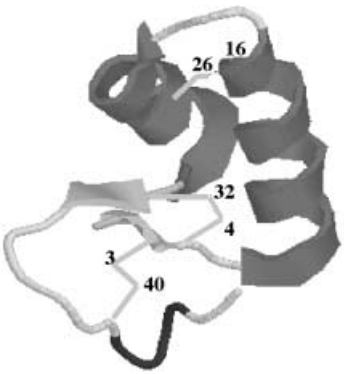


PL-form crambin

Thr-Thr-Cys-Cys-Pro-
Ser-Ile-Val-Ala-Arg-
Ser-Asn-Phe-Asn-Ala-
Cys-Arg-Leu-Pro-Gly-
Thr-Pro-Glu-Ala-Leu-
Cys-Ala-Thr-Tyr-Thr-
Gly-Cys-Ile-Ile-Ile-
Pro-Gly-Ala-Thr-Cys-
Pro-Gly-Asp-Tyr-Ala-
Asn[46]

Figure 1. Molecular structure of crambin (PDB accession no. 1AB1) and target amino acid sequence.

Protein Synthesis

A One-Pot Total Synthesis of Crambin**

*Duhee Bang and Stephen B. H. Kent**

Crambin (Figure 1) is a small protein isolated from the plant *Crambe abyssinica* that has no known biological function.[1,2] It contains three disulfide bonds, an α-helix, a β-sheet, and a reverse turn—all features typical of a globular protein molecule—and for that reason crambin has been studied extensively as a model. Ultrahigh resolution (0.54 Å) structures have been obtained by X-ray crystallography.[3,4] Also, crambin has been used as a model system for novel protein NMR techniques[5] and for computational studies of protein folding.[6,7] Because of difficulties with recombinant DNA expression of crambin in microbial systems,[8] presumably arising from its hydrophobic nature, experimental studies on the crambin molecule have been severely limited: no protein engineering of this interesting molecule has been reported.

Our goal is to use crambin as a model system for experimental studies. We envisioned that a complete control of the chemistry of the crambin molecule would enable us to illuminate fundamental principles of protein folding. Recently, we reported an effective total chemical synthesis of [V15A]crambin.[9] This synthesis used chemical ligation of unprotected peptides in a three-segment strategy to give the full-length 46-residue polypeptide chain that folded efficiently to form native crambin. However, the chemical tactics we used necessitated multiple intermediate purification and lyophilization steps, and was thus arduous and time-consuming. A more effective synthesis with fewer purification steps would facilitate rapid studies of chemical variants of the crambin molecule and would enable site-specific isotopic labeling for solid-state NMR and FT-IR spectroscopy, and other studies.

Here, we report a one-pot synthesis of crambin from three peptide segments that makes use of novel chemical tactics that enable all the ligation steps, and the folding of the full-length polypeptide chain, to be performed without purification or lyophilization of intermediate products. The synthesis makes use of a convergent strategy, with an improved protecting group for the Cys residue at the N-terminal of the middle segment. The resulting synthetic protein is correctly folded, is of exceptional purity, and is obtained rapidly and in high yield.

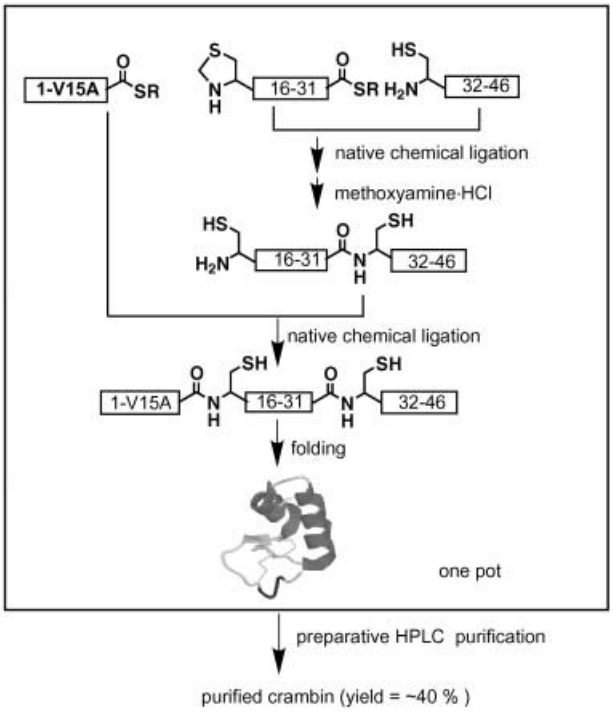
The synthetic strategy for the one-pot synthesis is shown in Scheme 1. In the one-pot synthesis, only a single, final purification step will be necessary for the production of high-purity crambin. In order to make the one-pot synthesis possible, we investigated and optimized the necessary reactions. In a three-segment ligation strategy (Scheme 1), the N-terminal Cys residue of the middle peptide segment must be protected in order to prevent cyclization caused by intramolecular reaction with the thioester moiety.[10,11] In our previous synthesis of crambin we used the acetamidomethyl

