



**SYNTHESIS OF BIVALENT SFTI-1 MICROPROTEIN  
BASED ON THE DIAMINOPIMELIC ACID SCAFFOLD  
AND BIOLOGICAL ACTIVITIES**

**BY**

**MISS SUJICHON THANGVICHIEEN**

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF SCIENCE (CHEMISTRY)  
DEPARTMENT OF CHEMISTRY  
FACULTY OF SCIENCE AND TECHNOLOGY  
THAMMASAT UNIVERSITY  
ACADEMIC YEAR 2020  
COPYRIGHT OF THAMMASAT UNIVERSITY**

**SYNTHESIS OF BIVALENT SFTI-1 MICROPROTEIN  
BASED ON THE DIAMINOPIMELIC ACID SCAFFOLD  
AND BIOLOGICAL ACTIVITIES**

**BY**

**MISS SUJICHON THANGVICHIE**

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF SCIENCE (CHEMISTRY)  
DEPARTMENT OF CHEMISTRY  
FACULTY OF SCIENCE AND TECHNOLOGY  
THAMMASAT UNIVERSITY  
ACADEMIC YEAR 2020  
COPYRIGHT OF THAMMASAT UNIVERSITY**

THAMMASAT UNIVERSITY  
FACULTY OF SCIENCE AND TECHNOLOGY

THESIS

BY

MISS SUJICHON THANGVICHEN

ENTITLED

SYNTHESIS OF BIVALENT SFTI-1 MICROPROTEIN BASED ON THE  
DIAMINOPIMELIC ACID SCAFFOLD AND BIOLOGICAL ACTIVITIES

was approved as partial fulfillment of the requirements for  
the degree of Master of Science (Chemistry)

on July 14, 2021

Chairman



(Jiraporn Arunpanichlert, Ph.D.)

Member and Advisor



(Assistant Professor Panumart Thongyoo, Ph.D.)

Member



(Assistant Professor Jariya Romisaiyud, Ph.D.)

Dean



(Associate Professor Nuttanont Hongwarittorn, Ph.D.)

Thesis Title	SYNTHESIS OF BIVALENT SFTI-1 MICROPROTEIN BASED ON THE DIAMINOPIMELIC ACID SCAFFOLD AND BIOLOGICAL ACTIVITIES
Author	Miss Sujichon Thangvichien
Degree	Master of Science (Chemistry)
Department/Faculty/University	Chemistry Faculty of Science and Technology Thammasat University
Thesis Advisor	Assistant Professor Panumart Thongyoo, Ph.D
Academic Year	2020

## ABSTRACT

The sunflower trypsin inhibitor (SFTI-1) is a protease inhibitor isolated from sunflower seeds. Its structure is consisted of 14 amino acids with a cyclic backbone. SFTI-1 demonstrated an inhibitory activity against matriptase, an over-expressed protease in cancerous cells. Also, it showed an inhibition against human  $\beta$ -trypsin, an enzyme involved in allergic and inflammatory disorders in human. To this work, a bivalent peptide inhibitor derived from SFTI-1 was designed and synthesized *via* a one pot 'double' macrolactamization. The newly synthesized bivalent scaffold was generated using the combination of 2,6-diaminopimelic acid scaffold grafted by two SFTI-1 inhibitory epitopes with minor modifications. This bivalent SFTI-1 showed a promising inhibitory activity against HT29 ( $IC_{50} = 53 \mu M$ ) and HeLa cells ( $IC_{50} = 76 \mu M$ ). Importantly, the linear scaffold of bivalent SFTI-1 was fluorescently labelled by FITC, and successfully applied to visualize HeLa cells using a confocal microscope.

**Keywords:** Bicyclic Inhibitor, Sunflower Trypsin Inhibitor (SFTI-1), Bio-imaging

## ACKNOWLEDGEMENTS

This thesis was carried out at Department of Chemistry, Faculty of Science and Technology, Thammasat University under the supervision of Assistant Professor Dr. Panumart Thongyoo.

Firstly, I would like to express my sincere thanks to my thesis advisor, Assistant Professor Dr. Panumart Thongyoo, Department of Chemistry, Faculty of Science and Technology, Thammasat University for his professional advice and invaluable support throughout the course of this research.

I am grateful to Dr. Peerada Yingyuad, Department of Chemistry, Faculty of Science, Kasetsart University and Assistant Professor Dr. Jariya Romsaiyud, Department of Chemistry, Faculty of Science, Ramkhamhaeng University for kindness in NMR, MS and MTT experiments.

Special thanks to Mr. Dusit Promrug and Assistant Professor Dr. Dumrongkiet Arthan, Department of Tropical Nutrition and Food Science, Faculty of Tropical Medical, Mahidol University for their help and instruments in bioimaging applications.

I would like to thank Dr. Jiraporn Arunpanichlert, Department of Chemistry, Faculty of Science and Technology, Thammasat University and Assistant Professor Dr. Jariya Romsaiyud, Department of Chemistry, Faculty of Science, Ramkhamhaeng University for suggestions and generous advices.

I am really thankful Medicinal Chemistry Research Unit, Thammasat University members for suggestion and all their help.

I would like to thank Department of Chemistry, Faculty of Science and Technology and Center of Scientific Equipment for Advanced Research, Thammasat University for instrumental support.

Furthermore, this research was financially supported by National Research Council of Thailand “FY2019 Thesis Grant for master degree student”

Finally, I most gratefully acknowledge my parents and my friends for all their support throughout the period of this research.

Miss Sujichon Thangvichien

## TABLE OF CONTENTS

	Page
ABSTRACT	(1)
ACKNOWLEDGEMENTS	(2)
LIST OF TABLES	(7)
LIST OF FIGURES	(9)
LIST OF SCHEMES	(13)
LIST OF ABBREVIATIONS	(15)
CHAPTER 1 INTRODUCTION	1
1.1 Peptides and Pharmaceutical usage	1
1.1.1 Peptides	1
1.1.2 Pharmaceutical usage	2
1.2 Serine protease and Catalytic mechanism	5
1.2.1 Serine protease	5
1.2.2 Catalytic mechanisms of serine protease	6
1.2.3 Human $\beta$ -tryptase	9
1.2.4 Matriptase	10
1.3 Enzyme Inhibitor	11
1.3.1 Bowman-Birk Inhibitor (BBI)	11
1.3.1.1 Sunflower Trypsin Inhibitor-1 (SFTI-1)	12
1.4 Objectives	13

CHAPTER 2 REVIEW OF LITERATURE	14
2.1 Chemical Peptide Synthesis	14
2.1.1 Amino acid	14
2.1.2 Protecting groups	18
2.1.3 Coupling reagents	20
2.1.3.1 Carbodiimides	20
2.1.3.2 Phosphonium salts	22
2.1.3.3 Uronium/Aminium salts	24
2.1.4 Peptide synthesis strategies	27
2.1.4.1 Solution phase peptide synthesis	27
2.1.4.2 Solid phase peptide synthesis (SPPS)	27
2.1.4.2.1 Boc strategy	28
2.1.4.2.2 Fmoc strategy	29
2.2 Structure modification of peptides	33
2.2.1 Bivalent inhibitor	33
2.2.2 Bicyclic peptide	37
2.2.3 SFTI-1 modification	41
2.2.3.1 The residues modification	42
2.2.3.2 Disulfide bridge modification	46
CHAPTER 3 RESEARCH METHODOLOGY	52
3.1 Chemicals and Instrument	52
3.1.1 Chemicals	52
3.1.2 Instruments	54
3.2 Experiments	55
3.2.1 The synthesis of tripeptide attached on scaffold <i>via</i> solution phase peptide synthesis	55
3.2.2 The synthesis of linear peptide chain <i>via</i> solid phase peptide synthesis (SPPS)	62

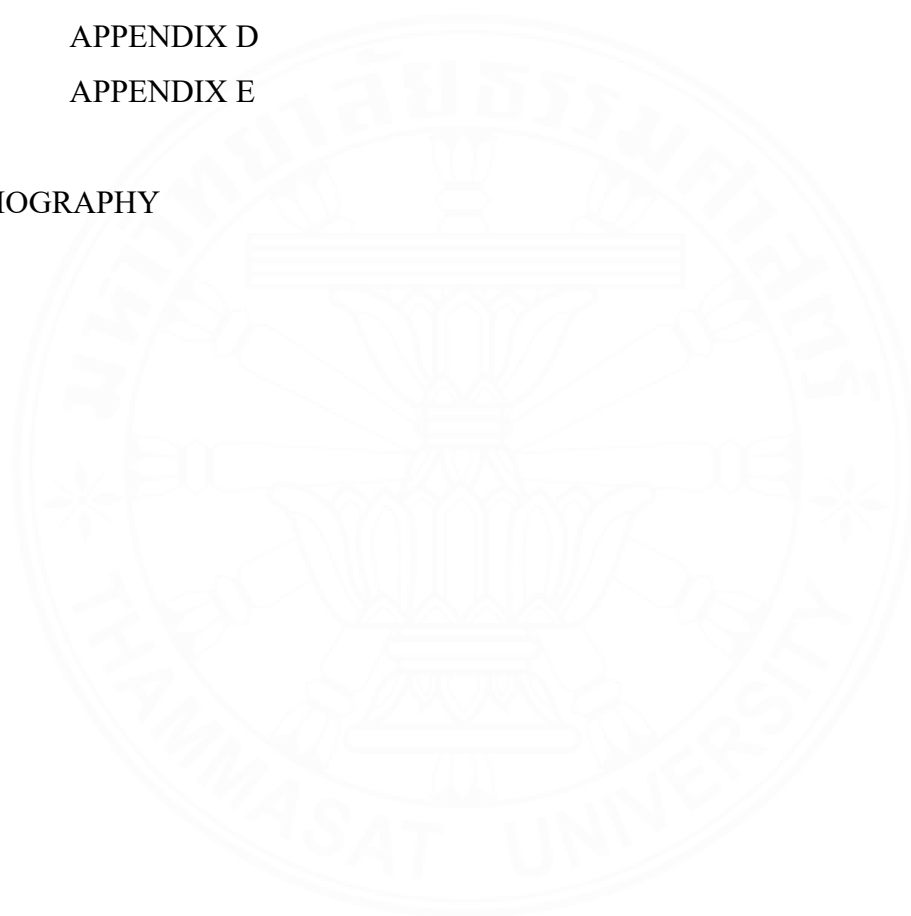
3.2.3 The ligation and cyclization	65
3.2.4 The synthesis of fluorescently labelled bivalent SFTI-1	68
3.2.5 Evaluation of cell viability	70
3.2.6 Fluorescence Imaging Protocol	70
 CHAPTER 4 RESULTS AND DISCUSSION	 71
4.1 The synthesis of tripeptide attached on scaffold <i>via</i> solution phase peptide synthesis	72
4.1.1 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-glycinyI)pimelic acid	72
4.1.2 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-arginyl(Pbf)-glycinyI) pimelic acid	81
4.1.3 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycinyI) pimelic acid	90
4.2 The synthesis of linear peptide chains <i>via</i> solid phase peptide synthesis (SPPS)	101
4.3 The ligation and cyclization	105
4.3.1 The peptide ligation	105
4.3.2 The cyclization	115
4.4.3 The global deprotection	118
4.4 The synthesis of fluorescently labelled bivalent SFTI-1	121
4.5 The biological activity and application	127
4.5.1 The anticancer activity	127
4.5.2 The bioimaging application	129
 CHAPTER 5 CONCLUSIONS	 131

REFERENCES	134
------------	-----

## APPENDICES

APPENDIX A	140
APPENDIX B	141
APPENDIX C	142
APPENDIX D	149
APPENDIX E	159

BIOGRAPHY	165
-----------	-----



## LIST OF TABLES

Tables	Page
1.1 Top peptide drugs by sales in 2019	4
2.1 Chemical structures of proteinogenic amino acids divided by polarity properties	15
2.2 The protecting groups, chemical structure and functional group protection on amino acid residues.	19
2.3 The structure of DCC, DIC, EDC	20
2.4 The structures of BOP, PyBOP, AOP, PyAOP	22
2.5 The structure of HATU, HBTU, HCTU	24
2.6 The structure of DMAP, HOAt, HOBt	26
2.7 An <i>in vitro</i> evaluation of quinazoline-base structure linked by trans-1,4-Diaminocyclohexane	35
2.8 An <i>in vitro</i> potency for $\beta$ -tryptase of dimeric inhibitor series	36
2.9 The inhibitory activities of native SFTI-1 and monocyclic SFTI-1 against several serine proteases	41
2.10 The amino acid sequences and inhibitory activities of native SFTI-1 and SFTI-1 analogues against trypsin	43
2.11 The amino acid sequences and inhibitory activities of native SFTI-1 and SFTI-1 analogues against chymotrypsin	43
2.12 The amino acid sequences and inhibitory activities of native SFTI-1 and SFTI-1 analogues against matrilysin	45
2.13 The amino acid residues of monocyclic SFTI-1 and Seleno SFTI-1	47
2.14 Trypsin inhibitory constants of monocyclic SFTI-1 and Seleno SFTI-1	47
2.15 Physicochemical properties and association equilibrium constant with Bovine $\beta$ -trypsin of SFTI-1 analogues containing carbonyl bridge	49
2.16 Relative matrilysin inhibitory activity of SFTI-1, SFTI-2 and SFTI-3	51
4.1 The used condition of <i>N,N'</i> -diamino-bis(glycyl)pimelic acid syntheses	73
4.2 $^1\text{H}$ -NMR of <i>N,N'</i> -diamino-bis(Fmoc-glycyl) pimelic acid	74
4.3 $^1\text{H}$ -NMR of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -glycyl) pimelic acid	79

4.4 $^1\text{H}$ -NMR of <i>N,N'</i> -diamino-bis(Fmoc-argenyl(Pbf)-glyciny)l pimelic acid	83
4.5 $^1\text{H}$ -NMR of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginy)l(Pbf)-glyciny)l pimelic acid	93
4.6 The developed condition of solution phase peptide synthesis (SPPS).	100
4.7 The conditions of ligation.	106
4.8 The development of ligation approach	107
5.1 The suitable synthesis condition of bivalent SFTI-1 for each steps	132



## LIST OF FIGURES

Figures	Page
1.1 Peptide bond formation	1
1.2 Peptide drugs approvals increased steadily over the last six decades.	2
1.3 Therapeutic indication	3
1.4 Primary structure of insulin	3
1.5 The peptide bond hydrolysis mechanism at the active site of serine protease enzyme	5
1.6 The substrate-binding pocket of the serine protease	7
1.7 The catalytic mechanism of serine protease	8
1.8 The Structure of Human $\beta$ -tryptase	9
1.9 The Structure of Matriptase	10
1.10 The Sunflower Trypsin Inhibitor (SFTI-1) structure	12
2.1 The amino acid structure	14
2.2 Structure of poly BHG2-62 and poly PHG2-54 inhibitors	34
2.3 Structure of IAPs antagonist based on quinazoline scaffold	35
2.4 Structure of monomeric and dimeric human- $\beta$ -tryptase inhibitors	36
2.5 The bicyclic peptide for therapeutic applications	37
2.6 The structure of bicyclic peptide coagulation factor XII inhibitor	38
2.7 The specificity profile and inhibitory activity of FXII618 against several serine protease enzymes	38
2.8 The structure of PTP1B inhibitor	39
2.9 The Live-cell confocal microscopy images of A549 cells after treatment with FITC-labeled monocyclic and bicyclic inhibitors	40
2.10 The structure of selenocysteine and cysteine amino acids	46
2.11 The structure of SFTI-1 analogues with carbonyl bridge	48
2.12 The structure of Dap, Dab , Orn and Lys amino acids	48
2.13 Structure of SFTI-1, SFTI-2 and SFTI-3	50

3.1	The structure of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l)pimelic acid	55
3.2.	The structure of peptide chain Fmoc-isoleucine-serine( <sup>t</sup> Bu)-lysine(Boc)-threonine( <sup>t</sup> Bu)-OH	63
4.1	The structure of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l) pimelic acid	72
4.2	The structure of <i>N,N'</i> -diamino-bis(Fmoc-glyciny)l)pimelic acid	74
4.3.	ESI-MS of <i>N,N'</i> -diamino-bis(Fmoc-glyciny)l)pimelic acid	76
4.4.	HPLC chromatogram of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -glyciny)l)pimelic acid	77
4.5	The structure of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -glyciny)l)pimelic acid	78
4.6	<sup>1</sup> H NMR spectrum of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -glyciny)l)pimelic acid	78
4.7	HR-ESI-MS of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -glyciny)l) pimelic acid.	80
4.8	The structure of <i>N,N'</i> -diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l) pimelic acid	82
4.9	<sup>1</sup> H-NMR spectrum of <i>N,N'</i> -diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l) pimelic acid	84
4.10	ESI-MS spectrum of <i>N,N'</i> -diamino-bis(Fmoc-argenyl(Pbf)-glyciny)l) pimelic acid	85
4.11	The structure of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -arginyl(Pbf)-glyciny)l) pimelic acid	86
4.12	<sup>1</sup> H NMR spectrum of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -arginyl(Pbf)-glyciny)l)pimelic acid	87
4.13	HR-ESI-MS ( <i>via</i> positive and negative ion mode) of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -arginyl(Pbf)-glyciny)l) pimelic acid.	88
4.14	HPLC chromatogram of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -argenyl(Pbf)-glyciny)l) pimelic acid	89
4.15	The structure of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l) pimelic acid	91
4.16	<sup>1</sup> H NMR spectrum of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l) pimelic acid	92
4.17	ESI-MS of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l) pimelic acid	94

4.18	The structure of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid	95
4.19	HPLC chromatogram of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> cysteinyl(Trt)-argenyl (Pbf)-glyciny)l pimelic acid	96
4.20	<sup>1</sup> H-NMR spectrum of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid	97
4.21	HR-ESI-MS of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid	98
4.22	The structure of Fmoc-isoleucine-serine( <sup>t</sup> Bu)-lysine(Boc)-threonine( <sup>t</sup> Bu)-OH	103
4.23	ESI-MS of Fmoc-isoleucine-serine( <sup>t</sup> Bu)-lysine(Boc)-threonyl( <sup>t</sup> Bu)-OH	104
4.24	The structure of <i>N,N'</i> -diamino-bis(Fmoc-isoleuciny)l-seriny)l( <sup>t</sup> Bu)-lysiny)l(Boc)-threonyl( <sup>t</sup> Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid	109 110
4.25	ESI-MS spectrum of <i>N,N'</i> -diamino-bis(Fmoc-isoleuciny)l-seriny)l( <sup>t</sup> Bu)-lysiny)l(Boc)-threonyl( <sup>t</sup> Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid	112
4.26	The comparison HPLC chromatogram of ligation reaction and peptide precursor	113 114
4.27	<sup>1</sup> H NMR spectrum of the corresponding ligated peptide.	117
4.28	ESI-MS of corresponding ligation by product.	
4.29	The HPLC chromatogram comparison between Fmoc deprotected ligation and cyclization reaction crude	118
4.30.	The HPLC chromatogram and MALDI-TOF mass spectrum of bismacrocylic SFTI-1	120
4.31	The structure of <i>N,N'</i> -bis(cyclic-isoleuciny)l-seriny)l-lysiny)l-threonyl-cysteinyl-arginyl-glyciny)l pimelic acid	123
4.32	The structure of fluorescently labelled <i>N,N'</i> -diamino-bis(isoleuciny)l-seriny)l-lysiny)l-threonyl-cysteinyl-arginyl-glyciny)l pimelic acid	124
4.33	The HPLC chromatogram of fluorescently labelled bivalent SFTI-1 reaction and precursors	126

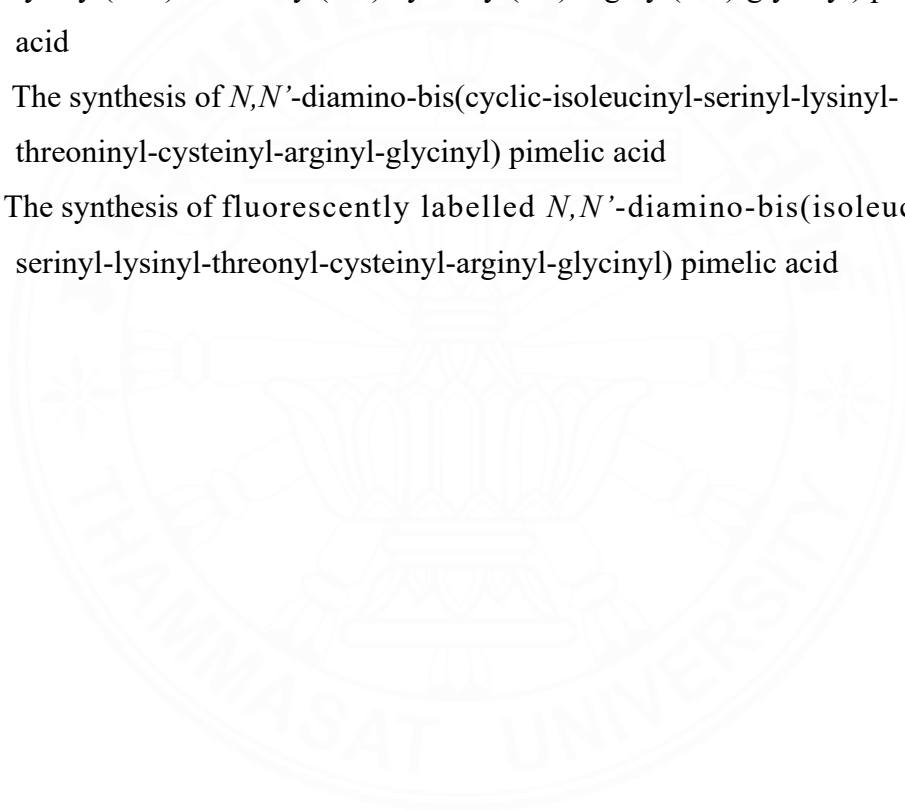
- 4.34 MALDI-TOF mass spectrum of fluorescently labelled *N,N'*-bis(isoleucinyl  
serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid 128  
129
- 4.35 The MTT assays of bivalent SFTI-1
- 4.36 The live-cell confocal microscopy image of HeLa cells after incubated with  
fluorescently labelled bivalent SFTI-1



## LIST OF SCHEMES

Schemes	Page
2.1 The proposed peptide bond formation by using DCC as a coupling reagents.	21 23
2.2 The proposed peptide bond formation by using BOP as a coupling reagents.	25
2.3 The proposed peptide bond formation by using HATU as a coupling reagent.	28
2.4 The mechanism of Boc deprotection	29
2.5 The mechanism of Fmoc deprotection	31
2.6 Fmoc-based solid phase peptide synthesis (SPPS)	56
3.1 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-glycinyI)pimelic acid	57
3.2 The synthesis of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -glycinyI) pimelic acid	58
3.3 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-arginyI(Pbf)-glycinyI) pimelic acid	59
3.4 The synthesis of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -arginyI(Pbf)-glycinyI) pimelic acid	60
3.5 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginyI(Pbf)-glycinyI) pimelic acid	61
3.6 The synthesis of <i>N,N'</i> -diamino-bis(NH <sub>2</sub> -cysteinyl(Trt)-arginyI(Pbf)-glycinyI) pimelic acid	62
3.7 The first amino loading of Fmoc-threonine( <sup>t</sup> Bu)-OH on 2-chlorotrityl chloride resin	63
3.8 Synthesis of Fmoc-isoleucine-serine( <sup>t</sup> Bu)-lysine(Boc)-threonine-( <sup>t</sup> Bu)-OH	65
3.9 Ligation reaction of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginyI(Pbf)-glycinyI) pimelic acid and Fmoc-isoleucine-serine( <sup>t</sup> Bu)-lysine(Boc)-threonine( <sup>t</sup> Bu)-OH	67
3.10 Cyclization and global deprotection reaction	69
3.11 The synthesis of fluorescently labelled <i>N,N'</i> -diamino-bis(isoleucinyl-serinyl-lysinyI-threonyI-cysteinyl-arginyI-glycinyI) pimelic acid	73
4.1 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-glycinyI)pimelic acid	81
4.2 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-arginyI(Pbf)-glycinyI) pimelic acid	90

4.3 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid	102
4.4 The synthesis of Fmoc-isoleucine-serine( <sup>t</sup> Bu)-lysine(Boc)-threonine( <sup>t</sup> Bu)-OH	108
4.5 The synthesis of <i>N,N'</i> -diamino-bis(Fmoc-isoleucinyl-serinyl( <sup>t</sup> Bu)-lysiny(Boc)-threoninyl( <sup>t</sup> Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid	116 119
4.6 The synthesis of <i>N,N'</i> -diamino-bis(cyclic-isoleucinyl-serinyl( <sup>t</sup> Bu)-lysiny(Boc)-threoninyl( <sup>t</sup> Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid	122
4.7 The synthesis of <i>N,N'</i> -diamino-bis(cyclic-isoleucinyl-serinyl-lysiny-threoninyl-cysteinyl-arginyl-glyciny)l pimelic acid	
4.8 The synthesis of fluorescently labelled <i>N,N'</i> -diamino-bis(isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny)l pimelic acid	



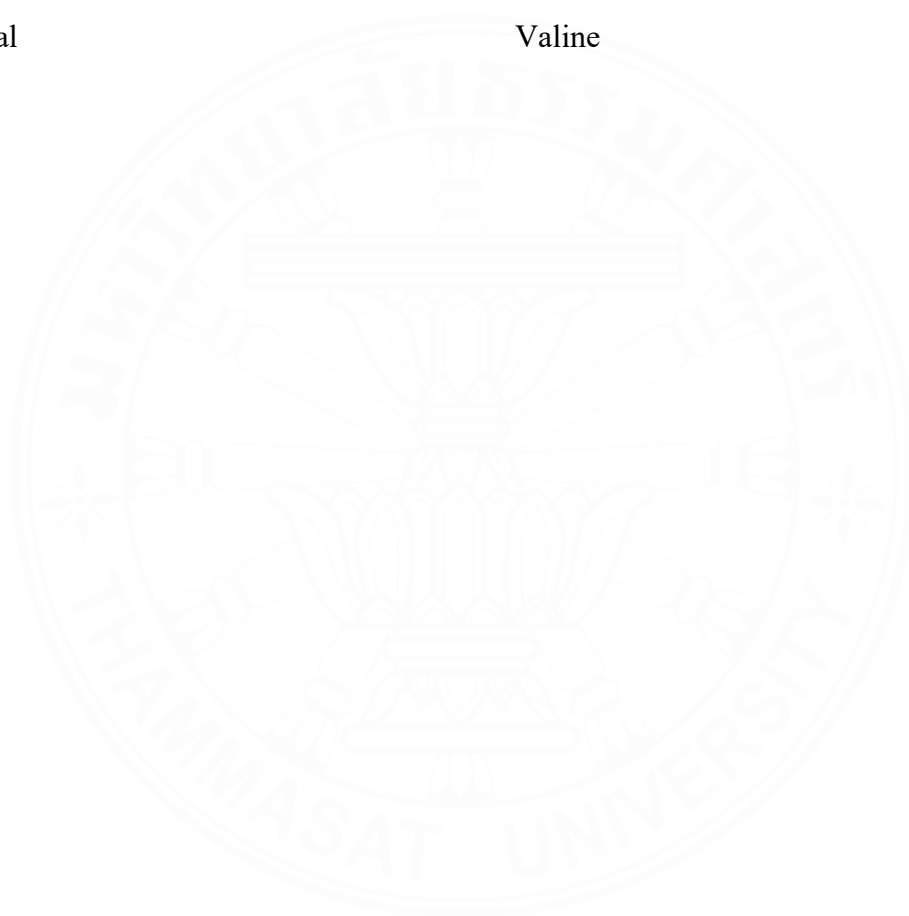
## LIST OF ABBREVIATIONS

Symbols/Abbreviations	Terms
$\alpha$	Alpha
$\beta$	Beta
$\mu\text{M}$	Micromolar
Abs	Absorbance
ACN	Acetonitrile
Ala	Alanine
AOP	7-(Azabenzotriazol-1-yl)oxytris(dimethylamino)phosphonium hexafluorophosphate
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
BBi	Bowman-Birk trypsin inhibitor
Bn/Bzl	Benzyl
Boc	<i>Tert</i> -butyloxycarbonyl
BOP	benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate
<sup>t</sup> Bu	<i>Tert</i> -butyl
CPP	Cell-penetrating peptide
Cys	Cystein
Dab	L-2,4-diaminobutyric acid
Dap	L-2,3-diaminopropionic acid
DCC	<i>N,N'</i> -Dicyclohexylcarbodiimide
DCM	Dichloromethane
DCU	<i>N,N'</i> -dicyclohexylurea
DIC	<i>N,N'</i> -Diisopropylcarbodiimide

DIPEA	<i>N,N'</i> -Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
EDC	<i>N</i> -(3-Dimethylaminopropyl)- <i>N'</i> -ethylcarbonate
EDT	Ethane-1,2-dithiol
EtOAc	Ethylacetate
ESI-MS	Electrospray ionization mass spectrometry
FITC	Fluorescein isothiocyanate
Fmoc	9-fluorenylmethyloxycarbonyl
Glu	Glutamic acid
Gly	Glycine
HATU	1-[Bis(dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i> ]pyridinium-3-oxid hexafluorophosphate
HBTU	(2-(1 <i>H</i> -benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HCTU	1-[Bis(dimethylamino)methylen]-5-chlorobenzotriazolium 3-oxide hexafluorophosphate
HF	Hydrogenfluoride
His	Histidine
HMPA	Hexamethylphosphorotriamide
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HR-ESI-MS	High-resolution electrospray ionization mass spectrometry
IC <sub>50</sub>	Half maximal inhibitory concentration
Ile	Isoleucine
K <sub>a</sub>	Association equilibrium constant

K <sub>i</sub>	Inhibitory constant
Leu	Leucine
Lys	Lysine
MALDI-TOF-MS	Matrix assisted laser desorption/ionization-time of flight mass spectrometry
Met	Methionine
MeOH	Methanol
mg	Milligram
MS	Mass spectrometry
nM	Nanomolar
NMR	Nuclear magnetic resonance spectroscopy
Orn	Ornithine
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran- 5-sulfonyl
Phe	Phenylalanine
PIP	Piperidine
Pro	Proline
PyAOP	[(7-azabenzotriazol-1- yl)oxy]tris(pyrrolidino)-phosphonium hexafluorophosphate
PyBOP	benzotriazol-1-yl-oxy- <i>tris</i> (pyrrolidino) phosphonium hexafluorophosphate
RCM	Ring closure metathesis
SDMI-1	SFTI-1-derived matriptase inhibitor-1
Sec	Selenocysteine
Ser	Serine
SFTI-1	Sunflower trypsin inhibitor-1
SPPS	Solid phase peptide synthesis
TFA	Trifluoacetic acid

THF	Tetrahydrofuran
Thr	Threonine
TIPS	Triisopropylsilane
t <sub>R</sub>	Retention time
Trt	Triphenylmethane
Tyr	Tyrosine
UV	Ultraviolet
Val	Valine



## CHAPTER 1

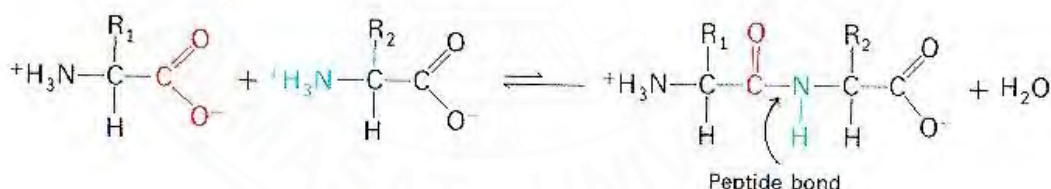
### INTRODUCTION

#### 1.1 Peptides and Pharmaceutical usage

##### 1.1.1 Peptides

Peptides<sup>[1-2]</sup> are essential biomolecules comprised of amino acids, which is widely found in all living organisms and play an important role in many biological activities. There are various types of peptides categorized according to their sources (such as plant peptides, bacterial/antibiotic peptides, fungal peptides, venom peptides) or functions (cancer/anticancer peptides, inflammatory peptides, respiratory peptides, vaccine peptides). The function of peptides is dependent on the amino acid sequences and the specific shape.

Peptides are short protein chain whereas the length is varied between two and fifty amino acids. The peptide bond is formed by condensation reaction between the amino group and the carboxyl group of amino acids as shown in **Figure 1.1**.



**Figure 1.1** Peptide bond formation

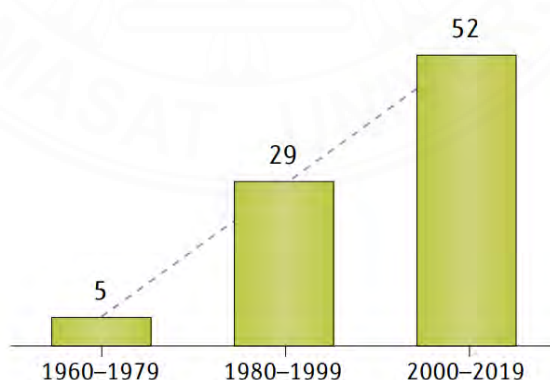
Peptides are used in a wide range of applications, particularly in the cosmetic, supplement and pharmaceuticals mainly due to their potential applications, and less side effects. One of the most popular peptides is collagen, which is usually used as supplements and cosmetics for improving skin health and reverse effect of aging.

### 1.1.2 Pharmaceutical usage

Nowadays, more than 7,000 naturally occurring peptides have been identified<sup>[3]</sup>, and most have shown an important role in many physiological processes in human body. Typically, peptides are high potency of action and high specificity. For this reason, the discovery of novel bioactive peptide has gained of great interest for designing a novel therapeutics.

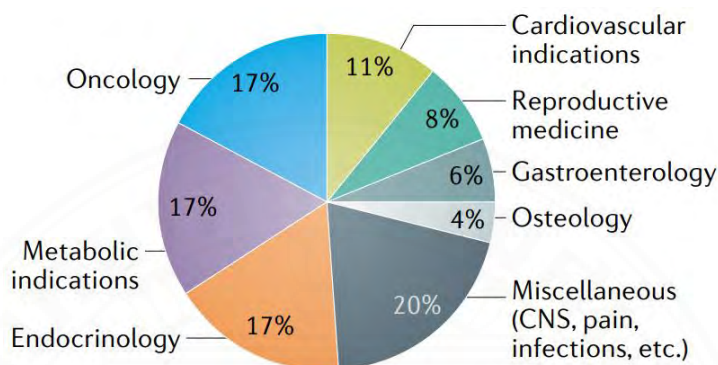
The utilization of peptides as pharmaceuticals has shown the great advantage relative to that of small molecules in terms of high potency, great specificity and selectivity with low toxicity. However, some limitations (poor solubility, low membrane permeability, short half-life, chemically and physically instable and low oral bioavailability<sup>[4,5]</sup>) have clearly been noticed. Currently, most of peptide based drugs (approximately 75%) are given as subcutaneous injection.

During the past decade, peptides in pharmaceutical research and development have gained of great interest (**Figure 1.2**). Currently, there are around 80 peptide medicines on the global market approved by US-FDA (U.S. Food and Drug Administration) and EMA (European Medicines Agency). In addition, this is expected to increase significantly with many peptide and protein drugs currently in clinical trials and preclinical development.<sup>[6]</sup>



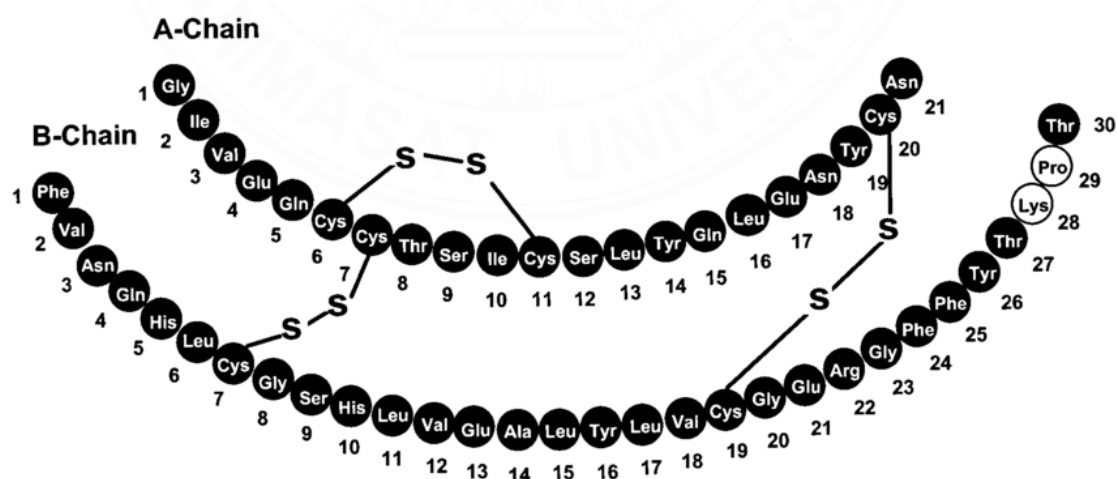
**Figure 1.2** Peptide drugs approvals increased steadily over the last six decades.<sup>[7]</sup>

A metabolic disease has become the main area for utilizing peptides as therapeutics<sup>[7]</sup>. However, peptides have been also used widely as therapeutics in diverse disease areas such as cardiovascular condition, gastroenterology, bone disease, dermatology and sexual dysfunction (**Figure 1.3**).



**Figure 1.3** Therapeutic indication<sup>[7]</sup>

The information from US-FDA, company financial reports and global sale analysis presents top peptide drugs by sales in 2019<sup>[7]</sup> as shown in **Table 1.1**. The great selling peptides have been insulin and analogues. The primary structure of insulin is shown in **Figure 1.4**



**Figure 1.4** Primary structure of insulin<sup>[8]</sup>

**Table 1.1** Top peptide drugs by sales in 2019<sup>[7]</sup>

Generic name	Most common brand names	Companies	Indication	First approval	Sales in 2019 (US\$ million)
Insulin and analogues	Ademelog, Apidra, Humulin, Humalog, Insuman, NovoMix, NovoRapid, Lantus, Levemir, Ryzodeg, Tresiba, Toujeo	Novo Nordisk, Eli Lilly, Sanofi	Diabetes	1982	25,000
Dulaglutide	Trulicity	Eli Lilly, Dainippon Sumitomo	Diabetes	2014	4,394
Liraglutide	Victoza, Saxenda	Novo Nordisk	Diabetes, obesity	2010	4,142
Leuprolide	Lupron, Eligard	AbbVie, Astellas, Takeda	Cancer	1985	2,022
Semaglutide	Ozempic, Rybelsus	Novo Nordisk	Diabetes, obesity	2017, 2019	1,694
Octreotide	Sandostatin	Novartis	Cancer	1988	1,585
Glatiramer	Copaxone, Glatopa	Teva, Sandoz	MS	1996	1,531
Teriparatide	Forteo	Eli Lilly	Osteoporosis	2002	1,405
Cyclosporine	Restasis	Allergan	Immune diseases, organ transplants	1983	1,189
Lanreotide	Somatuline	Ipsen	Acromegaly	2007	1,124
Carfilzomib	Kyprolis	Amgen	Multiple myeloma	2012	1,044
ACTH	Acthar	Mallinckrodt	IS, MS	1950	953
Linaclotide	Linzess, Constella	Allergan, Astellas Pharma	IBS-C	2012	877
Romiplostim	Nplate, Romiplate	Amgen, Kyowa Kirin	Chronic ITP <sup>*</sup>	2008	841
Goserelin	Zoladex	AstraZeneca	Cancer	1989	813
Etelcalcetide	Parsabiv	Amgen, Ono Pharma	Hyperparathyroidism	2017	693
Exenatide	Byetta, Bydureon	AstraZeneca	Diabetes	2005	659
Teduglutide	Gattex, Revestive	Takeda	Short bowel syndrome	2012	555
Vasopressin	Vasopressin	Endo Pharma	CDI	NR	532

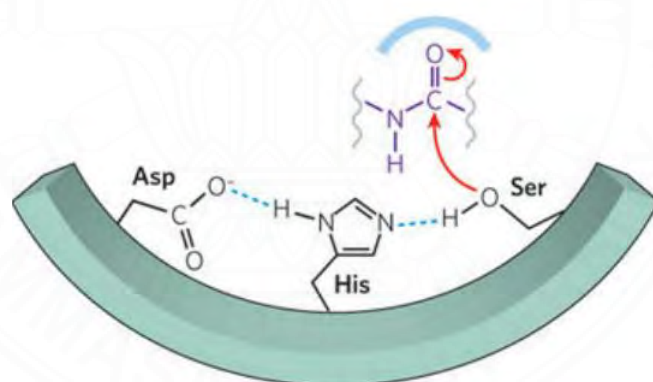
\*ACTH, adrenocorticotropic hormone; CDI, central diabetes insipidus; IBS-C, irritable bowel syndrome with constipation; IS, infantile spasms; ITP, immune thrombocytopenia; MS, multiple sclerosis; NR, not recorded.

## 1.2 Serine protease and Catalytic mechanism

### 1.2.1 Serine protease

Serine protease<sup>[9]</sup> is the group of enzyme, involving in a large number of essential biological processes such as digestion, immunity, fibrinolysis and blood coagulation. However, there are some serine proteases causing the disease; for example human  $\beta$ -tryptase involving in asthma formation, inflammatory disorder and other allergic disorders<sup>[10]</sup>, matriptase contributing to cancer development<sup>[11]</sup>

To date, there are about 50 families of serine protease which could be differentiated based on the amino acid sequence<sup>[12]</sup>. However, the common feature of enzyme in this group is that they shared the similar active site, comprising of serine, histidine, and aspartic acid (catalytic triad), all of which are involved in the peptide bond hydrolysis as displayed in **Figure 1.5**.



**Figure 1.5** The peptide bond hydrolysis mechanism at the active site of serine protease enzyme<sup>[9]</sup>

Furthermore, the serine protease was also classified based on the substrate specificity. For example trypsin-like serine protease, chymotrypsin-like serine protease and elastase-like serine protease<sup>[13]</sup>.

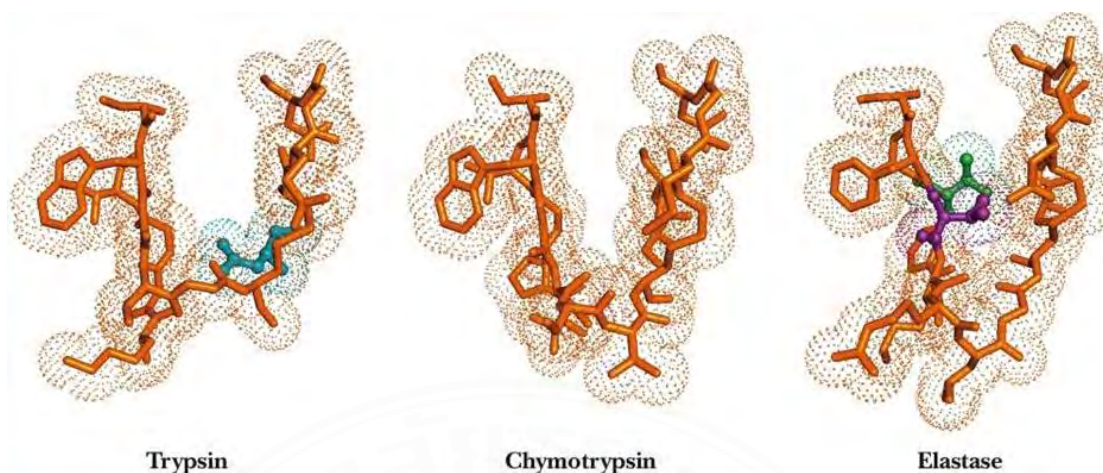
1. Trypsin-like serine protease preferentially cleaved the peptide bond of positively charged amino acid, such as Lysine (Lys) or Arginine (Arg).
2. Chymotrypsin-like serine protease specifically cleaved the peptide bond of aromatic amino acids, such as Tyrosine (Tyr), Phenylalanine (Phe) and Tryptophan (Trp).
3. Elastase-like serine protease preferentially cleaved the peptide bond of a small hydrophobic amino acid, such as Alanine (Ala), Glycine (Gly) and Valine (Val).

### 1.2.2 Catalytic mechanisms of serine protease

Generally, the protein molecule is consisted of two sides that are *N*-terminus and *C*-terminus. Generally, the enzyme preferentially cleaved the peptide bond from the *C*-terminus; for example chymotrypsin preferentially cleaved at the *C*-terminus side of Phenylalanine.

According to the primary sequence, the catalytic triad (serine-195, histidine-57 and aspartic acid-102) is not in close proximity. However, the polypeptide chain will possibly fold into the specific conformation, causing the catalytic triad located close to each other. As shown in **Figure 1.5**, the imidazole side chain of histidine-57 is in close proximity to the hydroxyl group of serine-195, resulting in the hydrogen bond formation between these two residues. Additionally, aspartic acid-102 was also in close proximity to the side chain of histidine at the other side, causing the hydrogen bond also.<sup>[14]</sup>

The binding site of the enzyme substrate was widely known as S1 pocket (**Figure 1.6**)<sup>[15]</sup>. After binding to the enzyme, the protein substrate was subsequently cleaved at the active site.

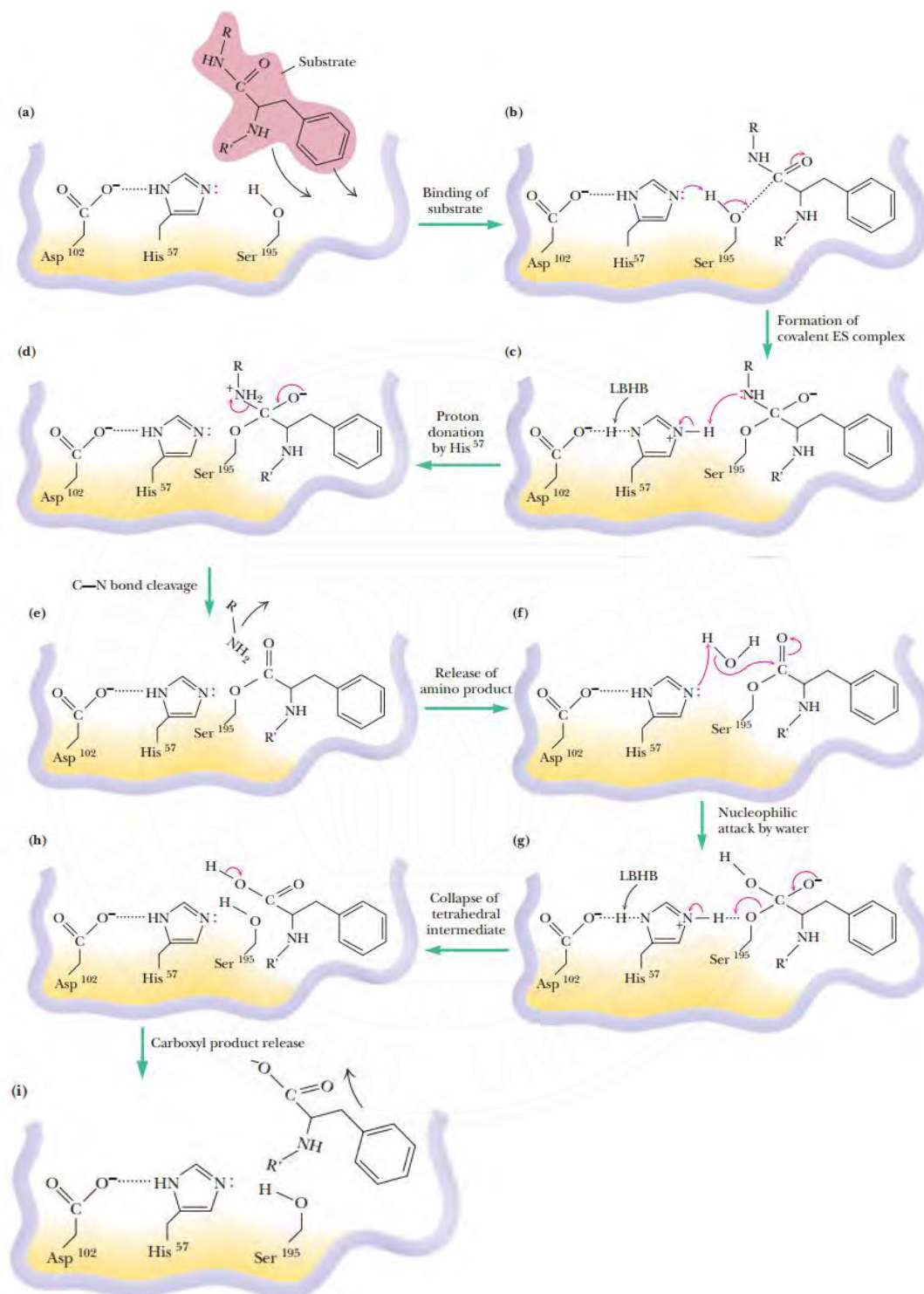


**Figure 1.6** The substrate-binding pocket of the serine protease<sup>[15]</sup>

Serine proteases usually follow a two-step hydrolysis reaction in which a covalently linked enzyme-peptide intermediate is formed with the loss of the amino acid.<sup>[16]</sup>

The first step is the acylation which is the rapid part of reaction. Firstly, the hydrogen of hydroxyl group at Ser195 is deprotonated by imidazole nitrogen at His57 to afford Ser-O<sup>-</sup> which subsequently attack the carbonyl of substrate *via* the nucleophilic addition (**Figure 1.7b**). Next, the proton of imidazolium ion at His57 is deprotonated to cause the collapse of the tetrahedral intermediate (**Figure 1.7c**). The transition state is stabilized by the amide hydrogen leading to the peptide bond cleavage (**Figure 1.7d**) which gives the amino acid as the product (**Figure 1.7e**).

The second step is the hydrolysis which is the slow step of the catalysis. The carboxyl moiety at the substrate at Ser195 was attacked by water molecule at carbonyl position causing the oxygen become the negative charge (**Figure 1.7f**), following by the formation and the stabilization of tetrahedral intermediate, leading to C-O bond cleavage (**Figure 1.7g**) which gives the carboxyl product (**Figure 1.7h**). After the peptide is cleaved, the enzyme is returned to its original states. (**Figure 1.7i**).

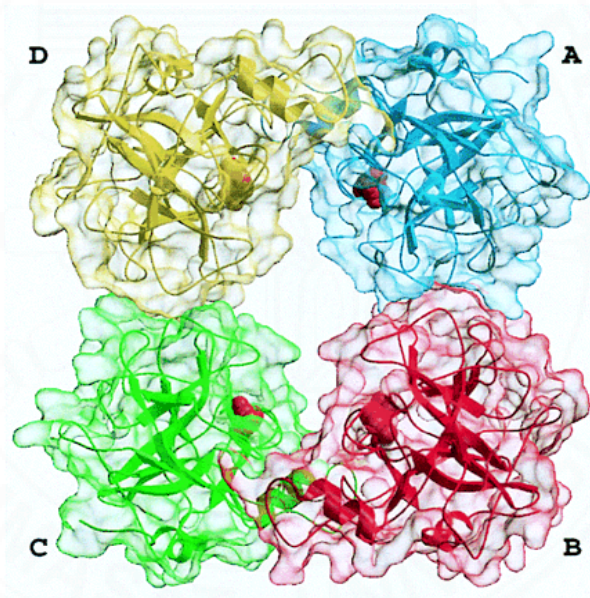


**Figure 1.7** The catalytic mechanism of serine protease<sup>[16]</sup>

### 1.2.3 Human $\beta$ -tryptase

Human  $\beta$ -tryptase<sup>[17]</sup> is classified as serine protease, involving in asthma formation, inflammatory disorder and other allergic disorders. This enzyme is tetramer consisted of four equivalent monomers with two domains where each monomer is interacted by hydrophobic and polar interactions. The four active sites are located next to the central pore as shown in **Figure 1.8**.

Human  $\beta$ -tryptase potentially hydrolyzed the peptide bond with the similar mechanism compared to those of serine protease mechanism with the specificity at the P1 position to be Arginine (R) and Lysine (K).



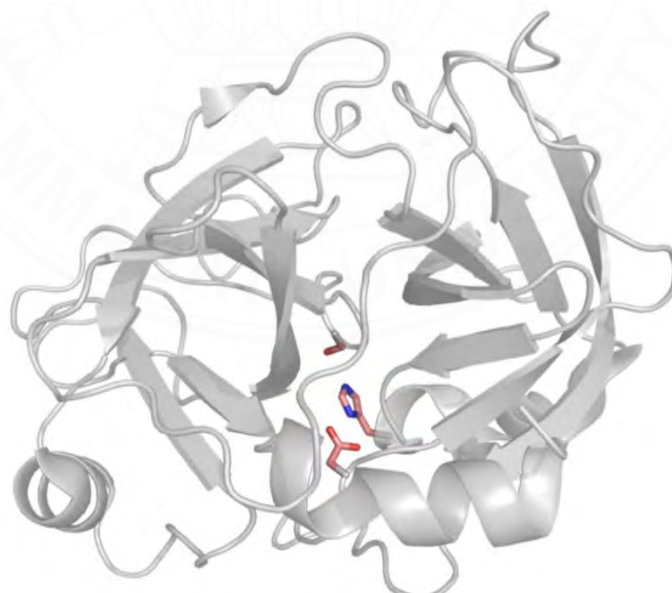
**Figure 1.8** The Structure of Human  $\beta$ -tryptase<sup>[18]</sup>

Human tryptase is the prominent protein in human mast cells, which are present in greater numbers along epithelial linings of respiratory tract. These enzymes are activated intracellularly and stored in secretory granules in catalytically active form. When mast cells interacted with an allergen, they subsequently release histamine and human  $\beta$ -tryptase which led to the constriction and swelling of the airway, causing asthma symptoms such as shortness of breath and/or chest tightness, wheezing, coughing and airflow limitation.

