Supplementary Material

Novel selective inhibitors of aminopeptidases that generate antigenic peptides

Athanasios Papakyriakou[a,b], Efthalia Zervoudi[c], Emmanuel A. Theodorakis[a], Loredana Saveanu[d], Efstratios Stratikos[c], Dionisios Vourloumis[b*]

[a] *Department of Chemistry & Biochemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0358, USA*
[b] *Chemical Biology and* [c] *Protein Chemistry Laboratories, National Center for Scientific Research "Demokritos", Agia Paraskevi Attikis, GR-15310, Greece*
[d] *Institut National de la Santé et de la Recherché Médicale, Unité 1013 and Université Paris Descartes, Sorbonne Paris Cité, Faculté de médecine, 75015 Paris, France,*

***Corresponding author:** Tel. +30 2106503624, E-mail. vourloumis@chem.demokritos.gr

Table of contents

Chemistry …… 2

Recombinant enzyme production and purification ……. 9

Enzyme activity assay ……. 10

Figure S1 ……. 11

HPLC analysis ……. 12

Figure S2 ……. 12

Computational methods …….. 13

Supplementary references ……... 14

Chemistry

Unless otherwise noted, all solvents and reagents for organic synthesis were obtained from commercial suppliers and were used without further purification. All reactions were carried out under a dry Ar atmosphere with anhydrous, freshly distilled solvents under anhydrous conditions unless otherwise noted. All reactions were stirred with Teflon-coated magnetic stir bars, and temperatures were measured externally. Reactions requiring anhydrous conditions were carried out in oven-dried (120 ^{o}C, 24 h) or flame-dried (vacuum <0.5 Torr) glassware. Yields refer to chromatographically and spectroscopically (^{1}H NMR) homogeneous materials. Reactions were monitored by thin-layer chromatography (TLC) carried out on 0.25 mm silica gel plates (60 F254, E. Merck). Silica gel (60, particle size: 0.040–0.063 mm, E. Merck) was used for flash column chromatography.

NMR spectra were recorded on a Bruker Avance DRX-500 instrument at 298 K using deuterated solvents as internal standards: [$CDCl_3$, 7.26 ppm (^{1}H) and 77.2 ppm (^{13}C); MeOD, 3.31 ppm (^{1}H) and 49.0 ppm (^{13}C)]. Chemical shift values (δ) are given within an accuracy of 0.01 ppm for ^{1}H NMR and 0.1 ppm for ^{13}C NMR data, while the coupling constants (J) are within 0.1 Hz. Multiplicities are designated as singlet (s), doublet (d), triplet (t), or multiplet (m). 2D NMR ^{1}H–^{1}H COSY and ^{1}H–^{13}C HMQC correlation spectra were also recorded at 298 K in order to assist the peak assignment. Quantitation of the final compounds was achieved using an internal standard of 2,5-dimethylfuran (DMFu, 0.1 mM in MeOD).[S1] High-resolution mass spectra (HRMS) were measured on Agilent 6224 Accurate Mass TOF LC/MS at the Faculty of Chemistry and Chemical Technology, University of Ljubljana.

(S)-methyl 4-amino-3-(2-(tert-butoxycarbonylamino)-4-phenylbutanamido)benzoate (3).
A mixture of methyl 3,4-diaminobenzoate (500 mg, 3.0 mmol) and (S)-2-(tert-butoxycarbonylamino)-4-phenylbutanoic acid (920 mg, 3.3 mmol), *O*-(benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate (HBTU, 2.27 g, 6.0 mmol) and *N*,*N*-diisopropylethylamine (DIEA, 1.57 mL, 9.0 mmol) in anhydrous DMF (12 mL, 0.25 mM) was stirred under Ar atmosphere at ambient temperature for 4 h. The mixture was then diluted in EtOAc (60 mL) and washed sequentially with aqueous HCl 1.0 N (20 mL), sat. $NaHCO_3$ (20 mL) and sat. NaCl (30 mL). Consequently, the solvent was evaporated under reduced pressure to provide the crude product, which was eluted from a gradient of 40% to 80% EtOAc in hexanes to yield **3** as colorless oil (1.05 g, 82%). R_f=0.55 (50% EtOAc/Hexanes). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.45 (1H, s), 7.78 (1H, s), 7.65 (1H, d, *J*=8.4 Hz), 7.26–7.20 (2H, m), 7.19–7.10 (3H, m), 6.58 (1H, d, *J*=8.4 Hz), 5.84 (1H, d, *J*=7.0 Hz), 4.30 (1H, s.br), 3.79 (3H, s), 2.78–2.64 (2H, m),

2.22–2.12 (1H, m), 2.04–1.94 (1H, m), 1.42 (9H, s). ^{13}C NMR ($CDCl_3$, 126 MHz) δ 171.9, 166.9, 156.5, 146.5, 140.8, 129.5, 128.6, 128.5 128.4, 126.2, 121.5, 119.3, 115.6, 80.5, 55.0, 51.7, 33.8, 32.0, 28.4

(S)-4-amino-3-(2-(tert-butoxycarbonylamino)-4-phenylbutanamido)benzoic acid (4).

