

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

Mechanism of action of a Janus-faced single-domain protein inhibitor simultaneously targeting two peptidase classes

Sergio Trillo-Muyo^a, Sergio Martínez-Rodríguez^b, Joan L. Arolas^{a,*} and F. Xavier Gomis-Rüth^{a,*}

Received (in XXX, XXX) Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXXX 20XX

DOI: 10.1039/b000000x

Protein inhibitors provide a physiological mechanism of regulation of proteolytic enzymes. While most single-domain inhibitors have one reactive site with which they target peptidases of a specific catalytic class, selected specimens inhibit two peptidase molecules simultaneously, thus giving rise to ternary complexes. To study such inhibition, we analyzed the function of one of these proteins, *sermetstatin*, which strongly binds as a dimer to serine proteinases (SPs) and a metallopeptidase (MP). In addition, we determined the structures of the isolated inhibitor dimer and its heterotetrameric complexes with the SP subtilisin and the MP snapalysin, which reveal that inhibition occurs through two independent distal reactive sites. These structures and the derived model for the heterohexameric complex provide for the first time a detailed view of the molecular mechanism of simultaneous inhibition of proteinases belonging to two distinct mechanistic classes by a single-domain protein.

Introduction

Due to the essentially irreversible nature of peptide-bond hydrolysis, proteolytic enzymes must be tightly regulated to prevent spatially and temporally deleterious activity. One such regulatory mechanism is provided by co-localizing protein inhibitors which mostly target one or several peptidases of a certain catalytic class through a single inhibitory region by blocking the active-site cleft of the enzyme^{1,2}. A special case is provided by inhibitors with two inhibitory sites which mimic the head architecture of the Roman god Janus and are capable of inhibiting two peptidase molecules simultaneously³. Such Janus-faced inhibitors (see Suppl. Table S1) are infrequent among the plethora of inhibitors described, and none of them has yet been structurally characterized in a ternary complex with two different targets.

To address this question, we focused on the bacterial genus *Streptomyces*, which occurs in soil and aquatic habitats and is the source of many medicinal, pharmaceutical, and industrial products of natural origin such as antibiotics⁴. These bacteria secrete a number of peptidases and cognate protein inhibitors thought to participate in morphological development and defense against extrinsic proteinases. Among these inhibitors are members of family I16 (according to MEROPS inhibitor database, <http://merops.sanger.ac.uk>; ⁵), which only occur within the bacterial order *Actinomycetales*. These inhibitors generally target serine proteinases (SPs) of the chymotrypsin- and subtilisin-type; some also bind tightly to *Streptomyces griseus* griselysin but apparently not the structural and functional relative thermolysin from *Bacillus thermoproteolyticus*, which is in

MEROPS metallopeptidase (MP) family M4⁵. This inhibition is exerted by the same region employed for SPs⁶, so that, although they are bi-functional, dual inhibition cannot occur simultaneously. Previously studied I16 inhibitors include *Streptomyces* subtilisin inhibitor (SSI) from *Streptomyces albogriseolus*, plasminostreptin from *Streptomyces antifibrinolyticus*, alkaline protease inhibitor API-2c' from *Streptomyces pseudogriseolus*, and trypsin inhibitor STI2 from *Streptomyces longisporus*^{5,7-11} (see Fig. 1a). Among them is also *sermetstatin*—also known as *Streptomyces caespitosus* neutral proteinase inhibitor⁸—, which is a 113-residue secreted single-domain molecule and the main object of study of the present work.

Inhibitory profile of sermetstatin

We produced *sermetstatin* by recombinant overexpression in *Escherichia coli* (see *Experimental procedures* in the *Electronic Supplementary Information [ESI]*) and found it to be a homodimer in solution, thus confirming earlier results⁸. We assessed its inhibitory function, which had been previously ruled out for the cysteine proteinase cathepsin B and the aspartic proteinase pepsin⁸, against a broad panel of model peptidases (Suppl. Table S2). To this end, we recombinantly produced and purified the MP snapalysin from *S. caespitosus*, which is a member of the metzincin clan of MPs—unrelated to griselysin and other M4 peptidases¹²⁻¹⁴. We found it is potently inhibited by *sermetstatin* with an apparent K_i of 6nM, which is consistent with reported results⁸. These inhibition studies were complemented by autolytic activation studies with recombinant wild-type prosnapalysin from *Streptomyces coelicolor* (the sequence of the propeptide of the *S. caespitosus* zymogen is not known). We found the zymogen underwent autolytic activation *in vitro* at the



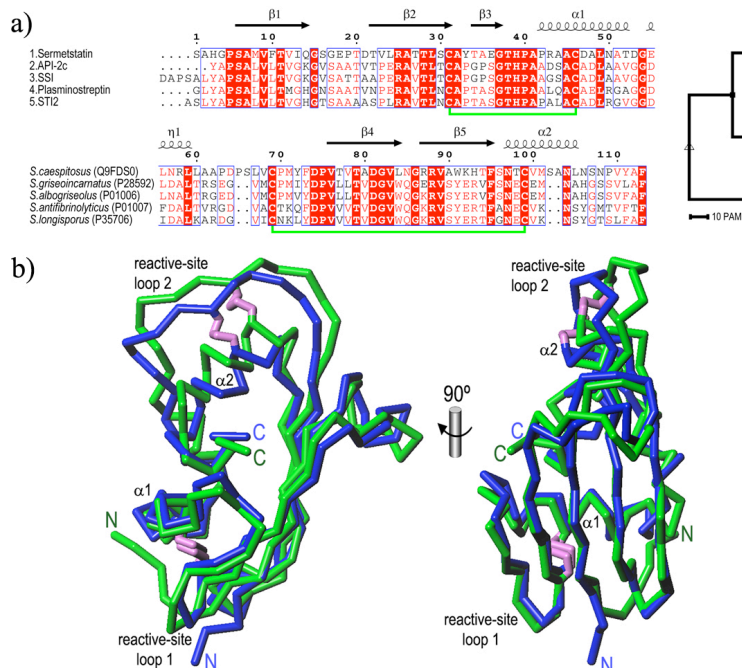


Figure 1. (a) (Left) Sequence alignment of representative members of the I16 family of peptidase inhibitors according to MEROPS database (<http://merops.sanger.ac.uk>). API-2c² stands for alkaline protease inhibitor-2c², SSI for *Streptomyces* subtilisin inhibitor, and ST12 for trypsin inhibitor ST12. The two disulfide bonds found in these proteins are depicted as green lines. The corresponding organism names (“S.” stands for *Streptomyces*) and UniProt database access codes are provided in the second block. (Right) Phylogenetic analysis of the five inhibitors (numbering corresponds to that of the upper block of the alignment) reflecting the large evolutionary distance of sermetstatin from the other four sequences. PAM=point accepted mutation. (b) (Left) Superposition of the α -traces of sermetstatin (green) and SSI (blue; PDB 2S1C) taken from their respective complexes with subtilisin in the front reference view and (right) a lateral view. 96 common α atoms deviating less than 3Å showed an *rmsd* of 1.48Å. The respective termini are labeled as the reactive-site loops 1 and 2, as well as scaffold-helices α 1 and α 2. Disulfide bonds are in pink.

trypsin (Suppl. Table S2). These results confirm that sermetstatin is a potent, though rather broad-spectrum inhibitor of subtilisin-type—and, partially, chymotrypsin-type—SPs through its other Janus face.

