ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

NANYANG
TECHNOLOGICAL
UNIVERSITY

CHEMICAL AND ENZYMATIC METHODS FOR SELECTIVE PEPTIDE AND PROTEIN MODIFICATION OR SYNTHESIS

TAN XIAOHONG

SCHOOL OF BIOLOGICAL SCIENCES

2008

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chemical and Enzymatic Methods for Selective Peptide and Protein Modification or Synthesis

Tan Xiaohong

School of Biological Sciences

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

2008

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

To my parents and my wife

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Acknowledgements

The first thank-you must go to my supervisor, Prof. Liu Chuanfa, for his guidance, support, assistance and understanding throughout the course of my PhD work. I particularly appreciate his enthusiasm and dedication to my research and also his allowing me to have the opportunity to work on such interesting projects. What I have learned from my supervisor will have a direct impact on my personal and professional life. His serious attitude toward science also deeply impresses me.

The next thank-you goes to my parents, wife and family for their unconditional love, encouragement, understanding and support.

I appreciate the time that my thesis examiners spent on reviewing my writing and their valuable feedbacks and comments.

I also thank the past and present members of our group for their immense help during the past few years, particularly Mr. Andre Wirjo, Dr. Zeng Yun, Ms. Zhang Xiaohong, Mr.Yang Renliang and Dr. Sathan Thennarasu.

I must thank all of my friends, especially Mr. Liu Xinyu, Mr. Xie Yi, Mr. Dai Liang, Dr. Shi Minglong, Mr. Ji Zhe, Dr. Liu Bin, Dr. Kan Chongbao and Ms. Tang Manli. I thank them for their friendship and support.

Finally, I would like to thank all the academic, teaching and technical staff in the School of Biological Sciences, Nanyang Technological University for their assistance during the past years.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Table of Contents

Acknowledgments……………………………………………………iii

Table of Contents……………………………………………..…..iv

Abstract………………………………………………………vii

Abbreviations…………………………………………………ix

Chapter 1: Introduction………………………………………...1

1.1 Selective Modification Reactions on the Side Chains of 20 Natural Amino Acid Residues…………………………………………………………..2

1.2 NT- or CT-specific Reactions for Single-site Protein Modification...11

1.2.1 NT-Modification…………………………………………......11

1.2.2 CT-Modification…………………………………………………13

1.3 Purpose of this Study………………………………………...........17

Chapter 2: Systematic Investigation of the Transamination Reaction for N-terminal Selective Modification……18

2.1 Introduction……………………………..18

2.2 Results and Discussion…………………………………………………20

2.2.1 Solid Phase Transamination Reaction Assay………………………..20

2.2.2 Transamination Reaction Assay in Solution………………………...27

2.2.3 An Application Example of the Transamination Reaction………...39

2.3 Summary and Conclusion……………………………………………......40

2.4 Experimental……………………………………………………………..41

2.4.1 Materials…………………………………………………………..41

2.4.2 High Performance Liquid Chromatography…………………………42

2.4.3 Mass Spectrometry………………………………………………..42

2.4.4 NMR Spectroscopy………………………………………………..42

2.4.5 Preparation of Dipeptide Library……………………………………42

2.4.6 Preparation of Dabsyl-Ala(β)-hydrazide…………………………..43

2.4.7 Peptide Synthesis……………………………………………………44

2.4.8 Transamination Reaction of the Membrane-bound Dipeptide

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Library……………………………………………………………44
2.4.9 Reaction of Transaminated Dipeptide Library with Dabsyl-Ala(β)–hydrazine………………………………………………………...44
2.4.10 Transamination Reactions of Model Peptides in Solution....……….45

Chapter 3: Development of an Enzymatic Method for the Synthesis of Peptide C^{α}-thioesters…………………..46

3.1 Introduction…………………………………………………………...46
3.2 Results and Discussion………………………………………………….54
3.3 Summary and Conclusion……………………………………………....66
3.4 Experimental……………………………………………………………67
3.4.1 Materials………………………………………………………….67
3.4.2 High Performance Liquid Chromatography……………………….67
3.4.3 Mass Spectrometry………………………………………………..67
3.4.4 Expression and Purification of Subtiligase………………………….68
3.4.5 Peptide Synthesis………………………………………………….69
3.4.6 Procedures for Enzyme Assays……………………………………...70
3.4.7 Chemical Ligation between Enzymatically Synthesized Peptide Thioesters and NT-Cysteinyl Peptides……………………………….73

Chapter 4: Development of Enzymatic and Chemical Methods for the Synthesis of Peptide C^{α}-thioacids………………..74

4.1 Introduction…………………………………………………………...74
4.2 Results and Discussion………………………………………………...79
4.2.1 An Enzymatic Method for the Synthesis of Peptide Thioacids……...79
4.2.2 Redesign of Subtiligase for the Synthesis of Peptide Thioacids.........89
4.2.3 A Simple Chemical Method for the Synthesis of Peptide Thioacids..96
4.3 Summary and Conclusion……………………………………………...107
4.4 Experimental…………………………………………………………...108
4.4.1 Materials…………………………………………………………108
4.4.2 High Performance Liquid Chromatography………………………..109
4.4.3 Mass Spectrometry………………………………………………109
4.4.4 Expression, Purification and Site-specific Mutation of Subtiligase..109

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.4.5 Peptide Synthesis…………………………………………………...109

4.4.6 Modification of NT-Cys Peptides by DTNP……...........................111

4.4.7 Thiol Quantification……………………………………………..112

4.4.8 Procedures for Enzyme Assays……………………………………113

4.4.9 Conversion of Peptidyl Thioesters to Thioacids through Hydrothioly-sis in Aqueous Buffer……………………………………………...116

4.4.10 Peptide Chemical Ligation Procedures……………………………116

Summary……………………………………………………120

References…………………………………………………….126

Publications…………………………………………………..139

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Abstract

This thesis aims to investigate and develop N- or C-terminus specific methods for selective protein modification.

Nowadays, protein modification methods rely mainly on specific reactions on the side chains of amino acid residues, such as Cys and Lys. However, because most proteins display multiple copies of the targeted residue on their surface and also for some other reasons, modifying a protein on a desired site is always problematic. Alternatively, if one can differentiate the N- or C-terminus of a protein from the internal side-chain functional groups, it may be possible to site-specifically modify a protein at one single site. In this thesis, I have conducted several projects related to protein/peptide NT- or CT-selective modifications.

Firstly, I investigated, in a systematic manner, the NT-specific transamination reaction on solid phase and in solution. For the solid phase study, the relative reactivity of membrane-bound dipeptides towards transamination was investigated by comparing the color intensities of the dyed spots through the use of a dye reagent. A series of solution reactions was then performed on selected model peptides. This is the first time that such a study was conducted which revealed the tendency of NT-amino acids toward transamination as well as the neighbouring group impact from the second residue. All of these findings will provide useful guidelines for the future use of transamination for protein modification and bioconjugation applications.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Secondly, on the CT-modification aspect, I focused on developing new chemical and enzymatic methods to produce peptidyl C^{α}-thioesters and C^{α}-thioacids, which are important building blocks for peptide and protein synthesis. Currently, the synthesis of peptide CT-thioesters and thioacids remains difficult. The methods I have developed are useful alternatives to the currently available techniques and will make these compounds readily accessible. The usefulness of our methods has been demonstrated in several applications.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Abbreviations

1. Reagents

Ac	Acetyl
ACA	N-Acetylcysteamine
Boc	t-Butyloxycarbonyl
PyBOP	Benzotriazole-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
Dabsyl	4-(Dimethylamino)azobenzene-4'-sulfonyl
DCC	Dicyclohexylcarbodimide
DCM	Dichloromethane
DIEA	N,N-Diisopropylethylamine
DMAP	4-(Dimethylamino)pyridine
DMF	N,N-Dimethylformamide
DMSO	Dimethylsulphoxide
DTNP	2-2'Dithiobis-(5-nitropyridine)
DTNB	5,5′-Dithiobis(2-nitrobenzoic acid)
EDTA	Ethylenediaminetetraacetic acid
Fmoc	9-Fluorenylmethoxycarbonyl
glc	Glycolate
MBHA	4-Methylbenzhydrylamine
MES	Mesityl-2,4,6-trimethylphenyl
MPA	Mercaptoproprionic acid
Npys	3-nitro-2-pyridinesulfenyl
SDS	Sodium dodecyl sulfate
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
TIS	Triisopropylsilane
TRIS	Tris(hydroxymethyl)aminomethane

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2×TY	2×Tryptone yeast

2. Others

CT	C-Terminal
Da	Dalton
ESI-MS	Electrospray ionization mass spectrometry
HPLC	High performance liquid chromatography
MW	Molecular weight
MALDI	Matrix-assisted laser desorption/ionization
NMR	Nuclear magnetic resonance
NT	N-Terminal
PCR	Polymerase chain reaction
PAGE	Polyacrylamide gel electrophoresis
RT	Room temperature
SPPS	Solid phase peptide synthesis
UV	Ultraviolet Rays

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3. The 20 Natural Amino Acids

Amino Acid	Abbreviation 3 Letter	1 Letter	Integral Mass	Structure
Alanine	Ala	A	71	CH_3 —NH-CH-CO—
Arginine	Arg	R	156	$CH_2CH_2CH_2$-NH-C(=NH)-NH_2 —NH-CH-CO—
Asparagine	Asn	N	114	CH_2—$CONH_2$ —NH-CH-CO—
Aspartic Acid	Asp	D	115	CH_2—CO_2H —NH-CH-CO—
Cysteine	Cys	C	103	CH_2—SH —NH-CH-CO—
Glutamic Acid	Glu	E	129	CH_2CH_2—CO_2H —NH-CH-CO—
Glutamine	Gln	Q	128	CH_2CH_2—$CONH_2$ —NH-CH-CO—
Glycine	Gly	G	57	H —NH-CH-CO—
Histidine	His	H	137	HN, N CH_2— —NH-CH-CO—
Isoleucine	Ile	I	113	$CH(CH_3)CH_2CH_3$ —NH-CH-CO—
Leucine	Leu	L	113	$CH_2CH(CH_3)_2$ —NH-CH-CO—
Lysine	Lys	K	128	$CH_2CH_2CH_2CH_2$-NH_2 —NH-CH-CO—
Methionine	Met	M	131	CH_2CH_2-S-CH_3 —NH-CH-CO—
Phenylalanine	Phe	F	147	CH_2— —NH-CH-CO—
Proline	Pro	P	97	—N-CH-CO—
Serine	Ser	S	87	CH_2—OH —NH-CH-CO—
Threonine	Thr	T	101	$CH(OH)CH_3$ —NH-CH-CO—
Tryptophan	Trp	W	186	NH CH_2— —NH-CH-CO—
Tyrosine	Tyr	Y	163	—OH CH_2— —NH-CH-CO—
Valine	Val	V	99	$CH(CH_3)_2$ —NH-CH-CO—

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 1: Introduction

Site-selective protein modification is an essential technique for the preparation of protein conjugates and for protein reengineering. At present, protein engineering is carried out mostly through mutagenesis at the gene level. On the other hand, selective enzymatic and chemical modifications on the amino acid side chains, N-terminal (NT-) amino groups and C-terminal (CT-) carboxylic acids are being used increasingly for a broad range of applications. For instance, by using these techniques proteins can be modified with isotopic tags to facilitate proteomics analysis by mass spectrometry[1], with fluorescent[2] or paramagnetic probes[3] for biophysical studies by fluorescence, electron paramagnetic resonance[4] or NMR spectroscopy, or with biocompatible polymers for use as diagnostic and therapeutic agents possessing improved pharmacokinetic properties. However, it remains a big challenge in the field of protein bioconjugation to develop strategies that can modify proteins once at a single desired site.

To make an original contribution to this area, I aim to study selective modification reactions on the N- and C-termini of peptides or proteins. This thesis consists of two parts: (1) systematic and extensive investigation of the transamination reaction in order to provide reliable and predictable information about this NT-specific reaction for protein or peptide modification; (2) developing novel enzymatic and chemical methods for the synthesis of peptide C^{α}-thioesters or thioacids, which are important and versatile building blocks in protein and polypeptide synthesis or can be used for bioconjugation at the C-terminus.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

In this introduction chapter, firstly I will describe the currently known side-chain modification reactions on the 20 natural amino acid residues in a systematic way. Secondly, the NT- and CT-specific modification reactions will be reviewed. Moreover, the advantages of NT- and CT-modifications over side-chain modifications will be discussed, which highlights the important contribution of my thesis to this area of research.

1.1 Selective Modification Reactions on the Side Chains of 20 Natural Amino Acid Residues

In vivo, selective side-chain modifications are usually achieved through specific enzymatic reactions, whereas *in vitro*, chemical methods are used to attach a functional moiety to a protein. The two are introduced separately in the following sections.

1. Specific enzymatic modification reactions, such as phosphorylation, glycosylation, sulfation, acetylation, methylation, farnesylation, ubiquitination and so on, play a defining role in cellular processes, including protein trafficking and localization, signal transduction, transcriptional regulation, and targeted protein destruction. These reactions are mainly carried out on the side chains of amino acid residues. Herein I list some examples for a short review.

For Lys residue, besides the most common acetylation, methylation, and ubiquitination[5, 6], hydroxylation was also reported[7], by which a 5-OH group was introduced and subsequently the Lys residue was glycosylated.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

For Arg residue, the N-methylation reaction often takes place on proteins such as histones[6, 8]. Another modification is citrullination or deimination, through which an arginine residue in a protein is changed to the amino acid citrulline[9]. This conversion has important consequences for the structure and function of proteins, since arginine is positively charged at physiological pH, whereas citrulline is uncharged. This increases the hydrophobicity of the protein, which may lead to protein unfolding.

For Cys residue, disulfide bond formation, S-acylation[10] and S-prenylation[11] are the most common modifications. S-palmitoylation can also be carried out on the sulfhydyl group of Cys residues[12].

For Asp residue, one possible *in vivo* modification is phosphorylation[13]. For Glu residue, a common modification reaction is carboxylation[14], which involves enzymatic fixation of CO_2 to the γ-methylene carbon of Glu residues. This converts a Glu residue to a Gla residue which can undergo bidentate chelation with divalent cations such as $[Ca^{2+}]$[15]. Furthermore, in the polyglutamylation reaction, the Gamma carboxyl group of Glu may form peptide-like bond with the amino group of a free Glu whose α-carboxy group can then be extended into a polyglutamate chain[16].

For Ser and Thr residues, both are most known to undergo not only phosphorylation reaction but also O-glycosylation reaction.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

For Tyr residue, phosphorylation is a common modification[13]. Additionally, the SO_3^- group of phosphoadenosine phosphosulfate can be transferred to side chains of Tyr in proteins[17].

For Asn and Gln residues, one famous modification reaction is deamidation, where an amide functional group is changed to carboxylate. This reaction is important in the degradation of proteins because it converts Asn and Gln to Asp and Glu respectively[18, 19]. In addition, for Asn residue, besides the deamination reaction, another main modification reaction is N-glycosylation[20]. Moreover, hydroxylase can act on the Asn residue to generate 3-OH-Asn[21]. Interestingly, Gln residue usually undergoes transglutamination[22] instead of N-glycosylation.

For His residue, the common posttranslational modifications are the phosphorylation reaction[13] and N-methylation. In addition, with the imidazole ring on its side chain, His is a quite active amino acid to act as a catalytic residue in many enzymes as well as a good ligand to bind metal ions such as Cu^{2+} and Ca^{2+}.

For Met residue, sulfoxide formation by the oxidation reaction is its most common modification[23].

For Pro, usually an inactive residue, C-hydroxylation, which is an oxidation reaction, can occur on the side chain to form the 3-OH-Pro or 4-OH-Pro[7]. The same can occur to the Gly residue even it only has one hydrogen group on its side chain[7].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2. Performing site-specific chemical modification reactions *in vitro* is challenging and difficult due to the variety of functionalities present in a typical protein. Nevertheless, there is an increasing need for such protein modification methods in the post-genome era when the study of proteins occupies the central stage of bioscience research. Some of the currently available methods are already routinely used to prepare protein bioconjugates for a wide range of applications[24, 25].

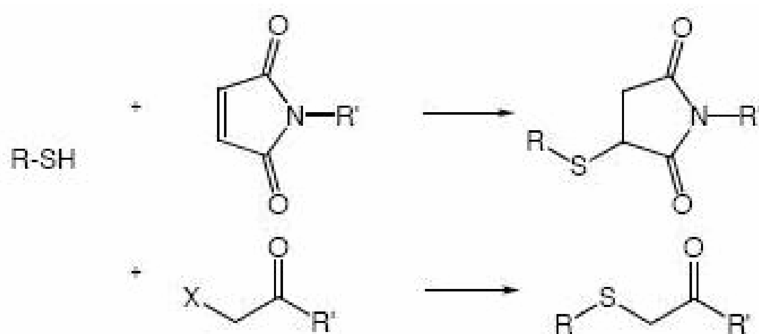
Among the 20 amino acid residues, some of them, like Leu, Ile, Val, Ala and Gly, cannot undergo specific chemical modification reactions on their side chains. On the other hand, various modification reactions on the side chains of the other amino acid residues were reported. Among them, the most widely used strategies rely on electrophilic reagents targeting the nucleophilic functional groups present on lysine, cysteine, aspartic acid and glutamic acid side chains. In fact, over 70% of the commercially available protein labeling reagents are reactive toward lysine or cysteine residues[26]. In the following section, I collected a number of common chemical reactions for modification on the side chains of the natural amino acid residues to make a relatively comprehensive review on this topic.

For Lys residue, the ε-amino group, as a strong nucleophile, is commonly used for protein modification. Usually, electrophilic reagents, such as N-hydroxysuccinimidyl-esters and isocyanates, preferentially react with the lysine ε-amino groups. For example, to stabilize carboxypeptidase for X-ray structure determination, glutaraldehyde was used for enzyme crosslinking through reaction on Lys residues[27]. This kind of crosslinking modification can also improve the stability of penicillin G acylase by using dextran dialdehyde[28]. Moreover, the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

thermal stability of enzymes, such as trypsin, can also be enhanced by crosslinking with glutaraldehyde or sucrose aldehydes (monomeric or polymeric), followed by stabilization of the aldimine bonds by reduction[29]. Also, functionally important groups can be introduced into proteins through the lysine residues by reductive alkylation with aldehydes[29]. In another elegant approach, researchers introduced metals into a catalytic antibody through the lysine residue by utilizing bisimidazole anhydride[30].

For Cys residue, maleimides, α-haloketones, iodoacetamides or other electrophilic groups preferentially react with thiols of the cysteine residues (Scheme 1.1).



Scheme 1.1 Common reagents to react with thiol groups of cysteines.

The modifications on the side chains of Cys residues were well documented. For example, the unique thiol group of SBL cysteine mutants was reacted with methanethiosulfonate (MTS) reagents to introduce a variety of groups which give new functionalities to the protein[31]. In another elegant work, a cationic cofactor was created in the modified protein by reaction between a cysteine residue and a thiopyridyl reagent[32]. Interestingly, the cysteine residues in some enzymes can be

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

converted to serine residues. For example, [Cys25]-papain was changed to [Ser25]-papain to give it a new enzymatic function[33].

Recently, formyglycine-generating enzyme (FGF), which oxidizes cysteine to formyglycine in eukaryotes[34, 35], was utilized by the Bertozzi group to introduce genetically encoded aldehydes into proteins[36]. FGF recognizes the peptides containing a highly conserved CXPXR submotif, where X is usually serine, threonine, alanine or glycine. Through this method, functionally important chemical compounds, such as PEG, can be covalently linked to proteins for quite broad applications.

Besides Lys and Cys residues, the two most common modification targets, selective modification reactions on the side chains of other amino acid residues were also intensively investigated.

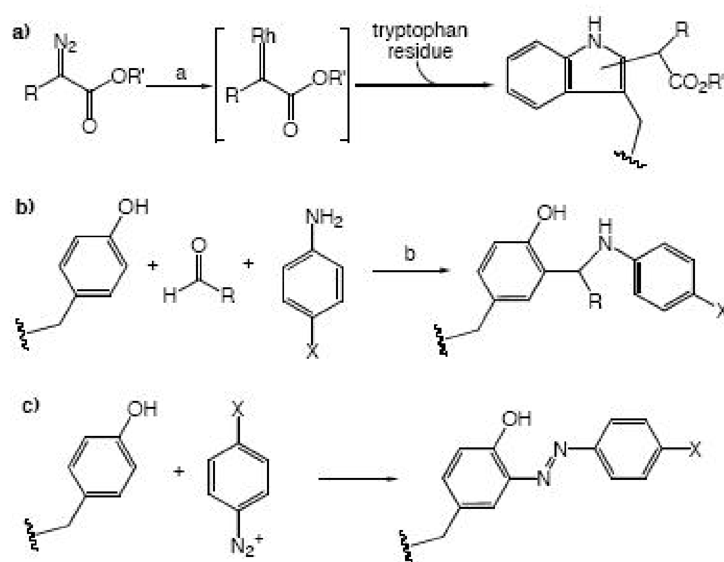
For Asp and Glu residues, because of the carboxyl group on their side chains, they can be coupled to amine groups to form amide bonds in the presence of dehydration reagents such as EDC and HOBT.

For Ser residue, one good example is that the catalytic serine residue of trypsin was activated with phenylmethyl sulfonyl fluoride and then displaced with sodium selenide yielding seleno-trypsin. The seleno-trypsin was devoid of hydrolytic activity but exhibited good glutathione peroxidase activity[37]. A similar approach was applied to subtilisin to generate thiol-subtilisin and seleno-subtilisin[38, 39].

For Trp residue, the indole functionality can react with the rhodium carbenoids[40], which is one of the first chemoselective methods that can attach new functionality

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

to Trp residues. Due to the low abundance of Trp residues in a protein, the strategy targeting Trp residues for specific protein modification reactions is a good alternative to the cysteine modification chemistry. For Tyr residue, a new modification method has been developed through the use of pi-allylpalladium complexes[41]. Moreover, other reactions to modify aromatic rings were developed on the side chains of both Tyr and Trp residues[40, 42, 43] (Scheme 1.2).



Scheme 1.2 Reactions to modify the aromatic amino acids. a) Trp modification; b) and c) Tyr modification.

Tyr, as well as Trp, can also undergo oxidative cross-linking. During this reaction, a tyrosyl radical is generated and couples to nearby functional groups, joining the two species through covalent bond formation[44].

For Arg residue, the guanidino groups can react with phenylglyoxal under mild conditions, pH 7 to 8, rapidly and selectively[45]. Moreover, the derivative is stable

under mildly acidic condition and can be reversibly removed in good yield at a higher pH.

For His residue, the most famous reagent for specific modification is diethylpyrocarbonate, which can be used for quantification of His residues in proteins[46].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

residues on its surface, except when the reactivity of the two cysteines is quite different[50].

Nevertheless, protein modification on a single site is often necessary for a particular application. For example, when an antibody is covalently conjugated to a drug or a ligand, it is always desirable to modify the antibody at only one site in order to maximally preserve its binding activity. But most traditional protein modification methods discussed above do not meet this requirement, and their use often results in product mixtures with considerable structural heterogeneity, which complicates subsequent purification and analytical efforts. For this reason, numerous new chemo- and regio-selective protein modification methods have been developed in recent years.

Perhaps the most exciting recent development in this area lies with the use of methods that exploit the protein synthesis machinery for direct incorporation of unnatural amino acids into proteins at the translation step. These methods rely on an amber stop codon and a corresponding suppressor tRNA that is charged with a desired unnatural amino acid which may already carry the biophysical probe of interest in place or contain a special chemical/photochemical reactivity for later use. This concept was first applied in cell-free protein synthesis systems with great success[51]. More recently, *in vivo* systems based on both prokaryotic and eukaryotic organisms have also been developed, which employ an orthogonal amber tRNA/aminoacyl-tRNA synthetase pair for each non-natural amino acid encoded by the amber stop codon[52]. Although the aforementioned new methods are capable of introducing a site-specific modification, they all require genetic manipulation,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

such as site-directed mutagenesis, on the target protein's gene sequence and are therefore not applicable to natural proteins which already exist.

Given the disadvantages of the side-chain modification methods mentioned above, it is necessary to develop strategies that can modify peptides or proteins by targeting their N- or C-terminus because such a strategy would provide a straightforward way to introduce a functional group into the target protein at one single site for the obvious reason that a protein contains only one N-terminus and one C-terminus.

1.2 NT- or CT-specific Reaction for Single-site Protein Modification

1.2.1 NT-Modification

In vivo, the common modification on N-terminus includes acetylation, formylation, methylation, myristoylation[53], pyroglutamate[54], glycation[55] and carbamylation[56]. These kinds of modifications are crucial for the biological activities of many proteins and peptides. For example, acetylation is essential to the biological stability and activity of many neuropeptides. Chemically, there are several strategies to achieve N-terminus-directed protein modification *in vitro* because of the unique reactive properties of the NT-amine. Firstly, the lower p*K*a value of NT-amino groups (compare to lysine side-chain ε-amine) can be used in directing acylation reactions to this site through careful control of the reaction pH[57]. However, the presence of a large number of competing lysine residues limits the selectivity of this reaction in most cases. Secondly, strategies have been developed

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

to target the N-terminus in combination with specific amino acid side chains. For example, an aldehyde group can be introduced into the protein N-terminus through the periodate oxidation of the 1,2-aminoalcohol moiety on the NT-Ser/Thr residue[58]. Moreover, the NT-cysteine residues can be modified through thiazolidine formation by using aldehyde reagents[59]. Also, they are able to react with thioesters or thioacids for peptide chemical ligation[60]. Thus far, NT-Ser, Thr, His and Trp, which contain weak-base nucleophiles such as an amine or hydroxyl group on their side chains spatially separated by two atoms from their α-amines, have been found to be most suitable to react with peptidyl thioesters or thioacids to form amide bonds in protein and peptide semi-synthesis[61]. For the NT-tryptophan residues, Pictet–Spengler reactions can also be used for selective protein modification[62].

Perhaps the most useful method for NT-modification is the transamination reaction. This classical transamination reaction represents a potentially general method that can work for almost all NT-amino acid residues. This reaction converts the NT-residue of a protein to a 2-oxoacyl moiety. Therefore, in principle, any protein with a free NT-amine can be transaminated and the resultant carbonyl from the 2-oxoacyl moiety can be used in a subsequent bioconjugation reaction through the formation of a hydrazone, oxime or thiazolidine linkage. However, there is a lack of information on whether and how readily a particular NT-amino acid can be converted to the 2-oxoacyl moiety in the context of a peptide chain as well as the possible influence from the neighbouring residues. Therefore, I have conducted a systematic and extensive investigation of the transamination reaction using a membrane-bound dipeptide library that includes all the 400 possible

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

dipeptide combinations of the 20 genetically encoded amino acids. The details of this study will be discussed in Chapter 2.

1.2.2 CT-Modification

The C-termini of proteins can be modified posttranslationally. For instance, the addition of a lipid anchor to the C-terminus allows the protein to be inserted into a membrane without having a transmembrane domain. Moreover, amidation[63], O-methylation[64] and sumoylation[65, 66] were also reported. However, some proteins or peptides are naturally CT-protected. For example, the melanocyte-releasing hormone α-MSH is naturally protected by carboxy-amidation and is degraded only by endopeptidases *in vivo*[67].

One well-known CT-modification is the expressed protein ligation (EPL)[68]. Expressed protein ligation is a novel protein semi-synthesis method that permits the *in vitro* ligation of a chemically synthesized CT-segment of a protein to a recombinant NT-segment fused through its C-terminus to an intein protein splicing element. The practical convenience of this method makes it well suited for probing the molecular basis of complex processes such as protein-protein interaction.

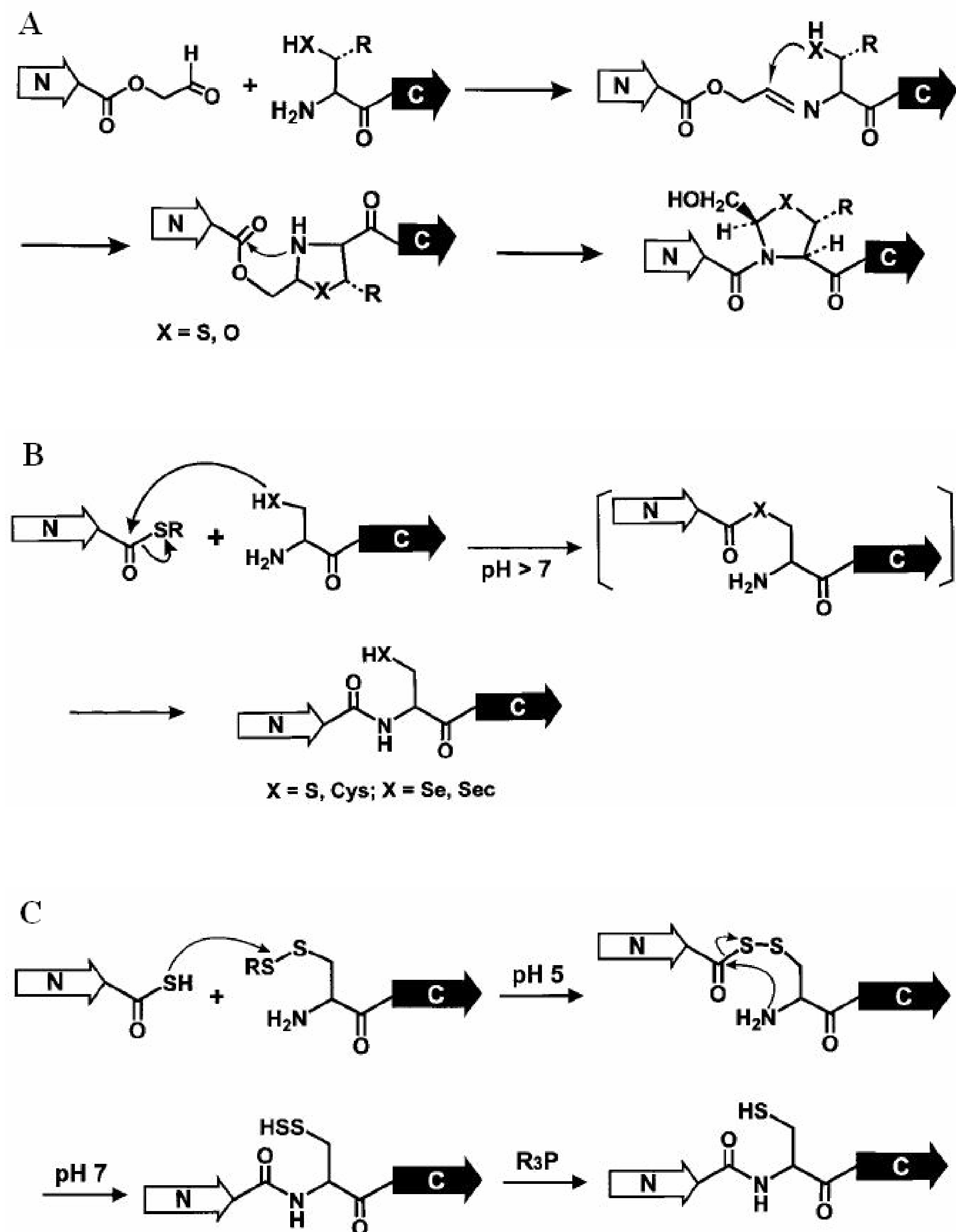
However, site-specific chemical modification methods on C-termini of proteins may encounter some difficulties. This is mainly because the chemical reactivity of the CT-carboxylic acids are almost the same as that of the side-chain carboxyl groups of Asp or Glu residues, which make it difficult to carry out a modification only on the C-termini in the presence of Asp or Glu residues in the targeted natural proteins. Nevertheless, for synthetic or semi-synthetic proteins or peptides, this

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

problem can be solved rather easily by the solid phase synthesis technique. The C-termini can be pre-modified since peptide synthesis is extended from C-terminus to N-terminus. Therefore, several kinds of functional groups can be introduced into the C-termini of peptides or proteins which act as important build blocks for chemical synthesis. To perform chemical ligation, a requirement for the CT-COOH is that it must be an ester, thioester or thioacid (Scheme 1.3 A, B and C, respectively, page 15)[61].

Scheme 1.3 lists the common strategies for orthogonal peptide ligation which consists of two reaction steps. The first step requires a nucleophile (or electrophile) proximally placed at an NT-amine segment and another compatible electrophile (or nucleophile) also proximally located at a CT-segment. The chemo-selective capture of the nucleophile and electrophile pair forms an ester or a thioester and brings the NT-amino group and CT-ester into such a close proximity. Therefore, in the second step a spontaneous intramolecular acyl transfer occurs to form an amide bond.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



Scheme 1.3 Peptide chemical ligation by (A) peptide CT-ester, (B) peptide CT-thioester and (C) peptide CT-thioacid.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Although orthogonal peptide ligation is a powerful method in protein and peptide synthesis, it has some shortcomings. One of them is that the desired cysteine residue has to be part of the produced protein at a suitable place. Of course, homocysteine can replace cysteine with subsequent methylation to form methionine[69]. In addition, the cysteine can also be desulfurized to alanine[70]. Alternately, ligation auxiliaries can be used that mimic an NT-cysteine for the ligation reaction, which can be removed after synthesis.

However, the main limitation lies with the difficulty to prepare peptidyl thioesters or thioacids. Usually, peptidyl thioesters are prepared by Boc chemistry or by Fmoc techniques but with lower yield[71]. As for the peptidyl thioacids, their chemical synthesis remains challenging. The only available approach for solid phase synthesis of peptide thioacids is based on Boc chemistry and no Fmoc-based methods have been developed. The Boc SPPS approach requires the use of a special benzhydryl linker and HF for final deprotection and cleavage. Moreover, it usually suffers from problems of low yield due to side reactions during the cleavage stage where cationic electrophiles generated from the various side chain protecting groups tend to attack the supernucleophilic thioacid.

