OPTIMIZATION OF THE POSTTRANSLATIONAL CLICK MODIFICATION OF PROTEINS

Milan VRABEL[1,*], Emine KAYA[2], Stefan PRILL[3], Veronika EHMKE[4] and Thomas CARELL[5,*]

Ludwig-Maximilian University Münich, Butenandtstrasse 5-13, D 81377 Münich, Germany; e-mail: [1] milan.vrabel@cup.uni-muenchen.de, [2] emine.kaya@cup.uni-muenchen.de, [3] stefan.prill@cup.uni-muenchen.de, [4] veronika.ehmke@org.chem.ethz.ch, [5] thomas.carell@cup.uni-muenchen.de

Received May 1, 2011
Accepted May 16, 2011
Published online August 22, 2011

Dedicated to Professor Antonín Holý on the occasion of his 75th birthday.

In order to develop efficient methods that would enable the synthesis of posttranslationaly modified proteins in a site-specific manner we have adopted the orthogonal pyrrolysyl-tRNA synthetase/tRNA pair to genetically encode various pyrrolysine analogs, which we were able to insert into the yellow fluorescent protein (YFP). These experiments showed that the alkene and alkyne containing amino acids **5** and **6** are superior substrates for the pyrrolysyl-tRNA synthetase and that they can be successfully incorporated into proteins. Using the Cu(I)-catalyzed Huisgen–Meldal–Sharpless click reaction, the alkyne containing YFP was finally glycosylated with various sugars. We confirmed the presence of the modified amino acids as well as the corresponding sugar modifications by HPLC-MS/MS mass spectrometry.
Keywords: Alkynes; Amino acids; Protein modification; Click chemistry; Triazoles; Azido compounds; Pyrrolysyl-tRNA synthetase; Pyrrolysine.

With few exceptions, the ribosomal biosynthesis of proteins in all living organisms is restricted to the 20 canonical amino acids. However, to fulfill the complex functions of the proteins, frequent posttranslational modifications as well as the presence of various cofactors are often required. The addition of new amino acids to the genetic code expands the range of various functions of the proteins but also provide new powerful tools to study their function and structure both *in vitro* and *in vivo*[1]. More than 50 unnatural amino acids have been already incorporated into proteins via a unique codon and an "orthogonal" tRNA/aminoacyl-tRNA synthetase (aaRS) pair, which do not crossreact with the host counterparts[2]. This was accomplished mainly by chemically aminoacylating a nonsense suppressor tRNA

doi:10.1135/cccc2011103

($tRNA_{CUA}$) with the desired non-natural amino acid and adding this aminoacyl tRNA to an *in vitro* translation/transcription system together with the gene of interest containing TAG mutation at the target site. Along this amber suppression methodology a frameshift suppression system using four-base codon-anticodon pairs was also developed[3]. An elegant alternative method for the incorporation of unnatural amino acids into proteins represents the pyrrolysine system. Pyrrolysine, as a lysine derivative with a bulky pyrroline ring, was discovered in the active site of monomethylamine methyltransferase from *Methanosarcina barkeri*[4]. This 22nd amino acid is directly esterified to its specific tRNA (PylT), which has the anticodon CUA complementary to the UAG codon, by pyrrolysyl-tRNA synthetase (PylRS). The broad substrate specificity of PylRS [5] enables the incorporation of a whole range of unnatural amino acids into various proteins, including amino acids containing bio-orthogonal chemical functionalities or posttranslationaly modified amino acids[6]. Directed mutagenesis of the PylRS can further extend the range of pyrrolysine analogs that can be incorporated into proteins using this system[7].

RESULTS AND DISCUSSION

In order to improve the existing methods for the posttranslational modification of proteins in a site-specific manner we systematically explored the orthogonal pyrrolysyl-tRNA synthetase/tRNA pair to genetically encode various unnatural amino acids bearing useful chemical functions that are bio-orthogonal. The targeted amino acids, for the site specific incorporation are all amenable to further modification using appropriate chemical reactions (Fig. 1).

Based on our initial modeling studies using the *Moloc* software we designed and synthesized (see experimental part) the series of pyrrolysine analogs **1–6** depicted in Fig. 2.

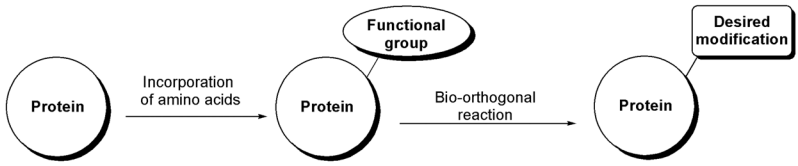


FIG. 1
Schematic representation of site-specific posttranslational protein modification

To test whether these unnatural amino acids **1–6** can serve as substrates for PylRS and to set up an assay that would report efficient stop-codon suppression we transformed *E. coli* with a plasmid containing the genes for the *Methanosarcina mazei* pyrrolysyl-tRNA synthetase, the corresponding *M. mazei* pyrrolysyl-$tRNA_{CUA}$ (PylT) and the C-terminal StrepII-tag modified yellow fluorescent protein (YFP) containing an amber TAG stop-codon. Finally, the plasmid contained an ampicillin resistance gene that allows selection of properly transfected *E. coli* cells. In order to study if one of the pyrrolysine analogs **1–6** would be inserted into the protein we added the amino acids at a final concentration of 5 mM to the growing transformed *E. coli* cells. The full-length YFP can be produced only in the presence of pyrrolysine analogs that are accepted by the PylRS. Only in this case, the PylT will be charged with the corresponding amino acid which enables the system reading through the stop-codon. In the absence of a charged PylT, stop-codon suppression will not be possible leading to the expression of only non-fluorescence, truncated versions of the YFP protein. Our experiments showed that the commercially available alloc-Lys derivative **5** as well as the alkyne derivative **6** are the most efficient substrates for the PylRS. Both were successfully incorporated into proteins using the described system[13].

