Chemical Synthesis of Proteins Containing Post-translational Modifications and PNA-templated Disulfide Bond Formation

Li Fupeng

Supervisor: Associate Professor Liu Chuan-Fa

School of Biological Sciences

A thesis submitted to the Nanyang Technological University in fulfillment

of the requirement for the degree of Doctor of Philosophy

To my family

Acknowledgements

During the period of my Ph.D study, there are numerous people who gave me help, support and encouragement.

First of all, I would like to show my greatest appreciation to my supervisor A/Prof. Liu Chuan-Fa, a respectable, responsible and resourceful scholar who has provided me with insightful ideas and instructive guidance in every stage of my postgraduate study. Obviously I could not have accomplished this thesis without his enlightening instruction, kindness and patience. His strict requirement and rigorous attitude towards science will continue to impact my research career in the future.

I shall extend my thanks to all the previous and present members of our lab, Dr. Tan Xiaohong, Dr. Zeng Yun, Dr. Yang Yi, Dr. Zhang Xiaohong, Dr. Lu Xiaowei, Mr. Andre Wirjo, Ms. Hou Wen, Ms. Ding Yingjie, Mr. Yang Renliang, Mr. Cao Yuan for their kindness, helpful discussion and encouragement.

My sincere appreciation also goes to Mr. Abdollah Allahverdi from Prof. Lars's lab for his collaboration in performing the nucleosomal array folding assay. Thanks also to my friends in SBS, especially Mr. Gao Xiang who has helped me a lot in the past four years.

At last, I'd like to express my gratitude to my parents for their selfless sacrifice, and also to my beloved wife – Chen Maohua and lovely son – Li Kechen: you are always at the centre of my life.

Table of Content

Abstract

The study of protein structure and function is a central topic of bioscience research. As the most important molecules of life, proteins often undergo post-translational modifications (PTM) in order to carry out their biological functions. A good example is the histone proteins whose PTMs include acetylation, methylation and ubiquitination on specific amino acid side chains. The understanding of how a protein exerts its function in either its native or post-translationally modified form can not only solve many problems in basic biology but also produce therapeutic breakthroughs in medicine. Protein chemical synthesis is a useful approach for the production of these modified proteins which are invaluable reagents for the structural and functional characterization of individual PTM events.

In my study, two popular chemical ligation methods were studied for ligation at valine residue: the thioester-mediated ligation and thioacid capture ligation. An N-terminal penicillamine was used to mediate the ligation reaction, similar to the Cys-mediated chemical ligation. Subsequent desulfurization gives a Val residue at the ligation site. However, the steric hindrance of the tertiary thiol group in penicillamine significantly slows down the ligation reaction as compared to ligation at a cysteine residue. Several model peptide thioesters and thioacids were prepared to ligate with the penicillyl peptides by the respective ligation methods. From the results, we can conclude that the thioacid capture ligation works very well for ligation at the sterically hindered Pen

compared with the thioester-mediated ligation. A modified histone H2B K120Ac was successfully synthesized by the thioacid capture ligation method. Moreover, a new strategy was developed for the purification of the ligation product.

Although chemical ligation method can be used to synthesize the Lys-acetylated histones as mentioned above, significant technical barrier exists for the wide adoption of such methods by the large research community. An alternative method is needed to address this problem. A direct method for site-specific protein modification was developed that allowed for the introduction of an acetyl-lysine mimic into recombinant proteins. The lysine of interest was mutated to cysteine first, and the free thiol of cysteine was reacted with N-vinylacetamide (NVA) by radical-mediated thiol-ene coupling to generate acetyl-4-thialysine (sLys (Ac)) which is an acetyl-lysine analog. This reaction was specific and efficient and could be used to prepare large quantities of site-specifically acetylated proteins. Moreover, three similar reagents, vinyl acetate, N-allylacetamide and N-methyl-N-vinylacetamide, were also used to create three other kinds of acetyl-lysine analogs. At last, several assays were utilized to assess the function of the acetyl-lysine analogs, which showed that sLys (Ac) was functionally equivalent or similar to the native Lys (Ac).

In the third part of my research, I have attempted to explore the idea of using Watson-Crick base pairing to bring together reactive groups for a particular reaction. Specifcally, PNA was used to mediate correct disulfide bond formation between two

different peptides. The results showed that this reaction system was highly selective with the desired S-S heterodimer formed exclusively. Since PNA synthesis is fully compatible with peptide synthesis, this method has an advantage over and can be more useful than the DNA-templated reaction systems.

Abbreviations

Acm	Acetamidomethyl
AUC	Analytical Ultracentrifugation
Boc	*tert*-butoxycarbonyl
BzSH	Benzyl mercaptan
DIEA	*N, N,*-diisopropylethylamine
DCM	Dichloromethane
DMF	*N, N,*-dimethylformamide
DTS	DNA-templated organic synthesis
DTT	Dithiothreitol
EPL	Expressed protein ligation
eq	equivalents
ESI	Electrospray ionization
Fmoc	9-Fluorenylmethoxycarbonyl
HAT	Histone acetyltransferase
HDAC	Histone deacetylases
HPLC	high performance liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MALDI	Matrix assisted laser desorption ionization
MESNA	Sodium 2-sulfanylethanesulfonate
MS	mass spectrometry
MTCA	Methylthio-carbonyl-aziridine
Npys	5-Nitro-2-pyridinesulfenyl
NVA	N-vinylacetamide
Pen	Penicillamine
PTM	Post-translational modification
PTS	PNA-templated organic synthesis

PNA	Peptide nucleic acid
PyBOP	Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate
SPPS	Solid-phase peptide synthesis
TCEP	Triscarboxyethylenphosphin
TCL	Thioacid capture ligation
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethanesulfonic acid
Thz	Thiazolidine
TOF	time-of-flight
Trt	Trityl
UV	Ultraviolet rays
VA-044	2,2'-Azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride

Chapter 1: General Introduction

1.1 Post-translationally modified proteins

Proteins are the most important molecules of life, charged with the task of performing almost all the biological functions in a living organism. The study of protein structure and function is a central topic in life science research. As we know, proteins often need to undergo post-translational modifications (PTMs) in order to carry out their biological functions. A good example is the histone proteins whose PTMs include acetylation, methylation, phosphorylation and ubiquitination on specific amino acid side chains (Rice and Allis 2001; Peterson and Laniel 2004). These histone modifications have crucial functions in the cell (Fig. 1) (Allfrey, Faulkner et al. 1964; Strahl and Allis 2000; Jenuwein and Allis 2001; Rice and Allis 2001). For example, the side chain amines of lysines at the N-termini of the histones are positively charged under physiological conditions. It is believed that electrostatic interactions between the positivel charged histone N-terminal tails and the negatively charged phosphate backbone in DNA helps to pack DNA tightly in chromatin. However, after the Lys side chain is neutralized by acetylation, the binding force would be greatly reduced which may loosen up the DNA in chromatin (Kimura, Matsubara et al. 2005; Shahbazian and Grunstein 2007). Furthermore, the acetyl group on lysine can serve as a docking site for many nuclear proteins such as transcription factors. Therefore, in many instances post-translational acetylation on histone can up-regulate the activity of the underlying DNA, promoting gene transcription and other DNA-related functions.

Histone methylation is another very important post-translational modification (Klose and Zhang 2007; Lan, Bayliss et al. 2007; Swigut and Wysocka 2007; Shilatifard 2008). This modification includes adding one, two or three methyl groups on the side chain amine of lysine. Lysine methylation may cause tighter bindings between histones and DNA and therefore may down-regulate or even turn off gene transcription. So acetylation and methylation on histones are the crucial mechanisms for gene regulation.

There are also several other post-translational modifications, like phosphorylation, SUMOylation, citrullination, ubiquitination and ADP-ribosylation (Strahl and Allis 2000; Peterson and Laniel 2004; Kamakaka and Biggins 2005). These modifications play important roles too. All together these histone modifications discussed above form the so-called epigenetic marks that serve to regulate key biological processes such as gene expression, DNA repair and recombination, and cell division. Clearly, the understanding of how a protein exerts its function in either its native or post-translational modified form can not only solve fundamental problems in basic biology but also produce therapeutic breakthroughs in medicine.

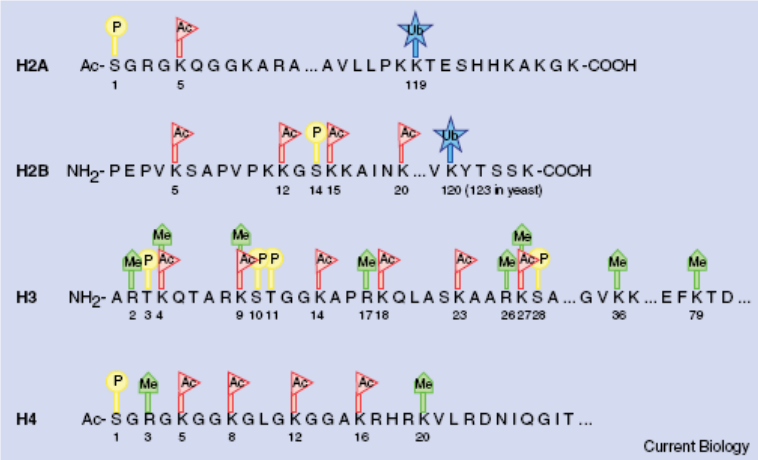


Figure 1. *Post-translational modifications of the core histones (adopted from* Peterson and Laniel 2004*).*

1.2 Chemical synthesis of proteins containing PTMs

Although many years of intensive research have found that histone post-translational modifications play essential roles in gene regulation, DNA repair and chromosome condensation, the exact effects of individual protein modification events remain to be elucidated. As a major limiting factor, the study of histone modification is often hindered by a lack of sufficient amounts of homogeneous protein samples containing the site-specific modification of interest. Such modified proteins would be invaluable reagents for discerning the structural and functional effects of a particular PTM via biophysical and biochemical means.

As known to all, the classical recombinant DNA technology can be used to produce proteins in large quantities. But there are several limitations for this technology. For instance, some proteins can be toxic to the cell used for protein production. It is also

often difficult to express large, multidomain proteins. Uncontrolled processing of nascent polypeptides in the host cell will also cause product heterogeneity. Moreover, since living cells are used for protein expression, such synthesis systems are inherently limited to natural amino acid building blocks (Morrow 1979; Dawson and Kent 2000). Therefore, these post-translationally modified proteins are not accessible by the conventional recombinant technology. Chemical synthesis on the other hand is a useful approach for the production of these modified proteins. In fact, the combined use of solid phase peptide synthesis and chemoselective peptide ligation methods has become a powerful enabling technology in protein structure-function studies.

1.2.1 Solid phase peptide synthesis

Since 1963, solid phase peptide synthesis (SPPS) has been widely used for the synthesis of peptides and small proteins containing a large variety of structural units (Merrifield 1963). In repetitive steps, the peptides can be synthesized through assembling the amino acid building blocks, some of which may be unnatural or modified amino acids such as D-amino acids, phosphorylated Tyr/Ser/Thr and acetylated/methylated Lys. As an early example of using chemical synthesis to study protein PTM, the Allfrey lab synthesized a H4 peptide (1-37) with acetylation on K12 and K16 by the classical SPPS, and the acetyl group could be removed by HDACs *in vitro* (Krieger, Levine et al. 1974). This experiment demonstrated that a synthetic peptide with Lys acetylation PTMs had the similar function as the native post-translationally modified protein. However, due to the possible interchain

aggregation of the long growing protected peptide chain, it is still a challenge to use SPPS for the synthesis of very large peptides or proteins. For example, the histone proteins which are around 120 amino acids in length would be difficult targets for SPPS. For this reason, the convergent chemical ligation approach to assemble a protein from intermediate peptide building blocks appears attractive from a strategical point of view.

1.2.2 Principle of chemical ligation

Since the discovery that peptide bond formation between an amino acid thioester and a cysteine derivative could take place via an intramolecular acyl transfer reaction in 1953 (Wieland, Bokelmann et al. 1953), this proximity driven, intramolecular reaction principle has been used to develop a number of chemical ligation methods. As shown in Fig. 2, the C-terminal **X** group of the peptide acyl donor chemoselectively reacts with the N-terminal **Y** group of the acyl receptor to form a new bond that brings the two components together in a capture reaction. After the capture step, the intramolecular acylation will occur to generate an amide bond. Several ligation methods have been developed by varying the pair of **X** and **Y** (Fig. 2).

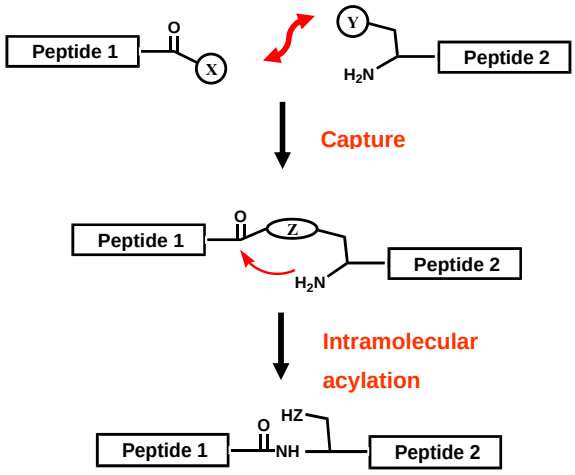


Figure 2. *Proximity-driven peptide bond formation.*

1.2.3 Thioester-mediated ligation

Among these chemical ligation methods (Tam, Yu et al. 1999), thioester-mediated chemical ligation is currently the most used one (Nilsson, Soellner et al. 2005). In this ligation reaction, the C-terminal thioester of an unprotected peptide reacts with the N-terminal Cys residue of another unprotected peptide. A thiol-thioester exchange reaction first takes place, which forms a new thioester between the acyl group of the first peptide and the side-chain thiol of the N-ter cysteine of the second peptide. This is followed by an intramolecular S→N acyl transfer for peptide bond formation. Derived from the reaction of Wieland (Wieland, Bokelmann et al. 1953), this peptide ligation method was named as native chemical ligation (NCL) (Fig. 3) (Dawson, Muir et al. 1994). Several factors affect the reaction rate of this ligation. First, tris(2-carboxylethyl)phosphine (TCEP) and a free thiol compound should be added in

the ligation buffer to maintain the reducing environment of the ligation reaction in order to prevent N,S-bisacylated byproducts and reduce disulfide formation, therefore improving the overall ligation yield (Tam, Lu et al. 1995). Second, the ligation should be conducted in a neutral or slightly basic pH environment, which proceeds faster than at weakly acidic pH (Hackeng, Griffin et al. 1999). Third, a proper thiol catalyst can accelerate the thiol-thioester reaction (Johnson and Kent 2006).

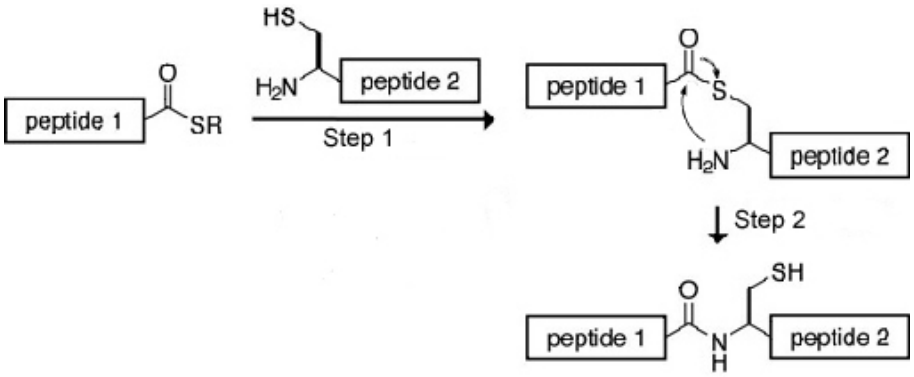


Figure 3. *Thioester-mediated chemical ligation. Step 1: Thiol-thioester exchange to form a new thioester intermediate. Step 2: Peptide bond formation by intramolecular S→N acyl transfer.*

1.2.4 Thioacid capture ligation

Although thioester-mediated ligation is the most popular ligation method for protein chemical synthesis, it also has some drawbacks. For example, the first step of the reaction scheme is not fast and is usually the rate-limiting step. In 1996, another ligation method through acyl disulfide-mediated intrarnolecular acylation, the thioacid

capture ligation, was reported. In this ligation method (Fig. 4), an unprotected C-terminal thioacid of peptide 1 reacts with an Npys modified cysteine residue of peptide 2, which forms a mixed acyl disulfide intermediate. After a rapid intramolecular S→N acyl transfer through a 6-member ring transition state, a peptide bond is formed. Finally, the thiolytic reduction of the resulting hydrodisulfide (S-SH) gives a native Cys residue at the ligation site (Liu, Rao et al. 1996). This ligation method is much faster than the above-mentioned thioester-mediated ligation.

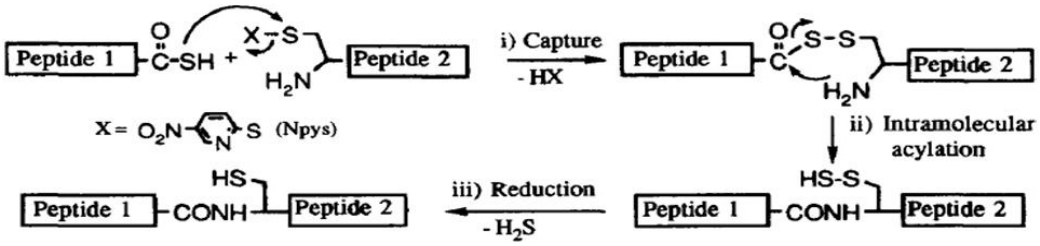


Figure 4. *Thioacid capture ligation. Step 1: Capturing through acyl disulfide formation. Step 2: Peptide bond formation by intramolecular acylation. Step 3: Thiolytic reduction to give the native Cys (adopted from* Liu, Rao et al. 1996*)*

1.2.5 Expressed protein ligation (EPL)

As mentioned above, a peptide thioester or thioacid can be ligated with another cysteinyl or Npys-cysteinyl peptide to form a longer peptide or protein. However, it's very difficult to synthesize the whole histone protein as there are more than 100 amino acids in histones. The expressed protein ligation (EPL) would be a useful approach to solve this problem. The EPL was first reported by Muir *et al* in 1998

(Muir, Sondhi et al. 1998). In EPL, a peptide fragment with a Cys at the N-terminus synthesized by SPPS is chemically ligated with an expressed protein thioester which is usually obtained using the intein-mediated protein splicing. EPL can also be the ligation between an N-terminal synthesized peptide thioester and a C-terminal expressed protein (Fig. 5). In the similar way, one can chemically synthesize a histone protein with PTMs by SPPS, and ligate it with an expressed protein or protein thioester. In principle, one can incorporate any modifications on histones site-specifically.

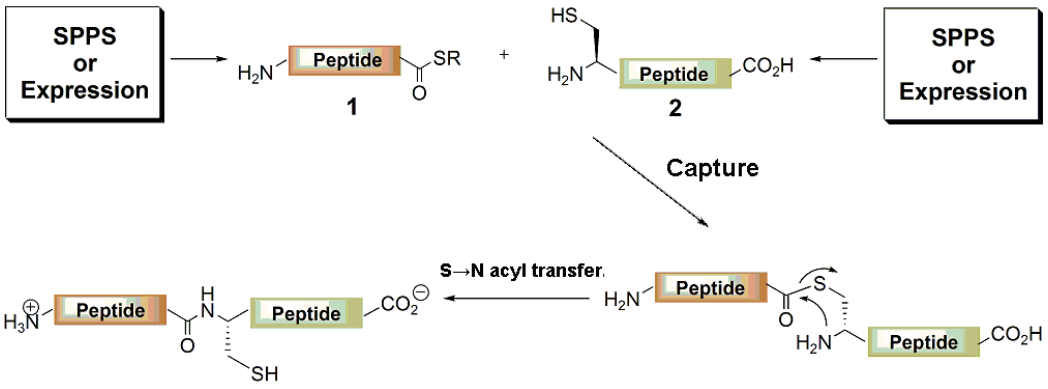


Figure 5. *Expressed protein ligation.*

In practice, this approach has several limitations. First of all, EPL is a cysteine-dependent ligation and there is no cysteine in H2A, H2B and H4, only one in H3. Second, if the modification is at the middle of the protein, it would not be practical to use this semi-synthetic approach as one would have to synthesize very long peptides. Third, it's very difficult to prepare large quantities of protein thioesters, because of the low expression level of the intein fusion protein and the low efficiency

of the cleavage step. Finally, the ligation product is usually very difficult to be separated from the truncated protein by HPLC.

1.3 Cysteine based site-specific installation of the modification group

An alternative method for introducing site-specific protein modification was to utilize the unique reactivity of the strongly nucleophilic side chain sulfhydryl of cysteine. For instance, a protein containing a cysteine can be selectively alkylated with 2-bromoethylamine to create 4-thialysine, a lysine analog (Fig. 6). The positively charged 4-thialysine side chain can be recognized by trypsin for the proteolytic action, which indicates the 4-thialysine is a functional equivalent to lysine (Kenyon and Thomas 1977).

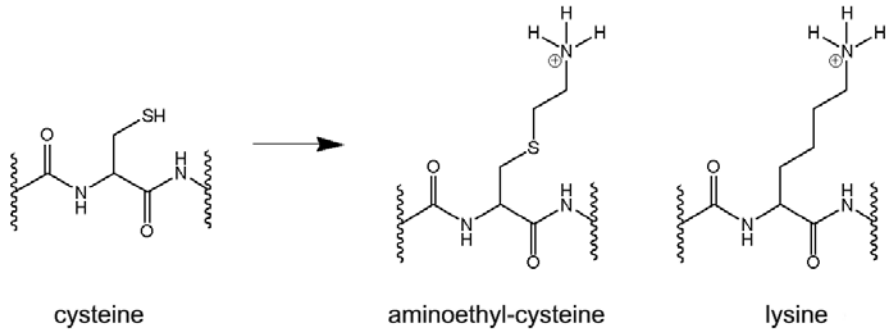


Figure 6. *4-Thialysine generated by the aminoethylation reaction.*

This method can also be extended to N-methyl-lysine analogs when using N-methylaminoethyl halides as the alkylating agents to synthesize the mono-, di- and

trimethylated aminoethylcysteine which are the methylated lysine analogs (Fig. 7). Through this method, it becomes very simple to prepare large quantities of site-specifically methylated histones by a one-step reaction. And the methylated lysine analog has the similar function in binding and enzymatic assays (Simon, Chu et al. 2007).

