

Orchestrating the biosynthesis of an unnatural pyrrolysine amino acid for its direct incorporation into proteins inside living cells

Michael Ehrlich^{[a]+}, Michael J. Gattner^{[a]+}, Bastien Viverge^[a], Johanna Bretzler^[a], David Eisen^[a], Michael Stadlmeier^[a], Milan Vrabec^[b], Thomas Carell^{*[a]}

^[a] Center for Integrated Protein Science at the Department of Chemistry, Ludwig-Maximilians Universität München, Butenandtstr. 5-13, 81377 Munich

^[b] Institute of Organic Chemistry and Biochemistry; Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10, Prague, Czech

* Corresponding author E-mail: Thomas.Carell@lmu.de

^[+] These authors contributed equally to this work.

This article was published online 02.04.2015 in Chemistry – A European Journal

DOI: 10.1002/chem.201500971 Print: vol. 21(21):7701 - 7704

Abstract: We here report the construction of an *E. coli* expression system able to manufacture an unnatural amino acid by an artificial biosynthesis. This can be orchestrated with incorporation into protein by amber stop codon suppression inside a living cell. In our case an alkyne bearing pyrrolysine amino acid was biosynthesized and incorporated site-specifically allowing orthogonal double protein labeling.

The incorporation of unnatural amino acids into proteins is a highly topical research field as it allows to modulate the properties of proteins by chemical means.^[1] Practically, the incorporation of amino acids that are amenable to further chemical modification by orthogonal “click”-reaction is interesting for the addition of e. g. polyethylene glycols, toxophores and fluorophores to proteins for diagnostic or medical applications.^[2] A recently introduced technology in this direction is based on the 22nd genetically encoded amino acid pyrrolysine (Pyl), which is biosynthesized from two L-lysine amino acids by the enzymes PylB, PylC and PylD (Figure 1 and SI Figure 1).^[3] The biosynthetic pathway was elucidated by feeding experiments^[4] and structural characterization of the involved enzymes.^[5] Pyl is subsequently loaded onto a special Pyl-tRNA (PylT) with the assistance of a specific aminoacyl-tRNA synthetase (PylRS). PylT delivers Pyl to the ribosome for incorporation into the nascent polypeptide chain in response to the presence of the amber stop codon (UAG) in the mRNA. Therefore the pyrrolysine system reprograms this stop codon for Pyl insertion.

As PylRS is rather promiscuous and can be adopted to bind unnatural amino acids using bio-engineering, the PylRS/pylT pair has enabled the incorporation of a large variety of Pyl derivatives featuring additional functional groups e. g. for click modification.^[6] One problem associated with this technology is, however, that the unnatural Pyl amino acids need to be synthesized in rather large quantities to reach the required millimolar concentrations in the culture medium necessary for efficient incorporation. These laborious syntheses can be circumvented with a host organism (*E. coli*) that produces the modified unnatural amino acid by itself.^[7] This in consequence requires the insertion and modulation of the Pyl biosynthetic machinery in the *E. coli* host system.^[4b, 4c] Thus we planned to study if the Pyl biosynthesis pathway allows *in vivo* construction and finally incorporation of the unnatural 3S-ethynylpyrrolysine (ePyl) amino acid, which may be amenable to subsequent click modification (Figure 1).

Since Pyl is derived from the PylB product 3R-methyl-D-ornithine and one L-lysine *via* the action of the enzymes PylC and PylD (SI Figure 1), we cloned the genes of both enzymes into a first plasmid (1) (Figure 2). A second plasmid (2) was prepared containing one gene copy of the PylRS aminoacyl-tRNA synthetase gene *pylS* and three copies of the *pylT* gene. A third plasmid (3) was used with the gene for the human carbonic anhydrase 2 (*hCA*).^[8] This manipulated *hCA* gene contains an amber stop codon (GGG → TAG) at the amino acid position G131.^[9] All three plasmids were co-transformed into *E. coli* cells. Plasmid 3 was selected on carbenicillin containing medium, while plasmid 2 was equipped with a chloramphenicol acetyltransferase (*CAT*) gene bearing an amber stop codon at either position A69 or D111 upstream of the active site. This plasmid represents the master reporter as it allows growth of *E. coli* cells in the presence of chloramphenicol only, if a Pyl derivative is biosynthesized and incorporated in *CAT* efficiently by stop codon suppression.

