Supplemental Information

Title

Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate-peptide complex structures

Authors and Affiliations

Sebastian Zoll1§, Stancho Stanchev1, Jakub Began1,3, Jan Škerle1,2, Martin Lepšík1, Lucie Peclinovská1,2, Pavel Majer1 and Kvido Strisovsky1*

1Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo n. 2, Prague, 166 10, Czech Republic

2Department of Biochemistry, Faculty of Science, Charles University, Prague, Czech Republic

3Department of Microbiology, Faculty of Science, Charles University, Prague, Czech Republic

§ Present address: Department of Biochemistry, University of Oxford, Oxford, United Kingdom

* corresponding author

Contact Information

All correspondence to be addressed to kvido.strisovsky@uochb.cas.cz

Running Title

Structures of rhomboid:substrate-peptide complexes
Supplemental Figures and Tables

Figure S1

Ac-AA-cmk 0h

V₀

0.5
1
2
3
4
Time [min]

Ac-AA-cmk 4h

V₀

0.5
1
2
3
4
Time [min]

Ac-AAA-cmk 0h

V₀

0.2
0.4
0.6
0.8
1
Time [min]

Ac-AAA-cmk 4h

V₀

0.2
0.4
0.6
0.8
1
Time [min]
Figure S1: Stability of peptidyl-CMKs in aqueous solution. Peptidyl-CMKs were diluted into the reaction buffer (see Experimental Procedures, Inhibitions Assays) to 1 mM concentration and incubated at 25°C for the indicated time. They were then separated by reversed phase liquid chromatography and analyzed by mass spectrometry. Qualitative and quantitative comparisons at different time points were based on exact mass spectra of the individual peaks, their retention times (*) and peak areas (both as peak labels in the graphs). All compounds except Ac-IATG-cmk were stable for at least 4 hours at pH 7.4 showing no signs of degradation. Ac-IATG-cmk was mostly stable at pH 6.0, showing only marginal formation of a degradation product.
Figure S2: Characterisation of peptidyl chloromethylketones (A). Ac-IAAA-cmk inhibits GlpG in a concentration and time-dependent manner, which is indicative of an irreversible inhibitor. (B) The tetrapeptidyl chloromethylketone Ac-IAAA-cmk binds specifically to GlpG in a S201 and H254 dependent manner and produces a higher mobility band on SDS PAGE. The inhibitor at 1 mM concentration was reacted with GlpG variants over 2.5 hrs and reaction mixture was analysed by SDS PAGE and MALDI-MS in positive mode using α-cyano-4-hydroxycinnamic acid matrix. Mass spectra show that only wild type GlpG forms a covalent adduct with the inhibitor resulting in a peak shift of 382 Da compared to unreacted wild type GlpG (left panel). The peak mass shift is accompanied by the formation of a discrete, higher mobility band on SDS PAGE (inset). In contrast, active site mutants S201A and H254A do not form stable adducts with Ac-IAAA-cmk (middle and right panels) and do not cause gel-shift.
Figure S3
Figure S3: Source data for the positional scanning mutagenesis of the P5 to P1 region of TatA shown in Fig. 2A. Autoradiography of 35S-Methionine-labelled TatA variants shows that mutations in positions P5 (T4), P4 (I5), P3 (A6), P2 (T7) and P1 (A8) alter cleavage efficiency by GlpG. Amino acid substitutions in the P1 (A8) position are least tolerated resulting in less cleavage product. S, Full-length substrate TatA; P1 TatA cleavage product 1; mut, inactive S201A GlpG mutant; WT, wild type
Figure S4: Source data for quantification of inhibitory properties of sequence variants of peptidyl CMKs shown in Fig. 2C. The same amino acids substitutions in peptidyl-cmks as in substrates are not tolerated by GlpG. S, full-length TatA substrate; P1, cleavage product 1; E, Enzyme (GlpG); E-I, Enzyme-inhibitor complex; P2, cleavage product 2.
Figure S5: Superposition of Ac-IATA-cmk (yellow) and Ac-TIATA-cmk (orange) in the active site of GlpG. An F_o-F_c simulated annealing omit map, calculated at 2.7 Å and contoured at 3 σ, is shown 2 Å around Ac-TIATA-cmk. While both peptidyl CMKs have the same overall orientation, no additional density for P5 threonine is observed.

Figure S6: Electrostatic surface potential of the cavity constituting the S1 subsite/water retention site. The whole cavity has a strongly negative electrostatic potential. (A) Side view of the cavity. The P1 alanine of Ac-IATA-cmk points into the S1 subsite that connects to the water retention site. The retention site is filled with three ordered water molecules (w1-w3) that can be found in the same position in most GlpG crystal structures. (B) Cross section view of the cavity from the top. Water molecule 3 (w3) lies above the clipping plane. Its position indicated by a dashed contour line. The electrostatic potential was calculated using Adaptive Poisson-Boltzmann Solver (ABPS) plugin as implemented in Pymol.
**Figure S7:** Source data for the mutational analysis of the S4 subsite shown in Fig. 4A. Autoradiography of $^{35}$S-Methionine-labelled TatA variants shows compensatory effect of mutations in the GlpG S4 subsite and the TatA P4 position on enzymatic activity. Dashed lines indicate gel ends. S, Full-length substrate TatA; P1 TatA cleavage product 1; wt, wild type; mut, inactive S201A GlpG mutant; FA, GlpG F146A; FI, GlpG F146I.
**Figure S8**

**A**

![Graph A](image)

**B**

<table>
<thead>
<tr>
<th>R-bonds</th>
<th>Dist [Å]</th>
<th>Ang [°]</th>
<th>Occ [%]</th>
<th>AL [ps]</th>
</tr>
</thead>
<tbody>
<tr>
<td>His254Nε2-Ser201Oy⁺</td>
<td>2.98±0.18</td>
<td>29.6±11.9</td>
<td>98.4</td>
<td>1094</td>
</tr>
<tr>
<td>Ser201N-Ala(P1)O⁻</td>
<td>3.28±0.14</td>
<td>22.8±7.8</td>
<td>42.4</td>
<td>40</td>
</tr>
<tr>
<td>Asn154Nε2-Ala(P1)O⁻</td>
<td>2.95±0.14</td>
<td>19.1±9.7</td>
<td>99.6</td>
<td>3320</td>
</tr>
<tr>
<td>Ser246O-Thr(P2)N⁺</td>
<td>3.02±0.17</td>
<td>21.2±10.2</td>
<td>97.2</td>
<td>748</td>
</tr>
<tr>
<td>Gly198O-Ala(P1)N⁺</td>
<td>2.94±0.16</td>
<td>28.5±13.8</td>
<td>93.4</td>
<td>312</td>
</tr>
<tr>
<td>Ala250N-Thr(P2)O⁻</td>
<td>2.96±0.15</td>
<td>32.9±12.5</td>
<td>91.4</td>
<td>234</td>
</tr>
<tr>
<td>Trp196O-Ala(P3)N⁺</td>
<td>3.19±0.18</td>
<td>24.4±11.4</td>
<td>82.6</td>
<td>62</td>
</tr>
<tr>
<td>Gly199N-Ala(P3)O⁻</td>
<td>3.30±0.15</td>
<td>32.1±10.7</td>
<td>35.4</td>
<td>34</td>
</tr>
</tbody>
</table>

