Supplementary Information

Fluorine teams up with water to restore inhibitor activity to mutant BPTI

Shijie Ye[1], Bernhard Loll[2], Allison Ann Berger[1], Ulrike Mülow[1], Claudia Alings[2], Markus Christian Wahl[2] and Beate Koksch[1]*
Department of Biology, Chemistry, and Pharmacy, Freie Universität Berlin,
1. Institute of Chemistry and Biochemistry – Organic Chemistry, Takustrasse 3,
2. Institute of Chemistry and Biochemistry – Structural Biochemistry, Takustrasse 6,
D-14195, Berlin, Germany
*Corresponding Author. Tel.: 0049-30-83855344; Fax: 0049-30-83855644.
E-mail address: Beate.Koksch@fu-berlin.de

Materials

Fmoc-L-amino acids were purchased from Fa. Gerhardt. Boc-L-Lys (Pbf) was purchased from Sigma-Aldrich. Fmoc-(*S*)-2-aminobutyric acid (Abu) from Bachem. (*S*)-2-amino-4,4-difluorobutyric acid (DfeGly) and (*S*)-2-Amino-4,4,4-trifluorobutyric acid (TfeGly) were synthesized according to literature procedures.[1,2]

H-Gly-2-ClTrt (0.69 mmol/g) resin and Fmoc-Ala-NovaSyn®TGT (0.2 mmol/g) resin were purchased from Novabiochem. 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (TBTU), and 1-hydroxybenzotriazole (HOBt) were purchased from Fa. Gerhardt. Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) was purchased from Novbiochem. 4-mercaptophenylacetic acid (MPAA, 97%) and 4-acetamidothiophenol (98%), and dithiothreitol (DTT) were from Alfa Aesar. Tris-(2-carboxyethyl)phosphine hydrochloride (TECP) were purchased from VWR. Dimethylformamide for peptide synthesis, *N*,*N*-diisopropylethylamine (DIPEA, 98+ %), *N*,*N*-diisoporopylcarbodiimide (DIC, 99 %), piperidine (99 %, extra pure), trifluoroacetic acid

(TFA, 99 %), triisopropylsilane (TIPS, 99 %), and 1,2-ethanedithiol (EDT, 95%), and acetonitrile for HPLC were purchased from Acros. Phenol (99%) and Thioanisole (99%) were purchased from Janssen Chimica. Deionized water for buffer solutions and HPLC was prepared using the MilliQ-AdvantageA10-System (Millipore). Water (solvent A) and acetonitrile (solvent B) for reversed-phase HPLC were supplemented with 0.1 % TFA (Uvasol, Merck). Dialysis membrane, Spectra/Por® was purchased from VWR. Amicon® Ultra centrifugal filter devices for concentration of protein were purchased from Millipore. The standard chemicals are purchased from Sigma-Aldrich and Fluka.

Wild-type BPTI for assay, plasmin from human plasma, and enzyme substrate: *N*-Benzoyl-L-tyrosine ethyl ester (BTEE, chymotrypsin substrate), D-Ile-Phe-Lys *p*-nitroanilide (human plasmin substrate), and N_{α}-Benzoyl-L-arginine 4-nitroanilide hydrochloride (trypsin substrate) were purchased from Sigma-Aldrich. Wild-type BPTI for protein crystallography was purchased from Roche Applied Science. α-chymotrypsin (Cat. 27272, Lot. S43979058) and β-trypsin (DPCC-behandelt, Cat. 93611, Lot. S39283258) were purchased from Fluka.

HPLC and ESI-TOF

Analytical reversed-phase (RP) HPLC was performed on LaChrom HPLC system (VWR & HITACHI) and Elite Lachrom RP-HPLC system (VWR & HITACHI). Capcell pak C18 column (Shisheido, Type: SG120 5μm, Size: 4.6 mmφ x 250 mm) and Phenomenex Luna C18 column (5 μm, 250 mm × 4.6 mm) were used for analytic RP-HPLC. Peptide fragments and peptide-α-thioesters were purified on a Knauer HPLC system. A Phenomenex C18 column (Gemini-NX 10μ 110A AXIA, Size: 250 x 21.20 mm) was used. Purity of peptide was confirmed by analytical HPLC and Electrospray-ionization time-of-flight high-resolution mass spectrometry (ESI-TOF). ESI-TOF was performed on an Agilent 6210 ESI-TOF (Agilent Technologies, Santa Clara, CA, U.S.A.).

Peptide synthesis and purification

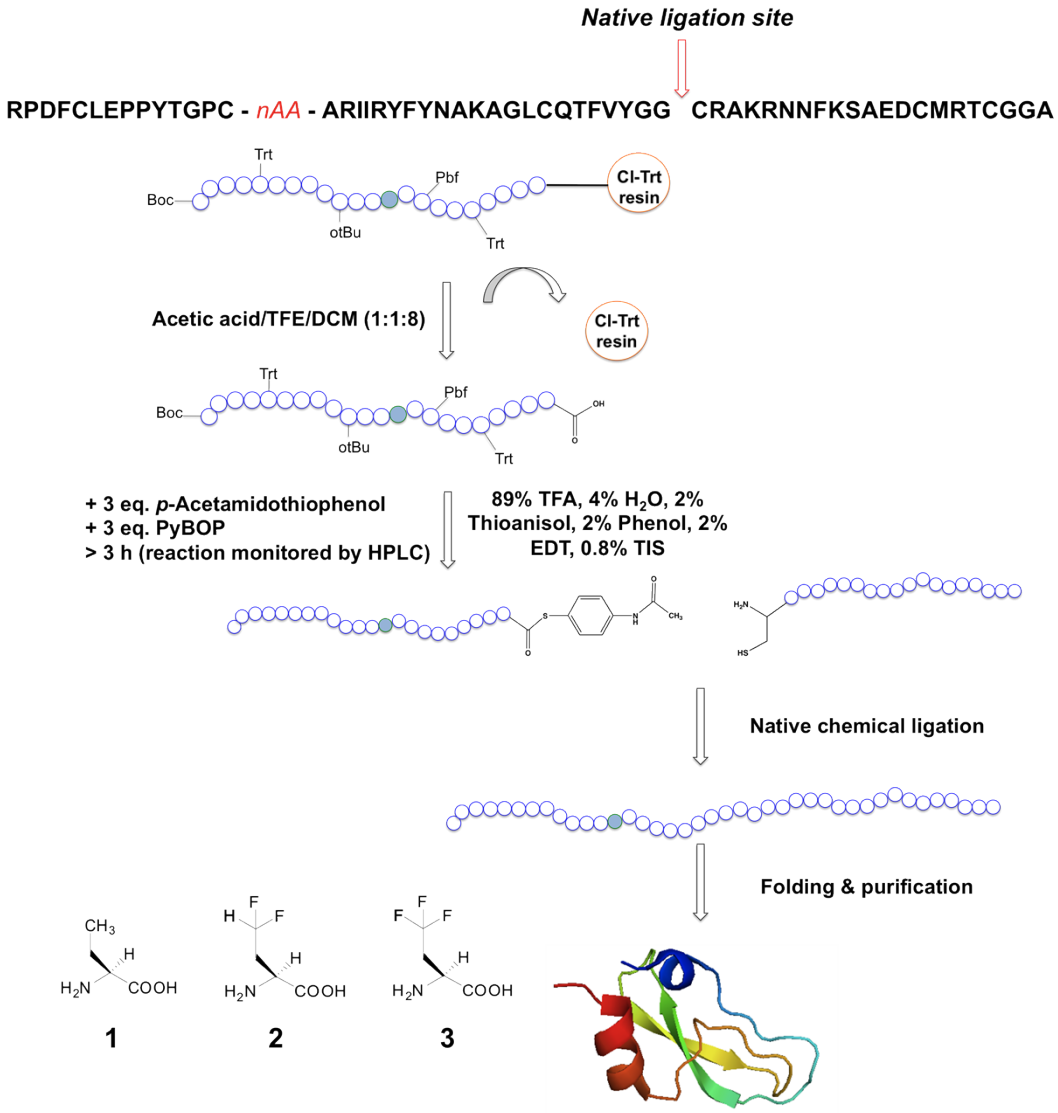


Fig. S1 Synthetic strategy of total chemical synthesis of BPTI containing noncanonical amino acids at position 15

Synthesis of BPTI 38-58 fragment

The C-terminal peptide segment BPTI 38-58 was synthesized on preloaded Fmoc-Ala-NovaSyn®TGT resin with a SyroXP-I peptide synthesizer (MultiSynTech GmbH) on a 0.05 mmol scale according to standard Fmoc/*t*Bu chemistry. For standard couplings, a four-fold excess of amino acids and coupling reagents (TBTU/HOBt) as well as an eight-fold excess of DIPEA relative to resin loading was used. All couplings were performed as double couplings for 30 minutes. Cleavage and removal of side chain protecting groups was carried out by treatment with 4 mL of cleavage cocktail containing 89% TFA, 2% phenol, 4% H_2O, 1% TIS, 2% thioanisol, 2% EDT. The peptides were precipitated with cold diethylether and purified by RP-HPLC using a gradient of 0-80% MeCH in H_2O at flow rate of 20 mL/minute over 30 minutes (yield 18%).