Because of all these problems, it is essential to develop new methods for the synthesis of peptidyl thioesters and thioacids. For this purpose, this thesis also aims to find new ways to prepare peptidyl thioesters and thioacids. This part of work is discussed in details in Chapter 3 and 4 respectively.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

1.2 Purpose of This Study

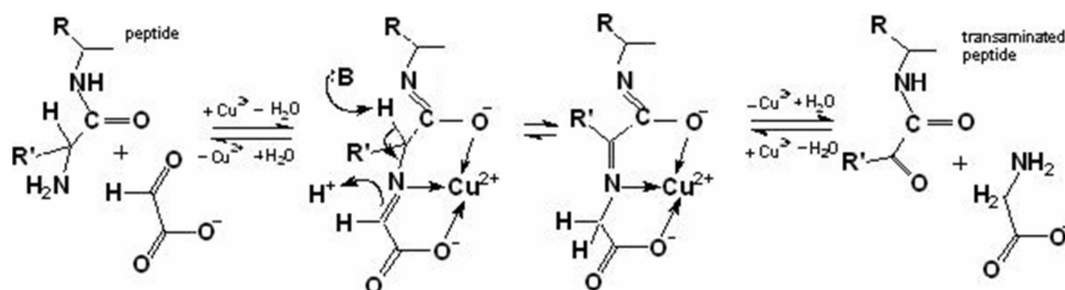
Because almost all peptides and proteins have only one N- and one C-terminus, modification targeting terminal residues is obviously an attractive way to introduce a functional group into a peptide or a protein at a single site as compared to the reactions on the amino acid side chains. This forms the rationale of our study. First, I aim to provide reliable information about the transamination reaction to make future use of this NT-specific protein modification reaction in a much more predictable way. Second, on the peptide CT-modification aspect, this study aims to develop new chemical and enzymatic methods to synthesize peptidyl thioesters and thioacids which are important derivatives to introduce new functional groups onto the C-termini of proteins and peptides.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 2: Systematic Investigation of the Transamination Reaction for N-terminal Selective Modification

2.1 Introduction

The transamination reaction is a highly selective protein modification technique which site-exclusively alters the N-terminus of peptide or proteins. With this property, transamination is more important than some of the other traditional techniques. The classical reaction represents a potentially general method for protein modification. As it requires a free amine alpha to a carboxyamide or carboxyl group, it is expected to be specific only to the NT-amine of a polypeptide chain[72, 73]. Typically, in the presence of glyoxylate and copper (II) or nickel (II) ions in a weakly acidic aqueous buffer, the NT-residue of a protein is converted to a 2-oxoacyl moiety (Scheme 2.1). Therefore, in principle, any protein with a free NT-amine can be transaminated and the resultant carbonyl from the 2-oxoacyl moiety can be used in a subsequent bioconjugation reaction through the formation of a hydrazone, oxime or thiazolidine linkage.



Scheme 2.1 The reaction mechanism of transamination reaction between a peptide and glyoxylic acid. Because of direct participation of the adjacent peptide bond, the transamination reaction is very specific to the N-terminal amino group.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Early examples of transamination included those with *Pseudomonas* cytochrome c-551[74], azurin[75], insulin[76], ribonuclease T_1[77] and carboxypeptidase[78] and a few model peptides[79, 80]. More recently, a biomimetic transamination method was applied to the modification of several proteins by using pyridoxal-5-phosphate as the transaminating reagent without the use of metal ions under very mild reaction conditions[81]. Successful transamination was also demonstrated with some solid-support bound peptides under the typical metal-ion catalysis conditions, although five of the ten NT-residues studied were protected at their side-chain functional groups[73]. Despite these successes, the number of proteins that have been modified by using this reaction remains limited, which clearly does not match up with the seemingly large potential of this rather simple technique. One major issue is that, although a wide range of NT-residues can be transaminated from the above studies, it is often difficult to predict whether and how readily a particular amino acid can undergo transamination in the context of a polypeptide chain. This has prompted us to conduct a systematic study of the transamination reaction using a membrane-immobilized dipeptide library to assess the propensities of each of the 20 amino acid residues to undergo transamination under the influence of the next amino acid residue. This library consists of all the 400 possible dipeptide combinations of the twenty genetically coded amino acids, each represented by one spot on the membrane. After being subjected to the standard treatment conditions with glyoxylate and copper (II) ion, the reactivity of each of the dipeptide spots towards transamination is directly visualized in a colorimetric assay with a dye-modified hydrazide. Our findings, some of which are also verified in solution reactions with a number of model peptides, will have some predictive value when using this technique for N-terminus-specific protein modification in future.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.2 Results and Discussion

2.2.1 Solid Phase Transamination Reaction Assay

The dipeptide arrays were synthesized on Whatman filtration papers using the spot synthesis technique developed by Ronald Frank[82]. At first, three identical dipeptide libraries were prepared for transamination study on solid phase. The arrangement of the library was in a matrix system containing 400 spots, each representing a dipeptide. The vertical letter row represents the NT-amino acid, while the horizontal letter column indicates the neighboring amino acids at the second position. The reaction solution used for transamination was a typical one: 3 M sodium acetate buffer (pH 5.5) containing 5 mM copper sulfate and 0.1 M glyoxylic acid. Two of the membrane-bound dipeptide libraries were incubated in this reaction solution for 1 h and 3 h respectively. In a separate experiment, the third library, as the control group, was treated for 1 h with the same acetate buffer containing 0.1 M glyoxylic acid but no copper (II) ion.

If the NT-amino acid residue is converted to the oxoacyl moiety during the reaction, the ketone or aldehyde (when the NT-residue is Gly) group should react easily with Dabsyl-Ala(β)-hydrazide to form a hydrazone, and the reactivity of each dipeptide was shown by the intensity and the size of the red dot formed. The β-Ala residue in this dye reagent would serve as a good spacer and reduce any steric hindrance that the bulky Dabsyl group might impose on the reacting hydrazide. To check whether Dabsyl-Ala(β)-hydrazide is a suitable reagent to detect membrane bound reactive carbonyls, I first synthesized a test peptide, H-Thr-Ala-Ala(β)-cellulose on a separate membrane. After oxidation by periodate,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

the NT-Thr was converted to oxoacetyl which was then subjected to react with the dye reagent Dabsyl-Ala(β)-hydrazide at 1 mM in 50% DMSO (0.1 M acetate buffer at pH 5.7, 1 h). As a result, the two treated spots was displayed as dark red dots (Figure 2.1B, only black and white pictures are presented in this report) whereas the other two untreated spots showed no color at all (Figure 2.1A).

Figure 2.1 Demonstration of the reactivity of Dabsyl-Ala(β)-hydrazine. A) two unoxidized peptide spots as the negative control; B) two oxidized peptide spots as positive results.

The color stayed on the two spots after extensive washing with organic solvents such as DMF and acidic buffers such as acetic acid solution, indicating covalent bonding of the dye to the membrane. The solvent system used in my study, 50% DMSO in 0.1 M acetate buffer at pH 5.7, was adopted from a previous report by Tam *et al.*,[83] in which they had demonstrated that mixtures of a water-miscible organic solvent such as DMSO and an aqueous buffer at a weakly acidic pH was most suitable for hydrazone formation. The above dipeptide libraries, after the transamination reaction, were treated with the red dye reagent in this solvent system for 22 h at 37 °C and washed. The results are shown in Figure 2.2 A, B and C.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

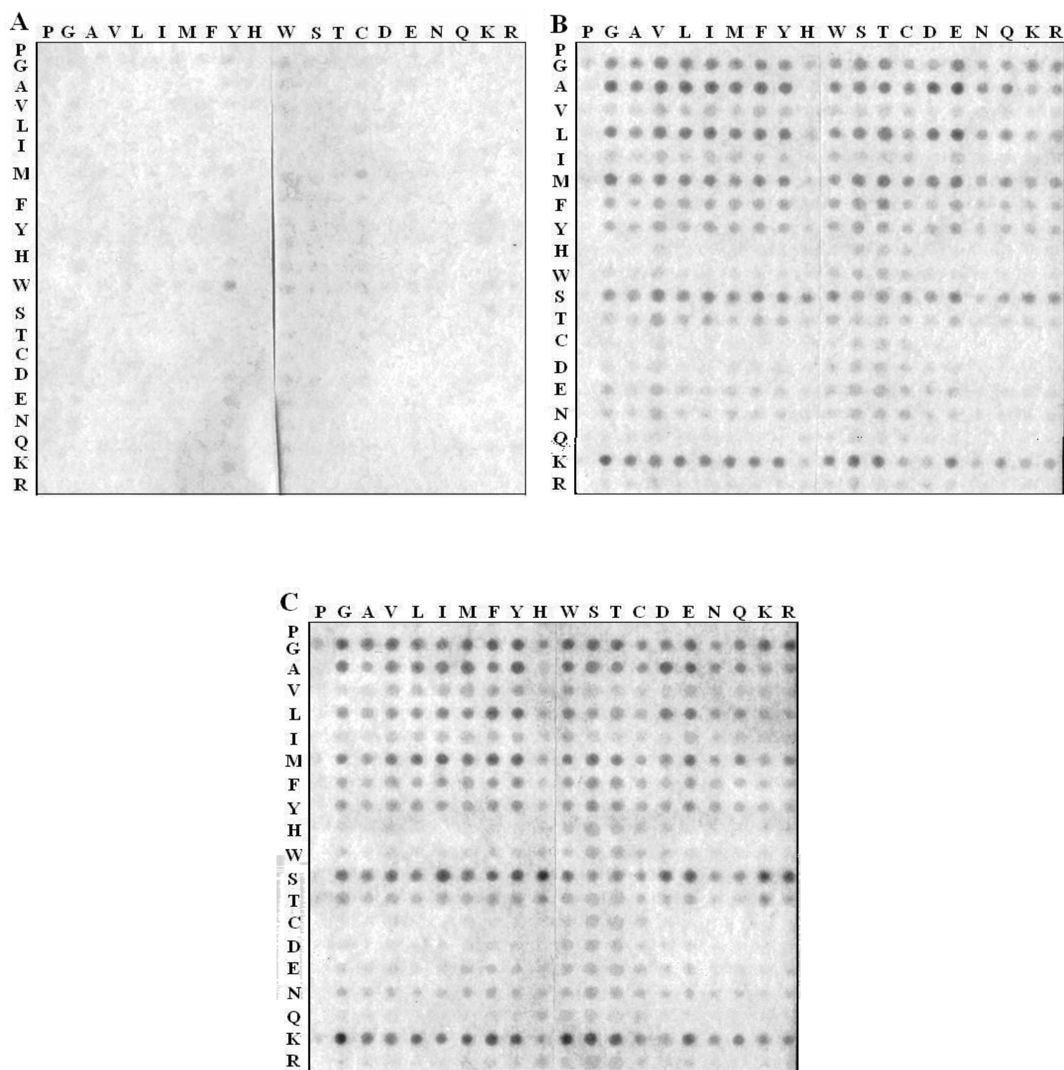


Figure 2.2 Dipeptide membrane libraries to demonstrate the trends of transamination reaction. The intensities of the spots indicate the extent of the reaction of the α-ketocarbonyl dipeptides with the Dabsyl-Ala(β)-hydrazine. A is the negative control; B and C indicate the transamination reaction for 1 h and 3 h respectively. The vertical letter column shows the NT-amino acid while the horizontal letter column indicates the neighboring amino acids (on the second position).

The results are summarized as in the following observations. First, when no cupric ion was added to the reaction solution, there was no transamination under the conditions used since the membrane showed no obviously colored spots (Fig.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.2 A), which also served as a negative control for my experiments. This is consistent with early findings that metal ions such as copper (II) and nickel (II) are required when using glyoxylate as the deamination reagent, as shown in the proposed mechanism (Scheme 2.1, page 18). The few very light spots on the membrane might be due to non-specific hydrophobic binding between the dye molecule and the membrane-bound dipeptides.

Second, when simply comparing the intensities of the colored spots on the membranes treated for 1 and 3 h (Fig 2.2 B and C), the darkest spots were those that displayed dipeptides with these NT-residues: Gly, Ala, Leu, Met, Ser and Lys. These amino acids are either small or contain long, unhindered side chains.

The next group is from these rows: Phe-X, Tyr-X, Val-X, Thr-X and Ile-X, Asn-X, Trp-X and perhaps Glu-X that displayed moderately colored or light-but-visible spots. It is not surprising that, being more hindered, an NT-Thr would react more slowly than the Ser. The aromatic residues Phe-X and Tyr-X also displayed reasonable reactivity in the solid phase transamination reaction. Val-X and Ile-X are expected to be less reactive than the first group, but the color on the spots were nevertheless clearly visible. And in a following experiment, when new strips of membrane containing the Val-X and Ile-X dipeptides were incubated with the transamination reaction solution for 10 h, very dark colors were indeed observed on the spots (Fig. 2.3, page 25).

For the last group, very light or no visible color was seen with the remaining rows of Asp-X, Cys-X, Arg-X, Gln-X, His-X and Pro-X. The row of Cys-X was not dyed in red possibly because the 1,2-aminothiol moiety reacted with the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

aldehyde group of glyoxylic acid for form a thiazolidine ring and it was confirmed by further study (see Figure 2.3, page 25). The apparent low reactivity of the Asn-X, Asp-X, Gln-X and Arg-X shown in the membrane study was possibly due to side reactions accompanying transamination that will be discussed later. Non-reactivity of the His-X was not unexpected as previous studies in solution reactions had shown that an NT-His could not undergo transamination[84]. It is assumed that the copper ion binds to histidine via the NT-amino group and the N-3 of the imidazole side chain to form a six-membered chelate ring[85]. It is also known that a NT-Pro cannot be transaminated due to its secondary amine nature.

We reasoned that the low color intensities on some of the spots might be due to slow kinetics of the solid phase reactions on the membrane as compared to solution reactions because of difficulties for the reagents penetrating the cellulose matrix. So, a new membrane was prepared to further evaluate the rows of dipeptides that displayed low color intensities on the first two membranes (Figure 2.3).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

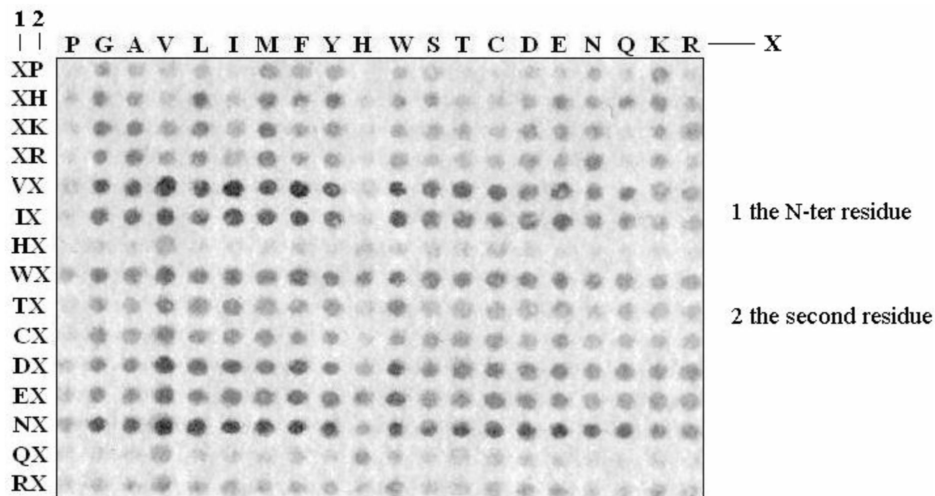


Figure 2.3 Transamination reaction for some selected peptides on a cellulose membrane.

For each dipeptide spot, X represents an amino acid residue, as labeled above, which is at the N terminal position or at the second position. For example, XP indicates a dipeptide in which the second residue is a fixed Pro residue while the first residue is any of the 20 amino acids, such as Pro or Gly or Ala. In contrast, VX indicates a dipeptide in which the first residue, i.e., the NT-amino acid, is a fixed Val residue while the second residue is any of the 20 amino acids, such as Pro or Gly or Ala. Reaction time is 10 h.

This new membrane was treated with the iodoacetamide firstly to block the thiol group of NT-Cys residue and then incubated in the transamination reaction solution for 10 h before reacting with the dye reagent. As expected, after this extended reaction, most of the less colored rows in the previous experiments showed up in much darker red color, especially for the rows of Val-X, Ile-X, Asn-X and Trp-X. The His-X row was still negative. What is puzzling is from the rows of Gln-X and Arg-X, which gave also negative results after the extended treatment. Again, the negative effect of a second Pro or His residue on transamination was confirmed on the new membrane.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Third, when the effect of the second amino acid residue on transamination was assessed, two residues, Pro and His, were found to have a clearly negative impact, as these two columns displayed much lighter color than did the other columns (Figure 2.2, page 22). The Pro column in particular had almost no colored spots, indicating that having a Pro as the second amino acid residue strongly inhibited the reaction (also confirmed by solution phase reaction results of peptide H-**GPEVK-G**-NH_2 in table 2.1, page 28). A possible explanation is that the presence of Pro at the second position would disallow the formation of the copper complex intermediate. As seen from Scheme 2.1 (page 18), enolization of the amide bond, presumably a prerequisite step for the formation of a stable copper complex intermediate, would not be possible for an X-Pro peptide bond because of its tertiary amide nature. Although an alternative copper complex involving the carbonyl oxygen of the keto form amide is also conceivable, the tertiary amide nature of an X-Pro peptide bond would still impose considerable steric hindrance disfavoring its formation. For the X-His dipeptides, a clear *albeit* less drastic inhibitory effect was observed. Except for Ser-His and Thr-His (confirmed by reactions in solution mentioned below, see Fig. 2.11, page 39), all spots on the His column were colored much lighter than the corresponding spots on most other columns. This inhibition effect might be caused by a competing chelate complex that engages the side chain imidazole ring of the His residue and the NT-amine, making it less available for transamination. As in the two exceptional cases of Ser-His or Thr-His, the chelating reaction could just engage the side chain hydroxyl of the Ser or Thr residue, saving the NT-amine for transamination. In addition, Ser and Thr, on the second position, were found to have a possible positive impact (Figure 2.2, page 22). To the best of our knowledge, no studies

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

have been reported on the influence of the second amino acid residue on transamination.

2.2.2 Transamination Reaction Assay in Solution

The above results revealed interesting trends on the reactivity of the dipeptides towards transamination. However, the solid phase method employed has its limitations. It can only detect the presence of ketone or aldehyde group reactive towards the dye-modified hydrazide reagent. It cannot reveal any side reactions nor can it reveal the nature of the side products. The transamination reaction under the conditions used is known to be prone of side reactions especially with some amino acids at the N-terminus. To verify my solid phase results and further study the peptide transamination reactions, a number of model peptides were synthesized for the study of transamination in solution. All the peptides were subjected to transamination reaction under the same condition, i.e., in 3 M acetate buffer at pH 5.5 containing 0.1 M glyoxylic acid and 5 mM $CuSO_4$. The results are summarized in Table 2.1.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Peptide	0 h		1 h			10 h		
	RT [a]	MS (m/z)	RT [a]	MS (m/z)	RP [c]	RT [a]	MS (m/z)	RP [b]
FEVKG	16.8	577.3	19.5	596.2	60	19.6	596.2	3
FFEVKG	22.3	725.4	26.5, 31.1	724.6, 780.3	0	31.1	780.5	0
PFEVKG	17.1	675.3	19.3	731.2	80	19.3	731.2	10
QLAEVKG	14.9	743.4	18.5, 21.2	742.6, 786.4	10	18.5, 21.1	742.5, 786.4	3
RTAEVKG	9.2	759.6	9.6	802.5	0	9.6	802.6	0
GFEVKG	16.8	635.2	18.5	652.0	60	18.7	652.0	50
TFEVKG	16.8	679.3	19.5	696.2	30	19.6	696.2	0
CFEVKG [c]	16	738.3	19.3, 21.2	781.4, 737.6	0	19.3, 21.2, 22.2	781.4, 737.5, 793.4	0
IYAEVKG	16.2	778.5	29.1	777.7	0	29.1	777.6	0
LYAEVKG	17	778.5	29.6	777.6	0	25.1, 29.7	851.5, 777.6	0
VFEVKG	18.6	677.4	33.3	676.3	60	33.2	676.4	0
VGFEVKG	20	734.5	29.4	733.6	35	29.4	733.6	0
VRLQAG	15.6	642.8	24.4	641.7	55	24.4	641.8	0
VVFEVKG	22.4	776.8	28.9, 36.3	849.5, 775.8	25	29.0, 36.3	849.5, 775.8	0
NFEVKG [d]	11.5	692.3	11.1,12.6	766.4,765.3	10	11.1,12.6	766.4,765.3	10
CFEVKG [d]	11.8	681.2	13.3,14.1	737.4, 737.4	0	13.3,14.1	737.4, 737.4	0
WFEVKG [d]	16.2	764.3	17.2,17.6	820.3	0	17.2,17.6	820.3	0
YFEVKG [d]	12.9	741.3	14.5,15.3,15.5	814.5,814.4,796.5	0	14.7,15.5	814.5,814.3	0
DFEVKG [d]	11.8	693.7	13.7,16.3	722.3,648.4	0	13.7,16.3	722.3,648.4	0
GPEVKG [d]	7.5	585.6	7.0	659.6	0	7.0	659.6	0
RFG [d]	8.8	378.4	10.2, 12.3	377.2,377.3,	30	10.2, 12.3	377.2,377.3,	22
KFG [d]	8.2	350.2	10.7	387.1	15	10.7	387.1	0

Table 2.1 MS and HPLC data of starting peptides and their main products after transamination reaction. All peptides are with C-ter-carboxyamide. Detailed experimental conditions are described in the Experimental Section.

[a] RT = retention time (min)

[b] RP = remaining percentage (%) of starting material peptide after reaction based on the remaining peptide peak area detected by RP-HPLC

[c] CFEVKG was treated by iodoacetamide before carrying out the transamination reaction

[d] HPLC gradient: 0-40%B in 20 min

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

From Table 2.1, one can make the following observations.

1. For peptides having Thr or an aliphatic amino acid residue (Leu, Val or Ile) at the N-terminus, the transamination product was obtained as the major product. Figure 2.4 (page 30) shows HPLC chromatograms of reactions of two peptides (TFEVKG and VGFEVKG). From previously published work and the solid phase data above, one can also deduce that Ser and other amino acids with non-functionalized aliphatic side chains would give similar patterns. These NT-amino acid residues, namely Gly, Ala, Leu, Ser, Thr, Val, Ile and Met, give straightforward transamination reaction with minimum side reactions. It is noteworthy that the solution reaction displayed a much faster reaction kinetic than did the solid phase reaction on membranes. For example, both Val and Thr are β-branched amino acids and the Val-X and Thr-X dipeptides were slow in transamination on the membrane; however, all the NT-Thr and Val model peptides were converted to desired products in substantial amount after 1 h reaction in solution.

For the glycinyl peptide, GFEVKG, no significant difference in conversion percentage was found between 1 h and 10 h reaction; apparently the reaction could proceed rapidly but an unfavorable equilibrium was also reached quickly, preventing further conversion of the starting material.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

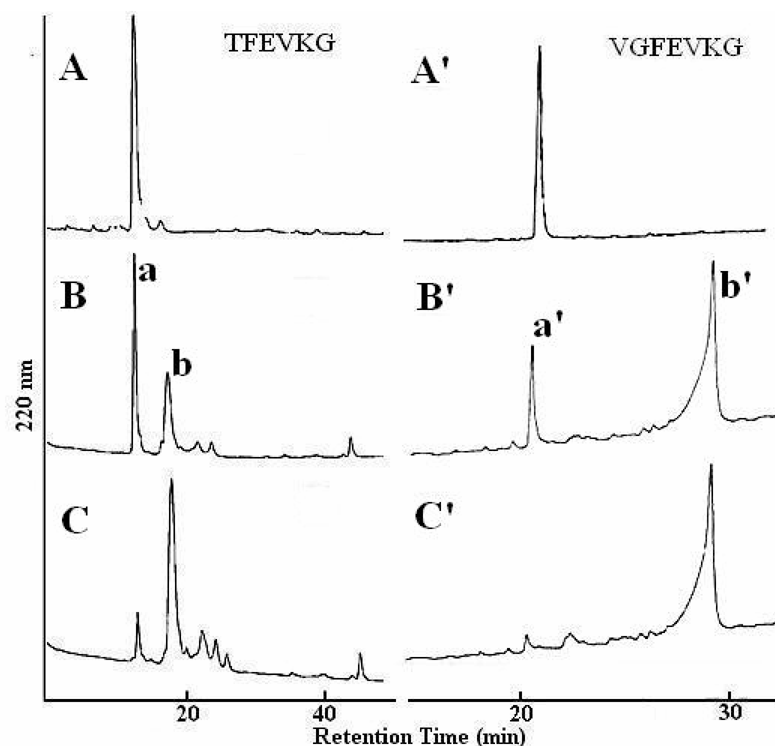


Figure 2.4 RP-HPLC analysis of transamination reactions of TFEVKG and VGFEVKG, with glyoxylic acid in 3 M acetate buffer, pH 5.5, at 23 ^{o}C. The concentration of the peptide is 1 mM, Cu^{2+} is 5 mM and glyoxylic acid is 0.2 M. **A** and **A'** indicate 0 h reaction as a control. **B** and **B'** indicate 1 h reaction. **C** and **C'** indicate 10 h reaction. Peptide starting materials are displayed as peak a and peak a' while the transamination products are displayed as peak b and peak b' for TFEVKG and VGFEVKG respectively. HPLC gradient: 0-40%B in 40 min.

2. For the peptide FFEVKG, the reaction was very fast, as after 1 h almost all starting material had disappeared; however the reaction was very complicated. At 1 h, HPLC analysis revealed many peaks as the intermediates with the transamination product being only a minor one and most of the peaks could not be identified. Interestingly all the intermediate peaks disappeared over time and after 10 h, only one major product was left, which had a molecular weight 56 Da higher than the starting material. To simplify the characterization of this product I synthesized a smaller tripeptide with the sequence H-Phe-Phe-Gly-NH_2 and

subjected it to transamination. A similar chromatogram was obtained (Figure 2.5) and again a major product with a +56 mass was obtained after 12 h incubation.

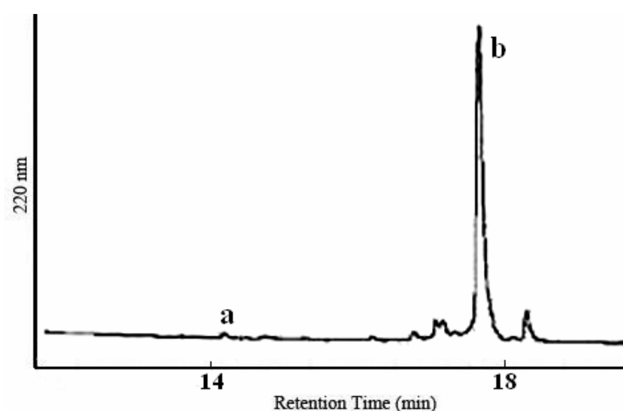


Figure 2.5 RP-HPLC analysis of transamination reaction of H-**FFG**-NH_2.

The reaction condition was the same as in Figure 2.4 except that the reaction was incubated at 37 ^{o}C for 12 h. The concentration of the peptide H-**FFG**-NH_2 was 1 mg/mL. Peak a is the remaining H-**FFG**-NH_2 and peak b is the product. HPLC gradient: 0-40%B in 20 min.

Preliminary NMR analysis showed that this product might have a lactone structure (Figure 2.6, 2.7).

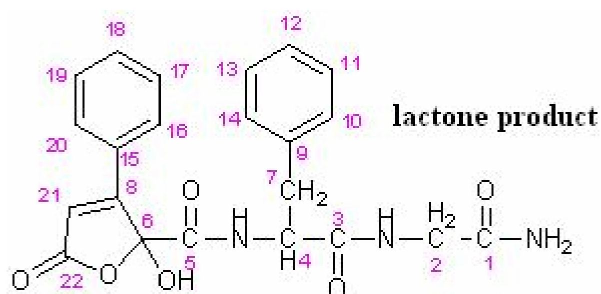


Figure 2.6 Possible structure of the lactone product (Figure 2.5 peak b) derived from H-**FFG**-NH_2 under the transamination react

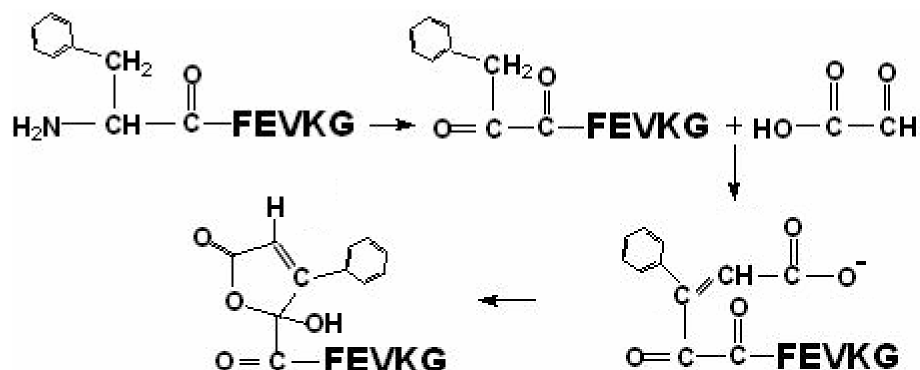
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Figure 2.7 NMR spectrum of the lactone product (Figure 2.5 peak b). A) ^{1}H NMR spectrum; B) ^{13}C NMR spectrum; C) DEPT NMR spectrum; D) ^{1}H-^{1}H COSY NMR spectrum.

^{1}H and ^{13}C NMR spectra were recorded on a Bruker DRX400 400MHz spectrometer at ca. 296 K and operating frequencies of 400.0 MHz. Proton chemical shifts (δ) is quoted in ppm to high reference frequency (0 ppm) of $SiMe_4$.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

With the lactone product structure derived from the NMR analysis, the possible reaction mechanism is displayed as below:



Scheme 2.2 Possible reaction pathway of FFEVKG under transamination reaction conditions.

From these observations, it appears that an NT-Phe gave very fast transamination reaction but the product formed was very unstable. This might be due to the fact that the methylene between the phenyl and the newly formed ketone is quite acidic and could engage in cross aldol addition reactions with glyoxylic acid.

3. Peptides that begin with Arg, Gln and Cys (side chain modified with iodoacetamide) displayed similar reaction patterns. When the heptapeptide QLAEVKG was subjected to the transamination reaction, at 1 h, HPLC analysis showed almost complete disappearance of the starting material and the appearance of several new peaks. A minor peak was found to be the transamination product as checked by ESI-MS. The main peak had a MW 43 Da higher than the starting material and persisted after 10 h reaction (Table 2.1, page 28). In a recently published study, it was indicated that a solid phase reaction of an NT-Gln with side

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

chain protected by trityl gave the transamination product as the major product in 6 h reaction[84]. It appears that the free carboxyamide group on the side chain of Gln was the cause of this side reaction. Similarly, several peaks were found after transamination reaction on RTAEVKG. One small peak whose molecular weights corresponded to the desired transamination product and the major one also had a mass change of +43 Da as compared to the starting material (Table 2.1, page 28). For CFEVKG, whose thiol group was blocked by using iodoacetamide, the transamination reaction again gave rise to several HPLC peaks after 10 h reaction, and one small peak was the desired product by ESI-MS analysis and the highest peak was again the +43 Da product (Table 2.1, page 28). A common feature of these amino acid residues was that they all contain a O=C-NH_2 or NH=C-NH_2 group. The exact nature of the +43 Da product was not determined at this moment.

4. Unlike Gln-X, the reactivity of Asn-X, Asp-X and Glu-X was relatively high (Fig. 2.3, page 25), and the reactions were complicated as well as with several unidentified side products. It is indicated that NT-Asp can have several products upon the transamination reaction[86]. The possible explanation is that after transamination, an active methylene is formed and due to the fact that there are strongly electron-withdrawing groups on each side of this methylene group, it can therefore lose a proton and combine with glyoxylate in an aldol condensation type of reaction.

5. The negative results with Pro-X and His-X on the membrane were confirmed by the transamination reactions in solution[84]. For HFEVKG, no transamination reaction occurred even after 10 h reaction. This further confirms that the NT-amino group of His was protected by the cupric ion in a six-membered chelate ring[85]

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

engaging the N-3 of the imidazole side chain. Interestingly, PFEVKG reacted slowly in the reaction buffer and after 20 h was almost completely converted to a product that gave a molecular ion 56 Da higher than the $[M+H]^+$ ion of the starting material (Table 2.1, page 28). The same finding was seen with a small tripeptide, PFG (Figure 2.8). NMR analysis indicated that the product, as also suggested by others, was a stable pyrrolidinium formed between the pyrrolidine ring of Pro and the aldehyde group of glyoxylic acid through dehydration (Figure 2.9 and 2.10, page 36 and 37, respectively).

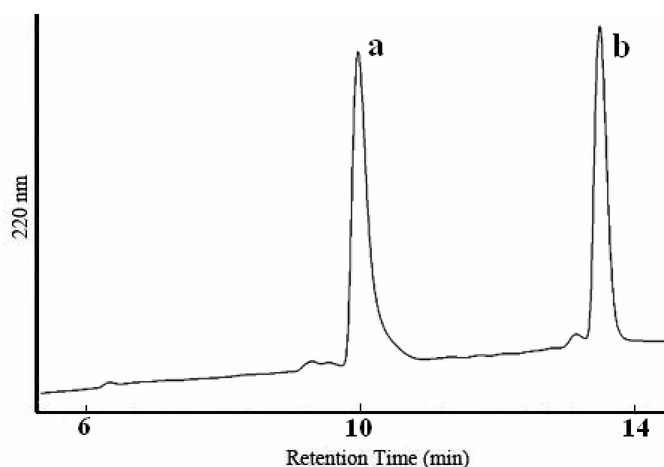


Figure 2.8 RP-HPLC analysis of H-**PFG**-NH_2 under transamination reaction conditions.