**C**

**Nucleophilic Attack**

**Complex**

Ser201Oy⁺-Ala(P1)C⁺: 3.13±0.18

**D**

![Graph D](image)

**E**

![Diagram E](image)

**Catalytic dyad and Oxyanion hole**

**S2' subsite**
Figure S8: Validation of the molecular dynamics model of Michaelis complex of the P4 to P3′ segment of the substrate with GlpG. (A) RMSD for protein (black) and substrate (red) backbone during the simulation. (B) H-bonds in the active site observed during the MD simulation. Dist, Distance; Ang, Angle; Occ, Occupancy; AL Average Lifetime; *Active site H-bonds/interactions present in Michaelis complex/tetrahedral intermediate/acyl enzyme; †H-bonds also present in the GlpG Ac-IATA-cmk complex crystal structure (C) Superposition of the Michaelis complex model with the DFP and Ac-IATA-cmk complex structures. The carbonyl oxygen of the P1 residue of the model adopts a similar orientation as in the DFP complex mimicking the tetrahedral intermediate (D) Ensemble of models showing a different orientation of the P2′ phenyl ring in the S2′ subsite as judged by monitoring the distance between P2′Cζ and P1′Ca (E) Superposition of the P2′ residue side chains of representative models of the Michaelis complex (major (92%) and minor (8%) ensembles) and the β-lactam inhibitor L29 thought to bind into the S2′ subsite. GlpG residues making van der Waals contacts to the phenyl ring of the P2′ residue are depicted in red, residues only contacting it in the models in orange.

Figure S9
Figure S9: A hydrophobic surface-exposed patch formed by the L1 loop is conserved in the same position among GlpG homologs. (A) Hydrophobicity distribution on the surface of HiGlpG with the L5 loop being omitted. The S4 subsite in HiGlpG resembles the one in EcGlpG. (B) Superposition of S4 subsite residues of EcGlpG, EcGlpG F146I and HiGlpG with bound peptidyl-cmks. The arrangement of hydrophobic residues in the S4 subsite of the HiGlpG apoenzyme is comparable to EcGlpG and allows for similar substrate interaction. (C) Multiple sequence alignment with 100 bacterial homologs of *E. coli* (EcGlpG) and *H. influenzae* (HiGlpG). EcGlpG and HiGlpG are framed by red boxes. The consensus sequence logo indicates that hydrophobic residues are conserved in the S4 subsite.
Table S1

Data collection and structure refinements statistics. Values in parentheses correspond to the highest resolution bin.

<table>
<thead>
<tr>
<th>GlpG Ac-IATA-cmk</th>
<th>GlpG Ac-FATA-cmk</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td><strong>Data collection</strong></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9201</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>48.99 - 2.1 (2.175 - 2.1)</td>
</tr>
<tr>
<td>Space group</td>
<td>P 6₃</td>
</tr>
<tr>
<td>Unit cell dimensions (Å)</td>
<td>97.99 97.99 66.40</td>
</tr>
<tr>
<td>Total reflections</td>
<td>162260 (16312)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>21305 (2138)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>7.6 (7.6)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.9 (99.9)</td>
</tr>
<tr>
<td>Mean I/sigma(I)</td>
<td>25.6 (4.3)</td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>32.0</td>
</tr>
<tr>
<td>R&lt;sub&gt;sym&lt;/sub&gt;</td>
<td>0.057 (0.485)</td>
</tr>
<tr>
<td>R&lt;sub&gt;meas&lt;/sub&gt;</td>
<td>0.061 (0.558)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement</th>
<th>Refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range (Å)</td>
<td>48.99 - 2.1</td>
</tr>
<tr>
<td>#R&lt;sub&gt;work&lt;/sub&gt;/R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.163/0.180</td>
</tr>
<tr>
<td>Protein residues</td>
<td>182</td>
</tr>
<tr>
<td>Number of atoms</td>
<td>1663</td>
</tr>
<tr>
<td>Total</td>
<td>Protein</td>
</tr>
<tr>
<td>Protein</td>
<td>S60 (Ac-IATA-cmk)</td>
</tr>
<tr>
<td>S60 (Ac-IATA)</td>
<td>150</td>
</tr>
<tr>
<td>Detergent</td>
<td>Water</td>
</tr>
<tr>
<td>Water</td>
<td>Detergent</td>
</tr>
<tr>
<td>Chloride</td>
<td>Detergent</td>
</tr>
<tr>
<td>B-factors (Å²)</td>
<td>1</td>
</tr>
<tr>
<td>Average</td>
<td>37.1</td>
</tr>
<tr>
<td>Protein</td>
<td>36.6</td>
</tr>
<tr>
<td>S60 (Ac-IATA)</td>
<td>42.9</td>
</tr>
<tr>
<td>Detergent</td>
<td>62.2</td>
</tr>
<tr>
<td>Water</td>
<td>45.9</td>
</tr>
<tr>
<td>Chloride</td>
<td>37.8</td>
</tr>
<tr>
<td>RMS deviation bonds (Å)</td>
<td>0.011</td>
</tr>
<tr>
<td>RMS deviation angles (°)</td>
<td>1.020</td>
</tr>
<tr>
<td>Ramachandran favored (%)</td>
<td>98.0</td>
</tr>
<tr>
<td>Ramachandran outliers (%)</td>
<td>0</td>
</tr>
</tbody>
</table>
**F146I Ac-FATA-cmk**

### Data collection

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>1.0</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>51.81 - 2.55 (2.641 - 2.55)</td>
</tr>
<tr>
<td>Space group</td>
<td>P 6₃</td>
</tr>
<tr>
<td>Unit cell dimensions (Å)</td>
<td>98.4 98.4 65.26</td>
</tr>
<tr>
<td>Total reflections</td>
<td>39119 (3855)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>11747 (1170)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.3 (3.3)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.0 (99.2)</td>
</tr>
<tr>
<td>Mean I/sigma(I)</td>
<td>15.9 (3.3)</td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>45.4</td>
</tr>
<tr>
<td><em>R&lt;sub&gt;s&lt;/sub&gt;ym</em></td>
<td>0.057 (0.397)</td>
</tr>
<tr>
<td><strong>R&lt;sub&gt;meas&lt;/sub&gt;</strong></td>
<td>0.068 (0.474)</td>
</tr>
</tbody>
</table>

### Refinement

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range (Å)</td>
<td>51.81 - 2.55</td>
</tr>
<tr>
<td>#R&lt;sub&gt;work&lt;/sub&gt;/R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.194/0.228</td>
</tr>
<tr>
<td>Protein residues</td>
<td>182</td>
</tr>
<tr>
<td>Number of atoms</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1531</td>
</tr>
<tr>
<td>Protein</td>
<td>1455</td>
</tr>
<tr>
<td>S16 (Ac-FATA-cmk)</td>
<td>32</td>
</tr>
<tr>
<td>Detergent</td>
<td>32</td>
</tr>
<tr>
<td>Water</td>
<td>11</td>
</tr>
<tr>
<td>Chloride</td>
<td>1</td>
</tr>
<tr>
<td>B-factors (Å²)</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>46.5</td>
</tr>
<tr>
<td>Protein</td>
<td>46.3</td>
</tr>
<tr>
<td>S16 (Ac-FATA-cmk)</td>
<td>44.5</td>
</tr>
<tr>
<td>Detergent</td>
<td>59.0</td>
</tr>
<tr>
<td>Water</td>
<td>46.5</td>
</tr>
<tr>
<td>Chloride</td>
<td>46.3</td>
</tr>
<tr>
<td>RMS deviations bonds (Å)</td>
<td>0.025</td>
</tr>
<tr>
<td>RMS deviation angles (°)</td>
<td>1.82</td>
</tr>
<tr>
<td>Ramachandran favored (%)</td>
<td>99.0</td>
</tr>
<tr>
<td>Ramachandran outliers (%)</td>
<td>0</td>
</tr>
</tbody>
</table>