[*] D. Bang, Prof. S. B. H. Kent
Institute for Biophysical Dynamics
Department of Biochemistry and Molecular Biology
Department of Chemistry, The University of Chicago
Chicago IL 60637 (USA)
Fax: (+1) 773-702-0439
E-mail: skent@uchicago.edu

[**] We acknowledge support from the NSF MRSEC at the University of Chicago (D.B., grant DMR 0213745). We thank Dr. Neeraj Chopra for helpful discussions.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

  DOI: 10.1002/ange.200353540



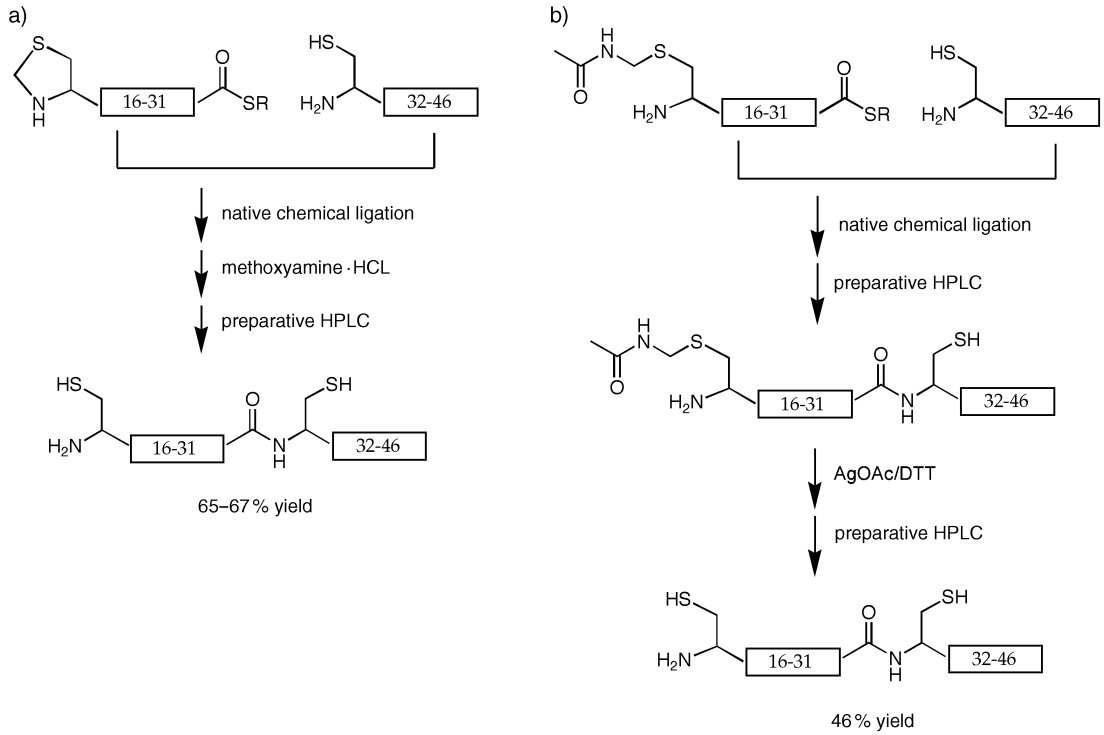
Scheme 1. Synthetic strategy for a one-pot synthesis of crambin by ligation of three unprotected peptide segments.

