DOI: 10.1002/ange.200600702

Kinetically Controlled Ligation for the Convergent Chemical Synthesis of Proteins**

*Duhee Bang, Brad L. Pentelute, and Stephen B. H. Kent**

In the past decade, total chemical synthesis has proved to be a robust and reproducible method for making proteins of defined tertiary structure and full biological activity.[1–5] This success has been made possible by the introduction of chemical ligation methods based on the chemoselective reaction of unprotected peptides under mild conditions in aqueous solution.[6] Our objective is the fully convergent synthesis of proteins from four or more peptide segments using modern native chemical ligation methods. Convergent synthesis is inherently parallel and is thus more efficient.[7] In a convergent approach, the two halves of the target sequence are prepared from multiple peptide segments and condensed in a final step to give the full-length polypeptide chain. In the fully convergent strategy reported herein for the chemical synthesis of a target polypeptide chain, novel synthetic routes are used to make the two halves of the target polypeptide chain. Thus, the left half of the target polypeptide is made by sequential C-to-N ligations in the presence of a potentially reactive thioester moiety at the C terminus, while the right half of the target sequence is made by unprecedented sequential ligations in the N-to-C direction.

Native chemical ligation, the thioester-mediated covalent joining of unprotected peptide segments at a cysteine residue,[8] is the most practical and most widely used chemoselective ligation method. Until now, most synthetic proteins have been prepared from just two peptide segment building blocks. Recent improvements have been focused on the preparation of proteins by sequential ligation of three or four peptide segments.[2,3,5,9,10] For syntheses involving multiple segments, existing strategies for the use of native chemical ligation are suitable only for sequential ligations towards the N terminus from a C-terminal Cys peptide segment. An N-terminal Cys residue can be reversibly protected (PG = protecting group) to give PG-Cys-peptide 1-($^{\alpha}$thioester), so that reaction with Cys-peptide 2 gives only the desired PG-Cys-peptide 1-peptide 2; removal of the protecting group to give

[*] Dr. D. Bang, B. L. Pentelute, Prof. S. B. H. Kent
Department of Chemistry
Institute for Biophysical Dynamics
Center for Integrative Science
The University of Chicago
929 East 57th Street, Chicago, IL 60637 (USA)
Fax: (+1) 773-702-0439
E-mail: skent@uchicago.edu

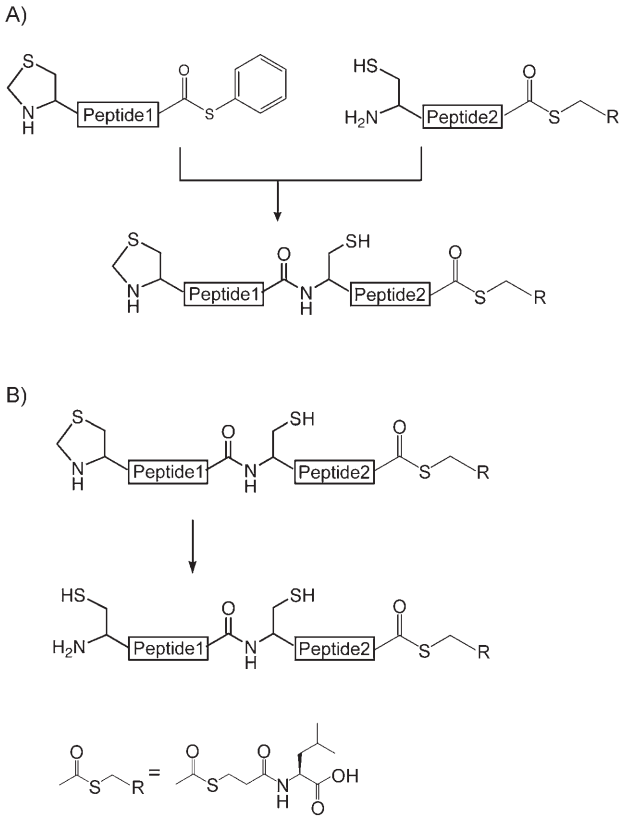
[**] This work was supported by funding from the DOE GTL program (grant DE-FG02-04ER63786 to S.B.K).