Reaction conditions were the same as in Figure 2.4 except that reaction time was 6 h. The concentration of the peptide was 1 mg/mL. Peak a is the H-**PFG**-NH_2 and peak b is the product. HPLC gradient: 0-30%B in 20 min.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

A

B

C

D

Figure 2.9 NMR spectrum of the pyrrolidinium product (Figure 2.8 peak b). A) ^{1}H NMR spectrum; B) ^{13}C NMR spectrum; C) DEPT NMR spectrum; D) ^{1}H-^{1}H COSY NMR spectrum.

^{1}H and ^{13}C NMR spectra were recorded on a Bruker DRX400 400MHz spectrometer at ca. 296 K and operating frequencies of 400.0 MHz. Proton chemical shifts (δ) is quoted in ppm to high reference frequency (0 ppm) of $SiMe_4$.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

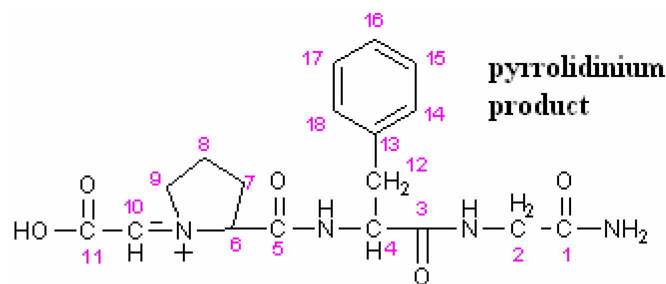
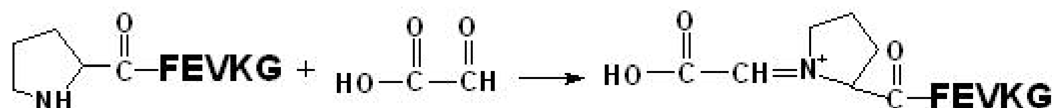


Figure 2.10 Possible structure of the product (Figure 2.8 peak b) derived from H-**PFG**-NH_2 under the transamination reaction condition.

Data of the ^{1}H NMR (D_2O): 7.35-7.18 (m, 5 H, Ph), 4.56 (t, 1 H, 6-H), 4.49 (t, 1 H, 4-H), 3.96-3.80 (m, 3 H, H-9, 10), 3.50-3.22 (m, 4 H, H-2, 12), 2.29-1.88 (m, 4 H, H-7, 8). Data of the ^{13}C NMR (D_2O):174.0 (C-11), 170.7 (C-3), 170.1 (C-5), 166.0 (C-1), 136.7 (C-11), 129.4 (Ph), 129.1 (Ph), 129.0 (Ph), 128.8 (Ph), 127.7 (Ph), 76.0 (C-10), 65.0 (C-6), 59.3 (C-4), 53.3 (C-2), 42.2 (C-9), 32.5 (C-12), 26.5 (C-7), 24.1 (C-8)

With the pyrrolidinium product structure derived from the NMR analysis, the possible mechanism is displayed as below:



Scheme 2.3 Possible reaction pathway of PFEVKG under transamination reaction condition.

6. We also checked the influence of the next amino acid residue in solution reaction. From the solid phase study on membranes, a His residue at the second position was found to have a strong retarding effect with the exception of H-**Ser-His-** and H-**Thr-His-**. Therefore, three peptides, H-**AHAL**-NH_2, H-**SHAL**-NH_2 and H-**THAL**-NH_2, were synthesized to study the transamination

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

reaction in solution. After transamination reaction and analysis by HPLC, the results agreed well with those from the solid-membrane reactions. From Fig. 2.11 (page 39), one can see that all three peptides can undergo the transamination reaction but with clear differences. For H-**SHAL**-NH_2 (upper trace in Fig. 2.11), the main product was the transaminated peptide, while for H-**THAL**-NH_2, the transamination reaction took place but at a slower rate (Fig. 2.11, middle trace). We reasoned that the latter one encountered steric hindrance because of the β-branched side chain of Thr. On the other hand, the reaction of H-**AHAL**-NH_2 was quite different with the above two. In this case, although the starting material was consumed very fast, the product peak was a broad hump (Fig. 2.11, lower trace). And MS data revealed the presence of a small amount of transaminated product but the exact nature of the major content of this hump could not be determined at this moment. One possible explanation is that this hump might be due to a chelate complex that engages the side chain imidazole ring of the His residue and the NT-amine, preventing it from undergoing transamination. While for the Ser-His or Thr-His, the chelating reaction could just engage the side chain hydroxyl of the Ser or Thr residue, saving the NT-amine for transamination.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

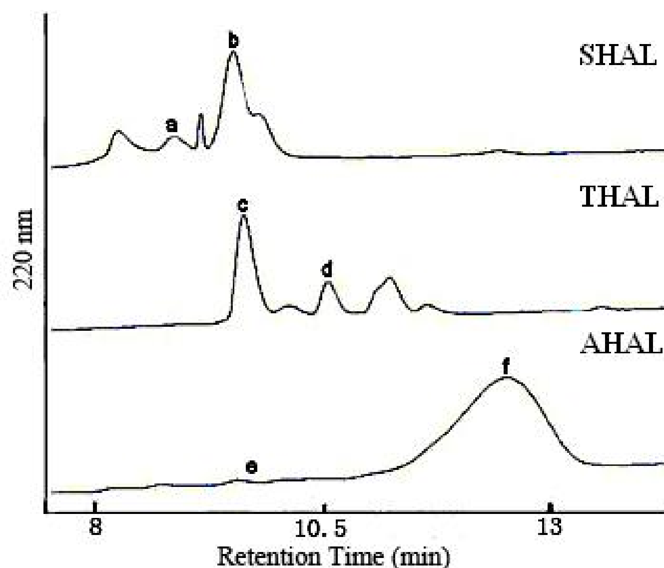


Figure 2.11 RP-HPLC analysis of H-**AHAL**-NH_2, H-**SHAL**-NH_2 and H-**THAL**-NH_2 under transamination reaction conditions.

The reaction condition is the same as that in Figure 2.4 except the reaction time was extended to 6 h. Peak a is H-**SHAL**-NH_2 (m/z $[M+H]^+$ found: 426.0), and peak b is the hydrated transaminated product (m/z $[M+H]^+$ found: 443.1). Peak c is H-**THAL**-NH_2 (m/z $[M+H]^+$ found: 440.1), and peak d is the transaminated product (m/z $[M+H]^+$ found: 439.2). Peak e is the H-**AHAL**-NH_2 (m/z $[M+H]^+$ found: 410.0), and peak f (m/z $[M+H]^+$ found: 409.0, 447.1), contains the transaminated product. HPLC gradient: 0-30%B in 20 min.

2.2.3 An Application Example of the Transamination Reaction.

With all the information discussed above, one can approximately predict the transamination reaction tendency for a given peptide. For example, I speculated that the peptide somatostatin, H-**AGCKNFFWKTFTSC**-NH_2, would be a perfect substrate for transamination reaction because the first two NT-residues (Ala-Gly) are both good for the transamination reaction. Therefore, somatostatin was synthesized using Fmoc SPPS. The somatostatin (m/z $[M+H]^+$ found: 1638.8, M_w

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

calcd: 1637.7), was dissolved in the transamination buffer to reach a final concentration of 5 mM. HPLC detection showed that the reaction went very fast and after only 30 min the starting material was almost completely converted to transaminated somatostatin (m/z $[M+H]^+$ found: 1637.7, M_w calcd: 1636.7). If the incubation was allowed for a longer time, such as 5 h, a new product appeared and it was the oxidized transaminated somatostatin which contain an intermolecular disulfide bond (m/z $[M+H]^+$ found: 1635.7, M_w calcd: 1634.7). This new product was purified and redissolved into 0.4 M sodium acetate buffer, pH 5.5, and 4 equiv. of biotin-hydrazine was added. Incubation was carried out at 37 ^{o}C for 25 h and finally the biotinylated oxidized somatostatin was generated (m/z $[M+H]^+$ found: 1876.8, M_w calcd: 1875.7).

2.3 Summary and Conclusion

I conducted a systematic solid phase study on the relative reactivity of membrane-bound dipeptides towards transamination by comparing the color intensities of the dyed spots. This was further complemented by a series of solution reactions performed on selected model peptides. My study revealed the complicated nature of the transamination reaction. Nevertheless, some interesting observations are noteworthy. Except for Pro and His, most amino acids at the N-terminus can readily undergo transamination. However, the nature of the side chains greatly influences the stability and therefore the destiny of the newly formed 2-oxoacyl functionality in the reaction milieu. For amino acids with non-functionalized aliphatic side chains and for Ser, Thr and Met, the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

transamination products are relatively stable and can be isolated as the major products. All other amino acids seem to generate a product that, to varying degrees, is unstable in the used reaction milieu and may lead to various side products, many of which are unidentified. This highlights the need for developing new transamination techniques that use mild reaction conditions and therefore cause less side reactions. The effect of the neighboring residues was also investigated, with the most interesting observation being that a Pro or His at position 2 would strongly inhibit the reaction. Our study will provide useful guidelines for the future use of transamination for protein modification and bioconjugation applications.

2.4 Experimental

2.4.1 Materials

Fmoc-Amino acids, Rink-amide-MBHA-resin, DCC (dicyclohexylca -rbodiimide), 2-chlorotrityl chloride resin, PyBOP (Benzotriazole-1-yl -oxytripyrrolidinophosphonium hexafluorophosphate), DMAP (4-dimethylamino -pyridine) and DMSO (dimethyl sulfoxide) were all obtained from GL Biochem (Shanghai, China). DMF (N, N'-dimethylformamide), DCM (dichloromethane), piperidine, TFA (trifluoroacetic acid), TIS (triisopropylsilane), DIEA (diisopropylethylamine), sodium acetate, methanol, hydrazine hydrate, triethylamine and dioxane were obtained from Merck. 1,3-diaminopropane, Dabsyl-Cl [4-(dimethylamino)azobenzene-4'-sulfonylchloride] and copper (II)

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

sulfate were purchased from Aldrich. All the commercial reagents and solvents were used as received.

2.4.2 High Performance Liquid Chromatography.

RP-HPLC was performed on an Agilent HPLC system with a Vydac C_{18} reverse-phase column (250 mm×4.6 mm, 5μm). The analysis was carried out using the mixture of two solvents – buffer A and buffer B as the mobile phase. Buffer A was deionized water containing 0.045% TFA and buffer B was 90% acetonitrile in deionized water (0.04% TFA). The analysis was carried out in a gradient of 0-40% of buffer B in buffer A over 40 min unless mentioned otherwise, using UV detection at 220, 254 and 280 nm. The mobile phase flow rate was 1 ml min^{-1} and the separation temperature was 23 ^{o}C.

2.4.3 Mass Spectrometry

ESI-MS spectra were obtained at the NTU SBS Mass Spectrometry Core using the Finnigan LCQ Deca XP MAX.

2.4.4 NMR Spectroscopy

NMR data were obtained at the NTU SBS NMR Core using the 400 MHz NMR machine from Bruker.

2.4.5 Preparation of Dipeptide Library

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

An 8×8 cm sheet of cellulose paper (Whatman #1 chromatography paper) was treated with the symmetrical anhydride of Fmoc-β-alanine (6 mmol), DCC (3 mmol) and DMAP (0.4 mmol) in 25 mL DMF at room temperature for 3 days[87]. Washed by DCM, DMF and methanol, the paper was treated with 20% piperidine in DMF for 1 h to remove the Fmoc group and the piperidine was washed 5-8 times until the wash was bromophenol blue negative. The paper was washed over methanol and dried. Then it was used for the 'spot-synthesis' of the library containing 400 dipeptides according to standard Fmoc machine protocols using a MultiPep synthesizer from Intavis Bioanalytical Instruments AG. A capping step was included after the first coupling step by immerging the membranes in 10% Ac_2O solution in DMF. When the synthesis was completed, the side chains were removed with 82.5% TFA, 5% phenol, 2.5% ethanedithiol and 5% H_2O for 2 h. The paper was washed with DCM, DMF and methanol, and air dried.

2.4.6 Preparation of Dabsyl-Ala(β)-hydrazide

2-chlorotrityl resin, 1 g (1.6 mmol/g), was soaked in dry DCM for 20 min. Hydrazine hydrate (16 mmol) and 10 ml dry DCM were added into the soaked resin and shaken for 1 h. Then the resin was washed by DCM to remove excess hydrazine. Fmoc-Ala(β)-OH was coupled to the resin by the standard Fmoc chemistry. After removing the Fmoc group and washing carefully, the resin was soaked in DMF/DCM (1:1) and Et_3N (1 equiv.) and Dabsyl-Cl (4 equiv.) were added and then was permitted to tumble overnight. The products were cleaved from resin by 50% TFA in DCM (v/v) for 2 h and purified by HPLC using a C_{18} column and checked by ESI-MS (m/z $[M+H]^+$ found: 391.1, M_w calcd: 390.1).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2.4.7 Peptide Synthesis

The peptides were synthesized using standard Fmoc chemistry[88] on Rink-Amide-MBHA-resin. The ratio of Rink-Amide-MBHA-resin to Fmoc amino acids, PyBOP and DIEA was 1:4:4:8. Before use, the Rink-Amide-MBHA-resin was soaked in DMF to swell for 20 min. Coupling reaction was carried out in DMF for 45-60 min. Fmoc group was removed by 20% piperidine (in DMF) for 30 min. After TFA cleavage, the peptides were purified by HPLC on a C_{18} reverse phase column (Vydac, Hesperia, CA) and characterized by ESI-MS.

2.4.8 Transamination Reaction of the Membrane-bound Dipeptide Library

The membranes displaying the dipeptide libraries were incubated in 3 M sodium acetate buffer (5 mM copper sulfate, 0.1 M glyoxylic acid, pH 5.5) for 1 h or 3 h. Another membrane was incubated in a plastic tray containing 20 mL of 5 mM iodoacetamide in 0.1 M sodium acetate (pH 5.6) for 5 min at RT and washed carefully by 25 mM sodium acetate buffer (pH 5.5). Dried with methanol, the membrane was incubated in transamination buffer for 10 h. The control one was incubated in the same transamination reaction solution for 1 h but without the addition of copper (II) ion. All reactions were stopped by the addition of 10 mM EDTA then the dipeptide libraries were washed by 25 mM sodium acetate buffer, pH 5.5.

2.4.9 Reaction of Transaminated Dipeptide Library with Dabsyl-Ala(β)-hydrazine

Dabsyl-Ala(β)-hydrazide was dissolved in 50% DMSO/0.1 M sodium acetate (pH 5.7) to reach a concentration of 1 mM. The transaminated dipeptide libraries

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

were treated with this Dabsyl-Ala(β)-hydrazide solution for 22 h at 37 °C. Then the membrane was washed carefully by DMF and acetic acid solution (pH 2.4) until the background is clean. The color development was monitored visually and recorded by photography.

2.4.10 Transamination Reactions of Model Peptides in Solution

Transaminations were achieved by treatment of peptides with a solution of glyoxylic acid in the presence of copper (II) sulfate and sodium acetate at pH 5.5 according to literature procedures with some modifications[84]. Transamination reactions were conducted with a freshly prepared aqueous solution of 0.2 M glyoxylic acid and 5 mM copper (II) sulfate pentahydrate in 3 M sodium acetate buffer, pH 5.5. The concentration of peptide was 1 mM. This reaction was carried out at room temperature (23 ^{o}C) and monitored by analytical RP-HPLC at different time intervals.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 3: Development of an Enzymatic Method for the Synthesis of Peptide C^{α}-thioesters

3.1 Introduction

A thioester functional group at the C-terminus of a protein or peptide is useful for protein or bioconjugate synthesis. Therefore, thioester formation is considered as one kind of CT-modification reaction. The thioester functionality can react with the 1,2-aminothiol moiety of an NT-Cys-peptide to form a native peptide bond[60, 89, 90] or be activated selectively with a heavy metal ion for coupling to an amine nucleophile[91, 92] which can be further used to introduce new function groups to the C-termini of peptides or proteins.

Traditionally, peptide C^{α}-thioesters are prepared by solid phase peptide synthesis using Boc chemistry[90, 93], although new methods have also been developed recently to allow for direct or indirect peptide thioester preparation using Fmoc chemistry with some success[94-96]. In my thesis here, I introduce a new enzymatic method which provides an alternative route for the synthesis of peptide thioesters from peptide esters that can be readily prepared with Fmoc chemistry[97].

Proteases have long been utilized as biocatalysts for organic transformations. For instance, exploitation of the native hydrolase activity of proteases has led to the development of new methods for kinetic resolution of racemic carboxylic acids or alcohols and for the processing of recombinant protein precursors by site-specific proteolytic cleavage.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Herein, our aim is to develop a novel enzymatic method to synthesize peptide thioesters. The serine proteases were chosen as the target enzymes due to their special biochemical properties. First of all, I introduce some basic knowledge about serine proteases.

For a serine protease, Ser, His and Asp residues form a hydrogen bonding system often referred to as the catalytic triad[98] and the backbone amide nitrogens of Ser and Gly residues compose the oxyanion hole[99]. For subtilisin, its oxyanion hole contains one backbone amide nitrogen of Ser-221 and the side chain of Asn-155[100]. These oxyanion holes are important for the stability of negatively-charged intermediates[101].

Typically, S_1, S_2, *etc.* are used for numbering the substrate binding subsites, proceeding from the scissile peptide bond to the N-terminus and S'_1, S'_2, *etc.* from the scissile peptide bond toward the C-terminus. The corresponding residues of the substrate are named P_1, P'_1, *etc.*[102]. The catalyzed cleavage reaction is between P_1 and P_1'. The understanding of optimized substrate sequence is essential for the enzymatic reaction design. For example, in my thesis, the peptide ester, Ac-His-Ala-Ala-Pro-Phe-glc-Phe-Gly-NH_2, was chosen to act as the model substrate for the target enzyme, subtiligase[103].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

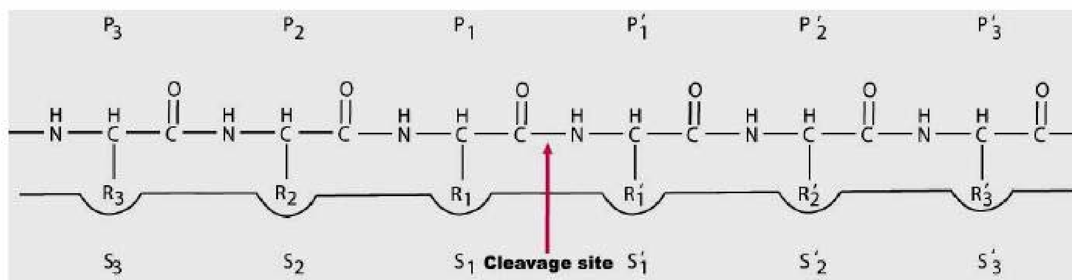


Figure 3.1 Schematic diagram of the subsites in a typical serine protease. S_1 and S_2 *etc.* are used for numbering the substrate binding subsites of enzymes. The corresponding residues of the substrates are named P_1 and P_2, *etc.*

The catalytic process contains two main steps: Enzyme acylation and subsequent deacylation. Enzyme acylation: a Ser residue attacks the substrate peptide bond; the first tetrahedral intermediate is formed and collapses to form the acyl-enzyme intermediate and releases the CT-part of the substrate polypeptide. Enzyme deacylation: the hydrolytic water molecule attacks the acyl-enzyme intermediate; the second tetrahedral intermediate is formed, which collapses and releases the NT-part of the substrate polypeptide.

The enzyme-substrate complex, termed as the Michaelis complex[99], was proposed as a common feature of all enzyme reactions to explain the phenomenon of saturation kinetics[104]. Figure 3.2 displays how subtilisin hydrolyzes a peptide substrate.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

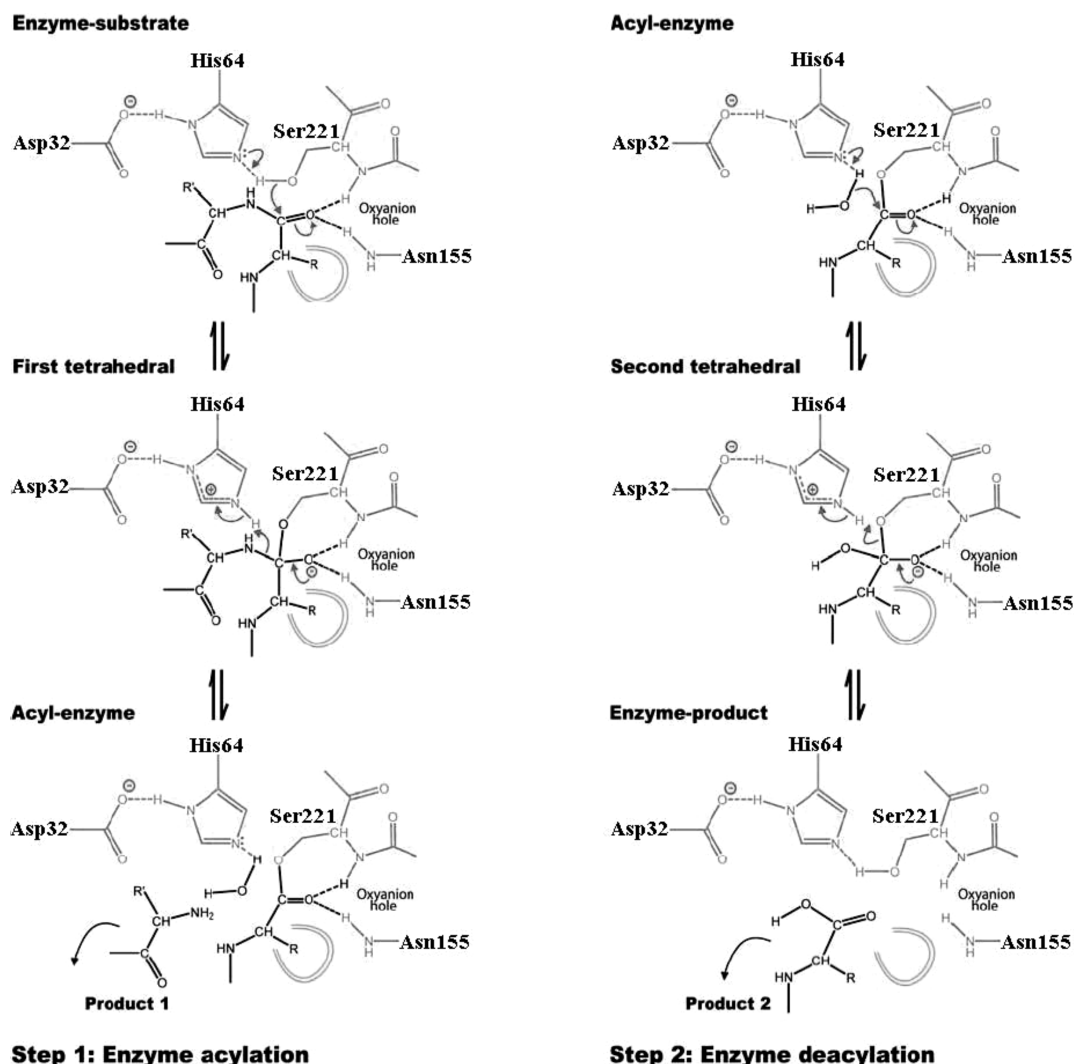
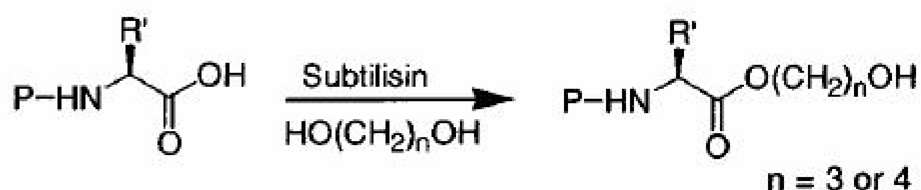


Figure 3.2 The two steps for displaying the catalytic mechanism of subtilisin: 1, enzyme acylation: attack of Ser-221 to the substrate peptide bond; first tetrahedral intermediate formation and collapse to form the acyl-enzyme intermediate and the release of the CT-part of the substrate polypeptide. 2, enzyme deacylation: attack of the hydrolytic water molecule onto the acyl-enzyme intermediate; second tetrahedral intermediate formation, collapse and release of the NT-part of the substrate polypeptide.

Proteases have long been utilized as biocatalysts for organic transformations. The most useful synthetic value of a protease, however, resides in its ability to catalyze the reverse hydrolytic or non-hydrolytic reactions. From figure 3.2, we can see that in the acyl-enzyme state of subtilisin, the water molecular was

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

deprotonated by the His64 and subsequently attacked the acyl-enzyme intermediate to form the second tetrahedral intermediate which finally collapsed to release the hydrolytic product. Therefore, if a nucleophile other than water is engaged in the enzyme deacylation step, other desired products should be acquired. With the understanding of the above mechanism, a good design[105], created by C. F. Liu and J. P. Tam, demonstrated how to use the subtilisin to act as a ligase to perform synthesis work (Scheme 3.1).



Scheme 3.1 Alcohol acts as the nucleophile in the ester bond formation reaction in the presence of subtilisin. P represents a peptide chain.

They reported that the bacterial serine protease subtilisin Carlsberg retained high catalytical activity in two aliphatic diols, 1,3-propanediol and 1,4-butanediol. In the presence of this enzyme, Boc-amino acids and unprotected peptides, solubilized in 1,3-propanediol or 1,4-butanediol containing 1-2.5% H_2O, were selectively esterified at their CR-COOH to the corresponding diol mono-esters in excellent yields[105].

In principle, if the alcohols are replaced by other nucleophiles such as thiols, the products should be thioesters. However, I tried to generate thioesters using

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

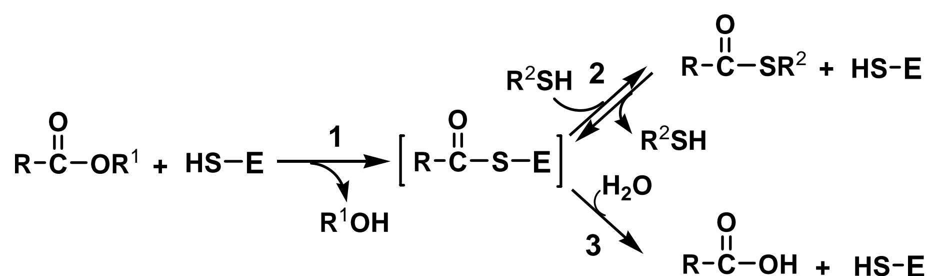
subtilisin but failed. As a result, to realize this goal, the crucial step is to choose a suitable protease to catalyze this reaction.

The nucleophile acts as the acceptor of the acyl group from the acyl-enzyme intermediate in a reaction that is under either thermodynamic or kinetic control. Although both mechanisms can be used, the kinetically-controlled approach has proven more practical, particularly for peptide synthesis. For such a reaction to be synthetically useful, the native hydrolase activity of the protease must be suppressed. This can be achieved in several ways: changing the medium in which the reaction is carried out, optimizing the substrate structure (e.g., the leaving group) to favor the desired reaction and engineering the enzyme to render it an efficient 'synthetase' rather than a hydrolase. Of all these approaches, enzyme engineering has the greatest potential where one of the most successful examples is the design of subtiligase[103]. Subtiligase is an engineered double mutant of the bacterial serine protease, subtilisin, in which the active site residue Ser221 is mutated to Cys and the Pro225 residue to Ala. Subtiligase still uses the same mechanism, through His64, to start the hydrolysis of the ester peptide[103]. As a result of the first mutation (S221C), subtiligase has all the structural features of a cysteine protease but lacks the normal amidase activity of the latter to catalyze the hydrolysis of otherwise susceptible peptide bonds[103]. However, since an ester bond is much less stable than a peptide amide bond, subtiligase is still capable of catalyzing the hydrolysis of an appropriately reactive peptide ester and therefore retains significant esterase activity. In addition, because the covalent radius of sulfur (S221C) is larger than that of oxygen, the second mutation (P225A) was introduced to move the γ-thiol of Cys221 away from the oxyanion hole and the catalytic histidine by 0.5-1.0 Å to make more room for the substrate binding[103],

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

which further increases the enzymatic activity towards an ester substrate. Most importantly, in the presence of a suitable amino terminus-free peptide nucleophile as the acyl acceptor, aminolysis of the peptide ester predominates over its hydrolysis, leading to almost exclusive formation of a ligation product by the kinetically controlled mechanism[103]. This work is a true demonstration of how rational protein engineering can be employed to convert a normally destructive protease to a synthetically useful peptide ligase. Notably, the synthetic value of subtiligase was successfully tested in the total synthesis of a functionally active ribonuclease A analog[106]. In the present report, I demonstrate that the synthetic utility of subtiligase can be extended to a new application, i.e., catalyzing the thiolysis of peptide esters for the formation of peptide thioesters, a class of compounds of increasing importance as these serve as key building blocks in the total synthesis of proteins.

At first, catalytic peptide ester-to-thioester transesterification by any serine or cysteine protease appeared straightforward, as thiolysis of the acyl-enzyme intermediate would directly lead to the formation of a peptide thioester product (Scheme 3.2).



Scheme 3.2 Transesterification catalyzed by a cysteine protease converting a peptide C^{α}-ester to thioester. R = peptide chain. HS-E = cysteine protease. Mechanism similar to aminolysis of peptide esters[103].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

However, unlike in peptide ligation where aminolysis of the acyl-E intermediate forms a more stable product than the peptide ester, the acyl donor, the thioester that would form in a thiolysis reaction is more labile than the ester substrate and therefore would be quickly hydrolyzed under the normal operation conditions. In fact, because of their higher enzymatic hydrolysis rates, peptidyl thioesters are traditionally used as sensitive substrates of various serine/cysteine peptidases and other proteases for the study of enzyme specificity and inhibition. Because of this, it would be difficult to use a natural protease for the transesterification reactions. Indeed, our attempts to use subtilisin and papain to convert a peptide ester substrate to the corresponding thioester were made with unsuccess (data not show). I therefore turned my attention to subtiligase. With its reduced hydrolytic activity and favorable aminolysis/hydrolysis profile, subtiligase generates a longer-lived acyl-E intermediate, which makes it easier to keep the reaction under kinetic control by manipulating the reaction conditions. In addition, not only a thiol is generally a better nucleophile than water, but also thiol is a soft nucleophile while water is a hard nucleophile. Because of the above reasons, a kinetically controlled reaction should favor thiolysis over hydrolysis and a less hydrolytically active subtiligase would make it possible to isolate the thioester product before it is hydrolyzed by the enzyme.

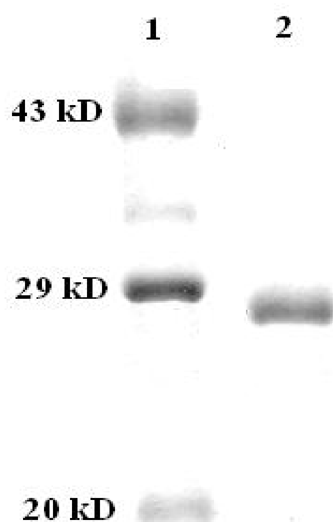
Herein for soft nucleophiles (or electrophiles), their nucleophilic (or electrophilic) atoms are usually large and their ions have diffuse charge. On the other hand, for hard nucleophiles (or electrophiles), their nucleophilic (or electrophilic) atoms are generally small and their ions have intense charge. In general, soft nucleophiles (Lewis bases), such as RHS, RS^- and I^-, have a greater tendency to attack an electron-deficient carbon. Hard nucleophiles (Lewis bases),

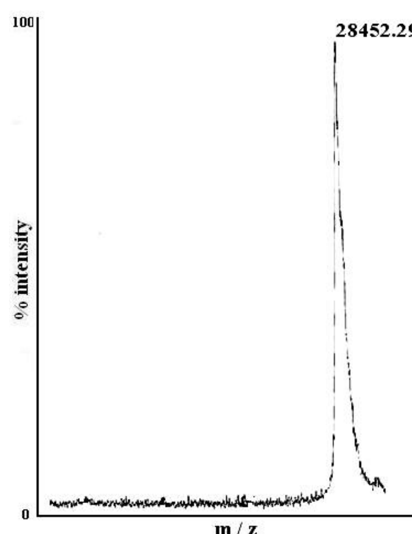
ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

such as H_2O, OH^- and RNH_2, have a greater tendency to abstract a proton. Compared with hard nucleophiles, soft nucleophiles are much easier to react with thioester bonds, which are soft electrophiles (Lewis acids), according to the Hard and Soft Acids and Bases (HSAB) principle[107].

3.2 Results and Discussion

To test this hypothesis, I produced a subtiligase variant that has a His_6 tag at its C-terminus to facilitate purification. The purified enzyme appeared to be homogeneous by SDS-PAGE with Coomassie staining (Figure 3.3 A), and mass spectrometry analysis by MALDI-TOF MS showed a molecular ion $[M+H]^+$ of m/z 28452.3 (Figure 3.3 B).





1, protein marker; **2,** subtiligase

Figure 3.3 A SDS-PAGE gel of subtiligase

Figure 3.3 B MALDI-TOF MS of subtiligase

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The presence of this His_6 tag did not affect the enzymatic activity, as tested in a model ligation reaction (Figure 3.4) between the glycolate ester peptide, Ac-His-Ala-Ala-Pro-Phe-glc-Phe-Gly-NH_2, and the peptide nucleophile, H-Ala-Phe-Ala-NH_2, which represent optimized acyl donor and acyl acceptor substrates, respectively, of subtiligase[103].

The above glycolate ester peptide, which contains an ester bond instead of an amide bond, was synthesized by Fmoc-based method using glycolate as a coupling unit (see 3.4.5, page 69). The –glc- indicates **–O-CH_2-CO-**.

