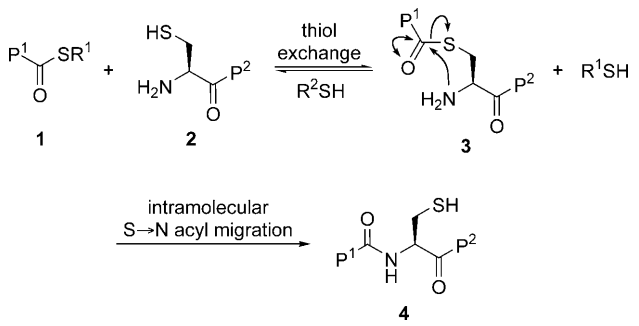
Protein Engineering

DOI: 10.1002/anie.200904472

A Pyrrolysine Analogue for Site-Specific Protein Ubiquitination**

*Xin Li, Tomasz Fekner, Jennifer J. Ottesen, and Michael K. Chan**

Native chemical ligation (NCL)[1] is the most widely used strategy for the convergent synthesis of proteins and large peptides.[2] It involves the chemoselective reaction between a coupling partner **1** armed with a C-terminal thioester and a peptide segment **2** bearing an N-terminal cysteine residue

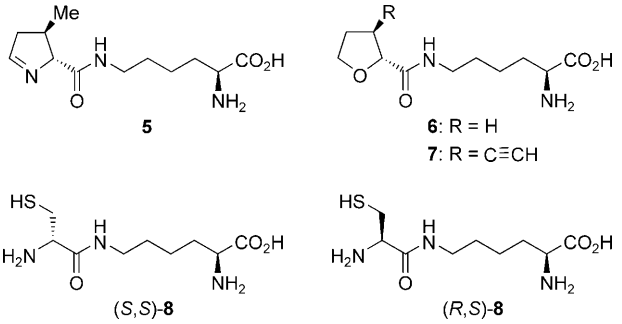


1 **2** **3** **4**

Scheme 1. Cysteine-based NCL.

(Scheme 1) to generate a product with a native backbone. Specifically, a reversible transthioesterification in the presence of an exogenous thiol[2] R^2SH gives an intermediate **3**, which in turn undergoes irreversible intramolecular S→N acyl transfer to give the final peptide **4**. Herein, we introduce a genetically encoded pyrrolysine analogue that places a ligation handle directly into a recombinant protein. We used NCL at this internal ligation site to generate a semisynthetic ubiquitinated protein.

The cellular machinery for the incorporation of pyrrolysine (**5**, Scheme 2),[3,4] the 22nd genetically encoded amino



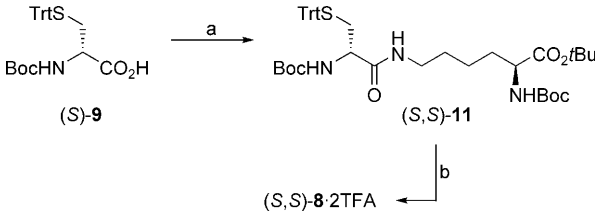
5 **6**: R = H **7**: R = C≡CH

(*S*,*S*)-**8** (*R*,*S*)-**8**

Scheme 2. Pyrrolysine (**5**) and analogues **6**–**8**.

acid, is sufficiently flexible to enable a number of other lysine derivatives to read through the amber stop codon.[5–9] We previously described **6**, a stable THF-based analogue of **5**.[10] We also introduced **7**, which, owing to the presence of a terminal alkyne functionality, can be used as a chemical handle to label proteins through click chemistry.[11] To further expand the range of available pyrrolysine analogues with unique and useful reactivity, we decided to test whether the D-cysteine-based analogue (*S*,*S*)-**8** (D-Cys-ε-Lys) could read through the UAG codon. We focused our attention on this cysteine isomer because our related readthrough studies of simple pyrrolysine analogues for protein click chemistry indicated that the presence of a lysine acyl substituent with the analogous sense of chirality to that found in **5** has a profound and beneficial influence on incorporation efficiency.[12] For comparison purposes, however, we also included in our studies the diastereomeric analogue (*R*,*S*)-**8** (L-Cys-ε-Lys).

