CELL BIOLOGY

Giant proteins in a giant cell: Molecular basis of ultrafast Ca^{2+}-dependent cell contraction

Jing Zhang[1]†, Weiwei Qin[1,2]†, Che Hu[1,3]†, Siyu Gu[1,2]†, Xiaocui Chai[1], Mingkun Yang[1], Fang Zhou[1], Xueyan Wang[1,2], Kai Chen[1], Guanxiong Yan[1], Guangying Wang[1], Chuanqi Jiang[1], Alan Warren[4], Jie Xiong[1]*, Wei Miao[1,2,5,6]*

The giant single-celled eukaryote, *Spirostomum*, exhibits one of the fastest movements in the biological world. This ultrafast contraction is dependent on Ca^{2+} rather than ATP and therefore differs to the actin-myosin system in muscle. We obtained the high-quality genome of *Spirostomum minus* from which we identified the key molecular components of its contractile apparatus, including two major Ca^{2+} binding proteins (Spasmin 1 and 2) and two giant proteins (GSBP1 and GSBP2), which act as the backbone and allow for the binding of hundreds of spasmins. The evidence suggests that the GSBP-spasmin protein complex is the functional unit of the mesh-like contractile fibrillar system, which, coupled with various other subcellular structures, provides the mechanism for repetitive ultrafast cell contraction and extension. These findings improve our understanding of the Ca^{2+}-dependent ultrafast movement and provide a blueprint for future biomimicry, design, and construction of this kind of micromachine.

INTRODUCTION

The actin-myosin and dynein/kinesin-tubulin systems are two well-known adenosine triphosphate (ATP)–dependent molecular mechanisms that facilitate cellular movement by the sliding of macromolecules (*1*–*3*). Less well known is the Ca^{2+}-dependent mechanism that enables ultrafast contraction in certain single-celled eukaryotes such as the ciliated protozoans *Vorticella* and *Spirostomum* (*4*–*6*). The giant heterotrich ciliate *Spirostomum* is characterized by its rod-like cell that is usually one to several millimeters in length. Upon contraction, the cell of *Spirostomum* typically twists and shortens its length up to 75% within 5 ms (*7*, *8*). In past decades, several studies have been carried out to elucidate the subcellular structures and biophysical features of this fascinating system (*7*, *9*–*13*).

The surface of *Spirostomum* contains numerous continuous spiral ciliated grooves that traverse its length (Fig. 1A) (*9*, *10*, *14*). There are two types of fibrillar systems beneath the cell membrane. One comprises microtubule-based fibrillar bundles that run parallel to the base of the spiral grooves and either attach to basal bodies or lie adjacent to them (Fig. 1B) (*10*, *14*). When the cell contracts, these bundles change in orientation and lie parallel to the grooves, resembling a compressed spring with high potential energy (*7*). It is hypothesized that the microtubule-based fibrillar bundles provide turgidity to the cell and may play an antagonistic role to cell contraction, thereby facilitating cell extension (*7*).

The other fibrillar system below the cell membrane of *Spirostomum* occurs all over the body and comprises fibrillar bundles arranged in a mesh-like pattern (*9*, *10*). This mesh-like filamentous structure was thought to be analogous to the myoneme that is found in another ultrafast contractile ciliated protozoan, *Vorticella*, and was thought to be the functional actuator for the ultrafast contraction of the cell (*9*). Further evidence for this is that both systems rely on Ca^{2+} rather than ATP for the contraction (*11*).

Little is known about the molecular components of the mesh-like contractile fibrillar system in *Spirostomum*. Limited knowledge come from *Vorticella*, in which proteins that are called spasmins and belong to the calcium-binding centrin family were identified in the myoneme of the stalk (*15*, *16*). Spasmins were therefore assumed to exist in the mesh-like contractile fibrillar system of *Spirostomum*, and this was verified by immunofluorescence (IF) staining methods (*7*, *13*). A molecular model of contraction based on spasmin-driven electrostatic repulsion was proposed (*17*). However, it was subsequently demonstrated that spasmin alone does not provide sufficient power for cell contraction in *Vorticella* (*5*, *18*). In 2003, Kilmartin (*19*) found that a protein named Sfi1, which is involved in yeast spindle pole body duplication, functions as a molecular backbone and has conserved centrin-binding sites. Considering that spasmin belongs to the centrin family, the existence in *Vorticella* of a backbone protein analogous to Sfi1, which provides binding sites for spasmins, has been suggested (*5*, *18*). However, no sequence of spasmin or a backbone component has been identified in *Spirostomum*.

In the present study, we sequenced the genome of *Spirostomum* with high quality from which calcium-binding spasmins was identified and characterized. Two giant proteins were found to function as a molecular backbone and therefore are key components of the mesh-like contractile fibrillar system. A model involving the coupling of different subcellular structures was proposed for the ultrafast contraction and extension cycle of *Spirostomum*. These findings greatly improve our understanding of ultrafast contraction in *Spirostomum* cells and provide a blueprint for future biomimicry including the design and construction of ultrafast contractile molecular machines both in vitro and in synthetic cells.

[1]Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China. [2]University of Chinese Academy of Sciences, Beijing 100049, China. [3]Harbin Normal University, Harbin 150025, China. [4]Department of Life Sciences, Natural History Museum, London SW7 5BD, UK. [5]State Key Laboratory of Freshwater Ecology and Biotechnology of China, Wuhan 430072, China. [6]CAS Center for Excellence in Animal Evolution and Genetics, Kunming 650223, China.
*Corresponding author. Email: miaowei@ihb.ac.cn (W.M.); xiongjie@ihb.ac.cn (J.X.)
†These authors contributed equally to this work.

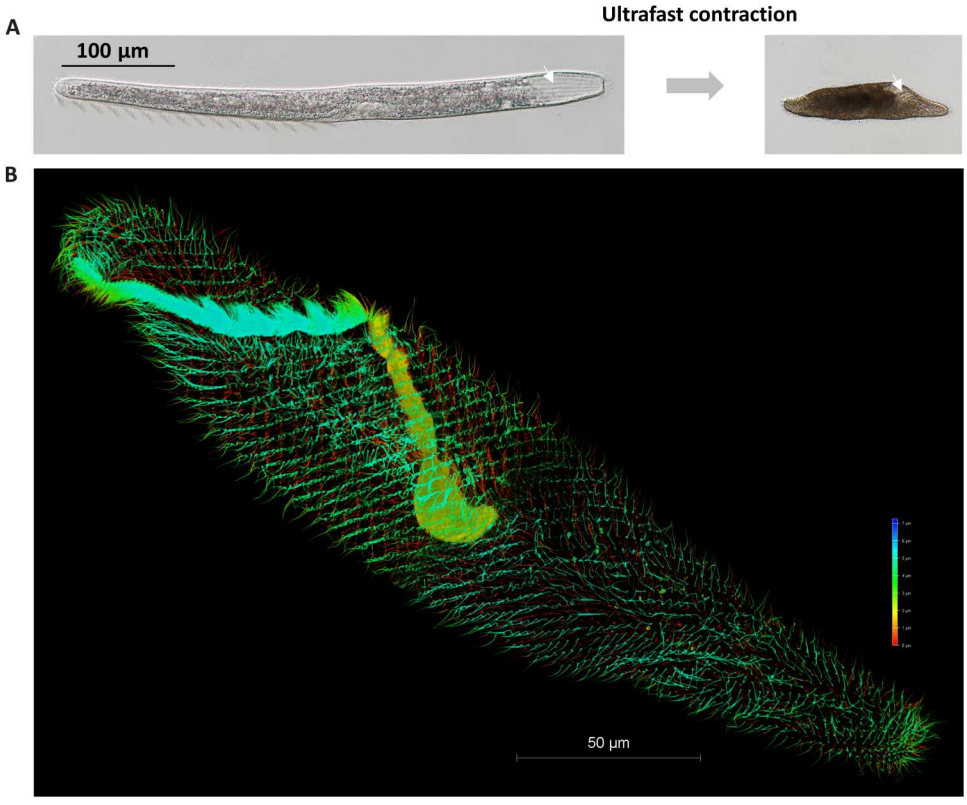

Fig. 1. The ultrafast contractile single-celled eukaryote, *S. minus*. (**A**) Extended state (normal state, left) and contracted state (right) of *S. minus*. The white arrow indicates a groove on the surface of *S. minus* cell. (**B**) Three-dimensional image of the microtubule-based fibrillar bundles stained using tubulin–Atto 488 (*73*), imaged using super-resolution microscopy. Gradient colors were used for different Z-depth.

RESULTS

Two calcium-binding spasmins were identified in the mesh-like contractile fibrillar system of *Spirostomum*

As a part of the Protist 10,000 Genomes Project (*20*), the macronuclear genome (hereafter genome) of *Spirostomum minus* (Fig. 1) was sequenced with high quality through a combination of Nanopore and Illumina technology. The genome of *S. minus* is 66.6 Mb (table S1). On the basis of the N50 and scaffold numbers, the quality of this genome is far higher than the previously published *Stentor coeruleus* genome, a species that belongs to the same class as *Spirostomum* (table S1) and is comparable to the genomes of two well-studied ciliates, *Paramecium tetraurelia* and *Tetrahymena thermophila* (table S1). To improve the accuracy of gene predictions, the results of different prediction tools were integrated and a total of 38,210 genes were annotated.

Since calcium-binding proteins, such as the centrin, calmodulin, and calretinin, usually harbor the EF-hand domain (*21*–*24*), all EF-hand domain–containing proteins were identified in the genomes of *S. minus*, *S. coeruleus*, and *P. tetraurelia*. There were 468, 1000, and 767 such proteins, respectively, in these three species. To identify the potential spasmin proteins, a phylogenetic tree was constructed using all these EF-hand proteins and incorporating three spasmin proteins previously reported in *Vorticella* and *Zoothamnium* (*15*, *25*). To identify the spasmin proteins in *S. minus*, we focus on a large clade in the phylogenetic tree of EF-hand proteins, which contains the three spasmin proteins in *Vorticella* and *Zoothamnium* (Fig. 2 and fig. S1) and shares two common features: (i) There is a disordered region in the N terminus of the protein sequences (in majority cases), a feature that is lacking in other EF-hand proteins that closely cluster with this clade (fig. S1); (ii) these proteins are longer than the closely clustered EF-hand proteins (96 amino acids versus 161 to 198 amino acids; fig. S1). Considering the mesh-like contractile fibrillar system that is present under the entire cell membrane, it is reasonable to hypothesize that the abundance of spasmins is high. Mass spectrometry (MS) was used to identify the proteins in *S. minus*, and 10 EF-hand proteins were found to have >10 matched peptides (Fig. 2). Immunoprecipitation-MS (IP-MS) was performed to narrow down the candidates of spasmin proteins, using the antibodies of the centrin and two giant backbone proteins of the mesh-like contractile fibrillar system [giant spasmin binding protein (GSBP) 1 and 2; see below]. A total of five spasmin candidates were identified (Fig. 2). These five candidates fall into two clusters. One contains two proteins (Fig. 2, purple dashed curve) and clusters closely with the ICL1e family of *Paramecium* (*26*), showing high abundance in whole-cell MS and IP-MS results, especially for the IP results using antibodies of the two giant backbone proteins (Fig. 2), and seems to be the major spasmin components in the mesh-like contractile fibrillar system. We named them Spasmin 1 and 2, respectively. Another cluster contains three candidates that are almost identical (overall sequence similarity: 96.5 to 97.7%), differing only in several amino acids in the N terminus, and thus, the MS peptides cannot be specifically assigned to each of these three candidates (Fig. 2, blue dashed curve). These three candidates