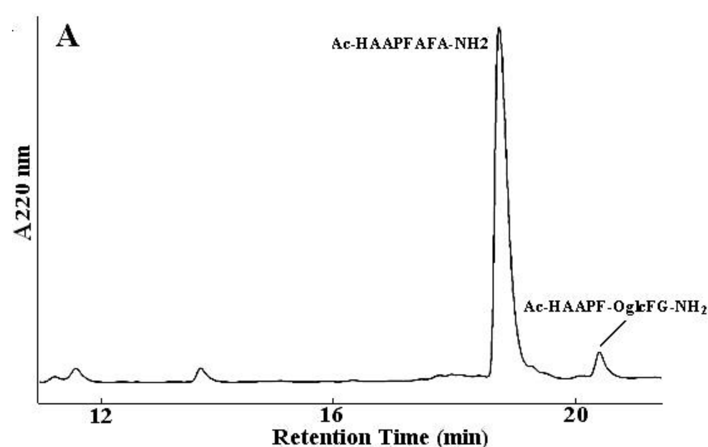


Figure 3.4 A HPLC data of model ligation reaction, which was catalyzed by subtiligase-His tag, between Ac-His-Ala-Ala-Pro-Phe-glc-Phe-Gly-NH_2 and the peptide nucleophile, H-Ala-Phe-Ala-NH_2. The ligation product is Ac-His-Ala-Ala-Pro-Phe-Ala-Phe-Ala-NH_2 (m/z $[M+H]^+$ found: 872.3, M_w calcd: 871.2).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

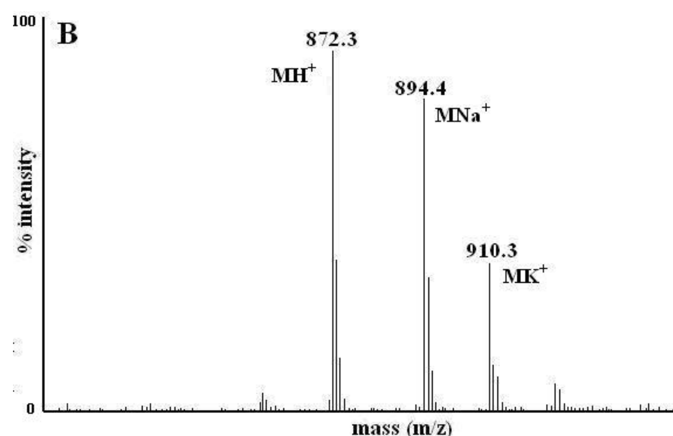


Figure 3.4 B ESI-MS for the ligation product (see Figure 3.4 A)

When the peptide ester substrate was subjected to enzymatic thiolysis at pH 7 with N-acetylcysteamine (ACA, Ac-$NHCH_2CH_2SH$), rapid formation of the thioester product was observed (Figure 3.5).

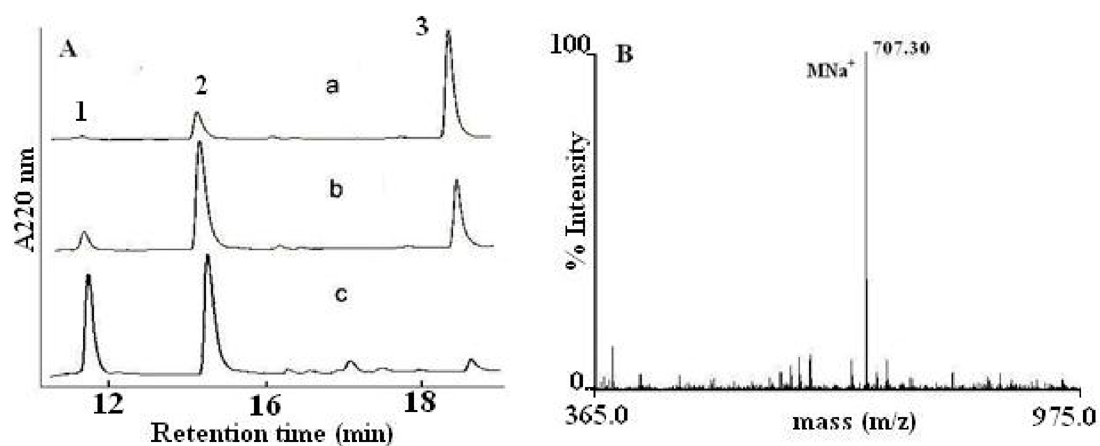


Figure 3.5 A) HPLC monitoring of subtiligase-catalyzed thiolysis of Ac-**HAAPF**-glc**FG**-NH_2 by ACA. a) Reaction at 2 min, b) at 5 min and c) at 18 min. Peak 1: hydrolysis product, Ac-**HAAPF**-OH; peak 2: the newly formed thioester product; and peak 3: the glc ester starting material. B) ESI-MS of the thioester product (m/z $[M+Na]^+$ found: 707.3, M_W calcd: 684.3). See experimental section for reaction conditions.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

As seen in Fig. 3.5, when approximately 75% of the substrate was consumed in 5 min, the thioester product, Ac-His-Ala-Ala-Pro-Phe-$SCH_2CH_2NHCOCH_3$, was formed in 65% yield with 10% being the hydrolysis product. However, the thiolysis/hydrolysis ratio decreased over time due to instability of the newly formed thioester toward enzymatic hydrolysis, and a significant increase of the thermodynamically-favored hydrolysis was observed at 18 min of reaction (Fig. 3.5 A, panel c).

This result clearly demonstrated the kinetically-controlled feature of the thiolysis reaction and that the thioester product could indeed be isolated in good yield under normal conditions of enzymatic reactions. In the absence of the enzyme, no thioester or hydrolysis product was formed. Moreover, the new formed thioester is relatively stable in the buffer because it showed no obvious hydrolysis under same reaction condition (mentioned in Figure 3.5 except no adding of subtiligase) for 1 h. I further found that this thiolysis reaction could be performed at a much lower pH, *albeit* at a significantly slower rate. For instance, thiolysis of the above peptide ester at pH 4.4 for 6 h gave about 78% thioester product and 13% hydrolysis with 9% ester substrate remaining unconverted. Lowering the pH of the reaction medium reduced the reaction rates of both thiolysis and hydrolysis. However, hydrolysis appeared to be affected to a greater degree. When the kinetic parameters of the normal hydrolysis reaction of the same peptide substrate were measured in the absence of ACA, a 82-fold drop in catalytic efficiency was observed as the pH was lowered from 7 to 4.6 (Table 3.1, page 59).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

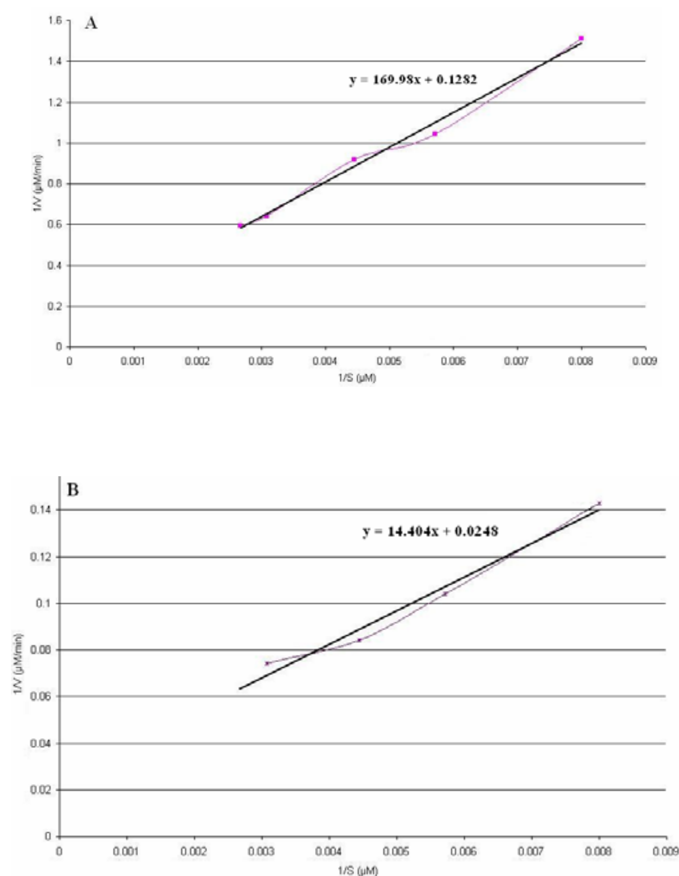
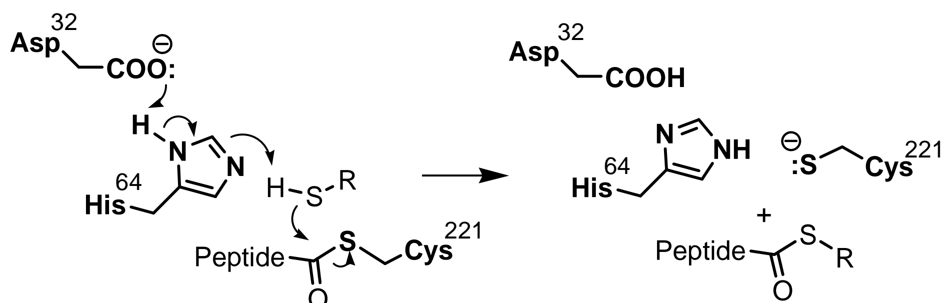


Figure 3.6 Lineweaver-Burk plots for thiolysis of Ac-**HAAPF**-glc**FG**-NH_2 with ACA catalyzed by subtiligase. The upper one, A, is incubated at pH 4.6 while B is at pH 7.0.

This was mainly due to a 51-fold decrease in the catalytic turn-over number (k_{cat}) as the binding constant (K_m) was reduced only by less than 2-fold. For the thiolysis reaction, there was a 12-fold decrease in catalytic efficiency, with k_{cat} decreasing by only about 5-fold. The improved thiolysis/hydrolysis profile at low pH may be due to the differences in pK_a values of ACA (pK_a 9.4)[108] and water (pK_a 15.7). Although lowering the pH would make it more difficult for the enzyme to activate both the thiol and water in the deacylation step, the low pK_a value of a mercaptan like ACA would still allow it to act as a good nucleophile at low pH and be deprotonated by the enzyme and compete more favorably with water for the interception of the acyl-E intermediate (Scheme 3.3).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



Scheme 3.3 Activation of an alkyl thiol at the active site of subtiligase in the deacylation step.

Thus, the thiolysis reaction was made even more favorable over hydrolysis under the kinetically controlled mechanism at low pH. The practical implication of my observations is that it offers more flexibility for one to choose the best pH condition for a given peptide ester substrate in this enzymatic thioester formation reaction, depending on its reactivity and solubility at different pH. For example, for a peptide ester substrate, whose solubility is much better in a weak acid buffer than in a neutral buffer, it is more suitable to perform the enzymatic thioester formation reaction under the weakly acidic condition.

Reaction	pH	K_m	k_{cat}	k_{cat}/K_m
Hydrolysis	7.0	0.38	142.2	3.7×10^5
	4.6	0.61	2.75	4.5×10^3
Thiolysis	7.0	0.58	51.5	8.9×10^4
	4.6	1.33	10.0	7.5×10^3

Table 3.1 Comparison of kinetic constants for the hydrolysis and thiolysis with ACA of Ac-His-Ala-Ala-Pro-Phe-glc-Phe-Gly-amide at pH 7 and pH 4.6. Kinetic constants were calculated by using Lineweaver-Burk double-reciprocal plot (Fig. 3.6). See 3.4.6 C for reaction conditions. Thiol used for thiolysis was ACA.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

I also determined the effect of the thiol concentration on the thiolysis reaction. The optimal concentration of ACA was 70 - 175 mM, which is about 280 to 700-fold excess compared to the substrate. Increasing the thiol concentration further would actually decrease the thiolysis/hydrolysis ratio. It is imperative to note that subtiligase-catalyzed thiolysis of the glycolate ester peptide produced almost exclusively the thioester product during the initial phase of the reaction (Fig. 3.5 A, page 56), even though a similar hydrolysis reaction conducted separately in the absence of a thiol nucleophile resulted in a comparable (at pH 4.6) or higher (at pH 7) catalytic efficiency than that of thiolysis (Table 3.1, page 59). Therefore, the thiol compound at a very low molar ratio to water was very effective in engaging the acyl-E thioester intermediate for the transthioesterification reaction at the deacylation step. The soft base nature of a thiol and soft acid nature of a thioester may also account for this preferred reactivity.

To assess to which extent the deacylation step is catalyzed by subtiligase, I prepared a peptide thioester (Figure 3.7 B) which is chemically equivalent to the acyl-E thioester intermediate (Figure 3.7 A) in the context of a peptide chain.

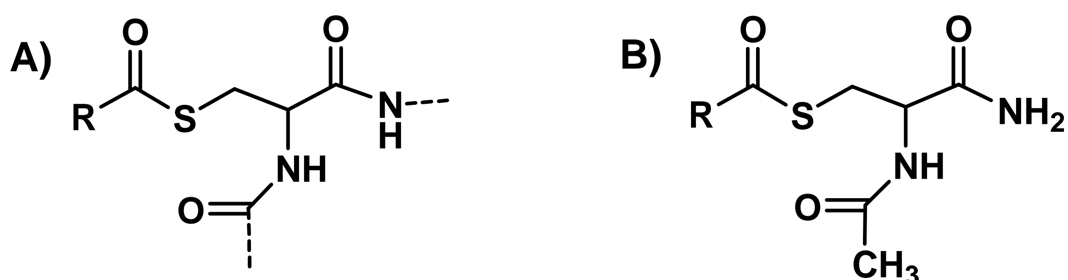


Figure 3.7 Structures of A) the acyl-E thioester intermediate and B) its chemically equivalent thioester formed with the side-chain thiol group of Ac-Cys-NH_2. R-CO- = Ac-His-Ala-Ala-Pro-Phe-.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

It turned out that, by itself, this peptide thioester could readily undergo transthioesterification with ACA at pH 4.6. As ACA was used in large excess (500-fold), this chemical transthioesterification can be treated as a first-order reaction and its rate constant k was determined to be ~ 0.013 min^{-1} (or $t_{1/2}$ ~ 53 min) (see 3.4.6 C for reaction condition). Using a peptide ester as the substrate for a serine or cysteine protease, deacylation of the acyl-enzyme intermediate is usually the rate-determining step[109, 110]. In this case, this deacylation step can also be viewed as a first-order reaction, considering the low [acyl-E] at any given moment and the high thiol concentration. Therefore, the k_{cat} of the enzyme catalyzed thiolysis reaction, being 10 min^{-1} at pH 4.6 (Table 3.1, page 59), is also the reaction rate constant of the rate-limiting deacylation step. A comparison of the rate constants observed for the two thioesters shown in Fig. 3.7 indicates that the peptidyl-enzyme thioester intermediate (Figure 3.7 A) undergoes thiol-thioester exchange ~ 800 times faster than does the chemically equivalent thioester (Figure 3.7 B) under the same conditions. This can be attributed to the catalysis brought about by the acid-base proton relay system at the enzyme's active site (Scheme 3.3, page 59).

I also tested the potential of several other thiols for use in the thiolysis reaction. As seen in Table 3.2, all the alkyl thiols tested gave modest to good yields of thiolysis products.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

N	Thiol[a]	pH 4.6[b]			pH 4.8[b]			MS[c]
		E%	T%	H%	E%	T%	H%	
1	ACA	26	65	9	8	74	18	707.3
2	MPA	50	47	3	5	80	15	694.2
3	MAA	32	57	11	4	64	32	680.0
4	TLA	47	44	9	8	57	35	694.3
5	MES	27	50	23	9	58	33	752.2

Table 3.2 Subtiligase-catalyzed thiolysis of Ac-**HAAPF**-glc**FG**-NH_2 by different thiols in a 2 h reaction.

Reaction conditions: 0.25 mM Ac-**HAAPF**-glc**FG**-NH_2 (with 0.25 mM Fmoc-Lys-OH as internal standard), 1.5 µM subtiligase, 175 mM thiol, 85 mM TCEP in 0.25 M sodium acetate buffer, at 25 °C for 2 h. a) MPA: mercaptoproprionic acid; MAA: mercaptoacetic acid; MES: 2-mercaptoethanesulfonate; TLA: thiolactic acid (racemic). b) Yields calculated using normalized HPLC absorption peak area at 220 nm. E: remaining peptide ester; H: hydrolysis; T: thiolysis. c) m/z of the $[M+Na]^+$ molecular ion from ESI-MS. All MS data agreed with calculated values.

The best results were obtained with ACA and mercaptoproprionic acid (MPA). Although mercaptoacetic acid (MAA) and 2-mercaptoethanesulfonate (MES) also gave satisfactory results, a significant degree of hydrolysis was observed, possibly due to the neighbouring group effect from the negative carboxylate or sulfonate moieties. However, an aromatic thiol, *p*-mercaptobenzenepropionic acid (MBPA) gave no thioester product (data not show). Therefore, my data suggest that simple alkyl thiols were very effective in capturing the acyl-E thioester intermediate, which is reminiscent of the biosynthetic intein-mediated thioester formation concept[111, 112].

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

To further demonstrate the synthetic value of subtiligase-catalyzed thiolysis, I prepared three other peptide substrates with the glc ester linkage at the C-termini.

N	Sequence	Yield	T/H ratio[a]	MW[b] calc.	$[M+H]^{+[b]}$ found
1	PKGTKDVAFGSF	42%	7/3 (60%)	1340.6	1341.7
2	TKGSAYSGKLEEFVQ	52.8%	3/2 (88%)	1730.8	1731.7
3	LEGQAAKEFIAWLVNGRAY	35%	1/1 (70%)	2223.1	2224.2

Table 3.3 Thioester formation of long glycolate ester peptides. All peptides were CT-esters of glc**FG**-NH_2. The thiol nucleophile was MPA. Conditions: 0.2 mM peptide, 2.5 μM subtiligase, 350 mM MPA, 175 mM TCEP in 0.5 M sodium acetate buffer at pH 5.1. Reaction time: 8 h for peptide 1; 16 h for peptide 2; 3 h for peptide 3. a)T/H: thiolysis/hydrolysis ratio. The value in brackets is how much ester peptide was consumed. b) Calculated molecular weight and found m/z value of the corresponding thioester products.

These peptides are complex in structure and contain amino acid residues with a diverse set of side chains. The data on the thiolysis reactions by MPA with these substrates are given in Table 3.3. It is evident from Table 3.3 that as long as the peptide glc esters can serve as substrates of subtiligase, the corresponding thioesters can be prepared in moderate yields. No intermolecular or intramolecular self-ligation was detected even though the N-termini of these peptides were not protected. Except for the competing hydrolysis reaction, the ester-to-thioester conversion proceeded smoothly and cleanly without other detectable side reactions. A representative HPLC trace showing the conversion of the glc ester peptide **2** to its thioester at pH 5.1 is presented in Figure 3.8.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

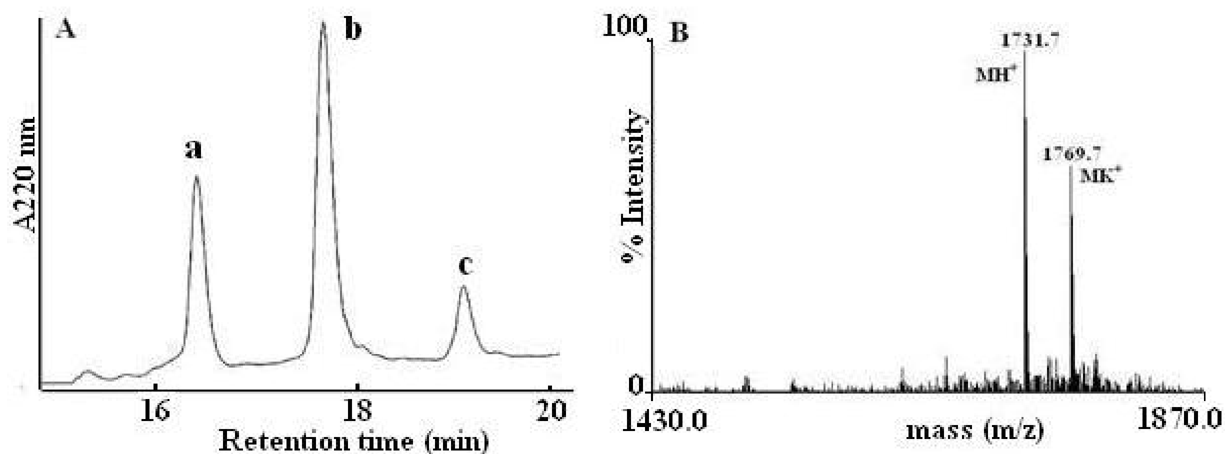


Figure 3.8 Subtiligase-catalyzed thiolysis of peptide 2 (from Table 3.3) by MPA. See Table 3.3 for reaction conditions. A) HPLC diagram for reaction at 16 h. Peak a: hydrolysis product (m/z $[M+H]^+$ found: 1644.1, M_w calcd: 1643.1); peak b: newly formed thioester; and peak c: starting material (m/z $[M+H]^+$ found: 1904.7, M_w calcd: 1903.6). HPLC gradient: 0-40% buffer B in 20 min. B) ESI-MS of the thioester product (M_W calcd. 1731.0).

The relatively high amount of hydrolysis product found in the cases of these peptide substrates compared to the first model reaction of Ac-**HAAPF**-glc**FG**-NH_2 by ACA is mainly due to the long reaction time required for these peptides which do not seem to be the ideal substrates for subtiligase.

The MPA thioester of peptide **2** (see Table 3.3), purified by HPLC, was tested for ligation with a 7-residue cysteinyl peptide, H-**CPFEVKG**-NH_2 (CG-7), at pH 7.5 for 20 h. The expected ligation product was obtained in 85% yield (Figure 3.9).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

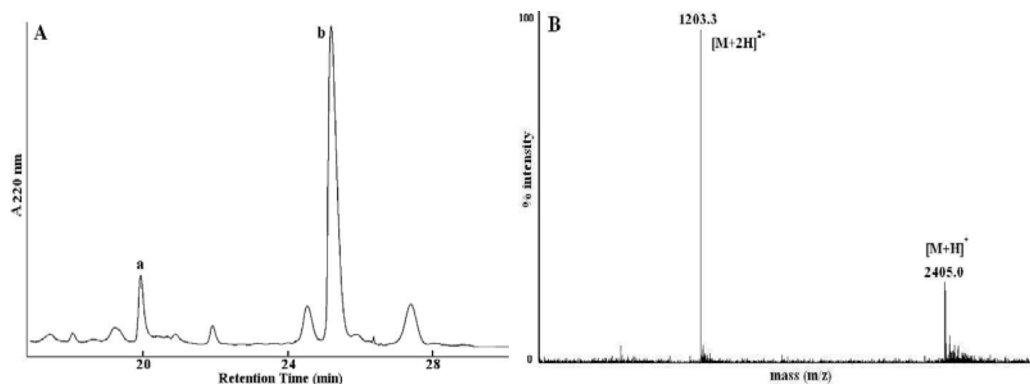


Figure 3.9 A) HPLC of ligation between H-**TKGSAYSGKLEEFVQ**-$COSCH_2CH_2COOH$ and CG-7. Peak a: H-**TKGSAYSGKLEEFVQ**-CO-$SCH_2CH_2NHCOCH_3$ and peak b: ligation product H-**TKGSAYSGKLEEFVQCPFEVKG**-NH_2 (m/z $[M+H]^+$ found: 2403.3, M_w calcd: 2402.2). B) ESI-MS of the ligation product (peak b). See experimental section for conditions.

Likewise, H-**TKGSAYSGKLEEFVQ**-CO-SCH_2CH_2COOH (2.5 mM) was reacted with H-**CGRGVPNGIPAEDSEQLASGQ**-NH_2 (**CQ**-21) (5 mM) in 0.2 M sodium phosphate buffer (pH 7.5) containing 4 M guanidine hydrochloride to give 60% ligation product (m/z $[M+H]^+$ found: 3708.7, M_w calcd: 3707.8).

It is necessary to note that because the second residue (P_2'-side) of H-**CPFEVKG**-NH_2 (**CG**-7) is a proline, **CG**-7 did not act as the acyl-acceptor substrate when it was used in subtiligase-catalyzed ligation with the glycolate ester of peptide **2**[103]. And the longer cysteinyl peptide **CQ**-21 was not soluble enough in the aqueous enzymatic ligation buffer used for subtiligase. Therefore, by converting the glycolate ester to a thioester, it was possible to overcome the substrate specificity (P'-side) and solubility restrictions imposed by the enzyme, which in a way expanded the scope of the enzymatic ligation method.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.3 Summary and Conclusion

Ester-to-thioester transesterification is an endergonic process and is thermodynamically unfavorable. Nevertheless, the reaction equilibrium can be shifted in favor of thioester formation at high concentrations of the thiol nucleophile; and the high activation energy for the reaction can be overcome through the catalysis of an enzyme. What is most critical, however, is to find a suitable reaction system that can minimize the competing, thermodynamically favoured hydrolysis reaction. Although anhydrous non-polar organic media have been employed for lipase-catalyzed preparation of simple aliphatic thioesters[113], the conditions - namely the solvent system and the enzyme - are not applicable to complex peptide molecules which are not substrates of lipases and are hydrophilic in nature. I demonstrated herein that subtiligase, a reengineered protease, was a suitable catalyst to convert a peptide ester to thioester in aqueous media by a kinetically controlled mechanism. To the best of my knowledge, this is the first time that a peptide thioester was prepared in such a way under conventional enzyme catalysis conditions. The relatively wide operating pH range makes it applicable to medium-sized peptides that are water-soluble. Of course, as for any enzymatic reaction, substrate specificity will be a major factor limiting the general use of this methodology. However, compared to other proteases, subtilisin has a relatively broad substrate specificity, which is paralleled by subtiligase[103]. This will make the enzymatic thiolysis reaction a useful alternative to the chemical methods now used for the synthesis of peptide thioesters. In conclusion, the subtiligase-catalyzed peptide ester-to-thioester transesterification reaction represents a significant addition to the repertoire of enzymatic methods for organic transformations.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.4 Experimental

3.4.1 Materials

All chemical reagents were of analytical grade, obtained from commercial suppliers and used without further purification. *B. amyloliquefaciens* (ATCC 23844) and plasmid pBS42 (ATCC 37279) were obtained from ATCC (USA). *Bacillus subtilis* RIK1285 is a gift from Dr. Fujio Kawamura of Rikkyo University, Japan.

3.4.2 High Performance Liquid Chromatography.

High Performance Liquid Chromatography (HPLC) was performed on an Agilent system (Agilent Tech., USA) with a Vydac C_{18} reverse-phase column (5 µm, 250 mm×4.6 mm). The analysis was carried out using the mixture of two solvents, A and B, as the mobile phase. Solution A was H_2O (deionized) containing 0.045% TFA and solution B was acetonitrile/H_2O (90/10) containing 0.04% TFA. All analyses were carried out in a concentration gradient 10-50% B in 40 min, except mentioned otherwise. The mobile phase flow rate was 1 mL-min^{-1} and the separation temperature was 23 ^{o}C. UV detection was carried out at 220 nm.

3.4.3 Mass Spectrometry.

ESI-MS data were obtained on a Finnigan LCQ Deca XP MAX instrument. MALDI-TOF-MS data were obtained on an Applied Biosystems 4700 Proteomic Analyzer 72.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

3.4.4 Expression and Purification of Subtiligase (Figure 3.3).

The gene of subtilisin was cloned from *B. amyloliquefaciens* (ATCC 23844) according to published procedures[114] except that a His_6 tag was added to the C-terminus using the two primers: ACCGAATTCGGTCTACTAAAATATTAT TCC (forward one) and ATTGGATCCTTAATGATGATGATGATGATGCTG AGCTGCCGCCTGTACGTT (reverse one). It was then subcloned as *EcoR* I-*BamH* I fragment into plasmid pBS42 (ATCC 37279), digested earlier with *EcoR* I and *BamH* I. Transformation of *E. coli* ER2738 was done by the competent cell method and positive clones were screened by digestion with *EcoR* I-*BamH* I. The S221C/P225A mutations for the subtiligase gene[103] were introduced into the wild-type subtilisin gene by using the Quick Change Site-Directed Mutagenesis Kit (Stratagene) and amplified by *E. coli* ER2738 transformation. The mutated gene was confirmed by DNA sequencing. *Bacillus subtilis* RIK1285, deficient in the production of neutral and alkaline proteases, is a gift from Dr. Fujio Kawamura[115]. Competent cells of *B. subtilis* RIK1285 were acquired and transformed by S221C/P225A mutant pBS42 as described before[116]. The *Bacillus subtilis* RIK1285 strains containing subtiligase gene were grown in 2×TY media, supplemented with 10 mM $CaCl_2$ and 12.5 μg/mL chloramphenicol at 37 ^{o}C with shaking for 24 h. Purification of subtiligase was performed using previously reported protocols[117, 118]. Expression cultures were clarified by centrifugation (8000g, 15 min, 4 ^{o}C) and three volumes of cold ethanol (-20 ^{o}C) was added to the supernatant and centrifugated for 15 min (8000g, 4 ^{o}C). The drained protein pellet from a 500-mL culture was redissolved in 80 mL of binding buffer containing 50 mM KH_2PO_4 (pH 6.5), 20 mM imidazole, 300 mM NaCl; the solution was clarified by centrifugation and 2 mL of Co^{2+} metal affinity resin (BD Talon^{TM}) was added.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

The solution was mixed at 4 ^{o}C on a rotary mixer for 3h; the resin was then packed in an Econo-Column (BioRad), washed with 100 mL binding buffer and eluted twice by 5 mL of elution buffer containing 50 mM KH_2PO_4 (pH 5.8), 250 mM imidazole, 300 mM NaCl. Purified subtiligase was desalted by PD-10 column (Amersham) with deionized H_2O. The concentrations of subtiligase were determined spectrophotometrically[119] and aliquots of enzymes were flash-frozen, lyophilized and stored at -80 ^{o}C until analyzed.

3.4.5 Peptide Synthesis.

All peptides were synthesized using standard Fmoc chemistry[120] on Rink-amide-MBHA-resin or Tentagel amide resin. The glycolate ester peptides were prepared by Fmoc-based synthesis protocols[97]. In a typical peptide synthesis experiment, the ratio of Rink-Amide-MBHA-resin to Fmoc amino acids, PyBOP and DIEA was 1:4:4:8. Coupling reaction was carried out in DMF for 45-60 min. Fmoc group was removed by 20% piperidine (in DMF). After sequence assembly, peptides were cleaved from the solid support using a cleavage solution of 95% TFA, 2.5% deionized H_2O and 2.5% TIS for 3 h. Precipitated by trituration with ice-cold Et_2O, the crude peptides were purified by reverse-phase HPLC (Vydac C_{18}). After lyophilization, peptides were obtained in powder form, which could be used freshly or stored at -20 ^{o}C until future use. The identities of all peptides were confirmed by ESI-MS analysis.

Ac-**HAAPF**-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 845.5, M_w calcd: 844.4). H-**PKGTKDVAFGSF**-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 1515.6, M_w calcd: 1513.7).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

H-**TKGSAYSGKLEEFVQ**-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 1904.7, M_w calcd: 1903.9).

H-**LEGQAAKEFIAWLVNGRAY**-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 2398.0, M_w calcd: 2396.2).

CG-7: H-**CPFEVKG**-NH_2 (m/z $[M+H]^+$ found: 778.5, M_w calcd: 777.4).

CQ-21: H-**CGRGVPNGIPAEDSEQLASGQ**-NH_2 (m/z $[M+H]^+$ found: 2084.5, M_w calcd: 2083.0).

For the synthesis of the acyl-E intermediate-equivalent peptide thioester Ac-Cys(Ac-**HAAPF**)-NH_2 (Fig. 3.7 B, page 60), Fmoc-Cys(Trt)-OH was first loaded to Rink-amide-MBHA resin, followed by removal of Fmoc and acetylation with Ac_2O. The trityl group was then removed by treating the resin with 2% TFA in DCM (containing 3% triisopropylsilane) for 5 x 2 min. Onto the free mercaptan of Cys was then coupled Fmoc-Phe-OH and the peptide chain was elongated using a protocol developed by Li *et al.*[121], and after final cleavage and purification the target thioester was obtained in *ca.* 15 % yield (m/z $[M+H]^+$ found: 728.3, M_w calcd: 727.3).

3.4.6 Procedures for Enzyme Assays.

All peptide substrates were HPLC-purified and lyophilized before use. The typical reaction volume for an enzymatic ligation or thiolysis was 10-30 μL.

A. Confirmation of enzymatic activity of the His_6-tagged subtiligase (Figure 3.4).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Enzymatic ligation of the glycolate ester peptide Ac-**HAAPF**-glc**FG**-NH_2 (1 mM) with the tripeptide H-**AFA**-NH_2 (10 mM) was performed in the presence of 4 μM subtiligase at 25 ^oC in 100 mM Tricine buffer (pH 8.0) containing 10 mM Tris[2-carboxyethyl]phosphine (TCEP). The reaction was stopped by adding 10% TFA/H_2O solution and analyzed by HPLC (Figure 3.4 A). The corresponding peaks were collected and detected by ESI-MS (Figure 3.4 B). The results indicated that subtiligase-His_6 was fully active to give near quantitative formation of the ligation product Ac-**HAAPFAFA**-NH_2 in 40 min, and no detectable hydrolysis of this product was observed after overnight incubation.

B. Thiolysis of Ac-**HAAPF**-glc**FG**-NH_2 with ACA (Fig. 3.5).

Ac-**HAAPF**-glc**FG**-NH_2 (0.25 mM) was subjected to thiolysis under the catalysis of subtiligase (1.5 μM) in a 0.25 M sodium phosphate buffer (pH 7) containing 175 ACA, 87.5 mM TCEP and 0.25 mM Fmoc-Lys-OH as an internal standard at 25 ^oC. The reaction was stopped at different time intervals by mixing an aliquot of the reaction solution with 10% aqueous TFA solution before subjecting it to reverse phase HPLC analysis. The thioester product was characterized by ESI-MS (m/z $[M+Na]^+$ found 707.3; M_W calcd: 684.3).