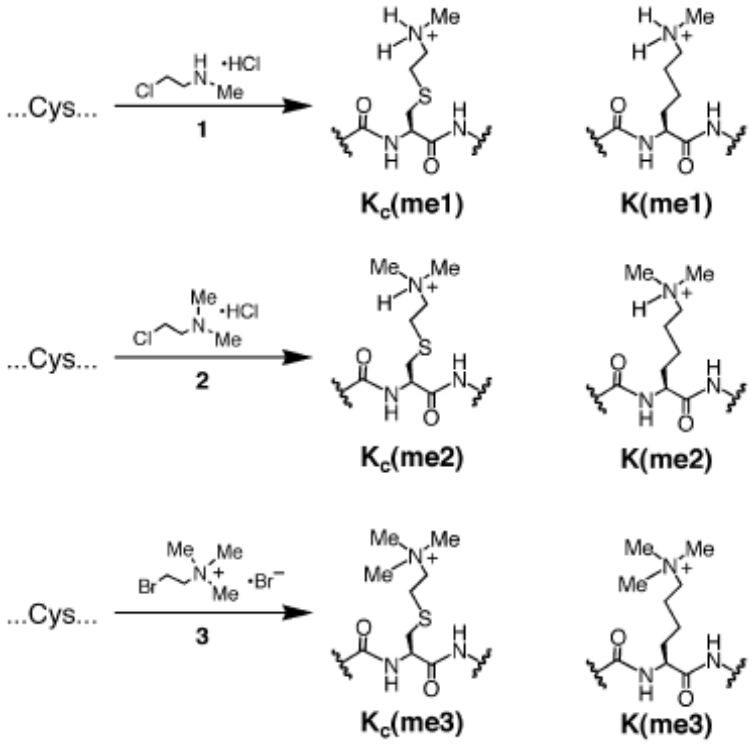


Figure 7. *Mono-, di- and trimethylated aminoethylcysteine (adopted from* Simon, Chu et al. 2007*).*

1.4 Objectives of my study

As a general goal of my study, I aim to make a contribution to the field of protein chemical biology by developing new chemical methods to synthesize post-translationally modified proteins which are difficult to be accessed by existing

technologies.

My first objective is to expand the application scope of thioacid capture ligation. Since thioacid capture ligation confers extremely high efficiency, my study is to test whether it can be used for ligation at Val residue and compare the efficiency between thioester-mediated ligation and thioacid capture ligation. Moreover, I want to apply this ligation method to synthesize an acetylated histone H2B K120Ac. This work will be introduced in chapter 2.

The second objective of my study is to develop, through exploiting the unique reactivity of the cysteine thiol, a direct protein modification method for the introduction of N^{ε}-acetyl-lysine analogs into proteins. A previously unexplored free radical addition reaction is utilized for this purpose. By using this method, large quantities of the Lys(Ac)-modified proteins can be prepared, overcoming the limitation of conventional total or semi protein synthesis methods. This part of work is presented in chapter 3.

In the last chapter, I would like to report a system for peptide heterodimerization through PNA-templated disulfide bond formation. By making use of Watson-Crick base pairing, interchain disulfide bonds between different peptides can be formed in a selective manner.

References

Allfrey, V. G., R. Faulkner, et al. (1964). "Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis." Proc Natl Acad Sci U S A 51: 786-794.

Dawson, P. E. and S. B. H. Kent (2000). "Synthesis of Native Proteins by Chemical ligation." Annu. Rev. Biochem. 69: 923-960.

Dawson, P. E., T. W. Muir, et al. (1994). "Synthesis of Proteins by Native Chemical Ligation." Sci. 266(5186): 776-779.

Hackeng, T. M., J. H. Griffin, et al. (1999). "Protein Synthesis by Native Chemical Ligation: Expanded Scope by Using Straightforward Methodology." Proc. Natl. Acad. Sci. USA 96: 10068-10073.

Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." Science 293(5532): 1074-1080.

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

Kamakaka, R. T. and S. Biggins (2005). "Histone variants: deviants?" Genes Dev 19(3): 295-310.

Kenyon, G.L., and Bruice, T.W. (1977). "Novel sulfhydryl reagents."Methods Enzymol 47: 407–430.

Kimura, A., K. Matsubara, et al. (2005). "A decade of histone acetylation: marking eukaryotic chromosomes with specific codes." J Biochem 138(6): 647-662.

Klose, R. J. and Y. Zhang (2007). "Regulation of histone methylation by demethylimination and demethylation." Nat Rev Mol Cell Biol 8(4): 307-318.

Krieger, D. E., R. Levine, et al. (1974). "Chemical studies of histone acetylation. Substrate specificity of a histone deacetylase from calf thymus nuclei." J Biol Chem 249(1): 332-334.

Lan, F., P. E. Bayliss, et al. (2007). "A histone H3 lysine 27 demethylase regulates animal posterior development." Nature 449(7163): 689-694.

Liu, C. F., C. Rao, et al. (1996). "Acyl Disulfide-Mediated Intramolecular Acylation for Orthogonal Coupling Between Unprotected Peptide Segments. Mechanism and Application." Tetrahedron Letters 37: 933-936.

Merrifield, R. B. (1963). "Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide." J. Am. Chem. Soc. 85: 2149-2154.

Morrow, J. F. (1979). "Recombinant DNA techniques." Methods in Enzymology 68: 3-24.

Muir, T. W., D. Sondhi, et al. (1998). "Expressed protein ligation: a general method for protein engineering." Proc Natl Acad Sci U S A 95(12): 6705-6710.

Nilsson, B. L., M. B. Soellner, et al. (2005). "Chemical synthesis of proteins." Annu. Rev. Biophys. Biomol. Struct. 34: 91-118.

Peterson, C. L. and M.-A. Laniel (2004). "Histones and histone modifications." Current Biology 14: R546-R551.

Rice, J. C. and C. D. Allis (2001). "Histone methylation versus histone acetylation: new insights into epigenetic regulation." Current Opinion in Cell Biology 13: 263-273.

Shahbazian, M. D. and M. Grunstein (2007). "Functions of site-specific histone acetylation and deacetylation." Annu Rev Biochem 76: 75-100.

Shilatifard, A. (2008). "Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation." Curr Opin Cell Biol 20(3): 341-348.

Simon, M. D., F. Chu, et al. (2007). "The site-specific installation of methyl-lysine analogs into recombinant histones." Cell 128(5): 1003-1012.

Strahl, B. D. and C. D. Allis (2000). "The language of covalent histone modifications." Nature 403(6765): 41-45.

Swigut, T. and J. Wysocka (2007). "H3K27 demethylases, at long last." Cell 131(1): 29-32.

Tam, J. P., Y.-A. Lu, et al. (1995). "Peptide synthesis using unprotected peptides through orthogonal coupling methods." Proc. Natl. Acad. Sci. USA 92: 12485-12489.

Tam, J. P., Q. Yu, et al. (1999). "Orthogonal Ligation Strategies for Peptide and Protein." Biopolymers 51: 311-332.

Wieland, T., E. Bokelmann, et al. (1953). Justus Liebigs Ann. Chem. 583: 129-149.

Chapter 2: Thioacid Capture Ligation at Valine

2.1 Introduction

As the conventional recombinant technology cannot be used to synthesize proteins with PTMs, chemical synthesis which combines solid phase peptide synthesis with chemoselective peptide ligation is a useful approach for the production of these modified proteins. However, the requirement for the rare amino acid cysteine as the ligation site limits the application scope of both thioester-mediated and thioacid capture ligation methods. In natural proteins, Cys is not an abundant amino acid (1.7 %) compared with Val (6.7%) or other amnio acids with higher frequency (Fig. 1) (Nilsson, Soellner et al. 2005). For example, there is no cysteine in the histone proteins, H2A, H2B and H4, and only one cysteine in H3. If we want to synthesize the modified (methylated or acetylated) histones, a new method that does not require the use of Cys at the ligation site must be developed.

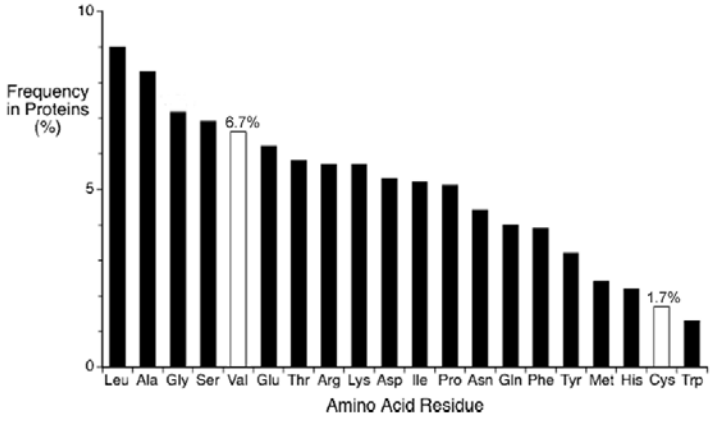


Figure 1. *Frequency of amino acid residue in proteins (adopted from* Nilsson, Soellner et al. 2005*).*

2.1.1 Thioester-mediated ligation at valine

Recently, several strategies have been developed to solve this problem (Canne, Bark et al. 1996; Offer and Dawson 2000). One strategy is to chemically convert cysteine to alanine by removing the free thiol from the cysteine by a desulfurization reaction after the ligation reaction. This approach can expand the application scope of the ligation method if X-Ala is present at a suitable location (Fig. 2) (Yan and Dawson 2001).

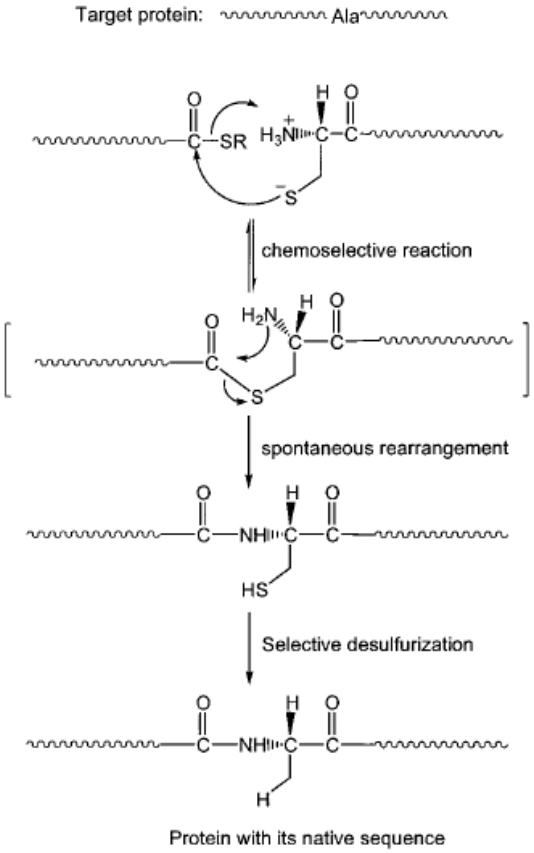


Figure 2. *Synthesis of proteins without Cys by thioester-mediated ligation combined with desulfurization (adopted from* Yan and Dawson 2001*).*

Similarly, penicillamine-mediated ligation also makes valine accessible as a ligation site (Chen, Wan et al. 2008; Haase, Rohde et al. 2008). The penicillamine is used as a precursor of valine and it can be used for ligation, as there is a free thiol on it which can mediate thiol-thioester exchange. After the ligation, the thiol group can be removed by desulfurization to give the natural valine at the ligation site (Fig. 3). Similar to the Cys→ Ala ligation strategy, this method can also extend the application scope of ligation method since valine is one of the most abundant amino acid residues in proteins.

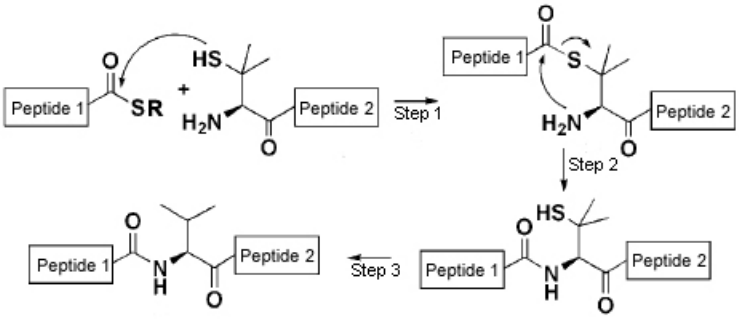


Figure 3. *Thioester-mediated ligation at valine. Step 1: Thiol exchange between the two peptides to form thioester intermediate. Step 2: Peptide bond formation by S-N acyl transfer. Step 3: The penicillamine was changed into valine by desulfurization (adopted from* Chen, Wan et al. 2008*).*

2.1.2 Strategies of our study: thioacid capture ligation at valine

Different from cysteine, the thiol of penicillamine is tertiary which would, because of steric hindrance, slow down the thiol-thioester exchange capture reaction in thioester-mediated ligation. As a matter of fact, ligation at valine using the thioester

approach exhibits very slow reaction kinetics, making it not practical for protein synthesis. Since thioacid capture ligation confers higher efficiency-as the thioacid group is a super-nucleophile and Npys is an excellent leaving group, we thought that the thioacid capture ligation might perform much better than the thioester-mediated ligation in this case.

In a previous study, only small peptides were used in model ligation reactions and the method has never been applied to synthesizing large modified proteins (Liu, Rao et al. 1996). In the present work, we have conducted a series of model studies using small synthetic peptide thioacids and shown that thioacid capture at valine is much more efficient than the corresponding thioester-mediated ligation at valine. Notably, ligation at the Pro-Val linkage, which is known to be very difficult for the thioester ligation method, has been achieved in respectable yields. We have also successfully synthesized an acetylated histone protein H2B K120Ac by this method (Fig. 4).

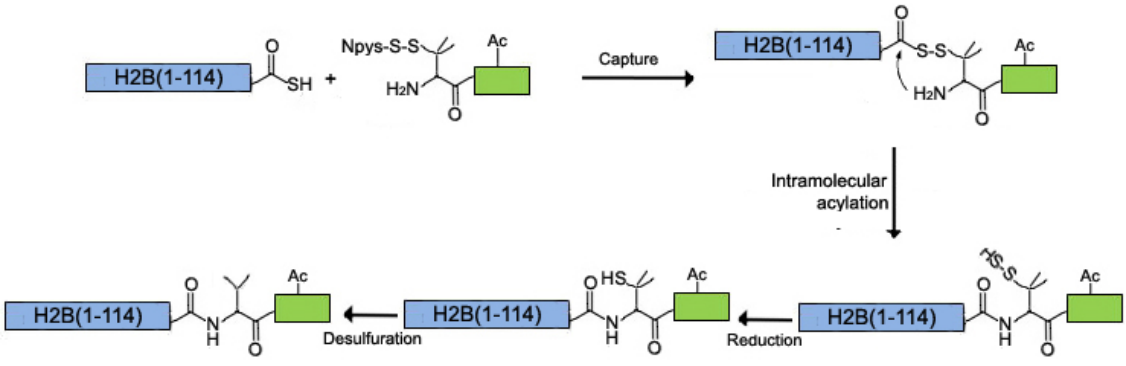


Figure 4. *Synthesis of H2B K120Ac by thioacid capture ligation at valine. Step 1: Capture to form disulfide bond. Step 2: Peptide bond formation by S→N acyl transfer. Step 3: Reduction to give the penicillamine. Step 4: Desulfurization to convert the penicillamine to valine.*

2.2 Results

2.2.1 Model peptide ligation by thioester-mediated and thioacid capture ligation methods

In order to apply the thioacid capture ligation at Val and compare the efficiency between thioester-mediated ligation and thioacid capture ligation, several model peptides with the sequence **EGTKXaa-SR** (R = H, thioacid; R = $CH_2CH_2CONH_2$, thioester; Xaa = Ala, Gly, Phe, Leu and Pro) have been synthesized. The peptide thioesters were synthesized by Boc SPPS. The peptide thioacids were prepared from corresponding thioesters by a hydrothiolysis reaction developed in our lab (Fig. 5) (Tan, Zhang et al. 2008).

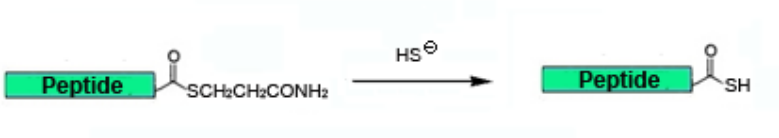
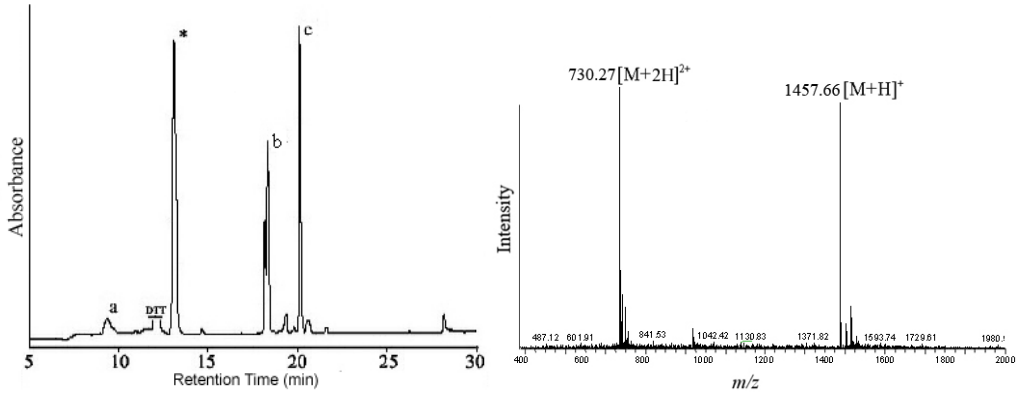


Figure 5. *Preparation of peptide thioacids by hydrothiolysis reaction.*

2.2.1.1 Model peptide ligation by thioacid capture ligation

At first, the peptide thioacid (**EGTKA-SH**) and Npys modified penicillyl peptide (**Pen(Npys)TK(Ac)YTSAK**) were dissovled in the ligation buffer (0.2 M phosphate, pH 6.0) in a ratio of acyl segment (1 mM) : amine segment (1.5 mM) = 1 : 1.5. After around 10 min at 37 ^{o}C, a yellow colour gradually appeared which indicated the release of Npys-H and therefore the capture reaction taking place through

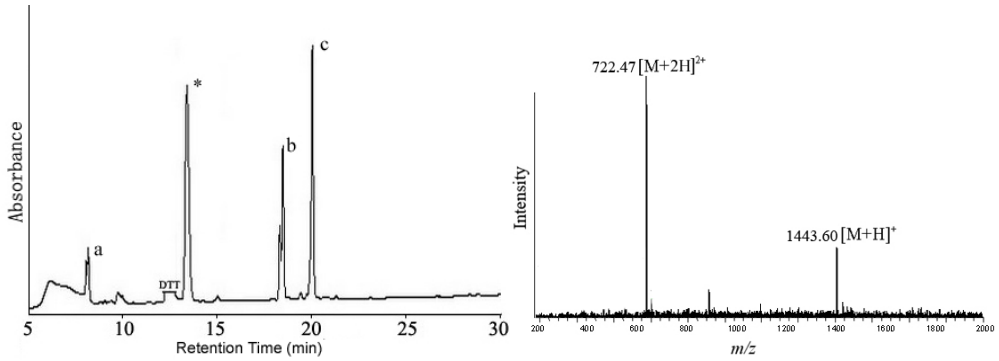
thio-disulfide exchange. This reaction was stopped after 1 h by adding DTT to the reaction mixture. As analyzed from the HPLC data (Fig. 6), the reaction gave a yield of about 80% on the basis of the thioacid componet.



A B

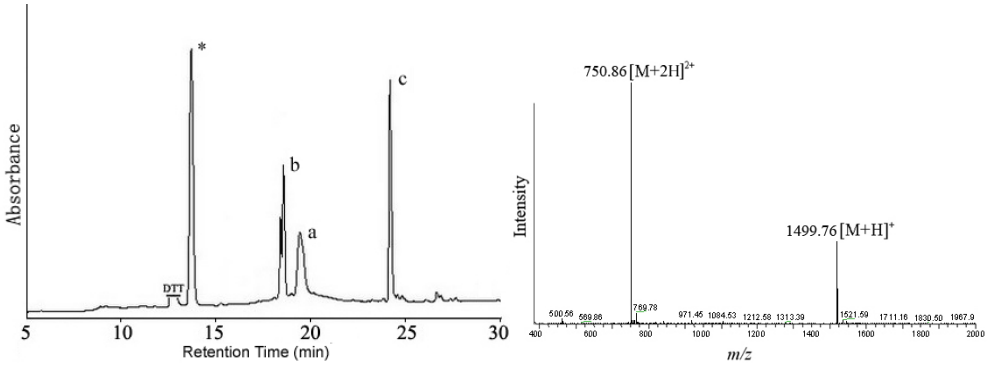
Figure 6. *Result of thioacid capture ligation between* ***EGTKA-SH*** *and* ***Pen(Npys)TK(Ac)YTSAK****. (A) C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKA-SH****; Peak* *******: Npys; Peak b:* ***PenTK(Ac)YTSAK****; Peak c: ligation product. (B) Mass spectrum of the ligation product determined by ESI-MS ([M+H]+ found: 1457.66, MW calcd: 1456.57). HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*

Similarly, the other four peptide thioacids with the last residue Gly/Leu/Phe/Pro were ligated with the same penicillyl peptide for 1 h, respectively. The results are shown in Figure 7, 8, 9 and 10. The thioacid capture ligation also performed well for these reactions with a yield of 87 %, 50 %, 83 % and 55 %, respectively. The yield is relatively low when the last residue is Leu or Pro, because the side chains of the two residues impose stronger steric hindrance compared with other residues.