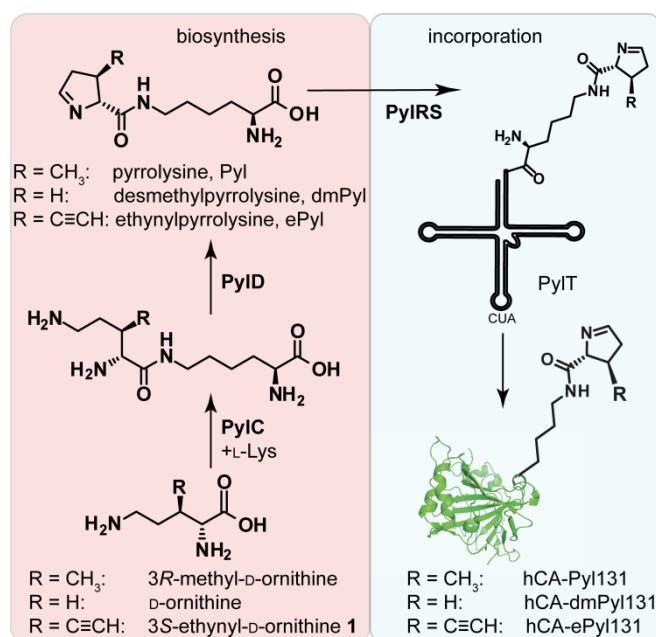


Figure 1. Developed methodology for biosynthesis and incorporation of the novel ethynylpyrrolysine (ePyl). The new amino acid 3S-ethynyl-D-ornithine **1** was converted to ePyl by hijacking the pyrrolysine biosynthesis machinery analog to the biosynthesis of pyrrolysine (Pyl) via biosynthesis enzymes PyIC and PyID. A site-specific incorporation into the protein human carbonic anhydrase 2 (hCA, PDB: 1BCD) by translation with aminoacyl-tRNA synthetase (PylRS)/Pyl-tRNA (PylT) pair at the amber stop codon was possible.

For selection of successfully transformed *E. coli* cells we added carbenicillin, chloramphenicol and initially D-ornithine. This system allowed us to test if D-ornithine is converted to desmethylpyrrolysine (dmPyl) by the action of PyIC and PyID followed by final incorporation into proteins via amber suppression.^[4b] Indeed, we observed growing *E. coli* showing that the biosynthesis operates and thus dmPyl is successfully incorporated into CAT.

We next used 3'-substituted D-ornithine derivatives and finally found growing *E. coli* cells in the presence of carbenicillin, chloramphenicol and 3S-ethynyl-D-ornithine **1**. This demonstrates that 3S-ethynyl-D-ornithine **1** is able to serve as a D-ornithine surrogate during biosynthesis.