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

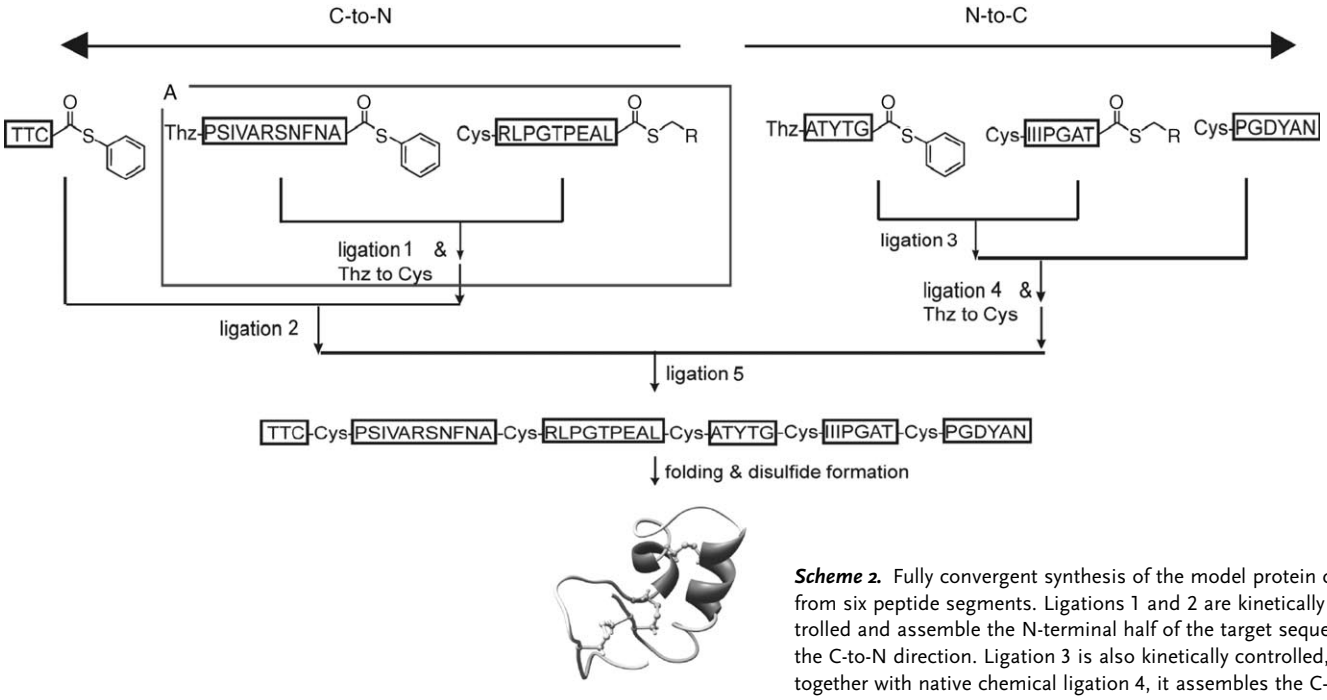
Cys-peptide 1-peptide 2 enables the synthesis to be continued by sequential ligation towards the N terminus of the target polypeptide chain.[3] No corresponding set of chemical tactics currently exists for protecting the thioester moiety of a peptide-(α-thioester), thus preventing not only N-to-C sequential assembly of the peptide segments by native chemical ligation but also fully convergent protein synthesis using this chemistry.[11]