In the absence of these amino acids the desired YFP fluorescence was not observed showing that in this case the full-length YFP is not produced. In the presence of the alternative amino acids **1–4** no fluorescence was observed, showing that these amino acids are not efficiently loaded onto the PylT. By comparing the chemical structures of the functional pyrrolysine

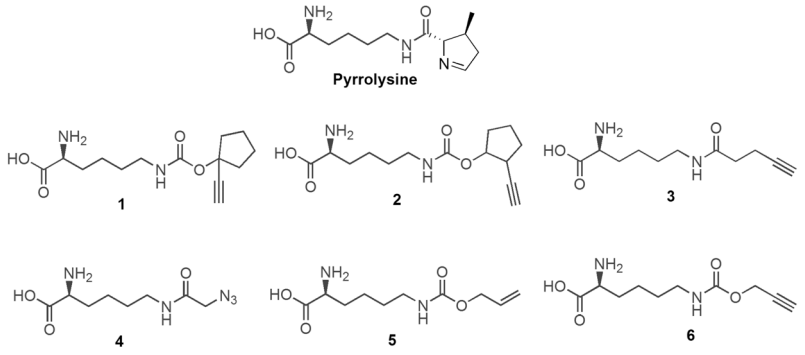


FIG. 2
Chemical structures of pyrrolysine and synthesized pyrrolysine analogs **1–6**

analogs **5** and **6** with the amino acids **1–4** we can conclude under the assumption that loading of the PylT is the rate limiting step, that the presence of the oxygen atom in the carbamate ester (urethane) moiety is essential for the recognition of these amino acid derivatives by PylRS (compare structures **3** and **6** in Fig. 2). This oxygen atom is most probably involved in hydrogen bonding in the active site of the PylRS that is necessary for the proper substrate recognition. The unsuccessful incorporation of compounds **1** and **2** (also containing carbamate ester functionality) is puzzling. A reason could be that they do not penetrate into the *E. coli* cells or that they are simply not accepted by the PylRS.

In order to enhance the amber suppression efficiency we next started to optimize the suppression system (Fig. 3a). Initially we increased the number of PylT gene copies on the plasmid and noted that by going from one copy to three gene copies the fluorescence strongly increased (Fig. 3b). This result shows that the amount of loaded PylT is a rate limiting factor. Increasing the amount of PylT enables the synthetase to charge more tRNAs, which increases the concentration of loaded PylT.

We next planned to create *E. coli* cells, which contain early on higher amounts of the synthetase PylRS and of the tRNA PylT. To this end, we implemented a system similar to the one published by Schultz and coworkers[8], in which they control both genes with the help of strong constitutive *E. coli* promoters (Fig. 4).

The production of PylRS was controlled by the GlnRS-promoter, for the cluster of three PylT genes we used the lysine-tRNA promoter proK and its corresponding terminator. For these studies we used two different plasmids,

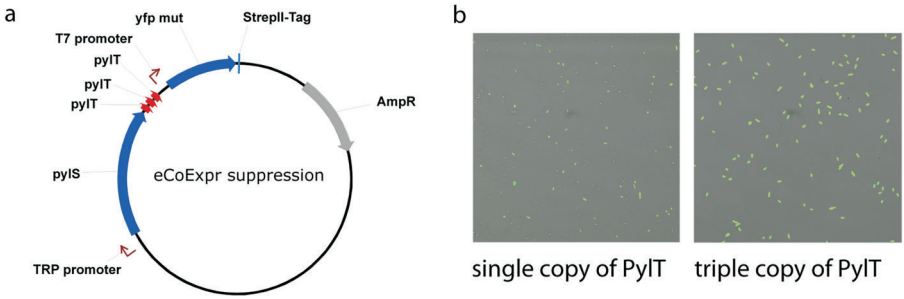


FIG. 3
Increased number of tRNA gene copies encoded on the expression plasmid (a). In the presence of more PylT in the *E. coli* cells an improved amber suppression efficiency could be obtained (b)

which were both simultaneously transformed into *E. coli*. Plasmid 1 contained the PylRS and PylT components, while plasmid 2 was used to insert the gene for the mutated YPF protein. Fluorescence analysis, however, did not show a large increase of the fluorescence indicating that for the pyrrolysine system these attempts do not increase the suppression efficiency in our hands.

The next goal was to use the incorporated amino acids for further click modification of the YFP protein. Initial attempts were focused on the use of the Cu-catalyzed azide-alkyne click reaction and the Cu-free click modification of the alkene amino acid with nitriloxides. This latter attempt was needed because we reasoned that more delicate proteins than the YFP studied here might unfold in the presence of large amounts of Cu-salts. The first method investigated was the Cu-free click reaction of nitriloxides with alkenes (method developed and used in our group for DNA modification)[9]. For the reaction we treated the alkene containing YFP protein with different concentrations of nitriloxides for various reaction times. However, we observed under no circumstances efficient labeling of the protein. In fact if we increased the reaction times all types of proteins reacted, even proteins that did not contain the alkene group, showing that the nitriloxides reaction is unspecific. In addition to the desired products we always observed the formation of by-products as a result of a nucleophilic attack of ε-amino groups of the lysine residues (or other nucleophiles of the protein) onto the