Structure of sermetstatin

We crystallized and solved the structure of sermetstatin by single-wavelength anomalous diffraction with the help of a selenomethione derivative (see *ESI* and Suppl. Table S3). The protein is an elongated α/β -sandwich of maximal dimensions 45x30x25Å. It consists of a five-stranded antiparallel β -sheet (β 3- β 2- β 1- β 4- β 5; connectivity -1,-1,+3,+1) twisted by $\sim 30^\circ$, whose concave face accommodates two α -helices (α 1 and α 2) and a 3_{10} -helix (η 1; see Figs. 1a and 2a). The overall structure is tethered by a large hydrophobic core provided by side chains of the β -sheet on its concave face, the two α -helices, and the C-terminal tail. The structure is further maintained by two disulfide bonds. The first (C^{31} - C^{46} ; sermetstatin mature-protein residues 55 numbered in superscript; see Fig. 1a and subtract 28 from UniProt Q9FDS0 numbers) links the first “scaffold helix” α 1 with β 2 and the subsequent loop connecting strands β 2 with β 3 (L β 2 β 3), which includes the “reactive-site loop 1” (C^{31} - A^{35}). The second disulfide bond (C^{69} - C^{99}) connects the second “scaffold helix” α 2 with

L α 2 β 4, which contains the “reactive-site loop 2” (P^{65} - F^{73}). Both reactive-site loops protrude from the molecular surface and are fully solvent accessible (Fig. 2a). The C-terminal carboxylate is anchored to the side chain of R²⁵ within β 2; by contrast, the N-terminus loosely protrudes from the molecular surface and becomes only rigid at P⁵, which leads to the first residue of strand β 1.

Two sermetstatin molecules intimately associate through the convex faces of their β -sheets to yield a crystallographically-related dimer (Fig. 2a), in agreement with the oligomeric state found in solution. The sheet planes are rotated by $\sim 35^\circ$ away from each other and the interaction surface spans $\sim 800\text{Å}^2$, *viz.* 25% of the total surface of a monomer. The interaction is mainly hydrophobic and made by 31 close contacts, which include symmetric hydrophobic interactions between ten residues provided by each sheet of either monomer, and only two salt bridges between R⁸⁸ of one protein chain and D⁸¹ of the other.

Structure similarity searches with sermetstatin against the Protein Data Bank (PDB) only identified the I16-family relative SSI, which is also a dimer, as being closely related. However, significant variations, in particular around reactive-site loop 2, the scaffold helices and the linking disulfide bonds, and, most importantly, the N-terminal segment—which is longer, differs in sequence (Fig. 1a), and is disordered until positions 4-7 in the different wild-type SSI structures reported (PDB access codes 3SSI¹⁸, 2SIC¹⁹, and 2TLD²⁰)—account for functional differences (see below). Consistently, the sermetstatin structure could not be solved by Patterson-search techniques using SSI as search model (see *ESI* and Fig. 1b). In addition to SSI, the isolated structure of plasminostreptin has been reported²¹ but the coordinates are not available.

concentration resulting from the purification strategy so fast that the intact zymogen could actually not be isolated. This cleavage also removed the N-terminal fusion protein employed for purification. Two initial cleavages were observed at bonds F59-F60 and G212-F213 (*S. coelicolor* prosnapalysin numbering in regular characters; see UniProt sequence database access code P0A3Z7), which removed not only the N-terminal part but also a 15-residue C-terminal fragment, as previously described for the MP ulilyisin from *Methanosarcina acetivorans*¹⁵. Subsequently, several cleavages occurred in the remaining part of the 49-residue propeptide to finally yield the mature 133-residue form spanning residues A80-G212. Consistent with an autolytic mechanism *in vitro*, mutant E164A, which affects the glutamate that plays the role of a general base/acid in the catalytic process in MPs^{12, 16, 17}, displays only very slow activation over a period of weeks. In nature, however, prosnapalysin activation is likely to occur under assistance of other proteinases.

Further to snapalysin, sermetstatin did not inhibit any other MP assayed, which included thermolysin and metzincins of the astacin, pappalysin, serralysin, fragilyisin, ADAM/adamalysin and matrix metalloproteinase families. Taken together with previous results showing no inhibition of the M4-family MPs *Pseudomonas aeruginosa* elastase and vimelysin, as well as the unassigned MP almelysin⁸, we conclude that sermetstatin is most likely a specific inhibitor of snapalysins through one of its Janus faces. In addition, sermetstatin potently inhibited subtilisin Carlsberg from *Bacillus licheniformis* (with an apparent K_i of 1.3nM) and proteinase K (Suppl. Table S2), as well as subtilisin BPN’ from *Bacillus amyloliquefaciens*⁸, all of them peptidases of biotechnological relevance. However, it only moderately inhibited chymotrypsin and elastase, and only weakly targeted