Herein, we report an approach based on the novel principle of "kinetically controlled ligation" that enables a highly practical, fully convergent strategy for the synthesis of a target protein. The major challenge was to control the intrinsic dual reactivity of a bifunctional Cys-peptide 2-(α-thioester) under native chemical ligation reaction conditions so that it would react with a peptide 1-(α-thioester) to yield only a single product (see Scheme 1 A).[12] Like most researchers who use native chemical ligation, until now we have routinely prepared relatively unreactive peptide-(α-$COSCH_2CH_2CO$)-Leu (i.e. alkyl) thioesters.[13] Typically, native chemical ligation of such a peptide-(α-thioalkylester) and a Cys-peptide has been carried out in the presence of thiophenol as a catalyst; the reactive species is assumed to be the peptide-(α-thiophenylester).[14] We conjectured that a preformed peptide-(α-thiophenylester) would react with a Cys-peptide so much faster than a standard peptide-(α-thioalkylester) that, in the same solution under competitive reaction conditions in the absence of exogenous thiophenol, this large rate difference (see Supporting Information) would make the standard peptide-(α-thioalkylester) effectively unreactive. This turned out to be the case.

To present concrete examples of the chemical challenges involved in the fully convergent synthesis of a protein from



Scheme 1. Key steps in a convergent synthesis using native chemical ligation. A) Kinetically controlled reaction of a peptide-(α-thiophenylester) with a Cys-peptide-(α-thioalkylester) to give a single ligation product. B) The conversion of a Thz-peptide-(α-thioalkylester) into a Cys-peptide-(α-thioalkylester), without damage to the thioalkylester moiety.



Scheme 2. Fully convergent synthesis of the model protein crambin from six peptide segments. Ligations 1 and 2 are kinetically controlled and assemble the N-terminal half of the target sequence in the C-to-N direction. Ligation 3 is also kinetically controlled, and together with native chemical ligation 4, it assembles the C-terminal half of the molecule in the N-to-C direction. Ligation 5 is the final native chemical ligation of the two halves of the target molecule 46-residue polypeptide, which is folded to give the protein crambin.

multiple peptide segments and to illustrate how we were able to effectively address those challenges, we undertook a synthesis of the model protein crambin from six peptide segments.[3,15–17] The sequence of the target molecule and the synthetic design are shown in Scheme 2.

The efficacy of our kinetic-control strategy for determining the outcome of a ligation reaction is shown in ligation 1 in the crambin synthesis (see Scheme 2, box A), with the reaction of Thz-Pro-Ser-Ile-Val-Ala-Arg-Ser-Asn-Phe-Asn-Ala-(α-thiophenylester) with Cys-Arg-Leu-Pro-Gly-Thr-Pro-Glu-Ala-Leu-(α-thioalkylester) in aqueous buffer (pH 6.8) in the absence of added thiophenol (see Figure 1 A and B). The ligation reaction was complete in one hour and gave the desired product Thz-Pro-Ser-Ile-Val-Ala-Arg-Ser-Asn-Phe-Asn-Ala-Cys-Arg-Leu-Pro-Gly-Thr-Pro-Glu-Ala-Leu-(α-thioalkylester) in near-quantitative yield. Only trace amounts of by-products from undesired reaction of the alkyl