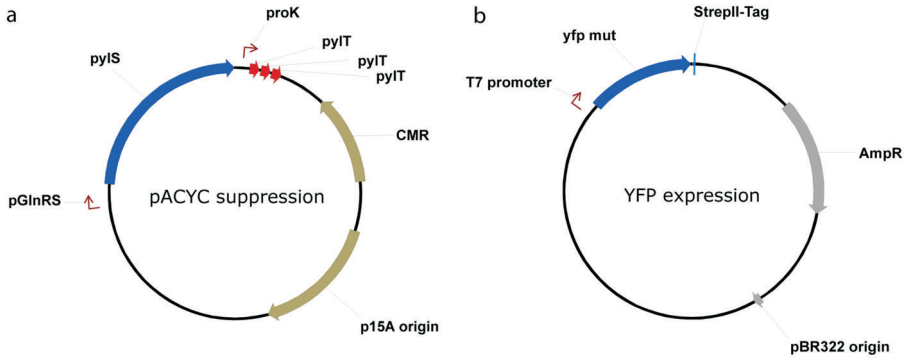


FIG. 4
A high concentration of both PylRS and PylT before biosynthesis of the target protein was achieved using the strong constitutive *E. coli* promoters pGlnRS and proK, respectively (a). The target protein (YFP) was encoded on a separate plasmid and its transcription was controlled by the T7-promoter (b)

carbon atom of the nitriloxide unit. In summary, we observed that the reaction of alkenes with nitriloxides is too unspecific for the site specific labelling of proteins (Fig. 5).

Our experience with the Cu(I)-catalyzed azide-alkyne cycloaddition reaction (CuAAC)[10] for DNA modification[11] prompted us to explore the possibility of using this method for protein modification. Considering the huge importance of glycosylation as posttranslational modification crucial for the proper function of many proteins[12] we decided to use various sugar azides as reaction partners in order to introduce these moieties into our alkyne containing YFP. Thus the purified YFP protein containing the corresponding alkyne unit was treated with galactose and mannose azides **7**, **8** and **9** as well as a solution of $CuSO_4$, sodium ascorbate and TBTA ligand. In addition we also used the biologically highly relevant sialic acid derivative **10** (Scheme 1). Using the Cu-catalyzed click reaction we observed that non-alkyne containing proteins did not react even not after prolonged reaction times with high amounts of added azides and Cu-salts. However, the alkyne

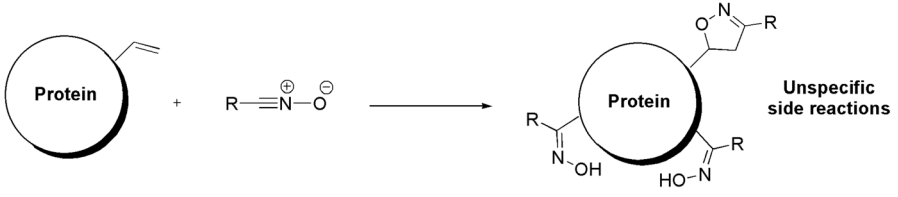
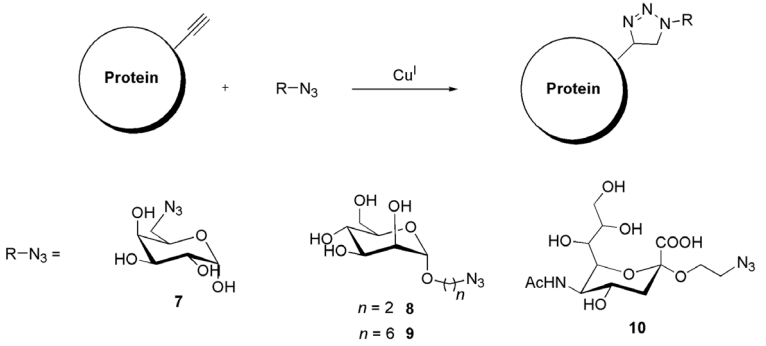


FIG. 5
Depiction of the unspecific nitriloxide click reaction with alkene containing protein



SCHEME 1
Depiction of the Cu-catalyzed click reaction with alkyne modified YFP and sugar azides **7**–**10**

containing protein reacted efficiently showing that this chemistry is well suited for the bioorthogonal modification of proteins[13].

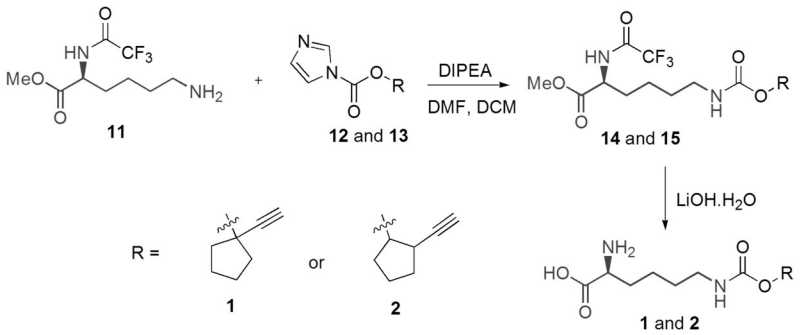
In order to prove the presence of the desired modifications we first digested the modified proteins with trypsin or chymotrypsin to generate the corresponding peptide fragments. The fragments of the digest were then analyzed by HPLC-MS/MS technique using a high resolution Orbitrap mass spectrometer. The observed masses were later on compared with the theoretically expected peptide fragments using the software "Xcalibur bioworks". These experiments clearly proved that in all cases the modified YFP protein contains not only the incorporated alkyne amino acids but also that these sites were properly glycosylated. As confirmed by the additional MS^n experiments the corresponding modifications were also placed in the correct sequence environment. The click efficiency was not 100% because next to the desired modified sugar-containing peptides we observed also the masses of the corresponding unmodified (un-clicked) alkyne-containing peptide fragments. However, based on the signal intensities comparison, we estimated that the reaction yield per alkyne unit is above 80% [13].

