

UCRL-JRNL-230072



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Biosynthesis of a Fully Functional Cyclotide inside Living Bacterial Cells

J. A. Camarero, R. H. Kimura, Y.-H. Woo, J. Cantor, A. Shekhtman

April 19, 2007

ChemBiochem

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Biosynthesis of a Fully Functional Cyclotide inside Living Bacterial Cells

Julio A. Camarero^{1*}, Richard H. Kimura¹, Youn-hi Woo¹, Jason Cantor¹ and Alexander Shekhtman²

¹*Biosciences and Biotechnology Division, Livermore National Laboratory, University of California, 7000 East Avenue, L-231, Livermore, CA 94550,* ²*Department of Chemistry, State University of New York, Albany, New York 12222, USA.*

*To whom correspondence should be addressed:

Julio A. Camarero

Lawrence Livermore National Laboratory, University of California, 7000 East Avenue, L-231, Livermore, CA 94550.

Phone: (925) 422 6807

Fax: (925) 422 3160

E-mail: camarero1@llnl.gov

Running Title: In vivo biosynthesis of a folded cyclotide

Abstract

The cyclotide MCoTI-II is a powerful trypsin inhibitor recently isolated from the seeds of *Momordica cochinchinensis*, a plant member of *cucurbitaceae* family. We report for the first time the *in vivo* biosynthesis of natively-folded MCoTI-II inside live *E. coli* cells. Our biomimetic approach involves the intracellular backbone cyclization of a linear cyclotide-intein fusion precursor mediated by a modified protein splicing domain. The cyclized peptide then spontaneously folds into its native conformation. The use of genetically engineered *E. coli* cells containing mutations in the glutathione and thioredoxin reductase genes considerably improves the production of folded MCoTI-II *in vivo*. Biochemical and structural characterization of the recombinant MCoTI-II confirmed its identity. Biosynthetic access to correctly-folded cyclotides allows the possibility of generating cell-based combinatorial libraries that can be screened inside living cells for their ability to modulate or inhibit cellular processes.

Keywords: Cyclotides, trypsin inhibitor, backbone cyclization, protein splicing, protein engineering, cyclic protein.

Cyclotides are a new emerging family of large plant-derived backbone-cyclized polypeptides ($\approx 28-37$ amino acids long) that share a disulfide-stabilized core (3 disulfide bonds) characterized by an unusual knotted structure (Figure 1) ^[1, 2]. In this motif, an embedded ring in the structure formed by two disulfide bonds and their connecting backbone segments is penetrated by the third disulfide bond. Cyclotides contrast with other circular polypeptides in that they have a highly defined three-dimensional structure, and despite their small size, can be considered as miniproteins. Their unique circular backbone topology and knotted arrangement of 3 disulfide bonds makes them exceptionally stable to thermal and enzymatic degradation ^[3]. Furthermore, their well defined structures have been also associated with a range of biological activities, including uterotonic ^[4], neurotensin antagonistic ^[5], hemolytic ^[6, 7], anti-HIV ^[8], antibacterial ^[7], insecticidal ^[9] and trypsin inhibitory ^[10] activities. Together, these characteristics make cyclotides ideal candidates to be used as molecular scaffolds for the development of stable peptide drugs ^[11, 12].

Although the chemical synthesis of several cyclotides using solid-phase peptide synthesis have been reported ^[7, 13-15], having access to biosynthetic cyclotides using recombinant DNA expression techniques offers the exciting possibility of producing large combinatorial libraries of highly stable miniproteins using the tools of molecular biology. This would allow the generation of cell-based libraries that could be screened either *in vitro* or *in vivo* for their ability to regulate cellular processes.

In nature, cyclotides are ribosomally synthesized as precursor proteins that are then subsequently processed into the corresponding circular polypeptide ^[9, 16]. Despite knowledge of the underlying gene sequences of different cyclotides, the mechanism that governs circularization is not totally

understood. Although the possibility of an autocatalytic split-intein-type mechanism has not yet been disproved, it is more likely the involvement of specialized proteolytic enzymes.

In an attempt to biosynthesize cyclotides inside living cells, our group recently reported the use of engineered protein splicing units for the *in vitro* cyclization and folding of cyclotides using conditions similar to those found inside living cells^[17]. Our biomimetic approach is based on the use of an intramolecular version of native chemical ligation (NCL)^[18] to produce the head-to-tail or backbone cyclization of the corresponding linear cyclotide precursor (Figure 2).