(Acm) group to protect the side-chain thiol of the N-terminal Cys. However, removal of the S-Acm group has proved to be problematic. For crambin we were not able to obtain a yield of more than 46% for the combined ligation and protection steps (Scheme 2b), even with a highly optimized protocol.[9]

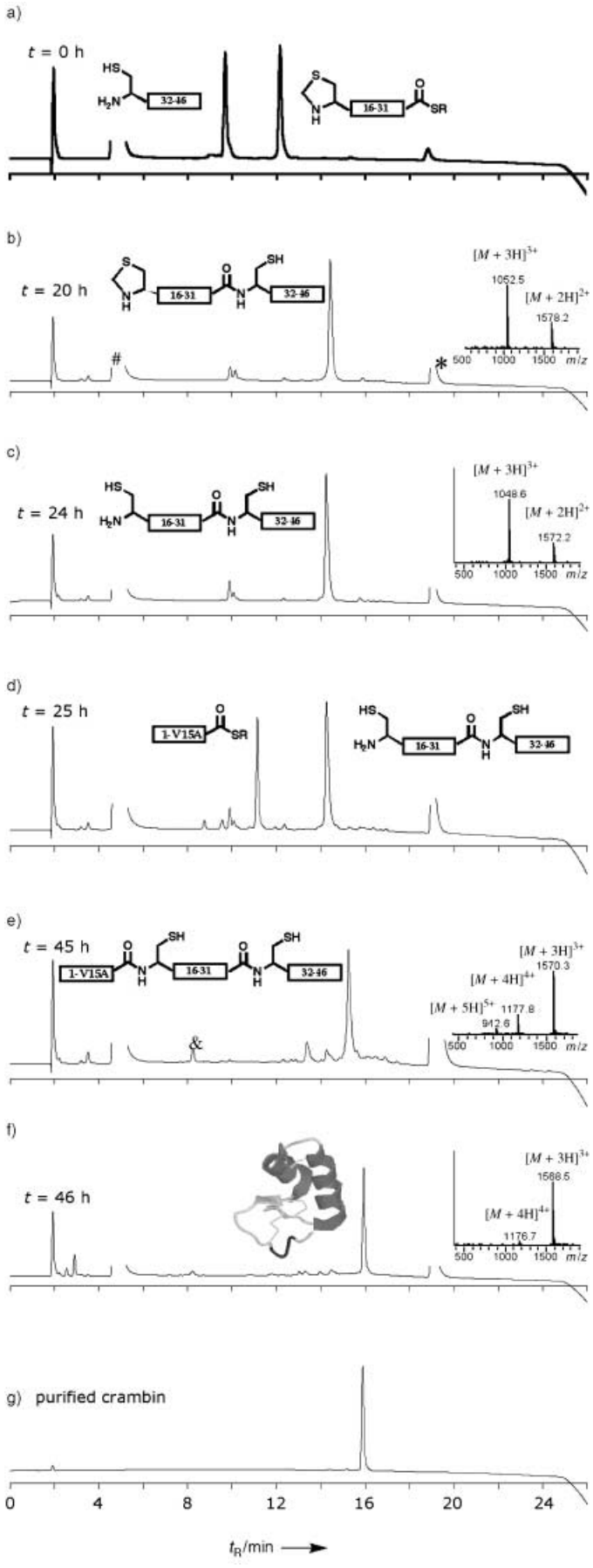
We explored the use of the 1,3-thiazolidine-4-carboxo (Thz) group (Scheme 2a) to protect the N-terminal Cys of the middle peptide segment. Facile conversion of a Thz-peptide to a Cys-peptide,[12] and use of the Thz group to protect an N-terminal cysteine in native chemical ligation[13] have been reported. Optimized stepwise solid-phase synthesis[9] of the segment Cram[Thz^{16}-31]$^{\alpha}$thioester on a 0.4-mmol scale gave a good yield ($\approx$50%, $\approx$350 mg) of the purified middle peptide segment.