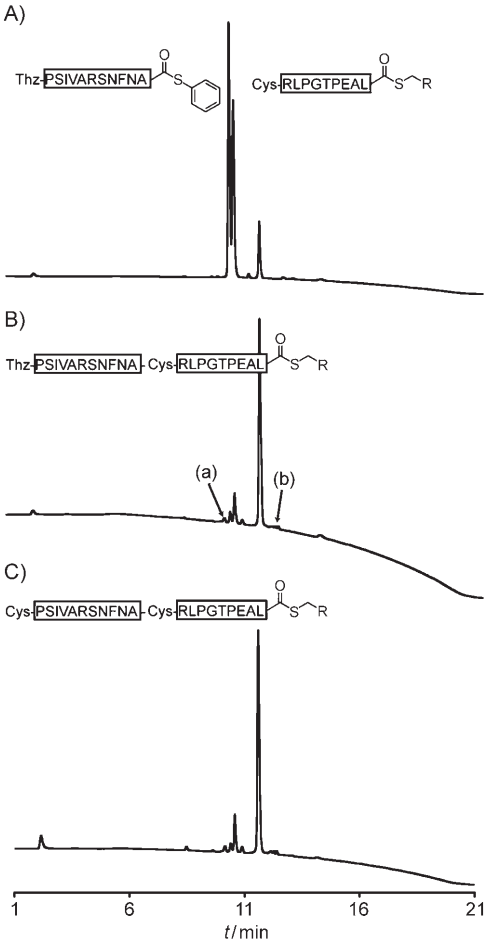


Figure 1. Analytical HPLC traces of the kinetically controlled ligation of [Thz4–Ala15]-(α-thiophenylester) and [Cys16–Leu25]-(α-thioalkylester). A) $t < 1$ min, B) $t = 1$ h. The ligation gave almost exclusively the [Thz4–Leu25]-(α-thioester). In part (B), arrow (a) indicates cyclized by-product (cyclic (Cys16–Leu25); the two peaks eluting just after this by-product are residual reactants), and arrow (b) indicates the by-product [Thz4–Leu25]–[Cys16–Leu25]-(α-thioalkylester). C) Conversion of [Thz4–Leu25]-(α-thioalkylester) into [Cys4–Leu25]-(α-thioalkylester) using 0.2 M methoxyamine hydrochloride added directly to the ligation reaction mixture. The reaction was complete in two hours.