Intramolecular NCL requires the presence of an N-terminal Cys residue and C-terminal α -thioester group in the same linear precursor^[19,20]. As shown in Figure 2, this was accomplished by producing a recombinant protein where the linear cyclotide precursor was fused in frame at its C- and N-terminus to a modified intein and a Met residue, respectively. This allows the generation of the required C-terminal thioester function^[21] and N-terminal Cys residue^[22]. This approach was successfully used for the *in vitro* biosynthesis of different Kalata B1-based cyclotides^[17].

In the present work we report the use of a similar approach for the *in vivo* biosynthesis of a fully folded and biologically active cyclotide inside living bacterial cells. For this purpose we used a modified gyrase intein^[23] instead. This bacterial-derived mini-intein is considerably smaller (≈ 27 kDa) than the previously employed vacuolar membrane ATPase (VMA) intein^[24] (≈ 57 kDa) and shows high levels of expression in bacterial expression systems. This ensures a higher level of expression for the corresponding cyclotide linear precursor. Incorporation of Met residue at the N-terminus of the cyclotide linear precursor sequence makes possible the generation of a N-terminal Cys residue by endogenous Met aminopeptidase (MAP) as the precursor intein fusion protein is translated *in vivo*^[25]. Initial attempts to use the gyrase intein for the cyclization

of the cyclotide KB1, however, were unsuccessful^[17] due to difficulties associated with purification of the corresponding folded cyclotide from a complex soluble cellular fraction. Despite the fact that KB1 was the first cyclotide to be discovered and it has been structurally well characterized, its biological function is not well understood and no binding partners have yet been identified. This made impossible to use of any type of affinity chromatography for its purification from the cytosolic fraction. In order to facilitate purification and determine whether folded cyclotides could be produced inside living cells, we decided to use the cyclotide MCoTI-II (*Momordica cochinchinensis* trypsin inhibitor II, Figure 1)^[10] instead. This cyclotide is a powerful trypsin inhibitor ($K_i \approx 30$ pM) that has been recently isolated from a plant of the squash-family and it. Moreover, its three-dimensional structure has also been recently determined by NMR by two independent groups^[26, 27] confirming its cyclic cystine knot (CCK) topology (Figure 1).

In vivo production of folded MCoTI-II was first explored by expressing several MCoTI-intein fusion proteins (**1** through **3**, Scheme 1) in *Escherichia coli* strain BL21(DE3). Each construct was designed to use a different N-terminal Cys residue for the cyclization site. Hence, constructs **1** and **2** used the naturally occurring Cys residues at the beginning and end of loop 6, respectively. This loop contains a highly flexible peptide sequence and it is not required for folding or biological activity^[28]. In fact, this loop is missing in MCoTI-III (a related acyclic trypsin inhibitor from the same plant) and not found in related linear trypsin inhibitors from other plants of the squash-family^[14]. Construct **3** used the native Cys residue at the end of loop 3. Fusion proteins **1** and **2** expressed at very high level in *E. coli* (≈ 40 mg/L). The expression level for construct **3**, however, was significantly less (≈ 4 mg/L). Affinity purification of the fusion proteins revealed that all the constructs had similar levels of *in vivo* cleavage ($\approx 60\%$, Figure 3a).

Treatment of the purified fusion proteins with reduced glutathione for 2 days at pH 7.4 yielded the folded MCoTI-II as one of the major components obtained from constructs **1** and **3**.

Construct **2**, however, gave a more complex crude reaction with only minor amounts of folded product when treated under the same conditions. This was attributed to the inefficient removal of the N-terminal Met residue by MAP. It has been described that the presence of Pro residue in position P2' considerably reduces the proteolytic efficiency of MAP ^[29].

Next, we tried the isolation of the folded MCoTI-II generated *in vivo* after expression of construct **1** in *E. coli* BL21(DE3) cells. This was accomplished by incubating the soluble fraction of a fresh cell lysate with trypsin-immobilized agarose beads. After extensive washing, the absorbed products were eluted with a solution of 8 M GdmCl and analyzed by HPLC. As shown in Figure 3b, the analytical HPLC trace revealed the presence of a major component, which had the expected mass for the fully folded MCoTI-II (Figure 3b inset). A fresh cell sample was also lysated in the presence of 20 mM ICH₂CONH₂ and purified using trypsin-agarose beads as described above. The presence of ICH₂CONH₂ quenches any cyclization/folding that may take place during the lysis step. In both cases the amount of folded MCoTI-II found in the soluble cellular fraction was identical, indicating that the folded cyclotide was produced *in vivo* and not during the lysis step. The intracellular concentration of folded MCoTI-II was estimated ≈100 nM. *In vivo* formation of folded MCoTI-II was further improved by expressing construct **1** in *E. coli* Origami 2(DE3) cells. These genetically engineered bacterial cells contain several mutations in the thioredoxin and glutathione reductases that favor the formation of disulfide bridges ^[30]. Hence, construct **1** was expressed very efficiently in Origami 2 cells, with expression levels similar to those obtained in BL21 cells. The expression level of folded MCoTI-II was estimated to be ≈33 μg/L of cell culture. Although the overall yield for the *in vivo* cyclization and folding

