ELECTRONIC SUPPLEMENTARY INFORMATION

Table of Contents

1. Experimental procedures

Production and purification of sermetstatin – A synthetic gene coding for sermetstatin from *Streptomyces caespitosus* (UniProt database code Q9FDS0), also known as *Streptomyces caespitosus* neutral proteinase inhibitor [1], was purchased (GenScript) and cloned into a modified pET-32a vector between the *Bgl*II and *Hind*III restriction sites. This vector attaches an N-terminal thioredoxin-His_6 fusion construct followed by a tobacco-etch-virus (TEV) protease recognition site. Sermetstatin was produced by heterologous overexpression in *Escherichia coli* Origami2 (DE3) cells (Novagen). These were grown at 37°C in Luria-Bertani (LB) medium containing 100µg/ml ampicillin and 10µg/ml tetracycline, induced at an OD_{550} of 0.6 with isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.25mM, and subsequently incubated overnight at 18°C. Cultures were centrifuged at 7,000x*g* for 30min at 4°C. Pellets were washed twice with buffer A (50mM Tris-HCl, 500mM NaCl, pH8.0) and resuspended in the same buffer further containing 20mM imidazole and supplemented with EDTA-free protease inhibitor cocktail tablets (Roche Diagnostics) and DNase I (Roche Diagnostics). Cells were lysed at 4ºC using a cell disrupter (Constant Systems) at a pressure of 1.35Kbar, and the cell debris was removed by centrifugation at 50,000x*g* for 1h at 4°C. The

supernatant was filtered (0.22µm pore size; Millipore) and incubated with nickel-nitrilotriacetic acid (Ni-NTA) resin (Invitrogen) previously equilibrated with buffer A, 20mM imidazole. The protein was eluted using buffer A, 350mM imidazole. Subsequently, the sample was dialyzed overnight at room temperature against buffer B (50mM Tris-HCl, 150mM NaCl, 0.5mM oxidized glutathione [GSSG], 3mM reduced glutathione [GSH], pH8.0) in the presence of His_6-tagged TEV protease at an enzyme:substrate ratio of 1:50 (w/w). TEV cleavage leaves an extra glycine residue at the N-terminus of the protein (termed G^{-1}; superscripted sermetstatin amino acid numbering hereafter corresponds to the mature protein without the 28-residue signal peptide [see UniProt Q9FDS0]). The digested sample was passed several times through Ni-NTA resin previously equilibrated with buffer A to remove His_6-containing molecules. The flow-through was collected, concentrated by ultrafiltration, and further purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column (GE Healthcare) previously equilibrated with buffer C (20mM Tris-HCl, 150mM NaCl, pH7.4). Previous column calibration revealed that the protein eluted as a dimer, and its identity and purity were assessed by mass spectrometry and 15% Tricine-SDS-PAGE stained with Coomassie blue. Ultrafiltration steps were performed with Vivaspin 15 and Vivaspin 500 filter devices with a 5-KDa cut-off (Sartorius Stedim Biotech). Protein concentration was determined by measuring the absorbance at 280 nm using a spectrophotometer (NanoDrop) and a calculated absorption coefficient $E_{0.1\%} = 0.86$.

The present production system based on a fusion with thioredoxin and expression in Origami2 cells, which is known to assist multi-disulfide proteins in folding correctly during intracellular biosynthesis, and further TEV-mediated removal of the fusion construct in selected concentrations of redox agents, yielded approx. 10mg of purified natively-folded protein per liter of cell culture. The selenomethionine variant of the protein was obtained in the same way as the wild-type except that cells were grown in minimal medium containing selenomethionine (Sigma) instead of methionine. A series of sermetstatin single- and double-point mutants, namely, H^3E, H^3R, $Y^{33}P+T^{34}G$, $M^{71}K$, $Y^{72}V+F^{73}Y$, $C^{31}S+C^{46}S$, and $C^{69}S+C^{99}S$, was generated using the QuickChange Site-Directed Mutagenesis kit (Stratagene). All constructs were verified by DNA sequencing and mutant variants were produced and purified as for the wild-type protein, with comparable yields except for $C^{31}S+C^{46}S$, which was obtained with an approx. 100-fold lower yield.

Production and purification of snapalysin – A synthetic gene coding for mature snapalysin from *S. caespitosus*, in which the N-terminal threonine had been replaced by methionine due to the cloning strategy, was purchased (GenScript) and cloned into a modified pET-28a vector between the *Nco*I and *Xho*I restriction sites. This vector attaches an N-terminal His_6 fusion tag, followed by a TEV protease recognition site. The protein was overproduced as inclusion bodies in *E. coli* BL21 (DE3) cells, which were grown at 37°C in LB medium supplemented with 30µg/ml kanamycin, induced at an OD_{550} of 0.8 with 1mM IPTG, and grown for a further 5h at 37°C. After centrifugation at 7,000x*g* for 30min at 4°C, the pellet was resuspended in buffer D (PBS, 1% Triton X-100, pH7.4) containing EDTA-free protease inhibitor cocktail tablets and DNase I. Cells were lysed at 4ºC using a cell