### 1.2.4 Matriptase

Matriptase is a transmembrane of trypsin-like serine protease. This enzyme is expressed in most epithelial cells and also found in mast cells, B cells and peripheral blood monocytes. It has an essential physiological role in activating growth and angiogenic factors, such as hepatocyte growth factor, protease activated receptor-2, and urokinase type plasminogen activator. However, matriptase is usually coexpressed with its cognate inhibitor, hepatocyte growth factor inhibitor or HAI-1. There is evidence indicating matriptase involved in the activation of hepatocyte growth factor and urokinase plasminogen activator, which has been implicated in metastasis and cancer development. Therefore, the over-expression of matriptase leads to epidermal tumor formation, and initiates multistage squamous cell carcinogenesis<sup>[11]</sup>.

Matriptase is isolated from epithelial derived cancer cell, such as cervical, ovarian, colorectal, breast and prostate. Notably, the reduction of matriptase levels leads to the suppression of ovarian primary tumor growth, prostate, and metastasis. The structure of matriptase is shown in **Figure 1.9**.



**Figure 1.9** The Structure of Matriptase<sup>[19]</sup>

### 1.3 Enzyme Inhibitor

An enzyme inhibitor<sup>[20]</sup> is a molecule that binding to active site of enzyme, which can decreases or totally inhibits the enzyme activity. Enzyme inhibitors usually occur naturally and are involved in the regulation of metabolism. Mostly, enzyme inhibitors are low molecular weight chemical molecules. However, there are some proteins as enzyme inhibitors that specifically bind to and inhibit an enzyme target.

The impact of enzyme inhibitor in drug discovery has become a fundamental approach to pharmaceutical industry. Drug bank online database reports more than 100 natural and synthetic protease inhibitors are currently marketed. Such inhibitors exhibit specific action in enzyme inhibition that can be used for the treatment of various disorders and diseases. The well-known enzyme inhibitor drugs such as aspirin are used for inflammation, pain, fever, or penicillins are used for bacterial infection.

#### 1.3.1. Bowman-Birk Inhibitor (BBI)

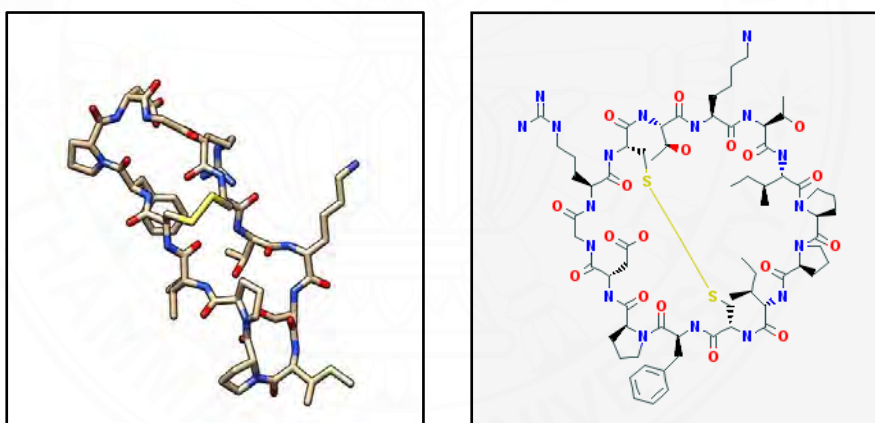
The Bowman-Birk inhibitor<sup>[21]</sup> is a family of protein consists of eukaryotic proteinase inhibitors that mainly inhibit serine protease. These inhibitors are primarily found in plant and in seed of legumes. According to Hellinger and Gruber<sup>[22]</sup>, there are 611 peptides was categorized as Bowman-Birk inhibitor and most of them usually found in plant. Despite differences in their sequence, members of this family share nine-amino acids binding loop motif CTP1SXPPXC (P1 is the inhibitory active site and X is various amino acids) and usually contains seven disulfide linkage. The central part of inhibitors contains a P1-P1' amino acid residues, which interact with the corresponding enzyme at S1 and S1' position. Even though the interaction is principally regulated by P1 residues, adjacent amino acids are also contribute inhibitory activity and specificity.

Members of this family usually have two homologous and independent binding loop located at the opposite site of the molecule, called 'double-head' inhibitors which can inhibit two enzymes with the same or different molecules. The first binding loop is usually related of trypsin inhibition with P1 position is occupied by lysine or

arginine. On the other hand, the second binding loop is involved with chymotrypsin inhibition and contains hydrophobic amino acid. Due to the presence of these two inhibitory loops, BBIs represent a high affinity toward trypsin and chymotrypsin.

### 1.3.1.1 Sunflower Trypsin Inhibitor-1 (SFTI-1)

Sunflower trypsin inhibitor-1 (SFTI-1)<sup>[23]</sup>, a member in Bowman-Birk inhibitor family, is a circular small peptide that is produced from sunflower seeds. This inhibitor is a cyclic peptide consisted of 14 amino acids (**Figure 1.10**) with only one disulfide bond. Even though this cyclized peptide containing only single disulfide bond is not related to other BBIs, they share an almost identical binding loop. Its inhibitory loop was comprised of Glycine, Arginine, Cysteine, Threonine, Lysine, Serine and Isoleucine (GRCTKSI), respectively.

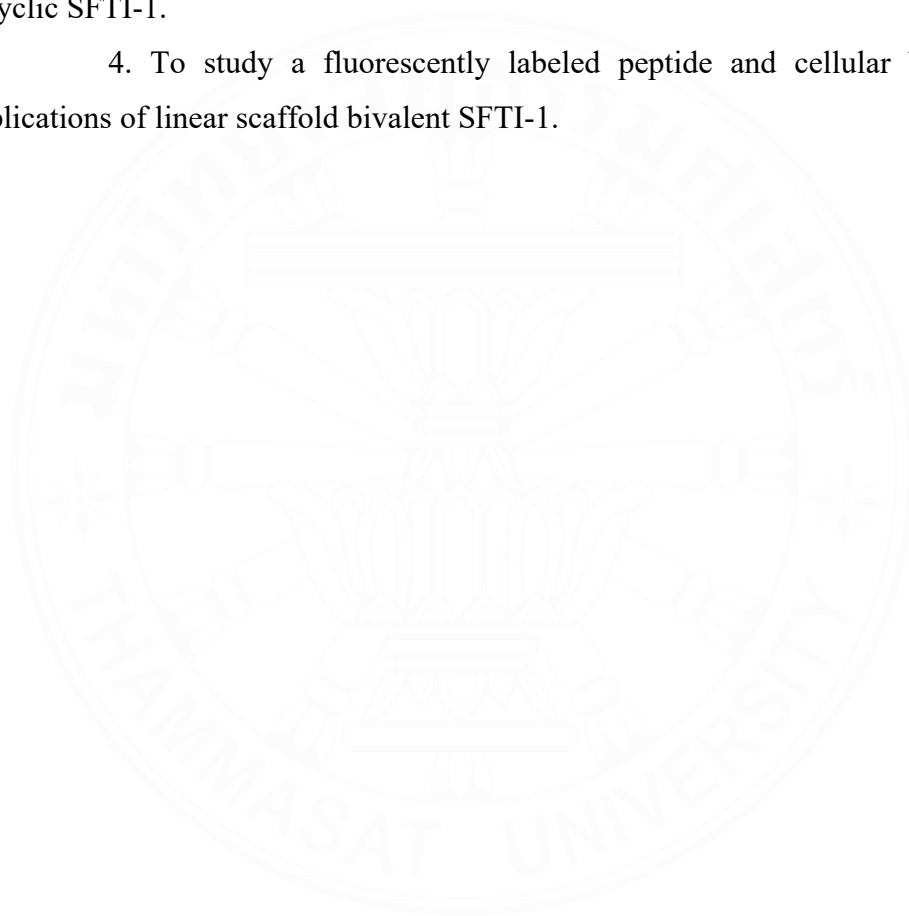


**Figure 1.10** The Sunflower Trypsin Inhibitor (SFTI-1) structure

The SFTI-1 showed remarkably potent inhibitory activity against serine protease trypsin with  $K_i$  of 0.1 nM<sup>[24]</sup>. Trypsin is one of the largest families of protease, which is involved in wide range key function of human physiology. Many of these protease are also inhibited by SFTI-1. With its small size (1.5 kDa) and compact structure, SFTI-1 has been recognized as an attractive engineer template to design potent, novel, and highly selective inhibitor, blocking diverse protease.

#### 1.4 Objectives

1. To design and synthesize novel bicyclic proteinase inhibitors based on the 2,6-diaminopimelic acid scaffold.
2. To develop the synthetic strategy of derived bivalent SFTI-1 inhibitor.
3. To assess a biological activity against a panel of cancer cell lines of bicyclic SFTI-1.
4. To study a fluorescently labeled peptide and cellular bio-imaging applications of linear scaffold bivalent SFTI-1.



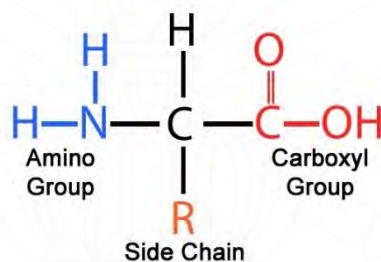
## CHAPTER 2

### REVIEW OF LITERATURE

#### 2.1 Chemical Peptide Synthesis

##### 2.1.1 Amino acid

An amino acid<sup>[25]</sup> is an organic compound that contains an amino group (-NH<sub>2</sub>) and carboxyl group (-COOH), covalently bonded to a central carbon atom as shown in **Figure 2.1**. It is also bonded to a hydrogen atom and side chain (R group) at the same carbon. The specific amino acid side chains give them different physical and chemical properties.

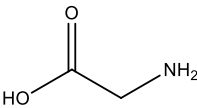
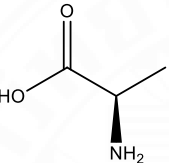
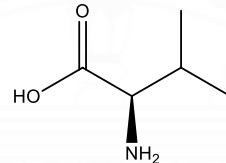
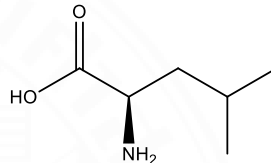
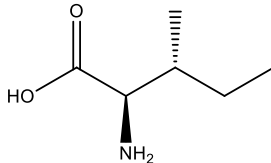
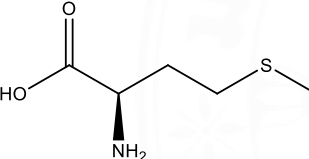
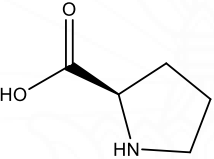
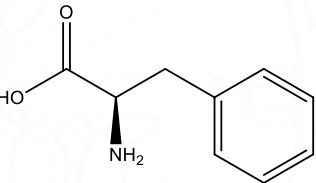
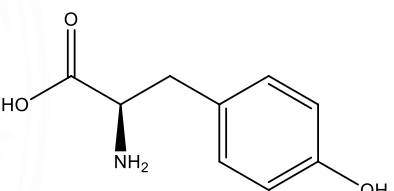
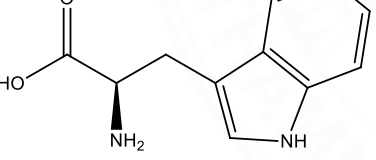
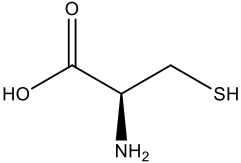
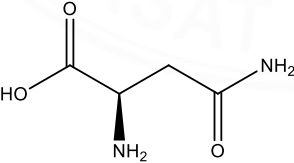
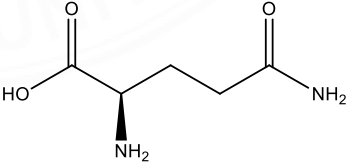


**Figure 2.1** The amino acid structure

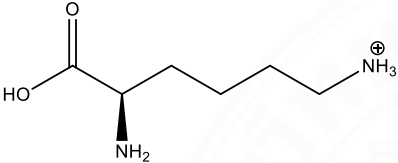
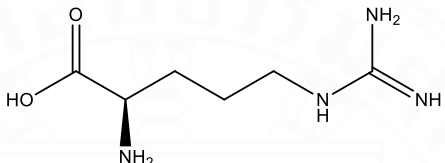
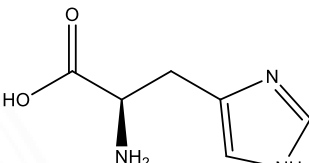
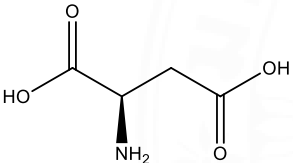
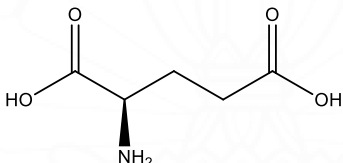
There are 500 amino acid have been discovered until present, most are metabolic intermediate isolated from bacteria<sup>[26]</sup>. However, only 20 amino acids that occur in living organisms are encoded by triplet codon in the genetic code relating to the synthesis of protein. These 20 amino acids were called standard amino acids. They can be classified as a group in many ways, such as nutrition requirement, pH level or the structure of its side chain, but the most popular manner is classified based on the polarity including polar, non-polar, positive and negative charged amino acids.

According to the International Nucleotide Sequence database, the amino acid can be described with one and three letter abbreviation in order to use the full name. The lists of proteinogenic amino acid abbreviation, structures and polarity properties are shown in **Table 2.1**.

**Table 2.1** The chart shows chemical structures of proteinogenic amino acids divided by polarity properties.

Groups	Names and Structures
Nonpolar amino acids	<div>      </div> <div> Glycine (Gly,G)      Alanine (Ala,A)      Valine (Val,V)      Leucine (Leu,L)      Isoleucine (Ile,I) </div>
	<div>     </div> <div> Methionine (Met,M)      Proline (Pro,P)      Phenylalanine (Phe,F)      Tyrosine (Tyr,W) </div>
	<div>  </div> <div> Tryptophan (Trp,W) </div>
Polar uncharged amino acids	<div>    </div> <div> Cysteine (Cys,C)      Asparagine (Asn,N)      Glutamine (Glu,Q) </div>

**Table 2.1** The chart shows chemical structures of proteinogenic amino acids divided by polarity properties. (Cont)

Polar positively charged amino acids	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p>Lysine (Lys,K)</p> </div> <div style="text-align: center;">  <p>Arginine (Arg,R)</p> </div> <div style="text-align: center;">  <p>Histidine (His,H)</p> </div> </div>
Polar Negatively charged amino acids	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p>Aspartic acid (Asp,D)</p> </div> <div style="text-align: center;">  <p>Glutamic acid (Glu,E)</p> </div> </div>

Generally, the 20 standard amino acids are classified base on their polarity propertied as follows:

1. Nonpolar amino acids contain aliphatic or aromatic rings on their side chains that make them hydrophobic properties. The members in this group consist of Glycine (G), Alanine (A), Valine (V), Leucine (L), Isoleucine (I), Methionine (M), Proline (P), Phenylalanine (F), Tyrosine (Y) and Tryptophan (W).

2. Polar uncharged amino acids have functional groups, such as sulfhydryl or amide groups with no charge on their side chains. The electron pairs of sulfur or nitrogen are available for hydrogen bond donors. This group is further included Cysteine (C), Asparagine (N) and Glutamine (Q).

3. Polar negatively charged amino acids or acidic amino acids, including Aspartic acid (D) and Glutamic acid (E). Each amino acid contains carboxylic acid on its side chain that provided its proton-donor properties. In an aqueous solution, dicarboxylic acid and mono amine groups are ionized, which give an overall charge to -1.

4. Polar positively charged amino acids or basic amino acids, for example Arginine (R), Histidine (H) and Lysine (K). Each side chain contains amine functional groups that can accept a proton. The presence of dibasic and monocarboxylic functionalities give them overall charge +1

### 2.1.2 Protecting groups

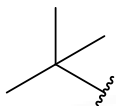
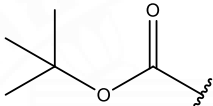
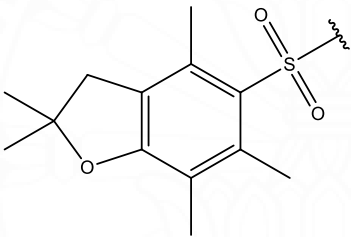
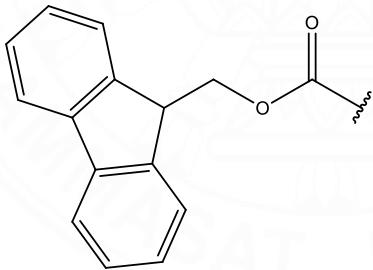
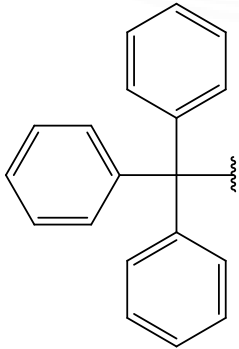
Amino acids contain various chemical functionalities include amino, carboxyl and many types of functional groups on their side chains. To synthesize the peptide bond, an amino group of one amino acid is covalently linked with a carboxyl group of another amino acid by condensation reaction to generate a new amide linkage. However, the unprotected amino acid can be self-polymerized. In addition, the reactive functional group at side chain can cause side reactions. Therefore, the protecting group is required for peptide syntheses to prevent undesired products.

The protecting group<sup>[27]</sup> is temporarily introduced into a functional group of amino acids to decrease their reactivity, which could prevent an unexpected reaction in subsequent steps. The properties of a good protecting group are as follows: (1) easy to remove in high yield under specific condition and (2) inert under synthetic conditions.

The most common *N*-amino protecting groups for peptide synthesis are 9-fluorenylmethoxycarbonyl (Fmoc) and *tert*-butyloxycarbonyl (Boc). Furthermore, there are diverse protecting groups for side chain protection; for example *tert*-butyl ('Bu), triphenylmethyl (Trt), Benzyl (Bn), Pbf, etc. The example of protecting groups is shown in **Table 2.2**.

It is well-known that each amino acid comprises of more than one functionality in a molecule, the nature of protecting groups must be carefully selected based on their different reactivity profile which could perform the selective deprotection of just one functional group while others remain unprotected. This is referred to as an orthogonal protecting group strategy. For example, Lysine contains two amino groups. The first one is protected with the Fmoc groups which can be selectively deprotected under basic condition, while the second is protected by Boc group which is sensitive under acidic condition.

**Table 2.2** The protecting groups of amino acid side chain.

Protecting group	structure	Functional group	Amino acids
<sup>t</sup> Bu		Hydroxyl Carboxyl	Serine (S) Threonine (T) Tyrosine (Y) Aspartic acid (D) Glutamic acid (E)
Boc		Amino	Lysine (K) Tryptophan (W) Histidine (H)
Pbf		Guanidino	Arginine (R)
Fmoc		$\alpha$ -amino	Standard amino acids
Trt		Sulfhydryl Carboxyl	Cysteine (C) Aspartic acid (D) Glutamic acid (E)

The chart shows all protecting groups, chemical structure and functional group protection on amino acid residues.

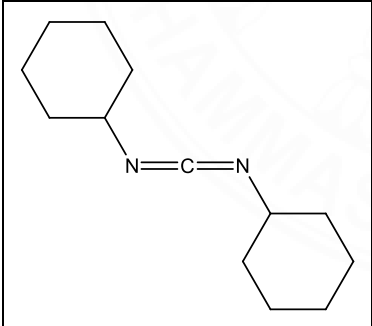
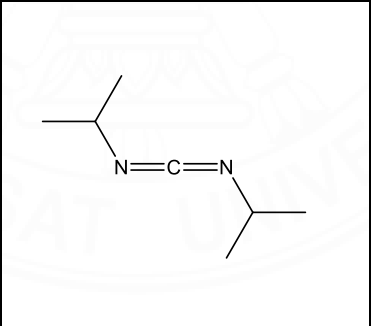
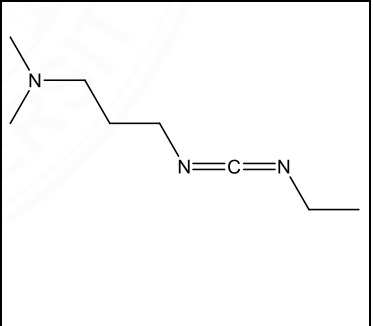
### 2.1.3 Coupling reagents

Generally, amide bond formation between an amine and carboxylic acid is relatively slow. The coupling reagent is usually required for carboxylic acid activation to effective amide bond formation<sup>[28]</sup>. There are several types of coupling reagents, including carbodiimides, phosphonium salts, aminium/uronium salts and others.

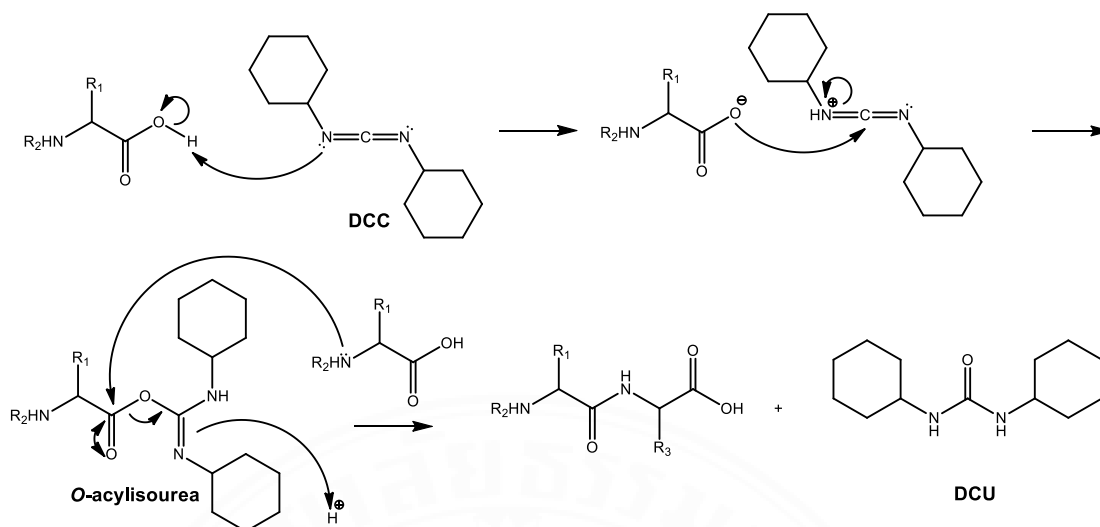
#### 2.1.3.1 Carbodiimides

The carbodiimides<sup>[28-30]</sup> have been widely used for the synthesis of peptide since the eighties to present. The commonly used carbodiimide are shown as follows; (1) *N,N'*-Dicyclohexylcarbodiimide (DCC), (2) *N,N'*-Diisopropylcarbodiimide (DIC) and (3) *N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbonate (EDC). Nowadays, all carbodiimides are often used and some efficient additives were also added to reduce racemization. The chemical structures of carbodiimide are shown in **Table 2.3**.

**Table 2.3** The chemical structure of carbodiimides

		
DCC	DIC	EDC

\*DCC, *N,N'*-Dicyclohexylcarbodiimide; DIC, *N,N'*-Diisopropylcarbodiimide; EDC, *N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbonate.



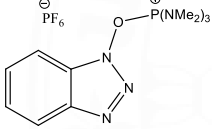
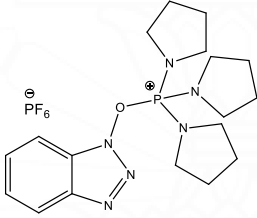
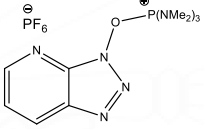
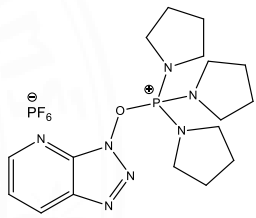
**Scheme 2.1** The proposed peptide bond formation by using DCC as a coupling reagents.

The carboxylic group of amino acid primarily reacted with carbodiimide to form a reactive intermediate *O*-acylisourea (**Scheme 2.1**) which further coupled with the second amino acid to generate a new amide bond. Notably, the disadvantage of using carbodiimide is the formation of side products, *N,N'*-dicyclohexylurea (DCU) in case of DCC and *N,N'*-diisopropylurea in case of DIC.

### 2.1.3.2 Phosphonium salts

The phosphonium salt<sup>[28-30]</sup> is a peptide coupling reagent which is more reactive than carbodiimides. Examples of coupling reagent in this group are: benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate (BOP), benzotriazol-1-yl-oxy-*tris*(pyrrolidino) phosphonium hexafluorophosphate (PyBOP), 7-(Azabenzotriazol-1-yl)oxytris(dimethylamino)phosphonium hexafluorophosphate (AOP), [(7-azabenzotriazol-1-yl)oxy]tris(pyrrolidino)-phosphonium hexafluorophosphate (PyAOP) as shown in Table 2.4.

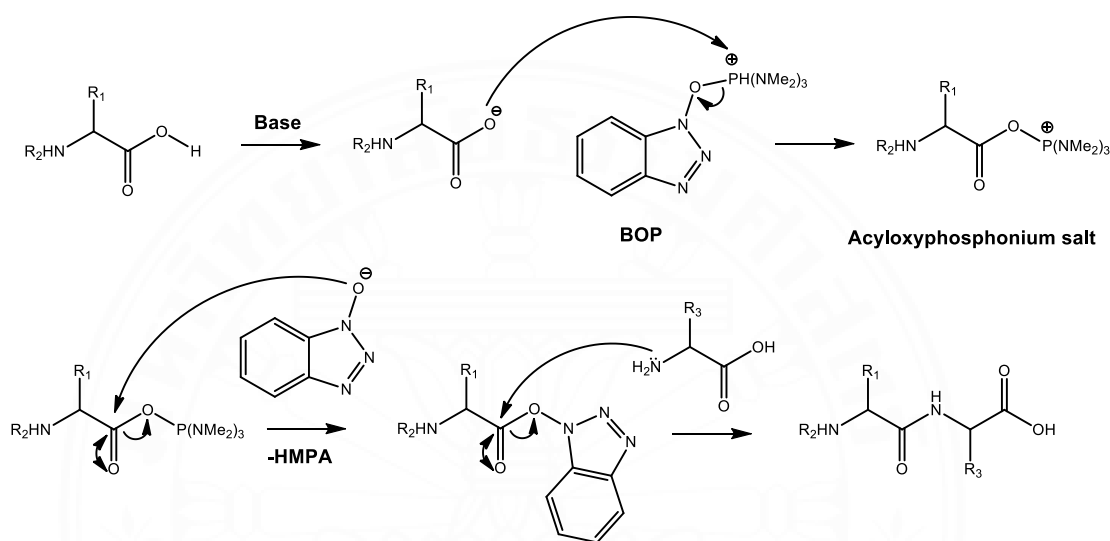
**Table 2.4** The chemical structure of phosphonium salts

			
BOP	PyBOP	AOP	PyAOP

\*BOP, benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate; PyBOP, benzotriazol-1-yl-oxy-*tris*(pyrrolidino) phosphonium hexafluorophosphate; AOP, 7-(Azabenzotriazol-1-yl)oxy tris(dimethylamino)phosphonium hexafluorophosphate; PyAOP, [(7-azabenzotriazol-1-yl)oxy]tris(pyrrolidino)-phosphonium hexafluorophosphate.

BOP was firstly used in 1975, which was developed to generate OBtu esters *in situ* to avoid racemization. This coupling reagent displayed the great advantage for peptide synthesis mainly due to the minimized side reaction from the dehydration of asparagine or glutamine. Unfortunately, BOP generally produced a highly toxic by product named HMPA (hexamethylphosphoramide) which is classified as a carcinogenic compound. Thus, it was replaced by the pyrrolidino derivative called PyBOP.

AOP is also well known as a peptide coupling reagent derived from HOAt (1-Hydroxy-7-azabenzotriazole) with great coupling efficiency. According to the similarity of AOP and BOP structures, it also caused formation the HMPA. Therefore, the structural related reagent named PyAOP is an alternative coupling reagent for the syntheses of peptides.



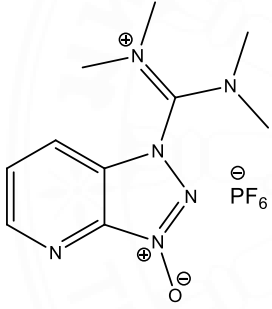
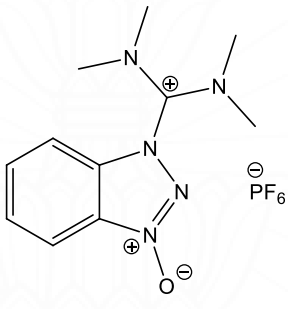
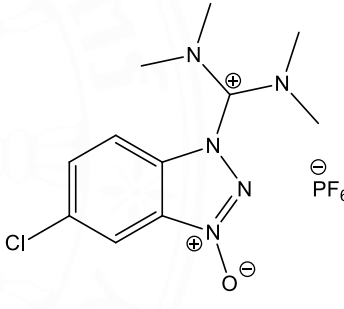
**Scheme 2.2** The proposed peptide bond formation by using BOP as a coupling reagents.

The activation of carboxylic acid by using phosphonium salt (**Scheme 2.2**) requires at least two equivalents of DIPEA. Then, a carboxylate subsequently reacts with phosphonium salt to generate the unstable intermediate acyloxyposphonium salt, which reacts immediately to form phosphonium salts active ester. Next, active species condensed with an amino group to formed new amide bond.

### 2.1.3.3 Uronium/Aminium salts

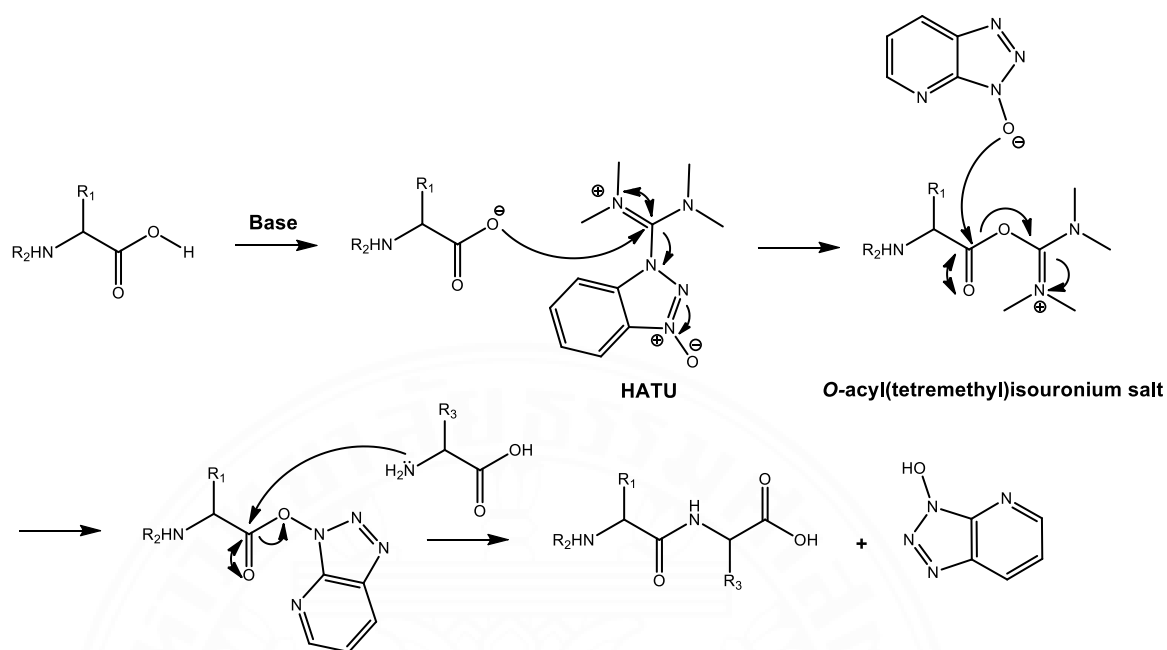
The uronium and aminium salts<sup>[28-30]</sup> are highly effective coupling reagent for peptide synthesis, which bear a positive carbon atom instead of phosphonium residue. The coupling reagents in this group are listed as follows; for example 1 - [Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-bipyridinium-3-oxide hexafluorophosphate (HATU), (2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and 1 - [Bis(dimethylamino)methylene]-5-chlorobenzotriazolium 3-oxide hexafluorophosphate (HCTU)

**Table 2.5** The chemical structure of uronium/aminium salts

		
HATU	HBTU	HCTU

\*HATU, 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-bipyridinium-3-oxide hexafluorophosphate; HBTU, (2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HCTU, 1-[Bis(dimethylamino)methylene]-5-chlorobenzotriazolium 3-oxide hexafluorophosphate.

The first uronium salt in this series is HBTU, derived from HOBt (hydroxybenzotriazole). The HBTU is an efficient coupling reagent with little racemization and is commonly used in standard Fmoc-based solid phase peptide synthesis. The other is HCTU, which is analogous to HBTU, possessing a chlorine at the sixth position in order to improve the coupling efficacy as shown in **Table 2.5**.

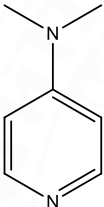
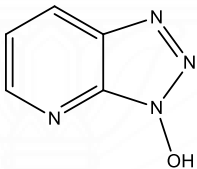
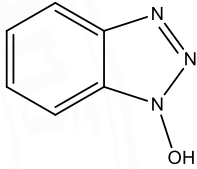


**Scheme 2.3** The proposed peptide bond formation by using HATU as a coupling reagent.

The mechanism of carboxylic acid activation by using uronium salts (**Scheme 2.3**) is similar to that of phosphonium salts. The first step is the generation of carboxylate, which subsequently reacts with uronium salts to form unstable *O*-acyl(tetramethyl)isouronium salt, which rapidly reacted with OAt anion to afford an intermediate OAt-active ester. Then, the addition of an amino group as nucleophile resulted in the amide bond formation.

In chemistry, racemization<sup>[31]</sup> is the reaction that effectively converts an optically active compound into optically inactive compound (racemic). Generally, all amino acids except glycine are optically active, possessing a stereogenic center at their  $\alpha$ -carbon. The racemization has become the major problem which converted an amino acid to have two different D and L configurations. Therefore, the minimization of amino acid racemization during coupling is critical in order to control the stereochemistry of the resulting peptide product.

**Table 2.6** The chemical structure of additives.

		
DMAP	HOAt	HOBt

\*DMAP, 4-Dimethylaminopyridine; HOAt, 1-Hydroxy-7-azabenzotriazole; HOBt, Hydroxybenzotriazole

The activation of amino acids has opened the possibility for racemization and the capacity racemic suppression is greatly related to used coupling reagents. An additive was also added for the syntheses of peptides as an activator to suppress the racemization which in turn increase the coupling efficacy. Examples of additive such as 4-Dimethylaminopyridine (DMAP), 1-Hydroxy-7-azabenzotriazole (HOAt) and Hydroxybenzotriazole (HOBt) are listed in **Table 2.6**.

### 2.1.4 Peptide synthesis strategies

Currently, the interest of pharmaceutical peptides has gained of great interest, which led to the development of new synthetic strategies. Peptides are effectively synthesized by using all of these synthetic strategies as follows; (1) solution phase peptide synthesis (2) Solid phase peptide synthesis (SPPS) or (3) a hybrid approach<sup>[32]</sup>.

#### 2.1.4.1 Solution phase peptide synthesis

Solution phase peptide synthesis<sup>[32-33]</sup>, or widely called liquid phase, is traditionally approach for peptide production in laboratories. The advantages of this approach are valuable for large-scale production and lower cost than other methods. However, this classical strategy is not suitable for larger peptides synthesis because of time consumption, low coupling yield, and essential purification after each step.

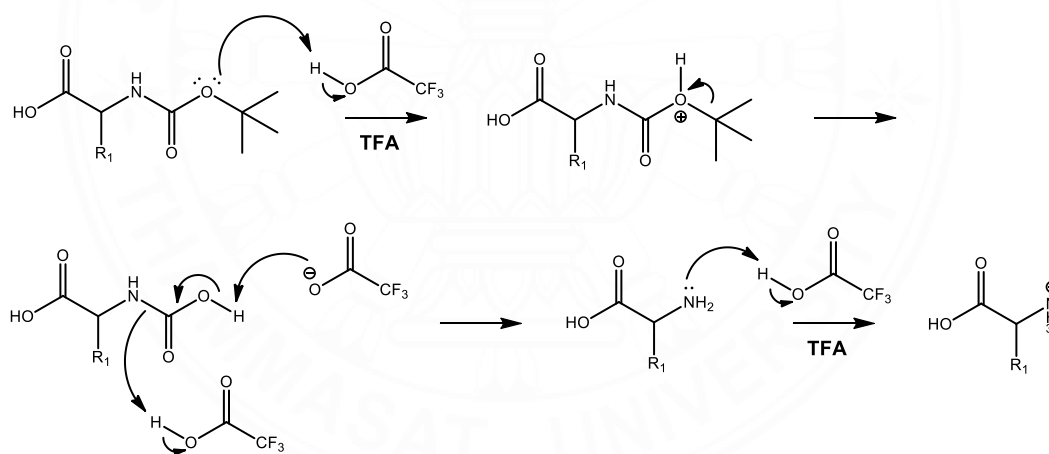
#### 2.1.4.2 Solid phase peptide synthesis (SPPS)

In the early 1960s, solid phase peptide synthesis (SPPS)<sup>[32-33]</sup> was firstly developed by Robert Merrifield by using polystyrene-based solid support for peptide synthesis. His pioneering work made for some massive leap forward in peptide synthesis technology. Currently, SPPS has become the standard technique for peptides and protein production. The SPPS method has several benefits, including the ability to synthesized large products, fast, high efficiency and lower time production (purification only at the end). However, this technique is suitable for small scale peptide synthesis and more expensive than the solution phase. Additionally, when the size of the peptide increases, it can possibly be aggregated during an amino acid assembly.

The solid phase peptide synthesis is a process which performed *via* step-by-step using selective deprotection of protecting groups of *N*-terminal amino group, and followed by the coupling of an amino acid. The synthesis is repeatedly assembled on the solid support for peptide chain elongation.

### 2.1.4.2.1 Boc strategy

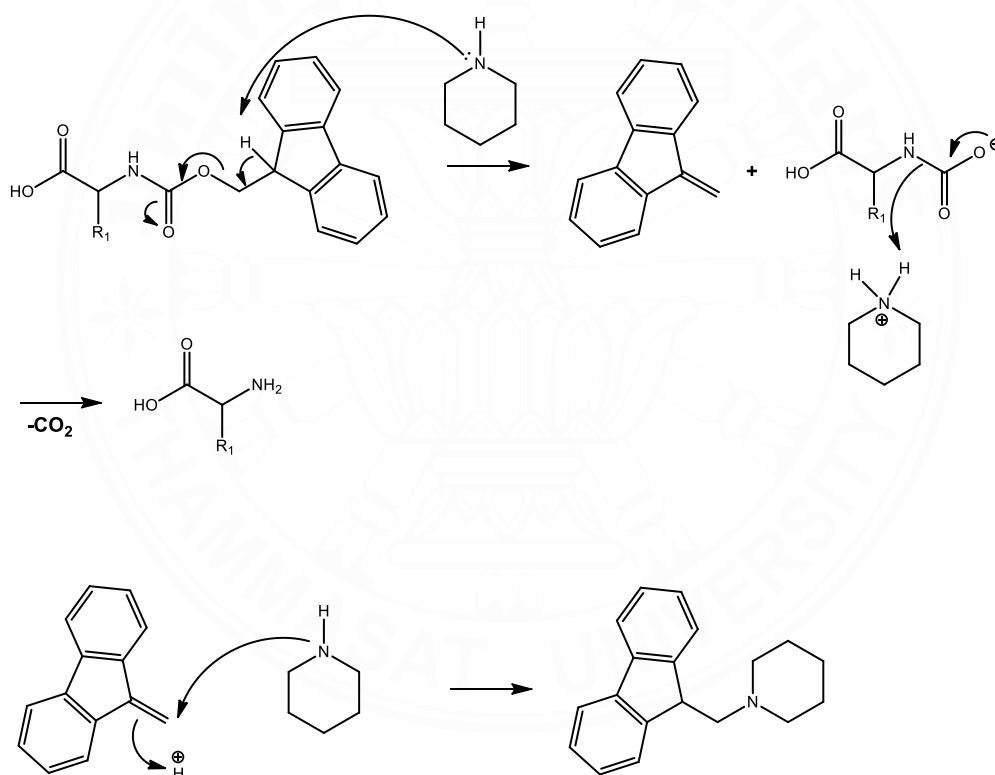
The Boc/Bzl SPPS<sup>[32-34]</sup> is a method for peptide synthesis stepwise from C to N terminus in which Boc is used as a temporary *N*-terminal protection. Benzyl-based protecting group is also utilized as a protecting group for the synthesis of peptide *via* Boc-strategy, and it was stable in moderate acid and sensitive in the presence of strong acid. Both Boc and benzyl based protecting groups are simultaneously cleaved by using hydrogen fluoride (HF). This strategy is preferred for the synthesis of the peptide containing base-sensitive protecting moieties. However, the *N*-protection and side chain protection groups are both acid labile. Therefore, the protecting groups on the side chain can also be deprotected during TFA treatment. Additionally, HF is a highly corrosive chemical, and is required for the global deprotection. The mechanism of Boc deprotection is shown in **Scheme 2.4**.



**Scheme 2.4** The mechanism of Boc deprotection

### 2.1.4.2.2 Fmoc strategy

The Fmoc-based solid phase peptide synthesis<sup>[32-34]</sup> was widely utilized more than Boc/Bzl SPPS in which 9-fluorenylmethoxycarbonyl (Fmoc) protecting group at *N*-terminal was used, and allowed the deprotection under mild basic conditions (usually in 20% piperidine in DMF). The mechanism of Fmoc deprotection is shown in **Scheme 2.5**. Examples of side chain protecting groups for Fmoc strategy are listed as follows; (1) Boc, (2) <sup>t</sup>Bu (3) Pbf and (4) Trt. The final peptide was globally deprotected by using TFA.



**Scheme 2.5** The mechanism of Fmoc deprotection

Typically, solid phase peptide synthesis (SPPS) approach<sup>[34]</sup> is consisted of three main steps as follows: (1) First amino acid loading (2) An amino acid assembly and (3) The peptide chain cleavage.

### **(1) First amino acid loading**

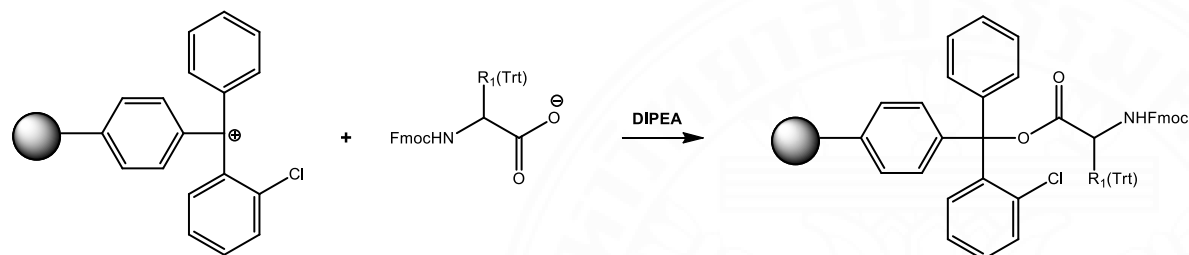
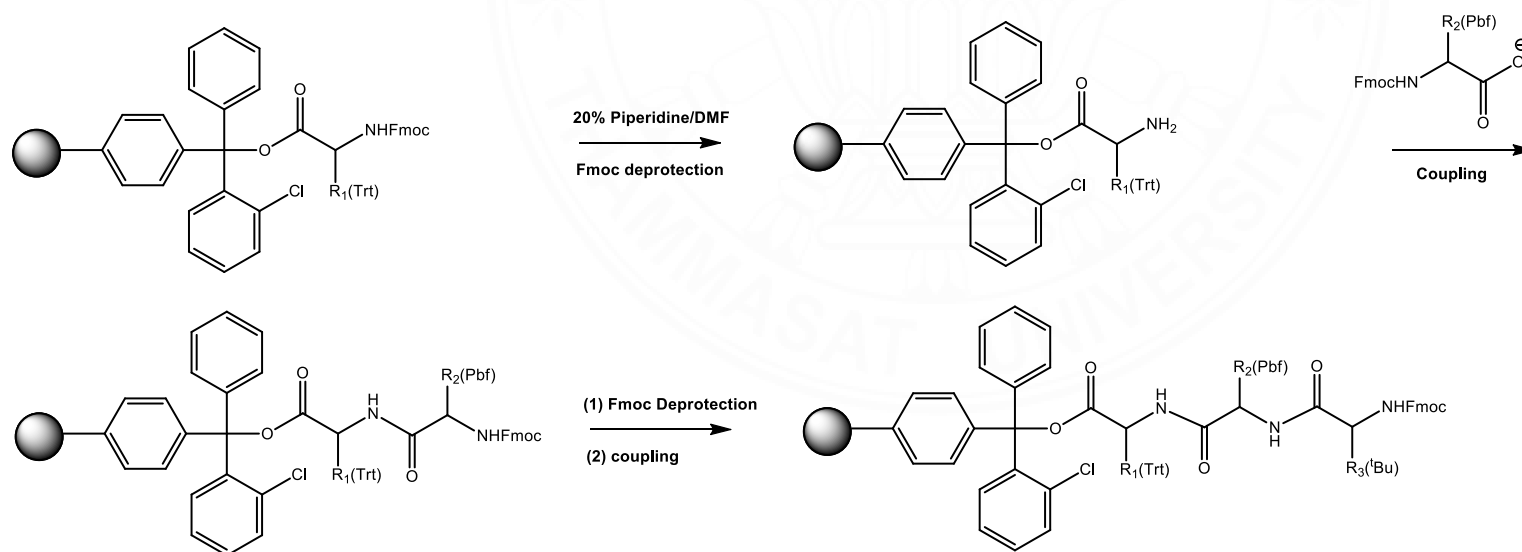
The first amino acid is initially attached to a solid support. As shown in **Scheme 2.6a**, the first amino acid is loaded onto 2-chlorotrityl resin. Importantly, the first amino acid attachment is the key reaction that significantly affected an overall yield.

### **(2) An amino acid assembly**

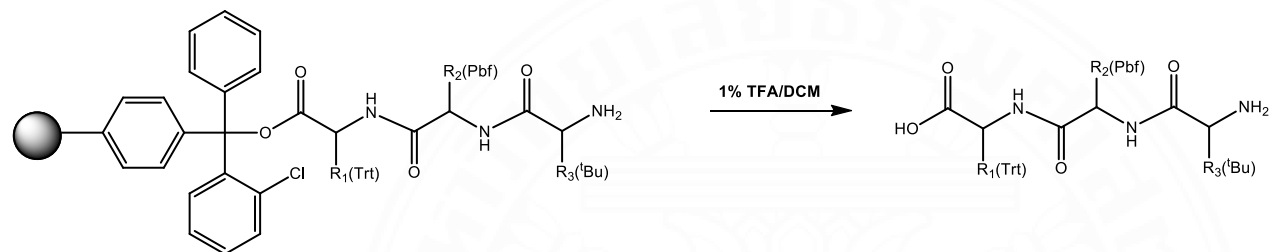
To this step, the peptide chain is assembled stepwise. The Fmoc group of the first amino acid is removed by piperidine to afford free amino group at *N*-terminus, which is subsequently coupled with carboxylic group of the second amino acid to yield the peptide bond on a solid support. The peptide chain is elongated by repeating the synthesis cycles as shown in **scheme 2.6b**. The unreacted reagents from each step and by-product are successfully removed, while the desired peptide is attached to the linker.

### **(3) The peptide chain cleavage**

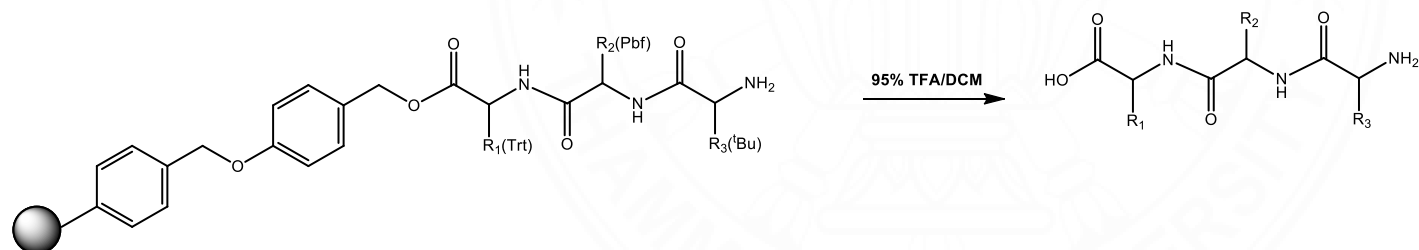
Finally, the peptide chain is globally cleaved from the solid support by using the deprotection cocktail which is suitable for individual solid support. For example, 2-chototrityl resin was cleave by using 1% TFA in DCM in order to obtain a fully protected peptide as shown in **Scheme 2.6c**. While, wang resin was cleaved by using 95% TFA to afford the fully unprotected peptide as shown in **Scheme 2.6d**

**(a) The first amino acid loading****(b) An amino acid assembly**

(c) The peptide chain cleavage from 2-chlorotrityl resin



(d) The peptide chain cleavage from wang resin



**Scheme 2.6** Fmoc-based solid phase peptide synthesis (SPPS)

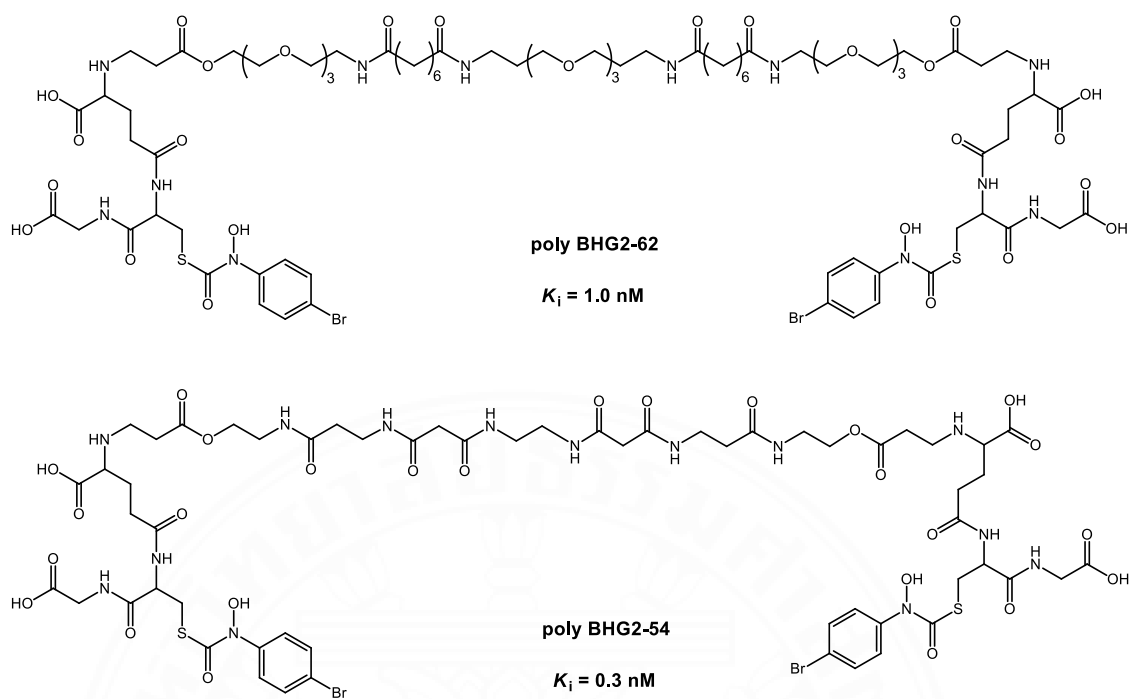
## 2.2 Structure modification of peptides

A structure modification (re-engineering) of peptides<sup>[35]</sup> is a useful technique for structural alteration of both new and known peptides in order to expand some physiochemical properties; for example, specificity, potency, cell-permeability, solubility, and toxicity. Generally, the most popular method to enhance bioactivity is modification monovalent counterparts to bivalent. The bivalent structure can be designed to homobivalent, which targeted two of the same receptor types. On the other hand, they were modified to be heterobivalent and binding with two different receptor types. This approach allowed the bivalent structures have more efficiency compared to their monovalent counterparts.

