

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**DESIGN AND SYNTHESIS OF
CYSTEINE-RICH PEPTIDES**

Yibo Qiu

School of Biological Sciences

2014

**DESIGN AND SYNTHESIS OF
CYSTEINE-RICH PEPTIDES**

Yibo Qiu

School of Biological Sciences

A thesis submitted to the Nanyang Technological University in partial
fulfillment of the requirement for the degree of Doctor of Philosophy

2014

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor Professor Ding Xiang Liu and my co-supervisor Professor James P. Tam for their great ideas and patient guidance throughout my PhD period. In this four-year study, they kept inspiring me with their immense knowledge, great motivation and trained me with a high standard. Their education makes the process an invaluable experience for me. It is my great pleasure to be one of their students and their training will always drive me to become better in my lifetime.

My appreciation also goes to Professor Chuan-Fa Liu for his encouragements and technical support during this project. He always offers me practical instructions and professional advices when I need help. I also would like to thank Professor Siu Kwan Sze and Professor Kathy Qian Luo for their help on the mass spectrometry and bioassays.

I am grateful to Dr. Xinya Hemu, Dr. Clarence T. T. Wong, Dr. Misako Taichi and Dr. Kien Truc Giang Nguyen. They are friends, mentors and partners. Their intelligence and attitude inspired me a lot and I enjoyed working with them. In addition, I would like to thank my FYP and URECA students for their input in this project. I would like to thank Dr. Ying Liao, Dr. Adrian Guan Ji How, Dr. Quoc Thuc Phuong Nguyen, Dr. Renliang Yang and other SBS friends who helped me during my experiment. Everyone is nice and helpful to me, which makes my stay here a happy and meaningful experience.

Last, I would like to thank my parents and my family for their constant support and understanding.

List of Conference Presentations

12th Chinese International Peptide Symposium, Shenyang China, 2012

Oral Presentation, Young Peptide Scientist Award

List of Publications

Qiu, Y., Wong, C. T. T., and Tam, J. P. (2012) Acid-catalyzed N-S acyl shift for preparing peptide thioesters using Fmoc chemistry. in *Proceedings of the 12th Chinese International Peptide Symposium*. Peptides, Chemistry and Biology. (Liu, K., Cheng, M., and Meng, Q. eds.), Chemical Industry Press, Shenyang, China **1**, 22-24

Tam, J. P., Wong, C. T. T., and **Qiu, Y.** (2012) Engineering of proteinaceous orally active bradykinin peptide antagonists. in *Proceedings of the 32nd European Peptide Symposium. Peptide 2012*. (Kokotos, G., Constantinou-Kokotou, V., and Matsoukas, J. eds.), John Wiley and Sons, Athens, Greece **18**, 130-131

Taichi, M., Hemu, X., **Qiu, Y.**, and Tam, J. P. (2013) A thioethylalkylamido (TEA) thioester surrogate in the synthesis of a cyclic peptide via a tandem acyl shift. *Org. Lett.* **15**, 2620-2623

Hemu, X., Taichi, M., **Qiu, Y.**, Liu, D. X., and Tam, J. P. (2013) Biomimetic synthesis of cyclic peptides using novel thioester surrogates. *Pept. Sci.* **100**, 492-501

Qiu, Y., Hemu, X., Liu, D. X., and Tam, J. P. (2014) Selective bi-directional amide bond cleavage of N-methylcysteinyl peptide. *Eur. J. Org. Chem.* **2014**, 4370-4380

Hemu, X., **Qiu, Y.**, and Tam, J. P. (2014) Peptide macrocyclization through amide-to-amide transpeptidation. *Tetrahedron* **70**, 7707-7713

Nguyen, G. K. T., Wang, S., **Qiu, Y.**, Hemu, X., Lian, Y., and Tam, J. P. (2014) Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **10**, 732-738

Table of Contents

| | |
|---|-------------|
| Acknowledgements | i |
| List of Conference Presentations..... | ii |
| List of Publications | iii |
| Table of Contents | iv |
| List of Figures..... | viii |
| List of Tables | xiii |
| Abbreviations | xiv |
| Abstract..... | xvii |
| Chapter 1 Introduction..... | 1 |
| 1. Cysteine-rich peptides..... | 1 |
| 2. Circular peptides | 7 |
| 3. Cyclotides | 9 |
| 3.1 Discovery of cyclotides..... | 9 |
| 3.2 Structure and classification of cyclotides..... | 9 |
| 3.3 Biosynthesis of cyclotides..... | 12 |
| 3.4 Biological activities of cyclotides | 18 |
| 4. CRPs as peptide biologics..... | 20 |
| 5. Amide-to-amide chemoenzymatic synthesis of cyclic peptides | 31 |
| 5.1 Intein-mediated synthesis of cyclic peptides..... | 31 |
| 5.2 Sortase-mediated synthesis of cyclic peptides | 35 |
| 6. Amide-to-amide chemical synthesis of cyclic peptides..... | 37 |
| 6.1 Solid-phase peptide synthesis | 38 |
| 6.2 Chemical ligation | 40 |
| 6.3 The thia zip assisted cyclization..... | 48 |
| 6.4 Preparation of thioesters for cysteine thioester ligation..... | 50 |
| 7. Oxidative folding of CRPs..... | 60 |
| 8. Aim | 66 |
| Chapter 2 Material and methods..... | 67 |
| 1. Chemical Synthesis of Cyclic CRPs..... | 67 |
| 1.1 General information | 67 |
| 1.2 MeCys-mediated N ^α -thioester formation in aqueous conditions..... | 68 |

| | |
|--|-----------|
| 1.3 MeCys-mediated synthesis of cyclic peptide SFTI-1 | 70 |
| 1.4 MeCys-mediated C ^α -thioester formation in strongly acidic conditions.. | 71 |
| 1.5 β-thiolactone formation..... | 76 |
| 1.6 TEBA-mediated N ^α -thioester formation in aqueous conditions | 77 |
| 1.7 TEBA-mediated synthesis of cyclic peptide SFTI-1 | 78 |
| 1.8 Synthesis of kalata B1 via a hydrazide linker | 79 |
| 2. Oxidative folding of CRPs..... | 81 |
| 2.1 Oxidative folding of kB1 in organic conditions..... | 81 |
| 2.2 ¹ H NMR characterization | 81 |
| 2.3 Synthesis of ET-1 | 81 |
| 2.4 Oxidative folding of ET-1 in organic solvent. | 82 |
| 3. Butelase-mediated synthesis of cyclic peptides | 82 |
| 3.1 Synthesis of peptide library XXXGIR-OH | 82 |
| 3.2 Optimization of butelase-mediated ligation reaction | 83 |
| 3.3 Ligation of KALVINHV and XXXGIR mediated by butelase 1 | 83 |
| 3.4 Synthesis of C-terminal substrate library KALVXNHV | 84 |
| 3.5 Butelase 1 mediated ligation of KALVXXHV and HIGGIR | 84 |
| 4. Design and engineering of cysteine-rich peptides | 85 |
| 4.1 Synthesis of SFBK, BKSF | 85 |
| 4.2 Synthesis of SFBK-L, BKSF-L, DALK | 86 |
| 4.3 Heat and pH stability assay | 87 |
| 4.4 Pepsin and trypsin stability assay..... | 88 |
| 4.5 Serum stability assay | 88 |
| 4.6 Trypsin inhibition assay | 88 |
| 4.7 Cytotoxicity..... | 88 |
| 4.8 Fluo-4 NW calcium assay | 89 |
| Chapter 3 Chemical Synthesis of Cyclic CRPs | 91 |
| 1. Introduction..... | 91 |
| 2. MeCys as a thioester surrogate for preparation of cyclic CRPs | 94 |
| 2.1 Synthetic strategy using MeCys peptides for preparing cyclic CRPs..... | 94 |
| 2.2 Synthesis of model peptides on the MeCys resin by Fmoc SPPS..... | 96 |
| 2.3 MeCys-mediated N ^α -thioester formation in aqueous conditions..... | 100 |
| 2.4 MeCys-mediated synthesis of cyclic peptide SFTI-1 | 111 |

| | |
|---|------------|
| 2.5 MeCys-mediated C ^α -thioester formation | 114 |
| 2.6 β-Thiolactone formation | 133 |
| 3. TEBA as a thioester surrogate for preparation of cyclic CRPs | 138 |
| 3.1 Design and rationale of using TEBA as a thioester surrogate | 138 |
| 3.2 Synthesis of model peptides on the TEBA resin by Fmoc SPPS..... | 138 |
| 3.3 TEBA-mediated N ^α -thioester formation in aqueous conditions | 141 |
| 3.4 TEBA-mediated synthesis of cyclic peptide SFTI-1 | 144 |
| 4. Hydrazide as a thioester surrogate for preparation of cyclic CRPs | 148 |
| 4.1 Synthesis of kalata B1 on the hydrazide resin by Fmoc chemistry..... | 149 |
| 4.2 One pot synthesis of kB1 | 149 |
| 5. Discussion | 162 |
| Chapter 4 Oxidative folding of cysteine-rich peptides | 169 |
| 1. Introduction..... | 169 |
| 2. Oxidative folding of kB1 in organic solvents | 170 |
| 2.1 Solvent effect | 172 |
| 2.2 Basicity effect..... | 177 |
| 2.3 Redox effect | 179 |
| 3. Oxidative folding of ET-1 in organic solvents | 183 |
| 4. Discussion | 187 |
| Chapter 5 Butelase-mediated synthesis of cyclic peptides | 191 |
| 1. Introduction..... | 191 |
| 2. N-terminal substrate specificity in butelase-mediated ligation..... | 194 |
| 2.1 Design and synthesis of peptide libraries | 194 |
| 2.2 Optimization of butelase-mediated ligation reaction | 197 |
| 2.3 N-terminal substrate specificity at the P1'' position..... | 199 |
| 2.4 N-terminal substrate specificity at the P2'' position..... | 199 |
| 2.5 N-terminal substrate specificity at the P3'' position..... | 200 |
| 2.6 Unnatural amino acids at the P1'' and P2'' positions..... | 203 |
| 3. C-terminal substrate specificity in butelase-mediated ligation | 205 |
| 3.1 Synthesis of C-terminal substrate library KALVXNHV | 205 |
| 3.2 C-terminal substrate specificity at the P1 and P2 positions | 208 |
| 4. Discussion | 210 |
| Chapter 6 Design and engineering of cysteine-rich peptides | 218 |

| | |
|---|------------|
| 1. Introduction..... | 218 |
| 2. Design and synthesis of engineered peptides..... | 220 |
| 2.1 Design of engineered bradykinin receptor antagonists | 220 |
| 2.2 Chemical synthesis of engineered peptides by Fmoc SPPS..... | 220 |
| 3. Characterizations of engineered peptides | 223 |
| 3.1 Heat stability assay..... | 223 |
| 3.2 pH stability assay | 223 |
| 3.3 Pepsin stability assay..... | 226 |
| 3.4 Trypsin stability assay | 226 |
| 3.5 Serum stability assay | 227 |
| 3.6 Trypsin inhibition assay | 230 |
| 3.7 Bradykinin receptor antagonism assay..... | 232 |
| 4. Discussion | 234 |
| Chapter 7 Conclusions..... | 240 |
| References | 242 |

List of Figures

- Figure 1.1 Examples of circular peptides from various species
- Figure 1.2 Three-dimensional structure of hedyotide B1 (hB1) showing the loops and disulfide bonds of cyclotide
- Figure 1.3 A proposed mechanism showing major steps involved in the biosynthesis of cyclotides
- Figure 1.4 A proposed model for AEP-mediated cyclization of kB1
- Figure 1.5 A schematic demonstration and modeling structure of cyclic NP-1
- Figure 1.6 A scheme showing the original and engineered α -conotoxin Vc1.1
- Figure 1.7 A scheme demonstrating the grafting of bioactive peptides into cyclotide scaffolds
- Figure 1.8 A scheme showing the design and sequence of bradykinin antagonists and the grafted peptide ckb-kal and ckb-kin
- Figure 1.9 Mechanism of intein-mediated protein splicing
- Figure 1.10 A scheme showing the Sortase-catalyzed anchoring of surface proteins to the bacterial cell walls
- Figure 1.11 Schematic demonstration of entropy-driven chemoselective ligation
- Figure 1.12 A scheme showing the mechanism of native chemical ligation
- Figure 1.13 Schematic demonstration of the thia zip cyclization mechanism
- Figure 1.14 Selected examples of safety catch methods for peptide thioester synthesis
- Figure 1.15 Peptide thioester formations through O/N-S acyl shift reactions
- Figure 1.16 “Safety switch” methods based on N-S acyl shifts
- Figure 3.1 Thioester surrogates containing the essential TEA moiety
- Figure 3.2 The “Cys-peptide-MeCys” strategy for synthesis of cyclic peptides
- Figure 3.3 Synthesis of the model peptides using MeCys resins
- Figure 3.4 The mechanism of 3-(1-piperidinyl)alanine formation on the C-terminal MeCys

- Figure 3.5 Reaction scheme and HPLC profiles of MeCys-mediated tandem thiol switch reaction at various pH
- Figure 3.6 Scheme, HPLC and ESI profile of TIGGIR-MES and CALVIN-NH₂ peptide ligation
- Figure 3.7 MS/MS profile of ligation product TIGGIRCALVIN-NH₂
- Figure 3.8 One-pot chemical ligation of TIGGIR-MeCys-Gly-NH₂ and CALVIN-NH₂ with external thiol MMA in pH 5 at 40 °C
- Figure 3.9 Comparison of N^α-Cleavage at a Gly-MeCys or a Gly-Cys bond
- Figure 3.10 MeCys-mediated synthesis of SFTI-1 through tandem thiol switch, thia zip cyclization and oxidation
- Figure 3.11 RP-HPLC profile of MeCys-mediated C^α-thioester formation
- Figure 3.12 MS/MS profile of TIGGIR-MeCys-TC
- Figure 3.13 RP-HPLC profile of TIGGIR-MeCys-TC in H₂O containing 0.1% TFA
- Figure 3.14 Scheme, RP-HPLC and MS profile of TIGGIR-MeCys-TC and CALVIN-NH₂ peptide ligation
- Figure 3.15 MS/MS profile of TIGGIR-MeCys-CALVIN-NH₂
- Figure 3.16 Bi-directional cleavage of the MeCys residue under different conditions
- Figure 3.17 RP-HPLC profile of TIGGIR-MeAla-Gly in strongly acidic conditions containing 5% TC, 0.1% TFMSA in TFA
- Figure 3.18 Summary of reactions containing different model peptides treated with 5% TC, 0.1% TFMSA in TFA
- Figure 3.19 The protonation ratio curve for DMS in non-aqueous TFMSA/TFA solutions
- Figure 3.20 A proposed mechanism for C^α- cleavage of *N*-methylated amino acid
- Figure 3.21 Mechanism for β-Thiolactone formation
- Figure 3.22 β-Thiolactone formation from TIGGIR-MeCys-TC at pH 4-7 phosphate buffer
- Figure 3.23 RP-HPLC of NEM alkylated TIGGIR-MeCys-TC in pH 5 phosphate buffer at 1 h and 12 h

- Figure 3.24 MS/MS profile of TIGGIR-MeCys β -thiolactone (m/z 715)
- Figure 3.25 RP-HPLC profiles of four different thioesters TIGGIR-MeAla-TC, TIGGIR-MeCys-MMA, TIGGIR-Cys-TC, and TIGGIR-Cys-MMA in pH 5 phosphate buffer
- Figure 3.26 Synthesis of model peptides on the TEBA resin
- Figure 3.27 TEBA-mediated tandem thiol switch reaction at various pH
- Figure 3.28 Comparison of pH-dependent MeCys and TEBA mediated tandem thiol switch reaction
- Figure 3.29 Scheme of TEBA-mediated synthesis of SFTI-1
- Figure 3.30 RP-HPLC and MS profiles of synthetic intermediates in the cyclization of SFTI-TEBA
- Figure 3.31 Scheme of one pot synthesis of kB1 by the hydrazide resin
- Figure 3.32 RP-HPLC and MALDI-TOF MS profiles of crude kB1 hydrazide after TFA cleavage
- Figure 3.33 RP-HPLC and MALDI-TOF MS profiles of synthetic kB1 hydrazide, kB1 azide, cyclic kB1 and native kB1
- Figure 3.34 RP-HPLC profiles of the cyclization and folding results with 1 mM and 0.5 mM peptide
- Figure 3.35 RP-HPLC profiles of the oxidative folding under different basicities
- Figure 3.36 RP-HPLC profiles of oxidative folding reactions with different concentrations of thiols
- Figure 3.37 Co-injection of native kB1 and synthetic kB1 in RP-HPLC
- Figure 3.38 MALDI-TOF MS and MS/MS profiles of synthetic kB1
- Figure 3.39 Scheme of native chemical ligation and the reversed thioester formation
- Figure 3.40 Comparison of intein-mediated protein splicing and TEA-mediated tandem thiol switch
- Figure 3.41 Schematic demonstrations of intein-catalyzed cyclization and TEA thioester surrogate-mediated cyclization
- Figure 4.1 RP-HPLC and MALDI-TOF MS profiles of native and reduced kB1

- Figure 4.2 RP-UPLC profiles of the oxidative folding of kB1 showing the effect of binary solvent mixtures
- Figure 4.3 RP-UPLC profiles showing the process of the oxidative folding of kB1
- Figure 4.4 RP-UPLC profiles of oxidative folding of kB1 showing the effect of different bases
- Figure 4.5 RP-UPLC profiles of oxidative folding of kB1 showing the effect of cysteamine and DMSO as redox agents
- Figure 4.6 RP-HPLC profiles of oxidative folding of kB1 with different thiols as reducing reagents
- Figure 4.7 ^1H NMR spectra comparison of the natural kB1 (red) and the synthetic one (blue)
- Figure 4.8 Schematic demonstrations of ET-1 isomers
- Figure 4.9 RP-HPLC and MALDI-TOF MS profiles of reduced and oxidized synthetic ET-1
- Figure 4.10 RP-HPLC profiles of ET-1 in different organic folding conditions containing 0-85% TFE
- Figure 5.1 Butelase-mediated peptide cyclization
- Figure 5.2 Scheme of butelase-mediated ligation of KALVINHV and XXXGIR
- Figure 5.3 Ligation yields of KALVINHV and XIGGIR (X=20 natural amino acids)
- Figure 5.4 Ligation yields of KALVINHV and LXGGIR (X=20 natural amino acids)
- Figure 5.5 Ligation yields of KALVINHV and LLXGIR (X=20 natural amino acids)
- Figure 5.6 Butelase-mediated peptide ligation with unnatural amino acids at the P1'' and P2'' position
- Figure 5.7 Molecular weights and molar ratio of amino acids in each batch
- Figure 5.8 RP-HPLC profile of C-terminal peptide library KALVXNHV
- Figure 5.9 Butelase-mediated intermolecular ligation of KALVIXHV and HIGGIR (X=Asn, Asp or D-form Asn)

- Figure 5.10 Butelase-mediated intermolecular ligation of KALVXNHV and HIGGIR (X=20 natural amino acids)
- Figure 5.11 Comparison of natural sequence preference and *in vitro* substrate specificity at various positions
- Figure 5.12 A scheme showing the mechanism of X-Cys ligation.
- Figure 5.13 Comparison of three chemoenzymatic approaches for peptide cyclization
- Figure 6.1 A scheme of SFTI-1 and its DALK grafted analogues
- Figure 6.2 Heat and low pH stability of engineered peptides
- Figure 6.3 Pepsin and trypsin stability of engineered peptides
- Figure 6.4 Serum stability of engineered peptides
- Figure 6.5 Trypsin inhibition activities of engineered peptides
- Figure 6.6 Bradykinin receptor antagonism activities of engineered peptides
- Figure 6.7 NMR structures of SFTI-1, SFBK and BKSF

List of Tables

| | |
|------------|--|
| Table 1.1 | General features of selected CRPs from plants to animals. |
| Table 1.2 | Representative disulfide bond patterns shared by CRPs from different species. |
| Table 1.3 | Selected examples demonstrating the grafting strategy using cysteine-rich peptide scaffolds. |
| Table 1.4 | Summary of various ligation methods. |
| Table 1.5 | “Safety switch” methods based on O-S acyl shifts. |
| Table 1.6. | Global oxidative folding conditions of CRPs or small disulfide-rich proteins. |
| Table 2.1 | Molecular weights of TIGGIR series peptides |
| Table 3.1 | Product summary of TIGGIR-Xaa-MeCys-Yaa-NH ₂ in strongly acidic conditions containing 5% TC, 0.1% TFMSA in TFA. |
| Table 3.2 | H_0 values of different acidic solution containing 0-0.5% TFMSA in TFA. |
| Table 3.3 | Cyclization conditions with different nitrite and thiol concentrations. |
| Table 4.1 | Oxidative conditions of kB1 with varying solvent combinations. |
| Table 5.1 | Peptide sequences in three combinatorial libraries. |
| Table 5.2 | Conditions with different ratios of butelase 1: KALVINHV: GIGGIR. |
| Table 6.1 | Sequence alignment of SFTI-1 and its DALK grafted analogs. |

Abbreviations

Standard abbreviations are used for the amino acids and protecting groups [IUPAC-IUB Commission for Biochemical Nomenclature (1985) *J biol. Chem.* 260:14-42].

| | |
|------------------|---|
| Acm | Acetamidomethyl |
| ACN | Acetonitrile |
| AEP | Asparaginyl endopeptidase |
| AMP | Antimicrobial peptide |
| Boc | <i>tert</i> -Butoxycarbonyl |
| BOP | Benzotriazol-1-yl-oxo-tris(dimethylamino)phosphonium hexafluorophosphate |
| CBD | Chitin-binding domain |
| CCK | Cyclic cystine knot |
| CS $\alpha\beta$ | Cysteine-stabilized α -helix β -sheet |
| CRP | Cysteine-rich peptide |
| DCC | <i>N,N'</i> -Dicyclohexylcarbodiimide |
| DCM | Dichloromethane |
| DIC | <i>N,N'</i> -Diisopropylcarbodiimide |
| DIEA | <i>N,N</i> -Diisopropylethylamine |
| DMAP | 4-Dimethylaminopyridine |
| DMF | <i>N,N</i> -Dimethylformamide |
| DTT | Dithiothreitol |
| DMSO | Dimethylsulfoxide |
| EDT | 1,2-Ethanedithiol |
| EDTA | Ethylenediaminetetraacetic acid |
| ER | Endoplasmic reticulum |
| ESI-MS | Electrospray ionization mass spectrometry |
| ET-1 | Endothelin-1 |
| FDA | Food and drug administration |
| Fmoc | 9-Fluorenylmethoxycarbonyl |
| GI | Gastrointestinal |
| Gnd-HCl | Guanidine hydrochloride |

| | |
|----------|---|
| GSH | Reduced glutathione |
| GSSG | Oxidized glutathione |
| HATU | <i>O</i> -(7-azabenzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate |
| HET | Hydroxyethylthiol |
| HF | Hydrofluoric acid |
| HOAt | 7-Aza-1-hydroxybenzotriazole |
| HOBt | 1-Hydroxybenzotriazole |
| HPLC | High performance liquid chromatography |
| kB1 | kalata B1 |
| LC-MS | Liquid chromatography mass spectrometry |
| MALDI MS | Matrix-assisted laser desorption ionization mass spectrometry |
| MeCys | <i>N</i> -methylated cysteine |
| MESNa | Mercaptoethanesulfonate sodium salt |
| MMA | Methyl mercaptoacetate |
| MW | Molecular weight |
| No. | Number |
| NMR | Nuclear magnetic resonance |
| NEM | <i>N</i> -Ethylmaleimide |
| PDI | Protein disulfide isomerase |
| PyBOP | (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate |
| RP-HPLC | Reverse-phase high-performance liquid chromatography |
| rcf | Relative centrifugal force |
| RTD | <i>Rhesus</i> theta-defensin |
| SFTI-1 | Sunflower trypsin inhibitor-1 |
| S-S | Disulfide bond |
| SPPS | Solid-phase peptide synthesis |
| TC | <i>p</i> -Thiocresol |
| TCEP | <i>tris</i> (2-carboxyethyl)phosphine |
| TEA | Thioethylamido |
| TEBA | Thioethylbutylamido |
| TFA | Trifluoroacetic acid |

| | |
|----------|---|
| TFE | Trifluoroethanol |
| TFMSA | Trifluoromethanesulfonic acid |
| TIS | Triisopropylsilane |
| TMAS | Tetramethylammonium sulfate |
| TMT | Thiomethylthiazolidine |
| TOF | Time-of-flight |
| Tris-HCl | Tris(hydroxymethyl)aminomethane hydrochloride |
| UPLC | Ultra performance liquid chromatography |

Abstract

My thesis focuses on the methodology and application of novel synthetic strategies for high-throughput preparation of the cyclic cysteine-rich peptides with therapeutic potential. Cyclic cysteine-rich peptides are macrocyclic peptides with intramolecular disulfide bonds. The end-to-end cyclic backbone together with multiple disulfide bonds provides conformational constraints to enhance the structural stability. My strategy involves synthesizing these peptides by chemical approaches using solid-phase peptide synthesis. A new cyclization method was developed employing an N-S acyl shift mechanism to mimic the natural production of peptide bonds. The oxidative folding process is also optimized by introducing organic solvents. With the novel cyclization and the organic folding strategy in a one-pot manner, I have successfully prepared cyclic cysteine-rich peptides in a shorter reaction time with an improved yield than with conventional methods. They could be used as a simple and high-throughput synthetic platform to prepare cyclic cysteine-rich peptides. In addition, the substrate specificity of a novel transpeptidase named butelase 1 was also investigated for ligation and cyclization reactions. With high efficiency and broad substrate specificity, it would be exploited as a novel chemoselective approach for peptide and protein engineering.

For applications, the cyclic peptide sunflower trypsin inhibitor-1 which possesses a cyclic backbone is modified by a grafting approach with bioactive peptides while retaining the cyclic peptide scaffold. New analogs with improved stability and new functions were developed as a proof-of-concept to advance the development of potential peptide biologics.

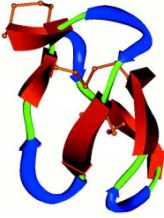

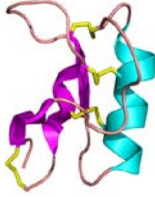
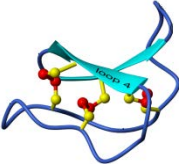
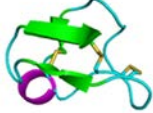
Chapter 1 Introduction

1. Cysteine-rich peptides

Among all amino acids, cysteine is unique as its thiol side chain can form disulfide bonds. Disulfide bonds place additional structural constraint to peptides and proteins and are important for their stabilities. Thus, cysteine-rich peptides (CRPs) have drawn much attention because they are generally stable for their multiple intramolecular disulfide bonds. Despite that the observed average frequency of cysteine is 3.28%,¹ CRPs could contain up to 50% cysteine. One of the smallest CRP is “scratcher” conotoxin tx3c from *Conus textile* venom. Three disulfide bonds are present in this 12-amino-acid peptide (H-CCRTCFGCTPCC-NH₂), which divide the sequence into three short loops. The disulfide bond between Cys1 and Cys6 links the two termini together and forms a pseudo-cyclic structure while the other two disulfide bonds (Cys2-Cys4, Cys3-Cys5) crossed each other.² Indeed, this peptide is very special in that half of its amino acids are involved in disulfide bonds. The constrained structure is considered to be highly evolved and important for its diffusion upon injection and rapid paralysis of prey.³

CRPs are diverse in occurrence and contain many classes. Examples include venom toxins, defensins, antimicrobial peptides, enzyme inhibitors, and growth factors (**Table 1.1**).⁴⁻⁸ Many organisms, such as snakes, spiders, marine snails, scorpions and plants, have been known to produce CRPs. In model plants *Oryza sativa* (rice) and *Arabidopsis thaliana*, it is estimated that CRPs represent 2-3% of the gene repertoires.⁹ Another example includes conotoxins from the venom of cone snails. Over 500 species have been identified in genus *Conus*. With each species expressing >100 distinct conotoxins,¹⁰ it is estimated that there are >50,000 conotoxins with ion-channel blocking activities.

Table 1.1 General features of selected CRPs from plants to animals.

| Group description | Species | Approx. Size (aa) | No of cysteine | Model structures* |
|-------------------|--------------------------------------|-------------------|----------------|---|
| Venom toxin | Snail Snake Spider Scorpion | 10-70 | 2-10 |  ω-Conotoxin MVIIA |
| Defensin | invertebrate vertebrate | 18-45 | 6-8 |  RTD-1 |
| | Plant | 45-54 | 8-10 |  NaD1 |
| Cyclotide | Plant | 28-37 | 6 |  kalata B1 |
| Enzyme inhibitor | Plant | 14-60 | 2-6 |  α-amylase inhibitor AAI |

*Model structures are from Protein Data Bank. PDB number: ω-conotoxin MVIIA (1OMG), RTD-1 (1HVZ), NaD1 (1MR4), kalata B1 (1KAL), α-amylase inhibitor AAI (1QFD).

In marine snails, spiders, scorpions and snakes, CRPs are secreted as toxins for predation as well as host defense. They include conotoxins from cone snails,³ atracotoxins,¹¹ hanatoxins,¹² raventoxins¹³ from spiders, pandinotoxins from scorpions¹⁴ and sarafotoxins from snakes¹⁵. Most are neurotoxins targeting specific ion channels and neurotransmitter receptors. Unlike plant CRPs, neurotoxic CRPs are hydrophilic, which helps them to exist in a concentrated solution in venom and diffuse rapidly after injection into preys' serum. Conotoxins are one of the most well-known neurotoxins. They contain 10-30 residues with one or more disulfide bonds,¹⁶ causing a rapid immobilization (seconds) of the preys once upon injection. These paralyzing effects are due to their strong binding to ion channels and receptors.

Conotoxins are target-specific and differ from family to family. α -Conotoxins inhibit nicotinic acetylcholine receptors¹⁷ while ω -conotoxins selectively bind to Ca^{2+} channels.¹⁸ Because of their high selectivity towards various ion channels, conotoxins have found great therapeutic potentials in pharmaceutical areas. For example, MVIIA from the ω -conotoxin family is a FDA (Food and Drug Administration) approved drug (Ziconotide¹⁹) used to treat severe and chronic pain. It is 100-1000 times more potent than morphine²⁰ and serves as a specific blocker of N-type voltage-gated Ca^{2+} channels.

Another major family of animal CRPs is defensins belonging to the anti-microbial peptide (AMP) group. Typically, they contain 18-45 residues with three or four disulfide bonds. They serve as defense peptides with a broad-spectrum of inhibition against bacteria, fungi, and viruses. The general mechanism of defensin function involves interaction of its hydrophobic patches with pathogens' membranes, leading to pore formations on the membrane surface and cell lysis of the pathogen.^{21,22} Animal defensins can be classified as α , β , or θ according to their disulfide connectivity.²³ *Rhesus* theta-defensin-1 (RTD-1), belonging to the θ -defensins, is unusual because of its cyclic backbone formed by ligation of two α -defensins.²⁴ The cyclic backbone confers additional conformational constraints on the peptide, which increase its anti-microbial activity by three-fold. The discovery of this cyclic defensin proves that host cells possess the machinery for synthesis of macrocyclic peptides and provides some clues for the biosynthesis process.

In plants, CRPs {e.g. plant defensins,²⁵ thionins,²⁶ heveins,²⁷ knottins (Database can be accessed at <http://knottin.cbs.cnrs.fr>),²⁸ lipid transfer proteins²⁹ and snakins^{30,31}} are produced as an innate defense against pathogens. Plants are sessile and continuously surrounded by predators and pathogens. Thus, they need to place barriers, both physical and chemical, to protect themselves from hazards. Defensins and other anti-microbial peptides, most of which are cysteine-rich, are abundantly produced to serve this purpose, as chemical responses to the environment. Due to their structural similarity with defensins from other organisms, these plant CRPs are simply named as plant defensins.²⁵ Typically, four or five disulfide bonds are present in plant defensins. Besides cysteine, plant defensins also contain a high ratio of hydrophobic and basic amino acids, which lead to disruption of bacterial and fungal membranes.

Cyclotides are another class of plant CRPs. They are gene-coded and possess end-to-end cyclic backbone. Although their exact function in plants has not yet been clarified, they display a wide range of pharmaceutical activity, such as anti-bacteria,³² anti-virus,^{33,34} anti-cancer,³⁵ and neurotensin antagonist,³⁶ making them promising scaffolds for drug design which will be discussed in the subsequent sections.

A third class of plant CRPs is enzyme inhibitors, which includes carboxypeptidase inhibitors, trypsin inhibitors and α -amylase inhibitors. The α -amylase inhibitor (AAI) from the seeds of *Amaranthus hypocondriacus* strongly inhibits α -amylase from the larvae of the red flour beetle, suggesting a defensive functions in plants.³⁷ Interestingly, AAI shares the same disulfide bond pattern as ω -conotoxin, although they are from distinct species (plant versus cone snails) and share little sequence homology and biological function (α -amylase inhibition versus Ca^{2+} channel blocking). This suggests that disulfide bond patterns might serve as structural templates to display functional epitopes.

Earlier studies of the single-disulfide-bond peptide oxytocin showed that disulfide bonds play a more important role in structure than in function.^{38,39} Similar conclusions were made from studies of θ -defensins. In fact, these studies found that disulfide bonds were not essential for its anti-microbial activities but important for structure and stability.⁴⁰ In CRPs, disulfide bonds confer structural rigidity to

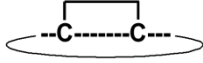
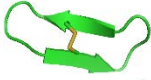
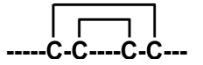

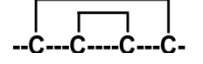

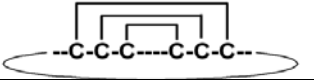


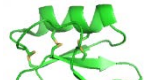
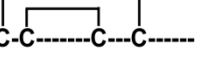

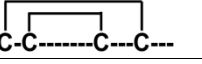
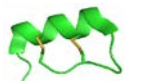
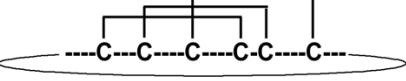

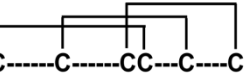

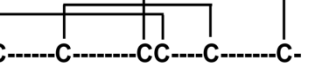

maintain the active conformation for receptor binding and improved stability against thermal and enzymatic degradation. Although sequences among various CRPs are quite different, disulfide bond patterns seem to be conserved and define the overall structure. Similar disulfide bond patterns are consistently observed across different species. Representative examples are summarized in **Table 1.2**.

The β -sheet “ladder” type of disulfide bond is well presented in CRPs from plants, crabs, porcines and *Rhesus macaques*. This bond type is characterized by the symmetric locations of one to three disulfide bonds between two β -sheets and examples include sunflower trypsin inhibitor-1 (SFTI-1), protegrins, tachyplesins and RTD-1. In overall structures, disulfide bonds link two antiparallel β -sheets and maintain constrained conformations.

The second type of ladder is the cysteine-stabilized α -helix β -sheet (CS $\alpha\beta$) motif, which was first named by Cornet and co-workers in 1995.⁴¹ After first being characterized from the three-dimensional structure of charybdotoxin,⁴² the CS $\alpha\beta$ motif was continuously discovered in other neurotoxins from bees, scorpions, and snakes.⁴³ This motif contains two disulfide bonds formed between CXC and CXXXC sequences which stabilize the α -helix and the paralleled β -sheets. Although they share the same motif, examples of peptides containing this motif are diverse in function. Neurotoxins with ion-channel modulating activities are the most-well-characterized. Endothelin-1 (ET-1), a peptide with the vasoconstriction effect, was also found to possess this motif. More recently, the CS $\alpha\beta$ motif was discovered in anti-microbial plant defensins although they contain two additional disulfide bonds, and the orientation of the β -sheet is different.⁴⁴

A third motif includes the cystine-knot which consists of three disulfide bonds, wherein two disulfide bonds (Cys1-Cys4, Cys2-Cys5) together with the short loops between them form a plane which is penetrated by the third one (Cys3-Cys6) forming a knot. The cystine-knot is the most frequently observed motif and this motif is commonly found in enzyme inhibitors, conotoxins, cyclotides, growth factors,^{37,45-47} and many plant AMPs.²⁸

Table 1.2 Representative disulfide bond patterns shared by CRPs from different species.

| Disulfide bond | | Representatives | | | | | |
|----------------|---|---------------------|---------------|------------|----------------|---|------|
| pattern | Disulfide motif | Family | Name | No. of S-S | Species | Structure | Ref. |
| β-sheet |  | Trypsin inhibitor | SFTI-1 | 1 | Plant |  | 48 |
| |  | AMP | Protegrin-1 | 2 | Porcine |  | 49 |
| |  | AMP | Tachyplesin-1 | 2 | Crab |  | 50 |
| Ladder |  | Defensin | RTD-1 | 3 | Rhesus macaque |  | 51 |
| CSαβ |  | Plant defensin | NaD1 | 4 | Plant |  | 52 |
| |  | Vasoconstrictor | ET-1 | 2 | Human |  | 53 |
| |  | Neurotoxin | SRTX-m | 2 | Snake |  | 54 |
| Cystine-knot |  | Cyclotide | kalata B1 | 3 | Plant |  | 55 |
| |  | Conotoxin | MVIIA | 3 | Cone snail |  | 45 |
| |  | α-amylase inhibitor | AAI | 3 | Plant |  | 56 |

Model structures are from Protein Data Bank. PDB number: SFTI-1 (1JBL), Protegrin-1 (1PG1), Tachyplesin-1 (1MA2), RTD-1 (1HVZ), NaD1 (1MR4), ET-1 (1EDN), SRTX-m (2LDF), kalata B1 (1KAL), ω-conotoxin MVIIA (1OMG), α-amylase inhibitor AAI (1QFD).

2. Circular peptides

Besides disulfide bonds, cyclic backbones are also shown to provide peptides with conformational constraints and resistance against protease degradation. Ribosomally synthesized circular peptides possess an end-to-end macrocyclic backbone with N- and C-termini linked by an amide bond. They are widely found in different organisms (**Figure 1.1**), including fungi, bacteria, mammals and plants.^{57,58} Fungi, or more specifically mushrooms, produce small cyclic toxins with 7-10 amino acids.⁵⁹ One of the most well-known examples is α -amanitin, a potent toxin with an oral LD₅₀ of 0.1 mg/kg.⁶⁰ Larger cyclic peptides with 35-78 amino acids are found in bacteria.⁶¹ Pilins, such as TrbC (78 residues)⁶² and bacteriocins, such as AS-48 (70 residues),⁶³ are the largest cyclic peptides identified so far. Smaller cyclic peptides with 6-11 amino acids are also observed in cyanobacteria.⁶⁴

θ -Defensins are cyclic peptides found in mammals.^{24,64} The θ -defensin gene is also present in human genome but not expressed because of a stop codon in the signal sequence. The putative human defensin peptide, named retrocyclin, was synthesized by solid-phase peptide synthesis, and has been shown to protect cells from HIV-1 infections *in vitro*.⁶⁵ Other cyclic peptides have been discovered in plants including cyclotides,⁶⁶ sunflower trypsin inhibitors⁶⁷ and Caryophyllaceae-type CP⁶⁸ (so-called orbitides⁶⁹). So far, plant cyclic peptides constitute the largest family of circular peptides, most being cyclotides, with >50,000 potential members.⁷⁰

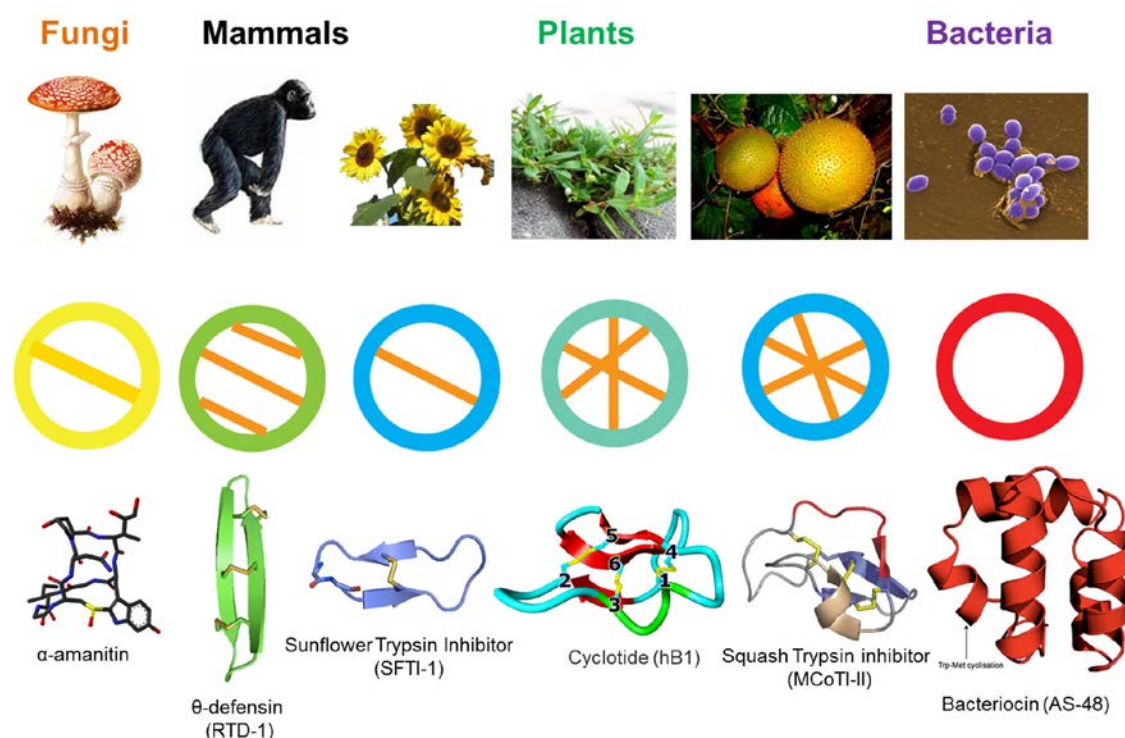


Figure 1.1 Examples of circular peptides from various species. Circular peptide producing species, schematic representations of peptide structures and three-dimensional peptide structures are aligned. In schematic representations, circles represent circular backbones. The yellow stick denotes a sulfoxide linkage between a 6-hydroxy-tryptophan and a cysteine, while orange sticks represent disulfide bonds. From left to right, α -amanitin (PDB No.: 1JBL) from *Amanita phalloides* (The death cap mushroom), RTD-1 (PDB No.: 1HVZ) from *Rhesus macaque* (Monkey), SFTI-1 (PDB No.: 1JBL) from *Helianthus annuus* (Sunflower), hB1 (unpublished structure by Sui and Tam, 2010) from *Hedyotis biflora*, MCoTI-II (PDB No.: 1IB9) from *Momordica cochinchinensis* (Red melon), AS-48 (PDB No.: 1E68) from *Enterococcus faecalis*.

3. Cyclotides

3.1 Discovery of cyclotides

The first cyclotide was discovered by Lorents Gran, a Norwegian physician, in the 1960s. During his relief mission to the Democratic Republic of Congo, he noticed that local women drank a herbal tea made from boiling leaves from the plant *Oldenlandia affinis* DC to accelerate childbirth during labor.⁷¹ Later, a peptide with the oxytocic activity was identified as the active component and named kalata B1 (kB1)⁷² after the local medicinal tea “kalata-kalata.” The structure of kB1 was resolved in 1995⁵⁵ and demonstrated as an end-to-end macrocyclic backbone with a cystine-knot motif. During this time, more and more cyclic peptides with similar structural motifs were discovered, such as circulins³⁴ and cyclopsychotride A³⁶. In addition, these cyclic peptides shared high sequence homologies. Thus, they were classified into a group called cyclotides (cyclo – peptides). Since then, over 200 cyclotides have been isolated mainly from plants belonging to the Rubiaceae (coffee),⁷³⁻⁷⁵ Violaceae (violet),⁷⁶ Fabaceae (legume),^{77,78} Solanaceae (potato)⁷⁹ and Cucurbitaceae (squash)⁸⁰ families. Recently, our laboratory has identified hundreds of novel cyclotides in medicinal plants, including the Fabaceae family⁷⁸, a large and important family in agriculture, and uncyclotide (linear form cyclotide) in Poaceae (rice) family⁸¹.

3.2 Structure and classification of cyclotides

Typically, cyclotides contain 28-37 amino acids and range from 2.6-3.5 kDa. Six cysteine residues are located within the structure, forming three disulfide bonds in a knotting structure (**Figure 1.2**). This cystine-knot structure feature, together with the cyclic backbone, is collectively called the cyclic cystine knot (CCK) motif,⁶⁶ which endows cyclotides with impressive stability. Based on the location of the six conserved cysteine residues, the cyclotide backbone can be divided into six loops. Among them, loops 1 and 4 are buried inside the cystine knot core and are constrained and conserved, while loop 6 is the cyclization site in the biosynthesis of cyclotides. Although cyclotides vary in loop size and sequence, several residues are quite conserved and critical. For example, Asn (or Asp) in loop 6 is believed to be essential for cyclization in cyclotide biosynthesis. In addition, Glu in loop 1 plays an

important role in the bioactivity of cyclotides.⁸² With the exception of loops 1 and 4, the other four loops (2, 3, 5, and 6) are more flexible and variable.

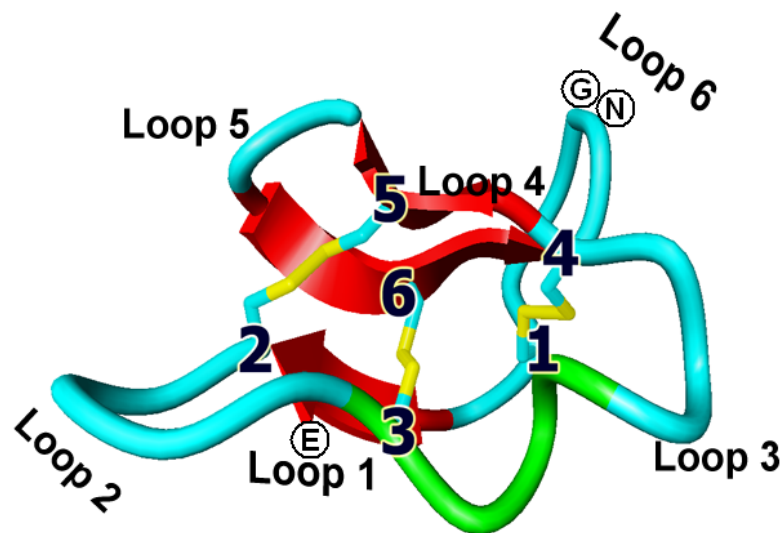


Figure 1.2 Three-dimensional structure of hedyotide B1 (hB1) showing the loops and disulfide bonds of cyclotide. The positions of cysteines are shown in bold Arabic numerals. β -sheet structures are colored by red and disulfide bonds are in yellow. Conserved residues are highlighted using circled letters.

Cyclotides can be classified into two subfamilies: Möbius and bracelet, according to the presence of the *cis*-proline residue in loop 5. In the Möbius subfamily, a twist is introduced due to the presence of the *cis*-proline in loop 5. In the bracelet subfamily, no twist exists in loop 5, and the overall shape looks like a bracelet. According to cybase (<http://www.cybase.org.au>), most of the cyclotides identified so far (about two-thirds) belong to the bracelet subfamily. It is believed that the two distinguished conformations are related to the oxidative folding process and result in dramatically different yield under *in vitro* folding conditions.⁸³ Besides the two major subfamilies, other cyclotide variants have been discovered. Recently, the trypsin inhibitor subfamily was identified. Although they do not share many conserved residues with the previous two subfamilies, the trypsin inhibitor subfamily has the same disulfide connectivity and displays similar cyclic backbone topology. Therefore, from a structural point of view, they are considered as one of the cyclotide subfamilies. MCoTI-II is the representative of this subfamily.⁸⁰ Recently, linear cyclotides (“uncyclotides”) were also isolated from Solanaceae, Poaceae and Rubiaceae families.^{74,79,81} Uncyclotides lack the cyclic backbone but share similar genetic origins and sequence homology with their cyclic counterparts. Therefore, they are also classified as a subfamily of cyclotides. Uncyclotides often lack the N-terminal Gly or C-terminal Asn, which are essential for backbone cyclization in cyclotides. It is believed that linear cyclotides are produced due to genetic diversification and post translational modifications.⁷⁴

3.3 Biosynthesis of cyclotides

Cyclotides have genetically-encoded sequences and are synthesized by ribosomes *in vivo*. The maturation of cyclotides involves two steps: oxidative folding and cyclization. Oxidative folding forms the correct disulfide connectivity, while the cyclization process joins the N- and C-termini of the peptide backbone together as an amide, which builds the CCK motif of cyclotides.

Cyclotides are post-translationally-processed products from linear precursors. Structures of precursor proteins vary from species to species. A typical representative cyclotide precursor (**Figure 1.3B**) contains five regions: an endoplasmic reticulum (ER) signal, an N-terminal propeptide (NTPP) followed by an N-terminal repeat

(NTR), a cyclotide domain, and a C-terminal propeptide (CTPP). The mature cyclotide sequence is encoded by the cyclotide domain. Most plants encode a single cyclotide copy in an individual gene, such as the cyclotide gene *oakI*⁸⁴ in *O. affinis* from the Rubiaceae family. However, multiple (as much as seven) copies⁸⁵ of the cyclotide domain are also observed in a single cyclotide precursor from the Cucurbitaceae family. The diversity of cyclotide genes does not limit to that. Our laboratory has discovered that the Albumin-1 gene is hijacked by the cyclotide gene to produce clotides in the plant *Clitoria ternatea*.⁷⁸ In this plant, the cyclotide domain replaced the chain b part of the Albumin-1 gene and formed a chimeric structure composed of a cyclotide domain and an Albumin-1 chain a domain. Similar gene structures are also found in another type of cyclic peptide SFTI-1.⁸⁶

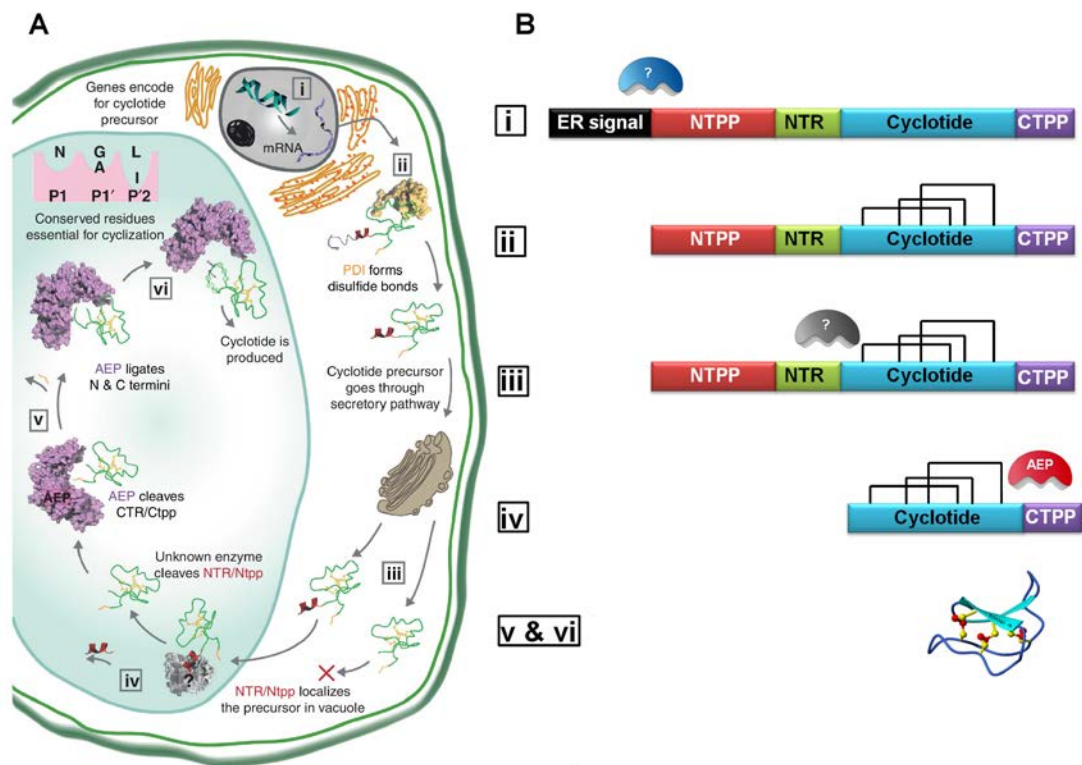


Figure 1.3 A proposed mechanism showing major steps involved in the biosynthesis of cyclotides. Panel A demonstrates the subcellular locations of each step in the cyclotide processing (Adapted from ref. ⁸⁷). Panel B illustrates the corresponding reactions occurring in each step. The process involves (i) a cyclotide gene was transcribed in the nucleus and translated into a precursor protein by ribosomes; (ii) ER signal targets the precursor to ER lumen, PDI helps to form the three disulfide bonds and removal of the ER signal sequence; (iii) NTPP/NTR localizes the folded precursor from ER to vacuole; (iv) removal of NTPP/NTR sequence by an unknown enzyme in vacuole; (v&vi) a transpeptidation reaction mediated by AEP to cleave the CTPP sequence and ligate the N- and C-termini.

The proposed biosynthesis pathway and major processing events occurring in this process are summarized in **Figure 1.3 A**. In the nucleus and ribosomes, cyclotide genes are transcribed and translated, respectively, as five-domain precursor proteins (22 kDa). The ER signal directs transfer of the cyclotide mRNA from the nucleus to the ER, where the oxidative folding forms the three conserved disulfide bonds. This process is suspected to be catalyzed by the protein disulfide isomerase (PDI). PDI ensures the correct disulfide connectivity by forming, breaking, and shuffling disulfide bonds of the substrate via a thiol-disulfide exchange reaction, although the exact mechanism is still unclear. *In vitro* studies have confirmed PDI facilitates formation of correct disulfide connectivity in both linear and cyclic cyclotides.⁸⁸ However, *in vivo* evidence of PDI-assisted oxidative folding in the biosynthesis of cyclotides is still missing. It is also believed that reduced glutathione (GSH) and oxidized glutathione (GSSG) help reshuffle the incorrect disulfide linkages to correct ones.⁸⁹ Gunasekera and co-workers⁹⁰ suggested that the NTR domain may play an essential role in folding based on its highly conserved sequence and helical structure, however, no conclusive evidence has yet been found. After the linear precursor is folded, the N- and C-termini of a cyclotide are brought into close proximity to allow the cyclization to occur.

With excision of the ER signal peptide, an 11 kDa precursor intermediate with correctly formed disulfide bonds is released from the ER, and then relocated to the vacuole for cyclization under the guidance of NTPP/NTR. Asparaginyl endopeptidases (AEPs) have been long suspected to be responsible for cyclization because of the highly conserved Asx (Asn or Asp) residues at the C-terminus of the cyclotide domain. Gillon and co-workers supported the hypothesis using mutagenesis methods in the model species tobacco and *Arabidopsis*.⁹¹ These non-cyclotide-producing plants expressed cyclic form cyclotides after being transfected with cyclotide genes. Single mutations of the C-terminal Asn residue in the cyclotide domain abolished the cyclization and only linear form cyclotides were produced. Besides mutations, gene silencing experiments to knock down the AEP expression also resulted in reduced or no production of cyclotides in transgenic plants.⁹² These results provide some ideas of the enzymes involved in the cyclization. Furthermore, the subcellular location of the cyclotide precursor *Oak1* is in vacuole⁹³, suggesting

that AEPs may be involved in the cyclization of cyclotides since they are known to be vacuolar processing enzymes.

The mechanism of AEP-mediated cyclotide cyclization has been proposed using prototypic cyclotide kB1 (**Figure 1.4**^{91,94}). Little is known about how the C-terminal propeptide is processed or how the cyclization reaction occurs. It is believed that cleavage of the CTPP and subsequent cyclization is accomplished by transpeptidation in one pot rather than separately. This is supported by the fact that cyclotides lacking the entire C-terminal propeptide were unable to cyclize; however, cyclization was recovered when the C-terminal tripeptide motif (GLP) was included.⁹¹ Thus, the C-terminal peptide GLP motif is considered to be important for the cyclization and is recognized by and binds to AEP at the S1', S2', and S3' positions. The recognition motif also contains tripeptide TRN, C-terminal residues of the cyclotide domain, occupying the S3, S2, and S1 binding sites, respectively.

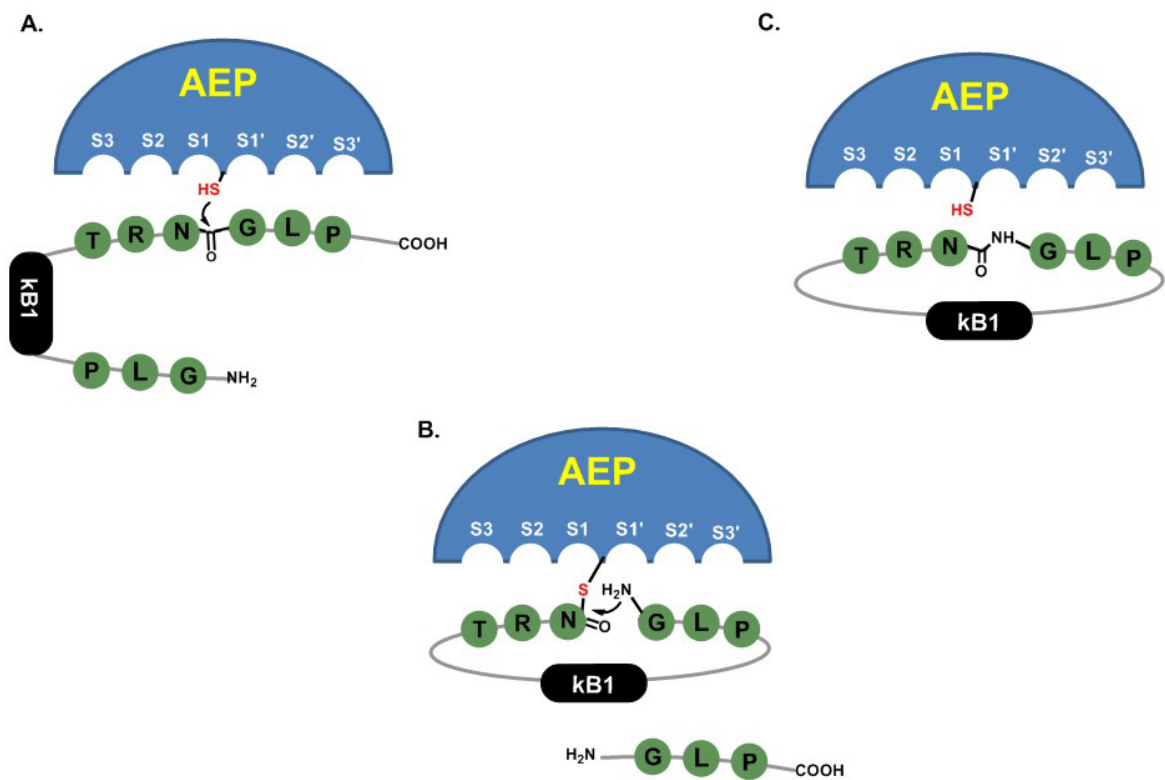


Figure 1.4 A proposed model for AEP-mediated cyclization of kB1.(Adapted from ref ⁹⁴) (A) AEP (Blue) recognizes the C-terminal motif (TRNGLP) in kB1 precursor and forms a thioacyl-enzyme intermediate with kB1 (black) to release the C-terminal propeptide (Green). (B) The N-terminal tripeptide GLP in kB1 peptide occupies the binding sites vacated by the C-terminal propeptide. Then, a ligation reaction occurs via S-N acyl shift. (C) The cyclic backbone of kB1 is formed and a mature kB1 is produced.

The process of AEP-mediated cyclization is similar as other transpeptidases from the cysteine protease family. Upon binding, the catalytic cysteine will be activated and serves as a nucleophile to attack the carbonyl group of Asn at S1 site. (**Figure 1.4A**) An acyl-enzyme intermediate is then formed between the AEP and the cyclotide domain via a thioester linkage, releasing the C-terminal propeptide GLPSLAA. (**Figure 1.4B**) Due to proximity-driven effects, the free amine of N-terminal Gly will initiate a nucleophilic attack on the thioester bond to resolve the acyl-enzyme intermediate and form a new amide bond. Thus, the cyclic backbone of a cyclotide is formed, releasing the cyclic peptide from the enzyme. (**Figure 1.4C**) Finally, a mature cyclotide is produced and secreted with its N- and C- termini linked by an amide and disulfide bonds correctly formed.

Besides AEPs, proteases such as trypsin are also able to mediate cyclization of cyclotides. Two cyclic peptides, MCOTI-II^{95,96} and SFTI-1⁹⁷, have been successfully synthesized *in vitro* employing this method. In this approach, a Lys residue is required to act as the ligation site rather than Asx in AEP-mediated cyclization, which corresponds to a Lys-Ile bond in MCOTI-II and Lys-Ser bond in SFTI-1. In nature, the ligation site of MCOTI-II and SFTI-1 is Asp-Gly, similar to the Asn-Gly ligation site in other cyclotides. Thus, trypsin is probably not the enzyme involved in their biosynthesis although it is able to mediate their cyclization. This process is thought to be a result of their trypsin inhibitory activities. Therefore, this approach is specific for certain trypsin inhibitors and not applicable for other cyclic peptides.

3.4 Biological activities of cyclotides

The very first cyclotide kB1 was reported by an assay-guided isolation procedure.⁷² Its oxytocic activity is the first biological activity people have discovered for cyclotides. Since then, more and more biological activities have been revealed for various cyclotides, including anti-microbial activities, anti-tumor, anti-HIV, hemolytic, anti-fouling, neurotensin antagonism and anti-insect larvae.^{32,34,36,98-100} Most biological activities of cyclotides are related to their toxicity on various organisms such as bacteria, fungi, viruses and insects. The wide range of organisms against which cyclotides exert inhibitory effects suggests their innate defensive function in plants.

In 1999, Tam and co-workers first reported the anti-microbial activities of cyclotides from both Möbius and bracelet subfamilies using synthetic kB1, circulin A and B, and cyclopsychotride. They showed broad-spectrum inhibition against Gram-positive and Gram-negative bacteria, as well as fungi at low micromolar concentrations. Later, more and more cyclotides were found to display anti-microbial activity; thus, this activity has become one of the major areas in cyclotide research. Subsequently, varv A and cycloviolacin O2 were found to show cytotoxic effects on tumor cells and normal cells with certain selectivity. They display a different inhibitory pattern from other clinically-proved anti-cancer drugs,¹⁰¹ which involves a rapid morphological modification on the tumor cell surface. Circulin A and B were the first two cyclotides found to possess anti-HIV properties but they were also highly cytotoxic to normal cells. Notably, mutations of hydrophobic residues in loop 2 and 3 to reduce the overall hydrophobicity drastically lowers the cytotoxicity but also reduces the anti-HIV effect.¹⁰²

As more and more cyclotides activities are found to be related to their hydrophobicity or membrane binding propensity, it seems that their propensity to bind and disrupt membranes are responsible for most of their biological activities.¹⁰³ Continuous hydrophobic patches (loop 2 and 5 for Möbius, loop 2 and 3 for bracelet subfamily¹⁰⁴) are exposed on the surface of cyclotides due to the compact CCK motif. Thus, the overall structure of cyclotides is highly defined, similar to proteins with amphipathic surfaces rather than random linear peptides. This shape helps cyclotides to attach and interact with membrane lipids through phosphatidylethanolamine.¹⁰⁵ Positively-charged residues are also important for membrane disruption, as most cyclotides with clustered positive charges are more membranolytic than negatively-charged or neutral cyclotides. The correlation of peptides' net positive charges and their anti-microbial activities has been revealed by Tam and co-workers in 2002.¹⁰⁶ It was suggested that the electrostatic interaction between the cationic residues and the negatively-charged microbial membranes disrupted the membrane surfaces and induced membrane lysis. Interestingly, no specific sequence motif has been found to be essential to the biological activity of cyclotides. Rather, their structure integrity seems more important for their function. Even those conserved amino acids important for cyclotide function, like Glu in loop 1, they do not act through the binding to

specific receptors but by maintaining structural integrity via hydrogen bonding.¹⁰⁷ Although most cyclotide biological activities are believed to be related to their membranolytic effects, this feature is not likely to be the mechanism behind all cyclotide functions, such as their oxytocic and neurotensin antagonistic activities. In these cases, it is more probably that cyclotides bind to specific receptors or ligands to take effect.

4. CRPs as peptide biologics

Biologics are a group of compounds with a wide range of medicinal functions.¹⁰⁸ Examples of biologics include sugars, proteins, peptides, and nucleic acids, as well as living tissue and stem cells. Among the different kinds of biologics, peptide biologics have drawn increasing attentions for their high affinity towards therapeutic targets and low toxicity resulting in fewer side effects.

However, developing peptide biologics as therapeutics has been a challenge because of their poor physiological stability, which results in short half-lives, low membrane permeability and difficulties for oral delivery. Compared to subcutaneous injection, oral administration of peptide drugs is often preferred because it is non-invasive, painless and convenient. Unfortunately, most peptide and protein biologics possess low oral bioavailability as they are easily denatured by heat, acid, and susceptible to proteolytic degradation. Thus, orally ingested peptide drugs are hardly able to reach the gastrointestinal (GI) tract for absorption. Moreover, the GI mucus hampers absorption of peptide and protein drugs. As a result, most peptide and protein biologics are often administered by injection, which limits their therapeutic application. Developing peptide biologics with the stability of small molecules and the specificity of protein biologics would fill the chemical space in between these two groups of therapeutics and serve as promising drug candidates.

In order to increase the stability and improve oral bioavailability of peptide biologics, many have attempted to utilize protein engineering to develop stable peptide biologics by rational design such as side chain cross-linking (stapled peptides¹⁰⁹), backbone modification by alkylation¹¹⁰ or introducing unnatural amino acids (substance P analog GR71251¹¹¹). Naturally occurring peptides with exceptional

stability, such as CRPs or more commonly cyclotides, provide some direction for new approaches including backbone cyclization and grafting bioactive peptide epitopes into highly stable peptidyl scaffolds. By introducing additional conformational constraints, peptide stability is improved as well as the receptor binding affinity which results in higher potency.

Backbone cyclization of peptides is known to confer peptides resistance against exopeptidases. Additionally, introducing conformational constraints through cyclization increases the receptor binding affinity and ligand selectivity. Examples include anti-microbial peptides,¹¹²⁻¹¹⁴ human salivary peptide histatin 1 (Hst1),¹¹⁵ cyclic α -conotoxin MII¹¹⁶ and Vc1.1¹¹⁷.

In 2000, Tam and co-workers first reported the cyclization of several anti-microbial peptides such as tachyplesins TP, protegrins PG and α -defensins NP. TP and PG are cationic peptides with two intramolecular disulfide bonds linking two anti-parallel β -sheets. By engineering, two ends in the peptide backbone were linked as an amide without additional linkers. The cyclic analogs bearing zero to three disulfide bonds showed similar or enhanced anti-microbial activity and cytotoxicity. Especially for those containing three disulfide bonds, the membranolytic selectivity was improved by 6- to 30-fold against tested microbes while the hemolytic effect decreased by 10-fold.¹¹⁴ Interestingly, activity of the cyclic analogs increased under high-salt conditions under which most anti-microbial peptides displayed greatly reduced efficacy. These results demonstrate that the backbone cyclization can be employed to increase the membrane selectivity of AMPs and specially reduce hemolytic effects without compromising anti-microbial efficacy. α -Defensin NP-1 is a 33-residue AMP with three anti-parallel β -sheets linked by three intramolecular disulfide bonds (**Figure 1.5**). In engineered analogs, one disulfide bond was replaced by an end-to-end cyclic backbone, which simplified the oxidative folding process. These new molecules maintained the anti-microbial activity but were less cytotoxic and salt-insensitive. Thus, these results demonstrate that the loss of conformational constraints due to deletion of disulfide bonds can be compensated by cyclic backbones while retaining the structural rigidity. Additionally, end-to-end cyclization creates continuous epitopes with charge clusters, conferring higher potency of anti-

microbial activity to original peptides. Together, the work by Tam and his co-workers suggests that cyclization can be used not only to increase stability but also to confer new functions to the desired peptide for drug applications.

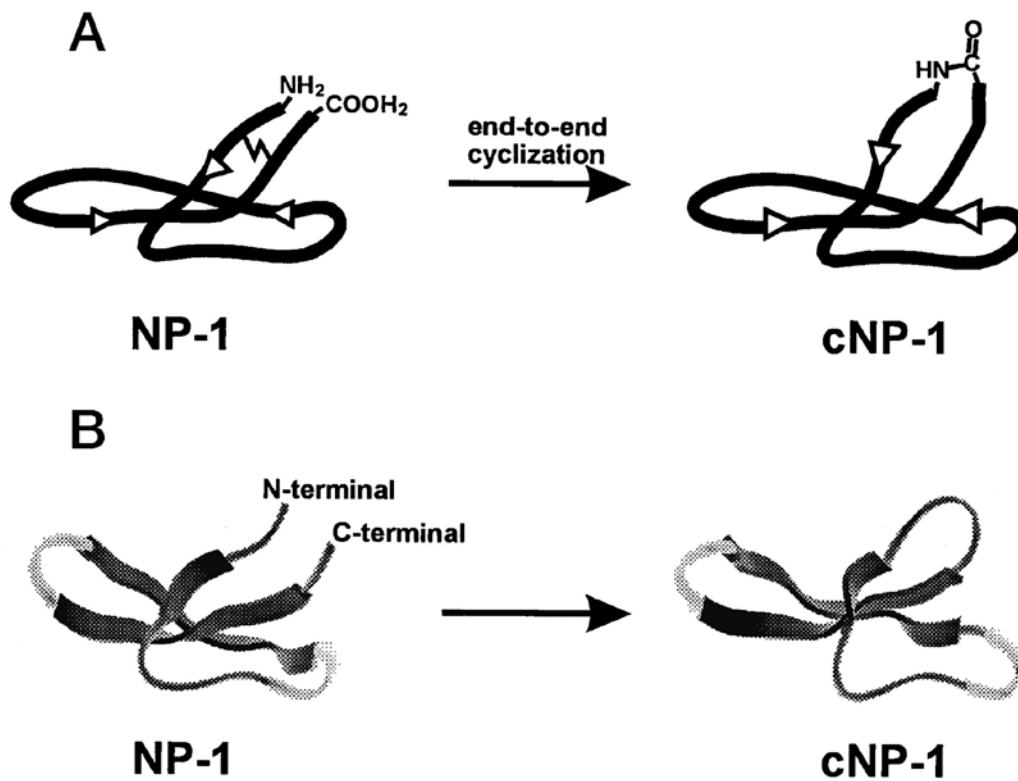


Figure 1.5 A schematic demonstration and modeling structure of cyclic NP-1. A, schematic strategy for cyclization of α -defensin NP-1.¹¹² In NP-1, the N- and C-termini are free amine and carboxylic acid. In cNP-1, the two ends are joined by an amide bond and one pair of disulfide bond is replaced by Gly; **B,** molecular modeling of NP-1 and cyclic NP-1 highlighting three β -strands. Disulfide bonds are not shown. Models are simulated based on the known structures of NP-1 and NP-2 using MacKerell protein potential parameter set.

Another example is human histatin 1 (Hst1), a 38-residue defensive peptide in saliva and a major factor for accelerating wound-healing. Thus, it is promising to develop synthetic histatins into novel therapeutic agents for the wound-treatment. To achieve improved clinical performance, backbone cyclization has been employed to increase its peptide stability and biological potency. Upon cyclization, the wound closure activity of Hst1 increased >1000 fold relative to its linear counterpart.¹¹⁵ Interestingly, the NMR and CD studies did not show obvious conformation constraints on the cyclic Hst1 compared with the linear form. It was hypothesized that structural differences due to cyclization only became manifest when these peptides were located in the binding domain of the receptor.¹¹⁸

Another example involves the cyclization of α -conotoxins using a series of linkers with various lengths (**Figure 1.6**). α -Conotoxins are a group of short peptides with two disulfide bonds and a typical α -helical structure. They have drawn much attention as potential drug leads for neuropathic pain treatments due to their potent nicotinic acetylcholine receptor antagonist activity. Two members from the α -conotoxin family have been successfully cyclized by a six-residue or seven-residue peptide linker. Cyclization dramatically protects the peptides from proteolytic degradation in serum and retains their nicotinic acetylcholine receptor antagonist activity. In the case of α -conotoxin Vc1.1, cyclic analogs possessed enhanced selectivity and potency for GABA_B receptor-mediated N-type Ca²⁺ channel currents compared to their linear counterparts, thereby being more potent on inhibiting neuropathic pain. Improved metabolic stability was also acquired by cyclization, which made cyclic Vc1.1 orally active. It should be noted that the influence of the peptide linker (composed of Gly and Ala) was not fully investigated and whether it contributes to the receptor binding affinity is not clear yet. Since the linkers are relatively long (six or seven residues) compared to the full length peptide (16 residues), the possible side effects of linkers is non-negligible and requires further study.

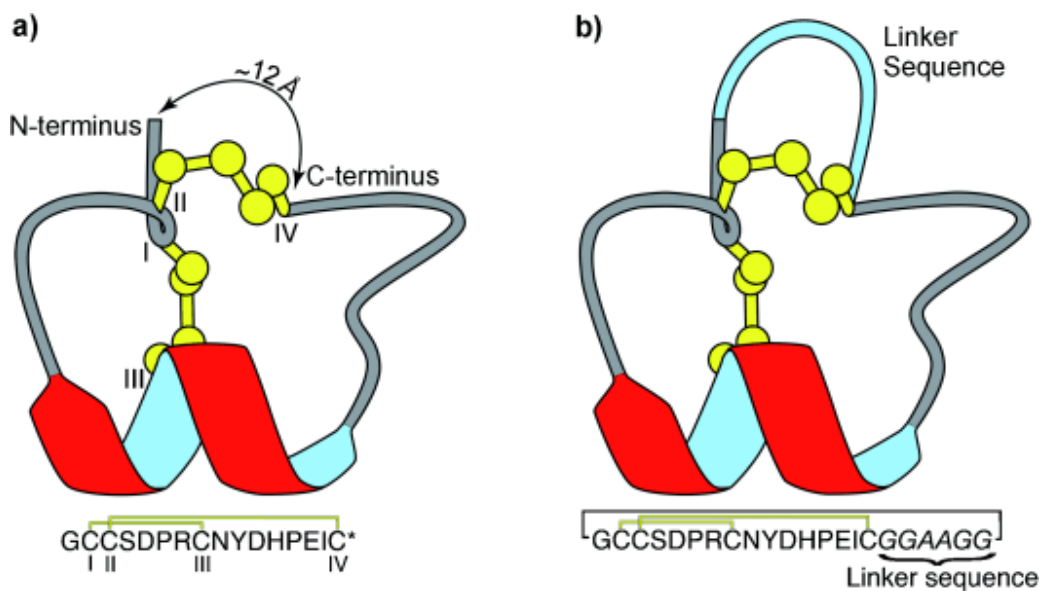


Figure 1.6 A scheme showing the original and engineered α -conotoxin Vc1.1. (Cited from ref ¹¹⁷) (a) Schematic structure and sequence of linear α -conotoxin Vc1.1 containing a α -helix (Red) and two disulfide bonds (Yellow). The distance between the N- and C- termini of Vc1.1 is estimated to be 12 Å. (b) Schematic structure and sequence of cyclic α -conotoxin Vc1.1 containing a peptide linker GGAAGG. The linker is highlighted by a blue line in the schematic drawing.

