Biosynthesis

DOI: 10.1002/ange.200503882

Biosynthesis of the Cyclotide Kalata B1 by Using Protein Splicing**

*Richard H. Kimura, Ahn-Tuyet Tran, and Julio A. Camarero**

Cyclotides are a newly emerging family of large-backbone cyclized polypeptides (≈30-residues long) characterized by a disulfide-stabilized core (3 disulfide bonds) with an unusual knotted structure.[1]

The core structural motif in cyclotides has been termed a cyclic cystine knot (CCK) and is characterized by a cystine knot that is embedded into a circular backbone topology.[1d] The cystine knot involves two disulfide bonds that form a ring that is penetrated by a third disulfide bond. The unique cyclic-backbone topology and knotted arrangement of the three disulfide bonds endow the cyclotides with exceptional stability and resistance to chemical, enzymatic, and thermal degradation.[2] Furthermore, their well-defined structures have been associated with a range of biological functions such as uterotonic activity, inhibition of trypsin and neurotension binding, cytotoxicity, anti-HIV, antimicrobial, and insecticidal activity.[1b–c] Together, these characteristics suggest that cyclotides are ideal molecular scaffolds for the development of stable peptide drugs.[1b]

Despite the fact that the chemical synthesis of circular peptides has been well explored and a number of different approaches involving the solid or liquid phases exist,[3] recent developments in the fields of molecular biology and protein engineering have now made possible the biosynthesis of cyclic peptides. This progress has been made mainly in two areas, nonribosomal peptide synthesis[4] and expressed protein ligation (EPL)/protein trans-splicing.[5] Access to biosynthetic cyclotides by using techniques of recombinant-DNA expression offers the exciting possibility of producing large combinatorial libraries of highly stable cyclic polypeptides. This would allow the generation of cell-based combinatorial libraries that could be screened either in vitro or in vivo for their ability to regulate cellular processes.

Herein, we describe the biosynthesis of the cyclotide Kalata B1 (KB1) by using an engineered intein. Our approach (Figure 1) is based on an intramolecular version of native chemical ligation (NCL).[6] NCL involves the chemoselective reaction between an N-terminal Cys residue of one peptide and an α-thioester group of a second peptide. Importantly, incorporation of these two groups into the same synthetic polypeptide leads to efficient circularization.[3]

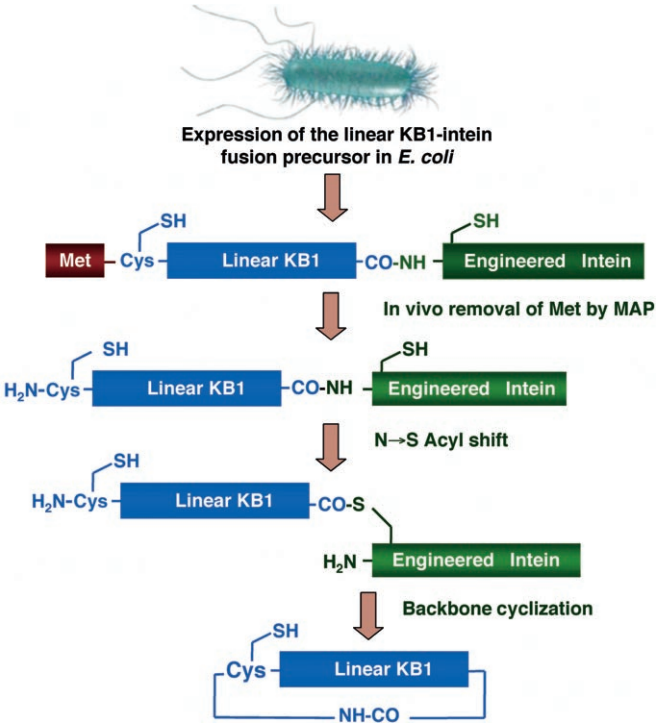

Figure 1. Backbone cyclization of a KB1 linear precursor by using an engineered intein.

To test this approach, we constructed several plasmids that encode different KB1 linear precursors (Figure 2) and

[*] Dr. R. H. Kimura, Dr. A.-T. Tran, Dr. J. A. Camarero
Chemical Biology and Nuclear Sciences Division
Lawrence Livermore National Laboratory
University of California
7000 East Avenue, Livermore, CA 94550 (USA)
Fax: (+1) 925-422-3160
E-mail: camarero1@llnl.gov

[**] This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48. We thank Dr. David Craik, University of Queensland, Australia, for kindly providing a sample of natural Kalata B1.