Compound **3** (470 mg, 1.1 mmol) was dissolved in 20 mL dioxane and then 22 mL aqueous LiOH 1.0 M (22 mmol, 20 eq.) were added under rigorous stirring. After 4 h at ambient temperature the mixture was diluted with 8 mL aqueous sat. NaCl and the organic solvent was evaporated. The remaining aqueous solution (pH~12) was extracted with EtOAc (20 mL) and then acidified with HCl 1.0 N until no further precipitation was observable (pH~2). The aqueous mixture was then extracted with EtOAc (2 × 30 mL) and the combined organic layers were dried with $MgSO_4$, filtered and washed with EtOAc. After evaporating the solvents and drying under vacuum, a colorless oil was obtained (415 mg, 92%) as the pure product **4** (>90% by NMR) and used without any further purification. R_f=0.75 (EtOAc). ^{1}H NMR (MeOD, 500 MHz) δ 7.74 (1H, s), 7.70 (1H, dd, *J*=8.5, 1.7 Hz), 7.29-7.14 (5H, m), 6.77 (1H, d, *J*=8.5 Hz), 4.15-4.09 (1H, m), 2.85-2.67 (2H, m), 2.18-1.96 (2H, m), 1.47 (9H, s). ^{13}C NMR ($CDCl_3$, 126 MHz) δ 174.8, 170.1, 158.3, 150.0, 142.3, 131.1, 130.8, 129.5, 129.5, 127.1, 122.0, 119.3, 115.8, 80.9, 56.5, 34.7, 33.2, 28.7.

General procedure for coupling of 4 with *O*-protected amino acids (Method A).

A mixture of **4** (50 mg, 0.12 mmol), the *O*-protected amino acid (0.18 mmol, 1.5 eq.), HBTU (95 mg, 0.25 mmol) and DIEA (70 µL, 0.40 mmol) in anhydrous DMF (0.5 mL, 0.25 mM) was stirred under Ar atmosphere at ambient temperature for 12 h. The reaction mixture was then diluted in EtOAc (25 mL) and washed sequentially with aqueous HCl 1.0 N (10 mL), sat. $NaHCO_3$ (10 mL) and sat. NaCl (15 mL). Consequently, the solvent was evaporated under reduced pressure to provide the crude product, which was purified using a gradient of EtOAc in hexanes to yield the desired product in yields of 72–92%.

(S)-methyl 2-(4-amino-3-((S)-2-(tert-butoxycarbonylamino)-4-phenylbutanamido)benzamido) propanoate (5).

Reaction of **4** with (S)-methyl 2-aminopropanoate (L-Ala-OMe) using Method A afforded **5** in yield of 92%. R_f=0.45 (70% EtOAc/Hexanes). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.64 (1H, s), 7.48-7.40 (2H, m), 7.32-7.25 (2H, m), 7.23-7.12 (4H, m), 6.60 (1H, d, *J*=8.2 Hz), 5.92 (1H, s), 4.77-4.68 (1H, m), 4.32 (1H, s), 3.72, (3H, s), 2.23-2.16 (1H, m), 2.10-1.97 (1H, m), 1.48 (3H, d, *J* =

7.3 Hz), 1.46 (9H, s). ^{13}C NMR ($CDCl_3$, 126 MHz) δ 174.4, 171.9, 166.8, 156.9, 145.2, 140.9, 128.5, 128.5, 128.5, 128.5, 126.9, 126.1, 125.5, 123.0, 121.9, 116.3, 80.5, 54.9, 52.6, 48.5, 34.2, 32.2, 28.4, 28.4, 28.4, 17.9

(S)-benzyl 2-(4-amino-3-((S)-2-(tert-butoxycarbonylamino)-4-phenylbutanamido)benzamido)-3-methylbutanoate (6).

Reaction of **4** with (S)-benzyl 2-amino-3-methylbutanoate (L-Val-OBn) using Method A afforded **6** in yield of 86%. R_f=0.80 (70% EtOAc/Hexanes). ^{1}H NMR (MeOD, 500 MHz) δ 7.61-7.53 (2H, m), 7.39-7.14 (10H, m), 6.80 (1H, d, *J*=8.2 Hz), 5.20 (1H, d, *J*=12.2 Hz), 5.12 (1H, d, *J*=12.2 Hz), 4.48-4.43 (1H, m), 4.16-4.11 (1H, m), 2.85-2.69 (2H, m), 2.25-1.98 (3H, m), 1.47 (9H, s), 0.98 (3H, d, *J*=6.8 Hz), 0.95 (3H, d, *J*=6.8 Hz).^{13}C NMR (MeOD, 126 MHz) δ 174.8, 173.3, 170.2, 158.3, 148.7, 142.4, 137.2, 129.5, 129.4, 129.3, 128.7, 128.3, 127.1, 123.0, 122.1, 116.1, 80.8, 67.7, 60.2, 56.5, 34.8, 33.2, 31.7, 28.7, 19.6, 19.2

(2S,3R)-methyl 2-(4-amino-3-((S)-2-(tert-butoxycarbonylamino)-4-phenylbutanamido) benzamido)-3-hydroxybutanoate (7).

Reaction of **4** with (2S,3R)-methyl 2-amino-3-hydroxybutanoate (L-Thr-OMe) using Method A afforded **7** in yield of 90%. R_f=0.40 (70% EtOAc/Hexanes). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.98 (1H, s), 7.51 (1H, d, *J*=8.4 Hz), 7.44 (1H, s), 7.38-7.18 (7H, m), 6.66 (1H, d, *J*=8.4 Hz), 6.15 (1H, s), 4.78-4.72 (1H, m), 4.43-4.34 (2H, m), 3.76 (3H, s), 2.27-2.03 (2H, m), 1.44 (9H, s), 1.23 (3H, d, *J*=5.8 Hz).

^{13}C NMR ($CDCl_3$, 126 MHz) δ 172.4, 172.3, 167.6, 156.8, 145.2, 140.8, 128.5, 127.5, 126.2, 125.2, 122.9, 121.7, 116.6, 80.7, 68.1, 58.3, 54.9, 52.6, 34.3, 32.2, 28.3, 20.0

(S)-methyl 2-(4-amino-3-((S)-2-(tert-butoxycarbonylamino)-4-phenylbutanamido) benzamido)-6-(tert-butoxycarbonylamino)hexanoate (8).