Grafting bioactive peptide epitopes into highly stable peptidyl scaffolds is another promising strategy for peptide engineering (**Figure 1.7**). Promising stable peptide scaffolds include cyclotides which are naturally occurring, ultra-stable peptides with membrane permeability with a wide range of valuable biological activities. Replacing some hydrophobic residues of kB1 with charge residues did not affect the overall cyclotide framework, and the engineered analog retained a similar conformation to the native kB1.¹¹⁹ Thus it further suggests the wide applicability of cyclotides as drug scaffolds.

Many laboratories have demonstrated this grafting strategy using stable CRPs such as toxins and cyclotides as scaffolds.¹²⁰⁻¹²⁶ In 1995, Vita and co-workers successfully grafted a metal binding site onto the scorpion toxin charybdotoxin. Nine residues were introduced into the 37-amino-acid sequence by chemical synthesis without modifying the original peptide conformation. The newly-produced molecule achieved high affinity towards Cu^{2+} with a $K_d=4.2 \times 10^{-8} \text{ M}^{120}$. Later, the same group developed a CD4 mimetic by substituting a short region of the scorpion toxin scyllatoxin with part of the CD4 sequence. The resulting chimeric peptide bound to HIV-1 envelop protein gp120 with CD4-like affinity and inhibited HIV-1 entry with IC_{50} at low micromolar to nanomolar concentrations.¹²¹ The 27-amino-acid mini-protein successfully mimicked CD4 (>40 kDa) and blocked HIV-1 entry efficiently without causing potential autoimmune responses.¹²⁷ This mini CD4 is now used as a commercially available alternative to full-length CD4 in various *in vitro* assays by many laboratories for HIV studies.

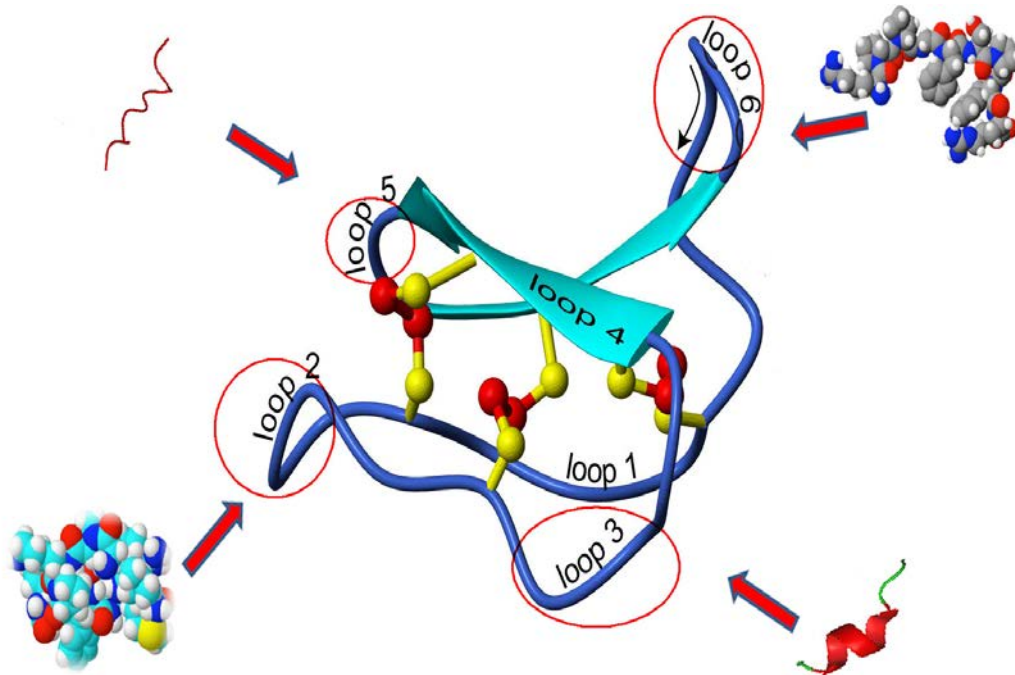


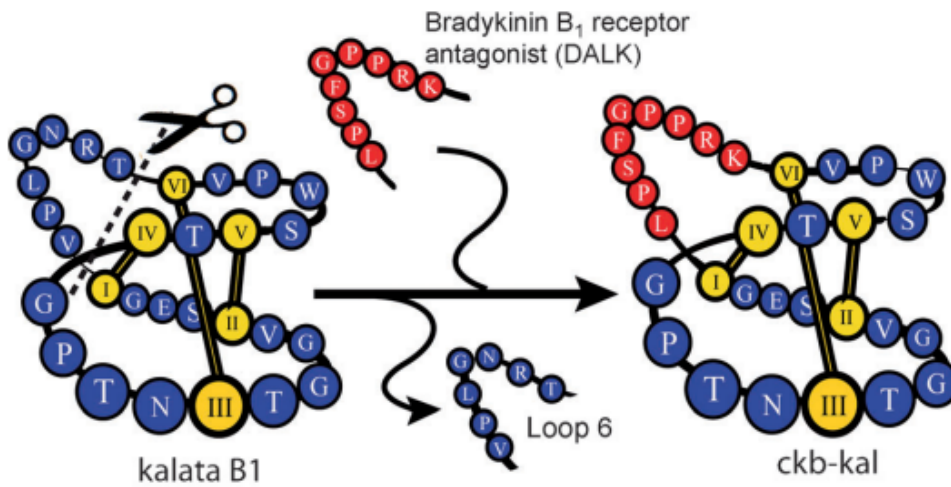
Figure 1.7 A scheme demonstrating the grafting of bioactive peptides into cyclotide scaffolds. The comparably flexible loop 2, 3, 5, 6 are amendable for inserting of or replacing with foreign peptides and sequences.

Table 1.3 Selected examples demonstrating the grafting strategy using cysteine-rich peptide scaffolds.

| Scaffold | Targets | Activity | Potency | Ref. |
|----------|---|------------------------------------|------------------------------|------|
| kB 1 | VEGF-A antagonist | Anti-cancer | IC ₅₀ =12 nM | 123 |
| | Melanocortin 4 receptor | Anti-obesity | K _i =29 nM | 128 |
| | Neuropilin-1 and -2 antagonist | Anti-tumor | EC ₅₀ =100 nM | 129 |
| | Bradykinin B ₁ receptor antagonist | Anti-chronic and inflammatory pain | p.o.=10 mg/kg | 125 |
| | MOG epitope (CNS protein) | Immune response modulation | s.c.=200 µg per time | 130 |
| MCoTI-I | CXCR4 antagonist | Anti-HIV | EC ₅₀ =20 nM | 124 |
| | Hdm2/HdmX antagonist | Anti-tumor | K _d = 2.3 ±0.1 nM | 131 |
| MCoTI-II | β-tryptase inhibitor | Anti-inflammation | K _i =1 nM | 132 |
| | α _v β ₆ binding peptide | Pancreatic cancer detection | K _d =3 nM | 133 |
| | Thrombin inhibitor | Cardiovascular disease treatment | K _d =670 nM | 134 |
| RTD-1 | RGD motif | Anti-platelet aggregation | IC ₅₀ =18 nM | 135 |

Recent applications of the grafting strategy using cyclotides as scaffolds are summarized in **Table 1.3**. A wide range of potential targets for drug development have been modified using the cyclotide framework. Our laboratory has developed an orally active peptide analgesic by grafting a bradykinin B₁ receptor antagonist into a proteinaceous natural product scaffold (**Figure 1.8**).¹²⁵ The bradykinin B₁ receptor stimulates chronic inflammatory responses; thus, their antagonists are believed to be useful for treatment of inflammatory pain.¹³⁶ Linear antagonists have limited clinical applications since they are easily digested by proteases. Two antagonists, DALK (nine residues) and DAK (seven residues), were grafted into the kB1 scaffold by replacing the loop 6 of it. The engineered bradykinin antagonists had superior bioavailability and displayed significant inhibition of pain response in animal models with oral administration compared to its linear analogs. These results suggest that the cyclotide scaffold confers impressive stability to the linear antagonist for survival in the GI tract and that the cyclotide framework is important for stronger oral activity.

Naturally stable CRPs demonstrate promising therapeutic potential, not only as active peptide biologics, but also as peptide scaffolds. As more and more peptide-based therapeutics emerge on the market, their applications as scaffolds for peptide therapeutics will enable greater advances in drug design.



| Peptides | Agonist/Antagonist | Sequences |
|---|---------------------------|------------|
| Bradykinin | B ₂ agonist | RPPGFSPFR |
| Kallidin | B ₂ agonist | KRPPGFSPFR |
| Kinestatin | B ₂ antagonist | QIPGLGPLR |
| des-Arg ¹⁰ -[Leu ⁹]-Kallidin(DALK) | B ₁ antagonist | KRPPGFSPL |
| des-Arg ⁹ -Kinestatin (DAK) | B ₁ antagonist | QIPGLGPL |



Figure 1.8 A scheme showing the design and sequence of bradykinin antagonists and the grafted peptide ckb-kal and ckb-kin. (Cited from ref ¹²⁵) The two bradykinin B₁ receptor antagonists DALK and DAK were grafted in the loop 6 of cyclotide scaffold kB1. The engineered bradykinin receptor antagonists are end-to-end cyclic with three disulfide bonds. The sequences of linear antagonists are highlighted in red and blue. Cysteines are highlighted in yellow.

5. Amide-to-amide chemoenzymatic synthesis of cyclic peptides

Macrocyclic peptides are ubiquitous in nature. However, the mechanism of how they are produced is poorly understood. Various organisms, such as fungi, bacteria, plants, and mammals, possess the machinery for generating peptide macrocycles. So far, little is known about their mechanisms of production and the cyclization process seems to differ from species to species. Thus, direct expression of cyclic peptides through conventional transgenetic approaches remains a challenge.

Ribosome-derived peptide macrocycles are processed by transpeptidases, which catalyze cleavage of the precursor sequence, followed by ligation of N- and C-termini to form a cyclic backbone. The transpeptidation reaction in this process involves the breaking and formation of amide bonds, thus we refer to it as the “amide-to-amide” cyclization. Currently, TraF,¹³⁷ PatG¹³⁸ and PCY1¹³⁹ are the only naturally occurring cyclases that have been isolated and characterized so far. TraF mediates the cyclization of Trb C by cleaving the C-terminal tetrapeptide precursor and facilitating amide bond formation between the N- and C-termini. Trb C is a 78-residue cyclic pilin subunit in bacteria and the largest cyclic peptide ever found. PatG mediates cyclization of much smaller peptides cyanobactins (6-10 amino acids) from cyanobacteria. PCY1 is the only cyclase isolated from eukaryotes. It is involved in the two-step biosynthesis of plant cyclic peptide Caryophyllaceae-type CP. However, none of these three cyclases have been exploited for chemoenzymatic synthesis of cyclic peptides because of their slow turnover rates and low catalytic efficiencies. Instead, other transpeptidases, such as subtiligase,¹⁴⁰ intein¹⁴¹ and sortase A¹⁴², which are not natural cyclases, have been successfully applied for peptide macrocyclization. Here, we will discuss the catalytic mechanisms and applications of intein- and sortase-mediated cyclization.

5.1 Intein-mediated synthesis of cyclic peptides

Intein is an intervening protein domain catalyzing protein splicing. This process involves excision of intein from the protein precursor and ligation of two flanking regions originally separated by intein, which is similar to the RNA splicing process. Like self-splicing introns, inteins are also auto-catalytic regions embedded in a

precursor sequence. So far, several hundreds of inteins have been identified among various unicellular organisms and they are highly conserved in sequence.¹⁴³

A common mechanism shared by all intein-mediated protein splicing is basically a similar process to the chemical ligation.¹⁴⁴⁻¹⁴⁶ (**Figure 1.9**) The entire mechanistic process can be characterized by four acyl shift reactions: first, an N-S/O acyl shift reaction occurs at the N-terminus of intein to afford a thioester/ester between the intein and N-extein. Second, the Cys/Thr/Ser located at the N-terminus of C-extein will initiate a nucleophilic attack on the newly formed thioester/ester. Through an S-S/O-O acyl shift, a branched intermediate is generated, which links the N-extein and C-extein. Third, intein mediates a self-cleavage from the intermediate by forming a succinimide through an N-N acyl shift. Fourth, a new amide bond is formed through a proximity-driven S/O-N acyl shift, resulting in a ligated protein.

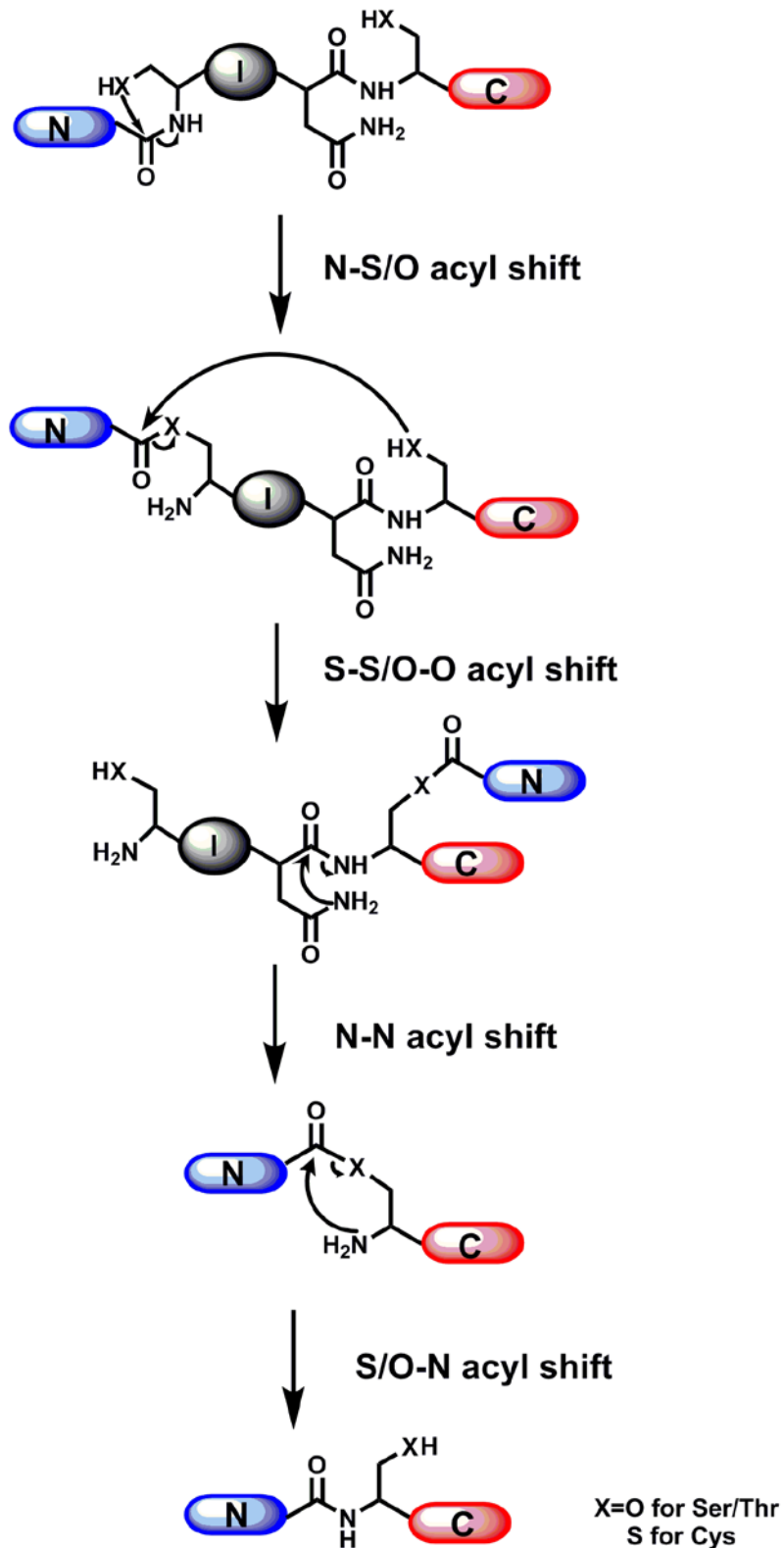


Figure 1.9 Mechanism of intein-mediated protein splicing. N-extein (blue) is labeled as N; intein (grey) is labeled as I; C-extein (red) is labeled as C.

Muir and co-workers first exhibited the application of intein-mediated splicing in a bacterial expression system for ligation reactions in 1998.¹⁴⁷ The ligation of a recombinant protein and a synthetic peptide with chemical modifications was achieved using the so-called expressed protein ligation system. In this system, an engineered intein was utilized which did not contain the endonuclease domain. The resulting intein molecule was smaller than the naturally occurring intein for expression convenience but still possessed the ability to mediate the first N-S acyl shift reactions to generate a C-terminal protein thioester. For purification purposes, the C-terminus of engineered intein was linked to a chitin-binding domain, which facilitated the separation of intein and the target protein thioester on a chitin column. Moreover, the C-terminal Asn of intein was mutated to Ala to prevent its self-cleavage from the chitin-binding domain. After expressing the target protein with the engineered intein, the fusion protein underwent an N-S acyl shift reaction catalyzed by the engineered intein. Subsequently, thiolysis by a large excess of external thiols resulted in a protein with a C-terminal thioester, which was readily captured by the N-terminal Cys of the synthetic peptide to form the ligation product.

This ligation scheme also applies to cyclization when the target peptide contains an N-terminal Cys. In this case, the N-terminal Cys will capture the C-terminal thioester intramolecularly to form a cyclic backbone. Muir and Camarero first applied this approach to cyclize a 57-residue Src homology 3 domain with enhanced biological activity in 1999.¹⁴⁸ For CRPs, not only the N-terminal Cys, but also the internal Cys will capture the C-terminal thioester to form thiolactones since multiple cysteines are present in the sequence. This is known as the thia zip reaction first proposed by Tam and co-workers in 1998.^{149,150} This process results in low yields in the *in vivo* intein-mediated cyclization due to the aggregation of thiol-thiolactone intermediates. Later, it was shown that *in vivo* intein-mediated synthesis of a peptide thioester followed by an *in vitro* cyclization under optimized conditions was a good solution since the cyclization reaction does not require intein catalysis. Using this approach, several cyclic CRPs have been successfully prepared by the intein-mediated cyclization, including cyclotides (kB1¹⁵¹ and MCOTI-II¹⁵²), θ -defensin RTD-1¹⁵³ and SFTI-1¹⁵⁴.

5.2 Sortase-mediated synthesis of cyclic peptides

Sortases are a group of transpeptidases discovered in bacteria that catalyze the anchoring of surface proteins to bacterial cell walls. Among them, sortase A is the most well-known, which was isolated from a Gram-positive bacterium *Staphylococcus Aureus* in 1999.^{155,156} This enzyme specifically recognizes a LPXTG sequence in bacterial surface proteins, also known as the sorting signal, and cleaves the bond between Thr and Gly to form a thioacyl enzyme intermediate. The N-terminal free amine of the oligoglycine peptide on the cell wall initiates a nucleophilic attack on this intermediate forming a new amide bond between the Thr and the oligoglycine (**Figure 1.10**). The proposed mechanism shows certain similarity with the proposed biosynthesis mechanism of cyclotides. First, both sortase A and AEP are cysteine proteases, which contain a Cys in the active site. Second, both involve transpeptidation reactions to break an amide bond in the substrate and form a new amide bond between the substrate and the N-terminal amino acid. Third, both recognize certain residues or motifs to initiate the transpeptidation reaction.

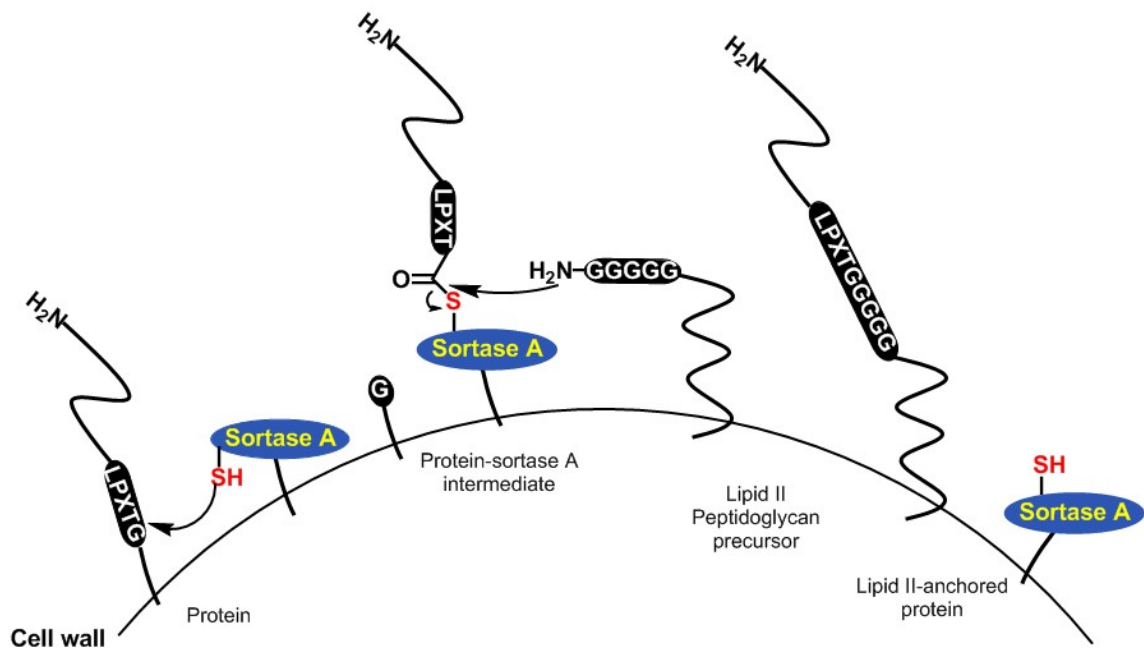


Figure 1.10 A scheme showing the Sortase-catalyzed anchoring of surface proteins to the bacterial cell walls. (Adapted from ref ¹⁵⁷)

The sortase-mediated chemoenzymatic ligation reaction has been exploited extensively in protein engineering for installations of various biological moieties into proteins. Ploegh and co-workers contributed a great deal of work in this field and demonstrated the feasibility of this approach for protein labeling, cell labeling, and incorporation of site-specific modifications.¹⁵⁷⁻¹⁵⁹ Although the intermolecular ligation reaction using sortase A has been exploited for a long time, intramolecular cyclization reaction was not investigated until 2009.¹⁴² Three structurally unrelated proteins (Cre Recombinase, eGFP, and UCHL3) were successfully cyclized by incubating a target protein containing an N-terminal Gly with sortase A in the absence of oligoglycine nucleophiles. These examples demonstrate the applicability of sortase-mediated cyclization of proteins (20-40 kDa). Smaller peptides, such as Hst1 (38 residues), kB1 (29 residues), α -conotoxin Vc1.1 (16 residues), and SFTI-1 (14 residues), have also been prepared by this method.^{118,160} Similarly, they must contain extra C-terminal LPXTG sequence and N-terminal oligoglycines. In addition, sortase-mediated intermolecular ligation was also employed to synthesize protein hydrazides¹⁶¹ and protein thioesters¹⁶², which are important building blocks in peptide cyclization reactions.

6. Amide-to-amide chemical synthesis of cyclic peptides

Besides chemoenzymatic approaches, cyclic peptides are more commonly prepared by chemical synthesis, which has the advantage of enabling modification and engineering of peptide macrocycles. Chemical synthesis of cyclic peptides provides access to investigate the structure-activity relationship of individual amino acids or the function of end-to-end cyclic backbone by mutating specific sequences or introducing perturbations in the peptide backbone. Moreover, in structural and functional studies, cyclic peptides from extraction and isolation are not reliable and sometimes lead to inconclusive results as they may contain other components from natural sources or bear undesired modifications during the *in vitro* purification process. In contrast, chemical synthesis produces pure and genuine compounds which provide more convincing and consistent results during various studies.

Chemical synthesis of peptides involves the formation of a peptide bond between an α -carboxylic acid and an α -amine in solution phase or solid phase. In this thesis, I

employed solid-phase peptide synthesis to synthesize linear and cyclic peptides chemically. In my approach, a transpeptidation scheme involving a series of acyl shift reactions is employed. It mimics the natural process of intein to break and form amide bonds, thus we refer it as the amide-to-amide chemical synthesis.

6.1 Solid-phase peptide synthesis

The chemical synthesis of peptides and proteins advanced in early twentieth century. Until 1950s, du Vigneaud and co-workers first reported the total synthesis of the naturally occurring 9-residue hormone oxytocin in solution phase.^{163,164} This is the first biologically active peptide that has ever been synthesized by chemical approaches. Ten years later, in 1963, Merrifield invented the revolutionary solid-phase peptide synthesis (SPPS) and published his work in the Journal of the American Chemical Society as the only author.¹⁶⁵ Since then, more than 500 bioactive peptides have been synthesized chemically in the following ten years and the SPPS has been shown to be a more efficient method compared with the solution-phase approach.

Unlike the solution-phase peptide synthesis, in the SPPS, all reactions are conducted on insoluble solid beads that is called resin and equipped with functional linkers. The cross-linked resins are made of polymers, such as polystyrene (PS), polyethylene glycol (PEG), and insoluble in reaction solvents. However, they can be swelled by the organic solvents to allow reagents to diffuse and react with the active sites. Thus, the peptide chain will be assembled on the resin rather than in the reaction solution. It greatly simplifies the laborious purification steps in solution-phase peptide synthesis that requires separation of products from the reaction solvents in each reaction step. In SPPS, the excessive reagents and all side products can be removed by washing the insoluble beads with reaction solvents and simple filtration, which improves the reaction yield as well.

The peptide chain is assembled in a C-terminus to N-terminus direction to avoid racemization, while in contrast the *in vivo* protein synthesis is on the N to C direction. All the amino acids' side chain containing functional groups such as $-OH$, $-SH$, $-NH_2$, $-COOH$ are protected temporarily during the peptide chain elongation. Since the reaction proceeds at a C-N direction, the N terminal amine of the amino acid is also temporarily protected before it is coupled to the C terminus of the existing peptide.

Therefore, the standard building block in SPPS is amino acids with N-terminal and side chain protecting groups and a free C-terminal carboxylic acid.

First, a protected amino acid is anchored to resin through the coupling reaction with the linker on the resin. The purpose of linkers is to link the synthetic peptide chain and protect the C-terminal carboxylic acid. Various linkers are available based on the resin types and the choice of the linker decides the C-terminal functional group in the synthetic peptide.

The coupling reaction is conducted under similar conditions as the solution-phase using coupling reagents in *N,N'*-dimethylformamide (DMF) or dichloromethane (DCM). Representative coupling reagent combinations include *N,N'*-dicyclohexylcarbodiimide (DCC), *N,N*-diisopropylcarbodiimide (DIC) with 1-hydroxybenzotriazole (HOBt), 7-Aza-1-hydroxybenzotriazole (HOAt) as additives, or (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP), *O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HATU) with *N,N*-diisopropylethylamine (DIEA). It is often achieved by activation of the inactive carboxylic acid into an active ester, which can readily react with the linker or the α -amine. The order of coupling efficiency is DCC/HOBt < DIC/HOBt << PyBOP < HATU.¹⁶⁶ An over-activation usually leads to racemization especially for the first amino acid.

Second, the amino acid-loaded resin is deprotected to remove the temporally N-terminal amine protecting group. The incoming protected amino acid will react with the newly free amine to form an amide bond, resulting in a dipeptide. The deprotection and coupling cycle is repeated until the desired peptide chain is assembled and immobilized on the solid support.

Third, the final cleavage reaction will remove the side chain protecting groups with concomitant release of unprotected peptide from the resin. The side chain protecting group and the resin linker are chosen to ensure the concomitant removal.

The two most common approaches of SPPS are Boc and Fmoc strategies. They represent two different kinds of N-amino protecting groups: *t*-butyloxycarbonyl and 9-fluorenylmethoxycarbonyl, respectively. Different deprotection and cleavage

cocktails are employed in these two methods. In Boc chemistry, the N-amino Boc protecting group is removed by 50% trifluoroacetic acid (TFA) under which the side chain protecting group is stable, whereas the final cleavage to remove the side chain protection and resin linkage is achieved by hydrofluoric acid (HF) or trifluoromethanesulfonic acid (TFMSA). Boc chemistry has been viewed as the most widely used method in SPPS until Sheppard and co-workers successfully adapted the Fmoc group to SPPS in 1978.^{167,168} The Fmoc protecting group was first developed by Carpino and Han for solution-phase synthesis in 1972¹⁶⁹ and it serves well for the mild deprotection purpose in SPPS approach. Instead of a strong acidic deprotection in Boc chemistry, Fmoc chemistry uses mild basic conditions usually containing 20% piperidine to remove the N-amino Fmoc protecting group, which makes it more suitable for synthesizing peptides carrying acid-sensitive moieties such as phosphate or glycan. Final removal of side chain protecting groups and cleavage of the peptide from the resin is performed with TFA in the presence of scavengers to prevent reversible alkylation.

The SPPS is the method of choice for peptide synthesis in this thesis due to its higher yields compared to solution phase peptide synthesis. The availability for automation and compatibility with unnatural peptide structures also make the SPPS a robust method to synthesize peptides.

6.2 Chemical ligation

Different from linear peptides, macrocyclic peptides has an end-to-end peptide backbone without N- or C-terminus. Thus, the preparation of peptide macrocycles requires a ligation reaction to link the two termini together after the linear precursor is synthesized by SPPS. The end-to-end macrocyclization is once considered as a challenge due to the great energy gap of bringing the N- and C- termini in proximity. Classic enthalpic ligation methods often involve partial or global protection of linear peptide precursor and a strong activation of the C-terminal moiety. The use of protection groups results in solubility problems and requires high dilution in the cyclization reaction. Thus, the reaction is usually performed in organic solvents to dissolve the protected peptides. Moreover, the strong activation of C-terminal moieties causes racemization of the C-terminal residue and oligomerization of

peptides. Due to these reasons, the preparation process is laborious and the yield is relatively low.

The emergence of entropic chemical ligation represents a breakthrough in ligation chemistry and makes the macrocyclization achievable. The difference between two ligation methods is that the entropic ligation overcomes the energy barrier by proximity-driven methods without over-activation of the C-terminal moiety. It is usually achieved through a proximity-driven O/S-N acyl shift to form an amide bond. Kemp and co-workers first introduced the entropic ligation concept to the peptide synthesis in 1981.¹⁷⁰ Using a prior thiol capture followed by an O-N acyl shift, they successfully ligated two unprotected peptide segments on an organic template. Because this O-N acyl shift was mediated by a 12-member-ring intermediate, the reaction rate was relatively slow. Chemoselective ligation methods using unprotected peptide segments in aqueous conditions began to emerge in 1990s. (**Figure 1.11**) The reaction exploits the super-nucleophilicity of $-OH/-SH$ groups to initiate a chemoselective capture on an electrophile that is normally an aldehyde or esters/thioesters without affecting the side chain functional groups. A five-member-ring acyl intermediate is then formed and spontaneously converts to a stable amide through a proximity-driven O/S-N acyl shift reaction.

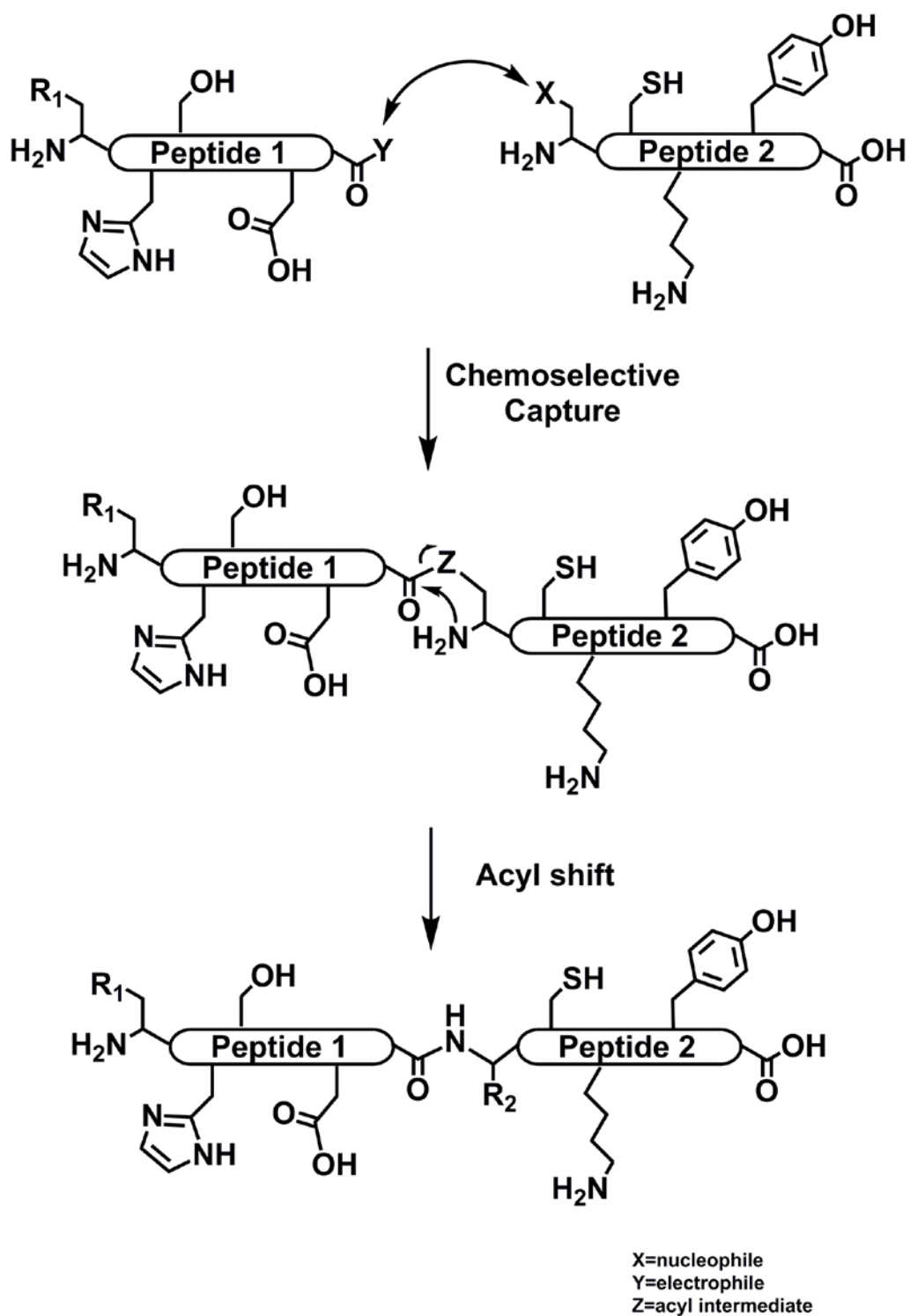


Figure 1.11 Schematic demonstration of entropy-driven chemoselective ligation.

In 1994, Liu and Tam reported a chemoselective ligation method using a peptide segment with N-terminal Cys/Ser/Thr and a C-terminal peptide segment bearing an aldehyde.¹⁷¹ This ligation method involved a 5-member-ring intermediate, which greatly accelerated the ligation rate. Parallely, Tam¹⁷² and Kent¹⁷³ reported a similar ligation reaction between an N-terminal cysteine and a C-terminal thioester which is known as Native Chemical Ligation (NCL) since it generates a native amide bond at the ligation site. They differ in the preparation of the C-terminal thioester: Tam and co-workers synthesized the peptide thioester directly using a mercaptoacetic acid (MPA) linker while Kent and co-workers prepared the thioester from a thioacid precursor.

In the ligation reaction, the free thiol group on the N-terminal cysteine captured the C-terminal thioester via a thiol-thioester exchange reaction to form a 5-member-ring intermediate. Subsequently intramolecular S-N acyl shift occurred spontaneously and a native peptide bond was generated. (**Figure 1.12**) No protection of side chain functional groups is required since this reaction is chemoselective.

The rapid capture between a thioester and a free thiol through the thiol-thioester exchange reaction was first reported by Wieland and co-workers, in which they described an amide bond formation in the reaction of Val-SPh and Cys.¹⁷⁴ This thiol-thioester exchange reaction is reversible and becomes the rate-determine step in the overall ligation reaction. The rate depends highly on the nature of the thioester as a leaving group. It is suggested that aryl thioesters are better leaving groups and lead to accelerated reaction rates in the NCL.^{172,175} Thus, excess of aryl thiols are usually added into the reaction mixture to facilitate the first thiol-thioester exchange and serve as radical scavengers and reductants as well. To best exploit the nucleophilicity of cysteinyl thiol side chain, the NCL reaction is usually performed at neutral aqueous conditions in presence of reducing reagents such as *tris*(2-carboxyethyl)phosphine (TCEP) to prevent disulfide bond formation.

The C-terminal amino acid in the peptide thioester also affects the ligation rate and yield. The ligation reaction proceeds faster when involving peptide thioesters bearing C-terminal amino acid with small side chain and low steric hindrance (Gly, Ala). In contrast, peptide thioesters containing steric hindered β -branched amino acid

(Ile, Val) at the C-terminus results in comparably slower reaction rate.¹⁷⁶ However, the slow rate of the most hindered Pro results from a different mechanism involving an $n-\pi^*$ interaction between the oxygen on the carbonyl group before Pro residue and the carbon atom on the carbonyl group of the C-terminal thioester.¹⁷⁷ These two atoms are brought to close proximity that favored the electronic interference, due to the backbone distortion by the *trans* conformation of Pro.

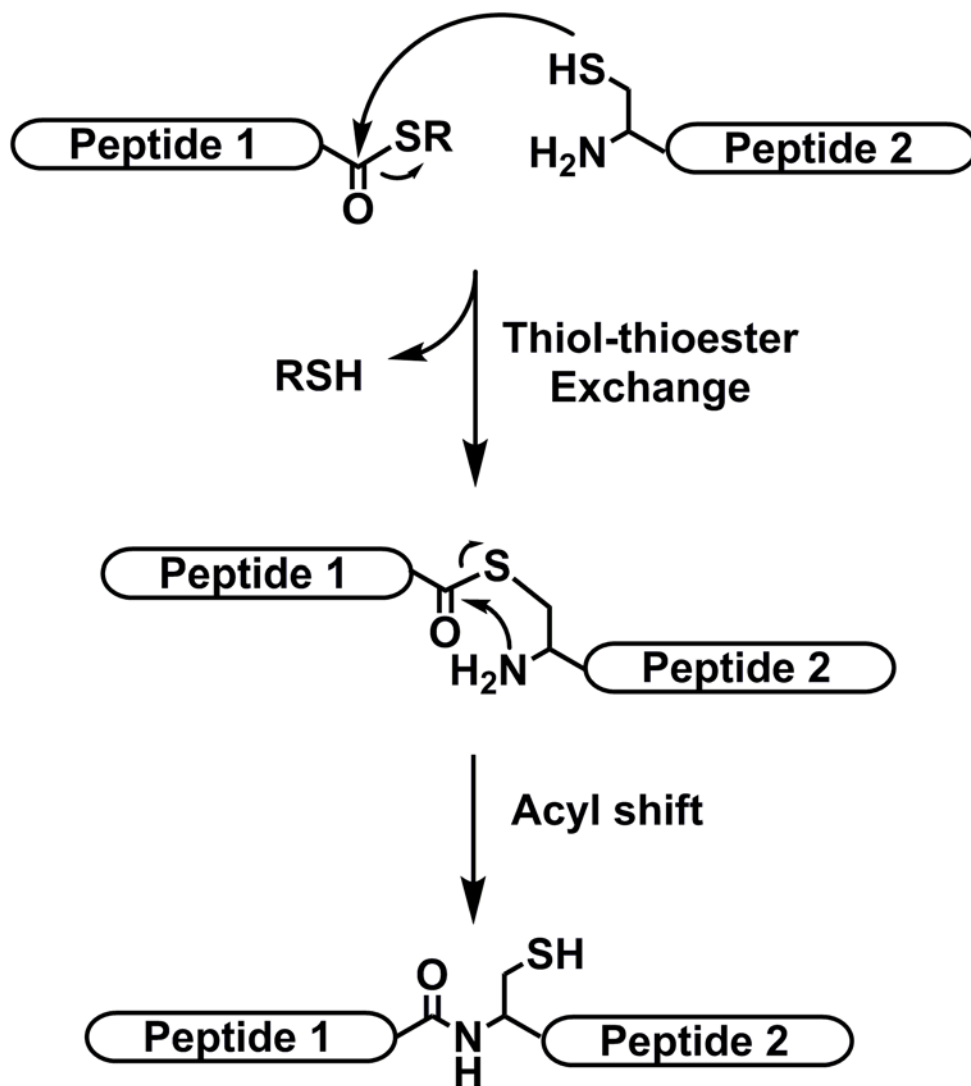


Figure 1.12 A scheme showing the mechanism of native chemical ligation.

Since a cysteine residue is regenerated at the ligation site in the thioester cysteine ligation method, it suits well for the synthesis of cyclic cysteine-rich peptides. With the peptide sequence possessing an N-terminal cysteine and a C-terminal thioester, an intramolecular ligation reaction occurs to produce an end-to-end macrocyclic peptide.

During 1997-1999, Tam and co-workers successfully synthesized various cyclotides including circulin A and B, kB1 and cyclopsychotride using Boc chemistry via the thioester cysteine ligation.^{149,150,178} It is the very first time that large cyclic peptides (ca. 30 residues) have been synthesized chemically. Camarero and Muir also successfully applied thioester cysteine ligation method to cyclize an unprotected peptide although with a smaller size (15 residues).¹⁷⁹ Then this method advances as the most useful tool and is adopted by other groups to synthesize a diverse range of cyclic peptides, including naturally occurring macrocycles like cyclotides, trypsin inhibitor MCoTI-I^{180,181} and engineered artificial peptide biologics.¹⁸² The chemical synthesis of Total chemical synthesis of cyclotides by Fmoc chemistry was first achieved by Leatherbarrow and co-workers using a sulfonamide linker which will be discussed in the “safety catch” section.¹⁸³ They synthesized a 34-residue trypsin inhibitor MCoTI-II with one-pot chemical cyclization and refolding.

Both Boc and Fmoc chemistry have been successfully applied to the synthesis of large cyclic peptides.^{90,184-188} The synthesis of macrocyclic peptides which was considered formidable is now readily achievable by chemical approaches due to the development of entropic chemical ligation methods. However, practical and efficient methods are yet to be developed for high-throughput production of cyclic CRPs.

Since the natural occurrence of cysteine is about 3%, most peptides do not contain cysteines for the cysteine thioester ligation. Ligation approaches involving other amino acids have thus been developed to expand the application and scope of thioester cysteine ligation. (**Table 1.4**) Most of the approaches mask the amino acid at the ligation site as a modified cysteine, and then regenerate the original residue by additional treatments such as methylation or desulfurization after the ligation. Currently, more than half of amino acids can serve as the ligation site for peptide ligation and cyclization.¹⁸⁹

Table 1.4 Summary of various ligation methods.

| Peptide segments | | Ligation site regenerated | Additional treatment after ligation | Ref. |
|-------------------------|-----------------------|---------------------------|-------------------------------------|---------|
| N-terminal | C-terminal | | | |
| Cys/Ser/Thr | ester aldehyde | Pseudo-Pro | none | 171,190 |
| Cys | thioester | Cys | none | 172,173 |
| His | perthioester | His | aryl disulfide-assisted ligation | 191 |
| Hcy | thioester | Met | s-alkylation | 192 |
| Cys | thioester | Ala | desulfurization | 193 |
| β -Mercapto-Phe | thioester | Phe | desulfurization | 194 |
| Penicillamine | thioester | Val | desulfurization | 195 |
| γ -Mercapto-Val | thioester | Val | desulfurization | 196 |
| Thiosugar Ser/Asp | thioester | Ser/Asp | removal of thiosugar | 197 |
| Cys | thioester | Pseudo-Lys | s-alkylation | 198 |
| <i>N</i> -Auxiliary Gly | thioester | Gly | deprotection | 199 |
| γ -Mercapto-Lys | thioester | Lys | desulfurization | 200 |
| Ser/Thr | ester salicylaldehyde | Ser/Thr | removal of <i>o</i> -auxiliary | 201-203 |
| Ser/Thr/Asn/Gly | thioester | Ser/Thr | Ag ⁺ -assisted ligation | 204 |
| Azido amino acids | phosphanyl thioester | Staudinger | modified staudinger reaction | 205,206 |

6.3 The thia zip assisted cyclization

Cyclizing large peptide macrocycles can be entropically unfavored due to the long distance (90 atoms for a 30-residue peptide) between the N- and C-termini. In 1998, Tam proposed the thia zip mechanism in macrocyclization of cysteine-rich peptides,^{149,150} during which the internal cysteines can help overcome the energy barrier and accelerate macrocyclization process through successive ring expansions (**Figure 1.13**). This reaction is initiated by an intramolecular thiol-thioester exchange reaction between the C-terminal thioester and the internal cysteine to form a thiolactone. Then a series of reversible thiol-thiolactone exchanges occur via the internal cysteines leading to successive ring expansion. Finally, when the N^α-amino thiolactone involving the N-terminal cysteine forms, the irreversible S-N acyl shift occurs spontaneously resulting in the formation of an amide bond. The thia zip reaction increases the macrocyclization rate by hundred fold compared with non-thia-zip cyclization (cyclization of peptides containing no internal cysteine).¹⁵⁰ The internal cysteines assist to form small thiolactone intermediates instead of a large 90-atom head-to-tail cyclic ring, which reduces the ring size and hence lowers the energy gap correlated with the cyclization. Therefore, the cyclization reaction can be performed with unprotected peptides at relatively high concentrations (3 mg/ml) in aqueous conditions. The thia zip cyclization overcomes the energy barrier and facilitates the synthesis of various naturally occurring and engineered large cyclic CRPs.^{96,112,183,207}

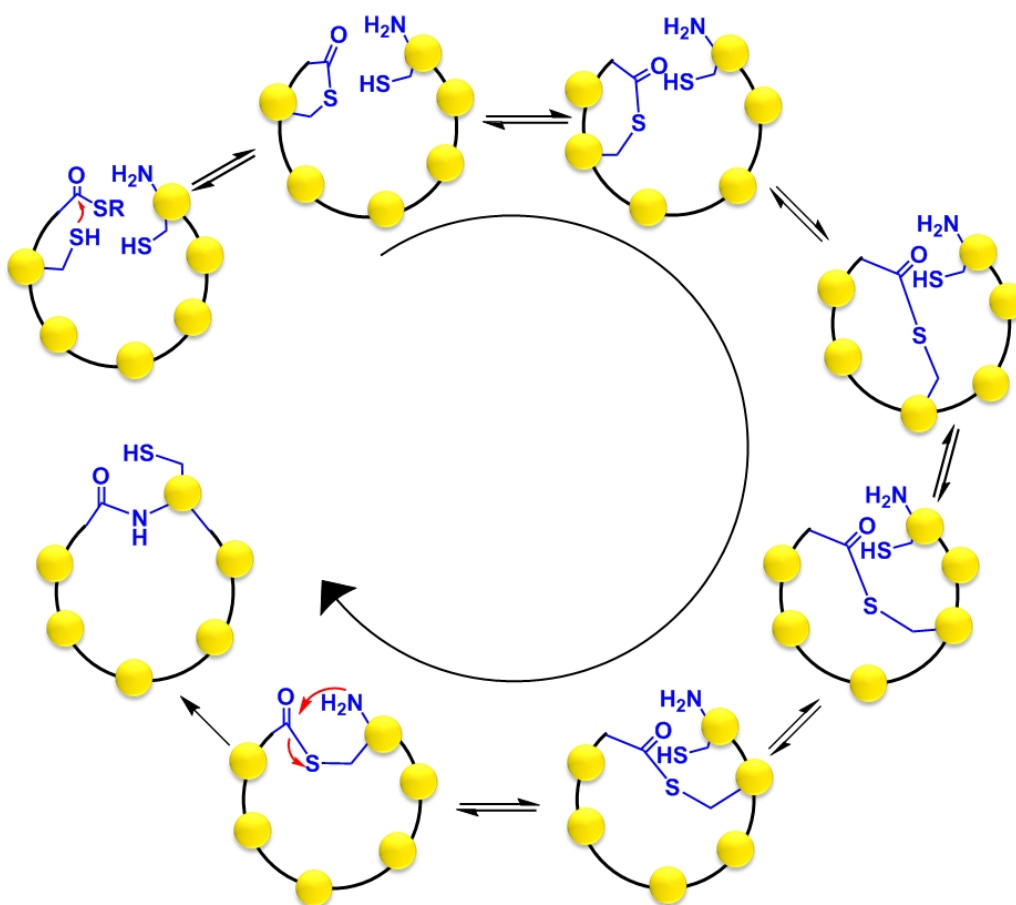


Figure 1.13 Schematic demonstration of the thia zip cyclization mechanism. The cysteine residues are highlighted in yellow dots. The reversible thiol-thioester and thiol-thiolactone exchange reactions lead to expansion of the ring size. The final S-N acyl shift irreversibly shifts the reaction equilibrium to the amide bond formation and generates the cyclic backbone.

6.4 Preparation of thioesters for cysteine thioester ligation

Thioesters are important building blocks for cysteine thioester ligation as well as other chemical reactions, especially in breaking and making peptide bonds. A practical method to prepare peptide thioesters efficiently thus becomes indispensable for the synthesis of peptide macrocycles. Direct syntheses of peptide thioesters are readily achievable by Boc chemistry using a $\text{SCH}_2\text{CH}_2\text{CO-MBHA}$ resin developed by Hojo and Aimoto²⁰⁸ and simplified by Zhang and Tam²⁰⁹. The peptide synthesis proceeds as the standard protocol and yields an unprotected peptide with the C-terminal *S*-alkyl thioester after final HF cleavage. As Fmoc chemistry advances to be the preferred method due to mild deprotection and cleavage conditions, there is increasing demand to synthesize peptide thioester by Fmoc chemistry. However, it is not possible to reproduce the Boc procedures in Fmoc chemistry since the thioester linkage is susceptible and intolerant to the piperidine treatment during the repetitive deprotection steps. Hence, there is a need to seek alternative methods for synthesizing C terminal peptide thioesters using Fmoc chemistry. Efforts have been devoted to develop Fmoc-compatible methods for the peptide thioester synthesis. Various approaches including using modified deprotection cocktails^{210,211}, post-synthesis thioesterification^{212,213}, and utilization of piperidine-resistant surrogate groups^{214,215} have been proposed. More methods employed the “safety catch” or “safety switch” strategy for preparing peptide thioesters. They share the same principle of employing thioester surrogates which are resistant to the piperidine deprotection during the peptide chain elongation process. However, the post-peptide assembly treatments to afford peptide thioesters are different in these two strategies.

6.4.1 The “safety catch” strategy

The “safety catch” strategy refers to those employing a stable thioester surrogate, which is the “safe” stage. Then by a simple activation, it can be transformed into a reactive functional moiety which will be subsequently caught by thiols to form a thioester via thiolysis, which is the “catch” stage. Sulfonamide is the prototype in this category, which was first developed by Kenner and co-workers in 1971²¹⁶ and modified by Ellman and others as versatile tools in biological and medicinal chemistry.²¹⁷⁻²¹⁹ In this approach, the *N*-sulfonamide linker served as a piperidine-resistant thioester surrogate during the peptide chain assembly. After that, selective *N*-

alkylation of the sulfonamide to an *N*-substituted sulfonamide made it vulnerable to thiolysis, which resulted in the unprotected peptide thioester after TFA cleavage. Various methods employing similar strategy have been developed to prepare peptide thioesters containing up to 35 amino acids.²²⁰⁻²²⁵ (**Figure 1.14**) Since the C-terminus of the peptide thioester is highly activated during the conversion process, the steric hindrance of C-terminal residue is not a problem in these methods. Various peptide thioesters with C-terminal Phe, Tyr, Val or Pro can be synthesized by safety catch methods. However, it also introduces some side reactions such as undesired alkylation or racemization during the activation step.

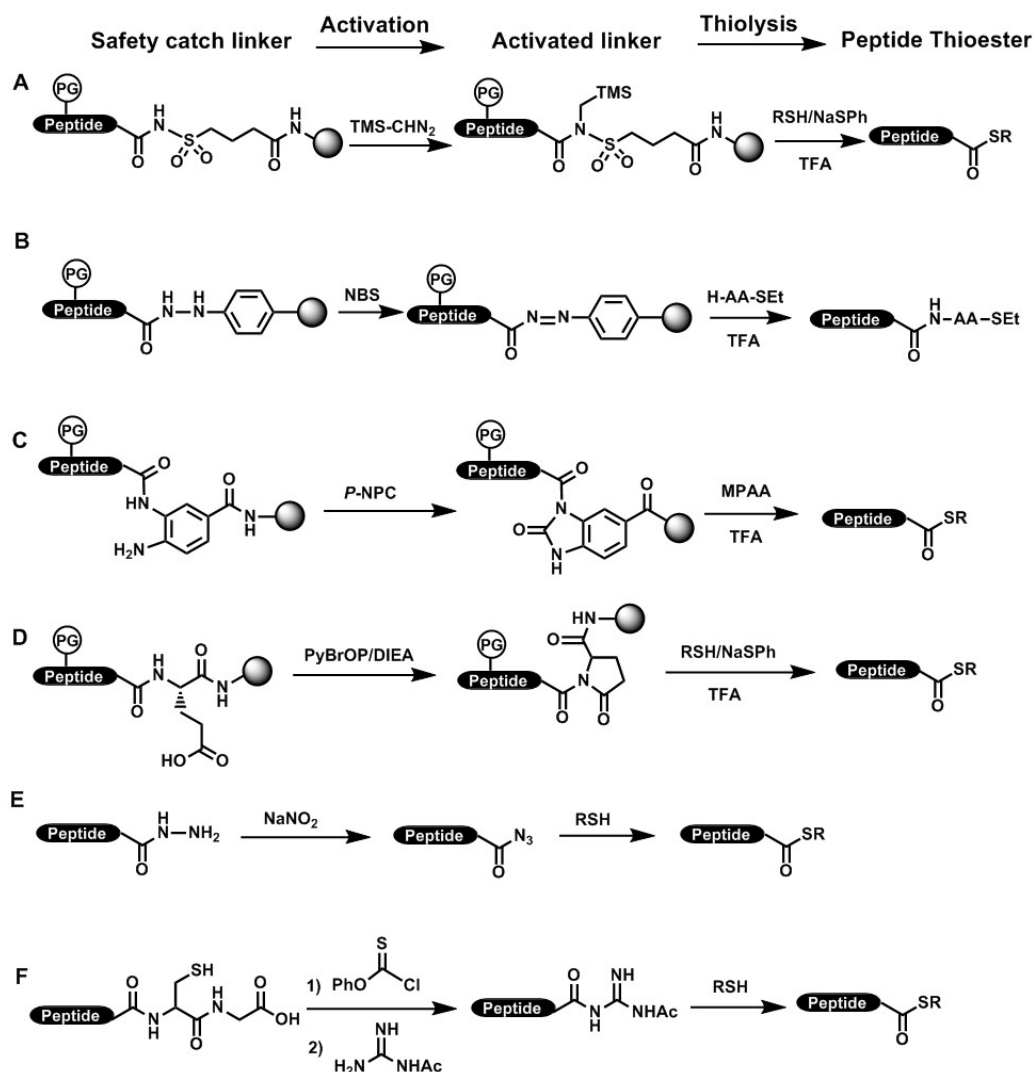


Figure 1.14 Selected examples of safety catch methods for peptide thioester synthesis. Safety catch linkers are first activated through various treatments, and then the activated linker will eventually be converted to thioester by thiolysis. **A.** Sulfonamide linker for peptide thioester synthesis.²²⁰ TMS-CHN₂: Trimethylsilyldiazomethane, (CH₃)₃SiCHN₂. **B.** Aryl hydrazine linker for peptide thioester synthesis.²²¹ NBS: *N*-Bromosuccinimide, C₄H₄BrNO₂. **C.** *N*-acylurea linker for peptide thioester synthesis.²²² *p*-NPC: *p*-nitrophenylchloroformate, ClCO₂C₆H₄NO₂, MPAA: 4-mercaptophenylacetic acid. **D.** Pyroglutamylimide linker for peptide thioester synthesis.²²³ PyBrOP: bromotripyrrolidinophosphonium hexafluorophosphate, C₁₂H₂₄BrF₆N₃P₂. **E.** Hydrazide linker for peptide thioester synthesis.²²⁴ **F.** Guanidine linker for peptide thioester synthesis.²²⁵

6.4.2 The “safety switch” strategy

The “safety switch” strategy represents a different mechanism in producing peptide thioesters compared with the “safety catch” strategy. Normally, the thioester formation in the “safety switch” strategy does not involve an alkylation and an intermolecular thiolysis, but employs an ester or an amide as the latent thioester which can undergo intramolecular O/N-S acyl shift reactions to afford peptide thioesters. Additionally, the acyl shift reaction is usually acid-catalyzed while basic conditions are more commonly used in the “safety catch” strategy. In “safety switch” methods, esters or amides (**Figure 1.15**) are linked to the resins as the C-terminus of target peptides using standard Fmoc SPPS procedures. After cleavage, the thiol functional moiety connected to the ester or amide linker will act as the nucleophile to attack the carbonyl group in the ester or amide. Peptide thioesters are subsequently formed in equilibrium through O/N-S acyl shift reaction. Eventually, the reversible thioester will be captured by external thiols RSH to afford stable thioesters.

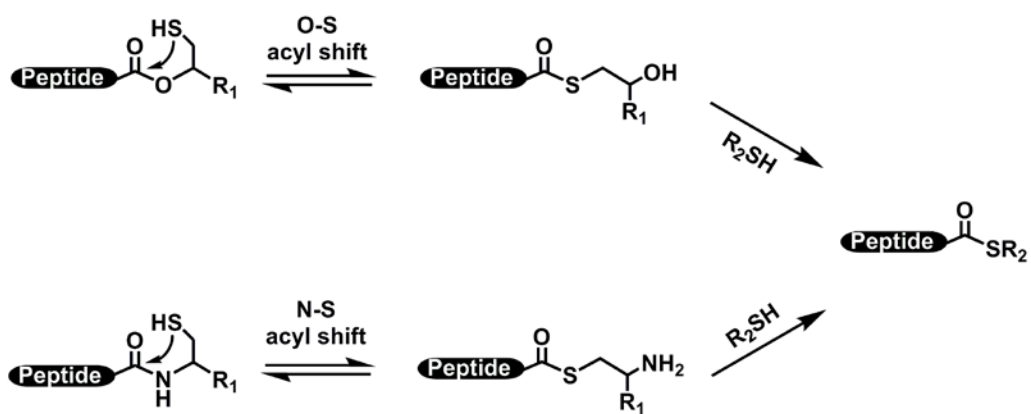
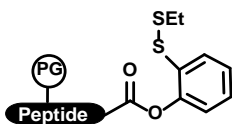
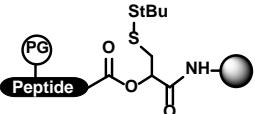
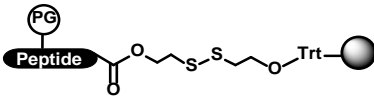
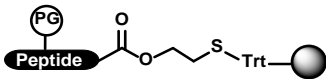
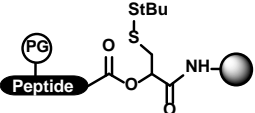


Figure 1.15 Peptide thioester formations through O/N-S acyl shift reactions. R₁ refers to any auxiliary groups. R₂SH represents external thiols.

Safety switch” strategies utilizing O-S acyl shift reactions were first reported by Danishefsky and co-workers in 2004.²²⁶ A disulfide protected mercaptophenyl ester was coupled to a fully protected pentapeptide acid by HATU/DIEA to afford the ester precursor in solution phase. Then the thiol nucleophile unmasked by reduction of the disulfide bond generated the thioester *in situ* via O-S acyl shift reaction, which was readily ligated with another peptide containing an N-terminal Cys. In the same year, Botti and co-workers described the peptide thioester method based on the O-S acyl shift reaction in solid phase.²²⁷ They synthesized a peptide ester bearing a disulfide protected thiol moiety at β position using standard Fmoc SPPS protocols. After cleavage, the peptide ester was used to ligate with an N-terminal cysteine fragment directly to generate the ligated peptide. The proposed mechanism suggested that under the reductive environment of the ligation reaction, the disulfide protection was removed, which initiated the O-S acyl shift reaction and generated a peptide thioester *in situ*. Subsequently, the newly produced thioester will be captured by the N-terminal cysteine fragment to allow the ligation reaction to occur. This method is simple and practical. No additional treatment is required after the first amino acid was coupled on the resin. However, the hydrolysis of the ester under neutral to basic conditions was significant (10-20%) which required large excessive N-terminal cysteine fragment in the ligation reaction and greatly reduced the ligation yield. To avoid the hydrolysis problem, our laboratory developed the hydroxyethylthiol resin, which generated peptide thioesters under acidic conditions containing TFMSA in TFA.²²⁸ Selected examples of methods based on O-S acyl shift reactions are summarized in **Table 1.5**.

Table 1.5 “Safety switch” methods based on O-S acyl shifts.

| Name | Structure | Condition | Ref |
|-----------------------|---|---|-----|
| Mercapto-phenyl ester |  | 1) TFA cleavage in solution phase 2) Direct NCL, excess MESNa, pH 7.4 | 226 |
| Mercapto-carboxyester |  | 1) TFA cleavage in solid phase 2) Direct NCL, excess PhSH, pH 6.5 | 227 |
| Dithiodiethanol ester |  | 1) TFA cleavage in solid phase 2) Direct NCL, 50 mM MESNa, pH 7 | 229 |
| Mercapto-ethylester |  | 0.1-0.2% TFMSA in TFA, 5% thiocresol | 228 |
| Alkyl oxoester |  | 1) 50mM TCEP•HCl, 100mM thiocresol, 1% DIEA in <i>t</i> BuOH/H ₂ O (95:5) 2) TFA cleavage | 230 |

Amides are more commonly employed in the “safety switch” strategy for the synthesis of thioesters. The amide-based thioester surrogates undergo an intramolecular N-S acyl shift reaction which mimics the protein splicing process to afford peptide thioesters, thus, switching from an amide to a thioester. The N-S acyl shift reaction occurred *in vivo* during the intein-mediated protein splicing but was not observed in Fmoc SPPS until early 2000. Different groups reported the observation of peptide amide and peptide thioester isomer with the same molecular weight after TFA cleavage of auxiliary groups.²³¹⁻²³³ The 4,5-dimethoxy-2-mercaptobenzyl (Dmmb) group was first developed by Vorherr and co-workers as a ligation auxiliary to facilitate the chemical ligation reaction at the Gly site in 2002.²³¹ During the removal of Dmmb group under acidic conditions, an isomerization was observed and suggested to be mediated by the N-S acyl shift reaction. Aimoto and co-workers confirmed the findings and developed the first thioester surrogate based on this mechanism in 2005.²³² They showed that the peptide containing a Dmmb group (N-peptide) would transform into a peptide thioester (S-peptide) under TFA cleavage. Later it was verified by C¹³ NMR, RP-HPLC and MS observation using a dipeptide model Fmoc-Gly(1-¹³C)-(Dmmb)Ala-OMe.²³⁴ The same type of isomerization was also observed by Danishefsky and co-workers when they used a similar 4,5,6-trimethoxy-2-mercaptobenzyl (Tmmb) auxiliary group to mediate cysteine-free ligation during synthesis of glycopeptides.²³³ Various methods have been then developed and representative examples are shown in **Figure 1.16**.

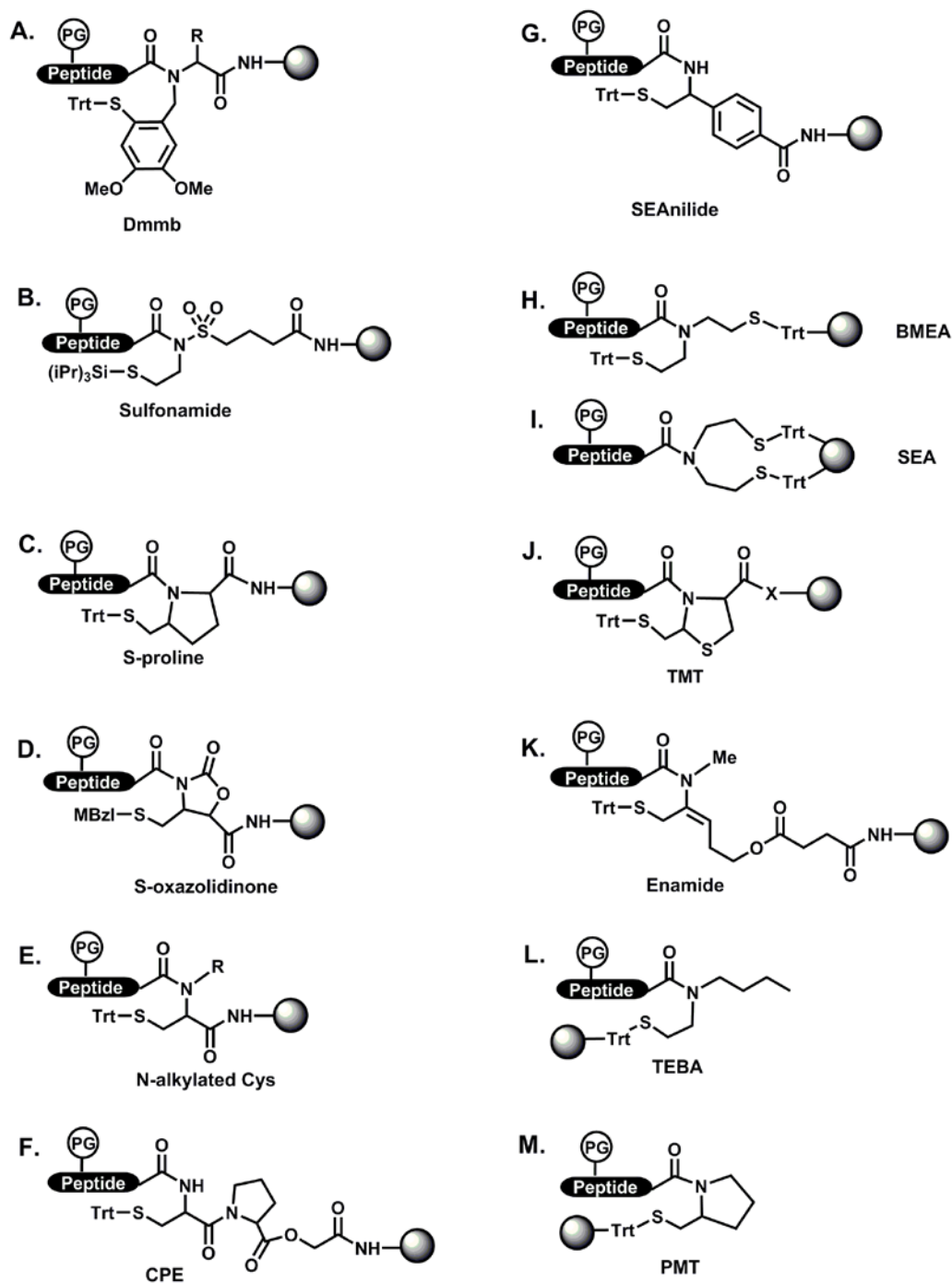


Figure 1.16 “Safety switch” methods based on N-S acyl shifts. Structures are rearranged to align the scissile peptide bond in the *cis* conformation to assist the nucleophilic attack. References for each example are: A. Dmmb²³², B. Sulfonamide²³⁵, C. S-proline²³⁶, D. S-oxazolidinone²³⁷, E. N-alkylated Cys²³⁸, F. CPE²³⁹, G. SEAnilide²⁴⁰, H. BMEA²⁴¹, I. SEA²⁴², J. TMT²⁴³, K. Enamide²⁴⁴, L. TEBA²⁰⁷, M. PMT²⁴⁵.

Hojo and co-workers reported a series of *N*-alkylated Cys linker as thioester surrogates. The alkylation at the amide nitrogen increased the ratio of *cis* amide bond to facilitate the N-S acyl shift and stabilized the thioester after the N-S acyl shift reaction. The reaction was achieved under acidic conditions (pH~1) with 5% mercaptopropionic acid at room temperature. Two-three days or longer was required for the completion of the reaction. Among all the analogs, *N*-methyl Cys and *N*-ethyl Cys gave highest yields while *t*butyl and Bn derivatives afforded lower yield due to the incomplete coupling of the following amino acid. This approach was applied to synthesize 95-residue chemokine CCL27²⁴⁶ and modified by Brik and co-workers in the total chemical synthesis of a ubiquitin thioester using *N*-methyl Cys as the thioester surrogate.²⁴⁷

Aimoto and co-workers developed an autoactivating unit CPE (cysteinyl proline ester) to induce N-S acyl shift reactions.²³⁹ Based on the diketopiperazine mechanism, the CPE unit trapped the newly formed thioester by an irreversible reaction. Thus, the thioester could not isomerize back to the amine form. Additionally, the diketopiperazine thioester can be converted to peptide thioesters by treating with excessive external thiols at pH 8.8, 37 °C or used directly in the chemical ligation reaction to form a ligated peptide in one pot.²⁴⁸ The proline could also be replaced by an *N*-substituted glycine to enhance the rate of the thioester formation.²⁴⁹

Liu's group and Melnyk's group independently reported a bis(2-sulfanylethyl)amino group named as BMEA and SEA, respectively. However, the way they prepared the linkers and their conditions for the thioester formation were different. Liu and co-workers applied this BMEA peptide for direct ligation with a cysteinyl peptide at microwave conditions pH 4-6 and successfully prepared a histone H3 protein. Melnyk and co-workers used reductive conditions at pH 7, 37 °C to conduct the ligation reaction. This thioester surrogate contained two thioethyl groups both of which could undergo the N-S acyl shift reaction regardless of the conformation of the scissile bond. However, the disulfide bond formation between the two thioethyl groups blocked the nucleophile and resulted in reduced yield and complicated product profiles. To overcome the disadvantage, Melnyk and co-workers converted the SEA group to thiazolidine thioesters by reacting with glyoxylic acid, or alkylthioesters through thiol-thioester exchange reaction for cysteine thioester

chemical ligation.^{250,251} It was also demonstrated that SEA group could be used for cysteine-free ligation through reactions with aspartic and glutamic acids to yield end-to-side chain or branched peptides.²⁵²

The N-S acyl shift reaction is also observed at cysteine residue although at an extremely slow rate (910 h for 90% completion in TFA)²³⁴. Efforts have been devoted to exploit the isomerization of cysteine residue and optimize the thioester formation. Besides *N*-alkylated Cys developed by Hojo and co-workers, Macmillan and co-workers utilized the labile linkage at Gly-Cys, His-Cys and Cys-Cys sites under heated acidic conditions to generate Gly, His or Cys peptide thioesters. They used increased temperature (60 °C) and large excess of MPA or MESNa (20%) to assist the N-S acyl shift reaction of cysteine.²⁵³⁻²⁵⁶ Recently, Offer and co-workers applied α -methylcysteine peptides directly in the ligation reactions in presence of large excess thiols. A 58-residue bovine pancreatic trypsin inhibitor (BPTI) and a 72 residue interleukin analog murine KC were successfully prepared by this method and it was fully compatible with Hmb and other protecting group, which would find great value in synthesizing peptides or proteins containing post-translational modifications.²⁵⁷

7. Oxidative folding of CRPs

Following macrocyclization, the oxidative folding completes the chemical synthesis of cyclic CRPs. The mechanism of protein folding, which involves formation of both the native three-dimensional structure and native disulfide bonds, was first explored in the 1960s by Anfinsen using ribonuclease as an example.^{258,259} This study revealed that the biologically active three-dimensional structure was determined only by the protein's amino acid sequence; the protein folding process was governed by a self-assembly principle. In the oxidative folding of peptides, the amino acid sequence itself may not contain sufficient information to direct a correct folding. As described earlier in the biosynthesis of mature peptides, the oxidative folding process usually occurs before the post-translational processing of precursor peptides and is facilitated by PDI, which was suggested to stabilize the native structure and assist with correct disulfide bond formation, respectively. During *in vitro* synthesis, the precursor peptide is missing and the enzyme is not involved. Thus, the oxidative folding process has to be optimized by adjusting experimental conditions.

Disulfide bonds play an important role in maintaining the native structure of CRPs and retaining their biological activity. The oxidative folding of CRPs mainly refers to the formation of correct disulfide bond connectivity, which stabilizes the secondary structural motifs and determines the native structure.

Disulfide bond formation is an oxidation reaction involving the formation of a covalent bond (S-S bond) between two sulfurs of two cysteine residues. Thus, peptides containing only one disulfide bond can be synthesized by simple oxidation reactions in neutral to basic conditions. Air oxidation is the simplest method to achieve the disulfide bond formation but has a comparatively long reaction time. To accelerate the oxidation process, strong oxidants such as iodine and $K_3Fe(CN)_6$ are utilized which reduces the reaction duration from days to minutes. The use of these reagents requires special handling otherwise they will result in the undesired oxidation of susceptible amino acids, such as Trp, Met, and Tyr. Tam and co-workers developed dimethyl sulfoxide (DMSO) as a mild oxidative reagent for disulfide bond formation.²⁶⁰ It led to efficient disulfide bond formation in aqueous buffers and has three advantages over other methods. First, the reaction was much accelerated compared with air oxidation, which required days to complete. Second, it was effective at a wide range of pH from acidic to basic, while other methods could only be performed at narrow pH ranges. This is especially useful for basic peptides which tend to aggregate and precipitate in neutral to basic folding buffer. Third, besides acting as an oxidant, DMSO is a good organic solvent which helps dissolve the hydrophobic peptides and serve as a structural chaperon assisting the disulfide bond formation.

For the oxidative folding of peptides containing multiple cysteines, direct oxidation is not suitable since there will be more than one isomer in the folding process (15 isomers for peptides with three-disulfide bonds). Thus, chemoselective oxidation methods were developed to avoid the formation of incorrectly bonded isomers. In chemoselective methods, the cysteine residues are selectively protected by orthogonal protecting groups. Utilizing stepwise deprotection, only one or two pairs of cysteine residues are allowed to form disulfide bond in each step, greatly reducing the number of possible isomers and leading to fewer isomers with incorrect disulfide linkages.

In 1999, the “2+1 strategy” was introduced by Tam’s group to synthesize the first cyclotides.³² In this strategy, two pairs of cysteines were selected to be protected by a HF-labile *S*-methylbenzyl group while the last pair of cysteines were protected by a HF-stable *S*-acetamidomethyl (Acm) group. The first two pairs of cysteine were deprotected after HF cleavage, while the Acm-protected cysteines were intact. Then, disulfide bond formation was conducted between the two pairs of cysteine and the correctly-folded product was collected to proceed. Acm-protecting groups were later removed by iodine/methanol, which allowed the last pair of cysteines to form a disulfide bond. This strategy reduced the number of isomers produced from 15 to three.