### 2.2.1 Bivalent inhibitor

Yanku Sang<sup>[36]</sup> reported two competitive bivalent inhibitors of a zinc metalloenzyme glyoxalase I (GlxI), involving in the catalyzation of glutathione-dependent inactivation of cytotoxic methylglyoxal in tumor tissue activities. Inhibitors were synthesized as bivalent moieties based on the transition state of analog *S*-(*N*-bromiphenyl-*N*-hydroxybarbomoyl) glutathione (BHG), cross-linking by optimized spacer unit.

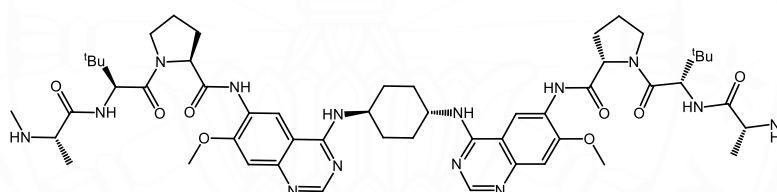
These two bivalent inhibitors (polyPHG2-62 and polyBHG2-54 (**Figure 2.2**) showed the binding ability to hGlxI with  $K_i = 0.1$  nM and 0.3 nM, respectively. Additionally, polyPHG2-62 exhibited higher potent than that of monomer BHG counterpart ( $K_i = 14$  nM).



**Figure 2.2** Structure of poly BHG2-62 and poly PHG2-54 inhibitors

Inhwan Bae and co-workers<sup>[37]</sup> synthesized new antagonists as an inhibitor of apoptosis proteins (IAPs), a group of proteins that block programmed cell death, by using a quinazoline-based structure. The antagonists were designed to be bivalent structure which consisted of three main parts; quinazoline scaffold, diamine linker, and tripeptide mimicked the key sequences of second mitochondria-derived activator of caspases (SMAC), which was known as natural IAP antagonists.

According to the investigation, the re-engineering of bivalent scaffold provided a strong cellular activity with no toxicity observed in normal mouse tissues. The great example of bivalent antagonist based on *trans*-1,4-diaminocyclohexane linker as shown in **Figure 2.3** inhibited an IAP family in MDA-MB-231 tumor tissue (with  $IC_{50} = 0.14$  nM) with higher cellular activity compared to that of monovalent ( $IC_{50} = 19$  nM). Additionally, a quinazoline-based antagonist showed the cellular potency over several cancer cell lines on the pancreas, ovary, breast, and bladder in the nanomolar range ( $IC_{50}$  0.1-27.8 nM) as shown in **Table 2.7**

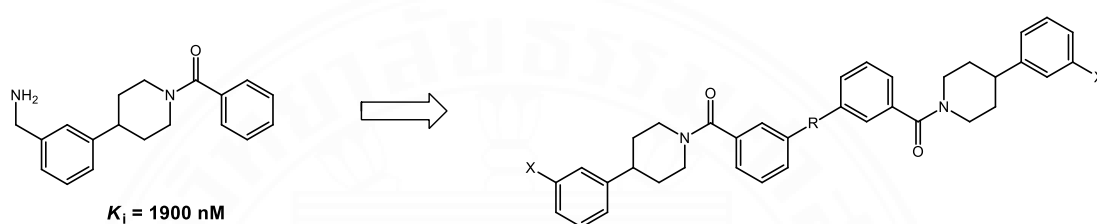


**Figure 2.3** Structure of IAPs antagonist based on quinazoline scaffold

**Table 2.7** An *in vitro* evaluation of quinazoline-base structure linked by *trans*-1,4-Diaminocyclohexane

Cell line	Tumor Type	$IC_{50}$ , nM
MDA-MB-468	Breast	27.8
MCF-7	Breast	19.2
OVCAR-3	Ovary	0.2
Bx-PC3	Pancreas	0.2
Panc-1	Pancreas	0.9
MaiPaCa-2	Pancreas	0.1
J82	Bladder	17.4

Guyan Liang <sup>[38]</sup> published the synthesis of dimeric inhibitors series that exhibited tight-binding characteristics with human  $\beta$ -tryptase (**Figure 2.4**). The bivalent  $\beta$ -tryptase inhibitors presented a high potential to span two adjacent substrate-binding sites with  $K_i$  value in subnanomolar concentrations (0.29-3 nM) as shown in **Table 2.8**, which was more potent than that of monomeric counterpart ( $K_i = 1900$  nM). The synergetic binding effect can significantly improve the binding affinity of bivalent inhibitors which were confirmed by more than 1000-fold enhancement in  $K_i$  values.



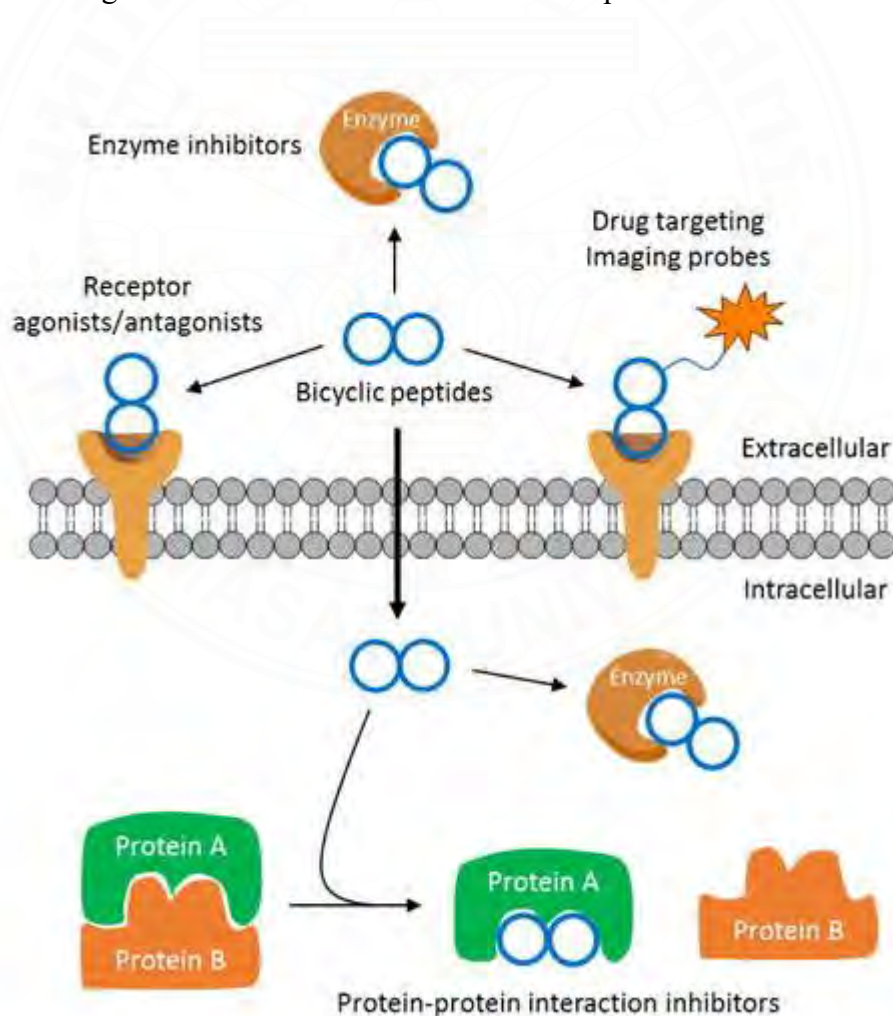
**Figure 2.4** Structure of monomeric and dimeric human- $\beta$ -tryptase inhibitors

**Table 2.8** An *in vitro* potency for  $\beta$ -tryptase of dimeric inhibitor series

X	R	$\beta$ -tryptase ( $K_i$ , nM)
		0.34
		2.0
		3.0
		0.29
		0.3

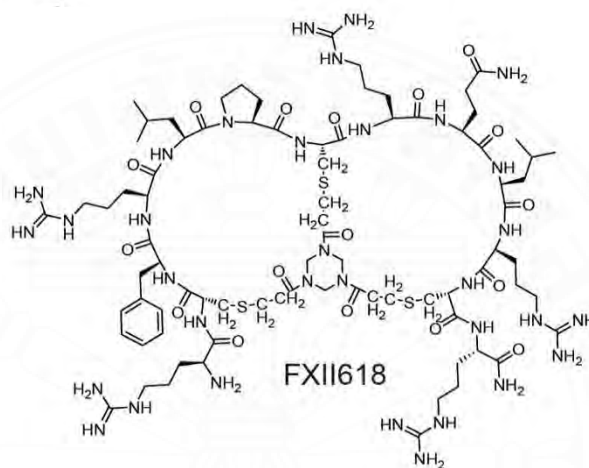
### 2.2.2 Bicyclic peptide

Bicyclic peptides<sup>[39]</sup> are of remarkable next-generation bioactive compounds as a potential drug targets with diverse applications; for example, enzyme inhibitors, receptors agonist/antagonist, or protein-protein interaction targets (**Figure 2.5**). Bicyclic scaffold development aimed to improve the drawbacks of the monocyclic peptide counterpart, particularly the susceptibility to proteolytic degradation. The increase of rigid conformation and the metabolic stability also improved their target affinity and selectivity. Moreover, the bicyclic system also permitted the two rings to function independently; such as, one ring binds with a specific target while another involved the cellular penetration.

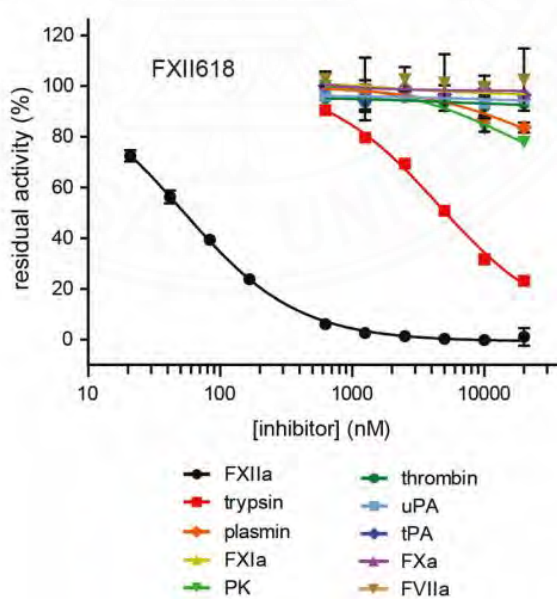


**Figure 2.5** The bicyclic peptide for therapeutic applications<sup>[39]</sup>

Vanessa Baeriswyl et al.<sup>[40]</sup> discovered the novel macrocycle peptide *via* bicyclic libraries screening (**Figure 2.6**). The synthetic inhibitor exhibited a high affinity against FXIIa (activated FXII) with an inhibitory constant  $K_i$  of 22 nM and selectivity of >2000-fold over other proteases (**Figure 2.7**). Additionally, the newly bicyclic peptide FXIIa inhibitor which was named FXII618 is 60-fold more potent and selective than the best FXIIa inhibitors ( $IC_{50} = 4.4 \mu M$ ).

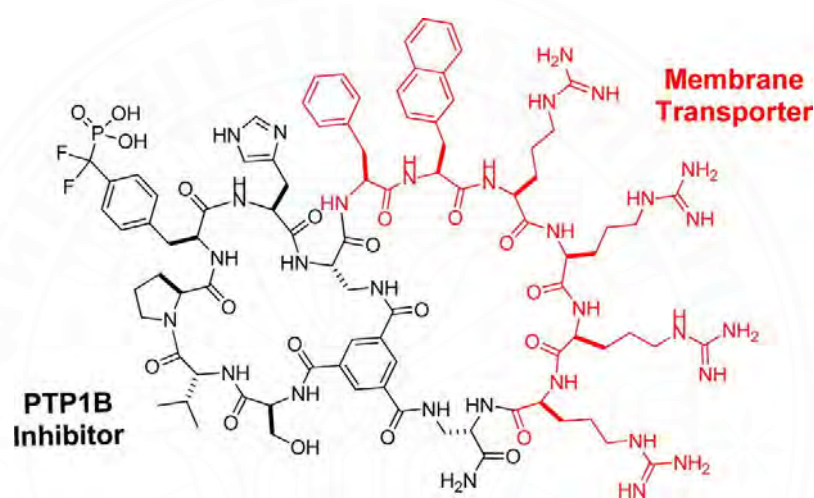


**Figure 2.6** The structure of bicyclic peptide coagulation factor XII inhibitor<sup>[40]</sup>



**Figure 2.7** The specificity profile and inhibitory activity of FXII618 against several serine protease enzymes.<sup>[40]</sup>

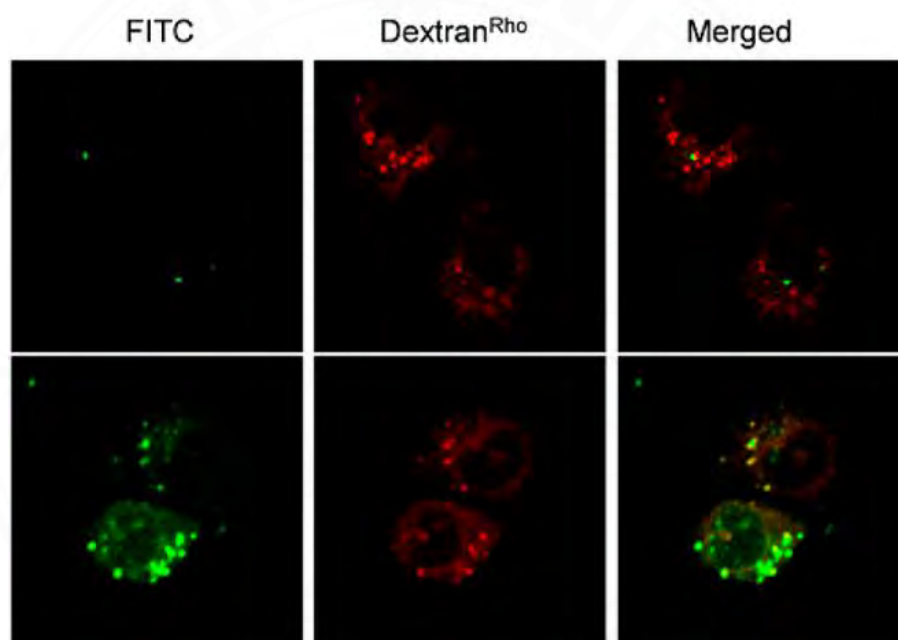
In 2014, Lian et al.<sup>[41]</sup> reported the discovery of a monocyclic peptide inhibitor against a protein-tyrosine phosphatase 1B (PTP1B) *via* a combinatorial screening. Unfortunately, the monocyclic inhibitor showed a poor membrane permeability. Therefore, this inhibitor was designed, and developed based on a bicyclic scaffold which is comprised of a cell-penetrating motif (Arg-Arg-Arg-Arg-Nal-Phe) where Nal is L-naphthylalanine, in the first ring whereas the second ring is comprised of target-binding sequence as shown in **Figure 2.8**.



**Figure 2.8** The structure of PTP1B inhibitor<sup>[41]</sup>

To confirm a cell-penetration ability, the bicyclic inhibitor was labelled with FITC and then treated with A549 lung cancer cells. The treated cell was detected by a live-cell confocal microscopy and obviously showed the fluorescence diffusion throughout the nucleus and cytoplasm more than that of monocyclic inhibitor as shown in **Figure 2.9**, indicating that this inhibitors efficiently entered to the mammalian cells.

Interestingly, the bicyclic inhibitor significantly improved the cell permeability while almost PTP1B inhibitory activity at the nanomolar concentration ( $IC_{50} = 30$  nM) were kept.



**Figure 2.9** The Live-cell confocal microscopy images of A549 cells after treatment with FITC-labeled monocyclic (top) and bicyclic inhibitors (bottom)<sup>[41]</sup>

### 2.2.3 SFTI-1 modification

The extremely high potent and very small size of SFTI-1 make this inhibitor an attractive template for peptide engineering. The main factor that has sparked interest in SFTI-1 is its natural target, the serine proteases involved in diverse critical roles in human physiology. Accordingly, SFTI-1 has gained of great interest in developing pharmaceutical inhibitors against numerous serine proteases<sup>[42-49]</sup> as shown in **Table 2.9**.

**Table 2.9** The inhibitory activities of native SFTI-1 and monocyclic SFTI-1 against several serine proteases

Enzyme	Bicyclic SFTI-1 (Native) K <sub>i</sub> (nM)	Monocyclic SFTI-1 K <sub>i</sub> (nM)
Trypsin	0.1	0.02698
Cathepsin G	0.15	NR
KLK14	25.1	NR
Matriptase	102	703
KLK15	143	NR
Matriptase-2	218	1365
Mesotrypsin	4960	NR
Chymotrypsin	7400	NR
Elastase	105,000	NR
Thrombin	136,000	NR

NR = not reported

### 2.2.3.1 The residues modification

The simple modification is the replacement of an amino acid at the P1 position, the specific recognition of the cleaved peptide bond. The P1 residue of SFTI-1 is Lys, which is favored by protease trypsin-like specificity but relatively weak in chymotrypsin and elastase (**Table 2.9**). Although Lys and Arg are positively charged amino acids, the replacement of Lys with Arg also exhibited the high affinity in the same magnitude (nM), but the inhibitory activities are lower by 7-15 fold (**Table 2.10**)<sup>[46,50]</sup>.

To redirect the inhibitor's activity to other serine proteases, such as the chymotrypsin family, only one amino acid sequence was changed at the P1 position from Lys to Phe without the need to alter its structure. According to several studies<sup>[42-43,51-52]</sup> (**Table 2.11**), a single lysine replacement to phenylalanine or 4-substituent L-phenylalanine improved the bioactivity against chymotrypsin up to 9800-fold compared to the parent structure.

**Table 2.10** The amino acid sequences and inhibitory activities of native SFTI-1 (Gray box) and SFTI-1 analogues against trypsin

SFTI-1 sequence position														K <sub>i</sub> (nM)	Fold change
1	2	3	4	5	6	7	8	9	10	11	12	13	14		
& <sup>1</sup> G	R	C(& <sup>2</sup> )	T	K	S	I	P	P	I	C(& <sup>2</sup> )	F	P	D& <sup>1</sup>	0.1	-
& <sup>1</sup> G	R	C(& <sup>2</sup> )	T	R	S	I	P	P	I	C(& <sup>2</sup> )	F	P	D& <sup>1</sup>	13.4	-13
G	R	C(& <sup>2</sup> )	T	R	S	I	P	P	I	C(&)	F	P	D	15.2	-15
& <sup>1</sup> G	T	C(& <sup>2</sup> )	T	R	S	I	P	P	I	C(& <sup>2</sup> )	N	P	N& <sup>1</sup>	0.7	-7

The P1 position is highlighted in blue color. Modified positions are marked in red color. The cyclization is indicated by (&)

**Table 2.11** The amino acid sequences and inhibitory activities of native SFTI-1 (Gray box) and SFTI-1 analogues against chymotrypsin

SFTI-1 sequence position														K <sub>i</sub> (nM)	Fold change
1	2	3	4	5	6	7	8	9	10	11	12	13	14		
& <sup>1</sup> G	R	C(& <sup>2</sup> )	T	K	S	I	P	P	I	C(& <sup>2</sup> )	F	P	D& <sup>1</sup>	7400	-
G	R	C	T	F	S	I	P	P	I	C	F	P	D	0.5	9800
G	R	C	T	4F	S	I	P	P	I	C	F	P	D	0.03	163000

\*4F = 4-substituent L-phenylalanine

The P1 position is highlighted in blue color. Modified positions are marked in red color. The cyclization is indicated by (&).

Fittler and co-workers<sup>[53-54]</sup> studied and developed a number of highly potent inhibitors of human matriptase based on the SFTI-1. The characteristic of inhibitors that impacted to the inhibitory activity of matriptase was shown in **Table 2.12**.

Firstly, the amino acid residues Ile and Arg at positions 10 and 12 were replaced with Phe and His. This re-engineered variant was named SFTI-1-derived matriptase inhibitor-1 (**SDMI-1**), which exhibited a higher inhibitory activity with  $K_i$  of 11.2 nM than native and monocyclic SFTI-1 ( $K_i = 703$  nM and 102, respectively). According to their studies, the additional favorable interactions between SDMI-1 and matriptase were observed. The side chain of Asp96 was re-oriented to form a bidentate hydrogen bond with a guanidinyll functionality of Arg10. Furthermore, the possibility for a proton donor-acceptor interaction between the backbone carbonyl oxygen of Asp96 and nitrogen of His12 was detected. According to these two reasons, the **SDMI-1** showed a higher inhibitory activity.

Secondly, the presence of a positive charge at the N-terminus is required for the maintenance of inhibitory activity. The glycine residue at position one was replaced with lysine (**SDMI-3**,  $K_i = 2.1$  nM), and exhibited a 5-fold increased an inhibitory activity compared to that of **SDMI-1** ( $K_i = 11.2$  nM).

Lastly, Compound **2** and **3** are the **SDMI-3** derivatives with head-to-tail and tail-to-side chain cyclized motif, respectively. The bicyclic **2** and **3** showed a similar inhibition constant ( $K_i = 4.1$  nM) compared with **SDMI-3** ( $K_i = 2.1$  nM). These results indicated that the cyclization did not significantly impact the inhibitory activity.

**Table 2.12** The amino acid sequences and inhibitory activities of native SFTI-1 (Gray box) and SFTI-1 analogues against matriptase

SFTI-1 sequence position														K <sub>i</sub> (nM)	Cpd
1	2	3	4	5	6	7	8	9	10	11	12	13	14		
& <sup>1</sup> G	R	C(& <sup>2</sup> )	T	K	S	I	P	P	I	C(& <sup>2</sup> )	F	P	D& <sup>1</sup>	102	<b>SFTI-1</b>
G	R	C(& <sup>2</sup> )	T	K	S	I	P	P	I	C(& <sup>2</sup> )	F	P	D	703	<b>αSFTI-1</b>
& <sup>1</sup> G	R	C(& <sup>2</sup> )	T	K	S	I	P	P	R	C(& <sup>2</sup> )	H	P	D& <sup>1</sup>	3.6	<b>1</b>
G	R	C(& <sup>2</sup> )	T	K	S	I	P	P	R	C(& <sup>2</sup> )	H	P	D	11.2	<b>SDMI-1</b>
K	R	C(& <sup>2</sup> )	T	K	S	I	P	P	R	C(& <sup>2</sup> )	H	P	D	2.1	<b>SDMI-3</b>
& <sup>1</sup> K	R	C(& <sup>2</sup> )	T	K	S	I	P	P	R	C(& <sup>2</sup> )	H	P	D& <sup>1</sup>	4.1	<b>2</b>
K(& <sup>1</sup> )	R	C(& <sup>2</sup> )	T	K	S	I	P	P	R	C(& <sup>2</sup> )	H	P	D& <sup>1</sup>	4.1	<b>3</b>

The P1 position is highlighted in blue color. Modified positions are marked in red color.

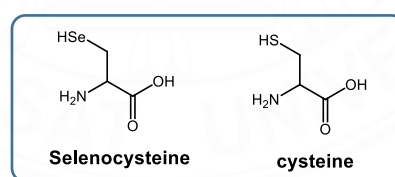
Compound 3 is tail-to-side-chain cyclized. The cyclization is indicated by (&)

αSFTI-1 = monocyclic SFTI-1 analogue

### 2.2.3.2 Disulfide bridge modification

To date, the remarkable high potency of SFTI-1 has inspired the development of new drug scaffold. In naturally occurring SFTI-1, the linkage is the presence of a disulfide bond, which plays an essential structural role for the proteinase inhibitory ability. However, the disadvantage of disulfide linkage is its redox instability. Several studies have used SFTI-1 as a framework and installed a non-natural linkage for the investigation of the structural change and its inhibitory activity. A variety of disulfide mimetics have been implemented in SFTI-1 scaffold; for example, diselenide, carbonyl, and dicarba bridge<sup>[42,55-57]</sup>

Xin Guo<sup>[58]</sup> published the synthesis and biological activity of monocyclic SFTI-1 analogues in which a disulfide bond was replaced with a diselenide bridge. The seleno SFTI-1 was synthesized by the Fmoc chemistry *via* solid-phase peptide synthesis using selenocysteine (Sec) instead of cysteine and using HBTU/HOBt as coupling reagents. Under the cleavage condition (trifluoroacetic acid: triisopropylsilane: thioanisole: water: dichloromethane 88:5:5:2:3), the cleavage, global deprotection, and diselenide bond formation (-Se-Se-) were simultaneously taken place in a single step without any oxidation reaction.



**Figure 2.10** The structure of selenocysteine and cysteine amino acids

**Table 2.13** The amino acid residues of monocyclic SFTI-1 and its derivatives.

The cyclization is indicated by (&)

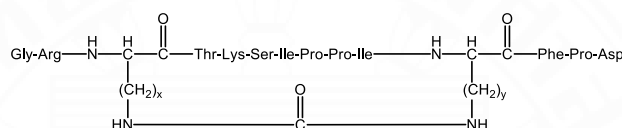
Bridge type	The amino acid sequences
diselenide	Gly-Arg-Sec(&)-Thr-Lys-Ser-Ile-Pro-Pro-Ile-Sec(&)-Phe-Pro-Asp
disulfide	Gly-Arg-Cys(&)-Thr-Lys-Ser-Ile-Pro-Pro-Ile-Cys(&)-Phe-Pro-Asp
absence	Gly-Arg-Ser-Thr-Lys-Ser-Ile-Pro-Pro-Ile-Ser-Phe-Pro-Asp

In the comparison (Table 2.14), the monocyclic SFTI-1 and [Ser<sup>3,11</sup>] αSFTI-1 analogues were also designed, and synthesized. The trypsin inhibition ability of the absence of disulfide bridge analogues ([Ser<sup>3,11</sup>] αSFTI-1) is significantly reduced three orders of magnitude, indicating that the disulfide bond is crucial for the structure and its inhibitory function. However, seleno αSFTI-1 exhibited the high potency in the same order of magnitude with an inhibitory constant of 6.5 nM, but the constant is reduced to 30% compared with that of monocyclic SFTI-1.

**Table 2.14** Trypsin inhibitory constants of monocyclic SFTI-1 and Seleno SFTI-1 αSFTI-1 = monocyclic SFTI-1 analogues

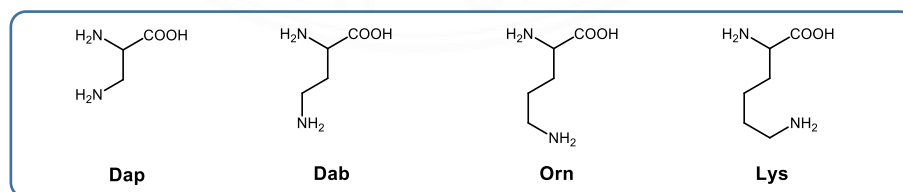
Synthetic peptide	Trypsin inhibitory constant [ $K_i$ ]
Seleno αSFTI-1	$6.50 \pm 0.49 \times 10^{-9}$
αSFTI-1	$1.96 \pm 0.20 \times 10^{-9}$
[Ser <sup>3,11</sup> ] αSFTI-1	$8.10 \pm 0.81 \times 10^{-6}$

Legowska et al.<sup>[59]</sup> reported the synthesis of monocyclic SFTI-1 analogues series (**Figure 2.11**). The disulfide bridge between Cys3 and Cys11 was replaced by the different size of carbonyl bridges. The positions of 3 and/or 11 were changed to Lysine (Lys), Ornithine (Orn), L-2,4-diaminobutyric acid (Dab), or L-2,3-diaminopropionic acid (Dap), which are the dibasic amino acid residues starting with Dap (one CH<sub>2</sub> group) up to Lys (four CH<sub>2</sub> groups). The SFTI-1 analogues were synthesized *via* a Fmoc chemistry solid-phases method using 2-cholorotrityl chloride resin, followed by cyclization on-resin. The new carbonyl bridge was achieved by the amino groups of the side chains of Dap, Orn, or Lys.



**Figure 2.11** The structure of SFTI-1 analogues with carbonyl bridge, where  $x, y = 1-4$  for Dap, Dab, Orn, and Lys, respectively

When the disulfide bond was replaced by the carbonyl bridge, all analogues presented the high inhibitory activity against Bovine  $\beta$ -trypsin with the same order of magnitude (**Table 2.15**) relative to the reference monocyclic SFTI1 ( $K_a = 9.9$  nM), indicating that the carbonyl bridge could be utilized as redox stability scaffold. According to the study, the length of the formed cycle did not significantly impact to the inhibitory activity.



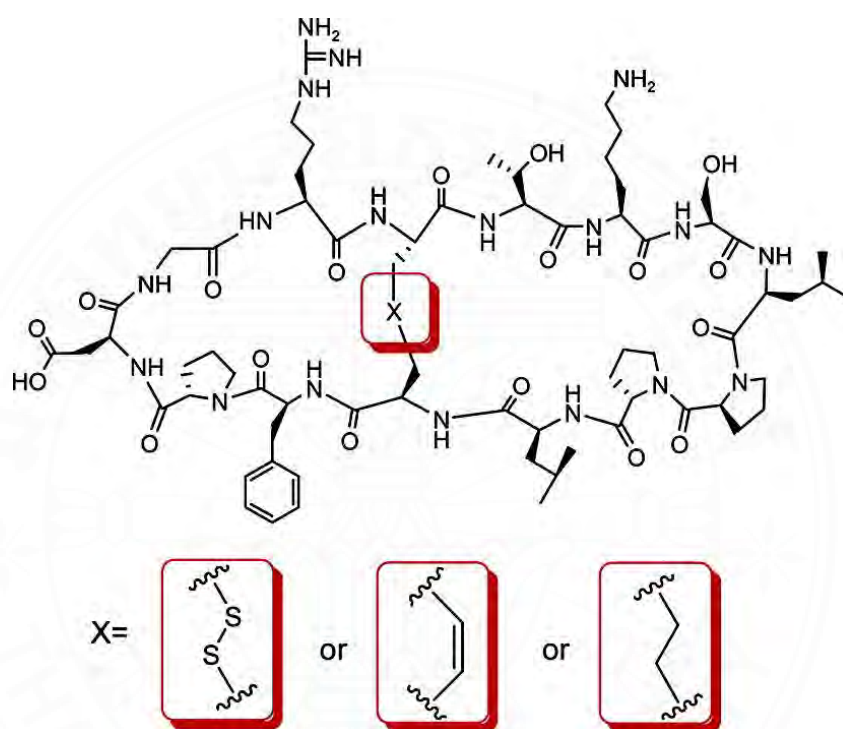
**Figure 2.12** The structure of L-2,3-diaminopropionic acid (Dap), L-2,4-diaminobutyric acid (Dab), Ornithine (Orn) and Lysine (Lys) amino acids

**Table 2.15** Physiochemical properties and association equilibrium constant ( $K_a$ ) with Bovine  $\beta$ -trypsin of SFTI-1 analogues containing carbonyl bridge

Synthetic peptide	Number of Chemical Bonds	$K_a$ (nM)
SFTI-1 native	5	$0.11 \pm 0.2$
([Ser <sup>3,11</sup> ])SFTI-1 dicyclic	6	$80.7 \pm 0.81$
[Dap <sup>3,11</sup> ] $\alpha$ SFTI-1	6	$34.6 \pm 0.27$
[Orn <sup>3,11</sup> ] $\alpha$ SFTI-1	10	$54.2 \pm 0.52$
[Lys <sup>3,11</sup> ] $\alpha$ SFTI-1	12	$59.1 \pm 0.51$
[Lys <sup>3</sup> ,Orn <sup>11</sup> ] $\alpha$ SFTI-1	11	$2.06 \pm 0.19$
[Orn <sup>3</sup> ,Lys <sup>11</sup> ] $\alpha$ SFTI-1	11	$1.03 \pm 0.12$
[Lys <sup>3</sup> ,Dab <sup>11</sup> ] $\alpha$ SFTI-1	10	$88.5 \pm 1.19$
[Lys <sup>3</sup> ,Dap <sup>11</sup> ] $\alpha$ SFTI-1	9	$3.89 \pm 0.71$
[Orn <sup>3</sup> ,Dap <sup>11</sup> ] $\alpha$ SFTI-1	9	$35.3 \pm 0.73$

$\alpha$ SFTI-1 = monocyclic SFTI-1 analogues

Sheng Jiang and co-workers<sup>[60]</sup> reported the design and synthesis of potent matriptase inhibitors based on the naturally occurring SFTI-1. The bicyclic inhibitors, SFTI-2 and SFTI-3 (**Figure 2.13**) were designed by mimicking an intracellular disulfide bridge with dicarba linker (olefin for SFTI-2 and ethylene for SFTI-3) in order to enhance the metabolic stability.



**Figure 2.13** Structure of SFTI-1, SFTI-2 and SFTI-3

The linear peptide precursor of SFTI-2 and SFTI-3 were synthesized by using Fmoc-based solid phase peptide synthesis. The cysteines 3 and 11 were replaced with L-allylglycine. Then, a fully protected linear peptide attached on resin was treated with Grubb's catalyst [(PCy<sub>3</sub>)-(Im(Mes)<sub>2</sub>)Ru=CHPh] to perform the ring closure metathesis (RCM). The protected olefin-containing macrocycle was cleaved from the resin and subsequently macrocyclization by using HATU as a coupling reagent. Finally, the protected bicyclic peptide was globally deprotected to afford SFTI-2. Then, SFTI-3, and SFTI-2 were treated with Pd-black as a catalyst for the hydrogenolysis.

**Table 2.16** Relative matriptase inhibitory activity of SFTI-1 and its derivatives.

compound	Bridge type	Relative matriptase inhibitory activity [ $K_i$ (nM)]
SFTI-1	Disulfide bridge	1
SFTI-2	Olefin bridge	25
SFTI-3	Ethylene bridge	2.5

The matriptase inhibitory activities were evaluated by using a binding assay as shown in **Table 2.16**. The ethylene bridge analog SFTI-3 showed an inhibitory potency similar to parent SFTI-1, whereas the olefin bridge analog SFTI-2 was 25-fold lower than its native.

## CHAPTER 3

### RESEARCH METHODOLOGY

#### 3.1 Chemicals and Instrument

##### 3.1.1 Chemicals

- (1) 2,6-Diaminopimelic acid
- (2) *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-glycine, Fmoc-Gly-OH
- (3) *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-*N'*-2,2,4,6,7-pentamethyl dihydrobenzofuran-5-sulfonyl-L-arginine, Fmoc-Arg(Pbf)-OH
- (4) *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-S-trityl-L-cysteine, Fmoc-Cys(Trt)-OH
- (5) *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-O-*tert*-butyl-L-threonine, Fmoc-Thr-OH
- (6) *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-O-*tert*-butyl-L-serine, Fmoc-Ser(<sup>t</sup>Bu)-OH
- (7) *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-*N*-epsilon-*tert*-butyloxy carbonyl-L-lysine, Fmoc-Lys(Boc)-OH
- (8) *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-L-isoleucine, Fmoc-Ile-OH
- (9) *N,N'*-Dicyclohexylcarbodiimide , DCC
- (10) *N,N'*-Diisopropylcarbodiimide , DIC
- (11) *N,N*-Dimethylpyridine-4-amine , DMAP
- (12) Hydroxybenzotriazole , HOBt
- (13) (2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate , HBTU
- (14) 1-[Bis(dimethylamino)methylen]-5-chlorobenzotriazolium 3-oxide hexafluorophosphate , HCTU
- (15) 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium-3-oxid hexafluorophosphate , HATU

- (16) *N,N*-Diisopropylethylamine , DIPEA
- (17) Trifluoroacetic acid , TFA
- (18) piperidine , PIP
- (19) 1,2-Ethanedithiol , EDT
- (20) Triisopropylsilane , TIS
- (21) 2-Chlorotrityl resin
- (22) Silica gel, 100-200 mesh
- (23) Nitrogen gas
- (24) Hexane
- (25) Ethyl acetate , EtOAc
- (26) Methanol , MeOH
- (27) Dichloromethane , DCM
- (28) Diethyl ether
- (29) Tetrahydrofuran , THF
- (30) Dimethylformamide , DMF
- (31) Methanol HPLC grade
- (32) Acetonitrile HPLC grade
- (33) Distilled water
- (34) Deionized water
- (35) Chloroform-d

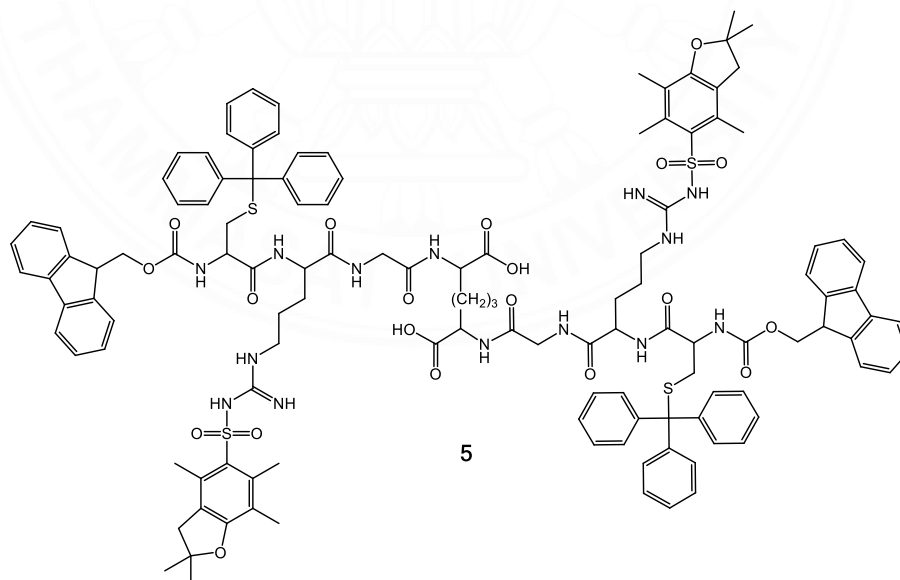
### 3.1.2 Instruments

- (1) Laboratory glassware
- (2) Glass chromatography columns
- (3) Silica thin layer chromatography plates
- (4) Glass filtration kit
- (5) Syringes and syringe filters
- (6) Cuvette
- (7) Heating mantle
- (8) Magnetic stirrer
- (9) Shaker, Stuart mini orbital shaker SSM1
- (10) Balance 4 digit, METTLER TOLEDO AB304-S
- (11) UV lamp
- (12) Rotary evaporator, BUCHI vacuum controller V-800 and rotavapor
- (13) Vacuum pump, EDWARDS model RV8
- (14) Freeze dryer, Flexi-Dry<sup>TM</sup> MP
- (15) Ultraviolet-visible spectrophotometer, SHIMADSU UV-1700 Pharma Spec
- (16) Analytical high-performance liquid chromatography, HPLC, SHIMADZU LC20AT and JASCO
- (17) Preparative high-performance liquid chromatography, Jasco preparative HPLC PU-4086-Binary pump, Jasco UV-4075 UV/Vis Detector
- (18) Nuclear magnetic resonance spectrometer (NMR), AVANCE NANOBAEY 400 MHz, Bruker
- (19) Electrospray ionization mass spectrometer, ESI-MS, Thermo Finnigan LCQ Advantage and Finnigan MAT LCQ
- (20) Matrix Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry (MALDITOF MS) (ID-Plus, Shimadzu and autoflexTOF/TOF, Bruker)
- (21) Microplate reader (Universal Microplate Analyzer, model AOPUS01 and AI53601, Packard BioScience, Connecticut)
- (22) Confocal laser scanning microscope, FV10i-DOC

### 3.2 Experiments

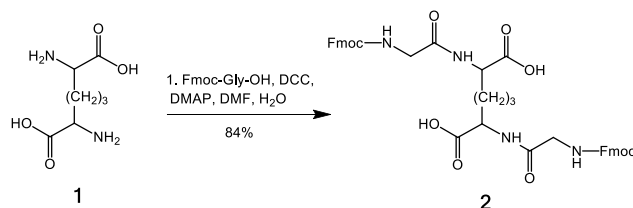
This research is focused on the syntheses of bicyclic SFTI-1. The microprotein was designed to be bicyclic peptide by using the SFTI-1 sequence, which showed the biological activity against human  $\beta$ -tryptase. The proposed structure is consisted of two inhibitory loops attached by 2,6-Diaminopimelic as a core. The syntheses can be classified into three main steps; (1) The synthesis of the three amino acids attached to 2,6-diaminopimelic acid by using solution phase peptide synthesis. (2) The amino acid assembly of linear peptide using solid phase peptide synthesis (SPPS). (3) The ligation and cyclization under an optimized high dilution condition. The synthesized bivalent inhibitor was assessed against cancer cell lines. Moreover, the bivalent SFTI-1 was labelled for the bioimaging application enabling to visualize the cellular macrocyclic peptide uptake.

#### 3.2.1 The synthesis of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycyl)pimelic acid *via* solution phase peptide synthesis



**Figure 3.1** The structure of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycyl)pimelic acid

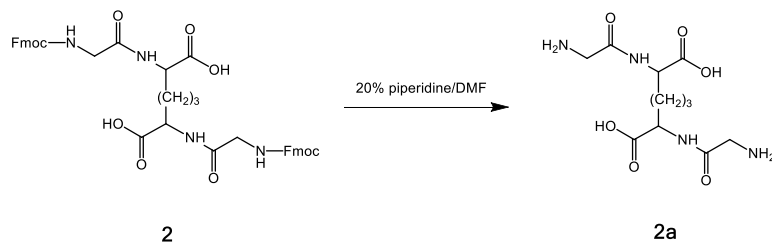
### 3.2.1.1 The synthesis of *N,N'*-diamino-bis(Fmoc-glyciny)l pimelic acid



**Scheme 3.1** The synthesis of *N,N'*-diamino-bis(Fmoc-glyciny)l pimelic acid

Fmoc-Gly-OH was firstly activated to an active form by reacting Fmoc-Gly-OH (2.5 eq. , 0.39 g , 1.32 mmol) with the mixture of DCC (2.5 eq. , 0.27 g , 1.32 mmol) and DMAP (2.5 eq. , 0.16 g , 1.32 mol). This reaction was dissolved by DMF (30 mL). The reaction was allowed at ambient temperature for 1 hour. To this reaction was slowly added by 2,6-diaminopimelic acid (1.0 eq. , 0.1 g , 0.53 mmol) dissolved in water (8 mL), The reaction was allowed to stir overnight and subsequently extracted by CH<sub>2</sub>Cl<sub>2</sub>. Then, the reaction was purified by using column chromatography (purified condition 30% EtOAc/Hexane) to afford *N,N'*-diamino-bis(Fmoc-glyciny)l pimelic acid in good yield (84%) as shown in **Scheme 3.1**. The structure was characterized by <sup>1</sup>H-NMR. <sup>1</sup>H-NMR (400 MHz), CDCl<sub>3</sub>) δ = 1.08-1.40 (m, 6H), 3.82-3.87 (m, 2H), 4.06-4.07 (d, *J* = 5.2, 4H), 4.19-4.23 (t, *J* = 7.2, 2H), 4.35-4.37 (d, *J* = 7.2, 4H), 7.27-7.31 (t, *J* = 7.4, 4H), 7.36-7.40 (t, *J* = 7.4, 4H), 7.57-7.59 (d, *J* = 7.4, 4H) and 7.73-7.75 (d, *J* = 7.4, 4H). Found 750.9 [M+H]<sup>+</sup> ; Calculated 748.2 [M]<sup>+</sup>.

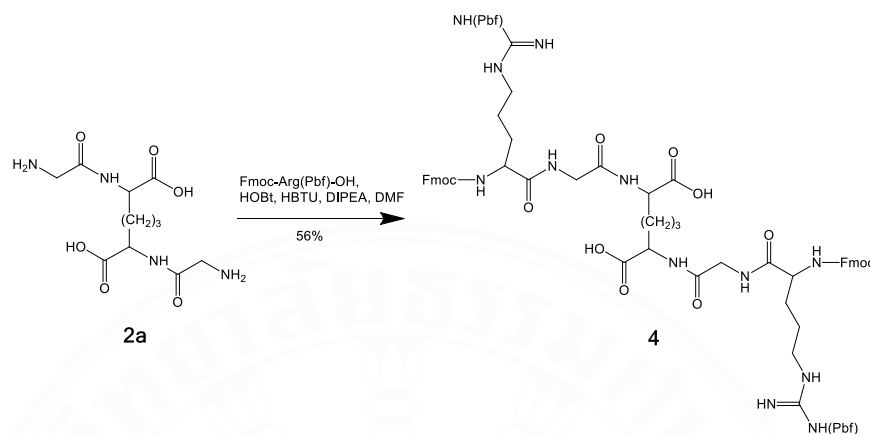
### 3.2.1.2 The synthesis of *N,N'*-diamino-bis(NH<sub>2</sub>-glycyl)pimelic acid



**Scheme 3.2** The synthesis of *N,N'*-diamino-bis(NH<sub>2</sub>-glycyl) pimelic acid

*N,N'*-diamino-bis(Fmoc-glycyl) pimelic acid ( 1 eq , 0.32 g, 0.42 mmol) was initially deprotected by using 20% piperidine in DMF for 30 min to yield *N,N'*-diamino-bis(NH<sub>2</sub>-glycyl) pimelic acid (**Scheme 3.2**). Then, the resulting product was purified by semi-preparative HPLC (gradient 10% MeOH/H<sub>2</sub>O to 100% MeOH in 30 minutes). The purity of resulting peptide was analysed by analytical HPLC (gradient 10% ACN/H<sub>2</sub>O to 100% ACN in 20 minutes), shows a peak at *t<sub>R</sub>* 12.673 minutes. The structure was characterized by <sup>1</sup>H-NMR. <sup>1</sup>H-NMR (500MHz, CDCl<sub>3</sub>) δ = 1.21-1.31 (m, 2H), 1.62-1.67 (m, 4H), 2.11-2.13 (q, J =12, 2H), 3.87 (s, 4H), 5.44 (S, 2H). Found 304.1956 [M+H]<sup>+</sup>, 339.1767 [M+Cl]<sup>-</sup> ; Calculated 304.14 [M]<sup>+</sup>.

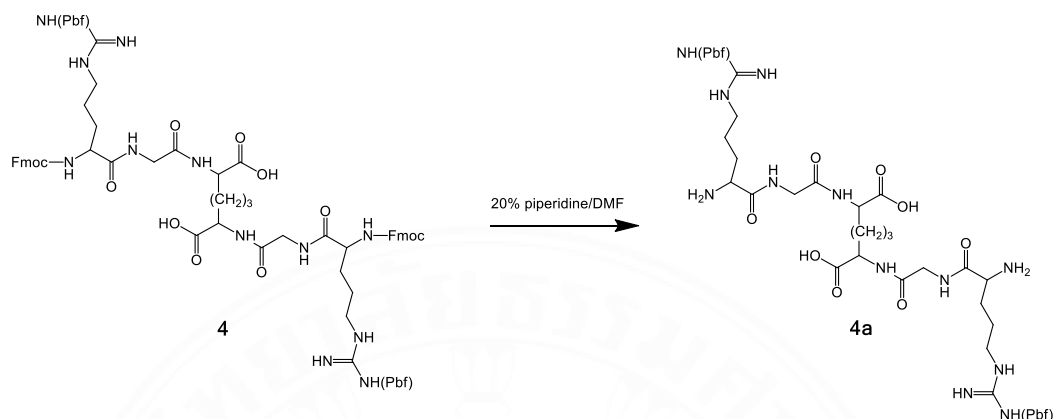
### 3.2.1.3 The synthesis of *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l pimelic acid



**Scheme 3.3** The synthesis of *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l pimelic acid

Next, Fmoc-Arg(Pbf)-OH was firstly activated to the active form of HOBt ester by the addition of solution Fmoc-Arg(Pbf)-OH (2.5 eq, , 0.69 g, 1.06 mmol) in DMF to the solution mixture of HOBt (2.5 eq, 0.18 g, 1.06 mmol), HBTU (2.5 eq, 0.40 g, 1.06 mmol) and DIPEA (2.5 eq, 180  $\mu$ L) in DMF. The reaction was stirred at ambient temperature for 1 h, and subsequently added by *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny)l pimelic acid dissolved in DMF. The reaction was allowed to stir at the ambient temperature for overnight and then purified by column chromatography (purified condition 50% EtOAc/Hexane) to yield *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l pimelic acid in a moderate yield (56%) as shown in **Scheme 3.3**. The structure was characterized by <sup>1</sup>H-NMR. <sup>1</sup>H-NMR(400 MHz), CDCl<sub>3</sub>)  $\delta$  = 1.05-1.37 (m, 6H), 1.42-1.47 (s, 12H), 1.55-1.60 (m, 4H), 1.89-1.92 (m, 4H), 1.98-2.02 (m, 2H), 2.08 (s, 6H), 2.50 (s, 6H), 2.55 (s, 6H), 2.95 (s, 4H), 3.36-3.46 (m, 4H), 4.19-4.22 (t,  $J$  = 6.6, 2H), 4.38-4.41 (m, 4H), 4.55-4.59 (m, 2H), 7.29-7.31 (t,  $J$  = 7.4, 4H), 7.36-7.40 (t,  $J$  = 7.4, 4H), 7.57-7.59 (m, 4H) and 7.73-7.75 (d,  $J$  = 7.4, 4H). Found 670.30 [M+CH<sub>3</sub>OH+Na]<sup>+</sup> and 1283.2 [2M+CH<sub>3</sub>OH+Na]<sup>+</sup>; Calculated fragment 616.32 (*N,N'*-diamino-bis(NH<sub>2</sub>-arginyl-glyciny)l pimelic acid).

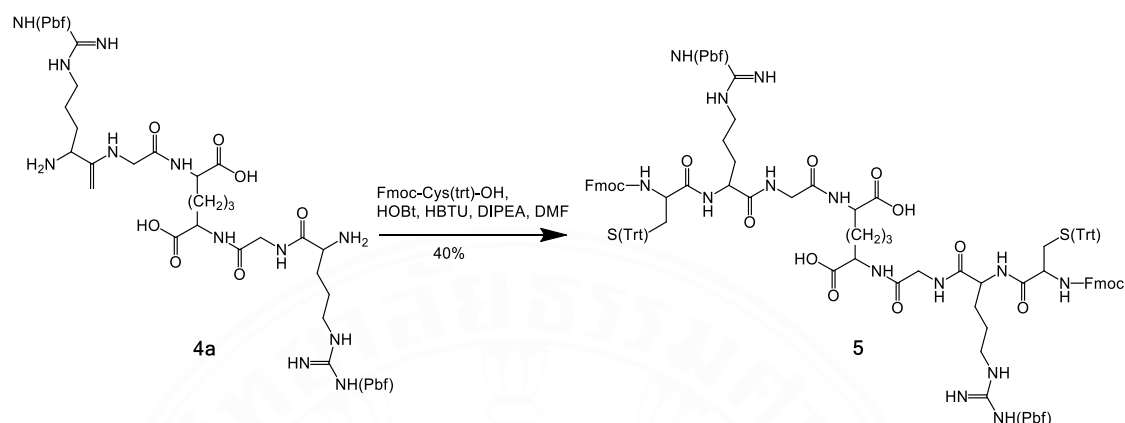
### 3.2.1.4 The synthesis of *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glyciny) pimelic acid



**Scheme 3.4** The synthesis of *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glyciny) pimelic acid

*N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny) pimelic acid (1.0 eq. , 0.37 g , 0.24 mmol) was firstly deprotected by using 20% piperidine in DMF for 30 minutes. Followed by purified by semi-preparative HPLC (gradient 10% MeOH/H<sub>2</sub>O to 100% MeOH in 30 minutes) to yield *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glyciny) pimelic acid (**Scheme 3.4**). The peptide purity was confirmed by analytical HPLC (gradient 10% ACN/H<sub>2</sub>O to 100% ACN in 20 minutes), shows a peak at *t<sub>R</sub>* 12.173 minutes. The structure was characterized by HR-ESI-MS. Found 1121.4576 [M+H]<sup>+</sup>, 1144.4171 [M+Na]<sup>+</sup> and 1157.5385 [M+K-2H]<sup>-</sup> ; Calculated 1120.50 [M]<sup>+</sup>.