Reaction of **4** with (S)-methyl 2-amino-6-(tert-butoxycarbonylamino)hexanoate (L-Lys(Boc)-OMe) using Method A afforded **8** in yield of 91%. R_f=0.45 (70% EtOAc/Hexanes). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.75 (1H, s), 7.46 (1H, d, *J*=8.4 Hz), 7.34-7.14 (7H, m), 6.61 (1H, d, *J*=8.4 Hz), 5.99 (1H, s), 4.81-4.66 (2H, m), 4.39 (1H, s), 3.70 (3H, s), 3.12-2.99 (2H, m), 2.84-2.69 (2H, m), 2.26-2.19 (1H, m), 1.94-1.73 (2H, m), 1.56-1.32 (23H, m). ^{13}C NMR ($CDCl_3$, 126 MHz) δ 174.0, 172.0, 167.0, 156.8, 156.3, 145.1, 140.9, 128.6, 127.3, 126.3, 125.2, 123.3, 122.0, 116.7, 80.7, 79.2, 54.9, 52.6, 52.5, 40.4, 34.5, 32.2, 31.8, 29.7, 28.5, 28.4, 23.0

(S)-methyl 2-(4-amino-3-((S)-2-(tert-butoxycarbonylamino)-4-phenylbutanamido) benzamido)-5-((Z)-2,3-bis(benzyloxycarbonyl)guanidino)pentanoate (9).

Reaction of **4** with (S,Z)-methyl 2-amino-5-(2,3-bis(benzyloxycarbonyl)guanidino)pentanoate (L-Arg$(Z)_2$-OMe) using Method A afforded **9** in yield of 72%. R_f=0.55 (70% EtOAc/Hexanes). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 9.38 (1H, s), 8.67 (1H, s), 8.53 (1H, s), 8.40 (1H, d, *J*=8.3 Hz), 7.92 (1H, s), 7.75 (1H, d, *J*=8.4 Hz), 7.48-7.09 (14H, m), 6.68 (1H, d, *J*=8.4 Hz), 5.87-5.75 (1H, m), 5.28-5.05 (4H, m), 4.77-4.71 (1H, m), 4.35-4.27 (1H, m), 4.04-3.93 (1H, m), 3.62 (3H, s), 2.82-2.67 (4H, m), 2.29-2.16 (2H, m), 1.91-1.60 (3H, m), 1.41 (9H, s). ^{13}C NMR ($CDCl_3$, 126 MHz) δ 172.3, 171.7, 167.0, 162.1, 156.6, 155.7, 151.6, 149.4, 145.2, 140.8, 135.1, 131.1, 129.6, 128.9, 128.8, 128.7, 128.5, 128.4, 128.3, 128.1, 128.0, 127.9, 126.9, 126.2, 125.8, 123.1, 121.1, 120.9, 116.0, 115.5, 111.5, 80.5, 69.0, 67.1, 54.9, 52.5, 52.3, 44.4, 34.1, 33.8, 32.0, 28.4, 25.1

(S)-methyl 2-(4-amino-3-((S)-2-(tert-butoxycarbonylamino)-4-phenylbutanamido) benzamido)-3-(4-tert-butoxyphenyl)propanoate (10).

Reaction of **4** with (S)-methyl 2-amino-3-(4-tert-butoxyphenyl)propanoate (L-Tyr(O-tBu)-OMe) using Method A afforded **10** in yield of 78%. R_f=0.50 (70% EtOAc/Hexanes). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.36 (1H, s), 7.44 (1H, s), 7.39-7.34 (1H, m), 7.33-7.27 (2H, m), 7.26-7.18 (3H, m), 7.08 (2H, d, *J*=8.2 Hz), 6.97-6.88 (3H, m), 6.68-6.58 (1H, m), 5.75-5.62 (1H, m), 5.04-4.97 (1H, m), 4.31 (1H, s), 3.71 (3H, s), 3.23-3.11 (2H, m), 2.86-2.73 (2H, m), 2.31-2.20 (1H, m), 2.10-1.98 (1H, m), 1.47 (9H, s), 1.33 (9H, s). ^{13}C NMR ($CDCl_3$, 126 MHz) δ 173.2, 171.5, 166.6, 156.6, 154.4, 145.0, 140.8, 131.2, 129.8, 128.7, 128.6, 126.8, 126.4, 125.7, 124.3, 123.6, 122.3, 116.5, 80.7, 78.6, 54.9, 54.0, 52.4, 37.3, 34.0, 32.2, 29.0, 28.5

(S)-benzyl 2-(4-amino-3-((S)-2-(tert-butoxycarbonylamino)-4-phenylbutanamido) benzamido)-3-(1H-indol-3-yl)propanoate (11).

Reaction of **4** with (S)-benzyl 2-amino-3-(1H-indol-3-yl)propanoate (L-Trp-OBn) using Method A afforded **11** in yield of 75%. R_f=0.45 (70% EtOAc/Hexanes). ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.56 (1H, s), 8.32 (1H, s), 7.47 (1H, d, *J*=7.8 Hz), 7.33-7.20 (10H, m), 7.19-7.04 (5H, m), 7.00 (1H, t, *J*=7.8 Hz), 6.85 (1H, d, *J*=6.8 Hz), 6.47 (1H, s), 6.19 (1H, d, *J*=8.4 Hz), 5.72 (1H, d, *J*=6.0) Hz, 5.13-5.00 (3H, m), 4.29-4.21 (1H, m), 3.37-3.26 (2H, m), 2.77-2.62 (2H. m), 2.21-2.11 (1H, m), 2.00-1.89 (1H, m), 1.41 (9H, s). ^{13}C NMR ($CDCl_3$, 126 MHz) δ 172.7, 172.0, 167.0, 156.6, 145.3, 140.9, 136.3, 135.5, 128.7, 128.7, 128.6, 128.5, 128.4, 127.5, 126.8, 126.3, 126.0, 123.5, 123.1, 122.1, 121.7, 119.5, 118.5, 116.1, 111.6, 109.7, 80.7, 67.3, 54.9, 53.4, 34.0, 32.1, 28.4, 27.6

General procedure for deprotection of *O*-Bn and *N*-Cbz group (Method B).