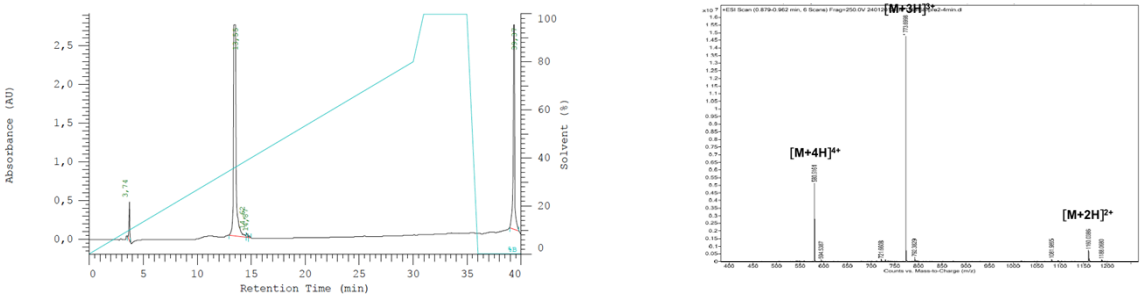


Fig. S2 Left, Analytic HPLC of purified BPTI 38-58 (retention time: 13.6 minute, 0-80% B in 30 minutes); right, ESI-MS analysis of BPTI 38-58 (observed mass: 1160.0386 $[M+2H]^{2+}$, 773.6998 $[M+3H]^{3+}$, 580.5161 $[M+4H]^{4+}$, calculated mass: 1159.5226 $[M+2H]^{2+}$, 773.3510 $[M+3H]^{3+}$, 580.2652 $[M+4H]^{4+}$).

Synthesis of BPTI 1-*15nAA*-37-α-thioester

The peptide fragment BPTI 16-37 was synthesized on preloaded H-Gly-2-ClTrt resin by use of a SyroXP-I peptide synthesizer on a 0.1 mmol scale according to standard Fmoc/*t*Bu chemistry. For standard couplings, a four-fold excess of amino acids and coupling reagents (TBTU/HOBt) was used. All couplings were performed as double couplings for 30 minutes. Products were analyzed by test cleavage using a cocktail containing 82% TFA, 3% DTT, 5% H_2O, 5% TIS, 5% thioanisol. The peptide from the test cleavage was subsequently analyzed by means of analytical HPLC and ESI-MS.

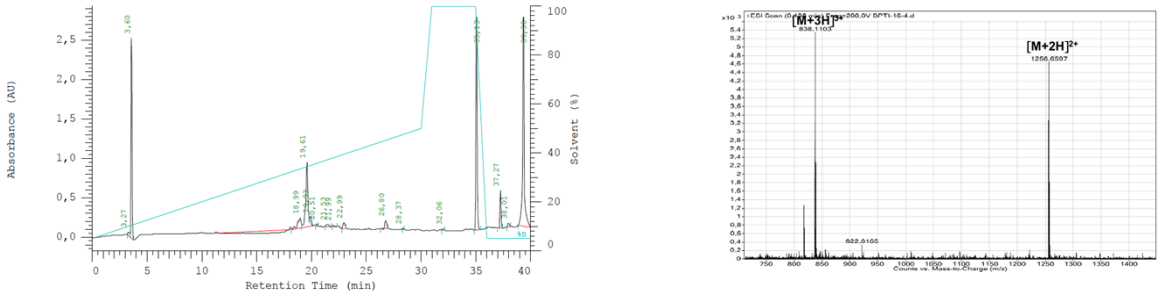


Fig. S3 Left, Analytic HPLC of crude BPTI 16-37 (retention time: 19.6 minute, 0-80% B in 30 minutes); right, ESI-MS analysis of crude BPTI 16-37 (observed mass: 1256.6597 $[M+2H]^{2+}$, 838.1103 $[M+3H]^{3+}$ calculated mass: 1256.1524 $[M+2H]^{2+}$, 837.7708 $[M+3H]^{3+}$)

The Fmoc-protected non-canonical amino acids were incorporated into peptides by means of manual coupling following the standard HOBt/DIC protocol. 1.25 equivalent of Fmoc-L-amino acid/HOBt/DIC relative to resin loading was used and incubated overnight. The efficiency of manual coupling was analyzed by test cleavage, HPLC, and ESI-MS (second coupling was applied in case the first coupling was not complete). The elongation of peptides was carried out on an Activo P11 Synthesizer. 10 equivalents of Fmoc-L-

amino acids/HOBt/DIC were used. Because the BPTI sequence is difficult to synthesize, the duration and number of coupling of each residue were carried out according to literature.[3] Boc-Arg(Pbf) was used for coupling the last residue, Arg 1.

The fully protected peptide fragments were cleaved from 2-ClTrt resins by treatment of the resin with a mixture containing acetic acid/trifluorethanol/DCM (1:1:8 v/v/v). To remove the acetic acid, the cleavage solution was extracted with half-saturated $NaHCO_3$ three times and the organic phases combined. Subsequent to concentration by evaporation, an aliquot was taken for a test cleavage (see conditions above). The peptide fragment from the test cleavage was analyzed by HPLC and ESI-MS. The α-thioesters of the peptide fragments were synthesized by incubation of the fully protected peptides with a solution containing 5 equivalents of *p*-acetamidothiophenol, DIPEA, and PyBOP in DCM. The reaction was incubated overnight. Aliquots were taken for test cleavage and HPLC/ESI-MS analysis. Upon completion, the organic solvent was evaporated under reduced pressure. Side chain protecting groups were removed by treating the peptide α-thioester with cleavage cocktail containing 89% TFA, 2% phenol, 4% H_2O, 1% TIS, 2% thioanisol and 2% EDT. The BPTI 1-*nAA15*-37-α-thioester is poorly soluble in H_2O/MeCN.[4] A suitable amount of DMSO was added to dissolve the peptide α-thioester. The purification was applied on RP-HPLC (C18 column) using a gradient of 5-70% MeCN in H_2O at a flow rate of 20 mL/minute over 30 minutes (yield 6-7%).

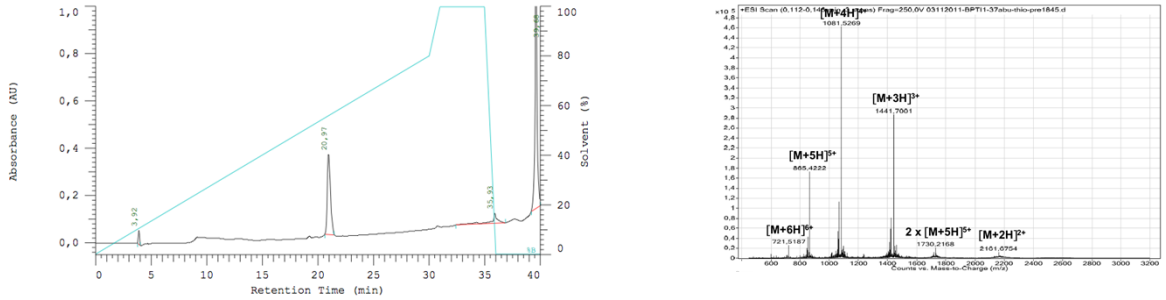


Fig. S4 Left, Analytic HPLC of purified BPTI 1-*Abu15*-37-SR (retention time: 21.0 minute, 5-70% MeCN in H_2O in 30 minutes); right, ESI-MS analysis of BPTI 1-*Abu15*-37-SR (observed mass: 2161.6754 $[M+2H]^{2+}$, 1441.7001 $[M+3H]^{3+}$, 1081.5269 $[M+4H]^{4+}$, 865.4222 $[M+5H]^{5+}$, 721.5187 $[M+6H]^{6+}$, calculated mass: 2161.1699 $[M+2H]^{2+}$, 1441.1158 $[M+3H]^{3+}$, 1081.0888 $[M+4H]^{4+}$, 865.0726 $[M+5H]^{5+}$, 721.0618 $[M+6H]^{6+}$).

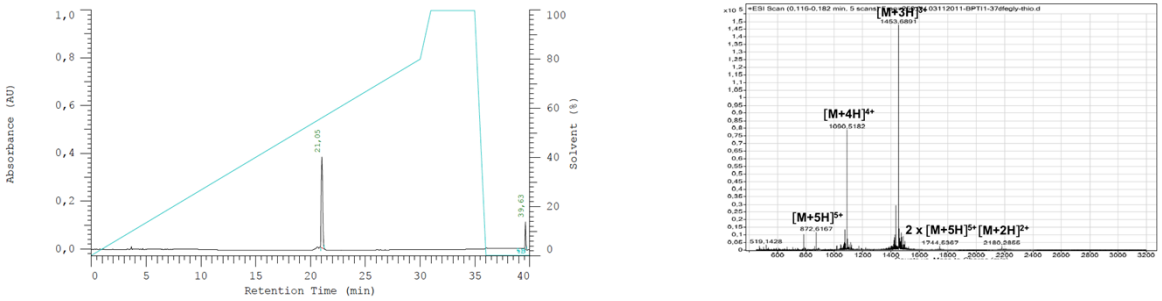


Fig. S5 Left, Analytic HPLC of purified BPTI 1-*DfeGly15*-37-SR (retention time: 21.1 minute, 5-70% MeCN in H_2O in 30 minutes); right, ESI-MS analysis of BPTI 1-*DfeGly15*-37-SR, (observed mass: 2180.2855 $[M+2H]^{2+}$, 1453.6891 $[M+3H]^{3+}$, 1090.5182 $[M+4H]^{4+}$, 872.8167 $[M+5H]^{5+}$, calculated mass: 2179.1299 $[M+2H]^{2+}$, 1453.0892 $[M+3H]^{3+}$, 1090.0688 $[M+4H]^{4+}$, 872.2566 $[M+5H]^{5+}$