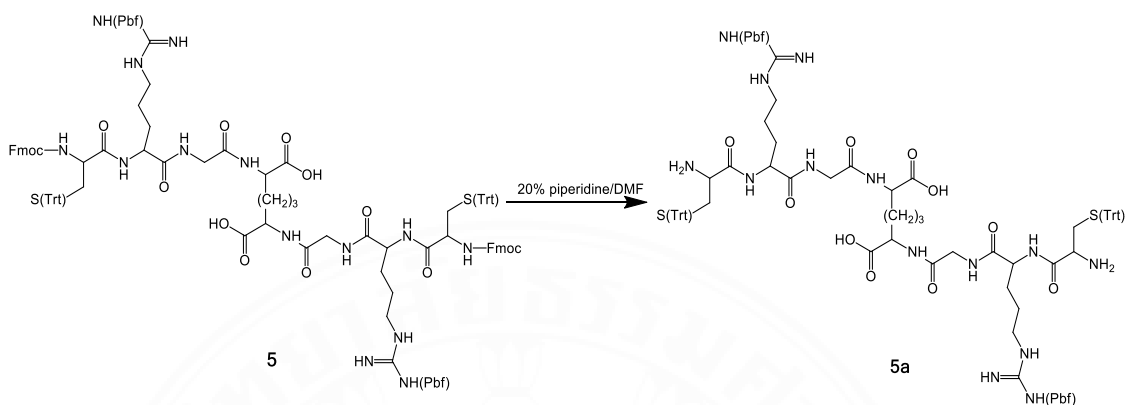
### 3.2.1.5 The synthesis of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)pimelic acid



**Scheme 3.5** The synthesis of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid

Fmoc-Cys(Trt)-OH was initially activated to the active form of HOBt ester by the addition of Fmoc-Cys(Trt)-OH (2.5 eq. , 0.35 g , 0.6 mmol) to the mixture of HOBt (2.5 eq. , 0.10 g , 0.6 mmol), HBTU (2.5 eq. 0.23 g , 0.6 mmol) and DIPEA (2.5 eq. 0.6 mmol , 0.10 mL). The reaction was stirred at the ambient temperature for 1 hour and then, was subsequently added by *N,N'*-diamino-bis(NH<sub>2</sub>-argenyl(Pbf)-glyciny) pimelic acid dissolved in DMF . The reaction was allowed to stir at ambient temperature for overnight, and then separated by column chromatography (purified condition 60% EtOAc/Hexane) to yield *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid in a modest yield (40%) as shown in **Scheme 3.5**. The structure was characterized by <sup>1</sup>H-NMR. <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  = 1.23-1.31 (m, 6H), 1.39-1.44 (s, 12H), 1.57-1.70 (m, 8H), 2.01-2.17 (m, 6H), 2.48-2.49 (d, *J* = 3.2, 6H), 2.53-2.54 (d, *J* = 3.2, 6H), 2.94 (s, 4H) , 3.26-3.33 (m, 4H), 3.71-3.93 (m, 4H), 4.17-4.19 (t, *J* = 6.4, 2H), 4.35-4.37 (m, 4H), 4.52-4.56 (m, 2H), 7.17-7.27 (m, 28H), 7.34-7.40 (m, 12H), 7.53-7.55 (m, 2H) and 7.70-7.74 (t, *J* = 7.3, 4H). Found 1327.8 [M+H]<sup>+</sup> ; Calculated fragment 1326.5 (*N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl-arginyl(Pbf)-glyciny) pimelic acid.

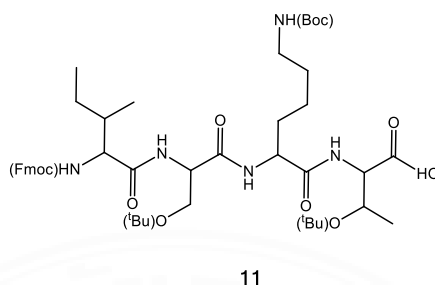
### 3.2.1.6 The synthesis of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid



**Scheme 3.6** The synthesis of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid

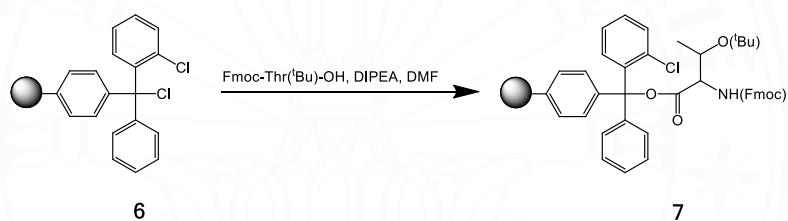
*N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid (1.0 eq. , 20 mg , 9  $\mu$ mol) was deprotected by 20% piperidine in DMF for 30 minutes and dried *in vacuo*, subsequently purified by semi-preparative HPLC (gradient 10% MeOH/H<sub>2</sub>O to 100% MeOH in 30 minutes) to yield *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid (**Scheme 3.6**). The purity of resulting peptide was analyzed by analytical HPLC (gradient 10% ACN/H<sub>2</sub>O to 100% ACN in 20 minutes), shows a peak at  $t_R$  22.576 minutes. The dimerization of Fmoc-Cys(Trt)-OH was confirmed by HR-ESI-MS. Found 927.6529 [M+2Na]<sup>2+</sup>, 945.4153 [M+2ACN-2H]<sup>2-</sup> ; Calculated 1810.74 [M]<sup>+</sup>.

### 3.2.2 The synthesis of linear peptide chain *via* solid phase peptide synthesis (SPPS)



**Figure 3.2.** The structure of peptide chain Fmoc-Ile-Ser(<sup>t</sup>Bu)-Lys(Boc)-Thr(<sup>t</sup>Bu)-OH

#### 3.2.2.1 The first amino acid loading on 2-chlorotrityl chloride resin



**Scheme 3.7** The first amino loading of Fmoc-Thr(<sup>t</sup>Bu)-OH on 2-chlorotrityl chloride resin

Fmoc-Thr(<sup>t</sup>Bu)-OH (1.5 eq. , 0.75 g , 1.89 mmol) was dissolved in DMF, following by the addition of DIPEA (1.5 eq., 1.89 mmol , 0.32 mL). To this solution was added by 2-chlorotrityl chloride resin (1.0 eq., 0.84 mmol/g, 1.5 g) under nitrogen atmosphere (**Scheme 3.7**). The reaction mixture was allowed to stir at ambient temperature for overnight. A resin was filtered, and subsequently washed by DMF (3×2 mL), DCM (3×2 mL) and diethyl ether (3×2 mL), respectively. The resin was dried in *vacuo*. The first amino acid loading was calculated by Fmoc test and found to be approximate 60%.

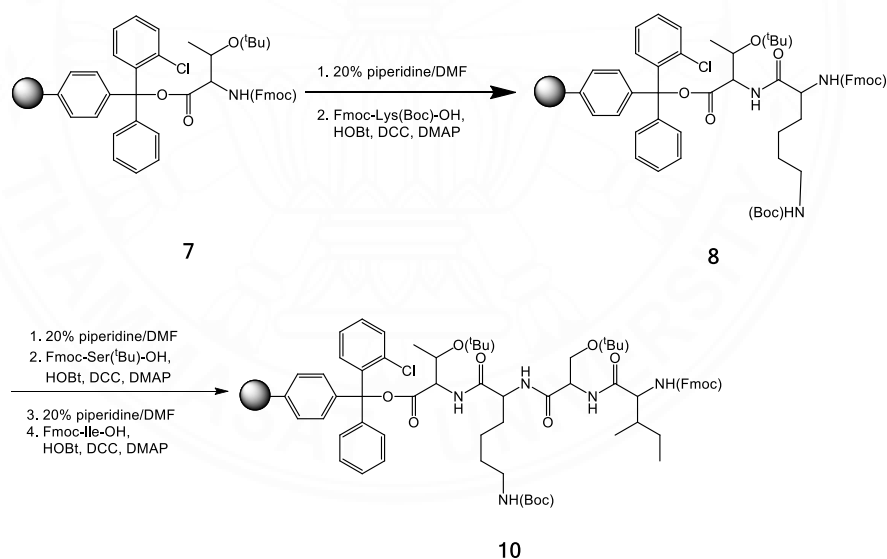
### 3.2.2.2 Calculate amino acid loading by using Fmoc test

Fmoc-Thr(<sup>t</sup>Bu)-O-resin (2 mg) was added into a small vial, then added 20% piperidine/DMF (3 mL) and allowed to shake for 45 minutes. Then, the reaction mixture was filtered and the solution was kept for UV measurement. The solution was measured the absorption intensity at 301 nm by using UV-Vis spectroscopy. The % loading was calculated by using the following equation,

$$\text{mmol amino acid/g of resin} = (\text{Abs.})/(1.65 \times \text{mg})$$

$$\% \text{ amino loading} = [(\text{mmol of amino acid/g of resin})/(\text{mmol/g of loading resin})] \times 100$$

### 3.2.2.3 Amino acid assembly *via* solid phase peptide synthesis (SPPS)



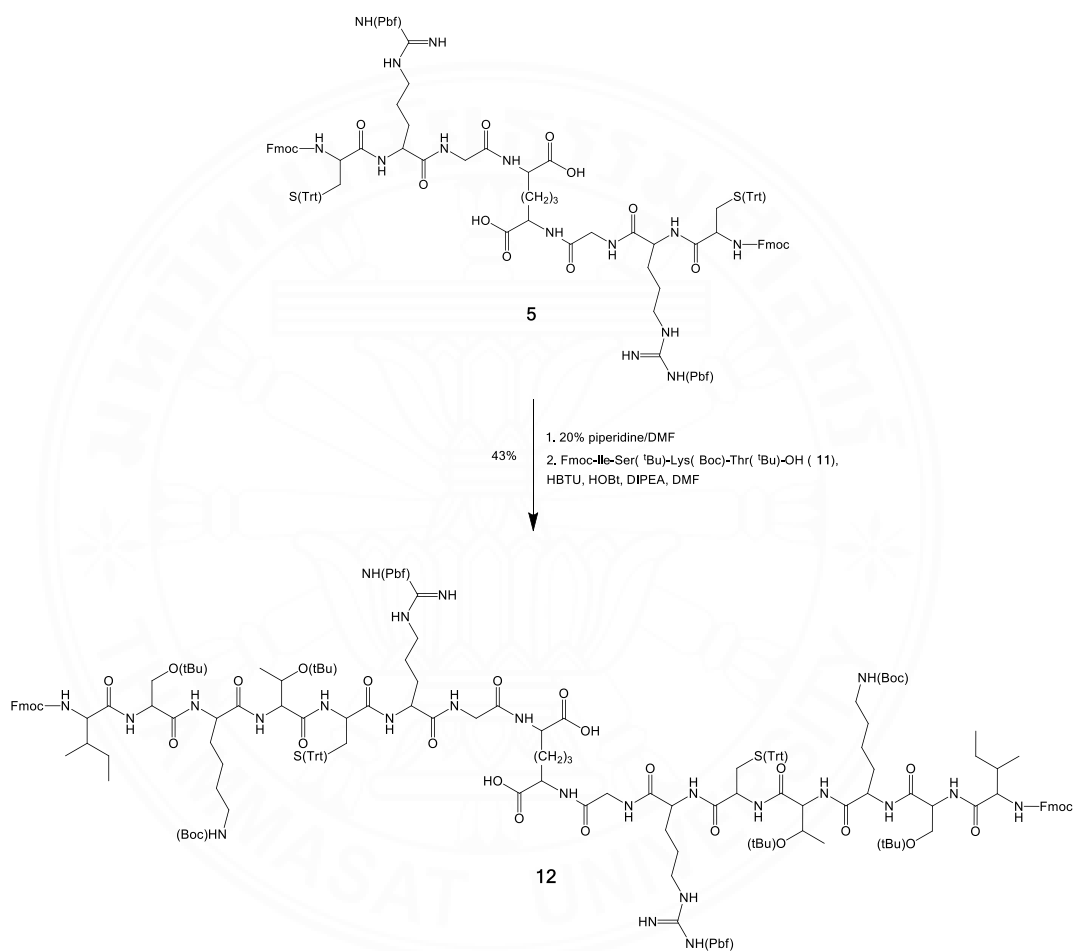
**Scheme 3.8** Synthesis of Fmoc-Ile-Ser(<sup>t</sup>Bu)-Lys(Boc)-Thr(<sup>t</sup>Bu)-OH

The preloaded Fmoc-Thr(<sup>t</sup>Bu)-O-resin (1 eq. , 0.2 g, 0.504 mmol amino/g) was swelled by DMF for 30 min, and then treated by 20% piperidine in DMF for 30 min to remove Fmoc protecting groups. Fmoc-Lys(Boc)-OH (1.5 eq. , 70.8 mg, 0.16 mmol) was coupled to a preloaded NH<sub>2</sub>-Thr(<sup>t</sup>Bu)-O-resin by using DCC (1.5 eq. , 31.2 mg, 0.16 mmol), DMAP (1.5 eq. , 18.5 mg, 0.16 mmol) and HOBT (1.5 eq. , 25.7 mg, 0.16 mmol) as coupling reagents. The reaction mixture was shaken for 60 min, filtered off, and washed by DMF (3×3 mL). The resin was treated by 20% piperidine in DMF for 30 min to remove all Fmoc protecting groups. The resin was drained off and washed with DMF (3×3 mL). The synthesis was repeated by using other amino acids; Fmoc-Ser(<sup>t</sup>Bu)-OH (1.5 eq. , 58 mg, 0.16 mmol) and Fmoc-Ile-OH (1.5 eq. , 53.5 mg, 0.16 mmol), respectively as shown in **Scheme 3.8**. After the synthesis was completed, the resin was drained off and washed with DMF (3×2 mL), DCM (3×2 mL) and diethyl ether (3×2 mL). A fully protected peptide was treated by 1% TFA in DCM to cleave a linear peptide chain from resin, following by draining off. The filtrate was evaporated to dryness under the reduced pressure and dried *in vacuo*. Found 904.9[M+Na]<sup>+</sup> ; Calculated 881.5[M]<sup>+</sup>.



### 3.2.3 Ligation and Cyclization

#### 3.2.3.1 The synthesis of *N,N'*-diamino-bis(Fmoc-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid



**Scheme 3.9** Ligation reaction of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid and Fmoc-IleSer(<sup>t</sup>Bu)Lys(Boc)Thr(<sup>t</sup>Bu)-OH

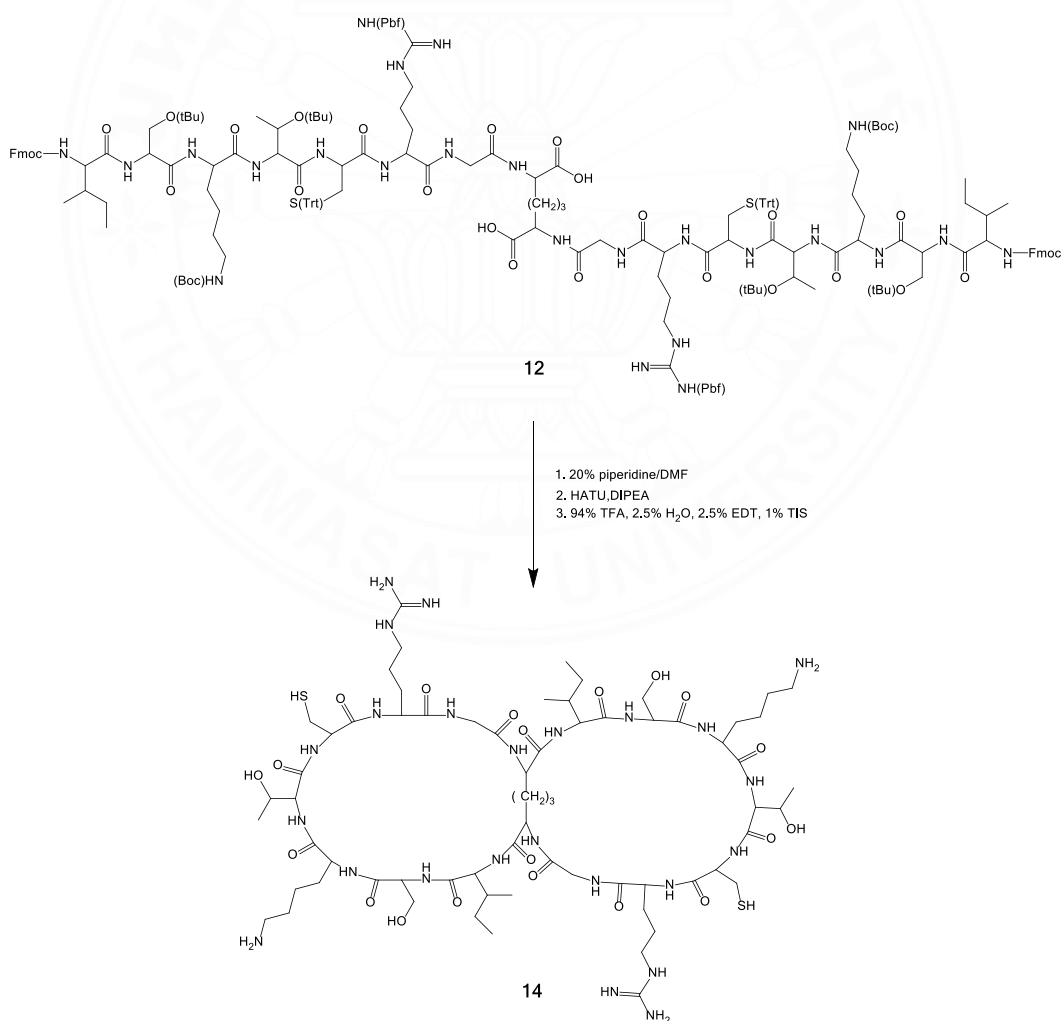
*N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid (1.0 eq. , 20 mg , 9  $\mu\text{mol}$ ) was deprotected by 20% piperidine in DMF for 30 minutes and dried *in vacuo*. Fmoc-IleSer(<sup>t</sup>Bu)Lys(Boc)Thr(<sup>t</sup>Bu)-OH (2.5 eq., 19.5 mg, 22.2  $\mu\text{mol}$ ) was activated by adding to the mixture of HBTU (2.5 eq. , 3.8 mg, 22.2  $\mu\text{mol}$ ) HOBt (2.5 eq. , 8.5 mg, 22.2  $\mu\text{mol}$ ) and DIPEA (2.5 eq. , 4  $\mu\text{L}$ , 22.2  $\mu\text{mol}$ ) in DMF, and allowed to stir for 1 hour. To this solution was slowly added (2 minutes : 1 drop) by *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid dissolved in DMF, and allowed to stir at ambient temperature overnight. The resulting product was separated by using column chromatography (gradient 10% EtOAc/Hexane to 50% MeOH/EtOAc) to afford *N,N'*-diamino-bis(Fmoc-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny)l(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid in modest yield (43%) as shown in **Scheme 3.9**. The ligation product was characterized by ESI-MS, showing the molecular mass of 1769.2  $[\text{M}+2\text{H}]^{2+}$  and 897.4  $[\text{M}+\text{ACN}+4\text{H}]^{4+}$  ; Calculated 3537.75  $[\text{M}]^{+}$ .

### 3.2.3.2 The synthesis of *N,N'*-diamino-bis(cyclic-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny)l(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid

*N,N'*-diamino-bis(Fmoc-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny)l(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid (1.0 eq. , 43.5 mg, 12  $\mu\text{mol}$ ) was treated by 20% piperidine in DMF for 30 minutes, which was further evaporated to dryness under reduced pressure, and subsequently dissolved in DMF 50 mL. To this solution DIPEA was added (5 eq., 10.3  $\mu\text{L}$ , 60.7  $\mu\text{mol}$ ), followed by slowly dropped HATU (1 drop per hour) dissolved in DMF (5 eq., 23 mg , 60.7  $\mu\text{mol}$ ) as shown in **Scheme 3.10**. The reaction mixture was allowed to stir at ambient temperature for overnight, followed by mixing with water 100 mL and freezedry to dryness.

### 3.2.3.3 The synthesis of *N,N'*-diamino-bis(cyclic-isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid

*N,N'*-diamino-bis(cyclic-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysinyl(Boc)-threnyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid was treated by reagent K (TFA 94% : EDT 2.5% : H<sub>2</sub>O 2.5% : TIS 1%) for 2 h. The reaction mixture was evaporated by using N<sub>2</sub> gas, and then purified by semi-preparative HPLC (gradient 10% MeOH/H<sub>2</sub>O to 100% MeOH in 30 minutes) to afford bivalent SFTI-1 as shown in **Scheme 3.10**, in which the molecular weight was verified by using MALDI-TOF spectrometry, showing the molecular mass of 1833.4335 [M+CHCA+H]<sup>+</sup> ; Calculated 1644.86 [M]<sup>+</sup>.

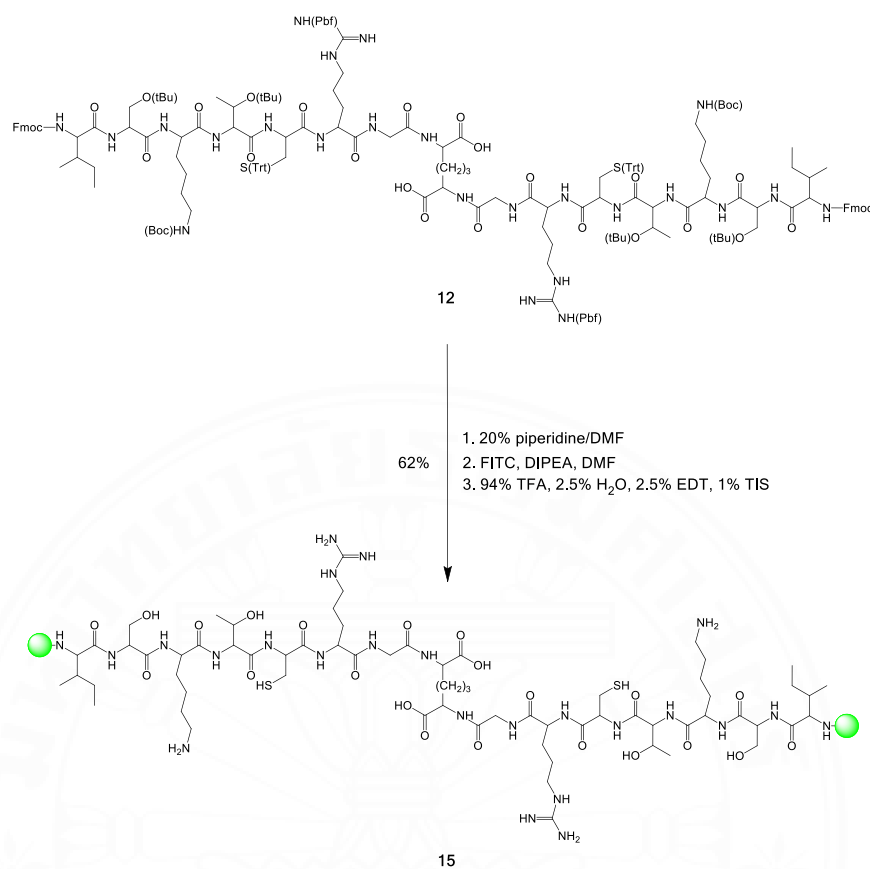


**Scheme 3.10** Cyclization and global deprotection reaction

### 3.2.4 The synthesis of fluorescently labelled *N,N'*-diamino-bis(isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glycinyl) pimelic acid

*N,N'*-diamino-bis(Fmoc-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glycinyl) pimelic acid (1.0 eq. , 5 mg , 1.6 μmol) was deprotected by 20% piperidine in DMF for 30 minutes and dried *in vacuo*. The reaction mixture of FITC (2.0 eq. , 1.8 mg , 3.2 μmol) and DIPEA (2.5 eq. , 1 μL , 4.1 μmol) were dissolved in DMF 2 mL. Next, the Fmoc deprotected peptide *N,N'*-diamino-bis(NH<sub>2</sub>-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glycinyl) pimelic acid, dissolved in DMF, was dropped into the reaction mixture in modest rate, and allowed to stir at ambient temperature overnight. Thereafter, the fluorescently labelled peptide was deprotected all protecting groups along the side chain by treating with reagent K (TFA 94% : EDT 2.5% : H<sub>2</sub>O 2.5% : TIS 1%) to yield fluorescently labelled *N,N'*-diamino-bis(isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glycinyl) pimelic acid (**Scheme 3.11**).

Then, the resulting labelled peptide was separated by semi-preparative HPLC (gradient 10% MeOH/H<sub>2</sub>O to 100% MeOH in 30 minutes) to yield fluorescently labelled *N,N'*-diamino-bis(isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glycinyl) pimelic acid in moderate yield (62%, 2.5 mg). The purity of resulting product was verified by analytical HPLC (gradient 10% ACN/H<sub>2</sub>O to 100% ACN in 20 minutes) and shows a peak at *t<sub>R</sub>* 22.387 minutes, and the molecular weight was characterized by MALDI-TOF. Found 616.002 [M+4H]<sup>4+</sup>, 845.114 [M+3Na]<sup>3+</sup> ; Calculated 2459.95 [M]<sup>+</sup>



**Scheme 3.11** The synthesis of fluorescently labelled *N,N'*-diamino-bis(isoleucinyl-serinyl-lysanyl-threonyl-cysteinyl-arginyl-glyciny) pimelic acid

### 3.2.5 Evaluation of cell viability

Cell viability was examined using an MTT assay. Human embryonic kidney (HEK293), human colorectal adenocarcinoma (HT29) and human cervical epithelial adenocarcinoma (HeLa) cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS), penicillin (100 units/mL), streptomycin (100 mg/mL) and *L*-glutamine (4 mM) at 37 °C, 5% CO<sub>2</sub>. Human breast adenocarcinoma (MCF7) was also grown as mentioned above except the medium containing 1% of insulin. For transfection, the cells were seeded up to  $1 \times 10^4$  cells/well in a 96-well plate to give 50-70% confluence to be used on the next day.

Before the cytotoxicity study, the growth medium was removed, and the cells were washed with PBS and replaced with 100  $\mu$ L of various concentrations of samples. The 96-well plate was incubated at 37 °C, 5% CO<sub>2</sub> for 48 h.

In the MTT experiment, the solution of the samples was removed and the cells were rinsed with PBS followed by incubation with 100  $\mu$ L of MTT containing medium (1 mg/mL) for 4 h. The formazan crystals formed in the living cells were dissolved using 100  $\mu$ L of dimethyl sulfoxide per well. The relative cell viability (in percent) was calculated based on the absorbance observed at 550 nm using a microplate reader (Universal Microplate Analyzer, model AOPUS01 and AI53601, Packard BioScience, Connecticut, USA) and compared to the non-treated cells. The viability of the non-treated cells was defined as 100%.

### 3.2.6 Fluorescence Imaging Protocol

HeLa cells (the human cervical cancer cell lines) were plated in 24-well plate at a density of  $2 \times 10^4$  mL<sup>-1</sup> for 24 hrs. The fluorescent probe was dissolved in DMSO and diluted to the indicated concentration (125  $\mu$ M). HeLa cells were incubated with the fluorescent probe (125  $\mu$ M) for 30 min. Thereafter, HeLa cells were washed 3 times by PBS buffer (pH 7.2). Then, HeLa cell imaging was performed under a confocal microscope.

## CHAPTER 4

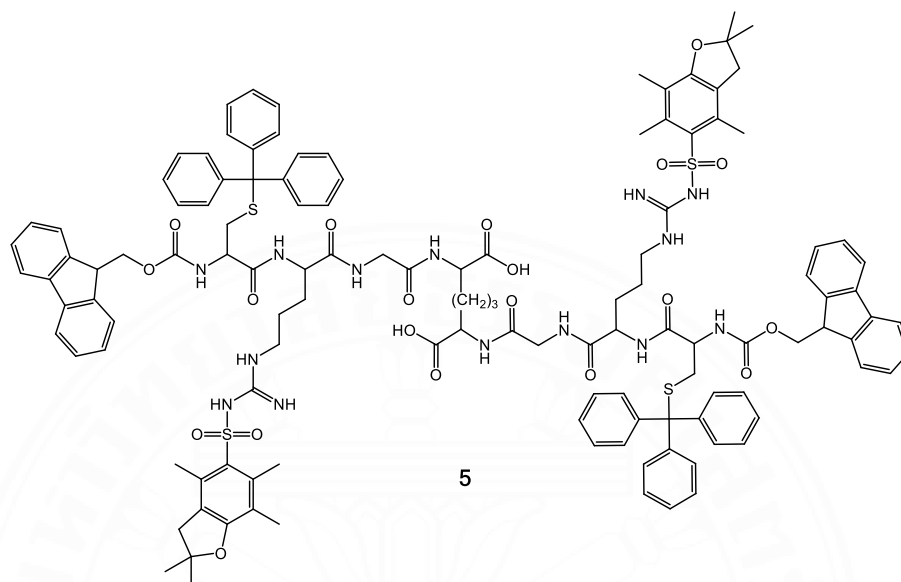
### RESULTS AND DISCUSSION

Bivalent SFTI-1 inhibitor was designed and subsequently synthesized to have the bicyclic scaffold by using 2,6-diaminopimelic acid as the main core. Each cyclic ring consisted of seven amino acids mimicked primarily according to the inhibitory loops of the SFTI-1 sequence includes glycine (Gly), arginine (Arg), cysteine (Cys), threonine (Thr), lysine (Lys), serine (Ser) and isoleucine (Ile), respectively.

Typically, 2,6-diaminopimelic acid is one of metabolites of *E. Coli*, which is involved in the production of the essential enzymes for living organisms. Its structure is consisted of seven carbons with two carboxyl groups and two amino groups. Interestingly, the three carbon in the middle mimicked the presence of a disulfide bridge of natural SFTI-1 and it is more metabolically stable. Therefore, 2,6-diaminopimelic acid is appropriate bivalent generator for the synthesis of bivalent SFTI-I scaffold.

This bivalent SFTI-1 inhibitor was synthesized by the combination of two peptide synthesis approaches. Firstly, the synthesis of the three amino acids attached to 2,6-diaminopimelic acid *via* solution phase peptide synthesis. Secondly, the amino acid assembly of the linear peptide using Fmoc solid phase peptide synthesis (SPPS). Finally, the resulting products from both solid and solution phase approaches were ligated, and followed by the double cyclization.

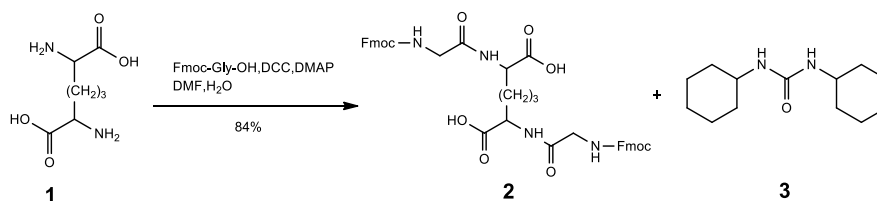
#### 4.1 The synthesis of tripeptides attached on scaffold *via* solution phase peptide synthesis (Fmoc-cysteinyl(Trt)-arginy(Pbf)-glyciny-pimelic acid)



**Figure 4.1** The structure of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginy(Pbf)-glyciny) pimelic acid

##### 4.1.1 The synthesis of *N,N'*-diamino-bis(Fmoc-glyciny)pimelic acid

To this work, 2,6-diaminopimelic acid was utilized as an important bivalent core, which was consecutively coupled with amino acids. Firstly, 2,6-diaminopimelic acid was coupled by Fmoc-Gly-OH. Fmoc-Gly-OH was initially activated to an active form and subsequently coupled with 2,6-diaminopimelic acid using standard coupling reagent DCC and DMAP (**Scheme 4.1**). Under this condition, the dimerization between 2,6-diaminopimelic acid was minimized. To minimize the solubility problem of 2,6-diaminopimelic acid, a small amount of water was utilized as a co-solvent with DMF (ratio 1:3) in order to increase the solubility and the coupling efficacy.

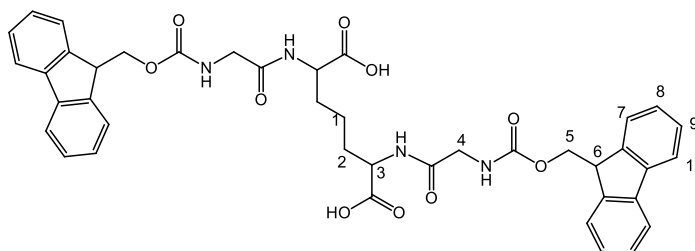


**Scheme 4.1.** The synthesis of *N,N'*-diamino-bis(Fmoc-glycyl)pimelic acid

Typically, DCC is the most frequently used coupling agent for peptide synthesis. However, the main disadvantage of using DCC is the formation of 1,3-dicyclohexylurea (DHU) which was inseparable by normal column chromatography. To resolve this problem, this coupling reagent (DCC) was replaced by DIC which is also classified as a carbodiimide type similar to DCC. Unfortunately, there is no desired product formed, and it is mainly due to the side reaction between reactive DIC and water which is used to increase the solubility during the activation. Accordingly, DCC and DMAP is applicable as coupling reagents for the first amino acid attachment to afford *N,N'*-diamino-bis(Fmoc-glycyl) pimelic acid in a good yield (84%). The used condition of *N,N'*-diamino-bis(glycyl)pimelic acid syntheses are summarized as shown in **Table 4.1**.

**Table 4.1** The used condition of *N,N'*-diamino-bis(glycyl)pimelic acid syntheses

Product	Coupling reagents	solvent	Yield (%)
Fmoc-Gly-pimelic acid (2)	DCC/DMAP	DMF /H <sub>2</sub> O (3:1)	84
	DIC	DMF /H <sub>2</sub> O (3:1)	-



**Figure 4.2** The structure of *N,N'*-diamino-bis(Fmoc-glycyl)pimelic acid

Name: *N,N'*-diamino-bis(Fmoc-glycyl)pimelic acid (**2**)

Chemical formula:  $C_{41}H_{40}N_4O_{10}$

ESI-MS: Found 750.9  $[M+H]^+$  ; Calculated 748.2  $[M]^+$ .

$^1H$ -NMR (400MHz,  $CDCl_3$ )  $\delta$  = 1.08-1.40 (m, 6H), 3.82-3.87 (m, 2H), 4.06-4.07 (d,  $J$  = 5.2, 4H), 4.19-4.23 (t,  $J$  = 7.2, 2H), 4.35-4.37 (d,  $J$  = 7.2, 4H), 7.27-7.31 (t,  $J$  = 7.4, 4H), 7.36-7.40 (t,  $J$  = 7.4, 4H), 7.57-7.59 (d,  $J$  = 7.4, 4H) and 7.73-7.75 (d,  $J$  = 7.4, 4H).

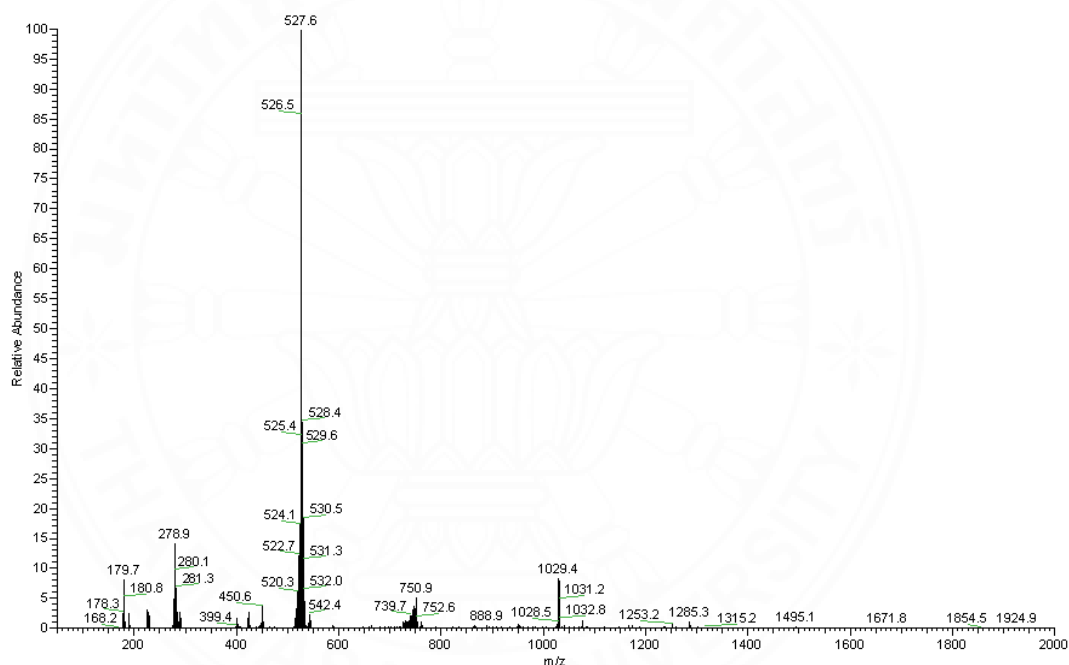
**Table 4.2**  $^1H$ -NMR of *N,N'*-diamino-bis(Fmoc-glycyl) pimelic acid (**2**)

Position number	Chemical shift (ppm)
1,2	1.08-1.40 (m, 6H)
3	3.82-3.87 (m, 2H)
4	4.06-4.07 (d, $J$ = 5.2, 4H)
5	4.35-4.37 (d, $J$ = 7.2, 4H)
6	4.19-4.23 (t, $J$ = 7.2, 2H)
7	7.57-7.59 (d, $J$ = 7.4, 4H)
8	7.27-7.31 (t, $J$ = 7.4, 4H)
9	7.36-7.40 (t, $J$ = 7.4, 4H)
10	7.73-7.75 (d, $J$ = 7.4, 4H)

*N,N'*-diamino-bis(Fmoc-glyciny)l pimelic acid was characterized by  $^1\text{H}$  NMR. The molecular weight was verified by ESI-MS. The result is shown in **Table 4.2**. According to  $^1\text{H}$ -NMR spectrum, it showed proton the resonance signals ( $\delta$ ) at  $\delta = 1.08$ - $1.40$  ppm and  $\delta = 3.82$ - $3.87$  ppm corresponding to the methylene protons ( $-\text{CH}_2-$ ) and  $-\text{CH}-$  of 2,6-diaminopimelic acid core, respectively. The methylene protons of glycine were observed at  $\delta = 4.06$ - $4.07$  ppm. The resonance signals at  $\delta = 4.35$ - $4.37$  ppm and  $\delta = 4.19$ - $4.23$  ppm corresponded to  $-\text{CH}_2-$  and  $\text{CH}$  of Fmoc protecting moiety. The aromatic protons of Fmoc moiety were also found at  $\delta = 7.57$ - $7.59$ ,  $7.27$ - $7.31$ ,  $7.36$ - $7.40$  ppm and  $7.73$ - $7.75$  ppm, respectively. Thus, this was clearly confirmed that 2,6-diaminopimelic acid scaffold was successfully grafted by Fmoc-Gly-OH. Moreover, the proton signal of indivisible 1,3-dicyclohexylurea was also identified.



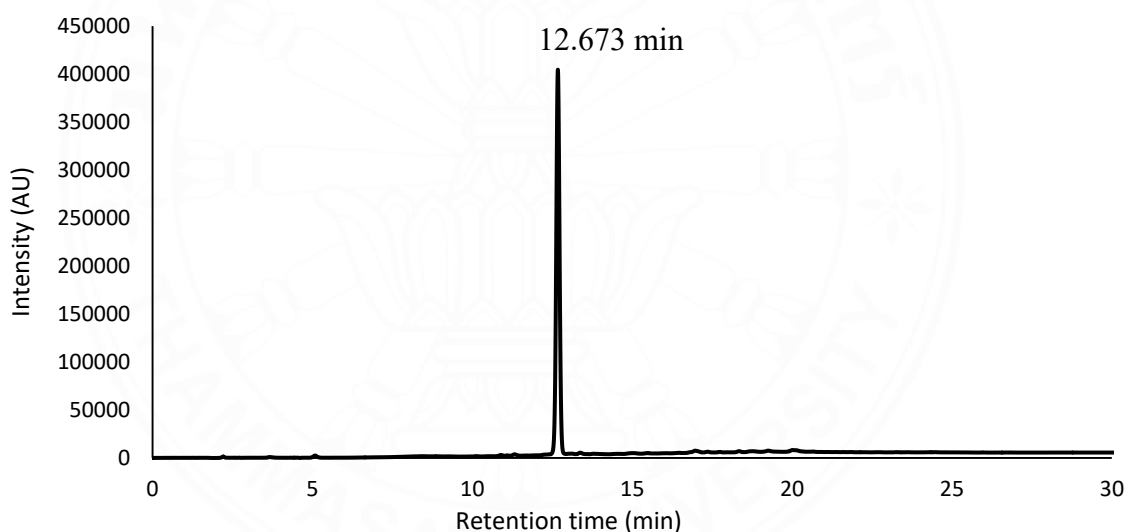
According to ESI-MS analyses (**Figure 4.3**), it showed molecular mass adduct ( $C_{41}H_{40}N_4O_{10}+H$ ) at 750.9  $m/z$  corresponding to  $[M+H]^+$ , while the molecular mass of *N,N'*-diamino-bis(Fmoc-glyciny)l pimelic acid was calculated to be 748.27  $m/z$  ( $C_{41}H_{40}N_4O_{10}$ ). Additionally, it showed  $[M+H]^+$  ion at  $m/z$  527.6 correspondings to the molecular formula  $C_{26}H_{30}N_4O_8$  (526.21 g/mol), indicating a fragment of *N,N'*-diamino-bis(Fmoc-glyciny)l pimelic acid with the loss of a Fmoc protecting group. The ESI-MS results clearly confirmed the double attachment of 2,6-diaminopimelic by Fmoc-Gly-OH.



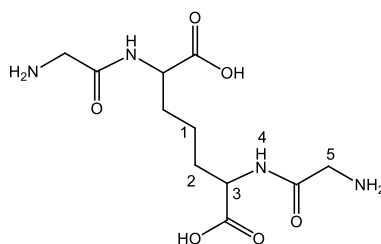
**Figure 4.3.** Positive mode mass spectrum of *N,N'*-diamino-bis(Fmoc-glyciny)l pimelic acid (**2**) Found 750.9 correspondings to molecular mass adduct  $[M+H]^+$ .

To eliminate 1,3-dicyclohexyl urea, semi-preparative HPLC technique has been applied for the isolation of *N,N'*-diamino-bis(Fmoc-glyciny)l)pimelic acid. By doing this, the deprotection of Fmoc moiety from *N,N'*-diamino-bis(Fmoc-glyciny)l)pimelic acid is initially required..

Firstly, *N,N'*-diamino-bis(Fmoc-glyciny)l)pimelic acid was treated with 20% piperidine in DMF to deprotect both Fmoc groups to yield *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny)l)pimelic acid. Then, the reaction mixture was subsequently purified by using semi-preparative HPLC. According to the HPLC results (**Figure 4.4**), the chromatogram was measured at 215 nm, which is the absorption wavelength for the peptide bond and showed a peak of *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny)l)pimelic acid at 12.673 min.



**Figure 4.4.** HPLC chromatogram of *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny)l)pimelic acid (**2**) measured at 215 nm.



**Figure 4.5** The structure of *N,N'*-diamino-bis(NH<sub>2</sub>-glycinyl) pimelic acid

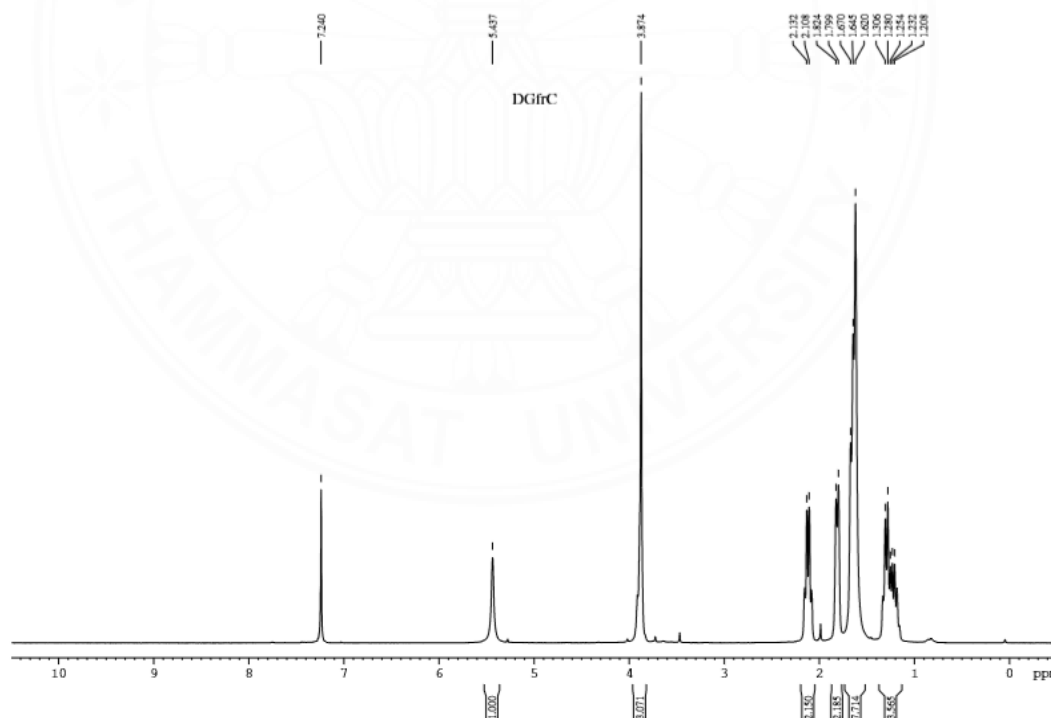
Name: *N,N'*-diamino-bis(NH<sub>2</sub>-glycinyl) pimelic acid (**2a**)

Chemical formula: C<sub>11</sub>H<sub>20</sub>N<sub>4</sub>O<sub>6</sub>

ESI-MS: Found 304.1956 [M+H]<sup>+</sup>, 339.1767 [M+Cl]<sup>-</sup>; Calculated 304.14 [M]<sup>+</sup>.

<sup>1</sup>H-NMR (500MHz, CDCl<sub>3</sub>) δ = 1.21-1.31 (m, 2H), 1.62-1.67 (m, 4H), 2.11-2.13 (m, J = 12, 2H), 3.87 (s, 4H), 5.44 (s, 2H)

HPLC (method 1): t<sub>R</sub> = 12.673 min



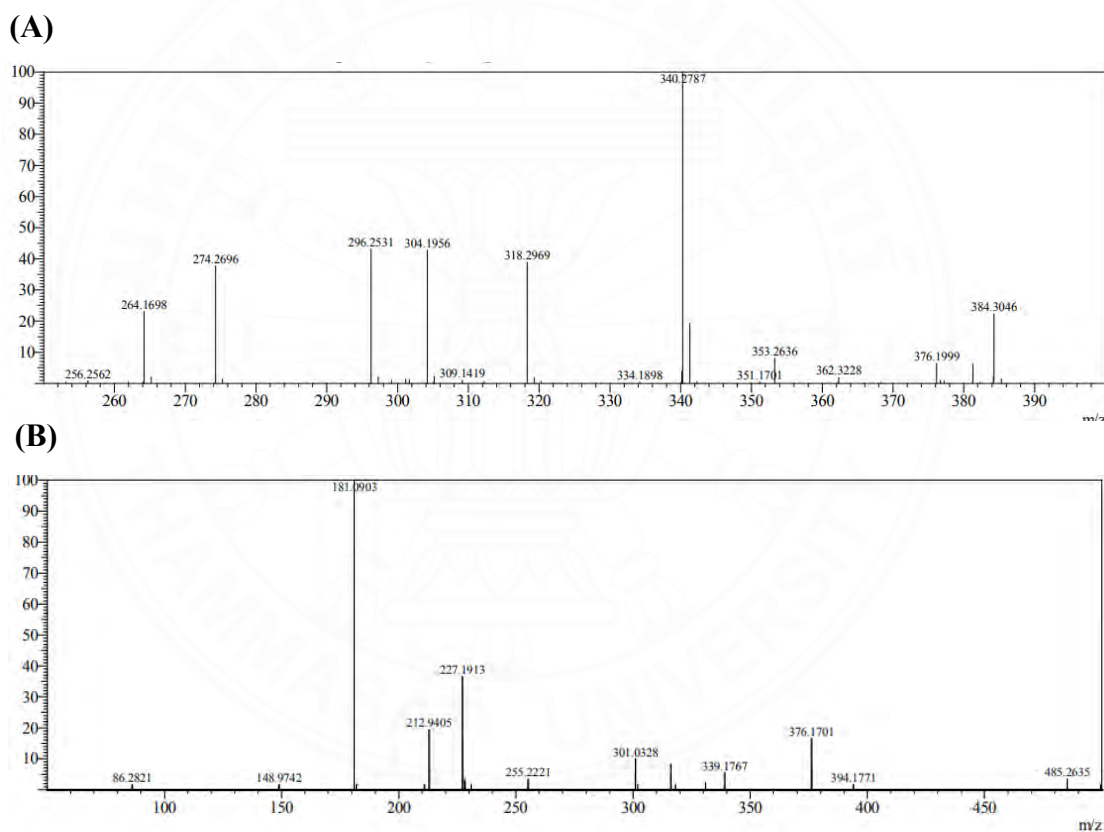
**Figure 4.6** <sup>1</sup>H NMR spectrum (500 MHz, CDCl<sub>3</sub>) of *N,N'*-diamino-bis(NH<sub>2</sub>-glycinyl) pimelic acid (**2a**)

**Table 4.3**  $^1\text{H}$ -NMR of *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny) pimelic acid (**2a**)

Position number	Chemical shift (ppm)
1	1.21-1.31 (m, 2H)
2	1.62-1.67 (m, 4H)
3	2.11-2.13 (q, J = 12, 2H)
4	5.44 (S, 2H)
5	3.97 (S, 4H)

The *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny) pimelic acid was characterized by  $^1\text{H}$ -NMR. The result is shown in **Table 4.3**. According to  $^1\text{H}$ -NMR spectrum (**Figure 4.6**), it showed the resonance signals ( $\delta$ ) at 1.21-1.31 ppm and 1.62-1.67 ppm corresponding to the methylene protons (-CH<sub>2</sub>-) of the 2,6-diaminopimelic acid core. Also, the resonance signal at  $\delta = 2.11$ -2.13 ppm, corresponding to -CH- of the 2,6-diaminopimelic acid were observed. The methylene protons of glycine were observed at  $\delta = 3.97$  ppm. In addition, the carboxylic acid protons were also identified at  $\delta = 5.44$  ppm.

The purity of *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny) pimelic acid was verified by using analytical HPLC along with HR-ESI-MS spectroscopy in both positive and negative ESI mode. According to HR-ESI-MS result (*via* positive mode as shown in **Figure 4.7A**), the molecular mass adduct was identified at  $m/z = 304.1956$  correspondings to  $[M+H]^+$ . Additionally, the molecular mass adduct at  $m/z = 339.1767$  corresponding to the  $[M+Cl]^-$  *via* negative mode (**Figure 4.7B**) was also observed. While the molecular mass of *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny) pimelic acid was calculated to be  $m/z = 304.14$  (C<sub>11</sub>H<sub>20</sub>N<sub>4</sub>O<sub>6</sub>).



**Figure 4.7** High resolution mass spectrum (HR-ESI-MS) positive and negative modes of *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny) pimelic acid (**2a**) ; Calculated 304.14  $m/z$

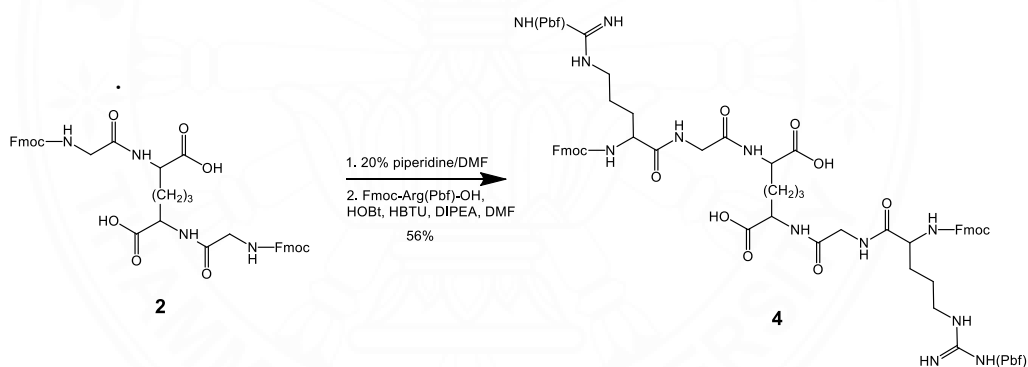
**(A)** Positive mode HR-ESI-MS showed  $m/z = 304.1956$  corresponding to molecular mass adduct  $[M+H]^+$ .

**(B)** Negative mode HR-ESI-MS showed  $m/z = 339.1767$  corresponding to molecular mass adduct  $[M+Cl]^-$ .

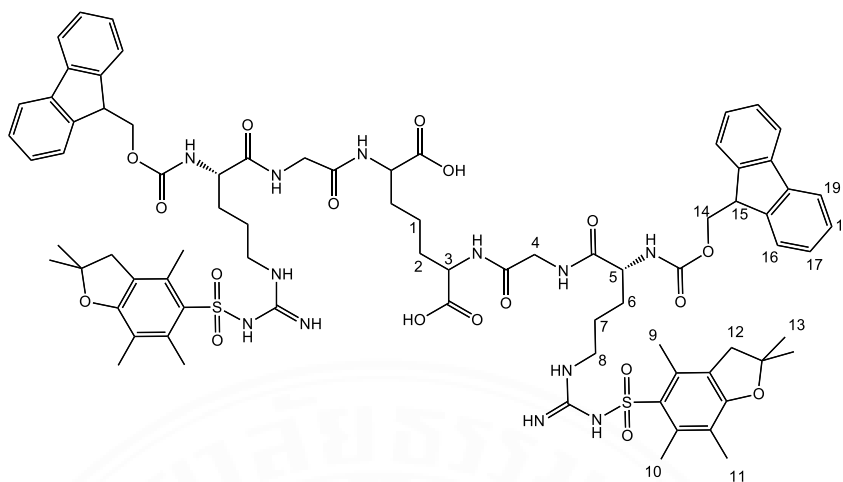
#### 4.1.2 The synthesis of *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glycyl) pimelic acid

As previously seen, the presence of 1,3 dicyclohexylurea in the reaction mixture as by product did not affect the formation of amide bond. With this notification, the purification after Fmoc deprotection step is not necessary.