The target pyrrolysine analogue (*S*,*S*)-**8** was prepared by coupling the N,S-protected cysteine derivative (*S*)-**9** with Boc-Lys-O*t*Bu (**10**)[11] to provide amide (*S*,*S*)-**11** in excellent yield (96%, d.r. > 99.9:0.1; Scheme 3). Full deprotection with trifluoroacetic acid (TFA)/Et_3SiH furnished (*S*,*S*)-**8** as its TFA salt. The diastereomer (*R*,*S*)-**8** was prepared in an analogous manner (see the Supporting Information).



(*S*)-**9** (*S*,*S*)-**11** (*S*,*S*)-**8**·2TFA

Scheme 3. Synthesis of the pyrrolysine analogue (*S*,*S*)-**8**: a) Boc-Lys-O*t*Bu (**10**), BOP, *N*-methylmorpholine, CH_2Cl_2, room temperature, 24 h, 96%; b) TFA, Et_3SiH, CH_2Cl_2, room temperature, 3 h, 90% (approximate yield). Boc = *tert*-butoxycarbonyl, BOP = (benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate, Trt = trityl (triphenylmethyl).

[*] X. Li,[+] Dr. T. Fekner,[+] Prof. M. K. Chan
The Ohio State Biophysics Program, Departments of Chemistry and Biochemistry, The Ohio State University
484 W 12th Avenue, Columbus, OH 43210 (USA)
Fax: (+1) 614-292-6773
E-mail: chan@chemistry.ohio-state.edu
Homepage: http://www.chemistry.ohio-state.edu/~chan/
Prof. J. J. Ottesen
Department of Biochemistry, The Ohio State University
484 W 12th Avenue, Columbus, OH 43210 (USA)
[+] These authors contributed equally.

[**] This research was supported by a grant from the US National Institutes of Health (GM061796) and an American Heart Association Great Rivers Affiliate Predoctoral Fellowship (0815449D) to X.L. We also thank the staff of the CCIC Mass Spectrometry and Proteomics Facility at OSU for protein analysis by mass spectrometry, Professor Michael Zhu (OSU) for providing *Rattus norvegicus* CaM cDNA, and Professor Bing Hao (UConn) for the human ubiquitin cDNA.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200904472.

To evaluate the UAG codon readthrough efficiency for the two diastereomers (*S*,*S*)-**8** and (*R*,*S*)-**8**, we employed the brightly emitting red fluorescent protein mCherry[13] as a reporter.[12] In brief, the Lys55 codon of this protein was mutated site specifically to UAG and inserted into the plasmid pPylST, which harbors the pyrrolysine tRNA (PylT) and synthetase (PylS) genes. *Escherichia coli* strain BL21-(DE3) transformed with this plasmid was grown in the Terrific Broth medium supplied with either (*S*,*S*)-**8** or (*R*,*S*)-**8** at varying concentrations. The results of the mCherry readthrough assays demonstrate that the presence of either isomer enables readthrough of the UAG codon, although (*S*,*S*)-**8** serves as a much better substrate in terms of incorporation efficiency (Figure 1).

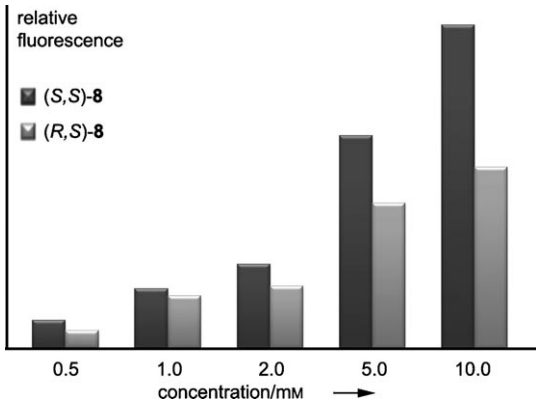


Figure 1. Dose-dependent readthrough of (*S*,*S*)-**8** and (*R*,*S*)-**8**.