disrupter at a pressure of 1.9Kbar and incubated for 30min at 37°C. The inclusion bodies were recovered by centrifugation at 7,000x*g* for 30min at 4°C and washed twice with buffer D, resuspended in buffer A, 8M urea, and incubated for 5h at room temperature under vigorous shaking. The sample was centrifuged at 50,000x*g* for 1h at 20°C, the supernatant was incubated with Ni-NTA resin previously equilibrated with buffer A, 8M urea, and the fusion protein was eluted using buffer E (50mM sodium phosphate, 250mM NaCl, 8M urea, pH4.0). The sample was dialyzed overnight at room temperature against buffer A plus 0.5mM GSSG, 1mM GSH, 10mM $CaCl_2$, 1mM $ZnS0_4$, and 0.4M L-arginine, and subsequently dialyzed overnight at 4°C against buffer C. The N-terminal extension of the snapalysin construct was removed by autolytic cleavage as determined by Edman degradation, yielding an N-terminal segment Q_{-3}+G_{-2}+P_{-1}+M_1 (italicized snapalysin amino acids subscripted numbering correspond to UniProt P56406 *plus* the additional N-terminal tag residues and the T_1M mutation; although this is a secreted protein, the signal-peptide sequence is unknown and the database entry corresponds to the authentic protein purified from culture supernatants of *S. caespitosus*). The protein was concentrated by ultrafiltration, and further purified by size-exclusion chromatography on a Superdex 75 10/300 column previously equilibrated with buffer C. Protein identity and purity were assessed by Edman degradation and 15% Tricine-SDS-PAGE stained with Coomassie blue. Ultrafiltration steps were performed with Vivaspin 15 and Vivaspin 500 filter devices with 5-KDa cut-off. Protein concentration was determined by measuring the absorbance at 280nm using a NanoDrop spectrophotometer and a calculated absorption coefficient $E_{0.1\%}$ = 1.7. This procedure yielded functional snapalysin at very low concentration (approx. 50μg per liter of cell culture), which was used for activity assays (see below).

In vitro *activation studies of prosnapalysin* – To study the transition between latent and mature snapalysins, a synthetic gene encompassing the chemical sequence of prosnapalysin from *Streptomyces coelicor*—the sequence of the pro-peptide of the zymogen from *S. caespitosus* is not known—lacking the 29-residue signal peptide and with the N-terminal alanine replaced with methionine (A30M; numbering in regular characters according to UniProt P0A3Z7) due to the cloning strategy, was purchased (GenScript) and cloned into the aforementioned modified pET-28a vector using the *Nco*I and *Xho*I restriction sites. Prosnapalysin active-site mutant E164A was obtained using the QuickChange Site-Directed Mutagenesis kit and verified by DNA sequencing. Both proteins were overproduced in *E. coli* BL21 (DE3) cells, which were grown at 37°C in LB medium supplemented with kanamycin to a final concentration of 30μg/ml. Thereafter, cells were induced at an OD_{550} of 0.8 with IPTG to a final concentration of 1mM and grown for a further 5h at 37°C. After centrifugation at 7,000x*g* for 30min at 4°C, the pellet was resuspended in buffer D containing EDTA-free protease inhibitor cocktail tablets and DNase I. Cells were lysed at 4ºC using a cell disrupter at a pressure of 1.9 Kbar and incubated for 30min at 37°C. The resulting inclusion bodies were recovered by centrifugation at 7,000x*g* for 30min at 4°C and washed twice with buffer D, resuspended in buffer A, 8M urea, and incubated for 5h at room temperature under vigorous shaking. The sample was

centrifuged at 50,000x*g* for 1h at 20°C, the supernatant was incubated with Ni-NTA resin previously equilibrated with buffer A, 8M urea, and the fusion protein was eluted using buffer E. The sample was then first dialyzed overnight at room temperature against buffer A plus 0.5mM GSSG, 1mM GSH, 10mM $CaCl_2$, 1mM $ZnS0_4$, and 0.4M L-arginine, and then dialyzed overnight at 4°C against buffer C. The fusion tag of prosnapalysin mutant E164A was cleaved overnight at room temperature in buffer C, 0.5mM GSSG, 3mM GSH by incubation with TEV proteinase (enzyme:substrate ratio of 1:50 [w/w]). Proteins were concentrated by ultrafiltration, and further purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column previously equilibrated with buffer C. The final yield for the wild-type and the mutant was approx. 2mg and 5mg, respectively, per liter of cell culture. Protein identity and purity were assessed by Edman degradation and 15% Tricine-SDS-PAGE stained with Coomassie blue. Cleavage of wild-type and mutant prosnapalysin was monitored by mass spectrometry, SDS-PAGE and Edman degradation of electro-blotted bands onto an Immun-Blot PVDF membrane (BioRad). Ultrafiltration steps were performed with Vivaspin 15 and Vivaspin 500 filter devices of 5-KDa cut-off. Protein concentration was determined by measuring the absorbance at 280 nm using a NanoDrop and a calculated absorption coefficient $E_{0.1\%}$ = 1.38 for both variants.

Proteolytic and inhibition assays – Ulilysin from *Methanosarcina acetivorans*, aeruginolysin from *Pseudomonas aeruginosa*, and fragilysin-3 from *Bacteroides fragilis* were produced and purified as previously described [2-5]. The vector coding for aeruginolysin was a kind gift from Ulrich Baumann (University of Cologne, Germany). Astacin and matrix metalloproteinases-1, -3, and -13 were kindly provided by Walter Stöcker (Johannes Gutenberg-University of Mainz, Germany) and Hideaki Nagase (Imperial College London, UK), respectively. Proteolytic activity of ADAM-17 was kindly measured by Jordi Malapeira and Joaquin Arribas (Vall d'Hebron Institute of Oncology, Spain) using the substrate vasorin in a cell-based assay as previously published [6]. Subtilisin Carlsberg from *Bacillus licheniformis*, thermolysin from *Bacillus thermoproteolyticus*, bovine pancreatic trypsin, bovine pancreatic chymotrypsin, porcine pancreatic elastase, and fungal proteinase K from *Engyodontium album* were purchased from Sigma. Evident proteolytic activity of all these proteinases and recombinant snapalysin from *S. caespitosus* was measured at 1µM enzyme concentration in buffer C at 37°C with the fluorescein conjugates BODIPY FL casein and DQ gelatin (65µg/ml, Invitrogen) as substrates by using microplates (Nunc) and a microplate fluorimeter (FLx800, BioTek) at λ_{ex} = 485nm / λ_{em} = 528nm. Increasing amounts of purified wild-type and mutant variants of sermetstatin were added to the assays to determine inhibition (see Suppl. Table S2). In addition, inhibition of wild-type sermetstatin against subtilisin and snapalysin was tested by using the substrates succinyl-A-A-P-F-*p*-nitroanilide (Sigma) and aminobenzoyl-V-K-F-Y-D-I-K(2,4-dinitrophenylamino), respectively,—kindly provided by Jean-Louis Reymond (University of Berne, Switzerland)—in buffer C, and the associated apparent inhibition constants (K_i) were derived from a Dixon plot.