C. Measurement of kinetic constants (Table 3.1, Figure 3.6).

Kinetic constants for the model peptide ester substrate at different pH were obtained from initial rate measurements of the hydrolysis or thiolysis reaction. The kinetic parameters obtained represent the average of five independent experiments. Each experiment used 5 different substrate concentrations (from 62.5

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

to 875 µM) of the ester substrate, Ac-**HAAPF**-glc**FG**-NH_2, and a fixed concentration of the subtiligase variant (0.75 µM) in a suitable reaction buffer. Fmoc-Lys-OH was included as an internal standard at the same concentration as the peptide ester substrate.

To ensure that the reaction rate measured was the initial velocity, the reaction - hydrolysis or thiolysis - was stopped at a time within which about 10% of the substrate was consumed. At this time, the amount of hydrolysis product was negligible in the thiolysis reaction.

The reaction buffer for hydrolysis at pH 4.6 and pH 7 was 0.35 M sodium acetate and potassium phosphate, respectively, each containing 10 mM TCEP. For thiolysis, 0.1 M ACA and 50 mM TCEP were included in the above buffers. The reaction was carried out at 25 ^{o}C and quenched by adding 10% aqueous TFA solution. The quenched reaction solution was then subjected to analysis by reversed-phase HPLC. The calculation was based on the corresponding normalized HPLC peak areas of the ester substrate, the thioester and hydrolysis products as a ratio to that of the internal standard Fmoc-Lys-OH.

D. Thioester formation with different thiols (Table 3.2).

Thiolysis of Ac-**HAAPF**-glc**FG**-NH_2 (0.25 mM) was conducted with 1.5 µM subtiligase in 0.25 M sodium acetate buffer (pH 4.6 or 4.8) containing 175 mM thiol compound, 87.5 mM TCEP and 0.25 mM Fmoc-Lys-OH at 25 ^{o}C. The reaction was stopped at 2 h by adding 10% TFA. All the thioester products were analyzed and confirmed by HPLC and ESI-MS.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

E. Thiolysis of long glc ester peptides with mercaptoproprionic acid (Table 3.3).

A glc ester peptide substrate (0.2 mM) was subjected to thiolysis by MPA (350 mM) in the presence of 2.5 μM subtiligase in 0.5 M sodium acetate buffer (pH 5.1) containing 175 mM TCEP at 25 ^{o}C. The reaction was quenched by adding 10% TFA at specified time (Table 3.3). All the products were detected by HPLC and ESI-MS.

3.4.7 Native Chemical Ligation between an Enzymatically Synthesized Peptide Thioester and NT-Cysteinyl Peptides (Fig. 3.9).

The MPA thioester of peptide **2** (table 3.3**),** H-**TKGSAYSGKLEEFVQ** -CO-SCH_2CH_2COOH (2.5 mM), was reacted with CG-7 (H-**CPFEVKG**-NH_2, 7.5 mM) in 0.2 M sodium phosphate buffer (pH 7.5) containing 30 mM ACA and 20 mM TCEP at 25 ^{o}C to give ca. 85% ligation product after 20 h based on HPLC analysis (Fig. 3.9).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Chapter 4: Development of Enzymatic and Chemical Methods for the Synthesis of Peptide C^{α}-thioacids

4.1 Introduction

The advent of chemoselective peptide ligation strategies marks a significant advance in synthetic peptide and protein chemistry. With the availability of these methods and together with solid phase peptide synthesis (SPPS), proteins of small-to-medium sizes are now within the reach of a synthetic chemist[61, 122]. This overcomes the limitation of traditional SPPS to produce large peptides (more than 60 amino acids in length) which usually suffers from problems such as low yield and purification difficulty[123].

Many of these peptide ligation methods make use of a thiol group to initiate the nucleophilic attack at the capture step to bring two reacting peptide segments together[89, 124, 125]. In these methods, peptide C^{α}-thioacids[124], as well as peptide C^{α}-thioesters[60, 89, 90] described in Chapter 3, are commonly used as building blocks. The sulfhydryl group of a thiocarboxylic acid (or thioacid) is a very soft and powerful nucleophile and as such has long been exploited in peptide synthesis.

The use of peptide thioacids has been well documented. In the early approaches, researchers usually coupled two partially deprotected peptides together in aqueous media, in the presence of a heavy metal ion such as silver ions and HOSu (N-hydroxysuccinimide)[126] (Scheme 4.1).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Boc-Gly-S–resin
↓ 1. Solid-phase peptide synthesis 2. HF
[GlyS[17]]-β-EP-(1–17)
↓ Citraconic anhydride
Cit-[Lys(Cit)[9], GlyS[17]]-β-EP-(1–17)
(Ia)

Boc(Bzl)-Glu resin
↓ 1. Solid-phase peptide synthesis 2. HF
Fmoc-β-EP-(18–31)
↓ 1. Citraconic anhydride 2. Piperidine
[Lys(Cit)[19,24,28,29]]-β-(18–31)
(IIa)

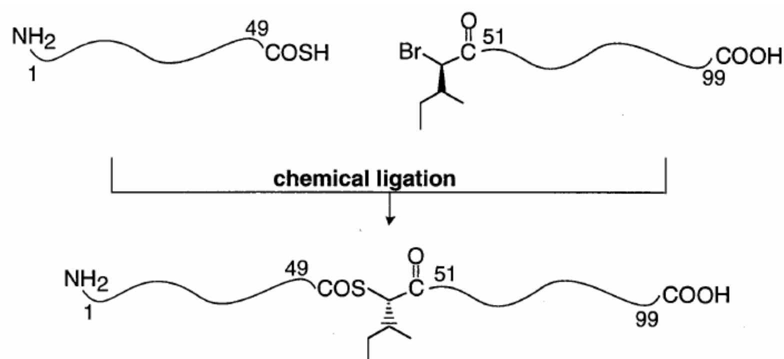
Ia + IIa → 1. $AgNO_3$/HOSu 2. AcOH

H-Tyr-Gly-Gly-Phe-Met-Thr-Ser-Glu-Lys-Ser-Gln-Thr-Pro-Leu-Val-
1 10
Thr-Gly-Phe-Lys-Asn-Ala-Ile-Ile-Lys-Asn-Ala-Tyr-Lys-Lys-Gly-Glu-OH
17 31

Scheme 4.1 Use of peptide thioacid to synthesize [Gly^{17}]-β-endorphin. The thioacid peptide, $GlyS^{17}$-β-endorphin-(1-17), which was synthesized by Boc based solid-phase method, and another synthetic peptide, β-endorphin-(18-31), were treated with citraconic anhydride to give peptides **Ia** and **IIa** respectively. Through catalysis by silver nitrate-HOSu in water, **Ia** and **IIa** were coupled together through an amide bond between Gly^{17} and Phe^{18}. After removing the citraconyl groups in aqueous acetic acid solution, [Gly^{17}]-β-endorphin was isolated in ~35% yield.

Later, a new strategy was developed based on the ligation between a completely deprotected CT-thioacid peptide and a NT-halogenated peptide[127, 128]. However, the limitation is that the newly-formed bond is a thioester bond instead of a stable amide bond (Scheme 4.2).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



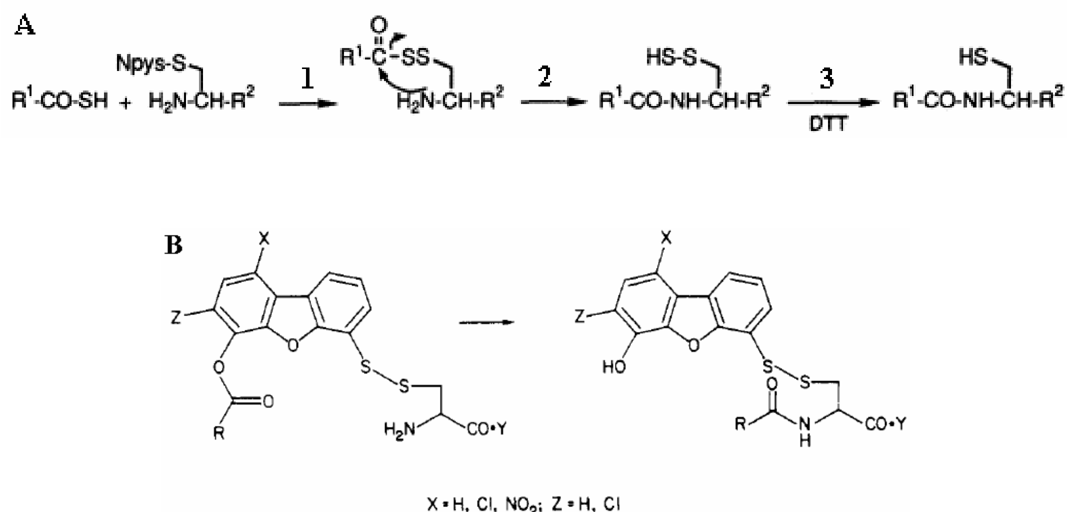
Scheme 4.2 Using of peptide thioacid to synthesize HIV-1 PR with a pseudopeptide bond. The unprotected HIV-1 protease thioacid peptide segment (1-49) and another segment (51-99) which was modified by bromomethylvaleric acid were joined together by chemical ligation.

The thioester pseudopeptide bond that connects the fragments apparently affected the biochemical properties of the newly formed protein compared with the native enzyme. For instance, when the functionally relevant amide -CONH- linkage between residues Gly49-Ile50 in the HIV-1 protease was replaced by an isosteric thioester -COS- bond, the new backbone-engineered enzyme had normal substrate specificity and affinity (K_m); however, the catalytic activity (K_{cat}) was reduced ≈ 3000-fold compared to the native enzyme[128].

This problem was overcome by new ligation methods that can form natural peptide bonds. One such method was developed by C.-F. Liu *et al.*[124]. It is a highly efficient orthogonal coupling approach for peptide bond formation using unprotected peptide segments. This method consists of three key steps (Scheme 4.3 A): 1. Capturing an Npys (3-nitro-2-pyridinesulfenyl) modified NT-Cys side-chain thiol of the amino segment with a C^{α}-thioacid of the acyl segment to form an acyl

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

disulfide; 2. A rapid intramolecular acylation to generate an amide bond; 3. A thiolytic reduction to produce the final product with a native Cys residue at the ligation site.



Scheme 4.3 A) The proposed mechanism of the mini thiol-capture ligation strategy. B) The intramolecular acylation reaction (Kemp reaction) mediated by a fused-ring template.

Conceptually similar to the early work of Kemp[129], the method is viewed as a 'mini' prior thiol-capture strategy since it does not use a fused-ring template to mediate the intramolecular acylation reaction (Scheme 4.3 B). Consequently, it is a particularly attractive peptide ligation method for protein synthesis. More recently, the amidation reaction from thioacid and azide has also received renewed attention and a thorough mechanistic study suggests its potential usefulness in peptide ligation[130, 131]

Although peptidyl thioacids have the great value to be versatile building blocks for protein synthesis, their chemical preparation remains challenging. The only available approach for solid phase synthesis of peptidyl thioacids is based on Boc

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

biological approaches such as random mutagenesis or gene shuffling for generating large diverse DNA libraries. On the other hand, either good three-dimensional structures of enzymes or amino acid sequence alignments can make rational enzyme redesign an achievable goal. There are three common methods for structure-base rational redesign: reshaping substrate-binding site[135-137], utilizing cofactor specificity[138] and re-engineering catalytic mechanisms[139-141]. The successful example for the subtiligase redesign in this thesis is based on re-engineering catalytic mechanisms.

The second approach to prepare peptidyl thioacids is a simple and convenient chemical approach. Both methods for thioacid preparation are discussed in details in the following sections.

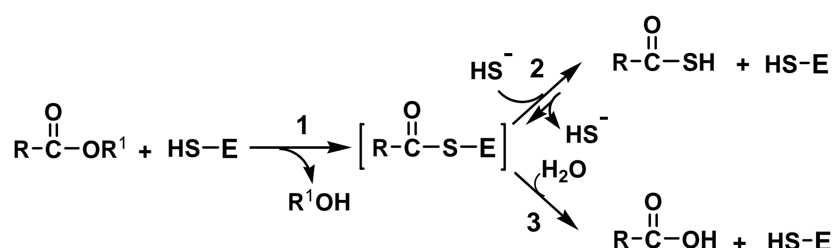
4.2 Results and Discussion

4.2.1 An Enzymatic Method for the Synthesis of Peptide Thioacids

In this part, I report an alternative strategy to Boc-based solid phase synthesis by using a simple and mild enzymatic method for the preparation of peptidyl thioacids. Our approach relies on the use of subtiligase - a peptide ligase reengineered from subtilisin[142] - to catalyze the thiolysis of a suitable peptide ester with the hydrosulfide anion. I have shown that peptide thioesters can be prepared through subtiligase-catalyzed transthioesterification of peptide glycolate esters with simple alkyl thiols[143] (see Chapter 3 for details). Now I have expanded this methodology to the synthesis of peptide thioacids. Subtiligase is a cysteine protease-like enzyme

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

that can catalyze the aminolysis of a suitable peptide ester for peptide ligation through an acyl-enzyme (acyl-E) intermediate[103]. Our approach to peptide thioacid synthesis is based on the rationale that the acyl-E intermediate, being a thioester in nature, can be intercepted by a nucleophile such as the hydrosulfide ion, which would lead to direct formation of a peptide thioacid. Being a soft base, the sulfhydryl ion would react effectively with and accept the acyl group from the soft-acidic acyl-E thioester intermediate (Scheme 4.4).



Scheme 4.4 Thiolysis catalyzed by subtiligase converting a peptide C-ester to thioacid. R = peptide chain. HS-E = subtiligase. R^1 = $-CH_2CO$-Phe-Gly-NH_2

Mechanistically, thioacid formation is similar to thioester formation. However, some profound differences would be expected between the two, mainly because of the unique properties of H_2S which, as a gas, has limited aqueous solubility. Its dissociation constant (pK_{a1} 6.9 and pK_{a2} ~17.1)[144] also dictates that it would exist mostly in the H_2S form at neutral to acidic pH and as HS^- ion under alkaline pH. In chapter 3, I found that the optimal pH for thioester formation with simple alkyl thiols was weakly acidic which could minimize the competing hydrolysis reaction. This pH condition may not be applicable to thioacid formation due to the fact that, when we adjust the pH value down [$(NH_4)_2S$ is basic, pH>11], in the process of pH adjustment, the effective concentration of the hydrosulfide ion will be decreased

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

consistently because of the unavoidable escape of the H_2S gas. There- fore, the reaction may need to be operated under slightly alkaline pH (pH>7). Also, the thiolysis product, thioacid, once formed, may be more stable than a thioester because, deprotonated at the weakly basic pH of the reaction milieu and without the P' components, it would be less recognized by the enzyme as a substrate.

Peptidyl thioacid formation was firstly investigated by using H_2S as the thiol resource. H_2S gas was directly injected into ddH_2O to give a final pH 5 solution. The concentration of factual sulfide ions was 68 mM determined by the Ellman's reagent[145] which was described in details in the experimental segment. Under this reaction condition, the starting material Ac-**HAAPF**-glc**FG**-NH_2 underwent thiolysis quite slowly (Fig. 4.2).

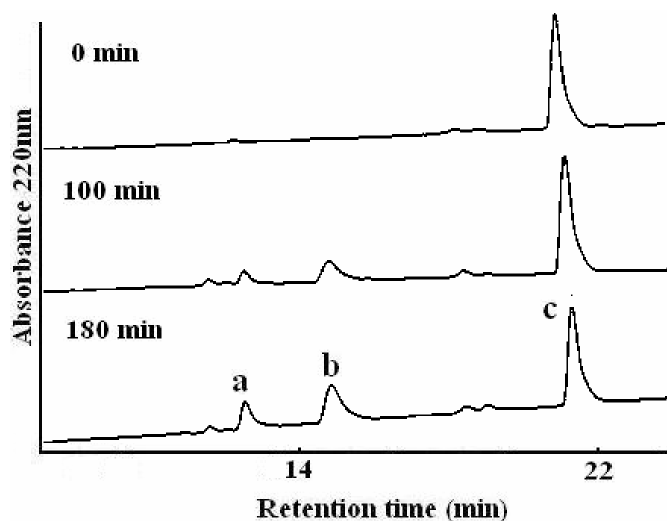


Figure 4.2 HPLC monitoring of the thioacid peptide formation catalyzed by subtiligase. Each time, 10 μL reaction solution was taken at the time point of 0, 100 and 180 min, respectively and quenched by 10 μL of 10% TFA and analyzed by HPLC. Peak a is hydrolysis product Ac-**HAAPF**-OH (m/z $[M+H]^+$ found: 584.5, M_w calcd: 583.6); peak b is the new formed Ac-**HAAPF**-SH (m/z $[M+H]^+$ found: 600.4, M_w calcd: 599.6); while peak c is the starting material Ac-**HAAPF**-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 845.3, M_w calcd: 844.4).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

From Figure 4.2, we can see that the reaction was quite slow. After 3 h only a small amount of the starting material was consumed to give a very little amount of product. To speed up the reaction, 1 M Tris-HCl buffer (pH 7) was used to replace water to dissolve more H_2S gas. Nevertheless, the increased pH value failed to optimize the reaction condition. The higher pH did accelerate the thiolysis rate, however, the hydrolysis rate was also increased. As a result, the ratio of thiolysis/ hydrolysis was not improved. The difficult in dissolving a high enough concentration of H_2S in aqueous buffer indicates that H_2S may not be a good thiol resource for thioacid formation.

As such, Na_2S, NaSH and $(NH_4)_2S$ were also tested in the thioacid formation reaction. Of these, $(NH_4)_2S$ was found to be the most convenient one to use and gave the best results. It was possible to prepare a $(NH_4)_2S$ solution with a high enough effective [HS^-] at pH 7 to 9 without too much escape of H_2S from the solution. Figure 4.3 shows thioacid formation at pH 8 through thiolysis of Ac-His-Ala-Ala-Pro-Phe-glc**FG**-NH_2 by $(NH_4)_2S$ as compared to thioester formation of the same substrate at the same pH with mercaptoacetic acid (MAA). As one can see from Fig. 4.3 (page 83), the reaction rates are comparable for thioacid and thioester formation reactions. For example, after 20 min, the starting material, Ac-His-Ala-Ala-Pro-Phe-glc**FG**-NH_2, was completely consumed in both reactions. However, there was more hydrolysis in the thioacid formation reaction than in the thioester formation reaction. It was further found that the hydrolysis product mainly resulted from direct hydrolysis of the ester substrate Ac-His-Ala-Ala-Pro-Phe-glc**FG**-NH_2, and not from secondary hydrolysis of the thioacid or thioester product. This observation was made by conducting a separate subtiligase-catalyzed hydrolysis assay on the ester substrate, the thioacid

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Ac-His-Ala-Ala-Pro-Phe-SH and thioester Ac-His-Ala-Ala-Pro-Phe-SCH_2COOH products, respectively, in 0.1 M sodium phosphate buffer (pH 8) without any thiol compound but containing 20 mM TCEP, and I found that the latter two were quite stable with less than 10% hydrolysis after 20 min while the glycolate ester peptide was completely hydrolyzed (data not shown).

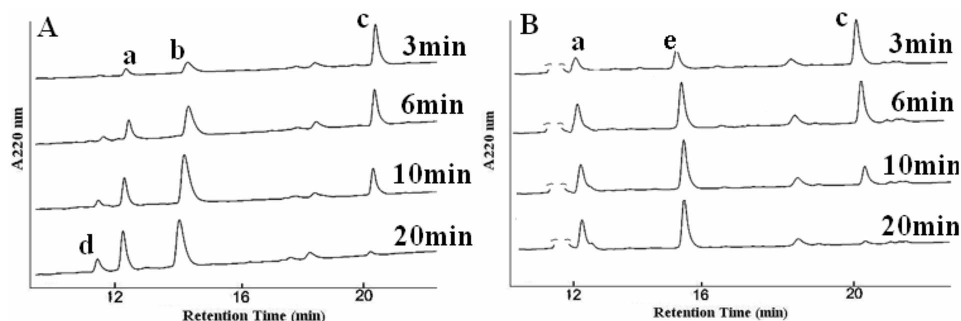


Figure 4.3 A) HPLC monitoring of subtiligase-catalyzed thiolysis of Ac-**HAAPF**-glc**FG**-NH_2 by 0.2 M $(NH_4)_2S$. B) Thiolysis by 0.2 M MAA. Peak a: hydrolysis product, Ac-**HAAPF**-OH (m/z $[M+H]^+$ found: 584.5, M_w calcd: 583.3); peak b: the newly formed thiol-acid product, Ac-**HAAPF**-SH (m/z $[M+H]^+$ found: 600.4, M_w calcd: 599.3); peak c: the ester starting material (m/z $[M+H]^+$ found: 845.5, M_w calcd: 844.4); peak e: the newly formed thioester product, Ac-**HAAPF**-S-MAA (m/z $[M+Na]^+$ found: 680.0, M_w calcd: 657.3); peak d is unknown, (m/z $[M+Na]^+$ found: 680.4 found). HPLC gradient is 10-35% Buffer B in Buffer A for 25 min.

It is worthy to note that the MAA thioester is more stable toward enzymatic hydrolysis than a similar thioester product formed with ACA and that more hydrolysis product was formed during thioester formation reaction with MAA than with ACA (Chapter 3). This may be due to the presence in MAA of a negatively charged α-carboxylate. The increased hydrolysis during thiolysis with the HS^- ion may be due to its poorer ability to enter the active site of the enzyme than other alkyl thiols. Still, it is remarkable that, at a very low molar ratio to water molecules,

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

the hydrosulfide anion is rather effective in engaging the acyl-E intermediate at the deacylation step.

Subsequently, we examined the influence of pH on the thioacid formation reaction. Three different pH values, pH 7.2, 8.2 and 8.8, were chosen and the results are shown in Figure 4.4. Having a pH below neutral was found impractical for maintaining a high enough thiol SH^- concentration owning to escape of H_2S. A pH that is too high (>9) would be deleterious to the enzyme and also lead to more hydrolysis.

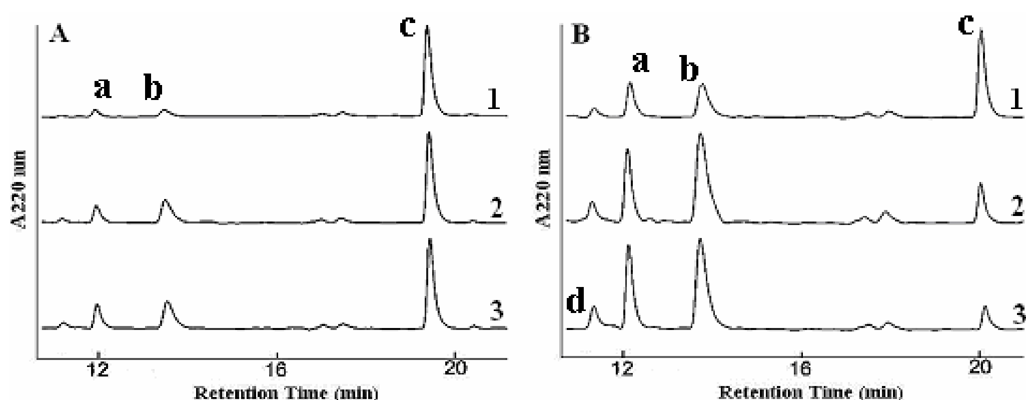


Figure 4.4 HPLC monitoring of subtiligase-catalyzed thiolysis of Ac-**HAAPF**-glc**FG**-NH_2 by *ca.* 0.36 M of $(NH_4)_2S$ for 2 min (panel A) or 8 min (panel B) at different pH values: trace 1 for pH 7.2; trace 2 for pH 8.2 and trace 3 for pH 8.8. Peak a: hydrolysis product, Ac-**HAAPF**-OH (m/z $[M+H]^+$ found: 584.5, M_w calcd: 583.3); peak b: the newly formed thioacid product, Ac-**HAAPF**-SH (m/z $[M+H]^+$ found: 600.4, M_w calcd: 599.3); peak c: the ester starting material (m/z $[M+H]^+$ found: 845.5, M_w calcd: 844.4); peak d is unknown, (m/z $[M+Na]^+$ found: 680.4 found). HPLC gradient: 10-35% buffer B in buffer A in 25 min.

As seen from Fig. 4.4, of the three pH values, pH 7.2 gave the slowest reaction. Reactions at pH 8.2 and 8.8 were much faster than that at pH 7.2, and the two were comparable in velocity with pH 8.8 giving a slightly faster reaction as judged by

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

the consumption rate of the substrate. However, more hydrolysis was observed at pH 8.8. For this reason, pH 8.2 was used in subsequent experiments. One should also note that a substantial amount of hydrolysis product was formed at all three pHs and that lowering the pH did not improve the thiolysis/hydrolysis ratio. All these may be due to that, at low pH, not only is it difficult to achieve a high overall thiol concentration, but the actual amount of the hydrosulfide ion is also greatly reduced. Again, the peptide acid (Fig. 4.4, peak a) appears to be a result of direct hydrolysis of the peptide ester as the ratio of thiolysis/hydrolysis remained almost unchanged over the reaction course in each case. From these results, one may conclude that the effective concentration of hydrosulfide anion nucleophile is crucial in determining the rate of the thiolysis reaction.

I also studied the effect of thiol concentration on thioacid formation reaction conducted with $(NH_4)_2S$ at pH 8.2. Within the range of the effective thiol concentration tested, which ranged from 0.038 to 0.3 M, we found that both the rate and yield of thioacid would increase with the increase of effective $[HS^-]$. The effective thiol concentrations were calculated by deducting the amount of TCEP from that of the total reducing reagents determined by Ellman's method (mentioned in the experimental section)[145]. At $[HS^-]$ = 0.15 M, ~15% of thioacid was obtained at 5 min of reaction, whereas at $[HS^-]$ = 0.3 M, nearly 40% of thioacid was formed. However, it became difficult to achieve and maintain an effective thiol concentration higher than 0.4 M without increasing the pH and/or buffer capacity, because of the rather basic nature of $(NH_4)_2S$ and the loss of H_2S that would result from pH adjustment. But increasing the pH would result in more hydrolysis, offsetting therefore the benefits of having a higher $[HS^-]$ (Figure 4.5).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

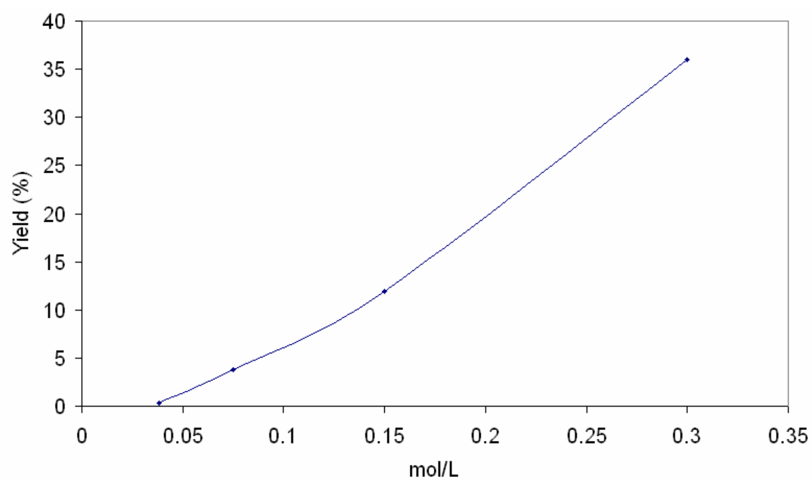


Figure 4.5 Thioacid formation by different concentration of $(NH_4)_2S$. Reaction condition: 1 μL of 2.5 mM ester, 0.3 μL of 45 μM subtiligase, 0.7 μL of 0.2 M sodium phosphate buffer, pH 8.2 and 18 μL thiol reaction solution including 0.038~0.3 M $(NH_4)_2S$ and indicated concentration of TCEP, respectively, in 0.58 M sodium phosphate buffer, pH 8.2 , 5 min.

To further demonstrate the synthetic value of subtiligase-catalyzed thiolysis for thioacid synthesis, I prepared three other peptide substrates with the glc ester linkage at the C-termini and tested them for thioacid formation (Table 4.1).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

N	Sequence	Yield [%]	T/H [a]	M_W [b] calc.	M_W [b] found
1	Ac-**KPGTVA**	57	3:1 (76%)	629.3	630.4
2	H-**TKGSAYSGKLEEFVQ**	44	1:1 (87%)	1658.8	1660.0
3	Ac-**KVLPNIQ**	56	13:5 (77%)	868.5	869.6

Table 4.1 Thioacid formation of other glycolate ester peptides. All peptides were CT-esters of glc**FG**-NH_2. a)T/H: thiolysis/hydrolysis ratio; values in brackets indicate the amount of ester peptide consumed. b) Calculated molecular weight and found ($[M+H]^+$) m/z value of the corresponding thioacid products. The thiol solution contained 0.2 M $(NH_4)_2S$ and 0.2 M TCEP in 0.58 M sodium phosphate buffer, pH 8.2. Reaction conditions: 1, 1 µL of 2.5 mM ester, 0.6 µL of 45 µM subtiligase, 0.4 µL H_2O and 18 µL reaction solution, 15 min; 2, 0.5 µL of 5 mM ester, 1.5 µL of 45 µM subtiligase, 18 µL reaction solution, 3 h; 3, 1 µL of 2.5 mM ester, 0.3 µL of 45 µM subtiligase, 0.7 µL H_2O, 18 µL reaction solution, 11 min.

The peptides listed in Table 4.1, plus the model peptide ester Ac-**HAAPF**-glc**FG**-NH_2, contain a diverse set of amino acid residues in the sequences. All these peptide esters were substrates of subtiligase which gave their corresponding thioacids in moderate yields. It was not necessary to protect the NT-amines. No intermolecular or intramolecular self-ligation was detected even though the N-terminus of one peptide (Table 4.1, peptide **2**) was not protected. Except for the competing hydrolysis reaction, the ester-to-thioacid conversion proceeded smoothly and cleanly without other detectable side reactions. A representative HPLC trace showing the conversion of the glc ester peptide **2** (Table 4.1) to its thioacid at pH 8 is presented in Figure 4.6.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

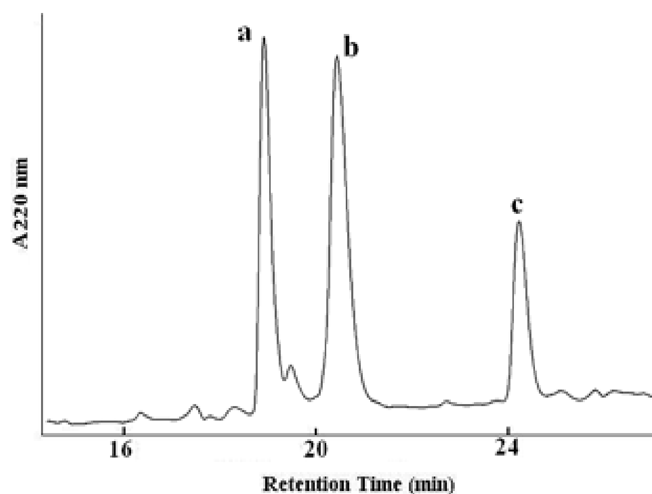


Figure 4.6 HPLC diagram for subtiligase-catalyzed thiolysis of peptide 2 (from Table 4.1) by $(NH_4)_2S$. See table 4.1 for reaction conditions. Peak a: hydrolysis product (m/z $[M+H]^+$ found: 1644.1, M_w calcd: 1642.8); peak b: newly formed thioacid (m/z $[M+H]^+$ found: 1660.0, M_w calcd: 1658.8); peak c: starting material (m/z $[M+H]^+$ found: 1904.7, M_w calcd: 1903.9). HPLC gradient: 10-40% buffer B in 30 min.

Thioacid products prepared by this enzymatic method are useful building blocks for chemical ligation. For instance, Ac-His-Ala-Ala-Pro-Phe-SH, the thioacid product of Ac-His-Ala-Ala-Pro-Phe-glc**FG**-NH_2, was tested for ligation with H-Cys(Npys)-Phe-Glu-Val-Lys-Gly-NH_2 using our “mini thiol-capture ligation” strategy, which gave almost quantitative yield of ligation product after 5 min of reaction (Fig. 4.7).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

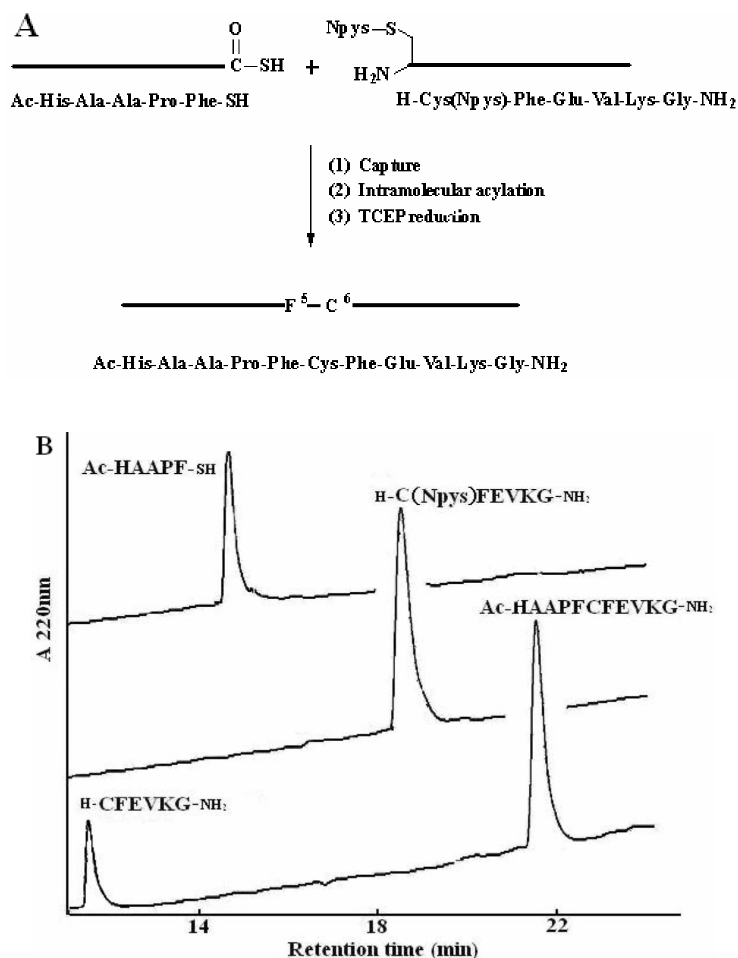


Figure 4.7 **A)** Synthesis of Ac-**HAAPFCFEVKG**-NH_2 by ligation of Ac-His-Ala-Ala-Pro-Phe-SH and H-Cys(Npys)-Phe-Glu-Val-Lys-Gly-NH_2. **B)** HPLC monitoring of the ligation reaction. Upper trace: thiol-acid peptide Ac-His-Ala-Ala-Pro-Phe-SH (m/z $[M+H]^+$ found: 600.4, M_w calcd: 599.3); middle trace: H-Cys(Npys)-Phe-Glu-Val-Lys-Gly-NH_2 (m/z $[M+H]^+$ found: 835.3, M_w calcd: 834.3); lower trace: TCEP-treated ligation mixture after 5 min. The peak at 21.7 min is the ligation product, Ac-**HAAPFCFEVKG**-NH_2 (m/z $[M+H]^+$ found: 1246.5, M_w calcd: 1245.6). The 12.3 min peak is H-**CFEVKG**-NH_2.

