Supporting Information

SI Material and Methods

Cells. PC3-M cells were kindly provided by Dr. Raymond Bergan, Northwestern University, Robert H Lurie Medical Research Center, Chicago, IL, USA. PC3-M cells were maintained in RPMI 1640 (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen).

Reagents. Synthetic oligonucleotides were obtained from Integrated DNA Technologies. Restriction enzymes and polymerases were purchased from New England Biolabs, and dNTPs, from Jena Bioscience. EBY100 yeast strain and yeast surface display (YSD) plasmid (pCTCON) were gifts from the laboratory of Dane Wittrup (MIT). Bacterial plasmid extraction and purification kits were obtained from RBC Bioscience, and yeast plasmid extraction kits, from Zymo Research. The methylotrophic yeast *P. pastoris* strain GS115, *Pichia* expression vector (pPIC9K), and fluorescein (FITC)-conjugated streptavidin were obtained from Invitrogen. Bovine trypsin, phycoerythrin (PE)-conjugated anti mouse antibody, and the substrates benzyloxycarbonyl-Gly-Pro-Arg-*p*-nitroanilide (*Z*-GPR-pNA), 4-nitrophenyl 4-guanidinobenzoate (pNPGB), and benzoyl-L-arginine-*p*-nitroanilide (L-BAPA) were obtained from Sigma-Aldrich. Mouse anti-c-Myc antibody (Ab-9E10) was obtained from Abcam. EZ-Link NHS-PEG4 biotinylation kit was purchased from ThermoFisher Scientific. Factor-XIa and its substrate S-2366 (Chromogenix) were obtained from Hematologic Technologies Inc. and Diapharma, respectively. Kallikrein-6 was a gift from the laboratory of Aubry Miller [German Cancer Research Center (DKFZ)], and its substrate BOC-Phe-Ser-Arg-AMC was obtained from Bachem. Lentiviral short hairpin RNA construct NM_002771.2-454s1c1 targeting human *PRSS3* was from Sigma-Aldrich. TaqMan™ assays Hs00605637_m1 detecting human *PRSS3* and Hs99999905_m1 for normalization against human GAPDH were from Applied Biosystems-Thermofisher.

Synthesis and Cloning of the DNA Encoding APPI$_{WT}$. The inhibitor domain of the amyloid precursor protein (APPI$_{WT}$) gene was constructed based on a published sequence (PDB id 1ZJD) by using codons optimized for both *S. cerevisiae* and *P.*

pastoris usage. The DNA sequence of $APPI_{WT}$ attached to a peptide linker (NH_3^+-$APPI_{WT}$-LPDKPLAFQDPS-COO^-) was generated by PCR-assembly using the following six overlapping oligonucleotides (5'-3'):

Oligo*1*:GATGGTATTTCGATGTTACTGAAGGTAAATGTGCTCCATTCTTCTATG GTGGTTGTGGTG;**Oligo*1′***:CCACAAACAGCCATACAATATTCTTCAGTATCGA AATTATTTCTATTACCACCACAACCACCATAGAAGAAT;**Oligo*2***:GAAGTTTGT TCTGAACAAGCTGAAACTGGTCCATGTAGAGCTATGATTTCTAGATGGTATT TCGATGTTACTG;**Oligo*2′***:GGAAAGCCAATGGTTTATCTGGCAAGGATCCAAT AGCAGAACCACAAACAGCCATACAATATTC;**Oligo*3***:GGTGGTTCTGGTGGTG GTGGTTCTGGTGGTGGTGGTCTGCTAGCGAAGTTTGTTCTGAACAAGCTG;**Ol igo*3′***:GAGCTATTACAAGTCCTCTTCAGAAATAAGCTTTTGTTCAGATGGATCT TGGAAAGCCAATGGTTTATC.

The synthetic insert gene was assembled by a set of three PCRs using Phusion DNA polymerase, with each paired reaction (**Oligo*X*/*X′***) serving as a template for the following reaction.

DNA sequence of APPI variants (caring combination of mutations) were assembled by the same methodology, but with oligonucleotides containing the respective mutations:

Oligo*2*^{a}:GAAGTTTGTTCTGAACAAGCTGAAACTGGTCCATGTAGAGCTGGTTT TTCTAGATGGTATTTCGATGTTACTG;**Oligo*2*^{b}**:GAAGTTTGTTCTGAACAAGCT GAAACTGGTCCATGTAGAGCTGGTATTTCTAGATGGTATTTCGATGTTACTG; **Oligo*2*^{c}**:GAAGTTTGTTCTGAACAAGCTGAAACTGGTCCATGTAGAGCTATGTT TTCTAGATGGTATTTCGATGTTACTG;**Oligo*1′*^{a}**:CCACAAACAGCCATACAATA TTCTTCAGTATCGAAATTATTTCTATTACCACCACAACCACCATAGACGAAT G were **Oligo*2*^{x}** replaces the original **Oligo*2*** and **Oligo*1′*^{a}** replaces the original **Oligo*1′*** in the respective PCR step.

The final PCR assembled fragment was gel-purified and cloned into the YSD vector (pCTCON; Fig. S1) via transformation of EBY100 yeast cells with a linearized vector (digested with *Nhe*I and *Bam*HI) and the PCR product, as previously described [1]. This simultaneous cloning and transformation occurs via the *in vivo* homologous recombination between the vector and the PCR insert to generate the YSD plasmid [1]. After sequence verification (DMSU, NIBN, BGU), the DNA construct of $APPI_{WT}$ served as the template for combinatorial library generation.