The ligation of Cram[Thz^{16}-31]$^{\alpha}$thioester and [Cys^{32}-46]Cram went to completion in 20 h (Figure 2b). Addition of 0.2 M methoxyamine·HCl to the crude ligation mixture reduced the pH to $\approx$4, and reaction for 2 h at ambient temperature resulted in quantitative conversion of the [Thz^{16}-46]Cram to the desired ligation product [Cys^{16}-46]Cram (Figure 2c). In order to evaluate the efficiency of this procedure, in preliminary studies we isolated the ligated Cys-peptide product by preparative HPLC. The yields of the recovered product were 65–67% (two trials). Thus, protection of the N-terminal Cys as a Thz residue was much more efficient, both in terms of recovered yield, manipulations required, and time, than with the commonly used Acm group (Scheme 2).

Next, we explored the possibility of carrying out the second ligation reaction directly on the crude product mixture after treatment with methoxyamine·HCl, that is, without purification of the deprotected peptide [Cys^{16}-46]Cram. The question was whether the thioester moiety would be sensitive



Scheme 2. Comparison of strategies for the protection of an N-terminal Cys residue. a) The strategy based on the thiazolidine resulted in better yield with less manipulation. b) The route based on the Acm group required a preparative-HPLC step before removal of the Acm group and gave a lower overall yield. DTT = dithiothreitol.



to the presence of methoxyamine. Model studies showed that a thioester peptide reacts only very slowly with methoxyamine in pH ≈ 7 buffer to form a small amount of the peptide-$^{\alpha}$(*N*-methoxy)carboxamide that does not take part in the ligation reaction. Thus, we conjectured that the second ligation could be conducted by adding to the crude product mixture a slight excess of Cram[1–15]$^{\alpha}$ thioester peptide over [Cys^{16}-46]Cram. In order to effect the ligation reaction in the same batch, we had to readjust the pH to ≈ 7, because the conversion of Thz to Cys by methoxyamine·HCl had been carried out at pH ≈ 4. Several trials were performed to optimize the pH-adjustment procedure (see Experimental Section). Addition of the third peptide segment Cram[1–15]$^{\alpha}$ thioester (in approximately 10% molar excess) to the solution resulting from the first ligation and Thz conversion (Figure 2d) at pH ≈ 7 resulted in quantitative ligation within 20 h (Figure 2e).

We expected that efficient folding of the product 46-residue polypeptide chain could be performed directly in this crude mixture, simply by adjusting the concentration of guanidinium·HCl (Gn·HCl) and the pH, and by addition of suitable redox reagents (if necessary). Thus, we added two volumes of pH 8.5, 0.1 M Tris buffer containing 8 mM Cys·HCl/1 mM cystine[14] to dilute the Gn·HCl to 2 M and to adjust the pH to ≈ 8, conditions previously shown[9] to be optimal for the folding of crambin and formation of disulfide bonds. As anticipated, folding and disulfide bond formation (Figure 2f) gave an essentially quantitative yield of the correctly folded crambin.

Now that we have optimized each reaction, can we carry out a "one-pot" total synthesis of the crambin molecule? Bringing it all together, we performed a one-pot, three-segment ligation and folding on a tens-of-milligrams scale: the first ligation, in pH 7.5 phosphate buffer containing 6 M Gn·HCl; addition of methoxyamine·HCl to convert Thz- to Cys-peptide at pH 3.8; readjustment of the solution pH to 7.3 and addition of the third segment to effect the second ligation; and, addition of two volumes of Tris buffer to dilute the Gn·HCl, followed by addition of cysteine/cystine redox reactants[14] to fold and form disulfide bonds. The folded crambin molecule was purified directly from the total crude