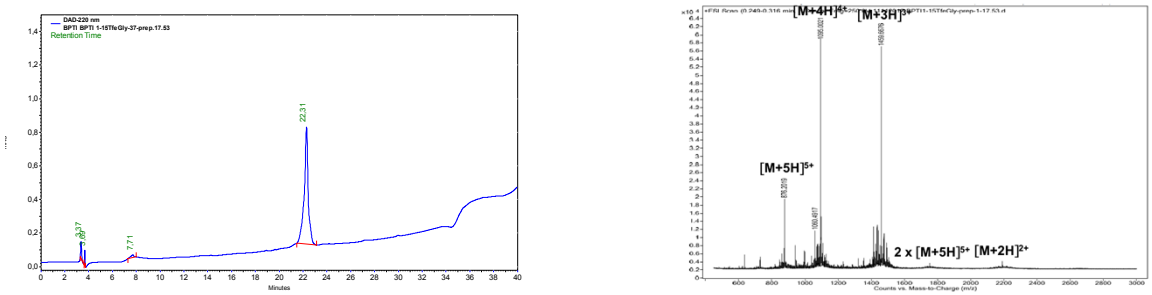


Fig. S6 Left, Analytic HPLC of purified BPTI 1-*TfeGly15*-37-SR (retention time: 22.3 minute, 0-80% B in 30 minutes); right, ESI-MS analysis of BPTI 1-*TfeGly15*-37-SR, (observed mass: 1459.6676 $[M+3H]^{3+}$, 1095.0021 $[M+4H]^{4+}$, 876.2019 $[M+5H]^{5+}$, calculated mass: 1459.0892 $[M+3H]^{3+}$, 1094.5688 $[M+4H]^{4+}$, 875.8566 $[M+5H]^{5+}$).

Native chemical ligation

Native chemical ligation (NCL) buffer containing 6 M guanidinium chloride (GdmCl) and 0.1 M Na_2HPO_4 was freshly prepared. MPAA and TCEP were added to final concentrations of 50 mM and 20 mM, respectively. 5 mL NCL buffer was degased with Argon and the pH was adjusted with 2 M NaOH to 7. Accurate amounts of Cys-peptide and peptide-thioester were added into this solution to give a final concentration of peptides between 1-5 mM. Reaction was monitored by HPLC/ESI-MS. 10 µL aliquots were taken at 0 h, 1 h, 2 h, 4 h, 8 h, and 16 h and diluted to 100 µL with MeCN/H_2O (v/v, 1:1, 0.1 % TFA) for analysis. After the NCL reaction was complete, NCL reaction mixture was directly applied onto a PD-10 desalting column (Sephadex G-25 M, GE Healthcare) to remove MAPP, which gives a huge peak in the HPLC. Subsequently, full-length peptides were purified by means of RP-HPLC (C 18 column) using a gradient of 5-70% MeCN in H_2O at flow rate of 20 mL/minute over 30 minutes.

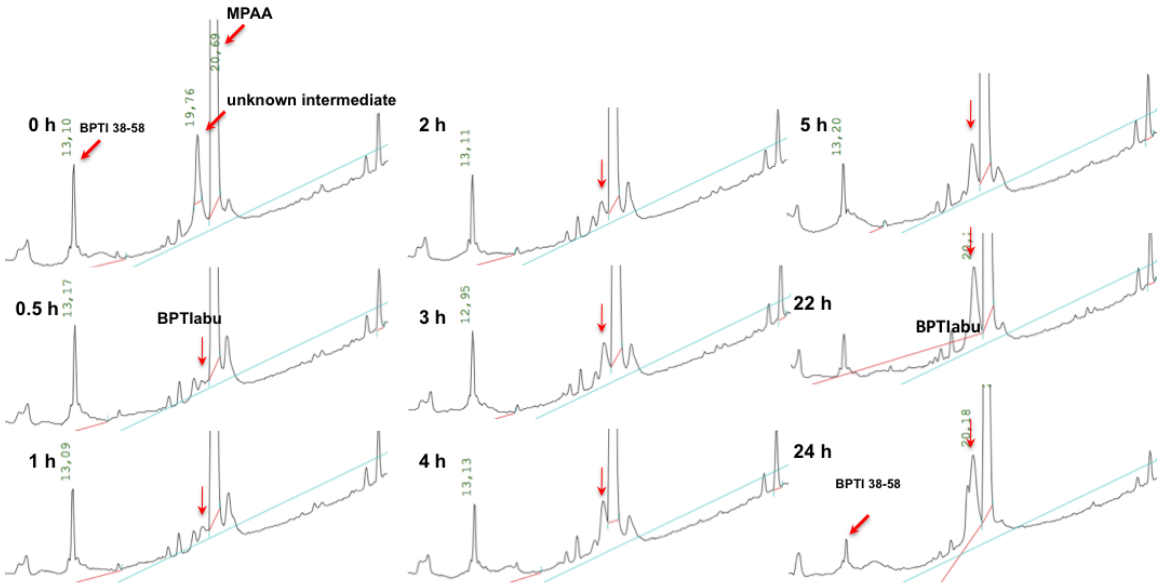


Fig. S7 Monitoring of native chemical ligation reaction by analytic RP-HPLC, (here BPTI 1-*Abu15*-37-SR was used for NCL), 10 µL aliquot of reaction mixture was taken and diluted with 90 µL H_2O/MeCN (1:1, 0.1% TFA) for analysis.

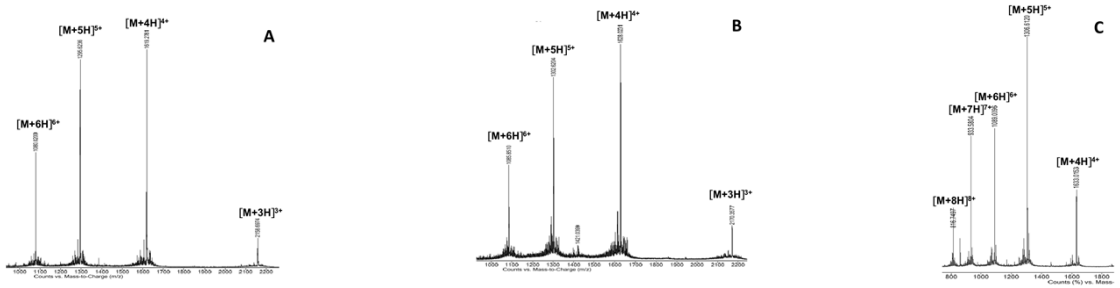


Fig. S8 ESI-MS analysis of unfolded full-length BPTI mutants, A) unfolded BPTI (Lys15Abu) observed mass: 2158.6974 $[M+3H]^{3+}$, 1619.2781 $[M+4H]^{4+}$, 1295.6236 $[M+5H]^{5+}$, 1080.0200 $[M+6H]^{6+}$, calculated mass: 2157.7122 $[M+3H]^{3+}$, 1618.5361 $[M+4H]^{4+}$, 1295.0304 $[M+5H]^{5+}$, 1079.3600 $[M+6H]^{6+}$; B) unfolded BPTI (Lys15DfeGly) observed mass: 2170.3577 $[M+3H]^{3+}$, 1628.0231 $[M+4H]^{4+}$, 1302.6204 $[M+5H]^{5+}$, 1085.8510 $[M+6H]^{6+}$, calculated mass: 2169.6855 $[M+3H]^{3+}$, 1627.5161 $[M+4H]^{4+}$, 1302.2144 $[M+5H]^{5+}$, 1085.3467 $[M+6H]^{6+}$; C) unfolded BPTI (Lys15TfeGly) observed mass: 1633.0153 $[M+4H]^{4+}$, 1306.6120 $[M+5H]^{5+}$, 1089.0096 $[M+6H]^{6+}$, 933.5804 $[M+7H]^{7+}$, 816.7497 $[M+6H]^{6+}$, calculated mass: 1632.0161 $[M+4H]^{4+}$, 1305.8144 $[M+5H]^{5+}$, 1088.3467 $[M+6H]^{6+}$, 933.0125 $[M+7H]^{7+}$, 816.5119 $[M+8H]^{8+}$.

Protein folding

Reduced mutant BPTI species were dissolved in folding buffer containing 0.6 M Tris-HCl, 6 mM EDTA and 6 M GdmCl at pH 8.7 to final concentration of 1 mg/ml. The solution was rapidly diluted 6-fold with water and stirred in an open-air container for 1-2 days. 10 µl aliquots were diluted to 100 µL with MeCN/H_2O (v/v, 1:1, 0.1 % TFA) for HPLC analysis.

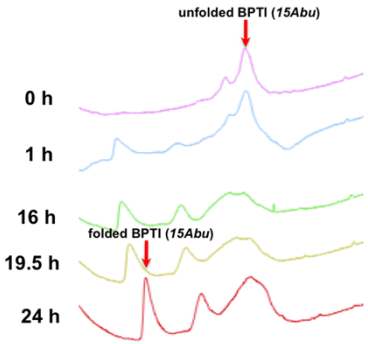


Fig. S9 Monitoring of BPTI folding process, folding analysis of BPTI (Lys15Abu) is shown here as an example.

Folded BPTI mutants were purified by RP-HPLC (C 18 column) using a gradient of 5-70% MeCN in H2O at flow rate of 20 mL/minute in 30 minutes. Purity was verified by HPLC/ESI-MS.