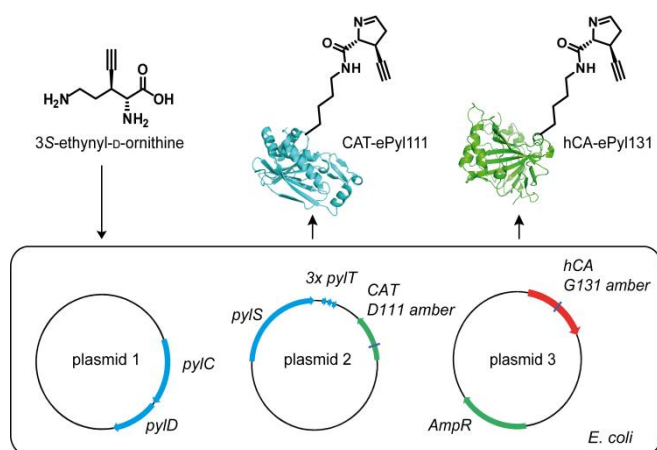
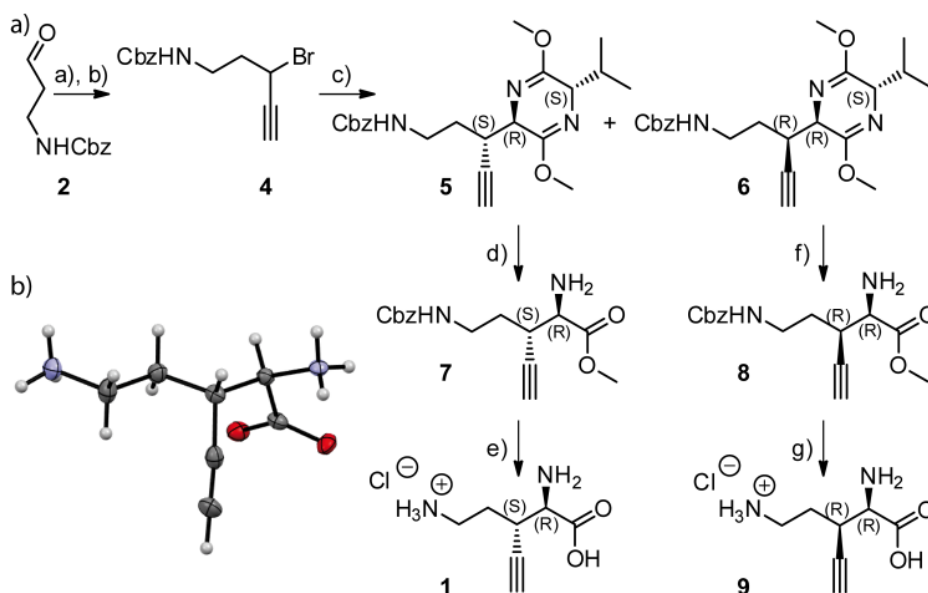


Figure 2. Biosynthetic conversion of 3S-ethynyl-D-ornithine **1 to ethynylpyrrolysine (ePyl) and incorporation into the amber codon bearing proteins chloramphenicol acetyltransferase (CAT) and human carbonic anhydrase 2 (hCA) using an *E. coli* expression system.** *E. coli* cells were transformed with following coding plasmids: plasmid 1: *pylC*, *pylD*; plasmid 2: *pylS*, 3x *pylT* and CAT-A69amber or CAT-D111amber; plasmid 3: hCA131amber and an ampicillin resistance gene (PDBs: 1BCD, 1Q23).

The synthesis of the starting material 3*S*-ethynyl-D-ornithine **1** is conveniently possible as depicted in Scheme 1. We started with Carboxybenzyl (Cbz) protected 3-amino-propionaldehyde **2** which was reacted with ethynylbromide in a *Grignard* reaction. The formed alcohol **3** was converted into the bromide **4** via an *Appel* reaction.^[10] Subsequent substitution of the bromide **4** with the deprotonated *S*-configured *Schöllkopf* reagent provided the two pyrazine derivatives with the corresponding *SR* (**5**) and *RR* (**6**) configuration in a diastereomeric mixture which could be separated using silica column chromatography. We were unable to detect the *SS* and *RS* products, showing the high stereo control provided by the *Schöllkopf* auxiliary.^[11] Sequential deprotection of both diastereoisomers to the methoxyesters **7** and **8** and subsequent hydrolysis provided the 3*S*,2*R* and 3*R*,2*R* configured ethynylornithines **1** and **9**. The absolute configuration was verified by a crystal structure of the 3*S*,2*R*-ethynylornithine **1** (Scheme 1b).