In summary we have designed and synthesized a series of pyrrolysine analogs **1–6** as new substrates for PylRS. We have used the *Mm*PylRS/*Mm*$tRNA_{CUA}$ system (pyrrolysine system) to genetically encode alkyne or alkene amino acids into proteins and we have shown that the alkyne units in the protein can be used for site-specific protein glycosylation using the Cu(I)-catalyzed Huisgen–Meldal–Sharpless click reaction. We next proved the correct incorporation of the amino acids as well as the following glycosylation pattern at the desired sites. The next step will be to improve the suppression yields and to develop alternative bio-orthogonal reactions that will allow the modification of proteins more efficiently under Cu-free conditions.

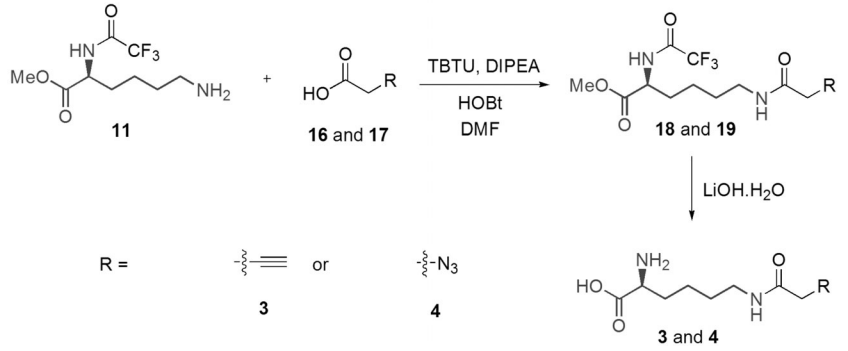
EXPERIMENTAL

Chemicals were purchased from Sigma–Aldrich, Fluka, ABCR or ACROS and used without further purification. Solvents used were of reagent grade and purified by standard methods. Reactions were monitored on Merck Silica 60 F254 TLC plates. Detection was done by irradiation with UV light (254 nm or 366 nm) and staining with *p*-anisaldehyde solution in ethanol or CAM staining solution. Flash column chromatography was performed on Silica 60 (Merck, 230–400 mesh). NMR spectra were recorded on the following spectrometers: Varian Oxford 200, Bruker AC 300, Varian XL 400 and Bruker AMX 600. Chemical shifts are given in ppm (δ-scale), coupling constants (*J*) in Hz. The spectrometers were calibrated using residual undeuterated solvents as internal reference. Mass spectra were recorded on the following machines: Finnigan MAT 95 (EI), Bruker Autoflex II (MALDI-Tof) and Thermo Finnigan LTQ-FT (ESI-ICR).

Compound **5** is commercially available. Synthesis of compound **6** was reported earlier[13]. Compounds **1–4** were synthesized as depicted in Schemes S1 and S2, respectively.



SCHEME S1
Synthesis of pyrrolysine analogs **1** and **2**



SCHEME S2
Synthesis of pyrrolysine analogs **3** and **4**

N-α-trifluoroacetyl-L-lysine methyl ester **11** was synthesized from commercially available *N*-ε-benzyloxycarbonyl-L-lysine methyl ester in two steps according to literature procedure[14].

Imidazole-1-carboxylic Acid 1-Ethynylcyclopentyl Ester (**12**)

To a solution of 1-ethynylcyclopentanol (1.56 ml, 13.62 mmol, 1.0 eq.) in DCM (80 ml) was added carbonyldiimidazole (7.95 g, 49.02 mmol, 3.6 eq.) and the mixture was stirred at r.t. for 24 h. The organic phase was washed with water (3 × 50 ml) and brine (50 ml). The organic phase was dried over $MgSO_4$. Evaporation of the solvent yielded product **12** (2.79 g, 13.61 mmol, 100%) as a colourless solid which was used in the next step without further purification. R_F 0.36 (isohexane/ethyl acetate 2:1). ^{1}H NMR (300 MHz, $CDCl_3$): 1.77–1.90 m, 4 H (CH_2); 2.23–2.34 m, 2 H (CH_2); 2.38–2.47 m, 2 H (CH_2); 2.71 s, 1 H (C≡CH); 7.06 s, 1 H

(CH_{Ar}); 7.40 s, 1 H (CH_{Ar}); 8.10 s, 1 H (CH_{Ar}). ^{13}C NMR (75 MHz, $CDCl_3$): 23.27 (CH_2), 40.37 (CH_2), 74.96 (C≡**C**H), 82.12 (**C**≡CH), 117.11 (CH_{Ar}), 130.55 (CH_{Ar}), 137.04 (CH_{Ar}). HR-EI MS: $[M]^+$ calculated 204.0893, found 204.0901.

6-(1-Ethynylcyclopentyloxycarbonylamino)-2-(2,2,2-trifluoroacetylamino)hexanoic Acid Methyl Ester (**14**)

N-α-trifluoroacetyl-L-lysine methyl ester **11** (800 mg, 3.12 mmol, 1.1 eq.) was dissolved in a mixture of DCM (10 ml) and DMF (4 ml). After addition of diisopropylethylamine (1.09 ml, 6.24 mmol, 2.2 eq.) and the reaction was stirred at r.t. for 10 min. Next **12** (582 mg, 2.84 mmol, 1.0 eq.) was added and the reaction mixture was stirred at r.t. for 25 h. Then it was diluted with DCM (10 ml) and washed with water (3 × 15 ml) and brine (15 ml). After drying over $MgSO_4$, the solvent was evaporated under reduced pressure. The yellow oil thus obtained was purified by silica gel column chromatography (isohexane/ethyl acetate 3:1). The corresponding fractions were collected and the solvent evaporated to yield **14** as a colourless solid (321 mg, 0.82 mmol, 29%). R_F 0.22 (isohexane/ethyl acetate 2:1). ^{1}H NMR (300 MHz, $CDCl_3$): 1.28–1.56 m, 4 H (CH_2); 1.71–2.01 m, 6 H (CH_2); 2.07–2.27 m, 4 H (CH_2); 2.56 s, 1 H (C≡CH); 3.12–3.20 m, 2 H (CH_2NHR); 3.78 s, 3 H (CH_3); 4.59 dt, J = 7.48, 5.39, 1 H ($CHCO_2$Me); 4.72 s br, 1 H (NH); 7.14 s br, 1 H (NH). ^{13}C NMR (75 MHz, $CDCl_3$): 22.05 (CH_2), 23.27 (CH_2), 23.29 (CH_2), 29.52 (CH_2), 31.28 (CH_2), 40.01 (CH_2), 40.64 (CH_2), 52.55 ($CHCO_2$Me), 52.95 (CH_3), 72.49 (C≡**C**H), 80.12 (**C**≡CH), 84.75 (cyclopentane-Cq), 115.62 (CF_3), 155.31 (CO), 171.22 (CO), 171.25 (CO_2R). ^{19}F NMR (376 MHz, $CDCl_3$): –76.17 s (CF_3). HR-ESI MS: $[M - H]^+$ calculated 391.1486, found 391.1483.