Cleavage of sermetstatin mutants –Wild-type sermetstatin, as well as mutants $C^{31}S+C^{46}S$ and $C^{69}S+C^{99}S$ (at 0.5mg/ml in 20mM Tris-HCl, 50mM NaCl, pH7.4), were incubated with equimolar

amounts of either subtilisin or snapalysin overnight at room temperature. The reactions were then assessed for cleavage in a Bruker Autoflex III mass spectrometer. Each sample was mixed at a 1:1 ratio (v/v) with a matrix solution of sinapinic acid (10mg/ml, Sigma) dissolved in 30% acetonitrile and 0.1% trifluoroacetic acid, and subsequently spotted onto the plate using the dried-droplet method. Mass spectra were acquired in linear mode geometry with >1,000 laser shots and using a protein mixture ranging from 5 KDa to 17.5 KDa (Protein Calibration Standard I, Bruker) as a calibrator.

Complex formation and purification – Commercial subtilisin Carlsberg from *B. licheniformis* was purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column previously equilibrated with buffer C. The subtilisin:sermetstatin complex was formed by incubation of equimolar amounts of purified wild-type inhibitor and protease, and it eluted as a 2+2 heterotetramer in calibrated size-exclusion chromatography. The complex was concentrated by ultrafiltration up to 5.5mg/ml using Vivaspin 15 and Vivaspin 500 filter devices of 30-KDa cut-off. Protein concentration was determined by measuring the absorbance at 280nm using a NanoDrop spectrophotometer and a calculated absorption coefficient $E_{0.1\%}$ = 0.89. To produce sufficient amounts of snapalysin for structural studies, inclusion bodies of snapalysin were obtained and purified as described above. Next, unfolded snapalysin dissolved in buffer E was diluted 1:9 against buffer A, 0.55mM GSSG, 1.1mM GSH, 11mM $CaCl_2$, 1.1mM $ZnS0_4$, 0.44M L-arginine and an equimolar amount of wild-type sermetstatin. The sample was incubated for 4h at 4°C and dialyzed overnight at 4°C against buffer C. The snapalysin:sermetstatin complex was concentrated by ultrafiltration, and further purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column previously equilibrated with buffer C and calibrated, which revealed that the complex was a 2+2 heterotetramer in solution. The sample was dialyzed overnight at 4°C against buffer F (20mM Tris-HCl, 1M NaCl, pH7.4) and concentrated by ultrafiltration to 4.5mg/ml using Vivaspin 15 and Vivaspin 500 filter devices of 10-KDa cut-off. The protein concentration was determined by measuring the absorbance at 280nm using a NanoDrop spectrophotometer and a calculated absorption coefficient $E_{0.1\%}$ = 1.32. Finally, the ternary complex between subtilisin, sermetstatin, and snapalysin was obtained by adding an equimolar amount of subtilisin to the snapalysin:sermetstatin complex. A further size-exclusion chromatography step in a calibrated Superdex 200 10/300 column previously equilibrated with buffer C revealed that the quaternary arrangement in solution was a 2+2+2 heterohexamer (see Suppl. Fig. S4). The complex was concentrated by ultrafiltration to 7.5mg/ml with Vivaspin 15 and Vivaspin 500 filter devices of 10-KDa cut-off. Protein concentration was determined by measuring the absorbance at 280nm using a NanoDrop spectrophotometer and a calculated absorption coefficient $E_{0.1\%}$ = 1.1.

Crystallization and X-ray diffraction data collection – Crystallization assays of sermetstatin, its two binary complexes with snapalysin and subtilisin, and the ternary complex with both peptidases were carried out by the sitting-drop vapor diffusion method. Reservoir solutions were prepared by a Tecan robot, and 200-nl crystallization drops, equivolumetric in protein solution and reservoir solution, were dispensed on 96x2-well MRC plates (Innovadyne) by a Cartesian nanodrop robot

(Genomic Solutions) at the IBMB/IRB joint High-Throughput Automated Crystallography Platform. Crystallization plates were stored in Bruker steady-temperature crystal farms at 4ºC and 20°C. Successful hits were scaled up to the microliter range with 24-well Cryschem crystallization dishes (Hampton Research) whenever possible.