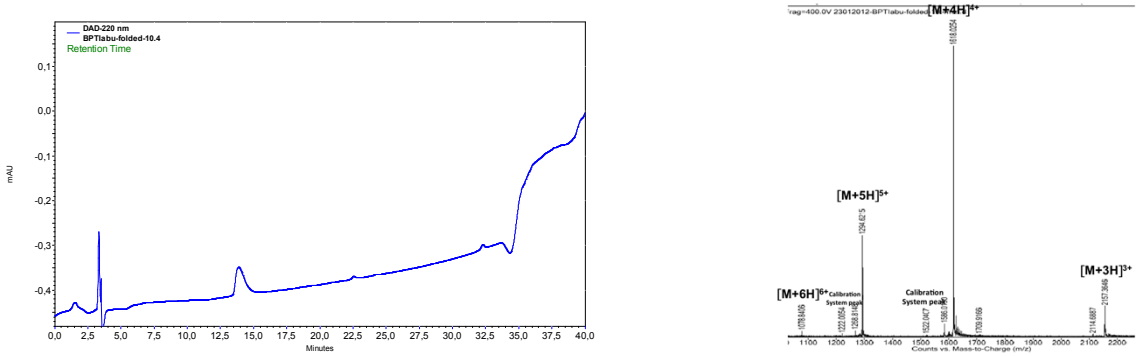


Fig. S10 Left, Analytic HPLC of purified folded BPTI (Lys15Abu) (retention time: 13.9 minute, 5-70% MeCN in H_2O in 30 minutes); right, ESI-MS analysis of BPTI (Lys15Abu) observed mass: 2157.3646 $[M+3H]^{3+}$, 1618.0254 $[M+4H]^{4+}$, 1294.6215 $[M+5H]^{5+}$, 1078.8406 $[M+6H]^{6+}$, calculated mass: 2156.0138 $[M+3H]^{3+}$, 1617.2623 $[M+4H]^{4+}$, 1294.0114 $[M+5H]^{5+}$, 1078.05108 $[M+6H]^{6+}$.

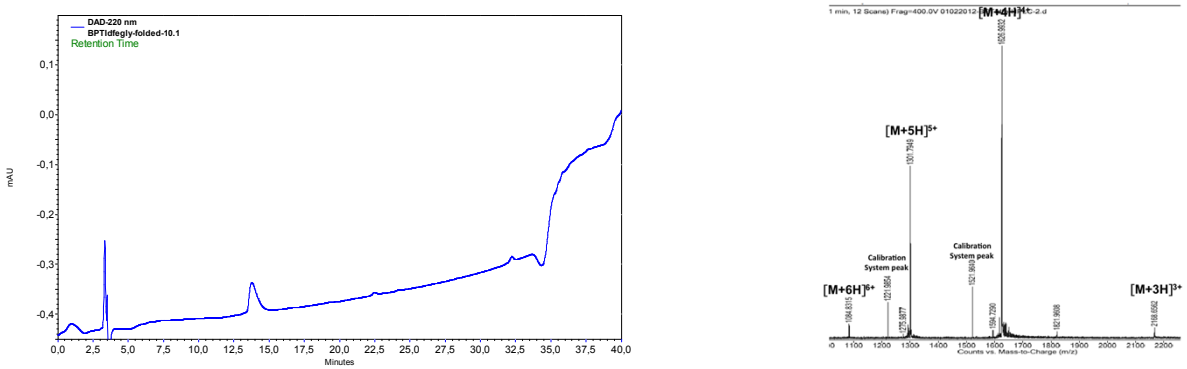


Fig. S11 Left, Analytic HPLC of purified BPTI (Lys15DfeGly) (retention time: 13.7 minute, 5-70% MeCN in H_2O in 30 minutes); right, ESI-MS analysis of BPTI (Lys15DfeGly) observed mass: 2168.6562 $[M+3H]^{3+}$, 1626.9932 $[M+4H]^{4+}$, 1301.7949 $[M+5H]^{5+}$, 1084.8315 $[M+6H]^{6+}$, calculated mass: 2167.9872 $[M+3H]^{3+}$, 1626.2423 $[M+4H]^{4+}$, 1301.1954 $[M+5H]^{5+}$, 1084.4975 $[M+6H]^{6+}$.

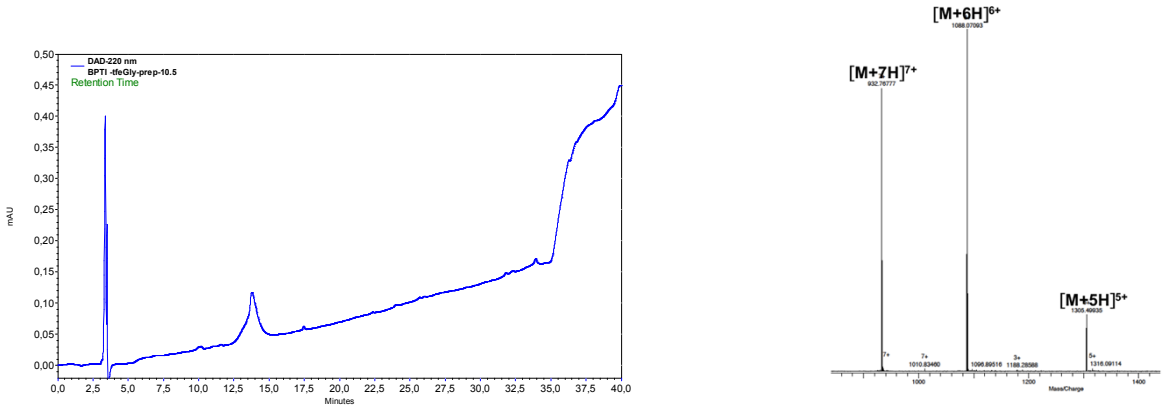


Fig. S12 Left, Analytic HPLC of purified BPTI (Lys15TfeGly) (retention time: 14.1 minute, 5-70% MeCN in H_2O in 30 minutes); right, ESI-MS analysis of BPTI (Lys15TfeGly) observed mass: 1305.4994 $[M+5H]^{5+}$, 1088.0709 $[M+6H]^{6+}$, 932.7678 $[M+7H]^{7+}$, calculated mass: 1304.6514 $[M+5H]^{5+}$, 1087.3775 $[M+6H]^{6+}$, 932.1818 $[M+7H]^{7+}$.

Tab. S1 Synthesis summary of mutant BPTI variants containing non-canonical amino acids at position 15

	C-terminal fragment	**N-terminal fragment-α-thioester**		**NCL**	**Protein folding**
Resin	*Fmoc-Ala-NovaSyn® TGT*	*Fmoc-Gly-2Cl-Trt*			
Yield	18%	Abu15	6%	26%	17%
		DfeGly15	6%	17%	13%
		TfeGly15	7%	20%	11%

Determination of protein concentration

Mutant BPTIs were dissolved in appropriate buffer for CD measurement or enzyme inhibition assay and dialyzed overnight against the same buffer. The concentration of BPTI variants and β-trypsin was determined by means of a Nanodrop 2000 by providing the

molecular weight and extinction coefficient. The extinction coefficient was calculated using the online ProtParam tool by giving the protein sequence (http://web.expasy.org/protparam/, ExPASy Bioinformatics Resource Portal). ε (BPTI) = 6335 $M^{-1}cm^{-1}$, ε (β-trypsin) = 37650 $M^{-1}cm^{-1}$.

CD measurements

CD spectra were recorded on a Jasco J-715 spectropolarimeter at 25 °C (Jasco PTC-348WI peltier thermostat). Overall peptide concentrations were 20 µm at pH 7.4 (10 mM Tris buffer). CD-spectra were obtained in the far-UV range (190–240 nm) using 0.1 cm Quartz Suprasil cuvettes (Hellma) equipped with a stopper. The nitrogen flow rate was set to 3.1 L/min. The spectra were recorded at 0.2 nm intervals with a 2 nm bandwidth and 3 second response time. After baseline correction, the measured ellipticities were converted to molar ellipticities per residue (degree cm^2 $dmol^{-1}$ $residue^{-1}$) by normalizing for the concentration of peptide, the number of amino acids and path length.

Denaturation measurements of mutant BPTI were performed in presence of 10 mM Tris-HCL, 6 M GdmCl and 10 mM Tris-HCl, 8 M urea at pH 2. Melting curves were recorded using the signal at 222 nm applying a heating rate of 3 K/min from 20 to 100 °C. Each sample was prepared three times and the melting curves were averaged.

Thermodynamic parameters were determined by non-linear least square fitting of the normalized CD melting curves with six parameters (a, b, ΔH_m, T_m, $[\theta]_F$, and $[\theta]_U$) as two-state folding-unfolding equilibrium.[5] The fits were performed in Microsoft Excel as follows:[6,7]

At any temperature, T, the folding constant, K, is:

$$K = [F] / [U] \qquad (1)$$

[F] and [U] are the concentration of the folded and unfolded proteins. The free energy of folding is

$$\Delta G = -RTInK \quad (2)$$

R is the gas constant = 1.986 cal $mol^{-1}K^{-1}$, T is the absolute temperature (Kelvin).