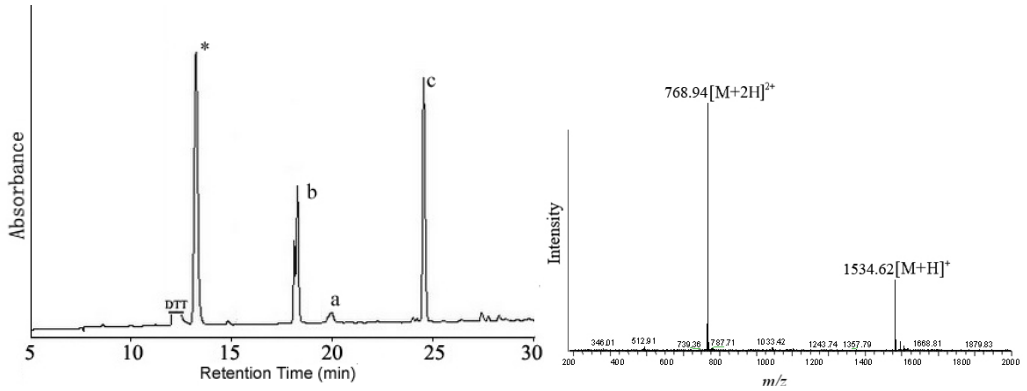
A B

Figure 7. *Result of thioacid capture ligation between* ***EGTKG-SH*** *and* ***Pen(Npys)TK(Ac)YTSAK****. (A) C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKG-SH****; Peak* *******: Npys; Peak b:* ***PenTK(Ac)YTSAK****; Peak c: ligation product. (B) Mass spectrum of the ligation product determined by ESI-MS ([M+H]*$^{+}$ *found: 1443.6, MW calcd: 1442.54). HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*



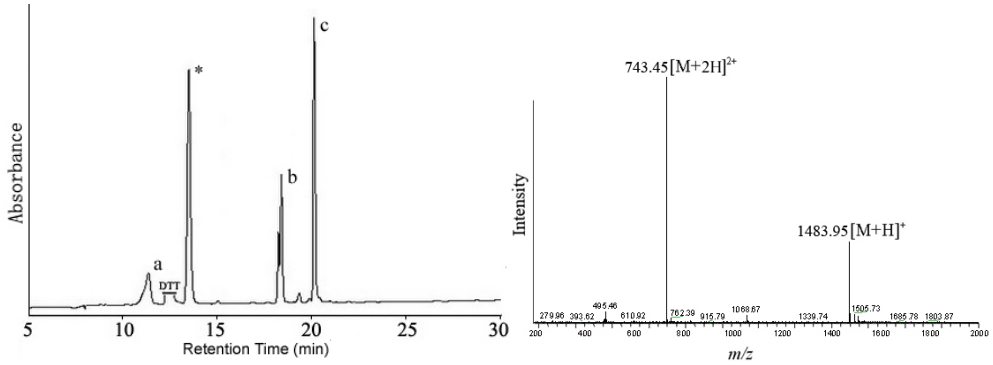
A B

Figure 8. *Result of thioacid capture ligation between* ***EGTKL-SH*** *and* ***Pen(Npys)TK(Ac)YTSAK****. (A) C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKL-SH****; Peak* *******: Npys; Peak b:* ***PenTK(Ac)YTSAK****; Peak c: ligation product. (B) Mass spectrum of the ligation product determined by ESI-MS ([M+H]*$^{+}$ *found: 1499.76, MW calcd: 1498.65). HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*



A B

Figure 9. *Result of thioacid capture ligation between* ***EGTKF-SH*** *and* ***Pen(Npys)TK(Ac)YTSAK****. (A) C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKF-SH****; Peak *: Npys; Peak b:* ***PenTK(Ac)YTSAK****; Peak c: ligation product. (B) Mass spectrum of the ligation product determined by ESI-MS ([M+H]*$^{+}$ *found: 1534.62, MW calcd: 1533.67). HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*

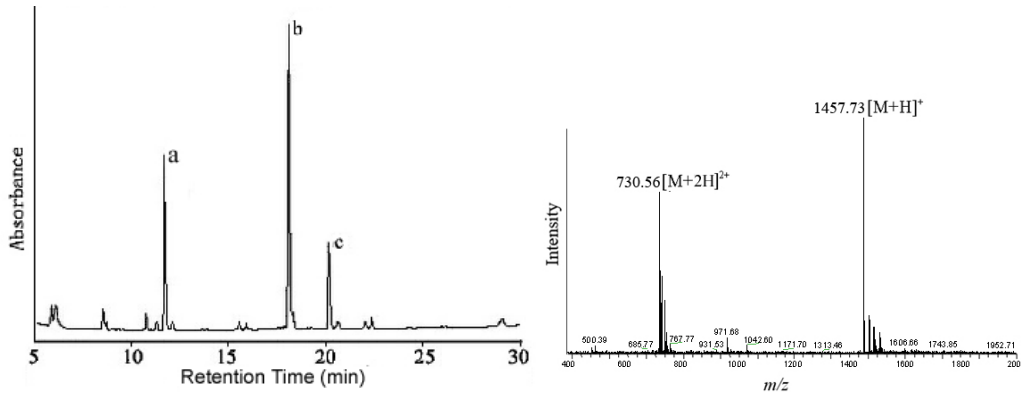


A B

Figure 10. *Result of thioacid capture ligation between* ***EGTKP-SH*** *and* ***Pen(Npys)TK(Ac)YTSAK****. (A) C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKP-SH****; Peak *: Npys; Peak b:* ***PenTK(Ac)YTSAK****; Peak c: ligation product. (B) Mass spectrum of the ligation product determined by ESI-MS ([M+H]*$^{+}$ *found: 1483.95, MW calcd: 1483.61). HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*

2.2.1.2 Model peptide ligation by thioester-mediated ligation

On the other hand, for thioester-mediated ligation, peptide thioester (**EGTKA-SR, R=$SCH_2CH_2CONH_2$**) and penicillyl peptide (**PenTK(Ac)YTSAK**) were first dissolved in the freshly prepared ligation buffer (100 mM NaH_2PO_4, 50 mM TCEP, 5% 4-mercaptobenzoic acid, pH 8.0), at a ratio of acyl peptide (1 mM) : amine peptide (1.5 mM) = 1 : 1.5, the reaction mixture was incubated at 37 ^oC and led to react for 2 h. This reaction only gave a yield of 8% (Fig. 10).



A B

Figure 10. *Result of thioester-mediated ligation between* ***EGTKA-SR*** *and* ***PenTK(Ac)YTSAK****. (A) C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKA-SR****; Peak b:* ***PenTK(Ac)YTSAK****; Peak c: ligation product. (B) Mass spectrum of the ligation product determined by ESI-MS ($[M+H]^+$ found: 1457.73, MW calcd: 1456.57). HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*

For the peptide thioester whose last residue was Gly, around 19 % of the starting material was reacted with penicillyl peptide to form the ligation product after 2 h of

reaction (Fig. 11). As the peptide is too small and the Gly is the easiest residue for ligation, about 20 % of peptide thioester was cyclized to form the cyclized peptide which decreased the ligation yield. The other four peptide thioesters with the last residue Leu/Phe/Pro were also ligated with the same penicillyl peptide, respectively. For Leu and Phe, the ligation yields were only around 6 % - 8 %. (Fig. 12 and 13). For Pro, no ligation product was detected (Fig. 14).

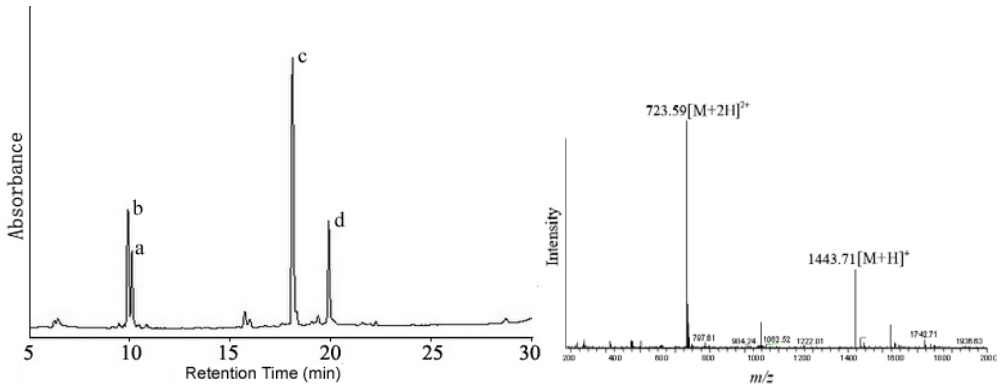
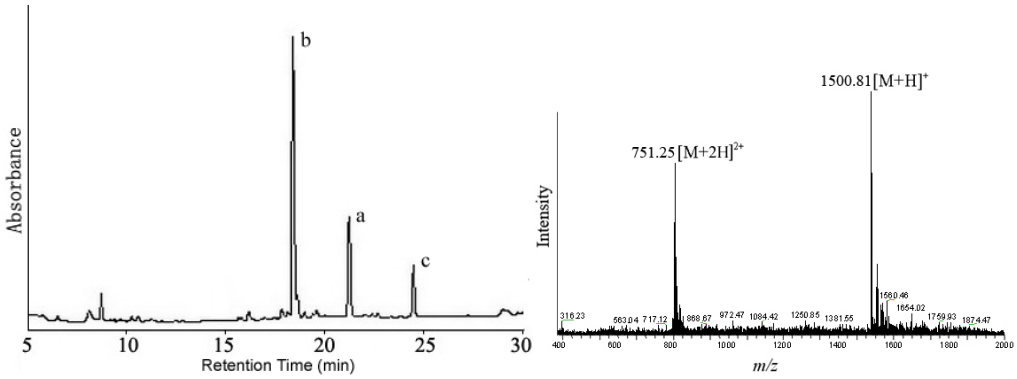
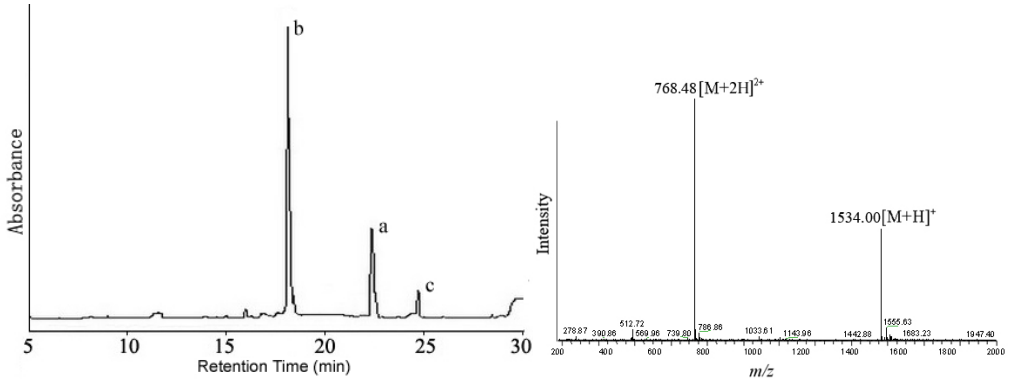


Figure 11. *Result of thioester-mediated ligation between* ***EGTKG-SR*** *and* ***PenTK(Ac)YTSAK****. (A) C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKG-SR****; Peak b: Cyclized peptide; Peak c:* ***PenTK(Ac)YTSAK****; Peak d: ligation product. (B) Mass spectrum of the ligation product determined by ESI-MS ($[M+H]^+$ found: 1443.71, MW calcd: 1442.54). HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*



A B

Figure 12. *Result of thioester-mediated ligation between* ***EGTKL-SR*** *and* ***PenTK(Ac)YTSAK****. (A) C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKL-SR****; Peak b:* ***PenTK(Ac)YTSAK****; Peak c: ligation product. (B) Mass spectrum of the ligation product determined by ESI-MS ($[M+H]^+$ found: 1500.81, MW calcd: 1499.65). HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*



A B

Figure 13. *Result of thioester-mediated ligation between* ***EGTKF-SR*** *and* ***PenTK(Ac)YTSAK****. (A) C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKF-SR****; Peak b:* ***PenTK(Ac)YTSAK****; Peak c: ligation product. (B) Mass spectrum of the ligation product determined by ESI-MS ($[M+H]^+$ found: 1534.62, MW calcd: 1533.67). HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*

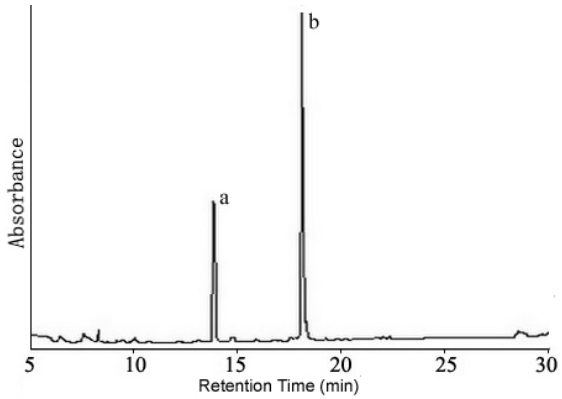


Figure 14. *Result of thioester-mediated ligation between* ***EGTKP-SR*** *and* ***PenTK(Ac)YTSAK****. C18 analytical HPLC profile of the ligation. Peak a:* ***EGTKP-SR****; Peak b:* ***PenTK(Ac)YTSAK****. HPLC condition: 0% to 30% of buffer B in buffer A in 30 min.*

2.2.1.3 Comparison between the two ligation method at Val

As shown in Table 1, thioacid capture ligation worked well for all residues studied and the reaction efficiency was much higher than thioester-mediated ligation. The most stricking difference is found when the last residue is Pro. We could not detect any ligation product in the thioester-mediated ligation. On the other hand, thioacid capture ligation gave a ligation yield of 55%. From the results, we can conclude that thioacid capture ligation works very well for ligation at the sterically hindered penicillamine. In fact, it is much more efficient than the thioester-mediated ligation.

Thioacid capture ligation			
Peptide thioacid	Penicillyl peptide	Product	Yield (1 h)
EGTKA-SH	**Pen(Npys)TK(Ac)YTSAK**	**EGTKAPenTK(Ac)YTSAK**	80%
EGTKG-SH	**Pen(Npys)TK(Ac)YTSAK**	**EGTKGPenTK(Ac)YTSAK**	87%
EGTKL-SH	**Pen(Npys)TK(Ac)YTSAK**	**EGTKLPenTK(Ac)YTSAK**	50%
EGTKF-SH	**Pen(Npys)TK(Ac)YTSAK**	**EGTKFPenTK(Ac)YTSAK**	83%
EGTKP-SH	**Pen(Npys)TK(Ac)YTSAK**	**EGTKPPenTK(Ac)YTSAK**	55%
Thioester-mediated ligation			
Peptide thioester	Penicillyl peptide	Product	Yield (2 h)
EGTKA-SR	**PenTK(Ac)YTSAK**	**EGTKAPenTK(Ac)YTSAK**	8%
EGTKG-SR	**PenTK(Ac)YTSAK**	**EGTKGPenTK(Ac)YTSAK**	19%
EGTKL-SR	**PenTK(Ac)YTSAK**	**EGTKLPenTK(Ac)YTSAK**	6.1%
EGTKF-SR	**PenTK(Ac)YTSAK**	**EGTKFPenTK(Ac)YTSAK**	7.6%
EGTKP-SR	**PenTK(Ac)YTSAK**	**EGTKPPenTK(Ac)YTSAK**	n.p.

Table 1. *Yields of the ligations. n.p.: no product observed.*

2.2.2 Synthesis of H2B K120Ac

2.2.2.1 Preparation of the H2B(4-117)-COSR and H2B(4-117)-COSH

To demonstrate the feasibility of thioacid capture ligation at Val for protein synthesis, we attempted the semi-synthesis of a model protein, histone H2B with acetylation at Lys 120 (H2B K120Ac) using both thioester-mediated ligation and thioacid capture ligation. For the synthesis of H2B K120Ac, V118 was selected as ligation junction. The pTWIN2 vector available from NEB contains a modified *Mth RIR1* intein which cannot undergo self-splicing and was used to genereate the protein thioester and thioacid (Xu and Evans 2001). The target gene coding for histone H2B(4-117) digested by NdeI and SapI was inserted into the pTWIN2 vector after an indentical

digestion, resulting in the fusion of the target gene for H2B(4-117) to the N-terminus of the Mth RIR1 intein. The engineered E. coli strain BL21(DE3) was transfected with pH2B(1-114)-TWIN2 to overexpress the target fusion protein H2B(4-117)-intein-CBD as shown in Fig. 15A (Lane 3). After thiol cleavage by MESNA or Na_2S, the thioester product H2B(4-117)-COSR or thioacid product H2B(4-117)-COSH was eluted out as shown in Fig 15A (Lane 4 and 5). Finally, after purification by C4 Semi-prep HPLC, the MW of the protein thioester and of the protein thioacid was confirmed by MALDI-MS (Fig. 15B and 15C).

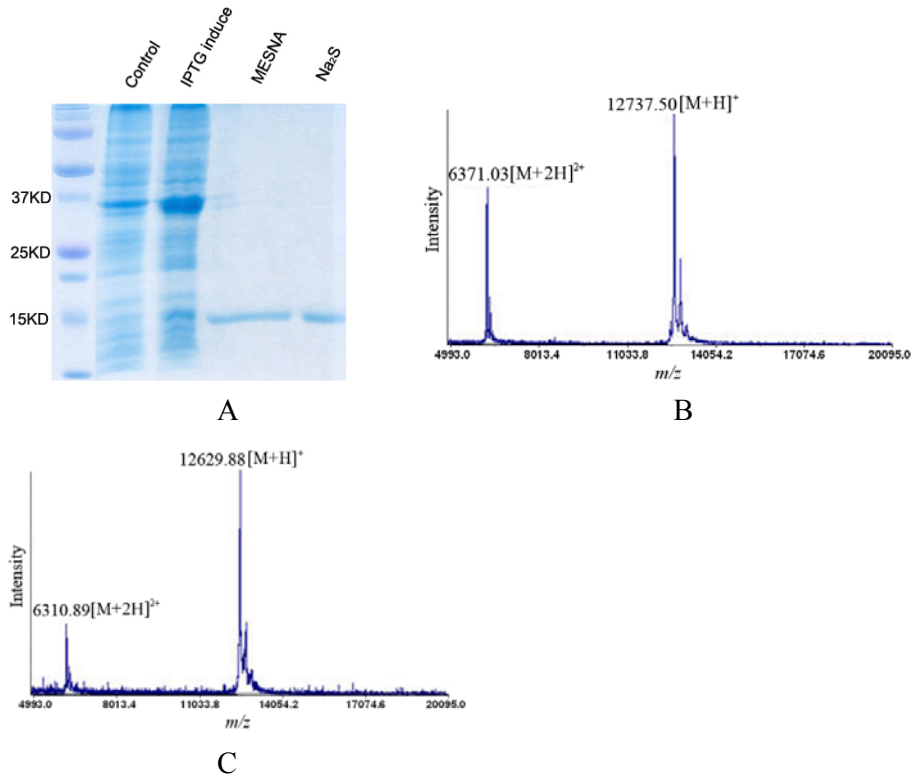
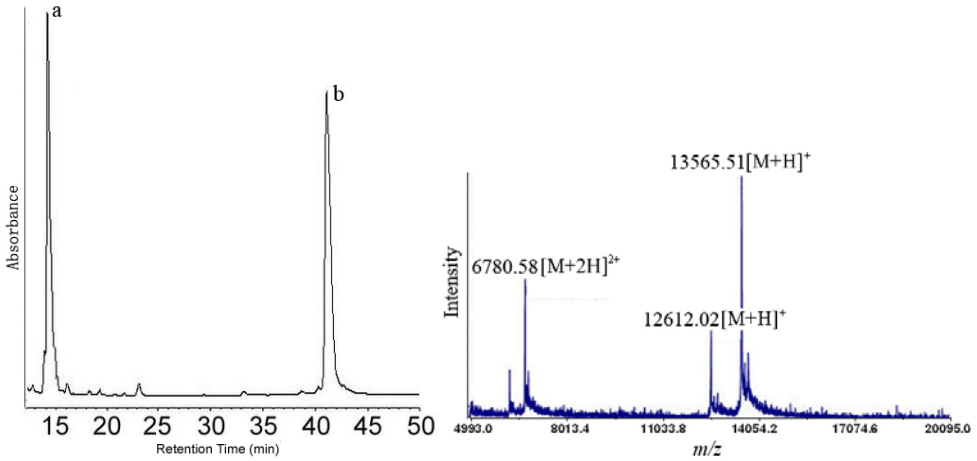


Figure 15. *(A) 15% SDS-PAGE Gel analysis of overexpression of H2B(4-117)-intein fusion protein after IPTG induction, H2B(4-117)-COSR cleaved by MESNA, and H2B(4-117)-COSH cleaved by Na_2S. (B) Mass spectrum of the H2B(4-117)-COSR determined by ESI-MS ($[M+H]^+$ found: 12737.50, MW calcd: 12737.71). (C) Mass spectrum of the H2B(4-117)-COSH determined by ESI-MS ($[M+H]^+$ found: 12629.88, MW calcd: 12630.71).*

2.2.2.2 Protein ligation by thioacid capture ligation at Val

In order to synthesize the H2B K120Ac, peptide **Pen(Npys)TK(Ac)YTSAK** was first mixed with H2B(4-117)-COSH in the ligation buffer (6 M Guanidinium HCl, 0.2 M phosphate, pH 6.0) at 37 ^{o}C. This reaction was stopped after 2 h by adding DTT to the reaction mixture. Although the ligation proceeded well, the ligation product could not be separated from the hydrolyzed protein H2B(4-117)-COOH because the two coeluted on HPLC (Fig. 16). The mass of hydrolyzed protein and the ligation product were both detected by MALDI-TOF MS.

To solve this problem in product purification, we realized that introduction of a large hydrophobic group onto the 2nd ligation component, such as Fmoc, may significantly delay the retenion time of the ligation product. Therefore, peptides **Pen(Npys)TK(Ac)YTSAK(Fmoc)** and **PenTK(Ac)YTSAK(Fmoc)** with Fmoc group at the side chain of the nonacetylated lysine were used for the synthesis of H2B K120Ac.