The desired compound (~0.1 mmol) was dissolved in MeOH (4 mL, ~0.025 mM) and the solution was degassed under Ar atmosphere. A catalytic amount of activated Pd/C 10% was added and degassed under Ar and then H_2. The mixture was stirred under H_2 at ambient temperature for 1 h, then filtered and washed with MeOH. The resulting product was acquired in yields of 90–95% and was used without any further purification.

General procedure for deprotection of *N*-Boc and *O*-tBu groups (Method C).

The desired compound (~0.1 mmol) was dissolved in a mixture of CH_2Cl_2/TFA (2:1 mL, ~0.03 mM) and stirred at ambient temperature for 30 min. Consequently, the solvents were evaporated and the product was treated with HCl 1.0 N (0.5 mL). The solvents were evaporated using toluene (2 × 2 mL) and dried under vacuum overnight to yield the hydrochloric salt of the terminal α-amine in yields of ~ 97% without any further purification.

(S)-methyl 2-(4-amino-3-((S)-2-amino-4-phenylbutanamido)benzamido)propanoate (12).

Deprotection of the *N*-Boc group of **5** using Method C yielded **12**. ^{1}H NMR (MeOD, 500 MHz) δ 7.72 (1H, s), 7.63 (1H, dd, J = 8.3, 1.8 Hz), 7.32–7.24 (4H, m), 7.21 (1H, m), 6.90 (1H, d, J = 8.3 Hz), 4.57 (1H, m), 4.22 (1H, m), 3.71 (3H, s), 2.83 (2H, *m*), 2.31 (1H, m), 2.26 (1H, m), 1.47 (3H, d, J = 7.2 Hz). ^{13}C NMR (MeOD, 126 MHz) δ 175.1, 169.6, 169.4, 146.9, 141.3, 129.7, 129.7, 129.3, 129.3, 128.5, 127.7, 127.5, 123.9, 122.5, 117.2, 54.8, 52.7, 50.1, 34.8, 32.2, 17.3 HRMS calcd for [M + H$^+$] $C_{21}H_{27}N_4O_4$, 399.2027; found: 399.2027

(S)-benzyl 2-(4-amino-3-((S)-2-amino-4-phenylbutanamido)benzamido)-3-methylbutanoate (13).

Deprotection of the *N*-Boc group of **6** using Method C yielded **13**. ^{1}H NMR (MeOD, 500 MHz) δ 7.68 (1H, s), 7.60 (1H, dd, J = 8.4, 1.8 Hz), 7.39–7.24 (9H, m), 7.21 (1H, m), 6.89 (1H, d, J = 8.4 Hz), 5.21 (1H, d, J = 12.2 Hz), 5.14 (1H, d, J = 12.2 Hz), 4.46 (1H, d, J = 6.9 Hz), 4.20 (1H, m), 2.83 (1H, m), 2.32 (1H, m), 2.29–2.18 (2H, m), 0.99 (3H, d, J = 6.8 Hz), 0.96 (3H, d, J = 6.8 Hz). ^{13}C NMR (MeOD, 126 MHz) δ 173.3, 170.1, 169.6, 147.8, 141.3, 137.2, 129.7, 129.7, 129.5, 129.5, 129.4, 129.4, 129.3, 129.3, 129.3, 128.6, 127.8, 127.6, 123.7, 122.0, 116.7, 67.8, 60.3, 54.8, 34.8, 32.2, 31.7, 19.6, 19.2. HRMS calcd for [M + H$^+$] $C_{29}H_{35}N_4O_4$ 503.2653; found 503.2648

(S)-2-(4-amino-3-((S)-2-amino-4-phenylbutanamido)benzamido)-3-methylbutanoic acid (14).

Deprotection of the *O*-Bn group of **6** using Method B, and subsequent deprotection of the *N*-Boc group using Method C yielded **14**. ^{1}H NMR (MeOD, 500 MHz) δ 7.69 (1H, s), 7.61 (1H, dd, J =

8.4, 1.8 Hz), 7.32–7.17 (5H, m), 6.86 (1H, d, *J* = 8.4 Hz), 4.47 (1H, d, *J* = 6.9 Hz), 4.21 (1H, m), 2.83 (1H, m), 2.37–2.21 (3H, m), 1.03 (6H, d, *J* = 6.8 Hz). ^{13}C NMR (MeOD, 126 MHz) *δ* 175.2, 169.9, 169.6, 147.3, 141.3, 129.7, 129.7, 129.3, 129.3, 128.5, 127.7, 127.6, 124.0, 122.2, 117.0, 59.8, 54.8, 34.8, 32.2, 31.8, 19.7, 18.9. HRMS calcd for [M + H$^+$] $C_{22}H_{29}N_4O_4$ 413.2183; found 413.2184

(2S,3R)-methyl 2-(4-amino-3-((S)-2-amino-4-phenylbutanamido)benzamido)-3-hydroxybutanoate (15).