*R<sub>sym</sub> = \frac{\sum |I - \langle I \rangle|}{\sum I}*

**R<sub>meas</sub> = \text{Redundancy independent } R_{sym}**

#R<sub>work</sub> = \frac{\sum |F_o - F_c|}{\sum F_o}

R<sub>free</sub> = \text{Cross-validation R-factor for 5% of test set of reflections omitted from model refinement}

RMS = \text{root mean square}
Supplemental Materials and Methods

Plasmids and mutagenesis
The plasmids encoding the library of *P. stuartii* full-length TatA mutants and the chimeric maltose-binding-protein-transmembrane-domain-thioredoxin (MBP-TMD-Trx) substrates and GlpG have been described (Strisovsky et al, 2009). All GlpG mutants have been constructed by Quick-Change mutagenesis or by overlap-extension PCR (Heckman & Pease, 2007) and their identity was verified by DNA sequencing of the full ORF.

Purification of GlpG core domain for crystallography
Published procedures (Vinothkumar et al, 2010; Wang et al, 2006) were modified as follows. After chymotryptic cleavage (Wang et al, 2006), which was carried out in imidazole-containing His-tag elution buffer, imidazole was removed using a HiPrep 26/10 desalting column (GE Healthcare). Omitting anion-exchange (Vinothkumar et al, 2010) purification the protein was concentrated (10 kDa MWCO, Vivaspin) and directly subjected to size exclusion chromatography using a HiLoad 16/600 Superdex 200 column equilibrated with 25 mM Tris (pH 8.0), 150 mM NaCl and 0.5 % (w/v) β-D-nonyl glucoside (NG, Anatrace). Peak fraction were >90% pure as judged by SDS-PAGE.

In vitro transcription and translation
Plasmids encoding TatA mutants (Strisovsky et al, 2009) were PCR-amplified with primers introducing SP6 polymerase promoter and ribosome binding site. The PCR products were directly *in vitro* transcribed with SP6 polymerase (New England Biolabs) and RNA was purified by lithium chloride precipitation (Pokrovskaya & Gurevich, 1994). At least 2500 ng of purified RNA was *in vitro* translated in a 25 µl reaction using wheat-germ extract (Promega) in the presence of radioactive \[^{35}S\]-L-Met.

In vivo assay of GlpG activity
The MBP-TatA-Trx encoding plasmids (Strisovsky et al, 2009) were transformed into wild-type *E. coli* MC4100 encoding endogenous GlpG or into its *glpG::tet* mutant derivative. Transformants were selected on 100 µg/mL ampicillin LB plates. Single colonies were used to inoculate overnight liquid cultures in LB with ampicillin. For the activity assay, the inocula were diluted to OD\(_{600}=0.05\) into fresh LB medium supplemented with appropriate antibiotics, grown at 37°C to OD\(_{600}\) of 0.8 and the expression of the substrate was induced by 1 mM IPTG. After 3 hours post induction the cells were spun at 5000g for 5 min at 4°C and the pellet was dissolved in reducing SDS-PAGE sample buffer. The substrate and cleavage products were analysed by SDS PAGE using 4-20% Tris-glycine gradient gels (Bio-Rad) and by immunoblotting using mouse monoclonal anti-polyHistidine peroxidase conjugate HIS-1 (1:2000, Sigma). HRP activity was detected using Luminata Forte Western HRP Substrate (Millipore).
Crystallisation and structure solution
The crystals of GlpG complexes with peptidyl CMKs grew within three days at 20°C in the following conditions: 10% (w/v) PEG 3000, 0.1 M Imidazole pH 8.0, 0.2 M Lithium sulphate (GlpG Ac-IATA-cmk); 10% PEG 8000 (w/v), 0.1 M CHES pH 9.5, 0.2 M sodium chloride (GlpG Ac-FATA-cmk); 30% (v/v) PEG 400, 0.1 M CHES pH 9.5 (GlpG-F146I AcFATA-cmk). Crystals of GlpG Ac-IATA-cmk and GlpG Ac-FATA-cmk were cryo-protected with 25% (v/v) glycerol before flash freezing in liquid nitrogen. No cryo-protectant was used for freezing of GlpG-F146I Ac-FATA-cmk crystals. All data sets were measured at 100K using synchrotron radiation. GlpG-AcIATA-cmk data were collected on a Rayonix 225 detector (Beamline 14.2, BESSY, Berlin, Germany), data sets of the other two complexes on a PILATUS 6M (Beamline ID29, ESRF, Grenoble). All datasets were indexed, integrated and scaled using the XDS package (Kabsch, 2010). For all complex structures, Molecular Replacement for initial phase calculation was performed with Phaser (McCoy et al, 2007) using the structure of unliganded GlpG (PDB-ID 2IC8) as a search model. The L5 loop next to the active site was omitted to avoid model bias. Simulated annealing and first rounds of refinement were carried out using PHENIX (Adams et al, 2010), manual model building was done in Coot (Emsley et al, 2010). Peptide structures were drawn in ChemBioDraw 13.0 (PerkinElmer), coordinate, library and link files generated with JLigand (Lebedev et al, 2012). Peptides were then added manually and fit to difference density by real space refinement. After addition of the peptide, final refinement cycles were carried out with REFMAC5 of the CCP4 suite (Murshudov et al, 2011; Winn et al, 2011). All crystals belong to the space group P63 with one protomer in the asymmetric unit. The final models comprise residues 90-271 of GlpG. Hydrogen bonds and hydrophobic contacts were calculated with LIGPLOT (Wallace et al, 1995) and the hydrophobicity surface scale with UCSF Chimera (Pettersen et al, 2004) according to Kyte and Doolittle (Kyte & Doolittle, 1982).

Modeling of the Michaelis complex and molecular dynamics
Modeling. The X-ray structure of the covalently bonded GlpG_Ac-IATA-cmk complex (this work) was used as a starting structure. All water molecules and chloride ions were stripped off. All lysine and arginine side chains and the N-terminus were treated as cations, while all aspartate and glutamate side chains and the C-terminus were considered anionic. In line with the experimental pH of 8, histidine side chains were monoprotonated. The location of the proton to Nδ(H145, H254) and Nε (H141, H150) was determined by visual inspection to maximize hydrogen bonding. The imidazole rings of H154 and H150 were moreover flipped. The covalent bonds between the catalytic dyad (S201, H254) and the Ac-IATA-cmk inhibitor were broken by 200 steps of steepest-descent gradient geometry optimization using ff10 force field of the AMBER 11 simulation package (Case et al, 2010). This completed model of the unprimed part of the substrate was subsequently extended by the P1' Ala and P2' Phe moieties which were modeled into the crevice in the GlpG surface visualized in PyMol (Schrodinger, 2012). To enable this, an alternative conformation of M208 side chain had to be used to avoid steric clashes. The P1', P2' and the surrounding GlpG residues were relaxed using minimization and a short molecular dynamics (MD) simulation (heating to 600 K and cooling down to 10 K during 2 ps in the presence of generalized Born (GB) implicit solvent [IGB=5])
(Onufriev et al, 2004). The model was further extended by the P3' Gly capped by N-methyl amide and relaxed similarly.