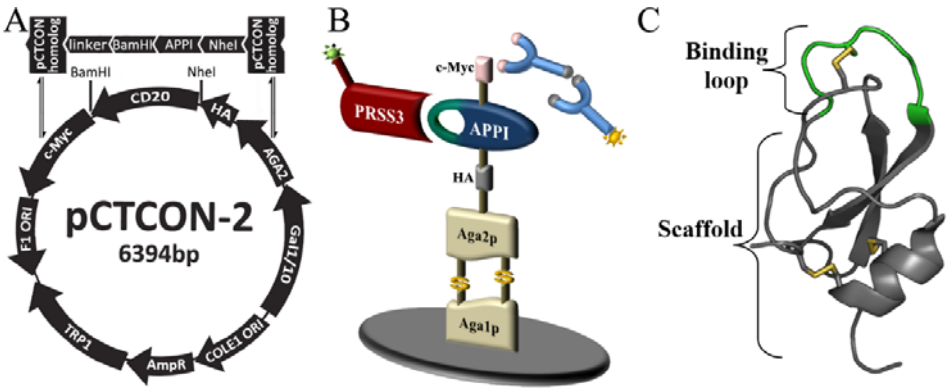

Fig. S1. YSD construct. (*A*) The YSD vector (pCTCON-2) is aligned with the general scheme of the insert: the insert gene consists of the APPI gene, which is flanked by two restriction sites (*Bam*HI and *Nhe*I) followed by a linker sequence (LPDKPLAFQDPS) on the 3' along with two pCTCON homologous sequences. (*B*) APPI is displayed on the yeast cell surface as a translational fusion to Aga2p that is linked to Aga1p by two disulfide bonds. Surface expression is detected using fluorescence-labeled antibodies binding to the C-terminal c-Myc epitope tag, while target binding is detected using fluorescence-labeled mesotrypsin. (*C*) 3-D structure of APPI (PDB id 1ZJD) shows a protease-inhibitory binding loop of characteristic canonical conformation (green), supported by a compact scaffold (gray) with a three-dimensional structure that is maintained by hydrophobic packing and three disulfide bonds (yellow).

Generation of combinatorial APPI library. After assembly and cloning of $APPI_{WT}$, the plasmid construct served as the template for the subsequent generation of the combinatorial library by using error-prone PCR. To generate a mutation frequency of ~3 mutations per clone, the PCR reaction was optimized to 15× PCR doublings of the 300-bp APPI fragment (including plasmid homologue regions) with low-fidelity Taq polymerase [2], 1% nucleotide analogues and 2 mM $MgCl_2$. The resulting mutated insert was amplified and transformed into yeast through homologous recombination to generate a library of about 9×10^6 in size, as estimated by dilution plating on selective SDCAA plates (15% agar, 2% dextrose, 1.47% sodium citrate, 0.429% citric acid monohydrate, 0.67% yeast nitrogen base and 0.5% casamino acids). Sequencing results (DMSU, NIBN, BGU) revealed an average mutation rate of 0-3 mutations per 300 bp (data not shown).

Production of Recombinant Trypsins. Recombinant human anionic trypsinogen, human cationic trypsinogen and human mesotrypsinogen and the catalytically inactive S195A mutant of mesotrypsinogen were expressed in *E. coli*, extracted from inclusion bodies, refolded, purified and activated with bovine enteropeptidase, as described in

previous work [3, 4]. Mesotrypsin and mesotrypsin-S195A were biotinylated for use in YSD screens, and biotin incorporation was quantified in a 4'-hydroxyazobenzene-2-carboxylic acid (HABA) assay, using the EZ-Link NHS-PEG4 biotinylation kit (ThermoFisher Scientific) according to manufacturer's instructions.

Construction and Cloning of the Expression Vector pPIC9K-APPI. The human cDNA of $APPI_{WT}$ was amplified by PCR using Phusion DNA polymerase with an upstream primer: 5′-AGCGTATACGTAGACTATAAGGATGACGAC GACAAAGAATTCGAAGTTTGTTCTGAACAAGCTG-3′ and a downstream primer: 5′-ATAGTTTAGCGGCCGCATGA TGGTGGTGATGGTGCCTAGGAATAGCAGAACCACAAACAGC-3′. The resulting construct included four restriction sites and two epitope tags (FLAG and HIS_6) as follows: *SnaBI*-FLAG-*EcoRI*-$APPI_{WT}$-*AvrII*-HIS_6-*NotI*. The obtained DNA fragment was digested with *SnaBI* and *NotI* and subcloned by using the same restriction sites of *Pichia* expression vector pPIC9K by standard methods (Fig. S2). Next, the recombinant expression plasmid was used as a template for the construction of the APPI variants as follows: cDNA of each variant was amplified by PCR with an upstream primer: 5′-CGGAGCGAATTCGAAGTTTGTTCTGAACAAGCTG-3′ and a downstream primer: 5′- CGCTACCCTAGGAATAGCA GAACCACAAACAGC-3′. The resulting construct included the restriction sites *EcoRI* and *AvrII*. The obtained DNA fragment was digested with *EcoRI* and *AvrII* and subcloned using the same restriction sites of the template vector. Finally, all the sequences of the recombinant expression plasmids were confirmed by DNA sequencing analysis (DMSU, NIBN, BGU, data not shown).

Expression vectors were linearized by *SacI* digestion and used to transform *P. pastoris* strain GS115 by electroporation. This resulted in insertion of the construct at the AOX1 (alcohol oxidase) locus of *P. pastoris*, thereby generating a His^+ Mut^+ phenotype. Transformants were selected for the His^+ phenotype on 2% agar containing regeneration dextrose biotin (RDB; 18.6% sorbitol, 2% dextrose, 1.34% yeast nitrogen base, 4×10^{-5} % biotin, and 0.005% each of L-glutamic acid, L-methionine, L-lysine, L-leucine, and L-isoleucine) and allowed to grow for 2 d at 30 °C. Cells were harvested from the plates and subjected to further selection for high copy number by their ability to grow on 2% agar containing 1% yeast extract, 2% peptone, 2% dextrose medium, and the antibiotic G418 (Geneticin, 4 mg/ml, Invitrogen).

To verify direct insertion of the construct at the AOX1 locus of *P. pastoris*, the genomic DNA of the highest APPI-expressing colony from each APPI variant was extracted [5] and amplified by PCR with an AOX1 upstream primer: 5′-GACTGGTTCCAATTGACAAGC-3′ and an AOX1 downstream primer: 5′-GCAAATGGCATTCTGACATCC-3′. The resulting linear DNA was gel purified, and its correct sequence was confirmed by DNA sequencing analysis (DMSU, NIBN, BGU, data not shown).