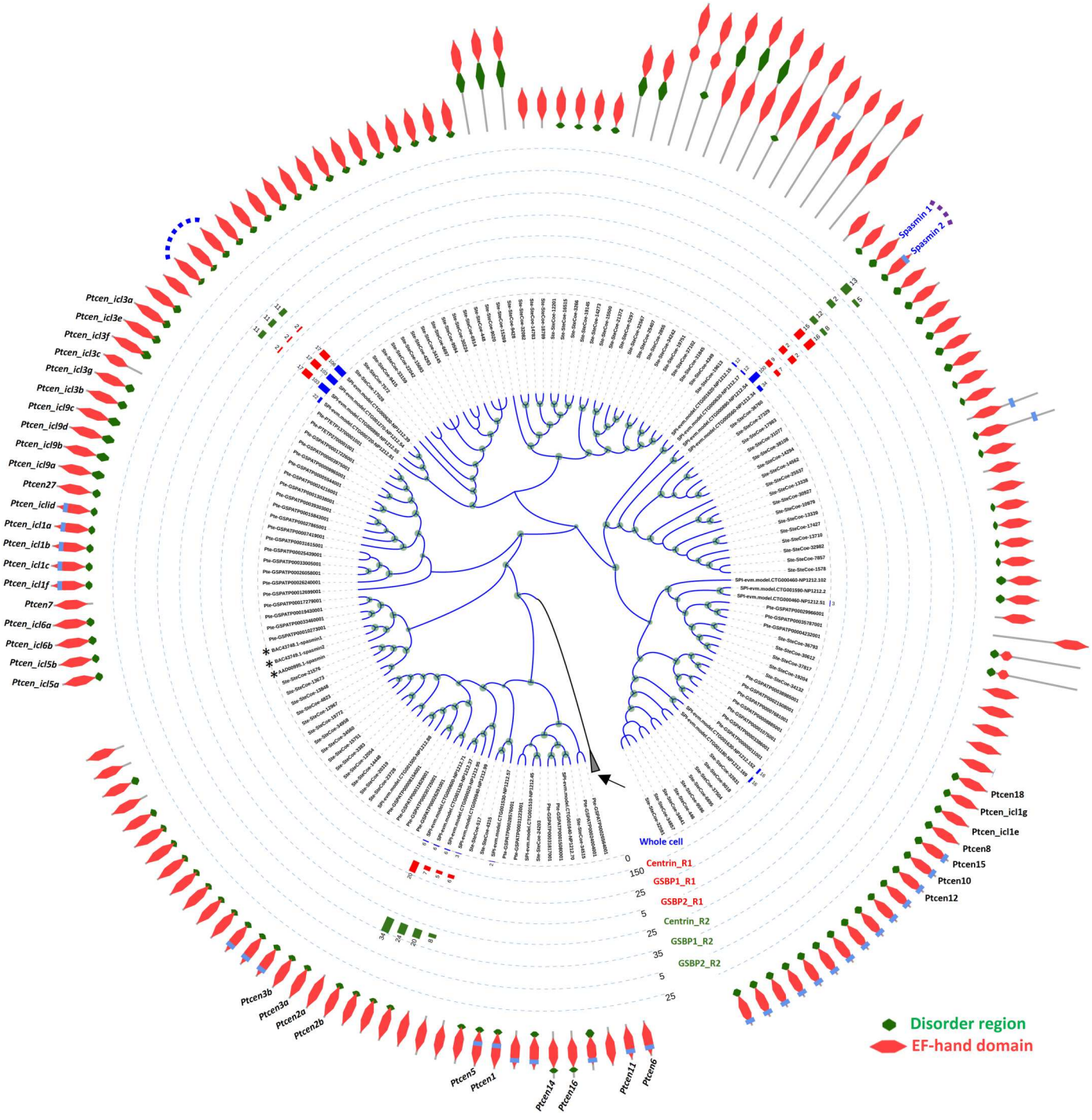

Fig. 2. Identification of spasmins in *S. minus*. Phylogenetic tree showing all EF-hand proteins in *S. minus*, *S. coeruleus*, and *P. tetraurelia*. Asterisks indicate the three previously reported spasmins in *Vorticella* and *Zoothamnium*. Only EF-hand proteins closely clustered with spasmins in *Vorticella* and *Zoothamnium* are shown; the other EF-hand proteins (2087) are collapsed (black arrow). Predicted domain organization is assigned to each EF-hand protein. All the gene names of EF-hand proteins in *Paramecium* are labeled. Taxon names prefixed with "SPI," "Ste," and "Pte" are genes from *S. minus*, *S. coeruleus*, and *P. tetraurelia*, respectively. Bar chart shows the abundance (matched peptide number) of EF-hand proteins in MS or IP-MS experiments. Whole-cell MS analysis using *S. minus* cells. Centrin, IP-MS results using the anti-centrin antibody (mouse); GSBP1, IP-MS results using the anti-GSBP1 antibody (rabbit); GSBP2, IP-MS results using the anti-GSBP2 antibody (rabbit); R1 and R2, two replicates of IP-MS experiments. Dashed curve, EF-hand proteins found in the IP-MS results of GSBP1 and/or GSBP2. Two major EF-hand proteins, which show high abundance in whole-cell MS, centrin, GSBP1, and GSBP2 IP-MS, are named Spasmin 1 and 2, respectively.

were only found in one IP-MS experiment (GSBP2 antibody) at low abundance (Fig. 2). Considering their high abundance in whole-cell MS and IP-MS using centrin antibodies, these three candidates may not function as spasmins or may be weakly associated with the two giant backbone proteins. However, further experiments need to verify whether these three candidates are spasmins.

Two giant proteins serve as spasmin-binding backbones —The key components of the mesh-like contractile fibrillar system

In yeast cells, centrins bind to a backbone protein Sfi1. This protein contains 21 motifs with the consensus sequence AX7LLX3F/LX2WK/R and is known as the centrin-binding site (*19*, *27*). It is noteworthy that tryptophan (W) is especially well conserved in these motifs. Using this motif or its relaxed version, a Sfi1-like protein was identified in *P. tetraurelia* (*28*). This inspired us to screen the spasmin-binding backbone in *S. minus*. AX7LLX3F/LX2WK/R motif–containing proteins and tryptophan-rich proteins were screened in all predicted protein sequences. Two high-profile tryptophan-rich proteins, each with a huge molecular weight and length, were found to harbor Sfi1-like motifs. These two proteins, which we named GSBP1 and GSBP2, have lengths of 13,125 (1458 kDa) and 17,974 (2008 kDa) residues, respectively (Fig. 3A).

Several lines of evidence confirmed that *GSBP1* and *GSBP2* are not artifacts of genome assembly or gene mis-prediction. The raw Nanopore long reads were mapped back to the assembled genome, and it was found that many single reads could span the whole gene body of *GSBP1* and *GSBP2*, suggesting that the assembly of these regions is correct (Fig. 4). Furthermore, transcriptomes were sequenced to recover the transcripts of these two genes. However, no transcriptome sequencing protocol, including the standard Illumina RNA sequencing (RNA-seq), single-cell SMRT-Seq2, and Nanopore full-length mRNA sequencing, recovered the transcripts when a poly(A) mRNA purification step was included, probably due to the excessive length of these two genes. Therefore, the protocol of transcriptome sequencing was modified to skip the poly(A) mRNA purification step, and library construction was directly performed using total RNA with the first-strand cDNA reverse-transcribed by random hexamer. On the basis of this strategy, the transcripts of *GSBP1* and *GSBP2* were successfully recovered (Fig. 4). At the protein level, prediction of GSBP1 and GSBP2 was also supported by high numbers of matched peptides in whole-cell MS experiments (Fig. 4). Last, the GSBP protein sequences were used to generate antibodies, and these two proteins were confirmed by Western blot (see next).

On the basis of the correct prediction of GSBP1 and GSBP2, we analyzed the features of these two proteins. GSBP1 and GSBP2 are homologs and share 54% sequence identity (fig. S2). GSBP1 and GSBP2 harbor 184 and 267 tryptophans, respectively (Fig. 3B). The 23 amino acids around each tryptophan, which is assumed to be a unit of centrin binding, were extracted, and it was found that they share a highly conserved motif with consensus sequence KGA/VM/LRXLXK/HKLAQ/DKPKDAXXXWR/K (Fig. 3B), suggesting tandem repeats in these proteins (Fig. 3B). This motif differed in most of sites from the one found in yeast, but two amino acids, the WR/K, are conserved. This would explain why homological search tools (e.g., BLAST) failed to find Sfi1 homologs in *S. minus*. An interesting phenomenon is that tryptophans seem to be uniformly distributed in both GSBP1 and GSBP2, and the interval between two neighboring tryptophan sites is 69 amino acids (Fig. 3C and fig. S3), i.e., three times a 23–amino acid unit. This finding leads us to question whether a tryptophan is necessary to

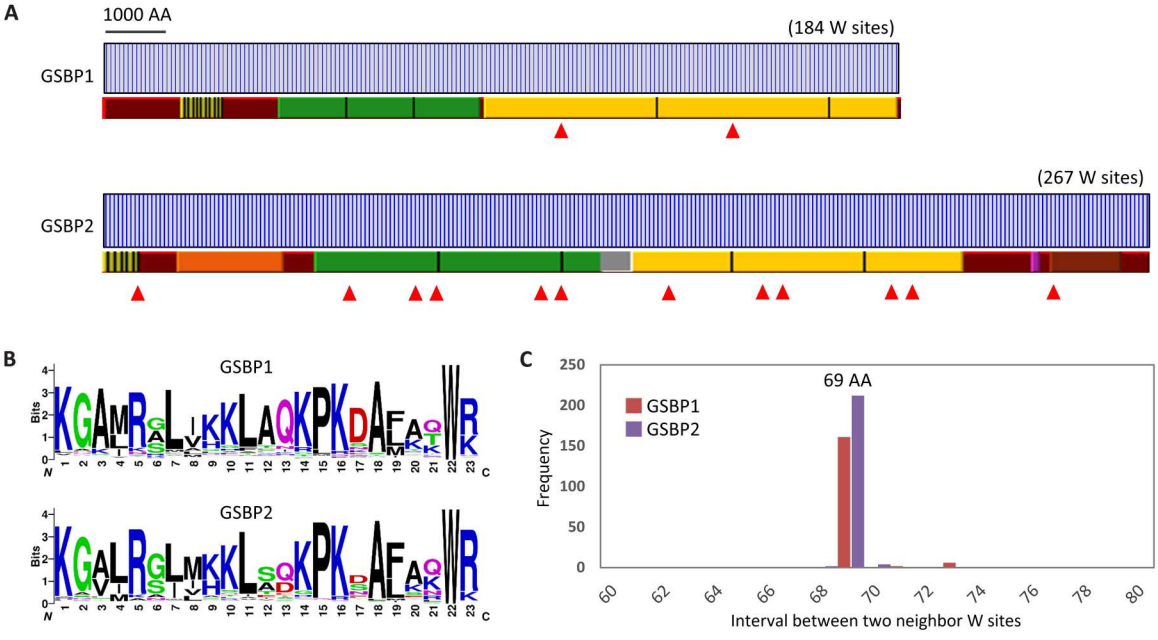

Fig. 3. Two giant proteins (GSBP1 and GSBP2) serve as spasmin-binding backbones. (**A**) Features of GSBP1 and GSBP2 in *S. minus*. Blue vertical bar indicates the tryptophan (W) sites. Repeat regions are indicated by bars with different colors. The same color means repeat regions (separated by black vertical bars) with high sequence identity. Filled red triangles indicate coiled-coil regions. (**B**) Sequence logos show the motif of 23 amino acids (AA) around the tryptophan sites. (**C**) The 69–amino acid intervals between two neighboring tryptophan sites.

