1 is essential for MSP-based cytoskeleton dynamics and sperm motility in *Ascaris*. (A) As_SRP-1 antiserum treatment of spermatozoa for 20 min caused MSP cytoskeleton disassembly and spermatozoon roundup, although adding 5 µg/mL As_SRP-1 rescued this phenotype. *Upper Left*, 1:50 preimmune serum treatment (control). *Insets*, higher magnification of outlined sperm. (Scale bars, 20 µm.) The sperm cytoskeletal disassembly was quantified (C). (B) An in vitro MSP fiber assembly assay where the extracts from spermatozoa treated with As_SRP-1 antiserum had fewer fibers and a slower fiber growth rate, whereas adding 5 µg/mL As_SRP-1 rescued the defects. *Upper Left*, MSP fibers assembled with the extract from spermatozoa treated with 1:50 preimmune serum (control). (Scale bars, 50 µm.) The area density and growth rates of assembled fibers were calculated by MetaMorph (D and E). The data shown in C–E are means ± SD ($n = 5$ experiments). * $P < 0.001$; ** $P < 0.0001$ (Student t test).

Fig. 3. Secreted As_SRP-1 blocks other sperm activation by inhibiting As_TRY-5 activity in SAS. (A) Secretion of As_SRP-1 during sperm activation increased. Sperm were pelleted by centrifugation after being treated with buffer alone (No SAS), heated SAS (SAS was heat-inactivated for 10 min before use) or SAS (0.5 μg/mL SAS) for 10 min, and then the supernatants were subjected to SDS/PAGE and Western blot with anti-As_SRP-1. (B) As_SRP-1 inhibited SAS-induced sperm activation. Purified native As_SRP-1 (0.5 μg/mL) was incubated with 0.5 μg/mL SAS for 10 min and then tested for sperm activation capability. The HKB buffer (No SAS) and SAS were used as negative and positive controls, respectively. (Scale bar, 10 μm.) (C) As_SRP-1 interacts physically with a serine protease present in SAS. Before incubation with SAS, As_SRP-1 was immobilized onto protein A beads through its antibody, and the collected supernatant lost its sperm activating activity (Top). Middle and Bottom, mock depletion, either As_SRP-1 antibody was replaced with IgG or As_SRP-1 was omitted for binding with beads before incubation with SAS. (Scale bar, 10 μm.) (D) Identification of As_TRY-5. Left: purified As_SRP-1 was incubated with the Con A eluate (see Fig. S8A) at 4 °C overnight, followed by SDS/PAGE and Western blot with anti-As_SRP-1. Arrowheads, complex formed between As_SRP-1 and a protease in the Con A eluate. Arrows, cleaved As_SRP-1. Right, De novo sequencing of the 90-kDa complex revealed peptide sequences that led to the identification of As_TRY-5. The 90-kDa complex was digested with trypsin, Asp-N and Lys-N before subjected to LC-MS/MS (SI Materials and Methods). Black, consensus sequence; blue, tryptic peptides; orange, Asp-N peptides; the underlined, sequences used for designing PCR primers to clone the cDNA encoding this protein, “...” indicates where peptide sequences are disconnected. BLAST search with these extended sequences revealed that they are highly homologous to trypsin-like protease protein 5 in *Brugia malayi* (XP_001894231). Sequence alignment of known nematode serine proteases helped place all except the first peptide (DIISTIPCPVESTFR) in relative order. (E) Comparison of the amino acid sequence of As_TRY-5 with other known serine proteases, including Ce_TRY-5 (*C. elegans*, NP_505421), Dm_TRY-5 (*Drosophila melanogaster*, NP_001163100), Hu_CAP-1 (human, NP_002764), and human trypsin (NP_002760). Red boxes, the catalytic triad (H, D and S); dark line, a predicted signal peptide; asterisk, identical amino acid; colon, amino acid with high similarity; dot, amino acid with less similarity. (F) Inhibitory effect of the secreted As_SRP-1 on sperm activation could be rescued by the addition of neutralizing antibody. The first batch of spermatids (upper images) was incubated with Con A eluate alone (Left), Con A eluate plus 1:50 As_SRP-1 antiserum (Center), or Con A eluate plus 1:50 preimmune serum (Right) at 37 °C for 10 min. For each of these three assay conditions, spermatids were activated (upper images) and the supernatants were collected after centrifugation to treat the second batch of spermatids (lower images). When secreted As_SRP-1 during the activation of the first batch of sperm (see A) was not neutralized by As_SRP-1 antiserum (Left and Right), the activation of the second batch of spermatids was inhibited, indicating the participation of As_SRP-1 in other sperm maturation. (Scale bar, 10 μm.)



because extracts from As_SRP-1 antiserum-treated spermatozoa resulted in a significant decrease in both the growth rate and area density of MSP fibers. Such motility defects disappeared in the extracts from spermatozoa that were perfused with the As_SRP-1 antiserum and the neutralizing As_SRP-1 (Fig. 2 B, D, and E). Together, these data indicate that As_SRP-1 at the outer PM of sperm pseudopod leading edge plays an essential role in regulating both MSP cytoskeleton assembly and cell motility. Our immunolabeling data suggest that As_SRP-1 regulates *Ascaris* sperm motility probably through protein tyrosine phosphorylation (Fig. S5).

Secreted As_SRP-1 Blocks Sperm Activation in Surrounding Spermatids. As shown in Fig. 1, the As_SRP-1 localization in *Ascaris* spermatids (in MOs) is different from that in spermatozoa (on the outer PM). A secretory signal peptide sequence is present at the N terminus of As_SRP-1 (Fig. 1F). Because the fusion of MOs with the PM during sperm activation is known to be

a regulated exocytosis process in *C. elegans* (11), exocytosed As_SRP-1 could be translocated from MOs to the outer PM during *Ascaris* sperm activation. Indeed, when *Ascaris* spermatids were activated by sperm-activating substance (SAS) (the extract from vas deferens) (12), the amount of As_SRP-1 in the medium increased dramatically as shown by Western blotting analysis using anti-As_SRP-1 antibody (Fig. 3A). In contrast, only a weak signal of As_SRP-1, most likely from rare, spontaneous activation, was detected in the medium of sperm that were either not subjected to SAS or subjected to heat-inactivated SAS. The secretion of a Serpin (As_SRP-1) during sperm activation (Fig. 3A) and the ability of proteases (6, 7, 12) to activate nematode sperm led us to test whether As_SRP-1 can inhibit SAS-induced sperm activation and whether the activity of a serine protease(s) in SAS is essential for sperm activation. Not surprisingly, we found that purified As_SRP-1 was able to inhibit SAS-induced sperm activation (Fig. 3B). Further experiments showed the inhibitory

target of As_SRP-1 was present in SAS but not on sperm itself because when sperm were treated with the mixture of As_SRP-1 and 0.5 or 5 $\mu\text{g}/\text{mL}$ of SAS, the sperm activation rate was lower than that from sperm treated with As_SRP-1 first, then followed by SAS addition (Fig. S6). The importance of a serine protease activity in SAS-induced sperm activation is also supported by our pharmacological studies in which the effect of various specific protease inhibitors on SAS was tested. These results demonstrate that only serine protease inhibitors, and not other inhibitors, prevent SAS-induced sperm activation (Fig. S7). Furthermore, we found that As_SRP-1 could interact with the predicted serine protease(s) in SAS using immunodepletion assays (Fig. 3C). We immobilized As_SRP-1 to protein A beads through As_SRP-1 antibody, incubated the As_SRP-1 beads with SAS and separated the SAS supernatant from the beads. If physical interaction occurs, the protease should be depleted from SAS by As_SRP-1 beads, causing the supernatant to lose its activity. This was indeed what we observed (Fig. 3C, *Top*). Meanwhile in the control experiments, Protein A beads with mock immobilization of As_SRP-1 through control IgG (Fig. 3C, *Middle*) or beads preloaded with As_SRP-1 antibody alone (Fig. 3C, *Bottom*) failed to deplete the activity from the SAS supernatant. Therefore, As_SRP-1 can physically bind to the serine protease(s) in SAS. Collectively, these data suggest that the activity of a serine protease(s) in SAS is critical for *Ascaris* sperm activation and the secreted As_SRP-1 likely inhibits sperm activation through its physical interaction with this protease(s).

Although our data (Fig. 3B and C and Fig. S7) and accumulating evidence (6, 7, 12) suggest that the activation of nematode sperm involves a serine protease(s), the identity of this protease(s) has been unknown. To identify the protease, we used conventional biochemical purification strategies to enrich the target protein by following its sperm activating activity (Fig. S8A). The Con A eluate showed strong activity in inducing sperm activation, and this activity could be inhibited by the serine protease inhibitor PMSF (Fig. S8B), suggesting that the fraction from Con A contains our target serine protease(s). This proteolytic fraction was shown to interact physically with As_SRP-1 (Fig. 3C), and the interaction was predicted to produce a large covalent protein complex containing the cleaved As_SRP-1 and target protease(s), according to the well-characterized Serpin–protease interaction mechanism (21). Indeed, as shown on both SDS/PAGE and Western blot, an $\sim 90\text{-kDa}$ band (Fig. 3D, *Left*) appeared after As_SRP-1 ($\sim 46\text{ kDa}$) was incubated with the Con A eluate. This 90-kDa band was then subjected to MS analysis and de novo peptide sequencing to identify the protease because its sequence was not present in existing databases. Using the pNovo algorithm (19), we obtained over a dozen high-quality peptide sequences that did not belong to any previously characterized protein (Fig. 3D, *Right*, and Fig. S9). A synthetic peptide was obtained for one of them, and its fragmentation spectra were found to be identical to those of the endogenous peptide (Fig. S9), thus validating the de novo sequencing results. We assembled these sequences into longer segments and found by BLAST search that they share homology with a trypsin-like serine protease (Fig. 3D, *Right*). Based on the peptides identified from de novo sequencing, we designed degenerative primers for RACE PCR and cloned the full-length cDNA (Fig. 3E). Sequence comparisons indicate that the protein encoded by this cDNA shares a high degree of homology, including a conserved catalytic triad, with other known serine proteases (Fig. 3E). We named this protein As_TRY-5 after its closest homolog, TRY-5, in *C. elegans*. Again, analysis of the original MS data against full-length TRY-5 using Mascot and pFind further confirmed the accuracy of sequence identifications made by pNovo (Figs. S10 and S11).