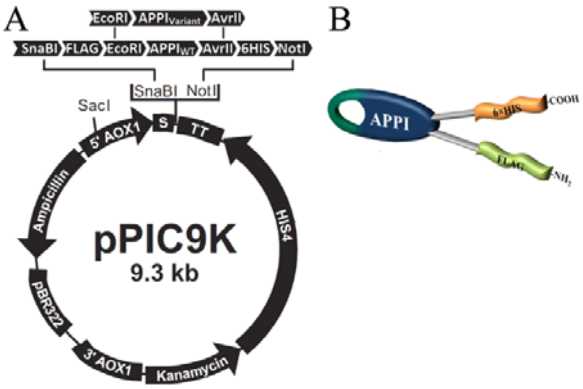

Fig. S2. Construction of the expression vector pPIC9K-APPI. (*A*) *Pichia* expression vector (pPIC9K) is aligned with the general scheme of the insert: *SnaBI*-FLAG-*EcoRI*-$APPI_{WT}$-*AvrII*-HIS_6-*NotI*. (*B*) APPI is expressed as a translational fusion to N-terminal FLAG and C-terminal HIS_6.

Large-Scale Expression and Purification of APPI. GS115-APPI clones were first inoculated into 50 mL of BMGY (1% yeast extract, 2% peptone, 0.23% potassium phosphate monobasic, 1.18% potassium phosphate dibasic, 1.34% yeast nitrogen base, 4×10^{-5} percent biotin and 1% glycerol) to an OD_{600} = 10.0, followed by scale up to 500 mL of BMGY until OD_{600}=10.0 was reached (overnight growth at 30 °C with shaking at 300 rpm). Cells were harvested by centrifugation and resuspended in 1 L BMMY (as for BMGY, but with 0.5% methanol instead of glycerol) to an OD_{600} of 5 to induce expression, and grown at 30°C with shaking at 300 rpm. Methanol was added to a final concentration of 2% every 24 h to maintain induction. Following 5 d of induction, the culture was centrifuged again, and the supernatant containing the secreted recombinant inhibitors was prepared for purification by nickel-immobilized metal affinity chromatography (IMAC).

The supernatant containing the recombinant APPI was filtered through a 0.22-μm Steritop bottle-top filter (Millipore). The filtered supernatant was adjusted to 10 mM imidazole and 0.5 M NaCl at pH 8.0 and left to stand overnight at 4°C. Thereafter, a

second filtration was performed to remove any additional precipitation. The resulting supernatant was loaded on a HisTrap 5-ml column (GE Healthcare) at a flow rate of 0.7 ml/min for 24 h, washed with washing buffer [20 mM sodium phosphate, 0.5 M NaCl, and 10 mM imidazole (pH 8.0)], and eluted with elution buffer (as for washing buffer, but with 0.5 M imidazole) in an ÄKTA pure instrument (Fig. S4*A*). The eluted inhibitors were concentrated, and the buffer was replaced with TB in a 3-kDa molecular weight cutoff (MWCO) Vivaspin concentrator (GE Healthcare). Gel filtration chromatography was performed using a Superdex 75 16/600 column (GE Healthcare) equilibrated with TB at a flow rate of 1 ml/min on an ÄKTA pure instrument (Fig. S4*A* and *B*). Gel filtration protein fractions were analyzed by SDS-PAGE on a 15% polyacrylamide gel under non-reducing conditions and tested for their ability to inhibit bovine trypsin catalytic activity (see experimental details below and Fig. S3). The correct mass of the pure proteins was validated using a MALDI-TOF REFLEX-IV (Bruker) mass spectrometer (Ilse Katz Institute for Nanoscale Science & Technology, BGU; data not shown). Purification yields for all APPI variants were 5–20 mg per one-liter culture flask.

Inhibition Studies. The concentrations of mesotrypsin, cationic trypsin, anionic trypsin and bovine trypsin were quantified by active site titration using pNPGB, which serves as both an irreversible trypsin inhibitor and a substrate [6]. Concentrations of FXIa and kallikrein-6 were determined by UV-Vis absorbance at 280 nm, with extinction coefficients (ε_{280}) of 214.4×10^3 M^{-1} cm^{-1} and 34.67×10^3 M^{-1} cm^{-1}, respectively. Concentrations of the chromogenic substrates Z-GPR-pNA and S-2366 were determined in an end-point assay from the change in the absorbance (plateau after complete hydrolysis) that is obtained by the release of *p*-nitroaniline. Concentrations of APPI variants were determined by titration with pre-titrated bovine trypsin and the substrate L-BAPA, as previously described [4].

The dissociation constant K_i of $APPI_{WT}$ and its variants, $APPI_{M17G}$, $APPI_{I18F}$ and $APPI_{F34V}$, in complex with mesotrypsin were determined according to the previously described methodology, with minor changes [4]. Later, this methodology was adjusted for measuring the dissociation constant of $APPI_{M17G/I18F/F34V}$ in complex with FXIa. Briefly, stock solutions of enzyme, substrate, and APPI proteins were prepared at 40× the desired final concentrations. Assays were performed at 37 °C in the presence of different concentrations of substrate and inhibitor in a Synergy2 microplate

spectrophotometer (BioTek). The concentrations of reagents are given in Fig. 3*A* and *B* and in Fig. S7 *A-E*. Assay buffer (296 µl), substrate (8 µl), and APPI (8 µl) were mixed and equilibrated in a 96-well microplate (Greiner) prior to the addition of enzyme (8 µl from 10 nM mesotrypsin or 5 nM FXIa). Here, the term 'assay buffer' refers to TB or FXIa buffer (FB; 50 mM Tris-HCl, pH 7.6, 150 mM NaCl, 5 mM $CaCl_2$ and 0.1% BSA), whereas 'substrate' represents Z-GPR-pNA and S-2366 for mesotrypsin and FXIa, respectively. Reactions were followed spectroscopically for 5 min, and initial rates were determined from the increase in absorbance caused by the release of *p*-nitroaniline (ε_{410} = 8480 M^{-1} cm^{-1}). Data were globally fitted by multiple regression to Eq. 1, the classic competitive inhibition equation, using Prism (GraphPad Software, San Diego CA).

$$V = \frac{K_{cat}[E]_0[S]}{K_m(1+[I]/K_i)+[S]} \qquad \text{(Eq. 1)}$$

where V is the velocity of product formation at the start of the reaction; K_m (Michaelis constant) and K_{cat} are the kinetic parameters for substrate hydrolysis; $[E]_0$ and [I] are the total concentrations of enzyme and inhibitor, respectively, and [S] is the substrate concentration.