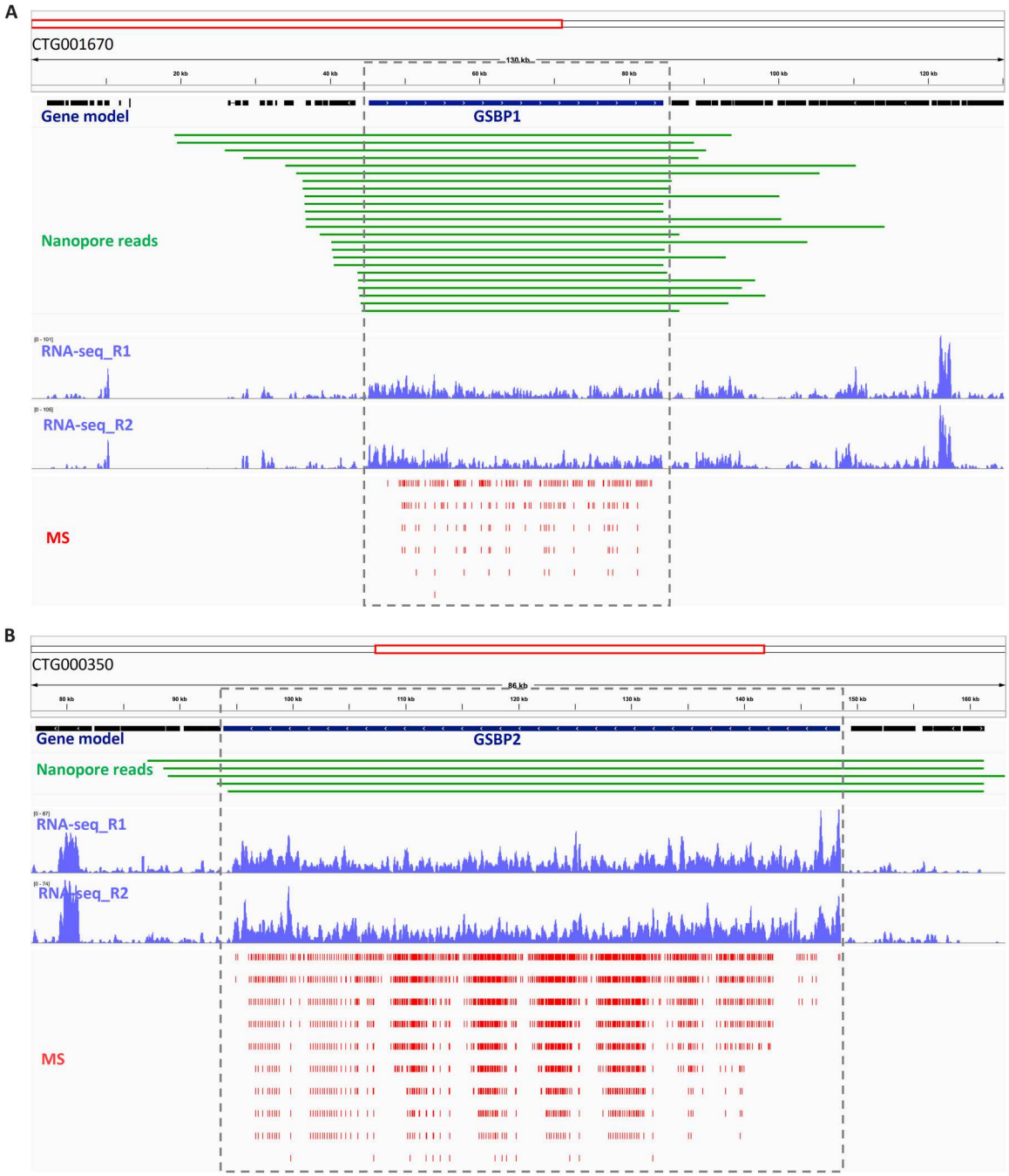

Fig. 4. Validation of two giant proteins, GSBP1 and GSBP2. (**A**) The validation of *GSBP1* gene model; (**B**) The validation of *GSBP2* gene model. The gene model was predicted as described in Materials and Methods. No introns were found in either *GSBP1* or *GSBP2*. Nanopore reads: The raw sequencing reads mapping back to genome assembly. Note that only reads that span the whole gene body are shown. RNA-seq: Read coverage of transcriptome sequencing of total RNA without poly(A) mRNA purification. R1 and R2, replicates 1 and 2. MS: The peptides matched to *GSBP1* and *GSBP2* in whole-cell protein MS identification. Note that *GSBP1* and *GSBP2* have many internal repeats, and if a peptide matched multiple loci, then all of these are shown.

the binding of spasmin. One possibility is that different spasmins, e.g., Spasmin 1 or 2 (Fig. 2), bind preferentially to different motifs. If this is true, GSBP1 and GSBP2 could bind up to 552 and 801 molecules of spasmin, respectively, and form huge protein complexes. Analysis of GSBP1 and GSBP2 showed that most regions of these proteins are α helices. Furthermore, several coiled coils were found, especially in GSBP2 in which 12 coiled-coil regions were identified (Fig. 3A). These findings indicate that the protein complexes resemble ropes that could be coiled as either a homomer or a heteromer.

Since GSBP1 and GSBP2 are too long to be structurally modeled, a small fraction of the protein sequence around the W site was used

to predict the GSBP-spasmin complex using AlphaFold2-multimer (*29*). The predicted structure shows that spasmin could bind to GSBP, similar to a hand grabbing a stick (Fig. 5A). On the basis of the sequence information, antibodies against GSBP1 and GSBP2 were generated in both mouse and rabbit. Using IF, colocalization between spasmin and GSBP1 or GSBP2 was found in the mesh-like contractile fibrillar system (Fig. 5B and fig. S4), suggesting an interaction between spasmin and GSBP1 or GSBP2. It is noteworthy that the fibrillar bundles in the mesh-like structure run in two orientations. One runs parallel to the grooves on the cell surface (Fig. 5B, blue arrowhead), whereas the other runs at an angle both to the first one and to the grooves (Fig. 5B, yellow arrowhead). It seems that the fibrillar bundles that lie parallel to the grooves are more continuous than the other bundles that do not extend around the entire cell in some regions (Fig. 5B, arrow). This indicates that assembly of the first fibrillar bundles may depend on elements under the grooves, whereas assembly of the second fibrillar bundles may solely depend on the assembly of the first fibrillar bundles. The findings of a previous study suggest that the mesh-like structure is organized as numerous parallelograms when the cell is fully extended (*7*). In the contracted state, as shown in Fig. 5B, it seems that the shrinking of individual parallelograms causes them to become round as the fibrillar bundles shorten and thicken in at least some regions. We measured the side lengths of the parallelograms in cells with different degrees of contraction (fig. S4). The side lengths are about 1.8 to 2.8 μm, which corresponds closely to the lengths of GSBP1 and GSBP2 (2.7 and 2.0 μm, respectively) if the structure of these proteins is a long α helix.

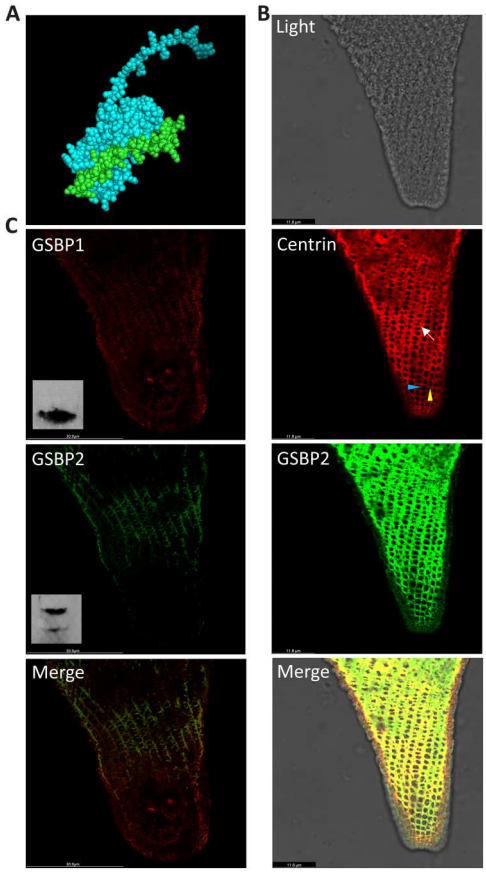

Fig. 5. Colocalization of spasmin and GSBP in *S. minus*. (**A**) AlphaFold2 prediction of partial GSBP sequences (green) and spasmin (cyan). (**B**) Colocalization of GSBP2 and spasmin. Red: IF using anti-centrin antibody (mouse) represents spasmins that have high abundance in the cell based on MS result. Green: IF using anti-GSBP2 antibody (rabbit). Blue and yellow arrowheads: Contractile fibrillar bundles running in different directions. Note that all layers of IF images are merged and shown. (**C**) Localization of GSBP1 and GSBP2. Red: IF using anti-GSBP1 antibody (mouse). Inset: Western blot result of GSBP1. Green: IF using anti-GSBP2 antibody (rabbit). Inset: Western blot result of GSBP2; the two bands are GSBP1 and GSBP2, respectively. Note that only one layer of IF images is shown.

The sequence used to generate GSBP1 antibody has only 28% identity with GSBP2, and the sequence used to generate GSBP2 antibody shows 38% identity with GSBP1. Only one band was found in the Western blot using GSBP1 antibody, whereas two bands, which correspond to GSBP1 and GSBP2, respectively, were found in Western blot using GSBP2 antibody (Fig. 5C and fig. S5). These results suggest that the GSBP1 antibody has specificity. Furthermore, IF staining showed that GSBP1 is distributed throughout the mesh-like contractile fibrillar system (Fig. 5C and fig. S4). IF staining using GSBP2 antibody shows similar results to GSBP1 (Fig. 5C). However, since the GSBP2 antibody does not have specificity, the question of whether GSBP1 and GSBP2 have differential localizations in the mesh-like structure is still not clear. In the GSBP1 IP-MS, we found high abundance of GSBP2, suggesting that there is an interaction between GSBP1 and GSBP2. It is highly possible that GSBP1, GSBP2, and spasmins form a large protein complex.

RNAi supports the role of GSBP in contraction of *Spirostomum*

RNA interference (RNAi) has been accomplished in only a few ciliates. The successful application of RNAi in *Stentor* (*30*, *31*), another giant ciliate that belongs to the same class as *Spirostomum*, suggests that it may be possible to investigate the GSBP function using this method. However, we found that *S. minus*, which was cultured either in a medium alone (sterilized mineral water plus wheat grains) or in a medium supplemented with bacteria cells (e.g., *Escherichia coli*), may have some abnormal cells. This makes the RNAi experiment more complicated. To obtain reliable results, we first carefully observed the cell division and cell death processes in *S. minus*. During cell division, cytokinesis occurs halfway along the length of the body axis when one daughter cell twists and then separates from the other (fig. S6). Cell death involves the shortening of the cell as it gradually becomes spherical before finally disintegrating (fig. S7).