4.2.2 Redesign of Subtiligase for the Synthesis of Peptide Thioacids

Comparing Figure 4.4 (thioacid formation, page 84) with Figure 3.5 (thioester formation, page 56), one can see that there indeed was more hydrolysis product in the thioacid formation than in the thioester formation under the same reaction conditions. In order to make the enzymatic thioacid formation reaction to prefer

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

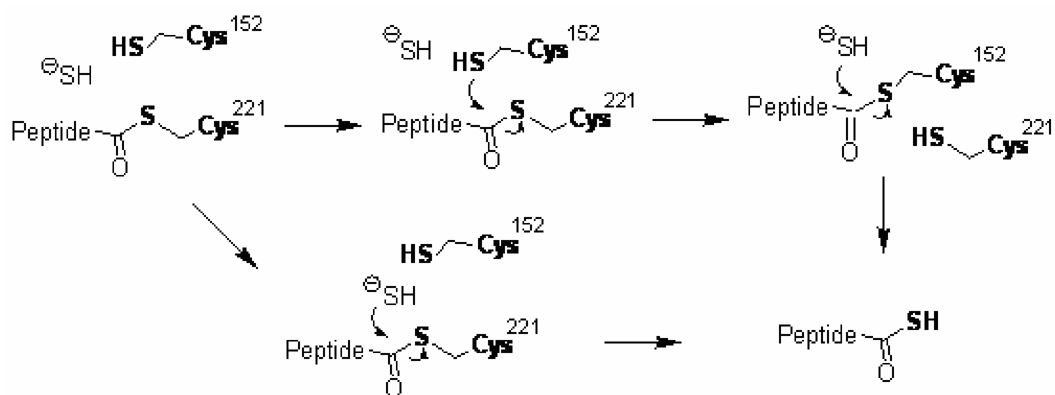
thiolysis to hydrolysis, some mutations were introduced into subtiligase. Several subtiligase variants, subtiligase P225G, D32C, S125C, A152C and N155C, were produced by our group. We speculate that, among them, except for designing P225G whose aim is to make much more room for the substrate binding compared with subtiligase (P225A), each of the other mutations was designed to introduce a new Cys residue which is spatially close enough to the catalytic Cys221 residue, and all of these mutations should possibly share the same mechanism to increase the ratio of thiolysis/hydrolysis (see Scheme 4.5 for details, page 91). All of them were tested in the thioacid formation reactions. However, only one, subtiligase A152C, showed some improved activity.

From Figure 3.2 (page 49), we can see how subtilisin uses a water molecule as a nucleophile to break the amide bond of a peptide, in which, water is deprotonated by the His64 residue to initiate the hydrolysis reaction. As a S221C/P225A mutation of subtilisin, subtiligase still uses the same mechanism, through His64, to start the hydrolysis of the ester peptide[103]. However, since the acyl-enzyme intermediate of subtiligase is a thioester in nature and water is a hard nucleophile compared with a free thiol, it is relatively difficult for a water molecule to attack the thioester bond, which reduces the hydrolysis ability of subtiligase.

On the other hand, as a soft nucleophile, a free thiol anion (HS^-) has the ability to react with the thioester in a thiol-thioester exchange reaction in aqueous solution to form a thioacid (proved and mentioned in the segment 4.2.3, page 96). Therefore HS^- can attack the thioester-enzyme intermediate by exchange reaction to form the thioacid. However, the reaction speed will be slower since it does not have the assistance of the enzyme to directly accelerate the thioacid formation

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

(Figure 3.7, page 60) and it may require more amount of A152C than subtiligase to form the same amount of thioacid. The possible catalytic mechanism of subtiligase A152C is shown in Scheme 4.5.



Scheme 4.5 Possible mechanism of subtiligase A152C in action. When the acyl-enzyme intermediate is formed, the peptidyl acyl group may migrate from Cys221 to Cys152 (the upper one) and this exchange prevents the newly formed Cys152-thioester from hydrolysis by subtiligase A152C since it is farther away from the enzymatic active site (His64) compared with the original Cys221-thioester. Moreover, the HS^- can still spontaneously attack the Cys152-thioester to release the thioacid product. Of course, it is also possible that subtiligase A152C can perform in the same way as subtiligase to generate the thioacid peptide.

We speculate that, when the acyl-enzyme intermediate is formed, the peptidyl acyl group migrates to Cys152 from Cys221 since these two residues are spatially close enough (4.3 Å, 3D structure information was acquired from NCBI database). Then, because the newly formed Cys152-thioester is farther away from the enzymatic active site (His64) than the original Cys221-thioester, the Cys152-thioester is less likely to be hydrolyzed by subtiligase A152C. Moreover, the HS^- still can spontaneously attack the Cys152-thioester to release the thioacid product which is relatively stable to subtiligase. Therefore, the new A152C mutant has one more pathway to produce the thioacid (the other one is the same as

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

subtiligase) while it has only one way to perform catalyzed hydrolysis which would be partially blocked by the thioester present on Cys152. As a result, the A152C mutant should give improved thiolysis/hydrolysis ratio.

This hypothesis was verified by performing the thioacid formation reactions catalyzed by the A152C mutant or subtiligase in the $(NH_4)_2S$ buffer. Compared with the subtiligase, A152C did improve the yield of thioacid (peak b), as one can see from the Fig. 4.8, trace 1 and 2. Moreover, after 20 min incubation, A152C generated much less hydrolysis than subtiligase (Fig. 4.8, trace 3 and 4). It must be mentioned that, in the above reactions, the concentration of A152C was 6-fold to that of subtiligase and this ratio was fixed in the following experiments.

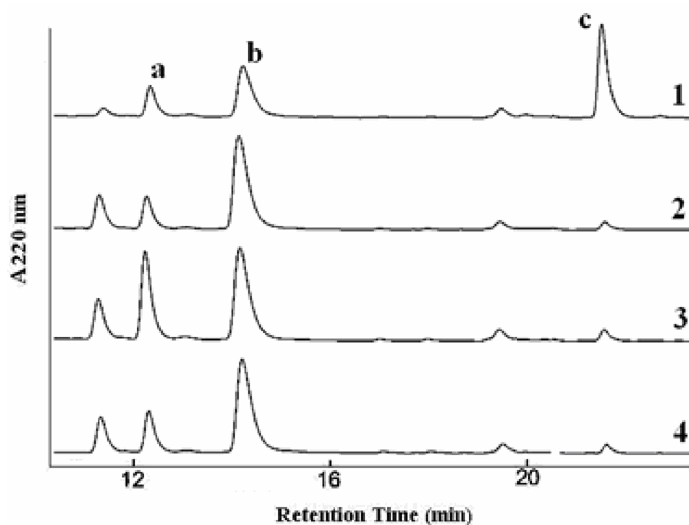


Figure 4.8 HPLC monitoring of enzymatic thioacid formation of Ac-**HAAPF**-glc**FG**-NH_2 with 0.4M of $(NH_4)_2S$ at pH 8. See experimental segment for detailed reaction conditions. Trace 1, by subtiligase for 4 min; trace 2, by A152C for 4 min; trace 3, by subtiligase for 20 min; trace 4, by A152C for 20 min. Peak a: hydrolysis product, Ac-**HAAPF**-OH (m/z $[M+H]^+$ found: 584.5, M_w calcd: 583.6); peak b: the newly formed thiol-acid product, Ac-**HAAPF**-SH (m/z $[M+H]^+$ found: 600.4, M_w calcd: 599.6); peak c: the ester starting material (m/z $[M+H]^+$ found: 845.1, M_w calcd: 844.4). HPLC gradient: 10-35% Buffer B for 25 min.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Form Fig. 4.8, one can clearly see that the ratio of thiolysis/hydrolysis was improved significantly with subtiligase A152C. Furthermore, I provided more evidence to support our hypothesis.

Firstly, I tested the hydrolysis reaction for the ester peptide, Ac-**HAAPF**-glc**FG**-NH_2, by subtiligase and A152C, respectively, at pH 8 for 10 min. Although the concentration of A152C was still 6-fold to that of subtiligase, A152C did hydrolyze the substrate much slower than subtiligase (Fig. 4.9). This confirmed that the A152C mutation almost abolished the hydrolase activity of subtiligase towards a peptide ester substrate.

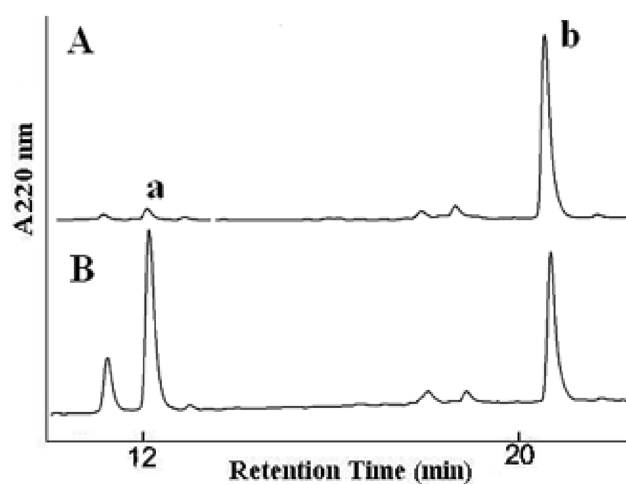


Figure 4.9 HPLC monitoring of enzymatic hydrolysis of Ac-**HAAPF**-glc**FG**-NH_2 by A152C (A) or subtiligase (B). See experimental segment for detailed reaction conditions. Reaction was performed at pH 8 for 10 min. Peak a: hydrolysis product, Ac-**HAAPF**-OH (m/z $[M+H]^+$ found: 584.5, M_w calcd: 583.6); peak b: the ester starting material (m/z $[M+H]^+$ found: 845.5, M_w calcd: 844.4). HPLC gradient: 10-35% Buffer B for 25 min.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Secondly, I compared subtiligase and A152C for ligation reaction between Ac-**HAAPF**-glc**FG**-NH_2 and H-**AFA**-NH_2 (Fig. 4.10). The ligation ability of A152C was also largely reduced.

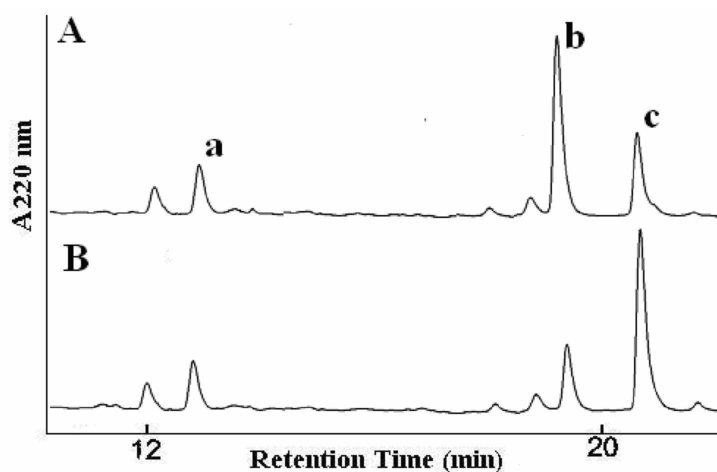


Figure 4.10 HPLC monitoring of enzymatic ligation between Ac-**HAAPF**-glc**FG**-NH_2 and H-**AFA**-NH_2 catalyzed by subtiligase (A) or A152C (B). See experimental segment for detailed reaction conditions. Reaction was performed at pH 8 for 45 min. Peak a: hydrolysis product, Ac-**HAAPF**-OH (m/z $[M+H]^+$ found: 584.5, M_w calcd: 583.6); peak b: the ligation product Ac-**HAAPFAFA**-NH_2 (m/z $[M+H]^+$ found: 872.3, M_w calcd: 871.4); peak c: the ester starting material (m/z $[M+H]^+$ found: 845.5, M_w calcd: 844.4). HPLC gradient: 10-35% Buffer B for 25 min.

The largely reduced efficiency of ligation activity as well as hydrolysis ability of the A152C mutant indicated that this mutation could limit the nucleophile, either H-**AFA**-NH_2 or H_2O molecule, to attack the acyl-enzyme intermediate to form the hydrolysis or ligation products. However, it was found that A152C could still engage a soft nucleophile like $[RS^-]$ and $[HS^-]$ for thioester and thioacid formation.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Thirdly, subtiligase and A152C were assayed in the thioester formation reactions by using ACA (N-acetylcysteamine) at pH 7 for 5 min, at this time point there were almost no hydrolysis products and reaction results directly reflected the yield of the thioester products. As shown in Figure 4.11, A152C generated a little bit more thioesters than subtiligase. This result led us to make the conclusion that in the reactions catalyzed by A152C, the reason that nucleophile-thiol anions, [RS^-] and [HS^-], can work but water and amine nucleophiles cannot is that [RS^-] and [HS^-] are both soft nucleophiles which can effectively react with a soft electrophile such as the acyl-enzyme thioester intermediate.

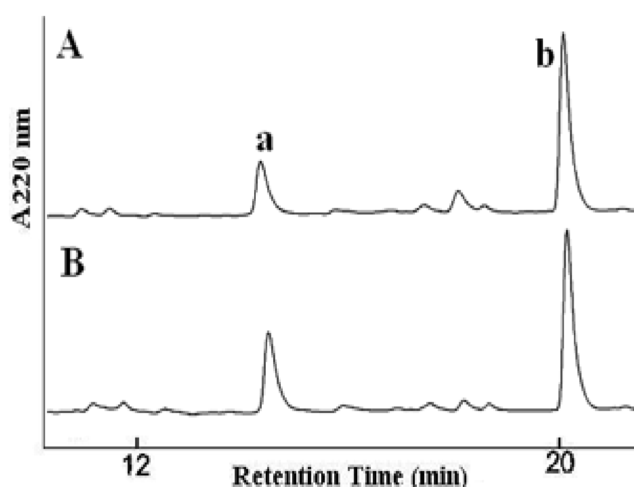


Figure 4.11 HPLC monitoring of enzymatic ACA-thioester formation catalyzed by subtiligase (A) or A152C (B). See experimental segment for detailed reaction conditions. Reaction was performed at pH 7 for 5 min. Peak a: thioester product, Ac-**HAAPF**-S-ACA (m/z $[M+H]^+$ found: 685.1, M_w calcd: 684.3); peak b: the ester starting material (m/z $[M+H]^+$ found: 845.5, M_w calcd: 844.4). HPLC gradient: 10-35% Buffer B for 25 min.

The above results provided strong, *albeit* indirect, evidence that for subtiligase A152C, the Cys152 residue participated in a thiol-thioester exchange relay reaction

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

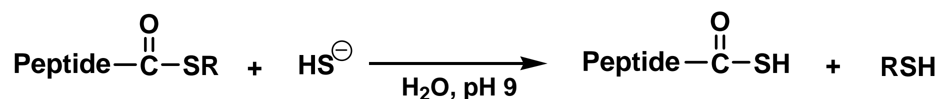
by accepting the acyl group from the active site Cys221 residue as proposed in Scheme 4.5 (page 91). However, further evidence such as structure study would be needed to support this conclusion. Nevertheless, all the evidence mentioned above indicated that the A152C is a good enzyme to produce peptidyl thioacids, although more amount is needed than subtiligase to produce the same amount of thioacids.

4.2.3 A Simple Chemical Method for the Synthesis of Peptide Thioacids

Although the subtiligase and its variants can be utilized in the synthesis of thioacids, the method has limitations. For example, subtiligase has substrate preference although with a wide range (subtilisin has a relatively broad substrate specificity, which is paralleled by subtiligase[103]). Moreover, the competing hydrolysis is also a problem. Additionally, longer peptides may have difficulties to be dissolved well in the enzymatic reaction buffers at the operating pH. Therefore, I designed another method to produce thioacids.

In this last segment, I report a simple and efficient method to produce peptide CT-thioacids through the hydrothiolysis of thioesters (Scheme 4.6). The peptide thioacids thus prepared can be used for peptide ligation by the mini thiol capture ligation strategy. More importantly, a subtiligase-insensitive thioacid group makes it possible to conduct sequential chemoenzymatic ligation through alternate use of mini thiol capture and enzymatic ligations.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



Scheme 4.6 Conversion of a peptide thioester to a thioacid through hydrothiolysis in aqueous buffer.

I firstly tested this reaction in a simple system with a small peptide thioester, Ac-**HAAPF**-S$(CH_2)_2$-COOH. Treating this thioester peptide (*ca.* 0.7 mM) with NaSH (*ca.* 120 mM SH^-) in 1 M phosphate buffer (pH 9) at 42 °C for 2 h led to near quantitative conversion to the thioacid (Fig. 4.12) with only a very small amount of hydrolysis (Fig. 4.12, peak c).

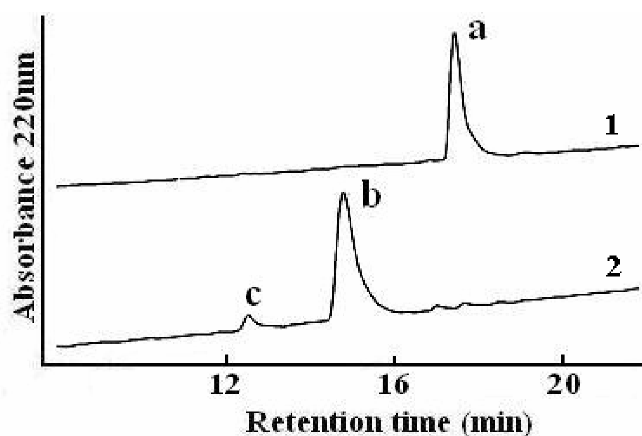


Figure 4.12 HPLC profile of the starting material Ac-**HAAPF**-S-$(CH_2)_2$-COOH (trace 1, peak a) (m/z $[M+H]^+$ found: 672.5, M_w calcd: 671.7), and of the hydrothiolysis reaction mixture at 2 h (trace 2). Peak b is the product, Ac-**HAAPF**-SH (m/z $[M+H]^+$ found: 600.4, M_w calcd: 599.6).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

When the reaction was conducted at 23 °C, the hydrothiolysis rate decreased significantly (Fig. 4.13 A). Increasing the pH (pH 8 to 10) also increased the reaction rate (Fig. 4.13 A). It became practically difficult to conduct the reaction when the pH value was adjusted to 7 or below, because of escape of H_2S from the reaction medium during the process of pH adjustment. Therefore for practical reasons, the reaction was usually performed at pH 8.5 - 9. Na_2S was also an excellent reagent for this reaction, which gave just a slightly slower reaction than did NaSH (Fig. 4.13 B). $(NH_4)_2S$ was comparable to Na_2S under the same reaction conditions (data not show). It is not surprising that similar results were obtained with these different sulfide compounds, as, at these pHs, the hydrosulfide ion HS^- ought to be the effective exchange agent for all these sulfide compounds considering the p*K*a value of H_2S (pK_{a1} 6.9, pK_{a2} 17.1[144]).

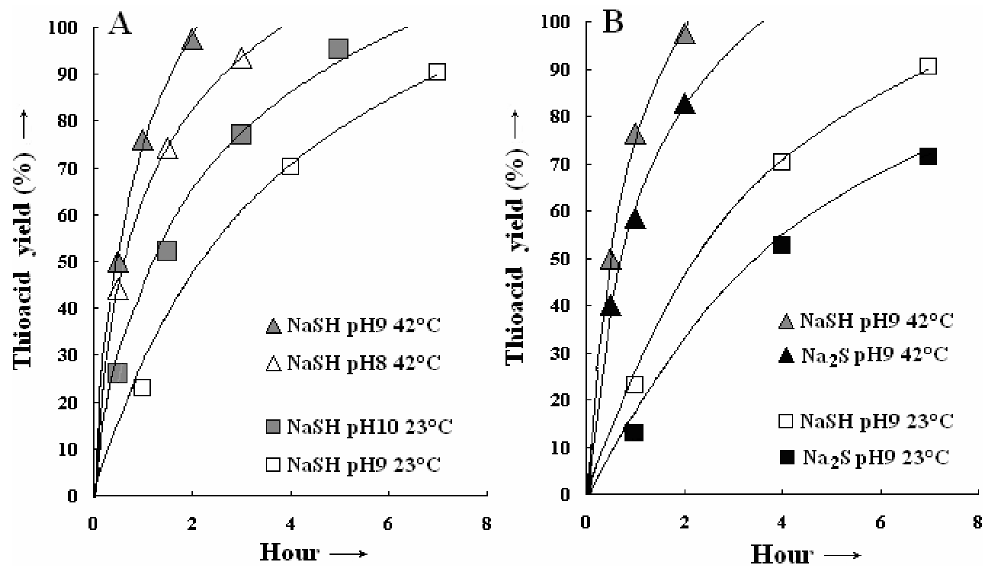


Figure 4.13 Thioacid formation under different reaction conditions. General reaction conditions: 0.7 mM Ac-**HAAPF**-S-$(CH_2)_2$-COOH, 120 mM NaSH or Na_2S (effective hydrosulfide ion) and incubated at 23 °C or 42 °C. Panel A shows the reaction under different temperatures or pH using NaSH, while B compares the reactions of NaSH and Na_2S.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

To demonstrate the general utility of this method, we prepared a number of peptide thioesters and tested these in the hydrothiolysis reaction. The results are summarized in Table 4.2.

No	Peptide thioester substrates [a]		New formed thiol-acid peptides		
	Sequence	Thioester component	Yield [b] [%]	M_W calc.	*m/z* $[M+H]^+$ found
1	H-**FSKLAV**	-S-$(CH_2)_2$-$CONH_2$	96 (4h)	679.4	680.4
2	H-**FSKLAI**	-S-$(CH_2)_2$-$CONH_2$	90 (4h)	693.4	694.5
3	H-**AFSKL**	-S-$(CH_2)_2$-$CONH_2$	95 (2h)	580.3	581.4
4	Ac-**LVKEI**	-S-$(CH_2)_2$-$CONH_2$	85 (8h) [c]	658.4	659.3
5	Ac-**HAAPF**	-S-$(CH_2)_2$-COOH	92 (2h)	599.3	600.4
6	H-**TKGSAYSGK LEEFVQ**	-S-$(CH_2)_2$-COOH	94 (3.5h)	1658.8	1660.0

Table 4.2 Thiol-acid formation reaction. [a] Reaction condition: thioester peptide (0.7 - 1 mM), NaSH (120 mM effective hydrosulfide ion in 1 M sodium phosphate buffer, pH = 9) at 42 °C except for entry 4. [b] Yield was based on HPLC ananlysis and numbers in the brackets are the reaction time. [c] Na_2S (120 mM effective hydrosulfide ion in 1 M sodium phosphate buffer, pH = 9) at 23 °C.

In table 4.2, peptide 1, 2 and 4 contain, at the C-terminus, a sterically hindered amino acid residue with a β-branched side chain. Yet the hydrothiolysis reactions with these peptides all proceeded cleanly to give the corresponding thioacids in excellent yields. The CT-residues in peptide 3, 5 and 6 are less hindered, and the reactions with these peptides took shorter time to complete. It is worth noting that all these peptides were unprotected on the side chains and therefore had good solubility in aqueous buffer. We also found that the reaction could be conducted

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

under denatured conditions, such as in the presence of high concentrations of urea or guanidine•HCl. For instance, hydrothiolysis of peptide 1 was also performed in the presence of 8 M urea, which gave clean formation of the thioacid product (data not show). This would be useful for long hydrophobic peptides which tend to aggregate in aqueous buffer under native conditions.

As a soft electrophile, a thioester is known to have the unique ability to react with a soft nucleophilic free thiol in a thiol-thioester exchange reaction in aqueous solution to form a new thioester. This transthioesterification was first reported by Wieland, *et al.* as early as in 1953 as the mechanism responsible for selective peptide bond formation between an amino acid- or peptide thioester and a cysteinyl peptide in aqueous buffer[146]. Later, it was also found in thiol-promoted deacylation of penicilloylated penicillin binding proteins[147] and in the spontaneous exchange between CoA (a thiol) and acetoacetylethanethiol (a thioester)[148]. Most notably, exploitation of this thiol-thioester exchange reaction has led to the development of a powerful method for protein synthesis, in which the CT-thioester of one peptide reacts, by a two-step mechanism similar to that originally proposed by Wieland[146], with the free NT-Cys residue of a second peptide to form a native Xaa-Cys peptide bond at the ligation site[60]. Chemoselective ligation with this method was first demonstrated with peptide C^{α}-thioaryl esters[60], and later with the more easily obtainable thioalkyl esters[89]. Subsequent studies have shown that the ligation reaction with alkyl thioesters can be catalyzed by other thiol additives through in situ transthioesterification of the thioester with these additives[149], and the best results were obtained with certain substituted thiophenols that have good solubility in aqueous buffer[150]. Depending on the nature of thioester and that of the thiol, the exchange reaction can occur in a wide pH range with a higher pH generally giving a

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

greater reaction rate. What we have demonstrated herein is that, when a hydrosulfide ion is used instead as the thiol source, the reaction leads to the formation of a thioacid as the hydrothiolysis product of a thioester.

Peptide thioacids are the essential acyl component building blocks for the 'mini' thiol capture ligation strategy[124]. This ligation approach works in two steps: (1) specific capture of the thioacid sulfhydryl of the acyl peptide by an Npys-modified NT-Cys residue of the amine component peptide to bring the two peptide components together with a covalent acyl disulfide linkage. The capture step is usually conducted in an acidic buffer and occurs almost instantaneously; (2) intramolecular S,N-acyl transfer via a 6-member ring intermediate, leading to the formation of an amide bond upon adjusting the pH to 5-6[124]. A natural Cys residue is regenerated upon addition of a reducing agent such as TCEP or DTT at the end of the reaction. A distinct feature of this ligation method is its high efficiency, which allows the entire ligation reaction to complete in 5 - 20 min for sterically unhindered peptide thioacids with, e.g., a CT- Ala or Phe residue. Although conceptually similar to the early work of Kemp[129], this method does not use a fused-ring template to mediate the intramolecular N-acylation reaction and therefore is much easy to use practically especially in view of the now convenient availability of peptide thioacids described in this report.

To test the chemical activity of the newly formed peptidyl thioacids, peptide ligation reaction between the newly formed peptide thioacid Ac-**HAAPF**-SH and H-**C**(Npys)-**GRGVPNGIPAEDSEQLASGQ**-NH_2 was conducted (Fig. 4.14).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

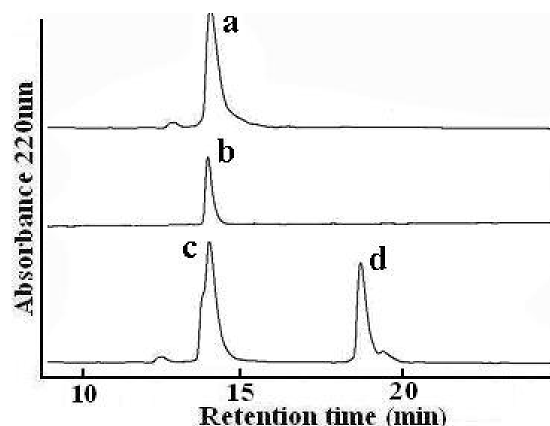


Figure 4.14 HPLC data of ligation reaction between Ac-**HAAPF**-SH and H-**C**(Npys)-**GRGVPNGIPAEDSEQLASGQ**-NH_2 (Npys: 3-nitro-2-pyridinesulfenyl). Peak a: H-**CGRGVPNGIPAEDSEQLASGQ**-NH_2 (m/z $[M+H]^+$ found: 2084.5, M_w calcd: 2083.0); peak b: Ac-**HAAPF**-SH; peak c: unconsumed a and b; peak d: ligation product Ac-**HAAPFCGRGVPNGIPAEDSEQLASGQ**-NH_2 (m/z $[M+H]^+$ found: 2650.6, M_w calcd: 2648.2).

Furthermore, I found that two newly formed peptidyl thioacids, H-**FSKLAV**-SH and H-**FSKLAI**-SH, were not good substrates for the mini thiol capture ligation reaction. After 20 min reaction at pH 6 (the same reaction condition as that in Fig. 4.14, see experimental section for details), no significant amount of ligation products was detected (data not show). This is mainly because this two peptide thioacids contain β-branched amino acids at the C-termini (Val or Ile residue) which sterically hindered the intramolecular S, N-acyl transfer via a 6-member ring intermediate. For extending the applications of these kinds of peptidyl thioacids with β-branched amino acid at the C-terminus. I coupled the aminoacetaldehyde dimethyl acetal by activating H-**FSKLAV**-SH with DNTP and the resulting acyl disulfide was active enough to acylate the amino group of aminoacetaldehyde, in which an amide bond was formed (Fig. 4.15). Through this reaction, the powerful functional aldehyde groups can be easily introduced into the C termini of peptides.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

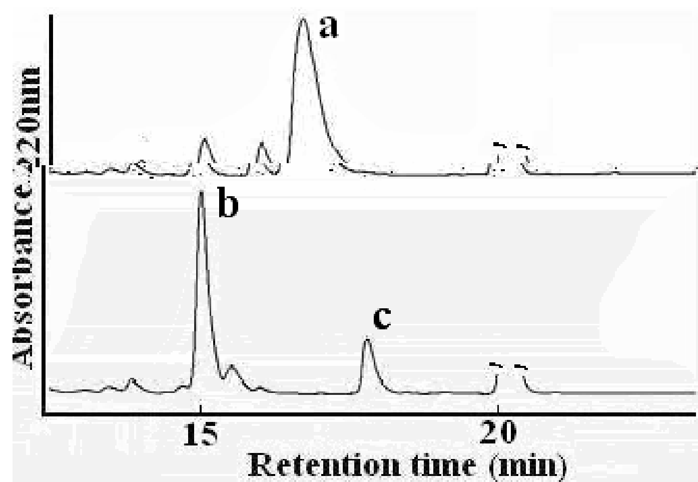
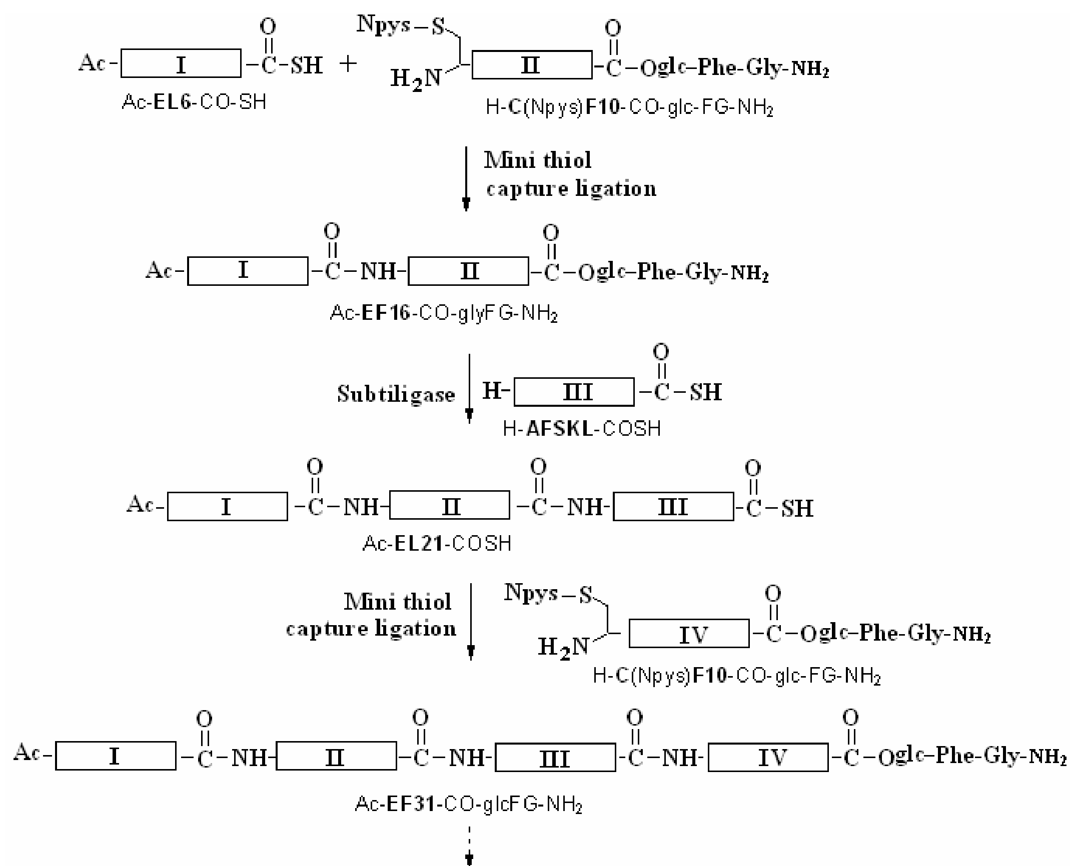


Figure 4.15 HPLC detection of the amide bond formation between the newly formed H-**FSKLAV**-SH and aminoacetaldehyde dimethyl acetal. Upper one is the reaction control and lower one is the reaction after 12 h. Peak a is H-**FSKLAV**-SH (m/z $[M+H]^+$ found: 680.4, M_w calcd: 679.4); peak b is the reaction product, H-**FSKLAV**-NH-CH(OCH_3)$_2$ (m/z $[M+Na]^+$ found: 773.1, M_w calcd: 750.4); peak c is unknown (m/z 668.6 found).