It should be noted that Eq. 1 assumes that the inhibitor concentration is not significantly reduced by its binding with the enzyme, thereby making the equation appropriate for measuring dissociation constants only for weak interactions. Although the dissociation constants that we calculated using Eq. 1 were relatively high (i.e., weak interactions; Table 1), we used inhibitor concentrations that were at least 10 times in excess of the enzyme concentrations, and therefore any reduction of the inhibitor concentration upon binding is negligible. Reported inhibition constants are average values obtained from three independent experiments.

Inhibition studies of (i) mesotrypsin with APPI variants M17G, M17G/I18F, M17G/F34V, I18F/F34V and M17G/I18F/F34V, (ii) cationic trypsin, anionic trypsin and kallikrein-6 with $APPI_{WT}$ and $APPI_{M17G/I18F/F34V}$, and (iii) FXIa with $APPI_{WT}$, were carried out in a similar manner, but the finding of slow, tight binding behavior in all three cases required a different kinetic treatment to that presented in Eq. 1. In tight binding kinetics, the reduction of the inhibitor concentration upon binding is significant (i.e., tight binding/strong interactions) and should be taken into consideration in the relevant equation (see equation 2 in the main paper, Fig. 3*C* and *F* and Fig. S7 *F-Q*). Briefly, tight binding experiments were conducted at a fixed concentration of Z-GPR-

pNA (145 μM); the inhibitor concentrations ranged between 5–80 nM, and the enzyme concentration (mesotrypsin, cationic trypsin or anionic trypsin) was 0.1 nM. Enzyme (8 µl), inhibitor (8 µl) and TB (144 µl) were preincubated at room temperature for 20-60 min; the reactions were then initiated by dilution of the enzyme/inhibitor mixture into a pre-equilibrated microplate (non-binding, 96 well; Greiner) containing TB (152 µl) and substrate (8 µl). The microplates were covered with lids and sealed with Parafilm to prevent evaporation. Reactions were run at 25 °C and were followed spectroscopically for 14 h so that reliable steady-state rates could be obtained. Conversion of substrate to product did not exceed 10% over the reaction time course.

Tight binding reactions of FXIa and kallikrein-6 were carried out in the same manner with minor changes, as follows: for FXIa, the substrate (S-2366) concentration was 600 μM, inhibitor concentrations ranged between 2–10 nM, enzyme concentration was 0.125 nM, assay buffer was FB, and the reactions were run at 37 °C and followed spectroscopically for 1 h. Reactions of kallikrein-6 were carried out at a fixed concentration of BOC-Phe–Ser–Arg-AMC (1 mM); inhibitor concentrations ranged between 5–50 nM, enzyme concentration was 1 nM, assay buffer was kallikrein buffer (KB; 50 mM Tris-HCl, pH 7.3, 100 mM NaCl and 0.2% BSA), and the reactions were run at 37 °C for 5 h and followed by fluorescent signal in a Tecan Infinite 200 PRO NanoQuant microplate reader set at 355 nm for excitation and 460 nm for emission.

Inhibition constants for tight binding reactions were calculated using Eq. 2, as described previously [4]:

$$\frac{(V_0 - V_i)}{V_i} = \frac{[I]_0}{K_i(1 + [S]_0/K_m)} \qquad \text{(Eq. 2)}$$

where V_i and V_0 are the steady-state rates in the presence and absence of inhibitor, K_m is the Michaelis constant for substrate cleavage, and $[S]_0$ and $[I]_0$ are the initial concentrations of substrate and inhibitor, respectively. Calculations were performed using K_m values of 24.66 ± 1.3 µM for mesotrypsin, 22.84 ± 1.9 µM for cationic trypsin, 10.69 ± 0.65 µM for anionic trypsin, 361.3 ± 12.1 µM for FXIa, and 329.3 ± 2.5 µM for kallikrein-6, as determined from at least three Michaelis–Menten kinetic experiments in our laboratory.

Hydrolysis Studies. The cleavage of intact APPI variants (between the residues Arg15-Ala16) in time course incubations with catalytically active mesotrypsin was monitored by HPLC as described previously, with minor modifications [3]. Briefly, mesotrypsin

was incubated with the APPI mutants in TB at 37 °C; inhibitor concentrations were 50 μM, and mesotrypsin concentrations were varied from 0.05 μM to 2.5 μM, as shown in Fig. S8. For HPLC analysis, aliquots of 30 μl were withdrawn from the hydrolysis reactions at periodic intervals (over 6 h), and samples were quenched immediately by acidification with 70 μl of 0.3 M HCl. Samples were resolved on a 50 × 2.0-mm Jupiter 4 μ 90-Å C_{12} column (Phenomenex) with a gradient of 0–100% acetonitrile in 0.1% trifluoroacetic acid (TFA) at a flow rate of 0.6 ml/min over 50 min. Intact inhibitors were quantified by peak integration of absorbance traces monitored at 210 nm. Initial rates were obtained by linear regression using a minimum of seven data points within the initial linear phase of the reaction (Fig. S8). Hydrolysis rates reported for each inhibitor represent the average of three independent experiments.