GSBP RNAi should result in the partial or full loss of cell contractility in *Spirostomum*. To check the contractility, different Ca^{2+} concentrations were used to treat *S. minus* cells and the contraction frequency was investigated. Ca^{2+} concentrations from 10^{-6} to 10^{-2} M in the culture medium could stimulate cell contraction while maintaining the viability of the cell. As the Ca^{2+} concentration was increased from 10^{-6} to 10^{-3} M, the average time of cell contraction was reduced from 38 to 16 s. Therefore, we checked the contractility of cells at 10^{-3} M Ca^{2+} (movie S1). The *E. coli* HT115 strain, which was transformed with the L4440 vector that carries the sequence of *GSBP1* and *GSBP2*, was used to feed *S. minus* cells and induce RNAi. HT115 cells without a vector or with an empty L4440 vector were used as controls.

In the RNAi experiments, three categories of abnormal cells were found. The first category featured cells that break down at either the anterior (Fig. 6A) or posterior (Fig. 6B) end of the cell, and a small part that may or may not contain the nuclear DNA depending on the position of the breakdown, and which eventually disappears from the culture. The second category featured cells with a distortion either at the anterior (Fig. 6C) or posterior end (Fig. 6D). The third category featured distorted cells that were sometimes entirely twisted (Fig. 6E). However, abnormal cells similar to those in the first and second categories could also be found in the controls (HT115 cells without vectors or with empty L4440 vectors) if the concentration of bacteria is high. The percentage of abnormal cells among HT115 cells, HT115 cells with empty L4440 vectors, HT115 cells with *GSBP1* target sequences, and HT115 cells with *GSBP2* target sequences showed similar levels, ranging from 0 to 8%, 0 to 10%, 0 to 10%, and 0 to 10%, respectively, in different replicates (table S1). These cells had the ability to contract in a 10^{-3} M Ca^{2+} solution. The abnormal cells in the third category were only found in the *GSBP1* and *GSBP2* RNAi experiments (0 to 4% in different replicates). When placed in 10^{-3} M Ca^{2+} solution, these cells did not contract in the course of microscopy observation (average time, 267 s), compared to the average interval contraction time of 16 s in wild-type cells. This result suggested the loss of the ability to contract in the abnormal cells (movie S2). Therefore, it is highly likely that these are the phenotype of *GSBP1* and *GSBP2* RNAi. To further confirm this, we checked the change of mesh-like contractile network using IF in E type cells. Under the *GSBP2* RNAi, disruption of the mesh-like structure could be clearly observed (Fig. 6, F and G). It seems that both the first and second fibrillar bundles were disrupted, and the remnants of the mesh-like structure after disruption resembled separated foci (Fig. 6, F and G, white arrowhead). The RNAi and IF results support the assertion that GSBP has an important role in cell contraction. Furthermore, because of the distortion of the RNAi cells, GSBP may also function in maintaining cell shape.

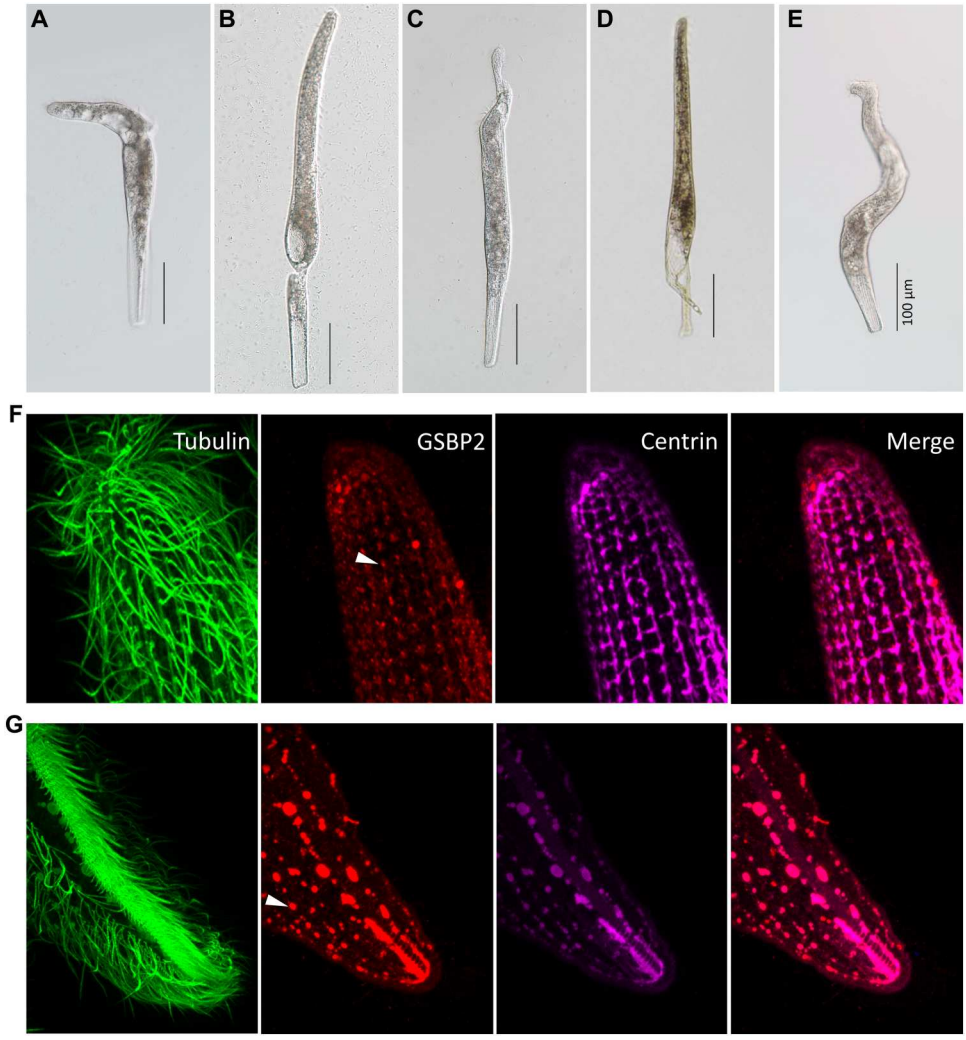

Fig. 6. RNAi of two giant proteins. (**A** and **B**) Cells breaking down at anterior or posterior, respectively. (**C** and **D**) Cells showing distortion at anterior and posterior, respectively. (**E**) Distortion and twisting of entire cell. Scale bars, 100 μm. IF images showing the changes of mesh-like structure in *GSBP2* RNAi cell using anti-GSBP2 antibody (rabbit). (**F**) Posterior region. (**G**) Anterior region.

A sophisticated system coupling different subcellular structures fulfill the contraction and extension cycle

DAPI (4′,6-diamidino-2-phenylindole) and IF staining using antibodies of different proteins revealed that *S. minus* has a sophisticated system that couples different subcellular structures to contract and extend the cell. The microtubule-based fibrillar bundles, which stain brighter than other regions of the cell in IF experiments, run parallel with the grooves (Fig. 7, A and B, yellow arrowhead, and figs. S8 and S9). The mesh-like structure, comprising GSBP and spasmin fibrillar bundles (Fig. 7, A and B, and figs. S8 and S9), is the key apparatus for cell contraction. The fibrillar bundles in the mesh-like structure fuse together at the posterior end of the cell and seems to attach to a ring structure (fig. S8A, arrow), possibly to leave a pathway for the expulsion of undigested food (fig. S10). The first fibrillar bundles (see above) clearly lie next to the microtubule-based fibrillar bundles in both extended and contracted cells (Fig. 7A), especially in the oral apparatus (fig. S8A, blue arrowhead). An interesting finding is that the ciliary basal bodies seem to locate

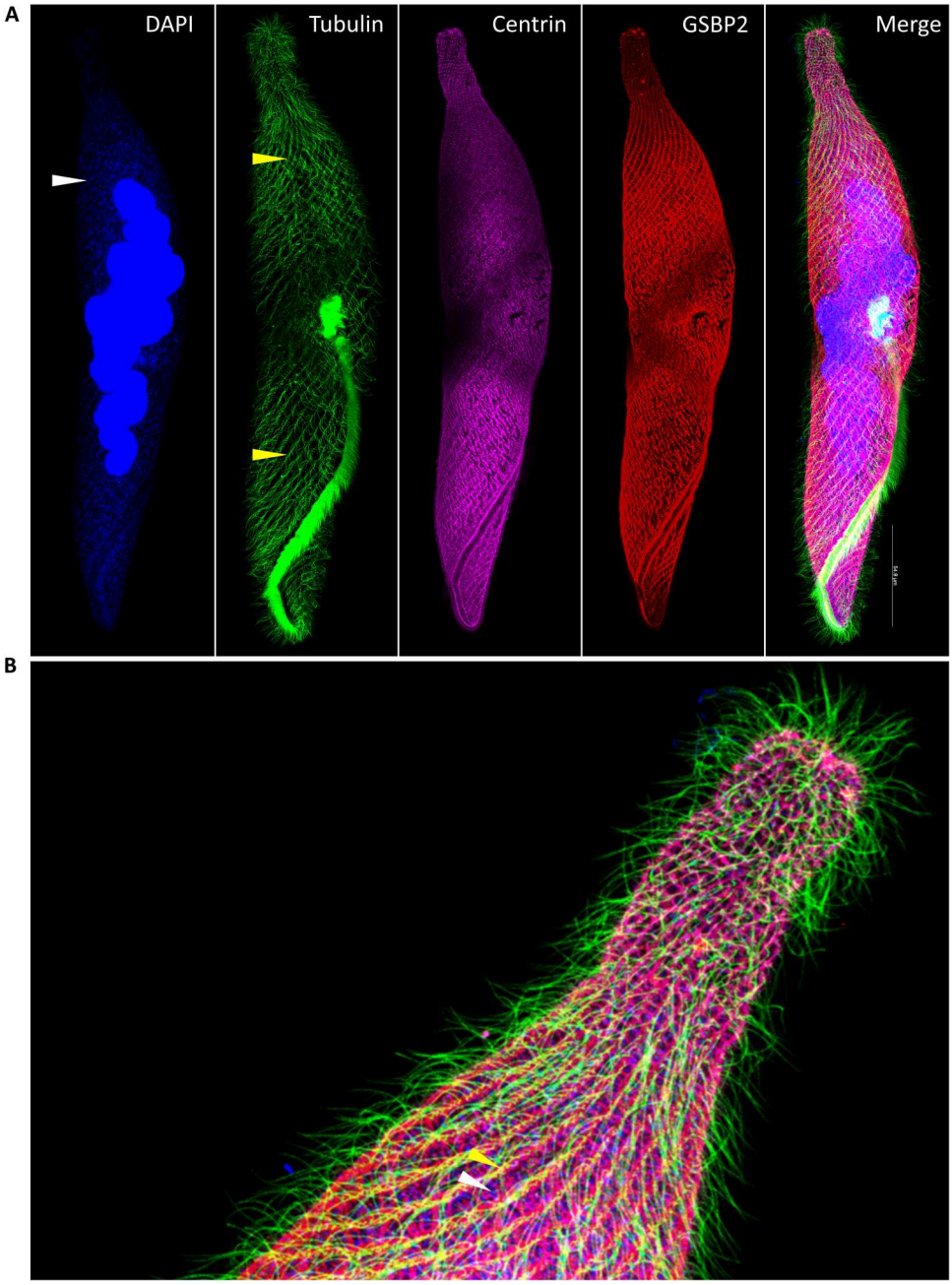

Fig. 7. Distribution patterns of subcellular structures in *S. minus*. (**A**) Super-resolution IF image shows the localization of tubulin, spasmin, and GSBP2. Blue: DAPI staining of DNA. Note that the foci (white arrowhead) indicate the mitochondria. Green: Tubulin-stained using tubulin–Atto 488; yellow arrowheads depict microtubule-based fibrillar bundles. Purple: IF using anti-centrin antibody (mouse). Red: IF using anti-GSBP2 antibody (rabbit). (**B**) Distribution patterns of the subcellular structures.

very near to the cross points in the mesh-like contractile fibrillar structure (Fig. 8 and fig. S11), strongly suggesting the interaction and coordination between these two structures. This finding supports the association among the mesh-like fibrillar bundles, the microtubule-based fibrillar bundles, and the ciliary basal bodies which has been reported previously (*9*, *14*). It is possible that the mesh-like fibrillar bundles (especially the first fibrillar bundles) are anchored to the microtubule-based fibrillar bundles through certain protein(s), and the assembly of mesh-like fibrillar bundles depends on the microtubule-based fibrillar bundles. The regions between two neighboring microtubule-based fibrillar bundles, two neighboring grooves, or two neighboring bundles in the first fibrillar system of the mesh-like structure are highly enriched with mitochondria (Fig. 7, A and B, white arrowhead, and fig. S8 and S9). The distribution patterns of the above subcellular structures strongly suggest that the mesh-like contractile structure couples with other structures to fulfill the contraction and extension cycle (discussed later).