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

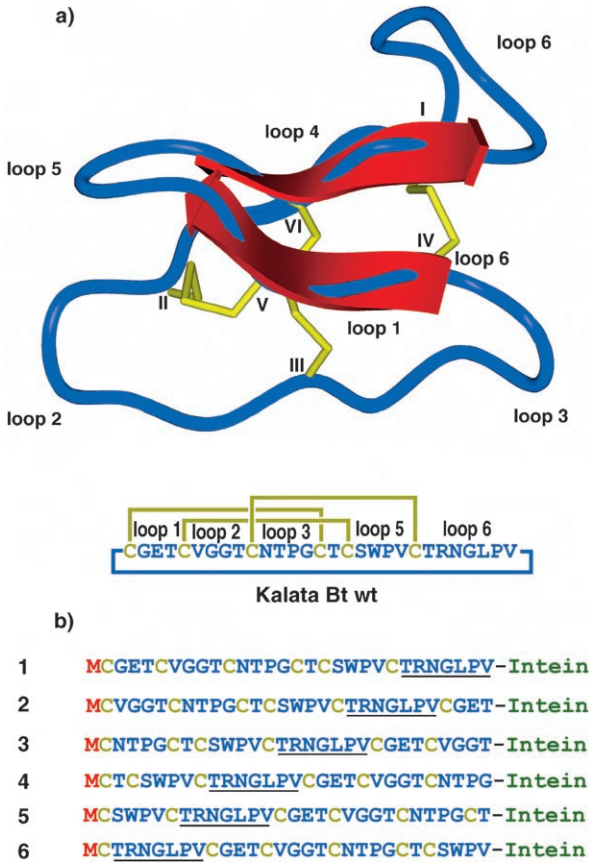

Figure 2. a) Structure (primary and tertiary) and disulfide connectivity of the cyclotide KB1. b) Sequences of the different linear precursors used for the backbone cyclization of KB1. The sequence corresponding to loop 6 has been underlined in all the linear precursors as a reference.

are fused in-frame at their C termini to a modified vacuolar membrane ATPase (VMA) intein.[5e] This allows the generation of the required C-terminal α-thioester function. Also, a Met residue was genetically introduced at the N terminus of the corresponding KB1 fusion protein. The Met residue is efficiently removed immediately after translation by the endogeneous Met aminopeptidase (MAP). This in vivo proteolytic event unmasks the N-terminal Cys residue required for NCL.[5a]

The KB1–intein fusion proteins (**1** through **6**, Figure 2b) were expressed in *E. coli* and purified by affinity chromatography. Analysis of the purified fusion proteins by SDS-PAGE revealed that the different linear precursors have different propensities for cleavage in vivo. Linear precursors **4** and **5** showed around 30% cleavage in vivo, **6** was almost completely cleaved (90%), and **1**, **2**, and **3** were cleaved to about 70% in vivo. Analysis of the soluble cell fractions that correspond to the constructs with high in vivo cleavage did not reveal the presence of any reduced or oxidized circular KB1. Analysis of the insoluble cellular fractions under reductive conditions, after being treated briefly at pH 12 to hydrolyze any residual thioester functionality, showed, on the other hand, the presence of the corresponding reduced linear carboxylated peptide. It is interesting to note that the linear precursors **1** and **6** also produced small amounts of reduced circular KB1 ($\leq 5\ \mu g\,L^{-1}$ of culture).

These results indicate that cleavage in vivo of KB1–intein fusion proteins is mostly triggered by internal Cys residues (rather than by the N-terminal Cys residues), which leads directly to backbone cyclization. These thiolactone intermediates have been shown to react by a "thia zip" mechanism that involves successive ring expansions through reversible thiol–thiolactone exchanges in the direction of the N terminus. The exchange continues until the thiolactone involving the N-terminal Cys residue is formed finally resulting in backbone cyclization.[7] It is very likely that some of these intermediates may aggregate before backbone cyclization can take place, thus explaining the low yield observed for the cyclization process in the bacterial cytoplasm.

We next tested the ability of the different linear precursors to cyclize in vitro. Cyclization was triggered by treatment of the purified KB1–intein fusion proteins with EtSH (3%) in column buffer solution (0.1 mM ethylenediaminetetraacetic acid, 50 mM sodium phosphate, 250 mM NaCl) at pH 7.2. In all the cases except **5**, the major product of the reaction was the reduced cyclic KB1. Among the different linear precursors, **1**, **3**, and **4** gave the best cyclization yields (Figure 3a). Precursor **6** also had a moderately good yield for the cyclization process. However, the fact that almost all of the precursor fusion protein was prematurely cleaved in vivo resulted in a small amount of circular KB1 being produced. Linear precursor **2** gave only a modest yield and although the reduced circular KB1 was still the major product, other minor products including the corresponding linear ethyl thioester precursor as well as thiolactone intermediates were also found in the cyclization crude mixture, even after 4 days of reaction. By far the worst result was obtained for the linear precursor **5**. In this case, cyclic reduced KB1 was found only as a minor product ($\leq 5\%$). Purified reduced KB1 was oxidatively folded in the presence of reduced glutathione (10 mM; GSH) in NH_4HCO_3 buffer solution (50 mM; pH 8.5) that contained *i*PrOH (50%) as described previously by Craik and co-workers[3c] Under these conditions, the folding of reduced KB1 is almost quantitative in less than 15 h (Figure 3b). The folding of reduced KB1 in the absence of organic cosolvents is relatively inefficient, being around 40% of that obtained by using *i*PrOH (50%). The use of Triton X-100 (0.5%) also improved the yield of folding (70% of that observed using *i*PrOH). The folded recombinant KB1 was shown by 2D NMR spectroscopy, HPLC, and ES-MS analysis to be identical to the natural product.