Bovine Trypsin Activity Assay. Assays were conducted at 37 °C in a Synergy2 plate reader spectrophotometer (BioTek). TB (185 μl), bovine trypsins (5 μl; 100 nM bovine trypsin), and APPI inhibitor (5 μl) were mixed and equilibrated prior to initiation of the reaction by the addition of the substrate, Z-GPR-pNA (5 μl; 1.5 mM). Reactions were followed spectroscopically for 5 min, and initial rates were determined from the increase in absorbance (410 nm) caused by the release of *p*-nitroaniline.

Far-UV Circular Dichroism Spectroscopy. CD spectra were recorded on a Jasco J-715 spectropolarimeter over a range of 190–260 nm using a quartz cuvette with a path length of 1 mm, a scanning speed of 50 nm/min and a data interval of 1 nm. Each sample of APPI variant was first analyzed at room temperature (20 ºC), and then left in the spectropolarimeter until 95 ºC was reached (denaturation); thereafter the sample was analyzed, cooled to 20 ºC (renaturation), and analyzed again. Proteins were analyzed in TB. Three scans of 50 μM protein solutions were averaged to obtain smooth data and background corrected with respect to protein-free buffer.

SI Results

Table S1. Sequencing analysis of randomly selected clones taken from S5 sort.

Region											Binding loop								β-strand	
											P5	P4	P3	P2	P1	P1'	P2'	P3'		
APPI			E	V	C	S	E	Q	A	E	T	G	P	C	R	A	**M**	**I**	S	R
Position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Mutation			G/Q												K		A	F		
Repeats			2												1		5	27		

Region									β-strand											
APPI	W	Y	F	D	V	T	E	G	K	C	A	P	F	**F**	Y	G	G	C	G	G
Position	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Mutation									R		V			V						
Repeats									1		1			5						

Region								α-helix											
APPI	N	R	N	N	F	D	T	E	E	Y	C	M	A	V	C	G	S	A	I
Position	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
Mutation									K	C		K							M
Repeats									1	1		1							1

* Numbers were adopted from the conventional BPTI numbering system as analogous residues of APPI

Table S2. Crystallographic data collection and refinement statistics

PDB ID	5C67
Complexes per ASU	2
Space group	$P4_12_12$
Unit cell, Å	78.15, 78.15, 243.73, 90°, 90°, 90°
Resolution, Å	1.83
Unique reflections	66,627
Completeness, %	96.7 (91.1)[a]
Multiplicity	6.0 (5.6)[a]
Mean I/σ	16.4 (3.1)[a]
R_{merge}	0.079 (0.796)[a]
R_{meas}	0.083 (0.910)
$R_{p.i.m.}$	0.032 (0.366)
R_{cryst}/R_{free} (%)	22.4/26.7
Average B-factor, $Å^2$	34.0
Protein atoms	4219
Water molecules	93
R.m.s.d. bonds, Å	0.0214
R.m.s.d. angles,	2.0083
ϕ,ψ angle distribution[b]:	
Favored regions (%)	96
Allowed regions (%)	4
Outliers (%)	0

[a]Values in parentheses are for the highest resolution shell.

[b]Ramachandran distributions are reported as defined by the PDB validation server/MolProbity.

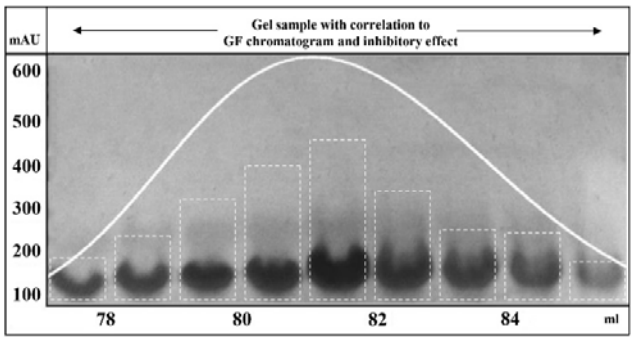

Fig. S3. Protein validation. A representative example of the non-reduced SDS-PAGE of $APPI_{WT}$ samples from gel-filtration (GF) with overlap on the GF chromatogram (solid line) and the inhibitory effect (dashed bars) on bovine trypsin catalytic activity. For the catalytic activity assay the inhibitor samples were diluted 1:1000 (inhibitory effect has no units i.e., normalized to the highest peak value).

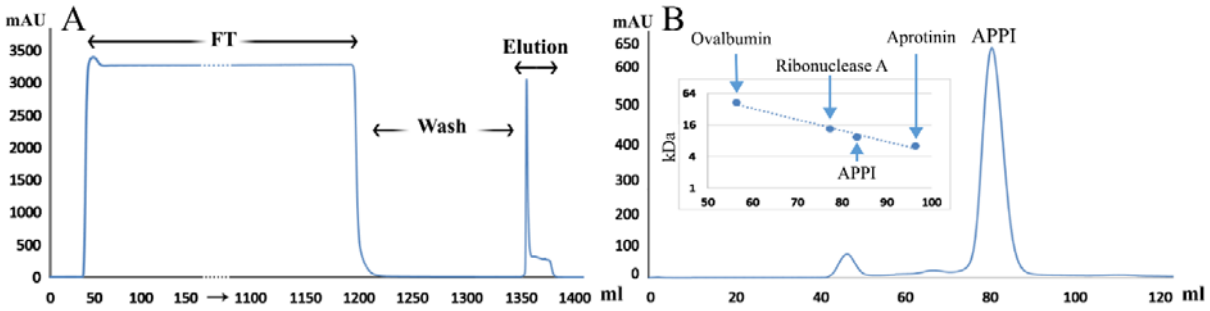

Fig. S4. (*A*) Representative nickel-IMAC purification of $APPI_{WT}$. The supernatant was loaded on a HisTrap (GE Healthcare) column for 24 h (Flowthrough; FT) using an ÄKTApure instrument (GE Healthcare), followed by washing and elution. (*B*) Gel filtration chromatography of $APPI_{WT}$. Eluted protein (2.5 ml) from the previous purification step was injected into a Superdex 75 16/600 column (GE Healthcare). The inset shows the elution time (ml) of the middle peak of different protein standards, including aprotinin (6.5 kDa), ribonuclease A (13.7 kDa), and ovalbumin (43 kDa). The Mw of the purified APPI was estimated to be ~9 kDa according to the standards.