Figure 2. A one-pot synthesis of crambin. Reactions were monitored by LC-MS. UV profile at $\lambda = 214$ nm and mass data corresponding to major product peaks are shown. The chromatographic separations were performed using a linear gradient (10–50%) of buffer B in buffer A over 20 min (buffer A = 0.1% trifluoroacetic acid (TFA) in water; buffer B = 0.08% TFA in acetonitrile). b) At $t = 20$ h, formation of the first ligation product [Thz^{16}-46]Cram ($t_r = 14.4$ min) was essentially complete. c) At t = 24 h, conversion of Thz- to Cys-peptide ($t_r = 14.2$ min) by methoxyamine·HCl was complete. Note the mass decrease of 12 Da and slightly earlier elution time. d) At $t = 45$ h, formation of the second ligation product Cram[1–46] was essentially complete ($t_r = 15.2$ min, obsd. mass = 4707.7 ± 0.6 Da; calcd. mass = 4708.4 Da.). e) At $t = 46$ h, the folding reaction was essentially complete to give crambin ($t_r = 15.8$ min, observed mass = 4702.7 ± 0.4 Da. calculated mass = 4702.4 Da.). Note that folded crambin eluted later than the linear polypeptide, reflecting the more hydrophobic nature of the protein molecule. # = thiophenol. * = diphenyldisulfide. & = Cram[1-V15A]-NHOCH$_3$.

products by preparative HPLC (Figure 2). The optimized crambin synthesis took 48 h total elapsed time and was successfully reproduced three times at different scales (Table 1). Overall yields from starting peptide segments were 35–45%. The covalent and tertiary structures of the synthetic protein were determined by liquid chromatography (LC)-MS (Figure 2g), 2D-NMR (see the Supporting Information), and X-ray crystallography.[15]

Table 1: Yields from one-pot syntheses of crambin.

Run	Cram[1-V15A]$^{\alpha}$thioester [mg] (mmol)	Cram[Thz^{16}-31]$^{\alpha}$thioester [mg] (mmol)	Cram[32–46] [mg] (mmol)	Purified crambin [mg] (mmol)	Yield [%]
1	50 (0.028)[a]	48 (0.026)	41 (0.026)	42 (0.009)	35
2	28 (0.016)[a]	28 (0.015)	22 (0.015)	26 (0.0055)	37
3	12 (0.007)[a]	11.2 (0.006)	9.5 (0.006)	12.5 (0.0027)	45

[a] Note that ≈10 mol% excess Cram[1-V15A]$^{\alpha}$thioester was required for the second ligation.

This one-pot synthesis of crambin has significant advantages over our previously reported synthesis. Starting from the three purified peptide segments, the previous synthesis took approximately ten days of uninterrupted manipulations, including multiple purifications and lyophilizations to provide purified crambin in ≈25% overall yield.[9] However, the current one-pot synthesis with only one final purification step gave a synthetic crambin product of identical purity in only two days elapsed time, with an improved overall yield of ≈40%. The one-pot synthesis also vividly demonstrated the synthetic utility of the Thz residue as a reversibly protected form of N-terminal Cys.

In summary, for the first time, we have demonstrated the total synthesis of a protein molecule by the native chemical ligation of three unprotected peptide segments, and the folding of the synthetic polypeptide chain, all in the same reaction mixture without purification of intermediates. This novel one-pot synthesis proved to be a substantially more efficient way of producing the crambin molecule. The synthetic work reported here may serve as a useful model for the synthesis of more complex protein targets. Successful application of this one-pot approach to total synthesis depends on near-quantitative ligation reactions to avoid residual reactants interfering with subsequent steps. In our experience, native chemical ligation reactions proceed in essentially quantitative yield, even with only very modest excesses of one reactant. Carry-over of residual amounts of unreacted segments containing Cys thiol groups could potentially interfere with folding of the full-length polypeptide product. However, reaction with thiols in such residual segments would not normally be expected to be part of a productive folding pathway and will be reversed by the excess thiol reactant used in the folding/disulfide-formation step. For these reasons, using the approach described here we anticipate that it will be straightforward to make molecules containing up to 120–150 amino acids (i.e. from three 40- to 50-residue segments). In principle a similar approach could be used to assemble larger proteins from more than three unprotected peptide building blocks, but at some as-yet undefined point increasing amounts of coproducts will interfere.