DISCUSSION

Ultrafast motility occurs in many organisms at different scales, such as the shrimp appendage, the trap-jaw ant mandible, jellyfish stingers, and cell contraction in certain ciliates, e.g., *Vorticella* and *Spirostomum* (*5*, *10*, *17*, *32*, *33*). The molecular mechanism underlining these ultrafast motilities is fascinating. *Spirostomum* is a giant single-celled eukaryote, usually millimeter scale in length. The cell of *Spirostomum* could contract and shorten its length up to 75% within several milliseconds. An intriguing question is how this cell is able to contract so fast. The centrin-like spasmin protein was thought to be the key component; however, homologs of spasmins are present in other ciliates that do not have ultrafast contraction. The finding of GSBP proteins in *S. minus* provides the answer. The *S. minus* cell uses the simple strategy of making the GSBP proteins longer. We BLASTP searched against the National Center for Biotechnology Information (NCBI) nonredundant protein database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using GSBP proteins but were unable to find GSBP homologs in any ciliates except in *Stentor* (*34*). GSBP homologs in *Stentor* are about 1000 to 4600 amino acids in length, far shorter than the GSBPs in *S. minus*. The >2-μm GSBP-spasmin complex, as well as its configuration by coiling together, may provide the power needed for ultrafast contraction. The GSBP-spasmin complex is somewhat analogous to the yeast Sfi1-centrin complex (*27*, *28*); however, our data suggest that both GSBP and spasmin have evolved distinctive features that differentiate them from yeast Sfi1 and centrin, respectively (e.g., Cdc31p, lacking the N-terminal disorder region). Although ultrafast motility is found in many organisms, the uniqueness of GSBP in *Spirostomum* indicates the multiple independent origins of GSBP-like backbone proteins in different organisms.

Spirostomum uses a sophisticated system to repeat the ultrafast contraction and extension cycle. The findings in this study provide a blueprint for biomimicry, design, and contraction of repeatable ultrafast contraction micromachines (*35*). Figure 9, which integrated the findings from previous studies (*9*, *14*) and our findings, emphasized the components of mesh-like contractile fibrillar bundles and functional coordination among different subcellular structures. The blueprint includes four core components (Fig. 9A). (i) The mesh-like contractile structure (Fig. 9, A and B) composed of GSBP and spasmins, may form coiled GSBP-spasmin complexes (Fig. 9B). These complexes bind Ca^{2+} and induce conformational changes, thereby providing the power for contraction. (ii) The tubulin-based (or microtubule-based) fibrillar bundles run spirally and parallel to grooves on the cell surface (Figs. 1B and 9A). The fibrillar bundles may act as a cytoskeleton to support the cell shape, and possibly also serve as anchoring sites, e.g., near the

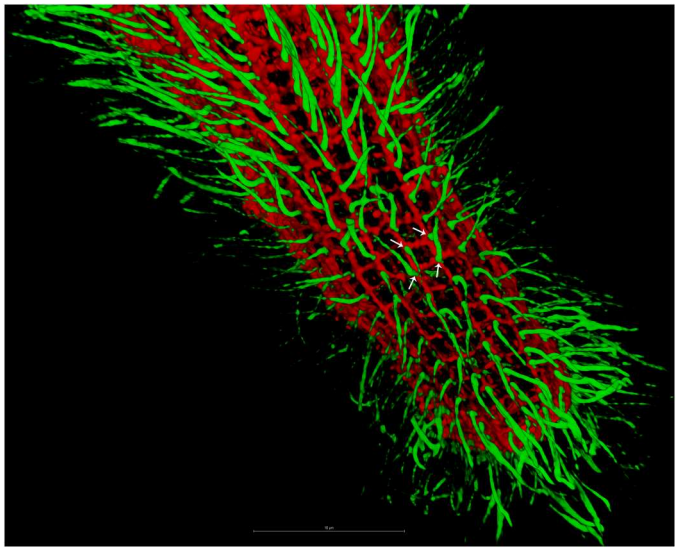

Fig. 8. Interaction between microtube-based and mesh-like fibrillar structures. Three-dimensional image shows the cilia (green) stained using tubulin–Atto 488 as in Fig. 1 and mesh-like contractile structure through IF using anti-GSBP1 antibody (rabbit). White arrows indicate the potential association between ciliary basal bodies and cross-points in mesh-like contractile structure.

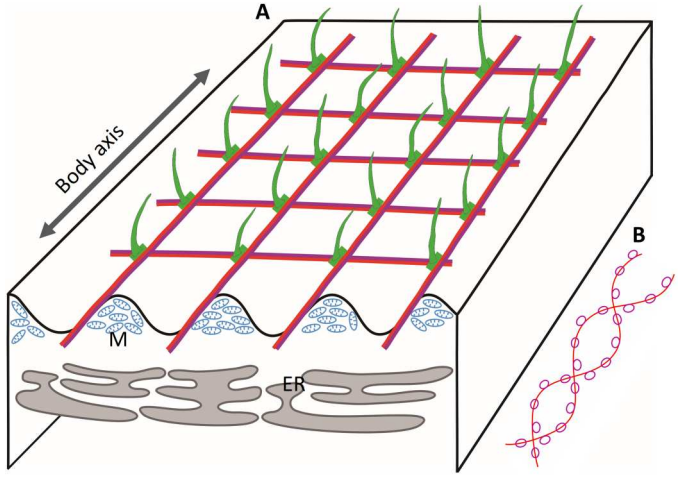

Fig. 9. Coupling of different subcellular structures to fit biological and physical needs for repeat ultrafast contraction and extension of *S. minus* cells. (**A**) Distribution patterns of four core components. Grooves on the cell surface are indicated by the wavy line. Green: Cilia and basal bodies. Fused red and purple lines: Mesh-like contractile fibers comprising the GSBP-spasmin complex. ER, endoplasmic reticulum; M, mitochondria. (**B**) Putative GSBP-spasmin protein complex. Red curve, GSBP protein; purple cycles, spasmins.

ciliary basal bodies (Fig. 8 and fig. S11) for the GSBP-spasmin complex. Upon contraction, the tubulin-based structure compresses like a spring and possibly stores some of the energy provided by contraction to power cell extension. (iii) The endoplasmic reticulum (ER) is located under the mesh-like contractile structure (Fig. 9A) (*9*). ER was recently suggested to be involved in energy dissipation during contraction by serving as "buffer vacuoles," thereby preventing cell damage (*12*). Another important role of the ER may be as Ca^{2+} storage sites (*36*), thus controlling the Ca^{2+} efflux [ryanodine receptor (RyR) (*37*)] and influx [ATP-dependent Ca^{2+} pump (*38*), e.g., sarco/endoplasmic reticulum Ca^{2+} ATPase (SERCA) (*39*)] for cell contraction and extension, respectively. In addition, ciliates usually have the alveolar sacs underneath their cell membranes, which could also be Ca^{2+} storage sites (*40*). However, the presence of contractile "myoneme"-like structures in Acantharia (unicellular eukaryotes within the Radiolaria), which do not have alveolar sacs, suggests that the ER is more likely the site of Ca^{2+} storage associated with the mesh-like contractile structure. Further experimental evidence is required to confirm this possibility. (iv) The mitochondria are located between neighboring grooves under the cell surface (Fig. 9A), providing ATP. Part of the ATP may be used for Ca^{2+} influx of ER through an ATP-dependent Ca^{2+} pump and to recycle Ca^{2+} in the extension process, when it dissociates with the GSBP-spasmin complex.

In general, coupling of these four components fits well with the biological and physical needs for repetitive ultrafast contraction and extension of the *Spirostomum* cell and may facilitate responses to environmental stimuli and communication with other individuals (*13*).

MATERIALS AND METHODS

Cell isolation, culture, morphological analysis, and species identification

S. minus was collected from the East Lake, Wuhan, China (30°32′30″N, 114°22′7″E). Single cells were isolated using a glass micropipette under a stereomicroscope (Zeiss Stemi 2000-C, Germany), and clonal cultures were maintained in the laboratory at 22°C. For *S. minus* cell culture, wheat grains were added to sterilized mineral water, which allowed a density of approximately 50 cells/ml. Identification of *S. minus* was based on observations of cells in vivo and following protargol staining. The identity was confirmed by analysis of the small subunit ribosomal DNA (SSU rDNA) sequence from the genome assembly, which showed 99.8% identity to a published *S. minus* rDNA sequence (HG939545.1) (*41*).

Genome sequencing and assembly

About 200,000 *S. minus* cells were harvested from clonal cultures by centrifuging at 1500*g*. Genomic DNA was prepared by the cetyltrimethylammonium bromide (CTAB) method followed by purification with a QIAGEN Genomic kit (QIAGEN, catalog no. 13343) for regular sequencing according to the standard operating procedure provided by the manufacturer. The quality of DNA was checked using 1% agarose gels. DNA purity was determined using a NanoDrop One UV-Vis spectrophotometer (Thermo Fisher Scientific, USA), i.e., OD260/280 (optical density at 260/280 nm) is between 1.8 and 2.0 and OD260/230 is between 2.0 and 2.2. The DNA concentration was further measured by Qubit 4.0 Fluorometer (Invitrogen, USA).

In most ciliates, the micronucleus is either haploid or diploid, whereas the macronucleus (MAC) is polyploid (*42*). No data about the copy number of *Spirostomum* MAC chromosomes have been reported; however, the closely related *Stentor* has ~60,000 copies of each chromosome in its MAC (*34*). Therefore, we anticipated that the MAC of *Spirostomum* also has a high ploidy level and used this natural enrichment of MAC to perform MAC genome sequencing as had previously been achieved in the oligohymenophorean ciliate *Ichthyophthirius multifiliis* (*43*).

DNA was sequenced using the long reads (Nanopore, https://nanoporetech.com/) and short reads (MGI, https://en.mgi-tech.com/) sequencing platforms. For Nanopore sequencing, library preparation was performed using the 1D Genomic DNA by ligation (SQK-LSK109, Oxford Nanopore) protocol following the manufacturer′s instructions. More than 5-kb fragments were selected to construct the sequencing library using Pippin HT (Sage Science, USA). The library was loaded on the PromethION flow cell and sequenced. For MGI sequencing, library preparation was performed using the DNA sample preparation kit (MGIseq) using the standard protocol recommended by the manufacturer. The library was sequenced on the MGIseq 2000 sequencer with 150–base pair (bp) paired-end.