process was rather small ($\approx 2\%$), the intracellular concentration of MCoTI-II produced inside Origami cells was estimated $\approx 5 \mu\text{M}$, which is high enough to carry out *in vivo* screening experiments ^[31, 32].

Biochemical and structural characterization confirmed the identity of the recombinantly produced MCoTI-II cyclotide. The trypsin inhibitory activity of recombinant MCoTI-II was measured by using a standard fluorescence assay (Figure 3c) ^[28]. The calculated K_i for porcine pancreatic trypsin was 25 pM , which is similar to the value reported in the literature for the natural product ^[28]. Recombinant MCoTI-II was also characterized by homonuclear 2D-NMR. Deviations in the chemical shifts for the H^αC and H^αN protons for the recombinant and natural cyclotide ^[26, 27] were uniformly less than 0.05 ppm (Figure 3d), confirming the native structure of recombinant MCoTI-II.

In summary, we have described for the first time the biosynthesis of a fully functional cyclotide inside living bacterial cells. This was accomplished by using an intein-mediated cyclization approach. The cyclized product folds spontaneously inside the bacterial cytoplasm to give the fully functional cyclotide. The use of engineered *E. coli* cell lines that promote disulfide bridge formation clearly improved the yield of MCoTI-II formation. Under these conditions, it was possible to reach intracellular concentrations up to $5 \mu\text{M}$ in folded cyclotide. These results open the possibility for the creation of cyclotide-based libraries that could be screened for their ability to attenuate or inhibit molecular processes inside living cells ^[31, 32].

Acknowledgments. We would like to thank Dr. Alexander R. Mitchell for useful discussions. 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 No. W-7405-Eng-48.