A B

Figure 16. *Result of thioacid capture ligation between H-Pen(Npys)-Thr-Lys(Ac)-Tyr-Thr-Ser-Ala-Lys-OH and H2B(1-114)-COSH. (A) C4 analytical HPLC profile of the ligation. Peak a: H-Pen-Thr-Lys(Ac)-Tyr-Thr-Ser-Ala-Lys-OH; Peak b: the hydrolyzed H2B(1-114) and ligation product eluted together. (B) Mass spectrum of the hydrolyzed protein ($[M+H]^+$ found: 12612.02, MW calcd: 12615.71) and ligation product determined by MALDI-MS ($[M+H]^+$ found: 13790.96, MW calcd: 13789.74). HPLC condition: 0% to 20% in 10 min, then to 60% in 40 min of buffer B in buffer A.*

The truncated protein H2B(4-117)-COSH was mixed with penicillyl peptide **Pen(Npys)TK(Ac)YTSAK(Fmoc)** containing an Fmoc in the ligation buffer (6 M Guanidinium HCl, 0.2 M phosphate, pH 6.0) at 37 ^{o}C for 2 h. The peak of the ligation product was then separated very well from the hydrolyzed H2B truncated protein H2B(4-117)-COOH, as there was an Fmoc on the last lysine. After ligation, the product was easily purified from the hydrolyzed protein (Fig. 17A). The result indicated that the ligation had occurred efficiently and the yield was ~67 %.

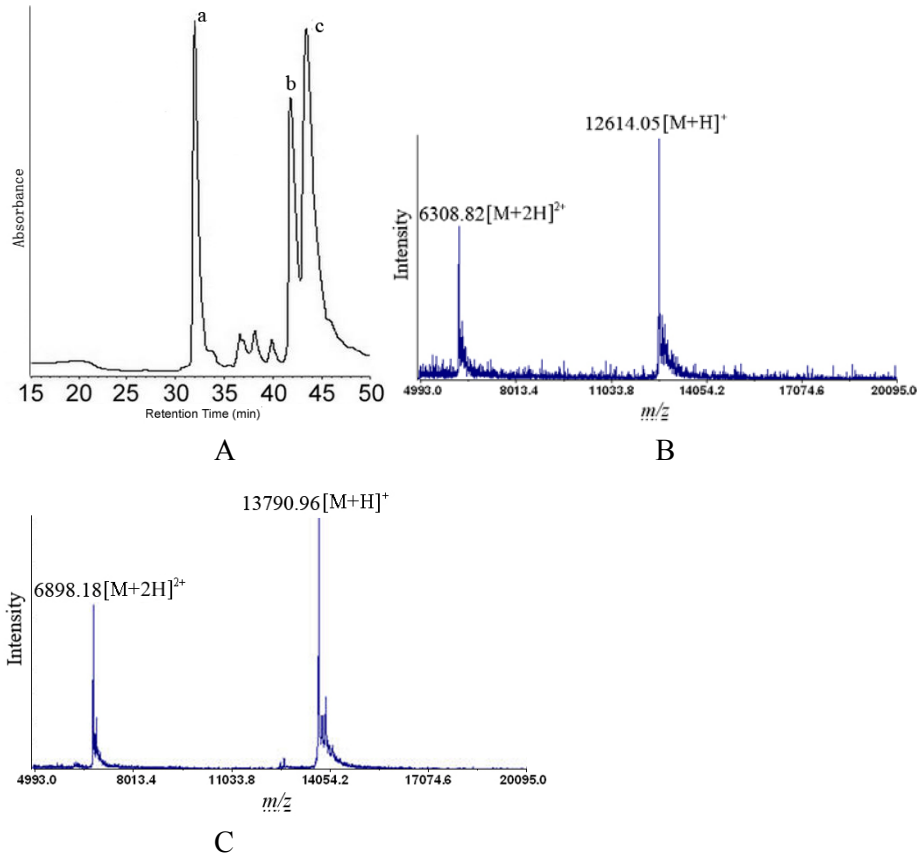


Figure 17. *Result of thioacid capture ligation between* ***Pen(Npys)TK(Ac)YTSAK(Fmoc)*** *and H2B(1-114)-COSH. (A) C4 analytical HPLC profile of the ligation. Peak a:* ***PenTK(Ac)YTSAK(Fmoc)****; Peak b: hydrolyzed H2B(1-114); Peak c: ligation product. (B) Mass spectrum of the hydrolyed H2B(1-114) determined by MALDI-MS ($[M+H]^+$ found: 12614.0, MW calcd: 12614.71). (C) Mass spectrum of the ligation product determined by MALDI-MS ($[M+H]^+$ found: 13790.96, MW calcd: 13789.74). HPLC condition: 0% to 20% in 10 min，then to 60% in 40 min of buffer B in buffer A.*

2.2.2.3 Protein ligation by thioester-mediated ligation at Val

The peptide **PenTK(Ac)YTSAK(Fmoc)** with an Fmoc was mixed with H2B(1-114)-COSR (R=MESNA) in the ligation buffer (6 M Guanidinium HCl, 100 mM NaH_2PO_4, 50 m M TCEP, pH 8.0) at 37 ^oC for overnight and 5 % of

4-mercaptobenzoic acid was used as catalyst. As one can see from Fig. 18A, only about 24% of ligation product was found after an overnight reaction. The longer reaction time in this protein ligation experiment may account for the higher yield than in the model peptide ligation study. The benzyl mercaptan was also tested and the result was shown in Fig. 19A. But in this case, no ligation product was found.

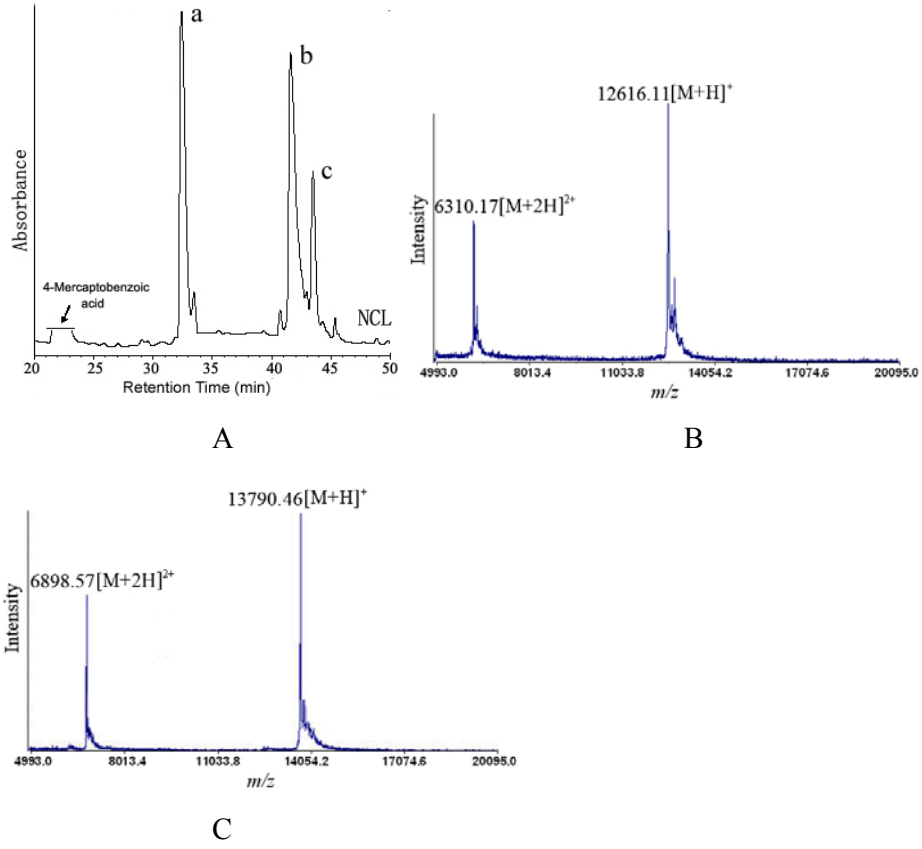
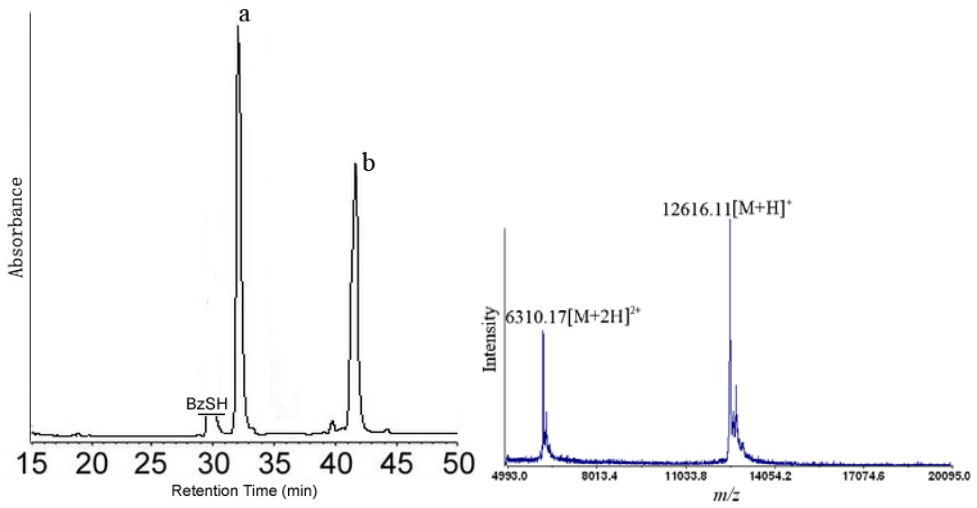


Figure 18. *Result of thioester-mediated ligation between* ***PenTK(Ac)YTSAK(Fmoc)*** *and H2B(1-114)-COSR when* 4-mercaptobenzoic acid was used as catalyst. *(A) C4 analytical HPLC profile of the ligation. Peak a:* ***PenTK(Ac)YTSAK(Fmoc)****; Peak b: hydrolyzed H2B(1-114); Peak c: ligation product. (B) Mass spectrum of the hydrolyed H2B(1-114) determined by MALDI-MS ($[M+H]^+$ found: 12616.11, MW calcd: 12614.71). (C) Mass spectrum of the ligation product determined by MALDI-MS ($[M+H]^+$ found: 13790.46, MW calcd: 13789.74). HPLC condition: 0% to 20% in 10 min，then to 60% in 40 min of buffer B in buffer A.*



A B

Figure 19. *Result of thioester-mediated ligation between* ***PenTK(Ac)YTSAK(Fmoc)*** *and H2B(1-114)-COSR when benzyl mercaptan* was used as catalyst*. (A) C4 analytical HPLC profile of the ligation. Peak a:* ***PenTK(Ac)YTSAK(Fmoc)****; Peak b: hydrolyzed H2B(1-114). (B) Mass spectrum of the hydrolyzed H2B(1-114) determined by MALDI-MS ($[M+H]^+$ found: 12616.11, MW calcd: 12615.71). HPLC condition: 0% to 20% in 10 min，then to 60% in 40 min of buffer B in buffer A. BzSH: Benzyl mercaptan.*

The above result indicated that thioacid capture ligation also performed well in the synthesis of the H2B protein analog (Fig. 17). On the other hand, the thioester-mediated ligation only got a yield of 24%, if 4-mercaptobenzoic acid was used as the catalyst (Fig. 18). If benzyl mercaptan was used as the catalyst, there was no ligation product (Fig. 19). This is because, as an aryl mercaptan, 4-mercaptobenzoic acid is a better leaving group than benzyl mercaptan – an alkyl thiol. Therefore, compared to the thioester-mediated ligation, thioacid capture ligation is clearly much better and much more applicable to protein synthesis when the

sterically hindered Val is used as the ligation site.

2.2.2.4 Purification of the ligation product

Although the peak of the ligation product with the hydrophobic Fmoc group present could be separated very well from that of the hydrolyzed protein H2B(4-117)-COOH, the removal of Fmoc from the ligation product was very difficult. The classical method for Fmoc group removal is to use 20% piperidine in DMF, but the H2B protein cannot dissolve in DMF and also the use of piperidine may cause some side reactions on a free protein. So this method cannot be applied here. Tetrabutylammonium fluoride in THF was also tried (Ueki and Amemiya 1987; Ueki, Nishigaki et al. 1993). However, the H2B protein cannot dissolve in THF either. Other solvent systems like DCM/TFE or 6 M Guanidine buffer also did not give positive results.

As seen in figure 16, when the peptide **Pen(Npys)TK(Ac)YTSAK** without Fmoc was used for ligation, the ligation product was coeluted with the hydrolyzed protein H2B(4-117)-COOH. A new method must be developed to purify the product. We took advantage of the fact that there was a free thiol in the ligation product which can be used for modification by disulfide bond to make it more separable from the hydrolyzed product (Fig. 20). Therefore, a purification petptide containing an activated S-S bond, TK^{Fmoc}-7, **Npys-SCH_2CO-TK(Ac)YTSAK(Fmoc)** was

designed and synthesized. The Npys modified thiol at the N-ter of the peptide $TK^{Fmoc-}7$ can react with the thiol of the Pen residue on the ligation product to form a disulfide bond. On the other hand, there is no free thiol on the hydrolyzed H2B(4-117). As a result, due to the hydrophobic effect of the Fmoc group on the purification peptide, the S-S linked conjugate between the ligation product and TK^{Fmoc}-7can be easily separated from the hydrolyzed H2B(4-117).

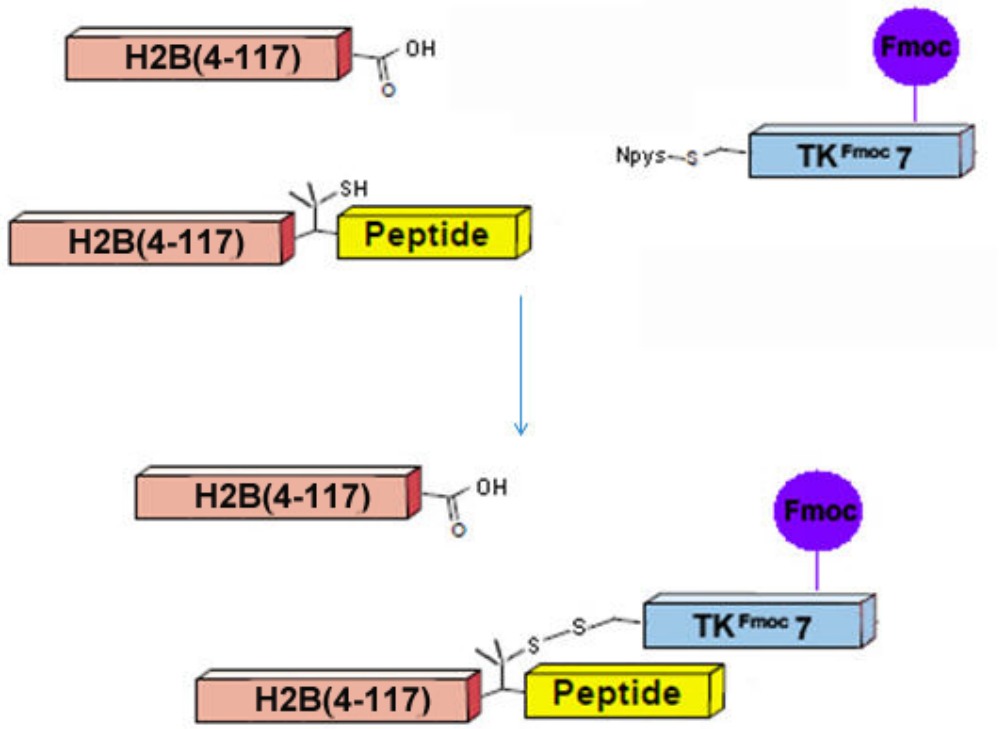
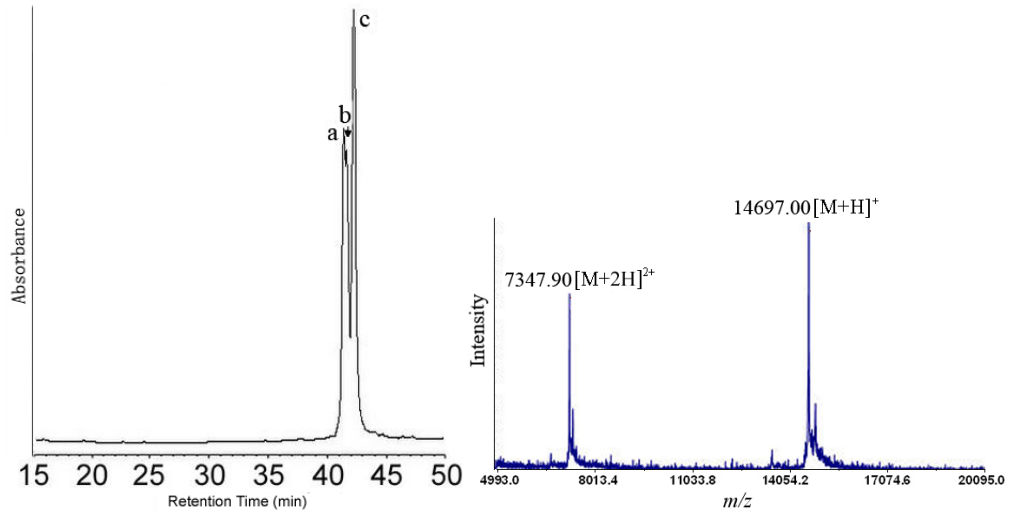


Figure 20. *A strategy for purification of ligation product.*

The result is shown in figure 21A. The mass of peak c (Fig. 21B) indicated that the disulfide bond was formed successfully between the ligation product and peptide $TK^{Fmoc-}7$. After isolation of peak c which corresponded to the S-S linked conjugate, the desired ligation product could be obtained after reduction of the S-S bond in a reduction reaction.

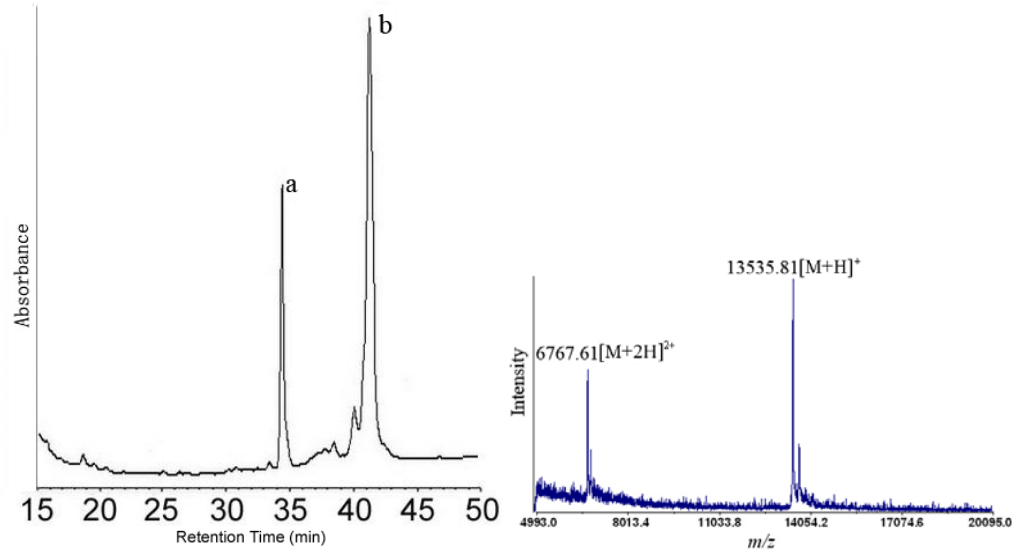


A B

Figure 21. *Purification of the ligation product. (A) C4 analytical HPLC profile of the reaction. Peak a: hydrolyzed H2B(1-114); Peak b:* ***Npys-SCH$_2$CO-TK(Ac)YTSAK(Fmoc)****; Peak c: S-S linked conjugate product. (B) Mass spectrum of the product determined by MALDI-MS ([M+H]$^+$ found: 14697.00, MW calcd: 14705.65). HPLC condition: 0% to 20% in 10 min，then to 60% in 40 min of buffer B in buffer A.*

2.2.2.5 Desulphurization and purification of H2B K120Ac

The ligation product being purified successfully contains a penicillamine residue from which the thiol must be removed to give the native valine. Raney nickel was used here to remove the sulfur. Because TCEP in the desulphurization buffer was used to reduce the disulfide bond between the ligation product and the peptide, after 5 h of the desulphurization reaction, the thiol on both the TK^{Fmoc}-7 peptide and the ligation product were removed successfully with high efficiency. The final product H2B K120Ac was purified by C4 Semi-Prep HPLC (Fig. 22A).



A B

Figure 22. *Purification of the desulfurization product. (A) C4 analytical HPLC profile of the reaction. Peak a: CH_3CO-* ***TK(Ac)YTSAK(Fmoc)****; Peak b: final product H2B K120Ac. (B) Mass spectrum of the final product determined by MALDI-MS ($[M+H]^+$ found: 13535.58, MW calcd: 13536.74). HPLC condition: 0% to 20% in 10 min，then to 60% in 40 min of buffer B in buffer A.*

2.2.3 The formation of histone octamer

This semisynthesized H2B K120Ac was used to form the octamer with recombinant H2A, H2B and H4 in equal molar amount (Luger, Mäder et al. 1997; Rhodes 1997). The histones were individually dissolved in an unfolding buffer (7 M Guanidinium HCl, 10 mM Tris-HCl pH7.5, 10 mM DTT) and then mixed together. The formation of histone octamer was achieved by serial dialysis from 7 M guanidinium-HCl to 2 M NaCl. The octamer was purified by gel filtration, the result is shown in figure 23. Only one large peak corresponding to the correctly formed histone octamer was

purified. No other significant peaks for the histone monomers or dimers were found. The octamer peak was analyzed by 15% SDS-PAGE, which is shown in Fig. 24. The result was the same as the octamer formed from the control proteins. The results show that the semisynthesized H2B K120Ac could bind with other three histones to form the correct octamer.