Firstly, *N,N'*-diamino-bis(Fmoc-glycyl) pimelic acid was treated with 20% piperidine in DMF to give *N,N'*-diamino-bis(NH<sub>2</sub>-glycyl) pimelic acid which was subsequently coupled with Fmoc-Arg(Pbf)-OH using standard coupling reagents HBTU/HOBt as coupling reagents in DMF to afford *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glycyl) pimelic acid in a moderate yield (56%) after the purification by column chromatography (50% EtOAc/Hexane) as shown in **Scheme 4.2**. In this step, 1,3-dicyclohexylurea can be separated from the reaction mixture.



**Scheme 4.2.** The synthesis of *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glycyl) pimelic acid)



**Figure 4.8** The structure of *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny) pimelic acid

Name: *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny)pimelic acid) (**4**)

Chemical formula:  $C_{79}H_{96}N_{12}O_{18}S_2$

ESI-MS: Found 654.4  $[M+K]^+$ , 670.3  $[M+Na]^+$

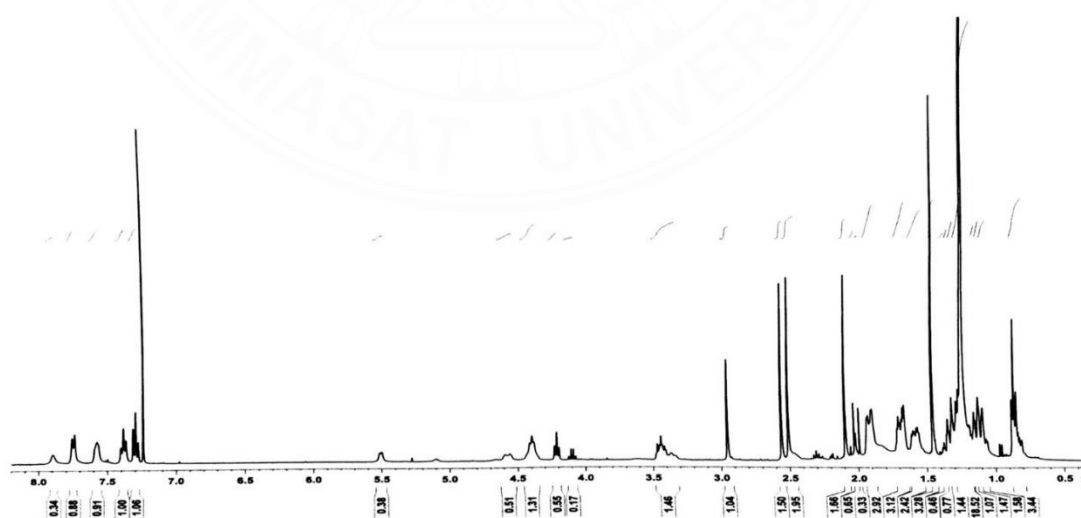
Calculated fragment:  $m/z = 616.32 [C_{23}H_{44}N_{12}O_8]^+$

$^1H$ -NMR (400MHz,  $CDCl_3$ )  $\delta$  = 1.05-1.37 (m, 6H), 1.42-1.47 (s, 12H), 1.55-1.60 (m, 4H), 1.89-1.92 (m, 4H), 1.98-2.02 (m, 2H), 2.08 (s, 6H), 2.50 (s, 6H), 2.55 (s, 6H), 2.95 (s, 4H), 3.36-3.46 (m, 4H), 4.19-4.22 (t,  $J = 6.6$ , 2H), 4.38-4.41 (m, 4H), 4.55-4.59 (m, 4H), 7.29-7.31 (t,  $J = 7.4$ , 4H), 7.36-7.40 (t,  $J = 7.4$ , 4H), 7.57-7.59 (m, 4H) and 7.73-7.75 (d,  $J = 7.4$ , 4H)

**Table 4.4**  $^1\text{H}$ -NMR of *N,N'*-diamino-bis(Fmoc-arginylyl(Pbf)-glycinylyl) pimelic acid (**4**)

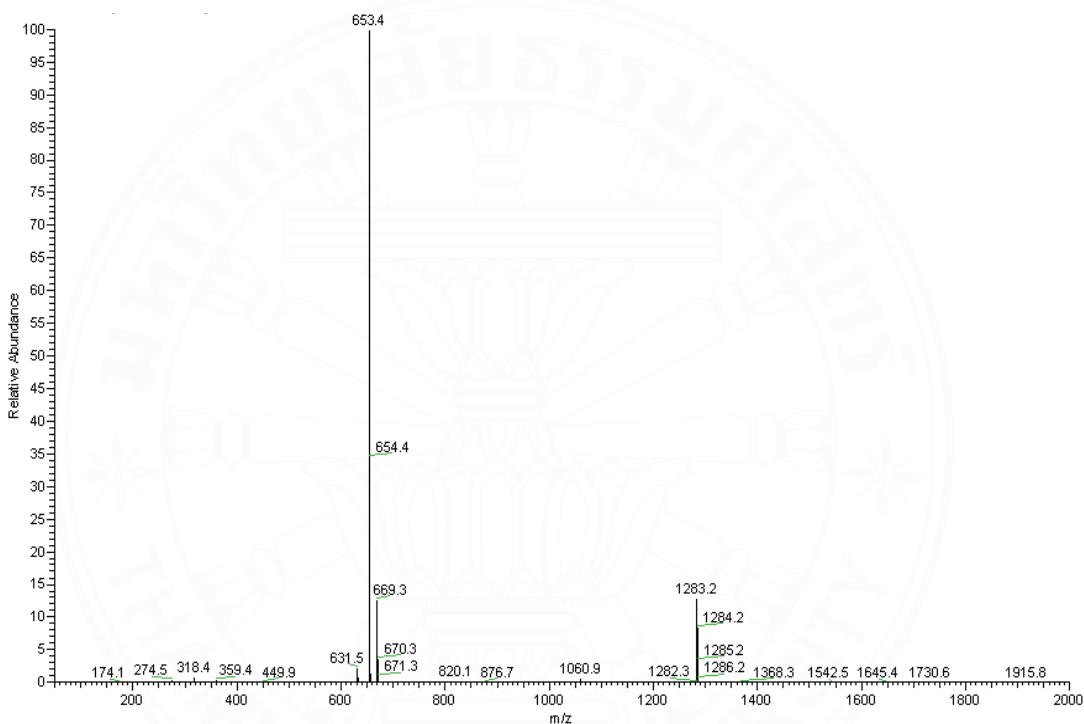
Position number	Chemical shift (ppm)
1,2	1.05-1.37 (m, 6H)
3	1.98-2.02 (m, 2H)
4	4.55-4.59 (m, 4H)
5	4.35-4.37 (d, $J = 7.2$ , 4H)
6	1.89-1.92 (m, 4H)
7	1.55-1.60 (m, 4H)
8	3.36-3.46 (m, 4H)
9	2.50 (s, 6H)
10	2.55 (s, 6H)
11	2.08 (s, 6H)
12	2.95 (s, 4H)
13	1.42-1.47 (m, 12H)
14	4.38-4.41 (m, 4H)
15	4.19-4.22 (t, $J = 6.6$ , 2H)
16	7.57-7.59 (m, 4H)
17	7.29-7.31 (t, $J = 7.4$ , 4H)
18	7.36-7.40 (t, $J = 7.4$ , 4H)
19	7.73-7.75 (d, $J = 7.4$ , 4H)

*N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l pimelic acid was characterized by using  $^1\text{H}$ -NMR (**Figure 4.9**). The result is shown in **Table 4.4**. According to the  $^1\text{H}$ -NMR spectrum, the proton signals ( $\delta$ ) were observed at 1.05-1.37 ppm and 1.98-2.02 ppm corresponding to the methylene protons ( $-\text{CH}_2-$ ) and  $-\text{CH}-$  of 2,6-diaminopimelic acid, respectively. Additionally, the methylene ( $-\text{CH}_2-$ ) protons of glycine moieties were observed at  $\delta = 4.55$ -4.59 ppm. To confirm the dimerization of *N,N'*-diamino-bis( $\text{NH}_2$ -glyciny)l pimelic acid with Fmoc-Arg(Pbf)-OH, the proton signals of arginine side chain which consists of three carbon aliphatic straight chain were observed at  $\delta = 1.55$ -1.60, 1.89-1.92 and 3.36-3.46 ppm. Moreover, the protecting groups (Fmoc and Pbf) were also observed as confirmed by  $^1\text{H}$ -NMR spectroscopy. The aromatic protons of Fmoc moieties were observed at  $\delta = 7.29$ -7.31, 7.36-7.40, 7.57-7.59 and 7.73-7.75 ppm, respectively. It also showed proton signals at  $\delta = 4.19$ -4.22 ppm and 4.38-4.41 ppm corresponding to  $-\text{CH}-$  and  $-\text{CH}_2-$  of Fmoc protecting groups at C15 and C14, respectively. Lastly, the proton signals of Pbf protecting groups at arginine residue were observed, and showed the signals at  $\delta = 2.50$  and 2.55 ppm corresponding to methyl moiety ( $-\text{CH}_3$ ) at ortho position at the Pbf moiety. The proton resonance signals of methyl ( $-\text{CH}_3-$ ) at other positions of Pbf moieties were also observed at 1.42-1.47 and 2.08 ppm. Finally, the resonance signal at 2.95 ppm corresponding to methylene proton of Pbf groups at C12 was also noticed.



**Figure 4.9**  $^1\text{H}$ -NMR spectrum (400 MHz,  $\text{CDCl}_3$ ) of *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l pimelic acid (**4**)

According to ESI-MS analyses (**Figure 4.10**), the results showed the molecular mass of *N,N'*-diamino-bis(Fmoc-arginylyl(Pbf)-glycinylyl) pimelic acid without any protecting group. The molecular mass of *N,N'*-diamino-bis(NH<sub>2</sub>-argenylyl-glycinylyl) pimelic acid was calculated to be  $m/z = 616.32$  corresponding to C<sub>79</sub>H<sub>96</sub>N<sub>12</sub>O<sub>18</sub>S<sub>2</sub>. The molecular mass adducts were also observed at  $m/z = 654.4$ , 670.3, 1283.2 corresponding to  $[M+K]^+$ ,  $[M+CH_3OH+Na]^+$  and  $[2M+CH_3OH+Na]^+$ , respectively.

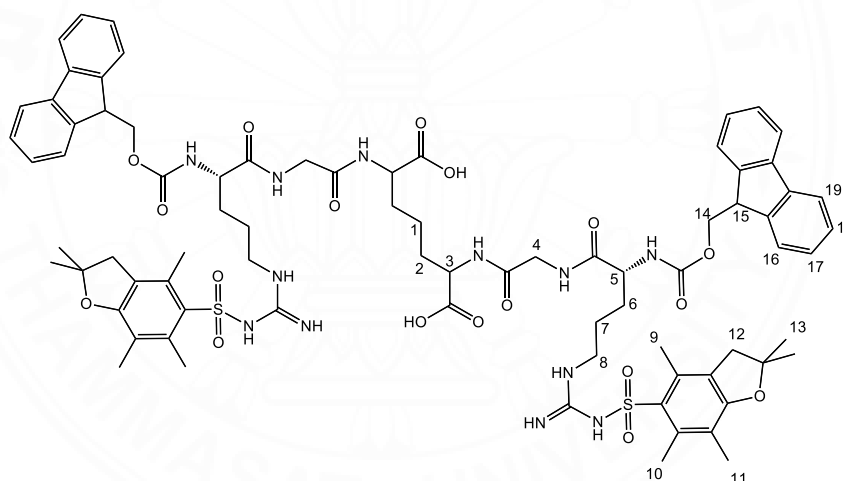


**Figure 4.10** ESI-MS spectrum (via positive mode) of *N,N'*-diamino-bis(Fmoc-argenylyl(Pbf)-glycinylyl) pimelic acid (**4**)

Found 654.4  $[M+K]^+$ , 670.3  $[M+Na]^+$

Calculated molecular mass of *N,N'*-diamino-bis(NH<sub>2</sub>-argenylyl-glycinylyl)pimelic acid  
 $m/z = 616.32$  [C<sub>79</sub>H<sub>96</sub>N<sub>12</sub>O<sub>18</sub>S<sub>2</sub>]<sup>+</sup>

Typically, electrospray ionization mass spectroscopy (ESI-MS) is a useful technique used in determining the molecular weight (MW) in peptide chemistry. The important feature of ESI is its ability to generate multiply charged species. Observing multiple peaks for the same peptide allows us to make multiple molecular weight calculations to obtain a more accurate molecular weight. However, according to ESI-MS results of *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l pimelic acid (**Figure 4.10**), only molecular mass adducts of *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l pimelic acid with all protecting group losses were observed. To measure an accurate molecular weight of the peptide product, the *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny)l pimelic acid was deprotected an Fmoc moiety by using 20% piperidine in DMF and subsequently purify by semi-preparative HPLC to obtain *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glyciny)l pimelic acid (**Figure 4.11**).



**Figure 4.11** The structure of *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glyciny)l pimelic acid

Name: *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glyciny)l pimelic acid (**4a**)

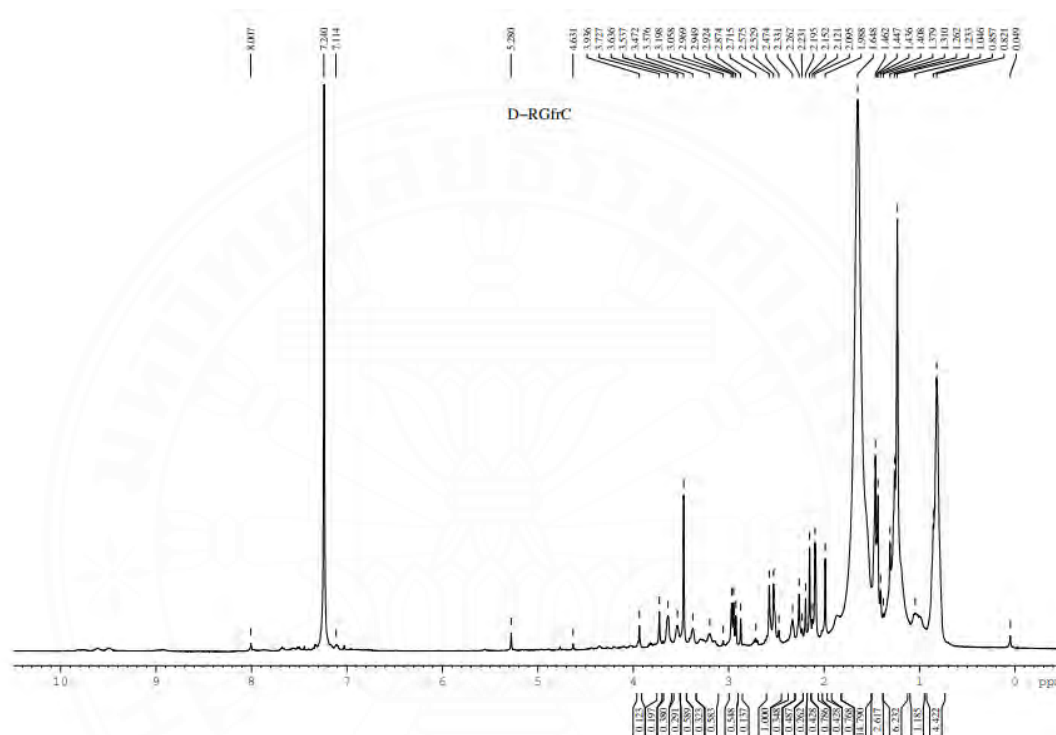
Chemical formula: C<sub>49</sub>H<sub>76</sub>N<sub>12</sub>O<sub>14</sub>S<sub>2</sub>

ESI-MS: Found 1121.4576 [M+H]<sup>+</sup>, 1144.4171 [M+Na]<sup>+</sup> and 1157.5385 [M+K-2H]<sup>-</sup> ;

Calculated 1120.50 [M]<sup>+</sup>

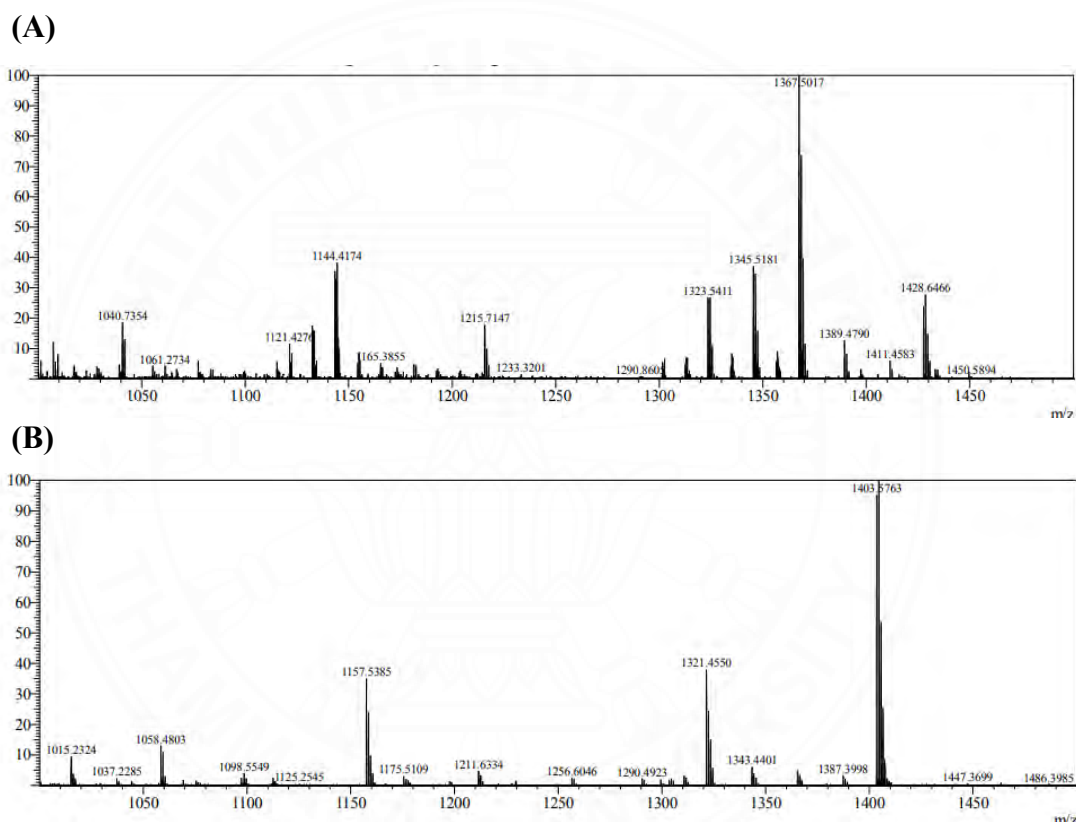
HPLC (method 1): t<sub>R</sub> = 12.173 min

Then, the resulting peptide was characterized by  $^1\text{H}$ -NMR and ESI-MS spectrometry. According to  $^1\text{H}$ -NMR results (**Figure 4.12**), we could identify the proton signals at 2.39 and 2.44 ppm, corresponding to all methyl group of Pbf moiety, and showed the similar pattern to that of *N,N'*-diamino-bis(Fmoc-arginy1(Pbf)-glyciny1) pimelic acid.



**Figure 4.12**  $^1\text{H}$  NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of *N,N'*-diamino-bis( $\text{NH}_2$ -arginy1(Pbf)-glyciny1)pimelic acid (**4a**)

According to HR-ESI-MS analyses, the results showed the molecular mass adduct at  $m/z = 1121.4576$  and  $1144.4171$ , corresponding to  $[M+H]^+$  and  $[M+Na]^+$  (**Figure 4.13A**). While, HR-ESI-MS (*via* negative mode) as shown in **Figure 4.12B**, it showed the molecular mass adducts of  $[M+K-2H]^-$  ion at  $m/z = 1157.5385$ . While, the molecular mass of *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glycidyl) pimelic acid was calculated to be  $m/z = 1120.50$  [ $C_{49}H_{76}N_{12}O_{14}S_2$ ].

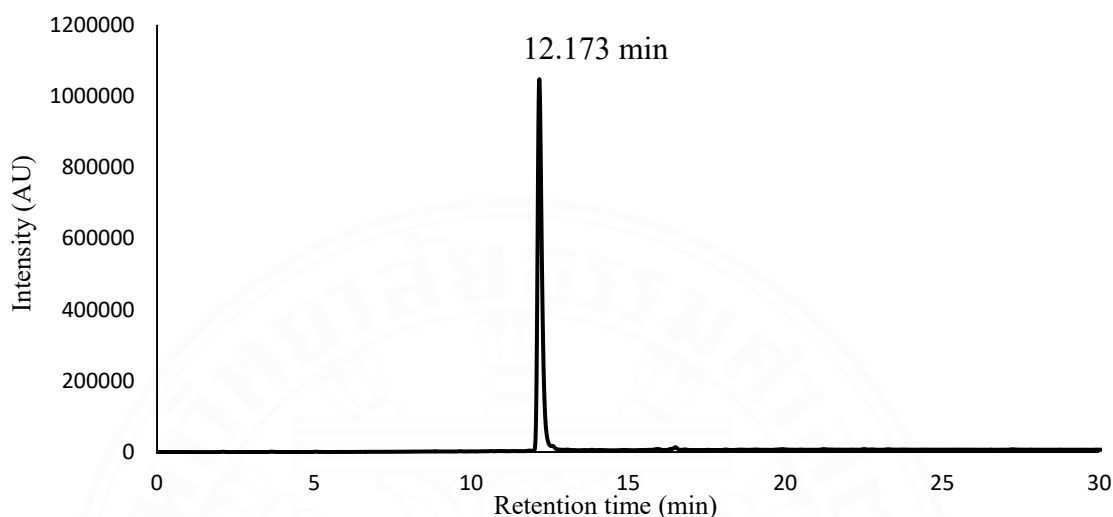


**Figure 4.13** HR-ESI-MS (*via* positive and negative ion mode) of *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glycidyl) pimelic acid (**4a**).

**(A)** ESI-MS spectrum (positive mode) showed  $m/z = 1121.4576$  and  $1144.4171$  corresponding to  $[M+H]^+$  and  $[M+Na]^+$

**(B)** ESI-MS spectrum (negative mode) showed  $m/z = 1157.5385$  corresponding to  $[M+K-2H]^-$

The HPLC chromatogram of purely *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glyciny) pimelic acid measured at 215 nm. showed a peak at 12.173 min as shown in **Figure 4.14**.

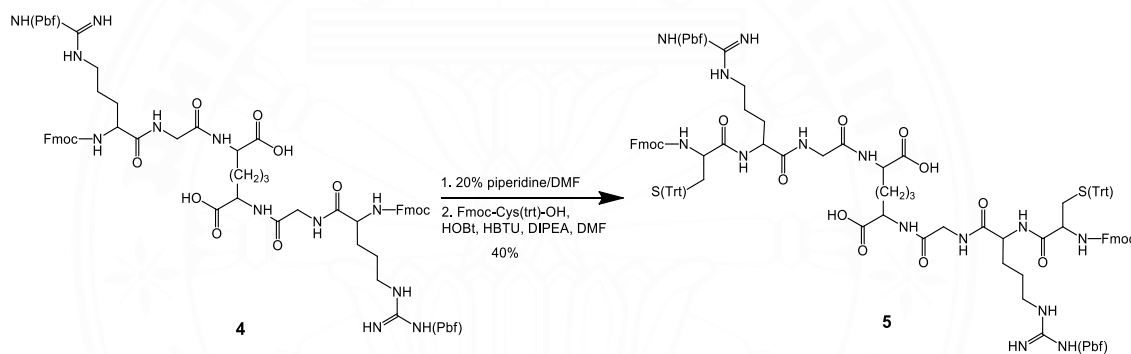


**Figure 4.14** HPLC chromatogram of *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glyciny) pimelic acid (**4a**) measured at 215 nm.

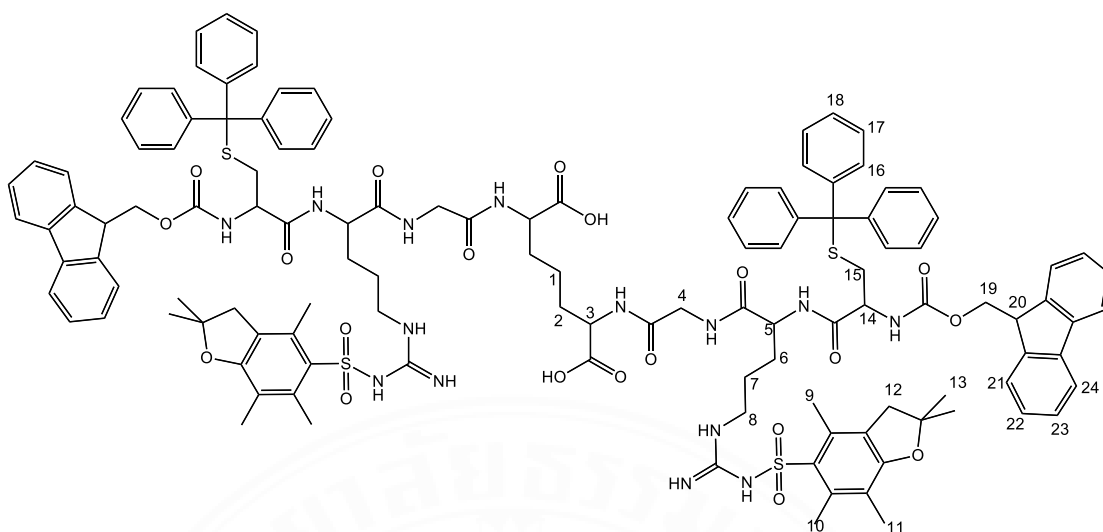
According to all spectroscopic results (<sup>1</sup>H-NMR, ESI-MS and analytical HPLC analyses) of both fully protected peptide (*N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny) pimelic acid) and Fmoc deprotected peptide (*N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glyciny) pimelic acid). It clearly confirmed that the synthesis of *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glyciny) pimelic acid was achieved perfectly. However, it is necessary to purify Fmoc deprotected peptide before further coupling in each step. In this case, we have done to clearly confirm the double attachment of an amino acid (the dimerization). Generally, the Fmoc deprotected crude peptide was usually coupled with the next amino acid residues without any further purification.

### 4.1.3 The synthesis of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl (Pbf)-glycyl) pimelic acid

To synthesize *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycyl) pimelic acid, *N,N'*-diamino-bis(Fmoc-arginyl(Pbf)-glycyl) pimelic acid was treated with 20% piperidine in DMF to afford two free amino groups (-NH<sub>2</sub>) of the arginine moieties which were subsequently coupled by Fmoc-Cys(Trt)-OH using standard coupling reagents (HBTU/HOBt) in DMF to yield *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycyl) pimelic acid in modest yield (40%) as shown in **Scheme 4.3**.



**Scheme 4.3** The synthesis of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycyl)pimelic acid



**Figure 4.15** The structure of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginylyl(Pbf)-glycinylyl) pimelic acid

Name: *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginylyl(Pbf)-glycinylyl) pimelic acid (**5**)

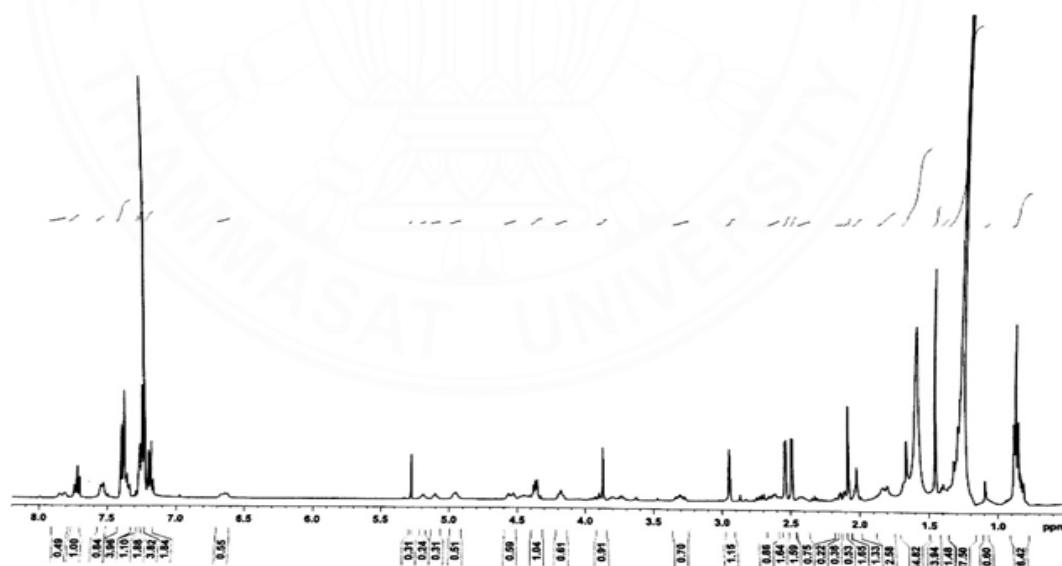
Chemical formula:  $C_{123}H_{134}N_{14}O_{20}S_4$

ESI-MS: Found 674.3  $[M+H+Na]^{2+}$ , 1326.7  $[M+H]^+$

Calculated fragment:  $m/z = 1326.5 [C_{55}H_{86}N_{14}O_{16}S_4]^+$

$^1H$ -NMR (400MHz,  $CDCl_3$ )  $\delta$  = 1.23-1.31 (m, 6H), 1.39-1.44 (s, 12H), 1.57-1.70 (m, 8H), 2.01-2.17 (m, 6H), 2.48 (s, 6H), 2.53 (s, 6H), 2.94 (s, 4H), 3.26-3.33 (m, 4H), 3.71-3.93 (m, 4H), 4.17-4.19 (t,  $J = 6.4$ , 2H), 4.35-4.37 (m, 4H), 4.52-4.56 (m, 2H), 7.17-7.27 (m, 30H), 7.34-7.40 (m, 12H) and 7.70-7.74 (t,  $J = 7.3$ , 4H)

The structure of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycyl) pimelic acid was verified by using  $^1\text{H}$ -NMR (**Figure 4.16**) and ESI-MS analyses. According to  $^1\text{H}$ -NMR results (**Table 4.5**), the proton signal of 2,6-diaminopimelic acid was observed at  $\delta = 1.23$ - $1.31$  ppm. The proton signals corresponding to methylene proton ( $-\text{CH}_2-$ ) of glycine and the methine proton ( $-\text{CH}-$ ) of arginine were observed at  $\delta = 3.71$ - $3.93$  ppm. The spectrum represented methylene protons ( $-\text{CH}_2-$ ) of arginine at  $\delta = 1.57$ - $1.70$  and  $\delta = 3.26$ - $3.33$  ppm and we also found the proton signals corresponding to methine protons of Pbf protecting groups (C9, C10, C11, C13) at  $\delta = 2.48$ ,  $2.53$ ,  $2.01$ - $2.17$  and  $1.39$ - $1.44$  ppm, respectively. To confirm the coupling between Fmoc-Cys(Trt)-OH and *N,N'*-diamino-bis( $\text{NH}_2$ -arginyl(Pbf)-glycyl) pimelic acid, the resonance signals at  $\delta = 4.52$ - $4.56$  ppm corresponding to the methine proton ( $-\text{CH}-$ ) of cysteine were observed. Furthermore, the methine protons of aromatic moiety of triphenylmethyl groups (Trt) and Fmoc groups on cysteine were shown at  $\delta = 7.17$ - $7.27$ ,  $7.34$ - $7.40$  and  $7.70$ - $7.74$  ppm, respectively.

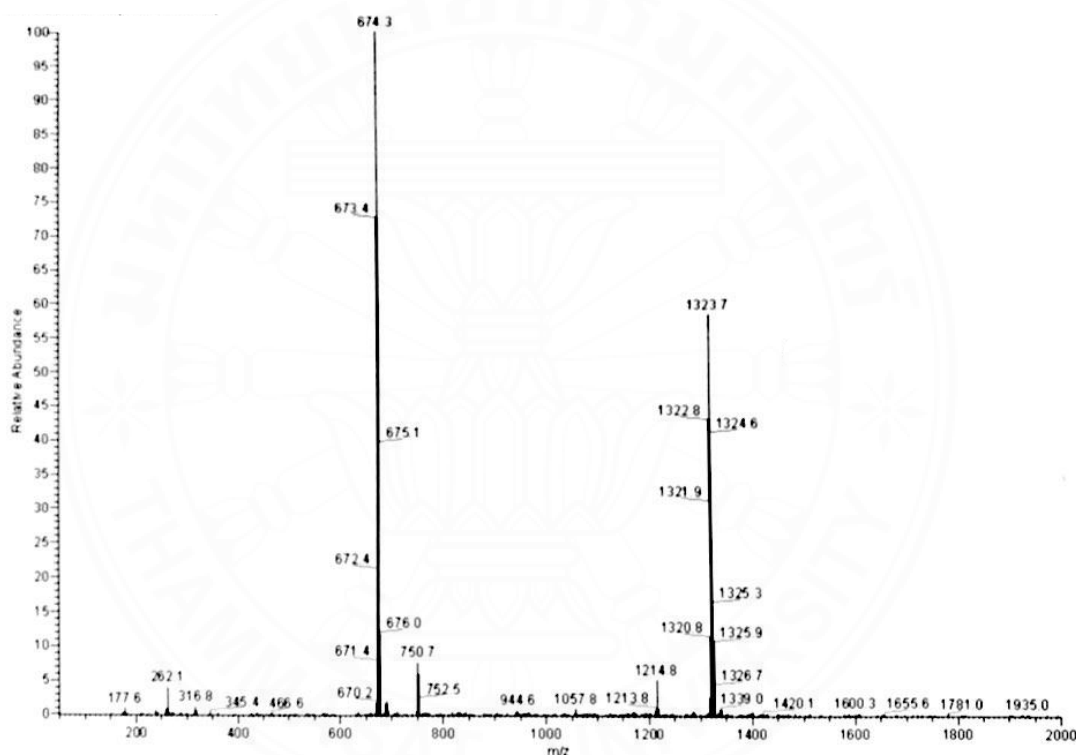


**Figure 4.16**  $^1\text{H}$  NMR spectrum (400 MHz,  $\text{CDCl}_3$ ) of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycyl) pimelic acid (**5**)

**Table 4.5**  $^1\text{H}$ -NMR of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid (**5**)

Position number	Chemical shift (ppm)
1,2	1.23-1.31 (m, 6H)
3	-
4,5	3.71-3.93 (m, 4H)
6,7	1.57-1.70 (m, 8H)
8	3.26-3.33 (m, 4H)
9	2.48 (s, 6H)
10	2.53 (s, 6H)
11	2.01-2.17 (m, 6H)
12	2.94 (s, 4H)
13	1.39-1.44 (s, 12H)
14	4.52-4.56 (m, 2H)
15	-
16,17,18	7.17-7.27 (m, 30H)
19	4.35-4.37 (m, 4H)
20	4.17-4.19 (t, $J = 6.4$ , 2H)
21,22,23	7.34-7.40 (m, 12H)
24	7.70-7.74 (t, $J = 7.3$ , 4H)

According to ESI-MS analyses (**Figure 4.17**), the results clearly indicated that losses of two Fmoc moieties and two Trt moieties of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid were found. The adducts of molecular mass were observed at  $m/z = 674.3$  and  $1326.7$  corresponding to  $[M+H+Na]^+$  and  $[M+H]^+$ , respectively. While, the molecular mass of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl-arginyl(Pbf)-glyciny)l pimelic acid was calculated to be  $m/z = 1326.52$  (C<sub>55</sub>H<sub>86</sub>N<sub>14</sub>O<sub>16</sub>S<sub>4</sub>).

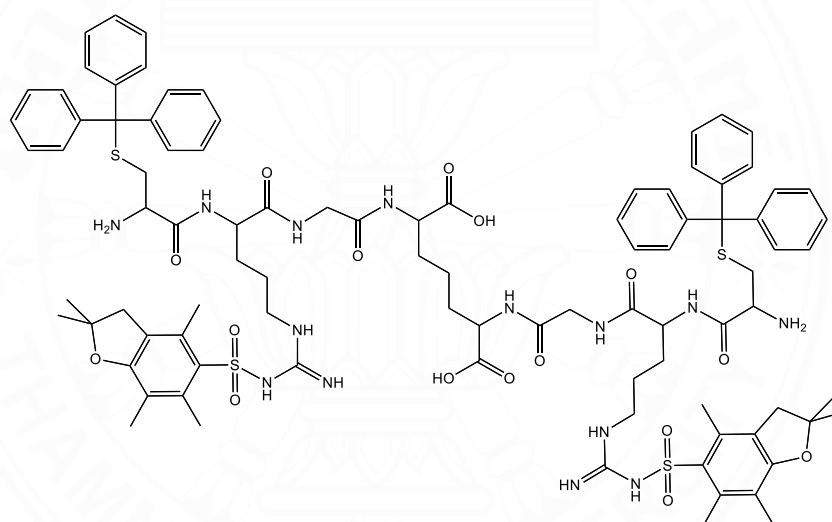


**Figure 4.17** ESI-MS (*via* positive mode) of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycidyl)pimelic acid (**5**).

Calculated fragment  $m/z = 1326.5$  [C<sub>55</sub>H<sub>86</sub>N<sub>14</sub>O<sub>16</sub>S<sub>4</sub>]<sup>+</sup>

ESI-MS: Found 674.3  $[M+H+Na]^{2+}$ , 1326.7  $[M+H]^+$

Similarly, according to ESI-MS results of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glycidyl)pimelic acid, it was found that only molecular mass adducts corresponding fragment with two Fmoc and two Trt moieties lossess (*N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl-glycidyl) pimelic acid) were observed. Therefore, the Fmoc protecting group of fully protected peptide was deprotected by using 20% piperidine/DMF to afford *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glycidyl) pimelic acid (**Figure 4.18**). Subsequently, the corresponding peptide was then characterized by using <sup>1</sup>H-NMR and HR-ESI-MS in order to identify the accurate molecular mass of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glycidyl) pimelic acid.



**Figure 4.18** The structure of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glycidyl) pimelic acid

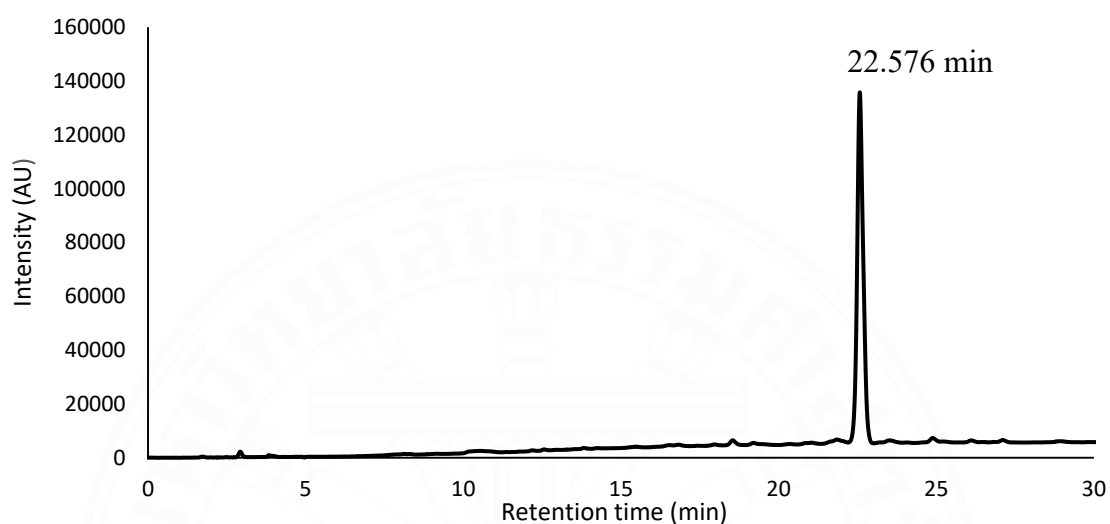
Name: *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glycidyl) pimelic acid (**5a**)

Chemical formula: C<sub>93</sub>H<sub>114</sub>N<sub>14</sub>O<sub>16</sub>S<sub>4</sub>

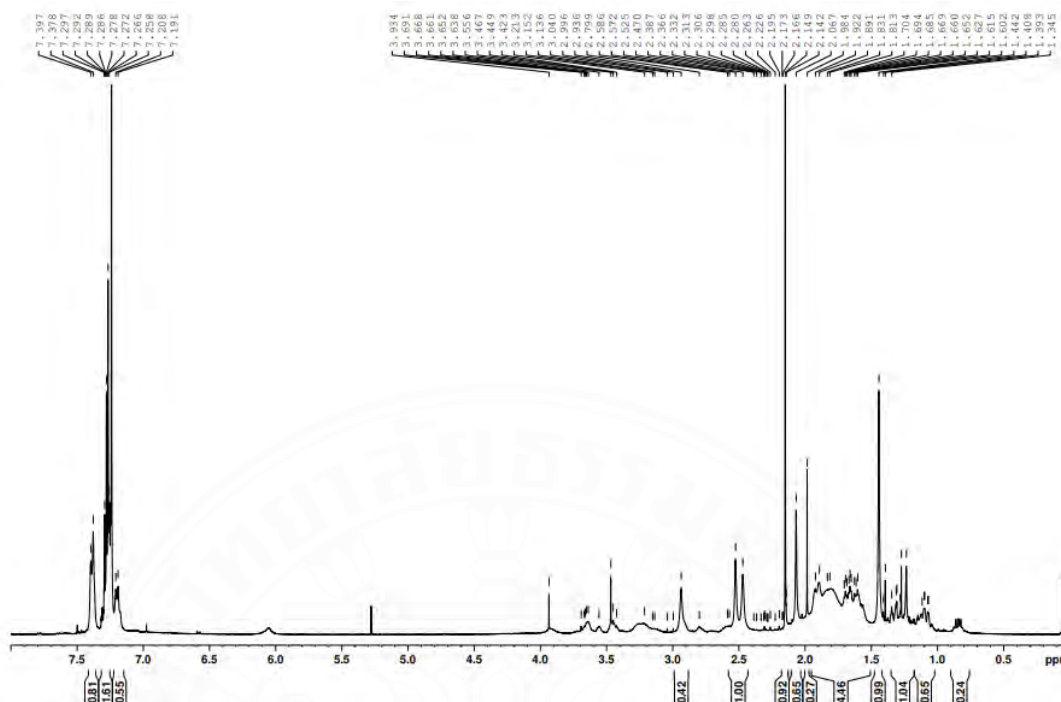
ESI-MS: Found 927.6529 [M+2Na]<sup>2+</sup>, 945.4153 [M+2ACN-2H]<sup>2-</sup>. Calculated 1810.74

HPLC (method 1): t<sub>R</sub> = 22.576 min

The purity was confirmed by analytical HPLC, and it was found that *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid showed a peak at 22.576 min measured at 215 nm (as seen in **Figure 4.19**).



**Figure 4.19** HPLC chromatogram of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl-arginyl(Pbf)-glyciny) pimelic acid (**5a**) measured at 215 nm.

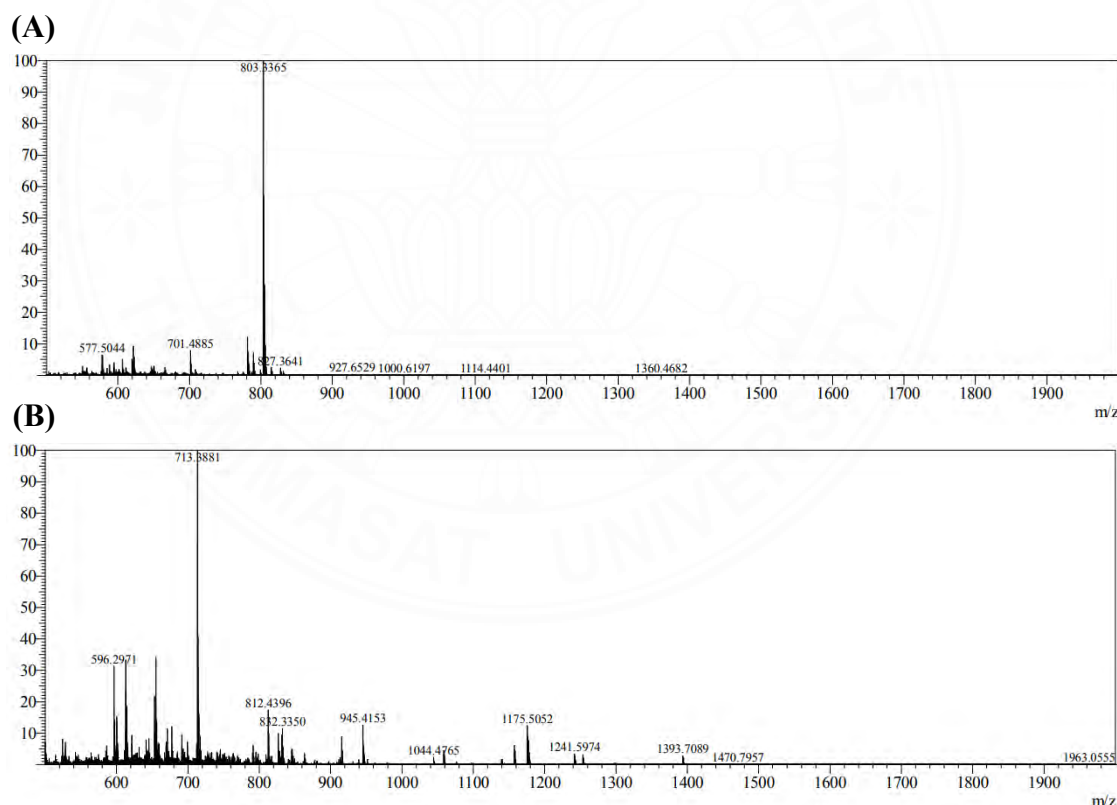


**Figure 4.20**  $^1\text{H}$ -NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of *N,N'*-diamino-bis( $\text{NH}_2$ -cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid (**5a**)

Due to the complex structure of *N,N'*-diamino-bis( $\text{NH}_2$ -cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid. The molecular masses of *N,N'*-diamino-bis( $\text{NH}_2$ -cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid with Fmoc protection and without Fmoc protection were analyzed by using HR-ESI-MS. According to  $^1\text{H}$ -NMR as seen in **Figure 4.20**, it showed the proton signals at  $\delta = 7.19\text{--}7.40$  ppm corresponding to the aromatic protons of Trityl (Trt) moieties which were used as the cysteine protecting groups. While, the  $^1\text{H}$ -NMR signal of Fmoc aromatic protons was disappeared. Meanwhile, the other parts (arginine residues, Pbf groups, glycine and 2,6-diaminopimelic acid core) are similar to that of fully protected peptide (*N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid).

According to HR-ESI-MS analyses (*via* negative mode as shown in **Figure 4.21B**), the molecular mass adduct of  $[M+2ACN-2H]^{2-}$  is showed obviously at  $m/z = 945.4153$ , indicating the molecular mass of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid, which was calculated to be  $m/z = 1810.74$  (C<sub>93</sub>H<sub>114</sub>N<sub>14</sub>O<sub>16</sub>S<sub>4</sub>)

On the other hand, the mass spectrum (*via* positive mode as seen in **Figure 4.21A**) indicated a small peak at  $m/z = 927.6529$  corresponding to  $[M+2Na]^{2+}$ . While the molecular mass of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid was calculated to be  $m/z = 1810.74$  (C<sub>93</sub>H<sub>114</sub>N<sub>14</sub>O<sub>16</sub>S<sub>4</sub>). However, the molecular mass adducts of the *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid with a Pbf group loss was calculated to be 1558.66 [C<sub>80</sub>H<sub>98</sub>N<sub>14</sub>O<sub>13</sub>S<sub>3</sub>], which was observed at  $m/z = 803.3365$  corresponding to  $[M+2Na]^{2+}$ .



**Figure 4.21** High resolution mass spectrum positive and negative ion mode of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid (**5a**)

(A) ESI-MS (positive mode) showed  $m/z = 927.6529$  corresponding to  $[M+2Na]^{2+}$

(B). ESI-MS (negative mode) showed  $m/z = 945.4153$  corresponding to  $[M+2ACN-2H]^{2-}$

To our synthetic strategy, the coupling efficacy of the second amino acid (Fmoc-Arg(Pbf)-OH) and the third amino acid (Fmoc-Cys(Trt)-OH) to the bivalent core were performed by using HBTU instead of DCC, enabling to increase the coupling efficacy. Typically, DMF has been considered as suitable solvent for peptide synthesis. However, the high b.p. of DMF is the major disadvantage. Recently, it has been studied the capacity of THF and ACN in a ratio 1:1 which are lower b.p. relative to that of DMF to yield the expected product in higher purity. Notably, the coupling reagents (HOBt and HBTU) were poorly soluble in THF/ACN compared to DMF, resulting in high dilution that caused a lower yield of an expected product (30%).

Interestingly, the attachment of amino acids to both sides of the bivalent scaffold is very challenging mainly due to the steric effect. To our approach, we found that the double attachment of the first amino acid (Fmoc-Gly-OH) and the second amino acid (Fmoc-Arg(Pbf)-OH) to the bivalent scaffold was achieved perfectly. While, the attachment of Fmoc-Cys(Trt)-OH to the bivalent core (*N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l) pimelic acid) was carried out, and a single attachment of Fmoc-Cys(Trt)-OH was also identified as a minor product.

The used conditions of solution phase peptide synthesis are summarized as shown in **Table 4.6**.

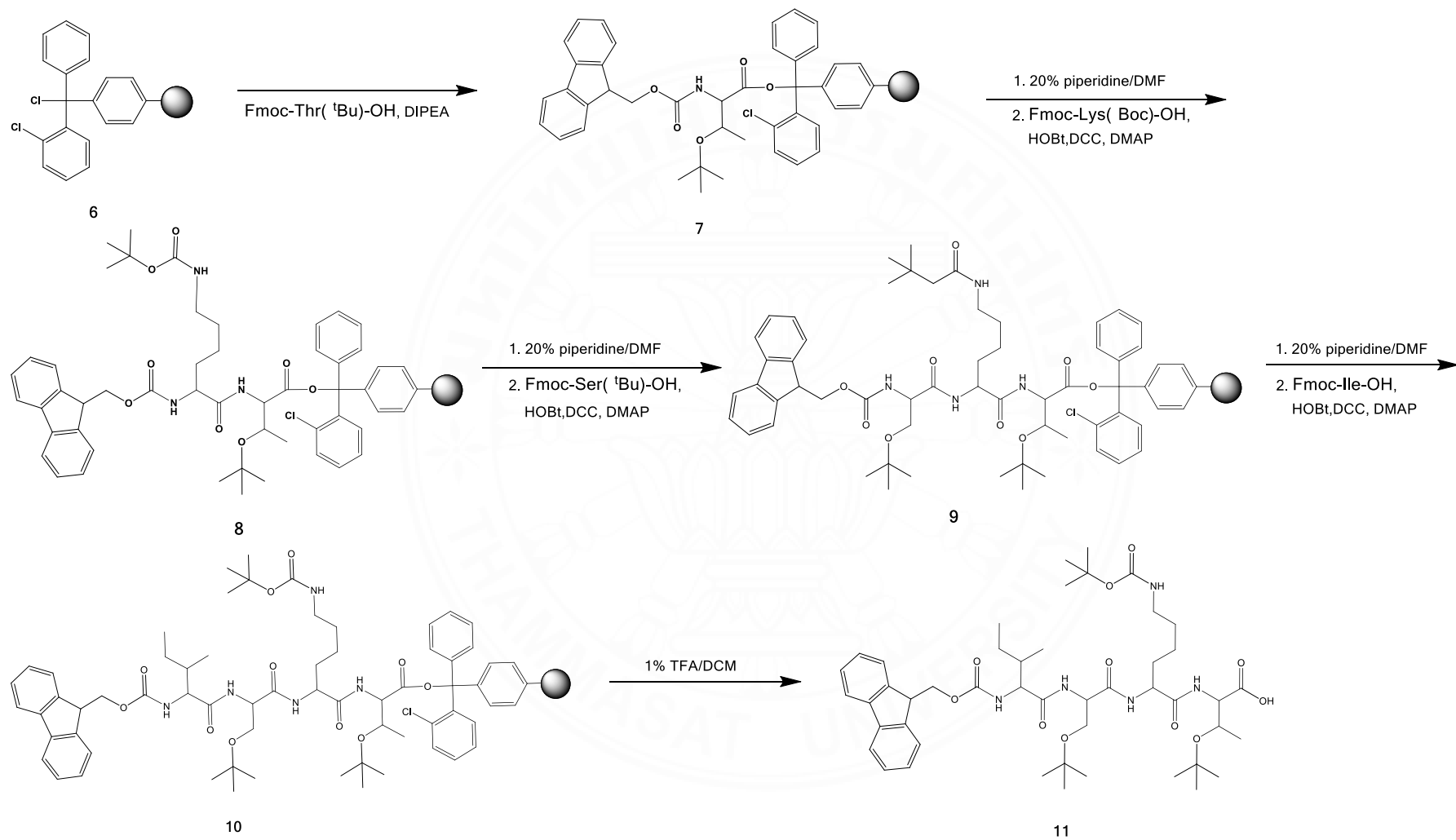
**Table 4.6** The developed condition of solution phase peptide synthesis (SPPS).