Another chemoselective method involves replacement of one pair of cysteines with selenocysteines. Since the redox potential of selenocystine (-381 mV) is much lower than the cystine (-323 mV) or the mixed selenocysteine-cysteine bond (-326 mV)²⁵⁸⁻²⁶³²⁶¹, diselenide bond formation is spontaneous and highly favored over the formation of a disulfide or the mix seleno-cys bond. Thus, it allows the chemoselective formation of a diselenide bond and increases the reaction rate as well. However, whether the introduction of unnatural amino acid selenocysteine is cytotoxic still requires further investigation.

Furthermore, due to the complicated reaction steps and unpredictable side reactions caused by iodine, the chemoselective strategy generally results in comparably low yields and is less frequently used in CRP synthesis now.

The global oxidation method has drawn much attention and has been used for folding of many CRPs successfully, such as cyclotides, conotoxins and defensins.^{185,262-267}

In this strategy, disulfide bonds are formed randomly, unlike the chemoselective strategy. Therefore, a redox pair like GSH/GSSG is required to form, break and shuffle the disulfide bonds. This process is mediated by thiol-disulfide exchange reactions between the redox reagents and the disulfide bonds in CRPs. This reaction is highly dependent on the proximity, reactivity, and accessibility of the thiol group and the disulfide bond.²⁶⁸ Thus, it is important for the peptide to adopt a native-like

conformation allowing the thiol groups to come into close proximity and form the correct disulfide bond. Additionally, most CRPs contain a buried cystine core in their native conformation which protects the correct disulfide bonds from shuffling with the redox reagents. In contrast, misfolded intermediates are often exposed to the surface and kept accessible to thiol groups until they form the native conformation with correct disulfide linkages, which is one of the most stable states.

This approach avoids the multiple purification steps needed in the chemoselective strategy, but cannot reduce the number of possible isomers formed. In order to improve yield of the correctly-folded peptide, oxidative folding conditions, such as peptide concentration, redox reagents, salt concentrations, temperature and addition of a co-solvent, have to be optimized.

Various global oxidative folding cocktails developed for different CRPs or proteins are listed in **Table 1.6**. Although they were invariably performed in aqueous buffered condition at pH 7.5-8.5 with GSH and GSSG as redox agents, there is no universal protocol for global oxidative folding. Several parameters such as salt concentration, additives and co-solvents are modified accordingly. In the oxidative folding of BPTI and ω -conotoxin MVIIC, salt concentration played an important role. High concentrations of Cl^- and SO_4^{2-} were added to distribute the highly positive charge of the peptides and prevent intermolecular aggregations.^{265,269} Temperature can also affect the folding efficiency to some extent. Kubo and Sakakibara showed that lower temperatures (5 °C) afforded better yield of correctly-folded peptides, while higher temperatures (21 °C) predominantly gave two incorrectly bonded isomers.²⁶⁵ For hydrophobic peptides, detergents and adjustment of pH can facilitate oxidative folding by minimizing hydrophobic interactions and intermolecular aggregations. The inclusion of urea or guanidine hydrochloride (Gnd-HCl) in slightly acidic conditions (pH 6) has been shown to greatly reduce polymerization and precipitation of human defensin and improved the folding yield.²⁶⁰

Table 1.6. Global oxidative folding conditions of CRPs or small disulfide-rich proteins.

| CRP | No of residues | Net charge at pH 7.0 | Hydrophilicity (%) ⁺ | Folding condition | Yield (%) | Rate ⁺⁺ (h) |
|--|----------------|----------------------|---------------------------------|--|-----------|------------------------|
| BPTI ^{269,270} | 58 | +5.7 | 33 | tris-HCl, pH 8.7, 200 mM KCl, 150 μ M GSSG | <80 | 168 |
| RNase A ²⁷¹ | 124 | +9.0 | 40 | tris-HCl, pH 8.0, 260 μ M DTT ^{OX} | -- | -- |
| Hirudin ²⁷² | 65 | -9.2 | 43 | NaHCO ₃ , pH 8.3, 60 μ M β -ME | 95-98 | 22 |
| kalata B1 ¹⁸⁴ | 29 | -0.3 | 17 | NaHCO ₃ , pH 8.5, 50% 2-propanol, 1 mM GSH | 40-50 | 48 |
| cycloviolacin O2 ¹⁸⁵ | 30 | +1.7 | 33 | tris-HCl, pH 8.5 35% DMSO 6% Brij 35 GSH:cysteamine (2:2) | 40 | 48 |
| MCoTI-II ⁹⁶ | 34 | +2.7 | 38 | NaHCO ₃ , pH 7.8, 1mM GSH | 95 | 24 |
| α -defensin-1 ²⁶⁴ | 93 | -3.2 | 44 | NH ₄ OAc, pH 7.7, 50% ACN, GSH:GSSG (5:1) | 21 | 40 |
| RTD-1 ²⁴ | 18 | +4.7 | 28 | NaHCO ₃ , pH 7.8, air oxidation | 36 | 18 |
| α -Conotoxin ImI ²⁶⁶ | 17 | +2.8 | 35 | tris-HCl, pH 8.7, 60% 2-propanol, GSH:GSSG (2:1) | 70 | 4 |
| MVIIC ²⁶⁵ | 27 | +5.7 | 38 | (NH ₄) ₂ SO ₄ , pH 7.7, GSH:GSSG (5:1) | 84 | 24 |
| EVIA ²⁷³ | 33 | -1.3 | 26 | [C ₂ mim][OAc]GSH:cysteamine(2:1) | 20 | 0.5 |

⁺ Ratio of hydrophilic residues / total number of residues

⁺⁺ Time to complete the oxidative folding reactions

More and more evidence has suggested that co-solvents or additives are sometimes required to help stabilize the native conformation and promote the accumulation of correctly folded peptides. Organic solvents, such as acetonitrile (ACN) and alcohols, have been successfully used for oxidative folding of hydrophobic peptides, such as human defensins, α -conotoxin, and cyclotides which are difficult to fold in normal aqueous buffer.^{184,264,266} Nishiuchi and co-workers reported that the oxidative folding yield of α -defensins could be improved from 27% to 89% by including 50% ACN in the folding reaction.

It should be noted that some peptides are more difficult to form correct disulfide bonds than others using this approach. As a result, this process is believed to be sequence- or structure-dependent. For example, the oxidative folding of endothelin-1, which contains an α -helical structure at the C-terminus, always resulted in an incorrect isomer besides the correct one. A Lys-Arg extension at the N-terminus greatly improved the ratio of correctly- to incorrectly-folded isomers by stabilizing the α -helical structure.²⁷⁴ Another example involves two subfamilies of the cyclotide. In the case of bracelet cyclotide cycloviolacin O2, addition of organic co-solvents did not lead to significant improvement in the oxidative folding yield, although they were effective in oxidative folding of Möbius cyclotides. Möbius cyclotides seem to be easier to fold than bracelet cyclotides; the twist structure in loop 5 is reported to account for the different global oxidative folding efficiency.⁸³

Although much work has been done to clarify the oxidative folding process, *in vitro* oxidative folding of CRPs still remains elusive and is difficult to achieve for peptide synthesis. These reactions are always accompanied by long reaction time and heterogeneous products. Thus, there is a need to develop novel and practical methods for fast and efficient oxidative folding of CRPs. One of the goals of this thesis was to develop a fast and efficient method for the oxidative folding of CRPs, and in particular, macrocycles such as cyclotides.

8. Aim

The major goal of my thesis is to develop practical methodology for chemical synthesis of cyclic cysteine-rich peptides (CRPs) and to design active peptide biologics by using cyclic CRPs as scaffolds.

The specific aim of my thesis are as follows:

1. The development and optimization of chemical synthesis for cyclic CRPs. The results are summarized in chapters 3 and 4, which mainly describe the synthesis of cyclic CRP kB1 and SFTI-1 using chemical ligations. A new approach mimicking the natural process of amide bond formation is reported. In this approach, the novel thioethylamido thioester surrogate served as a mini-intein to mediate the break and make of amide bonds with high efficiency. The whole cyclization process was simplified as a one-pot reaction. Moreover, the oxidative folding process was optimized as well by replacing the aqueous folding buffer with organic solvents. The method reported here is practical and productive, which can be applied to high-throughput synthesis of other cyclic CRPs.

2. The development and optimization of chemoenzymatic synthesis for cyclic CRPs. A novel cyclase butelase 1 which can mediate the cyclization of linear peptides was discovered and isolated from the plant *Clitoria ternatea*. The substrate specificity in the Butelase-mediated ligation was investigated and determined by a combinatorial library approach. The characterization of substrate specificity revealed the optimal conditions required for efficient transpeptidation reactions and suggested a minimal sequence requirement for butelase-mediated ligation reactions.

3. The design, synthesis and engineering of cyclic CRPs for peptide biologics development. Chapter 6 described the engineering and synthesis of a cyclic CRP SFTI-1 by chemical synthesis approaches. Enhanced stability against thermal and enzymatic degradation was obtained and the engineered peptide showed potential membrane permeability. This study can be used as a proof-of-concept research for the grafting strategy in drug design.

Chapter 2 Material and methods

1. Chemical Synthesis of Cyclic CRPs

1.1 General information

All Fmoc-amino acids, reagents and solvents were used without purification. Amino acids were purchased from Chem Impex International, Inc (Wood Dale, IL, USA) except Fmoc-MeCys(Trt)-OH and Fmoc-MeAla-OH, which were obtained from Advamacs (Lodz, Poland) and Novabiochem respectively. Resins including Rink amide MBHA resin, Wang resin and 2-chlorotrityl chloride (Cl-Trt(2-Cl)) resin were purchased from Novabiochem as well. Coupling reagents including HOAt, HOBt, BOP, PyBOP and HATU were purchased from GL Biochem (Shanghai, China) while DIC, DCC, DIEA and Kaiser test kit were from Sigma Aldrich (Milwaukee, WI, USA). Solvents including DMF, DCM, piperidine and acetonitrile (ACN) were obtained from Merck Biosciences. TFA was purchased from Alfa Aesar.

Peptide syntheses were performed manually or on a CEM Liberty 1 automated microwave peptide synthesizer. Most if not all manual syntheses were carried out on Rink amide MBHA resin (0.34 mmol/g) on a 0.1 mmol scale. The coupling reaction was performed using 4 eq. Fmoc-amino acid, 4 eq. PyBOP and 6 eq. DIEA while 20% v/v piperidine in DMF was used in the deprotection. Each step is monitored by Kaiser test using a mixture containing 10 μ l phenol (80% in ethanol), 10 μ l KCN (1 mM in pyridine) and 10 μ l ninhydrin (6% in ethanol). The Kaiser test solution was heated at 100 °C for 1 min. After each step, resins were washed by DCM and DMF. For secondary amine, acetaldehyde/chloranil test was used. It was conducted in 10 μ l 2% acetaldehyde, 10 μ l 2% chloranil in DMF, 5 min at 25 °C.

For automated synthesis, reagents used were as following: 0.2 M amino acid, 2 M DIEA (35% DIEA, 55% DMF, 10% DCM) as the activation base, 0.5 M PyBOP as activator, 20% piperidine in DMF with 0.1 M HOBt as deprotection. In each cycle, 5 eq. amino acid, 5 eq. coupling reagent and 10 eq. base was added into the reaction vessel to carry out the coupling reaction. Final cleavage was usually performed using

TFA/TIS/H₂O (95:2.5:2.5). The crude peptide was precipitated with cold diethyl ether, then was dried *in vacuo* and purified by preparative HPLC.

All if not all HPLC were performed on Shimadzu CBM-20A liquid chromatography (Tokyo, Japan) using a Grace or Phenomenex analytical column (4.6 mm x 250 mm) normally with a linear gradient of 10-60% ACN/0.1% TFA for 25 min at a flow rate of 1 ml/min, and a Grace or Phenomenex preparative column C18 column (22 mm x 250 mm) at a flow rate of 6 ml/min with a gradient of 10-40% for 80 min.

Molecular weights were measured with ABI QSTAR Elite electrospray ionization mass spectrometry (ESI-MS) or ABSCIEX 4800 plus MALDI-TOF/TOFTM Analyzer using α -Cyano-4-hydroxycinnamic acid (CHCA) matrix in matrix buffer containing 0.1% TFA, 29.9% H₂O and 70% ACN.

1.2 MeCys-mediated N ^{α} -thioester formation in aqueous conditions

1.2.1 Synthesis of model peptides 9a: TIGGIR-MeCys-Gly-NH₂ on the MeCys resin by Fmoc SPPS. (Refer to Figure 3.3)

Rink amide resin (294 mg, 0.1 mmol, 0.34 mmol/g) was swelled for 15 min in dichloromethane (DCM) (10 ml) and washed with *N,N*-dimethylformamide (DMF) and DCM (5 ml x 3). Fmoc-Gly-OH (119 mg, 0.4 mmol, MW 297.3) and 1-Hydroxy-7-azabenzotriazole (HOAt) (0.4 mmol, 54 mg, MW 136.1) in DMF (10 ml), *N,N'*-diisopropylcarbodiimide (DIC) (0.4 mmol, 63 μ l, MW 126.2, *d* 0.806 g/ml) were added to the resin, and the resin suspension was shaken for 1 h to afford resin **5**. Then the resin **5** was washed with DMF (5 ml x 3), treated with 20% piperidine in DMF (5 ml x 3), and washed with DMF and DCM (5 ml x 3). Fmoc-MeCys(Trt)-OH (240 mg, 0.4 mmol, MW 599.7), benzotriazol-1-yl-oxytritypyrrolidinophosphonium hexafluorophosphate (PyBOP) (208 mg, 0.4 mmol, MW 520.4) in DMF were added to the resin together with *N,N*-diisopropylethylamine (DIEA) (106 μ l, 0.6 mmol, MW 129.2, *d* 0.742 g/ml), which was shaken for 1 h to afford the resin **6**. The Fmoc group was then removed by 20% piperidine in DMF (5 ml x 3), and the resin was washed with DMF and DCM (10 ml x 3). Fmoc-Arg(Pbf)-OH (259 mg, 0.4 mmol, MW 648.6), *O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluoro-

phosphate (HATU) (152 mg, 0.4 mmol, MW 380.2) in DMF (5 ml), and DIEA (106 μ l, 0.6 mmol, MW 129.2, *d* 0.742 g/ml) were added to the resin and the coupling reaction was allowed to proceed for 1 h. Acetaldehyde/chloranil test was used to monitor the coupling reaction. A double or triple coupling was performed to make sure the coupling reaction was completed. Peptide assembly was carried out on resin **7** by using PyBOP (208 mg, 0.4 mmol, MW 520.4) and DIEA (106 μ l, 0.6 mmol, MW 129.2, *d* 0.742 g/ml). The protected peptide **8** (50 mg, 0.013 mmol) was cleaved and purified to afford the peptide **9a** (5.6 mg, 56%). MALDI-TOF MS was performed to detect the molecular weight. An *m/z* of 789.54 ($[M+H]^+$) was detected and the theoretical molecular weight is 789.43.

1.2.2 Synthesis of H-TIGGIR-MES (Refer to Figure 3.5)

The reaction buffer at pH 1-7 was prepared with either sulfuric acid or sodium phosphate. Diluted sulfuric acid with a concentration of 0.1 M and 5 mM was used as the pH 1 and 2 buffers respectively. Buffers at pH 3-5 were prepared using 0.1 M Na_2HPO_4 and NaH_2PO_4 while pH 6 and 7 buffers were prepared using phosphate buffer (0.2 M) containing 20 mM TCEP to prevent undesired disulfide formation. For each pH condition, TIGGIR-MeCys-Gly-NH₂ **9a** (0.158 mg, 0.2 mmol, MW 789) was dissolved in 40 μ l buffer to reach a final concentration of 5 mM with 50 eq. of external thiol sodium 2-sulfanylethanesulfonate (MESNa) (1642 mg, 10 mmol, MW 164.2) and incubated at 40 °C. The reaction was proceeded for 24 h and the process was monitored by RP-HPLC and MALDI-TOF MS.

1.2.3 Synthesis of H-TIGGIRCALVIN-NH₂, **13**: Ligation of TIGGIR-MES and CALVIN-NH₂ (Refer to Figure 3.6 and 3.7)

Peptide thioester TIGGIR-MES **11a** (0.24 mg, 0.3 μ mol, MW 739.3) and the synthetic peptide CALVIN-NH₂ **12** (0.57 mg, 0.9 μ mol, MW 630.8) were dissolved in 300 μ l pH 7.5 phosphate buffer (0.2 M) containing 20 mM TCEP, 20 mM methyl mercaptoacetate (MMA) (6 μ mol, 6 μ l 1 M solution, MW 106.1, *d* 1.187 g/ml), and 6 M Gnd-HCl. The reaction was allowed to proceed at 25 °C for 4 h to give a ligation product TIGGIRCALVIN-NH₂ **13** (theoretical molecular weight 1229.5, observed 1229.6 $[M+H]^+$). ESI-MS was used to measure the molecular weight.

1.2.4 One-pot ligation of TIGGIR-MeCys-Gly-NH₂ and peptide CALVIN-NH₂ (Refer to Figure 3.8)

Peptides TIGGIR-MeCys-Gly-NH₂ **9a** (0.24 mg, 0.3 μmol, MW 789.5) and CALVIN-NH₂ **12** (0.19 mg, 0.3 μmol, MW 630.8) were dissolved in 300 μl 0.2 M sodium phosphate buffer (pH 5) with 6 M Gnd-HCl and 20 mM TCEP. 50 eq. MMA (1.3 μl, 15 μmol, *d* 1.187 g/ml) was added into reaction mixture and the reaction was incubated at 40 °C for up to 24 h to give the ligation product TIGGIRCALVIN-NH₂ **13** (theoretical molecular weight 1229.5, observed 1229.6 [M+H]⁺).

1.2.5 N^α-Cleavage of Cys-Gly or MeCys-Gly bonds (Refer to Figure 3.9)

TIGGIR-Gly-MeCys-Gly-NH₂ (0.25 mg, 0.3 μmol, MW 846.0) or TIGGIR-Gly-Cys-Gly-NH₂ (0.249 mg, 0.3 μmol, MW 832.0) was treated with 300 μl pH 2 buffer containing 20% MMA (60 μl, 60 μl 1 M solution, 692.3 μmol, *d* 1.187 g/ml) (v/v) and 20 mM TCEP. The reaction was incubated at 40°C for up to 48 h.

1.3 MeCys-mediated synthesis of cyclic peptide SFTI-1 (Refer to Figure 3.10)

1.3.1 Synthesis of SFTI-MeCys-Gly-NH₂

Rink amide resin (294 mg, 0.1 mmol, 0.34 mmol/g) was used for synthesizing SFTI-1. The resin was treated with 20% piperidine in DMF for 5 min x 3 to remove the Fmoc groups. After the Fmoc groups were removed, Fmoc-Gly-OH (119 mg, 0.4 mmol, MW 297.3) was introduced onto the resin using DIC (63 μl, 0.4 mmol, MW 126.2, *d* 0.806 g/ml) and HOAt (54 mg, 0.4 mmol, MW 136.1). After that, we coupled the MeCys residue by using a cocktail containing Fmoc-MeCys(Trt)-OH (240 mg, 0.4 mmol, MW 599.7), PyBOP (208 mg, 0.4 mmol, MW 520.4) and DIEA (106 μl, 0.6 mmol, MW 129.2, *d* 0.742 g/ml) (4/4/6 eq.) in DMF for 30 min. 20% piperidine in DMF was used to remove the Fmoc deprotection group for 5 min x 3. The coupling and deprotection reaction were monitored by a Kaiser test. A protocol of Fmoc amino acid/HATU/DIEA (4/4/6 eq.) was employed for the coupling of amino acid after N-methyl amino acid. It was performed for 45 min and repeated twice. Acetaldehyde/chloranil test was used to monitor the coupling reaction. During the test, the yellow to transparent color on the resin indicated the absence of free secondary amine, which confirmed the completion of the reaction. Then the resin **7**

was subjected to a CEM Liberty 1 automated microwave peptide synthesizer to perform the peptide elongation reaction by using PyBOP/DIEA for coupling as described for coupling of MeCys residue. We modified the deprotection solution to 20% morpholine in DMF to minimize the aspartimide formation. After peptide chain assembly, resin **8** was washed with DMF, DCM and diethyl ether, and then dried *in vacuo*. 448 mg dry resin was obtained after the synthesis. Then the protected peptide was cleaved by a cocktail containing TFA/TIS/H₂O (9.5 ml/250 μ l/250 μ l) for 2 h at 25 °C. The synthesis afforded SFTI-MeCys-Gly-NH₂ **9b** (theoretical molecular weight 1707.1, observed 1706.8 [M+H]⁺) in 17% isolated yield.

1.3.2 One-pot cyclization of SFTI-MeCys-Gly-NH₂

SFTI-MeCys-Gly-NH₂ **9b** (1.7 mg, 1 μ mol) was dissolved in 1 ml 5 mM H₂SO₄ solution (pH 2). To the reaction, 50 eq. MESNa (8.2 mg, 50 μ mol, MW 164.2) was added together with TCEP (2.8 mg, 10 μ mol). Then the reaction mixture was incubated at 40 °C for 24 h. At 1 h, 3 h, 6 h, 12 h and 24 h, the reaction was monitored by RP-HPLC and the formation of the MES thioester **11b** thiolactones **14a** and **14b** was analyzed by MALDI-TOF MS. After the thioesterification was completed, the pH of the reaction was adjusted to 7.5 by 1 M NaOH to perform the thiazip cyclization. The mixture was gently stirred at 25 °C for 4 h to afford cyclic reduced SFTI-1 **15**. The cyclic reduced SFTI-1 **15** was purified by RP-HPLC with a Grace preparative column and lyophilized *in vacuo*. Then the reduced peptide **15** was oxidized by iodine to form the intramolecular disulfide bond. To a solution of **15** (1 mg, 0.66 μ mol) in 50% AcOH (1 ml) at 0 °C, 0.1 M I₂/MeOH (6.6 μ l, 0.66 μ mol) was added. The reaction mixture was stirred for 1 min, Then it was quenched with 1 M ascorbic acid (3.4 μ l, 3.3 μ mol), purified by RP-HPLC to give SFTI-1 (0.6 mg, 60%) (theoretical molecular weight: 1512.8, observed 1513.0 [M+H]⁺).

1.4 MeCys-mediated C ^{α} -thioester formation in strongly acidic conditions

1.4.1 Synthesis of TIGGIR-MeCys-TC (Refer to Figure 3.11 and 3.12)

TIGGIR-MeCys-Gly-NH₂ **9a** (0.16 mg, 0.2 μ mol) was dissolved in 170 μ l trifluoroacetic acid (TFA) (MW 114.0, *d* 1.489 g/ml) and then to the reaction solution, 20 μ l 1% trifluoromethanesulfonic acid (TFMSA) (MW 150.1, *d* 1.696 g/ml) and 10

μl thiocresol (TC) was added. 1% TFMSA was prepared by diluting 100 μl TFMSA in 9.9 ml TFA. The reaction was proceeded at 25 °C for 2 h. Cold dimethyl ether was used to precipitate the reaction mixture and then the precipitation was re-dissolved in *Milli-Q* H₂O. The solution was analyzed by analytical RP-HPLC immediately with a linear gradient 10-60% in 30 min.

1.4.2 Ligation reaction of TIGGIR-MeCys-TC and CALVIN-NH₂ (Refer to Figure 3.14)

Peptide thioester TIGGIR-MeCys-TC **17** (0.25 mg, 0.3 μmol , MW 839.2) was reacted with a synthetic peptide CALVIN-NH₂ **12** (0.57 mg, 0.9 μmol) to give ligation product TIGGIR-MeCys-CALVIN-NH₂ **18** in a phosphate reaction buffer (0.1 M, pH 7.5) containing 6 M Gnd-HCl. Reaction was performed at 25 °C for 2 h. The reaction mixture was reduced by TCEP before injection to HPLC. The ligation product was confirmed by tandem MS.

1.4.3 Synthesis of TIGGIR series peptides

Nine TIGGIR series peptides were synthesized using Fmoc SPPS. Rink amide resin (294 mg, 0.1 mmol, 0.34 mmol/g) or Wang resin (125 mg, 0.1 mmol, 0.8 mmol/g) was used for synthesizing peptides listed in (**Table 2.1**). The resins were first swelled in DCM for 15-30 min and then deprotected by 20% piperidine in DMF. After the Fmoc groups were removed, the first amino acid Fmoc-Gly-OH (119 mg, 0.4 mmol, MW 297.3) or Fmoc-Ala-OH was introduced onto the resin by HOAt (54 mg, 0.4 mmol, MW 136.1), and DIC (63 μl , 0.4 mmol, MW 126.2, *d* 0.806 g/ml). Peptide elongation was then carried out manually by using 4 eq. Fmoc amino acids, 4 eq. PyBOP (208 mg, 0.4 mmol, MW 520.4) and 6 eq. DIEA (0.4 mmol, 126 μl , MW 129.2, *d* 0.742 g/ml) in DMF for 30 min. When coupling an amino acid to the N-methyl amino acid, we used 4 eq. corresponding Fmoc amino acid (0.4 mmol), 4 eq. HATU (152 mg, 0.4 mmol, MW 380.2) and 6 eq. DIEA (106 μl , 0.6 mmol, MW 129.2, *d* 0.742 g/ml) in DMF for 45 min. The coupling reaction was repeated for 3-4 times and monitored by acetaldehyde/chloranil test. For TIGGIR-MeCys-OH, it was synthesized on Wang resin. The MeCys residue (240 mg, 0.4 mmol, MW 599.7) was introduced by HOAt (54 mg, 0.4 mmol, MW 136.1), and DIC (63 μl , 0.4 mmol, MW 126.2, *d* 0.806 g/ml). For TIGGIRC-TEBA, it was synthesized as the procedure

described in **1.17**. The following coupling reaction was the same as described above. After the peptide chain assembly and final Fmoc removal, the protected peptide was cleaved by a cocktail containing 4.75 ml TFA, 125 μ l TIS, and 125 μ l H₂O for 2 h. MALDI-TOF MS was used to measure the molecular weight which is summarized in **Table 2.1**. For TIGGIR-MeAla-Gly-NH₂ and TIGGIR-MeAla-(Gly)₆-NH₂, the synthesis procedure was the same as described above.

Table 2.1 Molecular weights of TIGGIR series peptides

| Peptide | Expected mass (m/z) [M+H] ⁺ | Observed mass (m/z) [M+H] ⁺ |
|---------------------------------|---|---|
| TIGGIR-MeCys-G-NH ₂ | 788.9 | 789.4 |
| TIGGIR-MeCys-A-NH ₂ | 803.0 | 803.6 |
| TIGGIR-MeCys-NH ₂ | 731.9 | 732.5 |
| TIGGIR-MeCys-OH | 732.9 | 733.4 |
| TIGGIRA-MeCys-G-NH ₂ | 860.0 | 860.7 |
| TIGGIRG-MeCys-G-NH ₂ | 846.0 | 846.6 |
| TIGGIRL-MeCys-G-NH ₂ | 902.1 | 902.6 |
| TIGGIRGCG-NH ₂ | 832.0 | 832.5 |
| TIGGIRC-TEBA | 835.1 | 835.5 |

1.4.4 TFMSA treatment of TIGGIR series peptides

Peptides listed in **Table 2.1** (0.2 μmol, 150 μg-180 μg) were dissolved in 170 μl TFA and then to the reaction solution, 20 μl 1% TFMSA and 10 μl TC was added. 1% TFMSA was prepared by diluting 100 μl TFMSA in 9.9 ml TFA. The reaction was proceeded at 25 °C for 2 h. Sample preparation was the same as described in section **1.9**. Solutions were analyzed by analytical RP-HPLC immediately with a linear gradient 10-60% in 30 min.

1.4.5 Determination of acidity function of TFMSA-TFA solution (Refer to **Figure 3.19** and **Table 3.2**)

We determined the acidity function of various TFMSA-TFA anhydrous solutions by measuring the ratio of protonated and unprotonated dimethyl sulfide (DMS) using ¹H-NMR.²⁷⁵ DMS will begin to be protonated by the addition of TFMSA in TFA and the ratio is dependent on the concentration of TFMSA. Thus, the ratio of protonated DMS can be converted to the acidity function of TFMSA-TFA solution.

A series of TFMSA-TFA solutions with various percentages of TFMSA (0-10% v/v) were prepared on ice. Together 5% (v/v) tetramethylammonium sulfate (TMAS) as an internal standard was added. 5% (v/v) DMS was mixed with the solution just before the measurement to avoid any oxidation or decomposition of it. ^1H -NMR was carried out using a Bruker AVANCE I spectrometer at 400 MHz. The ionization ratio I was calculated by the equation

$$I = \frac{\Delta\nu(\text{DMS}) - \Delta\nu(\text{B})}{\Delta\nu(\text{BH}^+) - \Delta\nu(\text{DMS})}$$

Using TMAS as an internal standard, the $\Delta\nu(\text{B})$ and $\Delta\nu(\text{BH}^+)$ can be corrected by the chemical shift of unprotonated and protonated TMAS as 0 and 309.56 Hz, respectively. Thus, the above equation was simplified as

$$I = \frac{\Delta\nu(\text{DMS})}{309.56 - \Delta\nu(\text{DMS})}$$

A plot showing the TFMSA concentration (vol%) against DMS protonation ratio (%) was generated by calculating ratio I in each TFMSA solution. A linear correlation between protonation ratio and acidity was found from 0-1% TFMSA-TFA solution.

Using the ionization ratio I , the acidity function H_0 was calculated by the equation

$$\log I = -mH_0 + pK_{\text{BH}^+}$$

Where the slope constant of Yates-McClelland equation $m = 1.26$, $pK_{\text{BH}^+} = 6.95$ ^{276,277}

$$H_0 = -\frac{1}{1.26} (6.95 + \log I)$$

1.5 β -thiolactone formation

1.5.1 Synthesis of β -thiolactone (Refer to Figure 3.22)

Peptide TIGGIR-MeCys-TC (0.25 mg, 0.3 μ mol, MW 839.2) was dissolved in 300 μ l phosphate buffer (pH 4-7). At various time intervals, 50 μ l aliquots were taken and analyzed by RP-HPLC and MALDI-TOF MS.

1.5.2 S-alkylation of TIGGIR-MeCys-TC (Refer to Figure 3.23)

Peptide TIGGIR-MeCys-TC (0.25 mg, 0.3 μ mol, MW 839.2) was dissolved in 300 μ l citric buffer (200 mM, pH 3.0) containing 20 mM *N*-ethylmaleimide (NEM) (6 μ mol, 750 μ g, MW 125.1) to reach a final peptide concentration of 1 mM. S-alkylation was performed at 25 $^{\circ}$ C for 1 h. Excessive NEM was removed by RP-HPLC purification to obtain TIGGIR-MeCys (NEM)-TC thioester.

1.5.3 Preparation of TIGGIR-MeCys-MMA, TIGGIR-Cys-MMA, TIGGIR-Cys-TC (Refer to Figure 3.25)

TIGGIR-MeCys-Gly-NH₂ **9a** (1.6 mg, 2 μ mol) was dissolved in 1700 μ l TFA and then to the reaction solution, 200 μ l 1% trifluoromethanesulfonic acid (TFMSA) and 100 μ l methyl mercaptoacetate (MMA) were added. 1% TFMSA was prepared by diluting 100 μ l TFMSA in 9.9 ml TFA. The reaction was proceeded at 25 $^{\circ}$ C for 2 h. Sample preparation was the same as described in section 1.9. The solution was purified RP-HPLC immediately to afford TIGGIR-MeCys-MMA (theoretical molecular weight: 820.4, observed 820.0 [M+H]⁺).

TIGGIR-Cys-TEBA (1.67 mg, 2 μ mol, MW 832.8) was dissolved in 400 μ l acetic acid at a final concentration of 5 mM together with 10% MMA. The reaction was incubated at 40 $^{\circ}$ C for 24 h. Acetic acid was then evaporated by a Thermo Scientific Speedvac SPD 1010. Sample preparation was the same as described in section 1.9. The solution was purified by RP-HPLC immediately to afford TIGGIR-C-MMA (theoretical molecular weight: 806.4, observed 806.0 [M+H]⁺).

TIGGIR-Cys-TEBA (1.67 mg, 2 μ mol, MW 832.8) was dissolved in 400 μ l acetic acid at a final concentration of 5 mM together with 10% TC. The reaction was

incubated at 40°C for 24 h. Acetic acid was then evaporated by a Thermo Scientific Speedvac SPD 1010. Sample preparation was the same as described in section 1.9. The solution was purified by RP-HPLC immediately to afford TIGGIRC-TC (theoretical molecular weight: 824.4, observed 824.1 [M+H]⁺).

1.5.4 Synthesis of TIGGIRC-TEBA and TIGGIR-TEBA (Refer to Figure 3.26)

For synthesis of TIGGIRC-TEBA, Cl-Trt(2-Cl) resin (84 mg, 0.1 mmol, 1.2 mmol/g) was used as the starting material. (2-Butylamino)ethanethiol (7.5 µl, 0.05 mmol) in DCM (5 ml) was added into the resin and shaken for 1 h at 25 °C. Then DIEA (53 µl, 0.3 mmol, MW 129.2, *d* 0.742 g/ml) in MeOH (5 ml) was added to block the unreacted Cl sites on resin. Subsequently, Fmoc-Cys(Trt)-OH (117 mg, 0.2 mmol, MW 585.3) was introduced onto the resin using HATU (76 mg 0.2 mmol, MW 380.2) and DIEA (53 µl, 0.3 mmol, MW 129.2, *d* 0.742 g/ml) in DMF. The resin was filtrated, and washed with DMF and DCM (5 ml x 3). An acetaldehyde/chloranil test was used to monitor the coupling reaction for detection of a secondary amino group. PyBOP/DIEA was used to couple the following amino acids. A repeat coupling or deprotection would be performed if necessary.

For synthesis of TIGGIR-TEBA, Cl-Trt(2-Cl) resin (336 mg, 0.4 mmol) was used. Fmoc-Arg(Pbf)-OH (519 mg, 0.8 mmol) was introduced onto the resin using HATU (304 mg 0.8 mmol, MW 380.2) and DIEA (212 µl 1.2 mmol, MW 129.2, *d* 0.742 g/ml) in DMF to afford **24**. The same procedure was carried out as described for TIGGIRC-TEBA. The resin was finally cleaved by 5 ml TFA/TIS/H₂O 90:5:5 to afford TIGGIR-TEBA **25a** (theoretical molecular weight 729.7, observed 730.4 [M+H]⁺).

1.6 TEBA-mediated N^α-thioester formation in aqueous conditions (Refer to Figure 3.27)

Seven conditions with pH ranged from 1 to 7 were tested. In each condition, TIGGIR-TEBA **25a** (0.158 mg, 0.2 µmol) was dissolved in 200 µl buffers (reaction buffers were prepared through the same procedure as described in 1.3 by diluted H₂SO₄ or 0.2 M sodium phosphate buffers) together with MESNa (1.6 mg, 10 µmol, MW 164.2). The final concentration of peptide and MESNa was 1 mM and 50 mM,

respectively. The reaction was performed at 40 °C for 24 h. The processes of reactions were monitored by RP-HPLC and MALDI-TOF MS.

1.7 TEBA-mediated synthesis of cyclic peptide SFTI-1 (Refer to Figure 3.29 and 3.30)

Cl-Trt(2-Cl) resin **22** (417 mg, 0.5 mmol, 1.2 mmol/g) was swollen for 30 min in DCM (15 ml) and washed with DMF and DCM (10 ml x 3). The resin **22** was re-suspended in 10 ml DCM. 2-(butylamino)ethanethiol (37.6 μ l, 0.25 mmol) was added to the suspension, and it was shaken for 1 h at 25 °C to afford resin **23**. Then a solution of DIEA (133 μ l, 0.75 mmol, MW 129.2, *d* 0.742 g/ml) in MeOH (10 ml) was added to the suspension and the reaction was shaken for 10 min to block the unreacted Cl site on the resin. After filtration, the resin was washed with DCM (10 ml x 3) and DMF (10 ml x 3). To the resin **23**, a pre-mixed solution containing 4 eq. Fmoc-Arg(Pbf)-OH (649 mg, 1 mmol, MW 648.8), 4 eq. HATU (380 mg, 1 mmol, MW 380.2) and 6 eq. DIEA (265 μ l, 0.75 mmol, MW 129.2, *d* 0.742 g/ml) was added. The coupling reaction was performed at 25 °C for 1 h and monitored by an acetaldehyde chloranil test. Then the resin **24** was subjected to an automatic microwave CEM Liberty 1 peptide synthesizer using PyBOP/DIEA. 20% morpholine in DMF was used for removal of Fmoc deprotection groups instead of 20% piperidine. We obtained 940 mg SFTI-TEBA-loaded resin. In the cleavage cocktail, we used TFA/TIS/thioanisole (95:2.5:2.5) to remove the solid support and side chain protection groups. Total of 330 mg crude peptide containing **27a** and **27b** were obtained and used directly for the following cyclization reaction without purification.

A crude peptide precipitation (dry weight 4.9 mg) containing SFTI-TEBA **27a** and **27b** was dissolved in 1 ml sodium phosphate buffer (0.1M, pH 3) to reach a concentration of 3 mM. Two conditions were performed with or without external thiols. The condition containing MESNa (24.7 mg, 150 μ mol, MW 164.2) was incubated at 40 °C for 18 h. To prevent inter- or intra-molecular disulfide formation, TCEP (2.9 mg, 10 μ mol, MW 250.2) was added. The reaction was monitored by RP-HPLC at both 220 and 260 nm. The formation of reaction intermediates such as thioesters **29** and thiolactones **28a** and **28b** was verified by MALDI-TOF MS. Then the reaction mixture was diluted with 2 ml sodium phosphate (0.1M, pH 3) to a final

peptide concentration of 1 mM. 1 M NaOH was used to adjust pH of the solution to 7. The mixture was gently stirred at 25 °C for 4 h to afford cyclic reduced SFTI-1 **30**. The cyclic reduced SFTI-1 **30** was purified by RP-HPLC with a Phenomenex preparative column and lyophilized *in vacuo*. Together with 20% DMSO (v/v), reduced cyclic SFTI-1 **30** was then dissolved in a pH 7.5 phosphate buffer at concentration of 1 mM. The oxidative folding reaction was performed at 25 °C and monitored by RP-HPLC. For conditions without MESNa, the cyclization reaction was similar except the absence of external thiol MESNa. After thia zip cyclization, 20% DMSO was added directly to the mixture to carry out the oxidative folding reaction. The global oxidative folding reaction was quenched with 1 M HCl to pH 2 for both conditions, and then purified by RP-HPLC with a Phenomenex preparative column to give native SFTI-1 **31** (theoretical molecular weight 1512.8, observed 1513.7 [M+H]⁺).

1.8 Synthesis of kalata B1 via a hydrazide linker

1.8.1 Preparation of hydrazide resin

Cl-Trt(2-Cl) resin (2 g, 1.2 mmol/g) was swelled for 30 min in DCM (10 ml). A mixture of DIEA (1.26 ml, 7.2 mmol, MW 129.2, *d* 0.742 g/ml) and hydrazine monohydrate (233 μl, 4.8 mmol, MW 50.06, *d* 1.032 g/ml) in DMF (10 ml) was added dropwise to the reaction and the suspension was stirred for 1 h at 25 °C. Then a solution of DIEA (636 μl, 3.6 mmol, MW 129.2, *d* 0.742 g/ml) in MeOH (10 ml) was added to the suspension and the reaction was shaken for 10 min to block the unreacted Cl site on the resin. The resin was filtered, washed with methanol for 3 times, then washed with DMF and DCM successively and dried *in vacuo* overnight. The substitution of the 2-chlorotrityl hydrazine resin was calculated by weight increment and the value is determined as 0.6 mmol/g.

1.8.2 Synthesis of peptide hydrazides (Refer to Figure 3.32)

The linear peptide hydrazides were prepared by standard Fmoc strategy. 2-chlorotrityl hydrazine resin (167 mg, 0.1 mmol) was used as the solid support for the synthesis. In the synthesis, the first amino acid Gly was coupled manually by using Fmoc-Gly-OH (119 mg, 0.4 mmol, MW 297.3), 4 eq. HATU (152 mg, 0.4 mmol,

MW 380.2) and 6 eq. DIEA (106 μ l, 0.4 mmol, MW 129.2, *d* 0.742 g/ml). The rest of the stepwise couplings were done by using BOP/DIEA (4/4 eq.) protocol in a CEM Liberty1™ automated microwave peptide synthesizer. After the assembling of the peptide, the cocktail containing TFA/thioanisole/TIS (90/5/5 v/v) was used to cleave the peptide from the solid support. Preparative RP-HPLC with a Grace C18 column (22 mm x 250 mm) was used to purify the peptide and followed by lyophilization.

1.8.3 Cyclization of kB1 (Refer to Figure 3.33)

The kalata B1 hydrazide was used as the starting material for cyclization of kalata B1. It involved two steps: azide formation step and cyclization step. In the azide formation step, 440 μ g kalata B1 hydrazide (0.15 μ mol) was dissolved in 300 μ l *Milli-Q* H₂O and 10 eq. NaNO₂ (1.5 μ mol, 7.5 μ l 200 mM solution) was added into the solution with the pH adjusted to 2 by 3 M HCl. The reaction mixture was allowed to proceed at 0-4 °C for 20 min. After the azide formation was completed, 100 eq. MMA (15 μ mol, 15 μ l 1 M solution, MW 106.1, *d* 1.187 g/ml) was added into solution and the pH was adjusted to 7 immediately by 1 M NaOH to initiate the cyclization step. This step was performed at 25 °C for 2 h. Both steps were monitored by RP-HPLC with a Grace analytical column. The reaction solution was reduced by 30 mM TCEP before RP-HPLC analysis. In various runs for optimization of the reaction conditions, peptide hydrazides concentration was kept constant at 0.5 mM, different concentrations of NaNO₂ or MMA were used.

1.8.4 Oxidative folding of kB1 (Refer to Figure 3.35)

After the cyclization was completed, the reaction solution proceeded to the oxidative folding directly without purification. The reaction mixture was diluted with 1200 μ l 2-propanol to peptide concentration of 0.1 mM and 10% DMSO (150 μ l), 2% (v/v) morpholine (30 μ l) were added successively. The oxidative folding reaction was performed at 25 °C for 2 h. Various conditions were investigated including the effect of morpholine and cysteamine. For conditions containing cysteamine, 100 eq. cysteamine (15 μ mol, 2 μ l 7.5 M solution) were added into the reaction mixture together with 10% DMSO.

2. Oxidative folding of CRPs

2.1 Oxidative folding of kB1 in organic conditions

The native kB1 was extracted from the plant *Oldenlandia affinis* and purified by RP-HPLC. 5 mg peptide was then dissolved in Tris-HCl buffer at pH 7.5. 10 eq. of TCEP were added into the solution and incubated at 25 °C for 1 h. The fully reduced peptide was purified by RP-HPLC on a Shimadzu CBM-20A module using a C18 Grace Vydac semi-preparative column (250 × 10 mm, 5 μm) and lyophilized.

Series of conditions were investigated including different concentration of solvent, base and redox agents. 2-propanol serves as the solvent here, the bases include pyridine, morpholine, imidazole, and the redox agents include cysteamine and DMSO. Each condition contains 100 μM reduced kB1, 10% (v/v) DMSO and 100 mM cysteamine. The total volume for each condition was topped up to 1 ml by 2-propanol. The concentration of 2-propanol varied in the opposite way to that of TFE. Different components involved in the oxidative folding reaction were added in the order of solvent, base, DMSO, cysteamine and peptide sample. At various intervals, 20 μl reaction mixtures was taken and quenched by 10% TFA, then monitored by Shimadzu Nexera UPLC-30A immediately.

2.2 ¹H NMR characterization

Synthetic and native kB1 (1 mg) were dissolved in 500 μl buffer containing 5% D₂O and 95% H₂O, pH 4.3. ¹H NMR was performed on a Bruker AVANCE 600 MHz NMR spectrometer equipped with cryo-probe at 298 K. The ¹H NMR spectra were processed and analyzed by the Topspin software.

2.3 Synthesis of ET-1

Rink amide MBHA resin (294 mg, 0.1 mmol, 0.34 mmol/g) was used for the synthesis of ET-1. The first amino acid Fmoc-Trp(Boc)-OH (211 mg, 0.4 mmol, MW 526.6) was introduced using DIC (63 μl, 0.4 mmol, M.W. 126.2, *d* 0.806 g/ml) and HOAt (54 mg, 0.4 mmol, MW 136.1). 20% piperidine in DMF was used to remove Fmoc groups for all amino acids except Phe⁸ of which the deprotection was achieved by 20% piperidine in DMSO at 50 °C. A BOP/DIEA combination with a ratio of 4:6

was used to couple the following amino acids as described above. Thioanisole instead of H₂O was used for cleavage as described in section **1.19**.

The crude peptide was dissolved in Tris-HCl buffer (pH 8.4) containing 8 M urea and DTT (30 mM) to reduce any disulfide bonds. The reaction was stirred for 2 h at 37 °C. The reaction sample was filtered and analyzed by analytical RP-HPLC with a linear gradient 10-60% in 30 min. After the completion of the reduction, ET-1 was purified by preparative RP-HPLC using a Phenomenex preparative column. The reduced ET-1 (7.23 mg) was obtained with 11% isolated yield (theoretic molecular weight: 2494.9, observed 2494.4 [M+H]⁺)

2.4 Oxidative folding of ET-1 in organic solvent.

Series of conditions were investigated including different concentration of TFE (0, 10, 20, 30, 40, 50, 60, 70, 80 and 85%). Each condition contains 100 μM reduced ET-1, 10% (v/v) DMSO and 100 mM cysteamine. The total volume for each condition was topped up to 1 ml by 2-propanol. The concentration of 2-propanol varied in the opposite way to that of TFE. Different components involved in the oxidative folding reaction were added in the order of solvent, base, DMSO, cysteamine and peptide sample. At 30 min, 200 μl reaction mixtures was taken and quenched by 10% TFA, then monitored by analytical HPLC immediately.

3. Butelase-mediated synthesis of cyclic peptides

3.1 Synthesis of peptide library XXXGIR-OH

Each peptide library was synthesized on Wang resin (1110 mg, 0.9 mmol/g, 1 mmol). The first few residues IGGIR, GGIR or GIR for the first, second and third library, respectively, were synthesized using PyBOP (2.34 g, 4.5 mmol) /DIEA (1.19 ml, 6.75 mmol, MW 129.2, *d* 0.742 g/ml). After that, the resin was split equally into 20 vessels to synthesize individual peptides parallelly. Each peptide library XIGGIR, LXGGIR, LLXGIR contained a single variation at the position X. 20 natural amino acids as well as D-amino acids were introduced at the corresponding position by PyBOP/DIEA with a ratio of 4:6. For Ac-KIGGIR, the N-terminus was acetylated with acetic anhydride and DIEA (2/1 eq.) for 15 min and followed by washing with

DMF and DCM. The final cleavage for each peptide was conducted in TFA/TIS (95:5) for 3 h. Peptides were then purified by RP-HPLC using a Phenomenex preparative column (250 × 10 mm, 5 μm). The purified peptides were lyophilized and obtained with >95% purity.

3.2 Optimization of butelase-mediated ligation reaction

Peptide substrates GIGGIR and KALVINHV were prepared in 10 mM and 1 mM stock in *Milli-Q* H₂O, respectively. For each condition, peptides substrates GIGGIR and KALVINHV were diluted to 1 mM and 0.1 mM, respectively, with a reaction buffer (pH 6) containing 10 mM phosphate, 5 mM β-mercaptoethanol and 1 mM EDTA. Butelase 1 was added then to a final concentration of 0.1 μM.

The reaction mixtures were incubated at 37 °C for 2 h and monitored by RP-HPLC. At various time intervals, 25 μl reaction samples were quenched by 250 μl 0.1% TFA. Then it was analyzed by RP-HPLC using a Phenomenex analytical column (150mm x 4.6mm). The products were verified by MALDI-TOF MS followed by MS/MS. The ligation yields were calculated based on the HPLC profile using the Lab Solution software.

3.3 Ligation of KALVINHV and XXXGIR mediated by butelase 1

An enzyme substrate ratio of butelase 1:KALVINHV:XXXGIR = 1:500:10000 was used for the butelase-mediated ligation reaction. The ligation reaction was performed individually with peptides KALVINHV and XXXGIR. In each condition, the final concentration of butelase 1, KALVINHV, XXXGIR were 0.1 μM, 0.05 mM and 1 mM respectively. The intermolecular ligation reaction was performed in an optimized condition with 1mM EDTA at pH 6.5, 42 °C. β-mercaptoethanol was not included in the reaction buffer. At 10 min and 2 h, 25 μl reaction aliquots were quenched with 250 μl 0.1%TFA. Then it was analyzed by RP-HPLC using a Phenomenex analytical column (150 mm x 4.6 mm). The products were verified by MALDI-TOF MS followed by MS/MS. The ligation yields were calculated based on the HPLC profile using the Lab Solution software.

3.4 Synthesis of C-terminal substrate library KALVXNHV

We used a combinatorial library method to synthesize peptide KALVXNHV. We divided the 20 peptides into 5 groups as summarized in **Figure 5.7**. For each group, Fmoc-Val-OH (135 mg, 0.4 mmol, MW 339.4) was introduced on Rink amide resin (297 mg, 0.1 mmol, 0.34 mmol/g) using HOAt (54 mg, 0.4 mmol, MW 136.1), and DIC (63 μ l, 0.4 mmol, MW 126.2, *d* 0.806 g/ml). At position X, four Fmoc amino acids were mixed in the ratio suggested in **Figure 5.7B** to make up a total amount of 1 mmol. They were dissolved in DMF together with PyBOP (0.55 g, 1 mmol, MW 520.4) / DIEA (0.26 ml, 1.5 mmol, MW 129.2, *d* 0.742 g/ml) to conduct the coupling reaction. The following amino acids were introduced using PyBOP/DIEA with a ratio of 4:6. The final cleavage for each group was conducted in TFA/TIS (95:5) for 3 h. Peptides were then purified by RP-HPLC using a Phenomenex preparative column (250 \times 10 mm, 5 μ m). The purified peptides were lyophilized and obtained with >95% purity.

3.5 Butelase 1 mediated ligation of KALVXXHV and HIGGIR

An enzyme substrate ratio of butelase 1:KALVXXHV:HIGGIR = 1:500:10000 was used for the butelase-mediated ligation reaction. The ligation reaction was performed individually with peptides KALVXXHV and HIGGIR. In each condition, the final concentration of butelase 1, KALVXXHV, HIGGIR were 0.1 μ M, 0.05 mM and 1 mM respectively. The intermolecular ligation reaction was performed in an optimized condition with 1mM EDTA at pH 6.5, 42 $^{\circ}$ C. β -mercaptoethanol was not included in the reaction buffer. At 10 min and 2 h, 25 μ l reaction mixtures were quenched with 250 μ l 0.1%TFA. Then it was analyzed by RP-HPLC using a Phenomenex analytical column (150mm \times 4.6mm). The products were verified by MALDI-TOF MS followed by MS/MS. The ligation yields were calculated based on the HPLC profile using the Lab Solution software.

4. Design and engineering of cysteine-rich peptides

4.1 Synthesis of SFBK, BKSF

Cl-Trt(2-Cl) resin (1.0 g, 1.2 mmol) was swelled with DCM for 30 min and washed with DCM (5 ml x 3). Hydrazine monohydrate (939 μ l, 12.0 mmol) and DIEA (627 μ l, 3.6 mmol, MW 129.2, *d* 0.742 g/ml) in DMF (20 ml) were added, and the suspension was shaken for 30 min at 25 °C. The resin was filtrated, washed with DMF (5 ml x 3). Fmoc-Leu-OH (1.70 g, 4.8 mmol, MW 353.4), BOP (2.12 g, 4.8 mmol, MW 442.3) in DMF (20 ml) and DIEA (1.25 ml, 7.2 mmol, MW 129.2, *d* 0.742 g/ml) were added to the resin and the suspension was shaken for 1 h. The reaction was repeated once. The resin was filtrated, and washed with DMF (5ml x 3), DCM (5 ml x 3), and diethyl ether (5 ml x 3), dried *in vacuo* (1.252g, 63%).

The following amino acids were introduced using BOP/DIEA with a ratio of 4:6 at 25°C for 30 min for coupling reaction and 20% piperidine in DMF for Fmoc deprotection. Standard cleavage using TFA, TIS H₂O afforded SFBK hydrazide (431 mg). MALDI-TOF MS was used to analyze the product (observed 1954.84, theoretical molecular weight: 1955.06 [M+H]⁺).

The crude SFBK hydrazide (150 mg, 76.7 μ mol) was dissolved in 0.2 M sodium phosphate buffer (pH 3, 20 ml) together with NaNO₂ (53.0 mg, 767 μ mol, MW 69.0). The reaction was performed at 0 °C for 20 min. MMA (686 μ l, 7.67 mmol, MW 106.1, *d* 1.187 g/ml) was added, and the pH was adjusted by addition of 1 M NaOH to 7. The reaction mixture was stirred for 7 h and TCEP (220 mg, 0.767 mmol, MW 250.2) was added. After the cyclization reaction was completed, it was quenched with 1 M HCl and purified by RP-HPLC using a Phenomenex preparative column to give reduced cyclic SFBK (26.5 mg, 18%). MALDI-TOF MS was used to analyze the product (observed 1923.03, theoretical molecular weight: 1923.03 [M+H]⁺).

Reduced cyclic SFBK (6.4 mg, 3.33 μ mol) was dissolved in 50% AcOH (3 ml) together with 0.1 M I₂/MeOH (33 μ l, 3.33 μ mol). The reaction mixture was stirred for 1 min, quenched with 1 M ascorbic acid (17 μ l, 16.7 μ mol), purified by RP-HPLC

using the preparative column to give SFBK (4.4 mg, 69%) (observed 1920.90, theoretical molecular weight: 1921.01 [M+H]⁺).

BKSF was synthesized using the same method as for SFBK. Standard cleavage using TFA, TIS H₂O afforded BKSF hydrazide (333 mg). MALDI-TOF MS was used to analyze the product (observed 1790.77, theoretical molecular weight: 1790.89 [M+H]⁺). Thioesterification afforded reduced cyclic BKSF (27.4 mg, 19%). MALDI-TOF MS was used to analyze the product (observed 1758.69, theoretical molecular weight: 1758.85). Then it was oxidized by iodine to give SFBK (4.2 mg, 74%). MALDI-TOF MS was used to analyze the product (observed 1756.50, theoretical molecular weight: 1756.83 [M+H]⁺).

4.2 Synthesis of SFBK-L, BKSF-L, DALK

Wang resin (119 mg, 0.1 mmol, 0.84 mmol/g) was swelled with DCM for 30 min and washed with DCM (5 ml x 3). A solution of Fmoc-Leu-OH (141 mg, 0.4 mmol) and 4-dimethylaminopyridine (DMAP) (12.0 mg, 0.1 mmol, MW 122.2) in DMF (15 ml) and DIC (62 µl, 0.4 mmol, MW 126.2, *d* 0.806 g/ml) were added, and the suspension was shaken for 2 h at 25 °C. The resin was filtrated, washed with DMF (5 ml x 3) to give Fmoc-Leu-Wang resin.

Peptide elongation was carried out starting from Fmoc-Leu-Wang resin using 4eq. BOP, 6 eq. DIEA at 25 °C for 30 min for coupling reaction and 20% piperidine in DMF for Fmoc deprotection. The protected resin was treated with TFA/TIS/H₂O (4.5 ml/250 µl/250 µl) for 2 h. The reduced SFBK-L (66.6 mg) was isolated in 33% yield. MALDI-TOF MS was used to analyze the product (observed 1941.84, theoretical value: 1941.04 [M+H]⁺).

Reduced SFBK-L (30.0 mg, 15.4 µmol) was dissolved in 50% AcOH (15.4 ml) together with 0.1 M I₂/MeOH (154 µl, 15.4 µmol). The reaction mixture was stirred for 1 min, quenched with 1 M ascorbic acid (77 µl, 77.0 µmol), purified by RP-HPLC give SFBK-L (18.1 mg, 61%). MALDI-TOF MS was used to analyze the product (observed 1940.07, theoretical molecular weight: 1939.02 [M+H]⁺).

Reduced BKSF-L was synthesized using the same procedure as that used in reduced SFBK-L. The reduced BKSF-L (29.9 mg) was isolated in 17% yield. MALDI-TOF MS was used to analyze the product (observed 1776.94, theoretical molecular weight: 1776.86 [M+H]⁺). Reduced BKSF-L was then oxidized by iodine to give BKSF-L (22.7 mg, 76%). MALDI-TOF MS was used to analyze the product (observed 1774.82, theoretical molecular weight: 1774.84 [M+H]⁺).

For synthesis of DALK, peptide elongation was carried out on Wang-Leu resin manually using Fmoc amino acid/HBTU/DIEA 4/4/6 eq. in DMF for 30 min. The protected peptide was treated with TFA/TIS/H₂O (4.5 ml/250 μ l/250 μ l) for 2 h. 30 mg of the crude peptide was purified by RP-HPLC to give DALK (12.2 mg, 40%). MALDI-TOF MS was used to analyze the product (observed 998.54, theoretical molecular weight: 998.57 [M+H]⁺).

4.3 Heat and pH stability assay

For the heat stability, all six peptides SFBK, BKSF, SFBK-L, BKSF-L, SFTI-1 and DALK were dissolved in a pH 7 phosphate buffer (0.1 M) at a final concentration of 0.1 mM. Each solution was subjected to 100 °C treatment in a PCR machine with heat-cap for 1 h. The same peptides were kept at 25 °C for the same duration as controls. At various time intervals, aliquots were taken and analyzed by Shimadzu Nexera UPLC. Fmoc-Gly-OH was used as an internal standard. The retention time and MS of the peptide were compared before and after heat treatment. Reactions were repeated 3 times. The peptide integrity was calculated based on the UPLC profile using the Lab Solution software.

For the pH stability, all six peptides SFBK, BKSF, SFBK-L, BKSF-L, SFTI-1 and DALK were subjected to 0.2 M aqueous HCl to a final concentration of 0.1 mM at 25 °C. The same peptides were dissolved in pH 7 phosphate buffer (0.1 M) for the same duration as controls. At various time intervals, aliquots were taken and analyzed by Shimadzu Nexera UPLC. Fmoc-Gly-OH was used as an internal standard. The retention time and MS of the peptide was compared before and after heat treatment. Reactions were repeated 3 times. The peptide integrity was calculated based on the UPLC profile using the Lab Solution software.

4.4 Pepsin and trypsin stability assay

Peptides (0.2 mg) together with pepsin or trypsin (8 μ l, 0.5 μ g/ μ l in *Milli-Q* H₂O) were dissolved in 1 ml sodium citrate buffer (100 mM, pH 2.5) or 1 ml ammonium bicarbonate (NH₄HCO₃) buffer (100 mM, pH 7.8) with an enzyme peptide ratio of 1:50 and incubated at 37 °C. At various time intervals, aliquots were taken and analyzed by Shimadzu Nexera UPLC. Fmoc-Gly-OH was used as an internal standard. The retention time and MS of the peptide were compared before and after heat treatment. Reactions were repeated 3 times.

4.5 Serum stability assay

We used 25% human serum (Sigma Aldrich) in Roswell Park Memorial Institute (RPMI) medium to perform the serum stability assay. To remove the lipid components, serum was centrifuged at an rcf of 18,000 g, and then 1 ml supernatant was allocated to a 1.5 ml Eppendorf tube and incubated at 37 °C for 15 min before use. Peptide stocks were diluted in serum to make a final concentration of 1mg/ml. At various time intervals, 100 μ l aliquots were removed and added to 200 μ l 90% ethanol to precipitate the serum proteins. Then the mixture was cooled at 4 °C then centrifuged at an rcf of 18,000 g for 10 min. The pellet was discarded and the supernatant was analyzed by Shimadzu Nexera UPLC. Fmoc-Gly-OH was used as an internal standard. The retention time and MS of the peptide were compared before and after heat treatment. Reactions were repeated 3 times. The peptide integrity was calculated based on the UPLC profile using the Lab Solution software.

4.6 Trypsin inhibition assay

To test the trypsin inhibition effect of peptides, we conducted a trypsin inhibition assay in 96-well plate with a total volume of 100 μ l per well. Each well contained 94 μ l of 100 mM ammonium bicarbonate solution and 1 μ l 0.5 μ g/ μ l trypsin. First, peptides were diluted to each well to reach a concentration of 0-100 μ M (0, 0.5, 1.0, 10.0, 50.0 and 100.0 μ M). The reaction plate was incubated at 37 °C for 15 min. Just before monitoring by the spectrometer, 5 μ l of 20 mM N α -Benzoyl-DL-arginine 4-nitroanilide hydrochloride (BAPNA) was added to each well by a multi-channel pipette and the mixture was mixed well quickly. This step has to be fast and the

microplate reader should be ready for measuring before adding the substrate. The process of the reaction was monitored by an ELISA microplate reader (Tecan-2) with an excitation wavelength of 405 nm and a reference wavelength of 600 nm, respectively, with a 10 nm bandwidth each. The reaction kinetic was measured over 10 min. The absorbance of each well at 405 nm was recorded and plotted against the time and concentration of each peptide. Negative control contained 94 μl 100 mM ammonium bicarbonate solutions, 1 μl 0.5 $\mu\text{g}/\mu\text{l}$ trypsin and 5 μl substrate without addition of peptides. In blank, the mixture contained 95 μl of a 100 mM ammonium bicarbonate solution, 5 μl substrate and trypsin was not added. All the reactions were repeated 3 times. The percentage of inhibition was calculated using the formula below:

$$\text{Trypsin activity} = \frac{A_{405 \text{ nm}}(\text{test}) - A_{405 \text{ nm}}(\text{blank})}{A_{405 \text{ nm}}(\text{negative control}) - A_{405 \text{ nm}}(\text{blank})} * 100\%$$

4.7 Cytotoxicity

Hela cells were seeded into a 96-well plate with 15,000 cells per well. SFBK and BKSF were each diluted in Dulbecco's Modified Eagle Medium (DMEM) and 6% Fetal Bovine Serum (FBS), and 100 μl were transferred to each well. The positive control was loaded with 100 μl of 0.2% Triton X-100 from Sigma Aldrich in 100% Phosphate Buffered Saline (PBS). The negative control contained 100 μl DMEM with 6% FBS without peptide. The 96-well plate was then incubated at 37°C for 24 h. After 24 h, 10 μl of PrestoBlue™ Cell Viability Reagent was added to each well for 2 h. The 96-wells plate was then loaded into a fluorescent microplate reader (Tecan-2) to record its fluorescence with an excitation wavelength of 560 nm and an emission wavelength of 590 nm at 10 nm bandwidth each. The fluorescence values were used to calculate the total viability of Hela cells.

4.8 Fluo-4 NW calcium assay

Hela cells were seeded into a 96-well plate with 30,000 cells per well. A dye loading buffer was prepared by adding 10 ml of Hanks' balanced salt solution with 20 mM HEPES buffer solution and 100 μl of 250 mM probenecid solution. DMEM was removed and 100 μl of the dye-loading buffer is added to each well on the 96-well

plate. 1 ml dye fluo-4 NW was added into each well. The 96-well plate was then incubated at 37 °C for 30 min. Then various concentrations of peptides (0.01, 0.1, 1 μM) were added and the reaction was incubated at 37 °C for another 30 min at 25 °C. Subsequently, the fluo-4 NW dye was removed and each well was washed by dye-loading buffer. The bradykinin B1 antagonists with various concentrations were then added into the well and incubated for 2 min followed by washing with PBS. Then 100 nM Des-Arg⁹-BK as the B1 agonist was added. The 96-well plate was immediately loaded into an inverted microscope (Nikon) to record its fluorescence with an excitation wavelength of 494 nm and an emission wavelength of 516 nm. Exposure time was set at 0.5 sec per wavelength. Fluorescence was recorded immediately after adding DALK, SFBK and BKSF, followed by Des-Arg⁹-BK into the well. The relative Ca²⁺ concentration was calculated by the fluorescence intensities.

Chapter 3 Chemical Synthesis of Cyclic CRPs

1. Introduction

Our laboratory has a long-standing interest in cyclic CRPs because of their impressive stability and potential oral availability. The detailed structure-function relationship and pharmacological evaluation of them require a large amount of samples with high purity, and which would be difficult to obtain by isolation and purification. In this chapter, I will describe my results on producing cyclic CRPs by chemical methods.

As described earlier in chapter 1, the chemical synthesis of macrocyclic peptides requires preparing a peptide thioester precursor. Conventional methods to generate peptide thioesters by Fmoc chemistry can be generally characterized into the “safety catch” and “safety switch” categories. Various thioester surrogates have been developed based on these two strategies. They normally require complicated preparation steps of a thioester linker or reaction conditions can be troublesome. Thus, a novel method to improve the design of a thioester linker and simplify its preparation would be desired.

Formation of amide bonds through intein-mediated transpeptidation inspired our design for a thioester linker. The intein-mediated transpeptidation formed an amide bond to link two proteins together through a series of acyl shift reactions. We envisioned that we could design a thioester surrogate (less than one amino acid), which can undergo the same acyl shift reactions, to mimic the intein-mediated process.

Cysteine residue (or sometimes serine and threonine) is a functional group in the intein-mediated transpeptidation and the thioethylamido (TEA) moiety of cysteine is the essential group which mediates the acyl shift reactions. Thus, our laboratory designed three thioester surrogates based on the TEA moiety which can mediate N-S acyl shift reactions to generate peptide thioesters (**Figure 3.1**). My thesis focused on the latter two linkers: *N*-methyl Cys (MeCys) and thioethylbutylamide (TEBA). The thiomethylthiazolidine (TMT) group can mediate N-S acyl shift reactions under strongly acidic conditions containing 0.25% TFMSA in TFA and generate a peptide

thioester concurrent with the deprotection of side chain protecting groups and resin linkage.²⁴³ However, the preparation of the resin can be tedious. To avoid multiple steps in the preparation process, I employed two other thioester surrogates MeCys and TEBA which are commercially available with minimal preparation.

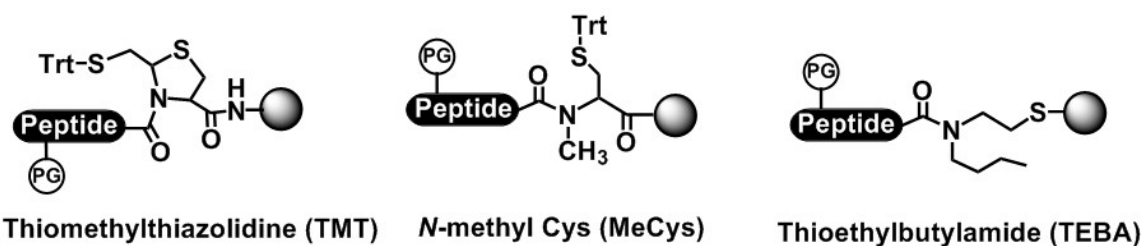


Figure 3.1 Thioester surrogates containing the essential TEA moiety. All the three thioester surrogates are Fmoc-compatible. PG represents side chain protecting groups in Fmoc SPPS.

2. MeCys as a thioester surrogate for preparation of cyclic CRPs

2.1 Synthetic strategy using MeCys peptides for preparing cyclic CRPs

A general design of “Cys-peptide-MeCys” is employed for our synthesis of cyclic peptides, which is represented in **Figure 3.2**. In such a design, a linear peptide **1**, with a Cys and a MeCys residue located at its N- and C-terminus, respectively, was used as the precursor. The C-terminal MeCys residue in peptide **1** can mediate acyl shift reactions for preparing a cyclic peptide from its linear precursor in a one-pot condition. First, an N-S acyl shift mediated by MeCys under acidic conditions forms a thioester **2**. Second, the N-terminal cysteinyl thiol will initiate a nucleophilic attack on the carbonyl group of the C-terminal thioester to form an end-to-end thiolactone **3** via thiol-thioester exchange. Finally, a proximity-driven S-N acyl shift reaction occurs spontaneously forming a macrocyclic peptide **4**. The scheme proposed here exploits the isomerization between an amide and a thioester to break the amide bond through an N-S acyl shift and form a new amide bond through an S-N acyl shift.

The linear peptide **1** can be conveniently prepared from commercially available Fmoc-MeCys(Trt)-OH as a thioester surrogate attached to the resin and followed by stepwise Fmoc SPPS. The synthesis will not follow the original sequence of the peptide precursor and cysteine must be chosen as the N-terminal amino acid. Since multiple cysteines are present in CRPs, we can choose the most unhindered Xaa-Cys junction as the ligation site and synthesize the peptide in the C- to N-terminus direction. As an example, we used the cyclic peptide SFTI-1 to demonstrate the feasibility of our “Cys-peptide-MeCys” strategy. SFTI-1 is a 14-amino-acid peptide with potent inhibitory effect against trypsin. It was isolated from sunflower seeds and contains one disulfide bond.⁶⁷

Among three acyl shift reactions involving in the synthesis process, the S-N acyl shift reaction is uncatalyzed and occurs spontaneously without any assistance. Thus, our efforts focused mainly on optimizing the N-S and S-S acyl shift reaction. To investigate mechanistic insights on MeCys-mediated N-S and S-S acyl shifts, model peptide TIGGIR-MeCys-Gly was prepared to determine the optimal conditions for forming a thioester.

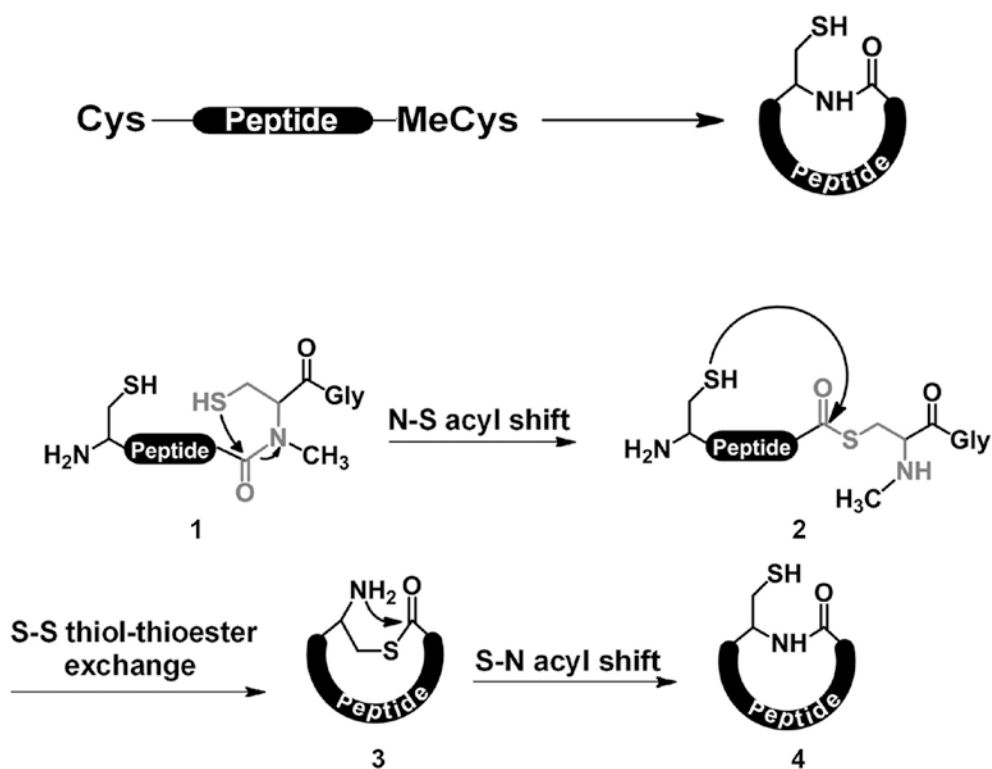


Figure 3.2 The “Cys-peptide-MeCys” strategy for synthesis of cyclic peptides. The model peptide bearing a Cys at the N-terminus and a MeCys at the C-terminus will undergo a series of acyl shift reactions to form the end-to-end cyclic backbone.

2.2 Synthesis of model peptides on the MeCys resin by Fmoc SPPS

To prepare the model peptide **9a** TIGGIR-MeCys-Gly-NH₂ and SFTI-1 precursor **9b**, we started to synthesize the MeCys resins (**Figure 3.3**). It commenced with the introduction of Fmoc-Gly-OH to the Rink amide resin using DIC/HOAt. The Gly residue was used as a spacer for reasons which will be discussed later in this section. After deprotecting the Fmoc group of Gly with 20% piperidine, Fmoc-*N*-methyl-Cys(Trt)-OH was coupled to the resin using PyBOP/DIEA to give the MeCys resin **6**. The following coupling reaction between Fmoc-Arg(Pbf)-OH and the MeCys residue was difficult due to the low reactivity of secondary amine in MeCys and hindrance of the bulky protecting group in Arg. Thus, a double or triple coupling by HATU/DIEA was required for completion of the reaction. The reaction process was monitored by the acetaldehyde/chloranil test to confirm the absence of free secondary amine on the resin.

For the synthesis of **9a**, the rest of peptide TIGGI was assembled on the Arg-preloaded resin **7** manually using PyBOP/DIEA as coupling reagents. Fully protected peptide was finally cleaved by TFA/TIS/H₂O to afford **9a** for kinetic studies of the N-S acyl shift reaction.

For the synthesis of **9b**, the Arg-preloaded resin **7** was subjected to an automatic microwave-assisted peptide synthesizer for peptide chain elongation. In SFTI-1 sequence cyclo-(G¹RCTKSIPPICFPD¹⁴), two cysteine junctions Arg-Cys and Ile-Cys were able to serve as the ligation site. To prevent the steric hindrance during preparation of thioesters and cyclization reactions, Arg was chosen over the β -branched amino acid Ile as the ligation site. Therefore, the peptide chain assembly followed the sequence CTKSIPPICFPDGR. The sequence contained a base-sensitive Asp-Gly (D-G) bond, which was prone to aspartimide formation during the deprotection of Fmoc groups. To suppress this accumulative side reaction, we used a modified deprotection solution containing 20% of a milder base morpholine instead of commonly used piperidine.²⁷⁸ The aspartimide formation was minimized to less than 1% after the completion of the synthesis. The unprotected peptide **9b** was obtained in 17% isolated yield based on the initial loading of the Rink amide resin.