Experimental Section

Peptides and peptide-$^{\alpha}$thioesters were made manually and stepwise by solid-phase peptide synthesis methods based on "in-situ neutralization" Boc chemistry[16] on -OCH_2-Pam-resins (free $^{\alpha}$carboxyl peptides) or on $HSCH_2CH_2CO$-Leu-OCH_2-Pam-resin[17] ($^{\alpha}$thioester peptides). Detailed procedures for peptide syntheses and for analytical and preparative HPLC were as described.[9] The first native chemical ligation reaction[18,19] was performed in 6 M Gn·HCl, 100 mM sodium phosphate, pH 7.5, 1% thiophenol (v/v), at a concentration of 2 mg mL^{-1} for each peptide. Conversion of Thz- to Cys-peptide was performed by adding 0.2 M methoxyamine·HCl, causing the pH of the reaction mixture to drop to ≈4. Subsequently, 0.5 M NaOH, 6 M Gn·HCl, 100 mM sodium phosphate, 1% thiophenol (v/v) solution was carefully added to the reaction mixture for adjustment to pH 7.0–7.3. The second ligation was performed by adding the peptide Cram[1–15]$^{\alpha}$ thioester to the reaction mixture at room temperature. The folding reaction was performed by adding two volumes of 0.1 M Tris, pH 8.5 buffer containing 8 mM Cys·HCl/1 mM cystine to give a solution that was pH ≈8 and 2 M Gn·HCl.

Received: December 16, 2003 [Z53540]

Keywords: crambin · native chemical ligation · protein engineering · protein folding

[1] M. M. Teeter, J. A. Mazer, J. J. L'Italien, *Biochemistry* **1981**, *20*, 5437–5443.
[2] C. H. VanEtten, H. C. Nielsen, J. E. Peters, *Phytochemistry* **1965**, *4*, 467–473.
[3] C. Jelsch, M. M. Teeter, V. Lamzin, V. Pichon-Pesme, R. H. Blessing, C. Lecomte, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3171–3176.
[4] M. M. Teeter, S. M. Roe, N. H. Heo, *J. Mol. Biol.* **1993**, *230*, 292–311.
[5] Y. Xu, J. Wu, D. Gorenstein, W. Braun, *J. Magn. Reson.* **1999**, *136*, 76–85.
[6] J. Shimada, E. L. Kussell, E. I. Shakhnovich, *J. Mol. Biol.* **2001**, *308*, 79–95.
[7] L. S. Caves, J. D. Evanseck, M. Karplus, *Protein Sci.* **1998**, *7*, 649–666.
[8] L. Lobb, B. Stec, E. K. Kantrowitz, A. Yamano, V. Stojanoff, O. Markman, M. M. Teeter, *Protein Eng.* **1996**, *9*, 1233–1239.
[9] D. Bang, C. Neeraj, S. B. Kent, *J. Am. Chem. Soc.* 2004, 126, 1377–1383.
[10] J. P. Tam, Y. A. Lu, *Protein Sci.* **1998**, *7*, 1583–1592.
[11] Cysteine residues located in *endo* positions of either segment are side-chain-unprotected but do not prevent the correct ligation reaction even if they react with the peptide thioester, since the resulting Cys side-chain thioesters cannot rearrange and are transesterified to the productive N-terminal cysteine thioester.
[12] M. Villain, J. Vizzavona, K. Rose, *Chem. Biol.* **2001**, *8*, 673–679.
[13] M. Villain, J. Vizzavona, H. Gaertner, *Proc. Second Int. Seventeenth Am. Peptide Symp.* **2001**, 107–108.
[14] In the case of crambin, experiments showed that the thiophenol/diphenyldisulfide mixture from the prolonged second ligation reaction forms an effective redox couple that leads to efficient folding and formation of the correct disulfides, even in the absence of added cysteine/cystine. However, it may be that for other globular proteins the more polar, charged cysteine/cystine

redox couple may be essential to prevent trapping of misfolded mixed-disulfide forms.
[15] D. Bang, A. A. Kossiakoff, S. B. Kent, V. Tereshko, unpublished results.
[16] M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. Kent, *Int. J. Pept. Protein Res.* **1992**, *40*, 180 – 193.
[17] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068 – 10073.
[18] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776 – 779.
[19] P. E. Dawson, S. B. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923 – 960.