**Molecular dynamics simulations.** A simplified setup was used to obtain a quick view of the dynamics of the whole system. The relaxation first included all the hydrogens (250 cycles of steepest descent) followed by the protein and non-prime side of the substrate side chains, and the entirety of the prime and non-prime side of the substrate (250 cycles of steepest descent plus 4750 cycles of conjugate gradient). The MD run consisted of i) gradual warming from 10 to 300 K in 50 ps using GB solvent model [IGB=5] (Onufriev et al, 2004) of dielectric constant of 20, ii) 200ps equilibration at 300 K and iii) 30 ns production run. The analyses comprised measurements of root-mean-square deviation (RMSD), selected distances and hydrogen bonds during the simulation.

**Modeling of the TatA transmembrane domain**
The software MODELLER 9.10 was used to calculate a model of *Providencia stuartii* TatA TMD. *P. stuartii* TatA shares high sequence homology with *E. coli* TatA, the NMR structure of which (PDB ID: 2LZR) served as a template for modeling. Different orientations of the TMD were then manually added in Coot.

**Chemical Synthesis**

**Materials**
All reagents were acquired from commercial sources and used without purification. Protected amino acids and amino acid derivatives were purchased from Iris Biotech, Marktredwitz, Germany or Sigma-Aldrich, St. Louis, MO, U.S.A.

**HPLC**
HPLC analyses were carried out on an HPLC system (PU1580, PU1575 Series, Jasco, Japan) with a Reprosil-100, C-18, 250 x 4.6, 5μm column (Watrex International, Inc, San Francisco, CA, USA), at a flow rate of 1 mL/min using various gradients (0.1%(v/v) aqueous trifluoroacetic acid (TFA) – acetonitrile) with UV detection at 210 nm.
Preparative HPLC separations were performed on an HPLC system (PU 986, PU975 Series, Jasco, Japan) using YMC-Pack, ODS-AM, 250x20mm, 5μm column (YMC America, Inc., Allentown, PA, U.S.A.), at flow rates of 7 to 10 mL/min using various gradients (0.1%(v/v) aqueous TFA – acetonitrile) with UV detection at 210 - 235 nm.

**Mass Spectrometry**
Mass-spectra were acquired on Waters Micromass ZQ ESCi multimode ionization mass-spectrometer and LTQ Orbitrap XL (Thermo Fisher Scientific) for HR-MS experiments. Both of mass-spectrometers are using ESI (+)-ionization method.
NMR spectroscopy

NMR spectra were acquired on a Bruker AV-400 MHz at room temperature.

Synthesis of peptidyl chloromethylketones

1. General procedures

1.1. Synthesis of protected N-α-acetyl peptides (Ac-Ile-Ala-Ala-OH, Ac-Ile-Ala-Thr(tBu)-OH, Ac-Ala-Ala-Ala-OH, Ac-Ala-Ala-OH, Ac-Thr(tBu)-Ile-Ala-Thr(tBu)-OH, Ac-Phe-Ala-Thr(tBu)-OH, Ac-Ser(tBu)-Ala-Thr(tBu)-OH, Ac-Ile-Asp(tBu)-Thr(tBu)-OH, Ac-Ile-Trp-Thr(tBu)-OH) (Bodanszky & Bodanszky, 1984)

1.1.1. Amide couplings

The amide couplings were performed in DMF solution, using either Cbz or Boc protected amino acids. C-terminal carboxyl groups of the amino components were protected as methyl esters. In a typical procedure 1 eq. of the carboxyl-component (protected amino) acid was combined with 1.2 eq. TBTU and 1.5 eq. DIEA in a minimal amount of DMF and was left to stir for 5 minutes at RT. Then 0.9 eq. of the amino-component (amino acid or peptide methyl ester) was added to the reaction mixture followed by 1.5 eq. of DIEA (or more) to assure basic pH of the solution. Reaction mixture was left to stir overnight at RT and then quenched by addition of 4-aminoethylmorpholine. The volatiles were removed in vacuum and the oily residue dissolved in ethyl acetate. The solution was washed three times with saturated aqueous NaHCO₃ solution, three times with 10% aqueous KHSO₄ and once with saturated brine. Ethyl acetate fractions were combined dried over MgSO₄, filtered and evaporated, yielding the crude product. Alternatively solid products were washed using the same solutions on a glass frit and dried in vacuo. The peptides Ac-Ile-Ala-Ala-OMe, Ac-Ala-Ala-Ala-OMe, Ac-Ala-Ala-OMe were prepared using Boc-chemistry whereas the remaining peptides were synthesized, using Cbz-chemistry due to presence of an acid labile t-butyl group in their side chains.

1.1.1.2. Deprotection of the α-amino-protecting groups (Boc or Cbz).

The removal of Boc group was achieved by mixture of 95% TFA and 5% water in an ultrasound bath for 10 minutes. Volatile reagents were then removed by flow of nitrogen and the residue after neutralization by DIEA used directly for the next coupling step.

Removal of Cbz group was performed by hydrogenolysis, using palladium hydroxide on carbon (Pearlman’s catalyst). The course of the reaction was monitored by analytical HPLC. The deprotection was usually complete in 30 – 120 min after which the catalyst was filtered off, the solvents evaporated and the product used directly for the next step.

1.1.1.3. Acetylation of the N-terminus
N-acetyloxysuccinimide was used for the N-terminal acetylation of the protected tripeptides. Reaction mixtures were worked up analogously as in the amide coupling steps.

1.1.1.4. Hydrolysis of the C-terminal methyl ester

The protected peptide (1 mmol) was dissolved in 6 mL of methanol. NaOH (1.5 ml, 5M aq. solution) was added and the mixture was stirred for up to 5 hours at room temperature, while being monitored by analytical HPLC. After completion the reaction mixture was diluted with water (5 mL) and methanol was evaporated in vacuum. The mixture was acidified with 10% KHSO₄ to pH 2-3 and the precipitated peptide with free terminal carboxyl group was filtered off, washed with water and dried in vacuum. Alternatively, peptides with higher aqueous solubility or oily consistency were extracted with ethyl acetate.

1.2. Synthesis of the amino acid chloromethylketones (CMK)

The Boc protected amino acid, triethylamine and ethyl chloroformate were dissolved in cooled (-5°C) dry THF. The mixture was stirred at 0°C for 30 min, while the white precipitate of triethylamine hydrochloride was formed. Then ~2 eq. of diazomethane ether solution (0.3 mol/L) were added. The mixture was stirred for 2 hours at 0°C after which the triethylamine hydrochloride was removed by filtration. The solvent and the excess of diazomethane were evaporated in vacuo and the yellow oily product (diazoketone) was dried, in order to remove the residues of triethylamine. Then, dried diazoketone was dissolved in a minimal volume of dry tetrahydrofuran, cooled to 0°C and treated with excess of 4M HCl in dry dioxane. Yellow color of the solution disappeared and the corresponding chloromethyl ketone precipitated as a hydrochloride with cold ether. The precipitate was filtered off, washed with ether and dried.