References

- [1] D. J. Craik, N. L. Daly, T. Bond, C. Waine, *J Mol Biol* **1999**, *294*, 1327.
- [2] K. J. Rosengren, N. L. Daly, M. R. Plan, C. Waine, D. J. Craik, *J Biol Chem* **2003**, *278*, 8606.
- [3] M. L. Colgrave, D. J. Craik, *Biochemistry* **2004**, *43*, 5965.
- [4] L. Gran, *Lloydia* **1973**, *36*, 174.
- [5] K. M. Witherup, M. J. Bogusky, P. S. Anderson, H. Ramjit, R. W. Ransom, T. Wood, M. Sardana, *J Nat Prod* **1994**, *57*, 1619.
- [6] N. L. Daly, D. J. Craik, *J Biol Chem* **2000**, *275*, 19068.
- [7] J. P. Tam, Y. A. Lu, J. L. Yang, K. W. Chiu, *Proc Natl Acad Sci U S A* **1999**, *96*, 8913.
- [8] K. R. Gustafson, R. C. Sowder, L. E. Louis E. Henderson, I. C. Parsons, Y. Kashman, J. H. Cardellina, J. B. McMahan, R. W. Buckheit, L. K. Pannell, M. R. Boyd, *J. Am. Chem. Soc.* **1994**, *116*, 9337.
- [9] C. Jennings, J. West, C. Waine, D. Craik, M. Anderson, *Proc Natl Acad Sci U S A* **2001**, *98*, 10614.
- [10] J. F. Hernandez, J. Gagnon, L. Chiche, T. M. Nguyen, J. P. Andrieu, A. Heitz, T. Trinh Hong, T. T. Pham, D. Le Nguyen, *Biochemistry* **2000**, *39*, 5722.
- [11] D. J. Craik, M. Cemazar, C. K. Wang, N. L. Daly, *Biopolymers* **2006**, *84*, 250.
- [12] D. J. Craik, M. Cemazar, N. L. Daly, *Curr Opin Drug Discov Devel* **2006**, *9*, 251.
- [13] N. L. Daly, S. Love, P. F. Alewood, D. J. Craik, *Biochemistry* **1999**, *38*, 10606.
- [14] L. Chiche, A. Heitz, J. C. Gelly, J. Gracy, P. T. Chau, P. T. Ha, J. F. Hernandez, D. Le-Nguyen, *Curr Protein Pept Sci* **2004**, *5*, 341.
- [15] R. J. Clark, N. L. Daly, D. J. Craik, *Biochem J* **2006**, *394*, 85.
- [16] J. L. Dutton, R. F. Renda, C. Waine, R. J. Clark, N. L. Daly, C. V. Jennings, M. A. Anderson, D. J. Craik, *J Biol Chem* **2004**, *279*, 46858.
- [17] R. H. Kimura, A. T. Tran, J. A. Camarero, *Angew Chem Int Ed Engl* **2006**, *45*, 973.
- [18] P. E. Dawson, S. B. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923.
- [19] J. A. Camarero, J. Pavel, T. W. Muir, *Angew. Chem. Int. Ed.* **1998**, *37*, 347.
- [20] J. A. Camarero, T. W. Muir, *J. Am. Chem. Soc.* **1999**, *121*, 5597.
- [21] T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249.
- [22] J. A. Camarero, D. Fushman, D. Cowburn, T. W. Muir, *Bioorg Med Chem* **2001**, *9*, 2479.
- [23] T. Klabunde, S. Sharma, A. Telenti, W. R. Jacobs, J. C. Sacchettini, *Nat. Struct. Biol.* **1998**, *5*, 31.
- [24] C. J. Noren, J. M. Wang, F. B. Perler, *Angew. Chem. Int. Ed.* **2000**, *39*, 451.
- [25] P. H. Hirel, M. J. Schmitter, P. Dessen, G. Fayat, S. Blanquet, *Proc. Natl. Acad. Sci. U S A* **1989**, *86*, 8247.
- [26] M. E. Felizmenio-Quimio, N. L. Daly, D. J. Craik, *J Biol Chem* **2001**, *276*, 22875.
- [27] A. Heitz, J. F. Hernandez, J. Gagnon, T. T. Hong, T. T. Pham, T. M. Nguyen, D. Le-Nguyen, L. Chiche, *Biochemistry* **2001**, *40*, 7973.
- [28] O. Avrutina, H. U. Schmoldt, D. Gaborjelcic-Geiger, D. Le Nguyen, C. P. Sommerhoff, U. Diederichsen, H. Kolmar, *Biol Chem* **2005**, *386*, 1301.
- [29] A. Ben-Bassat, K. Bauer, S. Y. Chang, K. Myambo, A. Boosman, S. Chang, *J Bacteriol* **1987**, *169*, 751.
- [30] P. H. Besette, F. Aslund, J. Beckwith, G. Georgiou, *Proc Natl Acad Sci U S A* **1999**, *96*, 13703.

- [31] T. M. Kinsella, C. T. Ohashi, A. G. Harder, G. C. Yam, W. Li, B. Peelle, E. S. Pali, M. K. Bennett, S. M. Molineaux, D. A. Anderson, E. S. Masuda, D. G. Payan, *J. Biol. Chem.* **2002**, *277*, 37512.
- [32] X. You, A. W. Nguyen, A. Jabaiah, M. A. Sheff, K. S. Thorn, P. S. Daugherty, *Proc Natl Acad Sci U S A* **2006**, *103*, 18458.

Figures and Schemes

- 1 **M**CGSGSDGGV**CPKILKK**CRRDSD**CPGAC**ICRGNGY-Intein
- 2 **M**CP**PKILKK**CRRDSD**CPGAC**ICRGNGY**CGSGSDGGV**-Intein
- 3 **MC**ICRGNGY**CGSGSDGGVCPKILKK**CRRDSD**CPGA**-Intein

Scheme 1. Sequences of the different precursors used in this work for the *in vivo* production of MCoTI-II. The sequence corresponding to loop 1, which is responsible for binding to the active site of trypsin, is in green as a reference.