The best crystals of wild-type and selenomethionine-derivatized sermetstatin appeared at 20°C with protein solution (10mg/ml in buffer C) and 100mM sodium citrate dihydrate, 200mM ammonium acetate, 10% (w/v) PEG3,350, pH5.6 as reservoir solution in microliter drops. The best crystals of the subtilisin:sermetstatin complex were obtained at 4°C with protein solution (5.5mg/ml in buffer C) and 100mM cacodylate, 200mM zinc acetate dihydrate, 10% (v/v) 2-propanol, pH6.5 as reservoir solution from microliter drops. Suitable crystals of the snapalysin:sermetstatin complex were obtained from 200-nanoliter drops at 20°C with protein solution (4.5mg/ml in buffer F) and 100mM HEPES, 200mM magnesium chloride, 15.0% (v/v) ethanol, pH7.5 as reservoir solution. No suitable crystals of the ternary complex could be obtained despite extensive trials. Crystals were cryo-protected by successive passages through reservoir solutions containing increasing amounts of glycerol (up to 25-30% [v/v]). Complete diffraction datasets were collected at 100K from liquid-N_2 flash-cryo-cooled crystals (Oxford Cryosystems 700 series cryostream) on MarCCD (beam line ID23-2, native sermetstatin, snapalysin:sermetstatin complex) and ADSC Q315R CCD (ID29, sermetstatin selenomethionine derivative; and ID23-1, subtilisin:sermetstatin complex) detectors at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) within the Block Allocation Group “BAG Barcelona”. Crystals of both native and selenomethionine-derivatized sermetstatin were trigonal, with one molecule per asymmetric unit (V_M = 3.2Å^3/Da; 62% solvent content). Crystals of subtilisin:sermetstatin were monoclinic, with two complexes per asymmetric unit (V_M = 3.6Å^3/Da; 65% solvent content). Crystals of snapalysin:sermetstatin were orthorhombic with six complexes per asymmetric unit (V_M = 2.9Å^3/Da; 58% solvent content). Diffraction data were integrated, scaled, merged, and reduced with programs XDS [7], XSCALE, and SCALA [8] within the CCP4 suite [9] (see Suppl. Table S3).

Structure solution and refinement – All attempts to solve the structure of unbound sermetstatin by Patterson search [10] using the coordinates of the related *Streptomyces* subtilisin inhibitor (SSI; Protein Data Bank [PDB] access code 3SSI; [11]) from *Streptomyces albogriseolus* as a searching model failed despite a sequence identity of 43%. Therefore, the structure was solved by single-wavelength anomalous diffraction using a selenomethionine derivative and program SHELXD [12]. Diffraction data of a crystal collected at the selenium absorption-peak wavelength, as inferred from a XANES fluorescence scan, enabled the program to identify all three selenium sites of the monomer present in the asymmetric unit. Subsequent phasing with SHELXE [12] resolved the twofold ambiguity intrinsic to a SAD experiment due to the difference in the values of the pseudo-free correlation coefficient of the two possible hands, which confirmed $P3_121$ as the correct space group. An initial model was built with program COOT [13] and refined against the selenomethionine dataset with program REFMAC5 [14]. The

resulting partial model was used to determine the native structure by Patterson-search methods with program PHASER [15]. One unambiguous hit was found, which yielded Z-scores for the rotation and translation functions of 20.4 and 46.0, respectively. Subsequently, manual model building with COOT and TURBO-Frodo [16] alternated with crystallographic refinement (including TLS refinement) with programs REFMAC5 and BUSTER/TNT [17] until completion of the model (Suppl. Table S3). The final model of sermetstatin comprised residues G^{-1}+S^{1}-F^{113} (chain A) *plus* ligands and solvent molecules (see Suppl. Table S3). The dimeric quaternary arrangement observed in solution (see above) was likewise found in the crystal formed by a crystallographic twofold axis.

The structure of the subtilisin:sermetstatin complex was solved by Patterson-search methods with program PHASER using the co-ordinates of subtilisin Carlsberg from *B. licheniformis* (Protein Data Bank [PDB] access code 1SBC; [18]) and sermetstatin as search models. Two unambiguous hits were found for each of the two molecules, which showed Z-scores for the rotation/translation functions of 20.8/22.8 and 18.7/46.2 (subtilisin) and 6.0/36.2 and 6.8/44.5 (sermetstatin). The quaternary arrangement in the crystal asymmetric unit is a 2+2 heterotetramer (chains A-D). Model building and refinement—with automatic non-crystallographic symmetry restraints—proceeded as before (Suppl. Table S3). The final model contained residues A_1-Q_{275} (subscripted subtilisin residue numbers according to the mature enzyme sequence; see UniProt P00780) and two calcium ions (Ca_{999} and Ca_{998}) for the two subtilisin moieties (chains A and C), G^{-1}+S^{1}-F^{113} for sermetstatin chain B, and G^{4}-F^{113} for sermetstatin chain D *plus* ligands and solvent molecules (see Suppl. Table S3).

The structure of the snapalysin:sermetstatin complex was likewise solved by Patterson search using the coordinates of *S. caespitosus* snapalysin (PDB 1KUH; [19]) and sermetstatin. Due to the presence of 12 molecules in the asymmetric unit in total, searches had to be split into two steps: first, four copies of each protein were searched for independently, enabling us to identify two complexes showing the same relative arrangement between proteinase and inhibitor. In a second search, one such complex was used as a search model, which yielded six solutions arranged as three 2+2 heterotetramers (chains A-D, E-H, I-L), with Z-scores for the respective rotation/translation functions of 6.7/15.3, 8.1/26.2, 7.3/29.8, 8.8/35.9, 5.5/37.8, and 6.6/35.7. Model building and refinement—with automatic non-crystallographic symmetry restraints—proceeded as before (Suppl. Table S3). The final model contained snapalysin residues G_{-2}+P_{-1}+M_1-G_{132} in chains A, C, and I; residues P_{-1}+M_1-G_{132} in chains G and K; and residues M_1-G_{132} in chain E. In addition, sermetstatin chains spanned residues S^{1}-F^{113} (chains B and L), A^{2}-F^{113} (chain F), S^{1}-A^{62}+L^{67}-F^{113} (chain D), S^{1}-L^{60}+L^{67}-F^{113} (chain H), and S^{1}-P^{65}+V^{68}-F^{113} (chain J). Further ligands and solvent molecules completed the model (see Suppl. Table S3). Due to the distinct packing environments of the molecules within the crystal, chains E and K were overall less well-defined by electron density that the other ten chains; this is indicated by the significantly higher average thermal displacement parameters for their protein parts (76.0 and 90.3Å^2, respectively) than the remaining molecules (44.8-64.5Å^2).