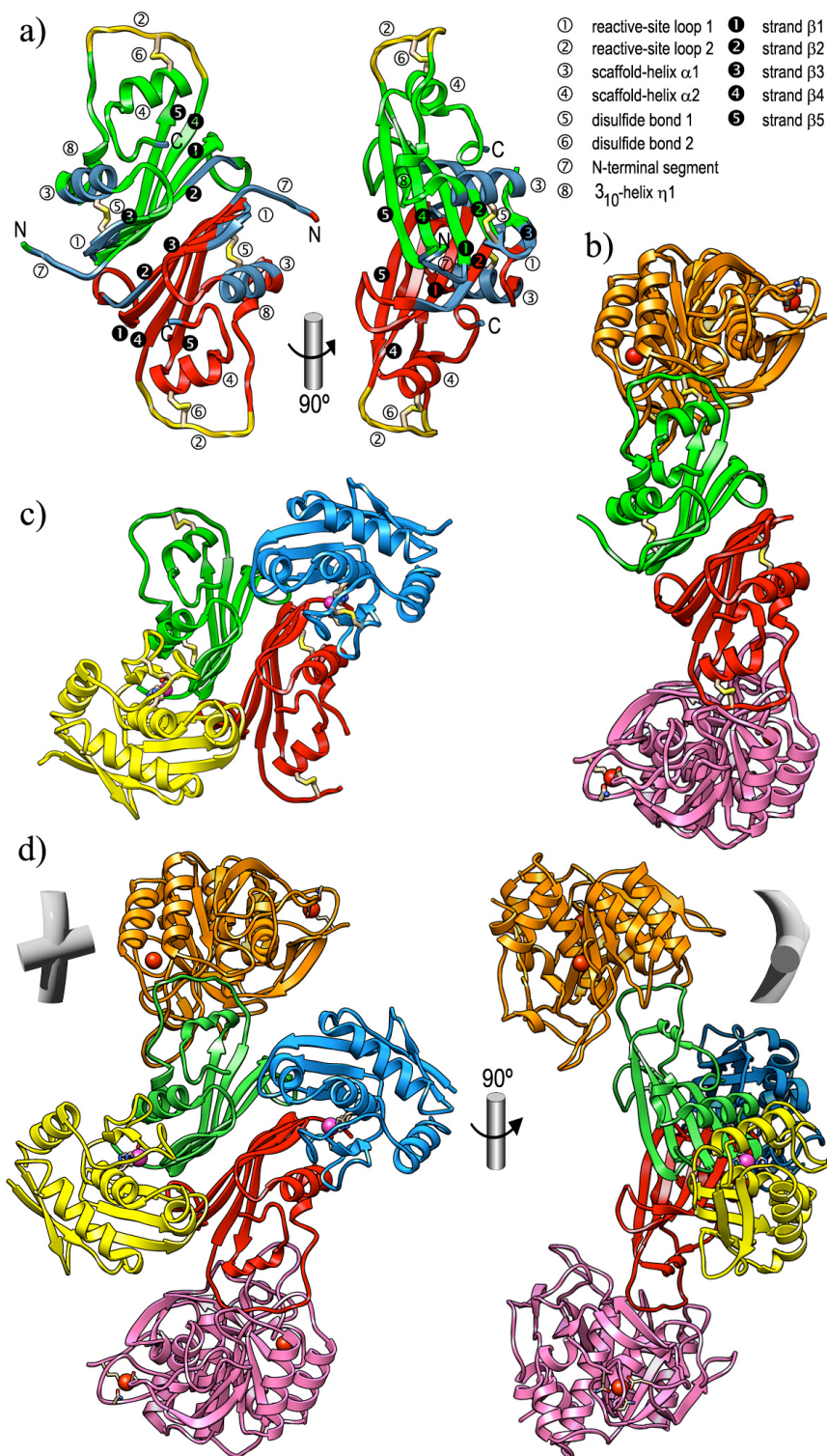


Figure 2. (a) (left) Structure of the unbound sermetstatin dimer (in red and green) in the reference orientation chosen and (right) after a vertical rotation of 90°. Characteristic segments are pinpointed by encircled numbers. Regions engaged in MP/SP binding are in blue/yellow. (b) Structure of the heterotetrameric subtilisin:sermetstatin and (c) snapalysin:sermetstatin complexes. Proteinase moieties are in orange/pink and blue/yellow, respectively. (d) Model of the heterohexameric inhibitor complex derived from (b) and (c) in two orientations.

Structure of the subtilisin: sermetstatin complex

The crystal structure of the complex between subtilisin Carlsberg and sermetstatin was determined by Patterson search

and reveals a 2+2 heterotetramer as anticipated by studies in solution, with a central non-cleaved inhibitor dimer—as present in the unbound structure—and one enzyme moiety bound to each inhibitor monomer (Fig. 2b and Suppl. Table S3). Only two segments of sermetstatin adopt a different conformation from the unbound structure: the flexible N-terminal segment until A⁷, and the segment preceding and including reactive-site loop 2 until C⁶⁹. The latter variation is due to the flip of the peptide bond L⁶⁰-A⁶¹. The only enzyme:inhibitor interaction is precisely made by reactive-site loop 2, adequately positioned by its scaffold helix, and the active-site cleft of the enzyme (Fig. 3a). The loop, which includes the reactive-site bond M⁷¹-Y⁷², is inserted wedge-like mimicking a substrate in extended, “canonical” conformation in the active-site cleft of the enzyme following the “standard mechanism” of inhibition, terms which were coined by Laskowski, Bode, Huber, and others^{1, 2, 22}. This mechanism has only been structurally proven to date for SP inhibitors in general and for the specific insect metalloproteinase inhibitor of thermolysin-family MPs²³. Rigidity to reactive-site loop 2 is conferred by the proline at position **P2** (active-site cleft and substrate sub-site nomenclature according to^{17, 24} in bold) and by disulfide bond C⁶⁹-C⁹⁹. On the non-primed side of the subtilisin cleft, the inhibitor main chain is anchored via a β -ribbon interaction to segment G₁₀₀-Y₁₀₄ (subtilisin mature-protein residue numbering in subscript; subtract 104 from UniProt P00780 numbers), which creates an upper rim of the cleft, through three hydrogen bonds. Below the cleft, subtilisin segment S₁₂₅-G₁₂₈ further tethers the inhibitor through three more inter-main-chain hydrogen bonds. V⁶⁸ and M⁷¹, respectively, occupy the two sub-sites of the active-site cleft important for specificity in subtilisin, **S4** and **S1**²⁵. The catalytic serine, S₂₂₁, hydrogen bonds through its O γ atom the inhibitor at M⁷¹ N and, downstream of **P1**, the polypeptide undergoes a kink that probably prevents cleavage at M⁷¹-Y⁷² and results in the main chain at the two subsequent residues being hydrogen-bonded in an antiparallel manner by N₃₁₈-T₂₂₀. Contrary to MPs, substrate position **P1**⁷ is not important for specificity in SPs in general, and Y⁷² points toward bulk solvent; F⁷³ in **P3**⁷, by contrast, approaches the enzyme and performs a stacking interaction with F₁₈₉ (Fig. 3a). The small interacting surface between enzyme and inhibitor—despite high inhibitory potency (see above)—means that upon superposition of the proteinase moieties of the two complexes within the heterotetramer, the inhibitor moieties display a relative rotation of ~10° around an axis horizontally traversing the active-site cleft