Scheme 1. a) Synthesis of 3*S*-ethynyl-D-ornithine **1** and 3*R*-ethynyl-D-ornithine **9**. Reagents and conditions a) $\text{HC}\equiv\text{CMgBr}$, THF, $-78\text{ }^\circ\text{C}$ to rt, 2 h; b) PPh_3 , CBr_4 , DCM, $0\text{ }^\circ\text{C}$ to rt, 5 h; 50% over 2 steps; c) (*S*)-2,5-dihydro-3,6-dimethoxy-2-isopropylpyrazine, *n*-BuLi, THF, $-78\text{ }^\circ\text{C}$, 30 min, 72% (*SR/RR*: 2/5); d) 3M HCl, MeOH, H_2O , rt, 2 d, 78%; e) 12 M HCl, H_2O , $80\text{ }^\circ\text{C}$, 15 h, 91%; f) 3M HCl, MeOH, H_2O , rt, 2 d, 86%; g) 12 M HCl, H_2O , $80\text{ }^\circ\text{C}$, 15 h, 11%. b) X-ray structure of **1**.

To investigate how efficiently 3*S*-ethynyl-D-ornithine **1** serves as a biosynthetic precursor for Pyl-derivatives we performed time dependent growth studies. Growth was observed in the presence of D-ornithine and 3*S*-ethynyl-D-ornithine **1** but interestingly not with 3*R*-ethynyl-D-ornithine **9** (SI Figure 2). We explain this result with the fact that PylC has a hydrophobic pocket to bind the methyl group of its natural 3*R*-methyl-D-ornithine substrate that is large enough to accommodate the 3*S*-ethynyl unit.^[5b] It is also interesting that *E. coli* cells show faster growth in the presence of 3*S*-ethynyl-D-ornithine compared to D-ornithine arguing that filling of the hydrophobic pocket enhances substrate affinity and hence the efficiency of the biosynthesis (SI Figure 3).

In order to prove that the designed *E. coli* cells are indeed able to manufacture 3*S*-ethynylpyrrolysine (ePyl) for subsequent incorporation into proteins, we next isolated the hCA protein from the *E. coli* cultures (SI Figure 4). The protein was conveniently purified in a single chromatographic step using a *p*-aminomethylbenzene-sulfonamide column exploiting the binding affinity to the Zn(II)-ion in the active site of hCA.^[12] When we added D-ornithine in our control experiments to the medium we typically isolated around 6 mg of hCA per L of culture. Upon addition of 3*S*-ethynyl-D-ornithine **1**, the yield increased dramatically to 75 mg per L arguing in combination with the results from the growth curves that the speed of the biosynthesis of ePyl is the determining factor.