A composite model of the ternary 2+2+2 complex between subtilisin, sermetstatin, and snapalysin was obtained with TURBO-Frodo starting from the subtilisin:sermetstatin complex after superposition of the dimeric inhibitor moieties of the binary complexes. Only the N-terminal segment, Lβ2β3, and segment β1-Lβ1β2-β2 had to be slightly remodeled to match the conformation found in the snapalysin:sermetstatin complex. These segments were subjected to geometric refinement with TURBO-Frodo.

Miscellaneous – Figures were prepared with programs TURBO-Frodo and CHIMERA [20]. Interaction surfaces (taken as half of the surface area buried at a complex interface) and close contacts (<4Å) were determined with CNS [21]. Structure similarities were investigated with DALI [22]. Model validation was performed with MOLPROBITY [23] and the WHATCHECK routine of WHATIF [24]. Sequence alignments and phylogenetic calculations were performed with MULTALIN [25]. The final coordinates of sermetstatin, the subtilisin:sermetstatin complex, and the snapalysin:sermetstatin complex have been deposited with the PDB at www.pdb.org (access codes 4HWX, 4HX2, and 4HX3). The model of the ternary complex is available from the authors upon request.

2. Acknowledgments

We thank Ulrich Baumann, Walter Stöcker, Hideaki Nagase, Jean-Louis Reymond, Jordi Malapeira, and Joaquin Arribas for providing reagents or performing assays. We are also grateful to Tibisay Guevara and the Automated Crystallography Platform at IBMB/IRB for assistance during crystallization experiments. This study was supported in part by grants from European, Spanish, and Catalan agencies (FP7-HEALTH-F3-2009-223101 “AntiPathoGN”; FP7-HEALTH-2010-261460 “Gums&Joints”; FP7-PEOPLE-2011-ITN-290246 “RAPID”; BIO2009-10334; BFU2012-32862; CSD2006-00015; a JAE postdoctoral contract from CSIC; an FPU Ph.D. fellowship from the Spanish Ministry for Science and Technology; a grant from the Andalusian Regional Government; Fundació “La Marató de TV3” grant 2009-100732; and 2009SGR1036). We acknowledge the help provided by ESRF synchrotron local contacts. Funding for data collection was provided in part by ESRF.

3. Supplemental References

1. K. Hiraga, T. Suzuki and K. Oda, *J. Biol. Chem.*, 2000, **275**, 25173-25179.
2. C. Tallant, R. García-Castellanos, J. Seco, U. Baumann and F. X. Gomis-Rüth, *J. Biol. Chem.*, 2006, **281**, 17920-17928.
3. T. Goulas, J. L. Arolas and F. X. Gomis-Rüth, *Proc. Natl. Acad. Sci. USA*, 2010, **108**, 1856-1861.
4. U. Baumann, S. Wu, K. M. Flaherty and D. B. McKay, *EMBO J.*, 1993, **12**, 3357-3364.
5. T. Hege, R. E. Feltzer, R. D. Gray and U. Baumann, *J. Biol. Chem.*, 2001, **276**, 35087-35092.
6. J. Malapeira, C. Esselens, J. J. Bech-Serra, F. Canals and J. Arribas, *Oncogene*, 2011, **30**, 1912-1922.
7. W. Kabsch, *Acta Crystallogr. sect. D*, 2010, **66**, 125-132.
8. P. Evans, *Acta Crystallogr. sect. D*, 2006, **62**, 72-82.
9. CCP4, *Acta Crystallogr. sect. D*, 1994, **50**, 760-763.
10. R. Huber, *Acta Crystallogr. sect. A*, 1965, **19**, 353-356.
11. Y. Mitsui, Y. Satow, T. Sakamaki and Y. Iitaka, *J. Biochem.*, 1977, **82**, 295-298.
12. G. M. Sheldrick, *Acta Crystallogr. sect. D*, 2010, **66**, 479-485.
13. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. sect. D*, 2010, **66**, 486-501.
14. G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta Crystallogr. sect. D*, 2011, **67**, 355-367.
15. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658-674.
16. C. Carranza, A.-G. Inisan, E. Mouthuy-Knoops, C. Cambillau and A. Roussel, in *AFMB Activity Report 1996-1999*, CNRS-UPR 9039, Marseille, 1999, pp. 89-90.
17. E. Blanc, P. Roversi, C. Vonrhein, C. Flensburg, S. M. Lea and G. Bricogne, *Acta Crystallogr. sect. D*, 2004, **60**, 2210-2221.
18. D. J. Neidhart and G. A. Petsko, *Protein Eng.*, 1988, **2**, 271-276.
19. G. Kurisu, T. Kinoshita, A. Sugimoto, A. Nagara, Y. Kai, N. Kasai and S. Harada, *J. Biochem.*, 1997, **121**, 304-308.
20. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
21. A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren, *Acta Crystallogr. sect. D*, 1998, **54**, 905-921.
22. L. Holm, S. Kaariainen, C. Wilton and D. Plewczynski, *Curr. Protoc. Bioinformatics*, 2006, **Chapter 5**, Unit 5 5.
23. I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. Bryan Arendall, 3rd, J. Snoeyink, J. S. Richardson and D. C. Richardson, *Nucl. Acids Res.*, 2007, **35 (Web Server issue)**, W375-W383.
24. G. Vriend, *J. Mol. Graph.*, 1990, **8**, 52-56.
25. F. Corpet, *Nucl. Acids Res.*, 1988, **16**, 10881-10890.
26. L. Migliolo, A. S. de Oliveira, E. A. Santos, O. L. Franco and M. P. de Sales, *J. Mol. Graph. Model.*, 2010, **29**, 148-156.
27. E. M. Meulenbroek, E. A. Thomassen, L. Pouvreau, J. P. Abrahams, H. Gruppen and N. S. Pannu, *Acta Crystallogr. sect. D*, 2012, **68**, 794-799.
28. T. A. Valueva, I. A. Parfenov, T. A. Revina, E. V. Morozkina and S. V. Benevolensky, *Plant. Physiol. Biochem.*, 2012, **52**, 83-90.
29. R. Bao, C. Z. Zhou, C. Jiang, S. X. Lin, C. W. Chi and Y. Chen, *J. Biol. Chem.*, 2009, **284**, 26676-26684.
30. Z. W. Xie, M. J. Luo, W. F. Xu and C. W. Chi, *Biochemistry*, 1997, **36**, 5846-5852.
31. M. Azarkan, S. Martinez-Rodriguez, L. Buts, D. Baeyens-Volant and A. Garcia-Pino, *J. Biol. Chem.*, 2011, **286**, 43726-43734.