The fraction unfolded at any temperature is f_U.

$$f_U = [U] / ([F] + [U]) \quad (3)$$

$$f_U = ([\theta]_o - ([\theta]_F) / ([\theta]_F + ([\theta]_U) \quad (4)$$

Where $[\theta]_o$ is the observed ellipticity at any time, $[\theta]_U$ represents the linear temperature dependence of the ellipticity of the fully unfolded proteins, $[\theta]_F$ is the linear temperature dependence of ellipticity of fully folded proteins:

$$[\theta]_t = f_U \times ([\theta]_F + [\theta]_U) + [\theta]_F \quad (5)$$

$$[\theta]_F = a \times T + [\theta]_F \text{ (at 0 K)} \quad (6)$$

$$[\theta]_U = b \times T + [\theta]_U \text{ (at 0 K)} \quad (7)$$

To fit the change of Circular dichroism at 222 nm as a function of temperature, T, Gibbs-Helmholtz equation is used to describe the folding as a function of temperature. The following equations are fit for unfolding of a monomer:

$$\Delta G = \Delta H (1 - T / T_M) - \Delta Cp ((T_M - T) + TIn (T / T_M)) \quad (8)$$

$$K = \exp (-\Delta G / (RT)) \quad (9)$$

$$f_U = K / (1 + K) \quad (10)$$

T_M is the temperature where $f_U = 0.5$ (melting temperature) and ΔCp is the change in heat capacity going from folded to unfolded state. ΔCp is usually assumed to be zero for the initial calculations of the thermodynamics for folding of monomeric proteins. Equations (5) – (10) were combined and data fitted directly.

Tab. S2 Fitting parameters of thermodynamic characterization

	BPTI species	*a*	*b*	$[\theta]_F$ (0K)[a]	$[\theta]_U$ (0K)[a]	T_m (0°)[b]	ΔH_m (kJ/ mol)[c]	$\Delta G°$ (kcal/ mol)[d]
in 6 M GdmCl	Wild-type	0.002	-0.007	-6.529	-1.625	66.48	45.70	11.99
	Lys15Abu	0.002	-0.012	-6.272	-1.129	63.44	42.37	11.26
	Lys15DfeGly	0.009	-0.028	-8.634	-0.208	68.77	46.78	12.42
	Lys15TfeGly	0.020	-0.040	-12.545	-0.248	68.27	46.00	12.05
in 8 M urea	Wild-type	0.010	-0.002	-6.595	-2.083	75.78	56.22	14.62
	Lys15Abu	0.013	-0.015	-8.862	-1.664	71.68	54.29	13.78
	Lys15DfeGly	0.011	-0.001	-8.238	-2.841	74.75	54.61	14.25
	Lys15TfeGly	0.011	-0.023	-7.691	-0.707	73.80	58.33	14.64

a. in 10^3 deg cm^2 $dmol^{-1}$ $residue^{-1}$, **b.** T_{den} is defined as the temperature at which the faction of folded is 0.5. Errors are typically not higher than 0.5 °C, **c.** Errors are typically not higher than 1 kcal/mol, **d.** $\Delta G°$ values were calculated for 1M standard at 25 °C, errors are not higher than 0.2 kcal/mol.

BPTI undergoes unfolding with a very small heat capacity change. An approximate value for ΔC_p was calculated afterwards from the ΔH_m against the melting temperature T_m (Van't Hoff plot). The slop of the plot was calculated to be 3.29 kcal $mol^{-1}K^{-1}$ for denaturation in the presence of 6 M GdmCl, and 1.31 kcal $mol^{-1}K^{-1}$ for denaturation in presence of 8 M urea, respectively (literature value: 3.0 at 25 °C, 2.6 at 50 °C, 2.1 at 75 °C, 1.3 at 100 °C).[5]

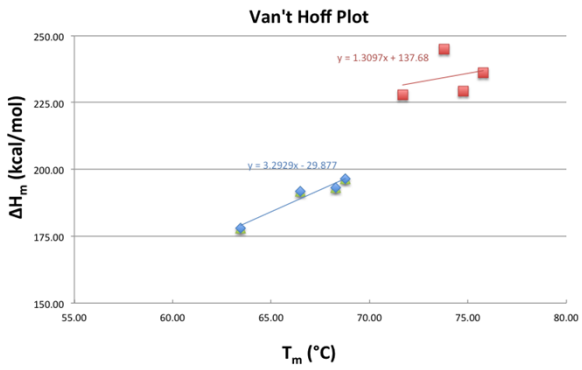


Fig. S13 Van't Hoff Plot of ΔH_m against T_m for BPTI species, the slope of which yield ΔC_p; denaturation in the presence of 6 M GdmCl (blue), denaturation in the presence of 8 M urea (red).

Enzymatic inhibitory assay

The inhibition capability of natural BPTI and synthetic BPTIs was tested by an enzyme inhibitory assay. N_α-Benzoyl-L-arginine 4-nitroanilide hydrochloride (Sigma-Aldrich) was dissolved in 1:1 (v:v) acetone and water to final concentration of 1 mM. The enzymatic activity and assay condition was tested with wild-type BPTI. Inhibitory assays were done by means of continuous spectrophotometric determination. Increasing concentration of BPTI were added to the constant concentration of proteases and well mixed (trypsin assay buffer contains 0.046 M Tris-HCl, pH 8.1 with 0.0115 M $CaCl_2$). The assay buffer conditions are based on the protocol from Sigma-Aldrich and Worthington Enzyme Manual. After incubation for 45 min, enzyme substrate was added into the mixture (rapidly mixed). Absorbance was recorded every 15 seconds over a period of 3 minutes by a BioPhotometer Eppdendorf, at 260 nm for the chymotrypsin assay and 405 nm for the plasmin and trypsin assays. The residual activity of enzyme was calculated

using standard enzyme assay procedures. The exctinction coefficient ε = 8800 $M^{-1}cm^{-1}$ was used for the trypsin assay.[8] The residual activity of enzyme was calculated by means of a standard inhibition assay protocol.

In order to calculate the association constant, the experimental data were fitted to the following equation.[9,10]

$$[E] = 0.5 \times \left\{ [E_0] - F \times [I_0] - K_a^{-1} + \sqrt{\left([E_0] + F \times [I_0] + K_a^{-1}\right)^2 - 4 \times [E_0] \times F \times [I_0]} \right\}$$

Where the $[E_0]$ and $[I_0]$ are total concentration of enzyme and inhibitor, $[E]$ is the residual concentration of enzyme, and F is the enzyme-inhibitor equimolarity factor ($F = 1$).

Protein crystallography

For protein crystallographic purposes trypsin, wild-type BPTI, and three synthetic BPTI species were extensively dialyzed in a buffer containing 25 mM Tris and 10 mM $CaCl_2$ pH 7.4 for 3 days, during which the buffer was changed twice. Trypsin was mixed with BPTI in a molar ratio of 1:1.1 to enable the formation of the complexes. The final concentrations were as follows: 56.0 mg/mL for β-trypsin-BPTI (wild-type), 25.6 mg/mL for β-trypsin-BPTI (Lys15Abu), 54.6 mg/mL for β-trypsin-BPTI (Lys15DfeGly), and 31.4 mg/mL for β-trypsin-BPTI (Lys15TfeGly). Crystal screening was performed on Cartesian Dispensing System (GENOMIC Solution®, a Harvard Bioscience Company) by means of sitting drop vapor diffusion technique using the $AmSO_4$ Suite Kit (*Qiagen*) in 96-well plates. Drops contained 100 nL of protein solution and 100 nL reservoir solution and incubated at room temperature. Well-formed crystals were soaked in 30% glycol plus reservoir solution and frozen in liquid nitrogen. Data was collected in Berlin at BESSYII, beamline 14-2.

Tab. S3 Data collection and refinement statistics

Data collection	Trypsin/BPTI wt	Trypsin/BPTI Abu	Trypsin/BPTI DfeGly	Trypsin/BPTI TfeGly
PDB entry	4Y0Y	4Y0Z	4Y10	4Y11
Space group	*I*222	*I*222	*I*222	*I*222
Wavelength [Å]	0.91841	0.91841	0.91841	0.91841
Unit cell a; b; c [Å]	74.8; 81.5; 124.2	74.9; 82.3; 123.4	74.7; 81.9; 123.6	75.0; 82.1; 123.6
α; β; γ [°]	90.0; 90.0; 90.0	90.0; 90.0; 90.0	90.0; 90.0; 90.0	90.0; 90.0; 90.0
Resolution [Å] [a]	50.00 - 1.25 (1.28 - 1.25)	50.00 - 1.37 (1.45 - 1.37)	50.00 - 1.37 (1.48 - 1.37)	50.00 - 1.30 (1.38 - 1.30)
Unique reflections	104119 (7297)	79271 (12329)	79871 (12620)	92992 (14647)
Completeness [a]	99.4 (94.8)	98.4 (95.6)	99.6 (98.6)	99.3 (97.9)
$\langle I/\sigma(I)\rangle$ [a]	18.7 (1.7)	16.4 (1.9)	19.2 (2.4)	16.8 (1.9)
R_{meas} [a, b]	0.043 (0.880)	0.067 (1.118)	0.072 (1.004)	0.052 (1.920)
CC1/2 [a]	100.0 (66.2)	99.9 (69.6)	100.0 (81.8)	99.9 (74.4)
Redundancy [a]	4.8 (3.3)	5.9 (5.9)	8.0 (8.0)	4.4 (4.1)
Refinement				
Non-hydrogen atoms	2698	2660	2601	2605
R_{work} [a, c]	0.138 (0.305)	0.134 (0.243)	0.138 (0.229)	0.134 (0.264)
R_{free} [a, d]	0.155 (0.319)	0.156 (0.285)	0.165 (0.263)	0.156 (0.292)
Average B-factor [Å^2]	20.7	21.5	19.4	21.0
Trypsin residues	243 / 21.6	243 / 21.8	243 / 20.2	243 / 22.5
BPTI residues	58 / 19.2	57 / 18.9	57 / 17.6	57 / 19.4
Water molecules	344 / 30.90	368 / 33.8	346 / 31.1	301 / 31.8
Buffer molecules	10 / 35.00	13 / 33.9	11 / 31.3	12 / 31.0
r.m.s.d. [e] bond length [Å]	0.016	0.012	0.017	0.017
bond angles [°]	1.721	1.421	1.690	1.687
Ramachandran outliers [%]	0.0	0.0	0.0	0.0
Ramachandran favoured [%]	98.7	98.4	98.6	98.3

[a] values in parentheses refer to the highest resolution shell.
[b] $R_{meas} = \Sigma_h [n/(n-1)]^{1/2} \Sigma_i \left| I_h - I_{h,i} \right| / \Sigma_h \Sigma_i I_{h,i}$.where I_h is the mean intensity of symmetry-equivalent reflections and n is the redundancy.
[c] $R_{work} = \Sigma_h \left| F_o - F_c \right| / \Sigma F_o$ (working set, no σ cut-off applied).
[d] R_{free} is the same as R_{cryst}, but calculated on 5% of the data excluded from refinement.
[e] Root-mean-square deviation (r.m.s.d.) from target geometries.