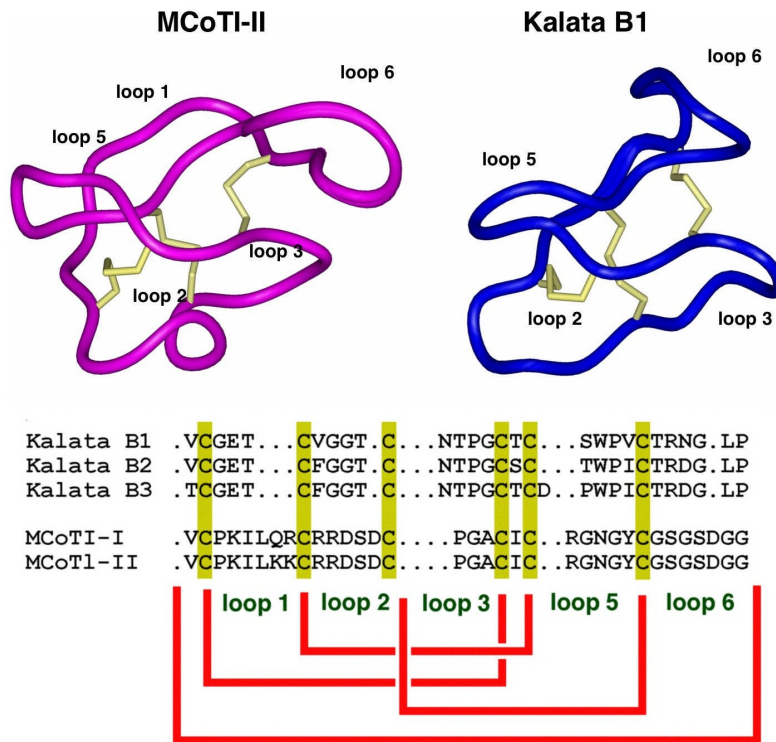


Figure 1. Primary and tertiary structure of cyclotides from the plants *Momordica cochinchinensis* (MCoTI-II) and *Oldenlandia affinis* (Kalata B1).