Product	Coupling reagents	solvent	Yield (%)
Fmoc-glycine-pimelic acid ( <b>2</b> )	DCC/DMAP	DMF /H <sub>2</sub> O (3:1)	84
	DIC	DMF /H <sub>2</sub> O (3:1)	-
Fmoc-arginine(Pbf)-glycine-pimelic acid ( <b>4</b> )	HBTU/HOBt	DMF	56
		ACN/THF (1:1)	30
Fmoc-cysteine(Trt)-arginine(Pbf)-glycine-pimelic acid ( <b>5</b> )	HBTU/HOBt	DMF	40
		ACN/THF (1:1)	20

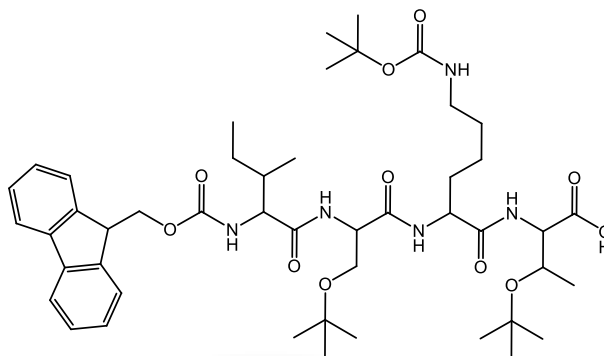
## 4.2 The synthesis of linear peptide chains *via* the solid phase peptide synthesis

To the amino acid assembly strategy, the linear peptide chain comprising of threonine (Thr), lysine (Lys), serine (Ser) and isoleucine (Ile) was assembled *via* Fmoc based solid-phase peptide synthesis (SPPS) approach. Typically, Fmoc strategy was widely used for peptide synthesis. This Fmoc based approach offered a very mild condition for the deprotection process by using usually 20% piperidine in DMF, enabling to remove the Fmoc group at N-terminus. The advantage of Fmoc protection is its stability under the acidic condition. This allowed the use of mild-acid tolerant protecting groups for an amino acid assembly. In this case, *tert*-butyl group (<sup>t</sup>Bu) and *tert*-butyloxycarbonyl group (Boc) were utilized as the side chain protecting groups. Lastly, trifluoroacetic acid (TFA) was used for the cleavage of resin.

According to **Scheme 4.4**, the precursor was synthesized on 2-chlorotrityl resin (**6**). This resin was utilized for an amino acid immobilization due to its acid-labile property that could effectively cleave under mild-acidic condition (1% TFA) to afford the free carboxyl linear peptide chain while all protecting groups in the side chain are kept. Fmoc-threonine(<sup>t</sup>Bu)-OH was initially loaded onto 2-chlorotrityl resin under N<sub>2</sub> condition to prevent the side reaction. It was found that the loading efficacy of Fmoc-threonine(<sup>t</sup>Bu)-OH was approximately 60% which was calculated based on the initial loading of 2-chlorotrityl resin. Then, the preloaded amino residue was treated with 20% piperidine in DMF to afford NH<sub>2</sub>-threonine(<sup>t</sup>Bu)-O-chlorotrityl resin, which was subsequently coupled with Fmoc-lysine(Boc)-OH by using DCC and DMAP as standard coupling reagents whereas HOBT was applied to prevent the racemization. Fmoc-lysine(Boc)-threonine(<sup>t</sup>Bu)-chlorotrityl resin (**8**) was obtained in a good yield. Then, the reaction was repeatedly synthesized as previous step by assembling with Fmoc-serine(<sup>t</sup>Bu)-OH and Fmoc-isoleucine-OH, respectively to afford Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-O-chlorotrityl resin (**10**). To obtain the free carboxyl groups linear peptide chain, The peptide attached on 2-chlorotrityl resin was subsequently cleaved from the solid support by treating with 1% TFA in DCM to yield Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH (**11**) with a total yield of 60%.



**Scheme 4.4.** The synthesis of Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH



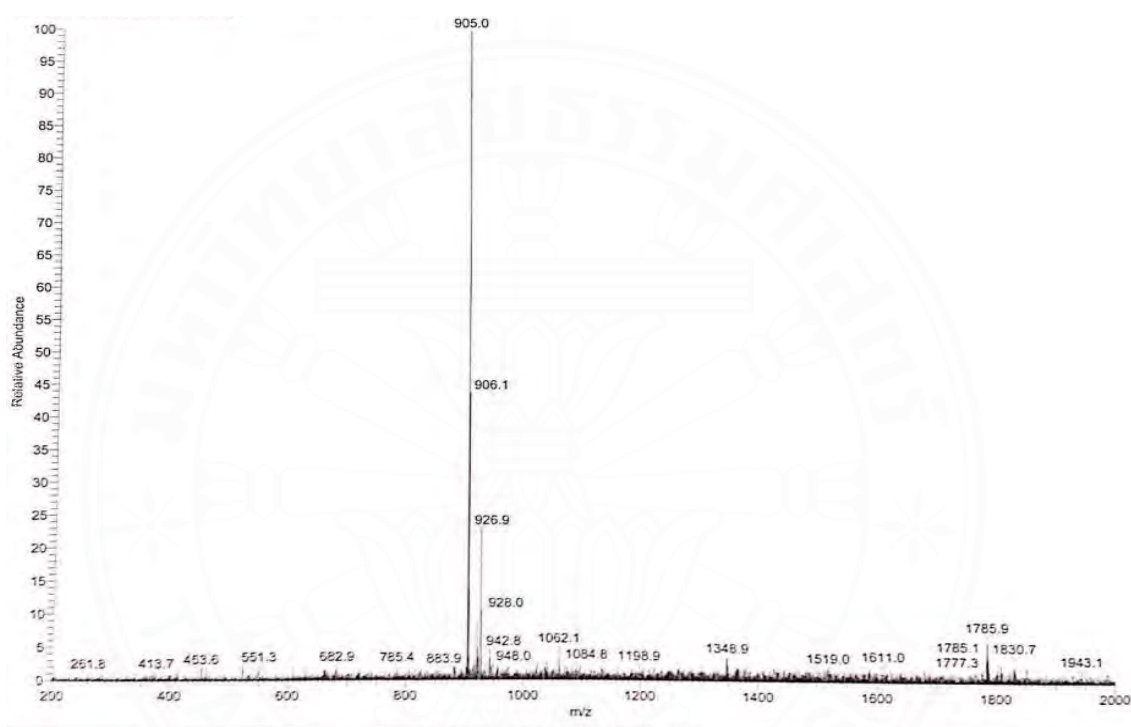
**Figure 4.22** The structure of Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH

Name: Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH (**11**)

Chemical formula: C<sub>47</sub>H<sub>71</sub>N<sub>5</sub>O<sub>11</sub>

ESI-MS: Found 905.0[M+Na]<sup>+</sup>, 1785.0 [2M+Na]<sup>+</sup> ; Calculated 881.52 [M]<sup>+</sup>

The linear peptide chain was characterized by ESI-MS analyses (**Figure 4.23**). The ESI-MS spectrum of Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH demonstrated the molecular mass at  $m/z = 905.0$  and 1785.0, corresponding molecular mass adduct  $[M+Na]^+$  and  $[2M+Na]^+$  in which the molecular mass was calculated to be 881.52.



**Figure 4.23** ESI-MS (*via* positive mode) of Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH (**11**).

### 4.3 The ligation and cyclization

#### 4.3.1 The peptide ligation

To our synthetic strategy, we combined two strategies for the peptide syntheses by using both solution phase and solid phase peptide synthesis (SPPS) for the chemical ligation step. The tripeptide attached with scaffold named *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid (**5**), was treated with 20% piperidine in DMF to give free two amino moieties at *N*-terminus of cysteine, which was subsequently coupled with free carboxyl groups of linear peptide chain named Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH (**11**) by using standard coupling reagents HBTU/HOBt and DIPEA as a base in DMF to afford *N,N'*-diamino-bis(Fmoc-isoleuciny)l-seriny)l(<sup>t</sup>Bu)-lysiny)l(Boc)-threony)l(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid with the modest yield (43%) as shown in **Scheme 4.5**.

Under this developed condition, the ligation was carried out by using HBTU as a coupling reagent to obtain a ligated product *N,N'*-diamino-bis(Fmoc-isoleuciny)l-seriny)l(<sup>t</sup>Bu)-lysiny)l(Boc)-threony)l(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid with an acceptable yield (43%). To get a better yield in a higher scale reaction, we performed a synthetic approach by using other uronium based coupling agent named HCTU. Unfortunately, HBTU showed cleaner reaction results relative to that of HCTU as confirmed by TLC experiment. To our study, we performed the ligation with two different conditions: (1) using DMF as a solvent and (2) using the combination of ACN/THF (1:1) as a solvent. According to the solution phase peptide synthesis as previously described, we found that DMF is the effective solvent for both coupling reaction and ligation as shown in **Table 4.7**.

**Table 4.7** The conditions of ligation.

Coupling reagents	Solvent	Yield (%)
HCTU	ACN/THF (1:1)	-
	DMF	31
HBTU/HOBt	ACN/THF (1:1)	23
	DMF	43

Additionally, we performed the ligation processes with different length of peptide precursors (from SPPS); such as (1) Fmoc-ISKTCRG-OH, (2) Fmoc-ISKTCR-OH, (3) Fmoc-ISKTC-OH and (4) Fmoc-ISK-T-OH as shown in **Table 4.8**.

To the first approach, the ligation process was performed by coupling 2,6-diaminopimelic acid scaffold with a 7 residue peptide chain (Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-cysteine(Trt)-arginine(Pbf)-glycine) synthesized by SPPS, and HOBt/HBTU is used as a coupling agent.

Unfortunately, the amide bond formation did not emerge probably due to the steric hindrance

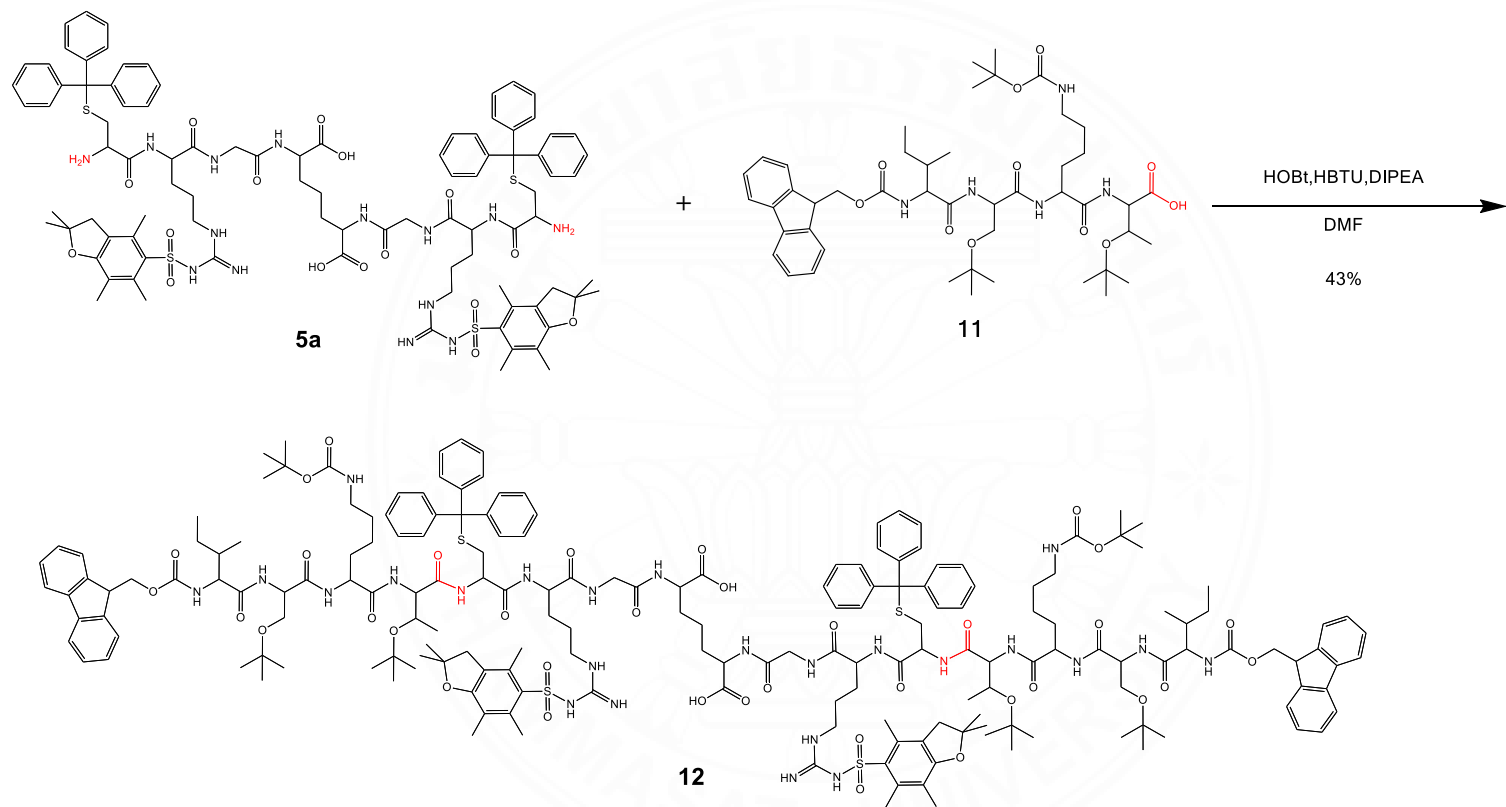
Secondly, we performed the ligation reaction between *N,N'*-diamino-bis(NH<sub>2</sub>-glyciny) pimelic acid and a 6-residue peptide chain (Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-cysteine(Trt)-arginine(Pbf)-OH). Unluckily, the ligation reaction did not take place. Additionally, the ligation between *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl-glyciny) pimelic acid and a 5-residue peptide chain (Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-cysteine(Trt)) did not take place also.

According to these results, it clearly indicated that only short peptide chain with 4 residues (Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)) was perfectly ligated to form an amide bond between cysteine and threonine.

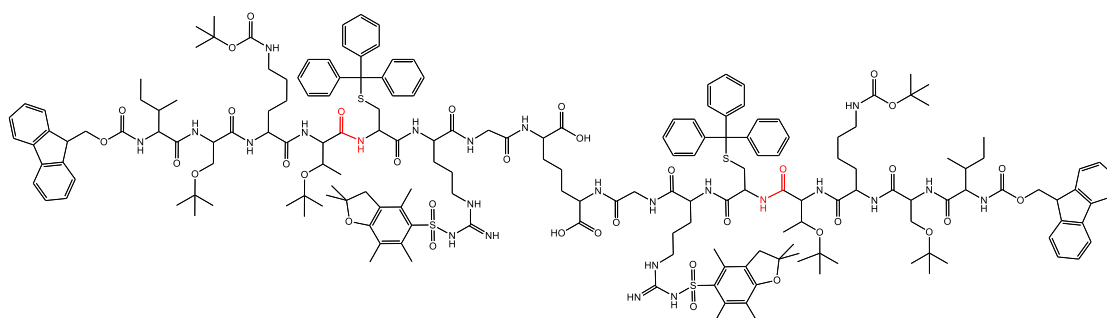
**Table 4.8** The development of ligation approach

Peptide precursors		results
Synthesis <i>via</i> solution phase	Synthesis <i>via</i> solid phase	
2,6-diaminopimelic acid	Fmoc-ISKTCRG-OH	-
NH <sub>2</sub> -G-pimelic acid	Fmoc-ISKTCR-OH	-
NH <sub>2</sub> -RG-pimelic acid	Fmoc-ISKTC-OH	-
NH <sub>2</sub> -CRG-pimelic acid	Fmoc-ISK-T-OH	<b>43%</b>





**Scheme 4.5.** The synthesis of *N,N'*-diamino-bis(Fmoc-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid



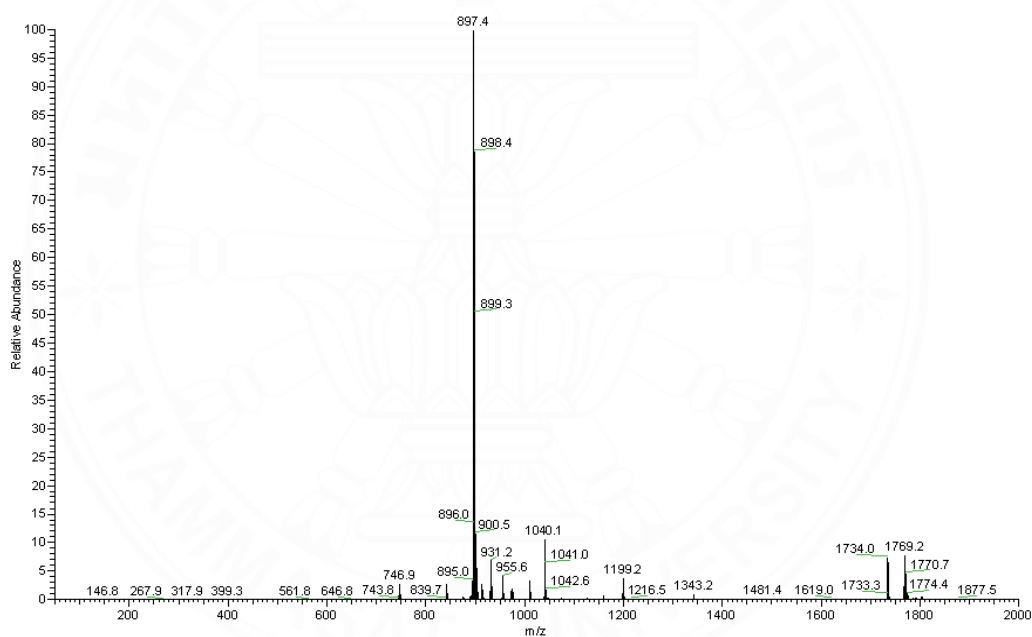
**Figure 4.24** The structure of *N,N'*-diamino-bis(Fmoc-isoleucyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid

Name: *N,N'*-diamino-bis(Fmoc-isoleucyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid (**12**)

Chemical formula: C<sub>187</sub>H<sub>252</sub>N<sub>24</sub>O<sub>36</sub>S<sub>4</sub>

ESI-MS: Found 897.4 [M+ACN+4H]<sup>4+</sup>, 1769.2 [M+2H]<sup>2+</sup> ; Calculated 3537.75 [M]<sup>+</sup>

The resulting ligation product was subsequently characterized by using ESI-MS analyses. According to ESI-MS spectrum (**Figure 4.25**), the molecular mass adducts were observed at  $m/z = 897.4$  and  $1769.2$  corresponding  $[M+ACN+4H]^{4+}$  and  $[M+2H]^{2+}$ , respectively. While the calculated molecular mass of fully protected ligation product, *N,N'*-diamino-bis(Fmoc-isoleucynyl-serinyl(*t*Bu)-lysiny(*Boc*)-threonyl(*t*Bu)-cysteinyl(*Trt*)-arginyl(*Pbf*)-glyciny) pimelic acid was found to be 3537.75 ( $C_{187}H_{252}N_{24}O_{36}S_4$ ). This clearly confirmed that the ligation of *N,N'*-diamino-bis(Fmoc-cysteinyl(*Trt*)-arginyl(*Pbf*)-glyciny) pimelic acid and Fmoc-isoleucine-serine(*t*Bu)-lysine(*Boc*)-threonine(*t*Bu)-OH was successful.

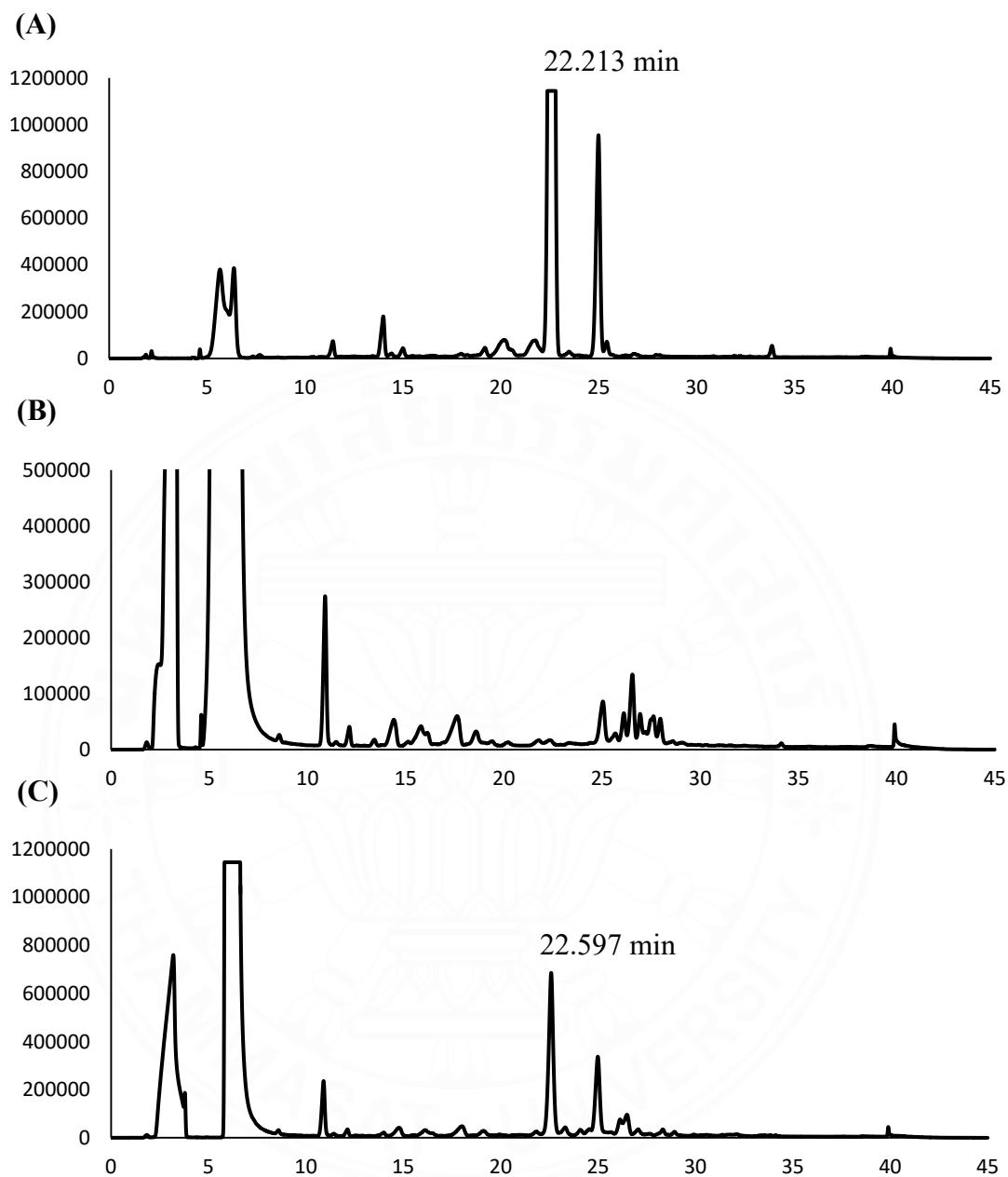


**Figure 4.25** ESI-MS spectrum (via Positive mode) of *N,N'*-diamino-bis(Fmoc-isoleucynyl-serinyl(*t*Bu)-lysiny(*Boc*)-threonyl(*t*Bu)-cysteinyl(*Trt*)-arginyl(*Pbf*)-glyciny) pimelic acid (**12**).

The progress of the ligation reaction is confirmed by HPLC experiment measured at 215 nm. As shown in **Figure 4.26 (A)** represented the HPLC chromatogram of Fmoc deprotected tripeptide precursor (*N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid), **(B)** showed the activation of the carboxyl moieties (Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH) and **(C)** demonstrated the ligation reaction, respectively.

According to **Figure 4.26C**, it showed the major peak at 22.597 min. which is closer to that of the tripeptide precursor (*t<sub>R</sub>* = 22.213 min.) as shown in **Figure 4.26A**. Then, the corresponding ligation product was characterized by <sup>1</sup>H-NMR and ESI-MS spectroscopy.





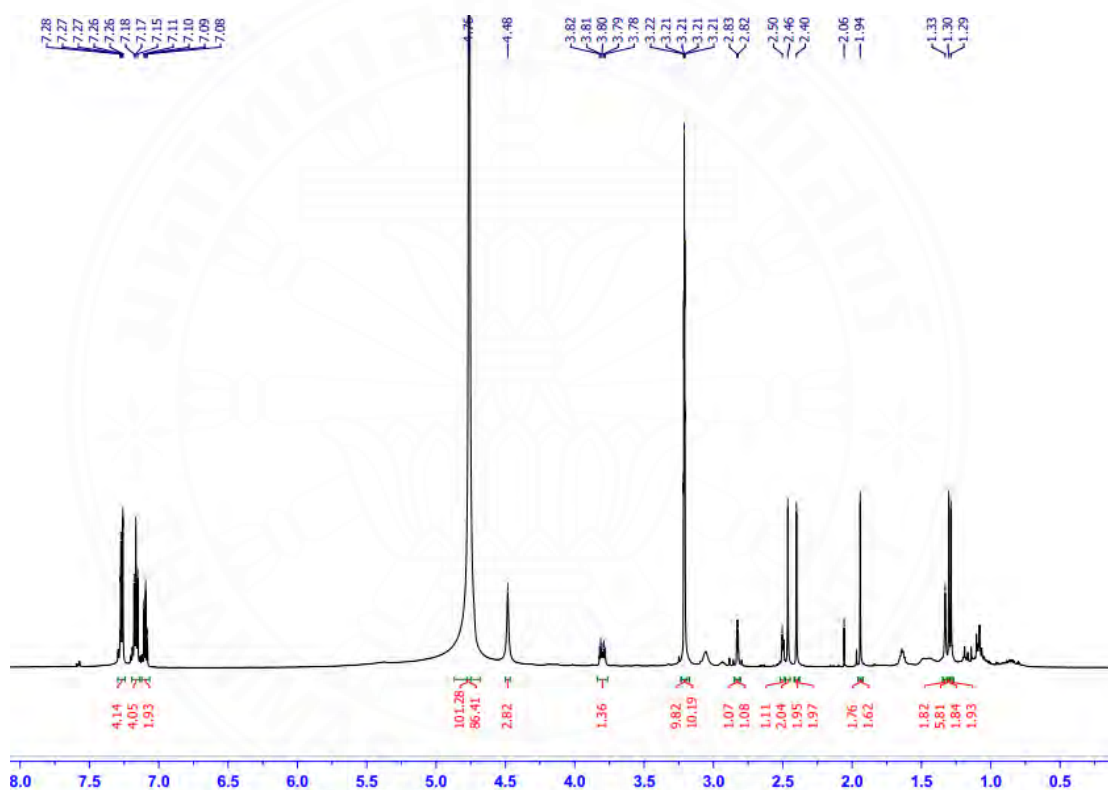
**Figure 4.26** The comparison HPLC chromatogram of ligation reaction and peptide precursor measure at 215 nm.

(A) Fmoc deprotection reaction crude of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid (**5**)

(B) carboxyl activation reaction of Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH (**11**)

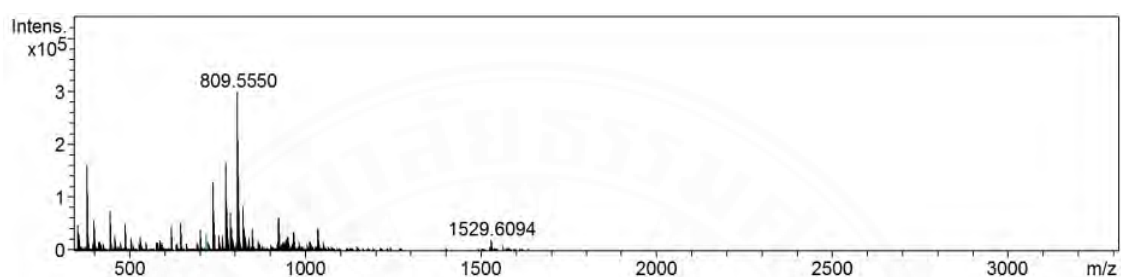
(C) the crude ligation reaction

The corresponding ligated peptide was purified and characterized by  $^1\text{H}$ -NMR and ESI-MS analyses. According to the NMR result as shown in **Figure 4.27**, it showed the resonance signal of trityl groups at cysteine at  $\delta = 7.083\text{--}7.275$  ppm. Additionally, resonance signals corresponding to methyl groups at ortho positions of Pbf groups at arginine sequence were also observed. The Boc and *tert*-butyl groups on lysine, threonine and serine sequence of peptide chain showed in the  $^1\text{H}$ -NMR spectrum at  $\delta = 1.25$  ppm.



**Figure 4.27**  $^1\text{H}$  NMR spectrum (600 MHz,  $\text{CDCl}_3$ ) of the corresponding ligated peptide.

According to ESI-MS analysis of corresponding ligated peptide as shown in **Figure 4.28**, it represented the molecular mass adduct at  $m/z = 809.5550$  and  $1529.6094$  corresponding to  $[M+ACN+4K]^{4+}$  and  $[M+H+Na]^{2+}$ , respectively. While fragment of ligated product with two Fmoc and one *tert*-butyl group loss was calculated to be  $3037.55$  amu ( $C_{153}H_{224}N_{24}O_{32}S_4$ ).

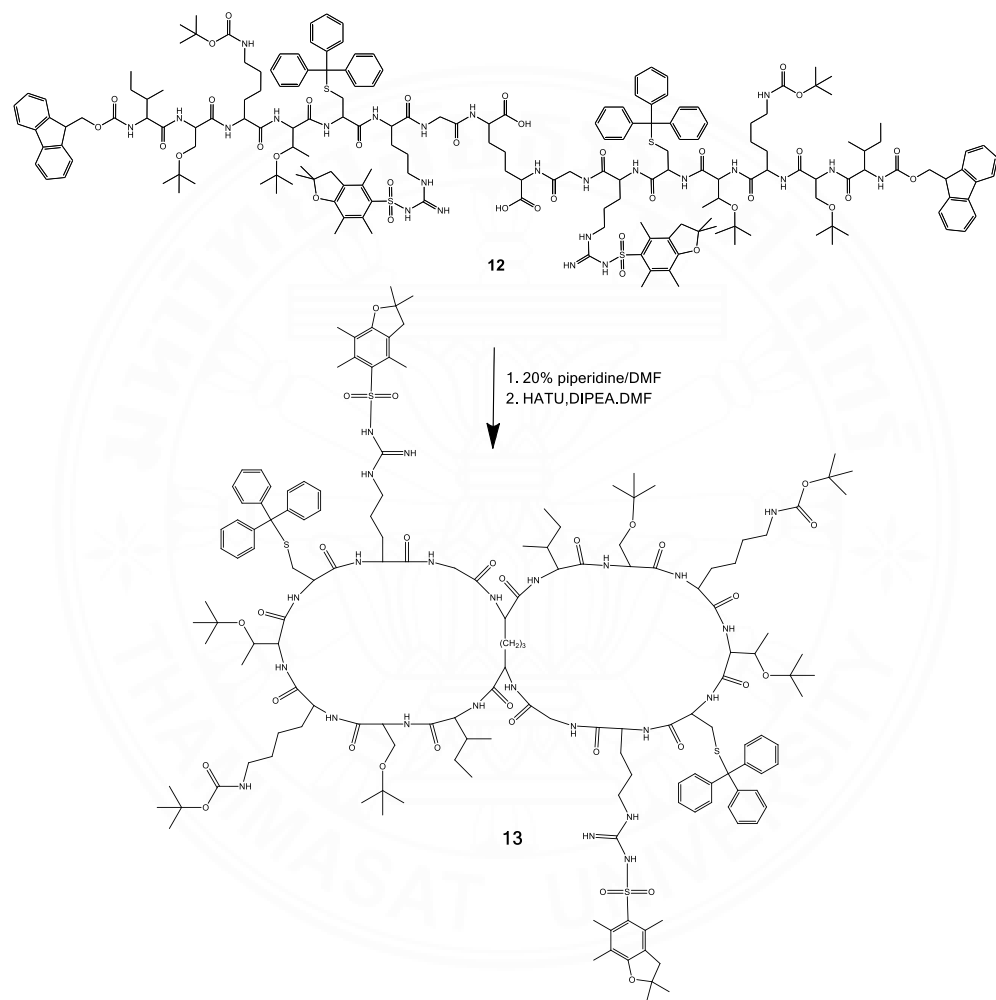


**Figure 4.28** Positive mode mass spectrum of corresponding ligation by product. Found  $809.5550$   $[M+ACN+4K]^{4+}$ ,  $1529.6094$   $[M+H+Na]^{2+}$ ; Calculated fragment  $3037.55$   $[C_{153}H_{224}N_{24}O_{32}S_4]^+$

### 4.3.2 The Cyclization

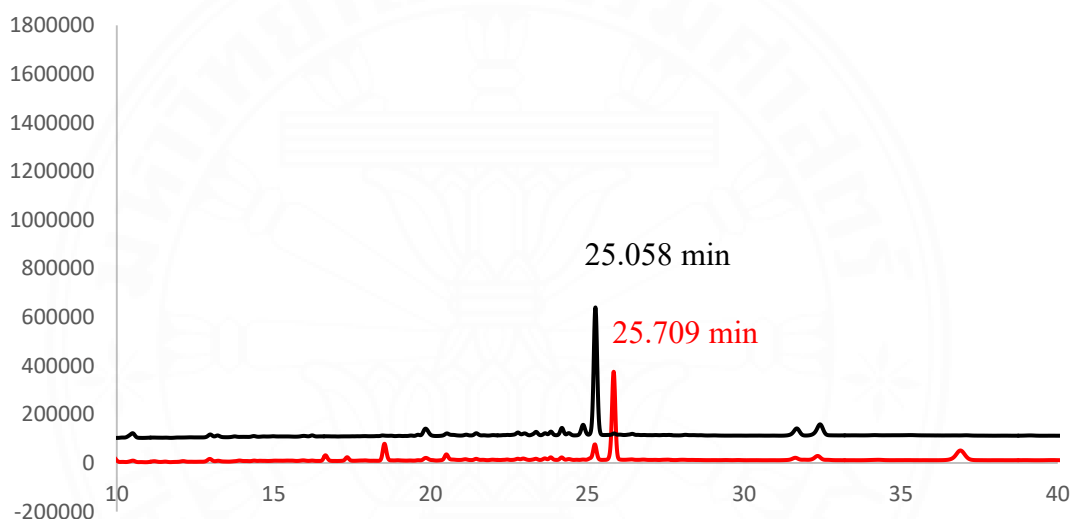
The main challenge in this research is that the double cyclization carried out in a one pot reaction. The ligation product *N,N'*-diamino-bis(Fmoc-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glycinyl) pimelic acid (**12**) was firstly deprotected by treating with 20% piperidine in DMF to afford *N,N'*-diamino-bis(NH<sub>2</sub>-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glycinyl) pimelic acid with free amino groups at *N*-terminus, which was subsequently intramolecularly cyclized at the carboxyl groups of 2,6 diaminopimelic scaffold by using high efficient coupling reagent HATU and DIPEA as a base in DMF as show in **Scheme 4.6** to afford *N,N'*-diamino-bis(cyclic-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glycinyl) pimelic acid (**13**).

To prevent an intermolecular cyclization, the reaction was performed under a very high dilution condition (approximately 1 mg/1 mL). To this reaction, the excess HATU (5 eq.) was dissolved in DMF, and slowly added to the reaction mixture of *N,N'*-diamino-bis(NH<sub>2</sub>-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glycinyl) pimelic acid and DIPEA (5 eq.) . This reaction mixture was allowed for three days to complete the double cyclization.



**Scheme 4.6.** The synthesis of *N,N'*-bis(cyclic-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny) pimelic acid

To confirm the double cyclization, the resulting reaction was examined by analytical HPLC. According to HPLC results, the comparison between Fmoc deprotected ligation (black line) and cyclization crude reaction (red line) was performed by detecting at wavelength 215 nm (for the peptide bond detection). It was found that the retention time of the linear peptide precursor is shown at 25.058 min, while the resulting double cyclization product is emerged at 25.709 min. According to the HPLC results, it showed that the cyclization reaction crude was relatively clean as shown in **Figure 4.29**.

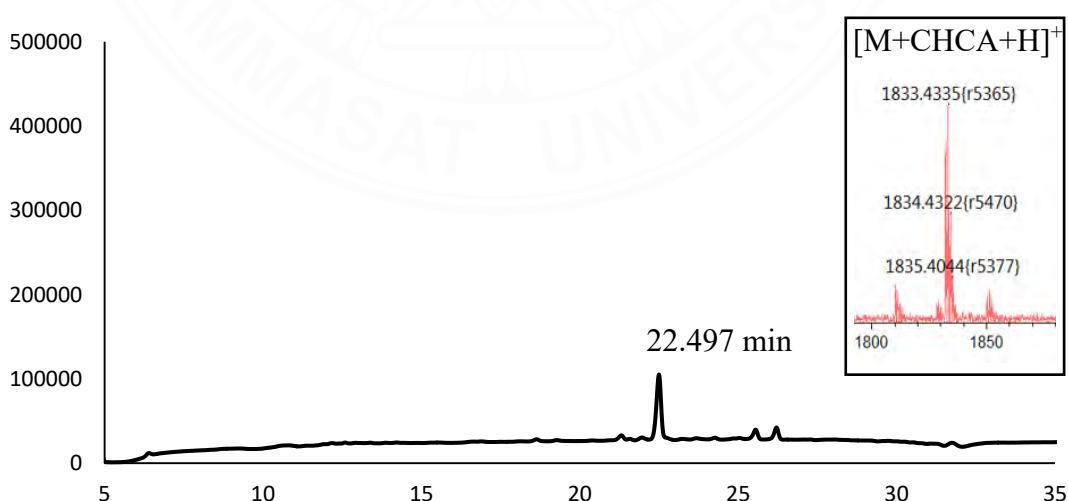


**Figure 4.29** The HPLC chromatogram comparison between Fmoc deprotected ligation (black line) and cyclization reaction crude (red line) measured at 215 nm.

### 4.3.3 The global deprotection

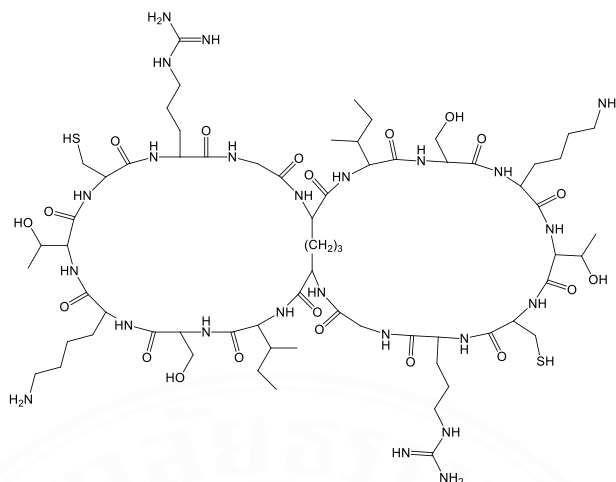
To receive the fully deprotected bicyclic peptide, the resulting bicyclic peptide was treated by the deprotection cocktail, comprising of 94% trifluoroacetic acid, 2.5% 1,2-Ethanedithiol (EDT), 2.5% H<sub>2</sub>O and 1% triisopropylsilane (TIS) to remove all protection groups within the side chain of bismacromolecular peptide followed by the purification using semi-preparative HPLC to yield *N,N'*-bis(cyclic-isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glycinyl) pimelic acid (**14**) with the overall yield of 1.6. The acid-labile protecting groups consisted of Pbf, Trityl, <sup>t</sup>Bu and Boc were completely deprotected under this condition with TFA (94%). However, the scavengers were required to prevent the re-attaching of protecting groups to the peptide chain. Water is a scavenger of *tert*-butyl groups. TIS is a scavenger of trityl and Pbf cations. Lastly, EDT is a scavenger of *t*-butyl cation and also suppresses the oxidation of a cysteine residue.

Finally, the bismolecular SFTI-1 peptide was characterized by MALDI-TOF spectrometry (**Figure 4.30**). The molecular ion peak was observed at 1833.4533 corresponding to  $[M+CHCA+H]^+$  in which the molecular mass was calculated to be  $m/z = 1644.86$  (C<sub>67</sub>H<sub>120</sub>N<sub>24</sub>O<sub>20</sub>S<sub>2</sub>). According to the analytical HPLC result, it was observed the peak at 22.497 min. as shown in **Figure 4.30**



**Figure 4.30.** The HPLC chromatogram and MALDI-TOF mass spectrum of bismacrocylic SFTI-1 (inset)

Ref. code: 25636009040087JCO



**Figure 4.31** The structure of *N,N'*-bis(cyclic-isoleucynyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid

Name: *N,N'*-bis(cyclic-isoleucynyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid (**14**)

Chemical formula:  $C_{67}H_{120}N_{24}O_{20}S_2$

MALDI-TOF MS: Found 1833.4533  $[M+CHCA+H]^+$  ; Calculated 1644.86  $[M]^+$

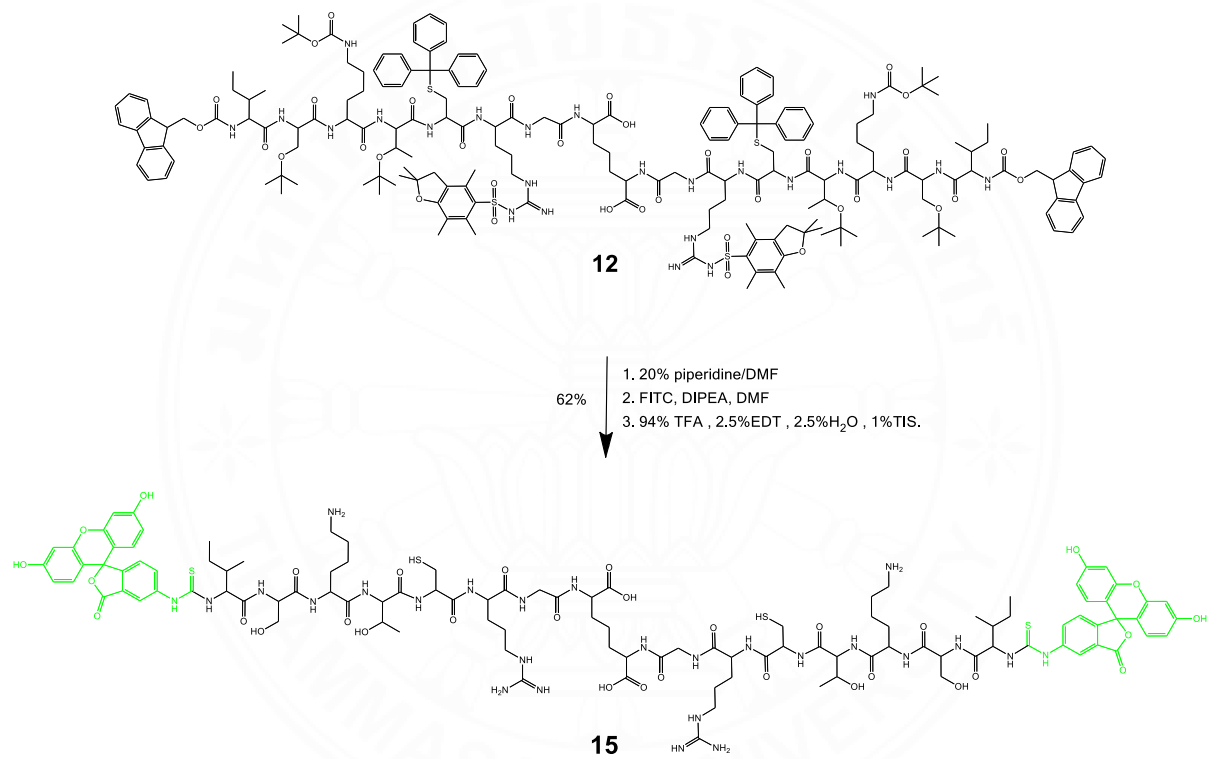
HPLC (method 3):  $t_R = 22.780$  min

To this strategy, we successfully synthesized the macromolecular peptide with the bicyclic scaffold *via* various steps, which resulted in a very low overall yield (approximately 1.6%).

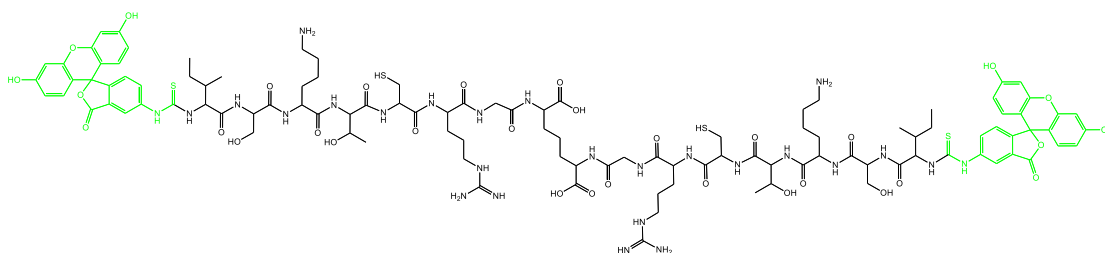
#### 4.4 The synthesis of fluorescently labelled bivalent SFTI-1

The peptide labelling with fluorescent dye is powerful tools for investigating the biology relevant interaction such as receptor-ligand binding, protein structure, protein-protein interaction, enzyme activity and imaging application.

Typically, the peptide was labelled with a fluorescence dye at lysine residue to monitor the cellular uptake. However, in the case of SFTI-1, Lys is an important residue which is affected the biological activity, and it is located in substrate specific P<sub>1</sub> position. With this reason, the labelling at Lys could affect its biological activity. Therefore, we designed to attach a fluorescent dye at *N*-terminus of isoleucine residues. To synthesize the fluorescently labelled bivalent SFTI-1, the resulting product from the ligation step *N,N'*-bis(Fmoc-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glycinyl) pimelic acid (**12**) was firstly treated with 20% piperidine in DMF to deprotect an Fmoc moiety at *N*-terminus to free the amino functionality, while all protecting groups in the side chain were kept. The fluorescent dye (FITC) was utilized for the peptide labeling (2 eq.) and DIPEA (2.5 eq.) were used as a base dissolved in DMF. Subsequently, all protecting groups were deprotected by using the deprotection cocktail as previously described to yield the fluorescently labelled bivalent SFTI-1 in moderate yield (62%) as shown in **Scheme 4.8**.



**Scheme 4.8** The synthesis of fluorescently labelled *N,N'*-diamino-bis(isoleucinyll-serinyll-lysinyll-threonyll-cysteinyl-arginyll-glycinyll) pimelic acid



**Figure 4.32** The structure of fluorescently labelled *N,N'*-diamino-bis(isoleuciny-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid

Name: fluorescently labelled *N,N'*-diamino-bis(isoleuciny-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid (**15**)

Chemical formula:  $C_{109}H_{146}N_{26}O_{32}S_4$

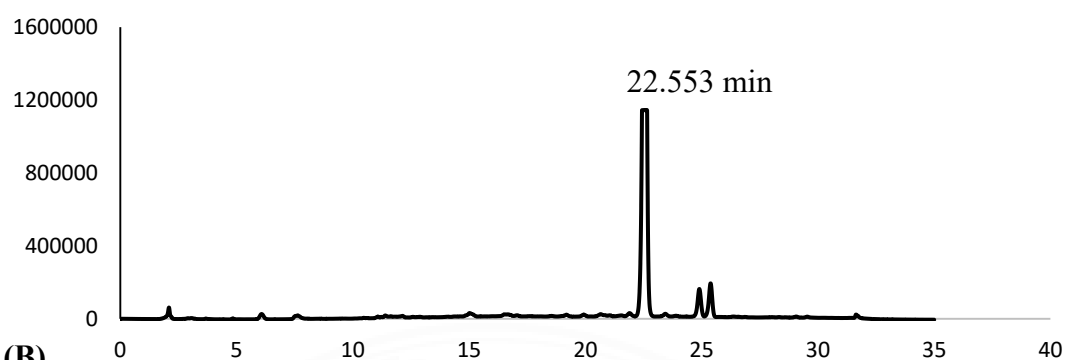
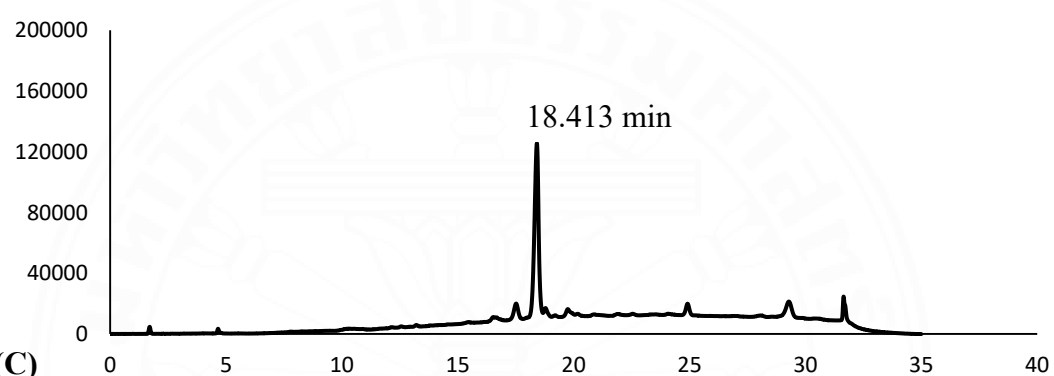
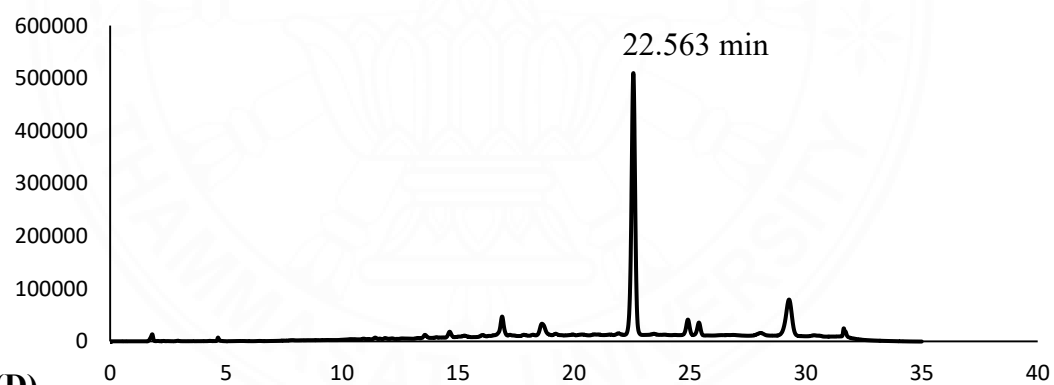
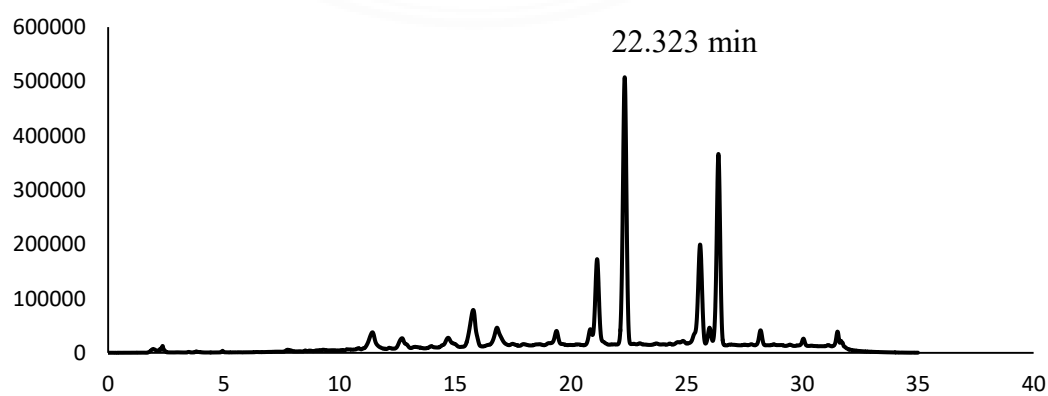
MALDI-TOF MS: Found 630.158  $[M+4H]^{4+}$ , 861.085  $[M+3Na]^{3+}$ .