Deprotection of the *N*-Boc group of **7** using Method C yielded **15**. ^{1}H NMR (MeOD, 500 MHz) *δ* 7.74 (1H, s), 7.64 (1H, dd, *J* = 8.3, 1.8 Hz), 7.33–7.25 (4H, m), 7.21 (1H, m), 6.90 (1H, d, *J* = 8.3 Hz), 4.66 (1H, m), 4.37 (1H, m), 4.19 (1H, m), 3.75 (3H, s), 2.85 (2H, m), 2.33 (1H, m), 2.27 (1H, m), 1.22 (3H, d, *J* = 6.4 Hz). ^{13}C NMR (MeOD, 126 MHz) *δ* 172.8, 169.8, 169.5, 147.7, 141.3, 129.7, 129.7, 129.3, 129.3, 128.4, 127.7, 127.6, 123.4, 122.2, 116.9, 68.6, 59.8, 54.8, 52.8, 34.8, 32.2, 20.5. HRMS calcd for [M + H$^+$] $C_{22}H_{29}N_4O_5$ 429.2133; found 429.2132

(S)-methyl 6-amino-2-(4-amino-3-((S)-2-amino-4-phenylbutanamido)benzamido)hexanoate (16).

Deprotection of both *N*-Boc groups of **8** using Method C yielded **16**. ^{1}H NMR (MeOD, 500 MHz) *δ* 7.73 (1H, s), 7.63 (1H, dd, *J* = 8.2, 1.8 Hz), 7.32–7.24 (4H, m), 7.21 (1H, t, *J* = 7.1 Hz), 6.89 (1H, d, *J* = 8.2 Hz), 4.59 (1H, m), 4.22 (1H, m), 3.72 (3H, s), 2.91 (2H, m), 2.83 (2H, *m*), 2.31 (1H, m), 2.26 (1H, m), 1.97 (1H, m), 1.86 (1H, m), 1.70 (2H, m), 1.51 (2H, m). ^{13}C NMR (MeOD, 126 MHz) *δ* 174.3, 169.8, 169.7, 147.5, 141.3, 129.7, 129.7, 129.3, 129.3, 128.5, 127.8, 127.5, 123.5, 122.2, 116.9, 54.7, 54.0, 52.8, 40.5, 34.8, 32.2, 31.8, 28.0, 24.0. HRMS calcd for [M + H$^+$] $C_{24}H_{34}N_5O_4$ 456.2605; found 456.2607

(S)-6-amino-2-(4-amino-3-((S)-2-amino-4-phenylbutanamido)benzamido)hexanoic acid (17).

Hydrolysis of the methyl ester of **8** as described for **4**, and subsequent deprotection of the two *N*-Boc groups using Method C yielded **17**. ^{1}H NMR (MeOD, 500 MHz) *δ* 7.63 (1H, s), 7.50 (1H, dd, *J* = 8.2, 1.8 Hz), 7.28–7.11 (5H, m), 6.81 (1H, d, *J* = 8.2 Hz), 4.42 (1H, m), 3.53 (1H, m), 2.79 (2H, m), 2.59 (2H, *m*), 2.11 (1H, m), 1.93 (2H, m), 1.78 (1H, m), 1.52–1.36 (4H, m). ^{13}C NMR (MeOD, 126 MHz) *δ* 179.5, 176.8, 169.3, 147.5, 143.2, 129.5, 129.5, 129.5, 129.5, 127.0, 126.8, 126.7, 123.9, 122.2, 116.4, 56.8, 56.6, 42.5, 40.4, 34.0, 33.8, 33.5, 24.3. HRMS calcd for [M + H$^+$] $C_{23}H_{32}N_5O_4$ 442.2449; found 442.2451

(S)-methyl 2-(4-amino-3-((S)-2-amino-4-phenylbutanamido)benzamido)-5-guanidino pentanoate (18).

Deprotection of the guanidine *N*-Cbz groups of **9** using Method B, and subsequent deprotection of the *N*-Boc group using Method C yielded **18**. ^{1}H NMR (MeOD, 500 MHz) δ 8.06 (1H, s), 7.99-7.93 (1H, m), 7.59-7.53 (1H, m), 7.34-7.22 (4H, m), 7.20-7.14 (1H, m), 4.46-4.37 (1H, m), 3.73 (3H, s), 3.53-3.47 (1H, m), 3.29-3.19 (2H, m), 2.93-2.81 (4H, m), 2.51-2.28 (2H, m), 2.15-1.69 (2H, m). ^{13}C NMR (MeOD, 126 MHz) δ 170.1, 168.2, 167.8, 158.6, 146.4, 141.3, 130.1, 129.6, 129.5, 129.4, 128.8, 128.1, 127.5, 126.9, 124.8, 123.6, 55.1, 54.2, 52.9, 41.9, 34.3, 32.2, 29.2, 26.6. HRMS calcd for [M + H^{+}] $C_{24}H_{34}N_7O_4$ 484.2667; found 484.2660

(S)-methyl 2-(4-amino-3-((S)-2-amino-4-phenylbutanamido)benzamido)-3-(4-hydroxyphenyl) propanoate (19).

Simultaneous deprotection of the *O*-tBu and *N*-Boc groups of **10** using Method C afforded **19**. ^{1}H NMR (MeOD, 500 MHz) δ 7.60 (1H, s), 7.52 (1H, d, J = 7.8 Hz), 7.33–7.17 (5H, m), 7.04 (2H, d, J = 8.0 Hz), 6.89 (1H, d, J = 8.2 Hz), 6.69 (2H, d, J = 8.0 Hz), 4.73 (1H, m), 4.18 (1H, m), 3.69 (3H, s), 3.14 (1H, m), 3.00 (1H, m), 2.83 (2H, m), 2.37–2.18 (2H, m). ^{13}C NMR (MeOD, 126 MHz) δ 174.0, 169.5, 157.3, 147.6, 141.3, 131.2, 131.2, 129.8, 129.8, 129.3, 129.3, 129.1, 128.4, 127.7, 127.6, 123.5, 122.1, 116.8, 116.8, 116.3, 116.3, 56.1, 54.7, 52.7, 37.5, 34.8, 32.2. HRMS calcd for [M + H^{+}] $C_{27}H_{31}N_4O_5$ 491.2289; found 491.2290

(S)-benzyl 2-(4-amino-3-((S)-2-amino-4-phenylbutanamido)benzamido)-3-(1H-indol-3-yl) propanoate (20).