We further tested whether the inhibitory effect of the secreted As_SRP-1 on sperm activation could be rescued by the addition of specific antiserum of As_SRP-1. We activated the first batch

of sperm with the Con A eluate, then collected the supernatant. When the supernatant was added to the second batch of sperm, no sperm activation was observed (Fig. 3F, *Upper and Lower Left*), probably because As_TRY-5 in the Con A eluate was inhibited by As_SRP-1 secreted from the first batch of sperm. As expected, when As_SRP-1 antiserum was added to neutralize As_SRP-1 secreted by the first batch of sperm, the resulting supernatant was able to activate the second batch of sperm (Fig. 3F, *Center, Upper and Lower*). As a control, the preimmune serum had no such effect (Fig. 3F, *Upper and Lower Right*). Thus, sperm-secreted As_SRP-1 during sperm activation blocks the activation of other sperm by inhibiting the glandular vas deferens-derived serine protease As_TRY-5.

Discussion

After the meiosis, spermatids are transcriptionally and translationally silent and, thus, sperm activation, motility acquisition, sperm competition, and fertilization are performed without new gene expression (22). Our ex vivo data provide evidence that motile spermatozoa are biochemically active in contributing a protein (As_SRP-1) to the seminal fluid and that this protein might coordinate both spermatozoon motility and sperm competition in vivo. On the one hand, for activated sperm in the uterus, As_SRP-1 is necessary for MSP cytoskeleton assembly and sperm motility acquisition (Fig. 2), thus improving the competitiveness of spermatozoa. Although the mechanism by which As_SRP-1 modulates cytoskeleton dynamics remains unclear, our data suggest that As_SRP-1 might act through protein tyrosine phosphorylation (Fig. S5), which has been known as a molecular switch in the regulation of MSP-based cell motility (14). On the other hand, for nonactivated sperm in the uterus from other males, As_SRP-1 irreversibly terminates the activity of a vas deferens-derived serine protease, As_TRY-5 (Fig. 3), thus inhibiting the activation of other spermatids. The spatially and temporally controlled encounter of nonmotile spermatids with the activating protease As_TRY-5 and the regulated release of the dual-function serine protease inhibitor As_SRP-1 during sperm activation constitute an elaborate opposing but complementary mechanism to coordinate sperm maturation and likely sperm competition in vivo.

Sperm competition in polyandrous species has been widely recognized as one of most potent driving forces in the evolution (23). Studies on sperm-competition mechanisms have focused on the physical traits of sperm [reviewed in (23, 24)], such as number of sperm inseminated, cell size, swimming velocity, and on seminal fluid produced by several accessory glands in the male body [reviewed in (25, 26)]. Several seminal fluid proteins in insects were involved in sperm competition by sperm displacement, sperm incapacitation or sperm ejection by females (27–30). Real-time live cell-imaging studies of sperm competition in transgenic flies with different fluorescent protein-labeled sperm support the sperm displacement mechanism, but not sperm incapacitation mechanism (31). Interestingly, the presence of sperm in addition to seminal fluid from second males would significantly enhance the magnitude of sperm displacement compared with that caused only by seminal fluid from spermless males (32). Seminal fluid is produced principally by accessory glands in the male body (25, 26), but our data show that sperm can secrete a component of this fluid. An interesting avenue for future investigation may be to determine whether other animal species besides nematodes also use sperm-secreted components mechanism to modulate sperm competition.

Some of our data from *Ascaris* not only agree with those from *C. elegans* but also further the understanding of *C. elegans* sperm activation. For example, it has been known for over three decades that nematode sperm can be activated in vitro by proteases, but the physiological relevance of this in vitro phenomenon is uncertain. A recent genetic study suggests that *C. elegans* sperm

activation involves protease activities regulated by SWM-1 (7), which contains two trypsin inhibitor-like domains. The purification and identification of As_TRY-5 and As_SRP-1 as important regulators of sperm activation in *Ascaris* provide unequivocal evidence that proteases and protease inhibitors indeed regulate sexual reproduction. Interestingly, proteolytic activity in seminal fluid is required for the activation of insect sperm in the female reproductive tract (33, 34). Nine and 8 out of 83 predicted seminal fluid proteins in *Drosophila* are proteases and inhibitors, respectively, implying their important roles in male fertility (35). Functional processing of fertilin, a metalloprotease associated with mammalian sperm, by convertase during sperm transit in the epididymis of mice is essential for sperm activation and male fertility (36). Lack of the serine protease inhibitor nexin-1 in the seminal fluid of mutant mice impaired male fertility (37). Therefore, proteolysis-mediated sperm activation might have broad phylogenetic conservation and the proteolytic activity outside of sperm is essential for male reproductive success. As_SRP-1 has dual functions in the modulation of nematode sperm maturation, providing insights for the fine-tuning of sperm function and male fertility before and postinsemination. In taxa outside Nematoda that produce flagellated sperm, regulated exocytosis is also required to create fertilization-competent sperm and to achieve reproductive success (38). Thus, sperm from different taxa might use this active secretion mechanism to alter their immediate environment to enhance their own competitiveness.

Materials and Methods

Ascaris sperm were obtained by dissecting males to recover seminal vesicles, which were processed to release seminal fluid into HKB buffer [50 mM Hepes, 70 mM KCl, 10 mM NaHCO₃ (pH 7.1)]. Spermatozoa were obtained by activating spermatids with the addition of SAS (vas deferens extract) (12). We observed sperm after various treatments as described in *SI Materials and Methods* using a DIC microscope (Axio Imager M2, Carl Zeiss) and MSP fibers assembled in vitro (13) using a phase-contrast microscope (Axio Observer, Carl Zeiss). All images were processed using MetaMorph (Universal Imaging). For additional details on fiber assembly in vitro, native protein purification, MS analysis, recombination protein expression, gene cloning, antibody preparation, immunodepletion, immunofluorescence, and cryo-immuno-EM assays, see *SI Materials and Methods*.

Note Added in Proof

While this paper was under review at PNAS, Smith and Stanfield (39) reported that *C. elegans* TRY-5, found in the male seminal fluid, is required for male-mediated sperm activation, consistent with our data described here.

ACKNOWLEDGMENTS. We thank Dr. Li-Lin Du for valuable ideas regarding the BLAST search; Dr. Jian Ren for advice with Fig. 3F; Ning Yang for help with Fig. 1E; Gail Ekman for advice in degenerate PCR; Dr. Taotao Wei for providing various protease inhibitors; Drs. Hengbin Wang and Yixian Zheng for critically reading the manuscript; and Wei Zhuang for excellent technical assistance. This research was supported by Grants 2012CB94502, 2010CB912303, and 30971648 (to L.M.), 2007AA02Z1A7 and 2010CB835203 (to M.-Q.D.), 2010CB912701 (to S.-M.H.), 30871226 and 31071180 (to Y.Z.) from the government of the People's Republic of China. L.M. is supported by the Chinese Academy of Sciences 100-Talents Program. S.W.L. was supported by National Institutes of Health Grant GM082932.

- Fraser LR (2010) The "switching on" of mammalian spermatozoa: molecular events involved in promotion and regulation of capacitation. *Mol Reprod Dev* 77:197–208.
- L'Hernault SW (2006) Spermatogenesis. *The C. elegans Research Community, WormBook*. Available at <http://www.wormbook.org>.
- Ward S, Carrel JS (1979) Fertilization and sperm competition in the nematode *Caenorhabditis elegans*. *Dev Biol* 73:304–321.
- LaMunyon CW, Ward S (1998) Larger sperm outcompete smaller sperm in the nematode *Caenorhabditis elegans*. *Proc Biol Sci* 265:1997–2002.
- Singson A, Hill KL, L'Hernault SW (1999) Sperm competition in the absence of fertilization in *Caenorhabditis elegans*. *Genetics* 152:201–208.
- Ward S, Hogan E, Nelson GA (1983) The initiation of spermiogenesis in the nematode *Caenorhabditis elegans*. *Dev Biol* 98:70–79.
- Stanfield GM, Villeneuve AM (2006) Regulation of sperm activation by SWM-1 is required for reproductive success of *C. elegans* males. *Curr Biol* 16:252–263.
- L'Hernault SW, Roberts TM (1995) Cell biology of nematode sperm. *Methods Cell Biol* 48:273–301.
- Zhu GD, et al. (2009) SPE-39 family proteins interact with the HOPS complex and function in lysosomal delivery. *Mol Biol Cell* 20:1223–1240.
- Ward S, Argon Y, Nelson GA (1981) Sperm morphogenesis in wild-type and fertilization-defective mutants of *Caenorhabditis elegans*. *J Cell Biol* 91:26–44.
- Washington NL, Ward S (2006) FER-1 regulates Ca²⁺-mediated membrane fusion during *C. elegans* spermatogenesis. *J Cell Sci* 119:2552–2562.
- Abbas M, Cain GD (1979) In vitro activation and behavior of the amoeboid sperm of *Ascaris suum* (Nematoda). *Cell Tissue Res* 200:273–284.
- Italiano JE, Jr., Roberts TM, Stewart M, Fontana CA (1996) Reconstitution in vitro of the motile apparatus from the amoeboid sperm of *Ascaris* shows that filament assembly and bundling move membranes. *Cell* 84:105–114.
- Miao L, Vanderlinde O, Stewart M, Roberts TM (2003) Retraction in amoeboid cell motility powered by cytoskeletal dynamics. *Science* 302:1405–1407.
- Yi K, Buttery SM, Stewart M, Roberts TM (2007) A Ser/Thr kinase required for membrane-associated assembly of the major sperm protein motility apparatus in the amoeboid sperm of *Ascaris*. *Mol Biol Cell* 18:1816–1825.
- Okamoto H, Thomson JN (1985) Monoclonal antibodies which distinguish certain classes of neuronal and supporting cells in the nervous tissue of the nematode *Caenorhabditis elegans*. *J Neurosci* 5:643–653.
- Arduengo PM, Appleberry OK, Chuang P, L'Hernault SW (1998) The presenilin protein family member SPE-4 localizes to an ER/Golgi derived organelle and is required for proper cytoplasmic partitioning during *Caenorhabditis elegans* spermatogenesis. *J Cell Sci* 111:3645–3654.
- Al Rawi S, et al. (2011) Postfertilization autophagy of sperm organelles prevents paternal mitochondrial DNA transmission. *Science* 334:1144–1147.
- Chi H, et al. (2010) pNovo: De novo peptide sequencing and identification using HCD spectra. *J Proteome Res* 9:2713–2724.
- Parkinson J, Whitton C, Schmid R, Thomson M, Blaxter M (2004) NEMBASE: A resource for parasitic nematode ESTs. *Nucleic Acids Res* 32(Database issue):D427–D430.
- Ye S, Goldsmith EJ (2001) Serpins and other covalent protease inhibitors. *Curr Opin Struct Biol* 11:740–745.
- Grunewald S, Paasch U, Glander HJ, Anderegg U (2005) Mature human spermatozoa do not transcribe novel RNA. *Andrologia* 37:69–71.
- Birkhead TR, Pizzari T (2002) Postcopulatory sexual selection. *Nat Rev Genet* 3:262–273.
- Wigby S, Chapman T (2004) Sperm competition. *Curr Biol* 14:R100–R102.
- Chapman T, Davies SJ (2004) Functions and analysis of the seminal fluid proteins of male *Drosophila melanogaster* fruit flies. *Peptides* 25:1477–1490.
- Avila FW, Sirot LK, LaFlamme BA, Rubinstein CD, Wolfner MF (2011) Insect seminal fluid proteins: Identification and function. *Annu Rev Entomol* 56:21–40.
- Price CS, Dyer KA, Coyne JA (1999) Sperm competition between *Drosophila* males involves both displacement and incapacitation. *Nature* 400:449–452.
- Chapman T, Neubaum DM, Wolfner MF, Partridge L (2000) The role of male accessory gland protein Acp36DE in sperm competition in *Drosophila melanogaster*. *Proc Biol Sci* 267:1097–1105.
- Snook RR, Hosken DJ (2004) Sperm death and dumping in *Drosophila*. *Nature* 428:939–941.
- den Boer SP, Baer B, Boomsma JJ (2010) Seminal fluid mediates ejaculate competition in social insects. *Science* 327:1506–1509.
- Manier MK, et al. (2010) Resolving mechanisms of competitive fertilization success in *Drosophila melanogaster*. *Science* 328:354–357.
- Gilchrist AS, Partridge L (2000) Why it is difficult to model sperm displacement in *Drosophila melanogaster*: The relation between sperm transfer and copulation duration. *Evolution* 54:534–542.
- Osanai M, Kasuga H (1990) Role of endopeptidase in motility induction in *Apyrene* silkworm spermatozoa - micropore formation in the flagellar membrane. *Experientia* 46:261–264.
- Friedländer M, Jeshtadi A, Reynolds SE (2001) The structural mechanism of trypsin-induced intrinsic motility in *Manduca sexta* spermatozoa in vitro. *J Insect Physiol* 47:245–255.
- Swanson WJ, Clark AG, Waldrip-Dail HM, Wolfner MF, Aquadro CF (2001) Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc Natl Acad Sci USA* 98:7375–7379.
- Cho C, et al. (1998) Fertilization defects in sperm from mice lacking fertilin beta. *Science* 281:1857–1859.
- Murer V, et al. (2001) Male fertility defects in mice lacking the serine protease inhibitor protease nexin-1. *Proc Natl Acad Sci USA* 98:3029–3033.
- Vacquier VD (1998) Evolution of gamete recognition proteins. *Science* 281:1995–1998.
- Smith JR, Stanfield GM (2011) TRY-5 is a sperm-activating protease in *Caenorhabditis elegans* seminal fluid. *PLoS Genet* 7(11):e1002375.