Alanine chloromethylketone (3-amino-1-chlorobutane-2-one) hydrochloride (Thomson & Denniss, 1973)

Starting materials: BocAlaOH (3 g, 16 mmol), ethylchloroformate (2.6 g, 2.28 mL, 24 mmol), triethylamine (2.42 g, 3.34 mL, 24 mmol), dry THF (15 mL), diazomethane solution (80 mL, 24 mmol) and 4M HCl in dry dioxane (40 ml). The product is a white crystalline substance.

Yield: 1.57g (81%).
ESI-MS: Monoisotopic mass: 121.03 calculated for C₄H₈ClNO, found: [M+H]^+ 122.1, 124.1(minor peak, because of the presence of 37Cl)

1H NMR (DMSO-d₆, 400 MHz); δ (ppm) = 1.42 (d, 3H, CH₃, J = 7.3 Hz), 4.25 (s, 1H, CH), 4.92 - 4.70 (m, 2H, COCH₂Cl), 8.46 (s, 3H, NH₃+). 13C NMR (101 MHz, DMSO-d₆) δ (ppm) = 199.01, 52.46, 46.86, 15.00.

Glycine chloromethylketone (3-amino-1-chloropropane-2-one) hydrochloride (Owen & Voorheis, 1976)
Starting materials: BocGlyOH (0.5 g, 2.85 mmol), triethylamine (0.43 mL, 3.05 mmol), ethylchloroformate (0.32 mL, 3.34 mmol), dry THF (5 ml), diazomethane solution (25 mL, 7.5 mmol). Diazomethane was removed by evaporation and the yellow oily product was dissolved in ether and washed with saturated NaHCO$_3$ solution, then dried over MgSO$_4$. The ether was evaporated; the crude product was dissolved in THF and treated with 4M HCl in dry dioxane. The product was precipitated at room temperature for 10 – 15 min with cold ether. The product is a pale beige crystals.

Yield: 0.214 g (52 %).

ESI-MS: Monoisotopic mass: 107,01 calculated for C$_3$H$_6$ClNO, found [M+H]$^+$ 108.12, 110.13 (minor peak, because of the presence of $^{37}$Cl).

$^1$H NMR (DMSO-d$_6$, 400 MHz); $\delta$ (ppm) = 4.0 (s, 2H, CH$_2$), 4.66 (s, 2H, COCH$_2$Cl), 8.44 (s, 3H, CONH).$^{13}$C NMR (101 MHz, DMSO-d$_6$) $\delta$ (ppm) = 196.21, 46.99, 45.19.

Valine chloromethylketone (3-amino-1-chloro-4-methylpentan-2-one) hydrochloride (Hauske et al, 2009)

Starting compounds: BOC-Val-OH (0.5 g, 2.3 mmol), triethylamine (0.42 ml, 3 mmol), ethylchloroformate (0.29 ml, 3 mmol), 5 ml dry tetrahydrofuran, diazomethane solution (28 ml, 8.4 mmol); 5h stirring at -5°C. The residue, dissolved in dry THF and cooled to 0°C, was treated with 1.6 ml 4 M HCl in dioxane dropwise. The mixture was stirred for 2h then cold ether was added, in order to precipitate the product as a pale yellow crystals. They were filtered, washed with ether and dried.

Yield: 0.15g (35 %)

ESI-MS: Monoisotopic mass: 149,06 calculated for C$_6$H$_{12}$ClNO, found [M+H]$^+$ 150.07 and 152.1(minor peak due to the presence of $^{37}$Cl).

$^1$H NMR (DMSO-d$_6$, 400 MHz); $\delta$ (ppm) = 0.86 (d, 3H, CH$_3$, J = 7.0 Hz), 1.01(d, 3H, CH$_3$, $^3$J = 7.0 Hz), 2.39 - 2.28 (m, 1H, CH), 4.2 (s, 1H, CH), 4.91-4.68 (m, 2H, COCH$_2$Cl), 8.5 (s, 3H, CONH).$^{13}$C NMR (101 MHz, DMSO-d$_6$) $\delta$ 198.87, 61.14, 47.91, 28.29, 18.75, 16.78 .

1.3. Coupling of the aminoacid C-terminal chloromethylketones (CMK) with the acetylated peptides (Jahreis et al, 1984; Thomson & Denniss, 1973)

Dried acetylated peptide (1 eqv) was dissolved in dry THF (or DCM) and the mixture was cooled down to -10°C. Then triethylamine (1 eqv) and ethylchloroformate (1 eqv) were added consequently and white precipitate of triethylamine hydrochloride was formed. The mixture was stirred at -10°C for 30 min and the 1 eqv of the aminoacid chloromethylketone hydrochloride was poured into the reaction mixture. 1 eqv of triethylamine, dissolved in 0.5 ml dry THF were added dropwise over 1 h. The temperature was maintained around -5°C. After the dropping of the triethylamine solution, the mixture was stirred 1 h at 0°C and pH was controlled (must be in the range 7.5 -8). The temperature was left to increase slowly to the ambient overnight. The THF was evaporated and the dry solid was
treated with 2 ml trifluoroacetic acid upon sonication, in order to remove the protecting tBu (20 min) and BOC-groups (3-4 min) (if it is necessary).

The TFA was evaporated under the stream of nitrogen and the residual oil-like substance was dissolved in 50 – 70 % CH₃COOH aqueous and was purified by preparative RP HPLC. The fractions contained the product were combined, frozen and lyophilized.

**Ac-Ile-Ala-Ala-Ala-CMK**

Starting compounds: Ac-Ile-Ala-Ala-OH (86 mg, 0.27 mmol), triethylamine (38 µl, 0.27 mmol) + (38 µl, 0.27 mmol), ethylchloroformate (26 µl, 0.27 mmol), anhydrous THF (6 ml), AlaCMK. HCl (43 mg, 0.27 mmol).

The sample for prepration was dissolved in the mixture acetic acid/water/methanol. The compound was isolated by RP HPLC, using gradient 15 – 50 % B. Inj. vol. 2 ml. The product was eluted in the 26 min.

Yield: 27 mg (24 %)

Analytical HPLC-control: Gradient 15 – 50 % B. One peak at 13.7 min

HR-MS: Monoisotopic mass 441.1881 calculated for [C₁₈H₃₁ClN₄O₅+Na]⁺, found 441.18749.

¹H NMR (DMSO-d₆, 400 MHz); δ (ppm) = 0.86 - 0.76 (m, 6H, CH₃), 1.30-1.11(m, 11H, CH₃, CH₂), 1.67 (m, 1H, CH), 1.85 (s, 3H, acetyl), 4.16-4.09 (m, 1H, CH), 4.29 – 4.17(m, 2H, CH), 4.37 - 4.30(m, 1H), 4.56 (s, 2H, COCH₂Cl), 7.9 (t, 2H, CONH, J= 7.1Hz), 8.08(d, 1H, CONH, J = 7.1 Hz), 8.33 (d, 1H, CONH, J= 7.1 Hz).

¹³C NMR (101 MHz, DMSO-d₆) δ (ppm) = 201.08, 171.85, 170.98, 169.31, 56.84, 52.16, 48.02, 47.37, 36.46, 24.37, 22.48, 17.70, 15.60, 15.31, 11.01.