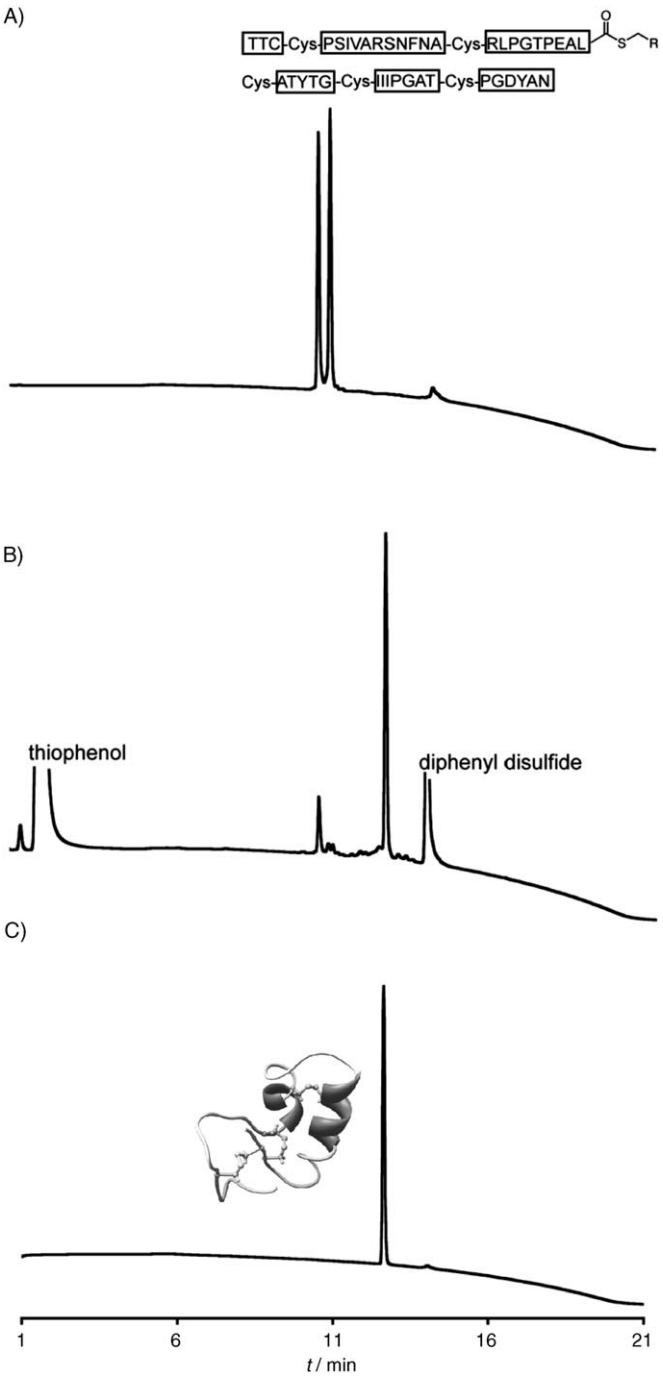


Figure 2. Analytical HPLC traces corresponding to the final step in the fully convergent synthesis of crambin, namely the native chemical ligation of the two halves of the target polypeptide chain [Thr1–Leu25]-(α-thioalkylester) and [Cys26–Asn46]. Trace (A) shows the starting point (i.e. before the addition of thiophenol). Overnight reaction under standard native chemical ligation conditions gave the full-length crambin polypeptide, which was then folded without purification. Trace (B) shows the total crude products from folding and disulfide formation of the unpurified full-length polypeptide ligation product obtained upon dilution to 1 M guanidinium hydrochloride in the presence of 8 mM Cys and 1 mM cystine. The correctly folded crambin molecule was formed in near-quantitative yield in 1 h. Trace (C) shows the folded crambin molecule after purification by preparative HPLC (observed mass: 4702.0 ± 0.8 Da; calculated mass: 4702.4 Da, as expected for three disulfides). Reverse-phase HPLC analyses were performed on a Vydac C4 column using a linear gradient (1–61%) of buffer B in buffer A over 15 min (buffer A: 0.1% trifluoroacetic acid (TFA) in water; buffer B: 0.08% TFA in acetonitrile). Disulfide bridges were not determined.

thioester were formed (see Figure 1B): cyclized by-product (<1.5%) and oligomer (≪1%). Similar results were obtained for the two other kinetically controlled ligations in the crambin synthesis (Scheme 2, ligations 2 and 3; see also Supporting Information).

The other key chemical transformation for the use of native chemical ligation in a fully convergent synthesis is to be able to convert a Thz-peptide-(α-thioalkylester) into a Cys-peptide-(α-thioalkylester) (Scheme 1B). This transformation was accomplished by treatment with 0.2 M methoxyamine hydrochloride. The thioester moiety was unaffected during the quantitative conversion from the Thz-peptide-(α-thioalkylester) into the Cys-peptide-(α-thioalkylester) (compare the products shown in Figure 1C with the starting mixture in Figure 1B).

The last step in the convergent synthesis was the native chemical ligation of the two halves of the polypeptide chain, followed by folding and disulfide formation, which gave the synthetic crambin protein in good yield (see Figure 2).

Several aspects of the kinetically controlled ligation reactions used in the convergent synthesis of crambin are worthy of comment. Ligation 2 (Scheme 2) involves reaction of a peptide-thiophenylester that contains an internal Cys residue with a free thiol group. In this instance, the presence of a free thiol did not interfere with the formation of the desired product (see Supporting Information); however, the ligation product underwent internal cyclization with the C-terminal thioester to form a thiolactone, which could be readily converted into a thioalkyl ester (see Supporting Information). The nature of the C-terminal amino acid residue in a peptide-(α-thioester) affects the relative reactivity of that peptide in native chemical ligation: hindered C-terminal amino acids such as Ile, Thr, or Val are much less reactive than, for example, a C-terminal Ala-(α-thioester). This needs to be taken into account in the design of a convergent synthesis and can be used to advantage to help control the outcome of the reaction as we have done in the example presented herein. We have also successfully used kinetically controlled ligation to join two peptides, both of which have a C-terminal alanine (i.e. ligation of a peptide-Ala-(α-thiophenylester) to a Cys-peptide-Ala-(α-thioalkylester)), with results that are comparable to those shown here (data not shown).