Encouraged by these results, we decided to explore the possibility of carrying out the cyclization and folding steps in a single-pot reaction. This was accomplished by treating **4** with different amounts and ratios of reduced and oxidized glutathione (GSH and GSSG, respectively). The best cleavage/cyclization conditions were accomplished by using GSH (100 mM; pH 7.2). Through the use of these conditions, cleavage of the linear intein precursor was almost quantitative in less than 18 h. Optimal conditions for oxidative folding of reduced KB1 were obtained by using a buffer solution that contained a GSH/GSSG ratio of 4:1 at pH 7.2 in *i*PrOH (50%). Hence, the purified linear precursor **4** was treated first

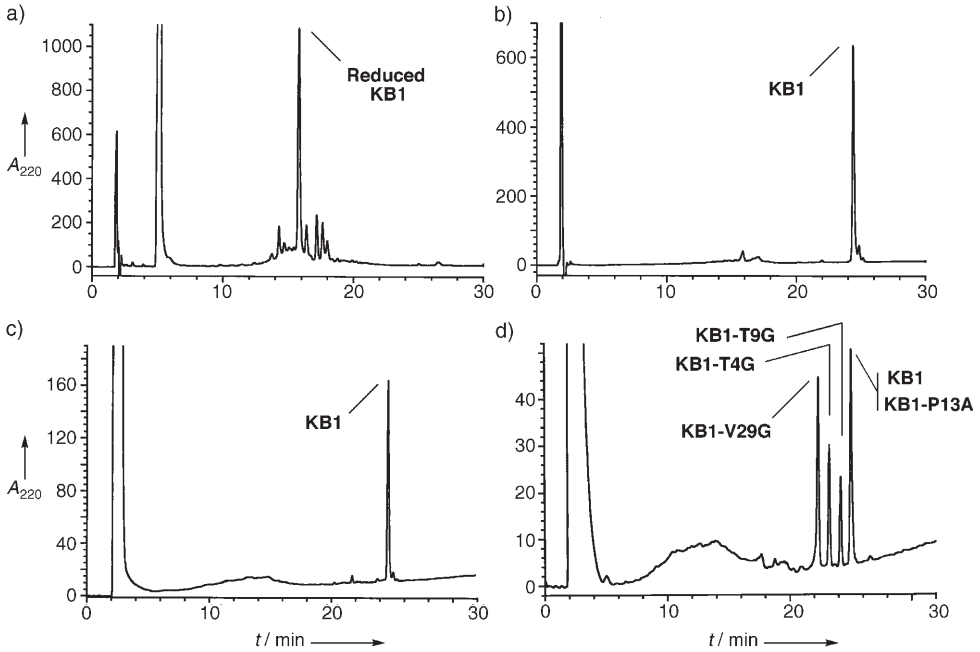

Figure 3. Cyclization and folding of native KB1. Analytical HPLC traces of: a) EtSH cleavage of purified linear precursor **4** after 2 days, b) oxidative refolding of reduced KB1 in GSH (10 mM), NH_4HCO_3 buffer (50 mM; pH 8.5) that contains *i*PrOH (50%) after 18 h, c) "one-pot" cyclization and folding of linear precursor **4** by using GSH/GSSG in *i*PrOH (50%; see text for details) and d) "one-pot" cyclization and folding of small libraries of KB1-based cyclotides by using the same procedure as described previously.

with GSH (100 mM; pH 7.2) for 18 h, and the crude reaction was then complemented with *i*PrOH (50% v/v) and GSSG (12 mM). Under these conditions, the folding reaction was extremely clean and efficient after 20 h (Figure 3c). When the oxidative cleavage was done in aqueous buffer at pH 7.2 in the presence of GSH (10 mM) and GSSG (2 mM) with no *i*PrOH, folded KB1 was also obtained with a yield of about 4% (based on the amount of linear precursor). Similar results were also obtained with **1**, **2**, **3**, and **6**.