**Ac-Ala-Ala-Ala-Ala-CMK (Jahreis et al, 1984; Thomson & Denniss, 1973)**

Starting compounds: Ac-Ala-Ala-Ala-OH (54.6 mg, 0.2 mmol), triethylamine (28 µl, 0.2 mmol) + (32 µl, 0.23 mmol), ethylchloroformate (19 µl, 0.2 mmol), anhydrous THF (3.5 - 4 ml), AlaCMK. HCl (36 mg, 0.23 mmol).

The compound was isolated by RP HPLC, using gradient 2 – 30 % B. Inj. vol. 2 ml. The product was eluted in the 25 min.

Yield: 40.9 mg (55 %)

Analytical HPLC-control: Gradient 2-100 % B. One peak at 13.1 min.

¹H NMR (DMSO-d₆, 400 MHz); δ (ppm) = 1.26 – 1.13 (m, 12H, CH₃), 1.83(s, 3H, acetyl), 4.21(h, 3H, CH, J= 7.1Hz), 4.33(p, 1H, CH, J = 7.1Hz), 4.56(d, 2H, COCH₂Cl, J = 1.7Hz), 7.91(d, 1H, CONH, J = 7 Hz), 8.05(dd, 2H, CONH, J= 14.5, 7.2 Hz), 8.28(d, 1H, CONH, J= 6.7 Hz).

**Ac-Ala-Ala-Ala-CMK (Jahreis et al, 1984; Thomson & Denniss, 1973)**
Starting compounds: Ac-Ala-Ala-OH (52 mg, 0.26 mmol), triethylamine (36 µl, 0.26 mmol) + (42 µl, 0.3 mmol), ethylchloroformate (25 µl, 0.26 mmol), dry THF (2.5 ml), AlaCMK. HCl (47.4 mg, 0.3 mmol).

The compound was isolated by RP HPLC, using gradient 2 – 30 % B. Inj. vol. 1.3 ml. The product was eluted in the 22 min.

Yield: 20 mg (25 %)

Analytical HPLC-control: Gradient 2-100 % B. One peak at 12.6 min.

$^1$H NMR (DMSO-d$_6$, 400 MHz); δ (ppm) = 1.17(d, 3H, CH$_3$, J = 7.1Hz), 1.21 (dd, 6H, CH$_3$, J = 7.1Hz, 2.9 Hz), 1.82 (s, 3H, acetyl), 4.21(h, 2H, CH, J = 7.2 Hz), 4.32 (p, 1H, CH, J = 7.1Hz), 4.56(d, 2H, COCH$_2$Cl, J = 1.6 Hz), 8.06 (dd, 2H, CONH, J = 11.9, 7.1 Hz), 8.28 (d, 1H, CONH, J = 6.7 Hz).

**Ac-Ala-Ala-CMK (Jahreis et al, 1984; Thomson & Denniss, 1973)**

Starting compounds: Ac-Ala-OH (88 mg, 0.67 mmol), triethylamine (94 µl, 0.67 mmol) + (107 µl, 0.77 mmol), ethylchloroformate (64 µl, 0.67 mmol), dry THF (2.5 ml), AlaCMK. HCl (121 mg, 0.77 mmol).

RP-HPLC isolation of the product: gradient 2 – 30 % B. Inj. vol. 1.3 ml. The product was eluting at 22 min.

Yield: 20 mg (13 %)

Analytical HPLC-control: Gradient 2-30 % B. One peak at 13.05 min.

$^1$H NMR (DMSO-d$_6$, 400 MHz); δ (ppm) = 1.23 – 1.11 (m, 6H, CH$_3$), 1.82 (s, 3H, acetyl), 4.27 – 4.13 (m, 1H, CH), 4.39 – 4.27 (m, 1H, CH), 4.65 – 4.50 (m, 2H, COCH$_2$Cl), 8.08 (t, 1H, CONH, $^3$J(H, H) = 6.4 Hz), 8.40 (dd, 1H, CONH, $^3$J(H, H) = 20.1, 6.9 Hz).

**Ac-Ile-Ala-Thr-Ala-CMK**

Starting compounds: Ac-Ile-Ala-Thr(tBu)-OH (50 mg, 0.125 mmol), triethylamine (19 µl, 0.138 mmol) + (20 µl, 0.144 mmol), ethylchloroformate (13.5 µl, 0.138 mmol), anhydrous DCM (0.6 ml), AlaCMK. HCl (23 mg, 0.144 mmol).

RP-HPLC purification of the product: Gradient 30 – 80 % B. Inj. vol. 1.25 ml. The product was eluting at 27 min.

Yield: 23 mg (40 %)

Analytical HPLC-control: Gradient 15-50 % B. One peak at 13.3 min.

HR-MS: Monoisotopic mass 449.21672 calculated for [C$_{19}$H$_{33}$ClN$_4$O$_6$+H$^+$]; found 449.21628.

$^1$H NMR (DMSO-d$_6$, 400 MHz); δ (ppm) = 0.86 – 0.71(m, 6H, CH$_3$), 1.04 (d, 3H, CH$_3$, $^3$J = 6.4 Hz), 1.13 – 1.06 (m, 2H, CH$_2$), 1.22 (d, 6H, CH$_3$, J = 7.1 Hz), 1.47 – 1.35 (m, 1H, CH), 1.85 (s, 3H, acetyl), 4.06 – 3.93 (m, 1H, CH), 4.21 – 4.10 (m, 2H, CH), 4.40 – 4.28 (m, 2H, CH), 4.65 – 4.52 (m, 2H, COCH$_2$Cl), 4.94 (s, 1H, OH), 7.6 (d, 1H, CONH, J = 7.9 Hz), 7.93 (d, 1H, CONH, J= 8.2 Hz), 8.23 (dd, 2H, CONH, J = 24.3, 6.9 Hz).
$^{13}$C NMR (101 MHz, DMSO-d$_6$) $\delta$ (ppm) = 201.17, 172.25, 171.12, 170.24, 169.33, 66.47, 58.04, 56.80, 52.37, 48.35, 47.45, 36.48, 24.36, 22.44, 19.78, 17.60, 15.70, 15.34, 10.98.

**Ac-Ile-Ala-Thr-Gly-CMK**

Starting compounds: Ac-Ile-Ala-Thr(tBu)-OH (50 mg, 0.125 mmol), triethylamine (17 µl, 0.125 mmol) + (17 µl, 0.125 mmol), ethylchloroformate (12 µl, 0.125 mmol), anhydrous THF (2 ml), GlyCMK. HCl (18 mg, 0.125 mmol).

RP-HPLC purification of the product: Gradient 2 – 30 % B. Inj. Vol. 1.5 ml (3 injections), The product was eluted at 36 min.

Yield: 12 mg (22 %)

Analytical HPLC-control: Gradient 2 – 100 % B. The only one peak is at 12 min.

HR-MS: Monoisotopic mass 435.20107, calculated for [C$_{18}$H$_{31}$ClN$_4$O$_6$+H]$^+$; found 435.20069.