The new tactics and chemistries described herein are a major advance in the chemical synthesis of proteins. We are successfully using fully convergent synthesis based on kinetically controlled ligation for an increasing number of protein targets, including erythropoietin, the HIV-1 protease, and a KChIP accessory protein. We believe that kinetically controlled ligation will form the basis for a truly practical convergent chemical synthesis of proteins.[18]

Received: February 23, 2006
Published online: April 26, 2006

[1] P. E. Dawson, S. B. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
[2] G. G. Kochendoerfer, S. Y. Chen, F. Mao, S. Cressman, S. Traviglia, H. Shao, C. L. Hunter, D. W. Low, E. N. Cagle, M. Carnevali, V. Gueriguian, P. J. Keogh, H. Porter, S. M. Stratton, M. C. Wiedeke, J. Wilken, J. Tang, J. J. Levy, L. P. Miranda, M. M. Crnogorac, S. Kalbag, P. Botti, J. Schindler-Horvat, L. Savatski, J. W. Adamson, A. Kung, S. B. Kent, J. A. Bradburne, *Science* **2003**, *299*, 884–887.
[3] D. Bang, S. B. Kent, *Angew. Chem.* **2004**, *116*, 2588–2592; *Angew. Chem. Int. Ed.* **2004**, *43*, 2534–2538.
[4] D. Bang, G. I. Makhatadze, V. Tereshko, A. A. Kossiakoff, S. B. Kent, *Angew. Chem.* **2005**, *117*, 3920–3924; *Angew. Chem. Int. Ed.* **2005**, *44*, 3852–3856.
[5] Z. Wu, J. Alexandratos, B. Ericksen, J. Lubkowski, R. C. Gallo, W. Lu, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11587–11592.
[6] M. Schnölzer, S. B. Kent, *Science* **1992**, *256*, 221–225.
[7] J. B. Hendrickson, *J. Am. Chem. Soc.* **1977**, *99*, 5439–5450.
[8] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.
[9] D. Bang, S. B. Kent, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5014–5019.
[10] L. E. Canne, P. Botti, R. J. Simon, Y. Chen, E. A. Dennis, S. B. Kent, *J. Am. Chem. Soc.* **1999**, *121*, 8720–8727.
[11] G. Casi, D. Hilvert, *Curr. Opin. Struct. Biol.* **2003**, *13*, 589–594.
[12] Under native chemical ligation reaction conditions, the Cys-peptide 2-(α-thioester) will itself undergo multiple reactions to form cyclic (Cys-peptide 2), Cys-peptide 2-Cys-peptide 2-(α-thioester), and other oligomerization and cyclization products. See: a) L. Zhang, J. P. Tam, *J. Am. Chem. Soc.* **1997**, *119*, 2363–2370; b) T. C. Evans, J. Benner, M.-Q. Xu, *J. Biol. Chem.* **1999**, *274*, 18359–18363.
[13] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073.
[14] P. E. Dawson, M. J. Churchill, M. R. Ghadiri, S. B. H. Kent, *J. Am. Chem. Soc.* **1997**, *119*, 4325–4329.
[15] M. M. Teeter, J. A. Mazer, J. J. L'Italien, *Biochemistry* **1981**, *20*, 5437–5443.
[16] C. Jelsch, M. M. Teeter, V. Lamzin, V. Pichon-Pesme, R. H. Blessing, C. Lecomte, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3171–3176.
[17] D. Bang, N. Chopra, S. B. Kent, *J. Am. Chem. Soc.* **2004**, *126*, 1377–1383.
[18] H. J. Musiol, S. Dong, M. Kaiser, R. Bausinger, A. Zumbusch, U. Bertsch, L. Moroder, *ChemBioChem* **2005**, *6*, 625–628.

Keywords: kinetics · native chemical ligation · peptides · protein synthesis