2-Amino-6-(1-ethynylcyclopentyloxycarbonylamino)hexanoic Acid (**1**)

Compound **14** (93 mg, 0.24 mmol, 1.0 eq.) was dissolved in a mixture of THF/MeOH/water (2:2:1, 9 ml) and LiOH·H_2O (30 mg, 0.71 mmol, 3.0 eq.) was added. The reaction mixture was stirred at r.t. for 4 h and the solvent was evaporated under reduced pressure. The residue was purified by silica gel column chromatography (chloroform/methanol 4:1). The product **1** was obtained as a colourless solid (31 mg, 0.11 mmol, 46%). R_F 0.11 (chloroform/methanol 4:1). ^{1}H NMR (400 MHz, DMSO-d_6): 1.19–1.39 m, 4 H (CH_2); 1.54–1.74 m, 6 H (CH_2); 1.94–2.01 m, 2 H (CH_2); 2.04–2.11 m, 2 H (CH_2); 2.90 t, J = 6.92, 2 H (CH_2NHR); 3.22 dd, J = 6.97, 5.08, 1 H ($CHCO_2$H); 3.28 s, 1 H (C≡CH). ^{13}C NMR (100 MHz, DMSO-d_6): 22.67 (CH_2), 23.25 (CH_2), 29.57 (CH_2), 30.88 (CH_2), 40.29 (CH_2NHR), 40.68 (CH_2), 54.53 ($CHCO_2$H), 75.14 (C≡**C**H), 79.38 (cyclopentane-C_q), 86.01 (**C**≡CH), 155.44 (CO), 171.65 (CO_2H). HR-ESI MS: calculated $[M + H]^+$ 283.1652, found 283.1655.

2-Trimethylsilanylethynylcyclopentanol (**20**)

Trimethylsilylacetylene (3.95 ml, 28.53 mmol, 1.2 eq.) was disolved in diethylether (50 ml) and cooled to –78 °C. *n*-Butyllithium (11.42 ml, 28.53 mmol, 1.2 eq., 2.5 M in hexane) was added dropwise over a period of 15 min. The reaction mixture was stirred at –78 °C for 20 min. Then BF_3·OEt_2 (3.62 ml, 28.53 mmol, 1.2 eq.) was added and the reaction mixture was again stirred at –78 °C for 20 min. Cyclopentene oxide (2.06 ml, 23.78 mmol, 1.0 eq.) in diethylether (40 ml) was added slowly and stirred at –78 °C for 4 h. The reaction mixture was diluted with saturated $NaHCO_3$ solution (500 ml) and warmed to r.t.. It was extracted with diethylether (3 × 250 ml) and the combined organic fractions were washed with water

(3 × 300 ml) and brine (300 ml). After drying over $MgSO_4$, the solvent was evaporated and the residue was purified by silica gel column chromatography (isohexane/ethyl acetate 12:1). The product **20** was obtained as a colourless oil (555 mg, 3.05 mmol, 13%). R_F 0.47 (isohexane/ethyl acetate 4:1). ^{1}H NMR (600 MHz, $CDCl_3$): 0.14 s, 9 H ($Si(CH_3)_3$); 1.53–1.59 m, 1 H (CHOH); 1.64–1.78 m, 4 H (CH_2); 1.99–2.10 m, 2 H (CH_2); 2.58 dt, J = 7.6, 5.7, 1 H (CHC≡C). ^{13}C NMR (75 MHz, $CDCl_3$): 0.18 ($SiCH_3$), 21.75 (CH_2), 31.08 (CH_2), 33.50 (CH_2), 40.66 (**C**HC≡C), 79.44 (C≡**C**Si), 85.70 (CHOH), 108.73 (**C**≡CSi). HR-EI MS: $[M]^+$ calculated 182.1121, found 182.1113.

Imidazole-1-carboxylic Acid 2-Trimethylsilanylethynylcyclopentyl Ester (**13**)

2-Trimethylsilanylethynylcyclopentanol **20** (629 mg, 3.50 mmol, 1.0 eq.) was dissolved in DCM (30 ml) and carbonyldiimidazole (2.04 g, 12.60 mmol, 3.6 eq.) was added. After stirring at r.t. for 24 h, the reaction mixture was washed with brine (20 ml). The organic phase was dried over $MgSO_4$ and the solvent was evaporated. Compound **13** was obtained as light yellow crystals (1.18 g, 4.28 mmol, 54%) and was used in the next step without further purification. R_F 0.28 (isohexane/ethyl acetate 2:1). ^{1}H NMR (600 MHz, $CDCl_3$): 0.13 s, 9 H (CH_3); 1.78–1.90 m, 4 H (CH_2); 2.09–2.15 m, 1 H (CH_2); 2.39–2.28 m, 1 H (CH_2); 2.69–2.99 m, 1 H (CHC≡C); 5.32–5.34 m, 1 H (CHOH); 7.11 s, 1 H (CH_{Ar}); 7.40 s, 1 H (CH_{Ar}); 8.10 s, 1 H (CH_{Ar}). ^{13}C NMR (75 MHz, $CDCl_3$): 0.03 ($SiCH_3$), 22.49 (CH_2), 31.46 (CH_2), 31.51 (CH_2), 37.70 (**C**HC≡C), 85.27 (CHOR), 87.05 (C≡**C**Si), 105.09 (**C**≡CSi), 117.09 (CH_{Ar}), 130.59 (CH_{Ar}), 137.03 (CH_{Ar}), 148.01 (CO). HR-ESI MS: $[M + H]^+$ calculated 277.1367, found 277.1371.