$^1$H NMR (400 MHz, DMSO-d$_6$) $\delta$ (ppm) = 0.85 – 0.75 (m, 6H, CH$_3$), 1.04 (d, J = 6.4 Hz, 3H, CH$_3$), 1.13 – 1.07 (m, 1H, CH$_2$), 1.22 (d, J = 7.1 Hz, 3H, CH$_3$), 1.41 (dq, J = 11.1, 3.7 Hz, 1H, CH$_2$), 1.74 – 1.60 (m, 1H, CH), 1.85 (s, 3H, acetyl), 4.03 (t, J = 5.4 Hz, 3H, CH, CH$_2$), 4.23 – 4.12 (m, 2H, CH), 4.35 (t, J = 7.1 Hz, 1H, CH), 4.54 (s, 2H, COCH$_2$Cl), 4.92 (s, 1H, OH), 7.65 (d, J = 8.4 Hz, 1H, CONH), 7.93 (d, J = 8.7 Hz, 1H, CONH), 8.24 – 8.07 (m, 2H, CONH).

$^{13}$C NMR (101 MHz, DMSO-d$_6$) $\delta$ = 198.65, 172.21, 171.02, 170.55, 169.28, 66.44, 58.08, 56.76, 48.25, 47.49, 46.91, 36.54, 24.36, 22.44, 19.63, 17.64, 15.32, 10.99.

**Ac-Ile-Ala-Thr-Val-CMK**

Starting compounds: Ac-Ile-Ala-Thr(tBu)-OH (55 mg, 0.137 mmol), triethylamine (18.4 µl, 0.137 mmol) + (18.4 µl, 0.137 mmol), ethylchloroformate (13.1 µl, 0.137 mmol), anhydrous DCM (2 ml), ValCMK. HCl (26 mg, 0.137 mmol).

RP-HPLC purification of the product: Gradient 15 – 50 % B, Inj. vol. 1.5 ml. The product was eluting at 34.5 min.

Yield: 23 mg (35 %)

Analytical HPLC-control: Gradient 15 – 50 % B. The only one peak is at 18.06 min.

HR-MS: Monoisotopic mass 477.24802 calculated for [C$_{21}$H$_{37}$ClN$_4$O$_6$+H]$^+$; found 477.24751.

$^1$H NMR (400 MHz, DMSO-d$_6$) $\delta$ (ppm) = 0.82 (td, J = 13.4, 7.1 Hz, 13H, CH$_3$, CH), 1.05 (d, J = 6.4 Hz, 3H, CH$_3$), 1.21 (d, 3H, CH$_3$), 1.68 (dq, J = 6.9, 4.2, 3.7 Hz, 1H, CH), 1.85 (s, 3H, acetyl), 2.16 (dq, J = 13.4, 6.7 Hz, 1H, CH), 4.00 (dd, J = 10.4, 5.3 Hz, 1H, CH), 4.15 (t, J = 8.0 Hz, 1H, CH), 4.24 (dd, J = 8.3, 3.9 Hz, 1H, CH), 4.32 (dt, J = 18.0, 6.7 Hz, 2H, CH), 4.58 (s, 2H, COCH$_2$Cl), 4.92 (d, J = 5.2 Hz, 1H, OH), 7.66 (d, J = 8.3 Hz, 1H, CONH), 7.96 (dd, J = 16.5, 8.2 Hz, 2H, CONH), 8.18 (d, J = 7.2 Hz, 1H, CONH).

$^{13}$C NMR (101 MHz, DMSO-d$_6$) $\delta$ (ppm) = 200.64, 172.24, 171.05, 170.59, 169.40, 66.54, 61.54, 58.00, 56.87, 48.15, 36.49, 28.60, 24.38, 22.45, 19.75, 19.33, 17.75, 15.37, 11.02.
Ac-Phe-Ala-Thr-Ala-CMK

Starting compounds: Ac-Phe-Ala-Thr(tBu)-OH (101 mg, 0.123 mmol), triethylamine (32 µl, 0.23 mmol) + (32 µl, 0.23 mmol), ethylchloroformate (22 µl, 0.23 mmol), anhydrous THF (3-4 ml), AlaCMK. HCl (36 mg, 0.23 mmol).

RP-HPLC purification of the product: Gradient 15-50 % B, Inj. vol. 1.2 ml (2 injections), The product was eluting at 30 min.

Yield: 14 mg (13 %)

Analytical HPLC-control: Gradient 15-50 % B. One peak at 15.22 min.

HR-MS: Monoisotopic mass 505.18298 calculated for [C_{22}H_{31}ClN_{4}O_{6}+Na]^+; found 505.18248.

$^1$H NMR (400 MHz, DMSO-d$_6$) δ (ppm) = 1.06 (d, J = 6.3 Hz, 3H, CH$_3$), 1.24 (t, J = 7.3 Hz, 6H, CH$_3$), 1.73 (s, 3H, acetyl), 2.76 – 2.63 (m, 1H, CH$_2$), 3.05 – 2.94 (m, 1H, CH$_2$), 4.06 – 3.97 (m, 1H, CH), 4.17 (dd, J = 8.2, 3.9 Hz, 1H, CH), 4.35 (t, J = 7.0 Hz, 2H, CH), 4.55 – 4.46 (m, 1H, CH), 4.59 (d, J = 5.0 Hz, 2H, COCH$_2$Cl), 7.19 (dt, J = 8.5, 4.0 Hz, 1H, aromatic), 7.26 (d, J = 4.3 Hz, 4H, aromatic), 7.65 (d, J = 8.2 Hz, 1H, CONH), 8.12 (d, J = 8.4 Hz, 2H, CONH), 8.30 (d, J = 10.8 Hz, 2H, CONH).

$^{13}$C NMR (101 MHz, DMSO-d$_6$) δ (ppm) = 201.18, 172.29, 171.44, 170.25, 169.24, 138.04, 129.09, 127.97, 126.17, 66.54, 57.95, 53.83, 52.40, 48.39, 47.45, 37.47, 22.40, 19.88, 17.62, 15.73.

Ac-Ser-Ala-Thr-Ala-CMK

Starting compounds: Ac-Ser(tBu)-Ala-Thr(tBu)-OH (104 mg, 0.24 mmol), triethylamine (32 µl, 0.23 mmol) + (33 µl, 0.24 mmol), ethylchloroformate (22 µl, 0.23 mmol), anhydrous THF (4.5 ml), AlaCMK. HCl (38 mg, 0.24 mmol).

RP-HPLC purification of the product: Gradient 15-50 % B, Inj. vol. 1.3 ml (2 injections), The product was eluting at 17 min.

Yield: 22 mg (22 %)

Analytical HPLC-control: Gradient 2-30 % B. One peak at 15.84 min.

HR-MS: Monoisotopic mass 423.16465, calculated for [C_{16}H_{27}ClN_{4}O_{7}+H]^+; found 423.16433.

$^1$H NMR (400 MHz, DMSO-d$_6$) δ (ppm) = 1.04 (d, J = 6.3 Hz, 3H, CH$_3$), 1.23 (t, J = 7.3 Hz, 6H, CH$_3$), 1.86 (s, 3H, acetyl), 3.55 (dd, J = 5.9, 3.4 Hz, 2H, CH$_2$), 8.22 (dd, J = 16.1, 6.9 Hz, 2H), 4.04 – 3.94 (m, 1H, CH), 4.19 – 4.10 (m, 1H, CH), 4.39 – 4.23 (m, 3H, CH), 4.67 – 4.50 (m, 2H, COCH$_2$Cl), 7.64 (d, J = 8.3 Hz, 1H, CONH), 8.01 (d, J = 7.6 Hz, 1H, CONH), 8.22 (dd, J = 16.1, 6.9 Hz, 2H, CONH).