Nanopore reads were de novo assembled with Nextdenovo (*44*) using the following parameters: read_cutoff = 0.3 k, nextgraph_options = -a 1 -q 10, genome_size = 100 m. The error rate of raw assembly from Nanopore reads is ~0.76%. Therefore, the MGIseq short reads were used to polish the above genome assembly using Nextpolish (*44*), which incorporates short reads mapping using Burrows-Wheeler alignment (BWA) (*45*). The completeness of genome was estimated using BUSCO (*46*).

Transcriptome sequencing

Total RNA was extracted using the Qiagen RNeasy Mini Kit (Qiagen, Valencia, CA, USA), and all the RNA samples had an RNA integrity number above 9.0. VAHTS Universal V6 RNA-seq Library Prep Kit for Illumina (Vazyme, Nanjing, China) was used for RNA library construction following the manufacturer′s protocol. No poly(A) mRNA purification was applied to the RNA samples. The total RNA was directly used for sequence library construction. Libraries were sequenced with the 2 × 150-bp paired-end (PE150) on an Illumina Novaseq 6000 sequencer (Illumina Corporation, San Diego, CA, USA). Adaptors of RNA-seq reads were removed using TrimGalore (https://github.com/FelixKrueger/TrimGalore) and then mapped to the assembled reference genome. Transcripts underwent reference-guided assembly using the TopHat and Cufflinks pipeline (*47*).

Gene annotation and bioinformatics analysis

The codon usage of the *Spirostomum* MAC genome was estimated using Codetta (*48*), and standard codon was found to be used by *Spirostomum.* Gene models were predicted by integrating the ab initio and homology-based methods (*49*). Assembled transcripts were used to train the ab initio gene prediction tools. For the ab initio approach, assembled transcripts were validated by aligning putative transcripts onto the assembled genome using PASA (*50*). The output models from the TransDecoder of PASA were used as a high-quality dataset for training ab initio gene predictors, including Augustus (*51*), GlimmerHMM (*52*), and SNAP (*53*). As Augustus

software could accept cDNA or protein evidence, assembled transcripts were also used as cDNA evidence in this program. For the homology-based method, all protein sequences of genome-sequenced ciliates were used as a reference database, and the assembled *Spirostomum* genome was aligned to this database for gene prediction using GenomeThreader (*54*), ATT (*55*), and Scipio (*56*). Last, an integrated set of gene models was created using Evidence Modeler (*57*) by merging all predicted gene models and the RNA-seq transcripts.

Homologs of predicted *Spirostomum* genes were searched against the NCBI nonredundant protein database using MMseqs2 (*58*). Protein domains and potential functional classifications were annotated using InterProScan, which integrates a series of search applications and databases, including disorder region information, domain information from Pfam, PROSITE, SMART, Gene3D, and KEGG KO lists and Gene Ontology information (*59*).

EF-hand domain–containing proteins were extracted based on the InterProScan annotation using the domain accession ID "IPR011992." In total, 2238 EF-hand proteins were extracted and the multiple sequence alignment was generated using MAFFT (auto model) (*60*). FastTree, which can handle large alignments, ran 100 to 1000 times faster than PhyML or RAxML with comparable accuracy (http://meta.microbesonline.org/fasttree/). Phylogenetic analyses were constructed using the approximate maximum likelihood–based FastTree software (version 2) with the default amino acid substitution model (*61*). iTolTree was used to visualize the phylogenetic tree and protein domain organization assignment (*62*) and IP-MS results.

Motif was searched using PatScanUI (*63*) or custom Perl scripts. Sequence logos were generated using WebLogo (*64*). Tandem repeats in GSBP proteins were analyzed using XSTREAM (*65*). Coiled coil was predicted using Waggawagga (*66*).

For validation of *GSBP* gene predictions, raw Nanopore reads were mapped back to the genome assembly using minimap2. MS peptides matched to GSBP proteins were translated to coordinates in scaffolds of genome assembly. IGV (*67*) was used to visualize different omic data.

Protein structures of spasmins and the GSBP-spasmin complex were modeled using AlphaFold2 or AlphaFold2-multimer (*29*, *68*) with the reduced_dbs option and with a modified database by incorporating predicted proteins from all ciliate genomes and transcriptomes from P10K (*20*).

MS analysis

Cells were harvested and immediately frozen in a −80°C cryogenic freezer. Protein extraction, trypsin digestion, and prefractionation were performed as described previously (*69*, *70*).

The digested peptides were dissolved in 0.1% formic acid (FA) and separated on an online nano-flow EASY-nLC 1200 system with a 75 μm × 15 cm analytical column (C18, 3 μm, Thermo Fisher Scientific) and analyzed on a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific). Peptides were eluted using a gradient of solvent B [0.1% FA in 80% acetonitrile (ACN)] that increased from 4 to 6% in 1 min, 6 to 10% in 2 min, 10 to 18% in 40 min, 18 to 33% in 10 min, climbed to 90% in 0.5 min, and were then held at 90% for 6.5 min, all at a constant flow rate of 300 nl/min. The mass spectrometer was operated in data-dependent acquisition mode with full scans [mass/charge ratio (*m/z*) range of 350 to 1800] at 60,000 mass resolution using an automatic gain control target value of 3×10^6. The top 20 most intense precursor ions were selected for following MS/MS fragmentation by higher-energy collision dissociation with normalized collision energy of 28% and analyzed with 15,000 resolution in the Orbitrap. The dynamic exclusion was set to 25 s, and the isolation width of precursor ions was set to 1.6 *m/z*. The maximum injection times were 20 and 50 ms for both MS and MS/MS, respectively. The intensity threshold was set to 5000.

All MS/MS spectra were analyzed by pFind software (version 3.1) (*71*) against the *S. minus* protein database combined with the reverse decoy database and common contaminants. Two missed cleavages were allowed for trypsin. Open-search algorithm in pFind was used for database searching. Methylation (K) and dimethylation (K) were set as variable modifications. The precursor and fragment ion mass tolerances were both 20 parts per million. The minimum peptide length was set at 6, while the estimated false discovery rate threshold for peptides and proteins was specified at a maximum of 1%.

Immunoprecipitation-MS

Cells were harvested and immediately frozen in a −80°C cryogenic freezer. The centrin, GSBP1, and GSBP2 polyclonal antibodies were incubated with protein A/G magnetic beads (DAIAN, M0134) at 4°C for 2 hours. For negative IP control, 1× phosphate-buffered saline (PBS) buffer was added with the same volume. Then, the beads that had been coated with antibody were washed five times with 1× PBS buffer. During this period, the *S. minus* cells were thawed and lysed in ice-cold lysis buffer (30 mM tris-HCl, 20 mM KCl, 2 mM $MgCl_2$, 1 mM phenylmethylsulfonyl fluoride, and 0.1% Triton X-100) supplemented with 1× cOmplete protease inhibitor (Roche Diagnostics, Indianapolis, IN, USA). The mixture was submitted to sonication for 4 min on ice. The remaining debris was removed by centrifugation at 15,000 rpm at 4°C for 1 hour. The supernatant was collected and incubated with the washed antibody-coated beads at 4°C for 2 hours. Last, the mixture of antibody-coated beads, magnetic beads, and antigens was washed five times with 1× PBST buffer (8 mM Na_2HPO_4, 136 mM NaCl, 2 mM KH_2PO_4, 2.6 mM KCl, and 0.1% Tween 20) and used for MS.

Immunofluorescence

Polyclonal primary antibodies of GSBP1 and GSBP2 were generated in both rabbit and mouse hosts. In detail, hydrophilic regions were selected to generate the antibodies of GSBP1 and GSBP2. An N-terminal fragment (265 to 454) containing 190 residues and a C-terminal fragment (17,623 to 17,823) containing 201 residues were used as antigens for GSBP1 and GSBP2, respectively. The fragment of GSBP1 showed about 28% identity to GSBP2, and the fragment of GSBP2 showed about 38% identity to GSBP1.

To generate the corresponding recombinant proteins of these two antigens, codon-optimized sequences were cloned into the pET32 expression vector and transformed into the *E. coli* BL21 DE3 strain. Recombinant proteins were purified by Ni-NTA (nitrilotriacetic acid) agarose. The protein molecular weight, purity, and concentration were identified by SDS–polyacrylamide gel electrophoresis. For rabbit antibody, 2 mg of purified protein was mixed with Freund's complete adjuvant (FCA) or Freund's incomplete adjuvant (FIA), followed by injection into Japanese White rabbit five times. For mouse antibody, 2 mg of purified protein was mixed with FCA or FIA, followed by injection into Balb/C mice four times.

Antiserum titer was tested by enzyme-linked immunosorbent assay. An affinity column was used for antibody purification and was made by coupling 1 mg of antigen to CNBr-activated Sepharose 4B (GE). The antibody was purified by incubating 10 ml of antiserum in the column for 12 hours and then eluted with the glycine buffer (0.15 M, pH 2.5). Commercial primary antibodies for β-tubulin (rat, Thermo Fisher Scientific Invitrogen, catalog no. MA180017) and centrin (mouse, Millipore-Sigma, catalog no. 04-1624) were used for IF.

Cells were transferred from cultures into a plate and washed twice with ddH_2O. The cells were permeabilized in 200 μl of 0.5× PHEM (60 mM Pipes, 25 mM HEPES, 10 mM EGTA, 2 mM $MgCl_2$) buffer at 4°C, incubated for 1 hour, and fixed with 200 μl of 1× PHEM–2% paraformaldehyde (PFA) for 15 min. Fixed cells were then treated with 300 μl of 2% PFA–1× PHEM for 1 hour and washed three times with 300 μl of 1× PBS for 10 min each.

IF experiments were done through combinations of primary antibodies from different hosts with corresponding secondary antibodies. To determine the localization of GSBP and spasmin, two primary antibodies were used, i.e., the GSBP polyclonal antibody (rabbit) and centrin antibody (mouse, Millipore-Sigma, catalog no. 04-1624). Cells were incubated in 200 μl of primary antibody solution, which was diluted (1:100) in 1× PBST (1× PBS containing 0.3% Triton X-100)–3% bovine serum albumin at 4°C for 12 hours, and washed three times with 300 μl of 1× PBS for 10 min each. Alexa Fluor 594–conjugated goat anti-rabbit immunoglobulin G (IgG) (Proteintech Group, catalog no. SA00006-4) and Alexa Fluor 488–conjugated goat anti-mouse IgG (Abcam, catalog no. ab150113) were used as the secondary antibodies. Cells were incubated in 200 μl of working solution in 1× PBST (1× PBS containing 0.05% Triton X-100) for 2 hours at room temperature (25°C) in a dark wet box and washed three times with 300 μl of 1× PBS for 10 min each. After washing, cells were stained with anti-fading agent containing DAPI (0.05 mg/ml).

To determine the localization of GSBP1 and GSBP2, primary antibodies of GSBP1 (mouse) and GSBP2 (rabbit) and secondary antibodies of Alexa Fluor 488–conjugated goat anti-mouse IgG and Alexa Fluor 594–conjugated goat anti-rabbit IgG, respectively, were used. The experimental conditions were the same as those for GSBP and spasmin described above.