Supporting Information

Zhao et al. 10.1073/pnas.1109912109

SI Materials and Methods

Isolation and Activation of *Ascaris suum* Sperm. *Ascaris* males were collected from the intestines of infected hogs at Zhongrui Pork Processors and Jinlue Meat Products. They were placed in worm buffer [PBS containing 100 mM NaHCO₃ (pH 7.0) 37 °C] and transported to the laboratory and stored at 37 °C overnight for recovery. Spermatids were obtained by dissecting males, removing the seminal vesicle, and extruding the seminal fluid into HKB buffer [50 mM Hepes, 70 mM KCl, 10 mM NaHCO₃ (pH 7.1)]. Spermatids were activated to extend pseudopods in HKB buffer with the addition of a glandular vas deferens extract (SAS) (1). Live cells were pipetted into chambers formed by mounting a 22 × 22-mm glass coverslip onto a glass slide with two parallel strips of two-sided tape and examined with an Axio Imager M2 microscope (Carl Zeiss) equipped with a 40 × differential interference contrast (DIC) objective lens with appropriate filters. Images were captured with a charge-coupled device (CCD) (Andor) and processed with MetaMorph software (Universal Imaging). Sperm not used immediately were pelleted at 10,000 g for 30 s, and the cell pellet was stored at –80 °C after the removal of supernatant.

Preparation of Sperm Extracts (S100) and Fiber Assembly. Frozen spermatozoa were thawed on ice for 1 h and subject to centrifugation at 16,000 × g for 10 min. The supernatant was subject to centrifugation at 100,000 g for 1 h at 4 °C. The supernatant (S100) was used for the fiber assembly assay with the addition of 1 mM ATP in KPM buffer (10 mM potassium phosphate, 0.5 mM MgCl₂, pH 6.8) (2).

Purification of the 46-kDa Protein Recognized by the 1CB4 Antibody from Sperm Extracts. The sperm extract was diluted 1:4 with KPM buffer and fractionated by ammonium sulfate precipitation (3). Each pellet was resuspended in KPM buffer and dialyzed against KPM buffer for 24 h. The different fractions were analyzed using SDS/PAGE and Western blots with the 1CB4 antibody. The protein recognized by 1CB4 (~46 kDa) was enriched in the 60–80% fraction of ammonium sulfate (AmSO₄), and this fraction was dialyzed against pH 7.8 KPM buffer and fractionated in 1 mL SP-Sepharose High Performance cation exchange columns (GE Healthcare) using an AKTA FPLC system (GE Healthcare). The fraction containing the protein recognized by 1CB4 was loaded onto a Heparin HP affinity chromatography column (GE Healthcare) and eluted with a 0–500 mM NaCl gradient in 10 mM KPM (pH 7.8). The eluate contained a single 46 kDa band, as resolved by SDS/PAGE.

Immunodepletion Assay. The As_SRP-1 polyclonal antibody (10 µg/mL) was bound to Protein A Sepharose 4B (GE Healthcare), and these were washed five times with HKB buffer. Washed beads were then incubated with As_SRP-1 (10 µg/mL) 4 °C for 1 h to immobilize As_SRP-1 onto the beads. After SAS (0.5 µg/mL) was added to the preloaded beads for 1 h with rotation, the supernatant was collected to examine its effect on sperm activation. For mock depletion, either As_SRP-1 antibody was replaced with IgG or As_SRP-1 was omitted for binding with beads before incubation with SAS.

Purification of Sperm Activator As_TRY-5 from SAS. The SAS was separated by AmSO₄ precipitation. The 40–60% cut fraction was loaded onto a HiTrap phenyl HP column (GE Healthcare; 5 mL) equilibrated with HKB buffer [50 mM Hepes, 70 mM KCl, 10 mM NaHCO₃ (pH 7.1)]; all of the columns mentioned below had been equilibrated with HKB buffer] containing 1 M AmSO₄, then this

fraction was sequentially eluted with HKB buffer containing 0.5 M AmSO₄ (50% eluate) or 0 M AmSO₄ (100% eluate), respectively. The 100% eluate contained sperm-activating activity, which was concentrated with an Amicon Ultra-4 centrifugal filter devices (10 K cutoff) and loaded onto a Superdex 200 column (GE Healthcare; 120 mL). The fractions containing the sperm-activating activity were pooled and loaded onto a HiTrap Q HP column (GE Healthcare; 5 mL), which was eluted with a 60 mL linear gradient of HKB buffer containing 0–0.8 M NaCl. The peak fractions were again pooled, concentrated and loaded onto a Mono Q 5/50 GL column (GE Healthcare; 1 mL), and the column was eluted with a 60 mL linear gradient of HKB buffer containing 0.2–0.8 M NaCl. The sperm-activating peak fractions eluted with 0.6–0.8 M NaCl. Therefore, these fractions were pooled, concentrated with an Amicon Ultra-4 centrifugal filter devices (10 K cutoff) and loaded onto a HiTrap Con A 4B columns (GE Healthcare; 1 mL), following by the elution with HKB buffer containing 0.5 M methyl- α -d-glucopyranoside. Finally, the eluate was concentrated and incubated with purified As_SRP-1 at 4 °C for 12 h. The incubated mixture showed a new shifted ~90-kDa band by SDS/PAGE, and this band was recognized by our anti-As_SRP-1 antibody in Western blots. The 90-kDa band was excised from the gel for MS analysis and de novo sequencing.

MS Identification of As_SRP-1 As the 1CB4-Recognizing Protein. MS data collection for de novo sequencing. A highly purified *Ascaris* protein sample recognized by the 1CB4 antibody consisted of three closely migrating proteins around 46 kDa. This sample was digested with trypsin, loaded onto a C18 reverse-phase column with a pulled tip (100 µm ID; 8 cm in length), and packed with 3-µm, 125-Å Aqua C18 resin (Phenomenex). LC-MS/MS analysis was performed over a 70-min run on an LTQ-Orbitrap-ETD mass spectrometer (Thermo-Fisher Scientific) connected to an Agilent 1200 quaternary HPLC Pump. The HPLC gradient was made of buffer A (acetonitrile/H₂O/formic acid, 5/95/0.1) and buffer B (acetonitrile/H₂O/formic acid; 80/20/0.1) at a constant flow rate of 0.1 mL/min. The flow was split between a waste line and the C18 column. The resulting flow rate at the tip of the column was about 200 nL/min, and the gradient was as follows: 0 min, A = 100.0%; 2 min, A = 90.0% B = 10.0%; 32 min, A = 50.0% B = 50.0%; 42 min, B = 100.0%; 47 min, B = 100.0%; 50 min, A = 100.0%; 70 min, A = 100.0%. Three sets of higher-energy collisional dissociation (HCD) data were generated and in each, full scans ($R = 30,000$) and data-dependent HCD (normalized collisional energy = 40%) MS2 scans ($R = 7,500$) were all analyzed in the Orbitrap. For dataset 1 (HCD only), the top 5 most intense +2 precursor ions detected in each full scan were isolated for MS2 using HCD. Dynamic exclusion was turned on with two repeat counts, 5 s repeat duration, exclusion list 500, and exclusion duration 15 s. For dataset 2 (normal-mass range HCD), only the most intense +2 or +3 precursor ion in each full scan was isolated to generate four MS2 spectra: HCD (mass range 100–2,000 m/z), CID in LTQ, high-resolution electron transfer dissociation (ETD) spectra mass analyzed in the Orbitrap, and ETD in LTQ. The dynamic exclusion was set at repeat count 1, exclusion list 500, and exclusion duration 10 s. For dataset 3 (low-mass range HCD), the two most intense +2 or +3 precursor ions in each full scan were isolated and for each four MS2 spectra were generated: HCD (mass range 50–2,000 m/z), CID in LTQ, high-resolution ETD (detected in the Orbitrap), and ETD in LTQ. The dynamic exclusion setting was the same as in dataset 2 except that exclusion duration was 15 s. All tandem mass spectra were extracted from RAW files using Xcalibur 2.0.7. Different types of MS2 spectra were separated by an in-house software

MS2Extractor. Only the HCD and high-resolution ETD data were used in de novo peptide sequencing analysis.