Tab. S4 Summary of distances between structural water molecules themselves and residues in the S_1 pocket of wild-type BPTI-trypsin complex

		A'	**D'**	**E'**
Structural water molecules	**A'**		7.1	5.0
	D'			4.9
	E'			
BPTI	$Lys15^{NZ}$	2.9	5.1	2.8
	$Pro13^{O}$		3.0	
β-trypsin	$Asp189^{OD2}$			2.7
	$Ser190^{OG}$	3.1		
	$Gln192^{NE2}$		3.0	
	$Trp211^{O}$	3.1		
	$Ser213^{N}$			4.0
	$Gly214^{O}$		2.9	2.9
	$Lys220^{O}$			3.8
	$Val227^{O}$	2.8		
	$Val227^{N}$	3.3		

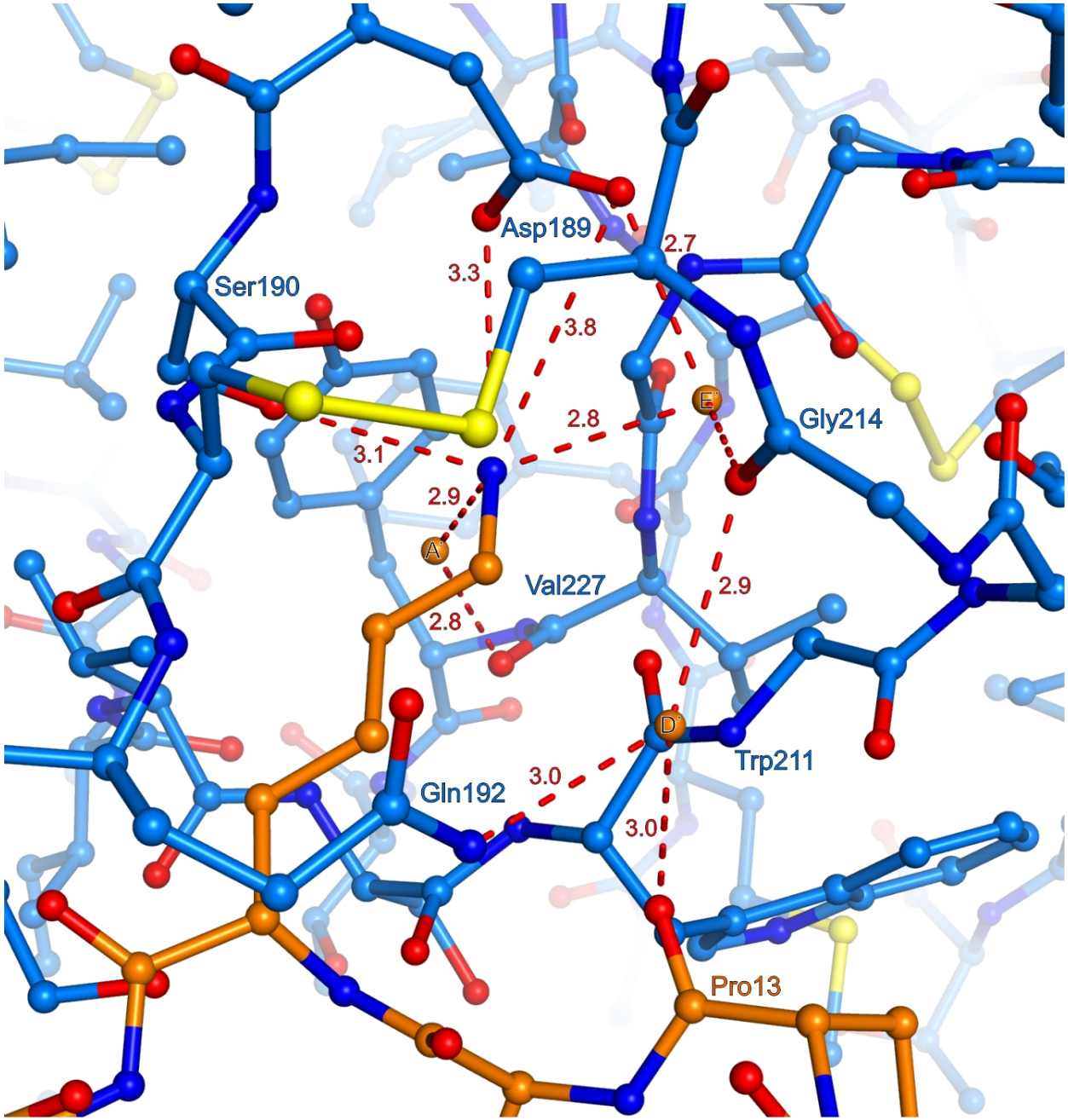


Fig. S14 schematic representation of hydrogen bond network of binding site in β-trypsin-BPTI (wild-type) complex. β-Trypsin is drawn in blue and BPTI in orange. Water molecules within the binding site are labeled A to E and drawn as orange spheres.

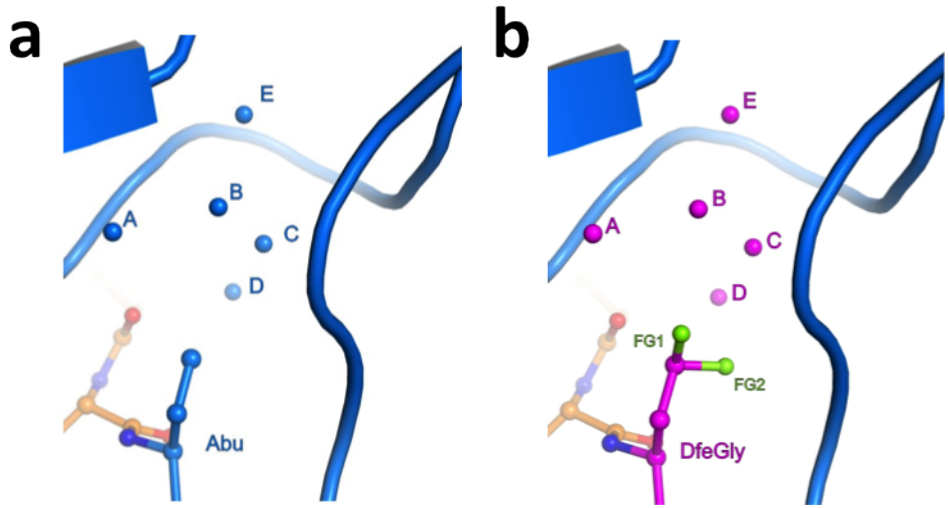


Fig. S15 Structural context of the P_1 side chain in S_1 binding pocket. β-Trypsin is shown in cartoon representation and the modified BPTI in stick representation. Structural water molecules were assigned unique identifiers A to E. **a**) Lys15Abu highlighted in light blue **b**) Lys15DfeGly shown in magenta with its two fluorine atoms in green.

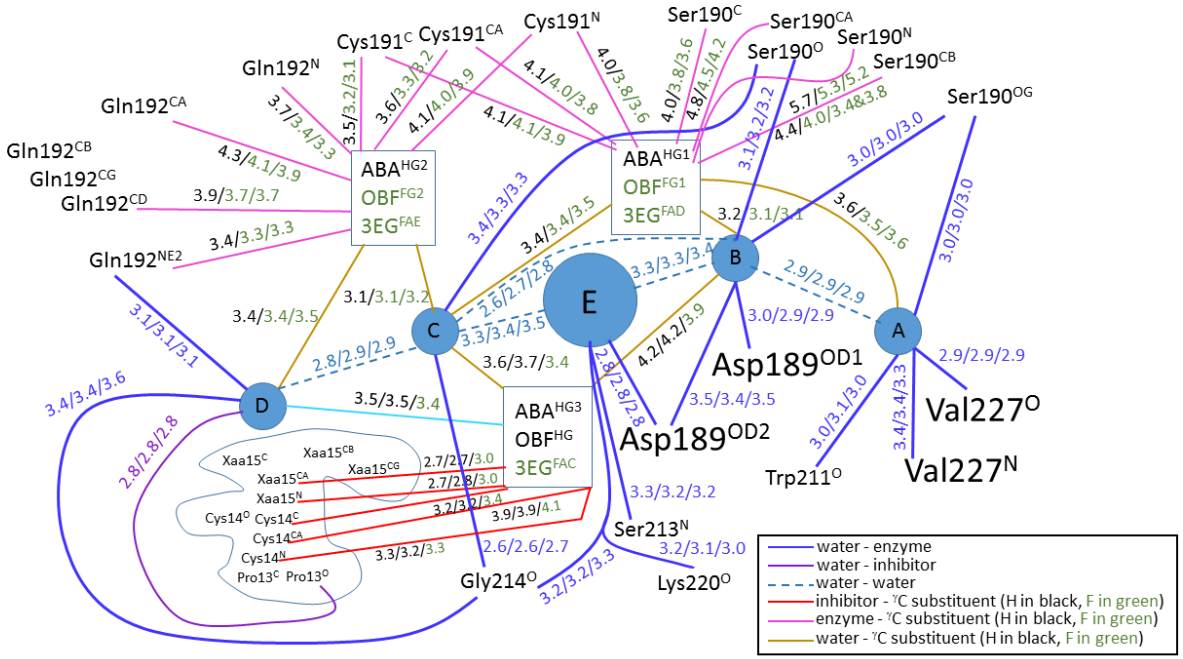
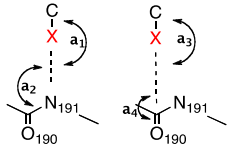
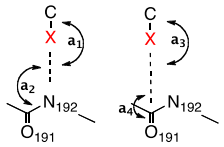
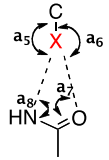
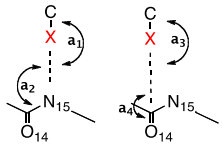