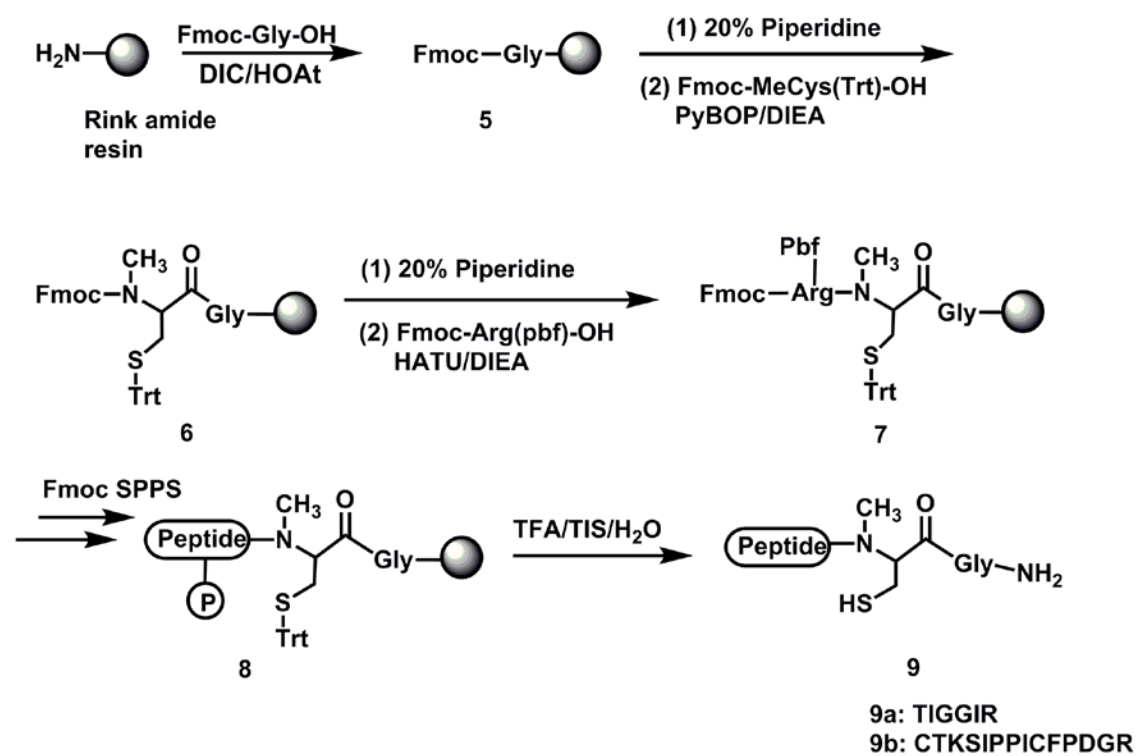


Figure 3.3 Synthesis of the model peptides using MeCys resins. P: side chain protecting groups. Model peptide **9a**: TIGGIR-MeCys-Gly-NH₂; **9b**: CTKSIPPICFPDGR-MeCys-Gly-NH₂.

The initial introduction of MeCys residue on Wang resin was not successful as it produced a significant amount of truncated peptides due to diketopiperazine (DKP) formation. Peptides containing a C-terminal proline or *N*-alkylated amino acids are prone to give a diketopiperazine product and it is especially problematic with benzyl ester type of resins like Wang resin.²⁷⁹ Therefore, Rink amide resin was chosen as the anchor for synthesizing the MeCys resin. We also used Gly as a spacer to prevent the DKP side reaction during synthesis. The other advantage of employing Gly as a spacer was that it could minimize the β -elimination during the synthesis process. Lukszo and co-workers reported the 3-(1-piperidinyl)alanine formation during synthesis of a C-terminal MeCys(Trt) peptide by Fmoc SPPS in 1996.²⁸⁰ The reaction contains two steps (**Figure 3.4**). The first reaction is a base-catalyzed β -elimination reaction induced by piperidine on the C-terminal cysteine residue to form a dehydroalanine derivative and the second reaction involves a nucleophilic addition of piperidine to form the side product 3-(1-piperidinyl)alanine. It resulted in a +51 Da side product in matrix-assisted laser desorption ionization mass spectrometry (MALDI-TOF MS) regardless of the protecting groups on Cys. We observed the same side product in both HPLC and MS profile when we coupled Fmoc-MeCys(Trt)-OH directly to the resin. The side product accumulated during the peptide chain assembly. Unfortunately, Lukszo and co-workers did not provide an efficient solution for minimizing this side reaction in their paper. We successfully suppressed the side reaction by using Gly as a spacer but did not resolve it completely. The β -elimination reaction occurred at the C-terminal cysteine that linked directly to the resin. By introducing a glycine as a spacer, the cysteine is no longer directly linked to the resin. Thus, the acidity of the α -carbon proton in cysteine is affected and the abstraction of α -carbon proton prevented to certain extent. The piperidinyl alanine product was reduced to <1% and 3% in our synthesis of peptide **9a** and **9b**, respectively. It suggested that this reaction did not only occur on the C-terminal cysteine but also on cysteine close to the C-terminus.

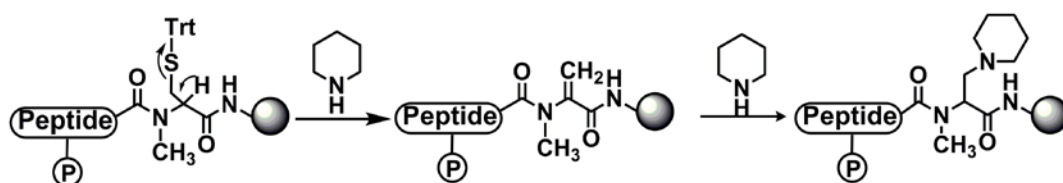


Figure 3.4 The mechanism of 3-(1-piperidinyl)alanine formation on the C-terminal MeCys. P represents side chain protecting groups. The MeCys is protected with Trt group in the scheme. MeCys(Trt) was converted to a dehydro-MeAla by the piperidine-induced β -elimination. Then the nucleophilic addition of alkenes resulted in the 3-(1-piperidinyl)alanine.

2.3 MeCys-mediated N^α-thioester formation in aqueous conditions

The thioester form and the amide form are in equilibrium under aqueous acidic conditions due to the reversibility of MeCys-mediated N-S acyl shift reaction. To trap the newly generated thioester and make the reaction irreversible, additional thiols are included to mediate an S-S acyl shift and afford a stable thioester that cannot be converted back to the amide form. To simplify the process and perform the reaction in a one-pot condition, we conduct a tandem thiol switch reaction including both the N-S and the S-S acyl shift.

2.3.1 Synthesis of H-TIGGIR-MES

The model peptide TIGGIR-MeCys-Gly-NH₂ **9a** was synthesized as described earlier to examine the suitable acidity of this reaction. It was in equilibrium with the peptide thioester **10a** through the N-S acyl shift reaction. To shift the equilibrium, we added sodium 2-sulfanylethanesulfonate (MESNa) which would react with via thiol-thioester exchange to form MES thioester **11a** (**Figure 3.5**). This reaction was quenched and monitored by RP-HPLC at different time intervals. It was found that the thioester formation equilibrated after 24 h at different pH. All reaction solutions with different pH were prepared with 0.1 M sodium phosphate buffers except pH 1 and 2 which were prepared using diluted sulfuric acid. To conduct the reaction, peptide **9a** was dissolved in various reaction buffers at a 5 mM concentration in presence of 50 eq. MESNa and the mixture was incubated at 40 °C. The optimal pH range for the tandem thiol switch reaction was pH 2-5 with the MES-thioester **11a** obtained in > 90% yield as shown in **Figure 3.5**. The best yield was obtained at pH 2, which afforded 95% **11a**. At low pH (pH < 2), no **9a** was detected suggesting the N-S acyl shift was favored. The presence of the thioester intermediate **10a** indicated that the S-S exchange reaction was the rate-limiting step in tandem thiol switch. At pH > 5, the starting material **9a** predominated which suggested that the N-S acyl shift reaction was the rate-limiting step in tandem thiol switch. It should be noted that side reactions occurred at either low or high pH. At pH 1, a deamidation reaction occurred at the C-terminal Gly residue of **10a**. We found that the deaminated product **10c** was 9% at 24 h. At pH 6 and 7, thioesters **11a** were not stable and hydrolyzed to side product TIGGIR-OH during prolonged treatment. Collectively, the MeCys-mediated thioester formation via tandem thiol switch could be performed at a broad pH range (pH 2-5)

with acceptable yield (>80%). The optimal pH for MeCys-mediated N^α-thioester formation was pH 2.

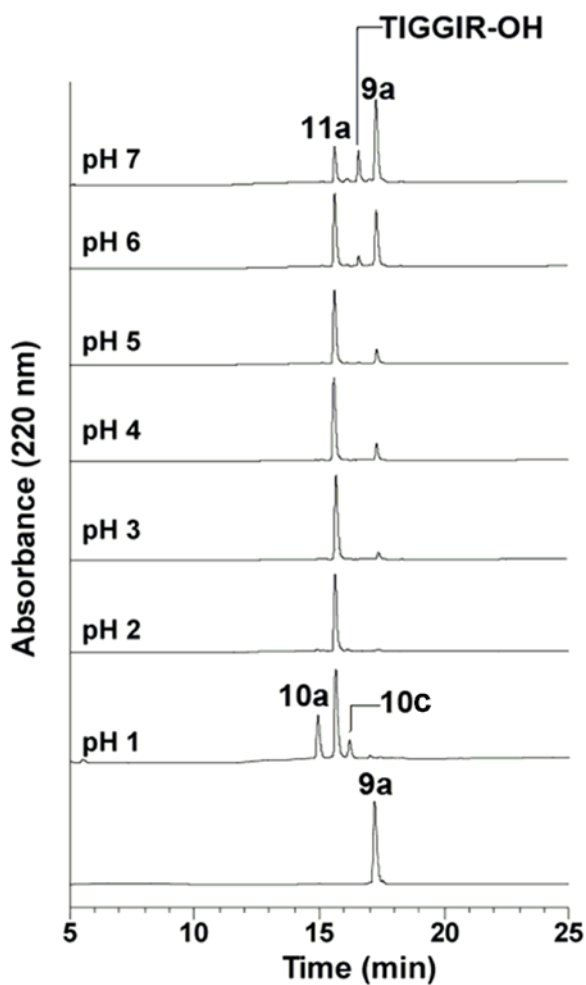
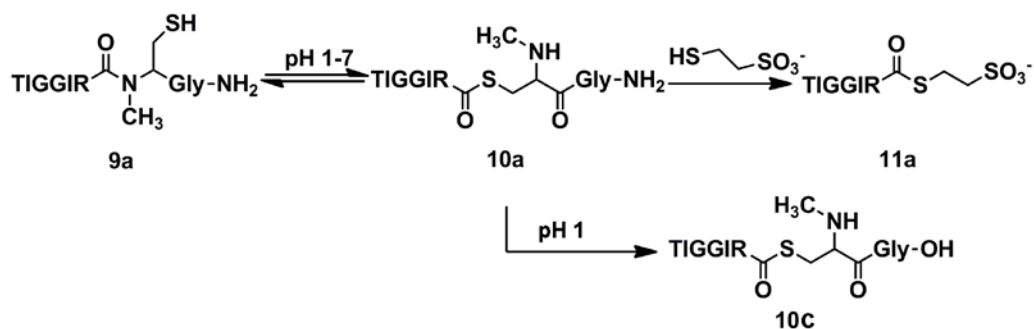


Figure 3.5 Reaction scheme and HPLC profiles of MeCys-mediated tandem thiol switch reaction at various pH. Products contained the amide form **9a**, the thioester form **10a**, the stable MES thioester **11a**, and two side products **10c**: deamidation form of **10a**; hydrolysis product TIGGIR-OH occurring at pH 1 and 7, respectively. The reaction mixture was quenched and subjected to HPLC at 24 h. HPLC gradient: 0-60% ACN in aqueous 0.1% TFA in 30 min with flow rate of 1 ml/min.

2.3.2 Ligation of TIGGIR-MES and CALVIN-NH₂

To confirm its validity, the thioester **11a** was then subjected to a chemical ligation with the model peptide Cys-Ala-Leu-Val-Ile-Asn-OH (CALVIN-NH₂) **12** (**Figure 3.6**). The reaction was conducted with 1 mM TIGGIR-MES and 3 mM peptide CALVIN-NH₂ in presence of 6 M Gnd-HCl, 50 mM external thiol methyl mercaptoacetate (MMA) and 20 mM TCEP at pH 7.5. Peptide **12** was used in excess to capture the thioester preventing the hydrolysis of thioesters. We monitored the reaction at various time intervals by RP-HPLC and ESI-MS. The thioester **11a** was able to ligate with peptide **12** smoothly to give the ligation product TIGGIRCALVIN-NH₂ **13** and the reaction completed after 4 h. We also observed significant amount of the TIGGIR-MMA at the early stage of the reaction (refer to 10 min HPLC profile). It was due to the thiol-thioester exchange between the thioester TIGGIR-MES and the external thiol MMA. Since MMA was more reactive than MES in the ligation reaction, we suspected that TIGGIR-MES was first converted to TIGGIR-MMA which was an intermediate in this ligation reaction. Both TIGGIR-MES (minor) and TIGGIR-MMA (major) reacted with the peptide CALVIN-NH₂ to give the ligation product **13**. The ligation product **13** was confirmed by MS/MS with each fragment identified. (**Figure 3.7**)

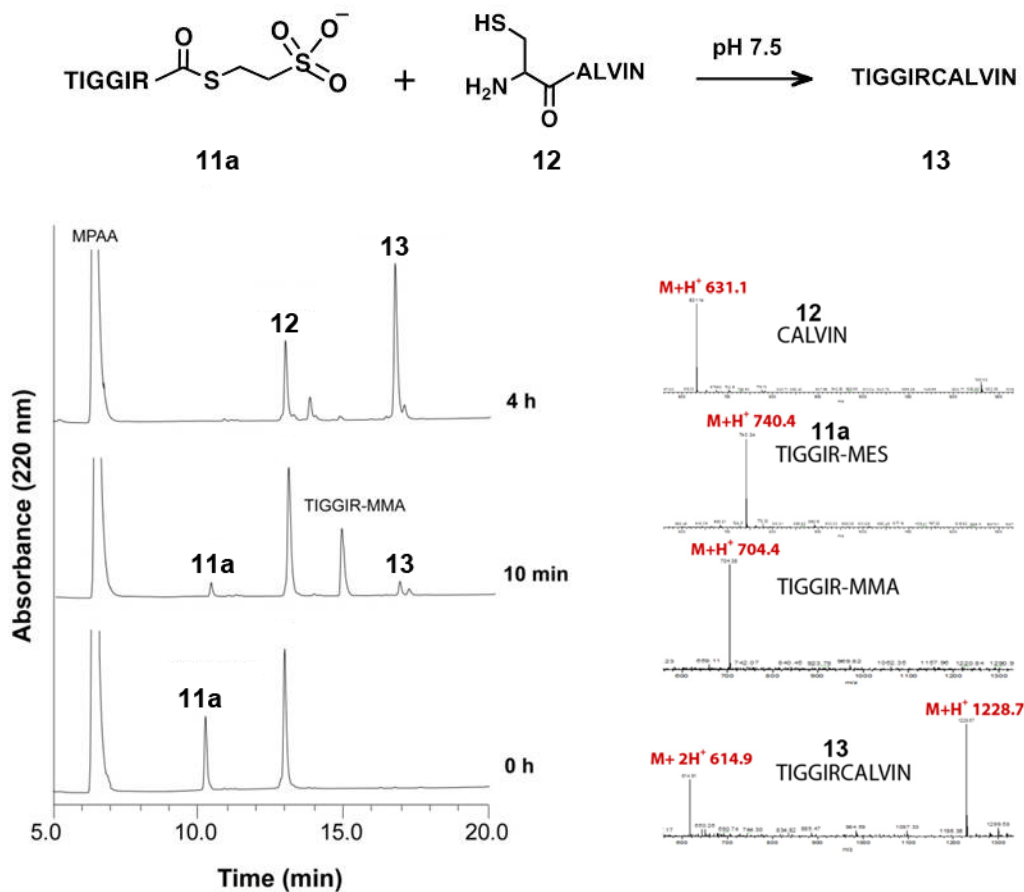


Figure 3.6 Scheme, HPLC and ESI profile of TIGGIR-MES and CALVIN-NH₂ peptide ligation.

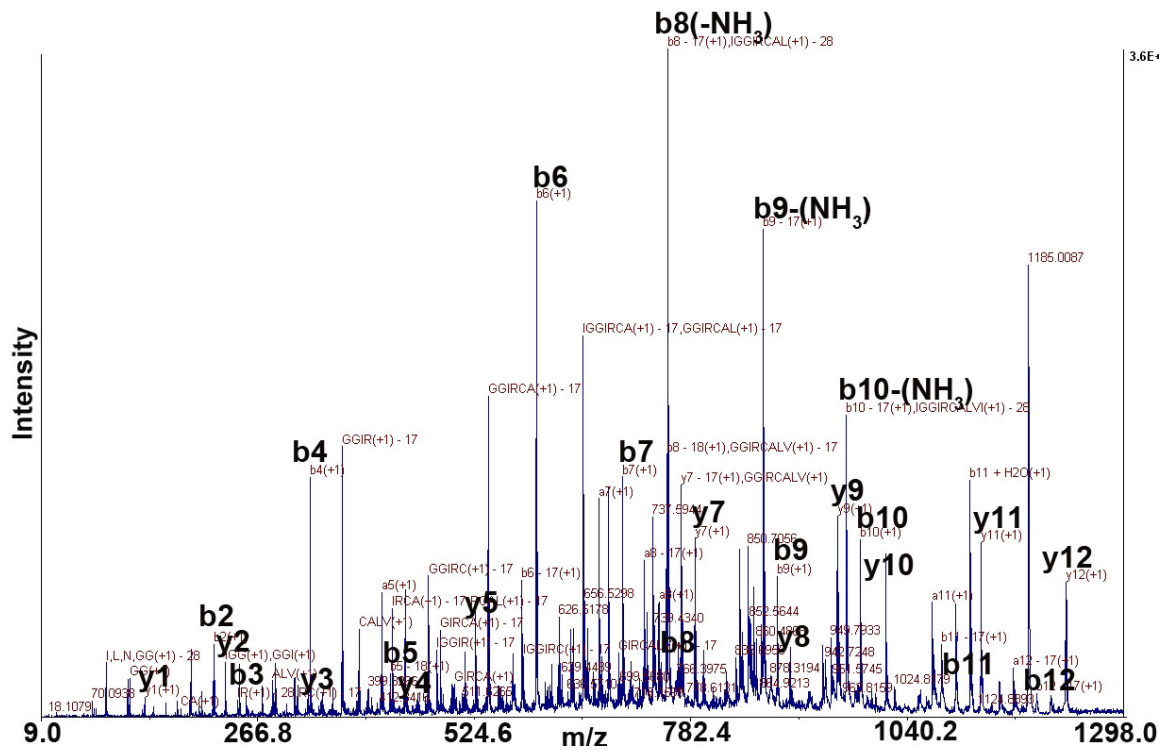


Figure 3.7 MS/MS profile of ligation product TIGGIRCALVIN-NH₂. Fragments are labeled on the top of the corresponding m/z peak.

2.3.3 One-pot ligation of TIGGIR-MeCys-Gly-NH₂ and peptide CALVIN-NH₂

Next, we tested whether the tandem thiol switch reaction could be used directly in the ligation reaction by conducting a one-pot *in situ* ligation. The starting material **9a** would undergo the N-S acyl shift reaction as described earlier. The newly generated thioester **10a** would not only be trapped by an external thiol but also by the cysteinyl peptide **12** to form the ligation product.

In this reaction, the peptide TIGGIR-MeCys-Gly-NH₂ **9a** was dissolved in 0.1 M phosphate buffer (pH 5) with 6 M Gnd-HCl and 20 mM TCEP to prepare a 1 mM solution. Then 3 eq. cysteinyl peptide **12** was added together with 50 eq. MMA and reaction mixture was incubated at 40 °C. Since pH 2-5 was shown to be optimal for the MeCys-mediated tandem thiol switch reaction and the effective pH for thioester-Cys ligation was neutral to basic, the reaction was performed at neutral to slightly acidic condition (pH 5). The ligation product **13** was able to form after 24 h with <30% yield (**Figure 3.8**). The one-pot ligation was shown to be slow under the proposed condition without microwave which is usually employed in the known one-pot *in situ* ligation conditions.^{241,281} The reason we did not utilize the microwave condition is that it may introduce side reactions such as fragmentation at Cys, aspartimide formation and bond breakage at Asp-Pro.

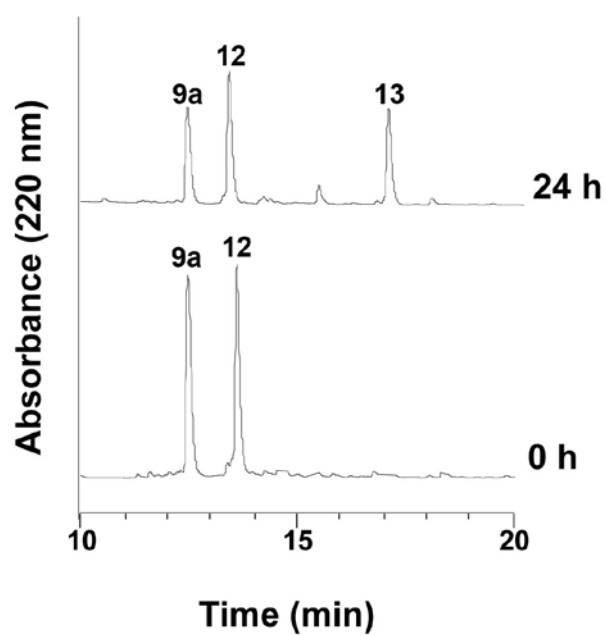
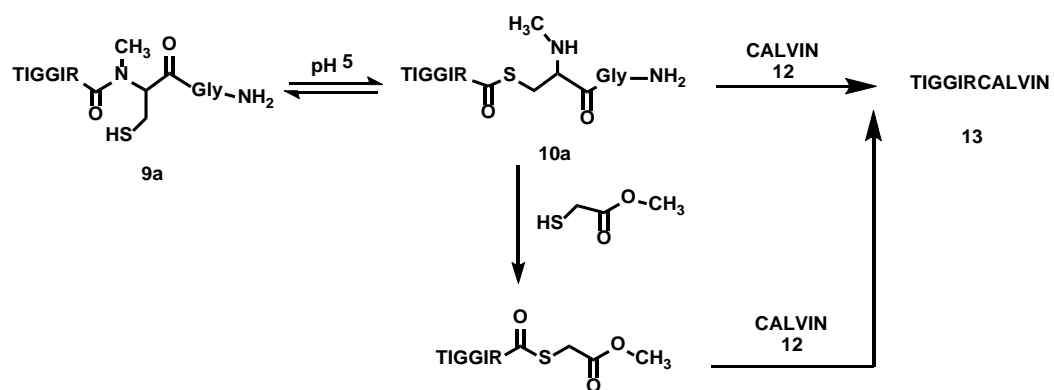


Figure 3.8 One-pot chemical ligation of TIGGIR-MeCys-Gly-NH₂ and CALVIN-NH₂ with external thiol MMA at pH 5 and 40 °C

2.3.4 Comparison of N^α-Cleavage at Gly-Cys or Gly-MeCys bond

Certain amide bonds are known to be fragile and susceptible to hydrolysis under aqueous acidic conditions. Particularly those able to undergo acyl shift reactions, such as bonds containing Cys, Ser or Thr residues, are prone to mediate rearrangement in certain circumstances¹⁴⁹. They are involved in the intein-mediated protein splicing for ligations of extein fragments^{145,146}. Our method exploiting the reversibility of amide bond isomerization to generate thioester utilized a special amino acid MeCys. Although Hojo and co-workers developed a series of thioester surrogates based on *N*-alkylated cysteine, detailed studies on isomerization rate of MeCys were missing. The isomerization of Cys was examined by Aimoto and co-workers using NMR.²³⁴ They reported the isomerization reaction of Cys in a Gly-Cys bond under TFA. Through the N-S acyl shift reaction, the Gly-Cys amide bond isomerized into a thioester bond with a slow rate of 0.09% per hour at room temperature. To assess the difference of Cys and MeCys in the thioester formation, two model peptides TIGGIR-Gly-MeCys-Gly-NH₂ and TIGGIR-Gly-Cys-Gly-NH₂ were synthesized to conduct the tandem thiol switch reaction.

Different from the previous conditions, 20% methyl mercaptoacetate (MMA) was used instead of 50 eq. MESNa for two reasons. First, MES thioesters are hydrophilic, thus the product and the starting material were poorly separated by RP-HPLC if we used MESNa as the external thiol. Second, since MMA is more reactive than MESNa, it would facilitate the tandem thiol switch of Gly-Cys bond, which was too slow to monitor in conditions using MESNa.

The tandem thiol switch contained two reactions: an N-S acyl shift and a thiol-thioester exchange reaction. The first N-S acyl shift followed the first-order rate equation, and since the external thiol was used in large excess for the second thiol-thioester exchange reaction, the overall rate for the tandem thiol switch reaction can be calculated as a first-order reaction in the initial stage. The yield of the thioester TIGGIR-Gly-MMA was calculated to compare the isomerization of Gly-Cys or Gly-MeCys bond (**Figure 3.9**).

Under our experimental conditions, MeCys was 150 fold faster than Cys in mediating the tandem thiol switch reaction based on the initial rates. The rate constant

was 0.31 min^{-1} for MeCys and 0.002 min^{-1} for Cys. In the reaction process, the Gly-Cys bond was comparably stable and did not result in >10% thioester after 3 h whereas 60% thioester was obtained in the MeCys peptide at the same time point. A prolonged treatment until 48 h did not lead to a completed reaction in the Cys peptide and side reactions such as disulfide formation began to accumulate even in presence of 20 mM TCEP. In contrast, the reaction completed within 9 h for the MeCys peptide. These results clearly show that MeCys facilitates the N-S acyl shift and greatly increases the isomerization rate compared with Cys. They also suggest that our approach is suitable for synthesizing cysteine-rich peptides. Under our experimental conditions, internal cysteine linkages were considerably safe without any undesired fragmentation.

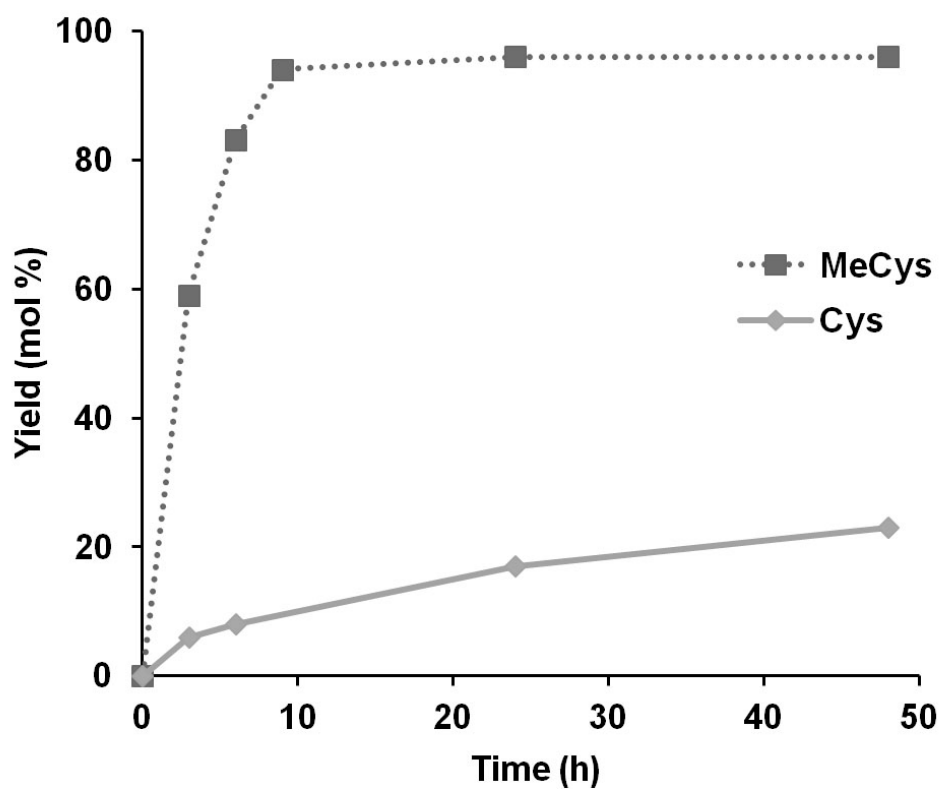


Figure 3.9 Comparison of N^α-Cleavage at a Gly-MeCys or a Gly-Cys bond. TIGGIR-Gly-MeCys-Gly-NH₂ and TIGGIR-Gly-Cys-Gly-NH₂ were treated with 20% MMA (v/v) in presence of 20 mM TCEP at pH 2, 40°C. The Y axis represented the yield of TIGGIR-Gly-MMA.

2.4 MeCys-mediated synthesis of cyclic peptide SFTI-1

In the synthesis of SFTI-1, we highlighted the “Cys-peptide-MeCys” strategy as described earlier. We envisioned that besides the end-to-end thiolactone, the S-S thiol-thioester exchange reaction would also form a smaller thiolactone with the internal cysteine residue in SFTI-1 intramolecularly. Such a thiolactone was in equilibrium with the end-to-end thiolactone which was expected to undergo a proximity-driven S-N acyl shift to form the cyclic SFTI-1 under basic conditions. The sequential thiolactone expansion was previously reported by our laboratory as a thia zip cyclization. The thia zip reaction greatly accelerated the cyclization of cysteine-rich peptides cyclopsychotride and cyclic conotoxin^{150,282} which contained six cysteines. Thus, the cyclization of SFTI-1 was suspected to undergo similar process and assisted by the thia zip mechanism.

Based on the results of model peptide **9a**, we determined the optimal pH range for the tandem thiol switch to form the MES thioester and demonstrated that the one-pot *in situ* ligation was slow. Therefore, in the synthesis of SFTI-1, we employed a two-step cyclization reaction (**Figure 3.10**). The tandem thiol switch of **9b** to give its MES thioester **11b** was performed in an aqueous solution at pH 2 which was shown to be the optimal pH to conduct this reaction based on previous results. Similarly, the reaction was completed after 24 h in presence of excessive external thiols (50 eq. MESNa) to facilitate the MES thioester formation. With the assistance of external thiols, SFTI-MES **11b** was obtained in 54% yield. No fragmented products were detected. Additionally, two thiolactones **14a** and **14b** were also produced in the tandem thiol switch reaction as we envisioned. They were derived from the thiol-thioester exchange reaction between the two cysteines in the sequence and the C-terminal thioester. We identified the thiolactones based on their molecular weight and signature absorbance at 260 nm RP-HPLC. The one closer to the cyclic SFTI-1 was assigned as the end-to-end thiolactone **14b** based on the retention time in HPLC. Then the thia-zip cyclization was performed at pH 7.5 at room temperature with 20 mM TCEP to prevent disulfide formation. Since only one internal cysteine was present in the SFTI-1 sequence, the cyclization reaction was not much accelerated which required 4 h for completion. The oxidative folding of cyclic SFTI was conducted using iodine oxidation under acidic conditions and quenched by ascorbic acid. This

reaction proceeded fast and efficient without any side products detected, which suggested that the direct oxidation was suitable for synthesizing peptides with one disulfide bond.

Collectively, the synthesis of native SFTI-1 by tandem thiol switch using methyl cysteinyl peptide precursors was practical and successful. The reaction required further optimization to accelerate the process.

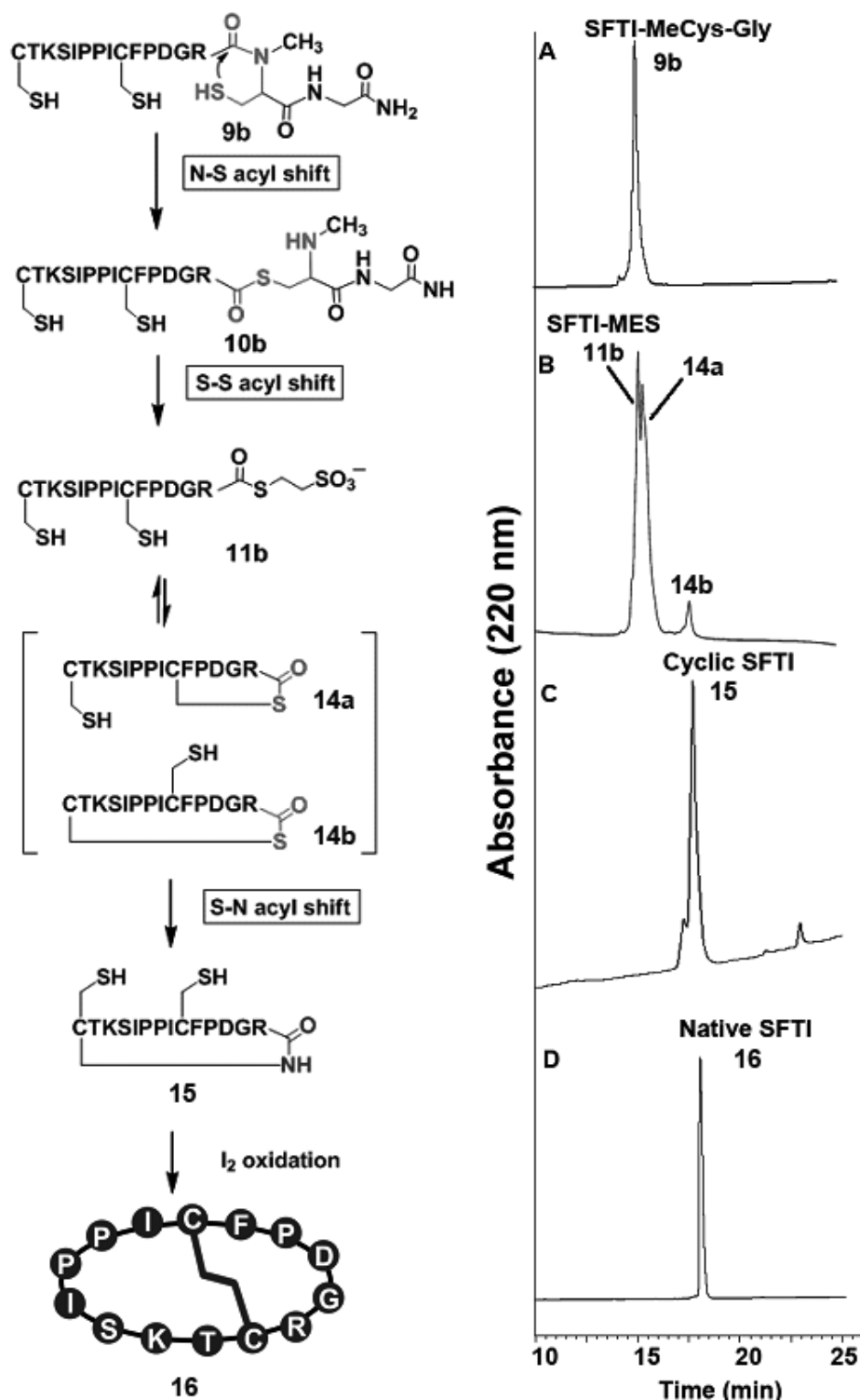


Figure 3.10 MeCys-mediated synthesis of SFTI-1 through tandem thiol switch, thia zip cyclization and oxidation. The tandem thiol switch reaction was conducted at pH 2, 40 °C while the cyclization was performed at pH 7.5, room temperature.

2.5 MeCys-mediated C^α-thioester formation

In previous papers, our laboratory has demonstrated two efficient Fmoc-compatible methods using a hydroxyethylthiol (HET) or thiomethylthiazolidine (TMT) linker for peptide thioester preparation in a condition containing a catalytic amount of TFMSA in TFA.^{228,243} Those linkers went through an intramolecular O-S or N-S acyl shift followed by an intermolecular S-S acyl shift to form a stable peptide thioester. In strongly acidic conditions, an excess amount of proton stabilized the thioester form by protonating the newly exposed secondary amine and drove the O-S or N-S acyl shift equilibrium to the thioester. Thus, the O-S or N-S acyl shift reaction was accelerated significantly and the peptide N^α-thioester was generated in 2 h with concurrent removal of side chain protection and resin.

2.5.1 Synthesis of TIGGIR-MeCys-TC

We envisioned that this scheme would also work for MeCys and accelerate the N-S acyl shift reaction herein. Therefore, a similar condition using 0.1% TFMSA in TFA with 5% thiocresol (TC) was applied to the MeCys-mediated tandem thiol switch at room temperature. In the reaction, TC served a dual role as both the scavenger for cleavage and the external thiol for the tandem thiol switch reaction. It was selected because of its lower pK_a, which made it less prone to be protonated than alkylthiols under strongly acidic conditions. To keep the condition consistent with what was reported in the paper, TIGGIR-MeCys-Gly Rink amide resin was used as the starting material instead of unprotected peptide.

Surprisingly, the strong acid condition displayed a different product profile compared with the aqueous conditions. Only little amount (<5%) of desired product N^α-thioester TIGGIR-TC (m/z 722) was detected by RP-HPLC (**Figure 3.11**). In contrast, a novel thioester (m/z 839), which we identified later as the C^α-thioester TIGGIR-MeCys-TC, was found to be the major product (75%). We named it C^α-thioester because it contained an additional C-terminal MeCys compared with N^α-thioester. The same product was observed when we used unprotected MeCys peptides or other MeCys resins with different C-terminal spacers rather than Gly. Thus, we confirmed the novel product with an m/z of 839 was not random impurity from the resin. Based on the MALDI-TOF MS and tandem mass spectrometry MS/MS profile,

we confirmed the product was TIGGIR-MeCys-TC with each fragmented mass identified (**Figure 3.12**).

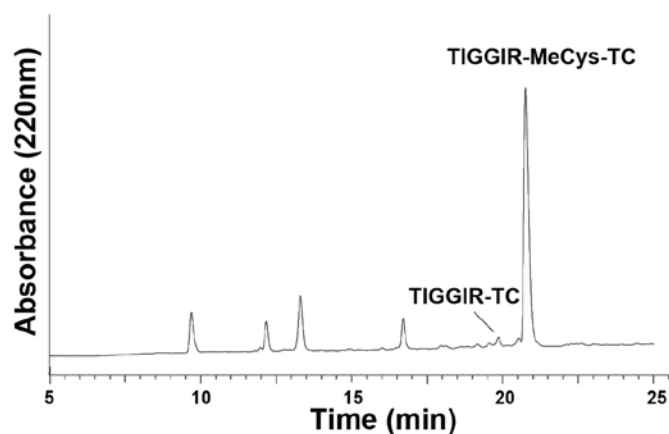


Figure 3.11 RP-HPLC profile of MeCys-mediated C^α-thioester formation. The TIGGIR-MeCys-Gly Rink amide resin was subjected to 0.1% TFMSA in TFA in presence of 5% TC. The reaction was allowed to proceed at room temperature for 2 h. Unprotected peptides were precipitated by diethyl ether and analyzed by RP-HPLC immediately.

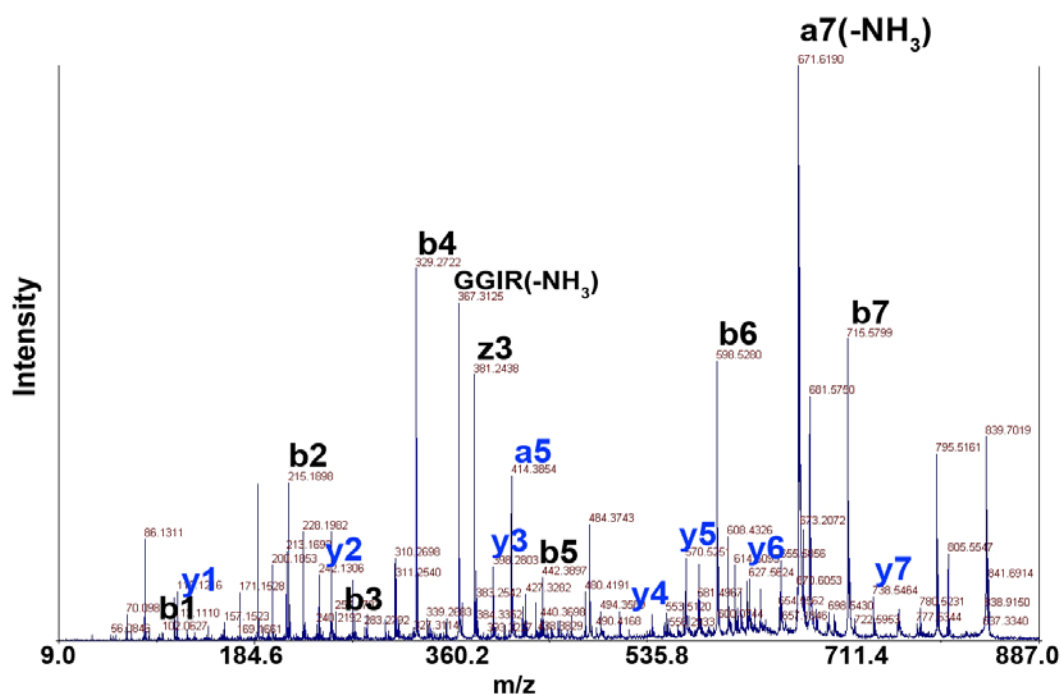


Figure 3.12 MS/MS profile of TIGGIR-MeCys-TC.

2.5.2 Characterization of TIGGIR-MeCys-TC

The C^α-thioester was shown to be in equilibrium between amide and thioester forms (also abbreviated as N and S form, **Figure 3.13**) in H₂O containing 0.1% TFA. After 12 h, 47% TIGGIR-MeCys-TC was converted to the thioester form which showed the same molecular weight but eluted earlier in RP-HPLC. The isomerization was due to the N-S acyl shift reaction mediated by the MeCys residue in the C^α-thioester TIGGIR-MeCys-TC. Interestingly, the newly formed thioester contained two thioester bonds (**Figure 3.13**). These results further confirmed that the C^α-thioester contained a MeCys residue which could mediate the N-S acyl shift reaction.

The C^α-thioester TIGGIR-MeCys-TC was also applied to a chemical ligation with the model peptide CALVIN-NH₂ **12**. The C^α-thioester **17** reacted with peptide **12** at pH 7.5 and afforded a 13-residue peptide TIGGIR-MeCys-CALVIN-NH₂ **18** with m/z 1345.8 (**Figure 3.14**). Peptide **18** was confirmed by MS/MS analysis with each fragment identified (**Figure 3.15**).

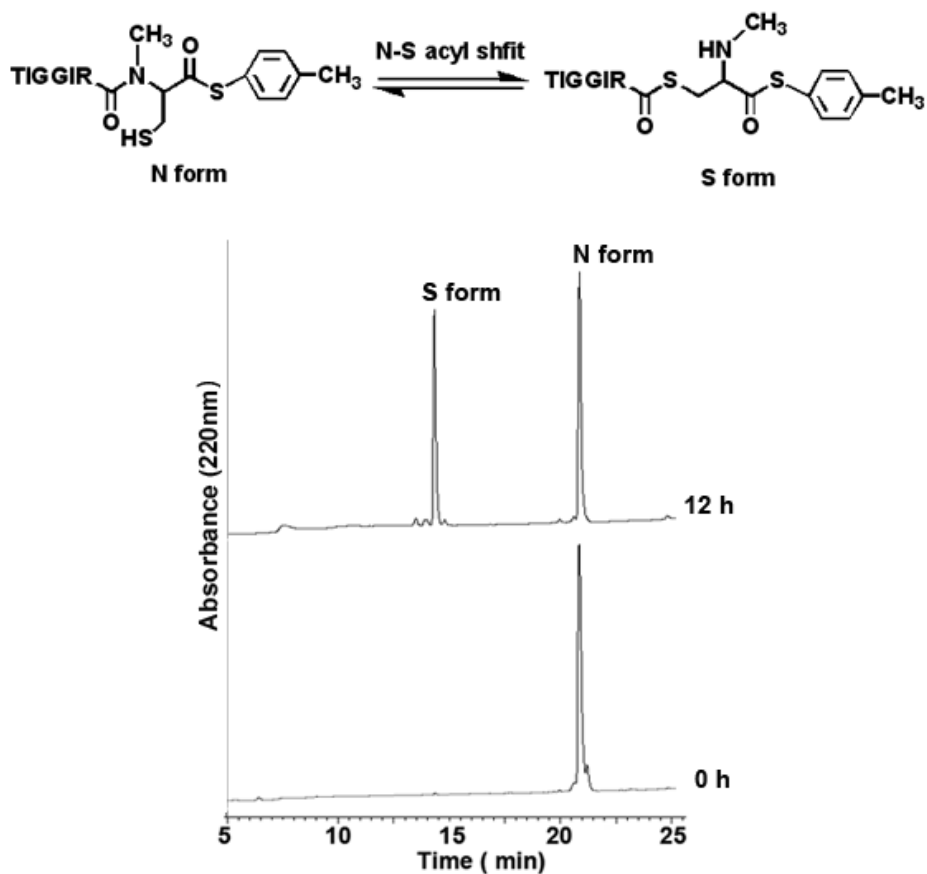


Figure 3.13 RP-HPLC profile of TIGGIR-MeCys-TC in H₂O containing 0.1% TFA. Reaction was monitored by RP-HPLC at 0 h and 12 h at room temperature. N form: TIGGIR-MeCys-TC amide form; S form: TIGGIR-MeCys-TC thioester form.

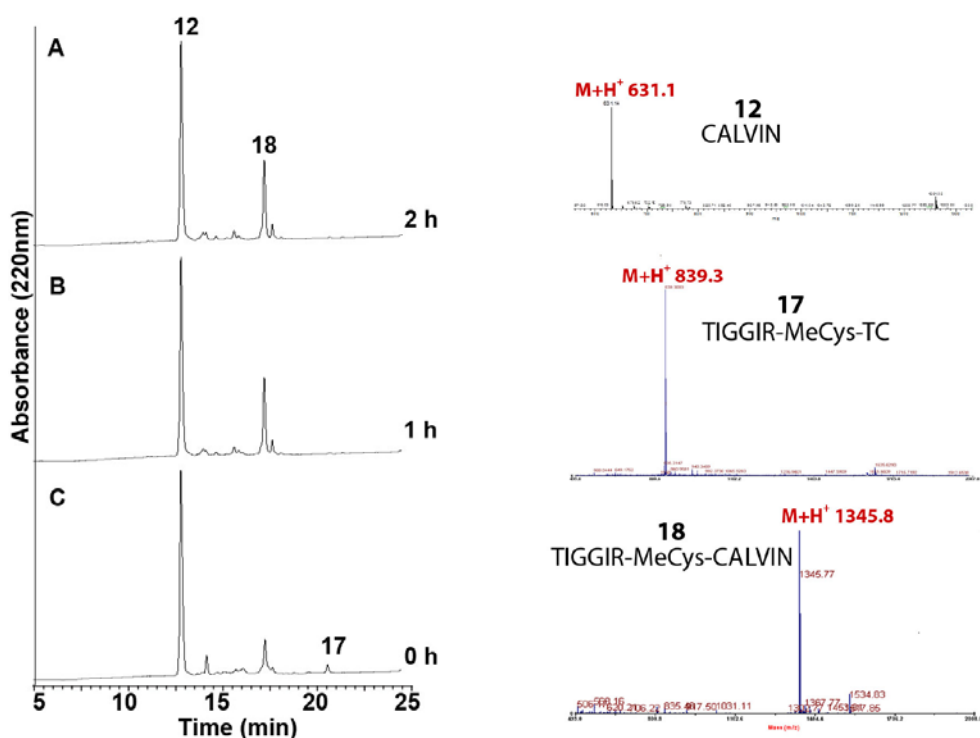
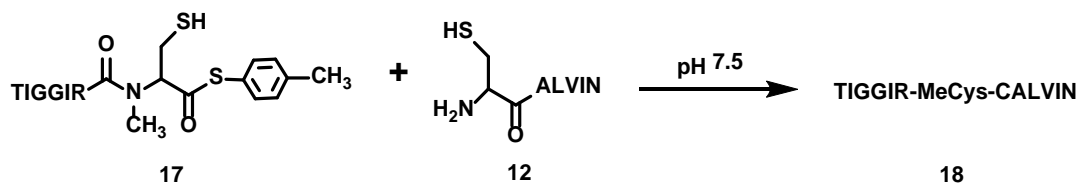


Figure 3.14 Scheme, RP-HPLC and MS profile of TIGGIR-MeCys-TC and CALVIN-NH₂ peptide ligation.

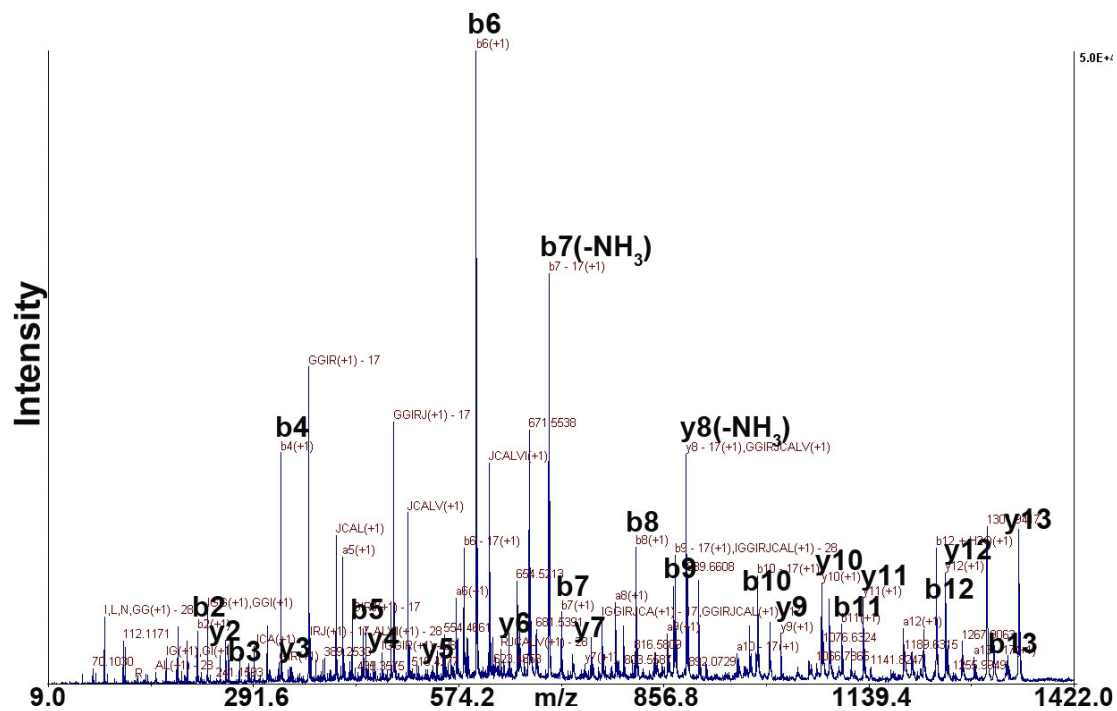


Figure 3.15 MS/MS profile of TIGGIR-MeCys-CALVIN-NH₂.

2.5.3 Sequence-independent C^α-cleavage

Under different reaction conditions, we observed two different thioester formations in the TIGGIR-MeCys-Gly peptide (**Figure 3.16**). The N^α-cleavage occurring at the Arg-MeCys bond produced a N^α-thioester at pH 1-7 in aqueous buffers, while C^α-cleavage occurring at the MeCys-Gly bond produced a C^α-thioester in strongly acidic conditions in presence of 5% TC. The production of the N^α-thioester from the N^α-cleavage was expected and it was due to the N-S acyl shift mechanism. However, the production of the C^α-thioester TIGGIR-MeCys-TC resulting from the cleavage at the MeCys-Gly bond was unusual. We did not expect the thioester to contain an additional MeCys residue although it was confirmed by evidences from different aspects. Sequence-dependent cleavage reactions of susceptible bond linkages such as Gly-Ser/Thr and Gly-Cys¹⁴⁹ bonds have been reported in literature but always occurred at the N-terminus of it. To determine whether the cleavage at MeCys-Gly bond is sequence-dependent, model peptides TIGGIR-Xaa-MeCys-Yaa-NH₂ were synthesized with different amino acids at Xaa and Yaa sites.

To investigate whether the cleavage was restricted to the glycine residue as the Yaa amino acid, we prepared three model peptides TIGGIR-MeCys-Yaa-NH₂ with different Yaa amino acids (Ala, Gly or none). In these peptides, there are no Xaa amino acid between TIGGIR and MeCys residue. These unprotected peptides were treated with 0.1% TFMSA in TFA in presence of 5% TC. We analyzed the products by RP-HPLC and summarized the product profile in **Table 3.1**.

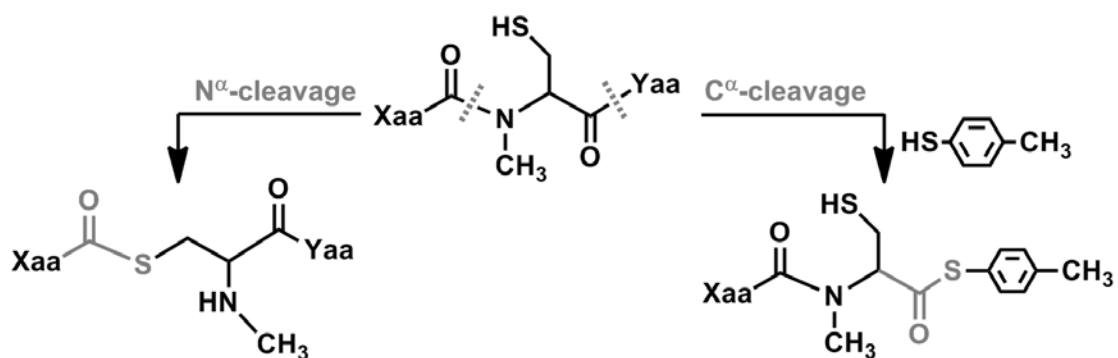


Figure 3.16 Bi-directional cleavage of the MeCys residue under different conditions. N^α-cleavage was observed at pH 1-5 in aqueous buffers. C^α-cleavage was observed in strongly acidic organic conditions containing 0.1% TFMSA and 5% TC in TFA

Two different types of products were observed in the reaction profile. One category including the S form of starting material and the N^α-thioester TIGGIR-TC resulted from the N^α-cleavage process. The other category containing TIGGIR-MeCys-OH, TIGGIR-MeCys-TC resulted from the C^α-cleavage process. S forms of both the TIGGIR-MeCys-TC and TIGGIR-MeCys-OH were also observed, which was likely produced during the RP-HPLC purification at acidic pH.

In all three peptides, products that derived from the C^α-cleavage reaction of a MeCys-Yaa bond accounted for >85% yield in total (89.3-97.0%). Irrespective of the C-terminal varying part, TIGGIR-MeCys-TC was invariably formed as the major product (**Table 3.1**, left column). These results clearly suggested that the C^α-cleavage of MeCys was not restricted to the Gly residue at the Yaa site and it also occurred when other amino acids were included.

To investigate whether amino acid at N-terminus of MeCys would affect the C^α-cleavage of MeCys, we varied the Xaa amino acid in model peptides TIGGIR-Xaa-MeCys-Gly-NH₂ with Xaa = Gly, Ala, Arg, and Leu and kept the Yaa residue as Gly (**Table 3.1**, right column). Four model peptides were treated with strong acid using the same condition described earlier. We did not observe substantial differences in C^α-cleavage yield when different Xaa amino acids were used. From Gly to Arg, comparable yields were obtained ranging from 94.1-98.0%. It should be noted that neither the basic residue Arg nor hydrophobic residue Leu significantly affected the C^α-cleavage of MeCys. Collectively, these results showed that the C^α-cleavage of MeCys was not sequence-dependent and ubiquitously occurred in different model peptides TIGGIR-Xaa-MeCys-Yaa with various Xaa and Yaa residues.

Table 3.1 Product summary of TIGGIR-Xaa-MeCys-Yaa in strongly acidic conditions containing 5% TC, 0.1% TFMSA in TFA.

| Type of cleavage | Product | Yield (%) | | | | | | |
|-------------------------------|-------------------------------|-----------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | | Yaa* | | | Xaa | | | |
| | | NH ₂ | Gly** | Ala** | ---*** | Gly | Ala | Leu |
| N^α-Cleavage | TIGGIR-Xaa-MeCys-Yaa (S form) | 5.3 | 1.6 | 5.2 | 1.6 | 4.8 | 0.5 | 1.0 |
| | TIGGIR-Xaa-TC | 5.4 | 1.4 | 2.5 | 1.4 | 1.1 | 1.8 | 1.0 |
| | Total | 10.7 | 3.0 | 7.7 | 3.0 | 5.9 | 2.3 | 2.0 |
| C^α-Cleavage | TIGGIR -Xaa-MeCys-OH (S form) | 18.0 | 7.7 | 28.5 | 7.7 | 5.0 | 2.6 | 3.6 |
| | TIGGIR -Xaa-MeCys-OH (N form) | 17.9 | 10.4 | 13.2 | 10.4 | 8.4 | 16.7 | 14.0 |
| | TIGGIR-Xaa-MeCys-TC (S form) | 6.1 | 3.9 | 2.2 | 3.9 | 3.5 | 8.9 | 6.0 |
| | TIGGIR-Xaa-MeCys-TC (N form) | 47.4 | 75.1 | 48.4 | 75.1 | 77.2 | 69.4 | 74.3 |
| | Total | 89.3 | 97.0 | 92.3 | 97.0 | 94.1 | 97.7 | 98.0 |

* For all Yaa peptides, there is no Xaa amino acid between TIGGIR and MeCys residue.

** Gly and Ala refer to Gly-NH₂ and Ala-NH₂

*** --- refers to no amino acid between TIGGIR and MeCys

2.5.4 C^α-cleavage in other model peptides

Since we have already proved the C^α-cleavage of MeCys is not dependent on either N- or C-terminal amino acid of MeCys, we would like to investigate whether it is assisted by its thioethyl side chain. Therefore, the model peptide TIGGIR-MeAla-Gly-NH₂ was prepared, in which the MeCys amino acid was replaced by a MeAla residue. TIGGIR-MeAla-TC was obtained in 75% yield when TIGGIR-MeAla-Gly-NH₂ was treated with 0.1% TFMSA and 5% TC in TFA (**Figure 3.17**). The result was similar to treating TIGGIR-MeCys-Gly-NH₂ with the same condition, suggesting that the C^α-cleavage was not dependent on the thioethyl side chain. To further confirm whether the C^α-cleavage was restricted to N-methylated residues at the proximal C-terminus, we synthesized a model peptide TIGGIR-MeAla-(Gly)₆-NH₂, in which the MeAla residue was followed by six Gly residues. Similarly, TIGGIR-MeAla-TC was obtained as the major product. It proved that the C^α-cleavage would occur at the C-terminal side of N-methylated residue no matter where the residue was located.

In the previous section, we have demonstrated the difference between Cys and MeCys in the N^α-cleavage. It was interesting to investigate their difference in the C^α-cleavage as well. We treated two model peptides TIGGIR-Gly-Cys-Gly-NH₂ and TIGGIR-Gly-MeCys-Gly-NH₂ with 0.1% TFMSA and 5% TC in TFA. For the MeCys peptide, C^α-thioester TIGGIR-Gly-MeCys-TC was produced in 75% yield after 1 h. However, for the Cys peptide, no C^α-thioester TIGGIR-Gly-Cys-TC was observed. Replacement of MeCys with Cys totally suppressed the C^α-cleavage. The prolonged treatment up to 170 h did not afford the C^α-thioester in >10% yield.

The results of various model peptides treated with 0.1% TFMSA and 5% TC in TFA were summarized in **Figure 3.18**. Peptides containing MeCys or MeAla produced C^α-thioester and it was not restricted to the C-terminal methylamino acids. Replacing MeCys with Cys abolished the C^α-cleavage completely. Collectively, these results suggested that the C^α-cleavage reaction was strongly accelerated by the N-alkylation of an amino acid under strongly acidic conditions.

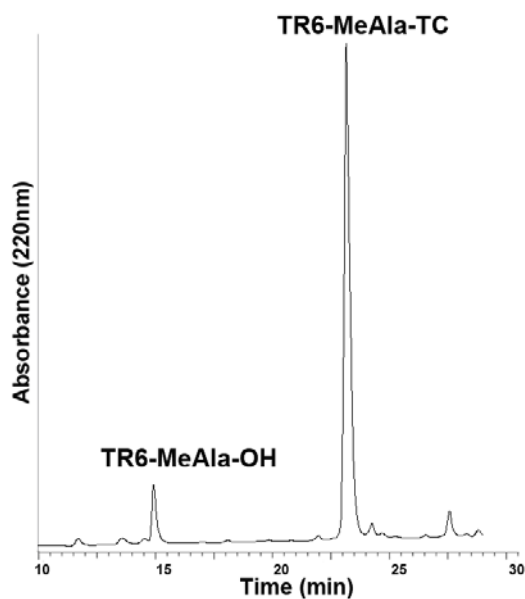


Figure 3.17 RP-HPLC profile of TIGGIR-MeAla-Gly in strongly acidic conditions containing 5% TC, 0.1% TFMSA in TFA.

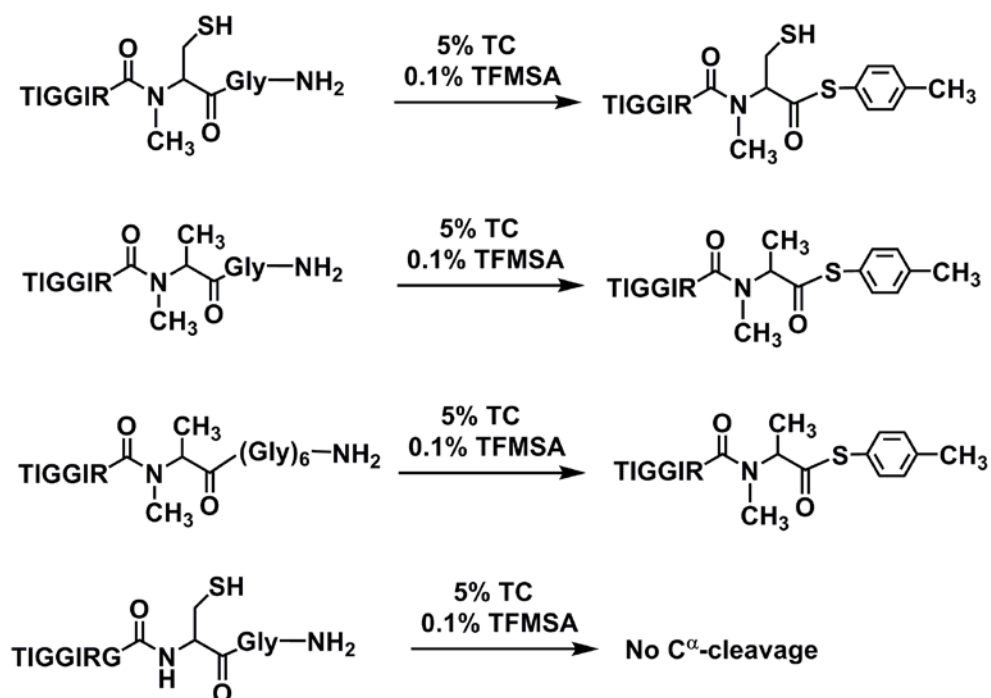


Figure 3.18 Summary of reactions containing different model peptides treated with 5% TC, 0.1% TFMSA in TFA.

2.5.5 Acidity function of 0.1% TFMSA in TFA

To compare the N^α- and C^α-cleavage and understand the mechanism of the C^α-cleavage, we determined the acidity function of 0.1% TFMSA in TFA. ¹H nuclear magnetic resonance (NMR) was used to measure the acidity by monitoring chemical shifts of dimethylsulfide (DMS) in various conditions.²⁷⁵ In different conditions containing various concentrations of TFMSA, the chemical shift of DMS would be an indicator for the amount of protons in the condition. DMS was not ionized in TFA but started to be protonated at higher acidity. By measuring the protonated and unprotonated forms of DMS in the NMR profile, we could obtain the ratio of protonation and calculate the acidity function based on it (**Figure 3.19**). The acidity of seven conditions containing 0.05-20% TFMSA were measured using this method. The acidic function H₀ of the solution containing 0.1% TFMSA in TFA was determined as -5 by our NMR titration results. It indicated that the C^α-cleavage occurred at H₀ -5. Since the N^α-cleavage was optimal at pH 2, thus, there was a seven log scale difference in the acidity of N^α-cleavage and C^α-cleavage reaction conditions.

**Protonation ratio of dimethyl sulfide in
non-aqueous TFMSA/TFA acidic solutions**

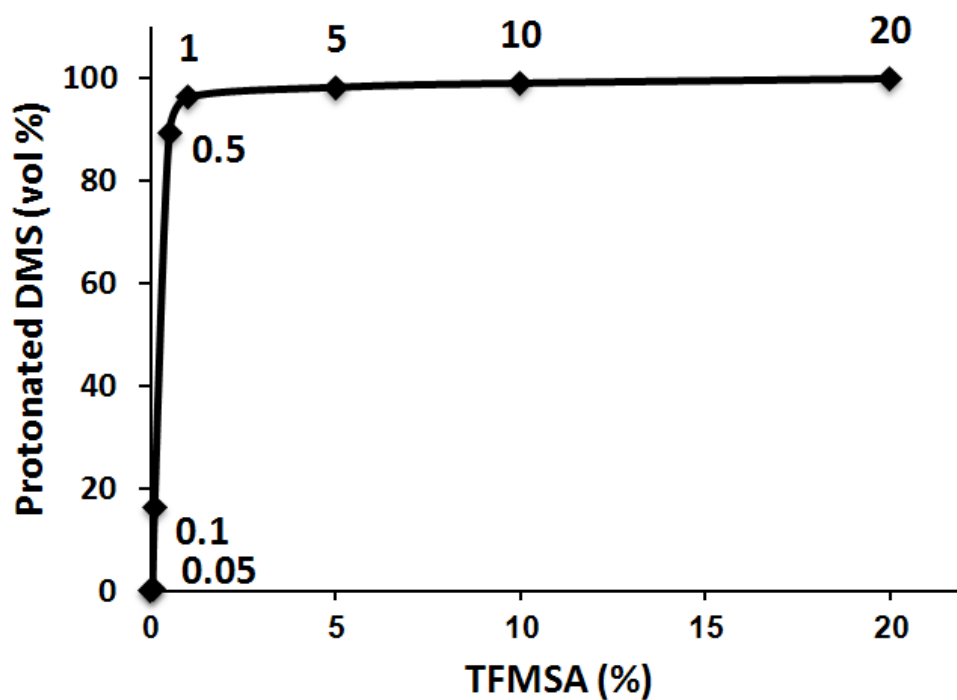


Figure 3.19 The protonation ratio curve for DMS in non-aqueous TFMSA/TFA solutions.

Table 3.2 H_0 values of acidic solutions containing 0-0.5% TFMSA in TFA.

| TFMSA (vol % in TFA) | $-H_0$ |
|----------------------|--------|
| 0.0 | 3.03 * |
| 0.05 | 3.70 |
| 0.1 | 4.95 |
| 0.5 | 6.25 |

* The value for acidity function H_0 value of TFA was referred from ref. ²⁸³.

2.5.6 Proposed mechanism for C^α-cleavage of N-methylamino acids

So far, we have proved that the C^α-cleavage reaction of MeCys residue in strongly acidic conditions was sequence-independent basing on peptides tested here. The amino acid at neither its N-terminal (Gly, Ala, Leu or Arg) nor C-terminal side (Gly, Ala or none) exerted significant effects on the C^α-cleavage as the C^α-thioester was ubiquitously formed as the major product. However, it remains to be determined if the same reaction will occur when the N- or C-terminus is other amino acids besides those listed here.

Furthermore, deletion of side chain thioethyl group did not affect the C^α-cleavage either since replacing MeCys with MeAla in model peptides also resulted in the formation of the C^α-thioester. It occurred at the C-terminal side of N-methylated amino acid and not restricted to N-methylated amino acid at proximal C-terminus. However, the C^α-cleavage reaction was strongly suppressed when the MeCys was replaced by Cys in model peptide, suggesting the importance of the N-methylation in this reaction. Thus, our results have shown that the C^α-cleavage was highly dependent on the presence of *N*-methylated amino acid in strongly acidic conditions and occurred specifically at the C-terminal side of it.

The susceptibility of *N*-methylated amino acids to racemization have been reported in strongly acidic conditions containing HBr in acetic acid²⁸⁴. This reaction was acid-catalyzed and has been suspected to be caused by the oxazolone formation. Similarities were found between our reaction and the reaction reported. Both of them were acid-catalyzed and occurred in strongly acidic conditions. The acidity function of our condition containing 0.1% TFMSA in TFA (H₀ -5) and of their condition containing 37% HBr in acetic acid was comparable. Therefore, we suspected that oxazolone formation was probably the mechanism for our observed C^α-cleavage reaction.

The proposed mechanism for the C^α-cleavage reaction was shown in **Figure 3.20**. The site-specific C^α-cleavage of MeCys or MeAla was mediated by an *N*-methyloxazolone intermediate which was later opened up by a nucleophile such as the thiocresol to form the corresponding C^α thioester. The oxazolone formation involved the carbonyl groups located at both the N- and C-terminal side of the *N*-methylated

amino acid. Under strongly acidic conditions, the carbonyl group of *N*-methyl amide bond would be strongly activated and susceptible to nucleophilic attack. Then an *N*-methyloxazolone was generated by attack on the activated carbonyl group of the adjacent amide bond. Thus, the peptide bond at the C-terminal side of *N*-methylated amino acid was cleaved. In presence of nucleophiles such as thiocresol and water, the *N*-methyloxazolone ring would be opened up by either thiocresol to form a TC thioester or water to form a hydrolyzed product.

The proposed mechanism corresponded well to our results and was supported by Urban *et al.*^{285,286} and Teixido *et al.*²⁸⁷. Both of them reported the C-terminal cleavage of *N*-alkylated amino acids during a prolonged treatment by TFA using protected and unprotected peptide, respectively. In their conditions, no external thiols were included, thus, a hydrolysis product was formed as the major product. They proved the importance of *N*-alkylation in activation of the amide bond which mediated the oxazolone formation. In addition, in our strongly acidic conditions, the thioethyl side chain was highly protonated, which reduced its nucleophilicity to mediate the N^α-cleavage. Thus, the C^α-cleavage was favored over the N^α-cleavage in the strongly acidic conditions.

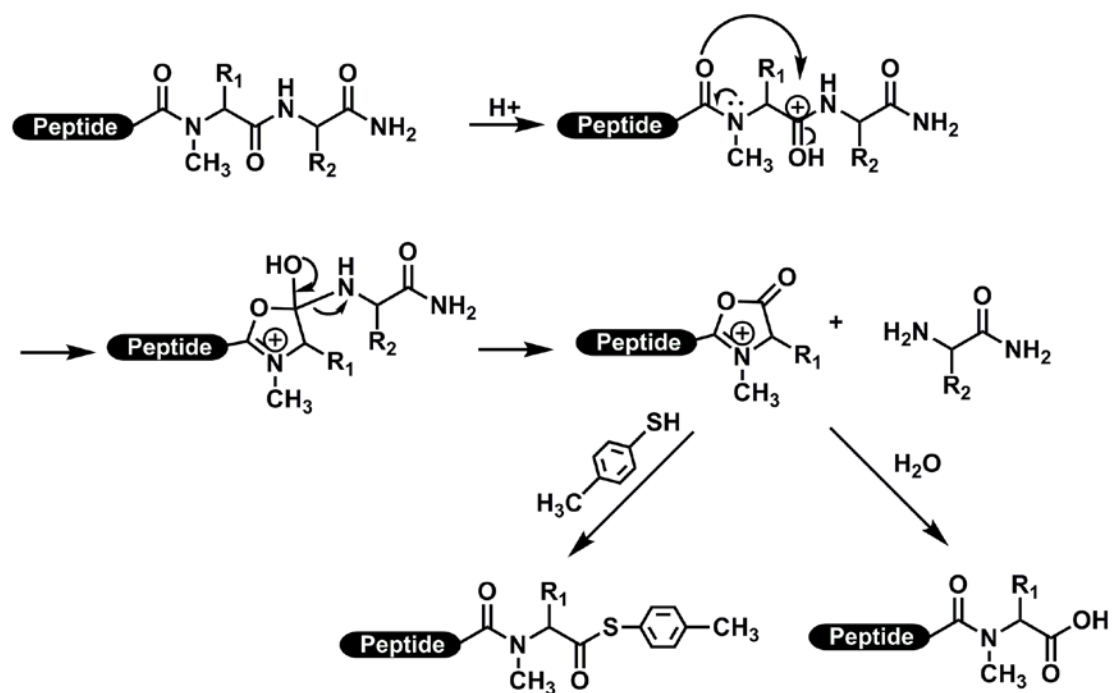


Figure 3.20 A proposed mechanism for C^α- cleavage of N-methylated amino acid.

According to Urban and co-workers, variations of N-methyl residue's alkyl side chains did not have any significant effects on the acid hydrolysis but the inductive effect of N-methylation in the N-methylated amino acid appeared to be important for activating the amide bond and led to the oxazolone formation. It was consistent with our results that MeCys and MeAla did not show significant differences in the C^α-thioester formation but replacing MeCys with Cys in model peptides suppressed the C^α-cleavage totally. It was also noteworthy that there was great difference between the rate of our condition and Urban's condition. All C^α-cleavage reactions were completed within 1 h in our conditions containing 0.1% TFMSA (H₀ -5) while in Urban's paper, the acid hydrolysis reaction of MeGly, MeAla or MePhe in TFA (H₀ -3) required 13.5-64.5 h for completion. It clearly demonstrated that the addition of TFMSA greatly facilitated the oxazolone formation by increasing the acidity of the reaction solution. The use of different nucleophiles could also contribute to the difference of the reaction rate. Since thiocresol was a better nucleophile than water, it accelerated the ring opening reaction of the N-methyloxazolone intermediate. It should be pointed out that in Urban's work, the acid hydrolysis of MeCys was not studied.

It should be noted that the N-methyloxazolone formation was not observed as a side reaction in our TMT-mediated thioester formation using 0.1% TFMSA in TFA. We also replaced the MeCys residue with a native secondary amino acid Pro in model peptides to check whether the C^α-cleavage occurred. Under the same reaction, no detectable C^α-thioester was observed up to 4 h. The N-methyloxazolone formation was not favored in these two secondary amines probably due to steric hindrance of the thiazolidine or proline ring in forming the bicyclic oxazolone intermediate.

2.6 β -Thiolactone formation

We were interested to determine the stability of C^α -thioesters TIGGIR-MeCys-TC since the thioethyl side chain and the thioester bond were in proximity. We supposed that a proximity-driven displacement reaction would be likely to occur and lead to a β -thiolactone formation especially when a good leaving group thiocresol was present (**Figure 3.21**). Indeed, a product with an m/z of 715 was observed by treating TIGGIR-MeCys-TC with aqueous phosphate buffers at pH 4 to 7 (**Figure 3.22**). We monitored the rapid conversion of TIGGIR-MeCys-TC to this product by RP-HPLC. At pH 4, TIGGIR-MeCys-TC was considerably stable and resulted in less than 5% product (m/z 715) after 80 min (**Figure 3.22A**) while at pH 7, TIGGIR-MeCys-TC and its product (m/z 715) degraded quickly after 1 h (**Figure 3.22D**). The conversion of TIGGIR-MeCys-TC to this m/z 715 product was clearly recorded at pH 5 and 6 and this product was obtained as the major product after 80 min and 1 h, respectively.