The incorporation of **8** into a target protein provides a chemical handle for branching through NCL at a specific site. A ubiquitinated protein is perhaps the most important example of a branched protein structure. It is generated by ubiquitination, a special posttranslational modification in which the C-terminal glycine residue (G76) of the small protein ubiquitin is attached to the ε-amino group of a lysine residue in a substrate through an isopeptide bond, the formation of which is catalyzed by a series of enzymes.[14,15] Protein ubiquitination plays an important role in many cellular processes, including protein degradation, cellular signaling, cell division and differentiation, and protein trafficking.[14–18] Biochemical and structural studies on ubiquitination require the isolation or generation of homogenous ubiquitinated proteins. However, isolation from an in vivo source is usually low yielding, and the obtained protein may contain other posttranslational modifications. On the other hand, the in vitro reconstitution of ubiquitination with purified proteins and enzymes often suffers from low productivity and limited availability of the specific ubiquitin ligases. In a pioneering study that highlights the use of NCL to generate ubiquitinated proteins, the Muir research group used NCL assisted by a ligation auxiliary to prepare ubiquitinated histone 2B (H2B) from three separate pieces.[19,20] Unfortunately, this method is convenient only for proteins that undergo ubiquitination near the termini, and the advanced synthetic chemistry techniques required limit its wider use as a general tool by researchers in biochemistry and structural biology.

Herein, we demonstrate that it is possible to generate a site specifically ubiquitinated protein in a single ligation step from two genetically encoded segments by taking advantage of our pyrrolysine analogue (*S*,*S*)-**8**. This analogue is an excellent mimic of the three key structural elements of the ubiquitination site. Namely, its lysine functionality is identical to the target lysine residue in the substrate protein, the D-Cys residue replaces the terminal Gly residue of ubiquitin, and an isopeptide bond between ubiquitin and the target lysine side chain is retained. Thus, the replacement of the Gly76 residue from ubiquitin with D-Cys is the only difference between the semisynthetic product and a natively ubiquitinated protein.

For our model studies, we chose calmodulin (CaM), a small 17-kDa protein that plays a central role in calcium signaling in eukaryotes. The reversible ubiquitination of CaM is catalyzed by E3-CaM (ubiquitin–calmodulin ligase, EC 6.3.2.21)[21–23] at Lys21[24] and leads to the production of ubiquitinated CaM (Ub-CaM).[24,25] Instead of targeting CaM for proteosome degradation,[26] the ubiquitination of CaM modulates its regulatory activities.

To generate (*S*,*S*)-**8**-containing CaM ((*S*,*S*)-**8**-CaM), *Rattus norvegicus* CaM(Lys21Pyl) cDNA was subcloned into pPylST (Figure 2a). The recombinant protein (*S*,*S*)-**8**-CaM produced in this way was purified by hydrophobic-interaction chromatography as described previously.[11] A quantity of 0.9 mg of (*S*,*S*)-**8**-CaM could be isolated from 50 mL of the culture supplied with 2.5 mM (*S*,*S*)-**8**. Significantly, MALDI-TOF MS analysis of the purified product (Figure 2c) demonstrated that the reactive Cys mimic remains intact throughout expression in a cellular system.

The truncated *Homo sapiens* ubiquitin Ub75 (containing residues 1–75) was produced as an Ub75/intein/CBD (chitin-binding domain) fusion protein and purified by chitin-affinity chromatography. On-column thiolysis was initiated with sodium 2-mercaptoethane sulfonate (MESNa) to generate the Ub75 thioester (Ub75-SR, $R = CH_2CH_2SO_3Na$), which was mixed with (*S*,*S*)-**8**-CaM in a 5:1 molar ratio to promote NCL (Figure 2a). The reaction mixture was incubated at room temperature overnight, and the ubiquitinated calmodulin product (Ub*-CaM) was separated from unreacted Ub75-SR and (*S*,*S*)-**8**-CaM by anion-exchange chromatography. Approximately 30% of the recombinant protein (*S*,*S*)-**8**-CaM was converted into ubiquitinated CaM (Figure 2b). This moderate yield is probably due to the NCL step rather than heterogeneity of (*S*,*S*)-**8**-CaM as a result of alternative readthrough, as no full-length CaM was observed on SDS-PAGE gels in the absence of the pyrrolysine analogue (*S*,*S*)-**8**. The identity of the ligation product was confirmed by MALDI-TOF mass spectrometry and tandem mass spectrometry (Figure 2c; see also Figure S2 and Table S1 in the Supporting Information).