To determine the localization of tubulin, spasmin, and GSBP, primary antibodies for β-tubulin (rat), centrin (mouse), and GSBP (rabbit) were used. The secondary antibodies used were DyLight 488–conjugated goat anti-rat IgG (Abcam, catalog no. ab98420), Alexa Fluor 680–conjugated goat anti-mouse IgG (Thermo Fisher Scientific Invitrogen, catalog no. A21057), and Alexa Fluor 594–conjugated goat anti-rabbit IgG, respectively. The experimental conditions were the same as for GSBP and spasmin described above. Fluorescence staining was observed at ×100 magnification, and images were recorded with a Leica TCS SP8 STED laser scanning confocal microscope (Leica Microsystems, Mannheim, Germany) and Leica LAS AF lite software (https://webshare.leica-microsystems.com/latest/core/widefield/).

Western blot

To prepare crude protein samples for Western blot, *S. minus* cells were isolated using glass micropipette and frozen in −80°C. The cells were fixed with 10% (w/v) trichloroacetic acid and incubated on ice for 10 min. Cell pellets were then lysed in SDS sample buffer and incubated at 100°C for 10 min.

Because of the high molecular weights, GSBPs should be hard to separate on the common polyacrylamide gel electrophoresis, and thus, three different gels were tested in this study, including SeaKem Gold agarose gel for high–molecular weight proteins (*72*), NuPAGE bis-tris precast polyacrylamide gel (4 to 12%, Thermo Fisher Scientific Invitrogen, catalog no. NP0321BOX), and 6% separation polyacrylamide gel. For electrophoresis conditions, two strategies were used to check the antibody specificity: (i) Electrophoresis was started with an initial voltage of 70 V for maintaining 2 hours, and increased to 130 V for 3.5 hours, for high–molecular weight proteins. Proteins were then transferred to a polyvinylidene difluoride (PVDF) membrane by electroblotting using 220 mA for 5 hours. (ii) Electrophoresis was started with a voltage of 70 V for low–molecular weight proteins and then transferred to a PVDF membrane by electroblotting using 220 mA for 2.5 hours. The results suggested that the 6% separation polyacrylamide gel worked for GSBPs.

The membranes were blocked with PBST buffer (8 mM Na_2HPO_4, 136 mM NaCl, 2 mM KH_2PO_4, 2.6 mM KCl, and 0.1% Tween 20) containing 5% skim milk powder for 1 hour at 25°C and then incubated with the primary polyclonal antibody (1:1000 dilution) overnight at 4°C. After washing three times with PBST buffer, the membranes were incubated with horseradish peroxidase–conjugated goat anti-rabbit IgG secondary antibody (1:5000 dilution) for 1 hour at 25°C. Last, the signal was visualized using a fluorescence scanner (Alpha Innotech FluorChem Q, USA) with WesternBright Sirius kits (Advansta, Menlo Park, CA, USA).

RNA interference

Target gene sequences were obtained by polymerase chain reaction (PCR) amplification from *S. minus* genomic DNA. PCR fragments were digested with Kpn I and Xho I and then ligated into the L4440 vector, which had been linearized with the same restriction enzymes. The resulting plasmids were transfected into *E. coli* HT115 cells by the heat shock method. The transformed HT115 cells were then cultured in liquid LB medium containing tetracycline and ampicillin at 37°C in a shaking incubator for 1 hour. Cultures were then spread onto LB agar plates containing the two aforementioned antibiotics. Plates were incubated at 37°C for approximately 12 hours until colonies appeared. Single colonies were inoculated into liquid LB medium containing the two antibiotics. When transformed HT115 cells were grown to early logarithmic phase (OD_{600} = ~0.4), they were induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 4 hours to produce double-stranded RNA (dsRNA). The bacterial cells were centrifuged to remove the media, and the total RNA was extracted using the TRIzol method. Reverse transcription PCR for *GSBP1* and *GSBP2* used the following primers: GSBP1-F, TACCTCGCATCCTCAAGGAC; GSBP1-R, GTACCTTGCCCTTCTTCACG; GSBP2-F, AGAAACTTGCTCAGAAGCCTAAG; GSBP2-R, GCGAGTTTCTTGACAAGAGACC. PCR products were analyzed by 1% agarose gel electrophoresis.

S. minus cells were washed and starved for 12 hours before the experiment began, with 50 or 200 cells per culture dish (Nest Biotechnology Co. Ltd., catalog no. 706001) containing 3 ml of sterilized mineral water. The HT115 strain of *E. coli* was used for the feeding vector, transformed with a plasmid containing the target

genes flanked by two opposing T7 sites for dsRNA production. HT115 cells were grown overnight at 37°C in a shaking incubator under the selection of both tetracycline and ampicillin. Cultures were grown to logarithmic phase (OD_{600}: 0.4 to 0.5) and then induced to produce dsRNA by the addition of 0.5 mM IPTG and incubated for a further 4 hours. Subsequently, bacterial cells were harvested by centrifugation and washed in sterile water, and the cell suspension was adjusted to OD_{600} = 2.2. These cells were then centrifuged at 20,000*g* for 3 min to remove the media. The resulting dry pellet was flash-frozen in liquid nitrogen and stored at −80°C for up to 1 week. Aliquots containing 200 μl of the bacterial culture were resuspended in mineral water daily and added to each culture dish containing *S. minus* cells. To avoid buildup of detritus, the *S. minus* cells were transferred into a new dish containing mineral water every other day. Control cells were fed with either untransformed HT115 cells or HT115 cells transformed with empty L4440 plasmids. All of the experiments were approved by the Animal Care and Use Committee of the Institute of Hydrobiology, Chinese Academy of Sciences.

Supplementary Materials

This PDF file includes:
Figs. S1 to S11
Tables S1 and S2
Legends for movies S1 and S2

Other Supplementary Material for this manuscript includes the following:
Movies S1 and S2

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. D. D. Hackney, The kinetic cycles of myosin, kinesin, and dynein. *Annu. Rev. Physiol.* **58**, 731–750 (1996).
2. S. S. Brown, Cooperation between microtubule-and actin-based motor proteins. *Annu. Rev. Cell Dev. Biol.* **15**, 63–80 (1999).
3. R. S. Adelstein, E. Eisenberg, Regulation and kinetics of the actin-myosin-ATP interaction. *Annu. Rev. Biochem.* **49**, 921–956 (1980).
4. W. B. Amos, Reversible mechanochemical cycle in the contraction of *Vorticella*. *Nature* **229**, 127–128 (1971).
5. H. E. Buhse, S. M. McCutcheon, J. C. Clamp, P. Sun, Vorticella. *eLS*, John Wiley and Sons, Ltd (2011).
6. H. Ishida, Y. Shigenaka, Cell model contraction in the ciliate *Spirostomum*. *Cell Motil. Cytoskeleton* **9**, 278–282 (1988).
7. L. X. Xu, M. S. Bhamla, Biophysical mechanism of ultrafast helical twisting contraction in the giant unicellular ciliate *Spirostomum ambiguum*. bioRxiv 854836 [Preprint]. 25 November 2019. https://doi.org/10.1101/854836.
8. T. C. Hamilton, D. Osborn, Measurements of contraction latencies to mechanical and electrical stimulation of the protozoan, *Spirostomum ambiguum*. *J. Cell Physiol.* **91**, 403–408 (1977).
9. R. Yagiu, Y. Shigenaka, Electron microscopy of the longitudinal fibrillar bundle and the contractile fibrillar system in *Spirostomum ambiguum*. *J. Protozool.* **10**, 364–369 (1963).
10. W. J. Lehman, L. I. Rebhun, The structural elements responsible for contraction in the ciliate *Spirostomum*. *Protoplasma* **72**, 153–178 (1971).
11. D. Osborn, J. C. Hsung, E. M. Eisenstein, The involvement of calcium in contractility in the ciliated protozoan, *Spirostomum ambiguum*. *Behav. Biol.* **8**, 665–677 (1973).
12. R. Chang, M. Prakash, Entangled architecture of rough endoplasmic reticulum (RER) and vacuoles enables topological damping in cytoplasm of an ultra-fast giant cell. bioRxiv 2021.12.13.472465 [Preprint]. 13 December 2021. https://doi.org/10.1101/2021.12.13.472465.
13. A. J. T. M. Mathijssen, J. Culver, M. S. Bhamla, M. Prakash, Collective intercellular communication through ultra-fast hydrodynamic trigger waves. *Nature* **571**, 560–564 (2019).
14. H. E. Finley, C. A. Brown, W. A. DANIEL, Electron microscopy of the ectoplasm and infraciliature of *Spirostomum ambiguum*. *J. Protozool.* **11**, 264–280 (1964).
15. J. J. Maciejewski, E. J. Vacchiano, S. M. Mccutcheon, H. E. Buhse Jr., Cloning and expression of a cDNA encoding a *Vorticella convallaria* spasmin: An EF-hand calcium-binding protein. *J. Eukaryot. Microbiol.* **46**, 165–173 (1999).
16. C. Q. Jiang, G. Y. Wang, J. Xiong, W. T. Yang, Z. Y. Sun, J. M. Feng, A. Warren, W. Miao, Insights into the origin and evolution of Peritrichia (Oligohymenophorea, Ciliophora) based on analyses of morphology and phylogenomics. *Mol. Phylogenet. Evol.* **132**, 25–35 (2019).
17. L. Mahadevan, P. Matsudaira, Motility powered by supramolecular springs and ratchets. *Science* **288**, 95–100 (2000).
18. H. Asai, T. Ninomiya, R. I. Kono, Y. Moriyama, Spasmin and a putative spasmin binding protein(s) isolated from solubilized spasmonemes. *J. Eukaryot. Microbiol.* **45**, 33–39 (1998).
19. J. V. Kilmartin, Sfi1p has conserved centrin-binding sites and an essential function in budding yeast spindle pole body duplication. *J. Cell Biol.* **162**, 1211–1221 (2003).
20. W. Miao, L. Song, S. Ba, L. Zhang, G. Guan, Z. Zhang, K. Ning, Protist 10,000 genomes project. *Innovation* **1**, 100058 (2020).
21. J. Elíes, M. Yáñez, T. M. C. Pereira, J. Gil-Longo, D. A. MacDougall, M. Campos-Toimil, An update to calcium binding proteins. *Adv. Exp. Med. Biol.* **1131**, 183–213 (2020).
22. J. L. Satisbury, Centrin, centrosomes, and mitotic spindle poles. *Curr. Opin. Cell Biol.* **7**, 39–45 (1995).
23. M. Maki, Y. Kitaura, H. Satoh, S. Ohkouchi, H. Shibata, Structures, functions and molecular evolution of the penta-EF-hand Ca^{2+}-binding proteins. *Biochim. Biophys. Acta.* **1600**, 51–60 (2002).
24. A. Lewit-Bentley, S. Réty, EF-hand calcium-binding proteins. *Curr. Opin. Struct. Biol.* **10**, 637–643 (2000).
25. T. Itabashi, K. Mikami, H. Asai, Characterization of the spasmin 1 gene in *Zoothamnium arbuscula* strain Kawagoe (protozoa, ciliophora) and its relation to other spasmins and centrins. *Res. Microbiol.* **154**, 361–367 (2003).
26. D. Gogendeau, C. Klotz, O. Arnaiz, A. Malinowska, M. Dadlez, N. G. de Loubresse, F. Ruiz, F. Koll, J. Beisson, Functional diversification of centrins and cell morphological complexity. *J. Cell Sci.* **121**, 65–74 (2008).
27. J. L. Salisbury, Centrosomes: Sfi1p and centrin unravel a structural riddle. *Curr. Biol.* **14**, R27–R29 (2004).
28. D. Gogendeau, J. Beisson, N. G. de Loubresse, J. P. Le Caer, F. Ruiz, J. Cohen, L. Sperling, F. Koll, C. Klotz, An Sfi1p-like centrin-binding protein mediates centrin-based Ca^{2+}-dependent contractility in *Paramecium tetraurelia*. *Eukaryot. Cell* **6**, 1992–2000 (2007).
29. R. Evans, M. O'Neill, A. Pritzel, N. Antropova, A. W. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim, O. Ronneberger, S. Bodenstein, M. Zielinski, A. Bridgland, A. Potapenko, A. Cowie, K. Tunyasuvunakool, R. Jain, E. Clancy, P. Kohli, J. Jumper, D. Hassabis, Protein complex prediction with AlphaFold-Multimer. bioRxiv 2021.10.04.463034 [Preprint]. 10 March 2022. https://doi.org/10.1101/2021.10.04.463034.
30. M. M. Slabodnick, J. G. Ruby, J. G. Dunn, J. L. Feldman, J. L. DeRisi, W. F. Marshall, The kinase regulator mob1 acts as a patterning protein for *Stentor* morphogenesis. *PLOS Biol.* **12**, e1001861 (2014).
31. W. Wei, C. Jiang, X. Chai, J. Zhang, C. C. Zhang, W. Miao, J. Xiong, RNA interference by cyanobacterial feeding demonstrates the SCSG1 gene is essential for ciliogenesis during oral apparatus regeneration in *Stentor*. *Microorganisms* **9**, 176 (2021).
32. S. N. Patek, W. L. Korff, R. L. Caldwell, Deadly strike mechanism of a mantis shrimp. *Nature* **428**, 819–820 (2004).
33. S. N. Patek, J. E. Baio, B. L. Fisher, A. V. Suarez, Multifunctionality and mechanical origins: Ballistic jaw propulsion in trap-jaw ants. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12787–12792 (2006).
34. M. M. Slabodnick, J. G. Ruby, S. B. Reiff, E. C. Swart, S. Gosai, S. Prabakaran, E. Witkowska, G. E. Larue, S. Fisher, R. M. FreemanJr., J. Gunawardena, W. Chu, N. A. Stover, B. D. Gregory, M. Nowacki, J. Derisi, S. W. Roy, W. F. Marshall, P. Sood, The macronuclear genome of *Stentor coeruleus* reveals tiny introns in a giant cell. *Curr. Biol.* **27**, 569–575 (2017).
35. S. Ryu, R. E. Pepper, M. Nagai, D. C. France, *Vorticella*: A protozoan for bio-inspired engineering. *Micromachines* **8**, 4 (2017).
36. R. D. Allen, Structures linking the myonemes, endoplasmic reticulum, and surface membranes in the contractile ciliate *Vorticella*. *J. Cell Biol.* **56**, 559–579 (1973).
37. R. Zalk, O. B. Clarke, A. des Georges, R. A. Grassucci, S. Reiken, F. Mancia, W. A. Hendrickson, J. Frank, A. R. Marks, Structure of a mammalian ryanodine receptor. *Nature* **517**, 44–49 (2015).
38. P. Caroni, E. Carafoli, An ATP-dependent Ca^{2+}–pumping system in dog heart sarcolemma. *Nature* **283**, 765–767 (1980).
39. C. M. Misquitta, D. P. Mack, A. K. Grover, Sarco/endoplasmic reticulum Ca^{2+} (SERCA)-pumps: Link to heart beats and calcium waves. *Cell Calcium* **25**, 277–290 (1999).