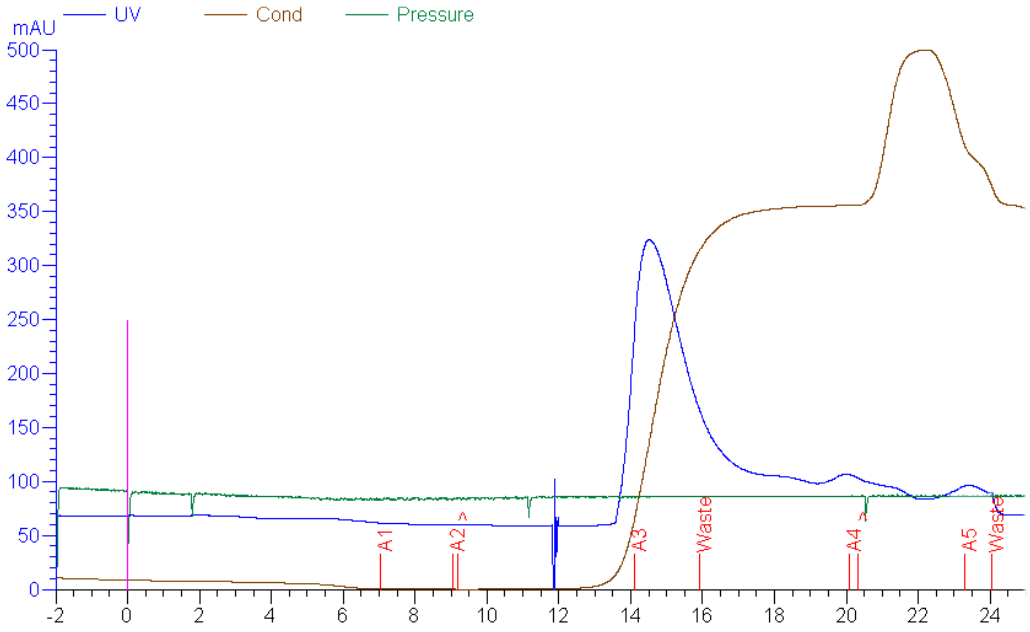


Figure 23. *Histone octamer purification by gel filtration.*

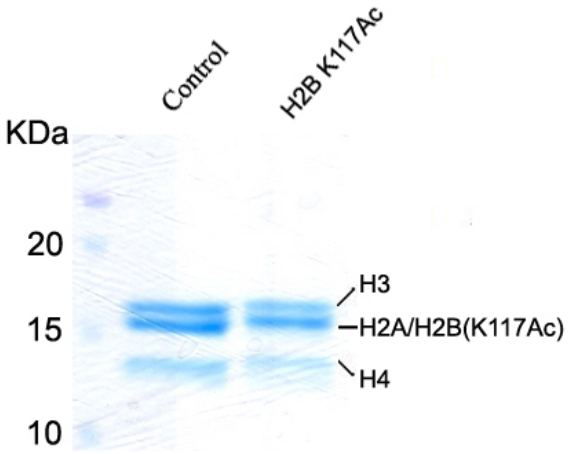


Figure 24. *15% SDS-PAGE gel analysis of histone octamer formation.*

2.2.4 The formation of nucleosome core particles

We also tested the binding of the octamer containing H2B K120Ac with a 185 bp DNA template. The results were analyzed by 5 % native PAGE gel stained by ethidium bromide (Fig. 25). Again, this octamer and the control octamer behaved in the same way in binding to the DNA. From this result, we can conclude that this modified octamer can assemble with the template DNA to form the nucleosome core particle. These preliminary results indicate that the function of the semisynthesized H2B K120Ac is identical to that of the natural H2B in the above two assays.

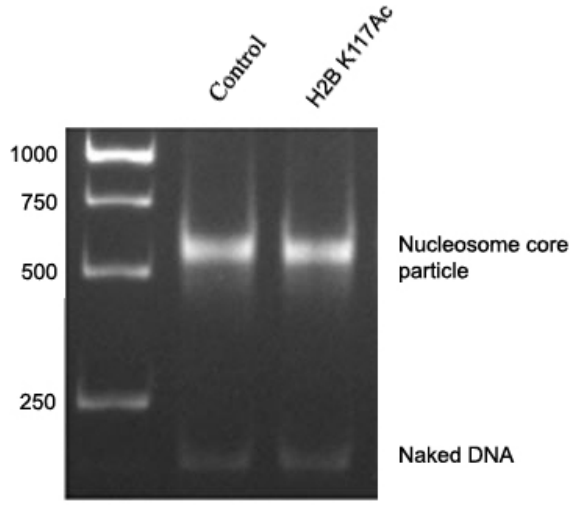


Figure 25. *Analysis of nucleosome core particle assembly on 5% native PAGE gel.*

2.3 Discussion and conclusion

In my research work, I have explored the possibility of peptide ligation at a sterically

hindered Val residue. I have used several model peptides to compare the ligation efficiency of two different ligation methods. In model peptide ligation, when using the thioester-mediated ligation method, only the peptide thioester with a C-ter Gly can achieve a yield of around 19 %. The other four peptides with a C-ter Ala/Leu/Phe/Pro, respectively, gave litte or no ligation product under the same conditions. On the other hand, the thioacid capture ligation was found to be faster compared with thioester-mediated ligation. The ligation yield was very high with a C-ter Gly (87 %), Phe (83 %) or Ala (80 %). It was also good with a C-ter Leu (50 %) or even Pro (55 %). Notably, when thioacid capture ligation at Val was applied to protein ligation, it gave a yield of around ~67% in the synthesis of H2B protein as compared with a yield of only ~24% for the thioester-mediated ligation.

The different mechanisms of thioester-mediated ligation and thioacid capture ligation may account for the difference in ligation efficiency between the two methods. In thioester-mediated ligation, thiol-thioester exchange is the rate-limiting step. The tertiary thiol group in penicillamine would impose considerable steric hindrance on its engagement with the thioester, making the ligation reaction much less efficient than in the case of Cys ligation. In thioacid capture ligation, the first step is a very efficient thiol exchange reaction between a thioacid and a highly activated disulfide, which is much more efficient that the thiol-thioester exchange reaction in thioester-mediated ligation, because the thioacid group is a super-nucleophile and Npys is an excellent leaving group. Although steric hindrance of the tertiary thiol of penicillamine would

slow down the thioacid capture reaction to some degree, its highly efficient feature would still allow the ligation to proceed at a reasonable rate. In addition, the acyl disulfide intermediate is an activated carboxyl derivative which can drive the intramolecular acylation reaction to occur in a very efficient way.

Since the ligation product was eluted together with the hydrolyzed protein H2B(1-114)-COOH when the peptide **Pen(Npys)TK(Ac)YTSAK** without Fmoc was used for ligation, a new purification method was developed. The idea was to modify the ligation product with a hydrophobic moiety via the formation of a S-S bridge. Because of the hydrophobic properties of this moiety, the resultant conjugate should exhibit a significantly different elution time in reversed phase HPLC as compared to the hydrolyzed H2B (1-114)-COOH. For this purpose, **NpyS-SCH$_2$CO-TK(Ac)YTS AK(Fmoc)**, a hydrophobic small peptide containing an active S-S bond, was prepared and used to modify the H2B ligation product to facilitate its purification. The disulfide formation step was very fast and was complete after a few minutes. Also, the disulfide bond could be easily reduced by TCEP to release the ligation product for desulfurization. As expected, the S-S linked conjugate of the hydrophobic peptide and the ligation product was easily separated from the unligated hydrolyzed H2B (1-114). This new method can be generally useful for the isolation of ligation products which are difficult to separate from the other reaction components, as it is commonly seen in the semisynthesis of proteins where a large protein domain is used to ligate with a small peptide fragment.

The successful synthesis of H2B K120Ac by thioacid capture ligation further validate the practical value of our protein semisynthesis strategy. Because valine is a very common amino acid found in proteins, we predict that our strategy will be used more frequently than cysteine–dependent ligation in the synthesis of proteins which are modified at the C-terminal tails. The availability of such semisynthetic proteins will make it easier to study the biological functions of the post-translational modifications found in many proteins, such as the histones. Taken together, our results indicate that thioacid capture ligation is particularly useful for ligation at sterically unfavourable junctions.

2.4 Materials and Methods

2.4.1 Synthesis of peptide thioesters

All resin-bound peptide thioesters are of the following structure: peptide-CO-SCH_2CH_2CO-MBHA resin. First of all, thio-functionalized MBHA resin was prepared by coupling of Trt-SCH_2CH_2COOH and detritylation with 2.5 % TFA containing 2% TIS in DCM for 3 times. Peptides were assembled on the thiol-derived resin using standard SPPS protocols by Boc chemistry. All amino acids were used in 4 eq of the resin, and were preactivated by 4 eq of PyBop and 8 eq of DIEA. After 1 h coupling of the first amino acid, the resin was washed with DCM (3 x), DFM (3 x)

and DCM (3 x). The successive α-amino group deprotection steps were performed in 30 % TFA/DCM 2 x 10 min after each amino acid coupling step. After the coupling of all the amino acids, the peptide thioesters were cleaved by 10 % TFMSA, 10 % methylphenylsufide, 80 % TFA for 1 h at room temperature, followed by diethyl ether precipitation and C18 Semi-prep HPLC purification. The molecular weights of the peptide thioesters were determined by ESI-MS. The synthesized peptide thioesters are shown as following:

H-*Glu-Gly-Thr-Lys-Gly*-$SCH_2CH_2CONH_2$

H-*Glu-Gly-Thr-Lys-Ala*-$SCH_2CH_2CONH_2$

H-*Glu-Gly-Thr-Lys-Leu*-$SCH_2CH_2CONH_2$

H-*Glu-Gly-Thr-Lys-Pro*-$SCH_2CH_2CONH_2$

H-*Glu-Gly-Thr-Lys-Phe*-$SCH_2CH_2CONH_2$

2.4.2 Synthesis of peptide thioacids

All the peptide thioacids were prepared by dissolving the freeze-dried peptide thioesters in the hydrothiolysis buffer (250 mM HEPES, 100 mM Na_2S, pH 8.0). The reaction was carried out for 4 h, followed by C18 Semi-prep HPLC purification. The molecular weights of the peptide thioacids were confirmed by ESI-MS. The peptide thioacids are shown as following:

H-*Glu*-*Gly*-*Thr*-*Lys*-*Gly*-SH

H-*Glu*-*Gly*-*Thr*-*Lys*-*Ala*-SH

H-*Glu*-*Gly*-*Thr*-*Lys*-*Leu*-SH

H-*Glu*-*Gly*-*Thr*-*Lys*-*Pro*-SH

H-*Glu*-*Gly*-*Thr*-*Lys*-*Phe*-SH

2.4.3 Synthesis of H2B C-ter penicillyl peptide

2.4.3.1 Preparation of Boc-Pen(Trt)-OH

The penicillamine was purchased from *Alfa Aesar* company. The free thiol and amine on the penicillamine must be protected. The free thiol was protected by trityl (Lipowska, Hansen et al. 2002). 1 eq of penicillamine and 1.2 eq of triphenylmethanol was dissolved in pure TFA, and this mixture was stirred for 1 h. The TFA was evaporated to dryness. DCM was used to redissolve the residue. DIEA was used to adjust pH to above 8, and 1.1 eq of Boc anhydride was added. After stirring for 6 h, 1M HCl was added slowly to adjust the pH to 3-4. The product was extracted by DCM for three times. After evaporation of DCM, the crude material can be used directly for the coupling.

2.4.3.2 Solid phase peptide synthesis

The penicillyl peptide **PenTK(Ac)YTSAK** and **PenTK(Ac)YTSAK(Fmoc)** were

synthesized by Boc chemistry on the PAM resin. When coupling the first amino acid to the resin, 5 eq of it was used, and it was preactivated with 5 eq of DCC and 1.25 eq of DMAP. After 12 h coupling, the resin was washed with DCM (3 x), DMF (3 x) and DCM (3 x). A 2nd coupling was done for 2 h again. Capping was performed by using 5 % acetic anhydride, 2.5 % DIEA in DMF/DCM (1:1) for 2 x 10 min. The resin was washed with DCM (3 x), DFM (3 x) and DCM (3 x). After the capping step, the successive α-amino group deprotection steps were performed in 30 % TFA/DCM 2 x 10 min. When coupling the lysine[117], the side chain acetylated lysine, Boc-Lys(Ac)-OH was used. The last building block, Boc-Pen(Trt)-OH, was coupled by the same method. After this, the peptide was cleaved by treatment with 10 % TFMSA, 10 % methylphenylsufide, 80 % TFA for 1 h at room temperature, followed by diethyl ether precipitation and C18 semi-prep HPLC purification. The molecular weight of the peptide was confirmed by ESI-MS.

2.4.3.3 Synthesis of Npys modified H2B C-ter penicillyl peptide

1 eq of the penicillyl peptide and 4 eq of the 2,2'-dithiobio-(5-nitropyridine) were dissolved in acetic acid/H2O (9:1). The reaction was performed for 4-6 h, followed by C18 Semi-prep HPLC purification. The molecular weight was confirmed by ESI-MS.

2.4.4 Model peptide ligation by thioester-mediated ligation

The freeze-dried peptides were dissolved in degassed buffer (100 mM NaH_2PO_4, 50

mM TCEP, 5% 4-mercaptobenzoic acid, pH 8.0), in a ratio of 1 eq of the peptide thioester (1 mM) to 1.5 eq of the penicillyl peptide (1.5 mM). The reaction was mixed gently and the reaction tube was incubated at 37 ^{o}C for 2 h, and the ligation reaction was monitored by C18 analytical HPLC.

2.4.5 Model peptide ligation by thioacid capture ligation

The freeze-dried peptides were dissolved in ligation buffer (0.2 M Na_2HPO_4/0.2 M NaH_2PO_4=87.7/12.3, pH 6.0), in a ratio of 1 eq of the peptide thioacid (1 mM) to 1.5 eq of Npys modified penicillyl peptide (1.5 mM). The reaction was mixed gently and incubated at 37 ^{o}C for 1 h, then monitored by C18 analytical HPLC.

2.4.6 Construction of truncated H2B expression plasmid pH2B(4-117)-TWIN2

The Xenopus H2B truncated gene was amplified by PCR using the primers

H2B T_F: 5'- GGTGGTCATATGGCCAAGTCCGCTCCAGC-3' and

H2B T_R:5'- GGTGGTTGCTCTTCCGCAAGCCTTGGTGCCCTCGGAC-3'.

The PCR product was purified and ligated into the T-easy vector (Promega). After digestion with *Nde*I and *Sap*I restriction enzymes, the product was purified and ligated into the indentically digested pTWIN2 vector (New England Biolabs). The correct insert was then comfirmed by DNA sequencing.

2.4.7 Overexpression and purification of H2B(4-117)-COSR and H2B(4-117)-COSH

The plasmid pH2B(4-117)-TWIN2 was transformed into *E.coli* BL21 (DE3) cells. The cells were grown in LB medium (containing 100 μg/ml Ampicillin) at 37 ^{o}C with shaking at 250 rpm to an OD_{600} of 0.6-0.8. The desired protein was induced by 200 μM IPTG at 37 ^{o}C for 3 h. After centrifuge at 6000 rpm for 10 min, cell pellets from 1 liter of cells were resuspended in 50 ml wash buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM beta-mercaptoethanol, pH 7.5). Cells were broken by sonication, and the debris was removed by centrifugation at 20,000g for 30 min. The pellet was dissolved in unfolding buffer (6 M Guanidinium HCl, 10 mM Tris-HCl, 10 mM DTT, pH 7.5), dialysed against 4 M urea buffer(4 M Urea, 50 mM HEPES, 1 mM EDTA, 1 M NaCl, 0.2 mM TCEP), then to 2 M Urea buffer (2 M Urea, 50 mM HEPES, 1 mM EDTA, 1 M NaCl, 0.5 % Tritone, 0.25 % TWEEN, 0.2 mM TCEP). The supernatant was mixed with 6 ml pre-equilibrated chitin beads (New England Biolabs) at 4 ^{o}C for 5 h. The beads were then poured into a column and washed with 40 ml 2 M Urea buffer. The fusion protein was cleaved by adding 4 ml 2 M urea buffer containing 100 mM 2-mercaptoethanesulfonic acid (MESNA) or 100 mM Na_2S and incubation at 25 ^{o}C overnight. The H2B(4-117)-COSR or H2B(4-117)-COSH was eluted by 10 ml 2 M Urea buffer. C4 semi-prep HPLC was used to purify the proteins. The molecular weight was confirmed by MALDI-MS. Affinity binding, cleavage and purification steps were also monitored by SDS-PAGE.

2.4.8 Protein ligation by thioester-mediated ligation

The freeze-dried peptide and H2B(4-117)-COSR were dissolved in degassed thioester-mediated ligation buffer (6 M Guanidinium HCl, 100 mM NaH_2PO_4, 50 m M TCEP, 5% 4-mercaptobenzoic acid (or benzyl mercaptan), pH 8.0), in a ratio of 1 eq of the H2B(4-117)-COSR (0.2 mM) to 8 eq of the penicillyl peptide (1.6 mM). The reaction was mixed gently and the reaction tube was incubated at 37 ^{o}C for 24 h. The ligation reaction was monitored by C4 analytical HPLC.

2.4.9 Protein ligation by thioacid capture ligation

The freeze-dried Npys modified peptide and H2B(4-117)-COSH were dissolved in thioacid capture ligation buffer (6 M Guanidinium HCl, 0.2 M Na_2HPO_4/0.2 M NaH_2PO_4=87.7/12.3, pH 6.0), in a ratio of 1 eq of H2B(4-117)-COSH (0.2 mM) to 8 eq of peptide (1.6 mM). The reaction was mixed gently and incubated at 37 ^{o}C for reaction. After 1 h it was monitored by C4 analytical HPLC and the product was purified by C4 semi-prep HPLC.

2.4.10 Purification of H2B K120Ac

The Npys modified peptide **NpyS-SCH$_2$CO- TK(Ac)YTSAK(Fmoc)** was synthesized by Boc chemistry to help purify the ligation product from the hydrolyzed

H2B(4-117) through disulfide bond formation. 1 eq of ligation mixture (0.2 mM) and 5 eq of peptide (1 mM) were dissolved in the buffer (6 M Guanidinium HCl, 0.2 M Na_2HPO_4/0.2 M NaH_2PO_4=87.7/12.3, pH 6.0) for 10 min. Then the mixture was purified by C4 semi-prep HPLC.

2.4.11 Desulphurization

280 mg of nickel sulfate was dissolved in 2 ml ddH_2O first. Raney nickel was prepared by slowly adding 40 mg $NaBH_4$ to the nickel sulfate solution. The black solid Raney nickel was filtered and washed with 100 ml ddH_2O. Freeze-dried difulfide bond containing product (1 mg) was dissolved in 50 ul desulphurization buffer (6 M Guanidinium chloride, 200 mM sodium phosphate, 50 mM TCEP) and mixed with Raney nickel (Yan and Dawson 2001; Pentelute and Kent 2007; McGinty, Kim et al. 2008). After 5 h, 1 ml of desulphurization buffer was added twice to wash out the nickel, and the supernatant was combinded together after centrifugation. The final product was purified by C4 semi-prep HPLC.

2.4.12 Expression and purification of recombinant histone proteins

The plasmids pET-3a containing the Xenopus laevis histone H2A, H2B and H4 gene and pET-3d containing H3 gene were obtained from Dr. C.A. Davey's lab. The plasmids were transformed into E. coli strain BL21(DE3)/pLysS $CaCl_2$-competent

cells. Cells were grown in 2 x TY medium, and when the OD_{600} reached 0.6-0.8, 0.5 mM IPTG was added to induce the protein overexpression for 3 h at 37 ^{o}C. The cell pellet was centrifuged and resuspended in the wash buffer (20 mM phosphate buffer, pH 7.0, 0.5 M NaCl, 1 mM EDTA). Microfluider was used to break the cells, followed by centrifugation at 20,000 g for 10 min to remove the cell debris. The inclusion body was dissolved in 6 M guanidinium HCl buffer for 4 h. After centrifugation to remove all the insoluble material, the supernatant was purified by C18 prep RP-HPLC. The purified protein was lyophilized and the molecular weight was determined by MALDI-MS.

2.4.13 The formation of histone octamer

The four histones in equal molar amount (around 1 mg each) were individually dissolved in unfolding buffer (7 M Guanidinium HCl, 10 mM Tris-HCl pH7.5, 10 mM DTT). For H3 20 mM DTT was added. After 30 min the four proteins were combined together. The combined protein solution was put into a dialysis bag and dialyzed against 600 ml of refolding buffer (2 M NaCl, 10 mM Tris-HCl pH 7.5, 1 mM Na-EDTA, 10 mM beta-mercaptoethanol). After 4 h, dialysis was continued against a fresh 600 ml of refolding buffer. After another 4 h, a fresh 600 ml of refolding buffer was used for dialysis overnight. The dialyzed sample was centrifuged at 20,000 g for 10 min at room temperature. The supernatant was concentrated by an Amicon concentrator device (MW cut-off of 10 kDa) and purified by size-exclusion

chromatography (using the 26/60 Sephacryl S-200 column). The fractions were collected and confirmed by 15 % SDS-PAGE.

2.4.14 The formation of nucleosome core particle

The 185 bp DNA (Schalch, Duda et al. 2005) was obtained by PCR using the primers: 601 F1:5'-CGGGATCCCGGCGCCCTGGAGAATCCCGGTGCC-3' and 601 R1:5'-CGCTCGAGCGAAGATCTTCCATGCACAGGATGTATATATCTGAC-3' and was purified using the QIAGEN PCR purification kit. 6 μM DNA, 5.4 μM histone Octamer, 2 M KCl, 10 μM DTT were mixed together. The mixture was dialyzed against the TCS-0.85 buffer (0.85 M KCl, 20 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT). After 2.5 h, dialysis was switched to TCS-0.65 buffer (0.65 M KCl, 20 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT), and after 2.5 h to TCS-0.45 buffer (0.45 M KCl, 20 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT), and after 2.5 h to TCS-0 buffer (20 mM Tris-HCl pH 7.5, 1 mM EDTA, 1 mM DTT) for dialysis overnight. The mononucleosome was concentrated by using an Amicon concentrator device (MW cut-off of 10 kDa). The formation of the mononucleosome was verified by separation on a 5 % native PAGE gel (5 % acrylamide, 0.15 % bis-acrylamide) and staining with ethidium bromide.

References

Canne, L. E., S. J. Bark, et al. (1996). "Extending the Applicability of Native Chemical Ligation." J. Am. Chem. Soc. 118.

Chen, J., Q. Wan, et al. (2008). "Native Chemical Ligation at Valine: A Contribution to Peptide and Glycopeptide Synthesis." Angew. Chem. Int. Ed 47: 1-5.

Haase, C., H. Rohde, et al. (2008). "Native Chemical Ligation at Valine." Angew. Chem. Int. Ed 47: 6807-6810.