One fundamental question is whether Ub*-CaM generated by NCL has the same functional properties as enzymatically prepared Ub-CaM. CaM is known to bind to phosphorylase kinase and increase its activity. The ubiquitination of CaM has been reported to lead to decreased affinity for

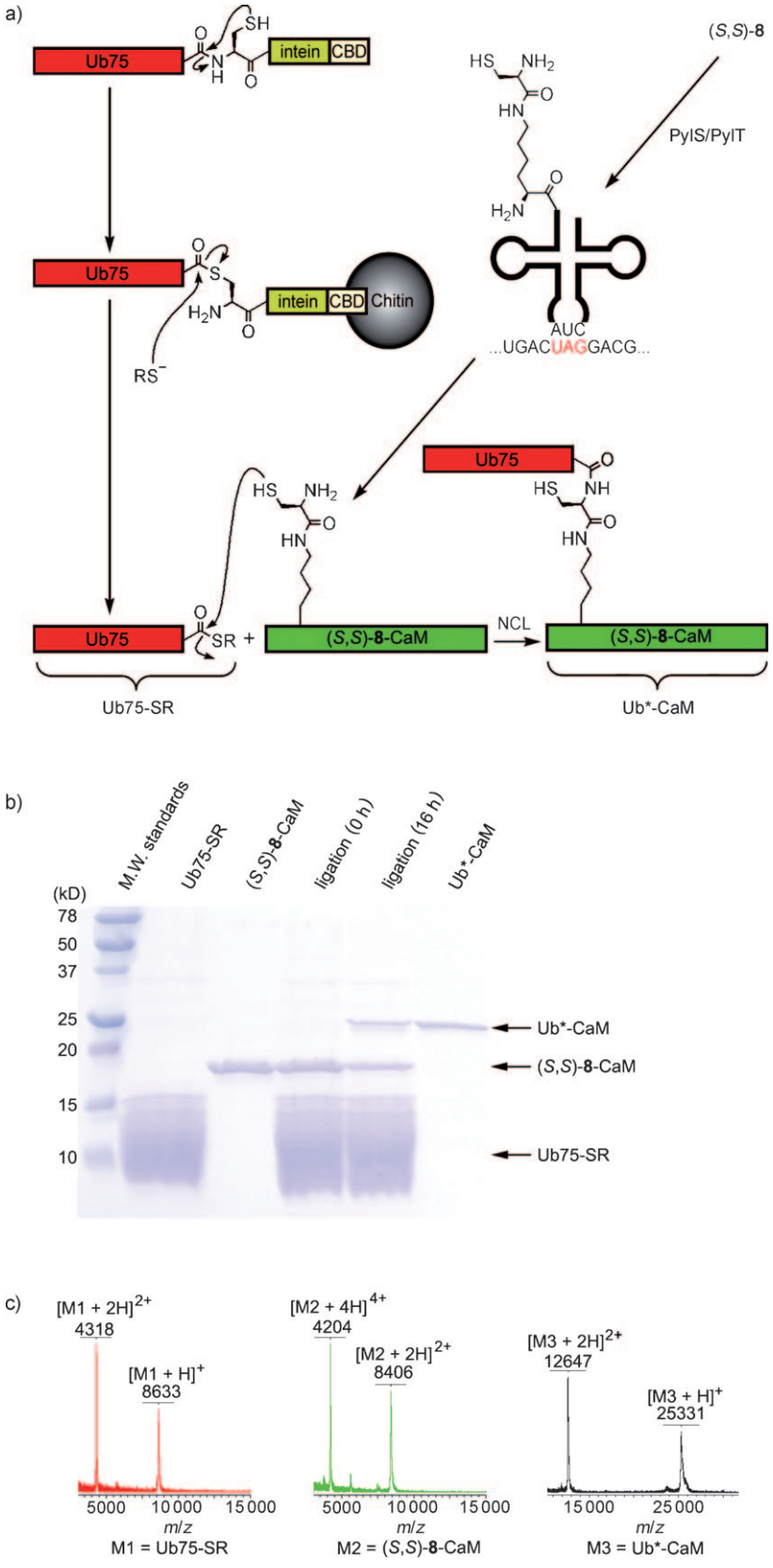


Figure 2. a) Overall reaction scheme for the generation of ubiquitinated CaM through NCL mediated by the pyrrolysine analogue (*S*,*S*)-**8**. b) The process was monitored by SDS-PAGE; the proteins were visualized by staining with Coomassie Blue. c) The identity of the coupling partners and their product was confirmed by MALDI-TOF mass spectrometric analysis.