The hydroxyl side chain of Thr, the guanidino side chain of Arg or the N-terminal α -amine could also act as nucleophiles to form a lactone, lactam, or the end-to-end macrocycle, respectively, which had the same molecular weight as the β -thiolactone (m/z 715). To rule out these possibilities, we alkylated the thiol side chain of the MeCys amino acid by N-ethylmaleimide (NEM). No β -thiolactone formation was observed after 12 h when treating the *S*-alkylated TIGGIR-MeCys-TC with aqueous buffer at pH 5 (**Figure 3.23**). No β -thiolactone formation was observed for up to 12 h. It clearly proved that the thioethyl side chain of MeCys participated in the β -thiolactone formation. The MS/MS fragmentation profile of TIGGIR-MeCys β -thiolactone also ruled out the possibilities of other lactones or lactams (**Figure 3.24**).

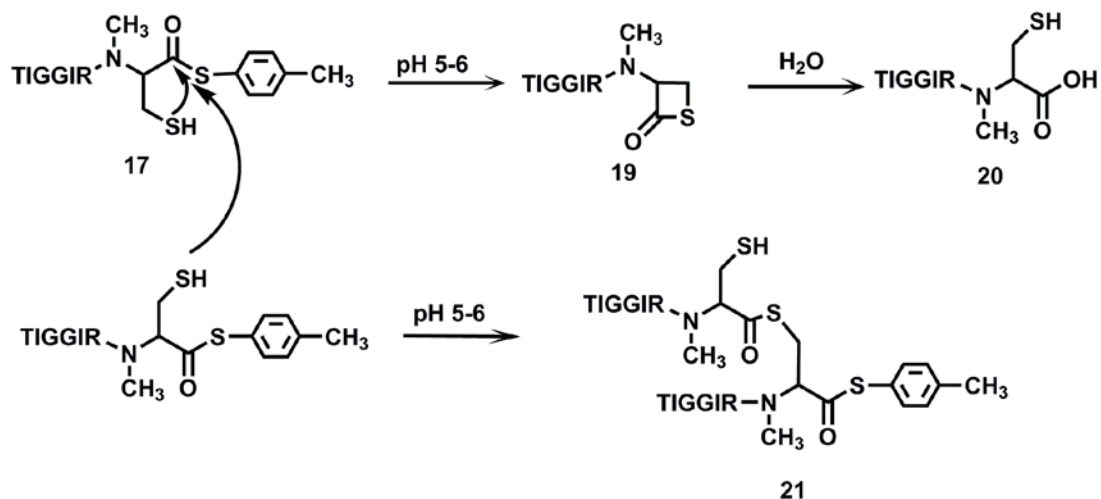


Figure 3.21 Mechanism for β -Thiolactone formation.

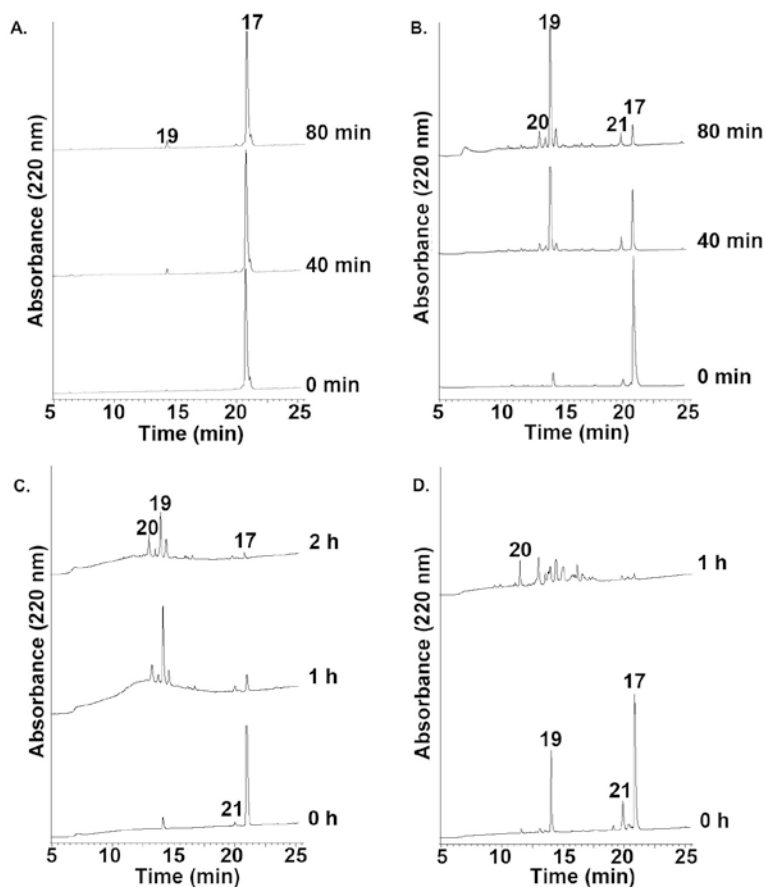


Figure 3.22 β -Thiolactone formation from TIGGIR-MeCys-TC at pH 4-7 phosphate buffer. A to D represented pH 4-7, respectively. 17: TIGGIR-MeCys-TC; 19: TIGGIR-MeCys β -thiolactone; 20: TIGGIR-MeCys-OH; 21: TIGGIR-MeCys diacyl-thioester.

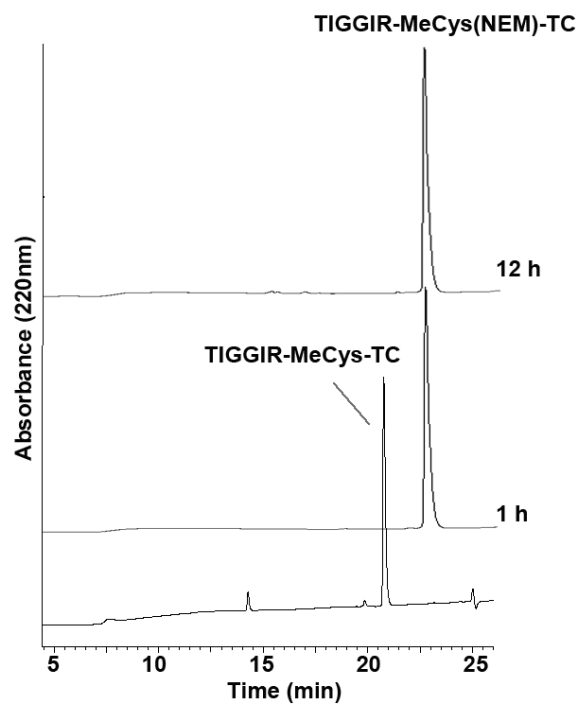


Figure 3.23 RP-HPLC of NEM alkylated TIGGIR-MeCys-TC in pH 5 phosphate buffer at 1 h and 12 h.

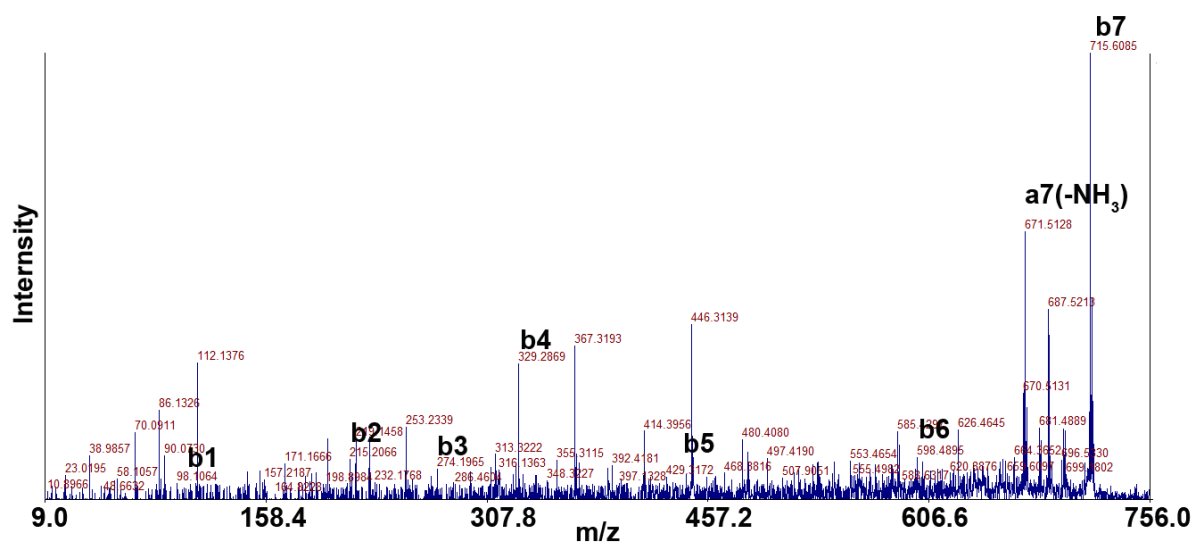


Figure 3.24 MS/MS profile of TIGGIR-MeCys β -thiolactone (m/z 715).

Replacing MeCys with Cys in model peptides resulted in the β -thiolactone formation with similar yield and rate (**Figure 3.25C**). To rule out possibilities of methyl group playing a role in this reaction, we secondly tested MeAla, a residue contains a methyl group but not a thiol side chain. Upon 15 h, no β -thiolactone formation was observed using the peptide TIGGIR-MeAla-TC as starting materials. Only a small amount of hydrolysis product was detected (**Figure 3.25A**).

Additionally, the C-terminal leaving group was also important for the β -thiolactone formation. When we converted the C-terminal leaving group from thiocresol to methyl mercaptoacetate (MMA), no β -thiolactone was obtained. Two MMA thioester TIGGIR-MeCys-MMA and TIGGIR-Cys-MMA were sufficiently stable in pH 5 phosphate buffer up to 24 h (**Figure 3.25B, D**) as monitored by RP-HPLC and MALDI-TOF MS. The β -thiolactone formation was detected only when the C-terminal leaving group was TC rather than MMA. TIGGIR-Cys-MMA was stable as long as 24 h in pH 5 phosphate buffer while 50% TIGGIR-Cys-TC was converted to the β -thiolactone after 1 h. These results could be explained by the difference of reactivity between an aryl thioester and an alkyl thioester as TC thioesters are better leaving groups compared with MMA thioesters.

Collectively, we proved that the β -thiolactone formation was highly dependent on both the thioethyl side chain of the Cys or MeCys residue and a good leaving group at the C-terminus, such as a C-terminal thiocresol.

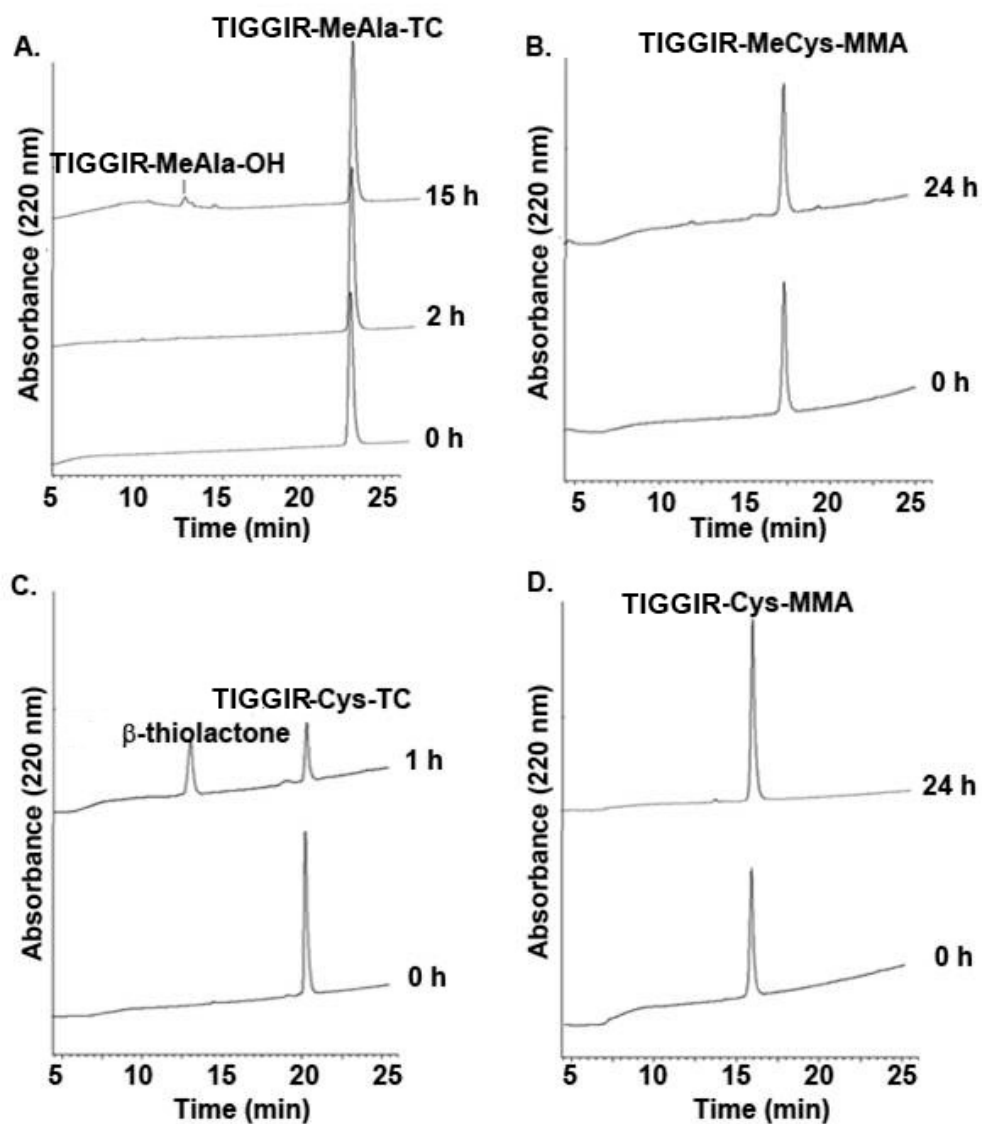


Figure 3.25 RP-HPLC profiles of four different thioesters TIGGIR-MeAla-TC, TIGGIR-MeCys-MMA, TIGGIR-Cys-TC, and TIGGIR-Cys-MMA in pH 5 phosphate buffer.

3. TEBA as a thioester surrogate for preparation of cyclic CRPs

3.1 Design and rationale of using TEBA as a thioester surrogate

Two limitations had been observed when using MeCys as a thioester surrogate. First was the low efficiency of introducing amino acids to the MeCys resin. Multiple couplings were required for completion due to the low reactivity of this secondary amine. Second was the formation of an undesired MeAla(Pip) side product due to a β -elimination during assembly of a peptide on the MeCys resin. Addition of a spacer residue Gly at the C-terminus could not resolve the problem completely.

To overcome these limitations, we optimized the design by eliminating the α -carbonyl moiety of MeCys residue. The new design employing a thioethylbutylamide (TEBA) as the thioester surrogate is simpler but retains the essential thioethylamido (TEA) group for acyl shift reactions. The TEBA thioester surrogate has several advantages over the MeCys linker. Without the carboxylic group, the β -elimination side reaction is avoided completely and the efficiency of the first amino acid introduction is increased due to the decreased steric hindrance during coupling. Another advantage of TEBA is that it is a commercially available material with a comparably cheap price. The preparation of TEBA resin could be achieved through a simple substitution reaction with the thiol group directly linked to and protected by the resin support. No additional steps to load protecting groups are required. As such, the TEBA thioester surrogate represents the simplest TEA linker for thioester preparation.

3.2 Synthesis of model peptides on the TEBA resin by Fmoc SPPS

The TEBA-Trt(2-Cl) resin **23** could be readily prepared by incubating the commercially available 2-(butylamino)ethanethiol with Cl-Trt(2-Cl) resin **22** in DCM (**Figure 3.26**). The reaction was completed in 1 h to afford an S-linked resin, which was used directly for synthesizing a model peptide TIGGIR. Unlike the multiple coupling reactions of the first amino acid in the synthesis of MeCys peptides, the Fmoc-Arg(Pbf)-OH was coupled to the TEBA resin by a single coupling reaction using HATU/DIEA. The coupling with TEBA was completed in 1 h while a triple or quadruple coupling was required for MeCys, suggesting that TEBA was more

reactive than MeCys in the synthesis. Peptide chain assembly was performed manually using Fmoc SPPS. Cleavage by TFA/TIS/Thioanisole afforded the unprotected peptides with TEBA linker without detection of the piperidine adduct side product. A mixture containing both amide **25a** and thioester **25b** forms (ratio ~6:4) was obtained after HPLC purification. It was likely because the amide form TIGGIR-TEBA **25a** was gradually converted to the thioester form **25b** in the acidic HPLC buffer. These results suggested that the TEBA-mediated N-S acyl shift was faster than MeCys since S form peptide was not detected in significant amount (<2%) after purification of peptides with a MeCys linker. It was noteworthy that the ratio was varying and also dependent on the incubation time of peptide in acidic buffers.

3.3 TEBA-mediated N^α-thioester formation in aqueous conditions

We used the same tandem thiol switch condition in TEBA-mediated N^α-thioester formation as in MeCys, which involved an N-S acyl shift reaction at pH 1-7 followed by an S-S thiol-thioester exchange reaction using MESNa (**Figure 3.27**). MESNa thioester **26** was obtained after 18.5-24 h at 40 °C in phosphate buffers with various pH (**Figure 3.27**). Similar to MeCys, TEBA-mediated tandem thiol switch reaction was pH-dependent. The N-S acyl shift was favored in acidic conditions (pH < 2) while the S-S thioester exchange reaction preferred to occur in neutral to basic conditions (pH > 6). The optimal pH for MeCys-mediated N^α-thioester formation was pH 2 and the tandem thiol switch reaction could be carried out at a wide pH ranges from 2-5.

The MeCys and TEBA mediated tandem thiol switch reactions were similar. Both were pH-dependent and shared a similar reaction mechanism. Despite that, the MeCys-mediated tandem thiol switch reaction required a more acidic condition (pH 2) than that of TEBA (pH 3) (**Figure 3.28**). This was due to the absence of electron-withdrawing carbonyl group on TEBA, which made it more basic than MeCys. In addition, TEBA served as a more reactive thioester surrogate than MeCys in the tandem acyl shift reaction. In TEBA-mediated tandem thiol switch, the thioester formation was completed within 18.5 h, while MeCys-mediated tandem reactions required 24 h to complete.

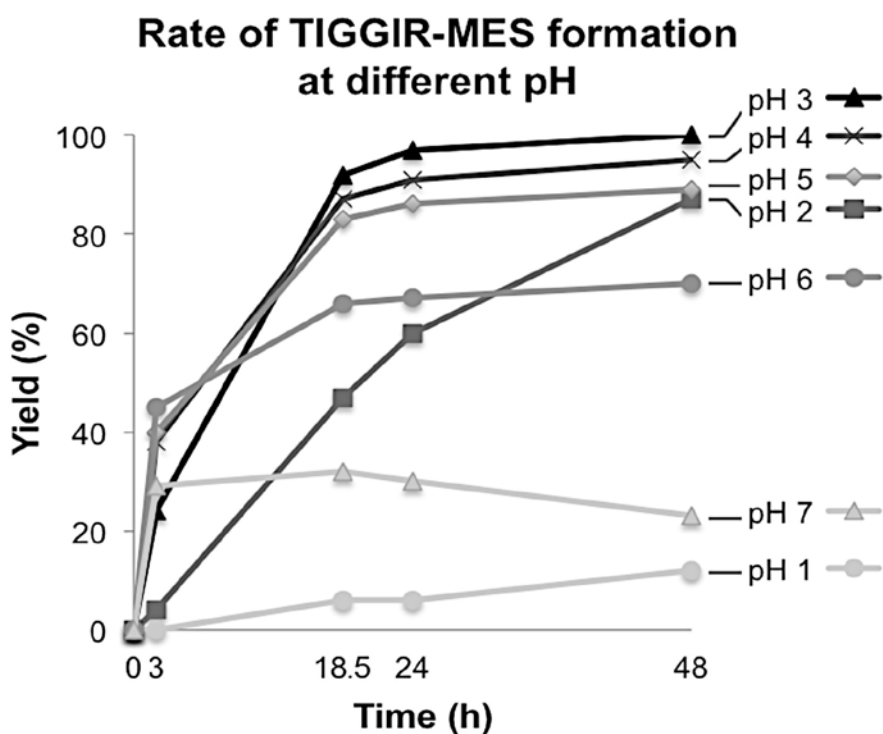
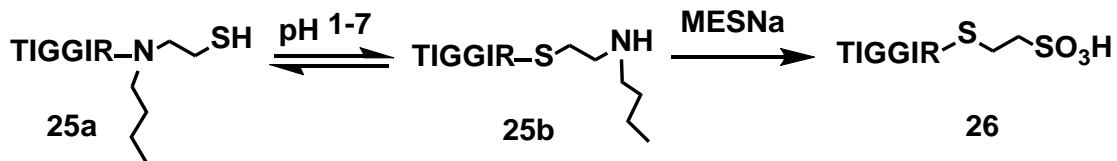


Figure 3.27 TEBA-mediated tandem thiol switch reaction at various pH. The scheme showed the tandem N-S, S-S acyl shift reactions to generate thioester TIGGIR-MES **26** from TIGGIR-TEBA **25a**. Yield of TIGGIR-MES thioester **26** formation at pH 1-7 was monitored at different time points at pH 1-7. The TEBA-mediated tandem thiol switch reaction was performed at 40 °C in phosphate buffer in presence of 5 mM peptide and 50 eq. MESNa.

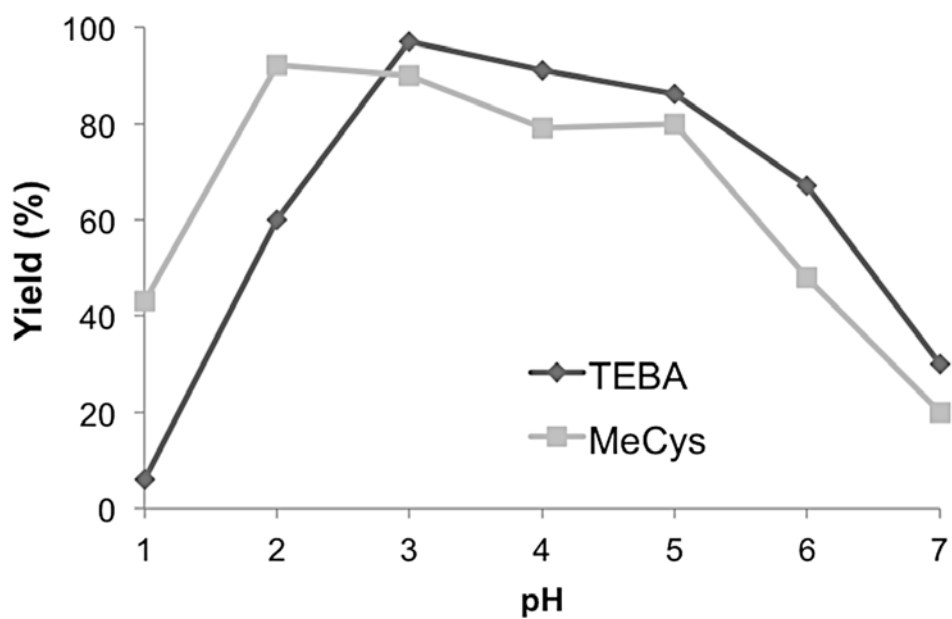


Figure 3.28 Comparison of pH-dependent MeCys and TEBA mediated tandem thiol switch reaction.

3.4 TEBA-mediated synthesis of cyclic peptide SFTI-1

We synthesized the 14-amino-acid cyclic peptide SFTI-1 as an example to demonstrate the TEBA-mediated synthesis of cyclic peptides. The sequence of SFTI-1 was shown in **Figure 3.29**. A similar synthetic approach as in the MeCys method was employed to afford unprotected peptides with a TEBA linker in 28% isolated yield. Two products with the same molecular weight as the desired TEBA peptide were obtained after HPLC purification, which are the amide form (N form) SFTI-TEBA **27a** and thioester form (S form) **27b**.

We used the same Cys-peptide-TEBA strategy for cyclization of SFTI-1 as in intein and MeCys mediated cyclization reaction. The unprotected peptides SFTI-TEBA **27a** and **27b** both served as the starting material in our cyclization reaction (**Figure 3.30A**). They were dissolved in phosphate buffer to reach a final concentration of 3 mM and then subjected to the cyclization reaction using two different approaches (**Figure 3.29**).

In the first approach, we used a two-step method involving a thiolysis by external thiols followed by thia zip cyclization, which was similar to the intein and MeCys approach. The reaction was conducted at pH 3 in presence of 50 eq. MESNa for up to 18 h at 40 °C. To prevent potential disulfide formation, TCEP was added to the reaction at a concentration of 10 mM. The thioester SFTI-MES **29** was observed as the major product after 15 h. Besides thioester SFTI-MES **29**, two thiolactones (**28a** and **28b**) and another TEBA thioester **27b** were isolated as well (**Figure 3.30B**). The two thiolactones, which were generated by the thiol-thioester exchange between the internal thiols and the C-terminal thioester, were identified based on mass spectrometry and RP-HPLC. Both of them possessed the same mass as the cyclic peptide with a distinct 260 nm absorbance, which was the signature absorbance for thioester bond. The one eluted closer to the cyclic product was identified as the 43-atom end-to-end thiolactone **28b** while the other one eluted earlier as the 19-atom thiolactone **28a**. Then the reaction mixture containing thioesters and thiolactones was adjusted to pH 7 at room temperature for cyclization reaction. Cyclic SFTI-1 **30** was afforded after 4 h and purified for the oxidative folding to generate the native SFTI-1 **31** in 33% yield.

With two cysteine residues in the sequence, we anticipated that external thiols would be unnecessary for the thioesterification reaction of SFTI-TEBA. Thus a one-step approach leading to a direct thiolactone formation without the external thiolysis step was performed (**Figure 3.29**). In this case, the cyclization of SFTI-TEBA was performed in similar experimental conditions as the previous approach but without MESNa or any external thiol. After 15 h, except for the thioester SFTI-MES **29**, all expected products including thioester **27b** and thiolactones **28a** and **28b** were detected (**Figure 3.30C**). In addition, a small amount of hydrolysis product, which eluted earlier than thiolactones **28a**, was observed in 5%. After the cyclization reaction at pH 7 for 4 h, the cyclic SFTI-1 **30** was generated and directly subjected to the oxidative folding with 20% DMSO without purification.

Both conditions with or without external thiols gave comparable cyclization yield which were 89% and 84%, respectively. However, conditions without external thiols required only one purification step from linear precursor to native product while the two-step approach required two. Thus the overall yield was 10% higher (43% vs 33% isolated yield) when no external thiols were included in the reaction.

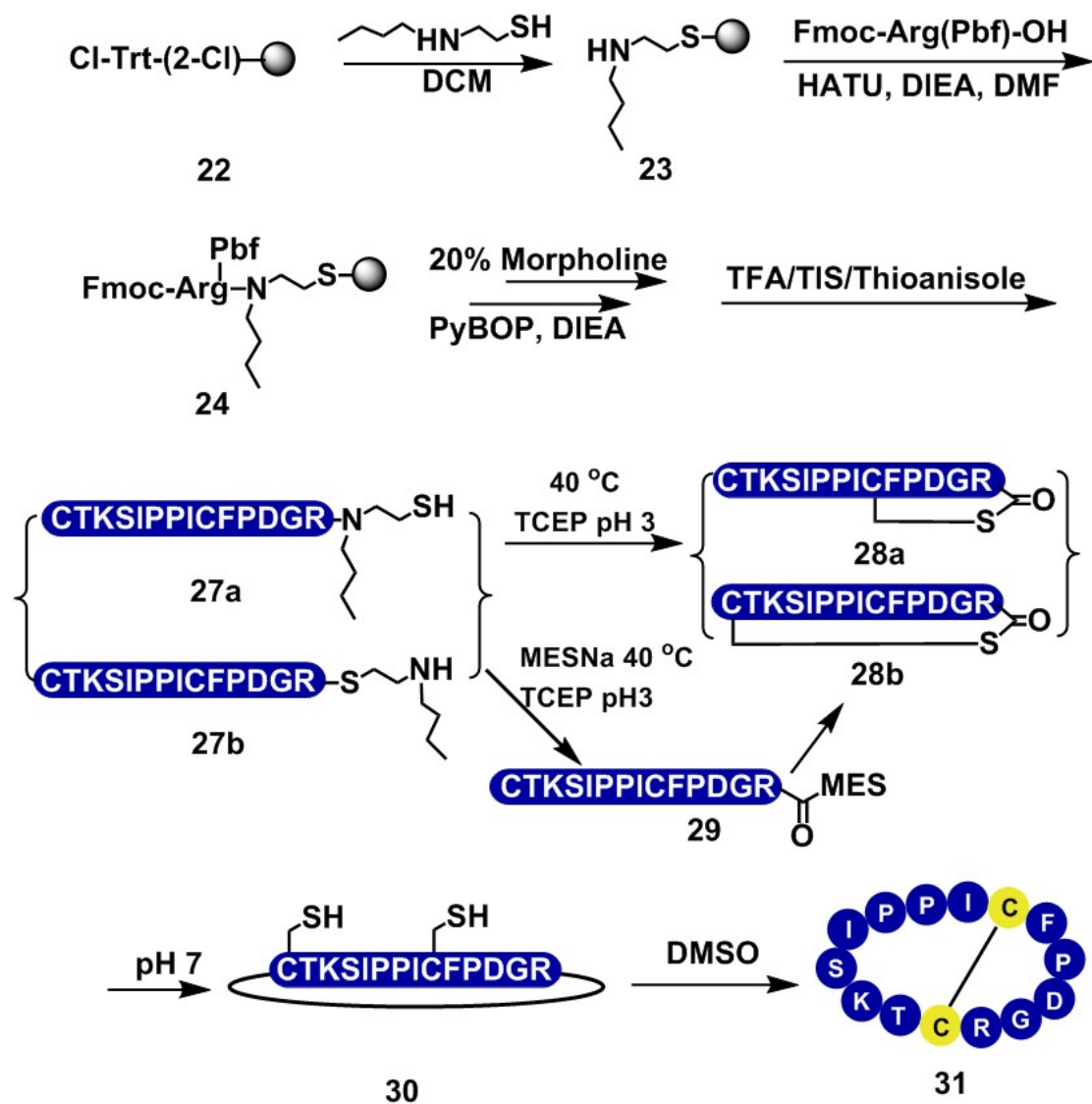


Figure 3.29 Scheme of TEBA-mediated synthesis of SFTI-1.

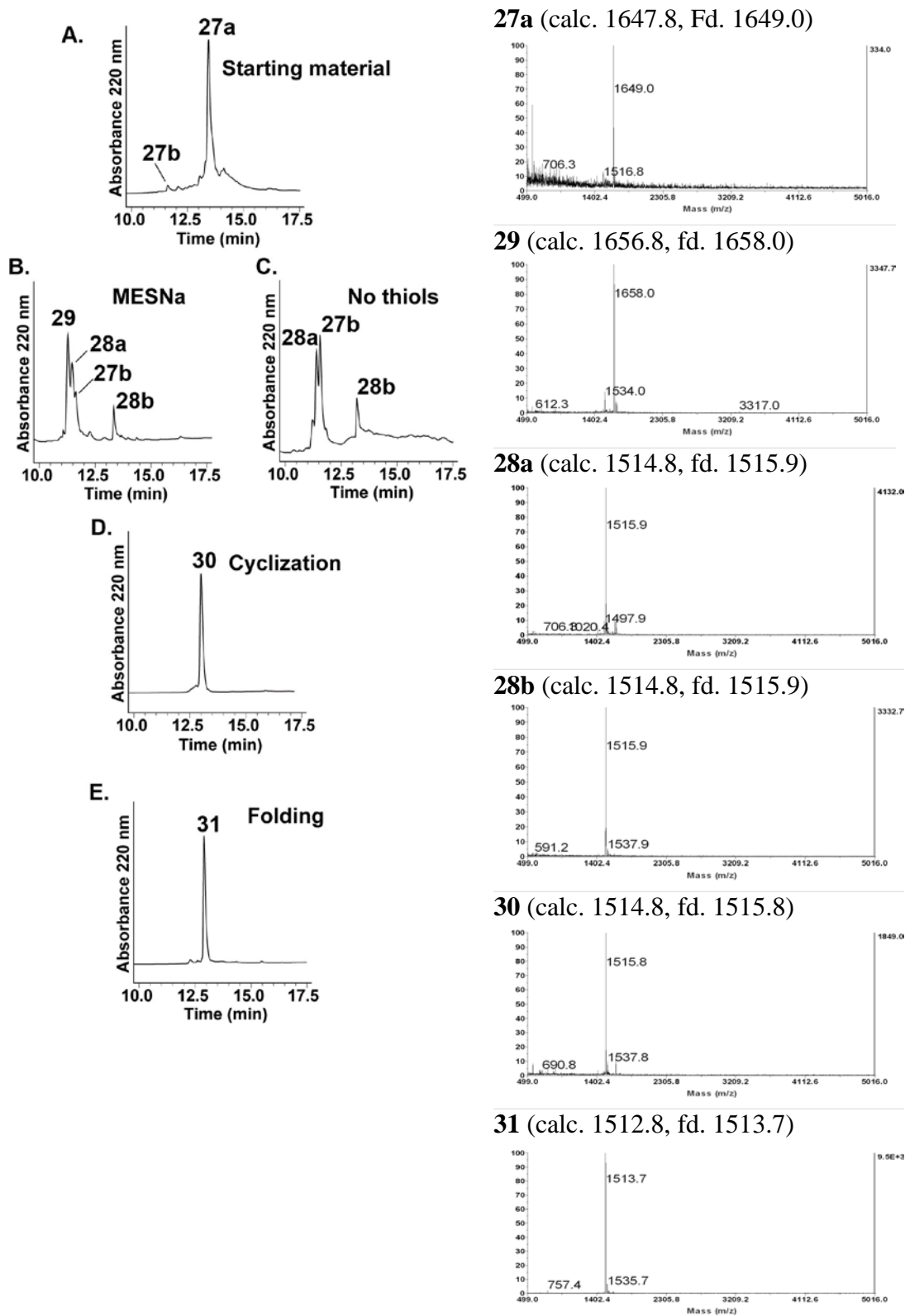


Figure 3.30 RP-HPLC and MS profiles of synthetic intermediates in the cyclization of SFTI-TEBA. A. starting material; B and C. thioesterification with and without MESNa; D. cyclization; E. folding.

4. Hydrazide as a thioester surrogate for preparation of cyclic CRPs

Lei Liu and co-workers recently reported a safety catch method for thioester formation using hydrazide as a thioester surrogate.²²⁴ The method generated thioester by conversion of a C-terminal hydrazide into a peptide azide and subsequent thiolysis of the acyl azide intermediate. The acyl azide was first introduced by Curtius as one of the earliest activation methods for peptide bond formation.^{288,289} Liu and co-workers successfully applied this reaction to thioester formation in acidic buffers at low temperature, which prevented the Curtius rearrangement from forming an isocyanate side product²⁹⁰.

Previous methods employing peptide hydrazide required necessary protecting at sensitive amino acids such as Cys or Lys sites. Liu's method used the unprotected peptide hydrazide which was activated by sodium nitrite at pH 2-3 and -10 °C. The low temperature and acidic conditions protected the peptide side chain functional groups from oxidation and allowed a smooth transition to peptide acyl azides. Then a large excess of thiols was added to generate a peptide thioester through the thioesterification and also reduced any oxidized Cys. This method is Fmoc-compatible and adopted by us for synthesis of cyclic CRPs to compare with the TEA series of thioester surrogates.

4.1 Synthesis of kalata B1 on the hydrazide resin by Fmoc chemistry

The hydrazide resin was synthesized by coupling the commercially available hydrazine to Cl-Trt(2-Cl) resin in DMF in the presence of excess DIEA (**Figure 3.31**). Since there were six cysteines in the kB1 sequence, the most unhindered Gly-Cys bond was selected as the cyclization site. Thus, the peptide chain elongation started with Gly by manually synthesis and was subsequently carried out in a microwave-assisted automatic peptide synthesizer by standard Fmoc SPPS. Final cleavage by TFA afforded the desired unprotected kB1 hydrazide (H-CTCSWPVCTRNGLPVCGETCVGGTCNTPG-NHNH₂). The product was analyzed and purified by RP-HPLC (**Figure 3.32A**). The MALDI-TOF MS confirmed the molecular weight of purified product (**Figure 3.32B**). The detected mass of kB1 hydrazide was 2929.4 while the expected mass was 2929.2 [M+H]⁺.

4.2 One pot synthesis of kB1

As shown in **Figure 3.31**, the one pot synthesis of kB1 included three steps: azide formation, cyclization and oxidative folding as shown.

For the azide formation, the unprotected peptide was dissolved in an acidic aqueous buffer (pH 2) and treated with excess oxidant NaNO₂ (10 eq.) at 0-4 °C. Low temperature was important for this step for avoiding side reactions. Thus, kB1 hydrazide and freshly prepared sodium nitrite were chilled on ice before conducting the reaction. After 20 min, the formation of kB1 azide was monitored by analytical HPLC and MALDI-TOF MS. The azide intermediate was unstable and the high-energy laser during mass spectrometry break the N=N double bond of azide and results in the loss of N₂ to generate a CO-N⁺ ion (M-N₂+H⁺) (**Figure 3.33**).

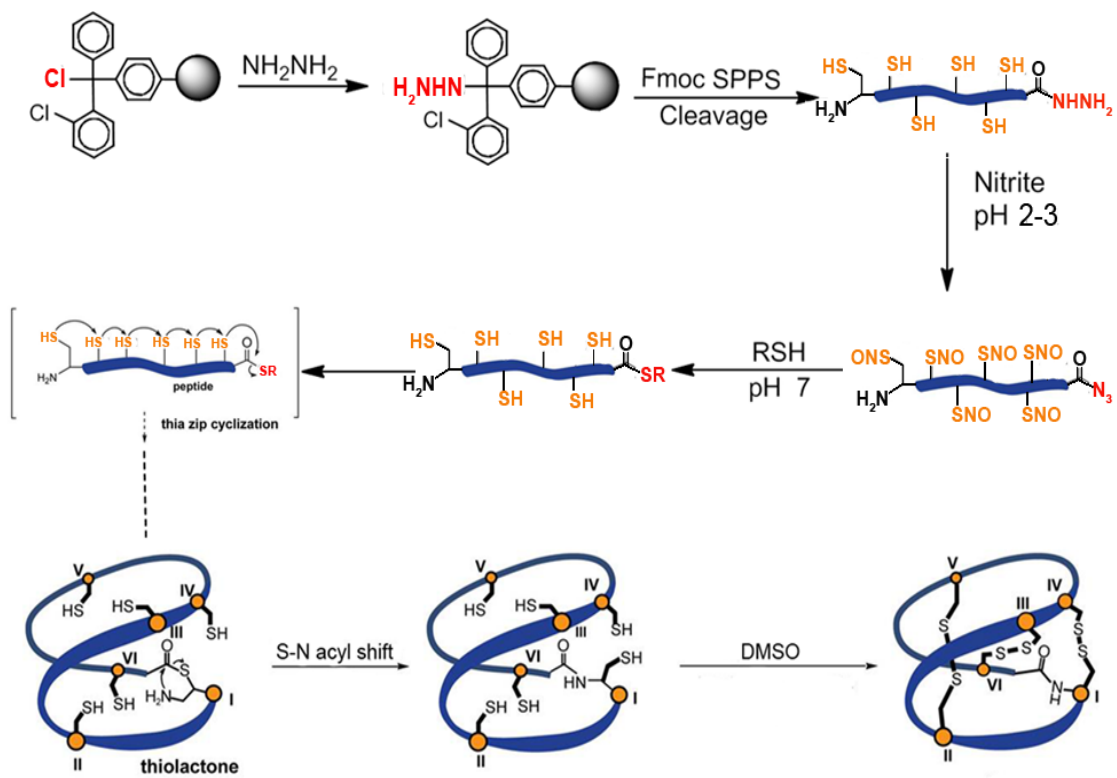


Figure 3.31 Scheme of one pot synthesis of kB1 by the hydrazide resin.

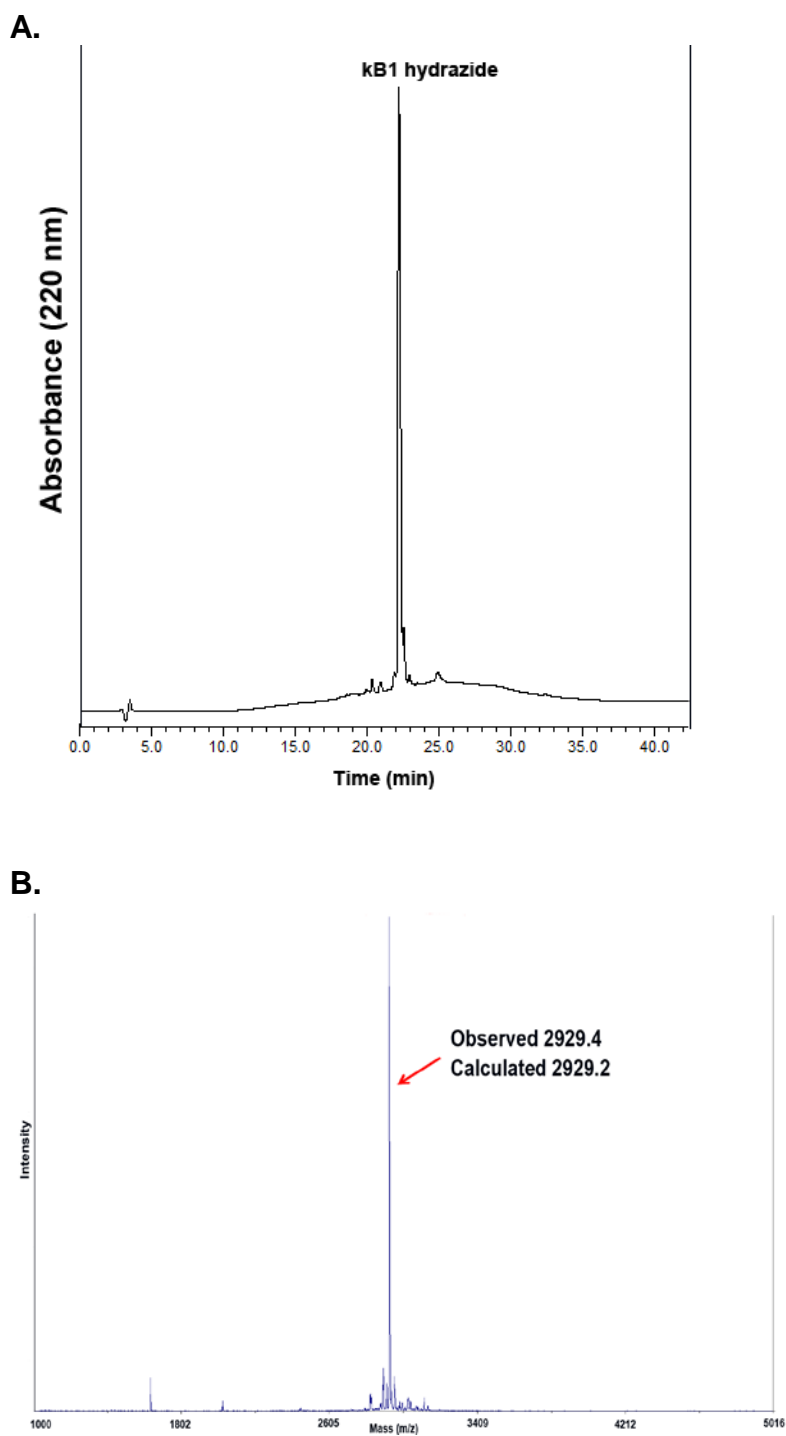


Figure 3.32 RP-HPLC and MALDI-TOF MS profiles of crude kB1 hydrazide after TFA cleavage. kB1 hydrazide eluted at 22.3 min with a linear gradient from 10-50% ACN in aqueous 0.1% TFA for 40 min.

Then external thiol MMA ($\text{HSCH}_2\text{COOCH}_3$, 100 eq.) was added into the reaction solution and the pH was adjusted to 7 immediately with 1 M NaOH to perform the cyclization reaction. A large amount of external thiols was necessary to protect the internal cysteines and capture the active acyl azide in this step. It must precede the pH adjustment to prevent undesired substitution reaction between the functional side chains, such as $-\text{NH}_2$, $-\text{SH}$, and the acyl azide at neutral pH. After the pH was adjusted, the reaction mixture was allowed to proceed at room temperature for 2 h to give the cyclic kB1 (**Figure 3.33**).

Subsequently, the reaction mixture was diluted 5 times with 2-propanol in presence of 10% DMSO to perform the oxidative folding (**Figure 3.33**). Then morpholine was added to a final concentration of 2% (v/v). The organic folding produced native kB1 with faster rate and higher efficiency, which will be discussed in chapter 4.

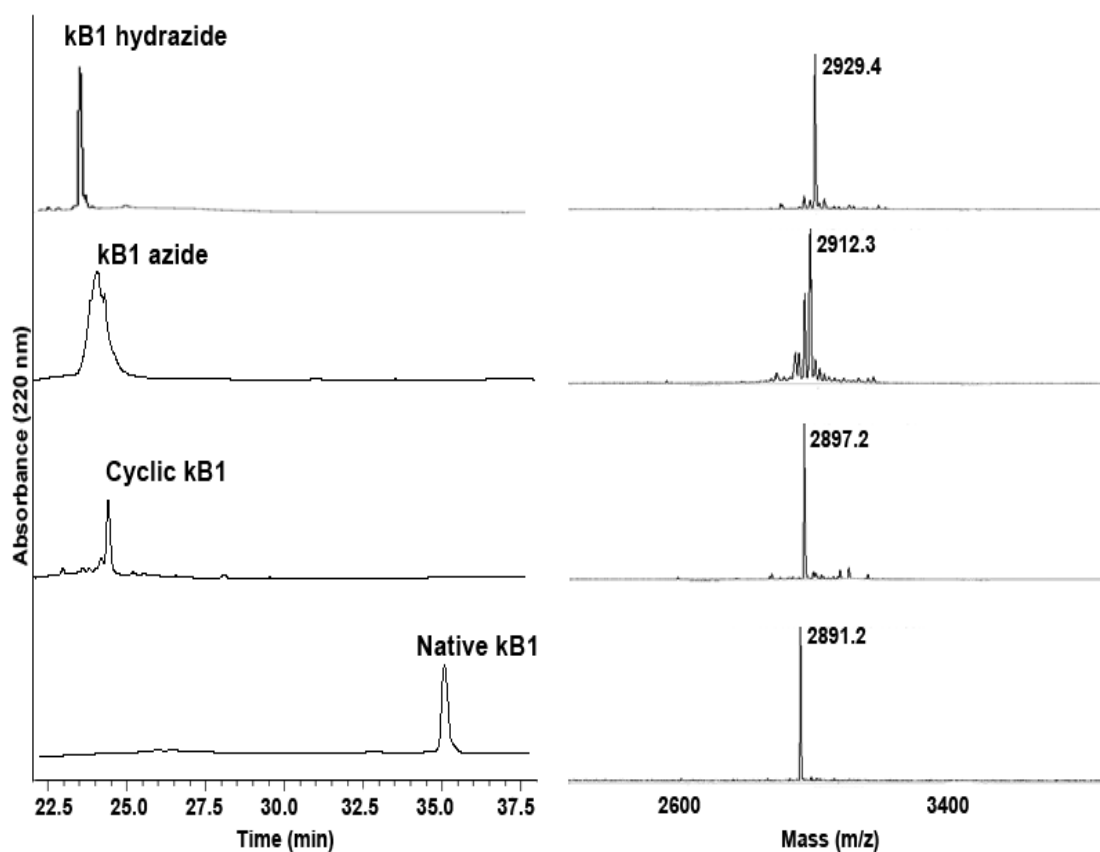


Figure 3.33 RP-HPLC and MALDI-TOF MS profiles of synthetic kB1 hydrazide, kB1 azide, cyclic kB1 and native kB1.

To optimize the one pot synthesis of kB1, several parameters in the reaction condition were investigated. First, the peptide concentration was adjusted to 0.5 mM instead of 1 mM reported in the Liu's paper. Unlike standard ligation or cyclization conditions, our one-pot approach used acidic *Milli-Q* H₂O instead of inorganic buffers containing denaturing reagents like Gnd-HCl. This was due to compatibility concern since the organic solvent involved in the later oxidative folding step was not compatible with the salty cyclization buffer. Therefore, to make sure the one-pot synthesis proceed smoothly and prevent peptide aggregation, the peptide was adjusted to a lower concentration compared to that used in denaturing conditions. As shown in **Figure 3.34**, cyclization with peptide concentration at 1 mM did not afford good yield under the one-pot condition. Although native kB1 was obtained under this condition, the peptide aggregation was serious after adding the external thiol MMA into the reaction mixture and the cyclization and oxidative folding reactions were not completed at 2 h and 1 h, respectively. Compared to 1 mM, 0.5 mM peptide concentration resulted in a better yield and no aggregation was observed. Both cyclization and oxidative folding reactions were completed at 2 h and 1 h, respectively.

Besides peptide concentration, the ratio of nitrite and external thiol was also investigated under the proposed condition (**Table 3.3**). At first, the thiol concentration was kept at 100 eq. while the nitrite concentration was adjusted. It was found that the cyclization yield remained below 25% when the nitrite concentration was 1 or 5 eq. After increasing the nitrite concentration to 10 eq., the cyclization yield was greatly increased to 90%. Then we kept the nitrite concentration at the optimal 10 eq. and evaluated various the thiol concentrations from 10 to 100 eq. There was not much difference among yields under the three conditions. Therefore, 10 eq. thiol was sufficient for the cyclization reaction.

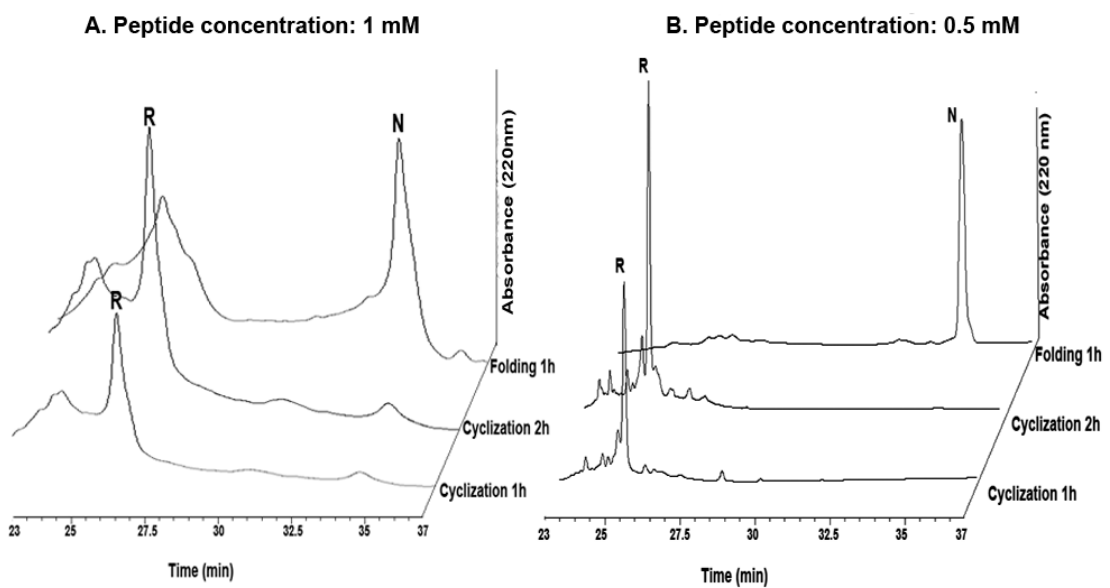


Figure 3.34 RP-HPLC profiles of the cyclization and folding results with 1 mM and 0.5 mM peptide. R and N represent the reduced and native form of peptide.

Table 3.3 Cyclization conditions with different nitrite and thiol concentrations.

| Run | Nitrite conc. (eq.) | Thiol conc. (eq.) | Yield (%) |
|-----|---------------------|-------------------|-----------|
| 1 | 1 | 100 | 18 |
| 2 | 5 | 100 | 25 |
| 3 | 10 | 100 | 90 |
| 4 | 10 | 50 | 89 |
| 5 | 10 | 10 | 85 |

For the oxidative folding condition, organic oxidative folding was employed which will be discussed in detail in chapter 4. First, the peptide concentration was reduced by diluting the reaction mixture with 2-propanol. This was for prevention of cross linking problem in oxidative folding. The final concentration of peptide for oxidative folding was 0.1 mM. The basicity was then investigated to optimize the oxidative folding condition. Since the cyclization reaction was performed at pH 7, it was envisioned that no additional bases were necessary. However, it was found that little amount of native form kb1 was produced in the oxidative folding after 1 h (**Figure 3.35**). Folding intermediates were found to be populated as the major products. Then 2% morpholine was added into the reaction mixture to increase the basicity. The amount of native form of kB1 was found to gain significantly after adding morpholine for 10 min. The yield was 81% after 1 h oxidative folding reaction with 2% morpholine. Therefore 2% morpholine was found to be necessary to maintain the basicity for oxidative folding reactions.

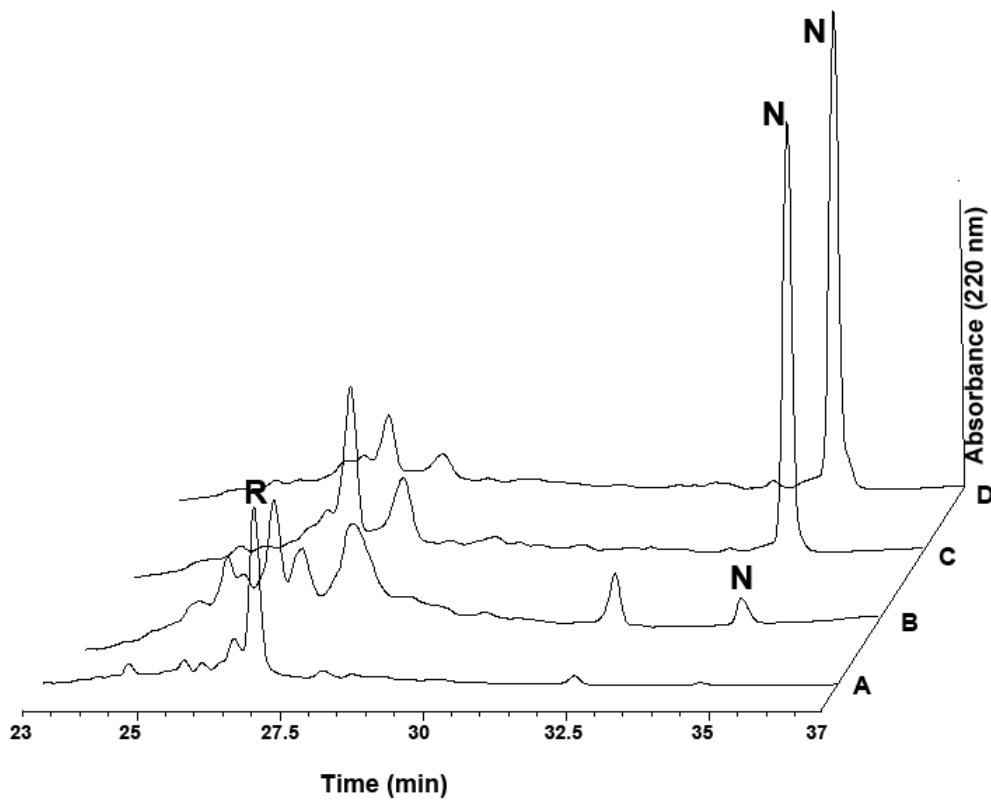


Figure 3.35 RP-HPLC profiles of the oxidative folding under different basicities.
A. cyclization after 1 h; **B.** oxidative folding without morpholine after 1 h; **C.** oxidative folding with morpholine 10 min; **D.** oxidative folding with morpholine 1 h.
R and **N** represent the reduced and native form of peptide, respectively.

In the one-pot synthesis, MMA served as a redox reagent in the oxidative folding reaction besides its catalyst role in the cyclization reaction. Although 10 eq. MMA was shown to be sufficient for the cyclization reaction, it was not optimal for the following oxidative folding reaction (**Figure 3.36**). The RP-HPLC profiles showed that both 10 eq. and 50 eq. MMA conditions led to population of some folding intermediates, resulting in less than 50% yield after 1 h in the oxidative folding (**Figure 3.36**). Prolonged treatments did not increase the yield. However, conditions with 100 eq. MMA were able to afford the native form of kB1 in 90% yield after 1 h. Folding intermediates were observed in less than 1% in this condition (**Figure 3.36**).

In summary, the one pot synthesis of kB1 was optimized as 0.5 mM peptide hydrazide reacted with 10 eq. sodium nitrite and 100 eq. MMA, which afforded native kB1 in 90% yield. The synthetic kB1 was characterized by co-injecting with the native kB1 standard. They were found to co-elute as shown in **Figure 3.37**.

The molecular weight of synthetic kB1 was confirmed by MALDI-TOF MS (**Figure 3.38A**). MS/MS analysis (**Figure 3.38B**) showed that it could not be fragmented suggesting its cyclic cystine knot structure.

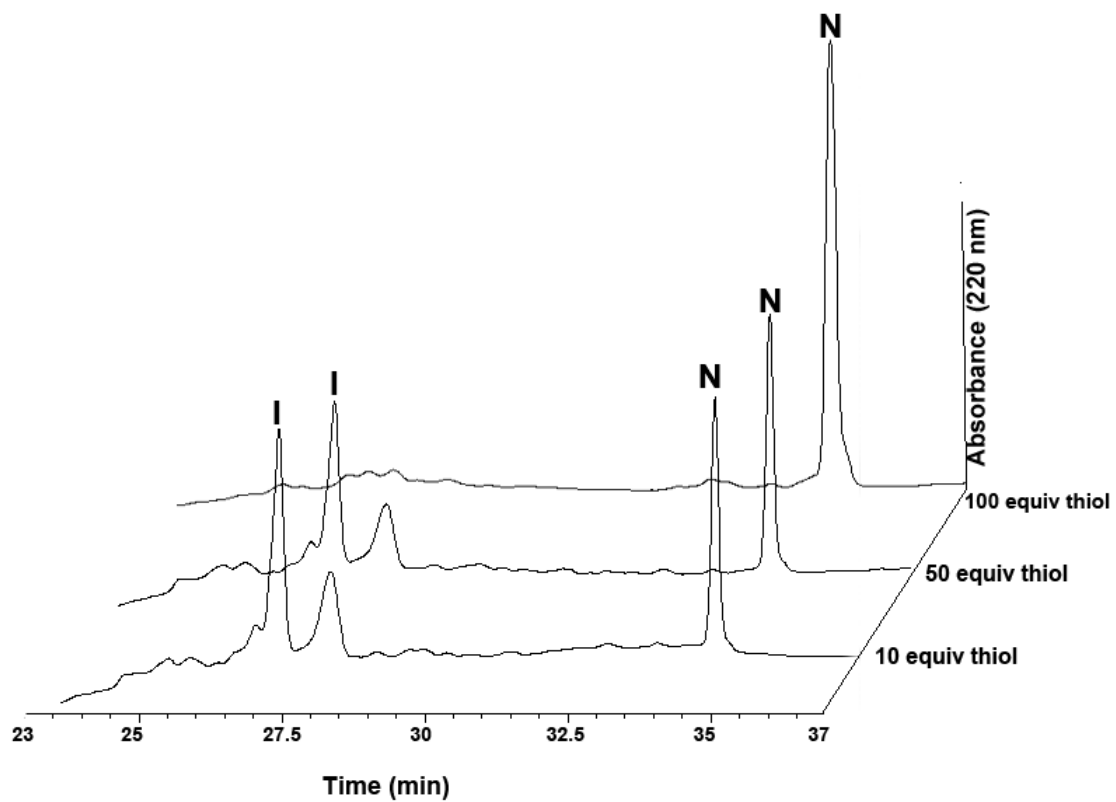


Figure 3.36 RP-HPLC profiles of oxidative folding reactions with different concentrations of thiols after 1 h. N and I represent the native form kB1 and the folding intermediate, respectively.

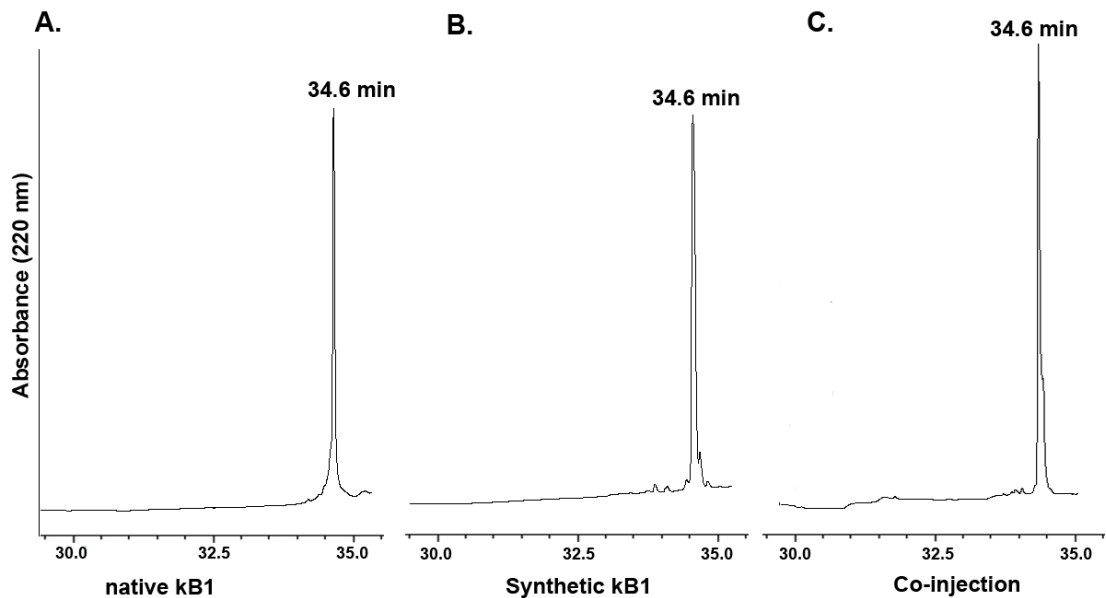


Figure 3.37 Co-injection of native kB1 and synthetic kB1 in RP-HPLC. **A.** native kB1 eluted at 34.6 min; **B.** synthetic kB1 eluted at 34.6 min; **C.** Co-injection of Native kB1 and synthetic kB1 eluted at 34.6 min. The HPLC gradient is 10-50% ACN in aqueous 0.1% TFA for 40 min.

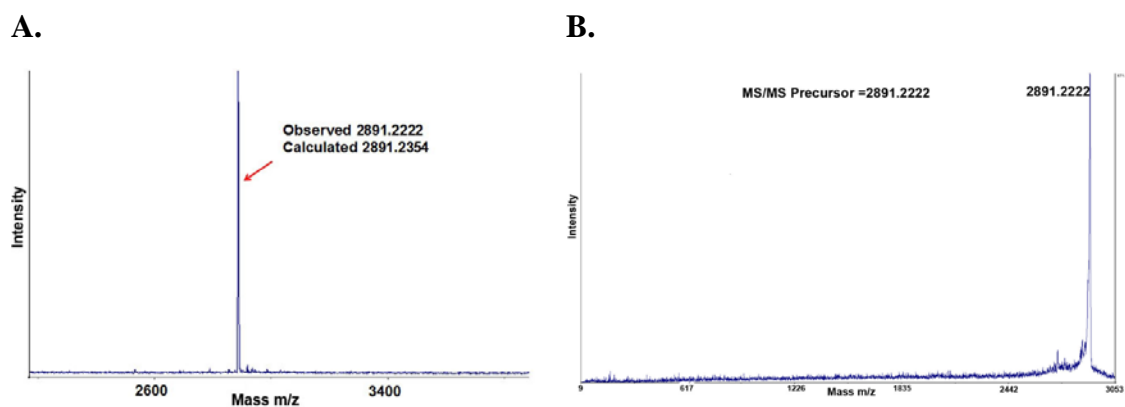


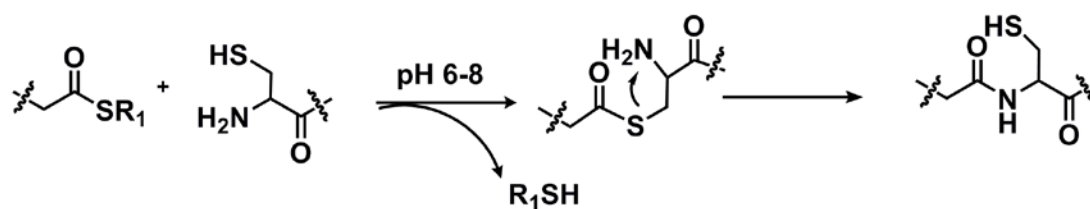
Figure 3.38 MALDI-TOF MS and MS/MS profiles of synthetic kB1. **A.** MS; **B.** MS/MS.

5. Discussion

The half-life of an amide bond is known to be 100 years²⁹¹, thus it is very difficult for a peptide bond to break and isomerize to a thioester bond in nature without enzymatic assistance. The isomerization of a special Xaa-Cys bond was exploited to enable the thioester formation by reversing the chemical ligation reaction. The idea comes from the chemoselective ligation between a thioester and a cysteine which forms an amide bond in basic conditions through an S-N acyl shift (**Figure 3.39**). By reversing the ligation reaction, the thioester can be regenerated from the amide via an N-S acyl shift. The common functional group in both the ligation and the thioester formation is the TEA moiety which mediates the formation of an amide or a thioester through the same five-member intermediate. Thus, we designed our thioester surrogates such as MeCys and TEBA based on the TEA moiety.

Since the isomerization reaction of Xaa-Cys bond is reversible and thermodynamically unfavorable under unassisted conditions, activation through heating at acidic conditions or conformational assistance are required. As we have shown in the previous section, MeCys and Cys displayed a 150 fold difference in the rate of thioester formation. This is because by alkylating the scissile amide bond, the *cisoid* conformation of amide bond is favored which facilitates the N-S acyl shift reaction. The fact that N-S acyl shift reaction is assisted by the *cisoid* amide was first discovered in intein-mediated protein splicing. Klabunde and co-workers found that the conserved residue assisting the initial acyl rearrangement was in a non-standard *cis* conformation and any residue which exists in a similar conformation and formed a similar hydrogen bonding can replace this residue.²⁹²

Native Chemical Ligation (S-N acyl shift)



Thioester formation (N-S acyl shift)

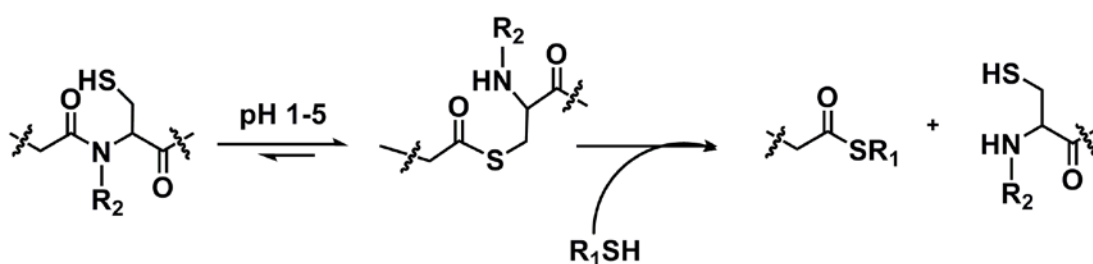


Figure 3.39 Scheme of native chemical ligation and the reversed thioester formation. R₁=alkyl or aryl groups; R₂=alkyl, aryl or auxiliary groups.

Our TEA series of thioester surrogates employ the same mechanism as intein to generate thioesters by a *cis* amide bond-assisted N-S acyl shift reaction. The catalytic role of intein to mediate the protein splicing involves four reactions: N-S, S-S, N-N and S-N acyl shifts. The TEA thioester surrogates mimic the role of intein by mediating three out of four acyl shift reactions (**Figure 3.40**). Two approaches share the common theme by exploiting the isomerization reaction of TEA moiety in a cysteine residue to form an amide bond. Instead of using intein with 198 residues, which requires protein expression to effect, our approach utilizes a chemical moiety smaller than one amino acid and requires a simple coupling reaction to realize the function. The feasibility of our TEA approach has been demonstrated successfully by preparing the naturally occurring cyclic CRP SFTI-1. There are also differences between these two approaches. Intein catalyzes the isomerization of amide bond at neutral pH whereas our TEA approach requires heating in acidic conditions.

We employed a Cys-peptide-TS (thioester surrogate) scheme to prepare cyclic peptides using an amide-to-amide approach. This approach mimicking the intein-mediated cyclization employs an amide bond as a latent thioester which can be utilized to form a new amide bond (**Figure 3.41**).

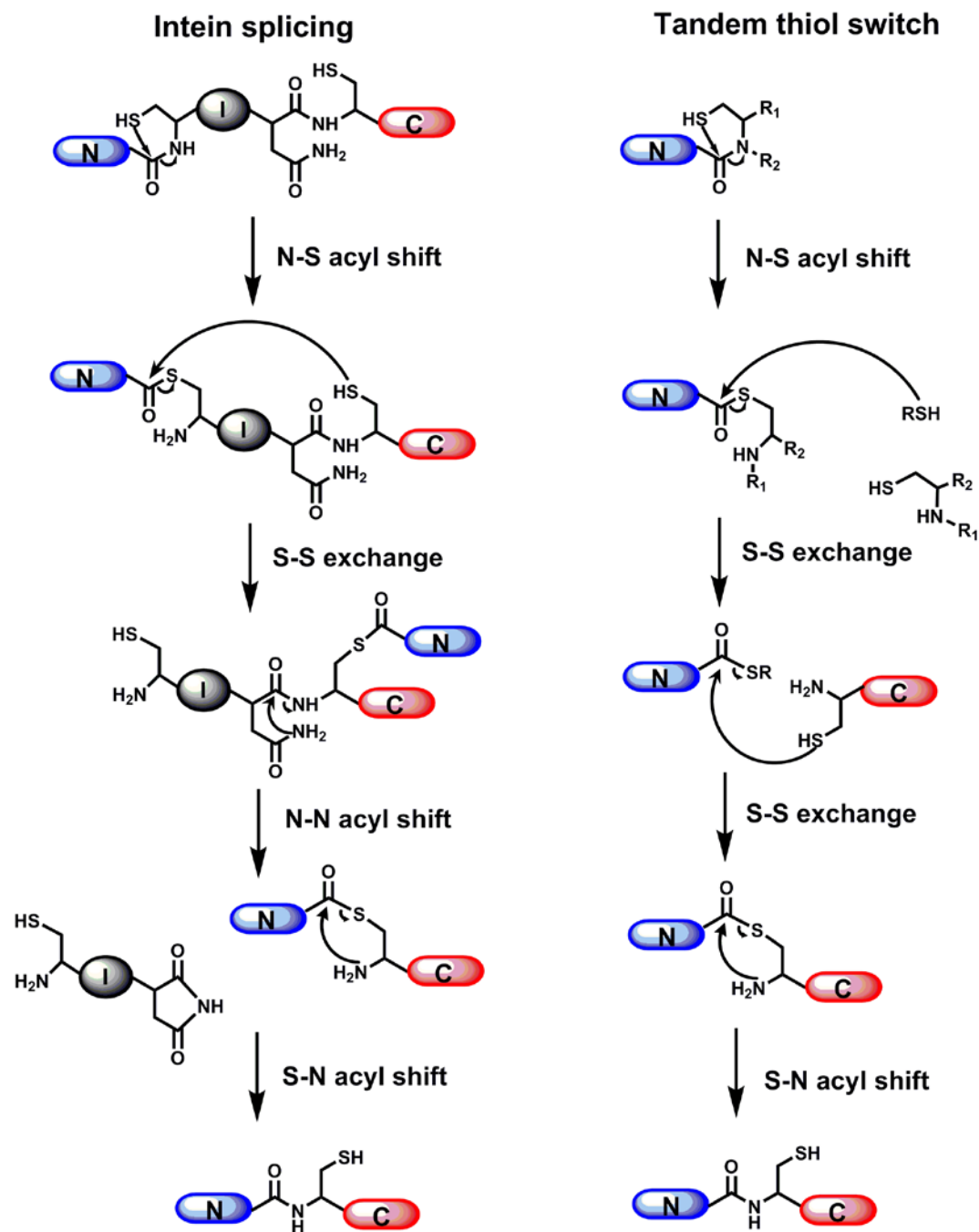


Figure 3.40 Comparison of intein-mediated protein splicing and TEA-mediated tandem thiol switch. I=intein; N=N-terminal peptide; C=C-terminal peptide; R1,R2=alkyl, aryl or auxiliary groups.

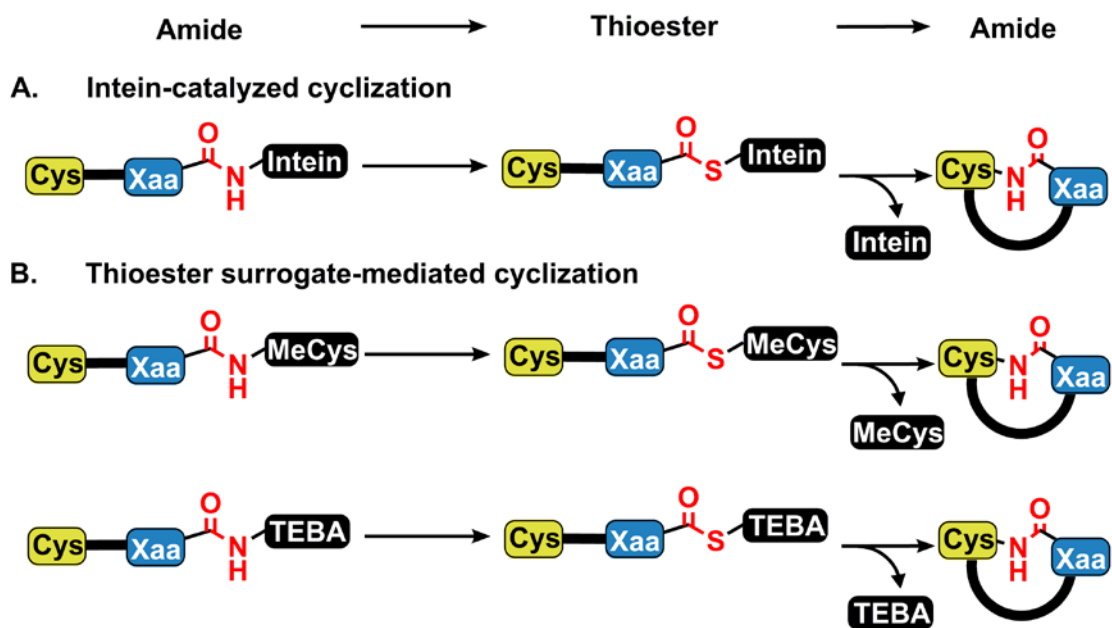


Figure 3.41 Schematic demonstrations of intein-catalyzed cyclization and TEA thioester surrogate-mediated cyclization.