Deprotection of the *N*-Boc group of **11** using Method C afforded **20**. ^{1}H NMR (MeOD, 500 MHz) δ 7.60 (1H, s), 7.53 (1H, d, J = 7.8 Hz), 7.50 (1H, dd, J = 8.2, 1.8 Hz), 7.33–7.13 (12H, m), 7.08–6.95 (3H, m), 6.89 (1H, d, J = 8.2 Hz), 5.07 (2H, s), 4.91 (1H, m), 4.16 (1H, m), 3.42–3.29 (2H, m), 2.81 (2H, m), 2.37–2.18 (2H, m). ^{13}C NMR (MeOD, 126 MHz) δ 173.8, 169.6, 169.5, 147.9, 141.3, 138.0, 137.1, 129.8, 129.8, 129.5, 129.5, 129.3, 129.3, 129.2, 129.2, 129.2, 128.8, 128.4, 127.8, 127.6, 124.4, 123.3, 122.5, 121.9, 119.9, 119.2, 116.7, 112.4, 110.8, 67.9, 55.8, 54.8, 34.8, 32.2, 28.4. HRMS calcd for [M + H^{+}] $C_{35}H_{36}N_5O_4$ 590.2753; found 590.2755

(S)-methyl 4-amino-3-(2-amino-4-phenylbutanamido)benzoate (21).

Deprotection of the *N*-Boc group of **3** using Method C afforded **21**. ^{1}H NMR (MeOD, 500 MHz) δ 7.97 (1H, s), 7.92 (1H, d, J=8.4 Hz), 7.33 (1H, d, J=8.4 Hz), 7.31–7.25 (4H, m), 7.22–7.16 (1H, m), 4.35–4.29 (1H, m), 3.89 (3H, s), 2.88–2.81 (2H, m), 2.45–2.34 (1H, m), 2.34–2.23 (1H, m). ^{13}C NMR

(MeOD, 126 MHz) δ 170.0, 167.3, 151.2, 141.3, 137.6, 130.1, 129.7, 129.4, 128.9, 128.3, 127.8, 127.6, 122.5, 55.0, 52.8, 34.5, 32.2

(S)-4-amino-3-(2-amino-4-phenylbutanamido)benzoic acid (22).

Deprotection of the *N*-Boc group of **4** using Method C afforded **22**. ^{1}H NMR (MeOD, 500 MHz) δ 7.97 (1H, s), 7.94 (1H, d, *J*=8.4 Hz), 7.33 (1H, d, *J*=8.4 Hz), 7.31–7.25 (4H, m), 7.22–7.16 (1H, m), 4.33–4.26 (1H, m), 2.88–2.80 (2H, m), 2.44–2.34 (1H, m), 2.34–2.23 (1H, m). ^{13}C NMR (MeOD, 126 MHz) δ 170.0, 168.4, 148.3, 141.2, 136.1, 130.5, 129.8, 129.4, 129.1, 128.3, 128.0, 127.6, 122.3, 55.0, 34.5, 32.2

Recombinant enzyme production and purification

ERAP1, ERAP2 and IRAP were produced by insect cells in suspension culture (Hi5 cells) after infection with recombinant baculovirus carrying the respective gene with an additional poly-histidine tag fused at the C-terminus for purification purposes. The expression systems for ERAP1 and ERAP2 have been previously described (ref. 17). For the construction of the recombinant baculovirus driving IRAP expression, soluble human recombinant IRAP was amplified from HeLa cDNA with the following primers: 5′–GCTAGCGCAACAAATGGGAAATTGTTT–3′ (underlined NheI) and 5′–GCGGCCGCCTA*ATGGTGATGGTGATGGTG*TCCCGACAGCCACCATGTGAGACT TTT–3′ (underlined NotI, italic 6xHis-tag). The PCR product was sequenced and cloned in pVL1392 in fusion with gp64 signal peptide and used to produce a recombinant baculovirus using the Baculogold kit (Invitrogen) according to the manufacturer's instructions. The enzyme was purified by Ni-NTA-Agar column from Hi5 cells culture supernatant two days post-infection with the recombinant baculovirus. The purity of all enzymes was found to be >95% by SDS-PAGE.

Enzyme activity assay

The enzymatic activity of ERAP1, ERAP2 and IRAP were determined by following the hydrolysis of the synthetic substrates L-Leucine-7-amido-4-methyl coumarin for ERAP1 and IRAP (L-AMC; Sigma) and L-Arginyl-7-amido-4-methyl coumarin for ERAP2 (R-AMC; Sigma), monitored by the release of the fluorogenic product, AMC, at excitation and emission wavelengths of 380 and 460 nm, respectively. All measurements were performed on a TECAN infinite® M200 microplate fluorescence reader. For evaluation of inhibitory activity, 2–7 nM of each enzyme (6.2 nM of ERAP1, 2.5 nM of ERAP2 and 2.6 nM of IRAP), diluted in 50 mM HEPES (pH 7.0), 100 mM NaCl buffer, was added in each well, along with 50 μM of substrate and various concentrations of each compound. The reaction was followed for 5–10 min and activity was calculated by measuring the slope of the time course.