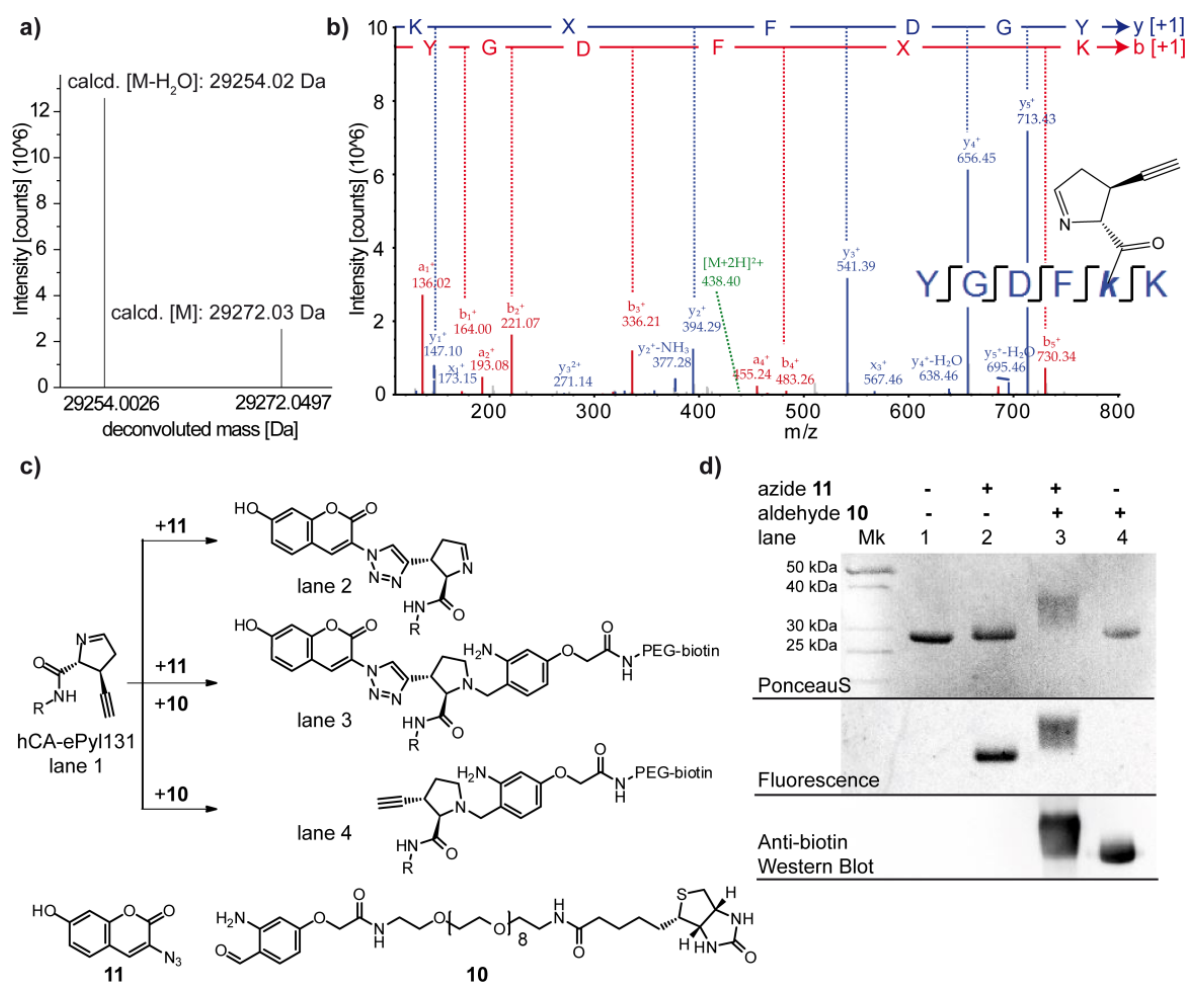


Figure 3. Site-specific incorporation of ePyl in hCA and orthogonal double labeling. a) Deconvoluted intact mass spectrum of hCA-ePyl131. Detected protein species: [M]: 29272.05 Da (calcd.: 29272.03 Da; relative abundance 20%), [M-H₂O]: 29254.00 (calcd.: 29254.02 Da; relative abundance 100%). b) Identification of ePyl peptide fragment after tryptic digestion: hCA sequence coverage of 88.03%. MS/MS spectra of peptide fragment YGDFFK (X = ePyl). Identified a- and b-ions are marked in red, identified x- and y-ions are marked in blue. Parent ion: [M+2H]²⁺_{obs.} = 438.71616 (ΔM = -0.01 ppm). c) Scheme of the single and/or double ePyl modification reaction with either azide **11** or the 2-aminobenzaldehyde reagent **10** or using both. d) Labeling of hCA-ePyl131. Protein bands were visualized for all samples after SDS-PAGE by fluorescence measurement. After transfer to a nitrocellulose membrane a PonceauS staining followed by Western Blot using an anti-biotin-HRP conjugate and chemiluminescent readout was performed.