32. J. K. Dattagupta, A. Podder, C. Chakrabarti, U. Sen, D. Mukhopadhyay, S. K. Dutta and M. Singh, *Proteins*, 1999, **35**, 321-331.
33. S. Capaldi, M. Perduca, B. Faggion, M. E. Carrizo, A. Tava, L. Ragona and H. L. Monaco, *J. Struct. Biol.*, 2007, **158**, 71-79.
34. P. Chen, J. Rose, R. Love, C. H. Wei and B. C. Wang, *J. Biol. Chem.*, 1992, **267**, 1990-1994.
35. J. Koepke, U. Ermler, E. Warkentin, G. Wenzl and P. Flecker, *J. Mol. Biol.*, 2000, **298**, 477-491.
36. E. M. Ragg, V. Galbusera, A. Scarafoni, A. Negri, G. Tedeschi, A. Consonni, F. Sessa and M. Duranti, *Febs J.*, 2006, **273**, 4024-4039.
37. J. A. Barbosa, L. P. Silva, R. C. Teles, G. F. Esteves, R. B. Azevedo, M. M. Ventura and S. M. de Freitas, *Biophys. J.*, 2007, **92**, 1638-1650.
38. K. N. Rao and C. G. Suresh, *Biochim. Biophys. Acta*, 2007, **1774**, 1264-1273.
39. Y. N. Sreerama and L. R. Gowda, *Biochim. Biophys. Acta*, 1997, **1343**, 235-242.
40. S. Terada, S. Fujimura and E. Kimoto, *Biosci. Biotechnol. Biochem.*, 1994, **58**, 376-379.
41. S. Norioka and T. Ikenaka, *J. Biochem.*, 1983, **94**, 589-599.
42. K. Nomura and N. Suzuki, *FEBS Lett.*, 1993, **321**, 84-88.
43. A. S. Tanaka, M. U. Sampaio, S. Marangoni, B. de Oliveira, J. C. Novello, M. L. Oliva, E. Fink and C. A. Sampaio, *Biol. Chem.*, 1997, **378**, 273-281.
44. S. Odani, T. Koide and T. Ono, *J. Biochem.*, 1986, **100**, 975-983.
45. J. E. Debreczeni, G. Bunkoczi, B. Girmann and G. M. Sheldrick, *Acta Crystallogr. sect. D*, 2003, **59**, 393-395.
46. J. Krahn and F. C. Stevens, *FEBS Lett.*, 1971, **13**, 339-341.
47. K. A. Wilson and M. Laskowski, Sr., *J. Biol. Chem.*, 1975, **250**, 4261-4267.
48. M. Renko, J. Sabotic, M. Mihelic, J. Brzin, J. Kos and D. Turk, *J. Biol. Chem.*, 2010, **285**, 308-316.
49. F. L. Tan, G. D. Zhang, J. F. Mu, N. Q. Lin and C. W. Chi, *Hoppe Seylers Z Physiol. Chem.*, 1984, **365**, 1211-1217.
50. N. Mallorquí-Fernández, S. P. Manandhar, G. Mallorquí-Fernández, I. Usón, K. Wawrzonek, T. Kantyka, M. Solà, I. B. Thogersen, J. J. Enghild, J. Potempa and F. X. Gomis-Rüth, *J. Biol. Chem.*, 2008, **283**, 2871-2882.
51. V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson, *Acta Crystallogr. sect. D*, 2010, **66**, 12-21.

4. Supplemental Tables

Suppl. Table S1. Janus-faced single-domain inhibitors with at least two reactive sites.