Both methods employed a peptide with an N-terminal cysteine residue which is obligatory for the cyclization reaction, whereas the intein or thioester surrogates are incorporated at the C-terminus of the linear peptide via an amide bond. Under uncatalyzed reactions, the amide bond will isomerize to a thioester bond and thus a C-terminal thioester is produced. Finally, a new amide bond will be formed between the N-terminal cysteine and the C-terminal thioester via ligation reactions with concurrent excision of intein or thioester surrogates. Examples include MeCys and TEBA thioester surrogates. Both of them mimic the intein-mediated peptide cyclization by mediating a series of acyl shift reactions to cyclize the peptide backbone and excise themselves from the peptide precursor. Advantages of the TEBA group as a thioester surrogate include its ease of synthesis and compatibility with Fmoc chemistry. Compared with MeCys, TEBA eliminates potential side reactions such as the β -elimination and piperidine addition during peptide synthesis and it was also shown to be more reactive due to the removal of the carbonyl group.

Cyclic CRPs have been prepared exclusively by Boc chemistry in 1990s.^{32,149,150,184,293} Although several methods have been reported in Fmoc chemistry for thioester formation, few of them have been applied successfully to prepare natural occurring cyclic CRPs. Representative examples include preparation of MCoTI-II by sulfonamide linkers¹⁸³, synthesis of cyclotides by peptide hydrazide¹⁸⁶ and production of cyclic defensins using the acyl urea linker¹⁸⁸. Most of them use safety catch type of thioester surrogates instead of safety switch type. We also prepared cyclic peptide kB1 by a safety catch thioester surrogate- hydrazide. My results showed that this method was fast, efficient and compatible to our one-pot scheme. However, careful handling and special expertise were required. The outcome could be unexpected due to potential side reactions. First, the activation step to convert a hydrazide to an azide must be performed at low temperature (better below 0 °C) to avoid irreversible Curtius rearrangement and oxidation on undesired side chain functional groups. Second, during the thiolysis step, $-\text{NH}_2$ or $-\text{OH}$ groups could form lactams or lactones by reacting with the active azide since it was a good leaving group and susceptible to nucleophilic attacks. These side products are irreversible and can not lead to cyclic products in the following cyclization reaction. The TEBA method prevented such side reactions efficiently by performing the reaction in acidic

conditions. Second, excessive thiols were required as scavengers to protect the side chain functional groups before adjusting the pH to neutral. In addition, the nitrite could lead to oxidation of $-SH$ into $-SNO$, which may result in disulfide formation reducing the number of reactive cysteines in the thia zip cyclization. Finally, the pH adjustment was crucial for the two-step conversion in hydrazide-mediated thioester formation. The reaction must be performed at a precise pH whereas the TEBA-mediated thioesterification can be carried out at a wide range of pH from 2-5. Collectively, the safety catch type of thioester surrogates which commonly involved a drastic activation step in the conversion may not be compatible with preparation of cyclic CRPs since it could lead to irreversible modifications such as oxidation of cysteines or other sensitive residues. In contrast, the TEA type of thioester surrogates, which permits a smooth transition from amide to thioester and mediates thioester formation at a wide pH range, are more applicable for synthesizing cyclic CRPs.

Chapter 4 Oxidative folding of cysteine-rich peptides

1. Introduction

In the synthesis of cysteine-rich peptides, the oxidative folding remains a bottleneck because of its unpredictability. *In vivo*, oxidative folding of proteins to form disulfide bonds is often characterized as a fast (less than one minute) and precise event.²⁹⁴ However, *in vitro* folding processes are characterized by slow kinetics of disulfide bond formation and product heterogeneity consisting of folding intermediates, isomeric and native proteins. Conventional global oxidative folding in aqueous conditions often takes hours or days to complete and the yields can be considerably low. Chemoselective oxidative folding can purposely form correct disulfide bonds by deprotecting one pair of cysteines at a time. However, low yields and side reactions, such as undesired oxidation at sensitive amino acids, impedes its wide application.

To mimic the oxidative folding *in vivo* and to improve the folding efficiency, we developed an organic folding system to accelerate the folding rates *in vitro*. In this system, four major components are introduced: organic base, organic solvent, structure enhancing co-solvent and the redox compound. Organic base is used to adjust a suitable basicity for the disulfide formation while organic solvent provides an environment in which the thiol-disulfide exchange reaction is 1000 times faster than in aqueous conditions as shown in literature^{295,296}. Thus with the help of organic base and solvent, the oxidative folding in our organic system would be greatly accelerated compared with the aqueous conditions. On the other hand, structure enhancing co-solvents and redox reagents confer the fidelity to the organic folding products. Using structure enhancing solvents such as TFE, the native secondary structure is formed preferentially thus brings the native pairing of cysteines in proximity to form the correct disulfide connections. Using organic-solvent-compatible reductant cysteamine and oxidant DMSO, the incorrect disulfide connection is able to be shuffled to the correct disulfide connection by thiol-disulfide exchange. Here, we describe the application of the organic folding system using two model peptides kB1 and

endothelin-1 (ET-1) as examples. Effects of solvent, basicity, redox reagents and chaperon-like solvent are investigated in this chapter.

2. Oxidative folding of kB1 in organic solvents

kB1, the prototypic cyclotide, contains three intramolecular disulfide bonds with a cyclic backbone, which forms a constraint cyclic cystine knot. For optimizing an oxidative folding condition for kB1, we used the fully reduced kB1 as starting material and refolded it under various conditions.

To prepare the reduced kB1, native kB1 from plant extractions was denatured to reduce the three disulfide bonds while its cyclic backbone was retained. The HPLC profile of native kB1 is shown in the panel **A** of **Figure 4.1**. Its molecular weight was determined by MALDI-TOF MS and is shown in the panel **B**. We detected a mass of 2891.39 Da consistent to the expected mass of 2891.2 Da. The purified native kB1 was subjected to TCEP reduction to remove all the disulfide bond linkages. After reduction, the reduced kB1 was purified by RP-HPLC. The profile showed that the retention time for kB1 shifted from 24.5 min to 18.8 min after reduction with a linear gradient of 10-100% ACN for 44 min, suggesting that the native kB1 was more hydrophobic than the reduced form. An increment by 6 Da in the molecular weight after reduction indicated that reduction of three disulfide bonds was achieved (panel **C** and **D**). The zoom-in MS profiles showed the difference between the two masses. Immediately after HPLC purification, the fully reduced peptide sample was lyophilized and subjected to the oxidative folding experiments.

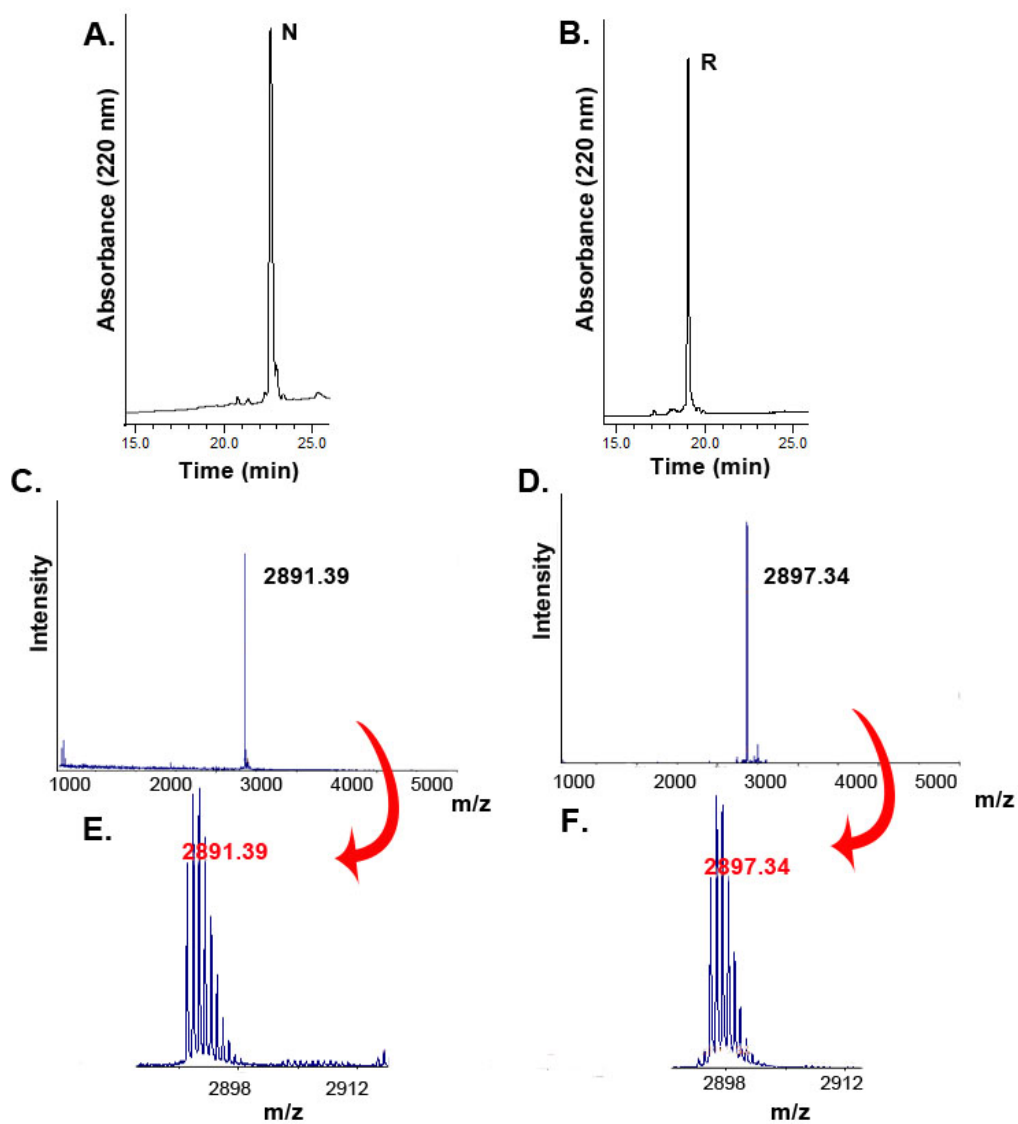


Figure 4.1 RP-HPLC and MALDI-TOF MS profiles of native and reduced kB1. Panel A and B show the RP-HPLC profiles of native and reduced kB1 with N and R representing native and reduced respectively. Panel C and D show the MS profiles of the peaks N and R accordingly. Panel E and F are the zoom-in profile of MS results in panel C and D which demonstrate the difference between these two masses.

2.1 Solvent effect

To test our hypothesis that kB1 could be oxidatively folded favorably in neat organic solvents, we employed a binary solvent system consisting of an organic base (pyridine or morpholine) and an alcohol (2-propanol) to mimic the *in vivo* folding conditions. Previous researches have explored the role of alcohols as co-solvents in oxidative folding of cyclotides. We and others demonstrated that 2-propanol was an effective co-solvent under aqueous conditions for promoting the oxidative folding of cyclotides.^{184,262} Pyridine and morpholine served a dual role as both a solvent and an organic base that provide suitable basic conditions to conduct the disulfide bond formation.

In the first trial, the reduced kB1 was subjected to oxidative folding conditions listed in **Table 4.1**. We used a binary solvent combination containing pyridine (or morpholine) and 2-propanol. All the conditions contained 100 μ M peptide, 10% (*v/v*) DMSO and 100 mM cysteamine but varied in molar ratios of pyridine, morpholine and 2-propanol. DMSO and cysteamine combination was used as the oxidizing and reducing agent, respectively, because GSH/GSSG was not compatible with the organic system due to solubility problem. Oxidative folding reactions were quenched by acid at different time intervals and monitored by RP-UPLC (Ultra Performance Liquid Chromatography) and MALDI-TOF MS.

Previous publications showed that the oxidative folding of kB1 in aqueous conditions required at least 48 h for completion.¹⁸⁴ Thus, we monitored the oxidative folding process of kB1 every 12 h. However, it was observed that the oxidative folding of kB1 was completed within 12 h under our conditions as the yield of correctly folded kB1 did not increase after the first time point 12 h. Thus, oxidative folding yields of different conditions at 12 h were calculated and compared. Among all conditions listed in **Table 4.1**, those without 2-propanol (Run 1, 3 and 6) did not afford a reasonable yield (<10%) at 12 h. The other two conditions containing 70% 2-propanol were able to achieve a yield of 88% and 91%, respectively, after 12 h. These results suggested that the presence of 2-propanol as an organic co-solvent was effective in improving the oxidative folding yield of kB1. It should be noted that a prolonged treatment of Run 1 in 48 h increased the yield to 51%, suggesting that

pyridine could also serve as an organic co-solvent. In addition, both pyridine and morpholine were compatible to the organic folding system and conditions containing 20% pyridine or 20% morpholine gave a comparable yield (88% and 91%) at 12 h. These results suggested that they were suitable organic bases for the organic folding system. Moreover, it proved that the oxidative folding of kB1 in neat organic solvents was successful and applicable.

The effect of this binary solvent system was further demonstrated by varying the 2-propanol concentration from 0-90%. As it is shown in **Figure 4.2**, all three conditions (B, C and D) used a binary solvent system containing pyridine and 2-propanol. Without 2-propanol, native form kB1 could not be formed in more than 10% yield. While in presence of 50% 2-propanol, the yield increased to 85%. Further increasing the 2-propanol to 90% led to incorrectly folded isomers and decreased the yield to 3%. This was due to the lack of base in the oxidative folding condition. Without organic base, the disulfide bond formation and thiol-disulfide exchange reaction were exceedingly slow. Thus, the native form kB1 was hardly formed at 12 h. In general, the yields using the binary solvent combination were superior to using pyridine or 2-propanol alone, suggesting that both base and co-solvent were important in our folding system.

Based on the above results, the oxidative folding was much faster in the neat organic conditions than in aqueous conditions. It was completed within 12 h while the corresponding aqueous condition requires 48 h to complete. Therefore, the reaction was monitored in shorter time scales to determine the reaction rate. We employed 20% pyridine and 70% 2-propanol as a model condition to test the reaction completion time. UPLC profiles at 0, 20, 40 and 60 min are shown in **Figure 4.3**. The reaction reached an 87% yield after 40 min. When the reaction was allowed to proceed for 60 min, the yield was slightly increased by 2%. These results demonstrated that the oxidative folding reaction was greatly accelerated in organic folding system which can be completed within 1 h.

Table 4.1 Oxidative conditions of kB1 with varying solvent combinations.

| Run | Pyridine (v/v %) | Morpholine (v/v %) | 2-Propanol (v/v %) | Yield* (%) |
|-----|---------------------|-----------------------|-----------------------|---------------|
| 1 | 90 | - | - | <10 |
| 2 | 20 | - | 70 | 88 |
| 3 | - | 90 | - | <10 |
| 4 | - | 20 | 70 | 91 |
| 5 | 45 | 45 | - | <10 |

*The yield is calculated accordingly based on the RP-HPLC profile of the reaction at 12 h.

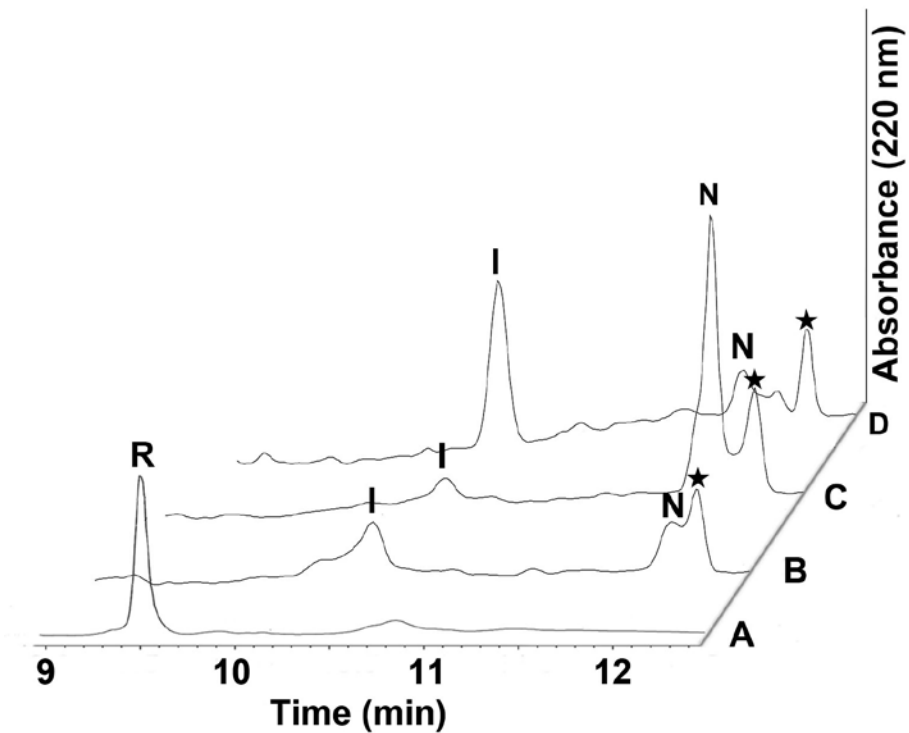


Figure 4.2 RP-UPLC profiles of the oxidative folding of kB1 showing the effect of binary solvent mixtures. A represents the reduced form of kB1. The three conditions: **B** contained 0% 2-propanol and 90% pyridine, **C** contained 50% 2-propanol and 40% pyridine and **D** contained 90% 2-propanol and 0% pyridine. All conditions contained 100 μ M peptide, 100 mM cysteamine and 10% (v/v) DMSO. **R** and **N** represents the reduced and native form of peptide, **I** denotes the folding intermediate while \star denotes the impurity from the plant extraction.

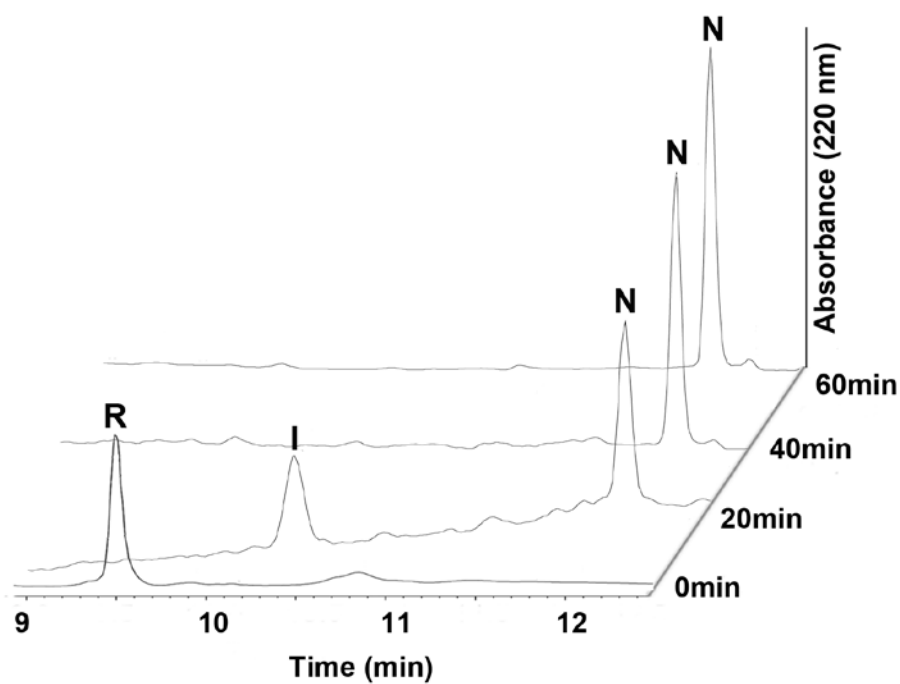


Figure 4.3 RP-UPLC profiles showing the process of the oxidative folding of kB1. **R** and **N** represent the reduced and native form of peptide, **I** represents the folding intermediate. The condition here contained 20% pyridine, 70% 2-propanol, 10% DMSO, 100 mM cysteamine with 100 μ M peptide.

2.2 Basicity effect

In aqueous conditions, the oxidative folding is invariably performed at pH 7.5-8.5 which is the optimal pH range for thiol-disulfide exchange reactions. Thus, maintaining suitable basicity is important for our non-aqueous conditions. To investigate the basicity requirement of the oxidative folding under organic conditions, three bases, pyridine, imidazole and morpholine with a pKa of 5.25, 7.05, 8.36, respectively, were chosen. All three bases worked efficiently in our system (**Figure 4.4**). Under all the conditions, >70% native form kB1 formed except for the condition containing 5% pyridine. The highest yield for each base was 82% for pyridine, 84% for imidazole and 88% for morpholine. For every single base, the increase of the yield was dependent on the base concentration. The yield increased when a higher percentage of base was used. For pyridine, 5% condition yielded 27% desired product after 1 h while the 30% and 50% conditions led to 72% and 82% yield, respectively. To accelerate the thiol-disulfide exchange rates, we increased the basicity of the oxidative folding conditions by adding bases with higher pKa such as imidazole and morpholine to the folding conditions. It was observed that lower concentrations of bases were required in imidazole and morpholine conditions. For imidazole, 5% condition was already sufficient for the oxidative folding reaction which led to 80% yield after 1h. Further increasing the imidazole concentration to 10% and 15% did not affect much the final yield which was 83% and 84%, respectively. However, the reaction rate was accelerated by increasing the base concentration. Under the 5% imidazole condition, it required 1 h to allow the oxidative folding to complete. In 10% and 15% conditions, less than 20 min was enough to accomplish the reaction. Similar results were obtained from morpholine. 88% yield was achieved in the 5% morpholine condition after 1 h. Further increasing the morpholine concentration to 30% did not result in an increase of the yield. High concentrations of organic bases such as morpholine may cause problems during work up of the reactions because they were difficult to evaporate. However, this was not a problem in our conditions since as low as 5% organic base was already sufficient to allow the oxidative folding to proceed smoothly.

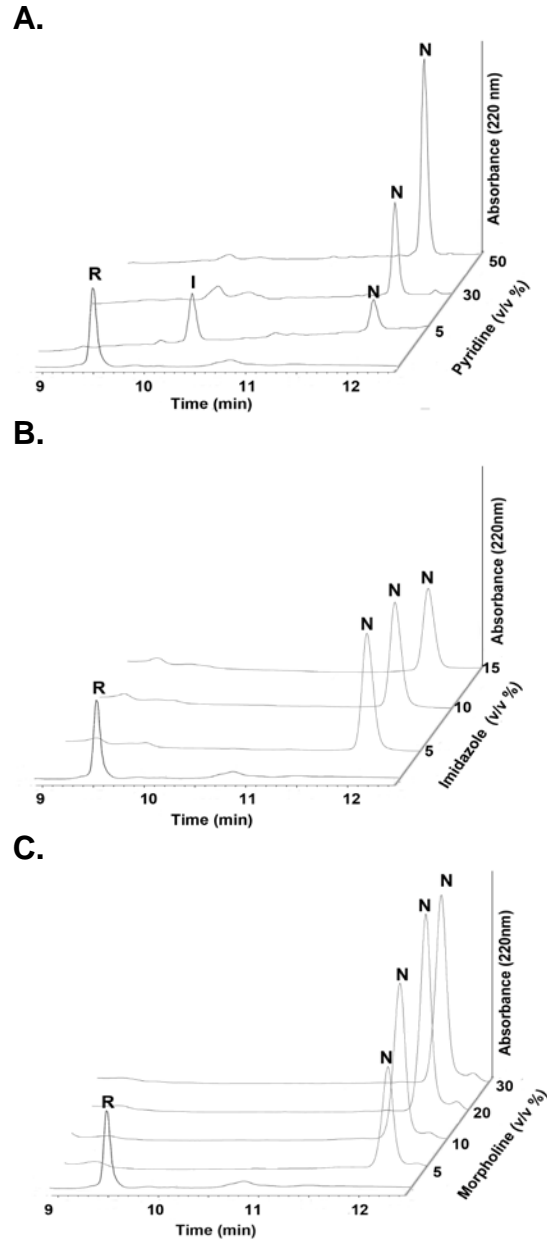


Figure 4.4 RP-UPLC profile of oxidative folding of kB1 showing the effect of different bases. **A.** Conditions contained different concentrations of pyridine from 5-50%; **B.** Conditions contained different concentrations of imidazole from 5-15%; **C.** Conditions contained different concentrations of morpholine from 5-30%. All conditions contained 100 μ M peptide, 100 mM cysteamine and 10% (v/v) DMSO. The profiles show the reaction performed for 1 h. **R** and **N** represent the reduced and native form of kB1.

2.3 Redox effect

In our experimental conditions, cysteamine served as a reductant and DMSO as an oxidant. To investigate the role of these redox agents for the non-aqueous oxidative folding conditions, conditions without cysteamine and DMSO were examined (**Figure 4.5**). Without the reductant cysteamine, an incorrect isomer was preferentially populated indicating that the oxidative folding involved a kinetic trapped product which required reshuffle to form the correctly folded product. In the absence of DMSO, it was found that 70% reduced peptide remains intact after 1 h oxidative folding. This result suggested that DMSO was an effective reagent for the disulfide formation while air oxidation usually required a long duration for oxidative folding of CRPs. Replacing cysteamine/DMSO with the conventional GSH/GSSG in the proposed organic solution led to poor yields due to the poor solubility of GSH/GSSG in organic solvents.

We also examined different thiols as reducing reagents. 2-mercaptoethanol and mercaptoacetic acid which contained a –OH or a –COOH group, respectively were compared with cysteamine which contained a –NH₂ group. Although they were able to mediate the thiol-disulfide exchange reaction, they were not as efficient as cysteamine (**Figure 4.6**). Mix-disulfides and reduced form kB1 were major products in conditions containing 2-mercaptoethanol and mercaptoacetic acid. 5% and 15% yields were obtained in the two conditions, respectively, whereas 90% yield was achieved in the cysteamine condition. This was probably because 2-mercaptoethanol and mercaptoacetic acid were too reducing thus could not build up a suitable redox potential with 10% DMSO.

The disulfide connectivity of the synthetic kB1 was confirmed by comparison with the native kB1 ¹H NMR spectra (**Figure 4.7**). It showed that the spectra of the synthetic sample and the native sample were identical, suggesting the same disulfide connectivity between the synthetic kB1 and the native kB1.

In summary, we have shown that the oxidative folding of kB1 in neat organic solvents was successful. With 2-propanol as a co-solvent, pyridine, imidazole or morpholine as a base and DMSO/cysteamine as redox reagents, the folding reaction can be completed in 1h with >80% yield.

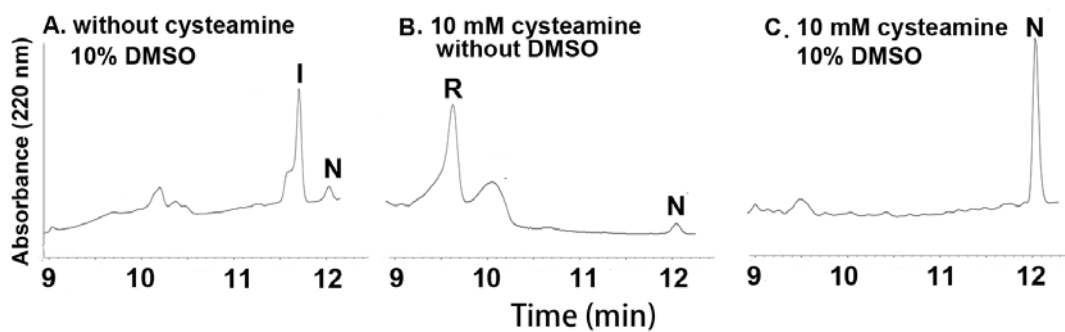


Figure 4.5 RP-UPLC profiles of oxidative folding of kB1 showing the effect of cysteamine and DMSO as redox agents. R and N represent the reduced and native forms of kB1 while I indicates the folding intermediate.

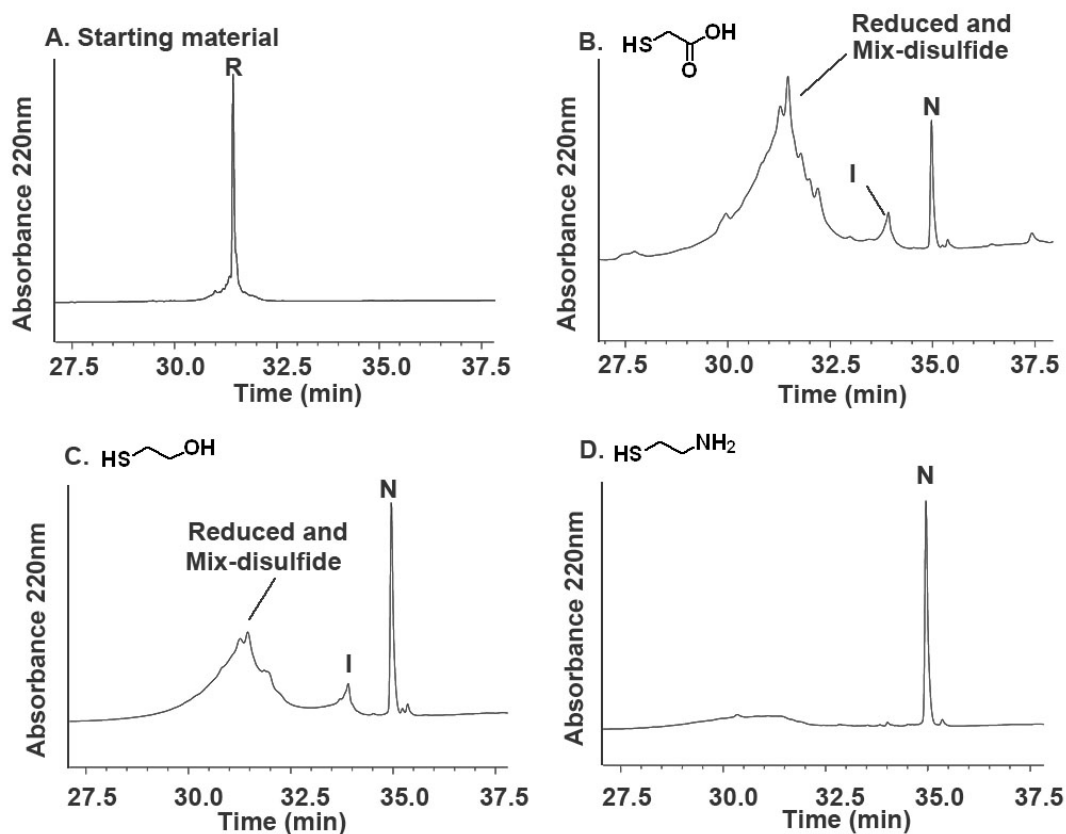


Figure 4.6 RP-HPLC profiles of oxidative folding of kB1 with different thiols as **reducing reagents**. **A**. Reduced kB1; **B**. Conditions containing 10 mM mercaptoacetic acid; **C**. Conditions containing 10 mM 2-mercaptoethanol; **D**. Conditions containing 10 mM cysteamine. **R** and **N** represent the reduced and native forms of kB1 while **I** indicates the folding intermediate. All conditions contained 100 μ M peptide, 100 mM thiol, 10% (v/v) DMSO, 85% 2-propanol and 5% morpholine.

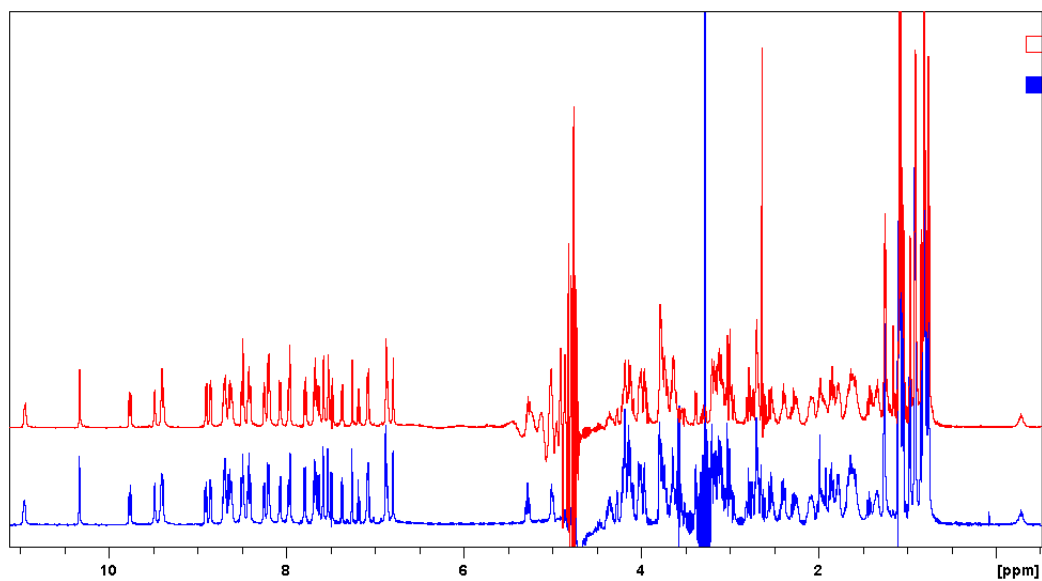


Figure 4.7 ^1H NMR spectra comparison of the natural kB1 (red) and the synthetic one (blue). Both peptides were dissolved in 90% $\text{H}_2\text{O}/10\%\text{D}_2\text{O}$ at pH 4.3. The spectra were obtained by a NMR spectrometer (600 MHz) at 298K.

3. Oxidative folding of ET-1 in organic solvents

Human endothelin-1 (ET-1) is known as a potent vasoconstrictor which modulates hormone levels and affects the nervous system. It composes of 21-amino acids including four Cys residues at positions 1, 3, 11, and 15. The native ET-1 contains a disulfide linkage of Cys1-Cys15 and Cys3-Cys11, which is named as the 1-4 isomer. A 1-3 isomer with a disulfide linkage of Cys1-Cys11 and Cys3-Cys15 is also obtained *in vitro* (**Figure 4.8**). In aqueous conditions, the oxidative folding of reduced ET-1 always resulted in a mixture of two products: the 1-4 isomer and the 1-3 isomer, at a 3:1 ratio. Kubo and co-workers increased the ratio of the 1-4 isomer over 1-3 isomer to 88:12 by extending the peptide N-terminus with a Lys-Arg dipeptide.²⁷⁴ They revealed that the dipeptide extension could form a salt bridge with the Asp⁸ residue stabilizing the C-terminal α -helix structure in native ET-1. Substitution of Asp to Asn disrupted the α -helix structure at the C-terminus and the improvement was abolished. Their results highly suggested that the C-terminal α -helix structure was important for forming the native disulfide bond pattern.

To examine whether our organic system could increase the ratio of native form ET-1, we refolded the fully reduced ET-1 with various concentrations of a structure enhancing solvent trifluoroethanol (TFE). TFE was known to promote the helix formation and stabilize the secondary structure. Thus, we hypothesized that the ratio of native form ET-1 would be improved in the organic folding system containing TFE as a co-solvent.

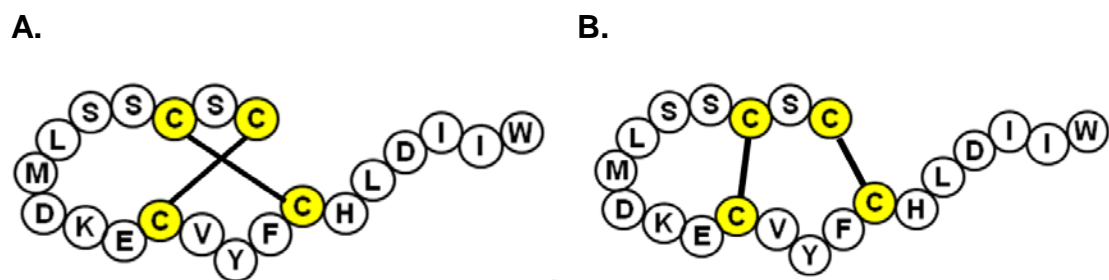


Figure 4.8 Schematic demonstrations of ET-1 isomers. **A.** 1-3 isomer; **B.** 1-4 isomer.

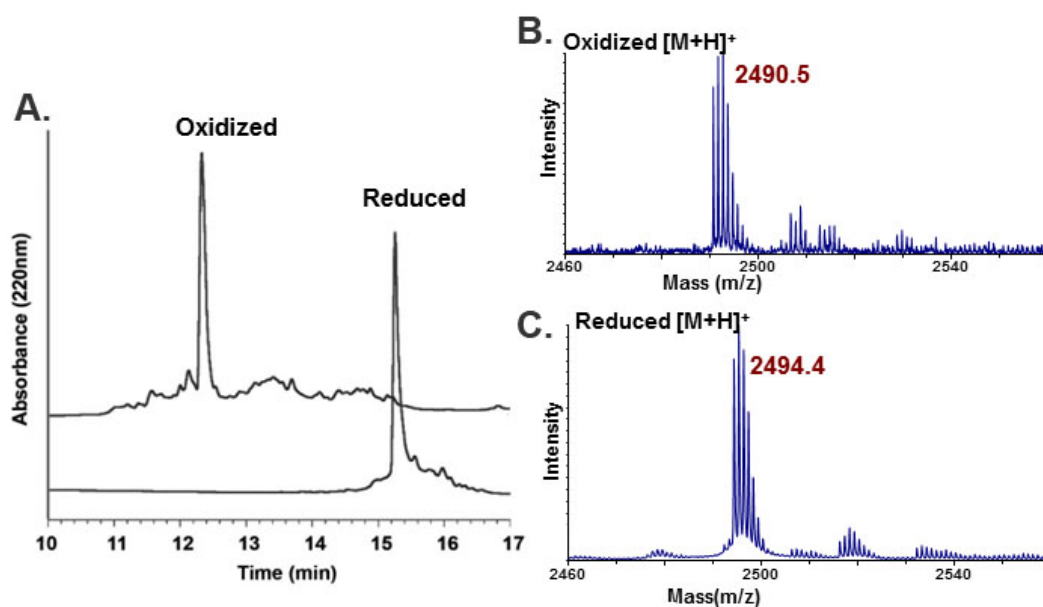


Figure 4.9 RP-HPLC and MALDI-TOF MS profiles of reduced and oxidized synthetic ET-1. **A.** HPLC profiles of reduced and oxidized ET-1 which eluted at the 15.3 min and 12.3 min, respectively; **B.** MS profile of oxidized ET-1 showing an m/z of 2490.5 while the theoretic molecular weight was 2491.2; **C.** MS profile of reduced ET-1 showing an m/z of 2494.4 while the theoretic molecular weight was 2495.3.

ET-1 was synthesized on Rink amide MBHA resin by Fmoc SPPS manually. After TFA cleavage and HPLC purification, we obtained the reduced form ET-1 in 11% yield (**Figure 4.9**). The reduced ET-1 was then subjected to organic folding conditions containing various concentrations of TFE. Based on the HPLC profile, the reduced ET-1 eluted later than the native ET-1 suggesting that the reduced form was more hydrophobic. The highly hydrophobic reduced ET-1 introduced problems for dissolving the peptide. Thus, in aqueous oxidative folding conditions, 8 M urea was often included to dissolve the peptide and high dilution was required.²⁹⁷ Under organic folding conditions, the solubility problem was successfully overcome with the assistance of DMSO. The peptide concentration could reach 100 μ M without aggregation.

Our results showed that the ratio of 1-4 isomer over 1-3 isomer was greatly improved from 3:1 to 13:1 by increasing the percentage of TFE (**Figure 4.10**). Although 2-propanol was effective in the oxidative folding of kB1, it did not promote the 1-4 isomer formation in oxidative folding of ET-1. A ratio of 3:1 was obtained similar to the aqueous conditions. We varied the TFE concentration from 0-85% with a 10% increment at each time. Conditions containing less than 50% TFE did not show any improvement in the ratio of 1-4 isomer. However, we found that the ratio of 1-4 to 1-3 isomers was highly influenced by high concentrations of TFE. In conditions containing 80 and 85% TFE, the ratio of 1-4 to 1-3 ET-1 isomer was increased to 12:1 and 13:1, respectively. These results proved that TFE stabilized the native form ET-1 and promoted its formation. The oxidative folding of ET-1 was greatly improved in organic solvent TFE compared with the aqueous conditions.

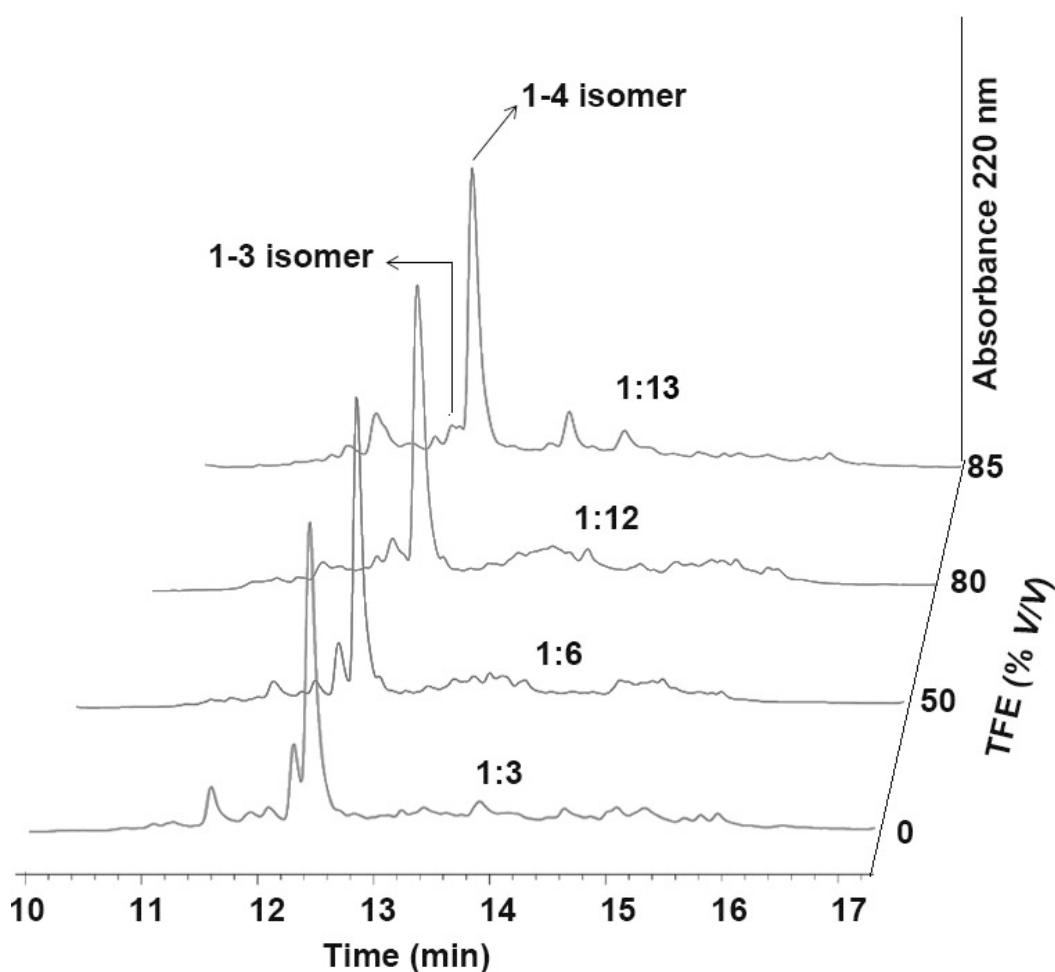


Figure 4.10 RP-HPLC profiles of ET-1 in different organic folding conditions containing 0-85% TFE. The ratio of 1-4 isomer over 1-3 isomer was increased from 3:1 to 13:1 with the increase of TFE concentrations. The folding condition contained 100 μ M peptide, 5% morpholine, 100 mM cysteamine, 10% DMSO and 0-85% TFE in 2-propanol for 30 min. The concentration of 2-propanol varies in the opposite way to that of TFE. Each reaction was monitored by RP-HPLC to quantify the ratio of 1-3 and 1-4 ET isomers.

4. Discussion

Biosynthesis of cyclic cysteine-rich peptides involves two key steps: formation of correct disulfide bonds and cyclization of the peptide backbone. For synthesis convenience, *in vitro* synthesis process is always the other way around by performing backbone cyclization before forming disulfide bonds. One disadvantage is that the cyclic backbone introduces structural constraints, which may result in a higher energy barrier between incorrectly folded peptide and the native one. Thus, it complicates the oxidative folding process and the folding yields are generally low. Recently, some CRPs have been folded successfully with the addition of organic co-solvents^{185,262,264,298} and it inspired us to perform the oxidative folding reaction in neat organic solvents which is a breakthrough in the oxidative folding methodology.

Our successful and highly efficient oxidative folding in neat organic solvents contradicts the conventional wisdom that the oxidative folding of peptides such as CRPs must be performed in aqueous solvents. The key to our success is the use of an organic solvent system which includes organic solvents (alcohols such as 2-propanol), organic bases (pyridine, imidazole or morpholine), and organic redox reagents (DMSO, cysteamine, and other thiols).

First, the reaction rate was greatly accelerated by exclusion of water in the oxidative folding condition. Disulfide exchange reaction is an S_N2 reaction which is affected by the solvation of nucleophiles. The rate of disulfide exchange reaction is reported to be 1000 times faster in polar aprotic solvents (DMSO) than polar protic solvents (water). The log of the rate constant increased linearly when increasing the proportion of DMSO in the water-DMSO mixture, suggesting that the solvation of thiolate by water is not significant and dependent solely on concentration of DMSO.²⁹⁵ Thus under the organic solvent condition, thiolate anion is a stronger nucleophile for the nucleophilic attack due to low degree of solvation. It was also proved in oxidative folding of kB1 using organic solvents; at least 48 fold improvements in the reaction rate were observed compared with aqueous conditions. Our organic folding condition afforded the correctly folded kB1 in less than 1 h while the aqueous condition required 48 h.

The disulfide exchange rate also depends on the basicity of the solution. By adding pyridine (pKa 5.3), imidazole (pKa 7.1) or morpholine (pKa 8.4) in the reaction, basicity of the condition can be adjusted. We observed an increase of the reaction rate with the increase of base concentration. The oxidative folding in 5% pyridine was slow which required 6 h for completion whereas a 1 h reaction was sufficient for completing the reaction in 50% pyridine. For imidazole and morpholine, although the yield did not increase much by including higher concentration of bases, the reaction rate increased greatly which required less than 20 min to be completed (**Figure 4.4**). In the one-pot synthesis of kB1 described in chapter 3, we also observed that the folding yield was as low as 5% after 1 h when no organic base was added. However, when we included 2% morpholine in the organic folding condition, the yield was increased to almost quantitative and the reaction was completed in 1 h (**Figure 3.35**). In practical use, due to undesired peptide degradation in prolonged treatments and the difficulty in removing organic bases during purification, we kept the base concentration as low as 5%.

Second, the peptide concentration was increased by minimizing aggregation problems in organic folding system. CRPs, especially those with hydrophobic patches, are prone to aggregation through hydrophobic interactions. Eisenberg and co-workers have demonstrated the aggregation could occur between short β -sheets.²⁹⁹ By analysing disease-related proteins at atomic level, they found that a 4–7-residue β -sheet segment was sufficient to form a fibril through hydrophobic interactions. Short β -sheet structures are commonly found in hydrophobic CRPs, which will lead to extensive aggregation and precipitation in aqueous buffers. Precipitation is also observed in hydrophilic CRPs, which is due to the intermolecular disulfide-crosslink-induced polymerization. Thus, aqueous conditions often required high dilution of peptides for formation of disulfide bonds. It greatly increases the reaction volume, which complicates the purification process and also results in a loss of peptide due to sticking extensively to the container surface. A peptide concentration of 10 μM is usually employed for aqueous folding whereas in our organic folding system, the peptide concentration can reach 100 μM or higher. This improvement was attributed to the use of DMSO. In our organic folding system, DMSO efficiently minimized aggregation due to β -sheet formation by preventing the hydrogen bonding. Reactions

without DMSO resulted in a low yield of correctly folded kB1 and precipitation after a prolonged treatment (**Figure 4.5C**). The polymerization via intermolecular disulfide crosslink was also avoided by using DMSO-cysteamine as a redox pair. The folding of kB1 and ET-1 in the organic folding system at a 100 μ M concentration proceeded smoothly without precipitation, indicating that this system is useful for oxidative folding of CRPs particularly those that are hydrophobic and aggregation-prone.

Third, the redox reagents catalyzed the formation of correctly folded peptide efficiently. Conventionally, the disulfide bond formation for oxidative folding of cysteinyl-containing peptides can be achieved through air oxidation with or without redox reagents. Some proteins such as ribonuclease A can be refolded efficiently in aqueous buffer to form the native disulfide connectivity preferentially. An amazing example is secretory leukocyte protease inhibitor (SLPI), which contains 8 intramolecular disulfide bonds.³⁰⁰ Theoretically, 46 million folding intermediates are possible during the oxidative folding process. Interestingly, the protein was able to form the native disulfide connectivity quantitatively by air oxidation without any redox reagents. It was proposed that the free cysteines in the protein sequence served as redox reagents for catalyzing the disulfide shuffling in non-native isomers. In contrast, simple air oxidation is generally not successful for most of CRPs containing two or three intramolecular disulfide bonds. Redox reagents are required to form native form peptides through the disulfide shuffling reaction.

Under our experimental conditions, DMSO/cysteamine combination served as efficient redox reagents. DMSO has been reported to be effective in disulfide bond formation as a mild oxidizing agent.²⁶⁰ It is able to mediate a faster and more efficient disulfide bond formation compared with air oxidation. Cysteamine, as a thiol catalyst, was also shown to be useful in the folding of cyclotide cycloviolacin O2.¹⁸⁵ In our results, the redox reagents were necessary for the accumulation of native form peptide. The yield significantly decreased in the absence of DMSO or cysteamine, suggesting that the dominant pathway for correct disulfide formation was mediated by thiol-disulfide exchange reactions (**Figure 4.5**).

We also examined other thiol catalysts with different functional groups. However, they were generally not suitable and prone to form mix-disulfide

intermediates instead of accumulating native form peptides. Reduced and oxidized glutathiones (GSH/GSSG) were shown to be crucial for the correct folding of many CRPs such as charybdotoxin, calciseptine and ω -conotoxin GS.³⁰¹⁻³⁰³ However, replacing DMSO/cysteamine by the conventional GSH/GSSG in our organic solution led to poor yields due to the poor solubility of GSH/GSSG in organic solvents. An advantage of the organic folding approach is that DMSO and cysteamine are considerably less expensive than GSH and GSSG.

Another reason for our success is attributed to the structure enhancement of the organic co-solvents. NMR structures have already shown that cyclotides like kB1 display an “inverse configuration” with the exposure of hydrophobic side chains forming a hydrophobic patch on the surface.³⁰⁴ Its hydrophobic character is also shown by the late retention time in RP-HPLC. It is likely that kB1 prefers to bear a native-like structure in hydrophobic organic solvent by inverting the hydrophobic side chains out. 2-propanol has been found to be effective in oxidative folding of other hydrophobic CRPs like α -conotoxin, hedyotide B1 and ω -conotoxin^{262,265,305,306} by preferentially stabilizing the more hydrophobic native form peptides. Thus it was employed in our conditions for oxidative folding of kB1. Our observation has shown that 2-propanol as the co-solvent increased the yield to >80%. 50-85% of 2-propanol could afford native form of kB1 with yields from 85-91%. The robustness provides opportunities for introducing other kinds of co-solvents when 2-propanol is not suitable. This is exemplified by the case of ET-1. By including high concentration of TFE, a α -helix inducing solvent, the ratio of native form ET-1 was greatly increased from 75% to 93%, suggesting that TFE was essential for formation of correct disulfide bonds in ET-1. Evidence from circular dichroism indicated that reduced ET-1 displayed a native-like α -helical structure in TFE instead of a random coil structure in aqueous buffer³⁰⁷, confirming that TFE facilitated the oxidative folding of ET-1 through promoting the native α -helical structure. Thus, organic co-solvents increased the fidelity of organic folding by preferentially stabilizing the native peptide structure.

Chapter 5 Butelase-mediated synthesis of cyclic peptides

1. Introduction

We have demonstrated the synthesis of cyclic peptides by chemical approaches in chapter 3 and 4. In chapter 5, I will describe the characterization of a novel cyclase which can be applied for cyclization. Chemoenzymatic methods using transpeptidases including intein, sortase, and subtiligase have been exploited by our laboratory and others for preparing peptide macrocycles.^{140-142,306} However, those transpeptidases are not naturally occurring cyclase and either the substrate or the enzyme has to be modified for cyclization. As described in chapter 1, the bioprocessing of a cyclotide from its precursor involves the joining of N- and C-termini by a cyclase, which is likely to be an asparaginyl endopeptidase (AEP). Based on screening of AEP, our laboratory has identified and isolated a novel transpeptidase butelase 1 from a cyclotide-producing plant *Clitoria ternatea*. Butelase 1 specifically recognizes Asx residue and its name comes from the local name for the plant bunga telang plus ligase.³⁰⁸

To investigate the requirement for butelase-mediated cyclization, model peptides were treated with butelase 1 to examine the cyclization efficiency. The characterization of the cyclotide genes from the Fabaceae family demonstrated a conserved propeptide sequence at the C-terminus. It involves highly conserved Asn, His, Val, Ile, Ala at P1 to P4' positions, respectively. Thus, five peptides kB1-NHVIA (GLPVCGETCVGGTCNTPGCTCSWPVCTR-NHVIA), kB1-NHVI, kB1-NHV, kB1-NH, kB1-N(amide) were prepared. They contained the same cyclotide sequence but varied in C-terminal propeptides. Treatment of kB1-NHV with butelase 1 produced >95% cyclic kB1 within 45 min. In contrast, model peptides with longer C-terminal propeptides (kB1-NHVIA and kB1-NHVI) displayed a decreased cyclization rate which required 2-3 h to complete. In addition, analogs with shorter C-terminal propeptides such as kB1-NH, kB1-N(amide) also displayed lower cyclization efficiency. Reactions yielded <10% cyclic kB1 after 4 h, and cyclization was not completed after 30 h. These kinetics studies of treating model peptides with butelase 1 revealed that a C-terminal tripeptide NHV was sufficient for butelase-mediated

cyclization. Butelase 1 successfully mediated the cleavage of a C-terminal HV sequence and ligation between the N- and C-terminus to form a cyclic kB1 (**Figure 5.1**).

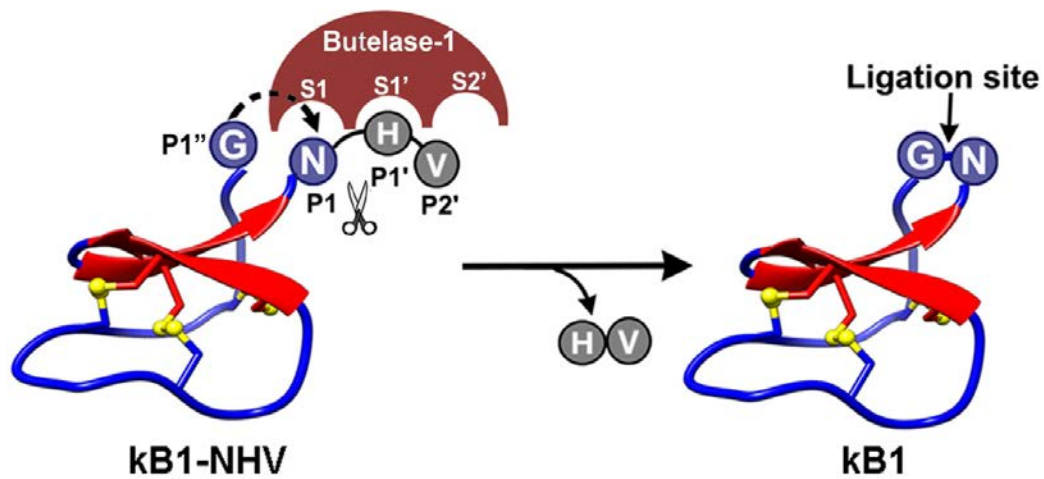


Figure 5.1 Butelase-mediated peptide cyclization. Butelase 1 (dark red) recognizes the C-terminal NHV propeptide in peptide kB1-NHV, cleaves at the Asn site and mediates the ligation between the N-terminal Gly and the C-terminal Asn to form cyclic kB1.

2. N-terminal substrate specificity in butelase-mediated ligation

To efficiently apply butelase 1 for peptide ligation and cyclization, a detailed understanding of substrate specificity is required. Although the C-terminal substrate specificity for butelase 1 is known at P1' and P2' positions, the N-terminal substrate specificity is not clear yet. Since the N-terminal substrate is the incoming nucleophile not the substrate cleaved by the enzyme, the corresponding site from N- to C-terminus is named as P1'', P2'', P3'' and so on.

2.1 Design and synthesis of peptide libraries

To investigate the optimal N-terminal substrate for butelase 1, three combinatorial libraries XXXGIR (X=any 20 natural amino acids) are synthesized. Each library contains a single position variant with all 20 natural amino acid at P1'', P2'' or P3'' position (**Table 5.1**). All 60 peptides were parallelly synthesized on 0.1 mmol Wang resin using Fmoc SPPS.

The design of the peptide libraries was based on previous research on cyclotide genes in Fabaceae family. Our laboratory revealed that in nature cyclotide substrates, hydrophobic amino acids especially branched amino acids like Ile, Leu and Val are highly conserved at the P2'' position while P1'' position favors a Gly residue.⁷⁸ Based on the putative specificity at P1'' and P2'' position, we synthesized the first library XIGGIR (X=any 20 natural amino acids) to investigate the substrate specificity at the P1'' position.

For the design of the second and third libraries, Leu was selected as the first and second amino acids (LXGGIR and LLXGIR) for two reasons. First, Leu was optimized for the P1'' and P2'' position based on our results. Second, the moderate hydrophobicity of Leu facilitated the analysis of the N-terminal peptide libraries as they were well separated from the C-terminal substrates and the ligated products.

For the C-terminal peptide, we used a model peptide KALVINHV-NH₂ and evaluated its ligation efficiency with XXXGIR. The design of the peptide sequence KALVINHV prevented self-cyclization due to steric hindrance of cyclic hexapeptides. The presence of NHV at the C-terminus allowed butelase 1 to recognize and mediate

an intermolecular transpeptidation with the library peptides. In addition, multiple Asn (N) residues in the model peptide were avoided to prevent undesired internal cleavage after prolonged reaction time. The presence of a positively-charge residue Lys (K) facilitated the purification and MS analysis of products.

After synthesis by Fmoc SPPS, these peptides were subsequently purified by RP-HPLC. The butelase-mediated ligation reaction was conducted using C-terminal peptide KALVINHV with different N-terminal peptides under optimized conditions for up to 2 h to investigate substrate specificity. The ligation yield for 10 min and 2 h were determined for each peptide of the libraries. In the ligation reaction profiles, ligation products were well separated from the starting materials, making the quantitation of the reaction possible. **Figure 5.2** showed the ligation reaction between substrate KALVINHV and peptide libraries XXXGIR with variations at P1'', P2'' or P3'' positions. Butelase 1 recognized the NHV motif in peptide KALVINHV and specifically cleaved at the Asn site to form an acyl-enzyme intermediate. Then the peptide XXXGIR would act as a nucleophile to resolve the acyl-enzyme intermediate and ligate with peptide KALVIN intermolecularly to form the ligation product KALVINXXXGIR.

Table 5.1 Peptide sequences in three combinatorial libraries.

| Library 1 | | Library 2 | | Library 3 | |
|-----------------------------------|--------|-----------------------------------|--------|-----------------------------------|--------|
| (H₂N-XIGGIR-OH) | | (H₂N-LXGGIR-OH) | | (H₂N-LLXGIR-OH) | |
| AIGGIR | MIGGIR | LAGGIR | LMGGIR | LLAGIR | LLMGIR |
| CIGGIR | NIGGIR | LCGGIR | LNGGIR | LLCGIR | LLNGIR |
| DIGGIR | PIGGIR | LDGGIR | LPGGIR | LLDGIR | LLPGIR |
| EIGGIR | QIGGIR | LEGGIR | LQGGIR | LLEGIR | LLQGIR |
| FIGGIR | RIGGIR | LFGGIR | LRGGIR | LLFGIR | LLRGIR |
| GIGGIR | SIGGIR | LGGGIR | LSGGIR | LLGGIR | LLSGIR |
| HIGGIR | TIGGIR | LHGGIR | LTGGIR | LLHGIR | LLTGIR |
| IIGGIR | VIGGIR | LIGGIR | LVGGIR | LLIGIR | LLVGIR |
| KIGGIR | WIGGIR | LKGGIR | LWGGIR | LLKGIR | LLWGIR |
| LIGGIR | YIGGIR | LLGGIR | LYGGIR | LLLGIR | LLYGIR |

2.2 Optimization of butelase-mediated ligation reaction

To investigate the optimized condition for the ligation reaction, five different peptide ratios were tested to determine the ligation efficiency (**Table 5.1**). Model peptides GIGGIR and KALVINHV were chosen as the N- and C-terminal substrate, respectively. The reaction was performed in a ligation buffer containing 1 mM EDTA and 5 mM β -mercaptoethanol at pH 6, 37 °C for 2 h in presence of varying ratios of butelase 1 and peptide substrates. After quenching the reaction with acid, HPLC and MS were performed to determine the ligation yield. We used 0.1 μ M butelase which was efficient for cyclization of kB1-NHV.

First, the butelase 1: KALVINHV ratio was kept constant at 1:500 while the N-terminal substrate GIGGIR was increased from 1000:10000. The ligation yields increased from 6.7% to 59.8% in a concentration-dependent manner. When the peptide KALVINHV was increased from 500:2000, the yield decreased significantly from 59.8% to 36.0%. Thus, the ratio of 1:500:10000 was shown to be optimal for the ligation reactions, achieving the highest yield of 59.8% at 2 h. For the following ligation reactions, the ratio of 1:500:10000 for butelase 1: KALVINHV: XXXGIR was applied. In addition, we also optimized the reaction by performing in a 10 mM phosphate buffer containing 1 mM EDTA at pH 6.5, 42 °C, which afforded higher yields and faster ligation rates.

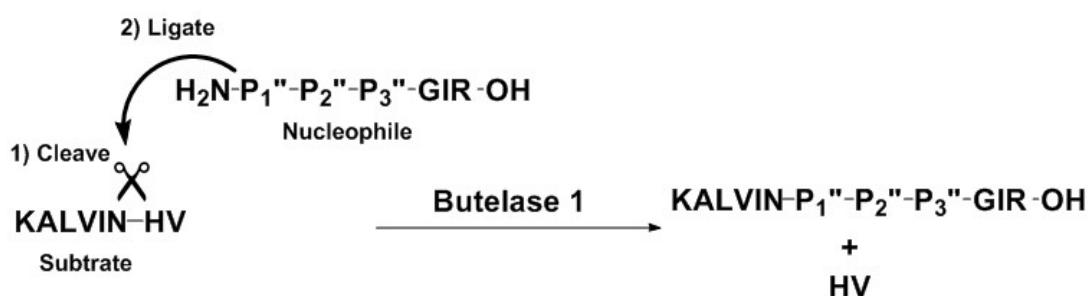


Figure 5.2 Scheme of butelase-mediated ligation of KALVINHV and XXXGIR. XXXGIR with variations at P1'', P2'' or P3'' was used as nucleophiles to ligate with peptide KALVINHV resulting the ligation product KALVINXXXGIR under the assistance of butelase 1.

Table 5.2 Conditions with different ratios of butelase 1: KALVINHV: GIGGIR.

| Butelase 1 | Ratio | | 2 h yield (%) |
|------------|--------------------------|-----------|---------------|
| | KALVINHV-NH ₂ | GIGGIR-OH | |
| 1 | 500 | 1000 | 6.7 |
| 1 | 500 | 2500 | 38.4 |
| 1 | 500 | 5000 | 51.3 |
| 1 | 500 | 10000 | 59.8 |
| 1 | 1000 | 10000 | 45.5 |
| 1 | 2000 | 10000 | 36.0 |

2.3 N-terminal substrate specificity at the P1'' position

To define the N-terminal substrate specificity at the P1'' position, we used KALVINHV as the C-terminal substrate and XIGGIR (X= any of the 20 natural amino acids) as the acceptor nucleophile, then evaluated their ligation efficiencies using HPLC. The reactions were performed at 42 °C, pH 6.5, using 0.1 μM butelase 1, 50 μM KALVINHV and 1 mM XIGGIR.

At the P1'' position, most of natural amino acids were accepted except for secondary amine Pro and acidic amino acids such as Asp and Glu (**Figure 5.3**). The intermolecular peptide ligations proceeded efficiently with the assistance of butelase 1. In most of ligation reactions, 60-80% yields were achieved within 10 min, together with 15-35% of starting materials and <5% hydrolyzed products. It should be pointed out that the ligation yields were higher than those in the optimization trials simply because the reaction was conducted in optimized conditions at higher pH and temperature.

2.4 N-terminal substrate specificity at the P2'' position

To define the N-terminal substrate specificity at the P2'' position, we synthesized the second library LXGGIR (X=any 20 natural amino acids) using the same approach.

Compared to the P1'' position, butelase 1 displayed a more strict requirement at the P2'' position. Hydrophobic amino acids, particularly Ile, Leu and Val were preferred over other amino acids (**Figure 5.4**). For Ile, Leu and Val, more than 80% of ligation yields were observed within 10 min while less than 30% yields were obtained for other amino acids. Peptides contained Cys at the P2'' position also afforded 72%. It was probably through a different mechanism which would be discussed later. It should be noted that at 2 h, there was a significant increase in ligation yields for Ala, Phe, His, Met, Asn, Gln, Ser, Thr, Trp and Tyr, especially for aromatic amino acids. In general, butelase 1 favored Leu, Ile, and Val at the P2'' position but also tolerated other amino acids to a less extent.

2.5 N-terminal substrate specificity at the P3'' position

To investigate whether the P3'' position shows strict requirements, we performed the intermolecular ligation of KALVINHV and LLXGIR (X=any 20 natural amino acids). The reactions were performed in the presence of 0.1 μ M butelase 1, 50 μ M KALVINHV, 1 mM LLXGIR as described before. Butelase 1 mediated efficient ligation reactions with all 20 peptides after 2 h. All 20 variants in the library achieved 60-90% yield except Ile with a slightly lower yield (37%) (**Figure 5.5**). No significant difference was observed. Butelase 1 displayed a broad N-terminal substrate specificity at the P3'' position. It was noteworthy that for some peptides the yield decreased at 2 h comparing with that at 10 min. This was because the butelase-mediated ligation was reversible and the recognition site was still capable of being cleaved after ligation. Collectively, butelase 1 displayed a broad specificity at the P3'' position and accepted all natural amino acids.

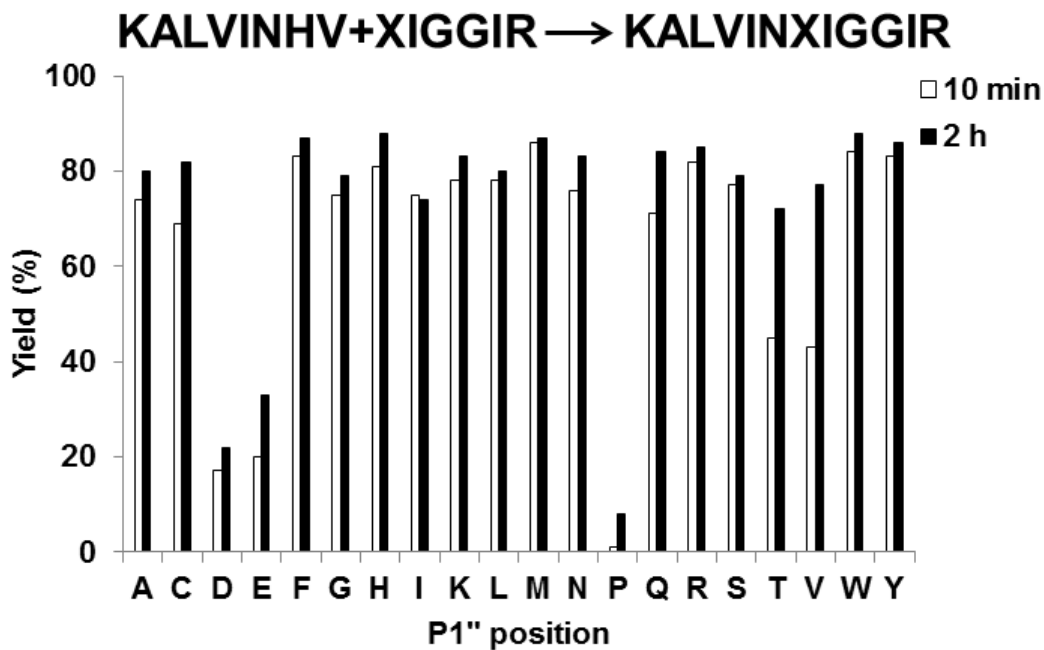


Figure 5.3 Ligation yields of KALVINHV and XIGGIR (X=20 natural amino acids). The yields were calculated based on triplicate experiments with less than 5% variation.

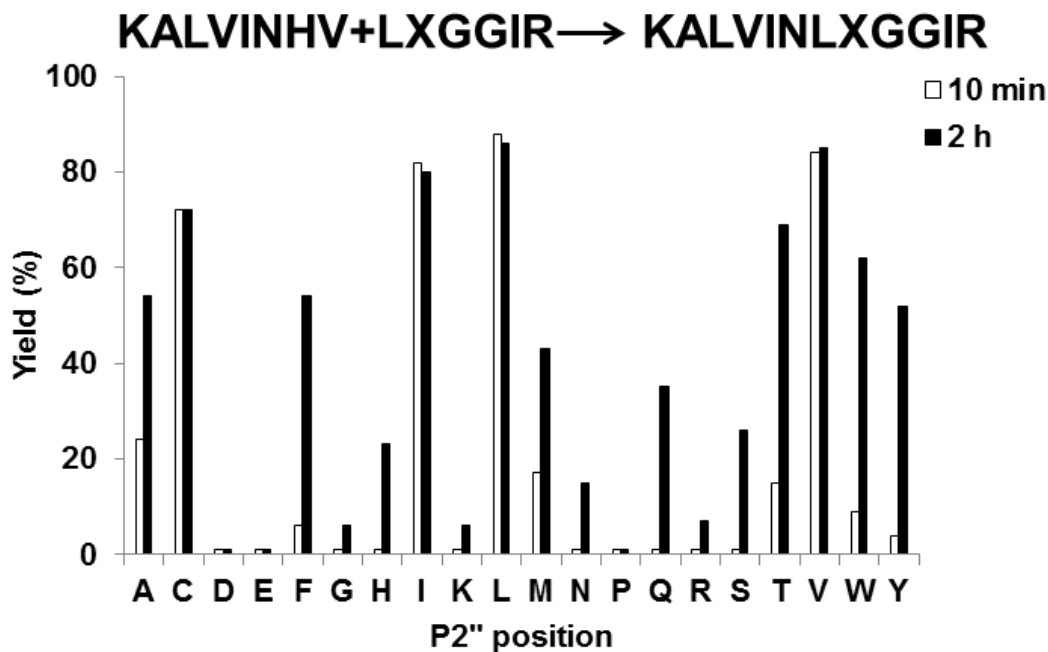


Figure 5.4 Ligation yields of KALVINHV and LXGGIR (X=20 natural amino acids). The yields were calculated based on triplicate experiments with less than 5% variation.



Figure 5.5 Ligation yields of KALVINHV and LLXGIR (X=20 natural amino acids). The yields were calculated based on triplicate experiments with less than 5% variation.

2.6 Unnatural amino acids at the P1'' and P2'' positions

We have demonstrated that butelase 1 was a promiscuous enzyme with broad substrate specificity. It was interesting to know whether unnatural amino acids were tolerated at P1'' and P2'' positions in the ligation reaction. We first replaced those amino acids which achieved high yields in the butelase-mediated intermolecular ligation to their D isomers. For the P1'' position, 5 unnatural amino acids including D-His, D-Lys, D-Arg, D-Phe, D-Trp were used to synthesize the model peptides XIGGIR. The ligation reaction was performed under the same condition described above. Interestingly, butelase 1 was able to recognize D-His, D-Lys and D-Arg at the P1'' position and facilitated the intermolecular ligation reaction although with less efficiency compared with their L-form counterpart. The ligation reaction between model peptides with D-His at the P1'' position achieved 46% yield after 2 h (**Figure 5.6**). In contrast, butelase 1 was unable to recognize D-Phe, D-Trp and D-Tyr at the P1'' position. Presumably, D isomers of basic residues are preferred at the P1'' position.

To investigate whether butelase 1 can distinguish the N-terminal α -amine and side chain ϵ -amine, a model peptide Ac-KIGGIR was synthesized with its N-terminal α -amine blocked by acetylation. With only side chain ϵ -amine free, the model peptide Ac-KIGGIR was not able to ligate with the model peptide KALVIN in presence of butelase 1. These results indicated that a free N-terminal α -amine was required for butelase 1 to perform the intermolecular ligation, highlighting the high specificity of butelase 1 that distinguishes the N-terminal α -amine from other side chain nucleophiles.

For the P2'' position, three residues (Ile, Leu and Val) are recognized by butelase 1 to facilitate the intermolecular ligation. Therefore, we synthesized peptides with D-Ile, D-Leu or D-Val at the P2'' position. Surprisingly, butelase 1 was able to recognize all three D- amino acids at the P2'' position. However, the efficiency decreased by 60% (**Figure 5.6**).