2-(2,2,2-Trifluoroacetylamino)-6-(2-trimethylsilanylethynyl-cyclopentyloxycarbonylamino)hexanoic Acid Methyl Ester (**15**)

Compound **13** (628 mg, 2.27 mmol, 1.0 eq.) was dissolved in DCM (6 ml) and *N*-α-trifluoroacetyl-L-lysine methyl ester **11** (698 mg, 2.73 mmol, 1.2 eq.) and diisopropylethylamine (871 µl, 5.00 mmol, 2.2 eq.) was added. The reaction mixture was stirred at r.t. for 40 h. The organic phase was washed with water (3 × 5 ml) and brine (5 ml), and dried over $MgSO_4$. The crude product was purified by column chromatography on silica gel (isohexane/ethyl acetate 4:1). The product **15** was obtained as a colourless oil (295 mg, 0.64 mmol, 28%). R_F 0.15 (isohexane/ethyl acetate 4:1). ^{1}H NMR (300 MHz, $CDCl_3$): 0.01 s, 9 H (CH_2); 1.31–2.13 m, 12 H (CH_2); 2.71–2.81 m, 1 H (CHC≡C); 3.10–3.19 m, 2 H (CH_2NHCO); 3.77 s, 3 H (CH_3); 4.51–4.61 m, 1 H (CHOR); 4.75 s br, 1 H (NH); 5.00–5.03 m, 1 H ($CHCO_2Me$). ^{13}C NMR (75 MHz, $CDCl_3$): 0.01 ($SiCH_3$), 21.97 (CH_2), 22.51 (CH_2), 29.41 (CH_2), 30.97 (CH_2), 31.39 (CH_2), 31.71 (CH_2), 37.69 (**C**HC≡C), 39.88 (CH_2NHCO), 52.55 (**C**HCO_2Me), 52.82 (CH_3), 81.52 (CHOR), 87.23 (C≡**C**Si), 107.59 (**C**≡CSi), 118.63 (CF_3), 156.31 (CO), 171.61 (CO), 171.57 (CO). ^{19}F NMR (376 MHz, $CDCl_3$): –76.15 s (CF_3). HR-ESI MS: $[M - H]^+$ calculated 463.1882, found 463.1874.

2-Amino-6-(2-ethynylcyclopentyloxycarbonylamino)hexanoic Acid (**2**)

Compound **15** (200 mg, 0.43 mmol, 1.0 eq.) was dissolved in a mixture of THF/MeOH/water (2:2:1, 9 ml). $LiOH·H_2O$ (72 mg, 1.72 mmol, 4.0 eq.) was added and the reaction mixture was stirred at r.t. overnight. The solvent was evaporated and the residue was purified by silica gel column chromatography (chloroform/methanol 4:1). Product **2** was obtained as a colourless solid (80 mg, 0.28 mmol, 66%). R_F 0.26 (chloroform/methanol 4:1). ^{1}H NMR

(400 MHz, D_2O): 1.24–1.34 m, 2 H (CH_2); 1.39–1.47 m, 2 H (CH_2); 1.52–1.80 m, 6 H (CH_2); 1.93–1.99 m, 2 H (CH_2); 2.43 s, 1 H (C≡CH); 2.63–2.68 m, 1 H (CHC≡C); 3.16 t, J = 6.82, 2 H (CH_2NHCO); 3.75 t, J = 6.14, 1 H (CHOR); 4.96–4.99 m, 1 H ($CHCO_2H$). ^{13}C NMR (100 MHz, D_2O): 21.56 (CH_2), 21.94 (CH_2), 28.47 (CH_2), 30.04 (CH_2), 30.97 (CH_2), 30.86 (CH_2), 36.16 (**C**HC≡C), 39.85 (CH_2NHCO), 54.66 (CHOR), 81.58 (**C**HCO_2H), 86.04 (C≡**C**H), 158.26 (CO), 174.72 (CO). HR-ESI MS: $[M - H]^+$ calculated 281.1507, found 281.1505.

6-Pent-4-ynoylamino-2-(2,2,2-trifluoroacetylamino)hexanoic Acid Methyl Ester (**18**)