40. H. Plattner, N. Klauke, Calcium in ciliated protozoa: Sources, regulation, and calcium-regulated cell functions. *Int. Rev. Cytol.* **201**, 115–208 (2001).
41. V. Boscaro, D. Carducci, G. Barbieri, M. V. X. Senra, I. Andreoli, F. Erra, G. Petroni, F. Verni, S. I. Fokin, Focusing on genera to improve species identification: Revised systematics of the ciliate *Spirostomum*. *Protist* **165**, 527–541 (2014).
42. J. O. Corliss, *The Ciliated Protozoa: Characterization, Classification and Guide to the Literature* (Elsevier, 2016).
43. R. S. Coyne, L. Hannick, D. Shanmugam, J. B. Hostetler, D. Brami, V. S. Joardar, J. Johnson, D. Radune, I. Singh, J. H. Badger, U. Kumar, M. Saier, Y. Wang, H. Cai, J. Gu, M. W. Mather, A. B. Vaidya, D. E. Wilkes, V. Rajagopalan, D. J. Asai, C. G. Pearson, R. C. Findly, H. W. Dickerson, M. Wu, C. Martens, Y. Van de Peer, D. S. Roos, D. M. Cassidy-Hanley, T. G. Clark, Comparative genomics of the pathogenic ciliate *Ichthyophthirius multifiliis*, its free-living relatives and a host species provide insights into adoption of a parasitic lifestyle and prospects for disease control. *Genome Biol.* **12**, R100 (2011).
44. J. Hu, Y. Han, Z. Wang, Z. Sun, S. Liu, D. Wang, in *Plant and Animal Genome XXVIII Conference*. (PAG, 2020).
45. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv: 1303.3997 [q-bio.GN] (16 March 2013).
46. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
47. C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, L. Pachter, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
48. Y. Shulgina, S. R. Eddy, A computational screen for alternative genetic codes in over 250,000 genomes. *eLife* **10**, e71402 (2021).
49. J. Xiong, W. T. Yang, K. Chen, C. Q. Jiang, Y. Ma, X. C. Chai, G. X. Yan, G. Y. Wang, D. X. Yuan, Y. F. Liu, S. L. Bidwell, N. Zafar, M. Hadjithomas, V. Krishnakumar, R. S. Coyne, E. Orias, W. Miao, Hidden genomic evolution in a morphospecies—The landscape of rapidly evolving genes in *Tetrahymena*. *PLOS Biol.* **17**, e3000294 (2019).
50. B. J. Haas, A. L. Delcher, S. M. Mount, J. R. Wortman, R. K. Smith Jr., L. I. Hannick, R. Maiti, C. M. Ronning, D. B. Rusch, C. D. Town, S. L. Salzberg, O. White, Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654–5666 (2003).
51. M. Stanke, M. Diekhans, R. Baertsch, D. Haussler, Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**, 637–644 (2008).
52. W. H. Majoros, M. Pertea, S. L. Salzberg, TigrScan and GlimmerHMM: Two open source ab initio eukaryotic gene-finders. *Bioinformatics* **20**, 2878–2879 (2004).
53. I. Korf, Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
54. G. Gremme, V. Brendel, M. E. Sparks, S. Kurtz, Engineering a software tool for gene structure prediction in higher organisms. *Inf. Softw. Technol.* **47**, 965–978 (2005).
55. X. Huang, M. D. Adams, H. Zhou, A. R. Kerlavage, A tool for analyzing and annotating genomic sequences. *Genomics* **46**, 37–45 (1997).
56. O. Keller, F. Odronitz, M. Stanke, M. Kollmar, S. Waack, Scipio: Using protein sequences to determine the precise exon/intron structures of genes and their orthologs in closely related species. *BMC Bioinformatics* **9**, 278 (2008).
57. B. J. Haas, S. L. Salzberg, W. Zhu, M. Pertea, J. E. Allen, J. Orvis, O. White, C. R. Buell, J. R. Wortman, Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. *Genome Biol.* **9**, R7 (2008).
58. M. Steinegger, J. Söding, MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).
59. P. Jones, D. Binns, H. Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S. Y. Yong, R. Lopez, S. Hunter, InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
60. K. Katoh, K. Misawa, K.-i. Kuma, T. Miyata, MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
61. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2–approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**, e9490 (2010).
62. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
63. M. Dsouza, N. Larsen, R. Overbeek, Searching for patterns in genomic data. *Trends Genet.* **13**, 497–498 (1997).
64. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
65. A. M. Newman, J. B. Cooper, XSTREAM: A practical algorithm for identification and architecture modeling of tandem repeats in protein sequences. *BMC Bioinformatics* **8**, 382 (2007).
66. D. Simm, K. Hatje, M. Kollmar, Waggawagga: Comparative visualization of coiled-coil predictions and detection of stable single α-helices (SAH domains). *Bioinformatics* **31**, 767–769 (2015).
67. H. Thorvaldsdóttir, J. T. Robinson, J. P. Mesirov, Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
68. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
69. S. C. Yu, M. K. Yang, J. Xiong, Q. Zhang, X. X. Gao, W. Miao, F. Ge, Proteogenomic analysis provides novel insight into genome annotation and nitrogen metabolism in *Nostoc* sp. PCC 7120. *Microbiol. Spectr.* **9**, e0049021 (2021).
70. M. K. Yang, X. H. Lin, X. Liu, J. Zhang, F. Ge, Genome annotation of a model diatom *Phaeodactylum tricornutum* using an integrated proteogenomic pipeline. *Mol. Plant.* **11**, 1292–1307 (2018).
71. H. Chi, K. He, B. Yang, Z. Chen, R. X. Sun, S. B. Fan, K. Zhang, C. Liu, Z. F. Yuan, Q. H. Wang, S. Q. Liu, M. Q. Dong, S. M. He, pFind–Alioth: A novel unrestricted database search algorithm to improve the interpretation of high-resolution MS/MS data. *J. Proteomics* **125**, 89–97 (2015).
72. M. L. Greaser, C. M. Warren, Electrophoretic separation of very large molecular weight proteins in SDS agarose, in *Electrophoretic Separation of Proteins* (Springer, 2019), pp. 203–210.
73. M. Zhang, M. Li, W. Zhang, Y. Han, Y. H. Zhang, Simple and efficient delivery of cell-impermeable organic fluorescent probes into live cells for live-cell superresolution imaging. *Light Sci. Appl.* **8**, 73 (2019).

Acknowledgments: We thank H. Long and X. Liu (Institute of Hydrobiology, Chinese Academy of Sciences) for help and suggestions about the Western blot and IF experiments; Z. X. Qiao, M. Wang, and S. Z. Jia (Analysis and Testing Center of Institute of Hydrobiology, Chinese Academy of Sciences) for help with genome and transcriptome sequencing and mass spectrometry analysis; J. Hu (GrandOmics Biosciences) for help in parameter tuning of NextDenovo; and Y. F. Wei (Huazhong University of Science and Technology) for tubulin staining using tubulin–Atto 488. We also thank the members of Protist 10,000 Genomes Project (P10K) consortium for helpful suggestions. The bioinformatics analysis was supported by the Wuhan Branch, Supercomputing Center, Chinese Academy of Sciences, China. The culture and maintenance of *S. minus* cells were supported by the National Aquatic Biological Resource Center (NABRC). **Funding:** This work was supported by the National Key R&D Program of China (2020YFA0907400), the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDPB18), and the Natural Science Foundation of China (32122015, 31872221, and 31900339). **Author contributions:** W.M. and J.X. designed the study. S.G., W.Q., J.Z., X.W., and X.C. collected samples and prepared DNA and RNA. J.X., J.Z., W.Q., C.H., S.G., M.Y., F.Z., K.C., G.Y., G.W., and C.J. performed the analysis and experiments, and J.X., W.M., J.Z., A.W., W.Q., S.G., and X.C. wrote the manuscript. All authors read and approved the final manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The raw sequencing reads, assembled genome, and gene predictions can be accessed through National Genomics Data Center (PRJCA009852, https://ngdc.cncb.ac.cn/).

Submitted 25 June 2022
Accepted 20 January 2023
Published 22 February 2023
10.1126/sciadv.add6550