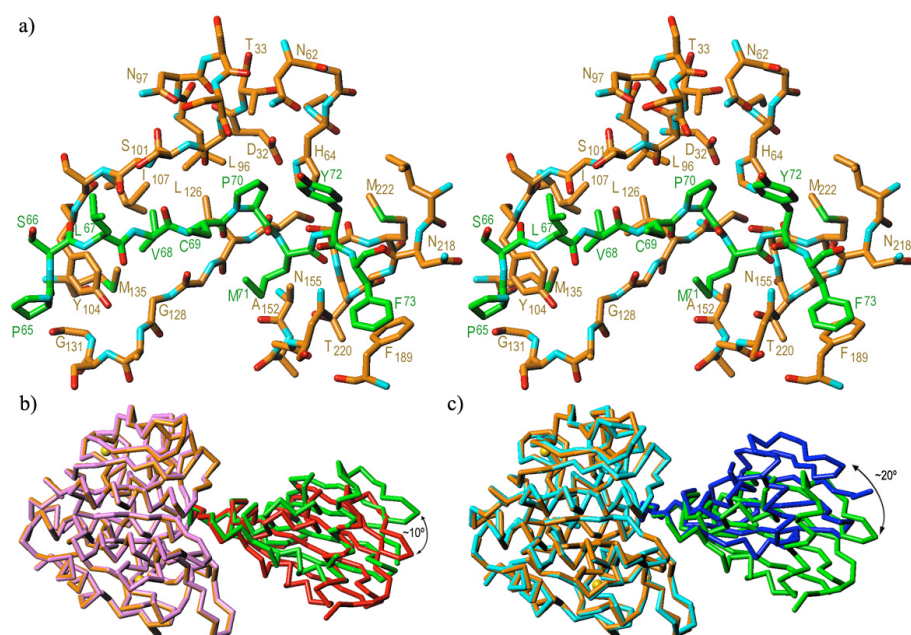


Figure 3. (a) Detail in cross-eye stereo of the interaction between sermetstatin reactive-site loop 2 (green carbon atoms and residue labels) and subtilisin Carlsberg (orange carbon atoms and residue labels) shown in standard orientation and in the reference front view of the inhibitor (see Fig. 2a). The catalytic triad is S₂₂₁-H₆₄-D₃₂. (b) Superposition of the two complexes within the heterotetrameric complex (chains A/B in orange/green and C/D in pink/red) after optimal fit of the proteinase moieties. The two calcium cations of chain A are shown as yellow spheres for reference. The view corresponds to (a) after a vertical and an in-plane rotation of 90° each. (c) Superposition of the sermetstatin:subtilisin Carlsberg complex (chains A/B in orange/green) onto the complex between SSI and subtilisin BPN' (PDB 2SIC) in blue/cyan after optimal fit of the proteinase moieties. View as in (b).

equivalent heterotetramers totaling 12 protein chains (A-L) were found in the crystal asymmetric unit, and the discussion hereafter is centered on tetramer ABCD (see *ESI*). The structure of the central dimeric sermetstatin core shows an intact inhibitor and almost exactly the same minor structural deviations from the unbound structure as seen in the subtilisin complex. However, unlike the unbound structure and the latter complex, reactive-site loop 2 is flexible and partially disordered in some of the polypeptide chains. Snapalysin, in turn, is in a virtually identical conformation to the unbound structure (PDB 1KUH; 26) throughout its entire structure except for loop A₃₆-A₄₀ (italicized *S. caespitosus* snapalysin residues and subscript numbers according to the mature MP sequence; see UniProt P56406), which is distal from the active-site cleft and thus the zone of interaction with sermetstatin. Within the complex, the proteinase is bound by a region on the opposite surface to the subtilisin binding site (Fig. 2c) through a large surface contributed to by several protein segments. This results in little

(Fig. 3b). In addition, it also explains why several SPs are targeted by the inhibitor, as no exosites are employed for binding that could confer specificity. Finally, superposition of the complex with the one reported between SSI and subtilisin BPN' (PDB 2SIC; 19) reveals very similar binding. As observed in the asymmetric unit of the subtilisin:sermetstatin structure, a relative reorientation between the sermetstatin and SSI inhibitor moieties—here of ~20°—is found upon superposition of the respective enzyme parts (Fig. 3c).

Redesign of inhibitory specificity

Inhibitory studies with wild-type sermetstatin revealed only weak activity against trypsin (Suppl. Table S2 and 8), possibly owing to the presence of a methionine in P1 (M⁷¹). In order to design a better inhibitor of the latter SP, we produced a mutant, M⁷¹K, to match the specificity of trypsin. While this mutant did not have any influence on the MPs tested, it showed a twofold increase in trypsin inhibition and a substantial decrease in the inhibition of hydrophobic S1-specific SPs of subtilisin type (subtilisin Carlsberg and proteinase K) and chymotrypsin type (chymotrypsin and elastase), i.e. we were able to alter the specificity of sermetstatin. In addition, we also replaced positions P1' and P2' of sermetstatin with the residues found in SSI (mutant Y⁷²V+F⁷³Y; see Fig. 1a) to assess the importance of these sites but could not observe any significant inhibitory difference with the wild-type (Suppl. Table S2).

Structure of the snapalysin: sermetstatin complex

The crystal structure of the complex between *S. caespitosus* snapalysin and sermetstatin was determined by multi-step Patterson search and also shows a heterotetrameric 2+2 quaternary structure (Fig. 2c and Suppl. Table S3), again in accordance with results in solution. Three such structurally-

flexibility at the inter-protein junction as seen upon superposition of the distinct complexes. The interaction between snapalysin (chain A) and sermetstatin (chain B) buries an interface of ~870Å² (~13% of the total surface of both enzyme and inhibitor). It includes 43 close contacts, among them 11 hydrogen bonds, and hydrophobic interactions between eight peptidase and seven inhibitor residues. Segments involved include the N-terminal tail (S¹-P⁵), reactive-site loop 1, helix α1 (at R⁴³-N⁵⁰), and Lβ4β5 (at N⁸⁵-G⁸⁶) of the inhibitor; and R₅₁-D₅₈, Q₇₁-Y₇₅, R₇₉-H₈₇, D₉₃-G₉₇, and G₁₀₅-P₁₀₈ of the proteinase moiety. This implies that not only the active-site cleft is targeted during inhibition but also exosites on the protease surface, and this, in turn, explains the specificity of this second Janus face for snapalysin.