4-Pentynoic acid (383 mg, 3.90 mmol, 2.0 eq.), 1-hydroxybenzotriazole (HOBt) (527 mg, 3.90 mmol, 2.0 eq.) and [(benzotriazol-1-yloxy)dimethylaminomethylene]dimethylammonium tetrafluoroborate (TBTU) (1.19 g, 3.71 mmol, 1.9 eq.) were dissolved in DMF (4 ml). Diisopropylethylamine (2.21 ml, 12.68 mmol, 6.5 eq.) was added and the reaction mixture was stirred at r.t. for 10 min. Subsequently, *N*-α-trifluoroacetyl-L-lysine methyl ester **11** (500 mg, 1.95 mmol, 1.0 eq.) was added and the reaction mixture was stirred at r.t. overnight. The reaction was diluted with ethyl acetate (10 ml) and washed with water (4 × 8 ml) and brine (8 ml). The organic layer was dried over $MgSO_4$ and the residual brown oil was purified by column chromatography (isohexane/ethyl acetate 2:1). The product **18** (314 mg, 0.93 mmol, 48%) was obtained as a colourless solid. R_F 0.19 (isohexane/ethyl acetate 1:1). ^1H NMR (300 MHz, $CDCl_3$): 1.22–1.61 m, 4 H (CH_2); 1.79–1.97 m, 2 H (CH_2); 2.00 t, J = 2.61, 1 H (C≡CH); 2.35–2.40 m, 2 H (CH_2); 2.48–2.54 m, 2 H (CH_2); 3.17–3.36 m, 2 H (CH_2NHR); 3.77 s, 3 H (CH_3); 4.53 dt, J = 7.75, 4.79, 1 H ($CHCO_2R$); 5.93 s br, 1 H (NH); 7.40 s br, 1 H (NH). ^{13}C NMR (75 MHz, $CDCl_3$): 14.86 (CH_2), 22.04 (CH_2), 29.00 (CH_2), 30.84 (CH_2), 35.32 (CH_2), 38.47 (CH_2NHR), 52.62 (**C**HCO_2Me), 52.83 (CH_3), 69.39 (C≡**C**H), 82.90 (**C**≡CH), 117.48 (CF_3), 156.99 ($COCF_3$), 171.23 (CO), 171.56 (CO). ^{19}F NMR (376 MHz, $CDCl_3$): –76.16 s (CF_3). HR-ESI MS: $[M + H]^+$ calculated 337.1370, found. 337.1373.

2-Amino-6-pent-4-ynoylaminohexanoic Acid (**3**)

To a solution of **18** (200 mg, 0.60 mmol, 1.0 eq.) in THF/MeOH/water (2:2:1, 9 ml) was added $LiOH·H_2O$ (75 mg, 1.79 mmol, 3.0 eq.) and the mixture was stirred at r.t. for 4.5 h until TLC showed complete consumption of the starting material. The reaction mixture was filtrated and the solvent evaporated. The RP chromatography (water/MeOH, 0 → 5 → 10 → 20%) gave **3** as a colourless solid (120 mg, 0.53 mmol, 89%). R_F 0.66 (1 M $(NH_4)_2SO_4$/isopropanol 1:1). ^1H NMR (400 MHz, DMSO-d_6): 1.24–1.58 m, 6 H (CH_2); 2.21–2.25 m, 2 H (CH_2C≡C); 2.31–2.35 m, 2 H (CH_2CONH); 2.71 t, J = 2.60, 1 H (C≡CH); 2.81–2.84 m, 1 H ($CHCO_2H$); 3.00 t, J = 6.84, 2 H (CH_2NHR). ^{13}C NMR (100 MHz, DMSO-d_6): 14.67 (CH_2), 23.60 (CH_2), 29.53 (CH_2), 34.58 (CH_2), 35.64 (CH_2), 39.04 (CH_2NH), 56.23 (**C**HCO_2H), 71.68 (C≡**C**H), 84.19 (**C**≡CH), 170.70 (CONH), 179.51 (CO_2H). HR-ESI MS: $[M - H]^+$ calculated 225.1245, found 225.1244.

6-(2-Azidoacetylamino)-2-(2,2,2-trifluoroacetylamino)hexanoic Acid Methyl Ester (**19**)

Azido acetic acid (101 mg, 1 mmol) was added to a solution of HOBt (135.1 mg, 1 mmol, 2 eq.) and TBTU (321.0 mg, 1 mmol, 2 eq.) in dry DMF (2 ml). To this solution was added diisopropylethylamine (420 mg, 566 µl, 6.5 eq.) and the mixture was stirred at r.t. for 10 min. *N*-α-trifluoroacetyl-L-lysine methyl ester **11** (128.1 mg, 0.5 mmol, 0.5 eq.) dissolved in dry DMF (2 ml) was added via syringe to the above solution and the reaction mixture was

stirred at r.t. overnight. Then it was diluted with AcOEt (75 ml) and washed with H_2O (20 ml) and brine (20 ml). The oranic phase was dried over Na_2SO_4 and the solvent evaporated under reduced pressure. The crude product was purified by column chromatography on silica gel using mixture of isohexane/ethyl acetate (3:2 to 1:1) as eluent. The corresponding fractions were collected and the solvent evaporated under reduced pressure to yield **19** as a colorless viscous oil (110 mg, 65%). ^{1}H NMR (400 MHz, CD_3OD): 4.45 dd, 1 H, J = 4.55 (CH); 3.86 s, 2 H (CH_2N_3); 3.74 s, 3 H (OCH_3); 3.23 t, 2 H, J = 6.97 (CH_2); 1.89–2.00 m, 1 H (CH_2); 1.73–1.85 m, 1 H (CH_2); 1.30–1.62 m, 4 H (CH_2). ^{13}C NMR (100 MHz, CD_3OD): 172.8, 170.2, 159.3, 117.5, 54.2, 53.2, 53.1, 40.2, 31.5, 29.8, 24.4. HR-ESI MS: $[M + H]^+$ calculated 340.1233, found 340.1228; $[M - H]^+$ calculated 338.1076, found 338.1106.

2-Amino-6-(2-azidoacetylamino)hexanoic Acid (**4**)

To a solution of **19** (70 mg, 0.21 mmol) in a mixture of THF/MeOH/H_2O 2:2:1 (3 ml) was added LiOH·H_2O (26.5 mg, 0.63 mmol, 3 eq.) and the reaction mixtre was stirred at r.t. for 4 h. The reaction mixture was then filtered, washed with THF/MeOH/H_2O 2:2:1 (10 ml) and the solvents removed under reduced pressure. The crude product was dissolved in minimal amount of H_2O and loaded onto short RP silica gel column. The column was washed with H_2O and the product was eluted using 10% of MeOH in H_2O. Compound **4** was obtained as a white solid after lyophilization (32 mg, 68%). ^{1}H NMR (400 MHz, D_2O): 3.98 s (CH_2N_3); 3.50 bs, 1 H (CH); 3.22 t, J = 6.90, 2 H (CH_2); 1.72 bs, 2 H (CH_2); 1.45–1.60 m, 2 H (CH_2); 1.35 bs, 2 H (CH_2). ^{13}C NMR (100 MHz, D_2O): 178.6, 170.1, 55.0, 51.8, 39.0, 32.0, 27.9, 21.9. HR-ESI MS: $[M + H]^+$ calculated 230.1253, found 230.1242; $[2 M + H]^+$ calculated 459.2428, found 459.2418.