For calculation of the *in vitro* IC_{50} values, experimental data were fit to the following equation using the GraphPad Prism software package:

$$Y=Bottom + (Top-Bottom)/(1+10^{((LogIC_{50}-X)*HillSlope)})$$

where: Y is the enzymatic activity and X the inhibitor concentration. Standard deviation was calculated from 3 separate experiments.

The activity of ERAP1 was also measured using the chromogenic substrate L-Leucine-para-nitroanilide (L-pNA) by following the absorbance of the enzymatic product para-nitroanilide (pNA) at 405 nm (extinction coefficient=9450 $M^{-1}cm^{-1}$) during incubation with ERAP1. Briefly, 1.5 μg ml^{-1} ERAP1 was incubated at room temperature with increasing concentrations of L-pNA (in the 0–10 mM range) in 50 mM HEPES, pH 7.0, 100 mM NaCl, for 5–10 min. The rate of hydrolysis was calculated by the slope of the time-dependent increase in absorbance. For Michaelis-Menten calculations, initial reaction rates were plotted versus different substrate concentrations and were fit to a standard Michaelis-Menten model (using the GraphPad Prism software).

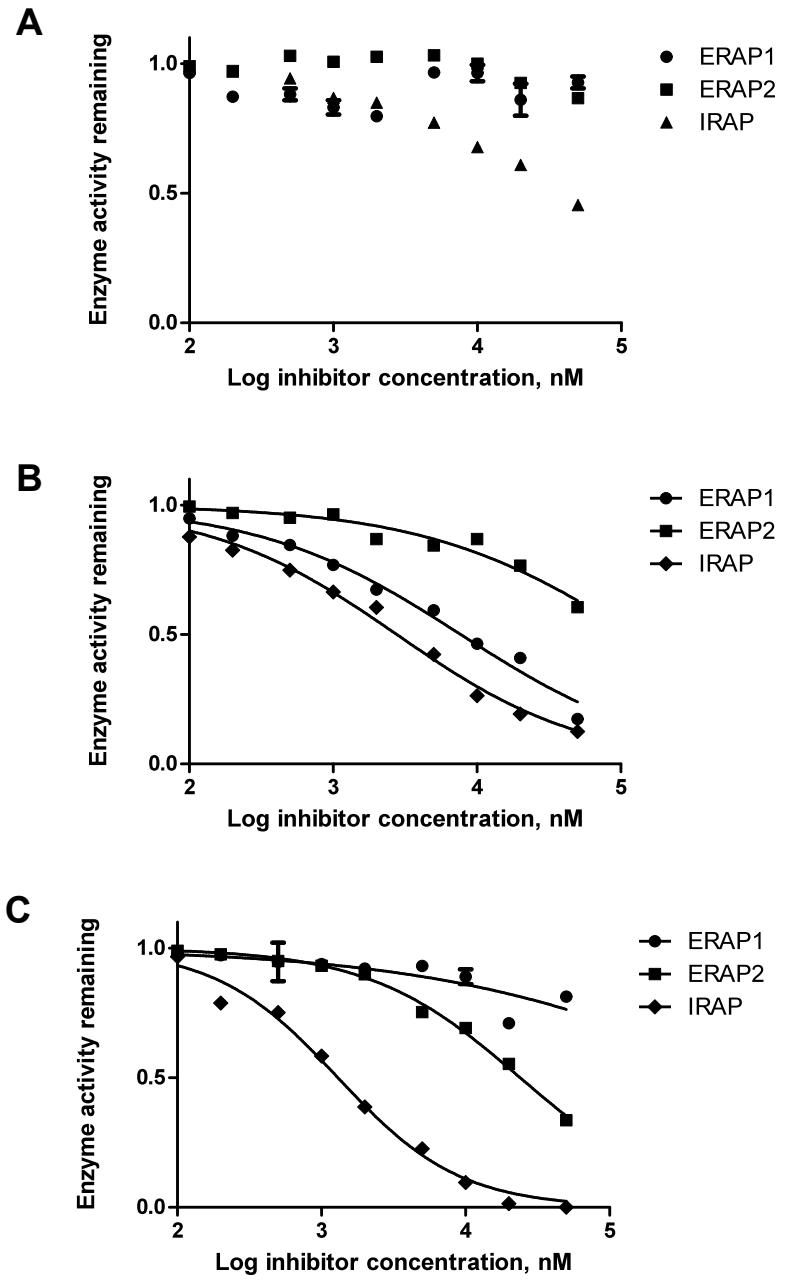

Figure S1. *In vitro* characterization of inhibitors **12** (A), **19** (B) and **20** (C). The inhibitory potency of each compound for the three enzymes was calculated by titrating increasing amounts of the compound while following the kinetics of hydrolysis of small fluorogenic substrates (L-AMC for ERAP1 and IRAP, and R-AMC for ERAP2)

HPLC analysis

Analysis of digestion products after incubation of compounds with ERAP1 was performed by RP-HPLC on a Chromolith C-18 analytical column (Merck). In brief, 10 μM of each peptide was incubated at room temperature with 50 nM ERAP1 in 50 mM HEPES pH 7.0, 100 mM NaCl buffer in a total volume of 200 μl for 7–30 min. After incubation, the reactions were terminated by adding 200 μl 0.1% trifluoroacetic acid and stored at -20°C until analysis. Before analysis, samples were centrifuged for 5 min at 10,000 g to remove precipitated protein. HPLC elution was performed using a 0–50% acetonitrile gradient, while following the absorbance at 257 nm.