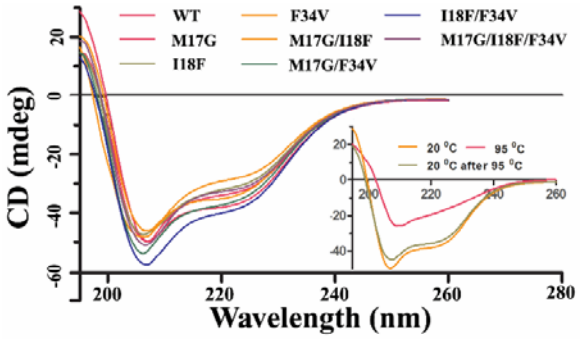

Fig. S5. Circular dichroism spectra. Absorbance was recorded over a range of 190–260 nm using a quartz cuvette with a path length of 1 mm. Three scans of 50 μM protein solutions were averaged to obtain smooth data, and background was corrected with respect to protein-free buffer. CD spectra of all the APPI clones were very similar to that of BPTI taken from the literature [7] (as well as to each other), which indicates correct 3D backbone folding for each APPI variant. The inset presents a representative example of $APPI_{WT}$ CD scans at room temperature (20 °C) and under denaturation (95 °C) and renaturation (at 20 °C following 95 °C incubation) conditions. The results show that all the APPI variants were completely renaturative after heating to 95 °C.

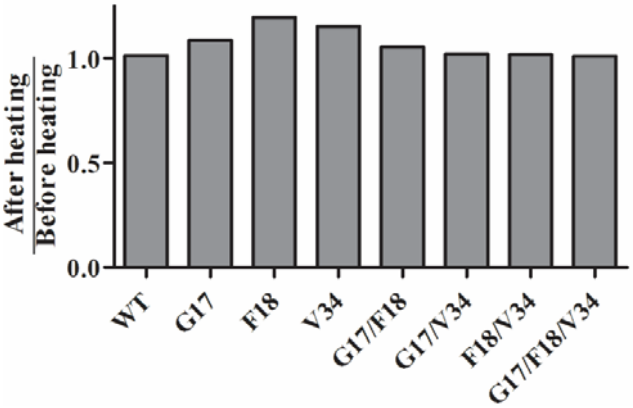

Fig. S6. Thermostability of APPI variants. Each variant (125 nM) was heated at 95 °C for 5 min and tested for its ability to inhibit bovine-trypsin (final concentrations of inhibitors and enzyme were 3.1 nM and 2.5 nM, respectively). Y axis represents the ratio of the % inhibitory effect of APPI after heating at 95 °C normalized by the % inhibitory effect of APPI before heating at 95 °C. Notably, there was no impairment of the inhibitory effect of all APPI variants, thus validating the high thermostability of these proteins.