We also explored the possibility of using this method to generate different KB1 mutants (KB1-V29G, KB1-T4G, KB1-T9G, KB1-P13A, KB1-W19A, KB1-W19R, and KB1-W19S).[8] Some of the mutants were designed to test the effect of replacement of different β-branched residues adjacent to Cys residues in loops 1, 2, and 6 by the highly flexible Gly residue (KB1-V29G, KB1-T4G, and KB1-T9G). Also, the replacement of the hydrophobic residue Trp in loop 5 by the more hydrophilic Arg and Ser residues (KB1-W19R and KB1-W19S) and the less hydrophobic Ala residue (KB1-W19A) was explored. Finally, the Pro residue in loop 3, which has been found to be important for defining the β turn in this loop,[9] was mutated to Ala (KB1-P13A).

The "one-pot" cyclization/folding reaction was very efficient for all the mutants tested with yields ranging from 10→60%, this therefore highlights the robustness of this structural fold for the construction of molecular libraries. The possibility of using this methodology to biosynthesize libraries based on the cyclotide KB1 was also explored. A small library that contained KB1 and the mutants KB1-V1G, KB1-T5G, KB1-T10G, and KB1-P14A were expressed in *E. coli* and then cyclized and folded as described above. The HPLC and ES-MS analysis of the single-pot cyclization/folding reaction revealed the presence of all of the five members of the library (Figure 3d). Quantification of the individual members by ES-MS and HPLC showed that all the cyclotides were produced with similar yields. Standard deviation for the yields of all the five members was within 30%.

In summary, we have described for the first time the biosynthesis of the cyclotide KB1 in *E. coli*. We have also shown that our biosynthetic approach very efficiently generates, in one single-pot reaction, natural KB1 as well as several mutants. This approach can also be used to generate cyclotide-based libraries that could be screened in vitro for biological activity. Furthermore, one can easily envision that by using cellular environments that are less reductive than *E. coli*'s cytoplasm, this biosynthetic method could be adapted for in vivo biosynthesis of cyclotides.

Received: November 3, 2005
Published online: December 28, 2005

Keywords: bioorganic chemistry · biosynthesis · circular peptides · protein engineering · protein ligation

[1] a) M. Trabi, D. J. Craik, *Trends Biochem. Sci.* **2002**, *27*, 132–138; b) D. J. Craik, N. L. Daly, J. Mulvenna, M. R. Plan, M. Trabi, *Curr. Protein Pept. Sci.* **2004**, *5*, 297–315; c) U. Goransson, E. Svangard, P. Claeson, L. Bohlin, *Curr. Protein Pept. Sci.* **2004**, *5*, 317–329; d) H. J. Vogel, D. I. Chan, *Structure* **2005**, *13*, 688–690.
[2] M. L. Colgrave, D. J. Craik, *Biochemistry* **2004**, *43*, 5965–5975.
[3] a) J. A. Camarero, J. Pavel, T. W. Muir, *Angew. Chem.* **1998**, *110*, 361–364; *Angew. Chem. Int. Ed.* **1998**, *37*, 347–349; b) J. P. Tam, Y. A. Lu, J. L. Yang, K. W. Chiu, *Proc. Natl. Acad. Sci. USA* **1999**,

96, 8913 – 8918; c) N. L. Daly, S. Love, P. F. Alewood, D. J. Craik, *Biochemistry* **1999**, *38*, 10606 – 10614.
[4] C. T. Walsh, *Science* **2004**, *303*, 1805 – 1810.
[5] a) J. A. Camarero, T. W. Muir, *J. Am. Chem. Soc.* **1999**, *121*, 5597 – 5598; b) T. C. Evans, J. Benner, M.-Q. Xu, *J. Biol. Chem.* **1999**, *274*, 18359 – 18381; c) J. A. Camarero, D. Fushman, D. Cowburn, T. W. Muir, *Bioorg. Med. Chem.* **2001**, *9*, 2479 – 2484; d) C. P. Scott, E. Abel-Santos, M. Wall, D. Wahnon, D. , S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13638 – 13643; e) C. J. Noren, J. Wang, F. B. Perler, *Angew. Chem.* **2000**, *112*, 458 – 476; *Angew. Chem. Int. Ed.* **2000**, *39*, 450 – 456; f) F. B. Perler, E. Adam, *Curr. Opin. Biotechnol.* **2000**, *11*, 377 – 383; g) T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 431 – 434.
[6] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776 – 779; b) J. P. Tam, Y. A. Lu, C. F. Liu, J. Shao, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12485 – 12489; c) P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923 – 960.
[7] J. P. Tam, Y. A. Lu, *Protein Sci.* **1998**, **7**, 1583 – 1592.
[8] Numbering of the mutations is always with reference to construct **1**.
[9] K. J. Rosengren, N. L. Daly, M. R. Plan, C. Waine, D. J. Craik, *J. Biol. Chem.* **2003**, *278*, 8606 – 8616.