Identification of *As_SRP-1* by de novo sequencing. Peptide sequences were deduced from normal- and low-mass range HCD spectra, as well as high-resolution ETD spectra using a de novo sequencing program called pNovo (4). The de novo sequencing results were pooled together, from which redundant peptide sequences and those with less than six amino acids or a *C*-score less than 0.5 were removed. The resulting 425 unique peptide sequences were BLAST searched against an *Ascaris* protein database translated from EST sequences (ASP_NEMBASE3_pro.fsa from www.nematodes.org). The PAM30 matrix, a word length of 2 aa and an E-value cutoff of 20,000 were used for BLASTP search. The BLAST results were tabulated and filtered, requiring >50% coverage (length of matched sequence/length of query peptide sequence) and >60% identity (number of identical amino acids within the matched sequence/length of matched sequence) for each retrieved match. For each protein found by BLAST search, the total number of unique peptide matches was calculated. With 60 unique peptide hits and 173 HCD spectra, *As_SRP-1* ascended to the top. This was followed by a copurified contaminant CTS-1 (citrate synthase) with 29 unique peptide hits obtained from 85 spectra.

Identification of *As_TRY-5* by de Novo Sequencing. The 90-kDa covalent complex between *As_SRP-1* and its target protease was excised from the SDS/PAGE gel, in-gel digested with trypsin, Asp-N, and Lys-N, separately. LC-MS/MS analyses of the resulting peptides were the same as described above except for the following changes in MS data collection. Each MS ($R = 30,000$) was followed by 10 MS2 on top 2 ions with HCD (100–2,000 m/z), low-mass range HCD (50–2,000 m/z), CID in LTQ, high-resolution ETD in Orbitrap, and ETD in LTQ for each ion, $R = 7500$ for HCD and high-resolution ETD. We used pNovo to identify peptides from the HCD and high-resolution ETD spectra of the 90 kDa band. Sequences belonging to *As_SRP-1* and contaminant proteins HSP90, keratins, trypsin, Lys-N, or Asp-N (these proteins were found in the sample by database search) were removed. Overlapping peptides found by pNovo were assembled into longer sequences using an in-house script. BLAST search matched these contiguous sequences to trypsin-like protease protein 5 [*Brugiya malayi*] with an E value as low as $8e-08$. Its closest homolog in *C. elegans* is TRY-5. Thus, the protease target of *As_SRP-1* is named *As_TRY-5* in *Ascaris suum*.

Molecular Cloning of *As_srp-1* and *As_try-5*. Total RNAs from *Ascaris* testes and vas deferens were prepared using TRIzol reagent (Invitrogen) and the poly(A)⁺ RNA was purified using Oligotex (QIAGEN). The first-strand cDNA was produced using a First-Choice RLM-RACE kit (Ambion). For the 5' RACE PCR template, the 5' adaptor provided in the kit was ligated to the CIP/TAP-treated RNA using T4 RNA ligase and the reverse transcription was performed using random decamers as primers. For the 3' RACE PCR template, reverse transcription was performed using a 3' adaptor as primer.

To obtain the full-length *As_srp-1*, the 5' RACE primers (P1: CATCATATCGACCAATTATGCGTTG; P2: TGGATATTGCCAACTTCCATTGA) and 3' RACE primers (P3: GTGGCATTGTGCAAGACCAACAC; P4: ATTCGTCGTGCGTTCAGCATGAC) were designed according to the *Ascaris* EST database. PCR amplification was performed with *PfuUltra* II fusion HS DNA polymerase (Stratagene) and the PCR products were cloned into *pEASY-Blunt* vectors (TransGen) for sequencing.

For *As_try-5* cloning, first we performed 3' RACE PCR using degenerative primers (2F: GGATCCGAYGARTTYGAYGARTGGGA; and 7F: GGATCCGGNGTNTGYGAYGAYGARGA) designed according to peptides deduced from de novo sequencing and obtained a 406 bp C-terminal sequence. The underlined sequences in primers show the restriction site for in-

creasing the melting temperature and specificity when the PCR was performed. Subsequently, 5' RACE PCR was performed using primers R1 (CAATCACCCAAACACAACAACAAG), R2 (TTTGTCAGTTTTTTGTTTCCTCCAC), R3 (GTTATGAGTGTTCCTCCGCAGTCG), and R4 (TGTTCTCCGCAGTCGGC-TTCTAA), and the PCR products were cloned into *pEASY-Blunt* vector (TransGen) for sequencing.

Recombinant *As_SRP-1* Expression in *E. coli*. The ORF of *As_srp-1* was amplified from cDNA and cloned into pET-28a vector (Novagen). The resulting plasmid was transformed into *E. coli* strain BL21 (DE3) for expressing recombinant *As_SRP-1*. *E. coli* lysates were analyzed by SDS/PAGE and Western blotting with either a polyclonal antibody against *As_SRP-1* (anti-P46) or the monoclonal 1CB4 to test the de novo sequencing result.

Production of a Polyclonal Antibody Against *As_SRP-1*. To generate a polyclonal antiserum against *As_SRP-1*, we injected 0.5 mg *As_SRP-1* purified from sperm extract (as described above) into a New Zealand white rabbit, and the rabbit was immunized five times. One week after the last injection, the rabbit was killed and serum was collected. The antibody was then purified from the antiserum using protein A-Sepharose HP columns (GE Healthcare).

Immunofluorescence Assay. Spermatids, spermatozoa, or fibers grown in vitro were fixed with 1.25% glutaraldehyde in buffer (HKB buffer for sperm and KPM buffer for fibers) at room temperature for 10 min in the culture chamber. Fixed samples were then blocked with 0.4% NaBH₄ three times, 10 min each, and permeabilized with 0.5% Triton X-100 in PBS (PBS) when necessary. After blocking with 2% BSA (BSA) at room temperature for 6 h, samples were incubated with 1CB4 or anti-*As_SRP-1* or anti-pY (Millipore) antibodies (1:200 in PBS with 2% BSA) overnight at 4 °C. After washing three times with PBS, samples were stained with Alexa Fluor 488-conjugated goat anti-mouse or Rhodamine-conjugated goat anti-rabbit secondary antibodies (Molecular Probes; 1:400 in PBS with 2% BSA) at room temperature for 1 h. Images were obtained with a confocal laser scanning microscope (Olympus).

MO Fusion Assay. Spermatids were isolated and incubated in HKB buffer no SAS, SAS, PMSF-treated SAS, or PHE-treated SAS for 10 min at 37 °C as described above. The treated cells were stained with FM1-43FX (Molecular Probes) at 5 µg/mL for 1 min to visualize the PM and MO upon fusion (5). Images were captured using a confocal laser scanning microscope (Leica).

Cryo-Immuno-EM. Ultrathin cryomicrotomy and immunogold labeling were performed according to the protocol of Tokuyasu (6, 7). Sperm were fixed for 30 min with 4% paraformaldehyde in HKB buffer containing 0.1% glutaraldehyde, and infused with 2.3 M sucrose overnight at room temperature for cryoprotection. Fixed sperm were then mounted and snap-frozen in liquid nitrogen on cryospecimen pins for cross-sectioning. Ultrathin cryosections (70–90 nm thick) were cut and transferred to 200-mesh grids with a Leica ultramicrotome. Grids containing these cryosections were then blocked with 5% goat serum for 30 min, followed by incubation with 1CB4 antibody (1:20) for 1 h. After rinsing four times for 5 min each, the grids were transferred to droplets of 15 nm immunogold-conjugated goat anti-mouse secondary antibody (Sigma; 1:20). The diluent for immuno-gold incubations and intermediate rinses was PBS containing 0.15% glycine and 1% BSA. Finally, antibody-labeled grids were fixed with 2.5% glutaraldehyde in PBS for 5 min, stained with 4% uranyl acetate after rinsing with water, transferred to 1% methylcellulose droplets and picked up using a platinum wire loop. Excess embedding solution was removed using filter paper and the grids were allowed to air dry. Images were captured on an FEI 20 transmission electron microscope (FEI).

1. Abbas M, Cain GD (1979) In vitro activation and behavior of the amoeboid sperm of *Ascaris suum* (Nematoda). *Cell Tissue Res* 200:273–284.
2. Italiano JE, Jr., Roberts TM, Stewart M, Fontana CA (1996) Reconstitution in vitro of the motile apparatus from the amoeboid sperm of *Ascaris* shows that filament assembly and bundling move membranes. *Cell* 84:105–114.
3. Buttery SM, Ekman GC, Seavy M, Stewart M, Roberts TM (2003) Dissection of the *Ascaris* sperm motility machinery identifies key proteins involved in major sperm protein-based amoeboid locomotion. *Mol Biol Cell* 14:5082–5088.
4. Chi H, et al. (2010) pNovo: de novo peptide sequencing and identification using HCD spectra. *J Proteome Res* 9:2713–2724.
5. Washington NL, Ward S (2006) FER-1 regulates Ca²⁺-mediated membrane fusion during *C. elegans* spermatogenesis. *J Cell Sci* 119:2552–2562.
6. Tokuyasu KT (1986) Application of cryoultramicrotomy to immunocytochemistry. *J Microsc* 143:139–149.
7. Tokuyasu KT (1989) Use of poly(vinylpyrrolidone) and poly(vinyl alcohol) for cryoultramicrotomy. *Histochem J* 21:163–171.