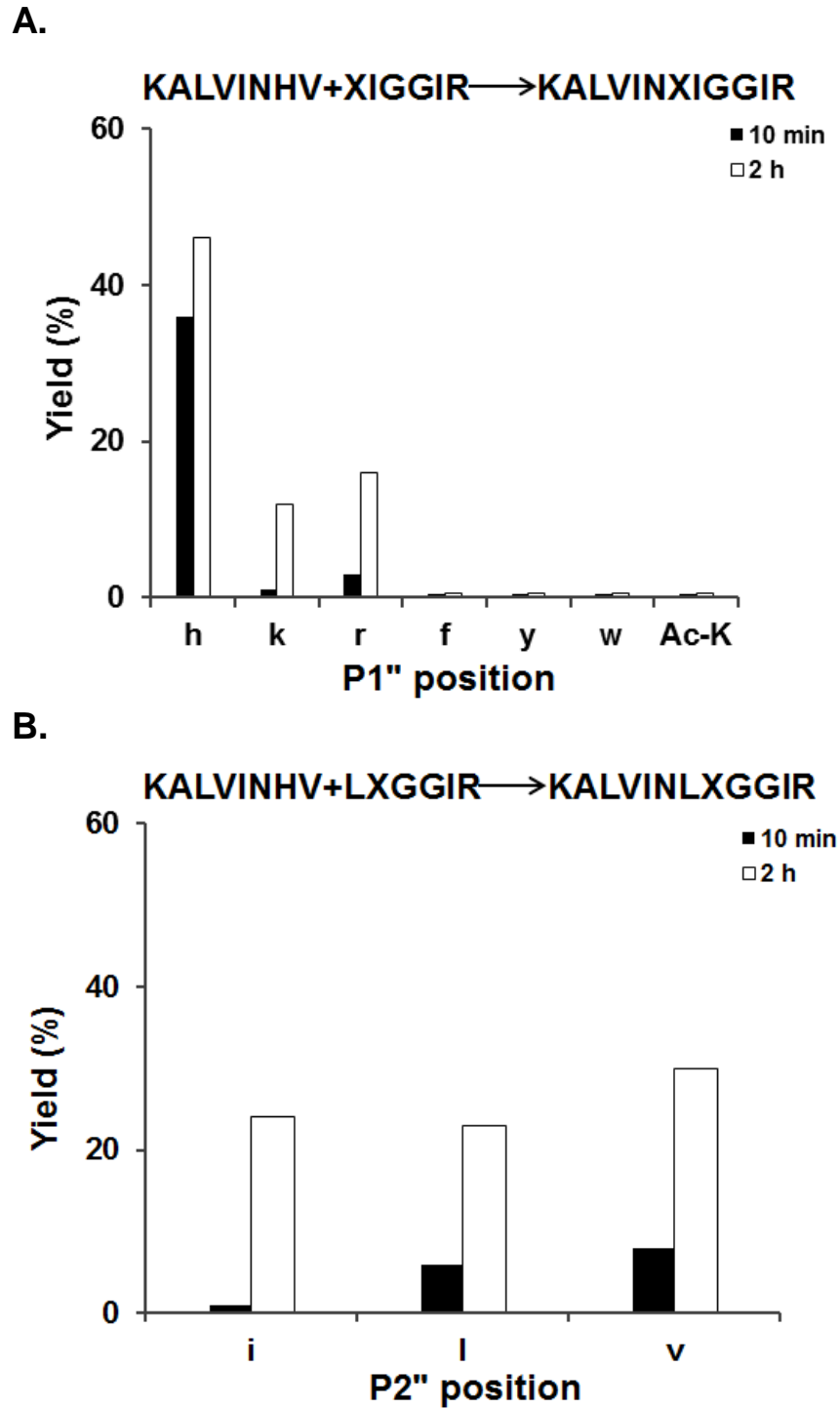


Figure 5.6 Butelase-mediated peptide ligation with unnatural amino acids at the **P1''** and **P2''** position. **A.** Ligation of KALVINHV and XIGGIR, X= D-His, D-Lys, D-Arg, D-Phe, D-Trp and α -amine-acetylated Lys; **B.** Ligation of KALVINHV and LXGGIR, X= D-Ile, D-Leu or D-Val.

3. C-terminal substrate specificity in butelase-mediated ligation

We have determined the C-terminal substrate specificity at P1' and P2' positions for butelase-mediated transpeptidation reactions. However, the requirement at P1 and P2 positions were not known yet. Therefore, we synthesized a C-terminal peptide library to investigate the substrate specificity.

3.1 Synthesis of C-terminal substrate library KALVXNHV

For the synthesis of C-terminal substrate library, we employed a combinatorial library approach to prepare the peptide substrates. The library KALVXNHV bearing a single positional variation at the P2 position was synthesized in parallel groups of 4 amino acids each (**Figure 5.7A**). We determined the group members based on the molecular weight and putative retention time of each peptide. Thus, each group contained four peptides with distinguished molecular weights. Since the coupling efficiency of amino acids varied from one to another, we adjusted the concentrations of amino acids according to the molar ratio described in an earlier work by Dooley and Houghten.³⁰⁹ (**Figure 5.7B**) It resulted in almost equal amount of each peptide in every group as illustrated in **Figure 5.8**.

A.

| KALVXNHV | MW | Group 1 | Group 2 | Group 3 | Group 4 | Group 5 |
|----------|-----|---------|---------|---------|---------|---------|
| K | 907 | 907 | | | | |
| H | 916 | | 916 | | | |
| R | 935 | | | 935 | | |
| N | 893 | | | | 893 | |
| Q | 907 | | | | | 907 |
| S | 866 | 866 | | | | |
| G | 836 | | 836 | | | |
| D | 894 | | | 894 | | |
| E | 908 | | | | 908 | |
| A | 850 | | | | | 850 |
| T | 880 | 880 | | | | |
| C | 882 | | 882 | | | |
| P | 876 | | | 876 | | |
| V | 878 | | | | 878 | |
| Y | 942 | | | | | 942 |
| M | 910 | 910 | | | | |
| I | 892 | | 892 | | | |
| L | 892 | | | 892 | | |
| W | 965 | | | | 965 | |
| F | 926 | | | | | 926 |

B.

| Amino acid | Molar ratio |
|--------------------|----------------|
| Single letter code | Fmoc synthesis |
| A | 0.75 |
| D | 1.20 |
| E | 1.00 |
| F | 0.60 |
| G | 0.70 |
| H | 0.78 |
| I | 2.29 |
| K | 1.00 |
| L | 0.77 |
| M | 0.71 |
| N | 1.20 |
| P | 0.86 |
| Q | 1.00 |
| R | 1.10 |
| S | 0.72 |
| T | 0.98 |
| V | 1.80 |
| W | 0.70 |
| Y | 0.71 |
| | <hr/> 18.87 |

Figure 5.7 Molecular weights and molar ratio of amino acids in each batch. **A.** Each group contains four peptides with distinguished molecular weights. The single letter represents the amino acid variant X in the peptide library KALVXNHV. **B.** The molar ratio of each amino acid in the coupling reaction.

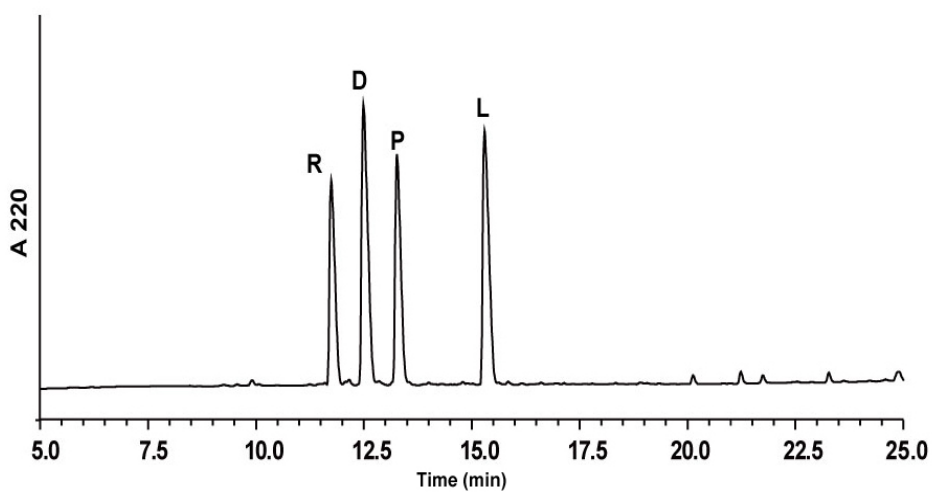


Figure 5.8 RP-HPLC profile of C-terminal peptide library KALVXNHV. The profile showed the TFA cleavage crude of group 3 which contained 4 peptides KALVXNHV with Arg, Asp, Pro or Leu at the X position.

3.2 C-terminal substrate specificity at the P1 and P2 positions

To define the C-terminal substrate specificity of butelase 1 at the P1 position, we performed the ligation reaction between KALVIXHV (X=N, n or D) and an acceptor peptide HIGGIR. The ligation reaction was carried out using the same condition as described above.

As illustrated in **Figure 5.9**, butelase 1 specifically recognized Asn at the P1 position in the ligation reaction with 83% yield within 10 min. Asp was also tolerated but with a lower efficiency as only 6% ligation yield could be obtained after 2 h. The D-form Asn, which was indicated as n in the figure, was not accepted by butelase 1 completely. No ligation product was observed after 2 h. A prolonged treatment to 4 h did not result in >1% ligation yield. These results demonstrated that butelase 1 was highly specific for L-form Asx amino acid at the P1 position.

The ligation reactions for the second library KALVXNHV bearing different amino acids at the P2 position were carried out using KALVXNHV and HIGGIR as the C-terminal and N-terminal substrate, respectively.

At the P2 position, butelase 1 displayed a broad tolerance for all 20 amino acids (**Figure 5.10**). 60-90% ligation yields were obtained within 10 min. No significant difference was observed among all 20 amino acids. Thus, there was a broad selectivity at the P2 position in butelase-mediated ligation reactions.

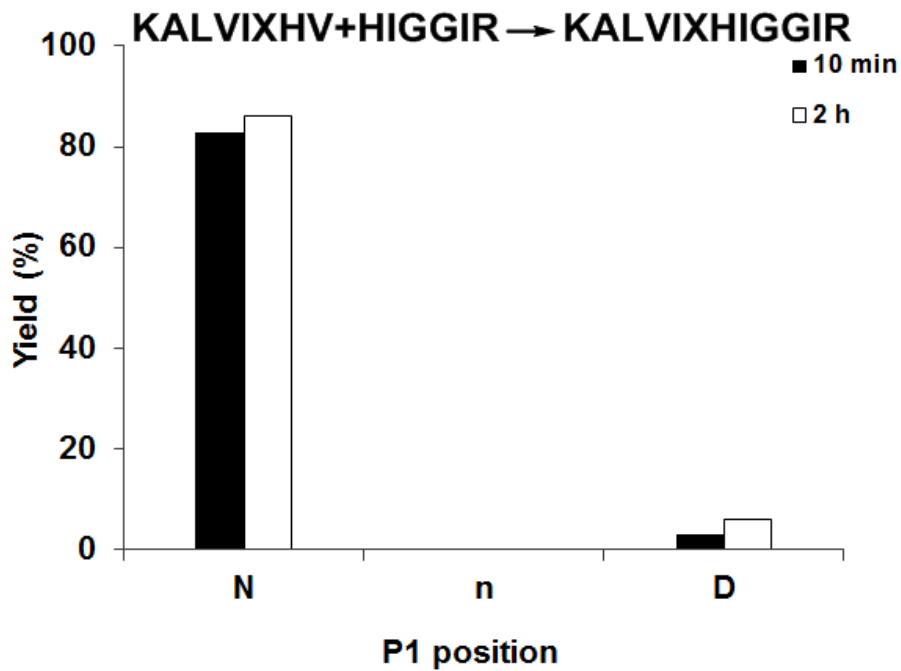


Figure 5.9 Butelase-mediated intermolecular ligation of KALVIXHV and HIGGIR (X=Asn, Asp or D-form Asn).

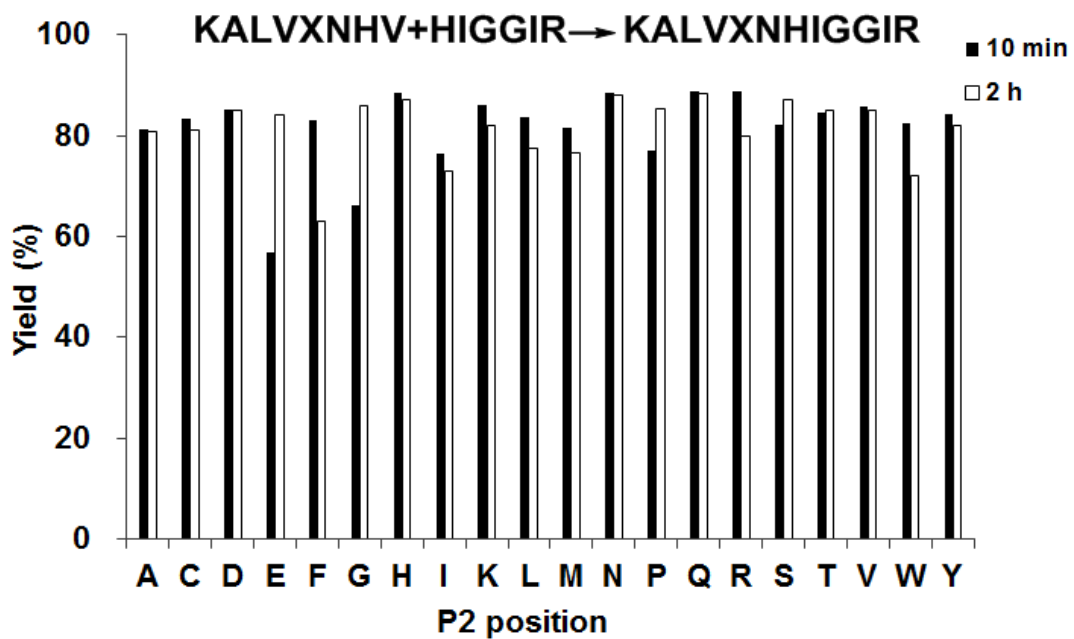


Figure 5.10 Butelase-mediated intermolecular ligation of KALVXNHV and HIGGIR (X=20 natural amino acids).

4. Discussion

As mentioned in chapter 1, natural peptide cyclases are rare as only three members have been isolated and characterized until now.¹³⁷⁻¹³⁹ Their low catalytic efficiencies limit their application as a tool for the synthesis of peptide macrocycles. The discovery of butelase 1, a novel cyclase from *Clitoria ternatea*, had sparked the potential of a new tool for ligation and cyclization of peptides. Our results on the substrate specificity in butelase-mediated ligation pave the way for applying butelase 1 for chemoenzymatic synthesis of cyclic peptides and site-specific labeling of proteins in protein engineering.

At the P1 position, butelase 1 highly preferred Asn or Asp to some extent. At the P1'' position, butelase 1 displayed a broad substrate selectivity for most amino acids however disfavored Asp, Glu and Pro. Almost all amino acids are tolerated except the secondary amine Pro and negatively charged Asp and Glu had shown negligible ligation reactions.

At the P2'' position, butelase 1 has shown stricter requirements where hydrophobic residues particularly those with branched side chains were preferred, such as Ile, Leu and Val. Other hydrophobic residues such as Met, Phe and Ala also gave moderate yields in the ligation reactions.

At the P3'' and P2 positions, butelase 1 displayed a broad substrate selectivity and high tolerance for all amino acid residues.

Our discovery on the substrate specificity in butelase-mediated intermolecular ligations was consistent with natural substrate sequences in the *Clitoria ternatea* with minor deviations (**Figure 5.11**). The stringent selectivity at the P1 and P2'' position corresponded well to the conserved residues at those positions in cyclotide sequences. In nature, the P1'' position preferred Gly whereas we had shown that it displayed a broad tolerance for amino acids except Pro, Asp and Glu. Nevertheless, it was explainable that Gly was preferred in nature since Gly was the most unhindered amino acid in the ligation reaction which has been demonstrated in the chemical synthesis. The P3'' position that tolerated a wide spectrum of amino acids, diverging

from its highly conserved residue proline was unexpected. We did not find a reasonable explanation yet. In fact, cyclization of peptides with disulfide constraints does not have to share the same rule as intermolecular ligation reactions. There should be higher constraints when ligating intramolecularly compared to intermolecularly. Therefore, the conserved GLP sequence in natural cyclotide substrate does not have to be essential for butelase-mediated ligation. Indeed, our results showed that Gly and Pro were not necessary at P1'' and P3'' positions. The highly conserved Asn at P1 position was confirmed by our *in vitro* study. Asn was favored over Asp with more than 100 fold efficiency. The results of P2 position were similar as that of the natural sequences. This position was highly variable without highly conserved residues.

What is interesting is that butelase 1 also tolerates D-amino acids at P1'' and P2'' positions. Butelase 1 was expected to display stereoselectivity against D amino-acids, as only L-amino acids are available in plants and D-amino acids normally should not be recognized by plant enzymes. However, our results had shown that butelase 1 did tolerate D-amino acids at both P1'' and P2'' positions. This could be explained by the possibility of horizontal gene transfer from prokaryotes to plants, which remained unverified. Gelvin and co-workers reported the infection of plant cells with *Agrobacterium* in 2000.³¹⁰ They found that a region of the Ti plasmid in the bacterium called the T-DNA was translocated into the genome of the plant cell, acquiring genetic advantages over time. This is an example showing how plant cells have a chance to encounter foreign gene materials. We supposed this may be one justification for why butelase 1 encounter D-amino acids. However, we do not have any solid evidence yet. Further structural studies will provide us more information on this unusual finding.

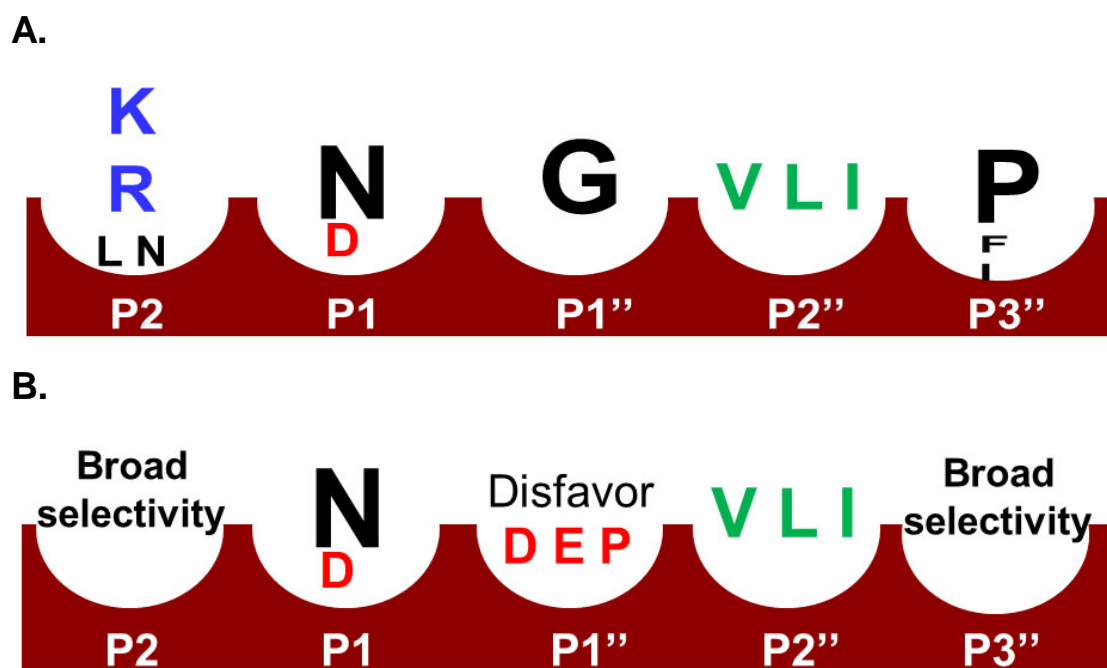


Figure 5.11 Comparison of natural sequence preference and *in vitro* substrate specificity at various positions. A. The preferred amino acids at various positions in natural cyclotide sequences from *Clitoria ternatea*. **B.** The substrate specificity at various positions in butelase-mediated intermolecular ligation reactions.

More than 50 novel cyclotides had been isolated or discovered at protein and gene levels in *Clitoria ternatea*. Their endogenous functions as plant defensive agents highly suggested that it was advantageous for the plant to have a cyclase with broad substrate tolerance to synthesize various cyclotides for its own defense. This is consistent with our findings that butelase 1 is a promiscuous enzyme with a broad substrate specificity at various positions. Except the highly conserved P1 position, at which butelase 1 favored Asn and did not accept D-amino acids, butelase 1 tolerated a wide range of amino acids at other positions. Even at the most strict P2'' position, other amino acids such as Ala, Met, Thr besides highly favored Ile, Leu and Val, also afforded more than 50% ligation yields after 2 h. The acceptor peptide LCGGIR from the P2'' library also obtained a relatively high ligation yield of 72 % at 10 min. It was probably due to an X-Cys ligation mechanism³¹¹, during which the thiol side chain of cysteine initiated a nucleophile attack on the acyl-enzyme intermediate to form an 8-member ring. Subsequently, an S-N acyl shift occurred spontaneously and formed a native peptide bond (**Figure 5.12**)

Chemoenzymatic synthesis approaches have been applied for preparation of macrocyclic peptides using sortase A and intein (**Figure 5.13A, B**). Sortase A, a versatile ligation tool for site-specific protein ligations, has been employed to cyclize histatin and kB1.^{118,160} However, its strict requirement of a C-terminal LPXTG signal motif and an N-terminal triglycine peptide would introduce extra sequences to the original peptide after cyclization, which may affect the overall conformation of the peptide. In addition, sortase-mediated cyclization reaction often requires high enzyme to peptide ratio (1:3) and a prolonged duration (usually overnight).

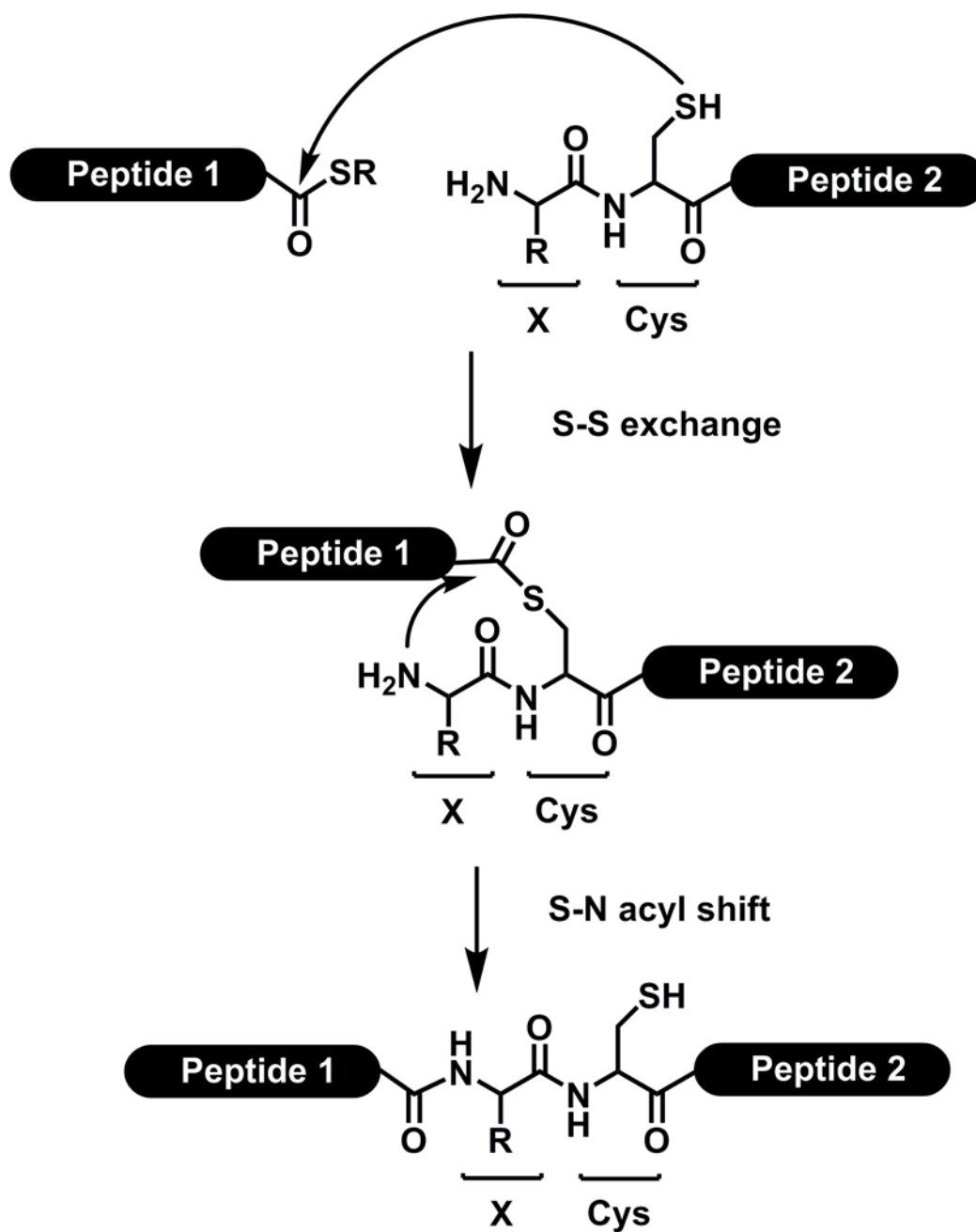


Figure 5.12 A scheme showing the mechanism of X-Cys ligation.

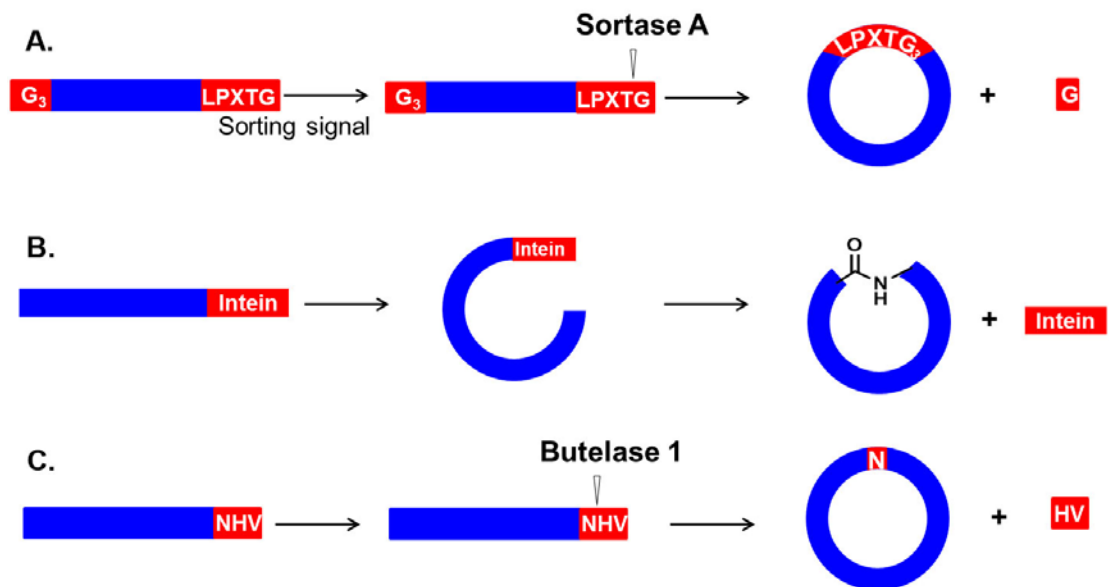


Figure 5.13 Comparison of three chemoenzymatic approaches for peptide cyclization. **A.** Sortase-mediated peptide cyclization required a C-terminal LPXTG sorting signal and an N-terminal triglycine as a nucleophile, resulting in a cyclic peptide with a LPXTGGG sequence. **B.** Intein-mediated peptide cyclization required the fusion of intein at the C-terminus of the target peptide. The cyclization reaction was achieved through acyl shift reactions with self-excision of intein. **C.** Butelase-mediated peptide cyclization required a C-terminal NHV propeptide, resulting in a cyclic peptide with a single Asn residue inside.

Intein-based peptide ligation and cyclization has been widely used for preparing cyclic peptides and their combinatorial libraries.^{151,153,154,312} As an N-terminal specific enzyme, it does not require any C-terminal recognition peptides. The cyclization reaction is achieved through acyl shift reactions like the chemical approach. To achieve the cyclization reaction, the target protein needs to be expressed as a fusion protein with the intein domain. This can be tedious and unproductive for cyclizing peptides since cyclic peptides are normally around 3 kDa whereas the intein domain is >30 kDa.

Butelase 1, which is capable of intermolecular transpeptidation of 14 to 58 residues and cyclize peptides within 1 h, proves to be more efficient.³⁰⁸ It requires a short recognition peptide NHV and produces cyclic peptides with a minimal introduction of extra sequences (**Figure 5.13C**). In addition, as we have demonstrated, its broad scope of substrate specificity allows cyclization of various cyclic peptides with high sequence diversity in nature. Unlike sortase A which strictly requires polyglycines as the N-terminal nucleophile, butelase 1 accepts a broad spectrum of amino acids as the nucleophile for transpeptidation reaction and introduces a single residue to the target peptide after cyclization. Its high catalytic efficiency and minimal recognition requirement for sequence will broaden its application as an alternative approach for peptide cyclization.

It is noteworthy that internal cleavage may occur when Asn is present in the sequence. We observed this phenomenon in our model peptide LLNGIR which contained an internal Asn residue and followed by Gly and Ile. This peptide was also recognized by butelase 1 and cleaved after the Asn residue. The ligation reaction of LLNGIR and KALVINHV resulted in more than one ligation product including KALVINLLN, LLNKALVINHV, LLNLLNGIR, LLNLLNLLNGIR besides the desired product KALVINLLNGIR. This result confirmed our findings that hydrophobic residues such as Ile in this case at the P2' position was important for the butelase recognition, For the P1' position, a wide range of amino acids was tolerated such as Gly in this case. Thus, peptides bearing a NXI/NXV/NXL sequence motif would be prone to cleavage by butelase 1. In addition, LLNKALVINHV was also observed as one of the products although in a small amount. In this case, KALVINHV acted as the acceptor nucleophile. The P2'' position was Ala instead of highly

preferred Ile, Val or Leu. It was consistent with our previous results that Ala in the P2'' position was also tolerated and obtained more than 50% ligation yield after 2 h. These results proved that butelase 1 accepted a wide range of substrate and thus internal Asn residues should be avoided for application of the butelase-mediated transpeptidation.

Our results on butelase-mediated ligations also suggested its potential as a tool for protein engineering, especially in its application for N- and C-terminal specific ligation of peptides, proteins or cells. With its high efficiency, minimal sequence requirements and broad substrate specificity, butelase 1 can find wide applications in chemoenzymatic ligation reactions such as bioconjugation of biomolecules, cyclization of functional proteins and labeling cell surface proteins. These have been achieved by sortase A by modifying peptide or proteins with an N-terminal oligoglycine or a C-terminal LPXTG sequence.^{158,313,314} Based on our substrate specificity, we could also engineer peptide or proteins with an N-terminal Xaa-Ile or a C-terminal NHV peptide to fulfill the same goal. Collectively, butelase 1 will be a versatile tool with wide scope and application for peptide and protein engineering.

Chapter 6 Design and engineering of cysteine-rich peptides

1. Introduction

Peptide therapeutics are emerging as a new class of drug candidates besides the small molecules (<1 kDa) and protein drugs (>10 kDa). The chemical space between 1-10 kDa contains valuable potential therapeutics combining the best of both worlds: the stability of small molecule and the specificity of protein biologics. This class of compounds are relatively stable for oral consumptions, like small molecule compounds, and highly specific for intramolecular targets, like protein drugs. Despite that, peptide derivatives are generally less toxic and free of auto immune issues. Their unique advantages highly suggest their potential applications in drug development.

Developing peptide therapeutics is both promising and challenging. Their applications in pharmaceutical area have been limited by their poor membrane permeability, susceptibility to enzyme digestions, rapid clearance *in vivo* and relatively high costs of production. To improve their deficiencies in pharmaceutical properties, conventional chemical modifications, including introducing non-natural amino acids, modifying amide bond with other functional groups such as N-methyl amide and developing peptide mimetics, are available but with limited efficiencies.³¹⁵ Compelling evidences suggest that the opportunity may exist in natural bioactive peptides with intrinsic favorable pharmaceutical properties. Through engineering of these peptides, we can achieve bioactive peptide-based drug leads with improved stabilities.

A method to increase the peptide stability is the grafting strategy by introducing a bioactive peptide into an intrinsically stable peptide scaffold. Cysteine-rich peptides are natural peptide scaffolds found in various organisms. They contain intramolecular disulfide bonds and have compact structures that confer them with chemical and enzymatic stability. These properties make them ideal for peptide grafting. Our laboratory has successfully demonstrated the strategy by grafting bradykinin B1 receptor antagonist DALK into a kalata B1 scaffold which was resistant to enzymatic and thermal degradation. This engineered peptide showed improved stability compared to the DALK peptide and served as a more potent bradykinin B1 receptor antagonist in oral administration.

Bradykinin is a positively charged nonapeptide, and an important mediator of a wide range of biological phenomena such as pain, inflammation, and vasodilation. Therefore, bradykinin receptor antagonists could be useful in pain treatment. Des-Arg¹⁰-[Leu⁹]-kallidin (DALK) is a prototype peptide antagonist which is known as one of the most potent bradykinin receptor antagonists. Here, we report the development of two bradykinin B1 receptor antagonists, SFBK and BKSF, by grafting the DALK peptide into a SFTI-1 scaffold.

Isolated from sunflower seeds, SFTI-1 is a cyclic peptide comprising 14 amino acids and is known as one of the smallest and most potent trypsin inhibitors. The cyclic backbone is divided into two loops by the disulfide bridge formed by the two cysteine residues, namely the trypsin inhibition (TI) loop and the secondary loop.³¹⁶ Being a cyclic peptide, SFTI-1 is resistant to proteolysis by exopeptidases and the ability to inhibit trypsin enhances its bioavailability and makes it an attractive scaffold for peptide grafting. Furthermore, possessing trypsin inhibition and bradykinin antagonist activity, the engineered peptide SFBK was expected to be bifunctional.

2. Design and synthesis of engineered peptides

2.1 Design of engineered bradykinin receptor antagonists

Bradykinin B1 receptor antagonist des-Arg¹⁰-[Leu⁹]-kallidin (DALK), a nonapeptide derived from bradykinin, was grafted into the SFTI-1 peptide sequence as shown in **Figure 6.1**. The secondary loop of SFTI-1 was substituted by DALK peptide to give a cyclic analog SFBK while the TI loop was substituted to give BKSF. Additionally, to compare the cyclic analogs, we also designed two linear analogs the BKSF-linear and SFBK-linear, abbreviated as BKSF-L and SFBK-L, respectively.

2.2 Chemical synthesis of engineered peptides by Fmoc SPPS

We have synthesized five analogs including three cyclic peptides and two linear ones as shown in **Table 6.1**. A linear DALK peptide was also synthesized as control. In all five analogs, the disulfide bond between the two cysteine residues was retained due to stability concern. Indeed, it was suggested that removal of the disulfide bond in SFTI-1 would lead to loss of its internal hydrogen bonding network and secondary structure hence, reducing its proteolytic stability.³¹⁶

Fmoc SPPS was employed to synthesize all peptide analogs. For the cyclic analogs, we used approaches described earlier in chapter 3. The peptide sequence was assembled on the resin together with a thioester surrogate including MeCys, TEBA or hydrazide. Then, the cyclic backbone was formed through transthioesterifications. Reduced cyclic peptides were subsequently subjected to DMSO or iodine oxidation to form the disulfide bond. Standard Fmoc SPPS was used to synthesize the linear analogs. The unprotected peptides were afforded after TFA cleavage and subjected to iodine oxidation to form the disulfide bond.

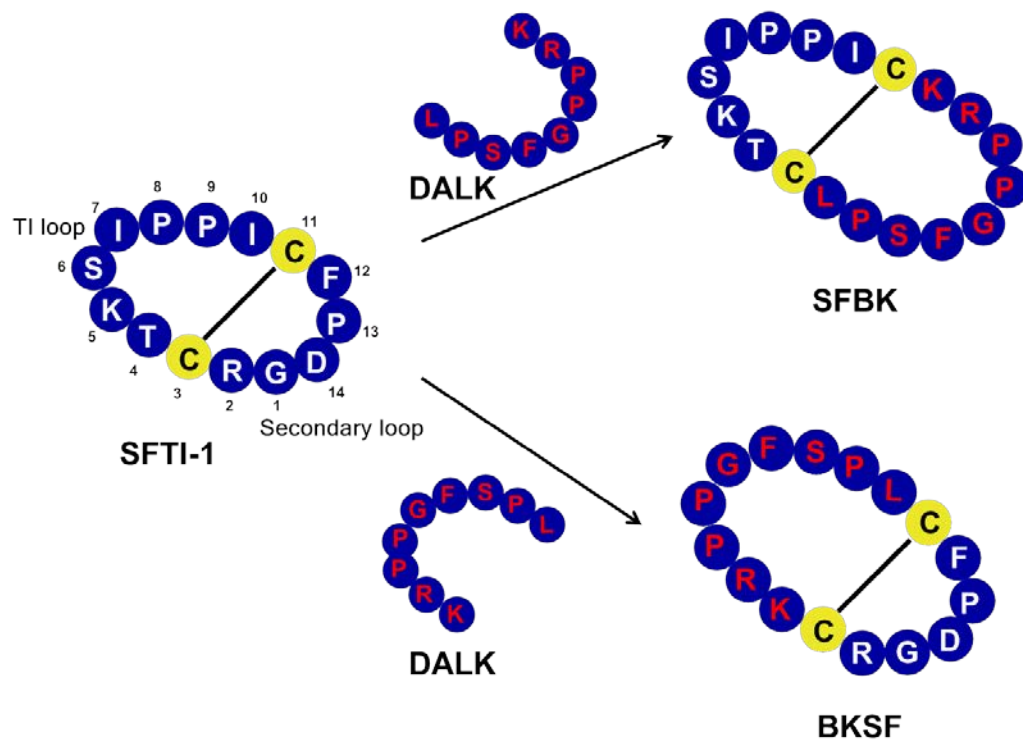


Figure 6.1 A scheme of SFTI-1 and its DALK grafted analogues. The DALK peptide replaced either the secondary loop or the trypsin inhibiting loop of SFTI-1 to produce engineered peptides SFBK and BKSF, respectively.

Table 6.1 Sequence alignment of SFTI-1 and its DALK grafted analogs.

| Peptide | Sequence | Total residues | Charge (+/-) | Mass | No. of Proline |
|----------------|--------------------------|----------------|--------------|--------|----------------|
| Cyclic analogs | | | | | |
| SFTI-1 | C-FPDGR-----C-TKSIPPI | 14 | 2/1 | 1513.7 | 3 |
| SFBK | C-KRPPGFSPL-C-TKSIPPI | 18 | 3/0 | 1921.0 | 5 |
| BKSF | C-FPDGR-----C-KRPPGFSPL | 16 | 3/1 | 1756.8 | 4 |
| Linear analogs | | | | | |
| DALK | KRPPGFSPL | 9 | 2/0 | 998.6 | 3 |
| SFBK-L | --C-TKSIPPI--C-KRPPGFSPL | 18 | 3/0 | 1939.0 | 5 |
| BKSF-L | GRCKRPPGFSPLCFPD | 16 | 3/1 | 1774.8 | 4 |

3. Characterizations of engineered peptides

To prove our hypothesis that the engineered peptide SFBK would obtain improved stability and act as a bifunctional molecule, a series of functional assays were performed on the synthetic analogs to investigate their stability and activities. *In vitro* assays that mimic the GI tract environment such as low pH stability, pepsin stability and trypsin stability assays were performed and monitored by RP-UPLC analysis at various intervals. Trypsin inhibition assay and bradykinin antagonism assay were carried out to investigate whether the engineered molecule possess the bifunctional activity.

In all assays, six peptides including DALK, BKSF, BKSF-L, SFBK, SFBK-L, and SFTI-1 were treated under the same experimental conditions to examine their effects. SFBK contained both the bradykinin receptor antagonist and trypsin inhibition loop while BKSF contained only the bradykinin receptor antagonist but not the trypsin inhibition loop. Their linear forms SFBK-L and BKSF-L were used to investigate the role of cyclic backbones. DALK and SFTI-1 were used as positive control for bradykinin receptor antagonism and trypsin inhibition, respectively.

3.1 Heat stability assay

All peptides were subjected to heat treatment to test their heat stability as proteins are generally denatured easily by high temperature. In the heating stability assay, peptides were dissolved in deoxygenated water and subjected to 100 °C treatment for up to 6 h. We monitored the process by checking the integrity of peptides using analytical RP-UPLC and MALDI-TOF MS at various time intervals (**Figure 6.2A**). The engineered peptide SFBK and BKSF were stable to 100 °C treatment remaining >90% intact after 6 h, which indicated that both of them were resistant to heat treatment. Other peptides also showed similar stabilities which withstood the heat treatment for up to 6 h. These results suggested that peptides tested here were heat stable. Macrocyclization was not crucial for heat stability.

3.2 pH stability assay

We tested the stability of peptides against acid hydrolysis in presence of 0.2 M HCl at 37°C which simulated the acidic environment of stomach. The hydrolysis

products were monitored by analytical RP-UPLC and MALDI-TOF MS. All peptides were stable which retained over 80% peptide integrities after 6 h. Both cyclic and linear peptides were generally stable. They remained intact after 6 h. These results suggested peptides tested here were stable under acidic denaturing.

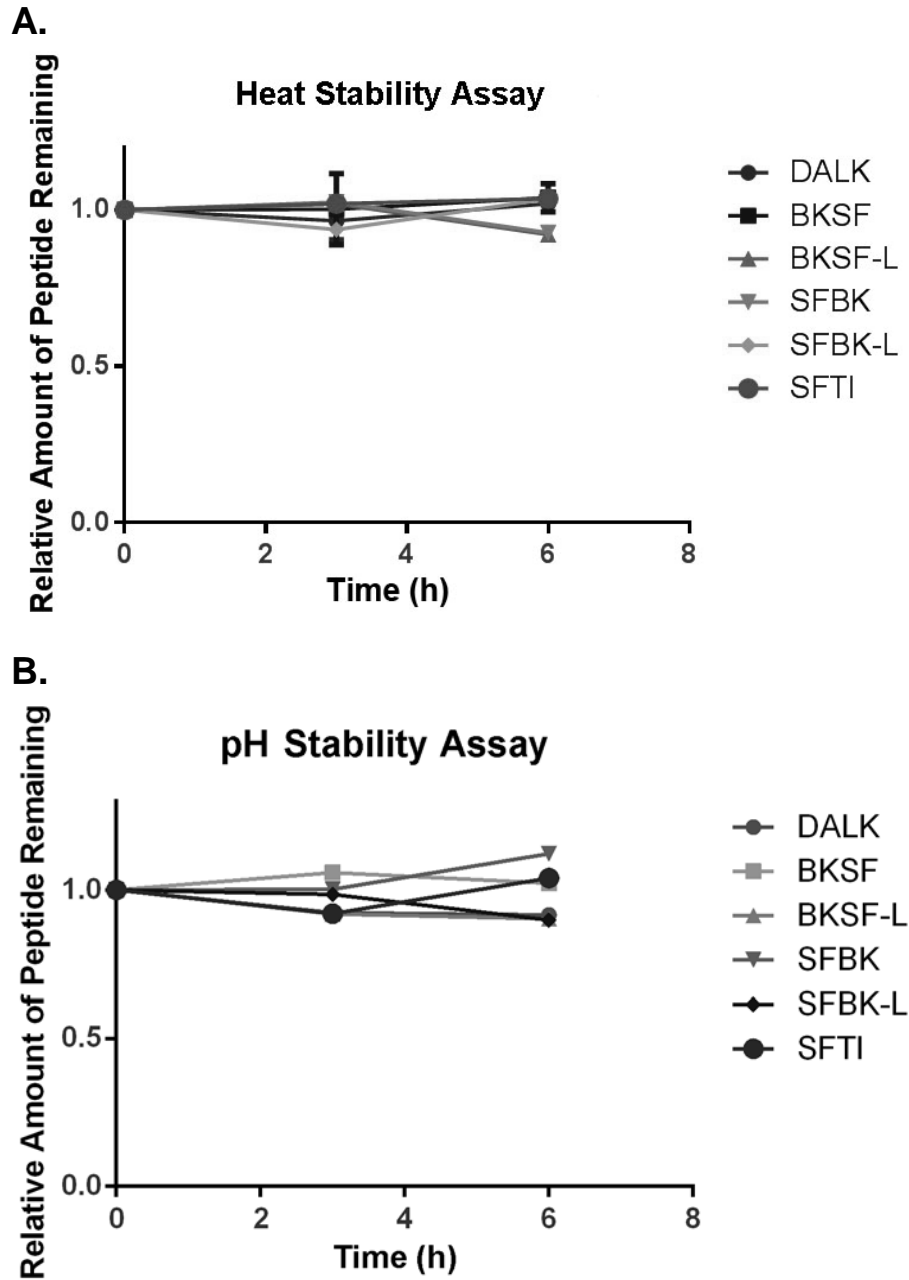


Figure 6.2 Heat and low pH stability of engineered peptides. The experiments were repeated in triplicates.

3.3 Pepsin stability assay

To evaluate the peptide stability against metabolic digestion, peptides were subjected to pepsin and trypsin digestion reactions at 37 °C. Digestion processes were monitored by analytical HPLC and MALDI-TOF MS.

Pepsin is the major enzyme in stomach and preferentially cleaves the peptide bond between a hydrophobic and an aromatic residue such as phenylalanine, tyrosine and tryptophan. DALK, SFBK and BKSF all contained a phenylalanine residue, thus, we are interested to know whether the compact structure could protect the peptide from pepsin digestion. In the pepsin stability assay, engineered peptides SFBK and BKSF exhibited resistance to pepsin degradation with >90% of the peptide intact after 6 h whereas linear peptide DALK was degraded rapidly under the same reaction condition with more than 50% peptide digested after 6 h (**Figure 6.3A**). The other two linear peptides SFBK-L and BKSF-L were also comparably unstable to pepsin digestion and degraded gradually during the treatment. These results suggested that the cyclic structure conferred pepsin resistance to the engineered cyclic peptides.

3.4 Trypsin stability assay

To investigate whether the engineered peptides were susceptible to cleavage in the small intestine trypsin digestion assay was carried out mimicking the digestive environment in the small intestine. Results showed that SFBK was resistant to trypsin whereas BKSF was completely digested into smaller fragments after 6 h which was confirmed by analytical HPLC and MALDI-TOF MS (**Figure 6.3B**). Trypsin recognizes basic residues such as lysine and arginine and hydrolyzes the peptide bond at their C-terminal side. Both Arg and Lys were present in the engineered peptide sequence, thus they were digested quickly by the trypsin. However, SFBK was resistant to trypsin digestion because it contained the trypsin inhibition loop. Similar results were obtained from SFTI-1 which was also stable to trypsin digestion for up to 6 h. SFBK-L, a linear peptide containing the trypsin inhibition loop, was also degraded rapidly in the trypsin digestion. It was probably due to the absence of the cyclic backbone, which disturbed the secondary structure. The hydrogen bond and structure integrity was important for the trypsin inhibition activity. Once the structure is disrupted, the activity cannot be retained.

3.5 Serum stability assay

All peptides were subjected to serum stability assays for comparison. The engineered peptides SFBK and BKSF showed exceptional stability in human serum for up to 6 h as they remain >90 % intact whereas the DALK peptide was degraded almost completely after 1 h (**Figure 6.4**). These results strongly suggested that after grafting into the SFTI-1 scaffold, the stability of DALK peptide was greatly improved in serum and allowed the peptide survive from protease degradation.

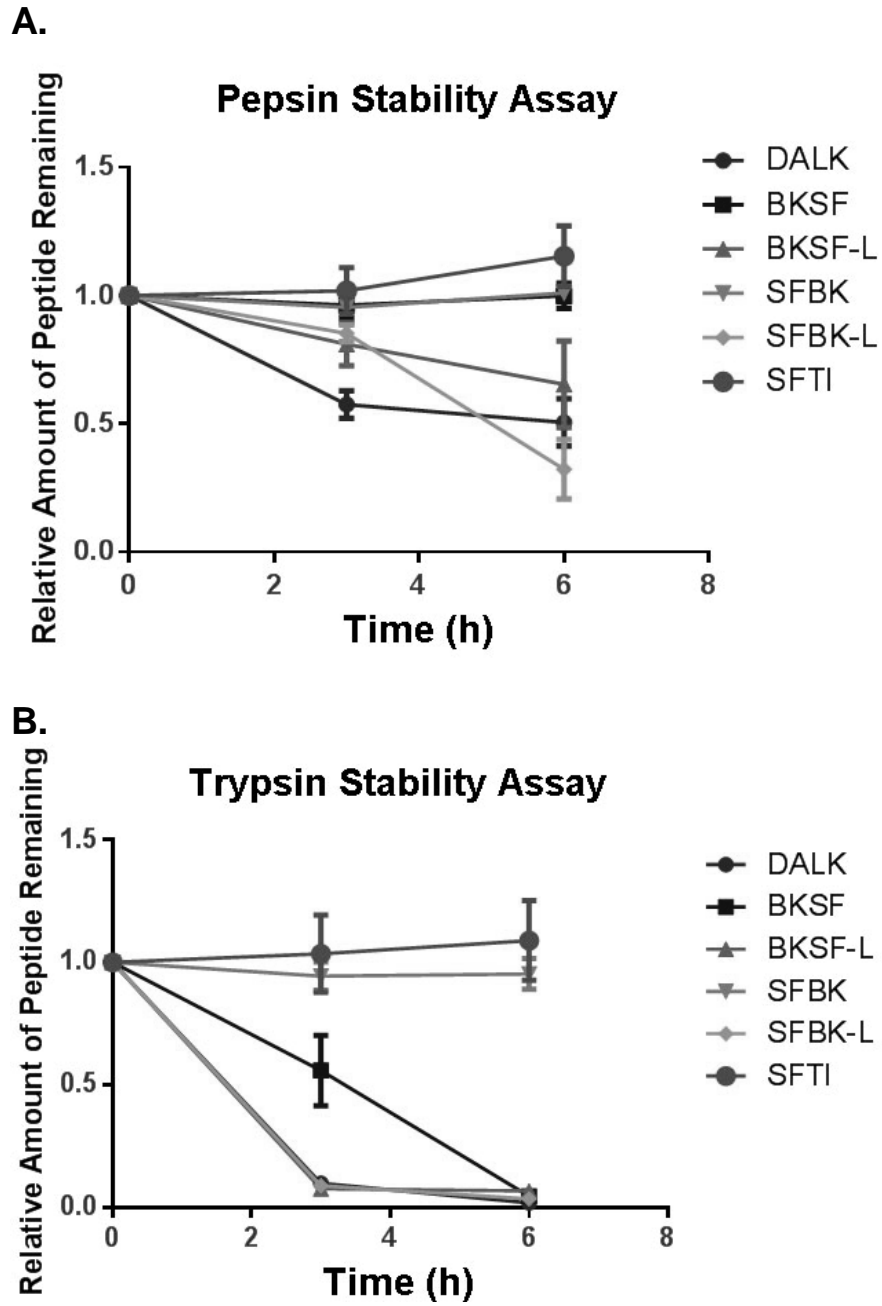


Figure 6.3 Pepsin and trypsin stability of engineered peptides. The experiments were repeated in triplicates.

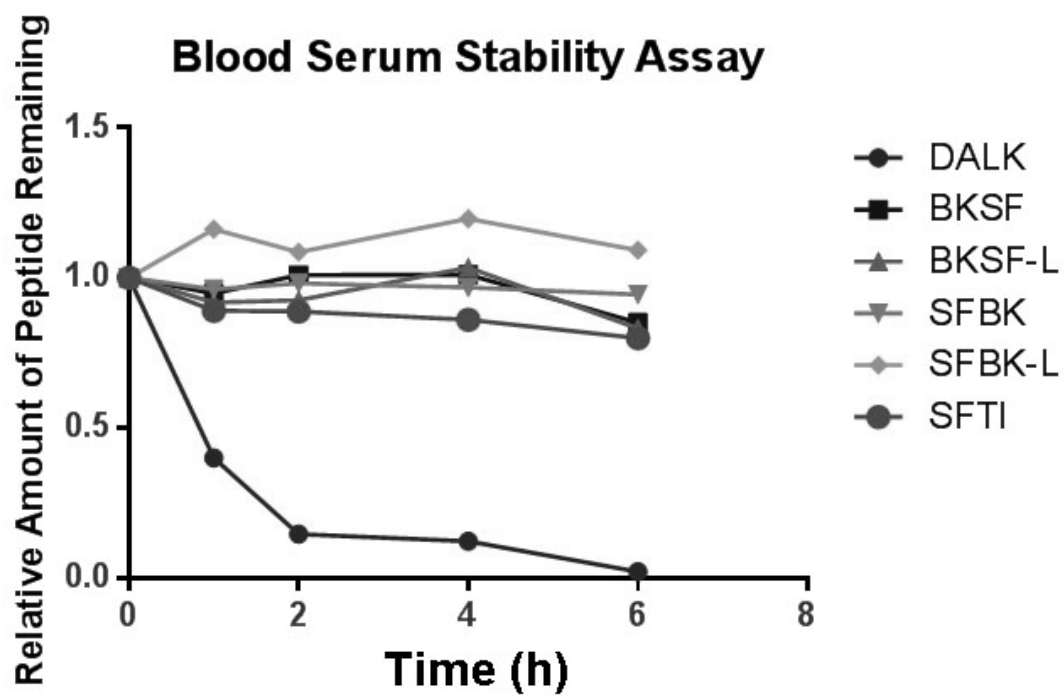


Figure 6.4 Serum stability of engineered peptides. The experiments were repeated in triplicates.

3.6 Trypsin inhibition assay

The trypsin inhibitory activity of SFTI-1 is contributed by its trypsin inhibition loop. It was reported that the Lys⁵-Ser⁶ in this loop of SFTI-1 served as a substrate of trypsin and inserted into the active site through formation of an extended β -sheet. While interacting with the catalytic site, the Lys-Ser bond was cleaved but then regenerated by the trypsin-mediated ligation. Consequently, SFTI-1 efficiently inhibited the catalytic activity of trypsin by occupying the active site.^{97,317}

Since the engineered peptide SFBK still contained the trypsin inhibition loop, we would like to investigate whether the trypsin inhibitory activity was retained. Engineered peptides were subjected to the trypsin treatment in presence of a chromogenic trypsin substrate BAPNA at 37 °C. The activity of trypsin was determined through measurement of the absorbance at 405 nm. As shown in **Figure 6.5**, SFBK exhibited a similar inhibitory effect as SFTI-1 with an IC₅₀ less than 0.5 μ M whereas no inhibitory activity was observed for BKSF up to 100 μ M. These results proved that SFBK retained the trypsin inhibition activity of SFTI-1 and was able to inhibit trypsin with similar potency. The other peptide BKSF without the trypsin inhibition loop was completely inactive in the trypsin inhibition assay, suggesting that the trypsin inhibition activity was dependent on the trypsin inhibition loop. It was noteworthy that SFBK-L was also unable to inhibit trypsin, which was probably due to the disruption of the β -sheet structure that was crucial for its binding to trypsin.

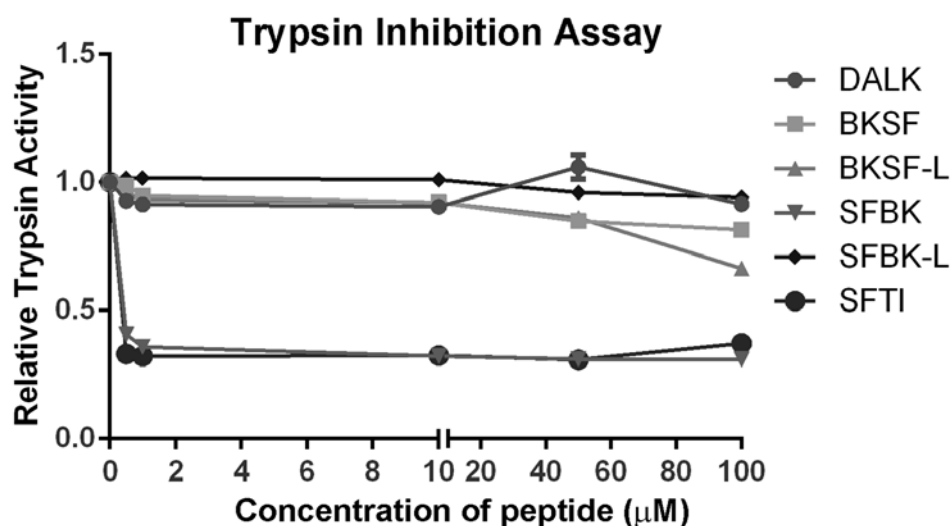


Figure 6.5 Trypsin inhibition activities of engineered peptides. The experiments were repeated in triplicates. The Y axis represents the activity of trypsin. The activity 1.0 was determined by incubating trypsin with the substrate alone without any other inhibitors. Two of the peptides SFBK and SFTI-1 showed an inhibition against trypsin, which lowered the trypsin activity to less than 0.5. The other four peptides did not show significant inhibitory effects for up to 100 µM.

3.7 Bradykinin receptor antagonism assay

To determine the activity of SFBK and BKSF as bradykinin B₁ receptor antagonists, an intracellular Ca²⁺ assay was carried out. Activation of the bradykinin B₁ receptor leads to the activation of phospholipase C and the release of inositol phosphates thus triggering an increase in the level of intracellular Ca²⁺. By adding bradykinin B₁ receptor antagonists, the activation would be suppressed and result in a decrease in the Ca²⁺ level.

Thus, we performed the assay by measuring the intracellular Ca²⁺ level in HeLa cells that were activated by a B₁ agonist (des-Arg⁹-BK). The cells were pre-loaded with a fluorescent dye fluo-4 NW which could be used to determine the intracellular Ca²⁺ level. Then measurements were performed before and after adding the engineered peptides for up to 1 μM. Cytotoxicity of the engineered peptides against HeLa cells was determined before the Ca²⁺ assay, which indicated that our peptides were not cytotoxic to HeLa cells for up to 100 μM.

Our results showed that both the engineered peptides SFBK and BKSF were able to suppress the increase of the Ca²⁺ level induced by the bradykinin B₁ receptor agonist in a concentration-dependent manner (**Figure 6.6**), suggesting that the bradykinin receptor antagonism activity was retained in the engineered peptides. Admittedly, the potency of engineered peptides was lower compared with the original antagonist DALK. The possible reason will be discussed in the discussion section.

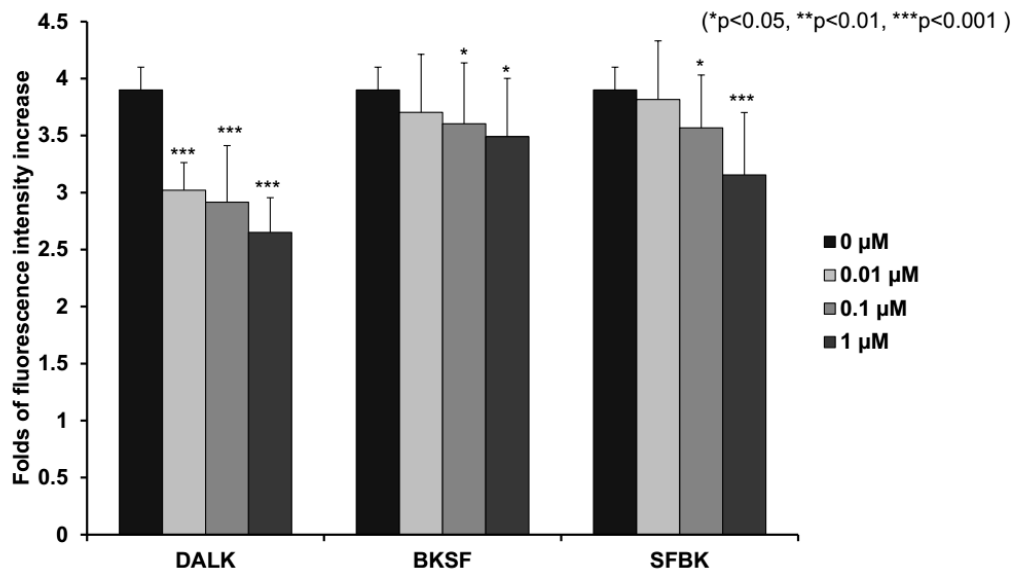


Figure 6.6 Bradykinin receptor antagonism activities of engineered peptides. The Y axis represents the fluorescence change in the HeLa cells, which can be used to indicate the intracellular Ca^{2+} level. Peptides were tested at a concentration from 0 to 1 μM . The increase of the fluorescence was due to the pretreatment of a B_1 agonist with HeLa cells. Peptides with an antagonism activity against bradykinin B_1 receptor would result in a decrease of the fluorescence.

4. Discussion

Currently, chronic pain is a global concern associated with multiple medical conditions such as trauma, infection or post-surgery. Undesirable side effects of common analgesics, such as non-steroidal anti-inflammatory drugs (NSAIDs) administered orally and opiates injected intravenously, continue to pose a problem in pain treatment. NSAIDs are known for their undesirable gastrointestinal side effects while morphine and opiates suffer from the problem of tolerance and addiction. To minimize the dependence on both NSAIDs and opiates, we seek to develop alternative analgesics for pain treatments.

Bradykinin is a pharmacologically active and short-lived peptide, belonging to the kinin group of proteins, and is the most potent mediator as an endogenous pain inducer. Full chain of bradykinin activates the B₂ receptors by the release of kinin during tissue injury, whereas its carboxypeptidase metabolite, [des-Arg⁹]-BK, activate B₁ receptors.¹³⁶ Due to the differences in receptor down regulation and ligand dissociation, the inflammatory phase is divided into the chronic phase and acute phase; the former is stimulated by activation of B₁ receptors, whereas the latter is stimulated by B₂ receptors. Evidence has shown that chronic pain could be treated with bradykinin antagonists, especially B₁ antagonists, in prolonged inflammatory pain.³¹⁸ Therefore, it will be attractive to treat inflammatory pain by engineering an orally active bradykinin B₁ antagonist as a substitute for NSAIDs and opiates.

Due to the short-lived nature of bradykinin antagonists, their application as an orally active therapeutic remains a challenge. To improve their *in vivo* stability and potency, many have developed several generations of bradykinin receptor antagonists mainly by introducing D-amino acids.³¹⁹ However, it may raise a problem of side effects and incomplete clearance *in vivo*.

Here, we described a novel approach to improve the peptide stability by grafting the bradykinin antagonist DALK into a cyclic peptide scaffold SFTI-1. Previously, the same strategy was demonstrated by our laboratory using a cyclotide scaffold kB1. Using SFTI-1 as a scaffold has two advantages. First, the ease of synthesizing the engineered peptides makes large-scale production easier and economically favorable as the SFTI-grafted peptide is much smaller and contained one disulfide bond. Second,

the additional trypsin inhibition activity conferred by the SFTI-1 scaffold when TI loop is retained makes the grafted peptide serve as a bifunctional molecule, which provides improved stability.

We performed acidic, enzymatic stability assays simulating the *in vivo* GI tract environment to evaluate the biological and chemical stability of the engineered peptides. In heat and acidic stability assays, the engineered peptides and their linear counterparts did not show significant differences and all peptides displayed remarkable stabilities in acidic and heating condition for 6 h. These conditions are generally harsh for protein drugs under which most proteins are prone to denaturation and cannot maintain their structures. Our results suggested comparably better stability of peptides under denaturing conditions which highlighted an advantage of peptide-based drugs over the protein-based drugs.

Peptides have to withstand enzymatic degradation to prevent them from being digested before absorption. Thus, we also examined the enzymatic stability of the engineered peptides compared with the original antagonist DALK. Pepsin and trypsin, the major proteases in stomach and small intestine, respectively, were chosen to investigate whether the engineered peptides were susceptible to enzymatic degradation. The results showed that the engineered peptide SFBK was exceptionally resistant against pepsin and trypsin digestion whereas the original peptide DALK was degraded rapidly. It proved that our grafting strategy greatly improved the peptide stability and enhanced its resistance against enzyme digestion. In contrast, the other grafted peptide lacking the trypsin inhibition loop did not show any improvement in terms of trypsin stability, suggesting the trypsin inhibition loop was important for the enhanced stability. Additionally, although the linear form peptide SFBK-L contained a trypsin inhibition loop, it was digested by pepsin and trypsin as well. It was reported that the trypsin inhibitory activity of SFTI-1 was highly dependent on its rigid structure. The open-up of the cyclic backbone disrupted hydrogen bonding and resulted in the loss of activity³¹⁶. We suspected that a weaker network of hydrogen bonding causing by the acyclic backbone of SFBK-L resulted in its susceptibility to enzymatic degradation.

Studies showed that bradykinin was rapidly degraded into a five-residue peptide in human blood serum by aminopeptidase P, dipeptidyl-peptidase IV and angiotensin-converting enzyme.³²⁰ The half-lives of bradykinin and des-Arg⁹-bradykinin in human serum were 27±10 and 643 ± 436 s, respectively.³²¹ Due to the sequence similarity of DALK and bradykinin, DALK was likely hydrolysed by similar enzymes in serum and degraded rapidly. Indeed, we observed a rapid degradation after incubating the DALK peptide in serum. In contrast, our engineered peptides exhibited impressive stability in human serum for >6 h with >90% peptide remained intact. Both the linear and cyclic form peptides displayed similar stabilities. High serum stability is an important factor for bioavailability of peptide drugs as it allows the peptide to circulate in the blood system and bind to specific targets. Through the grafting strategy, the serum stability of the bradykinin antagonist was remarkably improved.

Besides resistance to trypsin digestion, the engineered peptide SFBK also possessed a trypsin inhibitory activity (**Figure 6.5**). The inhibition potency was highly relied on the presence of the trypsin inhibition loop. Those peptides without a trypsin inhibition loop did not exhibit any trypsin inhibitory activity. SFBK, the engineered peptide containing both a trypsin inhibition loop and a bradykinin antagonist loop, displayed similar trypsin inhibitory activity to SFTI-1. It suggests that the SFTI-1 scaffold was tolerant to sequence substitutions and the trypsin inhibition activity was retained after grafting foreign sequences into the peptide scaffold. It should be noted that in the linear peptide SFBK-L, the absence of a cyclic backbone resulted in the loss of trypsin resistance and inhibition. The cyclic backbone appeared to be important to maintain the rigid structure, which has a fundamental role in its function.

The NMR structure resolved by our laboratory member Dr. Wang showed that the overall backbone structure of SFTI-1 was retained in SFBK, especially for the trypsin inhibition loop (**Figure 6.7**). The Pro⁸ of this loop existed in the same *cis*-conformation in both SFTI-1 and SFBK. The overlay structure also showed high similarity between them with almost identical trypsin inhibition loop, which explained the trypsin inhibition activity of SFBK. On the other hand, the other engineered peptide BKSF adopted a slightly different structure from the scaffold SFTI-1. The introduction of foreign sequences in the trypsin inhibition loop seems to disrupt the

hydrogen bonding in the cyclic backbone and constrained the peptide in a different orientation.

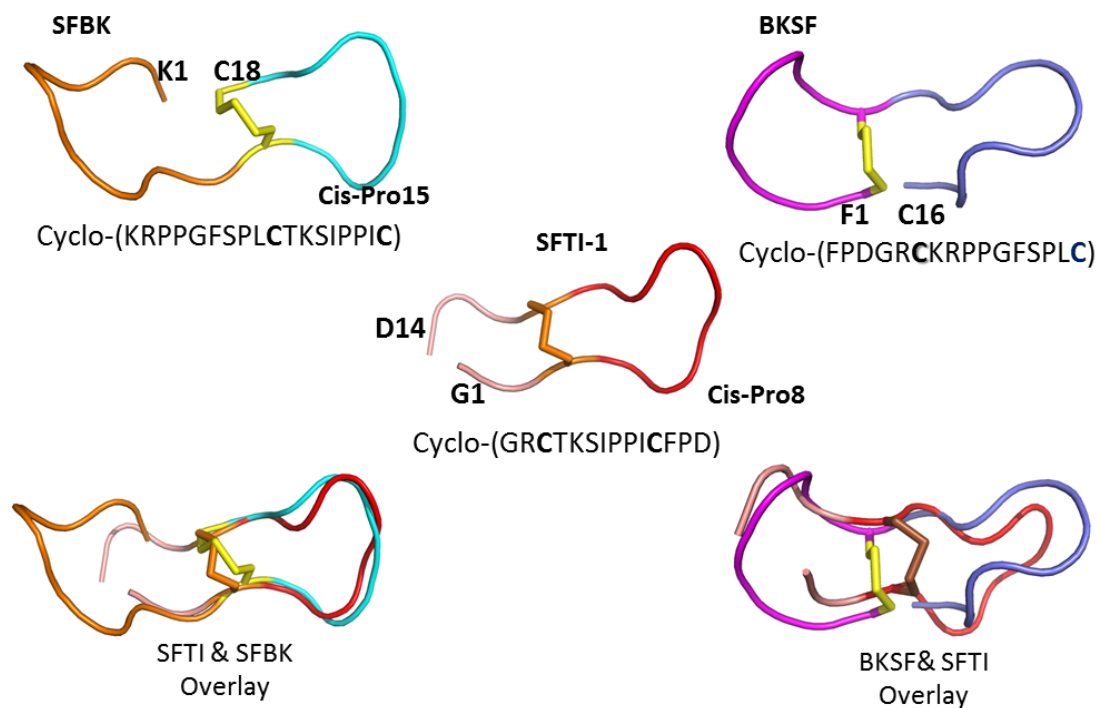


Figure 6.7 NMR structures of SFTI-1, SFBK and BKSF. The SFBK is in orange and green while the BKSF was colored magenta and blue. SFTI-1 was in orange and red. The overall structure of SFTI-1 was retained in SFBK except the longer grafted peptide part whereas BKSF displayed a more distorted structure compared with SFTI-1. Despite the structures appeared to be linear in the cartoon display mode, all structures are cyclic in fact.

Hess and Ha suggested that DALK may adopt a structure containing two reverse turns upon binding to the receptor.³²² The structure of SFBK and BKSF demonstrated a similar structure with two turns at the Pro-Pro-Phe and Pro-Ser site. It suggested the DALK peptide may adopt the active conformation after grafting. It explained their bradykinin antagonism effect we observed in the intracellular Ca^{2+} assay. Admittedly, the effect was not as potent as the original peptide DALK. It may be because the structure of DALK peptide was constrained by the SFTI-1 scaffold, which might interfere with the recognition and conformational change upon binding to the receptor. Unfortunately, due to the time limit, we were not able to look into it thoroughly in this thesis. In the future, detailed docking and molecular dynamics studies will be performed to investigate the receptor binding affinity.

Moreover, studies have shown that SFTI-1 was able to penetrate the mammalian cell membrane and enter into the cell.³²³ Recently, it was confirmed that it was membrane permeable through a non-specific endocytosis mechanism regardless of mutations in the secondary loop.³²⁴ Both suggested that after grafting, the engineered peptide is likely to be membrane permeable, which is essential for the oral bioavailability. Additionally, the hydrophobicity of peptides is known to be important for the cell-uptake. Our laboratory has successfully demonstrated that peptides with clusters of charge and hydrophobic residues such as tryptophan or proline were capable of entering cells.³²⁵ After engineering, the peptide SFBK contained 5 proline, 1 arginine and 1 phenylalanine, which made up 39% of the whole sequence. It highly suggested its cell penetrating property. In summary, with trypsin inhibition and bradykinin receptor antagonism activity, the engineered peptide SFBK was a promising bifunctional analgesics with potential oral bioavailability.

Chapter 7 Conclusions

My thesis focuses on developing practical methodologies for preparing cysteine-rich peptides and application of them as drug scaffolds. In this thesis, total syntheses of four cyclic peptides SFTI-1, kB1, SFBK and BKSF were achieved by various chemical approaches with high yields.

We successfully developed an Fmoc-compatible cyclization approach using thioethylamido (TEA) thioester surrogates. This functional moiety is essential to catalyze acyl shift reactions for the thioester formation. Based on the TEA structure, two thioester surrogates MeCys and TEBA were designed and employed for the synthesis of cyclic peptide SFTI-1. With the assistance of this zip cyclization mechanism, the cyclization and folding reaction were optimized in one pot with 43% overall yield. This method is simple, efficient and can apply for the large scale production of other cysteine-rich peptides.

In the synthesis of cysteine-rich peptides such as cyclotides, the oxidative folding process is the most challenging step. Innovatively, we conducted the peptide oxidative folding in neat organic solvents instead of aqueous buffers. The oxidative folding for two model peptides kB1 and ET-1 were optimized by varying organic solvents folding conditions. The organic folding greatly increased the yield and reaction rate of corresponding folding reactions. The optimized condition afforded the native form kB1 in 1 h with an 88% yield. We also demonstrated that the organic folding was able to combine with the cyclization reaction in one-pot for synthesis of cysteine-rich peptides.

In addition, a chemoenzymatic method was developed for synthesis of cyclic peptides using a novel transpeptidase butelase 1. The substrate specificity of this enzyme was investigated at P1, P2, P1'', P2'' and P3'' sites. The results showed that butelase 1 displayed broad substrate specificity with minimal sequence requirements. At P1'' and P2'' positions, even D-amino acids were tolerated resulting in moderate ligation yields. With the determined substrate specificity, butelase 1 can apply for peptide and protein labeling as well as peptide macrocyclization.

In last chapter of my thesis, I described the application of cysteine-rich peptides as drug scaffolds. Two cyclic peptide hybrids SFBK and BKSF were prepared by grafting a bradykinin B₁ receptor antagonist into an SFTI-1 scaffold. The engineered peptides particularly SFBK demonstrated improved stability against thermal and enzymatic digestions. They were stable in human serum for up to 6 h. *In vitro* characterization showed that the chimeric peptide SFBK possessed two functions as a trypsin inhibitor and a bradykinin B₁ receptor antagonist. The work reported here demonstrates the concept of grafting strategy for developing stable and orally active peptide therapeutics.