Lipowska, M., L. Hansen, et al. (2002). "Synthesis of new N2S2 ligands and Re(V)O(N2S2) analogues of 99mTc renal imaging agents. Characterization by NMR spectroscopy, molecular mechanics calculations, and X-ray crystallography." Inorganica Chimica Acta 339: 327-340.

Liu, C. F., C. Rao, et al. (1996). "Acyl Disulfide-Mediated Intramolecular Acylation for Orthogonal Coupling Between Unprotected Peptide Segments. Mechanism and Application." Tetrahedron Letters 37: 933-936.

Luger, K., A. W. Mäder, et al. (1997). "Crystal structure of the nucleosome core particle at 2.8 Å resolution." Nature 389: 251-260.

McGinty, R. K., J. Kim, et al. (2008). "Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation." Nature 453: 812-816.

Nilsson, B. L., M. B. Soellner, et al. (2005). "Chemical synthesis of proteins." Annu. Rev. Biophys. Biomol. Struct. 34: 91-118.

Offer, J. and P. E. Dawson (2000). "Nα-2-Mercaptobenzylamine-Assisted Chemical Ligation." 2: 23-26.

Pentelute, B. L. and S. B. H. Kent (2007). "Selective Desulfurization of Cysteine in the Presence of Cys(Acm) in Polypeptides Obtained by Native Chemical Ligation." Organic Letters 9: 687-690.

Rhodes, D. (1997). "Chromatin structure: The nucleosome core all wrapped up." Nature 389: 231-232.

Schalch, T., S. Duda, et al. (2005). "X-ray structure of a tetranucleosome and its implications for the chromatin fibre." Nature 436(138-141).

Tan, X.-H., X. Zhang, et al. (2008). "A Simple Method for Preparing Peptide C-Terminal Thioacids and Their Application in Sequential Chemoenzymatic Ligation." ChemBioChem 9: 1052-1056.

Ueki, M. and M. Amemiya (1987). "Removal of 9-Fluorenylmethyloxycarbonyl(Fmoc) Group with Tetrabutylammonium Fluoride." Tetrahedron Letters 28: 6617-6620.

Ueki, M., N. Nishigaki, et al. (1993). "One-pot Deprotection and Coupling of Peptides by the Fmoc Strategy." Chemistry Letters 22: 721-724.

Xu, M.-Q. and T. C. Evans (2001). "Intein-Mediated Ligation and Cyclization of Expressed Proteins." Methods 24(3): 257-277.

Yan, L. Z. and P. E. Dawson (2001). "Synthesis of Peptides and Proteins without Cysteine Residues by Native Chemical Ligation Combined with Desulfurization." J. Am. Chem. Soc. 123: 526-533.

Chapter 3: Cysteine-based Site-specific Installation of N^{ε}-acetyl-lysine Analogues into Synthetic Peptides and Recombinant Proteins by Thiol-ene Coupling

3.1 Introduction

Post-translational modification (PTM) is an important mechanism for the regulation of protein function. One such PTM with increasingly recognized significance is protein lysine acetylation (Norris, Lee et al. 2009). Like phosphoylation on Tyr, Ser or Thr, acetylation on Lys is also a reversible biochemical process. Whereas a lysine acetyltransferase adds an acetyl group onto the ε-amine of lysine residue, a deacetylase acts in the opposite way to remove it (Shahbazian and Grunstein 2007) (Fig. 4).

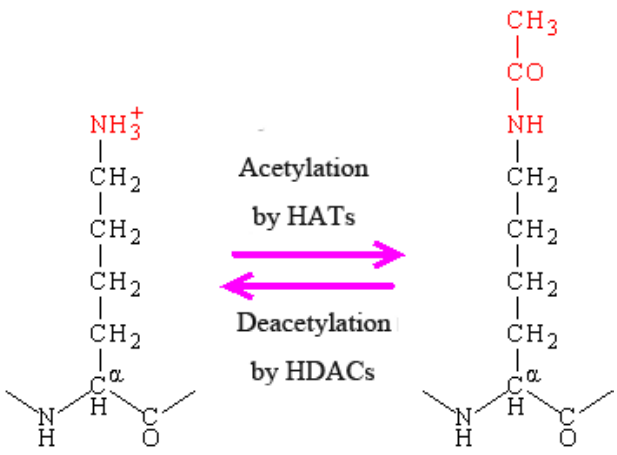


Figure 4. *Acetylation and deacetylation of the lysine on histones by enzymes.*

Initially discovered on histones (Allfrey, Faulkner et al. 1964), lysine acetylation has

also been observed in thousands of eukaryotic proteins (Kim, Sprung et al. 2006; Choudhary, Kumar et al. 2009; Wang, Zhang et al. 2010), pointing to its broad signaling function beyond the nuclear boundary (Spange, Wagner et al. 2009; Guan and Xiong 2011). There is mounting evidence that aberrant lysine acetylation is involved in numerous disease conditions, including cancer and neurological disorders. Therefore, modulating protein lysine acetylation and deacetylation reactions in the cell has significant therapeutic implications and enzymes responsible for these processes can be important drug targets (Hake, Xiao et al. 2004; Yang 2004). For instance, inhibitors of histone deacetylase (or lysine deacetylase) have shown potent antitumor activity with little toxicity *in vivo* in animal models (Minucci and Pelicci 2006; Marks and Breslow 2007). Clearly the study of lysine acetylation is of great importance and will lead to more therapeutic breakthroughs.

3.1.1 Methods for the preparation of N^{ε}- lysine acetylated proteins

Although lysine acetylation has long been recognized as an epigentic modification in histones, the exact effects of most individual protein acetylation events, especially those identified more recently, remain to be elucidated (Shahbazian and Grunstein 2007). A major difficulty in the study of lysine acetylation biology lies with the limited availability of homogeneous protein samples containing the acetylated lysine residue(s) of interest. Such materials would be invaluable reagents for discerning the

structural and functional effects of a particular Lys acetylation via biophysical and biochemical means (Shogren-Knaak, Ishii et al. 2006).

3.1.1.1 Stop codon suppression strategy

Several methods can be used to prepare site-specifically acetylated proteins, such as unnatural amino-acid mutagenesis using the amber stop codon/suppressor tRNA system (Neumann, Peak-Chew et al. 2008). In this system, an N^{ε}-acetyl-lysine is site-specifically incorporatd into recombinant proteins by an orthogonal N^{ε}-acetyl-lysyl-tRNA synthetase/tRNA-CUA pair. While the stop codon suppression strategy is a powerful method, it is currently not widely available.

Another method combines unnatural amino acid mutagenesis with chemical modification to introduce an acetyl-lysine analog into a protein. The phenylselenocysteine (PhSeCys) is site-specifically incorporatd into recombinant proteins. Followed by oxidative elimination and Michael addition of corresponding thiols, the metyl- or acetyl-lysine analogues could be synthesized (Fig. 5). However the chiral integrity of the modified amino acid is compromised in the process (Guo, Wang et al. 2008).

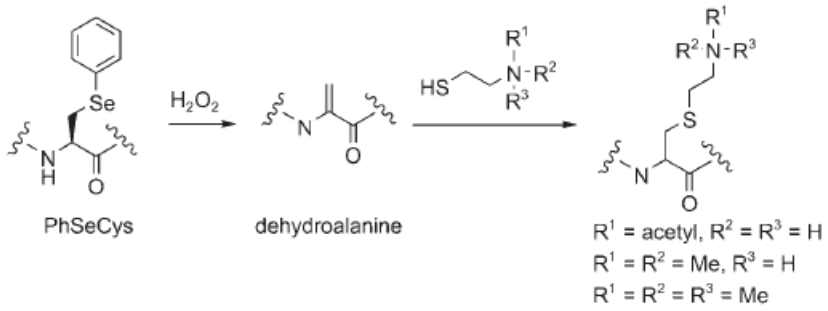


Figure 5. *Synthesis of metyl- or acetyl-lysine analogues by combination of unnatural amino acid mutagenesis and chemical modification (adopted from* Guo, Wang et al. 2008*).*

3.1.1.2 Cysteine based site-specific installation of modifications

An alternative strategy to prepare acetylated proteins is direct modication of natural proteins. Obviously, direct enzymatic Lys acetylation using lysine acetyl transferases is unrealistic given the often promiscuous as well as inefficient and incomplete nature of such enzymatic reactions. Chemical acetylation of a selected lysine among many Lys residues in a protein is also obviously not feasible. The unique reactivity of the thiol group of cysteine as a soft nucleophile has been exploited extensively for selective protein modification (Chalker, Bernardes et al. 2009). As mentioned in Chapter 1, the classic reaction of aminoethylation of Cys has long been used for converting a cysteine residue to 4-thialysine as a functional equivalent to lysine (Raftery and Cole 1966). This method was also extended to preparing N-methyl-lysine analogs when using N-methylaminoethyl halides as the alkylating agents. Unfortunately, this reaction system does not work for N-acetyl-thialysine by using N-acetyl-aziridine and N-acetyl-aminoethyl bromide or iodide for cysteine

alkylation (Huang, Holbert et al. 2010). An alternative way is using methylthiocarbonyl-arizidine (MTCA) to selectively alkylate Cys. The resultant methylthiocarbonyl-thiaLys (MTCTK) analog is shown to be a mimic of N^{ε}-acetyl-lysine in certain funcations, although the methylthio-carbamate moiety is electro-sterically rather different from the acetamide in N^{ε}-acetyl-lysine (Fig. 6).

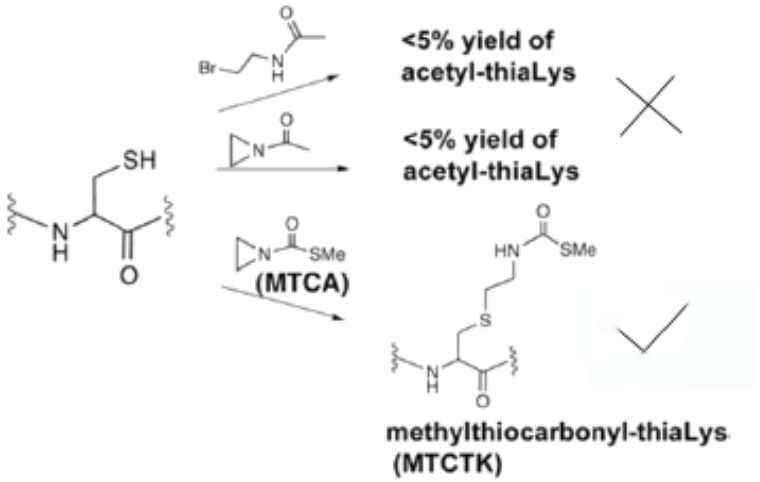


Figure 6. *Synthesis of the acetylated lysine mimic by methylthio-carbonyl-aziridine (MTCA) (adopted from* Huang, Holbert et al. 2010*).*

Clearly, similar to 4-thialysine and N-methyl-thialysine as ideal mimics of lysine and N-methyl-lysine, respectively, N-acetyl-thialysine [sLys(Ac)] would also be a perfect mimic of Lys(Ac) in its exhibited physicochemical and biochemical properties. However, since existing methods for cysteine modification are not applicable here, a new method must be discovered to obtain such a Lys(Ac) mimic.

3.1.2 Our strategy: site-specific installation of N^{ε}-acetyl-lysine analogs by thiol-ene coupling

In searching for ways of introducing an sLys(Ac) residue into proteins, we came across a radical reaction known as thiol–ene addition which might serve our needs (Griesbaum 1970). It is a classic reaction discovered over a century ago which yields an anti-Markovnikov addition thioether product. At the first step of the reaction, the free thiol is activated to generate a thiyl radical under the action of UV irradiation and /or a radical initiator. The thiyl radical then adds to a double bond to generate a carbon radical which can react with another thiol molecule in a radical transfer step. This radical transfer step is known as the rate-limiting step in the radical chain reaction. (Fig. 7) (Dondoni 2008; Hoyle and Bowman 2010).

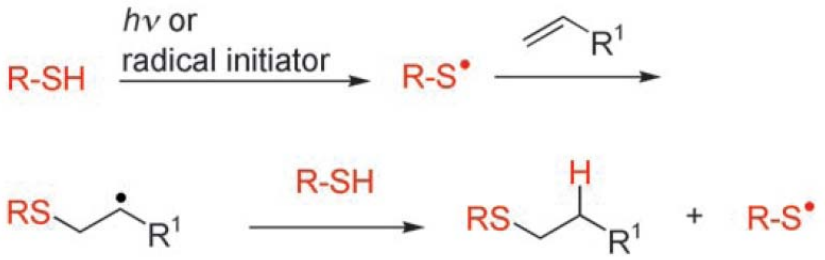


Figure 7. *The thiol-ene coupling reaction (adopted from* Dondoni 2008*).*

Over the years this reaction has found extensive use in polymer chemistry. More recently, it has also emerged as a useful click reaction for bioorganic functionalization (Floyd, Vijayakrishnan et al. 2009; Hoyle and Bowman 2010). We realized that thiol–ene coupling between the cysteine thiol and N-vinylacetamide (NVA) would directly generate the desired acetyl-thialysine (Fig. 8). In my study, this reaction was

found to be extremely specific and efficient. It could be finished in 1 or 2 h to give a yield of 95% or higher. Moreover, several assays were utilized to validate the function of the N^{ε}-acetyl-lysine mimic.

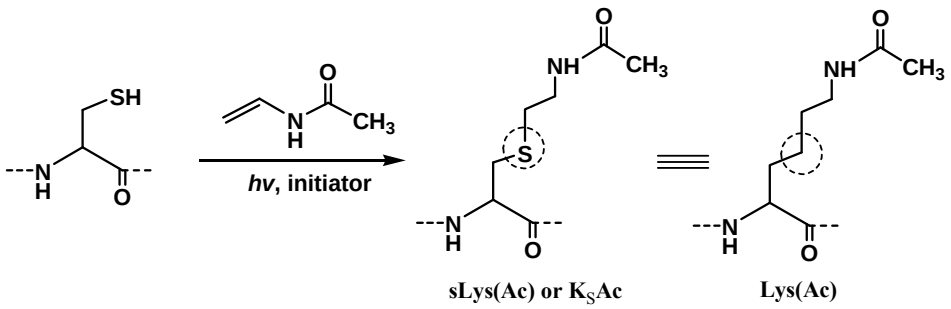


Figure 8. *Alkylation of Cys by N-vinylacetamide (NVA) to generate sLys(Ac), a close analog of the native Lys(Ac).*

Furthermore, several other reagents - vinyl acetate, N-allylacetamide and N-methyl-N-vinylacetamide, were also tested in the reaction with cysteine to generate other N^{ε}-acetyl-lysine analogues by the thiol-ene coupling method (Fig. 9). Although these analogues are structurally similar, they may differ from one another in their functions and be useful for better understanding the effects of protein modifications by the acetyl group.

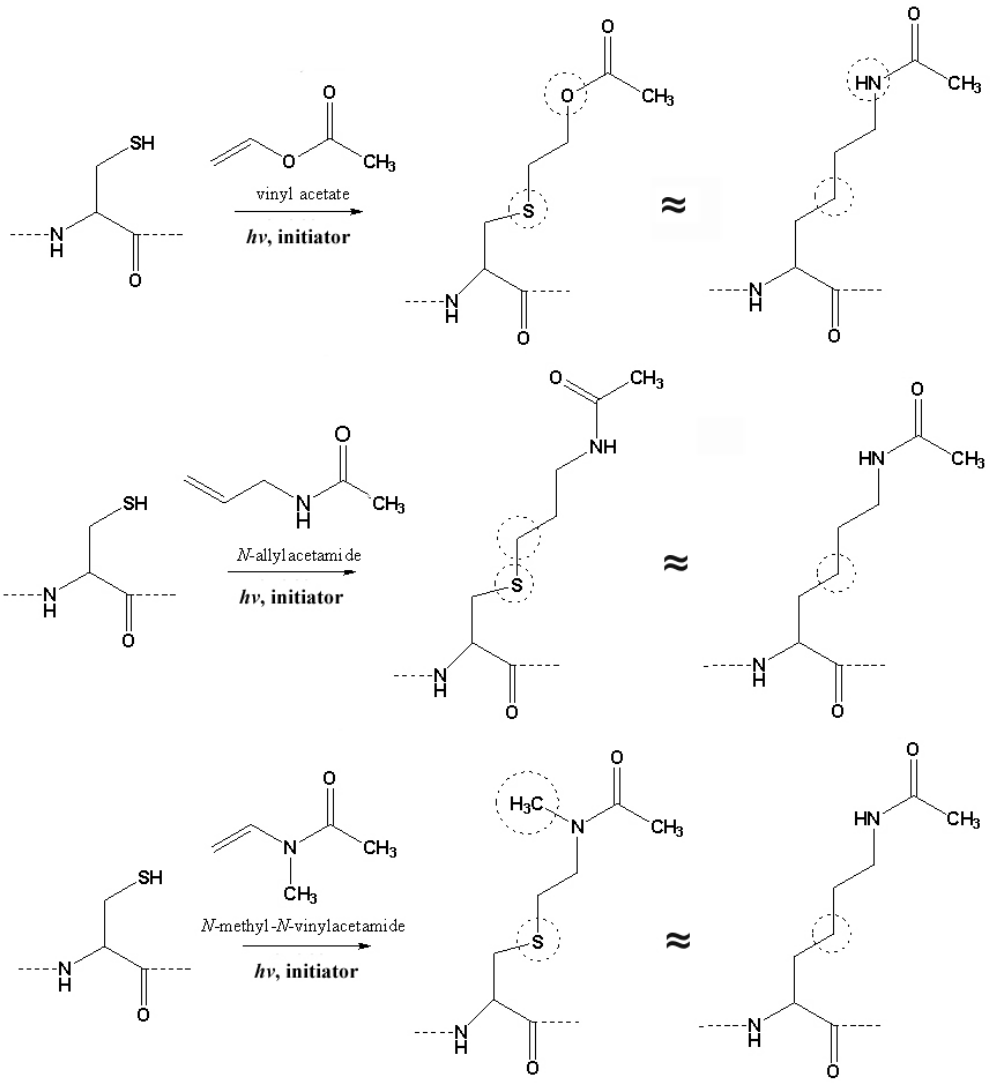


Figure 9. *Alkylation of Cys by vinyl acetate, N-allylacetamide, and N-methyl-N-vinylacetamide to generate different analogs of the native Lys(Ac).*

3.2 Results

3.2.1 Model study of the thiol-ene coupling reaction between NVA and benzyl mercaptan

At first, we used a small organic thiol compound, benzyl mercaptan (BzSH), as the substrate and examined the alkylation reaction under different conditions (Fig. 10). We found that the free radical reaction proceeded well in acetate buffer at pH 4 and in the presence of VA-044 as the initiator under UV irradiation at 365 nm. At a 1:1 ratio of NVA to BzSH and at low concentrations of the two reactants (at 5 or 10 mM), a 30-min reaction gave over 70% conversion of BzSH with the expected thiol acetamidoethylation product. The reaction at 10 mM also produced a minor side product which was due to telomerization, i.e., polymerization (in this case dimerization only) of NVA on the thiol. When NVA was used in 5-, 10- or even 20-fold excess to BzSH, the reaction was complete within 30 min but more telomerization products were formed. However, even under these conditions, the desired product still formed in predominant amount.

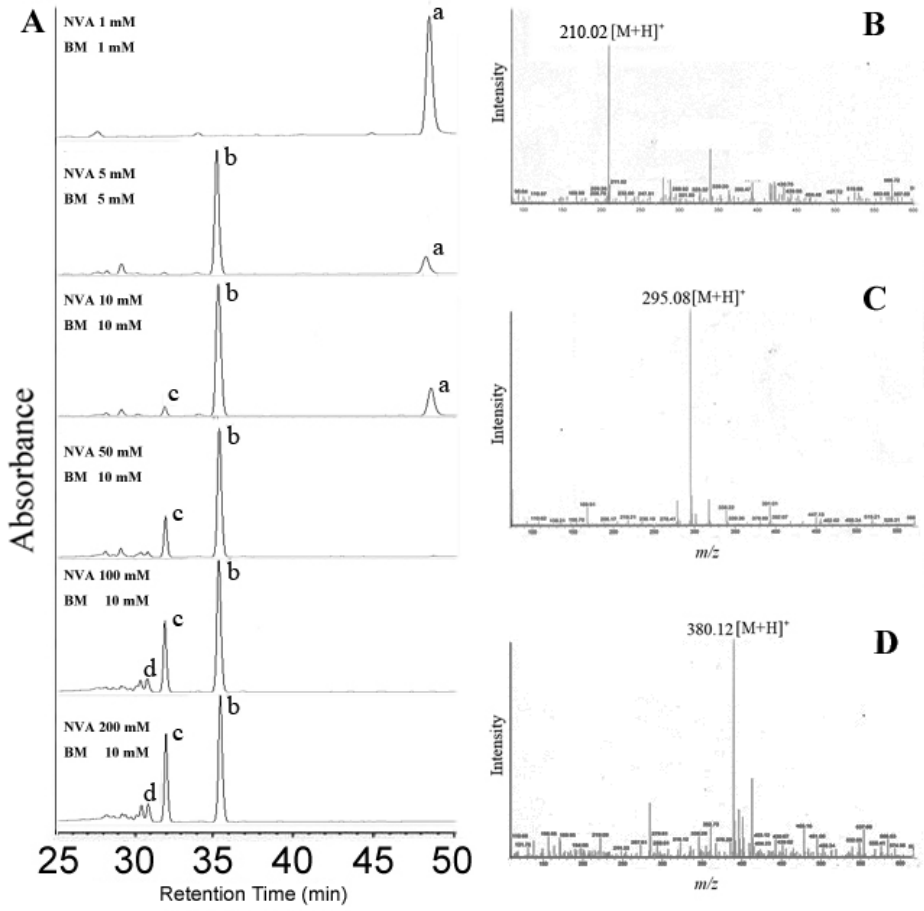


Figure 10. *C18 analytical HPLC profile of the thiol-ene coupling reaction between benzyl mercaptan and NVA at different ratio and concentration. (A) C18 analytical HPLC profile of these reactions. Peak a: Benzyl mercaptan; Peak b: thiol acetamidoethylation product (^{1}H NMR ($CDCl_3$): δ 7.37-7.30 (m, 5H), 6.58 (br, 1H), 5.64 (br, 1H), 4.12-4.11 (m, 1H), 3.77-3.65 (m, 2H), 2.77-2.70 (m, 1H), 2.67-2.56 (m, 1H), 1.96 (d, 3H)); Peak b: product (^{1}H NMR ($CDCl_3$): δ 7.34-7.27 (m, 5H), 5.74 (br, 1H), 3.73 (s, 2H), 3.42-3.37 (m, 2H), 2.61-2.58 (m, 2H), 1.97 (s, 3H)); Peak c: side product 1, NVA dimer was added to BM; Peak d: side product 2, NVA trimer was added to BM. (B) ESI-MS of peak b ($[M+H]^+$ found: 210.02, MW calcd: 209.09). (C) ESI-MS of peak c ($[M+H]^+$ found: 295.08, MW calcd: 294.14). (D) ESI-MS of peak d ($[M+H]^+$ found: 380.12, MW calcd: 379.19). HPLC condition: 0% to 50% of buffer B in buffer A in 50 min.*

A look into the mechanism of the classic radical thiol–ene addition reaction helps to explain the formation of the side products. The carbon radical formed at step 2 is usually much more likely to react with a thiol molecule in the critical rate-limiting step of radical chain transfer (step 3) which generates another thiyl radical. However, in the presence of a large excess of the ene compound (NVA), the carbon radical can also sometimes react with another NVA molecule in a phenomenon called telomerization, resulting in the formation of oligomerized NVA on the thiol (Fig. 11).