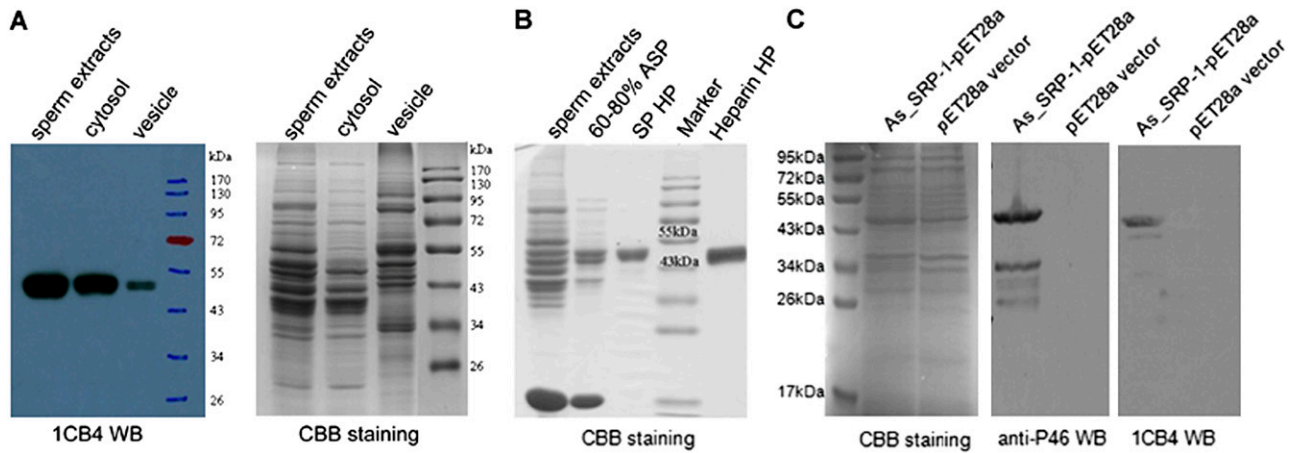


Fig. S1. Identification and confirmation of the 1CB4-recognizing protein as Serpin (As_SRP-1). (A) Single band recognized by 1CB4 was detected in sperm extracts, more specifically, in the cytosol and vesicle with SDS/PAGE followed by Western blot. (B) Scheme used to purify the *Ascaris* sperm protein recognized by 1CB4. ASP, ammonium sulfate precipitation; Heparin HP, heparin HP affinity chromatography; SH HP, cation exchange columns. (C) As_SRP-1 expressed in *E. coli* was detected by SDS/PAGE and Western blots with anti-As_SRP-1 polyclonal antibody (anti-P46) and monoclonal antibody 1CB4.

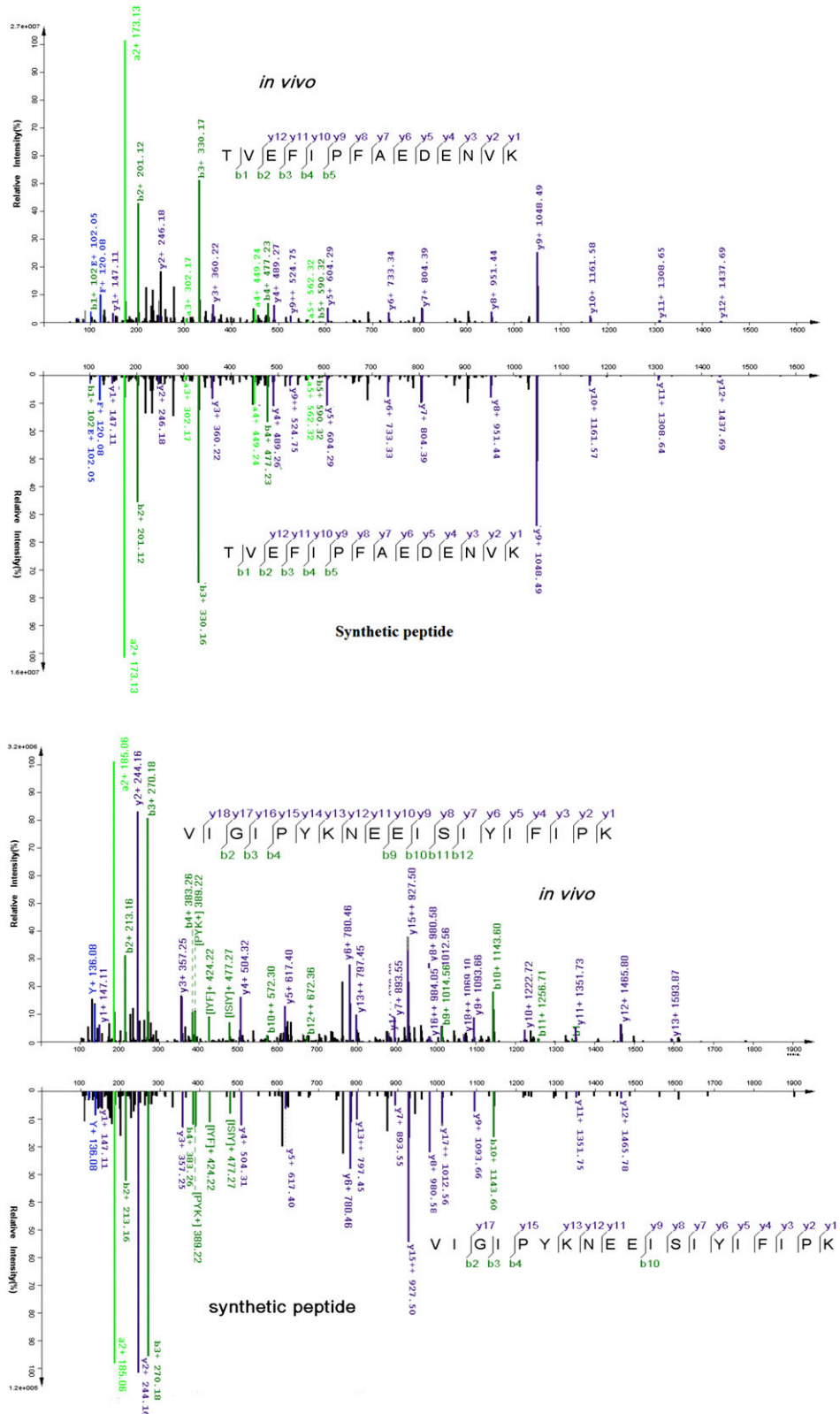


Fig. S2. Verification by synthetic peptides of the pNovo identification of two sequences originated from As_SRP-1. The peptide sequences were interpreted directly from the higher-energy collisional dissociation (HCD) spectra using the pNovo program. Peptides were synthesized for two of them. The original HCD spectra are labeled "in vivo." Displayed below each of them is an HCD spectrum of the synthetic peptide, in perfect agreement with the original spectrum.

As_SRP-1:

Peptide	Spectrum	Modification	pFind-e-value	Mascot-score	pNovo-result	pNovo-score
AFSMTAANFSGICAR	Ascaris25-28_HCD-2.1795.1795.2.dta	4,M(Oxidation);13,C(Carbamidomethyl);	1.87E-28	57.40	MSAFTAAANFSGICAR	0.507
FVEIMQEKFEGEIYTMDSKDPSETAK	Ascaris-25-28-hcd-cid-std2-top1only_090427.2442.2442.3.dta	/	1.77E-47	77.05	VFEIMQEKFEGEIYTMDSKDPSETAK	0.625
AIVNGCDDNTIINYSDIMHISQQPTTYTLNLANR	Ascaris-25-28-hcd-cid-std2-top1only_090427.3187.3187.3.dta	4,N(Deamidation);6,C(Carbamidomethyl);13,N(Deamidation);	1.98E-72	75.46	AIVNGCDDNTIINYSDIMHISQQPTTYTINIAR	0.562
KFVEIMQEK	Ascaris25-28_HCD-2.1572.1572.2.dta	/	6.66E-23	51.48	KFVEIMQEK	0.783
QFIANHPFLFALIKNETLFIHFGLK	Ascaris-25-28-hcd-cid-std2-top1only_090427.3914.3914.4.dta	15,N(Deamidation);	8.61E-50	47.63	/	/
FEGEIYTMDSKDPSETAKK	Ascaris-25-28-hcd-cid-std2-top1only_090427.1562.1562.3.dta	/	9.59E-24	32.58	FEGEIYTMDSKQVVEEK	0.619
EKFFPEENQR	Ascaris25-28_HCD-2.1524.1524.2.dta	/	2.29E-16	38.10	EKFFPEENQR	0.650
RAFSMTAANFSGICAR	Ascaris25-28_HCD-2.1860.1860.2.dta	14,C(Carbamidomethyl);	1.47E-17	58.25	/	/
IMVDMMR	mito_sun_90k_20100902.2433.2433.2.dta	/	4.49E-16	31.97	IMVDMMR	0.782
FGLHEIHTLNGSK	Ascaris25-28_HCD-2.1866.1866.2.dta	11,N(Deamidation);	8.51E-32	49.47	FGLHEIHTDGSK	0.592
TVEFLPFAEDENVK	mito_sun_90k_20100902.3024.3024.2.dta	/	4.69E-21	33.26	TVEFLPQVMDENVK	0.672
FEGEIYTMDSKDPSETAK	Ascaris-25-28-hcd-cid-std2-top1only_090427.1889.1889.3.dta	/	5.25E-38	74.29	FRCIYTMDSKDPSETIS	0.351
INQWLR	Ascaris25-28_HCD-2.1650.1650.2.dta	/	2.19E-11	43.95	INQWLR	0.662
DKLDEGNPDLIS	miaco-sun-aspn-20101106.2583.2583.2.dta	7,Q(Deamidation);	/	47.83	/	/
NEEISLYIFLPK	mito_sun_90k_20100902.3393.3393.2.dta	/	1.64E-21	30.87	NRICSIMKFIPIK	0.622
FFPEENQR	Ascaris25-28_HCD-2.1571.1571.2.dta	/	5.45E-14	46.42	FFPEENQR	0.641
VIGLPYKNEEISLYIFLPK	Ascaris-25-28-hcd-cid-std2-top1only_090427.3047.3047.3.dta	/	1.57E-33	51.02	VIGIPYKNEEISLYIFIPK	0.585
AFSMTAANFSGICARPTHIR	Ascaris-25-28-hcd-cid-std2-top1only_090427.2072.2072.3.dta	13,C(Carbamidomethyl);	1.05E-15	40.45	AFSMTAANFSGRRNIFIMMR	0.531
ALIEVNEGTAQAAVAIAEMVFK	Ascaris-25-28-hcd-cid-std2-top1only_090427.3442.3442.3.dta	/	1.38E-40	80.62	IAIEVNEGTAQAAVAIAEMVFK	0.566
KINQWLR	Ascaris-25-28-hcd-cid-std2-top1only_090427.1537.1537.2.dta	3,N(Deamidation);	3.29E-20	28.36	KIDQWLR	0.738
FVEIMQEK	Ascaris-25-28-hcd-cid-std2-top1only_090427.1379.1379.2.dta	5,M(Oxidation);	5.60E-15	24.71	VFEIMQEK	0.325
NETLFIHFGLK	Ascaris-25-28-hcd-cid-std2-top1only_090427.2719.2719.2.dta	/	1.25E-29	30.77	/	/
QFIANHPFLFALIK	Ascaris25-28_HCD-2.2534.2534.2.dta	0,Q(Gln->pyro-Glu);	1.19E-29	54.17	MIIFAHNAGPFAIK	0.576
VIGLPYK	mito_sun_90k_20100902.2372.2372.2.dta	/	2.07E-07	16.26	VIGIPYK	0.656
DMMRTVEFLPFAE	miaco-sun-aspn-20101106.3686.3686.2.dta	2,M(Oxidation);3,M(Oxidation);	1.25E-22	27.94	/	/
TVEFLPFAEDENVKIVIGLPYK	Ascaris25-28_HCD-2.2274.2274.2.dta	12,N(Deamidation);	4.20E-09	/	/	/
EKFFPEENQRIMV	miaco-sun-aspn-20101106.2987.2987.2.dta	/	9.38E-18	/	/	/
SNRPIKR	Ascaris25-28_HCD-2.368.368.2.dta	/	2.67E-09	/	/	/