Family (MEROPS)	Inhibitor name	Source	UniProt code	Class of targets	Targets	PDB code	Ref.
Kunitz (I3)	ApKTI	*Adenanthera pavonina*	P09941	SP [a] (and CP [a])	Trypsin and papain (very weakly)	-	26
	PKPI	*Solanum tuberosum*	Q66LL2	SP	Trypsin and chymotrypsin	3TC2	27, 28
	API	*Sagittaria sagittifolia*	Q7M1P4	SP	Trypsin, chymotrypsin and kallikrein	3E8L [b]	29, 30
	PPI	*Carica papaya*	P80691	SP	Trypsin	3S8K, 3S8J	31
	WCI	*Psophocarpus tetragonolobus*	P10822	SP	Chymotrypsin	2WBC	32
Bowman-Birk (I12)	MSTI	*Medicago scutellata*	P80321	SP	Trypsin	2ILN [b]	33
	sBBI	*Glycine max*	P01055	SP	Trypsin and chymotrypsin	1PI2, 1D6R [b]	34, 35
	LCTI	*Lens culinaris*	Q8W4Y8	SP	Trypsin and chymotrypsin	2AIH	36
	BTCI	*Vigna unguiculata*	P17734	SP	Trypsin and chymotrypsin	2G81 [c], 2OT6	37, 38
	HGI-III	*Dolichos biflorus*	Q9S9E3	SP	Trypsin and chymotrypsin	-	39
	CLTI	*Canavalia lineata*	Q7M1Q0	SP	Trypsin and chymotrypsin	-	40
	A-II	*Arachis hypogaea*	P01066	SP	Trypsin and chymotrypsin	Not deposited	41, 42
	TcTI	*Amburana acreana*	P83284	SP	Trypsin, plasmin and chymotrypsin	-	43
	I-2B	*Triticum aestivum*	P09863	SP	Trypsin	-	44
	II-4	*Triticum aestivum*	P09864	SP	Trypsin	-	44
	LBI	*Phaseolus lunatus*	P01056	SP	Trypsin and chymotrypsin	1H34	45, 46
	GBI	*Phaseolus vulgaris*	P01060	SP	Trypsin, chymotrypsin and elastase	-	47
Clitocypin (I48)	Clitocypin	*Clitocybe nebularis*	Q3Y9I4	CP	Papain and cathepsins L, V, S and K	3H6R, 3H6S [d]	48
Macrocypin (I85)	Macrocypin 1	*Macrolepiota procera*	B9V973	CP (and SP)	Papain and cathepsins L, V, S and K (Macrocypin 4 inhibits trypsin weakly)	3H6Q	48
Others	Trypsin inhibitor	*Trichosanthes kirilowii*	P01069	SP	Trypsin	-	49

[a] SP and CP stand for serine and cysteine proteases, respectively. [b] Complex with two molecules of trypsin. [c] Complex with trypsin. [d] Complex with cathepsin V.

Suppl. Table S2. Inhibitory activity of sermetstatin variants.

		Wild-type	$Y^{33}P+T^{34}G$	$M^{71}K$	$Y^{72}V+F^{73}Y$
Protease	Molar ratio (Inh./Enz.)	Inhibitory activity (%)	Inhibitory activity (%)	Inhibitory activity (%)	Inhibitory activity (%)
Metallopeptidases					
Snapalysin	1	99	1	90	96
	10	100	52	100	100
Astacin	100	0	0	0	0
Ulilysin	100	0	0	0	0
Aeruginolysin	100	0	0	0	0
Thermolysin	100	0	0	0	0
Fragilysin	100	0	0	0	0
ADAM-17 [a]	-	NI [b]	-	-	-
MMP-1,-3,-13 [a]	100	0	0	0	0
Serine proteinases					
Subtilisin	1	95	95	24	93
Trypsin	10	21	29	53	31
Chymotrypsin	10	69	47	29	56
Elastase	10	68	67	0	62
Proteinase K	1	82	75	5	77
	10	100	100	36	100

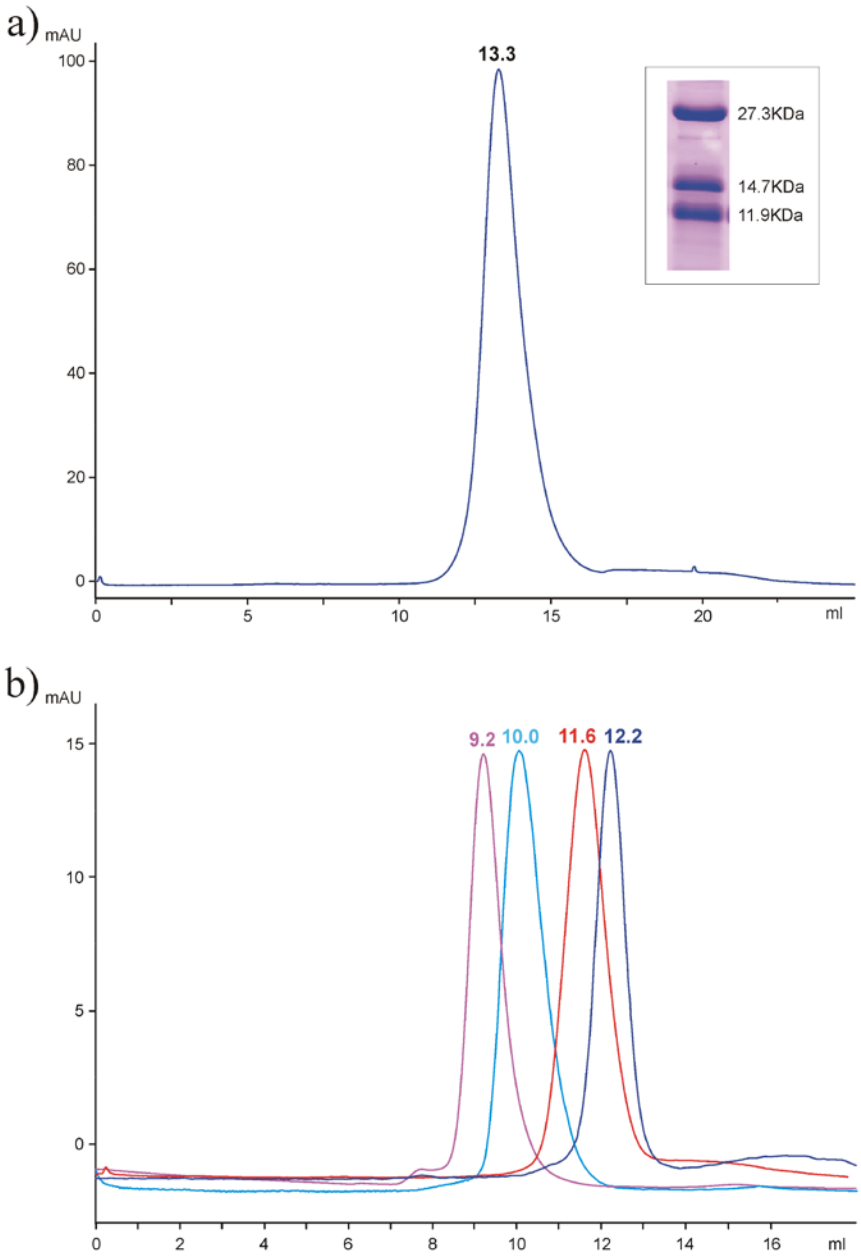
Values as mean of three independent measurements, SD within ±5. [a] ADAM stands for a disintegrin and metalloprotease, MMP for matrix metalloproteinase. [b] NI, no inhibition at 4 μM inhibitor concentration in a cell-based assay.