phosphorylase kinase and a decrease in the degree of activation of the enzyme.[24] To explore this issue, we evaluated the ability of wild-type (WT) CaM and NCL-derived Ub*-CaM to promote phosphorylase kinase activity in a coupled assay system. In our hands, the Ub*-CaM prepared through NCL exhibited a decreased ability to modulate phosphorylase kinase activity when compared to WT CaM. In particular, the degree of activation (ε) by Ub*-CaM was 85% lower than that observed with WT CaM (Figure 3a). This value compares well to the 75% decrease reported for enzymatically generated Ub-CaM.[24]

$$\varepsilon = (\text{activity}^{\text{CaM}} - \text{activity}^{\text{no CaM}})/\text{activity}^{\text{no CaM}} \quad (1)$$

Having demonstrated that Ub*-CaM behaves similarly to Ub-CaM in its effect on phosphorylase kinase activity, we sought to use Ub*-CaM to explore the effect of ubiquitination on the CaM-mediated regulation of other enzymes. One intriguing target was protein phosphatase 2B (PP2B), the only known protein phosphatase regulated by CaM. Such a study had the added appeal that no previous investigations had been performed, and thus it served as an excellent opportunity to explore a fundamental biochemical question. We assayed PP2B activities in the presence and absence of WT CaM or Ub*-CaM. PP2B dephosphorylation was activated by WT CaM as expected; however, surprisingly, ubiquitination had no effect on the ability of CaM to modulate PP2B activity (Figure 3b). The fact that ubiquitination of CaM has a different effect on these kinase and phosphatase systems opens the possibility that ubiquitination can modulate the CaM-mediated linkage between calcium signaling and phosphorylation.

In summary, we have shown that the simple lysine derivatives (*S*,*S*)-**8** and (*R*,*S*)-**8** are incorporated into a protein in response to the UAG codon when the pyrrolysine-incorporation machinery is present. The resulting protein incorporating (*S*,*S*)-**8** was successfully used to introduce ubiquitin through NCL. As this approach does not require the isolation of the key ligases needed for enzymatic ubiquitination, it should enable detailed biochemical studies of many ubiquitinated proteins that it would not otherwise be possible to characterize.

Received: August 10, 2009
Published online: October 30, 2009

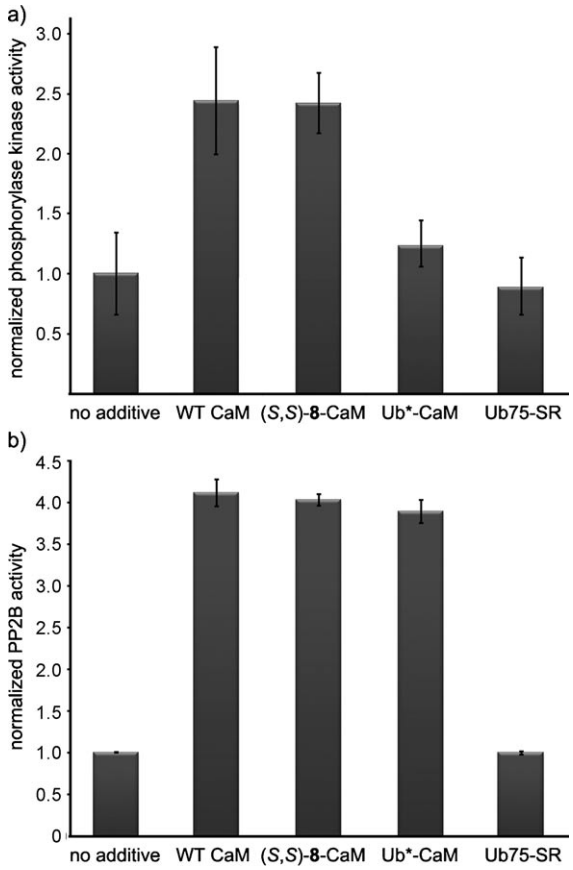


Figure 3. a) Phosphorylase kinase activity with phosphorylase b as the substrate and b) PP2B activity with *p*-nitrophenyl phosphate as the substrate in the presence of WT CaM, (*S*,*S*)-**8**-CaM, Ub*-CaM, or Ub75-SR. The activities are reported as normalized values against the wild-type enzyme in the absence of other activators. The error bars denote the standard deviations, each calculated from three independent measurements.