Fig. S3. Comparison of the de novo sequencing result and the retrospective database search result of As_SRP-1. After the identification and cloning of As_SRP-1, the protein sequence of As_SRP-1 was derived from its full-length cDNA and added to a database. The original MS data were searched against this database using Mascot and pFind, which reidentified As_SRP-1 as the top candidate. Most of the As_SRP-1 peptides identified by database search had been identified by pNovo. Highlighted in yellow are the identical regions between a pNovo-identified sequence and that by database search from the same spectrum.

```

-----|
ORIGINAL-SEQUENCE |
MASCOT----RESULT |
PFIND-----RESULT |
DENOVO----RESULT |
-----|

```

As_SRP-1:

```

MFHRI I A I G A I I C L F L Q T F C S L S A A S P C D N L N I Q L A D F A L R M L R E G G E K S S A I S P S S A I A L A M L Y V G A M N E T K E E I R R A I V N G C D D N T I I N Y S D I M H I S Q Q P T T Y T L N L A N R I Y T E K
-----|-----A I V N G C D D N T I I N Y S D I M H I S Q Q P T T Y T L N L A N R-----
-----|-----A I V N G C D D N T I I N Y S D I M H I S Q Q P T T Y T L N L A N R-----
-----|-----S P C D N I N I Q I A D F A I R-----A I V D G C D D N T I I N Y S D I M H I S Q Q P T T Y T I N I A N R-----
-----|-----D F A I R M I R E G G E K-----D I M H I S Q Q P T T Y T I N I A N R I Y T E K-----
SLV L Q K K F V E I M Q E K F E G E I Y T M D S D K P S E T A K K I N Q W L R N K T N N K I D K L I D E Q N P D L S D I L I I N A I Y F N G S W Q Y P F K T S S T I K E K F F P E E N Q R I M V D M M R T V E F L P F A E D E N V K V I G L P
-----|-----K F V E I M Q E K F E G E I Y T M D S D K P S E T A K K I N Q W L R-----D K L I D E Q N P D L S-----E K F F P E E N Q R I M V D M M R T V E F L P F A E D E N V K V I G L P
-----|-----K F V E I M Q E K F E G E I Y T M D S D K P S E T A K K I N Q W L R-----E K F F P E E N Q R I M V D M M R T V E F L P F A E D E N V K V I G L P
-----|-----V F E I M Q E K F E G E I Y T M D S D K P S E T A K-----E K F F P E E N Q R-----T V E F P Q T V M D E N V K-----
-----|-----K I D Q W I R-----T V E F I P F A E D E N V K-----
-----|-----K F V E I M Q E K F E G E I Y T M D S D K P S E T A K K I N Q W I R-----F F P E E N Q R-----D M M R T V E F I P F A E-----V I G I P
-----|-----F V E I M Q E K F E G E I Y K T M D S D K Q V E E K-----I M V D M M R-----P F A E D E N V K V I G I P
Y K N E E I S L Y I F L P K V K F G L H E I E H T L N G S K L L A L A K S C R K F G R V E V R I P K F T I K M N F P M K N A L Q K M S I R R A F S M T A A N F S G I C A R P T H I R N F V H K A L I E V N E G G T Q A A A V S A I E M V F K S N
Y K N E E I S L Y I F L P K-----F G L H E I E H T L N G S K-----R A F S M T A A N F S G I C A R P T H I R-----A L I E V N E G G T Q A A A V S A I E M V F K-----
Y K N E E I S L Y I F L P K-----F G L H E I E H T L N G S K-----R A F S M T A A N F S G I C A R P T H I R-----A L I E V N E G G T Q A A A V S A I E M V F K S N
-----|-----N R I C S I M K F P I K-----F G I H E I E H T I D G S K-----M S A F T A A N F S G I C A R-----I A I E V N E G S A Q A A A V S A I E M V F K-----
Y K N E E I S I Y I F I P K-----A F S M T A A N F S G R R N I F M M R-----A I I E V N E G G T Q A A A V S A I E M V F K-----
R P I K R Q F I A N H P F L F A L I K N E T I L F I G H F L K
-----|-----Q F I A N H P F L F A L I K N E T I L F I G H F L K
R P I K R Q F I A N H P F L F A L I K N E T I L F I G H F L K
-----|-----M I I F A H N A G P F A I I K-----
-----|-----N E T I L F I G H F I K-----
-----|-----Q F I A N H P F I F A I I K-----

```

Fig. S4. Sequence coverage of As_SRP-1 by de novo sequencing and database search. Peptide sequences identified by pNovo (green), Mascot (blue), or pFind (red) are mapped to the As_SRP-1 protein sequence. Yellow highlighting indicates the pNovo-identified sequence tags that match perfectly to the As_SRP-1 protein sequence.

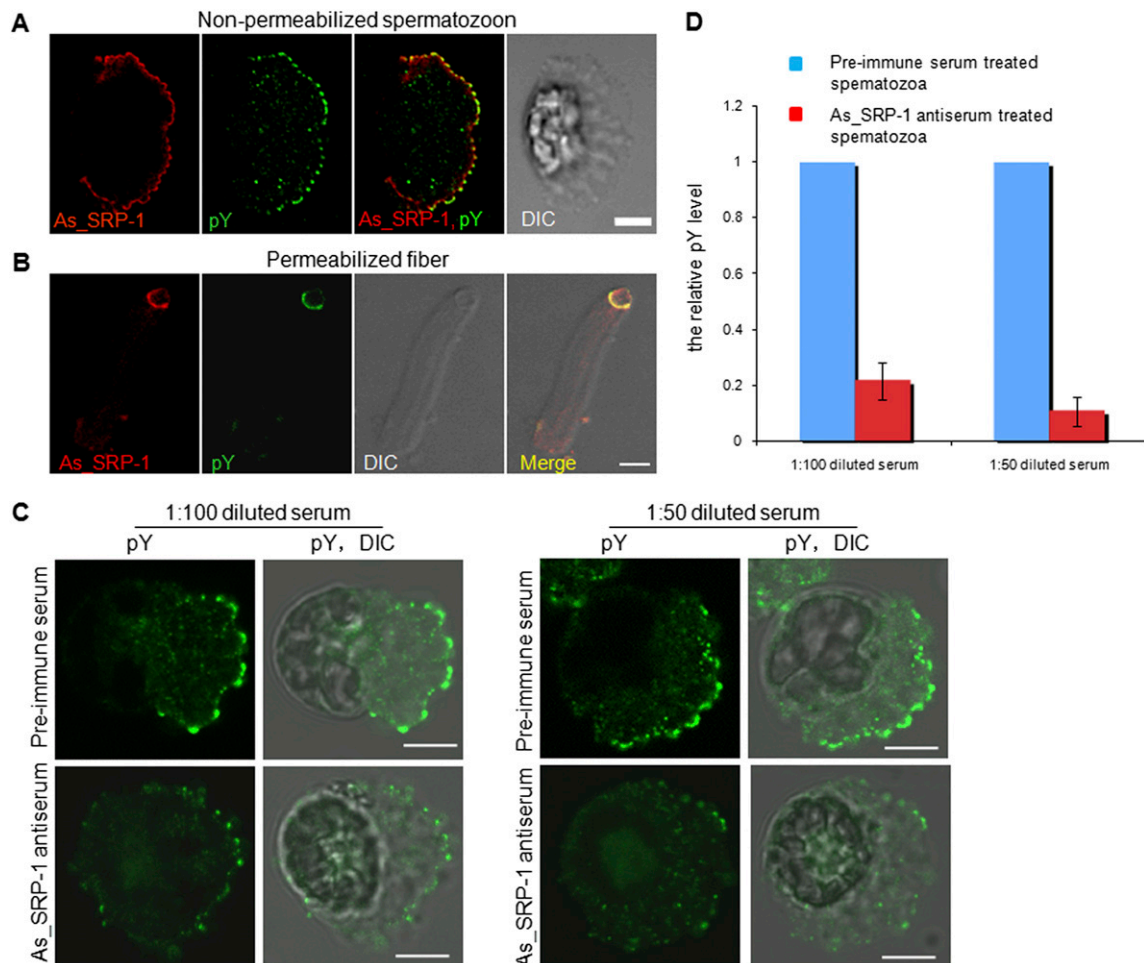


Fig. 55. *As_SRP-1* probably regulates *Ascaris* sperm motility through protein tyrosine phosphorylation (pY). (A) Colocalization of *As_SRP-1* (red) and pY (green) in nonpermeabilized spermatozoa. (Scale bars, 5 μm .) (B) Colocalization of *As_SRP-1* (red) and pY (green) in permeabilized fibers assembled in vitro. (Scale bars, 5 μm .) (C) *As_SRP-1* antiserum treatments (1:100 and 1:50) of spermatozoa for 10 min resulted in a decrease of pY signal (green) (confocal imaging) before spermatozoa showing cytoskeletal disassembly and roundup (DIC imaging). Preimmune treatments, control. Spermatozoa were not permeabilized for imaging. (Scale bars, 10 μm .) The relative fluorescence intensity of pY signal at the leading edge of spermatozoa was quantified in (D). All data are means \pm SD ($n = 3$ experiments).