We thank Bayer–Schering Pharma AG, the Excellence Cluster CiPSM, and SFB 749 for generous financial support. Further support from the Fonds of the German Chemical Industry is gratefully acknowledged. M.V. thanks the LMU Munich for a postdoctoral qualification fellowship.

REFERENCES

1. a) Wang L., Xie J., Schultz P. G.: *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225; b) Young T. S., Schultz P. G.: *J. Biol. Chem.* **2010**, *285*, 11039.
2. a) Hino N., Okazaki Y., Kobayashi T., Hayashi A., Sakamoto K., Yokohama S.: *Nat. Methods* **2005**, *2*, 201; b) Hendrickson T. L., de Crecy-Lagard V., Schimmel P.: *Annu. Rev. Biochem.* **2004**, *73*, 147; c) Sakamoto K., Hayashi A., Sakamoto A., Kiga D., Nakayama H., Soma A., Kobayashi T., Kitabatake M., Takio K., Saito K., Shirouzu M., Hirao I., Yokohama S.: *Nucleic Acids Res.* **2002**, *30*, 4692.
3. a) Hohsaka T., Ashizuka Y., Taira H., Murakami H., Sisido M.: *Biochemistry* **2001**, *40*, 11060; b) Hohsaka T., Sisido M.: *Curr. Opin. Chem. Biol.* **2002**, *6*, 809.
4. a) Hao B., Gong W., Ferguson T. K., James C. M., Krzycki J. A., Chan M. K.: *Science* **2002**, *296*, 1462; b) Atkins J. F., Gesteland R.: *Science* **2002**, *296*, 1409; c) Srinivasan G., James C. M., Krzycki J. A.: *Science* **2002**, *296*, 1459.
5. a) Polycarpo C. R., Herring S., Berube A., Wood J. L., Soll D., Ambrogelly A.: *FEBS Lett.* **2006**, *580*, 6695; b) Li W.-T., Mahapatra A., Longstaff D. G., Bechtel J., Zhao G., Kang P. T., Chan M. K., Krzycki J. A.: *J. Mol. Biol.* **2009**, *385*, 1156.

6. a) Nguyen D. P., Lusic H., Neumann H., Kapadnis P. B., Deiters A., Chin J. W.: *J. Am. Chem. Soc.* **2009**, *131*, 8720; b) Fekner T., Li X., Lee M. M., Chan M. K.: *Angew. Chem. Int. Ed.* **2009**, *48*, 1633; c) Neumann H., Peak-Chew S. Y., Chin J. W.: *Nat. Chem. Biol.* **2008**, *4*, 232.
7. a) Yanagisawa T., Ishii R., Fukunaga R., Kobayashi T., Sakamoto K., Yokoyama S.: *Chem. Biol.* **2008**, *15*, 1187; b) Chen P. R., Groff D., Guo J., Ou W., Cellitti S., Geierstanger B. H., Schultz P. G.: *Angew. Chem. Int. Ed.* **2009**, *48*, 4052; c) Gautier A., Nguyen D. P., Lusic H., An W., Deiters A., Chin J. W.: *J. Am. Chem. Soc.* **2009**, *132*, 4086.
8. a) Cheung A. Y., Söll D.: *J. Biol. Chem.* **1984**, *15*, 9953; b) Liu C. C., Schultz P. G.: *Annu. Rev. Biochem.* **2010**, *79*, 413.
9. Gutsmiedl K., Wirges Ch. T., Ehmke V., Carell T.: *Org. Lett.* **2009**, *11*, 2405.
10. a) Rostovtsev V. V., Green L. G., Fokin V. V., Sharpless K. B.: *Angew. Chem. Int. Ed.* **2002**, *41*, 2596; b) Tornoe C. W., Christensen C., Meldal M.: *J. Org. Chem.* **2002**, *67*, 3057.
11. a) Gramlich P. M. E., Wirges Ch. T., Manetto A., Carell T.: *Angew. Chem. Int. Ed.* **2008**, *47*, 8350; b) Gramlich P. M. E., Warncke S., Gierlich J., Carell T.: *Angew. Chem. Int. Ed.* **2008**, *47*, 3442; c) Gierlich J., Burley G. A., Gramlich P. M. E., Hammond D. M., Carell T.: *Org. Lett.* **2006**, *8*, 3639; d) Gierlich J., Gutsmiedl K., Gramlich P. M. E., Schmidt A., Burley G. A., Carell T.: *Chem. Eur. J.* **2007**, *13*, 9486; e) Gramlich P. M. E., Wirges Ch. T., Gierlich J., Carell T.: *Org. Lett.* **2008**, *10*, 249.
12. a) Chiba Y., Akeboshi H.: *Biol. Pharm. Bull.* **2009**, *32*, 786; b) Helenius A.: *Mol. Biol. Cell* **1994**, *5*, 253.
13. Kaya E., Gutsmiedl K., Vrábel M., Mueller M., Thumbs P., Carell T.: *ChemBioChem* **2009**, *10*, 2858.
14. Hao B., Zhao G., Kang P. T., Soares J. A., Fergoson T. K., Galucci J., Krzycki J. A., Chan M. K.: *Chem. Biol.* **2004**, *11*, 1317.