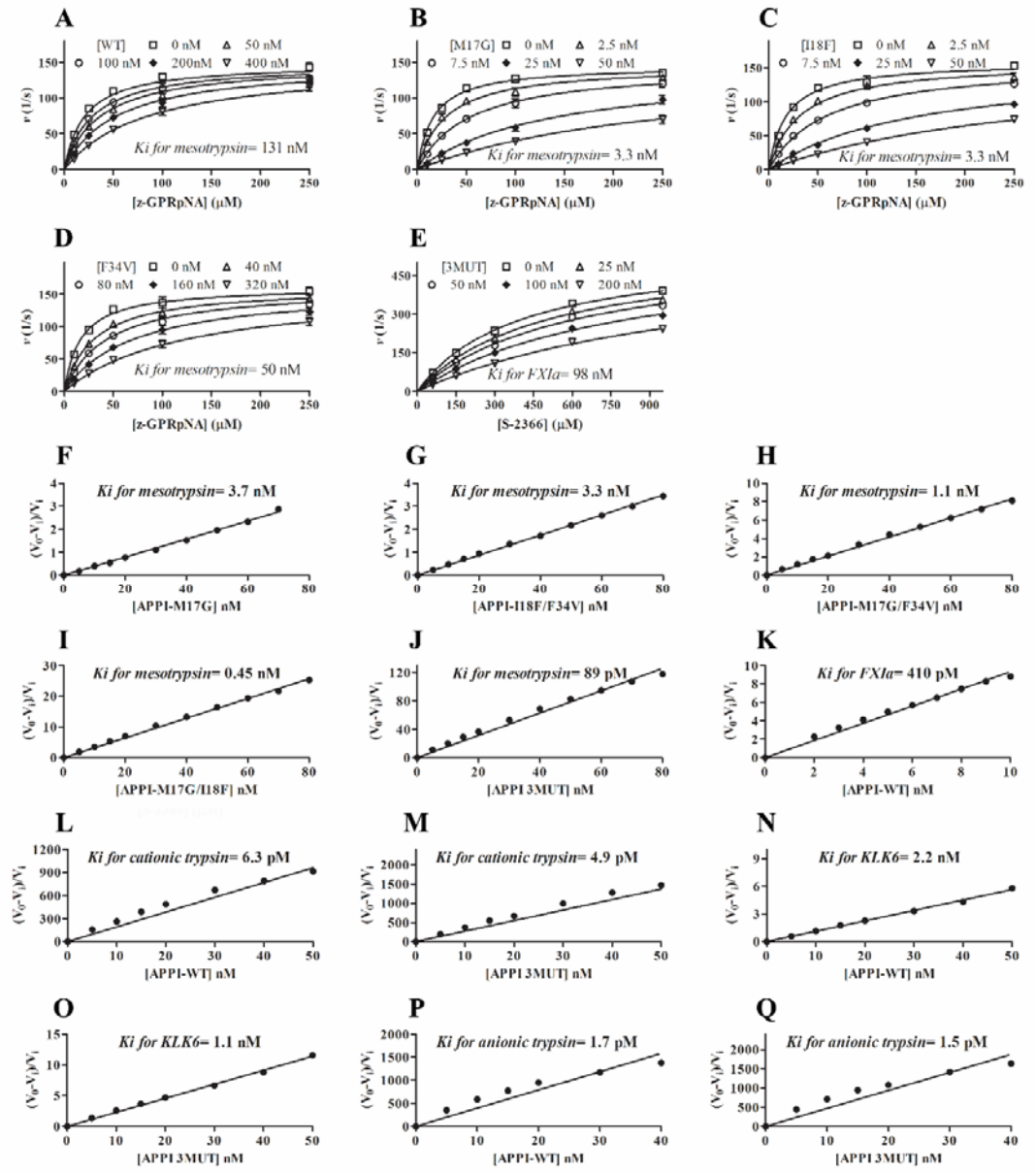

Fig. S7. Inhibition kinetics of mesotrypsin, anionic trypsin, cationic trypsin, kallikrein-6 and FXIa by APPI. (*A-D*) Competitive patterns of mesotrypsin inhibition by the respective APPI variant as noted on the upper left in each panel. (*E*) Competitive patterns of FXIa inhibition by $APPI_{M17G/I18F/F34V}$. APPI (inhibitor) concentration is given at the top of each plot. Data was fitted globally to the competitive inhibition equation using Prism, GraphPad Software. (*F-Q*) Slow, tight binding inhibition of mesotrypsin, anionic trypsin, cationic trypsin, kallikrein-6 or FXIa by APPI as noted on the figure. V_0 is the non-inhibited rate, and V_i is the rate in the presence of the respective APPI variant (noted in the X axis), which allows calculation of *Ki* using eq. 2 (see above).

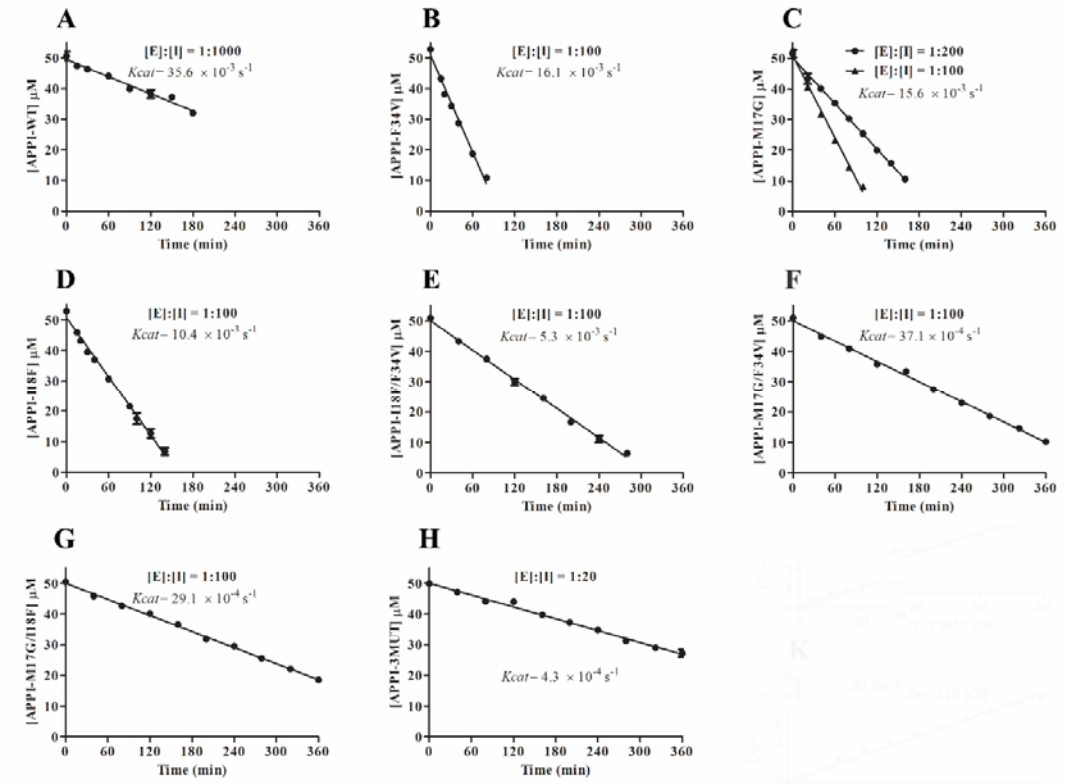

Fig. S8. Kinetics of APPI hydrolysis by mesotrypsin. Initial rates of hydrolysis, from which k_{cat} is calculated are shown. Disappearance of the respective intact APPI (noted in the Y axis) was quantified by integration of the HPLC peak in a different time courses as illustrated in each panel. Hydrolysis reaction contained 50 μM of APPI and 0.05-2.5 μM of enzyme, i.e., [E]:[I] ratios ranged from 1:1000 to 1:20.

Triple mutant cycle analysis of the interactions between APPI residues at positions 17, 18 and 34

For the APPI-mesotrypsin complex, each mutation had a different impact on the binding strength and the rate of hydrolysis. To assess the extent to which the effects of the mutations on the measured functional properties (K_i and k_{cat}) are independent (non-cooperative) or cooperative, we assessed the strength of the (direct or indirect) interactions between two residues *X* and *Y* in the protein (*P*) in a cycle that comprised the wild-type protein P_{XY}, two single mutants, P_{X0} and P_{0Y}, and the corresponding double mutant, P_{00} (*0* indicates a mutation). A measure of the strength of interaction is the coupling energy, $\Delta\Delta G_{int}$, which is given by:

$$\Delta\Delta G_{\text{int}} = \Delta G(P_{XY}) - \Delta G(P_{0Y}) - \Delta G(P_{X0}) + \Delta G(P_{00}) = -RT\ln\left(\frac{K_{XY}\times K_{00}}{K_{0Y}\times K_{X0}}\right) \qquad \text{(Eq. 3)}$$

where *R* is the gas constant, *T* is the absolute temperature, and $\Delta G(P_{XY})$, $\Delta G(P_{0Y})$, $\Delta G(P_{X0})$ and $\Delta G(P_{00})$ correspond to the free energies of binding (ΔG_a) or catalysis

(ΔG_{cat}). A coupling energy of zero (i.e., additivity of mutational effects) indicates that there is no interaction between *X* and *Y* with respect to the process (i.e., association or proteolytic stability) under consideration.