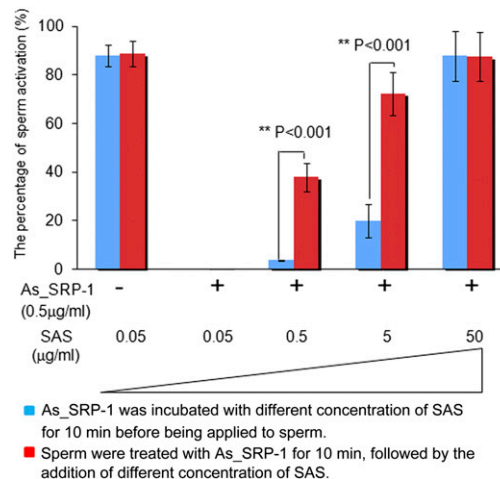


Fig. 56. The inhibitory target of *As_SRP-1* is present in SAS. When sperm were treated with the mixture of *As_SRP-1* and 0.5 or 5 $\mu\text{g/ml}$ of SAS (blue bars), the sperm activation rate was lower than that from sperm treated with *As_SRP-1* first, then followed by SAS addition (red bars). Data are means \pm SD ($n = 5$ experiments). $**P < 0.001$ (Student *t* test). No difference in sperm activation rate between blue and red bars at 0.05 and 50 $\mu\text{g/ml}$ of SAS was probably because either the SAS concentration was too low (0.05 $\mu\text{g/ml}$), thus, its sperm activating activity was blocked by 0.5 $\mu\text{g/ml}$ *As_SRP-1*, or the overwhelming concentration (50 $\mu\text{g/ml}$) of SAS abolished the inhibitory effect of *As_SRP-1*.

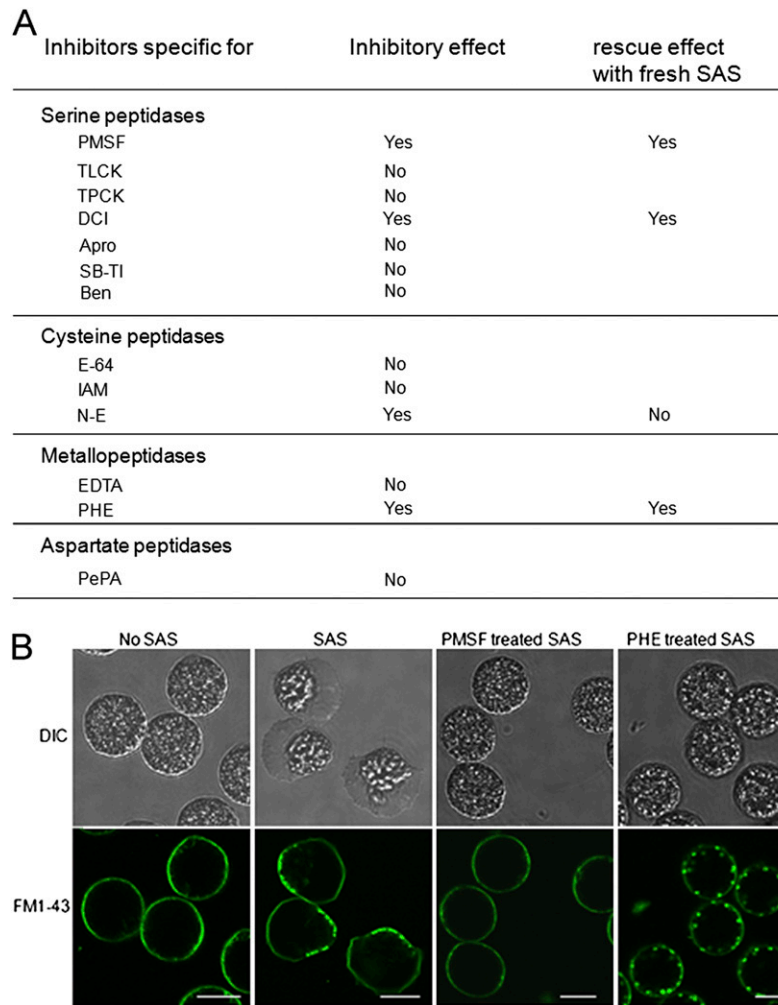


Fig. S7. Serine protease activity in SAS is necessary for sperm activation. (A) Summary of inhibitory effects of various protease inhibitors tested on SAS activity of sperm activation. The SAS (0.25mg/mL) was preincubated with various protease inhibitors [1 mM phenylmethanesulfonylfluoride (PMSF), 10 mM *N*- α -*D*-tosyl-L-lysine chloromethyl ketone (TLCK), 2.6 mM *N*-tosyl-L-phenylalanine chloromethyl ketone (TPCK), 1 mM 3,4-dichloroisocoumarin (DCI), 0.2 mg/mL aprotinin (Apro), 1 mg/mL Soybean trypsin inhibitor (SB-TI), 10 mM benzamidine (Ben) 0.15 mM E-64, 5 mM iodoacetamide (IAM) 25 mM *N*-ethylmaleimide (*N*-E), 2.5 mM EDTA, 5 mM *o*-phenanthroline (PHE), or 0.5 mM pepstatin A (PePA)], respectively, for 30 min on ice. The spermatids were incubated with treated SAS and checked by light microscopy. The results showed that serine protease, cysteine protease, and metalloprotease, but not aspartate protease activities in SAS might be important for sperm activation. For those inhibitors with inhibitory effect on sperm activation, the treated SAS was replaced by perfusion with fresh SAS to examine the rescue effects. The serine protease inhibitors (PMSF and DCI) and the metalloprotease inhibitor (PHE) inhibited sperm activation through their effect on SAS. The cysteine protease inhibitor *N*-E inhibited sperm directly and irreversibly (it could not be rescued by fresh SAS). (B) Examination of MO fusion. Spermatids were incubated with HKB buffer (no SAS), SAS, PMSF-treated SAS or PHE-treated SAS for 10 min. Lipophilic dye FM1-43FX that stains the outer leaflet of the PM was added to visualize the PM and MO upon fusion. *Upper*, DIC images; *Lower*, FM1-43FX staining. Spermatids (no SAS or PMSF-treated SAS) only displayed PM staining and spermatozoa (SAS) showed PM and many bright puncta (fused MOs) around the cell body periphery. Bright puncta were observed in the sperm incubated with PHE-treated SAS although pseudopod formation was inhibited. This assay further eliminated the metalloprotease activity in SAS as being responsible for sperm activation because PHE-treated SAS did induce spermatids to initiate MO fusion with the PM, although pseudopod formation was aborted. (Scale bar, 10 μ m.)

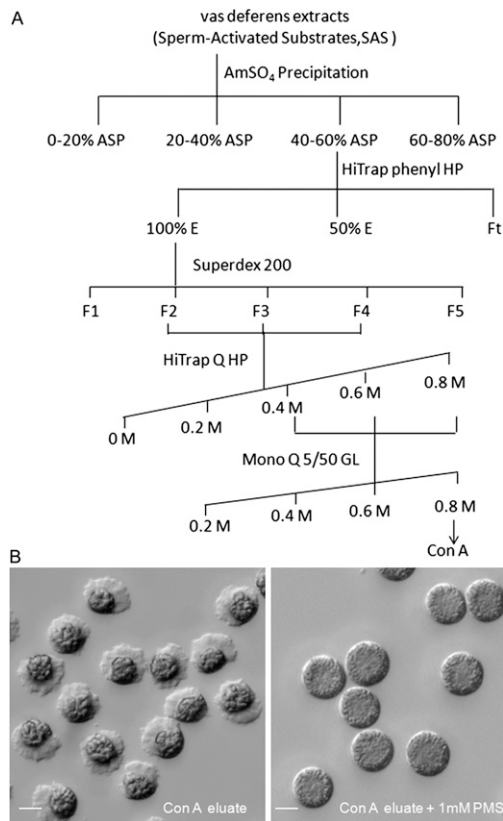


Fig. S8. Enrichment of the sperm activator, a serine protease(s) from SAS. (A) Flowchart of enrichment of sperm activator. (B) The Con A eluate showed strong sperm activating activity (left) and this activity was inhibited by a serine protease inhibitor PMSF (right). (Scale bar, 10 μ m.)