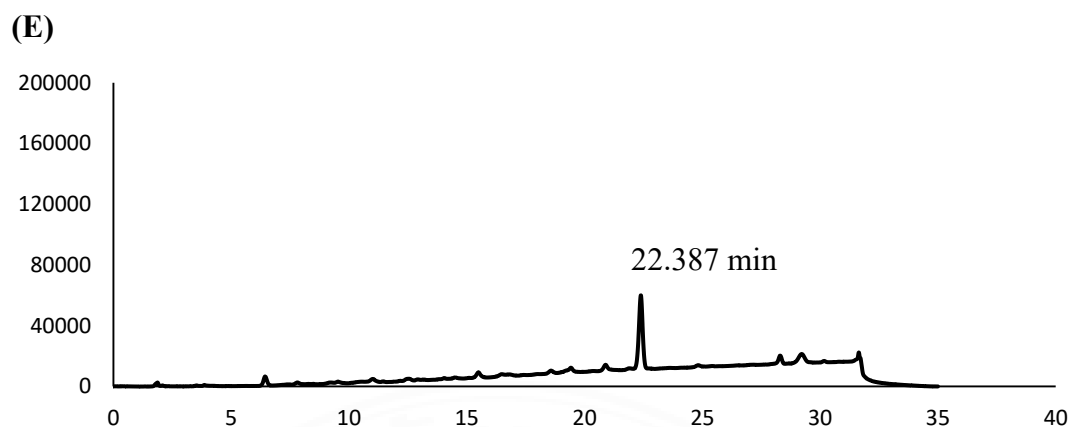
Calculated 2458.95  $[M]^+$

HPLC (method 2) :  $t_R = 22.323$  minutes

The labelling peptide reaction is traced by analytical HPLC measured at 215 nm. According to the results, the peak of an Fmoc deprotected peptide at 22.553 min. (**Figure 4.33A**) slightly shifted to 22.563 min. after the labelling reaction (**Figure 4.33C**). Additionally, the peak of standard FITC at 18.413 min. (**Figure 4.33B**) was not observed in a reaction mixture. The peak shifting and the fluorescence dye concentrations decreasing clearly confirmed that the peptide labelled reaction was successful.

Then, the reaction mixture was treated with deprotection cocktail (94:2.5:2.5:1 TFA:EDT:H<sub>2</sub>O:TIS) to remove all protection groups along the side chain without purification. The HPLC chromatogram of fluorescently labelled bismolecular SFTI-1 after the global deprotection is shown in **Figure 4.33D**. After purification, a peak of pure fluorescently labelled bivalent SFTI-1 represented at 22.387 min (**Figure 4.33E**).

**(A)****(B)****(C)****(D)**



**Figure 4.33** The HPLC chromatogram of fluorescently labelled peptide reaction and precursors measure at 215 nm.

**(A)** Fmoc deprotection reaction crude of ligation product

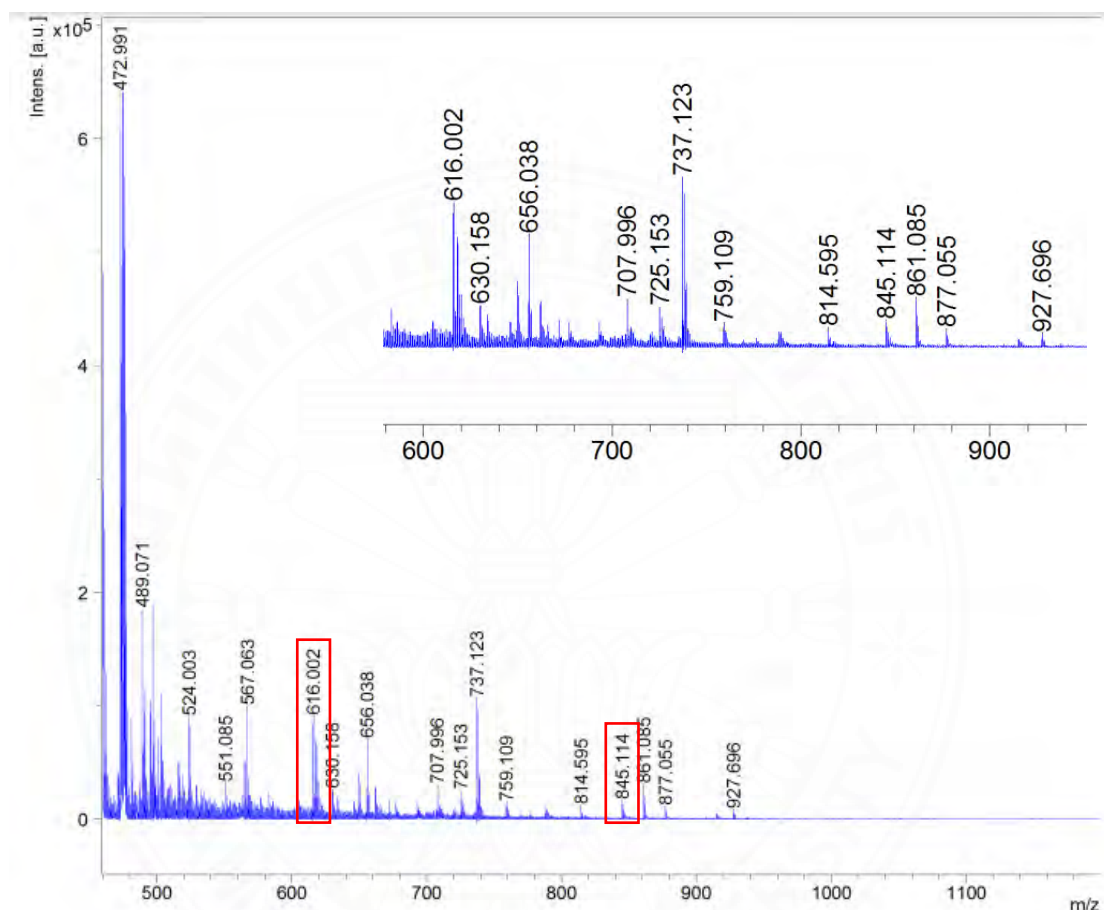
**(B)** Standard FITC

**(C)** Fluorescently labeled bivalent SFTI-1 reaction crude

**(D)** Global deprotection of fluorescently labelled bivalent SFTI-1 reaction crude

**(E)** Fluorescently labelled bivalent SFTI-1

The fluorescently labelled bivalent SFTI-1 was characterized by MALDI-TOF (**Figure 4.34**). It was found that the molecular mass adducts were observed at 616.002 and 845.114 corresponding to  $[M+4H]^{4+}$  and  $[M+3Na]^{3+}$ , respectively. The calculated molecular mass was found to be  $m/z = 2458.95$  ( $C_{109}H_{146}N_{26}O_{32}S_4$ ).



**Figure 4.34** MALDI-TOF mass spectrum of fluorescently labelled *N,N'*-diamino-bis(isoleuciny-l-serinyl-lysiny-l-threony-l-cysteinyl-arginyl-glyciny-l) pimelic acid (**15**). Found 616.002  $[M+4H]^{4+}$ , 845.114  $[M+3Na]^{3+}$ ; Calculated 2459.95  $[M]^+$

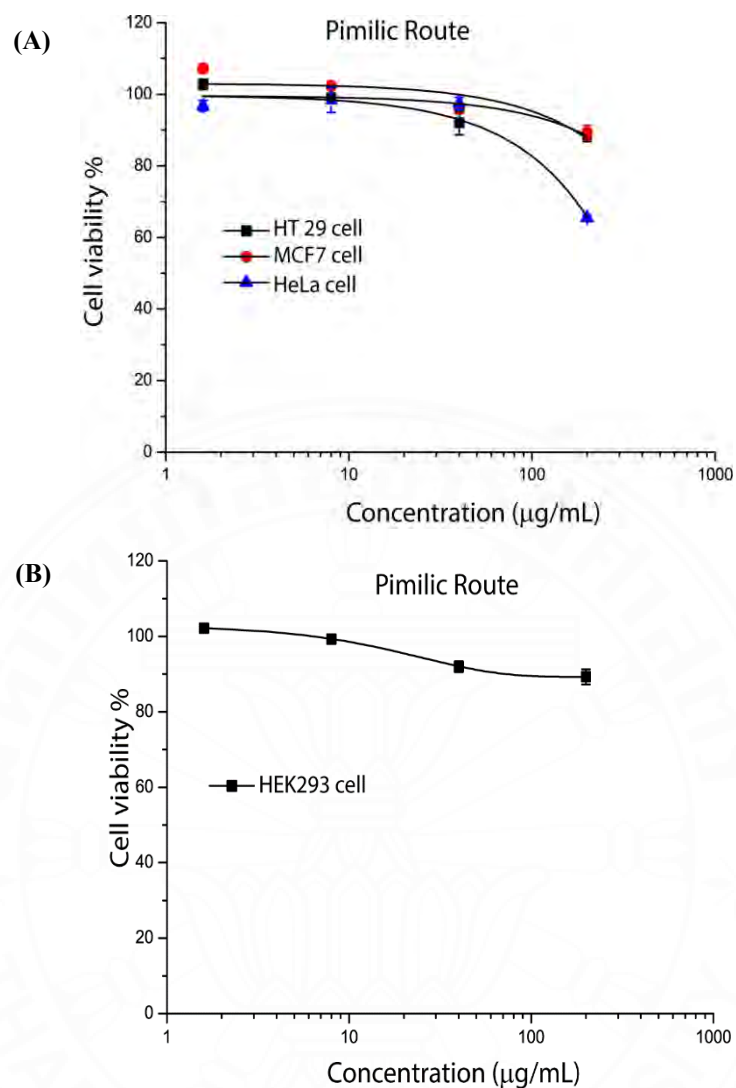
## 4.5 The biological activity and application

### 4.5.1 The anticancer activity

Although a synthetic bivalent SFTI-1 was designed based on naturally occurring SFTI-1 by increasing an inhibitory activity for anti-human  $\beta$ -trypsin, trypsin-like serine protease. However, a derived SFTI-1 also showed the bioactivity against other serine proteases. Matriptase is transmembrane serine protease, and is known to play an important role relating to cancer invasion and metastasis. Therefore the reduction of matriptase levels leads to the suppression of prostate, ovarian primary growth and metastasis<sup>[11]</sup>.

The inhibitory activity of derived bivalent SFTI-I inhibitor was assessed against three cancer cell lines (HT29, MCF7 and HeLa) by routine screening. Derived bivalent SFTI-I inhibitor displayed a promising anti-cancer activity in microgram/mL region against HT29, HeLa cells and MCF7 with the inhibitory activity = 53, 76 and >150  $\mu$ M (**Figure 4.35A**). Importantly, this bivalent SFTI-I was examined against human embryonic kidney 293 (**Figure 4.35B**), and showed no toxicity against normal cells (HEK293).

Nowadays, the peptide utilization has gained of great attention in therapeutic applications according to great potencies and high selectivity. Unfortunately, the utilization of peptides as pharmaceuticals was limited by their stabilities in the digestive system, which was resolved by using peptide cyclization. To our study, the derived bivalent SFTI-1 was designed to solve this problem by increasing the number of inhibitory activity and the cyclic scaffold which is believed to increase the stability against an enzymatic degradation. Therefore, this bioactive bismolecular SFTI-1 will be useful for drug design and therapy in the future.



**Figure 4.35** The MTT assays of bivalent SFTI-1 incubated with cancer cell lines and human normal cells measured at 550 nm.

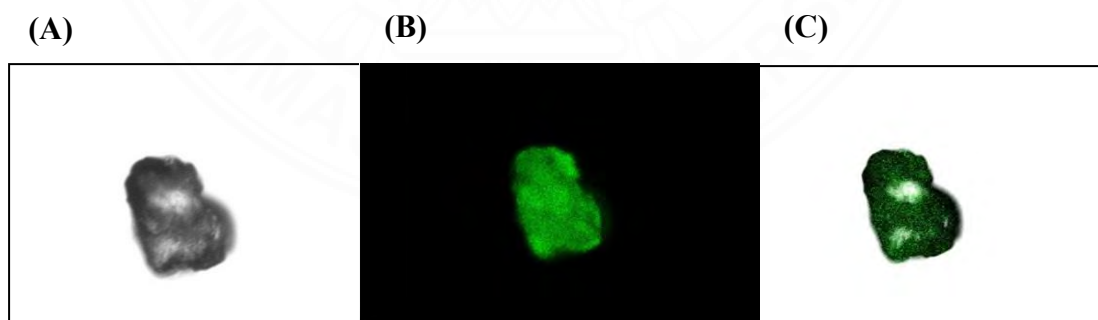
**(A)** Cell viability assays against HT29, HeLa cells and MCF7 cancer cell lines

**(B)** Cell viability assays against HEK293

#### 4.5.2 The bioimaging application

The SFTI-1, a plant-derived cyclic peptide, is one of cell-penetrating peptides that can translocate across the plasma membrane. Generally, characteristic of favorable cell-penetrating peptides is consisted of various disulfide bonds, positively charged amino acids and non-polar amino acids. Although the disulfide bridge of bivalent SFTI-1 was replaced with three carbon which is the part of 2,6-diaminopimelic acid. This bismolecular SFTI-1 structure is consisted of amino acids that had good intracellular properties. The arginine and lysine are positively charged amino acids. Also, isoleucine and glycine are non-polar amino acids. However, the mechanism of SFTI-1 entering the living cells remains unclear.

To confirm the penetration of synthesized bivalent SFTI-1 in living cells, the fluorescently labelled *N,N'*-diamino-bis(isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid was synthesized, and applied for the bioimaging experiment using HeLa cell lines. Briefly, HeLa cell was incubated by *N,N'*-diamino-bis(isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid (125  $\mu$ M). Interestingly, we found that fluorescently labelled peptide has clearly shown the ability for the cell internalization into the cancer cell lines (HeLa cell lines) as shown in **Figure 4.36**.



**Figure 4.36** The live-cell confocal microscopy image of HeLa cells after incubated with fluorescently labelled bivalent SFTI-I (125  $\mu$ M): BF-Mode (A), Fluorescent Mode (B), and Merge (C).

Due to the cellular bioavailability and the diverse functional groups (thiol, hydroxyl and amino groups) of bivalent SFTI-1, it obviously showed the great potential as the drug delivery candidates by conjugating with cargos (i.e. oligonucleotide, peptide sequence, or polysaccharide) for delivering through the cell membrane.



## CHAPTER 5

### CONCLUSIONS

A novel bicyclic SFTI-1 inhibitor was designed and successfully synthesized by using combination of 2,6-diaminopimelic acid as an important core grafted by two inhibitory loops mimicked native SFTI-1 epitopes. The synthetic pathway of bivalent SFTI-1 was divided into two major parts. Firstly, the synthesis of peptidyl pimelic acid, *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid (**5**), *via* solution phase peptide synthesis. Fully protected and Fmoc deprotected peptidyl pimelic acid of each step were identified by using <sup>1</sup>H-NMR, ESI-MS, HR-ESI-MS and HPLC to confirm the dimerization of 2,6 diaminopimelic acid scaffold and tripeptidyl moieties (glycine, arginine and cysteine, respectively). Secondly, an amino acid assembly of linear peptide chain (Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH (**11**)), by using Fmoc-based solid-phase peptide synthesis (SPPS). Both peptide segments were subsequently ligated to afford *N,N'*-diamino-bis (Fmoc-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid under our developed condition. Additionally, we performed the ligation processed with different length of peptide precursor from SPPS and it clearly indicated that only short peptide chain with 4 residues, Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH (**11**), was perfectly ligated to form an amide bond between cysteine and threonine with acceptable yield (43%). Finally, the challenge cyclization *via* one pot 'double' macrolactomization by using high efficient coupling reagent HATU under high dilution condition (1 mg/1mL) to yield *N,N'*-diamino-bis(cyclic-isoleucinyl-serinyl(<sup>t</sup>Bu)-lysiny(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid (**13**), followed by global deprotection to afford *N,N'*- bis(cyclic-isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny)l pimelic acid (**14**) with overall yield 1.6%.

**Table 5.1** The suitable synthesis condition of bivalent SFTI-1 for each steps

Reaction	Coupling reagents	solvents	Yield (%)
Fmoc-G-pimelic acid	DCC/DMAP	DMF/H <sub>2</sub> O (3:1)	84
Fmoc-RG-pimelic acid	HBTU/HOBt	DMF	56
Fmoc-CRG-pimelic acid	HBTU/HOBt	DMF	40
Ligation	HBTU/HOBt	DMF	43
Cyclization	HATU	DMF	N/A

N/A = Non-assessable

The synthetic bivalent SFTI-1 was evaluated anticancer activities against a panel of cancer cell lines, including HT29, MCF7 and HeLa corresponding to the colon, breast and cervical tumors, respectively. The bivalent SFTI-1 showed promising inhibitory activity against HT29 ( $IC_{50} = 53 \mu M$ ) and HeLa ( $IC_{50} = 76 \mu M$ ) and showed no toxicity against normal cells, confirmed by examination with human embryonic kidney 293 (HEK293).

Importantly, the linear scaffold of bivalent SFTI-1 (**12**) was successfully fluorescently labelled with fluorescent dye (FITC) to yield fluorescently labelled *N, N'*-diamino-bis(isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid (**15**), which was applied for bio-imaging experiment to confirm the penetration in living cells. According to our study, we found that fluorescently labelled peptide has clearly shown the ability for the cell internalization into the cancer cell lines (HeLa cell lines). The fluorescently labelled bivalent SFTI-1 also demonstrated a potential platform for cellular bio-imaging applications, particularly the detection of cancer cell lines.

This bivalent SFTI-1 will be further investigated for biological activities against other serine proteases (such as human  $\beta$ -tryptase). Furthermore, the P1 position of bivalent inhibitor could be slightly modified by replacing the lysine residue with arginine, valine and phenylalanine, which have ability to bind at S1 pocket of chymotrypsin-like and elastase-like serine proteases, respectively.



## REFERENCES

- [1] Hamley, I W. *introduction to Peptide Science*. Wiley (2020)
- [2] Jayne Leonard. "What to know about peptides for health" (2019)  
Available from: <https://www.medicalnewstoday.com/articles/326701>
- [3] Fosgerau, Keld, and Torsten Hoffmann. "Peptide therapeutics: current status and future directions." *Drug discovery today* 20.1 (2015): 122-128.
- [4] Marqus, Susan, Elena Pirogova, and Terrence J. Piva. "Evaluation of the use of therapeutic peptides for cancer treatment." *Journal of biomedical science* 24.1 (2017): 1-15.
- [5] Lewis, Richard J., et al. "Does nature do ion channel drug discovery better than us?." *Ion Channel Drug Discovery* 39 (2014): 297.
- [6] Tzotzos, Susan J. "Peptide Drugs of the Decade." (2020)
- [7] Muttenthaler, Markus, et al. "Trends in peptide drug discovery." *Nature Reviews Drug Discovery* (2021): 1-17.
- [8] Bowsher, Ronald R., et al. "Sensitive RIA for the specific determination of insulin lispro." *Clinical chemistry* 45.1 (1999): 104-110.
- [9] Neitzel, J. J. (2010) Enzyme Catalysis: The Serine Proteases . *Nature Education* 3(9):21
- [10] L.B. Schwartz, R.A. Lewis, D. Seldin, K.F. Austen Acid hydrolases and tryptase from secretory granules of dispersed human lung mast cells *J. Immunol.*, 126 (1981), pp. 1290–1294
- [11] Varela, Fausto A., Thomas E. Hyland, and Karin List. "Physiological functions and role of matriptase in cancer." *Extracellular Targeting of Cell Signaling in Cancer: Strategies Directed at MET and RON Receptor Tyrosine Kinase Pathways* (2018): 91.
- [12] Di Cera, Enrico. "Serine proteases." *IUBMB life* vol. 61,5 (2009): 510-5.
- [13] Ovaere, P., et al., *The emerging roles of serine protease cascades in the epidermis*. Trends Biochem Sci, 2009. 34(9): p. 453-63.
- [14] László Polgár . Catalytic Mechanisms of Serine and Threonine Peptidases *Handbook of Proteolytic Enzymes*, 2013, Pages 2524–2534

- [15] Schowen, r. L., Biochemistry by Reginald H. Garrett and Charles M. Grisham. J. Chem. Educ. 1997, 74 (2), 189-190.
- [16] Wieczorek, Rafal, et al. "Small and random peptides: An unexplored reservoir of potentially functional primitive organocatalysts. The case of seryl-histidine." *Life* 7.2 (2017): 19.
- [17] Pereira PJ<sup>1</sup>, Bergner A, Macedo-Ribeiro S, Huber R, Matschiner G, Fritz H, Sommerhoff CP, Bode W. "Human Beta-tryptase is a ring-like tetramer with active sites facing a central pore." *Nature* 392 (1998): 306-311
- [18] Sommerhoff, Christian P., et al. "The structure of the human  $\beta$ II-tryptase tetramer: Fo(u)r better or worse." *Proceedings of the National Academy of Sciences* 96.20 (1999): 10984-10991.
- [19] Friedrich, R., Bode, W. Crystal structure of MTSP1 (matriptase). Protein data bank
- [20] Mohan, Chandra & Long, Kevin & Mutneja, Manpreet.. Introduction to Inhibitors. (2014)
- [21] Gitlin-Domagalska, Agata, Aleksandra Maciejewska, and Dawid Dębowski. "Bowman-Birk Inhibitors: Insights into Family of Multifunctional Proteins and Peptides with Potential Therapeutical Applications." *Pharmaceuticals* 13.12 (2020): 421.
- [22] Hellinger, Roland, and Christian W. Gruber. "Peptide-based protease inhibitors from plants." *Drug discovery today* 24.9 (2019): 1877-1889.
- [23] QI, Rui-Feng, Zhan-Wu SONG, and Cheng-Wu CHI. "Structural features and molecular evolution of Bowman-Birk protease inhibitors and their potential application." *Acta biochimica et biophysica Sinica* 37.5 (2005): 283-292. 24.
- [24] Korsinczky, Michael LJ, et al. "Solution structures by 1H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant." *Journal of molecular biology* 311.3 (2001): 579-591.
- [25] Nelson DL, Cox MM. *Principles of Biochemistry* (4th ed.). New York: W. H. Freeman (2005)
- [26] Wagner, Ingrid, and Hans Musso. "New naturally occurring amino acids." *Angewandte Chemie International Edition in English* 22.11 (1983): 816-828.

- [27] Theodora W. Greene, Peter G. M. Wuts. *Protecting Groups in Organic Synthesis* (3 ed.). J. Wiley (1999)
- [28] El-Faham, Ayman, and Fernando Albericio. "Peptide coupling reagents, more than a letter soup." *Chemical reviews* 111.11 (2011): 6557-6602.
- [29] Albericio, Fernando, and Ayman El-Faham. "Choosing the right coupling reagent for peptides: a twenty-five-year journey." *Organic Process Research & Development* 22.7 (2018): 760-772.
- [30] aapptec Complete Peptide Product Source. "Coupling Reagents"  
Available from: <https://www.peptide.com/resources/solid-phase-peptide-synthesis/coupling-reagents/>
- [31] John D. Robert and Marjorie C. Caserio. *Basic Principles of Organic Chemistry, second edition*. W. A. Benjamin, Inc. (1977)
- [32] Saharsh Davuluri. "Peptide Synthesis: Solution Phase, Solid Phase or Hybrid?" (2017) Available from: <https://www.linkedin.com/pulse/peptide-synthesis-solution-phase-solid-hybrid-saharsh-davuluri>
- [33] Fields, Gregg B. "Introduction to peptide synthesis." *Current protocols in protein science* vol. Chapter 18 (2002)
- [34] Kates S. A., Albericio F. (ed): *Solid-Phase Synthesis, A practical guide*, Marcel Dekker Inc. (2000)
- [35] TISHLER, MAX. "Molecular modification in modern drug research." 1964.1-14
- [36] Sang, Yankui, et al. "Novel bivalent inhibitors with sub-nanomolar affinities towards human glyoxalase I." *Bioorganic & medicinal chemistry letters* 25.21 (2015): 4724-4727.
- [37] Bae, Inhwan, et al. "Design, synthesis and biological evaluation of new bivalent quinazoline analogues as IAP antagonists." *Bioorganic & Medicinal Chemistry Letters* 34 (2021): 127676.36. Sang, Yankui, et al. "Novel bivalent inhibitors with sub-nanomolar affinities towards human glyoxalase I." *Bioorganic & medicinal chemistry letters* 25.21 (2015): 4724-4727.
- [38] Liang, Guyan, et al. "Dimerization of  $\beta$ -tryptase inhibitors, does it work for both basic and neutral P1 groups?." *Bioorganic & medicinal chemistry letters* 22.9 (2012): 3370-3376.

- [39] Rhodes, Curran A., and Dehua Pei. "Bicyclic peptides as next-generation therapeutics." *Chemistry (Weinheim an der Bergstrasse, Germany)* 23.52 (2017): 12690.
- [40] Baeriswyl, Vanessa, et al. "A synthetic factor XIIa inhibitor blocks selectively intrinsic coagulation initiation." *ACS chemical biology* 10.8 (2015): 1861-1870.
- [41] Lian, Wenlong, et al. "Cell-permeable bicyclic peptide inhibitors against intracellular proteins." *Journal of the American Chemical Society* 136.28 (2014): 9830-9833.
- [42] de Veer, Simon J., Andrew M. White, and David J. Craik. "Sunflower trypsin Inhibitor-1 (SFTI-1): sowing seeds in the fields of chemistry and biology." *Angewandte Chemie International Edition* 60.15 (2021): 8050-8071.
- [43] Gitlin-Domagalska, Agata, Aleksandra Maciejewska, and Dawid Dębowski. "Bowman-Birk Inhibitors: Insights into Family of Multifunctional Proteins and Peptides with Potential Therapeutical Applications." *Pharmaceuticals* 13.12 (2020): 421.
- [44] Luckett, S., et al. "High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds." *Journal of molecular biology* 290.2 (1999): 525-533..
- [45] Yuan, Cai, et al. "Structure of catalytic domain of Matriptase in complex with Sunflower trypsin inhibitor-1." *BMC structural biology* 11.1 (2011): 1-10.
- [46] Gitlin, Agata, et al. "Inhibitors of Matriptase-2 Based on the Trypsin Inhibitor SFTI-1." *ChemBioChem* 16.11 (2015): 1601-1607.
- [47] de Veer, Simon J., et al. "Engineered protease inhibitors based on sunflower trypsin inhibitor-1 (SFTI-1) provide insights into the role of sequence and conformation in Laskowski mechanism inhibition." *Biochemical Journal* 469.2 (2015): 243-253.
- [48] de Veer, Simon J., et al. "Engineering potent mesotrypsin inhibitors based on the plant-derived cyclic peptide, sunflower trypsin inhibitor-1." *European journal of medicinal chemistry* 155 (2018): 695-704.
- [49] Dębowski, Dawid, et al. "Inhibition of human and yeast 20S proteasome by analogues of trypsin inhibitor SFTI-1." *PloS one* 9.2 (2014): e89465.
- [50] Chen, Xingchen, et al. "Potent, multi-target serine protease inhibition achieved by a simplified  $\beta$ -sheet motif." *Plos one* 14.1 (2019): e0210842.

- [51] Zabłotna, Ewa, et al. "Design of serine proteinase inhibitors by combinatorial chemistry using trypsin inhibitor SFTI-1 as a starting structure." *Journal of peptide science: an official publication of the European Peptide Society* 13.11 (2007): 749-755.
- [52] Łęgowska, Anna, et al. "Introduction of non-natural amino acid residues into the substrate-specific P1 position of trypsin inhibitor SFTI-1 yields potent chymotrypsin and cathepsin G inhibitors." *Bioorganic & medicinal chemistry* 17.9 (2009): 3302-3307.
- [53] Fittler, Heiko, et al. "Combinatorial tuning of peptidic drug candidates: high-affinity matriptase inhibitors through incremental structure-guided optimization." *Organic & biomolecular chemistry* 11.11 (2013): 1848-1857.
- [54] Fittler, Heiko, et al. "Potent inhibitors of human matriptase-1 based on the scaffold of sunflower trypsin inhibitor." *Journal of Peptide Science* 20.6 (2014): 415-420.
- [55] Guo, Xin, et al. "Synthesis and biological activity of seleno sunflower trypsin inhibitor analog." *Chemical biology & drug design* 68.6 (2006): 341-344.
- [56] Legowska, Anna, et al. "Analogues of trypsin inhibitor SFTI-1 with disulfide bridge substituted by various length of carbonyl bridges." *Protein and peptide letters* 17.10 (2010): 1223-1227.
- [57] Jiang, Sheng, et al. "Design and synthesis of redox stable analogues of sunflower trypsin inhibitors (SFTI-1) on solid support, potent inhibitors of matriptase." *Organic letters* 9.1 (2007): 9-12.

## **APPENDICES**

## APPENDIX A

### High performance liquid chromatography conditions

**Table A1** The conditions of high performance liquid chromatography

	Method 1	Method 2	Method 3
Column	ACE-C <sub>18</sub> 4.6 × 250 mm	ACE-C <sub>18</sub> 4.6 × 250 mm	ACE-C <sub>18</sub> 4.6 × 250 mm
Injection volume	10 µL	10 µL	10 µL
Flow rate	1.0 ml/min	1.0 ml/min	1.0 ml/min
Run time	37 min	45 min	45 min
Gradient	10-100% B in 20 min	10-100% B in 35 min	10-100% B in 20 min
Mobile phase	A : H <sub>2</sub> O B : ACN	A : H <sub>2</sub> O B : ACN	A : H <sub>2</sub> O B : ACN
Detection	UV 215 nm	UV 215 nm	UV 215 nm
Instrument	JASCO Analytical HPLC	JASCO Analytical HPLC	Shimadzu Analytical HPLC

## **APPENDIX B**

### **Sample Preparation for Analysis**

#### **- NMR sample preparation**

The sample compound (10-25 mg) was dried under vacuum for 60 minutes and dissolved with suitable deuterium solvent (for this research chloroform-d). The sample solution must completely dissolved, clear and without precipitation. Followed by transferred to NMR tube and submitted for analyses.

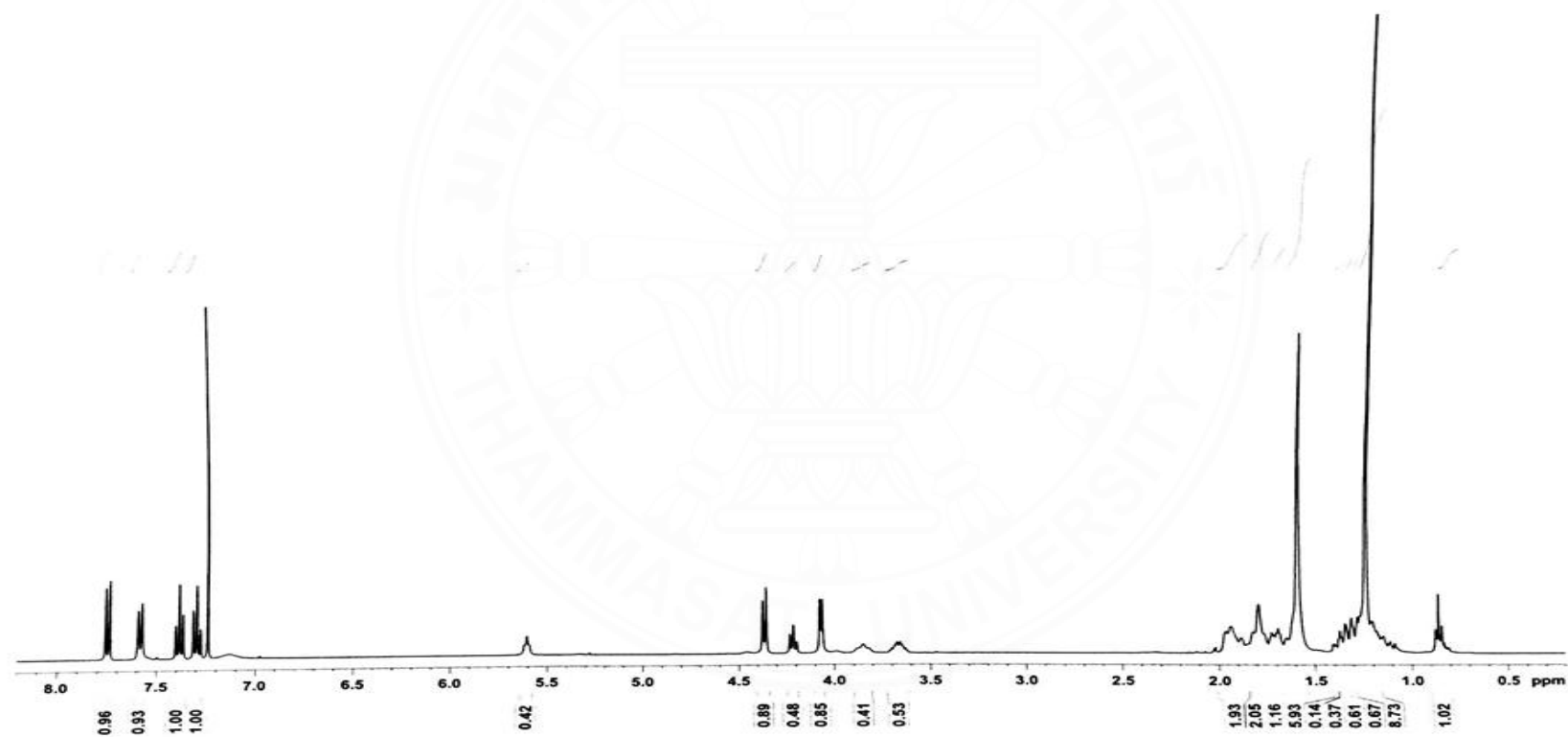
#### **- Mass spectrometry sample preparation**

The sample compound was dissolved by HPLC grade solvent such as methanol and acetonitrile. Then, the solution was filtered by using 0.22  $\mu\text{m}$  syringe filter and submitted for analyses.

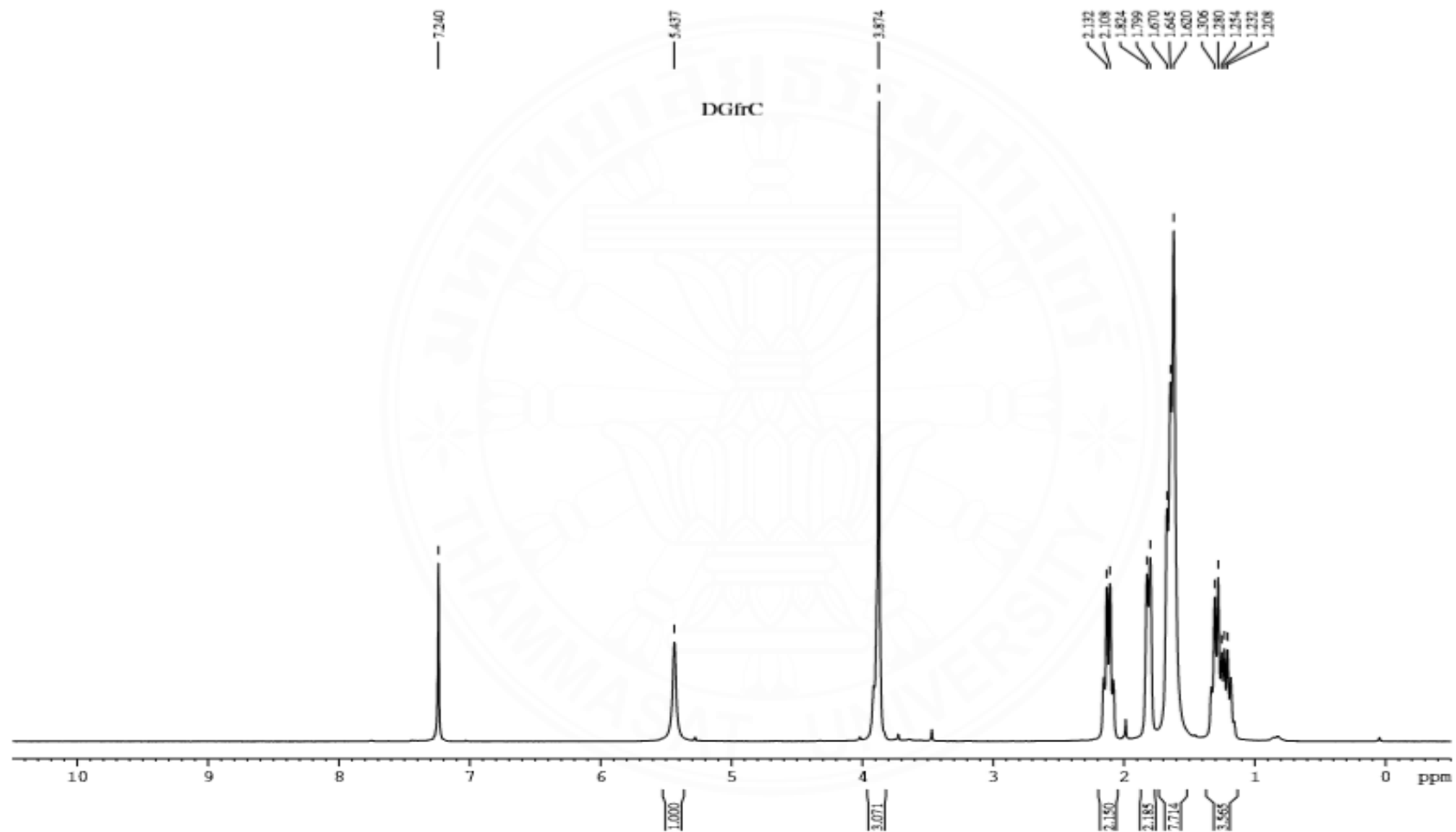
#### **- High performance liquid chromatography sample preparation and HPLC method**

The sample compound was dissolved by HPLC grade solvent such as methanol and acetonitrile. Then, the solution was filtered by using 0.22 or 0.45  $\mu\text{m}$  syringe filter and submitted for analyses. The detection wavelength is 215 and 280 nm.

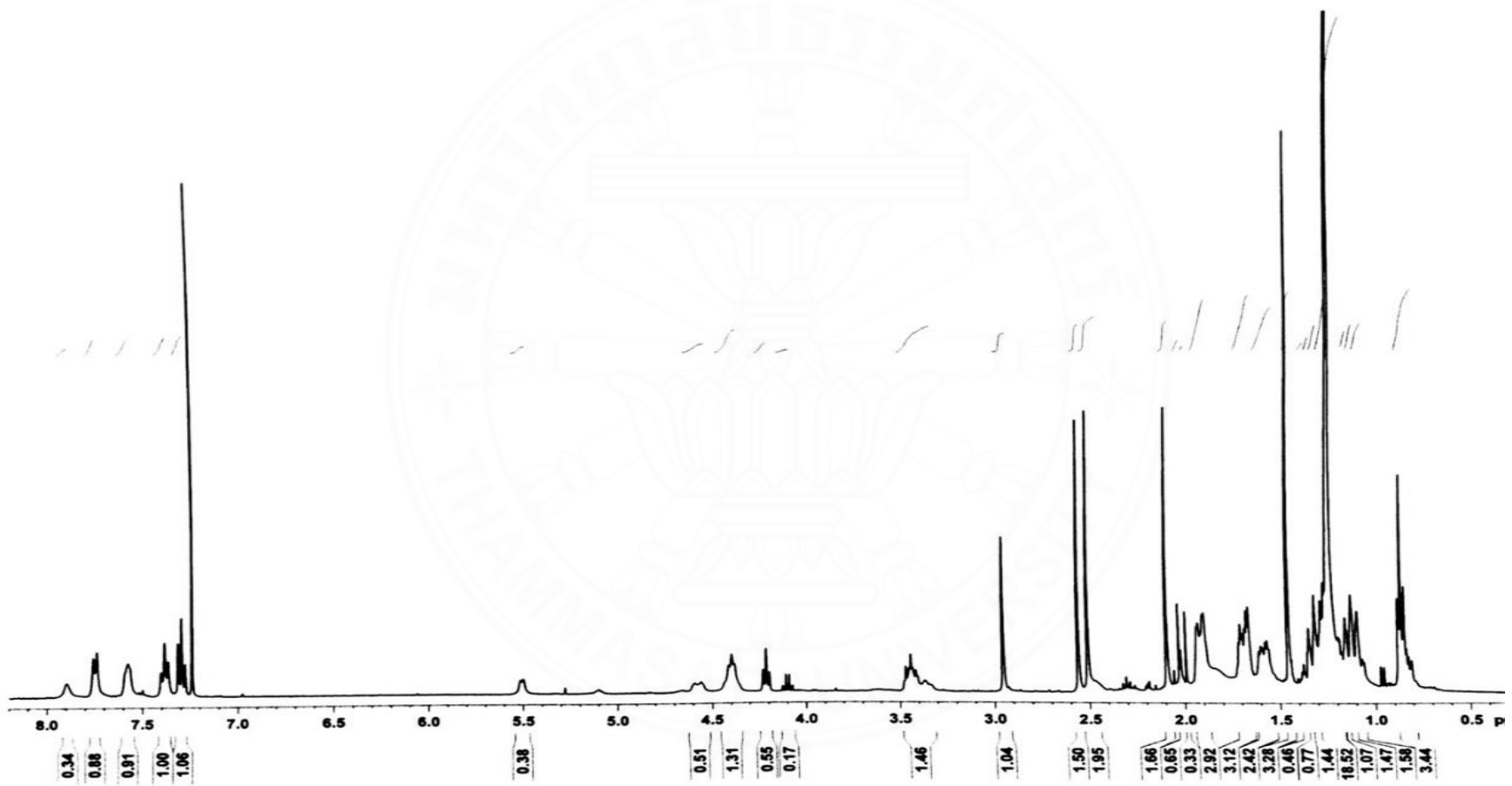
**APPENDIX C**  
**Nuclear magnetic resonance spectrum**



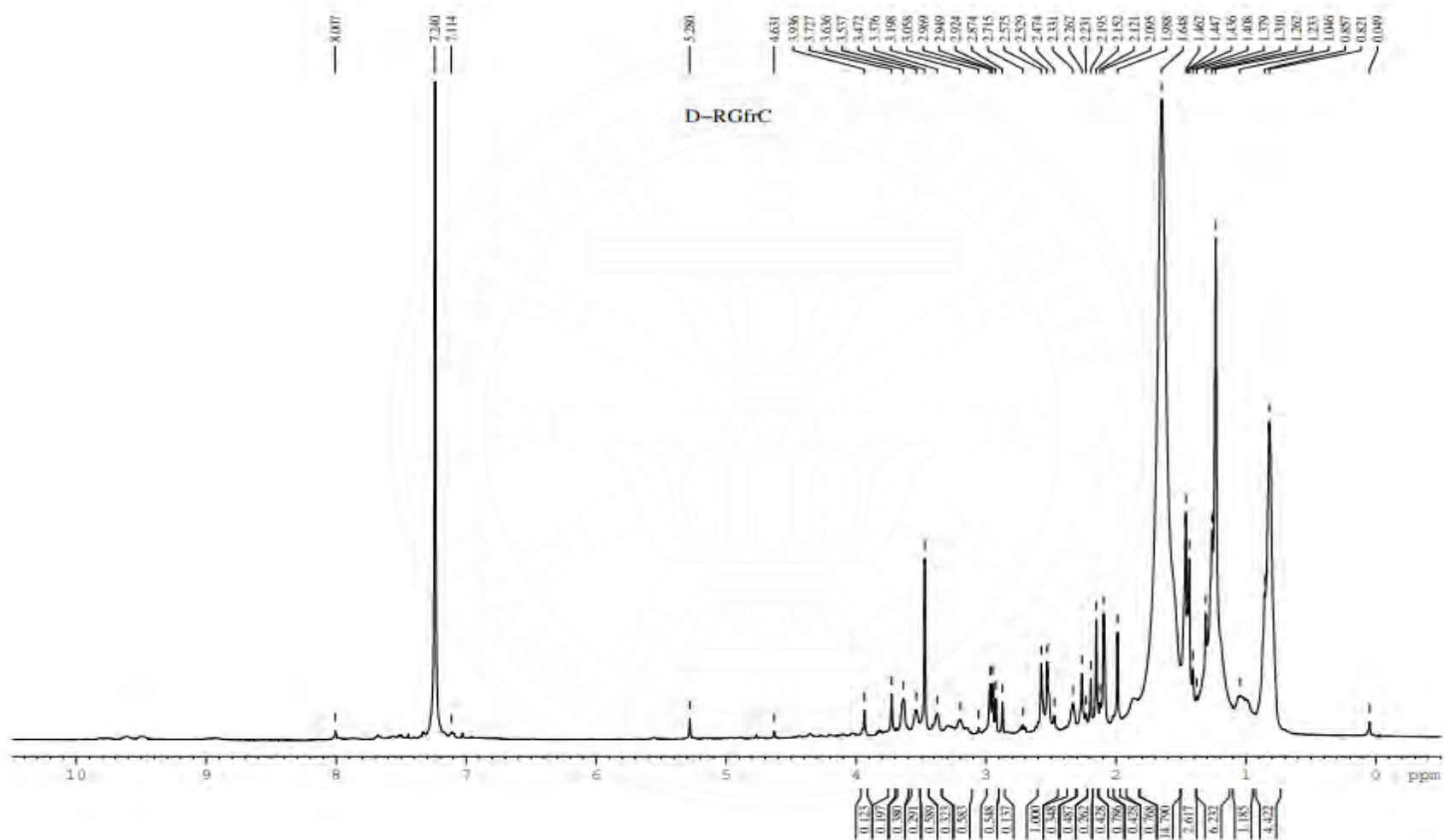
**Figure C1.** <sup>1</sup>H NMR spectrum (400 MHz, CDCl<sub>3</sub>) of *N,N'*-diamino-bis(Fmoc-glycyl)pimelic acid) (**2**)



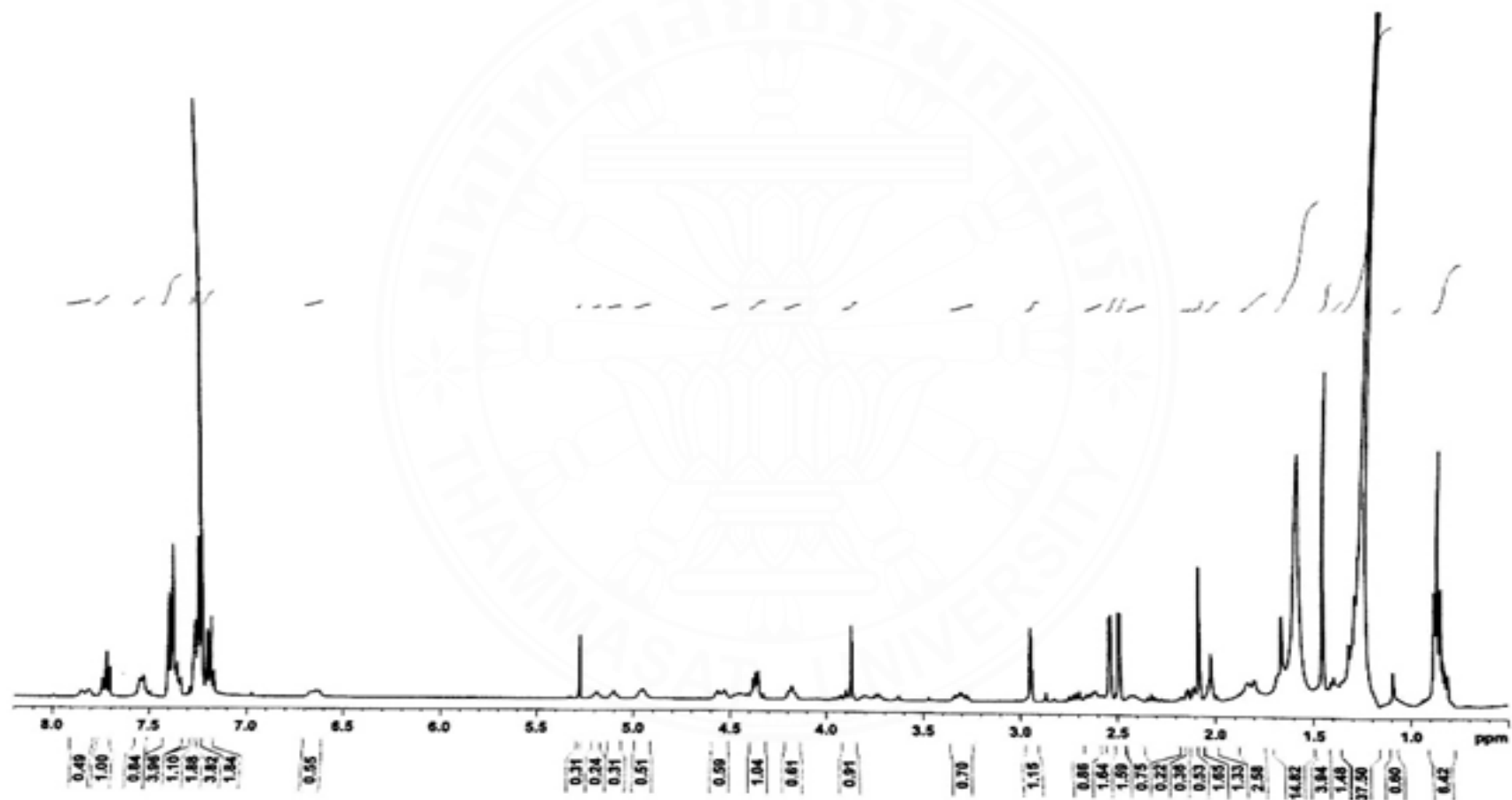
**Figure C2.**  $^1\text{H}$  NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of *N,N'*-diamino-bis( $\text{NH}_2$ -glyciny)l pimelic acid (**2a**)



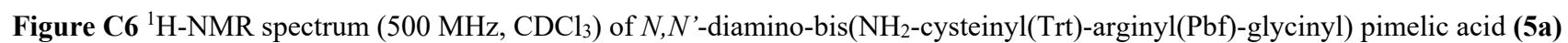
**Figure C3.**  $^1\text{H}$  NMR spectrum (400 MHz,  $\text{CDCl}_3$ ) of *N,N'*-diamino-bis(Fmoc-argenyl(Pbf)-glyciny)l) pimelic acid (4)

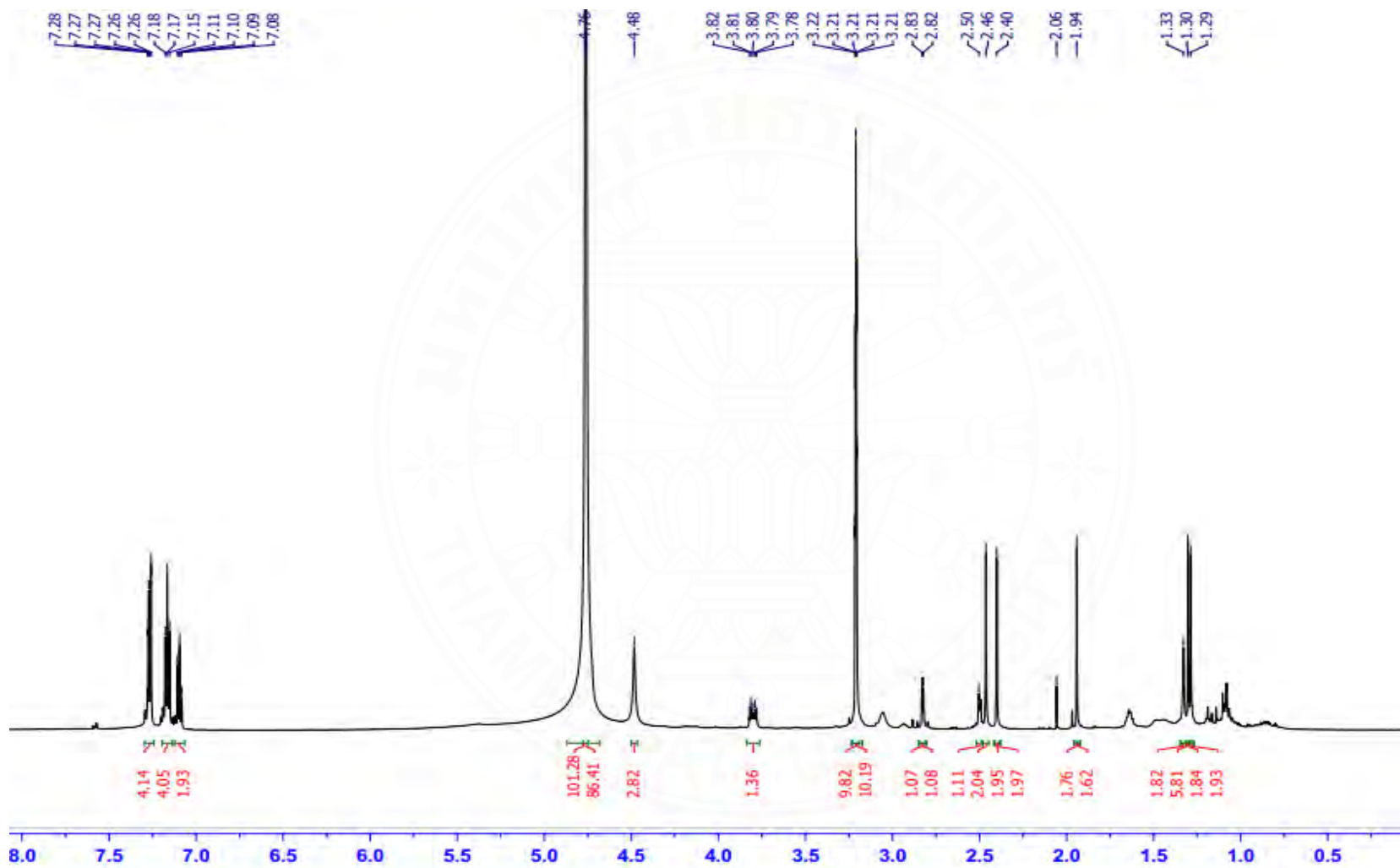


**Figure C4.**  $^1\text{H}$  NMR spectrum (500 MHz,  $\text{CDCl}_3$ ) of *N,N'*-diamino-bis( $\text{NH}_2$ -arginyl(Pbf)-glycyl)pimelic acid (**4a**)