Subsequently, the incorporation of ePyl into hCA enzyme was proven by intact protein mass spectrometric studies using a high resolution FT-ICR mass spectrometer. Indeed, the isolated hCA protein showed the correct molecular weight for the ePyl containing hCA-ePyl131 (Figure 3a; SI Figure 5). The position of ePyl incorporation was verified by the mass spectrometric analysis after tryptic digestion on an Orbitrap XL system (LC-MS/MS). This experiment demonstrates that the incorporation proceeded at the desired position (G131 → ePyl131, Figure 3b, SI Figure 6). In addition we verified the correct incorporation in CAT-ePyl111 (SI Figure 7). As further control we analyzed the isolated hCA-dmPyl131 protein that was generated in the presence of D-ornithine and observed correct molecular weight for dmPyl incorporation (G131 → dmPyl131; SI Figure 8). In summary the data show that our Pyl biosynthesis cluster containing *E. coli* cells produced an hCA protein variant with a G131 → ePyl131 mutation.

To demonstrate that the ePyl amino acid can be further manipulated inside a protein we next employed two orthogonal chemical modification reactions. We first used the chemistry developed by Geierstanger and coworkers^[13] to form a tertiary amine adduct by site-specific reaction at the pyrroline ring structure of Pyl with 2-aminobenzaldehyde (ABA) derivatives. We treated to this end the ePyl containing hCA with a biotinylated 2-aminobenzaldehyde derivative **10** shown in Figure 3c. The reagent was added to the hCA-ePyl131 protein solution and the protein was analyzed after 16 h reaction at r.t. by gel electrophoresis. Full conversion of the hCA-ePyl131 protein was observed based on migration differences of the labeled protein. The biotinylation of hCA-ePyl131 was verified in a Western blot using an anti-biotin antibody (Figure 3d).

Next, we studied the ability to click-react the alkyne unit by Cu(I) catalyzed reaction (CuAAC) with 7-hydroxycoumarin azide (**11**), which turns fluorescent after cycloaddition.^[14] Indeed, after treating the protein at r.t. for 16 h at pH 7.4 the product protein showed the expected blue fluorescence (Figure 3d). Tryptic digestion and LC-MS/MS analysis allowed us to unambiguously identify the coumarin labeled protein (SI Figure 9).

Finally, we investigated if the ePyl amino acid can be used for double labeling. To this end the hCA-ePyl131 protein was first treated with azide **11** (2 h, r.t. pH = 7.4) and subsequently reacted with the biotinylated 2-aminobenzaldehyde **10** (14 h, r.t., pH = 7.4). The obtained product protein showed blue fluorescence and presence of the inserted biotin label was proven by immunostaining with the anti-biotin antibody. Further LC-MS/MS analysis provided further support for the successful double labeling of the hCA protein (Figure 3d, SI Figure 10).

In summary we generated an *E. coli* expressions system which is able to biosynthesize an unnatural amino acid. This biosynthesis can be orchestrated with methods that allow the incorporation of the resulting unnatural amino acid into proteins by amber suppression. In this particular case we inserted an alkyne modified pyrrolysine amino acid which is amenable to two different orthogonal click reactions at one amino acid providing access to highly modified proteins.

Acknowledgements

We thank the Excellence Cluster CiPS^M and SFB749 (Project A4) as well as SFB1032 (Project A5) for financial support. Further support is acknowledged from the Fonds der Chemischen Industrie (predoctoral fellowship to M.S.). We thank Dr. Peter Mayer (LMU Munich) for X-ray analysis.

Keywords: pyrrolysine • unnatural amino acid • amber suppression • bioorganic chemistry • synthetic biology