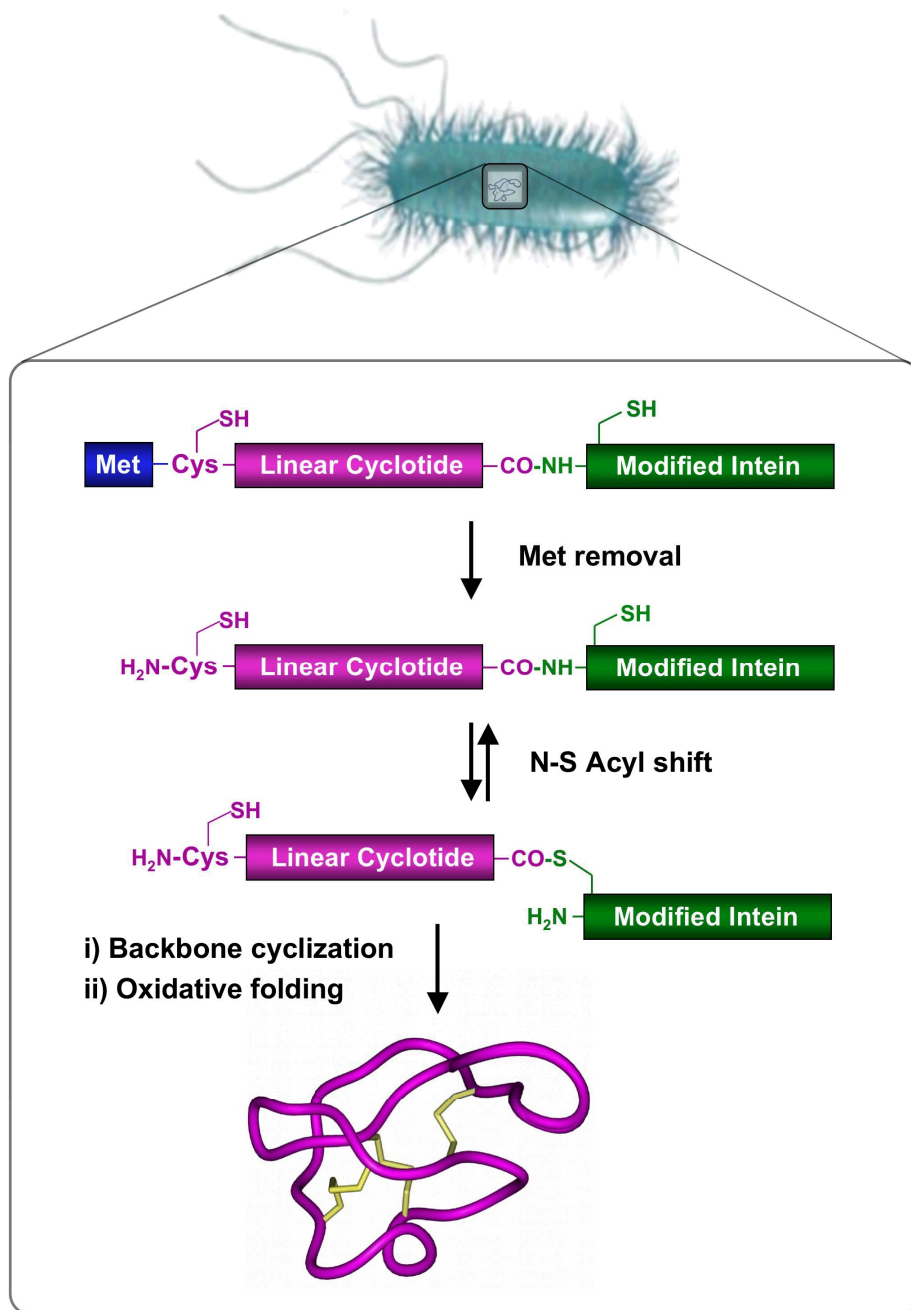


Figure 2. Biosynthetic approach for the in vivo production of cyclotide MCoTI-II inside live *E. coli* cells. Backbone cyclization of the linear cyclotide precursor is mediated by a modified protein splicing unit or intein. The cyclized product then folds spontaneously in the bacterial cytoplasm.

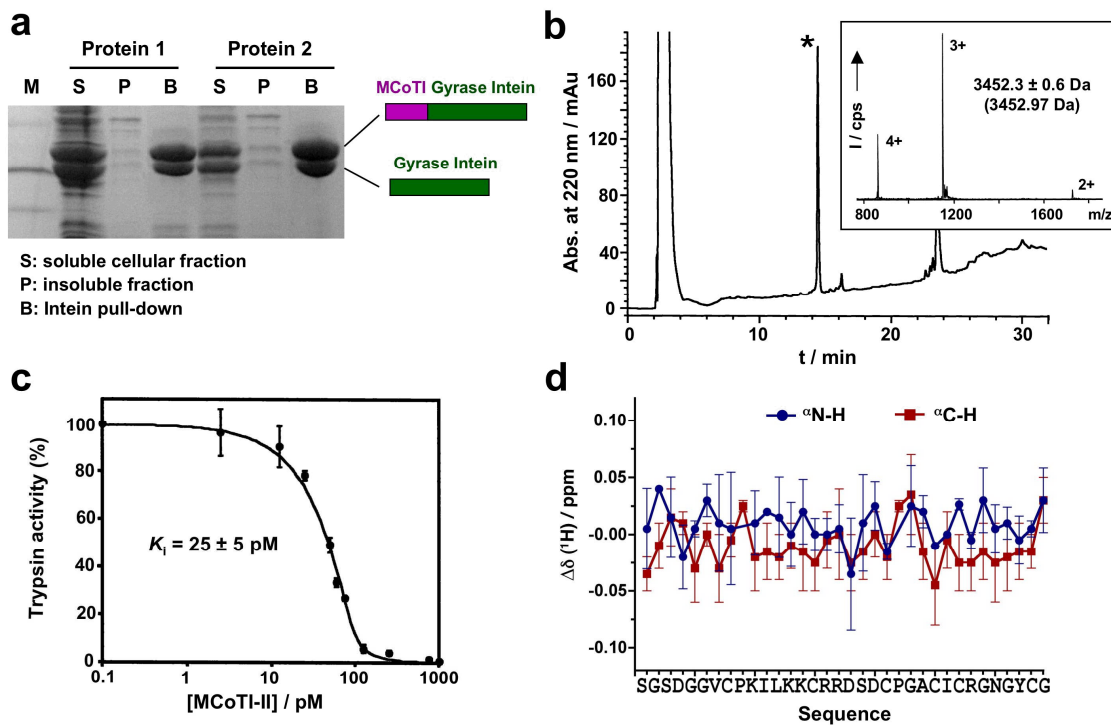


Figure 3. a) SDS PAGE analysis of the expression levels and *in vivo* cleavage of precursor proteins 1 and 2. b) Analytical reverse-phase HPLC trace of soluble cellular fraction from construct 1 after being purified on trypsin-immobilized agarose beads. The asterisk denotes the fully folded MCoTI-II. Inset, ES-MS analysis of affinity purified MCoTI-II. c) Inhibition of porcine trypsin by recombinant MCoTI-II. d) Summary of the ^1H -NMR assignments for the backbone protons: $\Delta\delta(^1\text{H})$ are the deviations in the chemical shifts of the main chain protons between the recombinant and the natural MCoTI-II cyclotide [26, 27].