***eLife’s* transparent reporting form**

We encourage authors to provide detailed information *within their submission* to facilitate the interpretation and replication of experiments. Authors can upload supporting documentation to indicate the use of appropriate reporting guidelines for health-related research (see [EQUATOR Network](http://www.equator-network.org/%20)), life science research (see the [BioSharing Information Resource](https://biosharing.org/)), or the [ARRIVE guidelines](http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.1000412) for reporting work involving animal research. Where applicable, authors should refer to any relevant reporting standards documents in this form.

If you have any questions, please consult our Journal Policies and/or contact us: [editorial@elifesciences.org](mailto:editorial@elifesciences.org).

**Sample-size estimation**

* You should state whether an appropriate sample size was computed when the study was being designed
* You should state the statistical method of sample size computation and any required assumptions
* If no explicit power analysis was used, you should describe how you decided what sample (replicate) size (number) to use

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn’t apply to your submission:

Information on the number of representative SecA proteins used to produce the phylogenetic tree in figure 1A can be found in the results section describing these results. Information on the number of independent modelling runs and the parameters used to produce the results depicted in figure 5B-F can be found in the methods section. Otherwise, this information is not applicable to the experiments described in this manuscript since they are primarily biochemical in nature.

**Replicates**

* You should report how often each experiment was performed
* You should include a definition of biological versus technical replication
* The data obtained should be provided and sufficient information should be provided to indicate the number of independent biological and/or technical replicates
* If you encountered any outliers, you should describe how these were handled
* Criteria for exclusion/inclusion of data should be clearly stated
* High-throughput sequence data should be uploaded before submission, with a private link for reviewers provided (these are available from both GEO and ArrayExpress)

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn’t apply to your submission:

For quantitative data (table 1, figure 2C, figure 5B-F, figure 2—figure supplements 1 and 2), experiments were reproduced at least three times independently (biological replicates), with the exception of the data depicted in figure 2A, which was reproduced twice. For qualitative data (figure 1D-E, figure 2B&D, figure 4, figure 5, figure 1—figure supplements 2 and 3, figure 2—figure supplement 3, figure 3—figure supplements 1 and 2, figure 4—figure supplement 1), experiments were reproduced at least twice. The x-ray crystal structure in figure 4A was solved with a single dataset. The SAXS data represent the results of a single SAXS run. However, a second run on a different SAXS line using full-length SecA was nearly indistinguishable from the results presented indicating that the results are robust. Where appropriate, this information is reported in the figure legends or in the appropriate methods section.

**Statistical reporting**

* Statistical analysis methods should be described and justified
* Raw data should be presented in figures whenever informative to do so (typically when N per group is less than 10)
* For each experiment, you should identify the statistical tests used, exact values of N, definitions of center, methods of multiple test correction, and dispersion and precision measures (e.g., mean, median, SD, SEM, confidence intervals; and, for the major substantive results, a measure of effect size (e.g., Pearson's r, Cohen's d)
* Report exact p-values wherever possible alongside the summary statistics and 95% confidence intervals. These should be reported for all key questions and not only when the p-value is less than 0.05.

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn’t apply to your submission:

Unless otherwise indicated, confidence intervals represent one standard deviation as described in the figure legends. The p-values presented in the results subsection describing figure 5F were determined used a chi-squared test as described in the figure legend.

(For large datasets, or papers with a very large number of statistical tests, you may upload a single table file with tests, Ns, etc., with reference to sections in the manuscript.)

**Group allocation**

* Indicate how samples were allocated into experimental groups (in the case of clinical studies, please specify allocation to treatment method); if randomization was used, please also state if restricted randomization was applied
* Indicate if masking was used during group allocation, data collection and/or data analysis

Please outline where this information can be found within the submission (e.g., sections or figure legends), or explain why this information doesn’t apply to your submission:

The work does not use this type of data. Thus, this information is not applicable to our submission.

**Additional data files (“source data”)**

* We encourage you to upload relevant additional data files, such as numerical data that are represented as a graph in a figure, or as a summary table
* Where provided, these should be in the most useful format, and they can be uploaded as “Source data” files linked to a main figure or table
* Include model definition files including the full list of parameters used
* Include code used for data analysis (e.g., R, MatLab)
* Avoid stating that data files are “available upon request”

Please indicate the figures or tables for which source data files have been provided:

The sequence alignment and tree data used to generate the phylogenetic tree depicted in figure 1B and the consensus depicted in figure 1C can be found in supplementary files 1 and 2. A summary of the peptides detected by MALDI-TOF are listed in supplementary file 3. X-ray crystal structure parameters are given in supplementary file 4. The x-ray crystallography data can be found in PDB with the accession code 6GOX. Parameters of the small-angle x-ray scattering data are given in supplementary file 5 and the SAXS data can be found in SASBDB under accession codes SASDDY9, SASDDZ9 and SASDE22. The results of fitting the SAXS data and the fitting parameters are given in supplementary file 6.