The present inhibition modus follows a novel mechanism for MP inhibitors: the N-terminal segment is inserted in a substrate-like manner in canonical conformation into the non-primed side of the active-site cleft, thus establishing a β-ribbon-type interaction with the β-strand above the active-site cleft at A₅₅-D₅₈ (see Fig. 4). The side chain of H³ is directed toward the peptidase moiety and approaches the zinc-bound catalytic solvent through a solvent bridge. To assess this interaction, we constructed two point mutants, H³R and H³E, which showed inhibitory activity indistinguishable from that of the wild type (data not shown), which is logical as the structure should be able to accommodate both changes simply through removal of the bridging solvent. The second main interacting region for inhibition between snapalysin and sermetstatin is provided by reactive-site loop 1, which penetrates cleft sub-site S1' with Y³³, thus matching the substrate specificity of the MP^{26, 27} (Fig. 4). The aromatic side chain approaches R₇₉, which creates the bottom of the hydrophobic S1' pocket of the enzyme through the aliphatic part of its side chain. Due to the intrusion of reactive-site loop 1, Y₉₅, which is in a position indistinguishable from that seen in the unbound structure²⁶ and is generally supposed to stabilize the

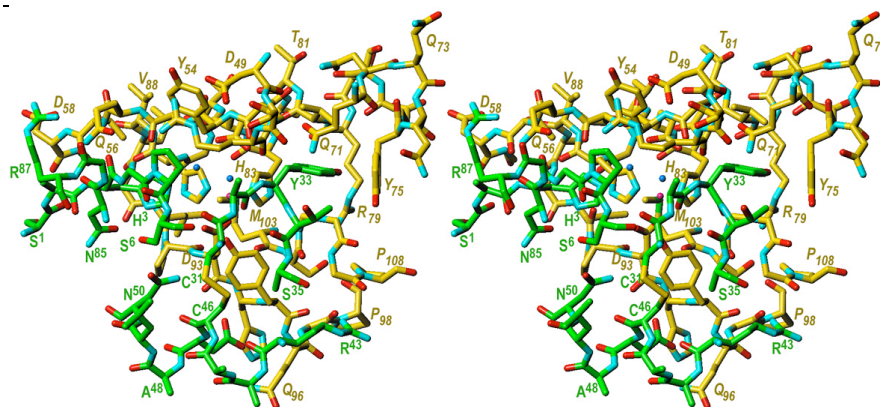


Figure 4. Detail in cross-eye stereo of the interaction of sermetstatin (chain B; green carbon atoms and residue labels) and snapalysin (chain A; yellow carbons and ochre residue labels). The peptidase is shown in standard orientation and its active-site zinc ion is depicted as a small magenta sphere. The two solvent molecules found in the active-site cleft are shown as blue spheres: one is the catalytic solvent molecule attached to the zinc ion (right) and the second one bridges H^{1'}N_{ε2} of sermetstatin and the former solvent molecule (left). Selected residues of either polypeptide chain are labeled for reference. The interaction of the snapalysin bulge segment (R₅₁-D₅₈) with regions of the second inhibitor molecule within the heterotetrameric 2+2 complex (chain D) is not shown for clarity.

tetrahedral reaction intermediate during catalysis¹², contacts the main-chain carbonyl of the loop at A³² through its side chain. To assess the importance of this loop *in vitro*, we constructed a double point mutant, Y³³P+T³⁴G, which mimics the sequence of SSI at this site (Fig. 1a), and found only residual inhibitory activity against snapalysin, while the behavior against SPs was unaltered (Suppl. Table S2). This supports an authentic key role in inhibition for this second Janus-face site, which is absent in SSI and, most likely, other I16 inhibitors studied, which are evolutionarily more distant from sermetstatin than from each other (Fig. 1a). In addition, secondary snapalysin:sermetstatin interaction sites are provided by sermetstatin helix α1, which interacts with the segment immediately downstream of the third zinc-binding residue of snapalysin (D₉₃-G₉₇), and Lβ4β5, which contacts the upper-rim strand of the peptidase on its non-primed side (Fig. 4). In particular, D⁴⁷ of α1 approaches the main-chain amide nitrogen atoms of both Y₉₅ and Q₉₆ through its side chain. This causes the side chain of the latter residue to be pulled out towards bulk solvent in the complex. Finally, the frontal bulge segment above the primed-side of the active-site cleft of snapalysin, R₅₁-D₅₈, interacts with three short segments of the second sermetstatin molecule (chain D) within the heterotetramer at V²³-T²⁸, G³⁷-P⁴⁰, and F¹¹³ through four hydrogen bonds. Overall, this type of enzyme:inhibitor interaction is only distantly reminiscent of the inhibitory mode of tissue inhibitors of metalloproteinases on their target matrix metalloproteinases²⁸ and of serralyisin inhibitors on their cognate serralyisin MPs²⁹. In these cases, N-terminal tails likewise bind in extended conformation but are found in the primed side of the cleft.

30 Disulfide bonds in sermetstatin

Disulfide bonds are a common feature of proteinase inhibitors that follow the standard mechanism, and we set out to assess the importance of the two that link the reactive-site loops to their subjacent scaffold helices by mutating the cysteine residues to serine. Wild-type sermetstatin, as well as mutants C³¹S+C⁴⁶S and C⁶⁹S+C⁹⁹S, were incubated with either subtilisin or snapalysin. While the wild-type protein was not cleaved by either peptidase, mutant C⁶⁹S+C⁹⁹S, which was obtained with similar yields to the wild-type protein (see ESI), was completely degraded by subtilisin but not snapalysin. This suggested that the mutant inhibitor still had sufficient structural integrity to maintain its MP site functional but not to prevent the SP site from collapsing. By contrast, mutant C³¹S+C⁴⁶S, which was obtained with a much lower yield than the wild-type, was readily degraded by both

peptidases, strongly suggesting that this disulfide is essential for the overall structural conformation of the inhibitor. Accordingly, the disulfide bonds of sermetstatin are required for its double inhibitory activity because they reduce the conformational flexibility of the main chain around the reactive site and contribute to maintaining the overall scaffold in a competent conformation.