Future work includes: 1. synthesis of other cysteine-rich peptides by the chemical and chemoenzymatic methods developed in this thesis. 2. Investigation of the oral bioavailability of the engineered peptide SFBK using *in vivo* abdominal constriction assay and tail-flick assay.^{326,327}

References

1. Miseta, A., and Csutora, P. (2000) Relationship between the occurrence of cysteine in proteins and the complexity of organisms. *Mol. Biol. Evol.* **17**, 1232-1239
2. Corpuz, G. P., Jacobsen, R. B., Jimenez, E. C., Watkins, M., Walker, C., Colledge, C., Garrett, J. E., McDougal, O., Li, W., Gray, W. R., Hillyard, D. R., Rivier, J., McIntosh, J. M., Cruz, L. J., and Olivera, B. M. (2005) Definition of the M-conotoxin superfamily: characterization of novel peptides from molluscivorous *Conus* venoms. *Biochemistry* **44**, 8176-8186
3. Olivera, B., Rivier, J., Clark, C., Ramilo, C., Corpuz, G., Abogadie, F., Mena, E., Woodward, Hillyard, D., and Cruz, L. (1990) Diversity of *Conus* neuropeptides. *Science* **249**, 257-263
4. Nelson, L. (2004) Venomous snails: one slip, and you're dead. *Nature* **429**, 798-799
5. Ganz, T. (2003) Defensins: antimicrobial peptides of innate immunity. *Nat. Rev. Immunol.* **3**, 710-720
6. Bulet, P., Stöcklin, R., and Menin, L. (2004) Anti-microbial peptides: from invertebrates to vertebrates. *Immunol. Rev.* **198**, 169-184
7. Ryan, C. A. (1990) Protease inhibitors in plants: genes for improving defenses against insects and pathogens. *Annu. Rev. Phytopathol.* **28**, 425-449
8. Marshall, E., Costa, L. M., and Gutierrez-Marcos, J. (2011) Cysteine-Rich Peptides (CRPs) mediate diverse aspects of cell–cell communication in plant reproduction and development. *J. Exp. Bot.* **62**, 1677-1686
9. Silverstein, K. A. T., Moskal, W. A., Wu, H. C., Underwood, B. A., Graham, M. A., Town, C. D., and VandenBosch, K. A. (2007) Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. *Plant J.* **51**, 262-280
10. Terlau, H., and Olivera, B. M. (2004) *Conus* venoms: a rich source of novel ion channel-targeted peptides. *Physiol. Rev.* **84**, 41-68
11. Nicholson, G. M., and Gaudins, A. (2002) Spiders of medical importance in the Asia–Pacific: Atracotoxin, latrotoxin and related spider neurotoxins. *Clin. Exp. Pharmacol. Physiol.* **29**, 785-794
12. Swartz, K. J., and MacKinnon, R. (1995) An inhibitor of the Kv2.1 potassium channel isolated from the venom of a Chilean tarantula. *Neuron* **15**, 941-949
13. Zeng, X.-Z., Xiao, Q.-B., and Liang, S.-P. (2003) Purification and characterization of raventoxin-I and raventoxin-III, two neurotoxic peptides from the venom of the spider *Macrothele raveni*. *Toxicon* **41**, 651-656
14. Rogowski, R. S., Collins, J. H., O'Neill, T. J., Gustafson, T. A., Werkman, T. R., Rogowski, M. A., Tenenholz, T. C., Weber, D. J., and Blaustein, M. P. (1996) Three

- new toxins from the scorpion *Pandinus imperator* selectively block certain voltage-gated K⁺ channels. *Mol. Pharmacol.* **50**, 1167-1177
15. Kloog, Y., Ambar, I., Sokolovsky, M., Kochva, E., Wollberg, Z., and Bdolah, A. (1988) Sarafotoxin, a novel vasoconstrictor peptide: phosphoinositide hydrolysis in rat heart and brain. *Science* **242**, 268-270
 16. McIntosh, J. M., and Jones, R. M. (2001) Cone venom—from accidental stings to deliberate injection. *Toxicon* **39**, 1447-1451
 17. Nicke, A., Wonnacott, S., and Lewis, R. J. (2004) α -Conotoxins as tools for the elucidation of structure and function of neuronal nicotinic acetylcholine receptor subtypes. *Eur. J. Biochem.* **271**, 2305-2319
 18. Nielsen, K. J., Schroeder, T., and Lewis, R. (2000) Structure–activity relationships of ω -conotoxins at N-type voltage-sensitive calcium channels. *J. Mol. Recognit.* **13**, 55-70
 19. Miljanich, G. (2004) Ziconotide: neuronal calcium channel blocker for treating severe chronic pain. *Curr. Med. Chem.* **11**, 3029-3040
 20. Bowersox, S. S., and Luther, R. (1998) Pharmacotherapeutic potential of omega-conotoxin MVIIA (SNX-111), an N-type neuronal calcium channel blocker found in the venom of *Conus magus*. *Toxicon* **36**, 1651-1658
 21. Lehrer, R. I. (2004) Primate defensins. *Nat. Rev. Micro.* **2**, 727-738
 22. Hancock, R. E. W., and Chapple, D. S. (1999) Peptide antibiotics. *Antimicrob. Agents Chemother.* **43**, 1317-1323
 23. Schneider, J., Unholzer, A., Schaller, M., Schäfer-Korting, M., and Korting, H. (2005) Human defensins. *J. Mol. Med.* **83**, 587-595
 24. Tang, Y.-Q., Yuan, J., Ösapay, G., Ösapay, K., Tran, D., Miller, C. J., Ouellette, A. J., and Selsted, M. E. (1999) A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated α -defensins. *Science* **286**, 498-502
 25. Thomma, B., Cammue, B., and Thevissen, K. (2002) Plant defensins. *Planta* **216**, 193-202
 26. Stec, B. (2006) Plant thionins – the structural perspective. *Cell Mol. Life Sci.* **63**, 1370-1385
 27. Van Parijs, J., Broekaert, W., Goldstein, I., and Peumans, W. (1991) Hevein: an antifungal protein from rubber-tree (*Hevea brasiliensis*) latex. *Planta* **183**, 258-264
 28. Broekaert, W. F., Cammue, B. P. A., De Bolle, M. F. C., Thevissen, K., De Samblanx, G. W., Osborn, R. W., and Nielson, K. (1997) Antimicrobial peptides from plants. *Crit. Rev. Plant Sci.* **16**, 297-323
 29. Molina, A., and García-Olmedo, F. (1993) Developmental and pathogen-induced expression of three barley genes encoding lipid transfer proteins. *Plant J.* **4**, 983-991

30. Segura, A., Moreno, M., Madueño, F., Molina, A., and García-Olmedo, F. (1999) Snakin-1, a peptide from potato that is active against plant pathogens. *Mol. Plant Microbe Interact.* **12**, 16-23
31. Berrocal-Lobo, M., Segura, A., Moreno, M., Lopez, G., Garcia-Olmedo, F., and Molina, A. (2002) Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. *Plant Physiol.* **128**, 951-961
32. Tam, J. P., Lu, Y. A., Yang, J. L., and Chiu, K. W. (1999) An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8913-8918
33. Gustafson, K. R., McKee, T. C., and Bokesch, H. R. (2004) Anti-HIV cyclotides. *Curr. Protein Pept. Sci.* **5**, 331-340
34. Gustafson, K. R., Sowder, R. C., Henderson, L. E., Parsons, I. C., Kashman, Y., Cardellina, J. H., McMahon, J. B., Buckheit, R. W., Pannell, L. K., and Boyd, M. R. (1994) Circulins A and B. Novel human immunodeficiency virus (HIV)-inhibitory macrocyclic peptides from the tropical tree *Chassalia parvifolia*. *J. Am. Chem. Soc.* **116**, 9337-9338
35. Svängård, E., Göransson, U., Hocaoglu, Z., Gullbo, J., Larsson, R., Claeson, P., and Bohlin, L. (2003) Cytotoxic Cyclotides from *Viola tricolor*. *J. Nat. Prod.* **67**, 144-147
36. Witherup, K. M., Bogusky, M. J., Anderson, P. S., Ramjit, H., Ransom, R. W., Wood, T., and Sardana, M. (1994) Cyclopsychotride A, a biologically active, 31-Residue cyclic peptide isolated from *Psychotria longipes*. *J. Nat. Prod.* **57**, 1619-1625
37. Chagolla-Lopez, A., Blanco-Labra, A., Patthy, A., Sánchez, R., and Pongor, S. (1994) A novel alpha-amylase inhibitor from amaranth (*Amaranthus hypocondriacus*) seeds. *J. Biol. Chem.* **269**, 23675-23680
38. Rudinger, J., and Jošt, K. (1964) A biologically active analogue of oxytocin not containing a disulfide group. *Experientia* **20**, 570-571
39. Pliska, V., Rudinger, J., Dousa, T., and Cort, J. H. (1968) Oxytocin activity and the integrity of the disulfide bridge. *Am. J. Physiol.* **215**, 916-920
40. Conibear, A. C., Rosengren, K. J., Daly, N. L., Henriques, S. T., and Craik, D. J. (2013) The cyclic cystine ladder in θ -defensins is important for structure and stability, but not antibacterial activity. *J. Biol. Chem.* **288**, 10830-10840
41. Cornet, B., Bonmatin, J.-M., Hetru, C., Hoffmann, J. A., Ptak, M., and Vovelle, F. (1995) Refined three-dimensional solution structure of insect defensin A. *Structure* **3**, 435-448
42. Bontems, F., Roumestand, C., Boyot, P., Gilquin, B., Doljansky, Y., Menez, A., and Toma, F. (1991) Three-dimensional structure of natural charybdotoxin in aqueous solution by ¹H-NMR Charybdotoxin possesses a structural motif found in other scorpion toxins. *Eur. J. Biochem.* **196**, 19-28

43. Kobayashi, Y., Takashima, H., Tamaoki, H., Kyogoku, Y., Lambert, P., Kuroda, H., Chino, N., Watanabe, T. X., Kimura, T., Sakakibara, S., and Moroder, L. (1991) The cystine-stabilized α -helix: a common structural motif of ion-channel blocking neurotoxic peptides. *Biopolymers* **31**, 1213-1220
44. Tamaoki, H., Miura, R., Kusunoki, M., Kyogoku, Y., Kobayashi, Y., and Moroder, L. (1998) Folding motifs induced and stabilized by distinct cystine frameworks. *Protein Eng.* **11**, 649-659
45. Kohno, T., Kim, J. I., Kobayashi, K., Kodera, Y., Maeda, T., and Sato, K. (1995) Three-dimensional structure in solution of the calcium channel blocker ω -conotoxin MVIIA. *Biochemistry* **34**, 10256-10265
46. Goransson, U., and Craik, D. J. (2003) Disulfide mapping of the cyclotide kalata B1. Chemical proof of the cystic cystine knot motif. *J. Biol. Chem.* **278**, 48188-48196
47. McDonald, N. Q., and Hendrickson, W. A. (1993) A structural superfamily of growth factors containing a cystine knot motif. *Cell* **73**, 421-424
48. Korsinczky, M. L. J., Schirra, H. J., Rosengren, K. J., West, J., Condie, B. A., Otvos, L., Anderson, M. A., and Craik, D. J. (2001) Solution structures by ¹H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant. *J. Mol. Biol.* **311**, 579-591
49. Fahrner, R. L., Dieckmann, T., Harwig, S. S. L., Lehrer, R. I., Eisenberg, D., and Feigon, J. (1996) Solution structure of protegrin-1, a broad-spectrum antimicrobial peptide from porcine leukocytes. *Chem. Biol.* **3**, 543-550
50. Laederach, A., Andreotti, A. H., and Fulton, D. B. (2002) Solution and micelle-bound structures of tachyplesin I and its active aromatic linear derivatives. *Biochemistry* **41**, 12359-12368
51. Trabi, M., Schirra, H. J., and Craik, D. J. (2001) Three-dimensional structure of RTD-1, a cyclic antimicrobial defensin from *Rhesus macaque* leukocytes. *Biochemistry* **40**, 4211-4221
52. Lay, F. T., Schirra, H. J., Scanlon, M. J., Anderson, M. A., and Craik, D. J. (2003) The three-dimensional solution structure of NaD1, a new floral defensin from *Nicotiana glauca* and its application to a homology model of the crop defense protein alfAFP. *J. Mol. Biol.* **325**, 175-188
53. Janes, R. W., Peapus, D. H., and Wallace, B. A. (1994) The crystal structure of human endothelin. *Nat. Struct. Mol. Biol.* **1**, 311-319
54. Mourier, G., Hajj, M., Cordier, F., Zorba, A., Gao, X., Coskun, T., Herbet, A., Marcon, E., Beau, F., Delepierre, M., Ducancel, F., and Servent, D. (2012) Pharmacological and structural characterization of long-sarafotoxins, a new family of endothelin-like peptides: Role of the C-terminus extension. *Biochimie* **94**, 461-470
55. Saether, O., Craik, D. J., Campbell, I. D., Sletten, K., Juul, J., and Norman, D. G. (1995) Elucidation of the primary and three-dimensional structure of the uterotonic polypeptide kalata B1. *Biochemistry* **34**, 4147-4158

56. Lu, S., Deng, P., Liu, X., Luo, J., Han, R., Gu, X., Liang, S., Wang, X., Li, F., Lozanov, V., Patthy, A., and Pongor, S. (1999) Solution structure of the major α -amylase inhibitor of the crop plant *Amaranth*. *J. Biol. Chem.* **274**, 20473-20478
57. Cascales, L., and Craik, D. J. (2010) Naturally occurring circular proteins: distribution, biosynthesis and evolution. *Org. Biomol. Chem.* **8**, 5035-5047
58. Thorstholm, L., and Craik, D. J. (2012) Discovery and applications of naturally occurring cyclic peptides. *Drug Discov. Today Technol.* **9**, e13-e21
59. Hallen, H. E., Luo, H., Scott-Craig, J. S., and Walton, J. D. (2007) Gene family encoding the major toxins of lethal *Amanita* mushrooms. *Proc. Natl. Acad. Sci. U.S.A.* **104** 19097-19101
60. Gosselin, R. E., and Gleason, M. N. (1976) *Clinical toxicology of commercial products: acute poisoning*, Williams & Wilkins
61. Montalbán-López, M., Sánchez-Hidalgo, M., Cebrián, R., and Maqueda, M. (2012) Discovering the bacterial circular proteins: bacteriocins, cyanobactins, and pilins. *J. Biol. Chem.* **287**, 27007-27013
62. Eisenbrandt, R., Kalkum, M., Lai, E.-M., Lurz, R., Kado, C. I., and Lanka, E. (1999) Conjugative pili of IncP plasmids, and the Ti plasmid T pilus are composed of cyclic subunits. *J. Biol. Chem.* **274**, 22548-22555
63. González, C., Langdon, G. M., Bruix, M., Gálvez, A., Valdivia, E., Maqueda, M., and Rico, M. (2000) Bacteriocin AS-48, a microbial cyclic polypeptide structurally and functionally related to mammalian NK-lysin. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11221-11226
64. Sivonen, K., Leikoski, N., Fewer, D., and Jokela, J. (2010) Cyanobactins—ribosomal cyclic peptides produced by cyanobacteria. *Appl. Microbiol. Biotechnol.* **86**, 1213-1225
65. Cole, A. M., Hong, T., Boo, L. M., Nguyen, T., Zhao, C., Bristol, G., Zack, J. A., Waring, A. J., Yang, O. O., and Lehrer, R. I. (2002) Retrocyclin: a primate peptide that protects cells from infection by T- and M-tropic strains of HIV-1. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1813-1818
66. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) Plant cyclotides: a unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.* **294**, 1327-1336
67. Luckett, S., Garcia, R. S., Barker, J. J., Konarev, A. V., Shewry, P. R., Clarke, A. R., and Brady, R. L. (1999) High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. *J. Mol. Biol.* **290**, 525-533
68. Tan, N.-H., and Zhou, J. (2006) Plant cyclopeptides. *Chem. Rev.* **106**, 840-895
69. Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K.-D.,

- Fischbach, M. A., Garavelli, J. S., Goransson, U., Gruber, C. W., Haft, D. H., Hemscheidt, T. K., Hertweck, C., Hill, C., Horswill, A. R., Jaspars, M., Kelly, W. L., Klinman, J. P., Kuipers, O. P., Link, A. J., Liu, W., Marahiel, M. A., Mitchell, D. A., Moll, G. N., Moore, B. S., Muller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J. T., Rebuffat, S., Ross, R. P., Sahl, H.-G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Sussmuth, R. D., Tagg, J. R., Tang, G.-L., Truman, A. W., Vederas, J. C., Walsh, C. T., Walton, J. D., Wenzel, S. C., Willey, J. M., and van der Donk, W. A. (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **30**, 108-160
70. Gruber, C. W., Elliott, A. G., Ireland, D. C., Delprete, P. G., Dessen, S., Göransson, U., Trabi, M., Wang, C. K., Kinghorn, A. B., and Robbrecht, E. (2008) Distribution and evolution of circular miniproteins in flowering plants. *Plant Cell* **20**, 2471-2483
71. Gran, L. (1970) An oxytocic principle found in *Oldenlandia affinis* DC. An indigenous, congolese drug "Kalata-Kalata" used to accelerate delivery. *Medd. Nor. Farm. Selsk.* **32**, 173-180
72. Gran, L. (1973) On the effect of a polypeptide isolated from "Kalata-Kalata" (*Oldenlandia affinis* DC) on the oestrogen dominated uterus. *Acta Pharmacol. Toxicol.* **33**, 400-408
73. Nguyen, G. K. T., Lim, W. H., Nguyen, P. Q. T., and Tam, J. P. (2012) Novel cyclotides and uncyclotides with highly shortened precursors from *Chassalia chartacea* and effects of methionine oxidation on bioactivities. *J. Biol. Chem.* **287**, 17598-17607
74. Nguyen, G. K. T., Zhang, S., Wang, W., Wong, C. T. T., Nguyen, N. T. K., and Tam, J. P. (2011) Discovery of a linear cyclotide from the bracelet subfamily and its disulfide mapping by top-down mass spectrometry. *J. Biol. Chem.* **286**, 44833-44844
75. Pinto, M. F. S., Fensterseifer, I. C. M., Migliolo, L., Sousa, D. A., de Capdville, G., Arboleda-Valencia, J. W., Colgrave, M. L., Craik, D. J., Magalhães, B. S., Dias, S. C., and Franco, O. L. (2012) Identification and structural characterization of novel cyclotide with activity against an insect pest of sugar cane. *J. Biol. Chem.* **287**, 134-147
76. Trabi, M., Svängård, E., Herrmann, A., Göransson, U., Claeson, P., Craik, D. J., and Bohlin, L. (2004) Variations in cyclotide expression in *Viola* species. *J. Nat. Prod.* **67**, 806-810
77. Poth, A. G., Colgrave, M. L., Philip, R., Kerenga, B., Daly, N. L., Anderson, M. A., and Craik, D. J. (2010) Discovery of cyclotides in the Fabaceae plant family provides new insights into the cyclization, evolution, and distribution of circular proteins. *ACS Chem. Biol.* **6**, 345-355
78. Nguyen, G. K. T., Zhang, S., Nguyen, N. T. K., Nguyen, P. Q. T., Chiu, M. S., Hardjojo, A., and Tam, J. P. (2011) Discovery and characterization of novel cyclotides originated from chimeric precursors consisting of albumin-1 chain a and cyclotide domains in the *Fabaceae* family. *J. Biol. Chem.* **286**, 24275-24287

79. Poth, A. G., Mylne, J. S., Grassl, J., Lyons, R. E., Millar, A. H., Colgrave, M. L., and Craik, D. J. (2012) Cyclotides associate with leaf vasculature and are the products of a novel precursor in *Petunia* (Solanaceae). *J. Biol. Chem.* **287**, 27033-27046
80. Hernandez, J. F., Gagnon, J., Chiche, L., Nguyen, T. M., Andrieu, J. P., Heitz, A., Trinh Hong, T., Pham, T. T., and Le Nguyen, D. (2000) Squash trypsin inhibitors from *Momordica cochinchinensis* exhibit an atypical macrocyclic structure. *Biochemistry* **39**, 5722-5730
81. Nguyen, G. K. T., Lian, Y., Pang, E. W. H., Nguyen, P. Q. T., Tran, T. D., and Tam, J. P. (2013) Discovery of linear cyclotides in monocot plant *Panicum laxum* of *Poaceae* Family provides new insights into evolution and distribution of cyclotides in plants. *J. Biol. Chem.* **288**, 3370-3380
82. Craik, D. J., Čemažar, M., Wang, C. K. L., and Daly, N. L. (2006) The cyclotide family of circular miniproteins: Nature's combinatorial peptide template. *Biopolymers* **84**, 250-266
83. Gunasekera, S., Daly, N. L., Clark, R. J., and Craik, D. J. (2009) Dissecting the oxidative folding of circular cystine knot miniproteins. *Antioxid. Redox Signal.* **11**, 971-980
84. Dutton, J. L., Renda, R. F., Waine, C., Clark, R. J., Daly, N. L., Jennings, C. V., Anderson, M. A., and Craik, D. J. (2004) Conserved structural and sequence elements implicated in the processing of gene-encoded circular proteins. *J. Biol. Chem.* **279**, 46858-46867
85. Mylne, J. S., Chan, L. Y., Chanson, A. H., Daly, N. L., Schaefer, H., Bailey, T. L., Nguyencong, P., Cascales, L., and Craik, D. J. (2012) Cyclic peptides arising by evolutionary parallelism via asparaginyl-endopeptidase-mediated biosynthesis. *Plant Cell* **24**, 2765-2778
86. Mylne, J. S., Colgrave, M. L., Daly, N. L., Chanson, A. H., Elliott, A. G., McCallum, E. J., Jones, A., and Craik, D. J. (2011) Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. *Nat Chem Biol* **7**, 257-259
87. Craik, D. J., and Malik, U. (2013) Cyclotide biosynthesis. *Curr. Opin. Chem. Biol.* **17**, 546-554
88. Gruber, C. W., Cemazar, M., Clark, R. J., Horibe, T., Renda, R. F., Anderson, M. A., and Craik, D. J. (2007) A novel plant protein-disulfide isomerase involved in the oxidative folding of cystine knot defense proteins. *J. Biol. Chem.* **282**, 20435-20446
89. Zito, E., Melo, E. P., Yang, Y., Wahlander, Å., Neubert, T. A., and Ron, D. (2010) Oxidative protein folding by an endoplasmic reticulum-localized peroxiredoxin. *Mol. Cell* **40**, 787-797
90. Gunasekera, S., Daly, N. L., Anderson, M. A., and Craik, D. J. (2006) Chemical synthesis and biosynthesis of the cyclotide family of circular proteins. *IUBMB Life* **58**, 515-524
91. Gillon, A. D., Saska, I., Jennings, C. V., Guarino, R. F., Craik, D. J., and Anderson, M. A. (2008) Biosynthesis of circular proteins in plants. *Plant J.* **53**, 505-515

92. Saska, I., Gillon, A. D., Hatsugai, N., Dietzgen, R. G., Hara-Nishimura, I., Anderson, M. A., and Craik, D. J. (2007) An asparaginyl endopeptidase mediates *in vivo* protein backbone cyclization. *J. Biol. Chem.* **282**, 29721-29728
93. Conlan, B. F., Gillon, A. D., Barbata, B. L., and Anderson, M. A. (2011) Subcellular targeting and biosynthesis of cyclotides in plant cells. *Am. J. Bot.* **98**, 2018-2026
94. Conlan, B. F., Gillon, A. D., Craik, D. J., and Anderson, M. A. (2010) Circular proteins and mechanisms of cyclization. *Biopolymers* **94**, 573-583
95. Thongyoo, P., Jaulent, A. M., Tate, E. W., and Leatherbarrow, R. J. (2007) Immobilized protease-assisted synthesis of engineered cysteine-knot microproteins. *ChemBioChem* **8**, 1107-1109
96. Thongyoo, P., Roque-Rosell, N., Leatherbarrow, R. J., and Tate, E. W. (2008) Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org. Biomol. Chem.* **6**, 1462-1470
97. Marx, U. C., Korsinczky, M. L. J., Schirra, H. J., Jones, A., Condie, B., Otvos, L., and Craik, D. J. (2003) Enzymatic cyclization of a potent Bowman-Birk protease inhibitor, sunflower trypsin inhibitor-1, and solution structure of an acyclic precursor peptide. *J. Biol. Chem.* **278**, 21782-21789
98. Jennings, C., West, J., Waite, C., Craik, D., and Anderson, M. (2001) Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10614-10619
99. Craik, D. J., Daly, N. L., Mulvenna, J., Plan, M. R., and Trabi, M. (2004) Discovery, structure and biological activities of the cyclotides. *Curr. Protein Pept. Sci.* **5**, 297-315
100. Göransson, U., Sjögren, M., Svängård, E., Claeson, P., and Bohlin, L. (2004) Reversible antifouling effect of the cyclotide cycloviolacin O2 against barnacles. *J. Nat. Prod.* **67**, 1287-1290
101. Lindholm, P., Göransson, U., Johansson, S., Claeson, P., Gullbo, J., Larsson, R., Bohlin, L., and Backlund, A. (2002) Cyclotides: a novel type of cytotoxic agents. *Mol. Cancer Ther.* **1**, 365-369
102. Ireland, D. C., Wang, C. K. L., Wilson, J. A., Gustafson, K. R., and Craik, D. J. (2008) Cyclotides as natural anti-HIV agents. *Pept. Sci.* **90**, 51-60
103. Burman, R., Gunasekera, S., Strömstedt, A. A., and Göransson, U. (2014) Chemistry and biology of cyclotides: circular plant peptides outside the box. *J. Nat. Prod.* **77**, 724-736
104. Wang, C. K., Colgrave, M. L., Ireland, D. C., Kaas, Q., and Craik, D. J. (2009) Despite a conserved cystine knot motif, different cyclotides have different membrane binding modes. *Biophys. J.* **97**, 1471-1481
105. Henriques, S. T., Huang, Y.-H., Castanho, M. A. R. B., Bagatolli, L. A., Sonza, S., Tachedjian, G., Daly, N. L., and Craik, D. J. (2012) Phosphatidylethanolamine-

- binding is a conserved feature of cyclotide-membrane interactions. *J. Biol. Chem.* **287**, 33629-33643
106. Tam, J. P., Lu, Y.-A., and Yang, J.-L. (2002) Correlations of cationic charges with salt sensitivity and microbial specificity of cystine-stabilized β -strand antimicrobial peptides. *J. Biol. Chem.* **277**, 50450-50456
 107. Göransson, U., Herrmann, A., Burman, R., Haugaard-Jönsson, L. M., and Rosengren, K. J. (2009) The conserved Glu in the cyclotide cycloviolacin O2 has a key structural role. *ChemBioChem* **10**, 2354-2360
 108. Covic, A., and Kuhlmann, M. K. (2007) Biosimilars: recent developments. *Int. Urol. Nephrol.* **39**, 261-266
 109. Schafmeister, C. E., Po, J., and Verdine, G. L. (2000) An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *J. Am. Chem. Soc.* **122**, 5891-5892
 110. Vagner, J., Qu, H., and Hruby, V. J. (2008) Peptidomimetics, a synthetic tool of drug discovery. *Curr. Opin. Chem. Biol.* **12**, 292-296
 111. Ward, P., Ewan, G. B., Jordan, C. C., Ireland, S. J., Hagan, R. M., and Brown, J. R. (1990) Potent and highly selective neurokinin antagonists. *J. Med. Chem.* **33**, 1848-1851
 112. Yu, Q., Lehrer, R. I., and Tam, J. P. (2000) Engineered salt-insensitive α -defensins with end-to-end circularized structures. *J. Biol. Chem.* **275**, 3943-3949
 113. Tam, J. P., Lu, Y.-A., and Yang, J.-L. (2000) Marked increase in membranolytic selectivity of novel cyclic tachyplesins constrained with an antiparallel two- β strand cystine knot framework. *Biochem. Biophys. Res. Commun.* **267**, 783-790
 114. Tam, J. P., Wu, C., and Yang, J.-L. (2000) Membranolytic selectivity of cystine-stabilized cyclic protegrins. *Eur. J. Biochem.* **267**, 3289-3300
 115. Oudhoff, M. J., Kroeze, K. L., Nazmi, K., van den Keijbus, P. A. M., van 't Hof, W., Fernandez-Borja, M., Hordijk, P. L., Gibbs, S., Bolscher, J. G. M., and Veerman, E. C. I. (2009) Structure-activity analysis of histatin, a potent wound healing peptide from human saliva: cyclization of histatin potentiates molar activity 1000-fold. *Faseb J* **23**, 3928-3935
 116. Clark, R. J., Fischer, H., Dempster, L., Daly, N. L., Rosengren, K. J., Nevin, S. T., Meunier, F. A., Adams, D. J., and Craik, D. J. (2005) Engineering stable peptide toxins by means of backbone cyclization: stabilization of the alpha-conotoxin MII. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13767-13772
 117. Clark, R. J., Jensen, J., Nevin, S. T., Callaghan, B. P., Adams, D. J., and Craik, D. J. (2010) The engineering of an orally active conotoxin for the treatment of neuropathic pain. *Angew. Chem., Int. Ed.* **49**, 6545-6548
 118. Bolscher, J. G., Oudhoff, M. J., Nazmi, K., Antos, J. M., Guimaraes, C. P., Spooner, E., Haney, E. F., Garcia Vallejo, J. J., Vogel, H. J., van't Hof, W., Ploegh, H. L., and

- Veerman, E. C. (2011) Sortase A as a tool for high-yield histatin cyclization. *Faseb J* **25**, 2650-2658
119. Daly, Norelle L., Clark, Richard J., and Craik, David J. (2006) Structural plasticity of the cyclic-cystine-knot framework: implications for biological activity and drug design. *Biochem. J.* **394**, 85-93
120. Vita, C., Roumestand, C., Toma, F., and Ménez, A. (1995) Scorpion toxins as natural scaffolds for protein engineering. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6404-6408
121. Vita, C., Drakopoulou, E., Vizzavona, J., Rochette, S., Martin, L., Ménez, a., Roumestand, C., Yang, Y. S., Ylisastigui, L., Benjouad, a., and Gluckman, J. C. (1999) Rational engineering of a miniprotein that reproduces the core of the CD4 site interacting with HIV-1 envelope glycoprotein. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 13091-13096
122. Veiseh, M., Gabikian, P., Bahrami, S.-B., Veiseh, O., Zhang, M., Hackman, R. C., Ravanpay, A. C., Stroud, M. R., Kusuma, Y., Hansen, S. J., Kwok, D., Munoz, N. M., Sze, R. W., Grady, W. M., Greenberg, N. M., Ellenbogen, R. G., and Olson, J. M. (2007) Tumor paint: a chlorotoxin: Cy5.5 bioconjugate for intraoperative visualization of cancer foci. *Cancer Res.* **67**, 6882-6888
123. Gunasekera, S., Foley, F. M., Clark, R. J., Sando, L., Fabri, L. J., Craik, D. J., and Daly, N. L. (2008) Engineering stabilized vascular endothelial growth factor-A antagonists: synthesis, structural characterization, and bioactivity of grafted analogues of cyclotides. *J. Med. Chem.* **51**, 7697-7704
124. Aboye, T. L., Ha, H., Majumder, S., Christ, F., Debyser, Z., Shekhtman, A., Neamati, N., and Camarero, J. a. (2012) Design of a novel cyclotide-based CXCR4 antagonist with anti-human immunodeficiency virus (HIV)-1 activity. *J. Med. Chem.* **55**, 10729-10734
125. Wong, C. T. T., Rowlands, D. K., Wong, C.-H., Lo, T. W. C., Nguyen, G. K. T., Li, H.-Y., and Tam, J. P. (2012) Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment. *Angew. Chem., Int. Ed.* **51**, 5620-5624
126. Koehbach, J., O'Brien, M., Muttenthaler, M., Miazzi, M., Akcan, M., Elliott, A. G., Daly, N. L., Harvey, P. J., Arrowsmith, S., Gunasekera, S., Smith, T. J., Wray, S., Goransson, U., Dawson, P. E., Craik, D. J., Freissmuth, M., and Gruber, C. W. (2013) Oxytocic plant cyclotides as templates for peptide G protein-coupled receptor ligand design. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 21183-21188
127. Martin, L., Stricher, F., Misse, D., Sironi, F., Pugniere, M., Barthe, P., Prado-Gotor, R., Freulon, I., Magne, X., Roumestand, C., Menez, A., Lusso, P., Veas, F., and Vita, C. (2003) Rational design of a CD4 mimic that inhibits HIV-1 entry and exposes cryptic neutralization epitopes. *Nat. Biotechnol.* **21**, 71-76
128. Eliassen, R., Daly, N. L., Wulff, B. S., Andresen, T. L., Conde-Frieboes, K. W., and Craik, D. J. (2012) Design, synthesis, structural and functional characterization of novel melanocortin agonists based on the cyclotide kalata B1. *J. Biol. Chem.* **287**, 40493-40501

129. Getz, J. A., Cheneval, O., Craik, D. J., and Daugherty, P. S. (2013) Design of a cyclotide antagonist of neuropilin-1 and -2 that potently inhibits endothelial cell migration. *ACS Chem. Biol.* **8**, 1147-1154
130. Wang, C. K., Gruber, C. W., Cemazar, M., Siatskas, C., Tagore, P., Payne, N., Sun, G., Wang, S., Bernard, C. C., and Craik, D. J. (2013) Molecular grafting onto a stable framework yields novel cyclic peptides for the treatment of multiple sclerosis. *ACS Chem. Biol.* **9**, 156-163
131. Ji, Y., Majumder, S., Millard, M., Borra, R., Bi, T., Elnagar, A. Y., Neamati, N., Shekhtman, A., and Camarero, J. A. (2013) *In vivo* activation of the p53 tumor suppressor pathway by an engineered cyclotide. *J. Am. Chem. Soc.* **135**, 11623-11633
132. Sommerhoff, C. P., Avrutina, O., Schmoldt, H.-U., Gabrijelcic-Geiger, D., Diederichsen, U., and Kolmar, H. (2010) Engineered cystine knot miniproteins as potent inhibitors of human mast cell tryptase β . *J. Mol. Biol.* **395**, 167-175
133. Kimura, R. H., Teed, R., Hackel, B. J., Pysz, M. A., Chuang, C. Z., Sathirachinda, A., Willmann, J. K., and Gambhir, S. S. (2012) Pharmacokinetically stabilized cystine knot peptides that bind α -v- β -6 integrin with single-digit nanomolar affinities for detection of pancreatic cancer. *Clin. Cancer Res.* **18**, 839-849
134. Getz, J. A., Rice, J. J., and Daugherty, P. S. (2011) Protease-resistant peptide ligands from a knottin scaffold library. *ACS Chem. Biol.* **6**, 837-844
135. Conibear, A. C., Bochen, A., Rosengren, K. J., Stupar, P., Wang, C., Kessler, H., and Craik, D. J. (2014) The cyclic cystine ladder of theta-defensins as a stable, bifunctional scaffold: a proof-of-concept study using the integrin-binding RGD motif. *ChemBioChem* **15**, 451-459
136. Marceau, F., and Regoli, D. (2004) Bradykinin receptor ligands: therapeutic perspectives. *Nat. Rev. Drug Discov.* **3**, 845-852
137. Haase, J., and Lanka, E. (1997) A specific protease encoded by the conjugative DNA transfer systems of IncP and Ti plasmids is essential for pilus synthesis. *J. Bacteriol.* **179**, 5728-5735
138. Lee, J., McIntosh, J., Hathaway, B. J., and Schmidt, E. W. (2009) Using marine natural products to discover a protease that catalyzes peptide macrocyclization of diverse substrates. *J. Am. Chem. Soc.* **131**, 2122-2124
139. Barber, C. J. S., Pujara, P. T., Reed, D. W., Chiwocha, S., Zhang, H., and Covello, P. S. (2013) The two-step biosynthesis of cyclic peptides from linear precursors in a member of the plant family Caryophyllaceae Involves cyclization by a serine protease-like enzyme. *J. Biol. Chem.* **288**, 12500-12510
140. Jackson, D. Y., Burnier, J. P., and Wells, J. A. (1995) Enzymic cyclization of linear peptide esters using subtiligase. *J. Am. Chem. Soc.* **117**, 819-820
141. Aboye, T. L., and Camarero, J. A. (2012) Biological synthesis of circular polypeptides. *J. Biol. Chem.* **287**, 27026-27032

142. Antos, J. M., Popp, M. W., Ernst, R., Chew, G. L., Spooner, E., and Ploegh, H. L. (2009) A straight path to circular proteins. *J. Biol. Chem.* **284**, 16028-16036
143. Perler, F. B. (2002) InBase: the intein database. *Nucl. Acids Res.* **30**, 383-384
144. Mathys, S., Evans Jr, T. C., Chute, I. C., Wu, H., Chong, S., Benner, J., Liu, X.-Q., and Xu, M.-Q. (1999) Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene* **231**, 1-13
145. Xu, M. Q., and Perler, F. B. (1996) The mechanism of protein splicing and its modulation by mutation. *Embo J* **15**, 5146-5153
146. Xu, M. Q., Southworth, M. W., Mersha, F. B., Hornstra, L. J., and Perler, F. B. (1993) In vitro protein splicing of purified precursor and the identification of a branched intermediate. *Cell* **75**, 1371-1377
147. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6705-6710
148. Camarero, J. A., and Muir, T. W. (1999) Biosynthesis of a head-to-tail cyclized protein with improved biological activity. *J. Am. Chem. Soc.* **121**, 5597-5598
149. Tam, J. P., and Lu, Y. A. (1998) A biomimetic strategy in the synthesis and fragmentation of cyclic protein. *Protein Sci.* **7**, 1583-1592
150. Tam, J. P., Lu, Y. A., and Yu, Q. T. (1999) Thia zip reaction for synthesis of large cyclic peptides: mechanisms and applications. *J. Am. Chem. Soc.* **121**, 4316-4324
151. Kimura, R. H., Tran, A.-T., and Camarero, J. A. (2006) Biosynthesis of the cyclotide kalata B1 by using protein splicing. *Angew. Chem., Int. Ed.* **45**, 973-976
152. Camarero, J. A., Kimura, R. H., Woo, Y.-H., Shekhtman, A., and Cantor, J. (2007) Biosynthesis of a fully functional cyclotide inside living bacterial cells. *ChemBioChem* **8**, 1363-1366
153. Gould, A., Li, Y., Majumder, S., Garcia, A. E., Carlsson, P., Shekhtman, A., and Camarero, J. A. (2012) Recombinant production of rhesus θ -defensin-1 (RTD-1) using a bacterial expression system. *Mol. BioSyst.* **8**, 1359-1365
154. Austin, J., Kimura, R., Woo, Y.-H., and Camarero, J. (2010) In vivo biosynthesis of an Ala-scan library based on the cyclic peptide SFTI-1. *Amino Acids* **38**, 1313-1322
155. Ton-That, H., Liu, G., Mazmanian, S. K., Faull, K. F., and Schneewind, O. (1999) Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12424-12429
156. Mazmanian, S. K., Liu, G., Ton-That, H., and Schneewind, O. (1999) *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285**, 760-763

157. Guimaraes, C. P., Witte, M. D., Theile, C. S., Bozkurt, G., Kundrat, L., Blom, A. E., and Ploegh, H. L. (2013) Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nat Protoc* **8**, 1787-1799
158. Popp, M. W., and Ploegh, H. L. (2011) Making and breaking peptide bonds: protein engineering using sortase. *Angew. Chem., Int. Ed.* **50**, 5024-5032
159. Strijbis, K., Spooner, E., and Ploegh, H. L. (2012) Protein ligation in living cells using sortase. *Traffic* **13**, 780-789
160. Jia, X., Kwon, S., Wang, C.-I. A., Huang, Y.-H., Chan, L. Y., Tan, C. C., Rosengren, K. J., Mulvenna, J. P., Schroeder, C. I., and Craik, D. J. (2014) Semiozymatic cyclization of disulfide-rich peptides using sortase A. *J. Biol. Chem.* **289**, 6627-6638
161. Li, Y. M., Li, Y. T., Pan, M., Kong, X. Q., Huang, Y. C., Hong, Z. Y., and Liu, L. (2014) Irreversible site-specific hydrazinolysis of proteins by use of sortase. *Angew. Chem., Int. Ed.* **53**, 2198-2202
162. Ling, J. J., Policarpo, R. L., Rabideau, A. E., Liao, X., and Pentelute, B. L. (2012) Protein thioester synthesis enabled by sortase. *J. Am. Chem. Soc.* **134**, 10749-10752
163. du Vigneaud, V., Ressler, C., Swan, C. J. M., Roberts, C. W., Katsoyannis, P. G., and Gordon, S. (1953) The synthesis of an octapeptide amide with the hormonal activity of oxytocin. *J. Am. Chem. Soc.* **75**, 4879-4880
164. du Vigneaud, V., Ressler, C., Swan, J. M., Roberts, C. W., and Katsoyannis, P. G. (1954) The synthesis of oxytocin. *J. Am. Chem. Soc.* **76**, 3115-3121
165. Merrifield, R. B. (1963) Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149-2154
166. Albericio, F., and Carpino, L. A. (1997) [7] Coupling reagents and activation. in *Meth. Enzymol.* (Gregg, B. F. ed.), Academic Press. pp 104-126
167. Atherton, E., Fox, H., Harkiss, D., Logan, C. J., Sheppard, R. C., and Williams, B. J. (1978) A mild procedure for solid phase peptide synthesis: use of fluorenylmethoxycarbonylamino-acids. *J. Chem. Soc., Chem. Commun.*, 537-539
168. Atherton, E., Bury, C., Sheppard, R. C., and Williams, B. J. (1979) Stability of fluorenylmethoxycarbonylamino groups in peptide synthesis. Cleavage by hydrogenolysis and by dipolar aprotic solvents. *Tetrahedron Lett.* **20**, 3041-3042
169. Carpino, L. A., and Han, G. Y. (1972) 9-Fluorenylmethoxycarbonyl amino-protecting group. *J. Org. Chem.* **37**, 3404-3409
170. Kemp, D. S. (1981) The amine capture strategy for peptide bond formation—an outline of progress. *Biopolymers* **20**, 1793-1804
171. Liu, C.-F., and Tam, J. P. (1994) Peptide segment ligation strategy without use of protecting groups. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6584-6588

172. Tam, J. P., Lu, Y. A., Liu, C. F., and Shao, J. (1995) Peptide synthesis using unprotected peptides through orthogonal coupling methods. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 12485-12489
173. Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779
174. Wieland, T., Bokelmann, E., Bauer, L., Lang, H. U., and Lau, H. (1953) Über Peptidsynthesen. 8. Mitteilung Bildung von S-haltigen Peptiden durch intramolekulare Wanderung von Aminoacylresten. *Justus Liebigs Ann. Chem.* **583**, 129-149
175. Johnson, E. C. B., and Kent, S. B. H. (2006) Insights into the mechanism and catalysis of the native chemical ligation reaction. *J. Am. Chem. Soc.* **128**, 6640-6646
176. Hackeng, T. M., Griffin, J. H., and Dawson, P. E. (1999) Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10068-10073
177. Pollock, S. B., and Kent, S. B. (2011) An investigation into the origin of the dramatically reduced reactivity of peptide-prolyl-thioesters in native chemical ligation. *Chem. Commun.* **47**, 2342-2344
178. Tam, J. P., and Lu, Y.-A. (1997) Synthesis of large cyclic cystine-knot peptide by orthogonal coupling strategy using unprotected peptide precursor. *Tetrahedron Lett.* **38**, 5599-5602
179. Camarero, J. A., and Muir, T. W. (1997) Chemoselective backbone cyclization of unprotected peptides. *Chem. Commun.*, 1369-1370
180. Chiche, L., Heitz, A., Gelly, J.-C., Gracy, J., Chau, P. T., Ha, P. T., Hernandez, J.-F., and Le-Nguyen, D. (2004) Squash inhibitors: from structural motifs to macrocyclic knottins. *Curr. Protein Pept. Sci.* **5**, 341-349
181. Nguyen, D. L., Barry, L. G., Tam, J. P., Heitz, A., Chiche, L., Hernandez, J.-F., and Pham, T.-C. (2002) Synthesis of MCoTI-I, a cyclic trypsin inhibitor from *Momordica cochinchinensis*. in *Peptides 2002, Proceedings of the 27th European Peptide Symposium* (Benedetti, E., and Pedone, C. eds.), Edizioni Ziino, Sorrento, Italy, 182-183
182. Clark, R. J., and Craik, D. J. (2010) Native chemical ligation applied to the synthesis and bioengineering of circular peptides and proteins. *Biopolymers* **94**, 414-422
183. Thongyoo, P., Tate, E. W., and Leatherbarrow, R. J. (2006) Total synthesis of the macrocyclic cysteine knot microprotein MCoTI-II. *Chem Commun (Camb)*, 2848-2850
184. Daly, N. L., Love, S., Alewood, P. F., and Craik, D. J. (1999) Chemical synthesis and folding pathways of large cyclic polypeptides: studies of the cystine knot polypeptide kalata B1. *Biochemistry* **38**, 10606-10614

185. Leta Aboye, T., Clark, R. J., Craik, D. J., and Göransson, U. (2008) Ultra-stable peptide scaffolds for protein engineering—synthesis and folding of the circular cystine knotted cyclotide cycloviolacin O2. *ChemBioChem* **9**, 103-113
186. Zheng, J.-S., Tang, S., Guo, Y., Chang, H.-N., and Liu, L. (2012) Synthesis of cyclic peptides and cyclic proteins via ligation of peptide hydrazides. *ChemBioChem* **13**, 542-546
187. Cowper, B., Craik, D. J., and Macmillan, D. (2013) Making ends meet: chemically mediated circularization of recombinant proteins. *ChemBioChem* **14**, 809-812
188. Gunasekera, S., Aboye, T., Madian, W., El-Seedi, H., and Göransson, U. (2013) Making ends meet: microwave-accelerated synthesis of cyclic and disulfide rich proteins via *in situ* thioesterification and native chemical ligation. *Int. J. Pept. Res. Ther.* **19**, 43-54
189. Tam, J. P., and Wong, C. T. T. (2012) Chemical synthesis of circular proteins. *J. Biol. Chem.* **287**, 27020-27025
190. Liu, C.-F., and Tam, J. P. (1994) Chemical ligation approach to form a peptide bond between unprotected peptide segments. Concept and model study. *J. Am. Chem. Soc.* **116**, 4149-4153
191. Zhang, L., and Tam, J. P. (1997) Orthogonal coupling of unprotected peptide segments through histidyl amino terminus. *Tetrahedron Lett.* **38**, 3-6
192. Tam, J. P., and Yu, Q. (1998) Methionine ligation strategy in the biomimetic synthesis of parathyroid hormones. *Biopolymers* **46**, 319-327
193. Yan, L. Z., and Dawson, P. E. (2001) Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **123**, 526-533
194. Crich, D., and Banerjee, A. (2007) Native chemical ligation at phenylalanine. *J. Am. Chem. Soc.* **129**, 10064-10065
195. Haase, C., Rohde, H., and Seitz, O. (2008) Native chemical ligation at valine. *Angew. Chem., Int. Ed.* **47**, 6807-6810
196. Chen, J., Wan, Q., Yuan, Y., Zhu, J., and Danishefsky, S. J. (2008) Native chemical ligation at valine: a contribution to peptide and glycopeptide synthesis. *Angew. Chem., Int. Ed.* **47**, 8521-8524
197. Brik, A., Yang, Y.-Y., Ficht, S., and Wong, C.-H. (2006) Sugar-assisted glycopeptide ligation. *J. Am. Chem. Soc.* **128**, 5626-5627
198. Johnson, E. C. B., Malito, E., Shen, Y., Pentelute, B., Rich, D., Florián, J., Tang, W.-J., and Kent, S. B. H. (2007) Insights from atomic-resolution X-ray structures of chemically synthesized HIV-1 protease in complex with inhibitors. *J. Mol. Biol.* **373**, 573-586

199. Nakamura, K., Kanao, T., Uesugi, T., Hara, T., Sato, T., Kawakami, T., and Aimoto, S. (2009) Synthesis of peptide thioesters via an N-S acyl shift reaction under mild acidic conditions on an N-4,5-dimethoxy-2-mercaptobenzyl auxiliary group. *J. Pept. Sci.* **15**, 731-737
200. Yang, R., Pasunooti, K. K., Li, F., Liu, X.-W., and Liu, C.-F. (2009) Dual native chemical ligation at lysine. *J. Am. Chem. Soc.* **131**, 13592-13593
201. Li, X., Lam, H. Y., Zhang, Y., and Chan, C. K. (2010) Salicylaldehyde ester-induced chemoselective peptide ligations: enabling generation of natural peptidic linkages at the serine/threonine sites. *Org. Lett.* **12**, 1724-1727
202. Wong, C. T. T., Lam, H. Y., Song, T., Chen, G., and Li, X. (2013) Synthesis of constrained head-to-tail cyclic tetrapeptides by an imine-induced ring-closing/contraction strategy. *Angew. Chem., Int. Ed.* **52**, 10212-10215
203. Zhao, J.-F., Zhang, X.-H., Ding, Y.-J., Yang, Y.-S., Bi, X.-B., and Liu, C.-F. (2013) Facile synthesis of peptidyl salicylaldehyde esters and its use in cyclic peptide synthesis. *Org. Lett.* **15**, 5182-5185
204. Zhang, L., and Tam, J. P. (1997) Metal ion-assisted peptide cyclization. *Tetrahedron Lett.* **38**, 4375-4378
205. Saxon, E., and Bertozzi, C. R. (2000) Cell surface engineering by a modified Staudinger reaction. *Science* **287**, 2007-2010
206. Nilsson, B. L., Kiessling, L. L., and Raines, R. T. (2000) Staudinger ligation: a peptide from a thioester and azide. *Org. Lett.* **2**, 1939-1941
207. Taichi, M., Hemu, X., Qiu, Y., and Tam, J. P. (2013) A thioethylalkylamido (TEA) thioester surrogate in the synthesis of a cyclic peptide via a tandem acyl shift. *Org. Lett.* **15**, 2620-2623
208. Hojo, H., and Aimoto, S. (1991) Polypeptide synthesis using the S-alkyl thioester of a partially protected peptide segment. Synthesis of the DNA-binding domain of c-Myb protein (142-193)-NH₂. *Bull. Chem. Soc. Jpn.* **64**, 111-117
209. Zhang, L., and Tam, J. P. (1999) Lactone and lactam library synthesis by silver ion-assisted orthogonal cyclization of unprotected peptides. *J. Am. Chem. Soc.* **121**, 3311-3320
210. Aimoto, S. (1999) Polypeptide synthesis by the thioester method. *Pept. Sci.* **51**, 247-265
211. Clippingdale, A. B., Barrow, C. J., and Wade, J. D. (2000) Peptide thioester preparation by Fmoc solid phase peptide synthesis for use in native chemical ligation. *J. Pept. Sci.* **6**, 225-234
212. Swinnen, D., and Hilvert, D. (2000) Facile, fmoc-compatible solid-phase synthesis of peptide C-terminal thioesters. *Org. Lett.* **2**, 2439-2442

213. Sewing, A., and Hilvert, D. (2001) Fmoc-compatible solid-phase peptide synthesis of long C-terminal peptide thioesters. *Angew. Chem., Int. Ed.* **40**, 3395-3396
214. Raz, R., and Rademann, J. r. (2011) Fmoc-based synthesis of peptide thioesters for native chemical ligation employing a *tert-butyl* thiol linker. *Org. Lett.* **13**, 1606-1609
215. Raz, R., and Rademann, J. (2012) Fmoc-based synthesis of peptide thioacids for azide ligations via 2-cyanoethyl thioesters. *Org. Lett.* **14**, 5038-5041
216. Kenner, G. W., McDermott, J. R., and Sheppard, R. C. (1971) The safety catch principle in solid phase peptide synthesis. *J. Chem. Soc. D: Chem. Commun.*, 636-637
217. Backes, B. J., Virgilio, A. A., and Ellman, J. A. (1996) Activation method to prepare a highly reactive acylsulfonamide "Safety-Catch" linker for solid-phase synthesis. *J. Am. Chem. Soc.* **118**, 3055-3056
218. Backes, B. J., and Ellman, J. A. (1999) An alkanesulfonamide "Safety-Catch" linker for solid-phase synthesis. *J. Org. Chem.* **64**, 2322-2330
219. Heidler, P., and Link, A. (2005) N-acyl-N-alkyl-sulfonamide anchors derived from Kenner's safety-catch linker: powerful tools in bioorganic and medicinal chemistry. *Bioorg. Med. Chem.* **13**, 585-599
220. Ingenito, R., Bianchi, E., Fattori, D., and Pessi, A. (1999) Solid phase synthesis of peptide C-terminal thioesters by Fmoc/t-Bu chemistry. *J. Am. Chem. Soc.* **121**, 11369-11374
221. Camarero, J. A., Hackel, B. J., de Yoreo, J. J., and Mitchell, A. R. (2004) Fmoc-based synthesis of peptide α -thioesters using an aryl hydrazine support. *J. Org. Chem.* **69**, 4145-4151
222. Blanco-Canosa, J. B., and Dawson, P. E. (2008) An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angew. Chem., Int. Ed.* **47**, 6851-6855
223. Tofteng, A. P., Sørensen, K. K., Conde-Frieboes, K. W., Hoeg-Jensen, T., and Jensen, K. J. (2009) Fmoc solid-phase synthesis of C-terminal peptide thioesters by formation of a backbone pyroglutamyl imide moiety. *Angew. Chem., Int. Ed.* **48**, 7411-7414
224. Fang, G.-M., Li, Y.-M., Shen, F., Huang, Y.-C., Li, J.-B., Lin, Y., Cui, H.-K., and Liu, L. (2011) Protein chemical synthesis by ligation of peptide hydrazides. *Angew. Chem., Int. Ed.* **50**, 7645-7649
225. Okamoto, R., Morooka, K., and Kajihara, Y. (2012) A synthetic approach to a peptide α -thioester from an unprotected peptide through cleavage and activation of a specific peptide bond by N-acetylguanidine. *Angew. Chem., Int. Ed.* **51**, 191-196
226. Warren, J. D., Miller, J. S., Keding, S. J., and Danishefsky, S. J. (2004) Toward fully synthetic glycoproteins by ultimately convergent routes: a solution to a long-standing problem. *J. Am. Chem. Soc.* **126**, 6576-6578

227. Botti, P., Villain, M., Manganiello, S., and Gaertner, H. (2004) Native chemical ligation through *in situ* O to S acyl shift. *Org. Lett.* **6**, 4861-4864
228. Eom, K. D., and Tam, J. P. (2011) Acid-catalyzed tandem thiol switch for preparing Peptide thioesters from mercaptoethyl esters. *Org. Lett.* **13**, 2610-2613
229. Tofteng, A. P., Jensen, K. J., and Hoeg-Jensen, T. (2007) Peptide dithiodiethanol esters for in situ generation of thioesters for use in native ligation. *Tetrahedron Lett.* **48**, 2105-2107
230. Zheng, J.-S., Xi, W.-X., Wang, F.-L., Li, J., and Guo, Q.-X. (2011) Fmoc-SPPS chemistry compatible approach for the generation of (glyco)peptide aryl thioesters. *Tetrahedron Lett.* **52**, 2655-2660
231. Vizzavona, J., Dick, F., and Vorherr, T. (2002) Synthesis and application of an auxiliary group for chemical ligation at the X-gly site. *Bioorg. Med. Chem. Lett.* **12**, 1963-1965
232. Kawakami, T., Sumida, M., Nakamura, K. i., Vorherr, T., and Aimoto, S. (2005) Peptide thioester preparation based on an N-S acyl shift reaction mediated by a thiol ligation auxiliary. *Tetrahedron Lett.* **46**, 8805-8807
233. Wu, B., Chen, J., Warren, J. D., Chen, G., Hua, Z., and Danishefsky, S. J. (2006) Building complex glycopeptides: development of a cysteine-free native chemical ligation protocol. *Angew. Chem., Int. Ed.* **45**, 4116-4125
234. Nakamura, K., Sumida, M., Kawakami, T., Vorherr, T., and Aimoto, S. (2006) Generation of an S-peptide via an N-S acyl shift reaction in a TFA solution. *Bull. Chem. Soc. Jpn.* **79**, 1773-1780
235. Ollivier, N., Behr, J.-B., El-Mahdi, O., Blanpain, A., and Melnyk, O. (2005) Fmoc solid-phase synthesis of peptide thioesters using an intramolecular N,S-acyl shift. *Org. Lett.* **7**, 2647-2650
236. Nagaike, F., Onuma, Y., Kanazawa, C., Hojo, H., Ueki, A., Nakahara, Y., and Nakahara, Y. (2006) Efficient microwave-assisted tandem N- to S-acyl transfer and thioester exchange for the preparation of a glycosylated peptide thioester. *Org. Lett.* **8**, 4465-4468
237. Ohta, Y., Itoh, S., Shigenaga, A., Shintaku, S., Fujii, N., and Otaka, A. (2006) Cysteine-derived S-protected oxazolidinones: potential chemical devices for the preparation of peptide thioesters. *Org. Lett.* **8**, 467-470
238. Hojo, H., Onuma, Y., Akimoto, Y., Nakahara, Y., and Nakahara, Y. (2007) N-alkyl cysteine-assisted thioesterification of peptides. *Tetrahedron Lett.* **48**, 25-28
239. Kawakami, T., and Aimoto, S. (2007) Sequential peptide ligation by using a controlled cysteinyl prolyl ester (CPE) autoactivating unit. *Tetrahedron Lett.* **48**, 1903-1905

240. Tsuda, S., Shigenaga, A., Bando, K., and Otaka, A. (2009) N→S acyl-transfer-mediated synthesis of peptide thioesters using anilide derivatives. *Org. Lett.* **11**, 823-826
241. Hou, W., Zhang, X., Li, F., and Liu, C.-F. (2011) Peptidyl N,N-bis(2-mercaptoethyl)-amides as thioester precursors for native chemical ligation. *Org. Lett.* **13**, 386-389
242. Ollivier, N., Dheur, J., Mhidia, R., Blanpain, A., and Melnyk, O. (2010) Bis(2-sulfanylethyl)amino native peptide ligation. *Org. Lett.* **12**, 5238-5241
243. Sharma, R. K., and Tam, J. P. (2011) Tandem thiol switch synthesis of peptide thioesters via N-S acyl shift on thiazolidine. *Org. Lett.* **13**, 5176-5179
244. Zheng, J.-S., Chang, H.-N., Wang, F.-L., and Liu, L. (2011) Fmoc synthesis of peptide thioesters without post-chain-assembly manipulation. *J. Am. Chem. Soc.* **133**, 11080-11083
245. Yang, R., Qi, L., Liu, Y., Ding, Y., Kwek, M. S. Y., and Liu, C.-F. (2013) Chemical synthesis of N-peptidyl 2-pyrrolidinemethanethiol for peptide ligation. *Tetrahedron Lett.* **54**, 3777-3780
246. Hojo, H., Murasawa, Y., Katayama, H., Ohira, T., and Nakahara, Y. (2008) Application of a novel thioesterification reaction to the synthesis of chemokine CCL27 by the modified thioester method. *Org. Biomol. Chem.* **6**, 1808-1813
247. Erlich, L. A., Kumar, K. S. A., Haj-Yahya, M., Dawson, P. E., and Brik, A. (2010) N-methylcysteine-mediated total chemical synthesis of ubiquitin thioester. *Org. Biomol. Chem.* **8**, 2392-2396
248. Kawakami, T., and Aimoto, S. (2009) The use of a cysteinyl prolyl ester (CPE) autoactivating unit in peptide ligation reactions. *Tetrahedron* **65**, 3871-3877
249. Kawakami, T., Kamauchi, A., Harada, E., and Aimoto, S. (2014) Enhancement in the rate of conversion of peptide Cys-Pro esters to peptide thioesters by structural modification. *Tetrahedron Lett.* **55**, 79-81
250. Dheur, J., Ollivier, N., and Melnyk, O. (2011) Synthesis of thiazolidine thioester peptides and acceleration of native chemical ligation. *Org. Lett.* **13**, 1560-1563
251. Dheur, J., Ollivier, N., Vallin, A. I., and Melnyk, O. (2011) Synthesis of peptide alkylthioesters using the intramolecular N,S-acyl shift properties of bis(2-sulfanylethyl)amido peptides. *J. Org. Chem.* **76**, 3194-3202
252. Boll, E., Dheur, J., Drobecq, H., and Melnyk, O. (2012) Access to cyclic or branched peptides using bis(2-sulfanylethyl)amido side-chain derivatives of Asp and Glu. *Org. Lett.* **14**, 2222-2225
253. Kang, J., Reynolds, N. L., Tyrrell, C., Dorin, J. R., and Macmillan, D. (2009) Peptide thioester synthesis through N→S acyl-transfer: application to the synthesis of a beta-defensin. *Org. Biomol. Chem.* **7**, 4918-4923

254. Kang, J., Richardson, J. P., and Macmillan, D. (2009) 3-Mercaptopropionic acid-mediated synthesis of peptide and protein thioesters. *Chem. Commun.*, 407-409
255. Richardson, J. P., Chan, C.-H., Blanc, J., Saadi, M., and Macmillan, D. (2010) Exploring neoglycoprotein assembly through native chemical ligation using neoglycopeptide thioesters prepared via N→S acyl transfer. *Org. Biomol. Chem.* **8**, 1351-1360
256. Adams, A. L., Cowper, B., Morgan, R. E., Premdjee, B., Caddick, S., and Macmillan, D. (2013) Cysteine promoted C-terminal hydrazinolysis of native peptides and proteins. *Angew. Chem., Int. Ed.* **52**, 13062-13066
257. Burlina, F., Papageorgiou, G., Morris, C., White, P. D., and Offer, J. (2014) In situ thioester formation for protein ligation using α -methylcysteine. *Chem. Sci.* **5**, 766-770
258. Anfinsen, C. B., Haber, E., Sela, M., and White, F. H. (1961) The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U.S.A.* **47**, 1309-1314
259. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* **181**, 223-230
260. Tam, J. P., Wu, C. R., Liu, W., and Zhang, J. W. (1991) Disulfide bond formation in peptides by dimethyl sulfoxide. Scope and applications. *J. Am. Chem. Soc.* **113**, 6657-6662
261. Besse, D., Siedler, F., Diercks, T., Kessler, H., and Moroder, L. (1997) The redox potential of selenocysteine in unconstrained cyclic peptides. *Angew. Chem., Int. Ed.* **36**, 883-885
262. Wong, C. T. T., Taichi, M., Nishio, H., Nishiuchi, Y., and Tam, J. P. (2011) Optimal oxidative folding of the novel antimicrobial cyclotide from *Hedyotis biflora* requires high alcohol concentrations. *Biochemistry* **50**, 7275-7283
263. Chino, N., Kubo, S., Nishio, H., Nishiuchi, Y., Nakazato, M., and Kimura, T. (2006) Chemical synthesis of human β -defensin (hBD)-1, -2, -3 and -4: optimization of the oxidative folding reaction. *Int. J. Pept. Res. Ther.* **12**, 203-209
264. Kubo, S., Tanimura, K., Nishio, H., Chino, N., Teshima, T., Kimura, T., and Nishiuchi, Y. (2008) Optimization of the oxidative folding reaction and disulfide structure determination of human α -defensin 1, 2, 3 and 5. *Int. J. Pept. Res. Ther.* **14**, 341-349
265. Kubo, S., Chino, N., Kimura, T., and Sakakibara, S. (1996) Oxidative folding of ω -conotoxin MVIIC: effects of temperature and salt. *Biopolymers* **38**, 733-744
266. Nielsen, J. S., Buczek, P., and Bulaj, G. (2004) Cosolvent-assisted oxidative folding of a bicyclic α -conotoxin ImI. *J. Pept. Sci.* **10**, 249-256
267. Steiner, A. M., and Bulaj, G. (2011) Optimization of oxidative folding methods for cysteine-rich peptides: a study of conotoxins containing three disulfide bridges. *J. Pept. Sci.* **17**, 1-7

268. Welker, E. (2001) Structural determinants of oxidative folding in proteins. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2312-2316
269. Weissman, J. S., and Kim, P. S. (1991) Reexamination of the folding of BPTI: predominance of native intermediates. *Science* **253**, 1386-1393
270. Kibria, F. M., and Lees, W. J. (2008) Balancing conformational and oxidative kinetic traps during the folding of bovine pancreatic trypsin inhibitor (BPTI) with glutathione and glutathione disulfide. *J. Am. Chem. Soc.* **130**, 796-797
271. Wedemeyer, W. J., Welker, E., Narayan, M., and Scheraga, H. A. (2000) Disulfide bonds and protein folding. *Biochemistry* **39**, 4207-4216
272. Chatrenet, B., and Chang, J. Y. (1993) The disulfide folding pathway of hirudin elucidated by stop/go folding experiments. *J. Biol. Chem.* **268**, 20988-20996
273. Miloslavina, A. A., Leipold, E., Kijas, M., Stark, A., Heinemann, S. H., and Imhof, D. (2009) A room temperature ionic liquid as convenient solvent for the oxidative folding of conopeptides. *J. Pept. Sci.* **15**, 72-77
274. Kubo, S., Chino, N., Nakajima, K., Aumelas, A., Chiche, L., Segawa, S.-i., Tamaoki, H., Kobayashi, Y., Kimura, T., and Sakakibara, S. (1997) Improvement in the oxidative folding of endothelin-1 by a Lys-Arg extension at the amino terminus: Implication of a salt bridge between Arg-1 and Asp8. *Lett Pept Sci* **4**, 185-192
275. Tam, J. P., Heath, W. F., and Merrifield, R. B. (1986) Mechanisms for the removal of benzyl protecting groups in synthetic peptides by trifluoromethanesulfonic acid-trifluoroacetic acid-dimethyl sulfide. *J. Am. Chem. Soc.* **108**, 5242-5251
276. Bonvicini, P., Levi, A., Lucchini, V., Modena, G., and Scorrano, G. (1973) Acid-base behavior of alkyl sulfur and oxygen bases. *J. Am. Chem. Soc.* **95**, 5960-5964
277. Bonvicini, P., Levi, A., Lucchini, V., and Scorrano, G. (1972) The acid-base behaviour of sulphides. *J. Chem. Soc., Perkin Trans. 2*, 2267-2269
278. Taichi, M., Yamazaki, T., Kimura, T., and Nishiuchi, Y. (2009) Total synthesis of marinostatin, a serine protease inhibitor isolated from the marine bacterium *Pseudoalteromonas sagamiensis*. *Tetrahedron Lett.* **50**, 2377-2380
279. Khosla, M. C., Smeby, R. R., and Bumpus, F. M. (1972) Failure sequence in solid-phase peptide synthesis due to the presence of an N-alkylamino acid. *J. Am. Chem. Soc.* **94**, 4721-4724
280. Lukszo, J., Patterson, D., Albericio, F., and Kates, S. A. (1996) 3-(1-Piperidiny)alanine formation during the preparation of C-terminal cysteine peptides with the Fmoc/t-Bu strategy. *Lett Pept Sci* **3**, 157-166
281. Otaka, A., Sato, K., Ding, H., and Shigenaga, A. (2012) One-pot/sequential native chemical ligation using *N*-sulfanylethylanilide peptide. *Chem. Rec.* **12**, 479-490
282. Hemu, X., Taichi, M., Qiu, Y., Liu, D. X., and Tam, J. P. (2013) Biomimetic synthesis of cyclic peptides using novel thioester surrogates. *Pept. Sci.* **100**, 492-501

283. Hyman, H. H., and Garber, R. A. (1959) The Hammett acidity function H_0 for trifluoroacetic acid solutions of sulfuric and hydrofluoric acids. *J. Am. Chem. Soc.* **81**, 1847-1849
284. McDermott, J. R., and Benoiton, N. L. (1973) N-methylamino acids in peptide synthesis. III. Racemization during deprotection by saponification and acidolysis. *Can. J. Chem.* **51**, 2555-2561
285. Urban, J. A. N., Vaisar, T., Shen, R., and Lee, M. S. (1996) Lability of N-alkylated peptides towards TFA cleavage. *Int. J. Pept. Protein Res.* **47**, 182-189
286. Polfer, N. C., Oomens, J., Suhai, S., and Paizs, B. (2005) Spectroscopic and theoretical evidence for oxazolone ring formation in collision-induced dissociation of peptides. *J. Am. Chem. Soc.* **127**, 17154-17155
287. Teixido, M., Albericio, F., and Giralt, E. (2005) Solid-phase synthesis and characterization of N-methyl-rich peptides. *J Pept Res* **65**, 153-166
288. Curtius, T. (1904) Verkettung von amidosäuren I. abhandlung. *J. Prakt. Chem.* **70**, 57-72
289. Curtius, T. (1902) Synthetische versuche mit hippurazid. *Ber. Dtsch. Chem. Ges.* **35**, 3226-3228
290. Curtius, T. (1890) Ueber Stickstoffwasserstoffsäure (Azoimid) N₃H. *Ber. Dtsch. Chem. Ges.* **23**, 3023-3033
291. Kahne, D., and Still, W. C. (1988) Hydrolysis of a peptide bond in neutral water. *J. Am. Chem. Soc.* **110**, 7529-7534
292. Klabunde, T., Sharma, S., Telenti, A., Jacobs, W. R., and Sacchettini, J. C. (1998) Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing. *Nat. Struct. Mol. Biol.* **5**, 31-36
293. Camarero, J. A., Pavel, J., and Muir, T. W. (1998) Chemical synthesis of a circular protein domain: evidence for folding-assisted cyclization. *Angew. Chem., Int. Ed.* **37**, 347-349
294. Tu, B. P., Ho-Schleyer, S. C., Travers, K. J., and Weissman, J. S. (2000) Biochemical basis of oxidative protein folding in the endoplasmic reticulum. *Science* **290**, 1571-1574
295. Singh, R., and Whitesides, G. M. (1990) Comparisons of rate constants for thiolate-disulfide interchange in water and in polar aprotic solvents using dynamic proton NMR line shape analysis. *J. Am. Chem. Soc.* **112**, 1190-1197
296. Singh, R., and Whitesides, G. M. (1993) Thiol—disulfide interchange. in *Sulphur-Containing Functional Groups (1993)*, John Wiley & Sons, Inc. pp 633-658
297. Tam, J. P., Liu, W., Zhang, J.-W., Galantino, M., Bertolero, F., Cristiani, C., Vaghi, F., and de Castiglione, R. (1994) Alanine scan of endothelin: Importance of aromatic residues. *Peptides* **15**, 703-708

298. Wu, X., Wu, Y., Zhu, F., Yang, Q., Wu, Q., Zhangsun, D., and Luo, S. (2013) Optimal cleavage and oxidative folding of α -conotoxin TxIB as a therapeutic candidate peptide. *Mar. Drugs* **11**, 3537-3553
299. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J. W., McFarlane, H. T., Madsen, A. O., Riek, C., and Eisenberg, D. (2007) Atomic structures of amyloid cross- β spines reveal varied steric zippers. *Nature* **447**, 453-457
300. Lin, C. C. J., and Chang, J.-Y. (2006) Pathway of oxidative folding of secretory leucocyte protease inhibitor: an 8-disulfide protein exhibits a unique mechanism of folding. *Biochemistry* **45**, 6231-6240
301. Lambert, P., Kuroda, H., Chino, N., Watanabe, T. X., Kimura, T., and Sakakibara, S. (1990) Solution synthesis of charybdotoxin (ChTX), A K⁺ channel blocker. *Biochem. Biophys. Res. Commun.* **170**, 684-690
302. Kuroda, H., Chen, Y., Watanabe, T., Kimura, T., and Sakakibara, S. (1991) Solution synthesis of calciseptine, an L-type specific calcium channel blocker. *Pept. Res.* **5**, 265-268
303. Nakao, M., Nishiuchi, Y., Nakata, M., Watanabe, T., Kimura, T., and Sakakibara, S. (1995) Synthesis and disulfide structure determination of conotoxin GS, a γ -carboxyglutamic acid-containing neurotoxic peptide. *Lett Pept Sci* **2**, 17-26
304. Wang, C. K., Hu, S. H., Martin, J. L., Sjogren, T., Hajdu, J., Bohlin, L., Claeson, P., Goransson, U., Rosengren, K. J., Tang, J., Tan, N. H., and Craik, D. J. (2009) Combined X-ray and NMR analysis of the stability of the cyclotide cystine knot fold that underpins its insecticidal activity and potential use as a drug scaffold. *J. Biol. Chem.* **284**, 10672-10683
305. Galanis, A. S., Albericio, F., and Grøtli, M. (2009) Enhanced microwave-assisted method for on-bead disulfide bond formation: synthesis of α -conotoxin MII. *Biopolymers* **92**, 23-34
306. Hemu, X., Qiu, Y., and Tam, J. P. (2014) Peptide macrocyclization through amide-to-amide transpeptidation. *Tetrahedron* **70**, 7707-7713
307. Tamaoki, H., Kyogoku, Y., Nakajima, K., Sakakibara, S., Hayashi, M., and Kobayashi, Y. (1992) Conformational study of endothelins and sarafotoxins with the cystine-stabilized helical motif by means of CD spectra. *Biopolymers* **32**, 353-357
308. Nguyen, G. K. T., Wang, S., Qiu, Y., Hemu, X., Lian, Y., and Tam, J. P. (2014) Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat Chem Biol* **10**, 732-738
309. Dooley, C., and Houghten, R. (1998) Synthesis and screening of positional scanning combinatorial libraries. in *Combinatorial peptide library protocols* (Cabilly, S. ed.), Humana Press. pp 13-24
310. Gelvin, S. B. (2000) Agrobacterium and plant genes involved in T-DNA transfer and integration. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **51**, 223-256

311. Lu, Y.-A., and Tam, J. P. (2003) X-Cys ligation. in *Proceedings of the 18th American Peptide Symposium. Peptide Revolution: Genomics, Proteomics and Therapeutics*, (M., C., and T.K., S. eds.), Biopolymers Boston, Massachusetts, U.S.A. **71**, 150-160
312. Young, T. S., Young, D. D., Ahmad, I., Louis, J. M., Benkovic, S. J., and Schultz, P. G. (2011) Evolution of cyclic peptide protease inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11052-11056
313. Mao, H., Hart, S. A., Schink, A., and Pollok, B. A. (2004) Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc.* **126**, 2670-2671
314. Williamson, D. J., Fascione, M. A., Webb, M. E., and Turnbull, W. B. (2012) Efficient N-terminal labeling of proteins by use of sortase. *Angew. Chem., Int. Ed.* **51**, 9377-9380
315. Northfield, S. E., Wang, C. K., Schroeder, C. I., Durek, T., Kan, M.-W., Swedberg, J. E., and Craik, D. J. (2014) Disulfide-rich macrocyclic peptides as templates in drug design. *Eur. J. Med. Chem.* **77**, 248-257
316. Korsinczky, M. L., Clark, R. J., and Craik, D. J. (2005) Disulfide bond mutagenesis and the structure and function of the head-to-tail macrocyclic trypsin inhibitor SFTI-1. *Biochemistry* **44**, 1145-1153
317. Laskowski, M., and Kato, I. (1980) Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **49**, 593-626
318. Huang, H., and Player, M. R. (2010) Bradykinin B1 receptor antagonists as potential therapeutic agents for pain. *J. Med. Chem.* **53**, 5383-5399
319. Stewart, J. M. (2004) Bradykinin antagonists: discovery and development. *Peptides* **25**, 527-532
320. Murphey, L. J., Hachey, D. L., Oates, J. A., Morrow, J. D., and Brown, N. J. (2000) Metabolism of bradykinin in vivo in humans: identification of BK1-5 as a stable plasma peptide metabolite. *J. Pharmacol. Exp. Ther.* **294**, 263-269
321. Cyr, M., Lepage, Y., Blais, C., Gervais, N., Cugno, M., Rouleau, J.-L., and Adam, A. (2001) Bradykinin and des-Arg9-bradykinin metabolic pathways and kinetics of activation of human plasma. *Am. J. Physiol. Heart Circ. Physiol.* **281**, H275-H283
322. Ha, S. N., Hey, P. J., Ransom, R. W., Bock, M. G., Su, D.-S., Murphy, K. L., Chang, R., Chen, T.-B., Pettibone, D., and Hess, J. F. (2006) Identification of the critical residues of bradykinin receptor B1 for interaction with the kinins guided by site-directed mutagenesis and molecular modeling. *Biochemistry* **45**, 14355-14361
323. Cascales, L., Henriques, S. T., Kerr, M. C., Huang, Y.-H., Sweet, M. J., Daly, N. L., and Craik, D. J. (2011) Identification and characterization of a new family of cell-penetrating peptides: cyclic cell-penetrating peptides. *J. Biol. Chem.* **286**, 36932-36943

324. D'Souza, C., Henriques, S. T., Wang, C. K., and Craik, D. J. (2014) Structural parameters modulating the cellular uptake of disulfide-rich cyclic cell-penetrating peptides: MCoTI-II and SFTI-1. *Eur. J. Med. Chem.* **88**, 10-18
325. Sadler, K., Eom, K. D., Yang, J.-L., Dimitrova, Y., and Tam, J. P. (2002) Translocating proline-rich peptides from the antimicrobial peptide bactenecin 7. *Biochemistry* **41**, 14150-14157
326. Collier, H. O., Dinneen, L. C., Johnson, C. A., and Schneider, C. (1968) The abdominal constriction response and its suppression by analgesic drugs in the mouse. *Br. J. Pharmacol. Chemother.* **32**, 295-310
327. Le Bars, D., Gozariu, M., and Cadden, S. W. (2001) Animal models of nociception. *Pharmacol. Rev.* **53**, 597-652