The free energy changes of catalysis ($\Delta\Delta G_{cat}$) and association ($\Delta\Delta G_a$) upon a single point mutation (i.e., $\Delta\Delta G$ of P_{XY} and P_{X0}) were calculated in a similar manner by using Eq. 4:

$$\Delta\Delta G = \Delta G(P_{X0}) - \Delta G(P_{XY}) = -RT \ln \frac{K_{X0}}{K_{XY}} \qquad \text{(Eq. 4)}$$

Previous studies of the free energy of coupling have suggested energy values of zero (within the error range) [8] as a basis for assuming additivity (no interaction between residues). Our results for the free energy of coupling for catalysis and association (ranging from -1.04 to 0.99 kcal/mol) suggest that there are some mutations that act independently (non-cooperatively) and others that act cooperatively (Fig. S9). For example, there is cooperativity between M17G and F34V with respect to proteolytic stability, with $\Delta\Delta G_{\text{int}}^{cat}$ = 0.39±0.09 kcal/mol (Fig. S9B), but there is no interaction between these mutations with respect to affinity ($\Delta\Delta G_{\text{int}}^{a}$ = -0.06±0.09 kcal/mol). F34V is advantageous only in combination with M17G; it confers no additional improvement in binding ($\Delta\Delta G_{\text{int}}^{a}$ = -0.59±0.08 kcal/mol, Fig. S9A) or proteolytic stability ($\Delta\Delta G_{\text{int}}^{cat}$ =0.07±0.08 kcal/mol, Fig. S9B) when combined only with I18F. Notably, M17G and F34V are independent and therefore energetically additive with respect to affinity ($\Delta\Delta G_{\text{int}}^{a}$ = -0.06±0.09 kcal/mol) when Ile is in position 18 (as in APPI-WT), but act cooperatively when Phe is in position 18 ($\Delta\Delta G_{\text{int}}^{a}$ = 0.99±0.04 kcal/mol). Most importantly, when combined to form the $\text{APPI}_{\text{M17G/I18F/F34V}}$ triple mutant, the three mutations (M17G, I18F, and F34V) act cooperatively in most cases with respect to affinity and in all cases with respect to proteolytic stability.

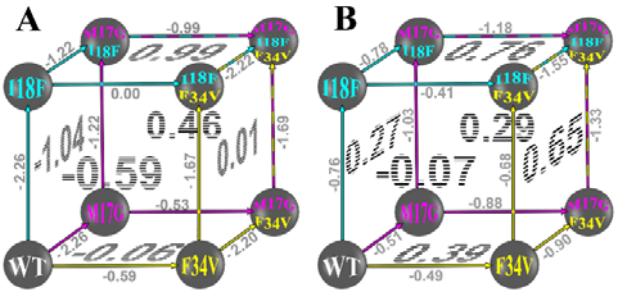

Fig. S9. 'Triple mutant cycle analysis cube' that summarizes the coupling in free energy changes attributable to mutation of residues 17, 18 and 34 on the APPI sequence

Each corner of the cube represents a different APPI variant, as annotated. (**A**) Values along each edge represent $\Delta\Delta G_a$ (kcal/mol), calculated using Eq. 4; each face of the cube represents $\Delta\Delta G^{a}_{\text{int}}$ (kcal/mol) of a double mutant cycle attributable to the corner variants, calculated using Eq. 3. Here, the equilibrium association constant that was used was approximated as the reciprocal of the measured inhibition constant *Ki*. (**B**) The figure shows the free energy changes as in panel *A*, but for catalysis, i.e., $\Delta\Delta G_{cat}$ and $\Delta\Delta G^{cat}_{\text{int}}$, respectively.

REFERENCES

1 Chao, G., Lau, W. L., Hackel, B. J., Sazinsky, S. L., Lippow, S. M. and Wittrup, K. D. (2006) Isolating and engineering human antibodies using yeast surface display. Nat Protoc. **1**, 755-768

2 Wilson, D. S. and Keefe, A. D. (2001) Random mutagenesis by PCR. Curr Protoc Mol Biol. **Chapter 8**, Unit8 3

3 Salameh, M. A., Soares, A. S., Hockla, A., Radisky, D. C. and Radisky, E. S. (2011) The P(2)' residue is a key determinant of mesotrypsin specificity: engineering a high-affinity inhibitor with anticancer activity. Biochem J. **440**, 95-105

4 Salameh, M. A., Soares, A. S., Hockla, A. and Radisky, E. S. (2008) Structural basis for accelerated cleavage of bovine pancreatic trypsin inhibitor (BPTI) by human mesotrypsin. J Biol Chem. **283**, 4115-4123

5 Looke, M., Kristjuhan, K. and Kristjuhan, A. (2011) Extraction of genomic DNA from yeasts for PCR-based applications. Biotechniques. **50**, 325-328

6 Chase, T., Jr. and Shaw, E. (1967) p-Nitrophenyl-p'-guanidinobenzoate HCl: a new active site titrant for trypsin. Biochem Biophys Res Commun. **29**, 508-514

7 Lees, J. G., Miles, A. J., Wien, F. and Wallace, B. A. (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. Bioinformatics. **22**, 1955-1962

8 Horovitz, A. and Rigbi, M. (1985) Protein-protein interactions: additivity of the free energies of association of amino acid residues. J Theor Biol. **116**, 149-159