Keywords: native chemical ligation · protein engineering · protein modifications · pyrrolysine · ubiquitination

[1] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776 – 779.
[2] S. B. Kent, *Chem. Soc. Rev.* **2009**, *38*, 338 – 351.
[3] B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, M. K. Chan, *Science* **2002**, *296*, 1462 – 1466.
[4] G. Srinivasan, C. M. James, J. A. Krzycki, *Science* **2002**, *296*, 1459 – 1462.
[5] T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto, S. Yokoyama, *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818 – 822.
[6] H. Neumann, S. Y. Peak-Chew, J. W. Chin, *Nat. Chem. Biol.* **2008**, *4*, 232 – 234.
[7] T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, S. Yokoyama, *Chem. Biol.* **2008**, *15*, 1187 – 1197.
[8] T. Kobayashi, T. Yanagisawa, K. Sakamoto, S. Yokoyama, *J. Mol. Biol.* **2009**, *385*, 1352 – 1360.
[9] D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters, J. W. Chin, *J. Am. Chem. Soc.* **2009**, *131*, 8720 – 8721.
[10] W. T. Li, A. Mahapatra, D. G. Longstaff, J. Bechtel, G. Zhao, P. T. Kang, M. K. Chan, J. A. Krzycki, *J. Mol. Biol.* **2009**, *385*, 1156 – 1164.
[11] T. Fekner, X. Li, M. M. Lee, M. K. Chan, *Angew. Chem.* **2009**, *121*, 1661 – 1663; *Angew. Chem. Int. Ed.* **2009**, *48*, 1633 – 1635.
[12] X. Li, T. Fekner, M. K. Chan, unpublished results.
[13] N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. Giepmans, A. E. Palmer, R. Y. Tsien, *Nat. Biotechnol.* **2004**, *22*, 1567 – 1572.
[14] C. M. Pickart, M. J. Eddins, *Biochim. Biophys. Acta Mol. Cell Res.* **2004**, *1695*, 55 – 72.
[15] C. M. Pickart, D. Fushman, *Curr. Opin. Chem. Biol.* **2004**, *8*, 610 – 616.
[16] R. L. Welchman, C. Gordon, R. J. Mayer, *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 599 – 609.
[17] D. Mukhopadhyay, H. Riezman, *Science* **2007**, *315*, 201 – 205.
[18] W. Li, Y. Ye, *Cell. Mol. Life Sci.* **2008**, *65*, 2397 – 2406.
[19] C. Chatterjee, R. K. McGinty, J. P. Pellois, T. W. Muir, *Angew. Chem.* **2007**, *119*, 2872 – 2876; *Angew. Chem. Int. Ed.* **2007**, *46*, 2814 – 2818.
[20] R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder, T. W. Muir, *Nature* **2008**, *453*, 812 – 816.
[21] H. P. Jennissen, M. Laub, *Biol. Chem. Hoppe-Seyler* **1988**, *369*, 1325 – 1330.
[22] M. Majetschak, M. Laub, C. Klocke, J. A. Steppuhn, H. P. Jennissen, *Eur. J. Biochem.* **1998**, *255*, 492 – 500.
[23] M. Majetschak, M. Laub, H. E. Meyer, H. P. Jennissen, *Eur. J. Biochem.* **1998**, *255*, 482 – 491.
[24] M. Laub, J. A. Steppuhn, M. Bluggel, D. Immler, H. E. Meyer, H. P. Jennissen, *Eur. J. Biochem.* **1998**, *255*, 422 – 431.
[25] H. P. Jennissen, G. Botzet, M. Majetschak, M. Laub, R. Ziegenhagen, A. Demiroglou, *FEBS Lett.* **1992**, *296*, 51 – 56.
[26] M. Laub, H. P. Jennissen, *Biochim. Biophys. Acta Mol. Cell Res.* **1997**, *1357*, 173 – 191.