55 Model for the ternary complex

All the structural information described has allowed us to construct a composite model of the ternary complex between subtilisin, sermetstatin, and snapalysin (Fig. 2d). This model shows a large inhibitory particle of maximal dimensions 120x85x65Å, which is rather asymmetric and resembles a distorted cross with a large vertical bar arched by ~55° and a short horizontal bar, which intersect at an angle of ~75°. The model confirms the heterohexameric 2+2+2 arrangement found in solution (see ESI and Suppl. Fig. S1); no steric hindrance is observed upon peptidase binding to the distal inhibitory sites, and the proteinase surfaces are at least ~15Å apart.

In summary, sermetstatin is a genuine Janus-faced dimeric single-domain inhibitor which has evolved from single-site standard-mechanism SP inhibitors of family I16 to give a protein capable of simultaneous inhibition of SPs in general and a specific MP through distinct but compatible sites.

Notes and references

- ⁷⁵ *a* Proteolysis Lab; Molecular Biology Institute of Barcelona; Spanish Research Council CSIC; Barcelona Science Park; c/ Baldiri Reixac, 15-21; 08028 Barcelona, Catalonia (Spain).
- b* Department of Physical Chemistry, Biochemistry, and Inorganic Chemistry; University of Almería; 04210 Almería (Spain).
- ⁸⁰ * Corresponding authors. E-mail: xgrcri@ibmb.csic.es and jlaeri@ibmb.csic.es; Fax: +34 934034979; Tel: +34 934020186.
- † Electronic Supplementary Information (ESI) available: Experimental procedures, acknowledgements, supplemental references, three supplemental tables and one supplemental figure. See DOI: 10.1039/b000000x/
1. W. Bode and R. Huber, *Biomed Biochim Acta*, 1991, **50**, 437-446.
2. W. Bode and R. Huber, *Eur. J. Biochem.*, 1992, **204**, 433-451.
- ⁹⁰ 3. R. Bao, C. Z. Zhou, C. Jiang, S. X. Lin, C. W. Chi and Y. Chen, *The Journal of biological chemistry*, 2009, **284**, 26676-26684.
4. Z. Zhou, J. Gu, Y. L. Du, Y. Q. Li and Y. Wang, *Curr. Genomics*, 2011, **12**, 404-416.
5. N. D. Rawlings, A. J. Barrett and A. Bateman, *Nucleic Acids Res.*, 2012, **40**, D343-D350.
- ⁹⁵ 6. T. Kumazaki, K. Kajiwara, S. Kojima, K. Miura and S. Ishii, *J. Biochem.*, 1993, **114**, 570-575.
7. M. Uyeda, S. Yoshida, K. Suzuki and M. Shibata, *Agr. Biol. Chem.*, 1976, **40**, 1237-1238.
- ¹⁰⁰ 8. K. Hiraga, T. Suzuki and K. Oda, *J. Biol. Chem.*, 2000, **275**, 25173-25179.
9. S. Murao and S. Sato, *Agr. Biol. Chem.*, 1972, **36**, 160-163.

-
10. J. E. Strickler, T. R. Berka, J. Gorniak, J. Fornwald, R. Keys, J. J. Rowland, M. Rosenberg and D. P. Taylor, *J. Biol. Chem.*, 1992, **267**, 3236-3241.
 11. A. Kakinuma, H. Sugino, N. Moriya and M. Isono, *J. Biol. Chem.*, 1978, **253**, 1529-1537.
 12. F. X. Gomis-Rüth, *J. Biol. Chem.*, 2009, **284**, 15353-15357.
 13. W. Bode, F. X. Gomis-Rüth and W. Stöcker, *FEBS Lett.*, 1993, **331**, 134-140.
 14. W. Stöcker, F. Grams, U. Baumann, P. Reinemer, F. X. Gomis-Rüth, D. B. McKay and W. Bode, *Prot. Sci.*, 1995, **4**, 823-840.
 15. C. Tallant, R. Garcia-Castellanos, J. Seco, U. Baumann and F. X. Gomis-Rüth, *J. Biol. Chem.*, 2006, **281**, 17920-17928.
 16. B. W. Matthews, *Acc. Chem. Res.*, 1988, **21**, 333-340.
 17. F. X. Gomis-Rüth, T. O. Botelho and W. Bode, *Biochim. Biophys. Acta*, 2012, **1824**, 157-163.
 18. Y. Mitsui, Y. Satow, T. Sakamaki and Y. Iitaka, *J. Biochem.*, 1977, **82**, 295-298.
 19. Y. Takeuchi, Y. Satow, K. T. Nakamura and Y. Mitsui, *J. Mol. Biol.*, 1991, **221**, 309-325.
 20. Y. Takeuchi, T. Nonaka, K. T. Nakamura, S. Kojima, K. Miura and Y. Mitsui, *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 4407-4411.
 21. N. Kamiya, M. Matsushima and H. Sugino, *Bull. Chem. Soc. Jpn.*, 1984, **57**, 2075-2081.
 22. M. Laskowski Jr. and I. Kato, *Annu. Rev. Biochem.*, 1980, **49**, 593-626.
 23. J. L. Arolas, T. O. Botelho, A. Vilcinskas and F. X. Gomis-Rüth, *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 10357-10360.
 24. I. Schechter and A. Berger, *Biochem. Biophys. Res. Commun.*, 1967, **27**, 157-162.
 25. Y. Takeuchi, S. Noguchi, Y. Satow, S. Kojima, I. Kumagai, K. Miura, K. T. Nakamura and Y. Mitsui, *Protein Eng.*, 1991, **4**, 501-508.
 26. G. Kurisu, T. Kinoshita, A. Sugimoto, A. Nagara, Y. Kai, N. Kasai and S. Harada, *J. Biochem.*, 1997, **121**, 304-308.
 27. S. Harada, T. Kinoshita, N. Kasai, S. Tsunasawa and F. Sakiyama, *Eur. J. Biochem.*, 1995, **233**, 683-686.
 28. F. X. Gomis-Rüth, K. Maskos, M. Betz, A. Bergner, R. Huber, K. Suzuki, N. Yoshida, H. Nagase, K. Brew, G. P. Bourenkov, H. Bartunik and W. Bode, *Nature*, 1997, **389**, 77-81.
 29. U. Baumann, M. Bauer, S. Letoffe, P. Delepelaire and C. Wandersman, *J. Mol. Biol.*, 1995, **248**, 653-661.