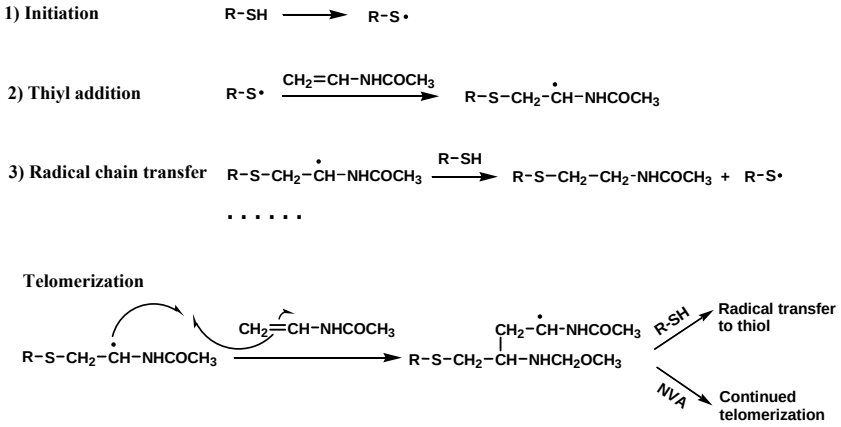


Figure 11. *Mechanism of formation of side products.*

3.2.2 Reaction of model peptides with NVA for the introduction of acetyl-lysine analog sLys (Ac)

The encouraging results obtained with benzyl mercaptan led us to further test this reaction on synthetic peptides. The sequences of the Cys-containing peptides were designed as following:

Peptide **1**: Ac-Phe-Gln-Pro-Lys-Cys-Gly-NH_2

Peptide **2**: H-Val-Gly-Cys-Ala-Glu-Lys-Ser-Leu-NH_2

Peptide **3**: H-Trp-Ala-Cys-Tyr-Lys-Ser-Leu-NH_2

Peptide **4**: Biotin-Gly-Lys-Gly-Gly-Ala-Cys-Arg-His-Arg-Lys-Val-Leu-Arg-Asp-Asn-NH_2

Peptide **5**: Biotin-Gly-Cys-Gly-Gly-Cys-Gly-Leu-Gly-Cys-Gly-Gly-Ala-Cys-Arg-NH_2

We reasoned that the inclusion of a second thiol compound in the reaction mixture might help to suppress the side reaction of radical telomerization. After numerous trials we found that glutathione (reduced) could perfectly serve this purpose and that a reaction mixture of 15 mM glutathione, 50 mM NVA and 5 mM VA-044 in 0.2 M acetate buffer (pH 4) or 0.2 M phosphate buffer (pH 6 or 7) was well suited for the alkylation of peptide and protein substrates. Using these conditions, the reaction on the synthetic peptides was basically completed in 30 min and no significant side reactions were found to occur when extending the reaction time to 1 h. For example, when 5 mM of peptide **1** was treated with this mixture in acetate buffer (pH 4) or phosphate buffer (pH 6), the desired alkylated product was obtained in near quantitative yield based on HPLC and MS analysis (Fig. 12). Under these conditions, no or little telomerization products were detected. Therefore, glutathione participated, together with the peptide substrate, in the critical radical chain transfer step – which is known as the rate-limiting step of this thiol–ene coupling reaction – to effectively quench the carbon free radical intermediate formed at the addition step of the thiyl radical to the NVA ethylene double bond and prevent it from reacting with another molecule of NVA. In fact, when peptide **4** was alkylated deliberately in the absence of

glutathione under otherwise identical conditions, products in which the Cys residue was alkylated by di-, tri- and tetrameric NVAs were observed in addition to the desired product. If the reducing agent TCEP was used to replace glutathione in the reaction, it led to desulfurization of the peptide substrate (Fig. 15). Remarkably, when 1.25 mM of this peptide **5** which contains 4 Cys residues was treated with the reaction mixture for 1 h, a very clean tetra-alkylating reaction gave the desired product in 95% yield (Fig. 16).

In general, the reactions were very clean as seen from HPLC data (Fig. 12, 13, 14, 15 and 16). The side reactions were very insignificant in nature. For example, formation of a minute amount (<3%) of disulfide cross-linked side product between the peptide substrate **1** and glutathione was detected from the reaction at pH 6 (Fig. 12-2). This seems due to the presence in the glutathione sample of a small amount of oxidized glutathione, which at higher pH can undergo disulfide exchange with the Cys thiol in the peptide substrate. At lower pH (e.g., pH 4), such an exchange reaction is inhibited. Therefore it is advised to use a new and pure glutathione (reduced) sample and perform degassing of the buffer used to prevent oxidation of the glutathione.

When peptides **3** and **5** were reacted at pH 4, a very small amount of a side product (< 5%) was detected; it was more hydrophilic with a MW that was 16 Da higher than that of the expected product. It appeared to result from oxidation of the thioether linkage to sulfoxide (Figs. 14 and 16).

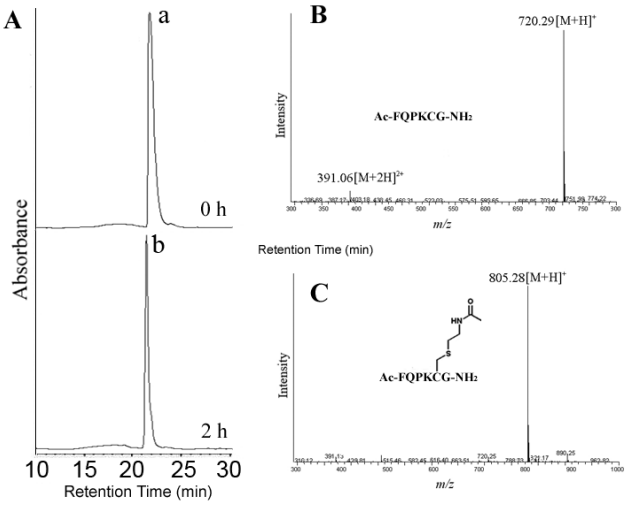


Figure 12-1. *Thiol-ene coupling reaction between peptide* ***1*** *and NVA at pH 4.0. (A) C18 analytical HPLC profile of the reaction. Peak a: peptide* ***1****; Peak b: product. (B) ESI-MS of peak a ($[M+H]^+$ found: 720.29, MW calcd: 719.34). (C) ESI-MS of peak b ($[M+H]^+$ found: 805.28, MW calcd: 804.39). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

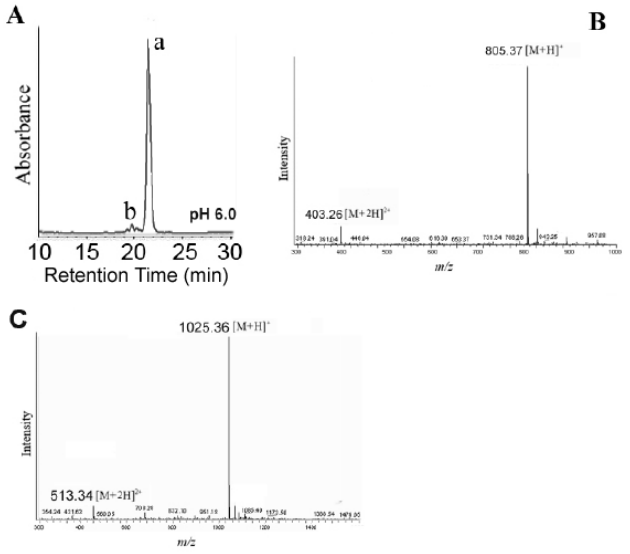


Figure 12-2. *Thiol-ene coupling reaction between peptide* ***1*** *and NVA at pH 6.0. (A) C18 analytical HPLC profile of the reaction. Peak a: product; Peak b: disulfide-linked side product between peptide* ***1*** *and glutathione. (B) ESI-MS of peak a ([M+H]+ found: 805.37, MW calcd: 804.39). (C) ESI-MS of peak b ([M+H]+ found: 1025.36, MW calcd: 1024.66). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

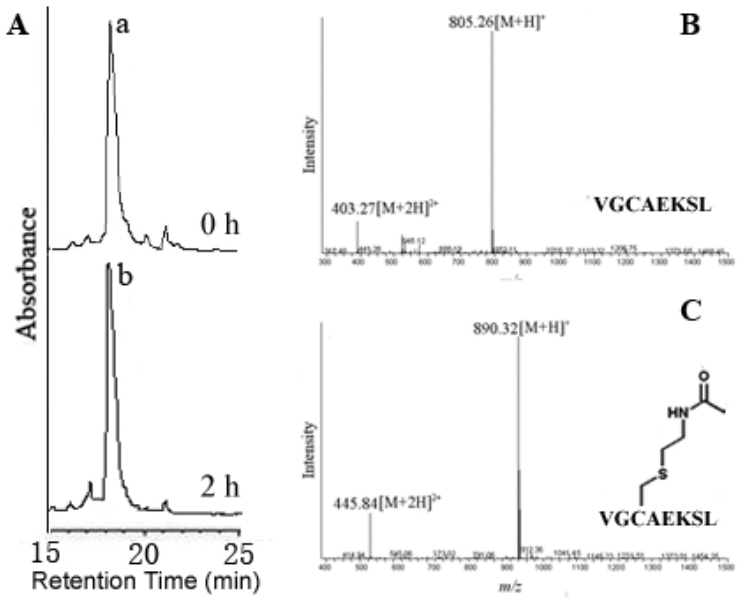


Figure 13. *Thiol-ene coupling reaction between peptide* ***2*** *and NVA. (A) C18 analytical HPLC profile of the reaction. Peak a: peptide* ***2****; Peak b: product. (B) ESI-MS of peak a ($[M+H]^+$ found: 805.26, MW calcd: 804.42). (C) ESI-MS of peak b ($[M+H]^+$ found: 890.32, MW calcd: 889.47). HPLC condition: 0 % to 30 % of buffer B in buffer A over 30 min.*

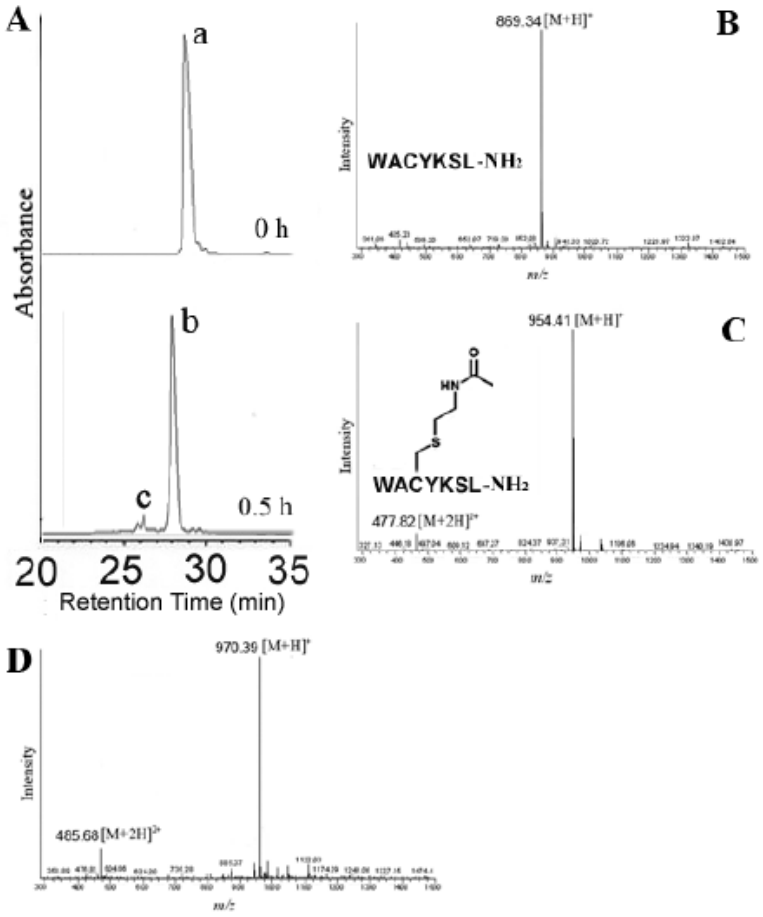


Figure 14. *Thiol-ene coupling reaction between peptide* ***3*** *and NVA at pH 4.0. (A) C18 analytical HPLC profile of the reaction. Peak a: peptide* ***3****; Peack b: expected product; Peak c:* —S— *oxidation product* —S(=O)—*. (B) ESI-MS of peak a ($[M+H]^+$ found: 869.34, MW calcd: 869.04). (C) ESI-MS of the peak b ($[M+H]^+$ found: 954.41, MW calcd: 954.09). (D) ESI-MS of the peak c ($[M+H]^+$ found: 970.39, MW calcd: 970.08). HPLC condition: 0 % to 30 % of buffer B in buffer A over 30 min.*

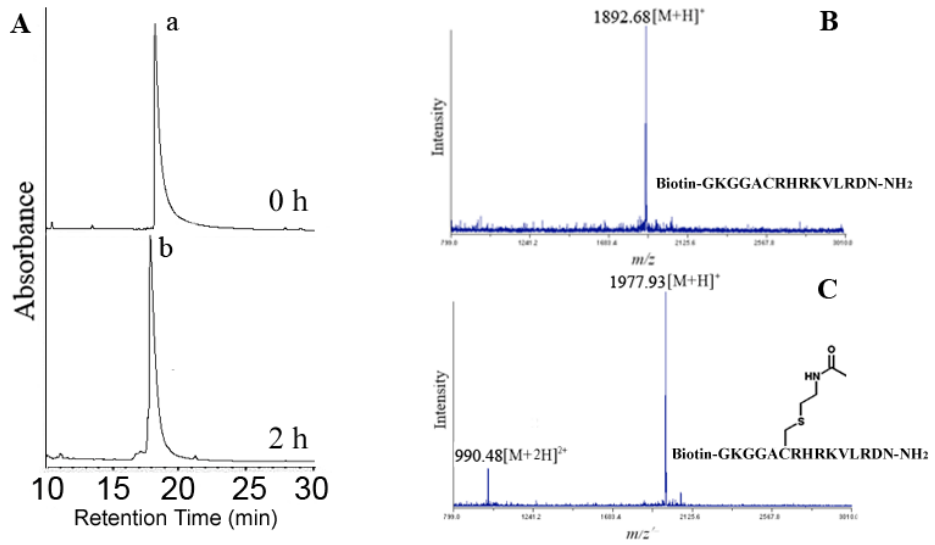


Figure 15-1. *Thiol-ene coupling reaction between peptide* ***4*** *and NVA. (A) C18 analytical HPLC profile of the reaction. Peak a: peptide* ***4****; Peak b: product. (B) MALDI-TOF MS of peak a ([M+H]*$^{+}$ *found: 1892.68, MW calcd: 1890.99). (C) MALDI-TOF MS of peak b ([M+H]*$^{+}$ *found: 1977.93, MW calcd: 1976.04). HPLC condition: 0 % to 30 % of buffer B in buffer A over 30 min.*

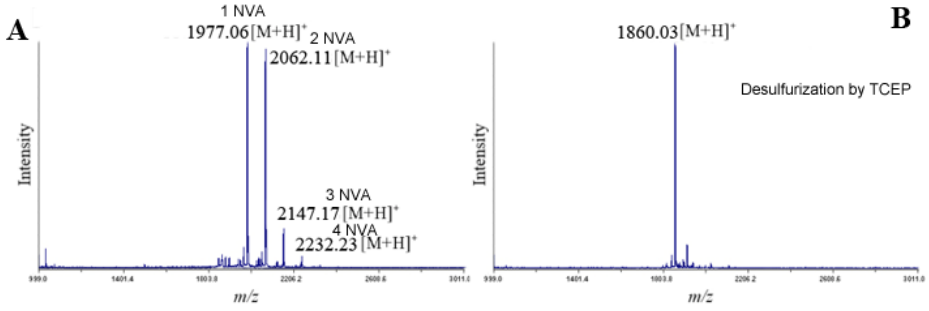


Figure 15-2. *Thiol-ene coupling reaction between peptide* ***4*** *and NVA under non-standard conditions. (A) MALDI-TOF MS of the products when no glutathione was added (other conditions were the same). (B) MALDI-TOF MS of the product if 15 mM TCEP was added to replace glutathione and the desulfurization reaction took place to give a product with a -32 MW.*

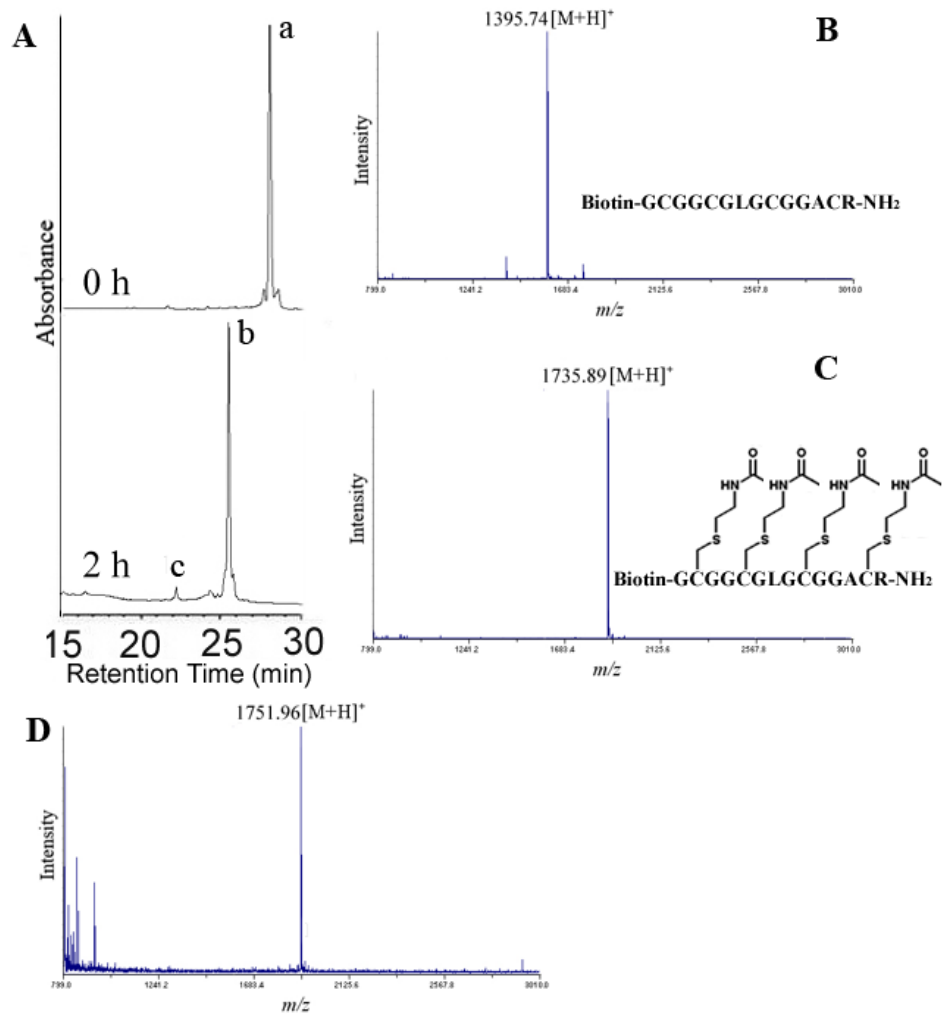


Figure 16. *Thiol-ene coupling reaction between peptide* **5** *and NVA. (A) C18 analytical HPLC profile of the reaction. Peak a: peptide* **5***; Peack b: expected product; Peak c:* —S— *oxidation product* —S(=O)—*. (B) MALDI-TOF MS of peak a ([M+H]$^+$ found: 1395.74, MW calcd: 1394.52). (C) MALDI-TOF MS of the peak b ([M+H]$^+$ found: 1735.89, MW calcd: 1734.72). (D) MALDI-TOF MS of the peak c ([M+H]$^+$ found: 1751.96, MW calcd: 1750.71). HPLC condition: 0 % to 30 % of buffer B in buffer A over 30 min.*

3.2.3 Introduction of sLys(Ac) into ubiquitin: Preparation of Ub K_S48Ac

This thiol-ene coupling reaction was also highly effective on protein substrates. First, a ubiquitin mutant containing a Cys at position 48 was prepared and subjected to this modification. The protein (0.5 mM) was used in its native folded state for alkylation in the same reaction mixture at pH 4 or 7. MS analysis clearly showed an almost quantitative conversion, in 2 h, of the Cys residue to sLys(Ac) with the expected +85 Da MW for the alkylated product (Fig.17). The protein remained soluble during the reaction, suggesting that no denaturing occurred and that the presence of 50 mM NVA and 5 mM VA-044 did not affect the structure of the folded protein. It is worth pointing out that it would be difficult to use a semisynthetic method to prepare such a modified protein since the modification site is found in the middle of the sequence.