**Figure C5.**  $^1\text{H}$  NMR spectrum (400 MHz,  $\text{CDCl}_3$ ) of *N,N'*-diamino-bis(Fmoc-cysteinyI(Trt)-argenyl(Pbf)-glycinyI)pimelic acid (5)

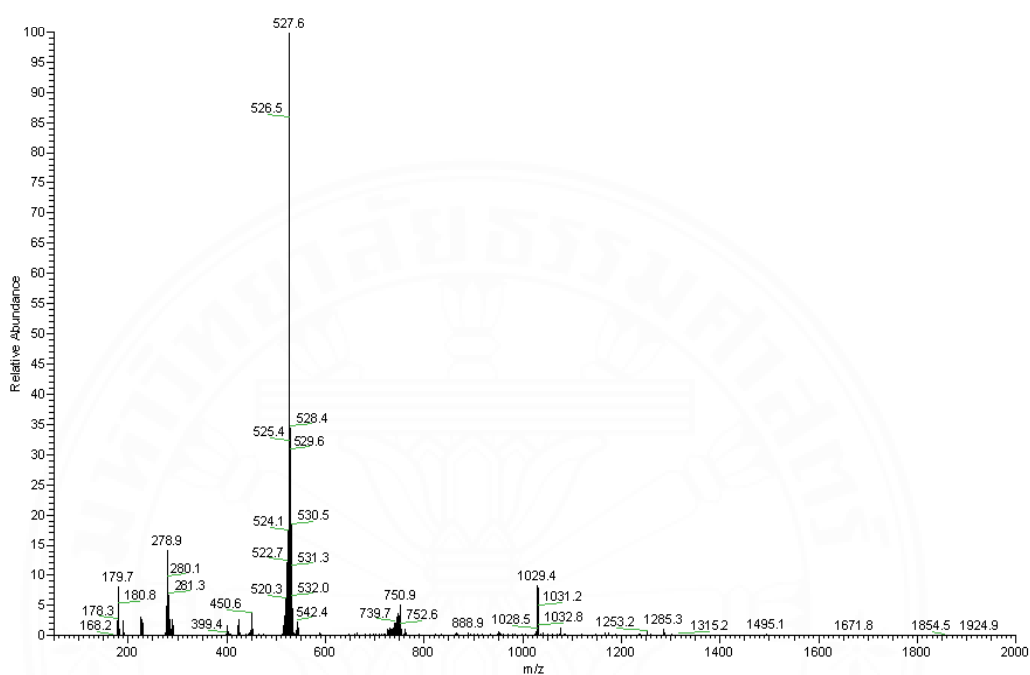




**Figure C7**  $^1\text{H}$  NMR spectrum (600 MHz,  $\text{CDCl}_3$ ) of the corresponding ligated peptide.

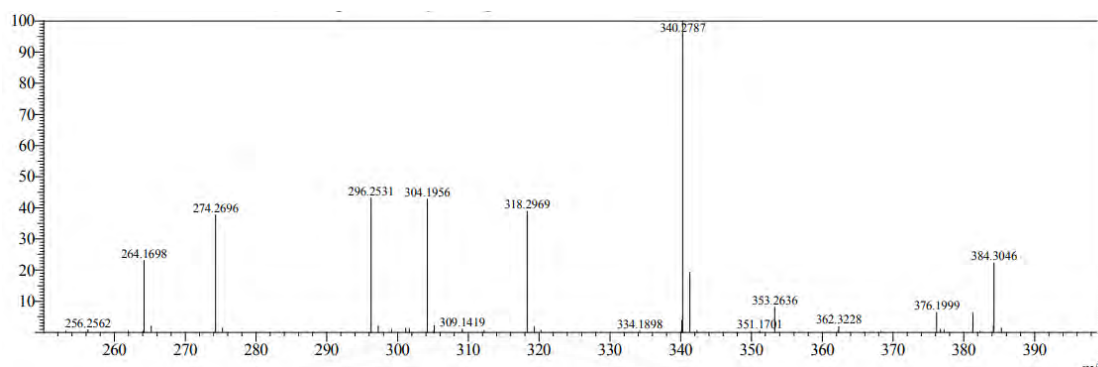
## APPENDIX D

## Mass spectrum

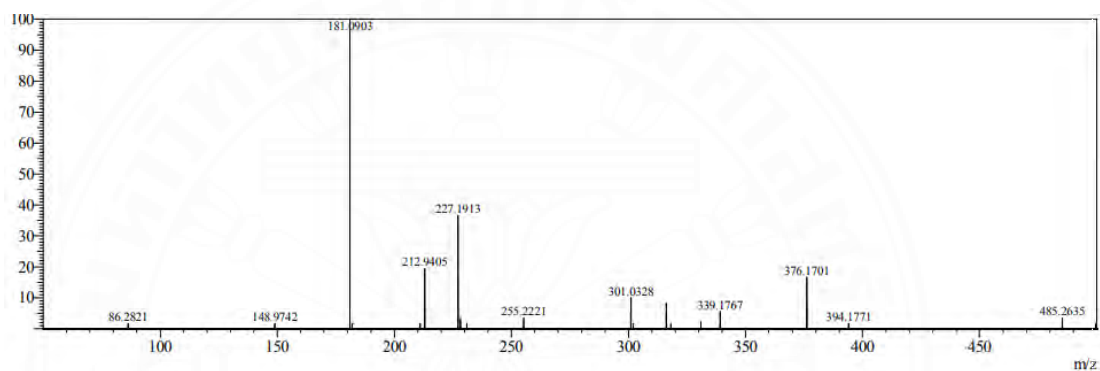


**Figure D1** Positive mode mass spectrum of *N,N'*-diamino-bis(Fmoc-glyciny)l) pimelic acid) (**2**) Found 750.9 [M+H]<sup>+</sup> ; Calculated 748.2 [M]<sup>+</sup>

(A)



(B)



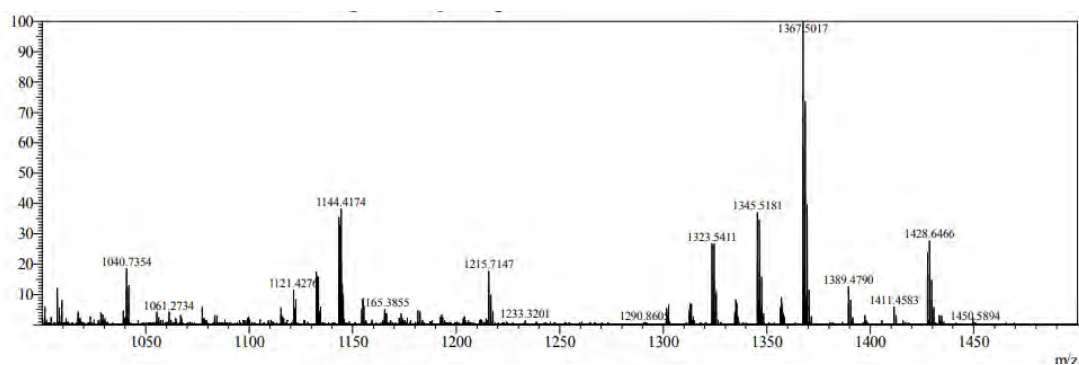
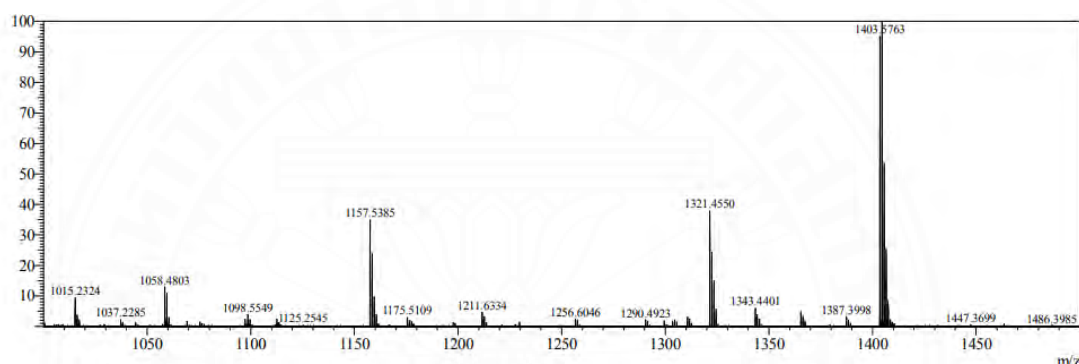
**Figure D2** High resolution mass spectrum (HR-ESI-MS) positive and negative modes of *N,N'*-diamino-bis( $\text{NH}_2$ -glyciny) pimelic acid (**2a**) ; Calculated 304.14  $[\text{M}]^+$

(A) Positive mode HR-ESI-MS. Found 304.1956  $[\text{M}+\text{H}]^+$ .

(B) Negative mode HR-ESI-MS Found 339.1767  $[\text{M}+\text{Cl}]^-$ .



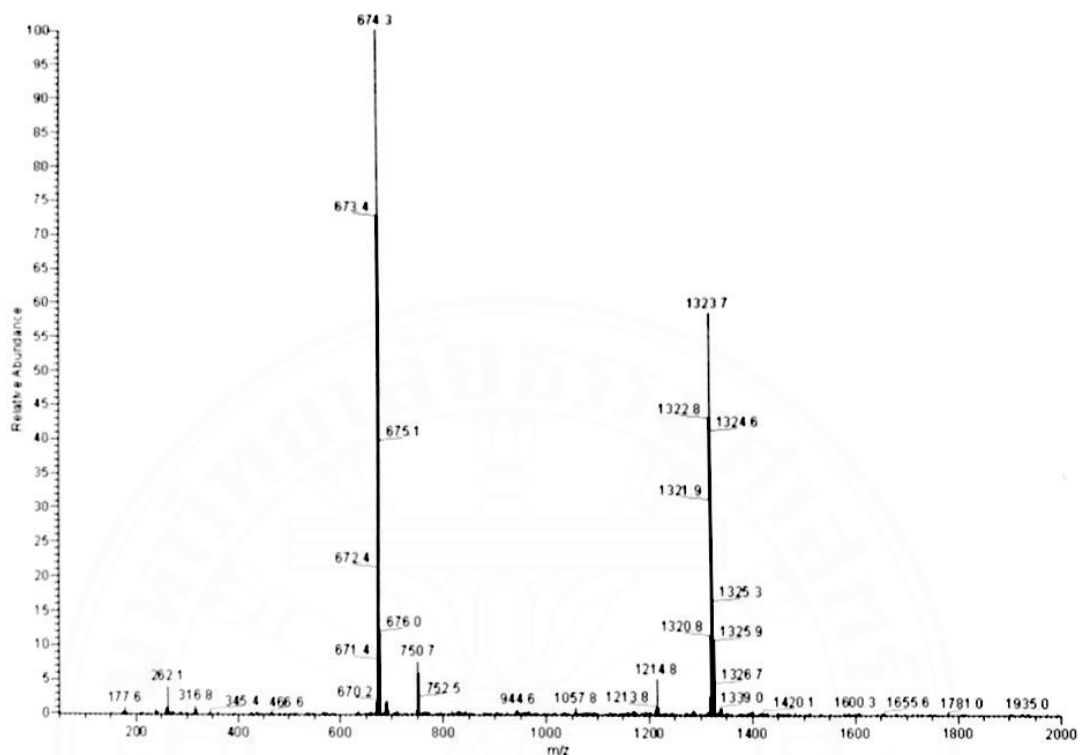
**Figure D3** Positive mode mass spectrum of *N,N'*-diamino-bis(Fmoc-argenyl(Pbf)-glycyl)pimelic acid) (**4**) Found 654.4  $[M+K]^+$ , 670.3  $[M+CH_3OH+Na]^+$ , 1283.2  $[2M+CH_3OH+Na]^+$ ; Calculated fragment 616.32  $[M]^+$

**(A)****(B)**

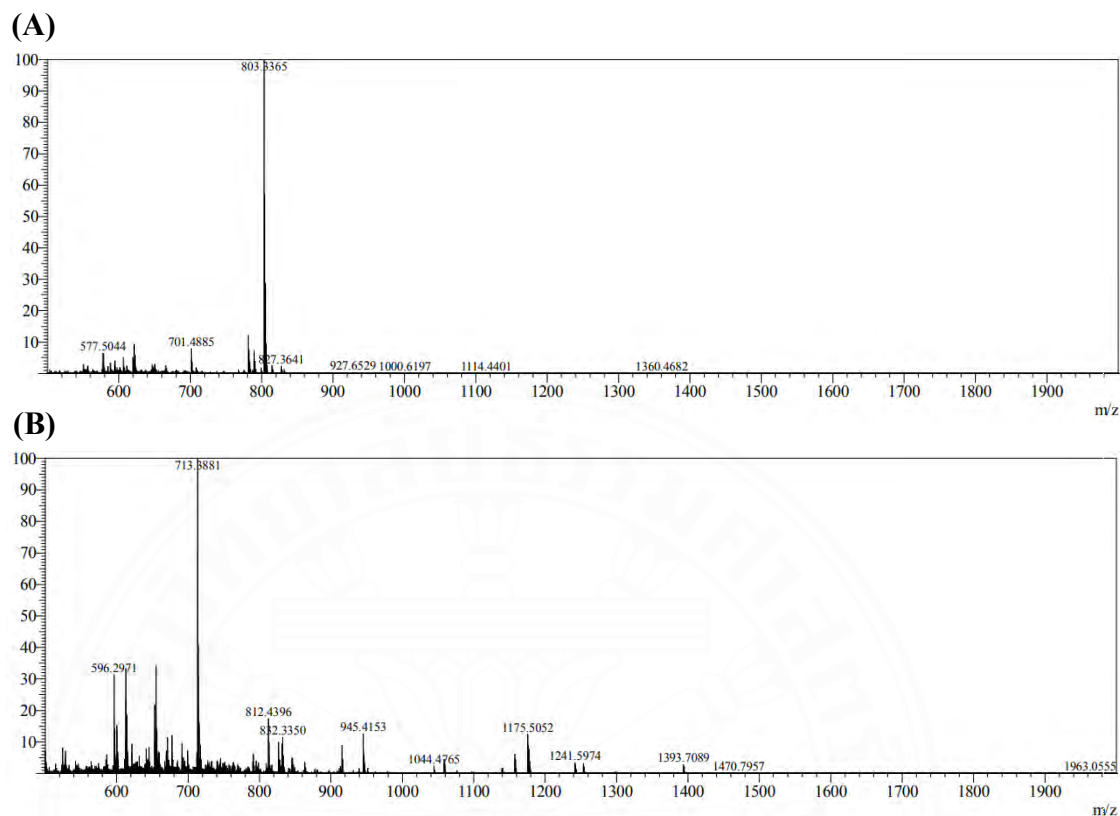
**Figure D4** HR-ESI-MS (*via* positive and negative ion mode) of *N,N'*-diamino-bis(NH<sub>2</sub>-arginyl(Pbf)-glycidyl) pimelic acid (**4a**) ; Calculated 1020.50 [M]<sup>+</sup>.

**(A)** Positive mode HR-ESI-MS. Found 1121.4576[M+H]<sup>+</sup> and 1144.4171[M+Na]<sup>+</sup>

**(B)** Negative mode HR-ESI-MS Found 1157.5385[M+K-2H]<sup>-</sup>



**Figure D5** (a) Positive mode mass spectrum of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-argenyl(Pbf)-glyciny)l) pimelic acid (**5**) Found 674.3  $[M+H+Na]^+$ , 1326.7  $[M+H]^+$  ; Calculated fragment 1326.52  $[M]^+$ .

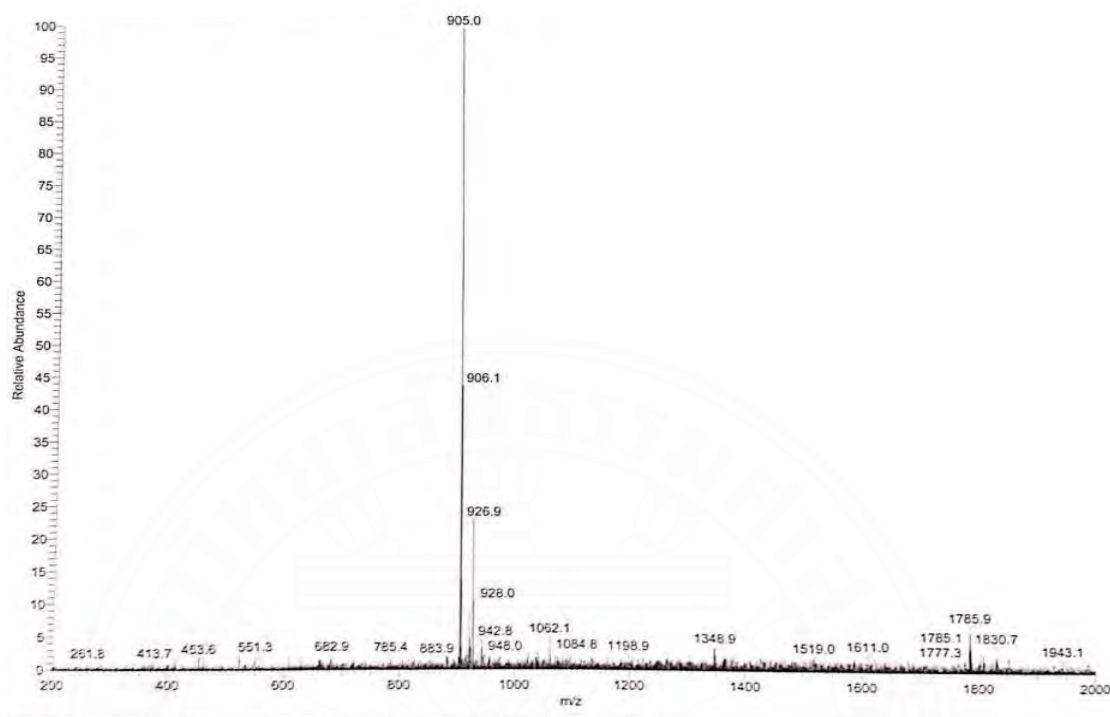


**Figure D6** High resolution mass spectrum positive and negative ion mode of *N,N'*-diamino-bis(NH<sub>2</sub>-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid (**5a**) ;

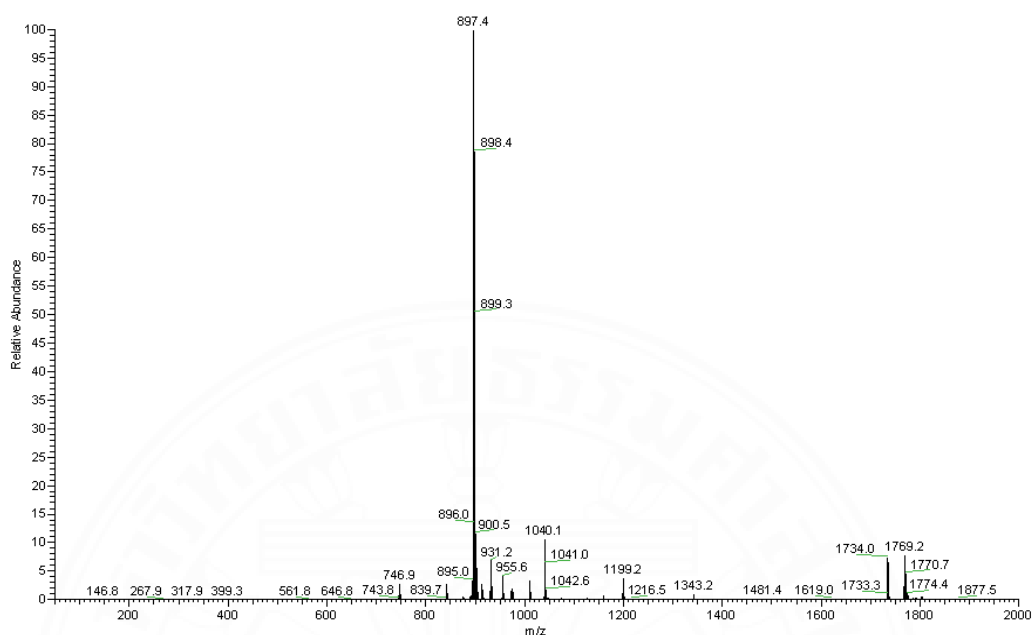
Calculated 1810.74 [M]<sup>+</sup>

**(A)** Positive mode HR-ESI-MS. Found 927.6529 [M+2Na]<sup>2+</sup>

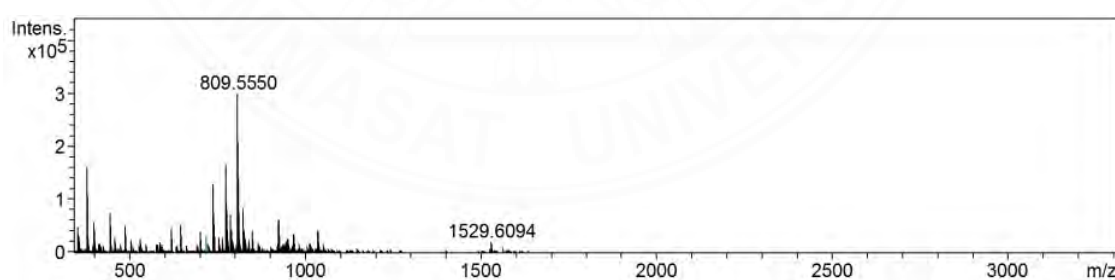
**(B)** Negative mode HR-ESI-MS Found 945.4153 [M+2ACN-2H]<sup>2-</sup>



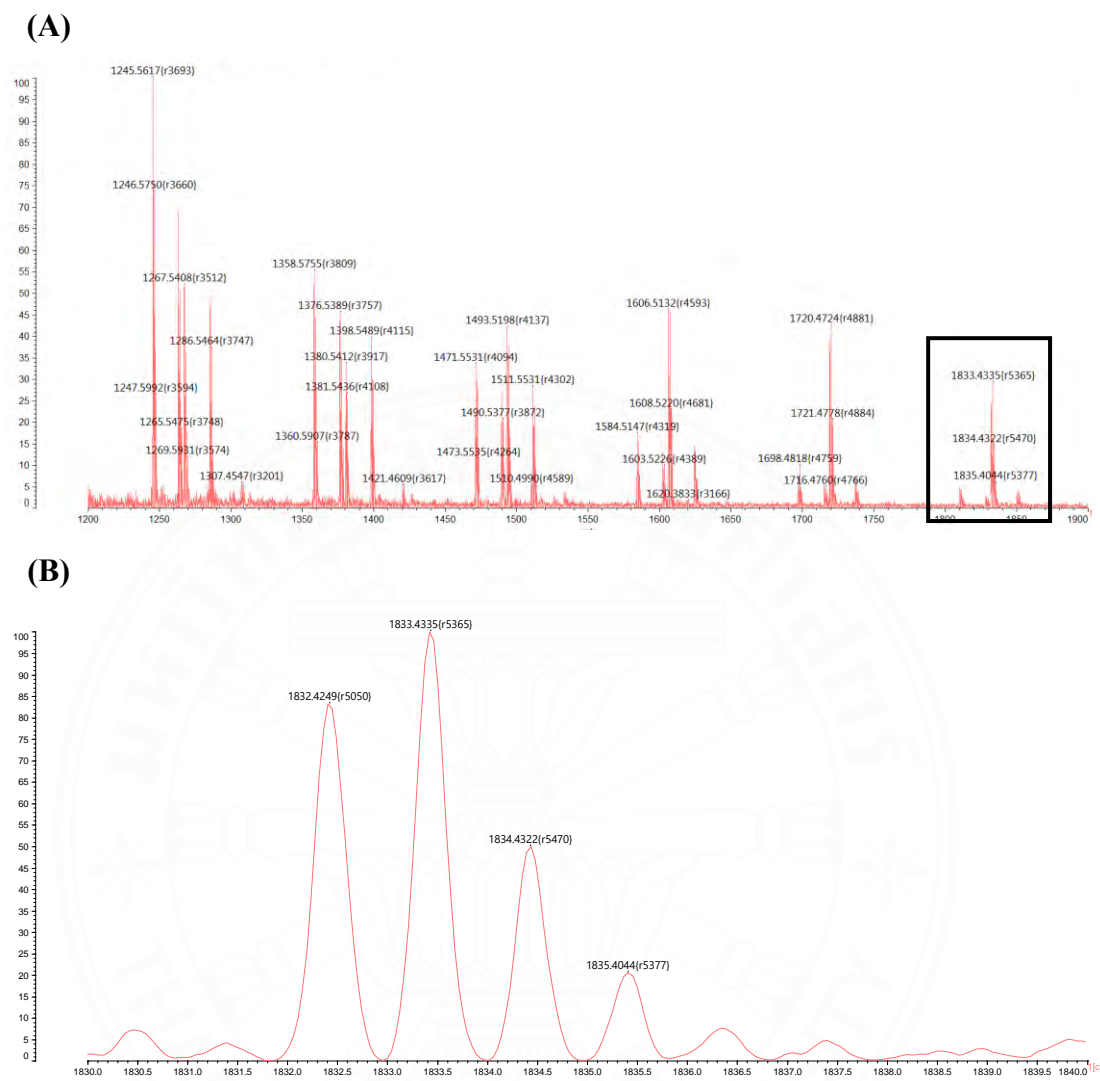
**Figure D7** Positive mode mass spectrum of Fmoc-Ile-Ser(<sup>t</sup>Bu)-Lys(Boc)-Thr(<sup>t</sup>Bu)-OH (11). Found 905.0 [M+Na]<sup>+</sup>, 1785.9 [2M+Na]<sup>+</sup> ; Calculated 881.52 [M]<sup>+</sup>



**Figure D8** Positive mode mass spectrum and structure of *N,N'*-diamino-bis(Fmoc-isoleucyl-serinyl(<sup>t</sup>Bu)-lysyl(Boc)-threonyl(<sup>t</sup>Bu)-cysteinyl(Trt)-arginyl(Pbf)-glycyl) pimelic acid (**12**). Found fragment 897.4  $[M+ACN+4H]^{4+}$ , 1769.2  $[M+2H]^{2+}$ ; Calculated 3537.75  $[M]^+$



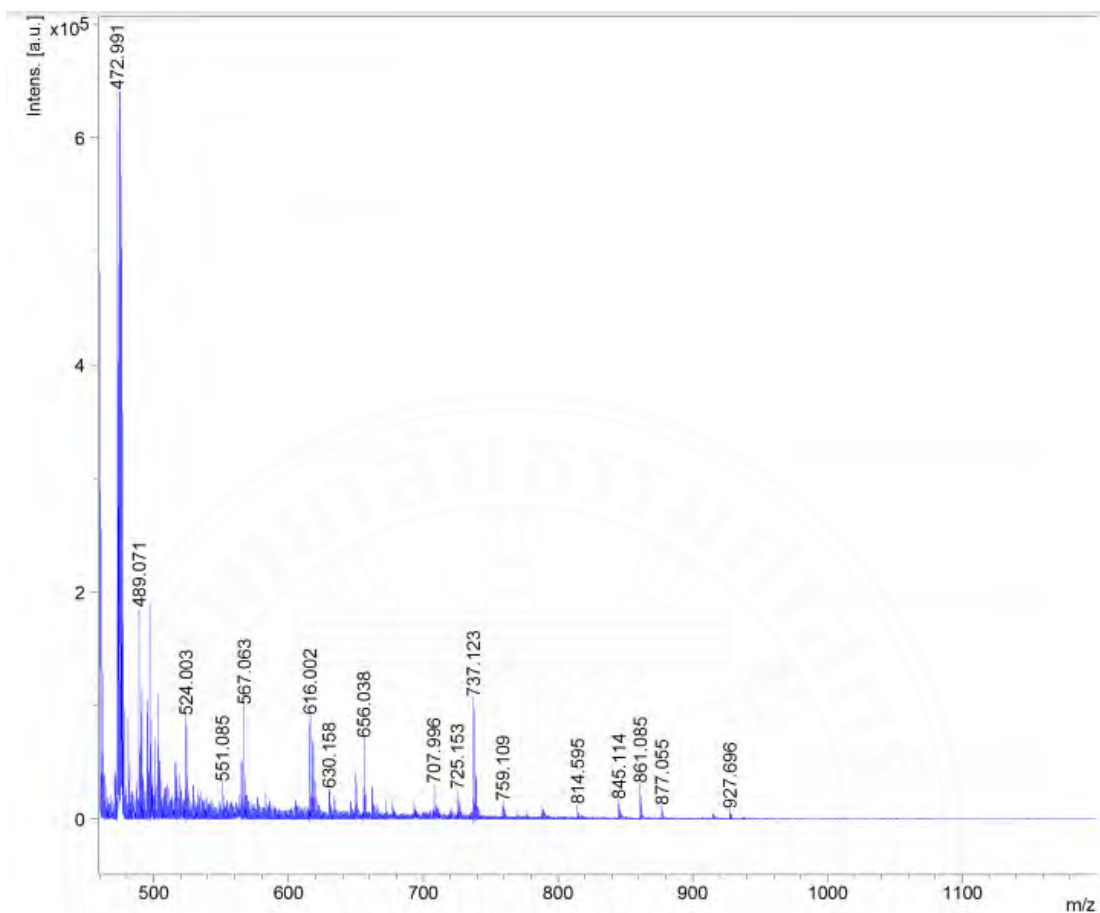
**Figure D9** Positive mode mass spectrum of corresponding ligation by product. Found fragment 809.5550  $[M+ACN+4K]^{4+}$ , 1529.6094  $[M+H+Na]^{2+}$ ; Calculated 3037.55  $[C_{153}H_{224}N_{24}O_{32}S_4]^+$



**Figure D10** MALDI-TOF mass spectrum of *N,N'*-diamino-bis(cyclic-isoleucinyl-serinyl-lysiny-threonyl-cysteinyl-arginyl-glyciny) pimelic acid (**14**). Found 1833.4335  $[M+CHCA+H]^+$  ; Calculated 1644.86  $[M]^+$

**(A)** MALDI-TOF mass spectrum wide range

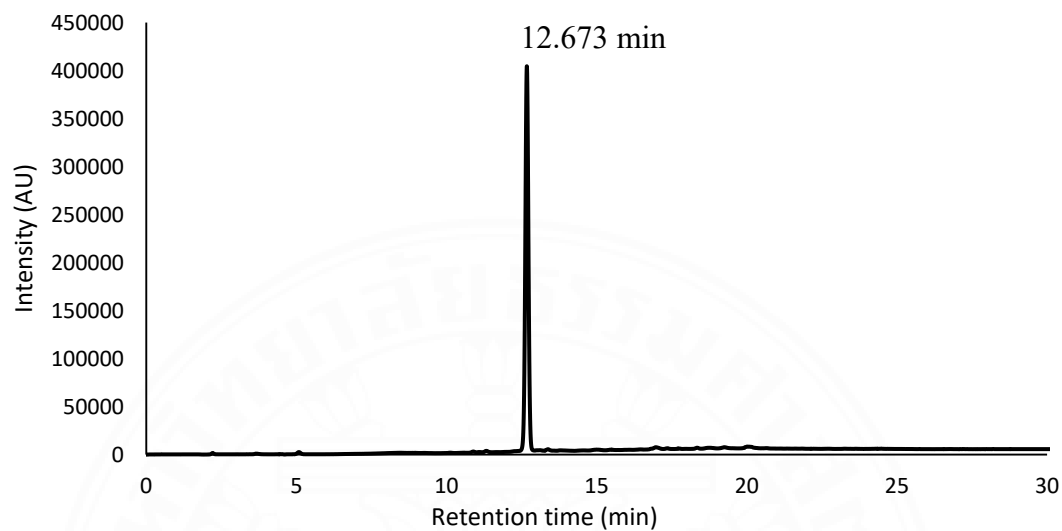
**(B)** MALDI-TOF mass spectrum expanded scale 1830-1840 m/z



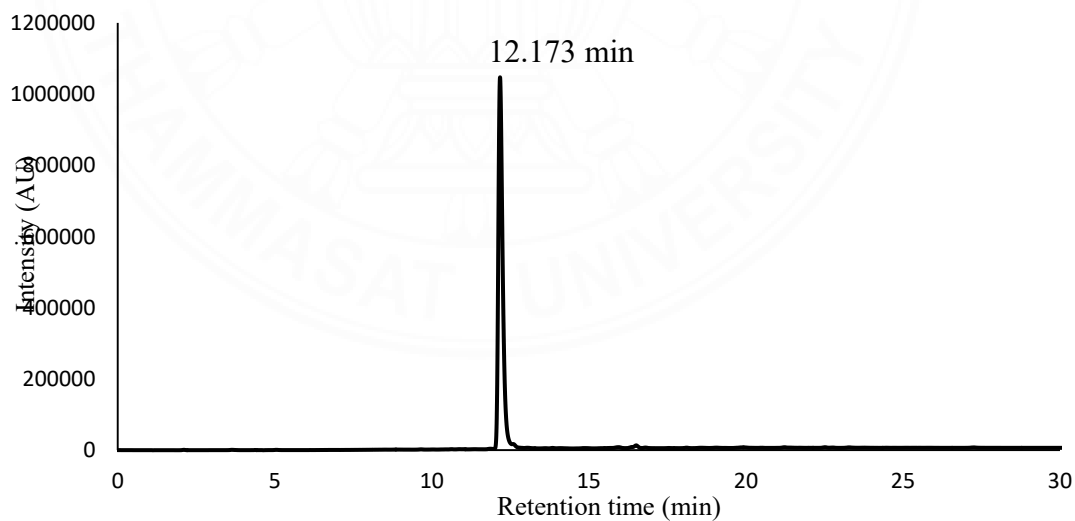
**Figure D10** MALDI-TOF mass spectrum and structure of fluorescently labelled *N,N'*-diamino-bis(isoleucinyll-serinyl-lysinyll-threonyll-cysteinyl-arginyl-glycinyll) pimelic acid (**15**). Found 616.002  $[M+4H]^{4+}$ , 845.114  $[M+3Na]^{3+}$ ; Calculated 2459.95  $[M]^+$

## APPENDIX E

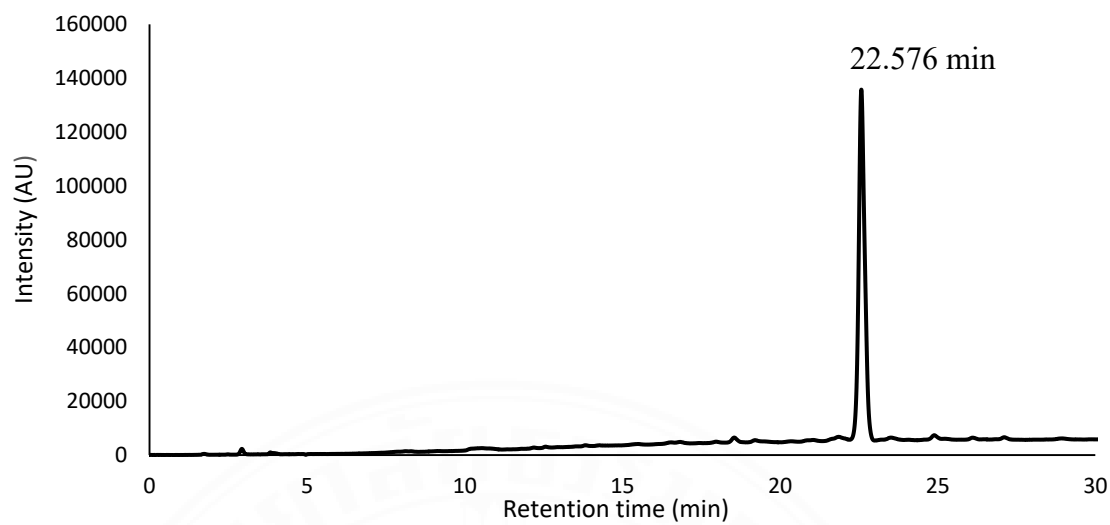
## High performance liquid chromatography chromatogram



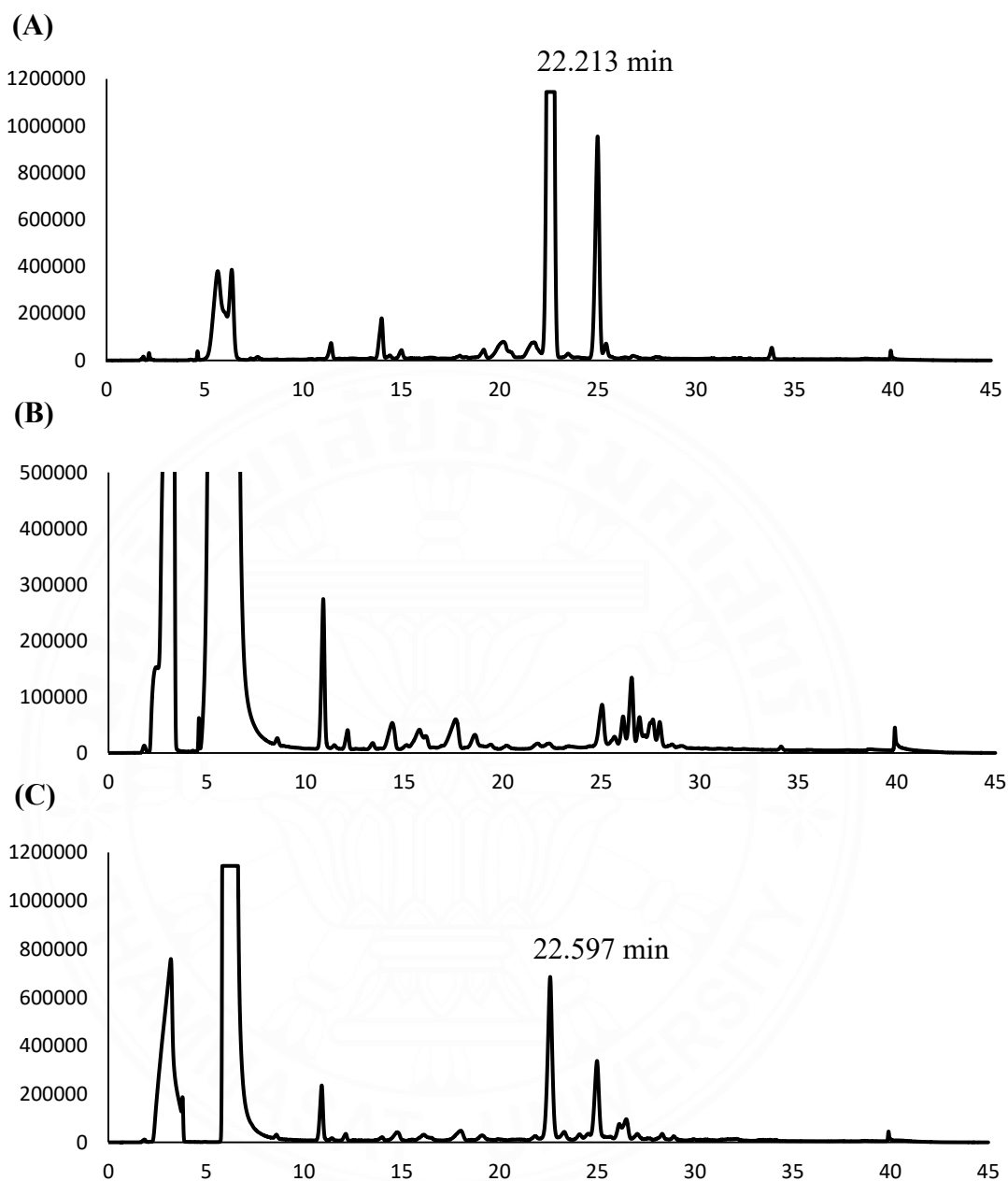
**Figure E1** HPLC chromatogram of *N,N'*-diamino-bis( $\text{NH}_2$ -glyciny)l) pimelic acid measured (**2a**) at 215 nm.



**Figure E2** HPLC chromatogram of *N,N'*-diamino-bis( $\text{NH}_2$ -argenyl(Pbf)-glyciny)l) pimelic acid (**4a**) measured at 215 nm.



**Figure E3** HPLC chromatogram of *N,N'*-diamino-bis(NH<sub>2</sub>-Cysteine(Trt)-argenyl(Pbf)-glyciny) pimelic acid (**5a**) measured at 215 nm.

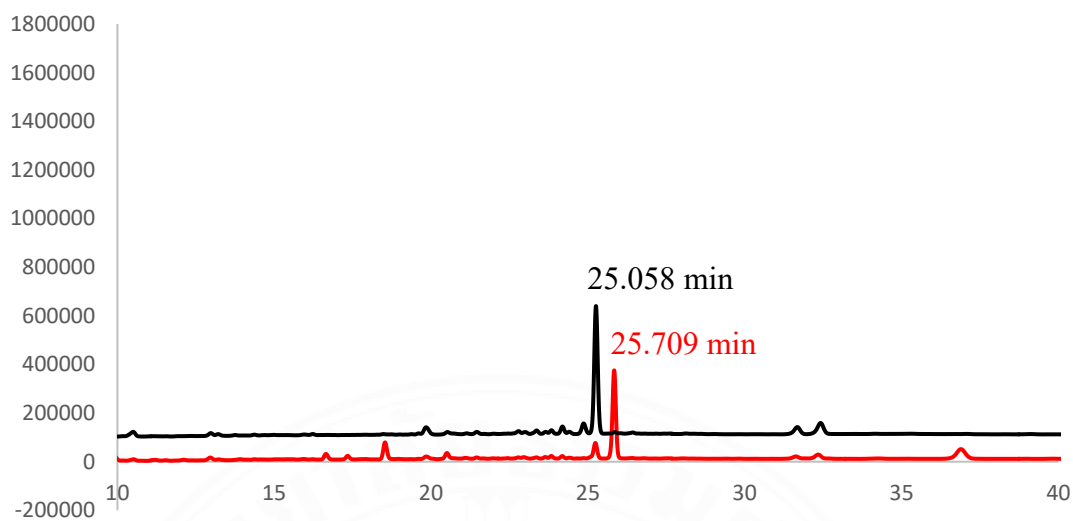


**Figure E4** The comparison HPLC chromatogram of ligation reaction and peptide precursor measure at 215 nm.

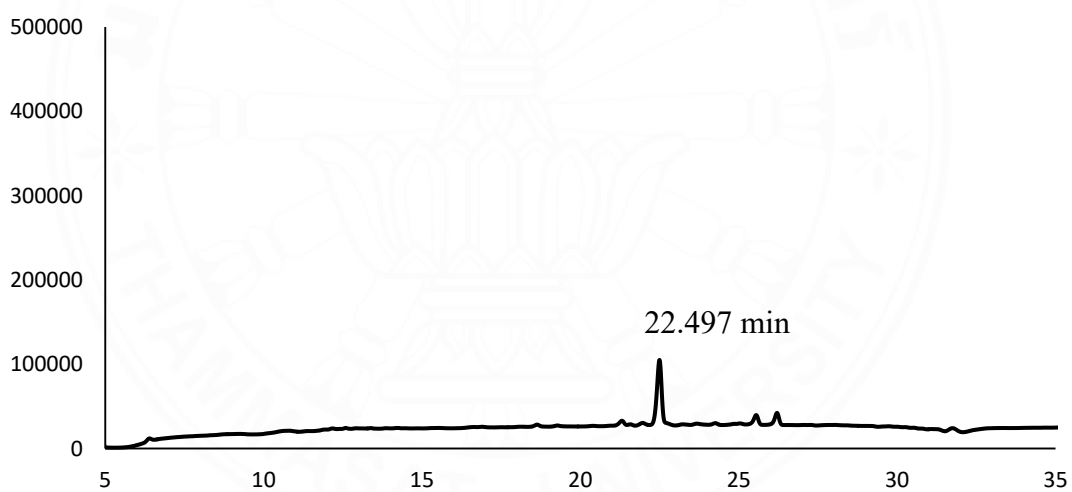
**(A)** Fmoc deprotection reaction crude of *N,N'*-diamino-bis(Fmoc-cysteinyl(Trt)-arginyl(Pbf)-glyciny)l pimelic acid) (**5a**)

**(B)** carboxyl activation reaction of Fmoc-isoleucine-serine(<sup>t</sup>Bu)-lysine(Boc)-threonine(<sup>t</sup>Bu)-OH (**11**)

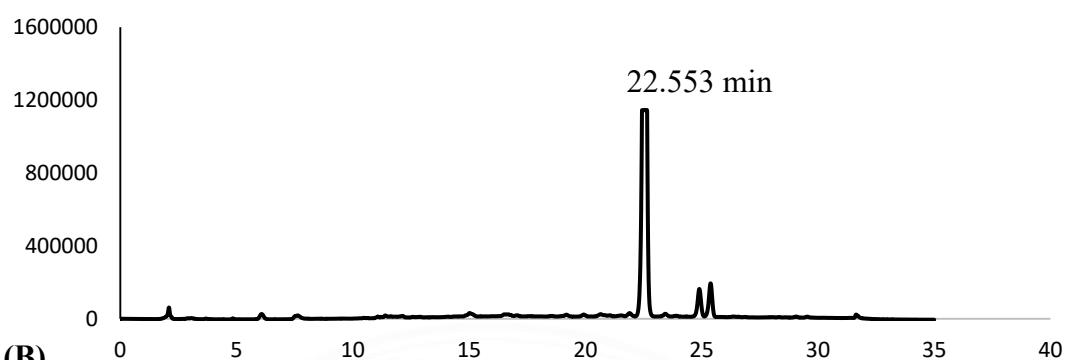
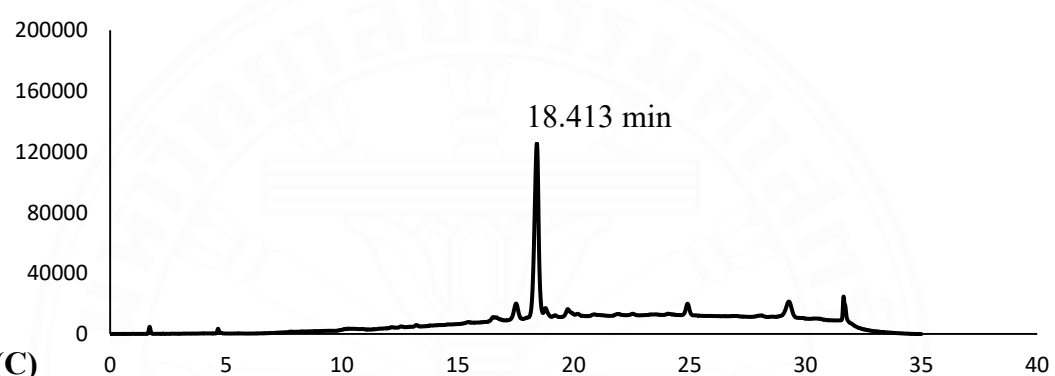
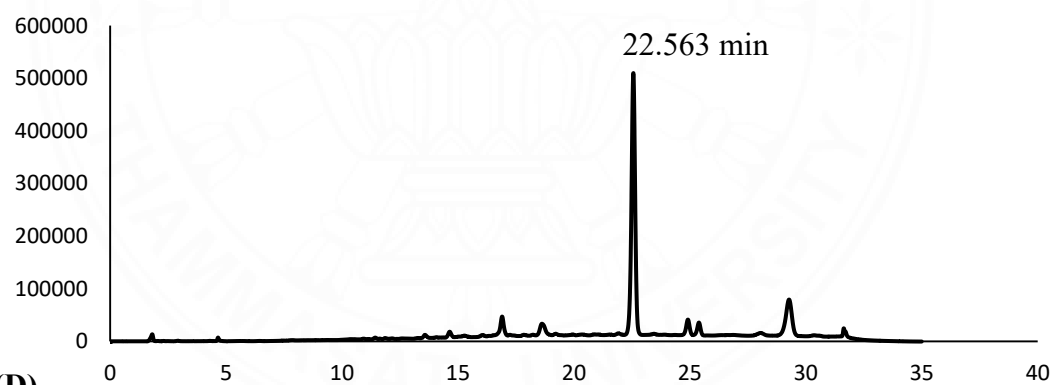
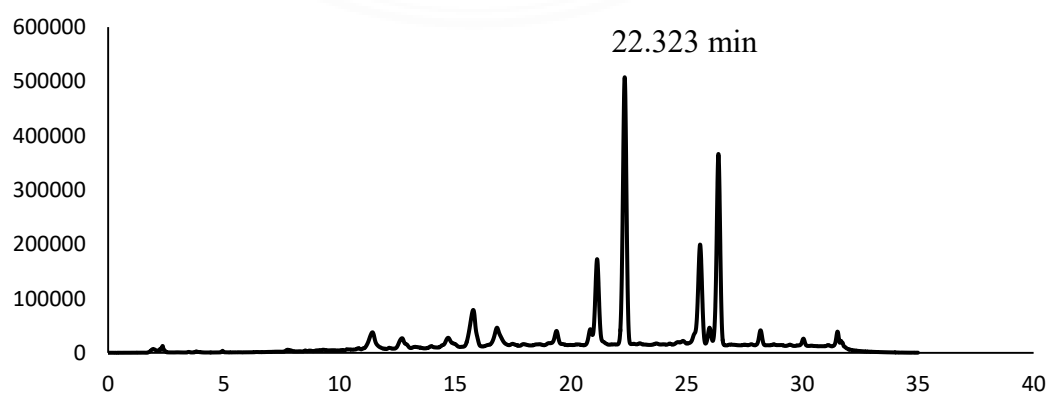
**(C)** the crude ligation reaction

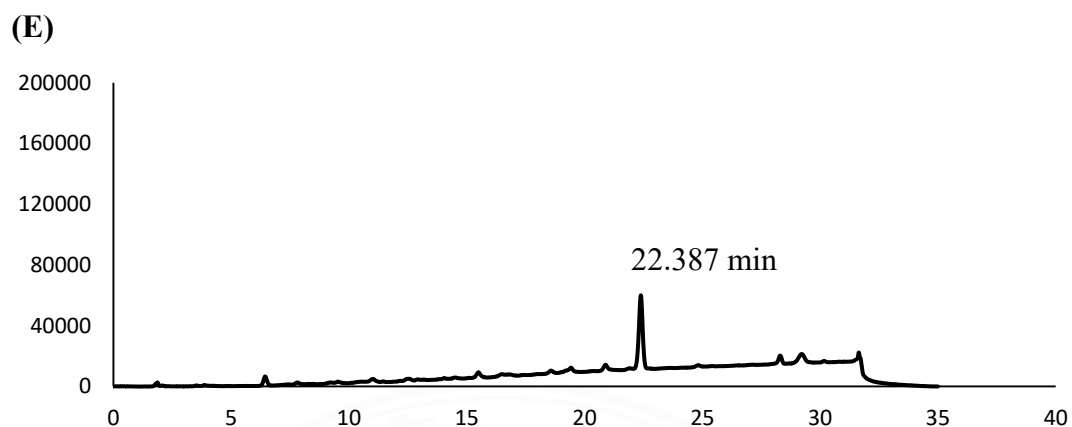


**Figure E5** The HPLC chromatogram comparison between Fmoc deprotected ligation (black line) and cyclization reaction crude (red line) measured at 215 nm.



**Figure E6** The HPLC chromatogram of bismacroyclic SFTI-1.

**(A)****(B)****(C)****(D)**



**Figure E7** The HPLC chromatogram of fluorescently labelled peptide reaction and precursors measure at 215 nm.

**(A)** Fmoc deprotection reaction crude of ligation product

**(B)** Standard FITC

**(C)** Fluorescently labeled bivalent SFTI-1 reaction crude

**(D)** Global deprotection of Fluorescently labelled bivalent SFTI-1 reaction crude

**(E)** Fluorescently labelled bivalent SFTI-1

## BIOGRAPHY

Name Miss Sujichon Thangvichien  
Date of Birth 21 April, 1995  
Educational Attainment 2016: B.Sc. (Chemistry)  
Department of Chemistry, Faculty of  
Science and Technology, Thammasat  
University

### Publication

- Sujichon Thangvichien, Pakonsiri Sontisiri, Warelee Saengdee and Panumart Thongyoo. Synthesis of bivalent SFTI-1 peptidomimic based on the diaminopimelic acid. Proceeding The 14<sup>th</sup> International Symposium of the Protein Society of Thailand. 2019:191-199.

Work Experiences 2017-2019: Teaching assistant at Department  
of Chemistry, Faculty of Science and  
technology, Thammasat University