Fig. S16 Diagrammatic overview (top view, as in Figure 5 in main text) of S1 subpocket contacts, based on high resolution crystal data, summarizing all water-enzyme, water-inhibitor, water-water, inhibitor-$^{\gamma}$C substituent, enzyme-$^{\gamma}$C substituent, and water-$^{\gamma}$C substituent contacts shorter than 4 Å (distances greater than this cut off are shown only in cases in which the comparison demands it). Hydrogen (black font) or fluorine (green font) substituents of the terminal $^{\gamma}$C group of Abu (ABA^{HG1}, ABA^{HG2}, and ABA^{HG3}), DfeGly (OBF^{FG1}, OBF^{FG2}, and OBF^{HG}), and TfeGly ($3EG^{FAD}$, $3EG^{FAE}$, and $3EG^{FAC}$) are represented by open boxes, structural waters by blue circles.

Tab. S5 Analysis of distances and angles between P_1 side chains to peptide backbone

			Atom identifier of BPTI residue 15								
			Abu			DfeGly			TfeGly		
			ABA^{HG1}	ABA^{HG2}	ABA^{HG3}	OBF^{FG1}	OBF^{FG2}	OBF^{HG}	3EG^{FAD}	3EG^{FAE}	3EG^{FAC}
Structural water molecules											
	A		**3.6**	5.1	4.3	**3.5**	5.5	4.5	**3.6**	5.5	4.3
	B		**3.2**	4.2	4.2	**3.1**	4.3	4.2	**3.1**	4.4	3.9
	C		**3.4**	**3.1**	**3.6**	**3.4**	**3.1**	**3.7**	**3.5**	**3.2**	**3.4**
	D		4.4	**3.4**	**3.5**	4.6	**3.4**	**3.5**	4.8	**3.5**	**3.4**
	E		5.9	6.2	6.2	6.0	6.4	6.3	6.2	6.5	6.1
Atom identifier of protein residues											
Distances (Å)	**Backbone**										
	β-trypsin	Asp189^{C}	6.8	7.5		6.4	7.3		6.3	7.3	
		Ser190^{N}	**5.7**	6.5		**5.3**	6.4		**5.2**	6.4	
		Ser190^{CA}	**4.8**	5.6		**4.5**	5.5		**4.2**	5.4	
		Ser190^{C}	**4.0**	4.5		**3.8**	4.3		**3.6**	4.3	
		Cys191^{N}	**4.0**	4.1		**3.8**	4.0		**3.6**	3.9	
		Cys191^{CA}	4.1	**3.6**		4.0	**3.3**		3.8	**3.2**	
		Cys191^{C}	4.1	**3.5**		4.1	**3.2**		3.9	**3.1**	
		Gln192^{N}	4.8	**3.7**		4.9	**3.4**		4.7	**3.3**	
		Gln192^{CA}	5.4	**4.3**		5.6	**4.1**		5.4	**3.9**	
	BPTI	Xaa15^{CA}			**2.7**			**2.7**			**3.0**
		Xaa15^{N}			**2.7**			**2.8**			**3.0**
		Cys14^{C}			**3.2**			**3.2**			**3.4**
		Cys14^{CA}			**3.9**			**3.9**			**4.1**
		Pro13^{O}			**3.3**			**3.2**			**3.3**
	Side chain	Ser190^{CB}	4.4	5.5		4.0	5.5		3.4/3.8	5.2/5.5	
		Cys191^{CB}	5.5	5.0		5.4	4.6		5.3	4.5	

	$Gln192^{CB}$	6.5	5.1		6.6	4.9		6.6	4.8	
	$Gln192^{CG}$	6.2	4.7		6.5	4.5		6.5	4.4	
	$Gln192^{CD}$	5.5	**3.9**		5.8	**3.7**		5.8	**3.7**	
	$Gln192^{NE}$	4.9	**3.4**		5.3	**3.3**		5.4	**3.3**	
Angle (°)										
Ser190-Cys191 backbone										
	a_1	112.2	102.4		112.9	105.8		112.2	98.0	
	a_2	82.1	96.2		77.5	96.5		78.5	98.4	
	a_3	126.4	96.6		129.4	99.4		128.9	92.0	
	a_4	80.6	75.3		82.7	75.5		83.3	75.3	
Cys191-Gln192 backbone										
	a_1	63.8	132.8		64.5	134.9		62.7	131.7	
	a_2	53.1	68.9		49.3	69.0		45.7	68.3	
	a_3	77.8	118.8		77.8	121.5		75.9	115.3	
	a_4	81.5	96.0		79.5	98.9		76.1	97.2	
Gln192 Side chain 	a_5	31.6	137.9		31.9	125.3		29.9	130.7	
	a_6	50.8	165.0		51.5	157.4		49.5	163.2	
	a_7	109.8	106.4		106.3	99.6		103.0	96.9	
	a_8	67.0	61.6		71.1	65.6		73.1	66.4	
Cys14-Xaa15 backbone										
	a_1			100.0			96.8			84.9
	a_2			100.0			98.6			96.5
	a_3			114.4			112.3			100.3
	a_4			107.1			105.9			103.8

* Side chain atoms of natural residues numbered according to IUPAC; side chains of noncanonical amino acids named according to identifier existing in PDB: ABA=Abu, OBF=DfeGly, 3EG=TfeGly; distances given between fluorine atom and carbon of main chain or side chain carbonyl group. Values in bold are the seven side chain-water and fifteen side chain-enzyme contacts shown in Figures 6 and 5, respectively, in the main text.

Tab. S6 Distances between P_1 side chain atoms and structural water molecules in the S_1 pocket, distances between water molecules, and distances between water molecules and amino acid side chains of trypsin

		Structural water molecules														
		A			B			C			D			E		
		Abu	**DfeGly**	**TfeGly**	**Abu**	**DfeGly**	**TfeGly**	**Abu**	**DfeGly**	**TfeGly**	**Abu**	**DfeGly**	**TfeGly**	**Abu**	**DfeGly**	**TfeGly**
Structural water molecules	**A**	-	-	-	2.9	2.9	2.9	5.0	5.2	5.2	6.7	6.9	6.8	5.5	5.7	5.7
	B	2.9	2.9	2.9	-	-	-	2.6	2.7	2.8	5.1	5.1	5.2	3.3	3.4	3.4
	C	5.0	5.2	5.2	2.6	2.7	2.8	-	-	-	2.8	2.9	2.9	3.3	3.4	3.5
	D	6.7	6.9	6.8	5.1	5.1	5.2	2.8	2.9	2.9	-	-	-	5.2	5.2	5.4
BPTI	Pro13^{O}										2.8	2.8	2.8			
β-trypsin	Asp189^{OD1}				3.0	2.9	2.9									
	Asp189^{OD2}				3.5	3.4	3.5							2.8	2.8	2.8
	Ser190^{O}				3.1	3.2	3.2	3.4	3.3	3.3						
	Ser190^{OG}	3.0	3.0	3.0	3.0	3.0	3.0									
	Gln192^{NE2}										3.1	3.1	3.1			
	Trp211^{O}	3.0	3.1	3.0												
	Ser213^{N}													3.3	3.2	3.2
	Gly214^{O}							2.6	2.6	2.7	3.4	3.4	3.6	3.2	3.3	3.3
	Lys220^{O}													3.2	3.1	3.0
	Val227^{O}	2.9	2.9	2.9												
	Val227^{N}	3.4	3.4	3.3												