- [1] C. H. Kim, J. Y. Axup, P. G. Schultz, *Curr Opin Chem Biol* **2013**, *17*, 412-419.
- [2] a) K. Lang, J. W. Chin, *Chem Rev* **2014**, *114*, 4764-4806; b) J. M. Harris, R. B. Chess, *Nat Rev Drug Discov* **2003**, *2*, 214-221; c) P. J. Carter, *Exp Cell Res* **2011**, *317*, 1261-1269.
- [3] a) B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, M. K. Chan, *Science* **2002**, *296*, 1462-1466; b) G. Srinivasan, C. M. James, J. A. Krzycki, *Science* **2002**, *296*, 1459-1462.
- [4] a) M. A. Gaston, L. Zhang, K. B. Green-Church, J. A. Krzycki, *Nature* **2011**, *471*, 647-650; b) S. E. Cellitti, W. Ou, H. P. Chiu, J. Grunewald, D. H. Jones, X. Hao, Q. Fan, L. L. Quinn, K. Ng, A. T. Anfora, S. A. Lesley, T. Uno, A. Brock, B. H. Geierstanger, *Nat Chem Biol* **2011**, *7*, 528-530; c) D. G. Longstaff, R. C. Larue, J. E. Faust, A. Mahapatra, L. Zhang, K. B. Green-Church, J. A. Krzycki, *Proc Natl Acad Sci U S A* **2007**, *104*, 1021-1026.
- [5] a) F. Quitterer, A. List, W. Eisenreich, A. Bacher, M. Groll, *Angew Chem Int Ed Engl* **2012**, *51*, 1339-1342; b) F. Quitterer, A. List, P. Beck, A. Bacher, M. Groll, *J Mol Biol* **2012**, *424*, 270-282; c) F. Quitterer, P. Beck, A. Bacher, M. Groll, *Angew Chem Int Ed Engl* **2013**, *52*, 7033-7037.
- [6] a) E. Kaya, K. Gutsmedl, M. Vrabel, M. Muller, P. Thumbs, T. Carell, *ChemBioChem* **2009**, *10*, 2858-2861; b) D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters, J. W. Chin, *J. Am. Chem. Soc.* **2009**, *131*, 8720-8721; c) T. Fekner, X. Li, M. M. Lee, M. K. Chan, *Angew. Chem., Int. Ed.* **2009**, *48*, 1633-1635; d) E. Kaya, M. Vrabel, C. Deiml, S. Prill, V. S. Fluxa, T. Carell, *Angew Chem Int Ed Engl* **2012**, *51*, 4466-4469; e) M. J. Gattner, M. Vrabel, T. Carell, *Chem. Commun. (Camb)*. **2013**, *49*, 379-381; f) W. Wan, J. M. Tharp, W. R. Liu, *Biochim Biophys Acta* **2014**, *1844*, 1059-1070.
- [7] a) R. A. Mehl, J. C. Anderson, S. W. Santoro, L. Wang, A. B. Martin, D. S. King, D. M. Horn, P. G. Schultz, *J. Am. Chem. Soc.* **2003**, *125*, 935-939; b) J.-E. Jung, S. Y. Lee, H. Park, H. Cha, W. Ko, K. Sachin, D. W. Kim, D. Y. Chi, H. S. Lee, *Chemical Science* **2014**, *5*, 1881-1885.
- [8] C. T. Supuran, *Nat. Rev. Drug Discov.* **2008**, *7*, 168-181.
- [9] S. Schneider, M. J. Gattner, M. Vrabel, V. Flügel, V. Lopez-Carrillo, S. Prill, T. Carell, *ChemBioChem* **2013**, *14*, 2114-2118.
- [10] R. Appel, *Angew Chem Int Ed Engl* **1975**, *14*, 801-811.
- [11] U. Schöllkopf, U. Groth, C. Deng, *Angew Chem Int Ed Engl* **1981**, *20*, 798-799.
- [12] P. L. Whitney, *Anal Biochem* **1974**, *57*, 467-476.
- [13] W. Ou, T. Uno, H. P. Chiu, J. Grunewald, S. E. Cellitti, T. Crossgrove, X. Hao, Q. Fan, L. L. Quinn, P. Patterson, L. Okach, D. H. Jones, S. A. Lesley, A. Brock, B. H. Geierstanger, *Proc Natl Acad Sci U S A* **2011**, *108*, 10437-10442.
- [14] K. Sivakumar, F. Xie, B. M. Cash, S. Long, H. N. Barnhill, Q. Wang, *Org Lett* **2004**, *6*, 4603-4606.