$^{13}$C NMR (101 MHz, DMSO-d$_6$) δ (ppm) = 201.19, 172.40, 170.32, 169.58, 66.44, 61.69, 58.16, 55.06, 52.44, 48.63, 47.47, 22.53, 19.90, 17.56, 15.67.

Ac-Gly-Ala-Thr-Ala-CMK
Starting compounds: Ac-Gly-Ala-Thr(tBu)-OH (32 mg, 0.074 mmol), triethylamine (10 µl, 0.074 mmol) + (10 µl, 0.074 mmol), ethylchloroformate (7 µl, 0.074 mmol), anhydrous THF (2 ml), AlaCMK. HCl (12 mg, 0.074 mmol).

RP-HPLC purification of the product: Gradient 2-30 % B, Inj. vol. 1.3 ml (2 injections), The product was eluting at 22.5 min.

Yield: 16 mg (55 %)
Analytical HPLC-control: Gradient 2-30 % B. One peak at 13.44 min.
HR-MS: Monoisotopic mass 415.13603, calculated for [C15H25ClN4O6+Na]+; found 415.135634.

^1H NMR (400 MHz, DMSO-d6) δ (ppm) = 1.06 (d, J = 6.3 Hz, 3H, CH₃), 1.22 (dd, J = 7.1, 2.4 Hz, 6H, CH₃), 1.84 (s, 3H, acetyl), 3.69 (d, J = 5.7 Hz, 2H, CH₂), 4.06 – 3.97 (m, 1H, CH), 4.14 (dd, J = 8.2, 4.1 Hz, 1H, CH), 4.34 (q, J = 6.9 Hz, 2H, CH), 4.66 – 4.50 (m, 2H, COCH₂Cl), 7.69 (d, J = 8.1 Hz, 1H, CONH).

^13C NMR (101 MHz, DMSO-d6) δ (ppm) = 201.68, 172.94, 170.81, 170.33, 169.49, 66.94, 58.71, 52.91, 48.87, 47.93, 42.56, 22.85, 20.44, 18.27, 16.15.

Ac-Ile-Asp-Thr-Ala-CMK

Starting compounds: Ac-Ile-Asp(tBu)-Thr(tBu)-OH (105 mg, 0.21 mmol), triethylamine (29 µl, 0.21 mmol) + (29 µl, 0.21 mmol), ethylchloroformate (20 µl, 0.21 mmol), anhydrous THF (1.5 ml), AlaCMK. HCl (33 mg, 0.21 mmol).

RP-HPLC purification of the product: Gradient 15 – 50 % B, Inj. vol. 1.4 ml. The product was eluting at 25 min.

Yield: 10 mg (10 %)
Analytical HPLC-control: Gradient 15 – 50 % B. One peak at 11.0 min.
HR-MS: Monoisotopic mass 515.18846, calculated for [C₂₀H₃₅ClN₄O₈+Na]⁺; found 515.18796.

^1H NMR (400 MHz, DMSO-d6) δ (ppm) = 0.90 – 0.69 (m, 6H, CH₃), 1.08-0.96 (m, 2H, CH₂), 1.26 – 1.11 (m, 6H, CH₃), 1.79 – 1.55 (m, 1H, CH), 1.86 (s, 3H, acetyl), 2.39 -2.30 (m, 1H, CH₂), 2.39 – 2.30 (m, 1H, CH₂), 4.00 (q, J = 7.2 Hz, 2H, CH), 4.16 – 4.07(m, 1H, CH), 4.26 – 4.18 (m, 1H, CH), 4.35 – 4.28 (m, 1H, CH), 4 .8.46 – 8.36 (m, 1H), 4.66 – 4.53 (m, 2H, COCH₂Cl), 7.51 (d, J = 8.1 Hz, 1H, CONH), 7.61 (d, J = 7.1 Hz, 1H, CONH), 8.02 (d, J = 8.1 Hz, 1H, CONH), 8.21 (d, J = 6.7 Hz, 1H, CONH), 12.37 (s, 1H, COOH).

^13C NMR (101 MHz, DMSO-d6) δ (ppm) = 203.34 , 172.02, 171.37, 170.62, 169.85, 168.31, 66.41, 62.37, 60.12, 53.62, 47.41, 36.41, 27.79, 24.42, 22.41, 19.75, 15.88, 15.26, 11.02 .

Ac-Thr-Ile-Ala-Thr-Ala-CMK

Starting compounds: Ac-Thr(tBu)-Ile-Thr(tBu)-OH (70 mg, 0.125 mmol), triethylamine (17.4 µl, 0.125 mmol) + (17.4 µl, 0.125 mmol), ethylchloroformate (12 µl, 0.125 mmol), anhydrous THF (1.8 ml), AlaCMK. HCl (20 mg, 0.125 mmol).
RP-HPLC purification of the product: Gradient 15 – 50 % B, Inj. vol. 1.5 ml. The product was eluting at 25 min.

Yield: 18 mg (26 %)
Analytical HPLC-control: Gradient 15 – 50 % B. One peak at 11.25 min.
HR-MS: Monoisotopic mass 550.26437, calculated for [C_{22}H_{40}ClN_{5}O_{8}+H] \^+; found 550.26395.
\[^1\text{H} \text{NMR} \ (400 \text{ MHz}, \text{DMSO-d}_6) \ \delta \ (\text{ppm}) = 0.90 – 0.72 \ (\text{m, 6H, CH}_3), \ 1.03 \ (\text{dd, } J = 9.7, 6.4 \text{ Hz, 8H, CH}_3, \ CH_2), \ 1.22 \ (\text{dd, } J = 7.1, 3.6 \text{ Hz, 6H, CH}_3), \ 1.77 – 1.66 \ (\text{m, 1H, CH}), \ 1.89 \ (\text{s, 3H, acetyl}), \ 4.06 – 3.90 \ (\text{m, 2H, CH}), \ 4.15 \ (\text{dd, } J = 8.2, 3.8 \text{ Hz, 1H, CH}), \ 4.29 – 4.18 \ (\text{m, 2H, CH}), \ 4.34 \ (\text{td, } J = 7.0, 2.8 \text{ Hz, 2H, CH}), \ 4.64 – 4.52 \ (\text{m, 2H, COCH}_2Cl), \ 4.92 \ (\text{s, 2H, OH}), \ 7.63 \ (\text{dd, } J = 14.9, 8.4 \text{ Hz, 2H, CONH}), \ 7.83 \ (\text{d, } J = 8.5 \text{ Hz, 1H, CONH}), \ 8.15 \ (\text{d, } J = 7.0 \text{ Hz, 1H, CONH }, \ 8.24 \ (\text{d, } J = 6.7 \text{ Hz, 1H, CONH}).
\[^{13}\text{C} \text{NMR} \ (101 \text{ MHz, DMSO-d}_6) \ \delta \ (\text{ppm}) = 201.20, \ 172.26, \ 170.78, \ 170.28, \ 169.52, \ 66.67, \ 58.03, \ 56.76, \ 52.52, \ 47.45, \ 36.80, \ 24.11, \ 22.54, \ 19.56, \ 17.63, \ 15.71, \ 15.31, \ 11.17.

Supplemental References