Tab. S7 Conformational analysis of noncanonical P_1 side chains

	Abu		DfeGly		TfeGly	
Bond length (Å)	C_γ-HG1	0.98	C_γ-FG1	1.32	C_γ-FAD	1.33
	C_γ-HG2	0.98	C_γ-FG2	1.34	C_γ-FAE	1.34
	Cγ-HG3	0.98	Cγ-HG	0.98	C_γ-HAC	1.35
	C_β-HB2	0.98	C_β-H1B	0.98	C_β-HBA	1.12
	C_β-HB3	0.98	C_β-H2B	0.98	C_β-HB	1.12
	C_β-C_γ	1.48	C_β-C_γ	1.48	C_β-C_γ	1.51
	C_α-C_β	1.54	C_α-C_β	1.53	C_α-C_β	1.53
Bond angle (°)	HG1-C_γ-HG2	109.47	FG1-C_γ-FG2	100.24	FAD-C_γ-FAE	107.00
	HG1-C_γ-HG3	109.41	FG1-C_γ-HG	112.90	FAD-C_γ-FAC	106.15
	HG2-C_γ-HG3	109.45	FG2-C_γ-HG	112.87	FAE-C_γ-FAC	104.75
	HG1-C_γ-C_β	109.51	FG1-C_γ-C_β	107.87	FAD-C_γ-Cβ	111.08
	HG2-C_γ-C_β	109.51	FG2-C_γ-C_β	108.80	FAE-C_γ-Cβ	114.44
	HG3-C_γ-C_β	109.46	HG-C_γ-C_β	113.26	FAC-C_γ-Cβ	112.84
	HB2-C_β-C_γ	109.10	H1B-C_β-C_γ	109.26	HBA-C_β-C_γ	107.94
	HB2-C_β-C_α	109.02	H1B-C_β-C_α	109.36	HBA-C_β-C_α	109.59
	HB3-C_β-C_γ	109.10	H2B-C_β-C_γ	108.98	HB-C_β-C_γ	107.95
	HB3-C_β-C_α	109.00	H2B-C_β-Cα	108.90	HB-C_β-C_α	110.94
	HB2-C_β-HB3	109.07	HB2-C_β-HB3	109.16	HBA-C_β-HB	108.87
	C_α-C_β-C_γ	111.52	C_α-C_β-C_γ	111.09	C_α-C_β-C_γ	111.46

Tab. S8 Analysis of potential interactions induced by methylene group of P_1 noncanonical residues

	Abu		DfeGly		TfeGly	
	HB2	HB3	H1B	H2B	HBA	HB
Cys191^{O}		2.7		2.8		2.6
Asp194^{N}		4.0		4.0		3.7
Ser195^{OG}	2.5	3.5	2.5	3.4	2.4	3.3
Ser195^{N}	3.6	3.5	3.6	3.4	3.6	3.2
Ser210^{O}	3.4		3.3		3.2	
ABA15^{O}/OBF15^{O}/3EG^{O}	3.2	2.5	3.2	2.5	3.3	2.3
ABA15^{N}/OBF15^{N}/3EG^{N}	2.6	3.3	2.6	3.3	2.6	3.4
ABA15^{HG1}/OBF15^{FG1}/3EG15^{FAD}	2.3	2.3	2.5	2.5	2.6	2.6
ABA15^{HG2}/OBF15^{FG2}/3EG15^{FAE}	2.8	2.3	3.1	2.4	3.3	2.7
ABA15^{HG3}/OBF15^{HG}/3EG15^{FAC}	2.3	2.8	2.4	2.8	2.7	3.3

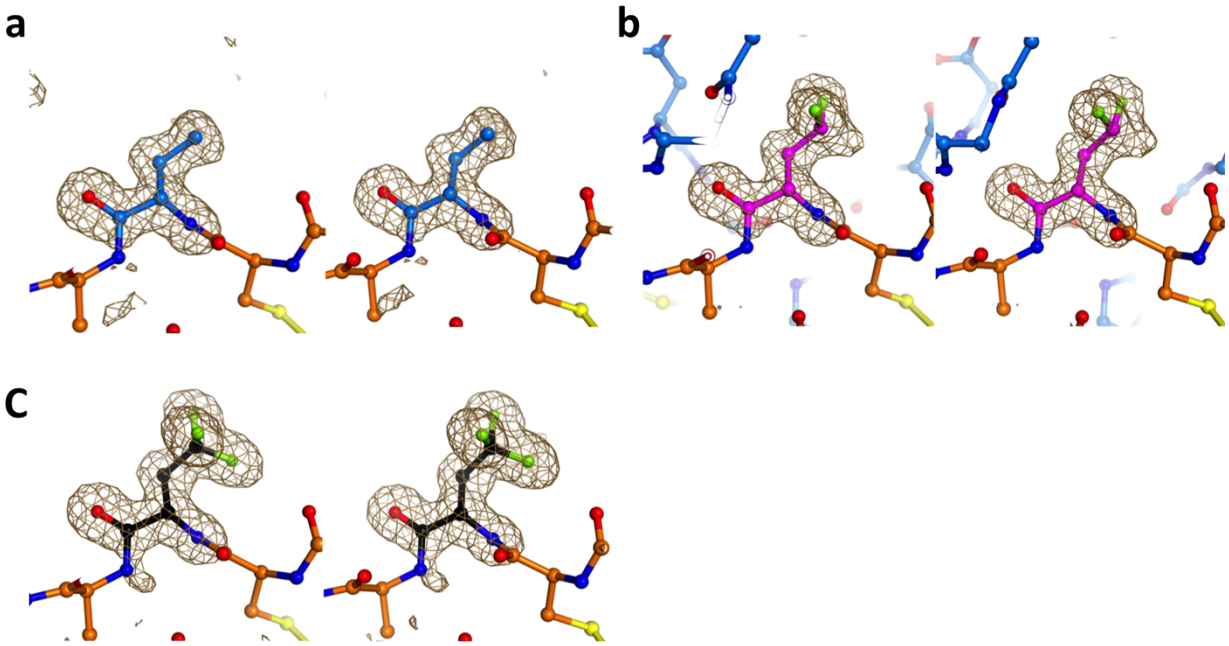


Fig. S17 Stereoview of the P_1 residue of BPTI that has been omitted for calculation of a simulated annealing omit map contoured at 3.0 σ shown as brown mesh. **a**) Lys15Abu. **b**) Lys15DfeGly and **c**) Lys15TfeGly

Table S9: Distances of fluorine atoms within the S1 binding pocket. Distance cut-off ≤ 3.5 Å.

Fluorine atom	**Hydrogen atom**	**Distance**
DfeGly		
FG2	Cys191^{HA}	2.7
	Gln192^{HE21}	3.5
	Gln192^{HE22}	3.2
FG1	Cys191^{HA}	3.5
	Val209^{HG12}	3.4
	Val209^{HG13}	3.4
	OBF^{HA}	2.5
TfeGly		
FAC	Trp211^{HA}	2.6
	3EG^{H}	3.2
	3EG^{HA}	3.0
FAD	Ser190^{HB2}	3.1
	Ser190^{HB3}	3.0
	Cys191^{HA}	3.4
	Val209^{HG11}	3.5
	Val209^{HG12}	3.4
	Val209^{HG13}	3.3
FAE	Cys191^{HA}	2.6
	Gln192^{H}	3.5
	Gln192^{HE21}	3.5
	Gln192^{HE22}	3.2
	3EG^{HA}	2.4

Reference:

1 Winkler, D. B., K. Synthesis of Enantiomerically Pure D- and L-Armentomycin and its Difluoro Analogues from Aspartic Acid. *Synthesis*, 1419-1421, doi:10.1055/s-1996-4409 (1996).

2 Tsushima, T. *et al.* Fluorine containing amino acids and their derivatives. 7. Synthesis and antitumor activity of α- and γ-substituted methotrexate analogs. *Tetrahedron* **44**, 5375-5387, doi:http://dx.doi.org/10.1016/S0040-4020(01)86044-0 (1988).

3 Barany, G. *et al.* in *Techniques in Protein Chemistry* Vol. Volume 7 (ed R. Marshak Daniel) 503-514 (Academic Press, 1996).

4 Sewing, A. & Hilvert, D. Fmoc-Compatible Solid-Phase Peptide Synthesis of Long C-Terminal Peptide Thioesters. *Angewandte Chemie International Edition* **40**, 3395-3396, doi:10.1002/1521-3773(20010917)40:18<3395::aid-anie3395>3.0.co;2-g (2001).

5 Makhatadze, G. I., Kim, K.-S., Woodward, C. & Privalov, P. L. Thermodynamics of bpti folding. *Protein Science* **2**, 2028-2036, doi:10.1002/pro.5560021204 (1993).

6 Greenfield, N. J. Analysis of the kinetics of folding of proteins and peptides using circular dichroism. *Nat. Protocols* **1**, 2891-2899, doi:http://www.nature.com/nprot/journal/v1/n6/suppinfo/nprot.2006.244_S1.html (2007).

7 Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protocols* **1**, 2876-2890, doi:http://www.nature.com/nprot/journal/v1/n6/suppinfo/nprot.2006.202_S1.html (2007).

8 Gravett, P. S., Viljoen, C. C. & Oosthuizen, M. M. J. A steady-state kinetic analysis of the reaction between arginine esterase E-I from Bitis gabonica venom and synthetic arginine substrates and the influence of pH, temperature and solvent deuterium isotope. *International Journal of Biochemistry* **23**, 1085-1099, doi:http://dx.doi.org/10.1016/0020-711X(91)90149-H (1991).

9 Buczek, O., Koscielska-Kasprzak, K., Krowarsch, D., Dadlez, M. & Otlewski, J. Analysis of serine proteinase–inhibitor interaction by alanine shaving. *Protein Science* **11**, 806-819, doi:10.1110/ps.3510102 (2002).

10 Krowarsch, D. *et al.* Interscaffolding additivity: binding of P1 variants of bovine pancreatic trypsin inhibitor to four serine proteases. *Journal of Molecular Biology* **289**, 175-186, doi:http://dx.doi.org/10.1006/jmbi.1999.2757 (1999).