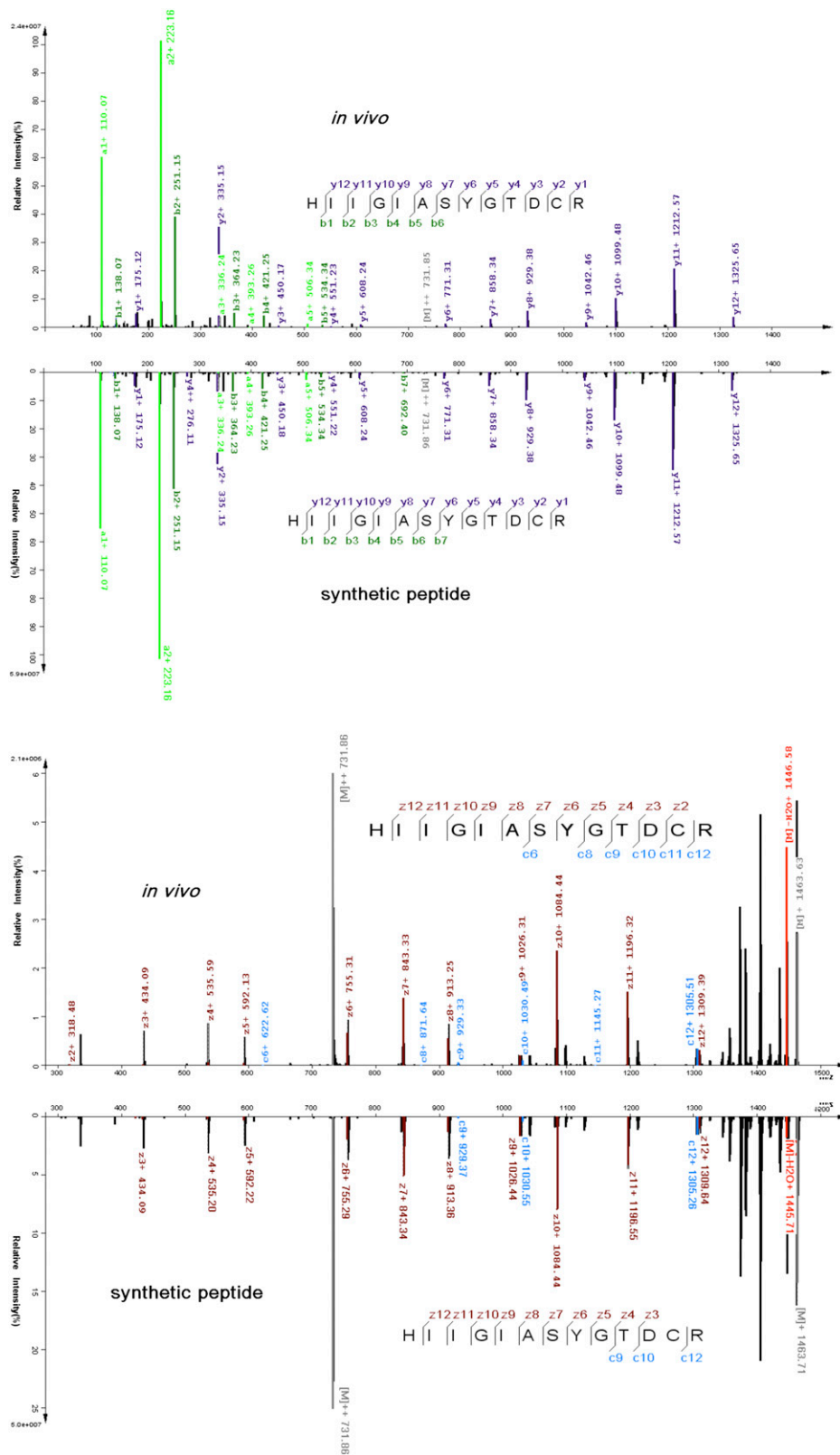


Fig. S9. Fragmentation spectra of a synthetic peptide validating the pNovo identification of a sequence originated from *As_TRY-5*. The HCD (*Upper*) and ETD (*Lower*) spectra of a synthetic peptide, which was synthesized according to a pNovo sequence identification, are fully consistent with the interpretation of the original spectra (labeled *in vivo*) by pNovo.

As_TRY-5:

Peptide	Spectrum	Modification	pFind-e-value	Mascot-score	pNovo-result	pNovo-score
DIRPFNNFICKYTGVCDE	miao-sun-aspN-20101106.3369.3369.2.dta	6,N(Deamidation);	/	32.36	/	/
DISDAIILQRN	miao-sun-aspN-20101106.2971.2971.2.dta	/	1.88E-12	34.75	/	/
TIEDISDAIILQR	mito_sun_90k_20100902.4204.4204.2.dta	/	6.66E-46	56.13	ITEDISDAIILQR	0.749
TIQMALDYRDEFDEWEK	mito_sun_90k_20100902.3054.3054.3.dta	4,M(Oxidation);	2.57E-24	11.72	TIQMTDFEDRYDIAQSEK	0.731
KHIGIASYGTDCR	mito_sun_90k_20100902.2141.2141.2.dta	13,C(Carbamidomethyl);	1.13E-13	22.67	KHCHYEMIGIVF	0.530
DEWEKSELAGGIFT	miao-sun-aspN-20101106.4552.4552.2.dta	/	/	70.44	EDWEKISEIANIFT	0.737
DAIILQRN	miao-sun-aspN-20101106.2462.2462.2.dta	/	9.77E-13	15.56	GEIILQRN	0.776
TQLASWGSASR	mito_sun_90k_20100902.2235.2235.2.dta	/	1.64E-20	63.76	QTASWGSASR	0.775
IFPAFLAQLCLASK	mito_sun_90k_20100902.3302.3302.2.dta	11,C(Carbamidomethyl);	8.87E-19	58.10	MKPAFIAGPICIAASK	0.806
HIIGIASYGTDCR	mito_sun_90k_20100902.2268.2268.2.dta	12,C(Carbamidomethyl);	6.09E-26	68.45	HIIGIASYGTDCR	0.816
LSELAGGIFTDIRPFNNFICK	mito_sun_90k_20100902.3335.3335.3.dta	16,N(Deamidation);20,C(Carbamidomethyl);	5.39E-25	37.74	ISEIAGGIFTDIRPFNNFICK	0.720
IDLSTIPCIESTFR	mito_sun_90k_20100902.3082.3082.2.dta	8,C(Carbamidomethyl);	5.12E-21	70.26	EVISTIPCIESTFR	0.732
DCRTIQMAL	miao-sun-aspN-20101106.2621.2621.2.dta	2,C(Carbamidomethyl);7,M(Oxidation);	1.78E-16	30.69	DCRTIQMAL	0.749
DLSTIPCIESTFR	miao-sun-aspN-20101106.3320.3320.2.dta	7,C(Carbamidomethyl);	1.49E-16	31.24	SIDTIPCIESTFR	0.688
DDNKTIEDIS	miao-sun-aspN-20101106.2426.2426.2.dta	/	/	11.20	/	/
DASRMKTITLTKI	miao-sun-aspN-20101106.2817.2817.3.dta	/	8.35E-36	50.53	/	/
DYRDEF	miao-sun-aspN-20101106.2630.2630.2.dta	/	/	11.59	/	/
DILSEYILDDNKIE	miao-sun-aspN-20101106.3807.3807.2.dta	/	/	52.08	/	/
DASRMKTITLTKIDLSTIPCIESTFR	miao-sun-aspN-20101106.3694.3694.3.dta	21,C(Carbamidomethyl);	2.01E-37	33.88	KTWHGHRITITIKIDISTICRTMKISIPN	0.595
DEFDEWEKSELAGGIFT	miao-sun-aspN-20101106.5047.5047.2.dta	/	/	61.02	EDFDEWEKISEISNIMV	0.653
GDLAIVELDAK	mito_sun_90k_20100902.2634.2634.2.dta	/	8.16E-16	32.98	GDLAIVEIDAK	0.732
DIRPFNNFICKYTGVC	miao-sun-aspN-20101106.3441.3441.2.dta	10,C(Carbamidomethyl);16,C(Carbamidomethyl);	1.85E-16	25.76	KCSPFRISICRYNNGVM	0.664
YTGVCDEEDILSEYILDDNK	mito_sun_90k_20100902.3203.3203.2.dta	5,C(Carbamidomethyl);	3.10E-18	41.85	EDDDCRTYISEYIIDDNK	0.685
DSGAGMVYSNKRKHIGIASYGTDCRTIQMAL	miao-sun-aspN-20101106.2852.2852.4.dta	26,C(Carbamidomethyl);31,M(Oxidation);	4.61E-08	/	/	/
DIRPFNNFICKYTGVC	miao-sun-aspN-20101106.3413.3413.2.dta	6,N(Deamidation);10,C(Carbamidomethyl);16,C(Carbamidomethyl);	1.46E-05	/	/	/
DECQRNQIKALTSKIRNVIIPKVF	miao-sun-aspN-20101106.3864.3864.4.dta	3,C(Carbamidomethyl);	0.00394	/	/	/

Fig. S10. Comparison of the de novo sequencing result and the retrospective database search result of As_TRY-5. After the identification and cloning of As_TRY-5, the original MS data were searched against a database containing this protein using Mascot and pFind, and As_TRY-5 was reidentified as the top candidate. Most of the As_TRY-5 peptides identified by database search had been identified by pNovo. Highlighted in yellow are the identical regions between a pNovo-identified sequence and that by database search from the same spectrum.

```

-----|
ORIGINAL-SEQUENCE |
MASCOT----RESULT |
PFIND-----RESULT |
DENOVO-----RESULT |
-----|

```

As_TRY-5:

MAESLRRIVVITMISLSHGFSARLSPNEAKIFSEYCGLSKNERNCPHVGLAMHGRFARNFEASWAVHLHSAYPGLEADCGGTLITEQHILTAACFFNEICYNSPTKLLKNNTWLAS

WTVYHSGECIPMSEDECQRNQIKALTSKIRNVIIPKVFIDEKCTRGLAIVELDAKIFPAFLAQPCLCLASKNNKIPLKTQLASWGSASRMKTITTLTKIDLSTIPCP IESTFRDVICVN

-----GDLAIVELDAKIFPAFLAQPCLCLASK-----TQLASWGSASRMKTITTLTKIDLSTIPCP IESTFR-----

-----DECQRNQIKALTSKIRNVIIPKVF-----GDLAIVELDAKIFPAFLAQPCLCLASK-----TQLASWGSASRMKTITTLTKIDLSTIPCP IESTFR-----

-----IFPAFLAQPCLCLASK-----TQLASWGSASR-----IDISTIPCP IESTFR-----

-----GDIAIVEIDAKIFPAFLAQP I-----DLSTIPCP IESTFR-----

-----GDIAIVEIDAK-----KTWHGHRITITKIDISTICRTMKSIPN-----

ESQDQNMCRGDSGAGMVYSNKGKRIIGIASYGTDCRTIQMALDYRDEFDEWEKLSLAGGIFTDIRPFNNFICKYTGVCDEDEDILSEYILDDNKTIEDISDAIILQRN

-----KHIIGIASYGTDCRTIQMALDYRDEFDEWEKLSLAGGIFTDIRPFNNFICKYTGVCDEDEDILSEYILDDNKTIEDISDAIILQRN-----

-----DSGAGMVYSNKGKRIIGIASYGTDCRTIQMALDYRDEFDEWEKLSLAGGIFTDIRPFNNFICKYTGVCDEDEDILSEYILDDNKTIEDISDAIILQRN-----

-----HIIGIASYGTDCR-----DIRPFNNFICKYTGVC-----TIEDISDAIILQR-----

-----TIQMALDYRDEFDEWEK-----YTGVCDEDEDIIEYIIDDHK-----DAIILQRN-----

-----TIQMTDFEDRYDIAQSEK-----DIRPFNNFICKYTGVC-----DDNKTIEDISDAIILQRN-----

-----KHCHYEMIGIIVF-----EDWEKISEIANIFT-----DDNKTIEDISGEIILQRN-----

-----DCRTIQMAL-----ISEIAGGIFTDIRMINNFICK-----

-----EDFDEWEKISEISNIMV-----EDDCRITYIIEYIIDDNK-----

Fig. S11. Sequence coverage of As_TRY-5 by de novo sequencing and database search. Peptide sequences identified by pNovo (green), Mascot (blue), or pFind (red) are mapped to the As_TRY-5 protein sequence. Yellow highlighting indicates the pNovo-identified sequence tags that match perfectly to the As_TRY-5 protein sequence.