Moreover, since a negatively charged thiocarboxylate is unlikely to be recognized as the substrate for subtiligase, the hydrothiolysis reaction can hence render a subtiligase-sensitive thioester bond to a subtiligase-insensitive thioacid group, allowing a thioacid peptide with a suitable free amino terminus to be used as the amine nucleophile for subtiligase-catalyzed peptide ligation[142, 151] while keeping the thioacid functionality intact. This makes it possible to perform sequential chemoenzymatic ligation in alternate mini thiol capture and enzymatic ligation steps. To demonstrate the feasibility of this sequential chemoenzymatic ligation scheme, a model peptide was synthesized by three consecutive steps of chemical and enzymatic ligation (Scheme 4.7) in the N-to-C direction.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library



Scheme 4.7 Synthesis of a model peptide by sequential chemoenzymatic ligation. Reaction details are given in the following text content.

When the thioacid peptide I, Ac-**EL6**-CO-SH, was ligated with the Npys-modified peptide glycolate ester, H-**C**(Npys)**F10**-CO-glc-**FG**-NH_2 (peptide II), ~ 60% ligation product was obtained after 20 min ligation reaction (Figure 4.16, trace 1, page 106). Since the newly formed ligation product, Ac-**EF16**-CO-gly**FG**-NH_2, contains a substiligase-recognizable glycolate ester moiety, it could then be used in a second, subtiligase-catalyzed enzymatic ligation step with peptide III which contains a CT-thioacid, H-**AFSKL**-COSH. The ligation product, Ac-**EL21**-COSH, was obtained in ca. 75% yield after 2 h reaction (Fig.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.16, trace 2, page 106). Only a small amount of the thioacid was hydrolyzed (trace 2, peak g) during the course of the enzymatic reaction. On the contrary, when the corresponding peptide thioester, H-**AFSKL**-CO-$SCH_2CH_2CONH_2$, was used as the amine nucleophile, almost all the thioester linkage was hydrolyzed to give the peptide C^{α}-carboxylic acid (Data not show). This result demonstrated the advantage of having a CT-thioacid instead of a thioester for such a tandem chemoenzymatic ligation scheme, because the surviving thioacid group would allow another step of mini thiol capture ligation. Indeed, the new thioacid peptide, Ac-**EL21**-COSH, was ligated again with a fourth peptide segment, H-**C**(Npys)**F10**-CO-glc-**FG**-NH_2, to give Ac-**EF31**-CO-glc**FG**-NH_2 in 90% yield after 20 min reaction (Fig. 4.16, trace 3, page 106). One should note that the free thiol from the Cys residue in Ac-**EL21**-COSH did not interfere with mini thiol capture ligation because it is less nucleophilic and less reactive than the thioacid at the operating pH of the capture step. With the presence of the CT-glycolate ester bond in the new ligation product, in principle this sequential chemoenzymatic ligation scheme can still continue to proceed with a new enzymatic ligation step.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

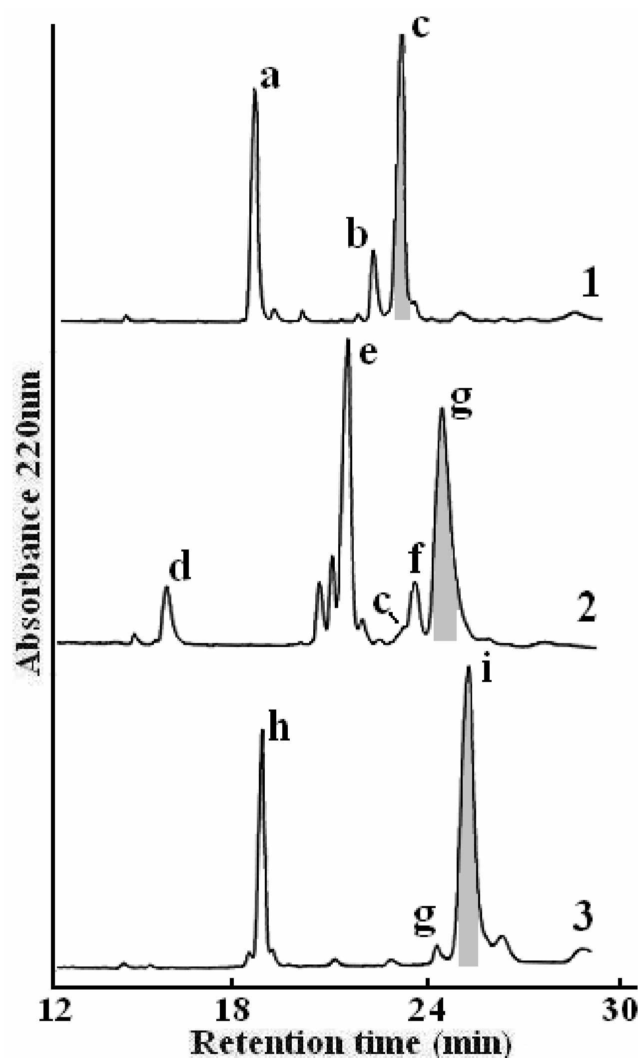


Figure 4.16 HPLC monitoring of the synthesis of Ac-**EF31**-CO-glc**FG**-NH_2 (Ac-**ELNKLLCSDSEHAAPFAFSKLCSDSEHAAPF**-glc-**FG**-NH_2) by three sequential steps of chemical and enzymatic ligation. Step 1 (trace 1): peak c is the product Ac-**EF16**-CO-gly**FG**-NH_2 (Ac-**ELNKLLCSDSEHAAPF**-glc-**FG**-NH_2 **(m/z** $[M+H]^+$ found: 2076.9, M_w calcd: 2075.9), generated by chemical ligation between Ac-**EL6**-COSH (Ac-**EKNKLL**-SH, peak b) and H-**C**(Npys)**F10**-CO-glc-**FG**-NH_2 (H-**C**(Npys)-**SDSEHAAPF**-glc-**FG**-NH_2). peak a is the H-**CSDSEHAAPF**-glc-**FG**-NH_2. Step 2 (trace 2): peak g is the Ac-**EL21**-COSH (Ac-**ELNKLLCSDSEHAAPFAFSKL**-SH **(m/z** $[M+H]^+$ found: 2377.9, M_w calcd: 2377.1), produced by ligation between Ac-**EF16**-CO-gly**FG**-NH_2 and H-**AFSKL**-COSH (peak d) in the presence of subtiligase. Peak e is the Ac-**ELNKLLCSDSEHAAPF**-OH and peak f is Ac-**ELNKLLCSDSEHAAPFAFSKL**-OH. Step 3 (trace 3), peak i is the Ac-**EF31**-CO-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 3667.6, M_w calcd: 3666.7), which was chemically ligated by Ac-**EL21**-COSH (peak g) and H-**C**(Npys)**F10**-CO-glc-**FG**-NH_2. peak h is identical to peak a.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

4.3 Summary and Conclusion

This project aims to develop novel methods to generate peptidyl thioacids.

Firstly, I have demonstrated herein that subtiligase is a viable catalyst to convert a peptide ester to a thioacid in aqueous media under conventional enzyme catalysis conditions. Since the glycolate ester peptides can be easily synthesized by using Fmoc SPPS and considering that subtiligase has a relatively broad substrate specificity[103], this enzymatic thiolysis reaction represents a useful alternative to the only available Boc-chemistry based method for the synthesis of peptide thioacids. To the best of our knowledge, this is the first time that such an enzymatic method was used for the preparation of a peptide thioacid. Subtiligase was originally designed as a peptide ligase for peptide bond formation, the results presented here and previously show that it can also function as a thiolase for thioacid and thioester formation. Although many enzyme redesign examples have been reported wherein the active site of an enzyme is engineered to catalyze a new chemical reaction[152], it is uncommon that the same reengineered enzyme functions to catalyze multiple reactions. The products of the enzymatic thiolysis reaction reported herein are the difficult-to-synthesize peptide thioacids, which again illustrates the power and versatility of enzymes as biocatalysts for organic transformations.

Secondly, to improve the thiolysis/hydrolysis ratio of subtiligase, we redesigned this enzyme and among several derivatives, the A152C mutant distinctly reduced its hydrolysis activity and still bears good thiolysis ability. As a result, it is a better enzyme than subtiligase for enzymatic thioacid formation reaction. Moreover, we proposed the reactive mechanism of A152C and furthermore provided some solid

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

evidences to support it. However, the enzymatic methods for thioacid formation have some shortcomings such as substrate specificity, small range of operating pH, temperature and ion concentration.

Finally, to overcome the problems encountered in the enzymatic thioacid formation reactions, I developed a novel chemical thiol-thioester exchange method to generated thioacids, which is a hydrothiolysis reaction of peptide thioesters. The reaction is highly efficient when conducted in an aqueous solution at alkaline pH. Since peptide thioesters can be easily obtained from the various currently available synthetic methods, this reaction can potentially become the method of choice for the preparation of peptide thioacids and provide the key input building blocks for peptide and protein synthesis by using, e.g., mini thiol capture ligation. Furthermore, the thioacid group is shown to be orthogonal to subtiligase-catalyzed enzymatic ligation. This makes it possible to conduct sequential chemical and enzymatic ligations of unprotected peptide segments. The synthesis of a model medium-sized peptide by a three-step ligation scheme has demonstrated the feasibility of this sequential chemoenzymatic ligation strategy and points to its potential in protein synthesis. The results obtained herein are likely to stimulate the development of novel protein synthesis strategies based on the use of peptide thioacids in the future.

4.4 Experimental

4.4.1 Materials

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

All chemical reagents were of analytical grade, obtained from commercial suppliers and used without further purification.

4.4.2 High Performance Liquid Chromatography

High Performance Liquid Chromatography (HPLC) was performed on an Agilent system (Agilent Tech., USA) with a Vydac C_{18} reverse-phase column (5 μm, 250 mm×4.6 mm). The analysis was carried out using the mixture of two solvents, A and B, as the mobile phase. Solution A was H_2O (deionized) containing 0.045% TFA and solution B was acetonitrile/H_2O (90/10) containing 0.04% TFA. The mobile phase flow rate was 1 ml-min^{-1} and the separation temperature was 25 ^{o}C. UV detection was carried out at 220 nm.

4.4.3 Mass Spectrometry

ESI-MS data were obtained on a Finnigan LCQ Deca XP MAX instrument for compounds of MW < 2000 Da. MALDI-TOF-MS data were obtained on an Applied Biosystems 4700 Proteomic Analyzer 72 for compounds of MW > 2000 Da.

4.4.4 Expression, Purification and Site-specific Mutation of Subtiligase.

This content is the same as the 3.4.4. The site-specific mutations were introduced into the subtiligase by using the Quick Change Site-Directed Mutagenesis Kit (Stratagene).

4.4.5 Peptide Synthesis.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

All peptides were synthesized using standard Fmoc chemistry[120] on Rink-amide-MBHA-resin or Tentagel amide resin. The glycolate ester peptides were prepared by Fmoc-based synthesis protocols[97]. For the synthesis of the H-**FSKLAV**-S-$(CH_2)_2$-$CONH_2$, H-**FSKLAI**-S-$(CH_2)_2$-$CONH_2$ and H-**AFSKL**–S -$(CH_2)_2$ $CONH_2$, Trt-S-$(CH_2)_2$-COOH was first loaded to Rink-amide-MBHA resin, followed by removal of trityl group by treating the resin with 2% TFA in DCM (containing 2% triisopropylsilane) for 5 x 3 min. Onto the free thiol group was then coupled Fmoc-Val-OH, Fmoc-Ile-OH and Fmoc-Leu-OH respectively by PyBOP and the peptide chain was elongated using a protocol developed by Li *et al.*[121], and after final cleavage and purification the targeted thioester was obtained in *ca.* 15% yield. The crude peptides were purified by reverse-phase HPLC (Vydac C_{18}). After lyophilization, peptides were obtained in powder form, which could be used freshly or stored at -20 ^oC until future use. Thioester peptides Ac-**LVKEI**-S-$(CH_2)_2$-$CONH_2$ and Ac-**ELNKLL**-S-CH_2CO-PheGly-NH_2 were prepared by Boc chemistry on MBHA resin and the peptides were cleaved from the resin using TFMSA/TFA. The identities of all peptides were confirmed by MS analysis. Calculated MW calculation is based on isotopic mass.

H-**CFEVKG**-NH_2 (m/z $[M+H]^+$ found: 681.0, M_w calcd: 680.3)

H-**CGRGVPNGIPAEDSEQLASGQ**-NH_2 (m/z $[M+H]^+$ found: 2084.5, M_w calcd: 2082.9)

Ac-**HAAPF**-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 845.5, M_w calcd: 844.4)

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

H-**TKGSAYSGKLEEFVQ**-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 1904.7, M_w calcd: 1903.9)

Ac-**KPGTVA**-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 875.4, M_w calcd: 874.4)

Ac-**KVLPNIQ**-glc**FG**-NH_2 (m/z $[M+H]^+$ found: 1114.6, M_w calcd: 1113.6)

H-**CSDSEHAAPF**-glc-**FG**-NH_2 (m/z $[M+H]^+$ found: 1324.6, M_w calcd: 1323.5)

H-**FSKLAV**-S-$(CH_2)_2$-$CONH_2$ (m/z $[M+H]^+$ found: 751.3, M_w calcd: 750.4)

H-**FSKLAI**-S-$(CH_2)_2$-$CONH_2$ (m/z $[M+H]^+$ found: 765.3, M_w calcd: 764.4)

H-**AFSKL**-S-$(CH_2)_2$-$CONH_2$ (m/z $[M+H]^+$ found: 652.3, M_w calcd: 651.3)

Ac-**LVKEI**-S-$(CH_2)_2$-$CONH_2$ (m/z $[M+H]^+$ found: 688.4, M_w calcd: 687.4)

4.4.6 Modification of NT-Cys Peptides by DTNP (2-2'Dithiobis-5-nitropyridine)

In a typical reaction, 4 equiv. of DTNP was dissolved in 400 µL acetic acid solution [(acetic acid: water = 3:1 (v/v)] and subsequently 1 equiv. of desired peptide (pre-dissolved in a small volume of 50% ACN/H_2O) was added respectively followed by vigorous stirring for 5 h. Another 600 µL pure acetic acid was added and followed by lyophilization. Then 100 µL of H_2O (0.1% TFA) was used to resuspend the dried powder, followed by 10 min sonication. The supernatant was collected after centrifugation twice at 4000 rpm for 10 min, and HPLC was utilized to purify the desired product which was confirmed by MS.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

H-**C**(Npys)-**GRGVPNGIPAEDSEQLASGQ**-NH_2 (m/z $[M+H]^+$ found: 2238.2, M_w calcd: 2237.0)

H-**C**(Npys)-**SDSEHAAPF**-glc-**FG**-NH_2 (m/z $[M+H]^+$ found: 1478.4, M_w calcd: 1477.6)

4.4.7 Thiol Quantification

5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB or Ellmans reagent[145]) has been used widely for the quantification of thiols[153]. In our case, sodium sulfide was dissolved in the degassed phosphate buffer (0.5 M, pH 8) to prepare the fresh solutions at concentrations of 12.5 - 100 mM. Then 1 μL of each of fresh sodium sulfide solutions at different concentrations was added into 999 μL of DTNB solution (2 mM in 25 mM phosphate buffer, pH 8) respectively. After 5 min, the standard absorbance-concentration plot was obtained by UV measurement at 412 nm after deducting the absorption of the control solution (1 μL phosphate buffer (0.5 M, pH 8) mixed with 999 μL of the DTNB solution.

Y=0.0223X-0.21 [Y is the $A_{412mm,}$ while X (mM) is the concentration of the hydrosulfide ion]

The concentration of hydrosulfide in the reaction buffer for hydrothiolysis was determined by using the above absorbance-concentration standard equation. For buffers containing >100 mM hydrosulfide ion, dilution was taken before reaction with Ellman's reagent.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

For the samples containing TCEP, The effective thiol concentrations were calculated by deducting the amount of TCEP from that of the total reducing reagents determined by above method.

4.4.8 Procedures for Enzyme Assays.

All peptide substrates were HPLC-purified before the following reactions. The typical reaction volume for an enzymatic reaction was 20-30 μL.

A. Thioacids formation catalyzed by subtiligase using H_2S as thiol source (Fig. 4.2).

H_2S gas was generated by the adding of 5% HCl solution dropwise on solid NaHS and subsequently the generated H_2S gas was introduced into ddH_2O and the final solution reached to pH 5 after 3 h. The concentration of factual sulfide ion was 68 mM detected by above method (4.4.7). Reaction condition: 3 μL Ac-**HAAPF**-glc**FG**-NH_2 (5 mM) and 0.5 μL subtiligase (45 μM) were mixed into above 26.5 μL reaction solution containing 68 mM sulfide ion, pH 5.0. During the whole procedure, the reaction system was kept in a small tube and, above the solution surface, the fresh H_2S gas was continuously filled into the tube. The reactions were stopped at different time intervals by mixing an aliquot of the reaction solution with 10% aqueous TFA solution before subjecting it to reverse phase HPLC analysis.

B. Thiolysis of Ac-**HAAPF**-glc**FG**-NH_2 with 0.2 M MAA and $(NH_4)_2S$ (Fig. 4.3)

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

1 μL of Ac-**HAAPF**-glc**FG**-NH_2 (2.5 mM) and 0.3 μL of 45 μM subtiligase were mixed with 18.7 μL of thiolysis buffer (pH 8) containing 0.58 M sodium phosphate buffer, 0.2 M TCEP and 0.2 M MAA or 0.2 M $(NH_4)_2S$. The reactions were stopped at different time intervals by mixing an aliquot of the reaction solution with 10% aqueous TFA solution before subjecting it to reverse phase HPLC analysis.

C. Thiolysis of Ac-**HAAPF**-glc**FG**-NH_2 with $(NH_4)_2S$ at different pH value (Fig. 4.4)

1 μL of Ac-**HAAPF**-glc**FG**-NH_2 (2.5 mM) and 0.3 μL of 45 μM subtiligase were mixed into 18.7 μL of thiolysis buffer containing 0.58 M sodium phosphate buffer, 0.36 M $(NH_4)_2S$ and indicated amount of TCEP was used to adjust the solution to pH 7.2, 8.2 or 8.8. The reaction was stopped at different time intervals by mixing an aliquot of the reaction solution with 10% aqueous TFA solution and followed by HPLC analysis.

D. Thiolysis of different glycolate ester peptides with $(NH_4)_2S$ (Table 4.1, Fig. 4.6)

1 μL of glycolate ester peptide (2.5 mM) and 0.3 μL of 45 μM subtiligase were mixed into 18 μL of thiolysis buffer containing 0.75 M sodium phosphate buffer, 0.4 M $(NH_4)_2S$ and 0.2 M of TCEP, pH 8. The reaction was stopped at different time intervals by mixing an aliquot of the reaction solution with 10% aqueous TFA solution and followed by HPLC analysis.

E. Thioacid formation of Ac-**HAAPF**-glc**FG**-NH_2 with $(NH_4)_2S$ by subtiligase or subtiligase-A152C (Fig. 4.8).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

1 μL of ester peptide (2.5 mM) and 0.3 μL of 45 μM subtiligase or 0.5 μL of 160 μM A152C were mixed into 18 μL of thiolysis buffer containing 0.75 M sodium phosphate buffer, 0.4 M $(NH_4)_2S$ and 0.2 M TCEP, pH 8. The reaction was stopped at different time intervals by mixing an aliquot of the reaction solution with 10% aqueous TFA solution and followed by HPLC analysis.

F. Enzymatic hydrolysis of Ac-**HAAPF**-glc**FG**-NH_2 by subtiligase or A152C (Fig. 4.9).

1 μL of ester peptide (2.5 mM) and 0.3 μL of 45 μM subtiligase or 0.5μL of 160 μM A152C were mixed into 18 μL of hydrolysis buffer containing 0.2 M tricine and 0.02 M TCEP, pH 8. The reaction was stopped after 10 min by adding 10% aqueous TFA solution and followed by HPLC analysis.

G. Enzymatic ligation of Ac-**HAAPF**-glc**FG**-NH_2 with H-**AFA**-NH_2 by subtiligase or A152C (Fig. 4.10).

1 μL of ester peptide (2.5 mM), 0.5 μL H-**AFA**-NH_2 (10 mM) and 0.3 μL of 45 μM subtiligase or 0.5μL of 160 μM A152C were mixed into 18 μL of ligation buffer containing 0.2 M tricine and 0.02 M TCEP, pH 8. The reaction was stopped after 45 min by adding 10% aqueous TFA solution and followed by HPLC analysis.

H. Enzymatic ACA-thioester formation of Ac-**HAAPF**-glc**FG**-NH_2 catalyzed by subtiligase or A152C (Fig. 4.11).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

1 μL of Ac-**HAAPF**-glc**FG**-NH_2 (2.5 mM) and 0.5 μL of 45 μM subtiligase or 0.5μL of 160 μM A152C were mixed into 18 μL of thiolysis buffer (pH 7) containing 0.5 M sodium phosphate buffer, 0.115 M TCEP and 0.23 M ACA. The reaction was stopped after 5 min by adding 10% aqueous TFA solution and followed by HPLC analysis.

4.4.9 Conversion of Peptidyl Thioesters to Thioacids through Hydrothiolysis in Aqueous Buffer.

A. Thioacid formation by exchange reaction between thioesters and NaSH (Fig. 4.12 and Table 4.2).

1μL of 12.5 mM thioester (produced by thioester formation reaction catalyzed by subtiligase, see chapter 3, or by solid phase synthesis) was mixed with 19 μL of 120 mM NaSH (in 1 M sodium phosphate buffer, pH = 9) for incubation at 42°C.

B. Thioacid formation by exchange reaction between thioester and NaSH or Na_2S at different pH values and reaction time (Fig. 4.13).

1μL of Ac-**HAAPF**-S-(CH2)$_2$-COOH, 12.5 mM, was mixed with 19 μL of 120 mM NaSH or Na_2S (dissolved in 1 M pH 8 sodium phosphate buffer followed by pH adjustment using 1 M pH 7 sodium phosphate solution to pH 8, 9 or 10 respectively) and incubated at 23 °C or 42 °C.

4.4.10 Peptide Chemical Ligation Procedures.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

A. Chemical ligation between an enzymatically synthesized peptide thioacid and H-**C**(Npys)**FEVKG**-NH_2 (Fig. 4.7)

2 μL of 1.5 mM Ac-**HAAPF**-SH (dissolved in HPLC buffer A), prepared by subtiligase catalysis from corresponding ester peptide, was mixed with 6 μL of 1.5 mM H-**C(Npys)FEVKG**-NH_2 (in buffer A) and immediately a yellow color was developed. After 2 min, 22 μL of 0.5 M sodium acetate (pH 6, degassed) was added into the reaction solution and the incubation was permitted for another 10 min. 2 μL of 1 M TCEP was subsequently added and 2 min later 8 μL of 50% TFA was added to quench the reaction and totally 40 μL reaction solution was injected into HPLC system for analysis.

B. Ligation reaction between new formed peptide thioacid Ac-**HAAPF**-SH and H-**C**(Npys)-**GRGVPNGIPAEDSEQLASGQ**-NH_2 (Fig. 4.14).

2μL of Ac-**HAAPF**-SH (2.5 mM) was mixed with 6 μL H-**C**(Npys)-**GRGVPNGIPAEDSEQLASGQ**-NH_2 (5 mM in HPLC buffer B) and immediately a yellow color developed. After 2 min, 22 μL of 0.5 M sodium acetate (pH 6, degassed) was added into the reaction solution and the incubation was permitted for another 10 min. 2 μL of 1 M TCEP was subsequently added and 2 min later 8 μL of 50% TFA was added to quench the reaction and totally 40 μL reaction solution was injected into HPLC system for analysis.

C. Amide-bond formation reaction between newly formed peptide CT-thioacid and aminoacetaldehyde dimer (Fig. 4.15).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

2μL of the newly formed H-**FSKLAV**-SH or Ac-**HAAPF**-SH (2.5 mM in HPLC buffer A) and 50 μg DTNP were suspended in 18 μL HPLC buffer A followed by sonication for 5 min. 10 μL of aminoacetaldehyde dimethyl acetal was added into the reaction solution and was permitted incubated at room temperature for 12 h (for H-**FSKLAV**-SH) or 40 min (for Ac-**HAAPF**-SH). Before the HPLC detection, 10 μL of 50% TFA was added to quench the reaction and reaction was monitored by HPLC and ESI-MS.

D. Synthesis of Ac-**EF31**-CO-glc**FG**-NH_2 by sequential chemoenzymatic ligation.

Step 1: 6 μL of 20 mM H-**C(Npys)F10**-CO-glc**FG**-NH_2 (dissolved in 50% ACN/H_2O, 0.01% TFA) and 2 μL of 20 mM Ac-**EL6**-COSH in 50% ACN/H_2O (0.01% TFA) were mixed together, and immediately a yellow color developed. After 2 min, 22 μL of 0.5 M sodium acetate (pH 6, degassed) was added. 20 min later, 2 μL of 1 M TCEP was added for two more min incubation. Finally, 8 μL of 50% TFA/H_2O was added to quench the reaction, and the entire reaction solution was injected into HPLC system to isolate the ligation product, Ac-**EF16**-CO-glc**FG**-NH_2.

Step 2: Ac-**EF16**-CO-glc**FG**-NH_2 (1 mM) was ligated with H-**AFSKL**-COSH (3 mM) under the catalysis of subtiligase (5 μM) in 0.15 M Tricine buffer (pH 8.0) containing 7.5 mM TCEP. After 2 h at RT, the reaction was quenched by addition of 10 μL of 50% TFA, and the ligation product, Ac-**EL21**-COSH, was purified by HPLC.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Step 3: Ac-**EL21**-COSH (1.25 mM) and H-**C(Npys)F10**-CO-glc**FG**-NH_2 (3.75 mM) were mixed together in 50% ACN/H_2O (0.01% TFA). The following procedures were the same as in step 1. The ligation product, Ac-**EF31**-CO-glc**FG**-NH_2, was purified by HPLC.

Ac-**EL6**-COSH: Ac-**EKNKLL**-SH

H-**C(Npys)F10**-CO-**glcFG**-NH_2: H-**C(Npys)-SDSEHAAPF-glc-FG**-NH_2

Ac-**EF16**-CO-**glcFG**-NH_2: Ac-**ELNKLLCSDSEHAAPF-glc-FG**-NH_2

Ac-**EL21**-COSH: Ac-**ELNKLLCSDSEHAAPFAFSKL**-SH

Ac-**EF31**-CO-**glcFG**-NH_2:
Ac-**ELNKLLCSDSEHAAPFAFSKLCSDSEHAAPF-glc-FG**-NH_2

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Summary

Because natural proteins usually contain only one single free N- and C-terminus, the modification targeting the N- or C-terminus is obviously a more straightforward way to introduce a functional moiety at a single site into a protein than the reactions carried out on the amino acid side chains. Therefore, in this thesis, we focused on developing specific enzymatic and chemical modification reactions on protein or peptide N- and C-terminus.

1. On NT-modification, I investigated the transamination reaction, a highly selective protein modification technique, which site-exclusively changes the N-terminus of peptide or proteins to a 2-oxoacyl moiety. In principle, any protein with a free NT-amine can be transaminated and the resultant carbonyl from the 2-oxoacyl moiety can be used in a subsequent bioconjugation reaction through the formation of a hydrazone, oxime or thiazolidine linkage. Despite some successful applications, the number of proteins that have been modified by using this reaction remains limited, which clearly does not match up with the seemingly large potential of this rather simple technique. One major issue is that, although a wide range of NT-residues can be transaminated from the above studies, it is often difficult to predict whether and how readily a particular amino acid can undergo transamination in the context of a polypeptide chain. Therefore, in my thesis, I performed a systematic and extensive investigation on the transamination reaction. First, I carried out a systematic solid phase study on the relative reactivity of membrane-bound dipeptides towards transamination by comparing the color intensities of the dyed spots. Second, a series of solution reactions were performed on selected model peptides to verify the findings from the above solid phase

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

studies. This work revealed the complicated nature of the transamination reaction. Nevertheless, some interesting observations are noteworthy. Except for Pro and His, most amino acids at the N-terminus can readily undergo transamination. However, the nature of the side chains greatly influences the stability and therefore the destiny of the newly formed 2-oxoacyl functionality in the reaction milieu. For amino acids with non-functionalized aliphatic side chains and for Ser, Thr and Met, the transamination products are relatively stable and can be isolated as the major products. All other amino acids seem to generate a product that is unstable, to varying degrees, in the reaction milieu and may lead to various side products, many of which are unidentified. This highlights the need for developing new transamination techniques that use mild reaction conditions and therefore cause fewer side reactions. The effect of the neighboring residues was also investigated, with the most interesting observation being that a Pro or His at position 2 would strongly inhibit the reaction. Our study will provide useful guidelines in the future use of the transamination reaction for protein modification and bioconjugation applications.

2. On CT-modification, I focused on developing new methods to produce peptidyl C^{α}-thioesters and C^{α}-thioacids, which are important building blocks in the total or semi-synthesis of peptides and proteins. Currently, peptidyl C^{α}-thioesters are prepared by solid phase peptide synthesis using Boc chemistry[90, 93], or Fmoc chemistry[94-96]. For the peptidyl C^{α}-thioacids, the only available approach for solid phase synthesis of peptidyl thioacids is based on Boc chemistry and no Fmoc based methods have been developed. The Boc approach requires the use of a special benzhydryl linker and HF for final deprotection and cleavage[132, 133]. To

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

make peptidyl C^{α}-thioesters and C^{α}-thioacids more accessible, I developed new enzymatic or chemical methods for their synthesis.

For the peptidyl C^{α}-thioesters, I demonstrated that subtiligase, a reengineered protease, was a suitable catalyst to convert a peptide ester to thioester in aqueous media by a kinetically controlled mechanism. To the best of our knowledge, this is the first time that a peptide thioester was prepared in such a way under conventional enzyme catalysis conditions. The relatively wide operating pH range makes it applicable to medium-sized peptides that are water-soluble. Of course, as for any enzymatic reaction, substrate specificity will be a major factor limiting the general use of this methodology. However, compared to other proteases, subtilisin has a relatively broad substrate specificity, which is paralleled by subtiligase[103]. This will make the enzymatic thiolysis reaction a useful alternative to the chemical methods now used for the synthesis of peptide thioesters.

For the peptidyl C^{α}- thioacids, I developed two new approaches, enzymatic and chemical methods, to generate peptidyl thioacids. Firstly, subtiligase was also found to be a viable catalyst to convert a peptide ester to a thioacid in aqueous media. Since the glycolate ester peptides can be easily synthesized by using Fmoc SPPS and considering that subtiligase has a relatively broad substrate specificity[103], this enzymatic thiolysis reaction represents a useful alternative to the only available Boc-chemistry based method for the synthesis of peptide thioacids. Although many enzyme redesign examples have been reported wherein the active site of an enzyme is engineered to catalyze a new chemical reaction[152], it is uncommon that the same reengineered enzyme functions to catalyze multiple reactions. The products of the enzymatic thiolysis reaction reported herein are the

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

difficult-to-synthesize peptide thioacids, which again illustrates the power and versatility of enzymes as biocatalysts for organic transformations. Moreover, to improve the thiolysis/hydrolysis ratio of subtiligase, we redesigned this enzyme and among several derivatives, the A152C has distinctly reduced hydrolysis activity but still bears moderate thiolysis ability. As a result, it is a good alternative to subtiligase for enzymatic thioacid formation reaction. Furthermore, we proposed a reactive mechanism of A152C, which involved a thioester relay reaction from the original active-site Cys221 to the new mutation-introduced Cys152. This mechanism was supported by some solid evidence from the hydrolysis, aminolysis and hydrothiolysis reactions.