ELECTRONIC SUPPLEMENTARY INFORMATION

Table of Contents

1. Experimental procedures	1
<i>Production and purification of sermetstatin</i>	1
<i>Production and purification of snapalysin</i>	2
<i>In vitro activation studies of prosnapalysin</i>	3
<i>Proteolytic and inhibition assays</i>	4
<i>Cleavage of sermetstatin mutants</i>	4
<i>Complex formation and purification</i>	5
<i>Crystallization and X-ray diffraction data collection</i>	5
<i>Structure solution and refinement</i>	6
<i>Miscellaneous</i>	8
2. Acknowledgments	8
3. Supplemental References	9
4. Supplemental Tables	11
<i>Supplemental Table S1</i>	11
<i>Supplemental Table S2</i>	12
<i>Supplemental Table S3</i>	13
5. Supplemental Figures	14
<i>Supplemental Figure S1</i>	14

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 inhibitor¹, 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-His₆ 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 OD₅₅₀ 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 7,000xg 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 4°C using a cell disrupter (Constant Systems) at a pressure of 1.35Kbar, and the cell debris was removed by centrifugation at 50,000xg for 1h at 4°C. The

supernatant was filtered (0.22 μ m 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 His₆-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 His₆-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, H³E, H³R, Y³³P+T³⁴G, M⁷¹K, Y⁷²V+F⁷³Y, C³¹S+C⁴⁶S, and C⁶⁹S+C⁹⁹S, 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 C³¹S+C⁴⁶S, 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 His₆ 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 30 μ g/ml kanamycin, induced at an OD₅₅₀ of 0.8 with 1mM IPTG, and grown for a further 5h at 37°C. After centrifugation at 7,000xg 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 4°C 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 7,000xg 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 50,000xg 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 CaCl₂, 1mM ZnSO₄, 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+M₁* (italicized snapalysin amino acids subscripted numbering correspond to UniProt P56406 *plus* the additional N-terminal tag residues and the *T₁M* 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 OD₅₅₀ of 0.8 with IPTG to a final concentration of 1mM and grown for a further 5h at 37°C. After centrifugation at 7,000xg 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 4°C 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 7,000xg 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 50,000xg 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 CaCl₂, 1mM ZnSO₄, 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 fragilyisin-3 from *Bacteroides fragilis* were produced and purified as previously described²⁻⁵. 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 published⁶. 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 $\lambda_{ex} = 485\text{nm}$ / $\lambda_{em} = 528\text{nm}$. 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 C³¹S+C⁴⁶S and C⁶⁹S+C⁹⁹S (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 CaCl₂, 1.1mM ZnSO₄, 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 4°C 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-N₂ 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 ($V_M = 3.2\text{\AA}^3/\text{Da}$; 62% solvent content). Crystals of subtilisin:sermetstatin were monoclinic, with two complexes per asymmetric unit ($V_M = 3.6\text{\AA}^3/\text{Da}$; 65% solvent content). Crystals of snapalysin:sermetstatin were orthorhombic with six complexes per asymmetric unit ($V_M = 2.9\text{\AA}^3/\text{Da}$; 58% solvent content). Diffraction data were integrated, scaled, merged, and reduced with programs XDS⁷, XSCALE, and SCALA⁸ within the CCP4 suite⁹ (see Suppl. Table S3).

Structure solution and refinement – All attempts to solve the structure of unbound sermetstatin by Patterson search¹⁰ 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 SHELXD¹². 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 SHELXE¹² 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 P3₁21 as the correct space group. An initial model was built with program COOT¹³ and refined against the selenomethionine dataset with program REFMAC5¹⁴. The

resulting partial model was used to determine the native structure by Patterson-search methods with program PHASER¹⁵. 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-Frodo¹⁶ alternated with crystallographic refinement (including TLS refinement) with programs REFMAC5 and BUSTER/TNT¹⁷ until completion of the model (Suppl. Table S3). The final model of sermetstatin comprised residues G⁻¹+S¹-F¹¹³ (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 A₁-Q₂₇₅ (subscripted subtilisin residue numbers according to the mature enzyme sequence; see UniProt P00780) and two calcium ions (Ca₉₉₉ and Ca₉₉₈) for the two subtilisin moieties (chains A and C), G⁻¹+S¹-F¹¹³ for sermetstatin chain B, and G⁴-F¹¹³ 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+M₁-G₁₃₂ in chains A, C, and I; residues P₋₁+M₁-G₁₃₂ in chains G and K; and residues M₁-G₁₃₂ in chain E. In addition, sermetstatin chains spanned residues S¹-F¹¹³ (chains B and L), A²-F¹¹³ (chain F), S¹-A⁶²+L⁶⁷-F¹¹³ (chain D), S¹-L⁶⁰+L⁶⁷-F¹¹³ (chain H), and S¹-P⁶⁵+V⁶⁸-F¹¹³ (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 than the other ten chains; this is indicated by the significantly higher average thermal displacement parameters for their protein parts (76.0 and 90.3 Å², respectively) than the remaining molecules (44.8-64.5 Å²).

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, L β 2 β 3, and segment β 1-L β 1 β 2- β 2 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 D Biol Crystallogr*, 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 D Biol Crystallogr*, 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.
52. F. X. Gomis-Rüth, T. O. Botelho and W. Bode, *Biochim. Biophys. Acta*, 2012, **1824**, 157-163.