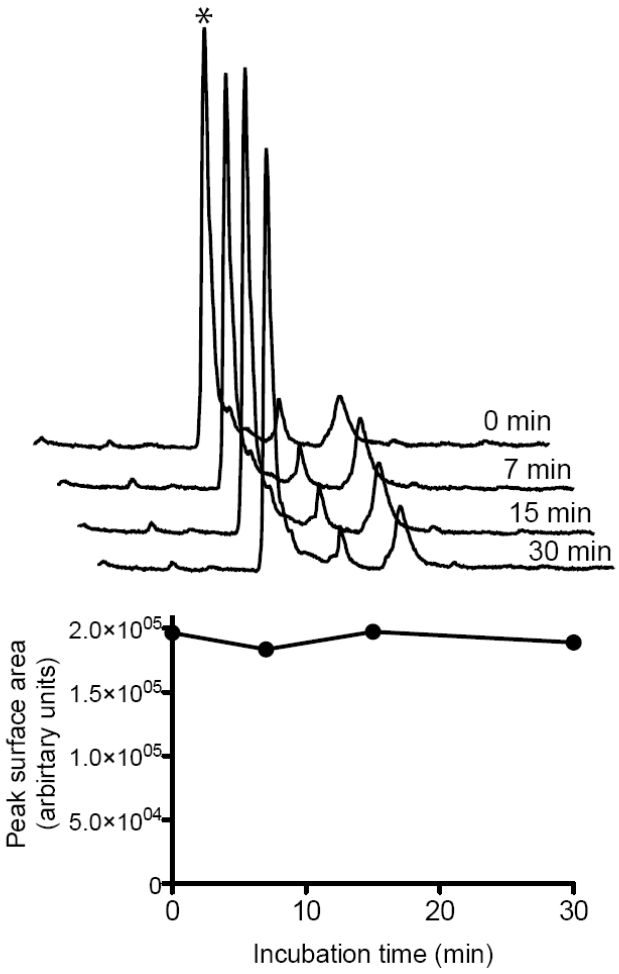

Figure S2. *Top panel,* RP-HPLC chromatograms of inhibitor **16** before and after incubation with 50 nM ERAP1. *Bottom panel*, time course of the surface of the chromatographic peak that corresponds to inhibitor **16** (marked with a star) indicating stability versus degradation by ERAP1.

Computational methods

In an effort to gain insight into the structure-activity relationships of the designed inhibitors, molecular docking was employed. The crystallographic structures of ERAP1[15] and ERAP2[16] in the closed conformation were retrieved from the Protein Data Bank with accession codes 2YD0 and 3SE6, respectively. To prepare the structures for the docking calculations a set of programs from AMBER 11 package were used.[S2] Hydrogen and missing atoms were added, and the improved ff99SB[S3] force field parameters were applied using the XLEaP module. The missing loop regions that are ~30 Å away from the active-site zinc ion were not modeled. Acidic and basic residues were set to their ionized state, except for the histidine residues that were set to their neutral state. Force field parameters and the charge distribution of the Zn^{2+} coordination sphere were taken from ref. S4. To optimize the position of the modeled atoms and relax the structures from crystallographic contacts a two-step minimization procedure was carried out in implicit solvent using a pairwise generalized Born model implemented in Amber (*igb*=1).[S5] In the first round of minimization, only hydrogen atoms were allowed to relax by applying positional restraints on all heavy atoms, while in the second round, positional restraints were applied only on the C^{α} atoms. The non-bonded interactions were not truncated using the effectively infinite default value for the cutoff, the positional restraints were applied with a force constant of 50 kcal mol^{-1} $Å^{-2}$, and the convergence criterion for the energy gradient was decreased to 10^{-3} kcal mol^{-1} $Å^{-1}$. A homology model for IRAP was obtained as described in ref. 17, but using both ERAP1 and ERAP2 energy-minimized structures as templates.

The initial conformations of the designed inhibitors were generated from SMILES representations using the program Omega 2.3 with default parameters.[S6,S7] Input files for docking were prepared using AutoDockTools 1.5.4[S8] Initially, non-polar hydrogens were merged and then either Kollman or Gasteiger charges were applied for the protein and ligand atoms, respectively. The search space was defined by a grid box centered close to the active-site zinc and comprised 121×101×121 grid points of 0.25 Å spacing. For each complex, 100 docking rounds were calculated with AutoDock 4.2 using the Lamarckian genetic algorithm with the default parameters from AutoDock 3.[S9,S10] The maximum number of energy evaluations was set to 5 million and the results were clustered using a tolerance of 2.0 Å. Visual examination of the results and rendering of the figures was performed using VMD 1.9.[S11] Calculations were carried out using Intel Xeon servers running Linux 2.6.32 kernels.

Supplementary References

(S1) Gerritz, S. W.; Sefler, A. M. 2,5-Dimethylfuran (DMFu): An Internal Standard for the "Traceless" Quantitation of Unknown Samples via (1)H NMR. *J. Comb. Chem.* **2000,** *2*, 39-41

(S2) Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26,* 1668–1688.

(S3) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725.

(S4) Papakyriakou, A.; Spyroulias, G. A.; Sturrock, E. D.; Manessi-Zoupa, E.; Cordopatis, P. Simulated interactions between angiotensin-converting enzyme and substrate gonadotropin-releasing hormone: novel insights into domain selectivity. *Biochemistry* **2007**, *46* 8753-8765.

(S5) Tsui, V.; Case, D.A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers (Nucl. Acid. Sci.)* **2001**, *56*, 275–291.

(S6) OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.

(S7) Hawkins, P. C. D.; Skillman, A. G.; Warren, G. L.; Ellingson B. A.; Stahl, M. T. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **2010**, *50*, 572–584.

(S8) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. *J. Comput. Chem.* **2009**, *16,* 2785–2791.

(S9) Morris, G. M.; Goodsell, D. S.; Halliday, R.S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.

(S10) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. A Semiempirical Free Energy Force Field with Charge-Based Desolvation. *J. Comput. Chem.* **2007**, *28*, 1145–1152.

(S11) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 27–38.