Secondly, I developed a simple chemical thiol-thioester exchange method to generated thioacids, which is a hydrothiolysis reaction of peptide thioesters. The reaction is highly efficient when conducted in an aqueous solution at alkaline pH. Since peptide thioesters can be easily obtained from the various currently available synthetic methods, this reaction can potentially become the method of choice for the preparation of peptide thioacids and provide the key input building blocks for peptide and protein synthesis by using, e.g., mini thiol capture ligation. Furthermore, the thioacid group is shown to be orthogonal to subtiligase-catalyzed enzymatic ligation. This makes it possible to conduct sequential chemical and enzymatic ligations of unprotected peptide segments. The synthesis of a model medium-sized peptide by a three-step ligation scheme has demonstrated the feasibility of this sequential chemoenzymatic ligation strategy and points to its potential in protein synthesis. The results obtained herein are likely to stimulate the development of novel protein synthesis strategies based on the use of peptide thioacids in the future.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

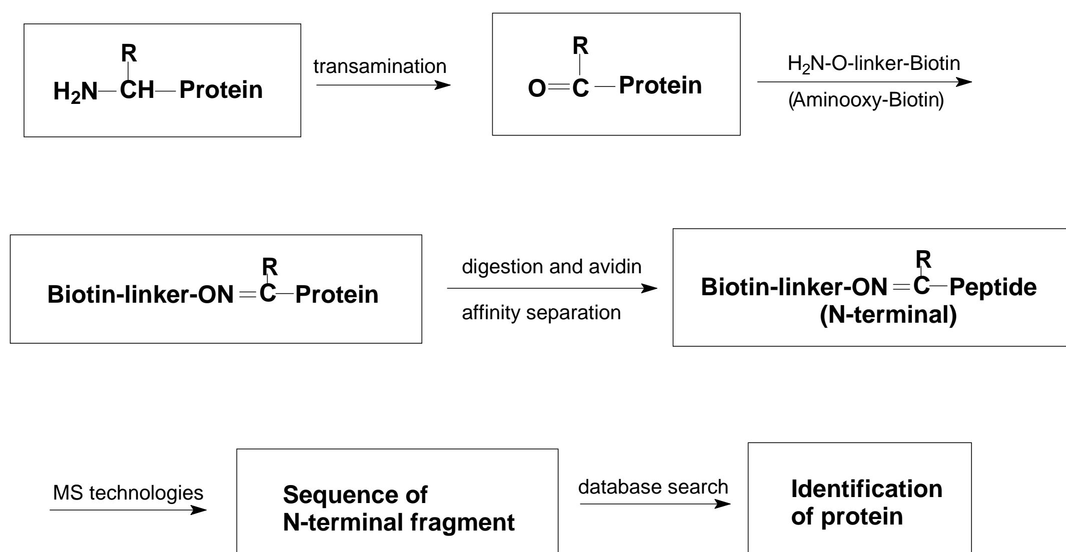
In this thesis, I described detailed studies on specific modifications of peptide N-terminus by transamination reaction and C-terminus by thiolysis reactions. Not only will these techniques have application in synthetic chemistry, but also they can be potentially useful in proteomics and in structural biology.

One possible application is using transamination reaction to design a proteomics project which is a one-protein-one-peptide approach. In this case, transamination reaction can convert the N-terminal residue of a protein specifically to a 2-oxoacyl moiety, which can be used to attach a reagent that has a nucleophilic group. Therefore, proteins are only labeled at their N-termini so that following trypsin digestion and affinity isolation, only one peptide fragment per protein, i.e., the N-terminal fragment, is picked up for LC/MS analysis. With the detailed information of transamination reaction acquired in chapter 2 about the reaction tendency of different N-terminal amino acid residues under the transamination reaction conditions, a database search can indentify many proteins. The procedure is illustrated in the flowchart (scheme 4.8, page 125).

Another possible application is to utilize these technologies in the NMR structure analysis. When we combine the enzymatic ligation with the mini-thiol capture ligation method and also the traditional chemical ligation method, proteins, especially those of medium size, can be efficiently synthesized. As a result, it becomes much easier to modify a protein at any target residue. For example, MTSSL (1-oxyl-2,2,5,5-tetramethyl-η_3-pyrroline-3-methyl) was used as a sulfhydro- specific labeling to introduce spin-labels into proteins for NMR analysis[154]. By our method, MTSSL can be quite easily introduced into desired

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Cys sites of peptide fragments or even at other residues where MTSSL is coupled by chemical approaches at the stage of solid-phase synthesis.



Scheme 4.8 A one-protein-one-peptide approach to identify proteins. Using transamination reaction to specifically pick up the N-terminal segments of proteins is followed by LC/MS analysis to identify the proteins. In our case, one labeled peptide represents one protein. Therefore, the analysis of MS data is much easier and more simplified than that of conventional proteomic methods using LC/MS.

I wish the results in my thesis can attract interest of the laboratories involved in this area of research.

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

References

1. Gygi, S. P. et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17, 994-9 (1999).

2. Hutter, H. Fluorescent reporter methods. Methods Mol Biol 351, 155-73 (2006).

3. Wagner, T. E. & Hsu, C. J. Paramagnetic labeling studies. 1. A new method for labeling proteins with a stable paramagnetic nitroxide radical. Analytical Biochemistry 36, 1-5 (1970).

4. Biswas, R., Kühne, H., Brudvig, G. W. & Gopalan, V. Use of EPR spectroscopy to study macromolecular structure and function. Science progress 84, 45-67 (2001).

5. Brooks, C. L. & Gu, W. Ubiquitination, phosphorylation and acetylation: The molecular basis for p53 regulation. Current Opinion in Cell Biology 15, 164-171 (2003).

6. Zhang, K. et al. Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. Molecular & cellular proteomics : MCP 1, 500-508 (2002).

7. Vranka, J. A., Sakai, L. Y. & Bachinger, H. P. Prolyl 3-hydroxylase 1, enzyme characterization and identification of a novel family of enzymes. J Biol Chem 279, 23615-21 (2004).

8. Khorasanizadeh, S. The Nucleosome: From Genomic Organization to Genomic Regulation. Cell 116, 259-272 (2004).

9. György, B., Tóth, E., Tarcsa, E., Falus, A. & Buzás, E. I. Citrullination: A posttranslational modification in health and disease. International Journal of Biochemistry and Cell Biology 38, 1662-1677 (2006).

10. Bijlmakers, M. J. & Marsh, M. The on-off story of protein palmitoylation. Trends Cell Biol 13, 32-42 (2003).

11. Glomset, J. A., Gelb, M. H. & Farnsworth, C. C. Prenyl proteins in eukaryotic cells: a new type of membrane anchor. Trends Biochem Sci 15, 139-42 (1990).

12. Bijlmakers, M. J. & Marsh, M. The on-off story of protein palmitoylation. Trends in Cell Biology 13, 32-42 (2003).

13. J. A. Hoch, T. J. S. Two Component Signal Transduction (ASM, Washington, 1995).

14. Wallin, R., Stanton, C. & Hutson, S. M. Intracellular maturation of the gamma-carboxyglutamic acid (Gla) region in prothrombin coincides with release of the propeptide. Biochem J 291 (Pt 3), 723-7 (1993).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

15. Furie, B., Bouchard, B. A. & Furie, B. C. Vitamin K-dependent biosynthesis of γ/-carboxyglutamic acid. Blood 93, 1798-1808 (1999).

16. Boucher, D., Larcher, J. C., Gros, F. & Denoulet, P. Polyglutamylation of tubulin as a progressive regulator of in vitro interactions between the microtubule-associated protein Tau and tubulin. Biochemistry 33, 12471-12477 (1994).

17. Chapman, E., Best, M. D., Hanson, S. R. & Wong, C. H. Sulfotransferases: Structure, mechanism, biological activity, inhibition, and synthetic utility. Angewandte Chemie - International Edition 43, 3526-3548 (2004).

18. Stephenson, R. C. & Clarke, S. Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. Journal of Biological Chemistry 264, 6164-6170 (1989).

19. Schmidt, G. et al. Gin 63 of Rho is deamidated by Escherichia coli cytotoxic necrotizing factor-1. Nature 387, 725-729 (1997).

20. Trombetta, E. S. & Parodi, A. J. N-glycan processing and glycoprotein folding. Advances in Protein Chemistry 59, 303-344 (2001).

21. McNeill, L. A. et al. Hypoxia-inducible factor asparaginyl hydroxylase (FIH-1) catalyses hydroxylation at the beta-carbon of asparagine-803. Biochem J 367, 571-5 (2002).

22. Guyot, N., Zani, M. L., Maurel, M. C., Dallet-Choisy, S. & Moreau, T. Elafin and its precursor trappin-2 still inhibit neutrophil serine proteinases when they are covalently bound to extracellular matrix proteins by tissue transglutaminase. Biochemistry 44, 15610-15618 (2005).

23. Moskovitz, J. Methionine sulfoxide reductases: ubiquitous enzymes involved in antioxidant defense, protein regulation, and prevention of aging-associated diseases. Biochim Biophys Acta 1703, 213-9 (2005).

24. Foley, T. L. & Burkart, M. D. Site-specific protein modification: advances and applications. Current Opinion in Chemical Biology 11, 12-19 (2007).

25. Wallace, C. J. A. Peptide ligation and semisynthesis. Current Opinion in Biotechnology 6, 403-410 (1995).

26. Antos, J. M. & Francis, M. B. Transition metal catalyzed methods for site-selective protein modification. Current Opinion in Chemical Biology 10, 253-262 (2006).

27. Quiocho, F. A. & Richards, F. M. Intermolecular Cross Linking of a Protein in the Crystalline State: Carboxypeptidase-A. Proc Natl Acad Sci U S A 52, 833-9 (1964).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

28. Kazan, D., Ertan, H. & Erarslan, A. Stabilization of Escherichia coli penicillin G acylase against thermal inactivation by cross-linking with dextran dialdehyde polymers. Appl Microbiol Biotechnol 48, 191-7 (1997).

29. Venkatesh, R. & Sundaram, P. V. Modulation of stability properties of bovine trypsin after in vitro structural changes with a variety of chemical modifiers. Protein Eng 11, 691-8 (1998).

30. Nicholas, K. M. et al. A cofactor approach to copper-dependent catalytic antibodies. Proc Natl Acad Sci U S A 99, 2648-53 (2002).

31. DeSantis, G. & Jones, J. B. Chemical modification of enzymes for enhanced functionality. Current Opinion in Biotechnology 10, 324-330 (1999).

32. Haring, D., Lees, M. R., Banaszak, L. J. & Distefano, M. D. Exploring routes to stabilize a cationic pyridoxamine in an artificial transaminase: site-directed mutagenesis versus synthetic cofactors. Protein Eng 15, 603-10 (2002).

33. Clark, P. I. & Lowe, G. Conversion of the active-site cysteine residue of papain into a dehydro-serine, a serine and a glycine residue. Eur J Biochem 84, 293-9 (1978).

34. Cosma, M. P. et al. The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. Cell 113, 445-456 (2003).

35. Dierks, T. et al. Multiple sulfatase deficiency is caused by mutations in the gene encoding the human Cα-formylglycine generating enzyme. Cell 113, 435-444 (2003).

36. Carrico, I. S., Carlson, B. L. & Bertozzi, C. R. Introducing genetically encoded aldehydes into proteins. Nature Chemical Biology 3, 321-322 (2007).

37. Liu, J. Q., Jiang, M. S., Luo, G. M., Yan, G. L. & Shen, J. C. Conversion of trypsin into a selenium-containing enzyme by using chemical mutation. Biotechnology Letters 20, 693-696 (1998).

38. Häring, D., Schüler, E., Adam, W., Saha-Möller, C. R. & Schreier, P. Semisynthetic enzymes in asymmetric synthesis: Enantioselective reduction of racemic hydroperoxides catalyzed by seleno-subtilisin. Journal of Organic Chemistry 64, 832-835 (1999).

39. Häring, D. & Schreier, P. Novel biocatalysts by chemical modification of known enzymes: Cross-linked microcrystals of the semisynthetic peroxidase seleno-subtilisin. Angewandte Chemie - International Edition 37, 2471-2473 (1998).

40. Antos, J. M. & Francis, M. B. Selective tryptophan modification with rhodium carbenoids in aqueous solution. Journal of the American Chemical Society 126, 10256-10257 (2004).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

41. Tilley, S. D. & Francis, M. B. Tyrosine-selective protein alkylation using π-allylpalladium complexes. Journal of the American Chemical Society 128, 1080-1081 (2006).

42. Joshi, N. S., Whitaker, L. R. & Francis, M. B. A three-component Mannich-type reaction for selective tyrosine bioconjugation. Journal of the American Chemical Society 126, 15942-15943 (2004).

43. Schlick, T. L., Ding, Z., Kovacs, E. W. & Francis, M. B. Dual-surface modification of the tobacco mosaic virus. Journal of the American Chemical Society 127, 3718-3723 (2005).

44. Kodadek, T., Duroux-Richard, I. & Bonnafous, J. C. Techniques: Oxidative cross-linking as an emergent tool for the analysis of receptor-mediated signalling events. Trends in Pharmacological Sciences 26, 210-217 (2005).

45. Takahashi, K. The reaction of phenylglyoxal with arginine residues in proteins. Journal of Biological Chemistry 243, 6171-6179 (1968).

46. Krysteva, M. A., Mazurier, J., Spik, G. & Montreuil, J. Comparative study on histidine modification by diethylpyrocarbonate in human serotransferrin and lactotransferrin. FEBS Letters 56, 337-340 (1975).

47. Gentry, C. L. et al. The effect of halogenation on blood-brain barrier permeability of a novel peptide drug. Peptides 20, 1229-1238 (1999).

48. Abbruscato, T. J. et al. Blood-to-central nervous system entry and stability of biphalin, a unique double-enkephalin analog, and its halogenated derivatives. Journal of Pharmacology and Experimental Therapeutics 276, 1049-1057 (1996).

49. Yamamoto, Y. et al. Site-specific PEGylation of a lysine-deficient TNF-alpha with full bioactivity. Nat Biotechnol 21, 546-52 (2003).

50. Ratner, V., Kahana, E., Eichler, M. & Haas, E. A general strategy for site-specific double labeling of globular proteins for kinetic FRET studies. Bioconjug Chem 13, 1163-70 (2002).

51. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. A general method for site-specific incorporation of unnatural amino acids into proteins. Science 244, 182-8 (1989).

52. Xie, J. & Schultz, P. G. A chemical toolkit for proteins--an expanded genetic code. Nat Rev Mol Cell Biol 7, 775-82 (2006).

53. Selvakumar, P. et al. Potential role of N-myristoyltransferase in cancer. Progress in Lipid Research 46, 1-36 (2007).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

54. Cummins, P. M. & O'Connor, B. Pyroglutamyl peptidase: An overview of the three known enzymatic forms. Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology 1429, 1-17 (1998).

55. Schalkwijk, C. G. Therapeutic interventions in the glyc(oxid)ation pathway. Immunology, Endocrine and Metabolic Agents in Medicinal Chemistry 7, 57-68 (2007).

56. Jaisson, S. et al. Carbamylation differentially alters type I collagen sensitivity to various collagenases. Matrix Biology 26, 190-196 (2007).

57. Baker, D. P. et al. N-terminally PEGylated human interferon-β-1a with improved pharmacokinetic properties and in vivo efficacy in a melanoma angiogenesis model. Bioconjugate Chemistry 17, 179-188 (2006).

58. Geoghegan, K. F. & Stroh, J. G. Site-directed conjugation of nonpeptide groups to peptides and proteins via periodate oxidation of a 2-amino alcohol. Application to modification at N-terminal serine. Bioconjugate Chemistry 3, 138-146 (1992).

59. Tam, J. P., Yu, Q. & Miao, Z. Orthogonal ligation strategies for peptide and protein. Biopolymers - Peptide Science Section 51, 311-332 (1999).

60. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. H. Synthesis of proteins by native chemical ligation. Science 266, 776-779 (1994).

61. Tam, J. P., Xu, J. & Khee, D. E. Methods and strategies of peptide ligation. Biopolymers - Peptide Science Section 60, 194-205 (2001).

62. Li, X., Zhang, L., Hall, S. E. & Tam, J. P. A new ligation method for N-terminal tryptophan-containing peptides using the Pictet-Spengler reaction. Tetrahedron Letters 41, 4069-4073 (2000).

63. In, Y., Fujii, M., Sasada, Y. & Ishida, T. Structural studies on C-amidated amino acids and peptides: Structures of hydrochloride salts of C-amidated Ile, Val, Thr, Ser, Met, Trp, Gln and Arg, and comparison with their C-unamidated counterparts. Acta Crystallographica Section B: Structural Science 57, 72-81 (2001).

64. Zhu, B. T. Catechol-O-methyltransferase (COMT)-mediated methylation metabolism of endogenous bioactive catechols and modulation by endobiotics and kenobiotics: Importance in pathophysiology and pathogenesis. Current Drug Metabolism 3, 321-349 (2002).

65. Cheng, J., Bawa, T., Lee, P., Gong, L. & Yeh, E. T. H. Role of desumoylation in the development of prostate cancer. Neoplasia 8, 667-676 (2006).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

66. Boggio, R. & Chiocca, S. Viruses and sumoylation: recent highlights. Current Opinion in Microbiology 9, 430-436 (2006).

67. Marks, N., Stern, F. & Kastin, A. J. Biodegradation of alpha-MSH and derived peptides by rat brain extracts, and by rat and human serum. Brain Res Bull 1, 591-3 (1976).

68. Muir, T. W. Semisynthesis of proteins by expressed protein ligation. Annual Review of Biochemistry 72, 249-289 (2003).

69. Tam, J. P. & Yu, Q. Methionine ligation strategy in the biomimetic synthesis of parathyroid hormones. Biopolymers 46, 319-327 (1998).

70. Yan, L. Z. & Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. Journal of the American Chemical Society 123, 526-533 (2001).

71. Shin, Y. et al. Fmoc-based synthesis of peptide-(α)thioesters: Application to the total chemical synthesis of a glycoprotein by native chemical ligation. Journal of the American Chemical Society 121, 11684-11689 (1999).

72. Dixon, H. B. & Weitkamp, L. R. Conversion of the N-terminal serine residue of corticotrophin into glycine. Biochem J 84, 462-8 (1962).

73. Papanikos, A., Rademann, J. & Meldal, M. α-ketocarbonyl peptides: A general approach to reactive resin-bound intermediates in the synthesis of peptide isosteres for protease inhibitor screening on solid support. Journal of the American Chemical Society 123, 2176-2181 (2001).

74. Dixon, H. B. & Moret, V. Removal of the N-Terminal Residue of a Protein after Transamination. Biochem J 94, 463-9 (1965).

75. Metzler, D. E., Olivard, J. & Snell, E. E. Transamination of Pyridoxamine and Amino Acids with Glyoxylic Acid. J. Am. Chem. Soc. 76, 644-648 (1954).

76. Bunzli, H. F. & Bosshard, H. R. Modification of the single arginine residue in insulin with phenylglyoxal. Hoppe-Seyler's Zeitschrift fur Physiologische Chemie 352, 1180-1182 (1971).

77. Fields, R. & Dixon, H. B. Micro method for determination of reactive carbonyl groups in proteins and peptides, using 2,4-dinitrophenylhydrazine. Biochemical Journal 121, 587-589 (1971).

78. van Heyningen, S., Tipton, K. F. & Dixon, H. B. Transamination of the N-terminal residue of carboxypeptidase. Biochemical Journal 108, 508-509 (1968).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

79. Levitzki, A., Anbar, M. & Berger, A. Specific oxidation of peptides via their copper complexes. Biochemistry 6, 3757-65 (1967).

80. Fleming, L. & Crosbie, G. W. Non-enzymic transamination between glycine and glyoxylate. Biochim Biophys Acta 43, 139-40 (1960).

81. Gilmore, J. M., Scheck, R. A., Esser-Kahn, A. P., Joshi, N. S. & Francis, M. B. N-terminal protein modification through a biomimetic transamination reaction. Angewandte Chemie - International Edition 45, 5307-5311 (2006).

82. Frank, R. The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. J Immunol Methods 267, 13-26 (2002).

83. Tam, J. S. a. J. P. Unprotected Peptides as Building Blocks for the Synthesis of Peptide Dendrimers with Oxime, Hydrazone, and Thiazolidine Linkages. J. Am. Chem. Soc. 117, 3893 - 3899 (1995).

84. Papanikos A, R. J., Meldal M. alpha-Ketocarbonyl peptides: a general approach to reactive resin-bound intermediates in the synthesis of peptide isosteres for protease inhibitor screening on solid support. J Am Chem Soc. 123, 2176-81 (2001).

85. Sugiura, Y. Proton-magnetic-resonance study of copper(I) complexes with peptides containing sulfhydryl and imidazole groups as possible model ligands for copper proteins. Eur J Biochem 78 (1977).

86. Sarkar B, D. H., Webster D. Removal by transamination and scission of residues from the peptide representing the copper-transport site of serum albumin. Biochem J. 173, 895-97 (1978).

87. Tam JP, R. C., Liu CF, Shao J. Specificity and formation of unusual amino acids of an amide ligation strategy for unprotected peptides. Int J Pept Protein Res 45, 209-16 (1995).

88. Field, G. B. a. N., R. L. Solid phase peptide synthesis utilizing 9-fluorenylmethoxy -carbonyl amino acids. Int. J. pept. Protein Res 35, 161-214 (1990).

89. Tam, J. P., Lu, Y. A., Liu, C. F. & Shao, J. Peptide synthesis using unprotected peptides through orthogonal coupling methods. Proceedings of the National Academy of Sciences of the United States of America 92, 12485-12489 (1995).

90. Muir, T. W., Dawson, P. E. & Kent, S. B. H. Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. Methods in Enzymology 289, 266-298 (1997).

91. Aimoto, S. Polypeptide synthesis by the thioester method. Biopolymers - Peptide Science Section 51, 247-265 (1999).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

92. Liu, C. F., Rao, C. & Tam, J. P. Orthogonal ligation of unprotected peptide segments through pseudoproline formation for the synthesis of HIV-1 protease analogs. Journal of the American Chemical Society 118, 307-312 (1996).

93. Hackeng, T. M., Griffin, J. H. & Dawson, P. E. Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. PNAS 96, 10068-10073 (1999).

94. Ohta, Y. et al. Cysteine-derived S-protected oxazolidinones: Potential chemical devices for the preparation of peptide thioesters. Organic Letters 8, 467-470 (2006).

95. Camarero, J. A. & Mitchell, A. R. Synthesis of proteins by native chemical ligation using Fmoc-based chemistry. Protein and Peptide Letters 12, 723-728 (2005).

96. Botti, P., Villain, M., Manganiello, S. & Gaertner, H. Native chemical ligation through in situ O to S acyl shift. Organic Letters 6, 4861-4864 (2004).

97. Suich, D. J., Ballinger, M. D., Wells, J. A. & DeGrado, W. F. Fmoc-based synthesis of glycolate ester peptides for the assembly of de novo designed multimeric proteins using subtiligase. Tetrahedron Letters 37, 6653-6656 (1996).

98. Blow, D. M. The tortuous story of Asp ... His ... Ser: structural analysis of alpha -chymotrypsin. Trends Biochem Sci 22, 405-8 (1997).

99. Henderson, R. Structure of crystalline alpha-chymotrypsin. IV. The structure of indoleacryloyl-alpha-chyotrypsin and its relevance to the hydrolytic mechanism of the enzyme. Journal of Molecular Biology 54, 341-354 (1970).

100. Matthews, D. A., Alden, R. A. & Birktoft, J. J. X ray crystallographic study of boronic acid adducts with subtilisin BPN (Novo). A model for the catalytic transition state. Journal of Biological Chemistry 250, 7120-7126 (1975).

101. Bone, R., Sampson, N. S., Bartlett, P. A. & Agard, D. A. Crystal structures of α-lytic protease complexes with irreversibly bound phosphonate esters. Biochemistry 30, 2263-2272 (1991).

102. Schechter, I. & Berger, A. On the size of the active site in proteases. I. Papain. Biochemical and Biophysical Research Communications 27, 157-162 (1967).

103. Wells, J. A. Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. Biochemistry 30, 4151-4159 (1991).

104. Johnson, C. H. & Knowles, J. R. The binding of inhibitors to alpha-chymotrypsin. Biochemical Journal 101, 56-62 (1966).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

105. Liu, C. F. & Tam, J. P. Subtilisin-catalyzed synthesis of amino acid and peptide esters. Application in a two-step enzymatic ligation strategy. Organic Letters 3, 4157-4159 (2001).

106. Jackson, D. Y. et al. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. Science 266, 243-247 (1994).

107. Chattaraj, P. K., Lee, H. & Parr, R. G. HSAB principle. Journal of the American Chemical Society 113, 1855-1856 (1991).

108. Barany, G. & Merrifield, R. B. Kinetics and mechanism of the thiolytic removal of the dithiasuccinoyl (Dts) amino protecting group. Journal of the American Chemical Society 102, 3084-3095 (1980).

109. Schellenberger, V. & Jakubke, H. D. Protease-catalyzed kinetically controlled peptide synthesis. Angewandte Chemie (International Edition in English) 30, 1437-1449 (1991).

110. Zaks, A. & Klibanov, A. M. Enzymatic catalysis in nonaqueous solvents. Journal of Biological Chemistry 263, 3194-3201 (1988).

111. Chong, S. et al. Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene 192, 271-281 (1997).

112. Muir, T. W., Sondhi, D. & Cole, P. A. Expressed protein ligation: A general method for protein engineering. Proceedings of the National Academy of Sciences of the United States of America 95, 6705-6710 (1998).

113. Weber, N. et al. Copolymeric polythioesters by lipase-catalyzed thioesterification and transthioesterification of α,ω-alkanedithiols. Applied Microbiology and Biotechnology 70, 290-297 (2006).

114. Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A. & Chen, E. Y. Cloning, sequencing, and secretion of Bacillus amyloliquefaciens subtilisin in Bacillus subtilis. Nucleic Acids Research 11, 7911-7925 (1983).

115. Murayama, R., Akanuma, G., Makino, Y., Nanamiya, H. & Kawamura, F. Spontaneous transformation and its use for genetic mapping in Bacillus subtilis. Bioscience, Biotechnology and Biochemistry 68, 1672-1680 (2004).

116. Wilcks, A. et al. Replication mechanism and sequence analysis of the replicon of paw63, a conjugative plasmid from bacillus thuringiensis. Journal of Bacteriology 181, 3193-3200 (1999).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

117. Carter, P. & Wells, J. A. Engineering enzyme specificity by 'substrate-assisted catalysis'. Science 237, 394-399 (1987).

118. Radisky, E. S., King, D. S., Kwan, G. & Koshland Jr, D. E. The role of the protein core in the inhibitory power of the classic serine protease inhibitor, chymotrypsin inhibitor 2. Biochemistry 42, 6484-6492 (2003).

119. Matsubara, H., Kasper, C. B., Brown, D. M. & Smith, E. L. Subtilisin Bpn'. I. Physical Properties and Amino Acid Composition. J Biol Chem 240, 1125-30 (1965).

120. Fields, G. B. & Noble, R. L. Solid phase peptide synthesis utilizing 9-fluorenyl -methoxycarbonyl amino acids. International Journal of Peptide and Protein Research 35, 161-214 (1990).

121. Li, X., Kawakami, T. & Aimoto, S. Direct preparation of peptide thioesters using an Fmoc solid-phase method. Tetrahedron Letters 39, 8669-8672 (1998).

122. Nilsson, B. L., Soellner, M. B. & Raines, R. T. Chemical synthesis of proteins. Annual Review of Biophysics and Biomolecular Structure 34, 91-118 (2005).

123. Schnolzer, M., Alewood, P., Jones, A., Alewood, D. & Kent, S. B. H. In situ neutralization in Boc-chemistry solid phase peptide synthesis. Rapid, high yield assembly of difficult sequences. International Journal of Peptide and Protein Research 40, 180-193 (1992).

124. Liu, C. F., Rao, C. & Tam, J. P. Acyl disulfide-mediated intramolecular acylation for orthogonal coupling between unprotected peptide segments. Mechanism and application. Tetrahedron Letters 37, 933-936 (1996).

125. Liu, C. F. & Tam, J. P. Chemical ligation approach to form a peptide bond between unprotected peptide segments. Concept and model study. Journal of the American Chemical Society 116 (1994).

126. Blake, J. & Li, C. H. New segment-coupling method for peptide synthesis in aqueous solution: application to synthesis of human [Gly17]-beta-endorphin. Proc Natl Acad Sci U S A 78, 4055-8 (1981).

127. Dawson, P. E. & Kent, S. B. H. Convenient total synthesis of a 4-helix TASP molecule by chemoselective ligation. Journal of the American Chemical Society 115, 7263-7266 (1993).

128. Baca, M. & Kent, S. B. H. Catalytic contribution of flap-substrate hydrogen bonds in 'HIV-1 protease' explored by chemical synthesis. Proceedings of the National Academy of Sciences of the United States of America 90, 11638-11642 (1993).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

129. Kemp, D. S., Galakatos, N. G., Bowen, B. & Tan, K. Peptide synthesis by prior thiol capture. 2. Design of templates for intramolecular O,N-acyl transfer. 4,6-Disubstituted dibenzofurans as optimal spacing elements. Journal of Organic Chemistry 51, 1829-1838 (1986).

130. Shangguan, N., Katukojvala, S., Greenberg, R. & Williams, L. J. The reaction of thio acids with azides: A new mechanism and new synthetic applications. Journal of the American Chemical Society 125, 7754-7755 (2003).

131. Soellner, M. B., Nilsson, B. L. & Raines, R. T. Reaction mechanism and kinetics of the traceless Staudinger ligation. Journal of the American Chemical Society 128, 8820-8828 (2006).

132. Yamashiro, D. & Li, C. H. New segment synthesis of alpha-inhibin-92 by the acyl disulfide method. International Journal of Peptide and Protein Research 31, 322-334 (1988).

133. Blake, J. & Hao Li, C. New segment-coupling method for peptide synthesis in aqueous solution: Application to synthesis of human [Gly17]-β-endorphin. Proceedings of the National Academy of Sciences of the United States of America 78, 4055-4058 (1981).

134. Canne, L. E., Walker, S. M. & Kent, S. B. H. A general method for the synthesis of thioester resin linkers for use in the solid phase synthesis of peptide-α-thioacids. Tetrahedron Letters 36, 1217-1220 (1995).

135. Harris, J. L., Peterson, E. P., Hudig, D., Thornberry, N. A. & Craik, C. S. Definition and redesign of the extended substrate specificity of granzyme B. Journal of Biological Chemistry 273, 27364-27373 (1998).

136. Cahoon, E. B., Lindqvist, Y., Schneider, G. & Shanklin, J. Redesign of soluble fatty acid desaturases from plants for altered substrate specificity and double bond position. Proceedings of the National Academy of Sciences of the United States of America 94, 4872-4877 (1997).

137. Hornung, E., Walther, M., Kühn, H. & Feussner, I. Conversion of cucumber linoleate 13-lipoxygenase to a 9-lipoxygenating species by site-directed mutagenesis. Proceedings of the National Academy of Sciences of the United States of America 96, 4192-4197 (1999).

138. Bonagura, C. A., Bhaskar, B., Sundaramoorthy, M. & Poulos, T. L. Conversion of an engineered potassium-binding site into a calcium- selective site in cytochrome c peroxidase. Journal of Biological Chemistry 274, 37827-37833 (1999).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

139. Kuroki, R., Weaver, L. H. & Matthews, B. W. Structural basis of the conversion of T4 lysozyme into a transglycosidase by reengineering the active site. Proceedings of the National Academy of Sciences of the United States of America 96, 8949-8954 (1999).

140. Vita, E. et al. Rational engineering of a miniprotein that reproduces the core of the CD4 site interacting with HIV-1 envelope glycoprotein. Proceedings of the National Academy of Sciences of the United States of America 96, 13091-13096 (1999).

141. Liu, E., Prasad, L., Delbaere, L. T. J., Waygood, E. B. & Lee, J. S. Conversion of an antibody into an enzyme which cleaves the protein HPr. Molecular Immunology 35, 1069-1077 (1998).

142. Abrahmsen, L. et al. Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. Biochemistry 30, 4151-4159 (1991).

143. Tan, X. H., Wirjo, A. & Liu, C. F. An Enzymatic Approach to the Synthesis of Peptide Thioesters: Mechanism and Scope. Chembiochem 8, 1512-1515 (2007).

144. Licht, S., Forouzan, F. & Longo, K. Differential densometric analysis of equilibria in highly concentrated media: determination of the aqueous second acid dissociation constant of H2S. Analytical Chemistry 62, 1356-1360 (1990).

145. Ellman, G. L. Tissue sulfhydryl groups. Arch Biochem Biophys 82, 70-7 (1959).

146. Wieland, T., Bokelmann, E., Bauer, L., Lang, H. U. & Lau, H. Justus Liebigs Ann. Chem. 586, 129-149 (1953).

147. Pellon, G. Deacylation of penicilloylated penicillin-binding proteins from Micrococcus Luteus: kinetic study of thiol-promoted release of the bound penicilloyl moiety. Biochimica et biophysica acta 883 389-395. (1986).

148. Ellison, P. A. & Engel, P. C. Does butyl-CoA dehydrogenase ctalyze thiol-thioester exchange? Biochemical Society Transactions 15, 236-237 (1987).

149. Dawson, P. E., Churchill, M., Ghadiri, M. R. & Kent, S. B. H. Modulation of Reactivity in Native Chemical Ligation through the Use of Thiol Additives. Journal of the American Chemical Society 119 4325-4329 (1997).

150. Johnson, E. C. B. & Kent, S. B. H. Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction. . J. Am. Chem. Soc. 128, 6640-6646 (2006).

151. Chang, T. K., Jackson, D. Y., Burnier, J. P. & Wells, J. A. Subtiligase: A tool for semisynthesis of proteins. Proceedings of the National Academy of Sciences of the United States of America 91, 12544-12548 (1994).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

152. Penning, T. M. & Jez, J. M. Enzyme redesign. Chemical Reviews 101, 3027-3046 (2001).

153. Albeck, A., Fluss, S. & Persky, R. Peptidyl epoxides: Novel selective inactivators of cysteine proteases. Journal of the American Chemical Society 118, 3591-3596 (1996).

154. Tamm, L. K. & Liang, B. NMR of membrane proteins in solution. Progress in Nuclear Magnetic Resonance Spectroscopy 48, 201-210 (2006).

ATTENTION: The Singapore Copyright Act applies to the use of this document. Nanyang Technological University Library

Publications

1. **An Enzymatic Approach to the Synthesis of Peptide Thioesters: Mechanism and Scope.** Xiao-hong Tan, Andre Wirjo, and Chuan-Fa Liu. *ChemBioChem,* 2007, 8, 1512-1515

2. **Subtiligase as a hydrothiolase for the synthesis of peptidyl thioacids.** Xiao-Hong Tan, Ren-Liang Yang, Andre Wirjo, Chuan-Fa Liu. (accepted by *Tetrahedron Letters*)

3. **A simple method to prepare peptide C-terminal thioacids and application in sequential chemoenzymatic ligation.** Xiao-Hong Tan, Xiao-Hong Zhang, Ren-Liang Yang and Chuan-Fa Liu. (accepted by *ChemBioChem*)

4. **A systematic investigation of the transamination reaction by membrane-immobilized dipeptide arrays.** Xiao-Hong Tan, Yun Zeng and Chuan-Fa Liu. (submitted)