4. Supplemental Tables

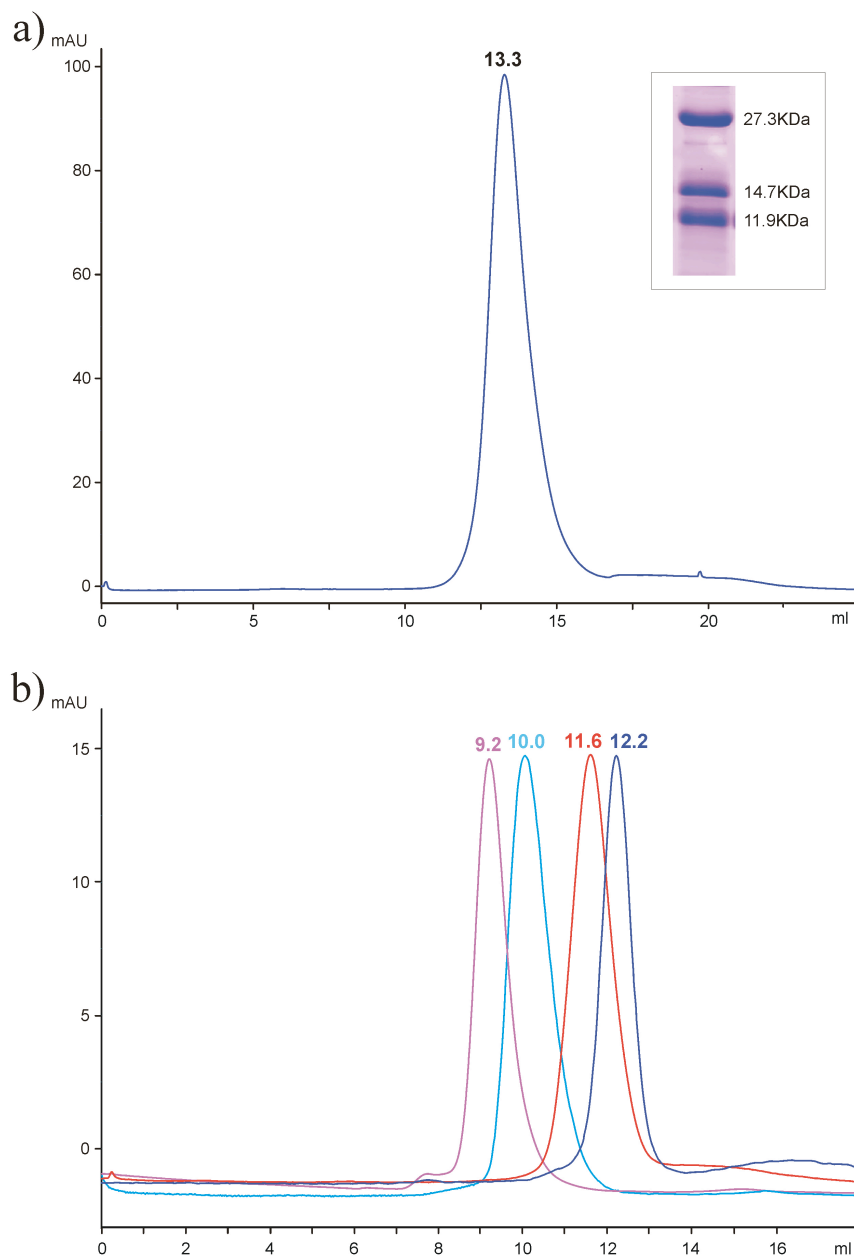
Suppl. Table S1. Janus-faced single-domain inhibitors with at least two reactive sites.								
Family (MEROPS)	Inhibitor name	Source	UniProt code	Class of targets	Targets	PDB code	Ref.	
Kunitz (I3)	ApKTI	<i>Adenanthera pavonina</i>	P09941	SP ^a (and CP ^a)	Trypsin and papain (very weakly)	-	26	
	PKPI	<i>Solanum tuberosum</i>	Q66LL2	SP	Trypsin and chymotrypsin	3TC2	27, 28	
	API	<i>Sagittaria sagittifolia</i>	Q7M1P4	SP	Trypsin, chymotrypsin and kallikrein	3E8L ^b	29, 30	
	PPI	<i>Carica papaya</i>	P80691	SP	Trypsin	3S8K, 3S8J	31	
	WCI	<i>Psophocarpus tetragonolobus</i>	P10822	SP	Chymotrypsin	2WBC	32	
Bowman-Birk (I12)	MSTI	<i>Medicago scutellata</i>	P80321	SP	Trypsin	2ILN ^b	33	
	sBBI	<i>Glycine max</i>	P01055	SP	Trypsin and chymotrypsin	1PI2, 1D6R ^b	34, 35	
	LCTI	<i>Lens culinaris</i>	Q8W4Y8	SP	Trypsin and chymotrypsin	2AIH	36	
	BTCl	<i>Vigna unguiculata</i>	P17734	SP	Trypsin and chymotrypsin	2G81 ^c , 2OT6	37, 38	
	HGI-III	<i>Dolichos biflorus</i>	Q9S9E3	SP	Trypsin and chymotrypsin	-	39	
	CLTI	<i>Canavalia lineata</i>	Q7M1Q0	SP	Trypsin and chymotrypsin	-	40	
	A-II	<i>Arachis hypogaea</i>	P01066	SP	Trypsin and chymotrypsin	Not deposited	41, 42	
	TcTI	<i>Amburana acreana</i>	P83284	SP	Trypsin, plasmin and chymotrypsin	-	43	
	I-2B	<i>Triticum aestivum</i>	P09863	SP	Trypsin	-	44	
	II-4	<i>Triticum aestivum</i>	P09864	SP	Trypsin	-	44	
	LBI	<i>Phaseolus lunatus</i>	P01056	SP	Trypsin and chymotrypsin	1H34	45, 46	
	GBI	<i>Phaseolus vulgaris</i>	P01060	SP	Trypsin, chymotrypsin and elastase	-	47	
	Clitocypin (I48)	Clitocypin	<i>Clitocybe nebularis</i>	Q3Y9I4	CP	Papain and cathepsins L, V, S and K	3H6R, 3H6S ^d	48
	Macrocyprin (I85)	Macrocyprin 1	<i>Macrolepiota procera</i>	B9V973	CP (and SP)	Papain and cathepsins L, V, S and K (Macrocyprin 4 inhibits trypsin weakly)	3H6Q	48
	Others	Trypsin inhibitor	<i>Trichosanthes kirilowii</i>	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 (%)
<i>Metallopeptidases</i>					
Snalysin	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
<i>Serine proteinases</i>					
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 μ M inhibitor concentration in a cell-based assay.					

Suppl. Table S3. Crystallographic data.				
	Sermetstatin		Subtilisin:sermetstatin	Snalypalin:sermetstatin
Dataset	Native	Selenomethionine (absorption peak) ^a	Native	Native
Space group	P3 ₁ 21	P3 ₁ 21	C2	P2 ₁ 2 ₁ 2 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 (0.217)	0.195 (0.242)
No. of protein atoms / solvent molecules / neutral ligands / ionic ligands	834 / 126 / 1 glycerol / 3 CH ₃ COO ⁻		5,483 / 305 / 1 diglycerol, 3 2-propanol, 2 glycerol / 4 Ca ²⁺ , 2 Zn ²⁺ , 2 K ⁺ , 1 Cl ⁻ , 1 (CH ₃) ₂ AsOO ⁻ , 1 CH ₃ COO ⁻ , 1 PO ₄ ³⁻	11,043 / 275 / 2 glycerol / 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		744 / 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).