Suppl. Table S3. Crystallographic data.

	Sermetstatin		Subtilisin:sermetstatin	Snapalysin:sermetstatin
Dataset	Native	Selenomethionine (absorption peak) [a]	Native	Native
Space group	$P3_121$	$P3_121$	C2	$P2_12_12_1$
Cell constants (a, b, c, in Å; β in ° if ≠90)	71.04, 71.04, 52.40	70.81, 70.81, 52.25	183.97, 83.62, 77.62; 110.78	116.54, 121.81, 130.67
Wavelength (Å)	0.8726	0.9791	0.9763	0.8726
No. of measurements / unique reflections	134,041 / 12,234	218,877 / 10,514	312,908 / 52,346	211,0198 / 51,602
Resolution range (Å) (outermost shell) [b]	39.9 – 1.90 (2.00 – 1.90)	26.4 – 2.00 (2.11 – 2.00)	48.6 – 2.25 (2.37– 2.25)	89.1 – 2.70 (2.77– 2.70)
Completeness [/ Anom. completeness] (%)	99.3 (98.9)	99.8 (100.0) / 99.7 (100.0)	99.9 (99.7)	99.4 (95.8)
R_{merge} [c]	0.038 (0.596)	0.074 (0.757)	0.097 (0.751)	0.091 (0.704)
$R_{r.i.m.}$ (= R_{meas}) [c] [/ $R_{p.i.m.}$ [c]]	0.040 (0.625) / 0.012 (0.185)	0.077 (0.793) / 0.023 (0.235)	0.106 (0.842) / 0.043 (0.368)	0.105 (0.845)
Average intensity (<[<I> / σ(<I>)]>)	40.3 (4.9)	35.4 (8.0)	13.2 (2.0)	14.7 (1.9)
B-Factor (Wilson) (Å^2) / Aver. multiplicity	32.7 / 11.0 (11.1)	29.5 / 20.8 (21.8)	41.9 / 6.0 (4.9)	51.9 / 4.1 (3.2)
Resolution range used for refinement (Å)	∞ – 1.90		∞ – 2.25	∞ – 2.70
No. of reflections used (test set)	11,844 (479)		51,487 (797)	50,699 (782)
Crystallographic R_{factor} (free R_{factor}) [c]	0.187 (0.214)		0.177 (0.218)	0.195 (0.242)
No. of protein atoms / solvent molecules /	834 / 126 /		5,487 / 359 /	11,043 / 275 /
neutral ligands	1 glycerol		1 diglycerol, 3 2-propanol, 2 glycerol	2 glycerol
ionic ligands	3 CH_3COO^-		4 Ca^{2+}, 2 Zn^{2+}, 2 K^+, 1 Cl^-, 1 $(CH_3)_2AsOO^-$, 1 CH_3COO^-, 1 PO_4^{3-}	6 Zn^{2+}
Rmsd from target values				
bonds (Å) / angles (°)	0.010 / 1.14		0.010 / 1.12	0.010 / 1.06
Average B-factors for protein atoms (Å^2)	40.4		51.1	58.2
Main-chain conformational angle analysis [d]				
Residues in				
favored regions / outliers / all residues	109 / 1 / 112		746 / 1 / 764	1,406 / 2 / 1,436

[a] Friedel-mates were treated as separate reflections. [b] Values in parentheses refer to the outermost resolution shell.
[c] For definitions, see Table 1 in [50]. [d] According to MOLPROBITY [51].

5. Supplemental Figures



Suppl. Figure S1 – Size-exclusion chromatography of the ternary complex. **(a)** The complex between subtilisin, sermetstatin and snapalysin was loaded onto a Superdex 200 10/300 column. The complex eluted after 13.3ml, which corresponds to 102KDa and reveals the formation of a 2+2+2 heterohexamer in solution (theoretical mass 107.8KDa). The column was calibrated with the following protein standards: aldolase (158KDa; elution after 12.5ml), conalbumin (75KDa; 13.9ml), ovalbumin (43KDa; 14.8ml), and cytochrome C (12.4KDa; 17.8ml). The presence of the three molecules (subtilisin 27KDa, snapalysin 15KDa, and sermetstatin 12KDa) was further confirmed by SDS-PAGE (see inset) after trichloroacetic acid-precipitation of the collected sample. **(b)** The isolated inhibitor and the binary and ternary complexes were loaded onto a Superdex 75 10/300 column for comparison. The sermetstatin eluted after 12.2ml (1+1), the snapalysin:sermetstatin complex after 11.6ml (2+2), the subtilisin:sermetstatin complex after 10.0ml (2+2), and the ternary complex after 9.2ml (2+2+2).