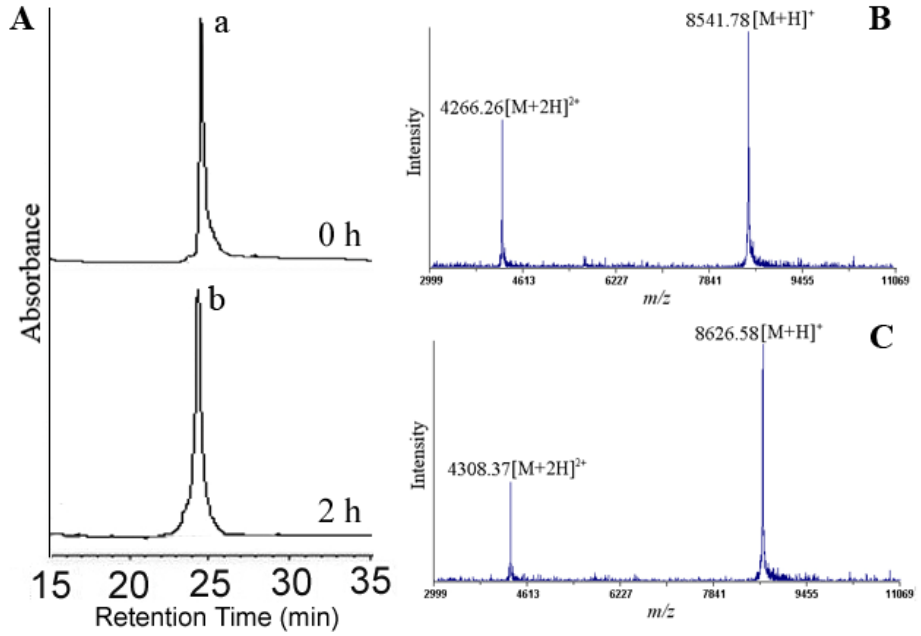


Figure 17-1. *Thiol-ene coupling reaction between ubiquitin K48C and NVA at pH4.0. (A) C18 analytical HPLC profile of the reaction. Peak a: ubiquitin K48C; Peak b: product. (B) MALDI-TOF MS of peak a ($[M+H]^+$ found: 8541.78, MW calcd: 8539.88). (C) MALDI-TOF MS of peak b ($[M+H]^+$ found: 8626.58, MW calcd: 8624.93). HPLC condition: 0 % to 30 % in 15 min, then to 50 % in 20 min of buffer B in buffer A.*

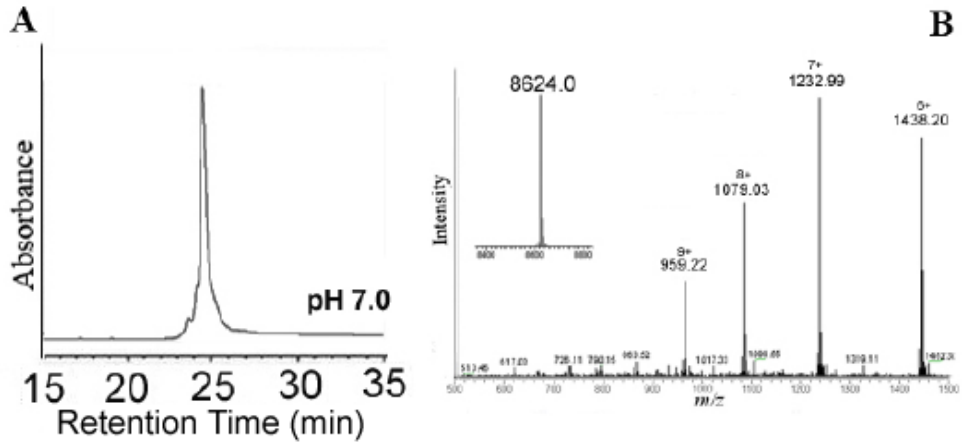


Figure 17-2. *Thiol-ene coupling reaction between ubiquitin K48C and NVA at pH 7.0. (A) C18 analytical HPLC profile of the reaction. (B) The raw and deconvoluted mass of product determined by ESI-MS. (MW found: 8624.0, MW calcd: 8624.93). HPLC condition: 0 % to 30 % in 15 min, then to 50 % in 20 min of buffer B in buffer A.*

3.2.4 Introduction of sLys(Ac) into histone H4 and H3: Preparation of H4 K_S16Ac and H3 K_S27Ac

Two other proteins, histone H4 K16C and H3 K27C, were also modified with excellent results. In these cases, 6 M Gdn·HCl was included in the alkylation reaction mixture. For the alkylation of H4 K16C at pH 4 or 7, the oxidation product was formed in significant amount (Fig. 18). When dimethylsulfide was added in the reaction mixture, this side product was minimized to about 5%.

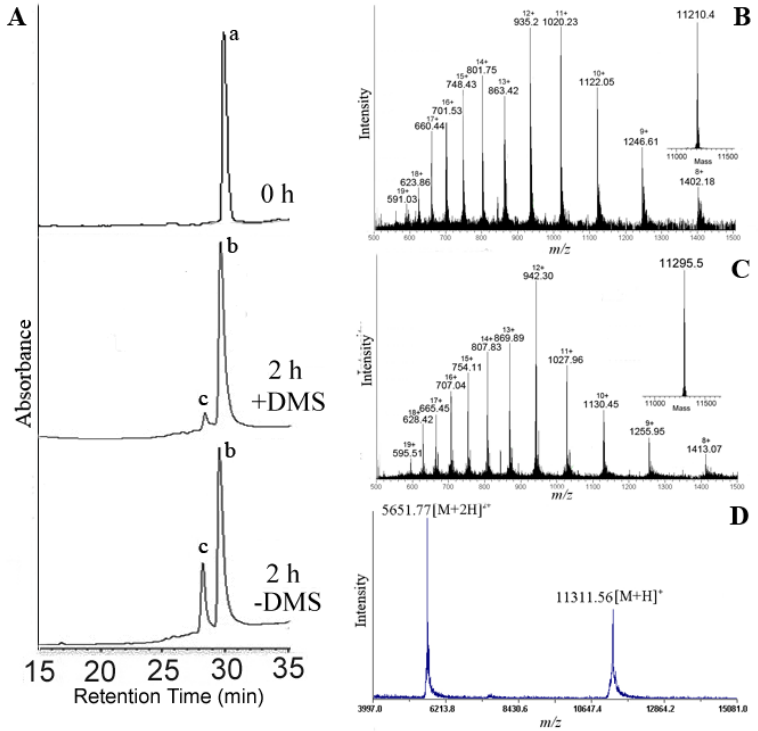


Figure 18-1. *Thiol-ene coupling reaction between H4 K16C and NVA at pH 4.0. (A) C4 semi-prep HPLC profile of the reaction with or without dimethyl sulfide. Peak a: H4 K16C; Peak b: expected product; Peak c: oxidation product. (B) The raw and deconvoluted mass of the starting material H4 K16C in peak a determined by ESI-MS. (MW found: 11210.4, MW calcd: 11211.16). (C) The raw and deconvoluted mass of peak b determined by ESI-MS. (MW found: 11295.5, MW calcd: 11296.21). (D) MALDI-TOF MS of peak c ($[M+H]^+$ found: 11311.56, MW calcd: 11312.2). HPLC condition: 0 % to 40 % in 20 min, then to 60 % in 20 min of buffer B in buffer A.*

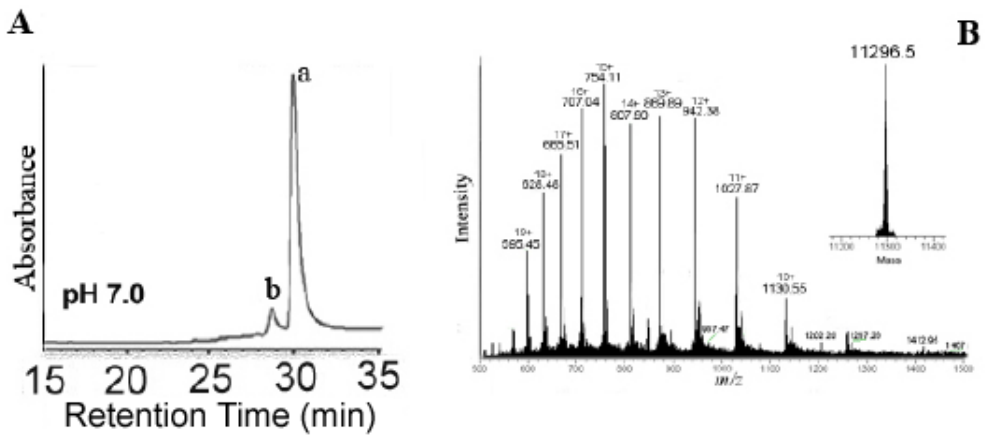


Figure 18-2. *Thiol-ene coupling reaction between H4 K16C and NVA at pH 7.0. (A) C4 semi-prep HPLC profile of the reaction. Peak a: expected product; Peak b: oxidation product. (B) The raw and deconvoluted mass of peak a determined by ESI-MS. (MW found: 11296.5, MW calcd: 11296.21). HPLC condition: 0 % to 40 % in 20 min, then to 60 % in 20 min of buffer B in buffer A.*

The reaction on histone H3 K27C was shown in Fig. 19 and 6 M Gdn·HCl was also included in the alkylation reaction mixture. ESI-MS analysis clearly showed an almost quantitative conversion (>90%), in 2 h, of the Cys residue to sLys(Ac) with the expected +85 Da MW for the alkylated product (Fig. 19).

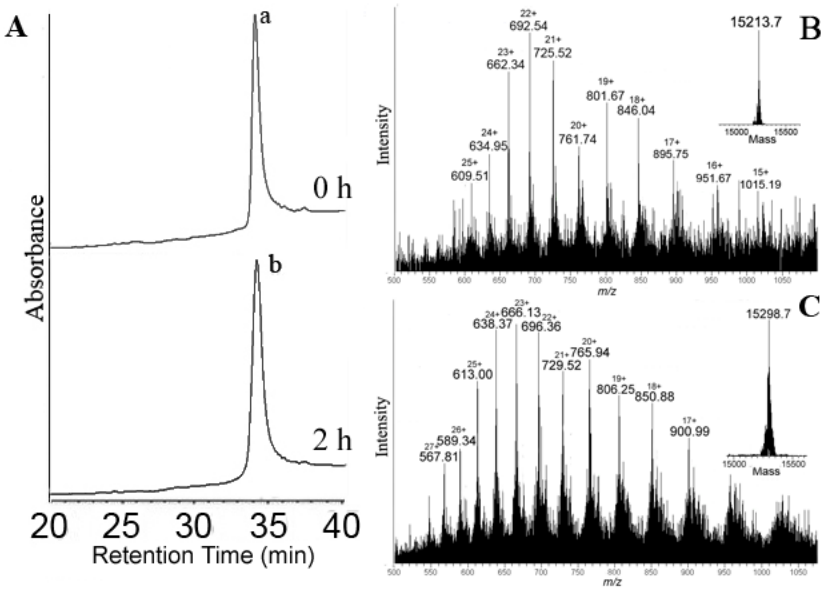


Figure 19. *Thiol-ene coupling reaction between H3 K27C and NVA. (A) C4 semi-prep HPLC profile of the reaction. Peak a: H3 K27C; Peak b: product. (B) The raw and deconvoluted mass of peak a by ESI-MS. (MW found: 15213.7, MW calcd: 15213.92). (C) The raw and deconvoluted mass of peak b determined by ESI-MS. (MW found: 15298.7, MW calcd: 15298.97). HPLC condition: 0 % to 40 % in 20 min, then to 60 % in 20 min of buffer B in buffer A.*

3.2.5 Functional characterization of N^{ε}-acetyl-lysine analog

The generated sLys(Ac) was shown to be a good functional mimic of the natural Lys(Ac). First, histone protein H4 K_S16Ac was recognized by a specific anti-H4 K16Ac antibody whereas the unmodified H4 K16C was not recognized by the same antibody (Fig. 20).

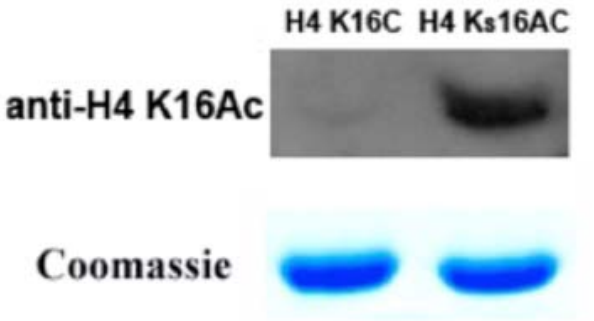


Figure 20. *Western blot with the specific Histone H4 AcK16 antibody on H4 AcK16 mimic.*

Next, an enzymatic test was conducted to investigate whether the acetyl-lysine mimic could be recognized by a histone deacetylase and used as substrate for deacetylation. SIRT2, a class III NAD-dependent deacetylase, was used for the deacetylation reaction of the alkylated peptide **1** and its native counterpart Ac-FQPKK(Ac)G (Sakaguchi, Herrera et al. 1998; Liu, Scolnick et al. 1999; Barlev, Liu et al. 2001). Using an HPLC assay (Fig. 21), we showed that the sLys(Ac) residue in the alkylated peptide **1** was susceptible to enzymatic deacetylation, albeit to a lesser degree compared to its native counterpart, which might be due to the fact that the side chain of sLys is slightly longer than that of native Lys.

The deactylases are not very efficient on deacetylation of synthetic peptide substrates and require relatively large amount of the enzyme and long reaction time for the deacetylation reaction. The results showed that the deacetylation rate of N-acetyl-thialysine peptide was about 1/3 of that on the control peptide containing the native N-acetyl-lysine (Fig. 22). In contrast, the methylthiocarbamate modification

reported earlier was not recognized as substrate by the deacetylase (Huang, Holbert et al. 2010).

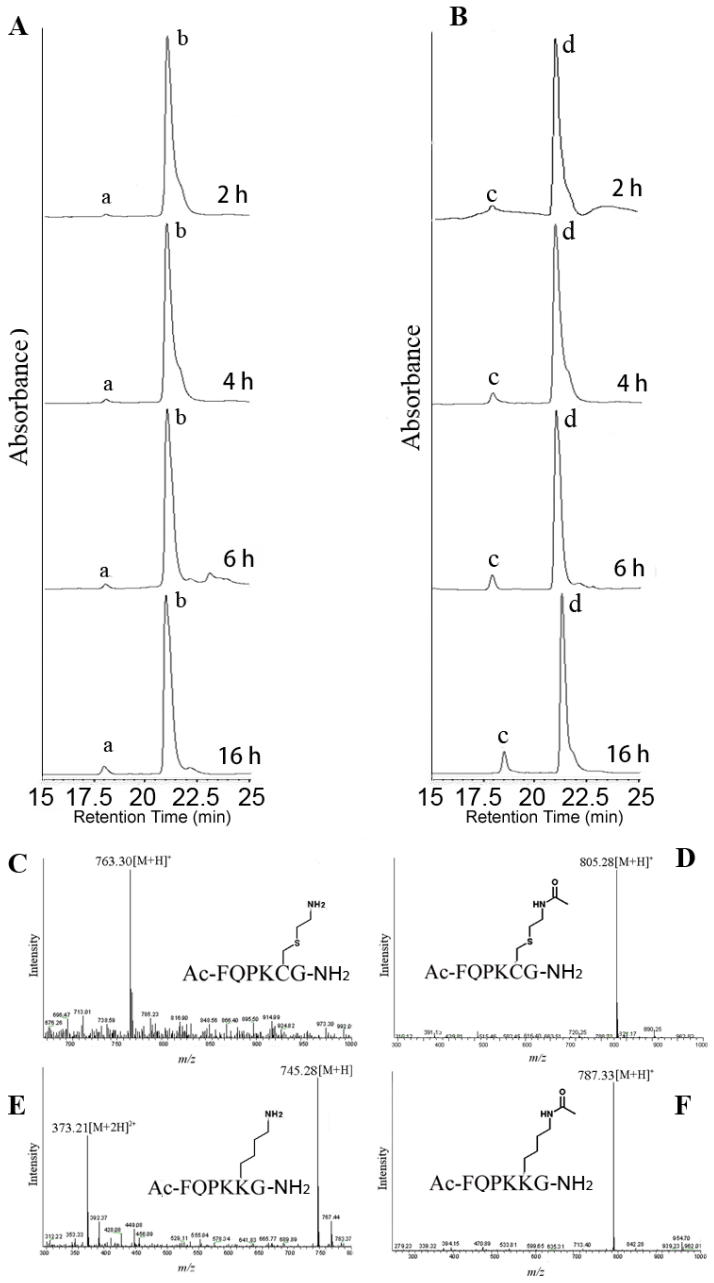


Figure 21. *Results of SIRT2 mediated deacetylation assay. (A) C18 analytical HPLC profile of peptide 1 with the installed acetyl-lysine analog from SIRT2 assay. Peak a: deacetylated product; Peak b: Ac-FQPKK$_S$(Ac)G-NH$_2$. (B) C18 analytical HPLC profile of the control peptide from SIRT2 assay. Peak c: deacetylated product; Peak d:*

control peptide. (C) ESI-MS of peak a ([M+H]$^+$ found: 763.30, MW calcd: 762.38). (D) ESI-MS of peak b ([M+H]$^+$ found: 805.28, MW calcd: 804.39). (E) ESI-MS of peak c ([M+H]$^+$ found: 745.28, MW calcd: 744.43) (F) ESI-MS of peak d ([M+H]$^+$ found: 787.33, MW calcd: 786.44). HPLC condition: 0 % to 30 % in 30 min of buffer B in buffer A.

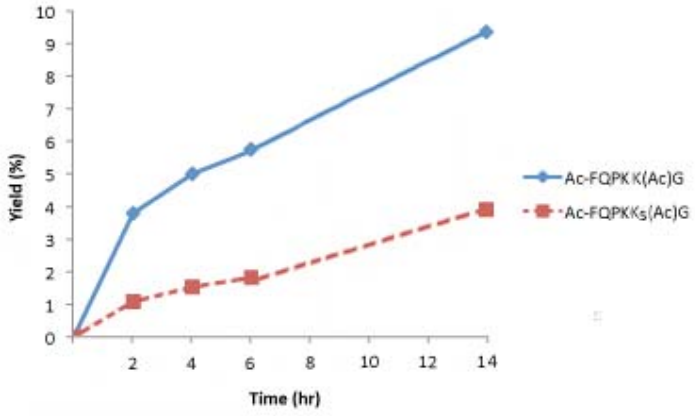


Figure 22. *Time course of the deacetylated product formation.*

3.2.6 Mg^{2+}-induced compaction of nucleosome arrays

Acetylation of Lys16 in histone H4 is known to inhibit the folding of nulceosome arrays and hence the formation of the compact 30-nm chromatin fiber (Shogren-Knaak, Ishii et al. 2006). H4 K_S16Ac and three other control H4 proteins (H4 K16Ac, H4 K_S16 and H4-WT) were incorporated respectively, together with H3, H2A and H2B, into histone octamers. While H4 K16Ac is the native control of H4 K_S16Ac, H4 K_S16 is used to see whether S-aminoethyl-cysteine or sLys (K_S) is functionally equivalent to the native Lys residue at position 16 of H4 in the chromatin compaction assay. H4 K16Ac was prepared using a semi-synthetic approach and H4

K_S16 was synthesized by alkylating H4 C16 with 2-bromoethylamine (Fig. 23-1) and the result of ESI-MS of the aminoethylation product is shown in Fig. 23-2. The four different octamers were then individually combined with the 12-177-601 DNA to assemble into the 12-nucleosome sting nucleosomal arrays.

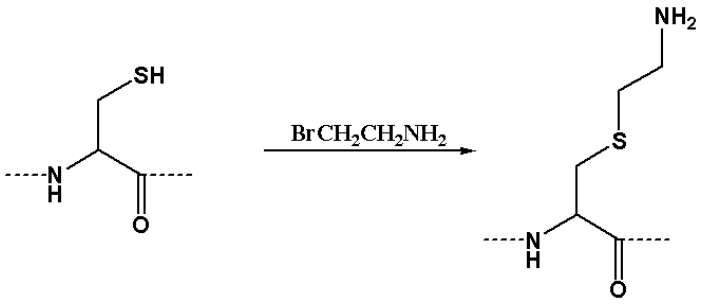


Figure 23-1. *Synthesis of H4 K_S16 by aminoethylation of H4 C16*

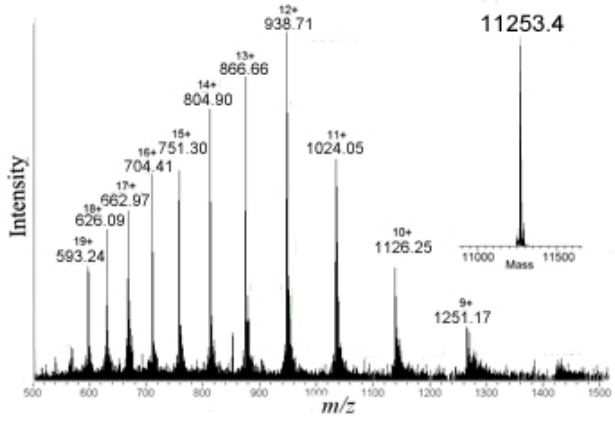


Figure 23-2. *The raw spectrum and deconvoluted mass of the alkylated product H4 KS16 as determined by ESI-MS (MW found: 11253.4, MW calcd: 11254.2).*

The array solutions with or without Mg^{2+} were loaded to a Beckman analytical ultracentrifuge and the AUC measurements of sedimentation coefficient ($S_{20,W}$) were analyzed by the van Holde-Weischet method (van Holde and Weischet 1978). The results clearly showed that, in the absence of Mg^{2+}, the four arrays showed similar distributions of sedimentation coefficient ($S_{20,\,w}$ = 34-36S) (Fig. 24, open symbols).

When wild-type H4 or its equivalent H4 K_S16 arrays were incubated in a buffer containing 1.0 mM $MgCl_2$, they folded into a significantly more compact state and the sedimentation coefficient distributions were shifted to 52-53S. These results are consistent with a previous study (Dorigo, Schalch et al. 2003; Shogren-Knaak, Ishii et al. 2006). On the other hand, when treated with 1.0 mM $MgCl_2$, the K16Ac and K_S16Ac arrays folded into a less compact state as compared with the WT array and the sedimentation coefficient was in the range of 44-45S (Fig. 24). By this method, we demonstrated that the K_S16Ac of interest produced an identical effect as the native K16Ac in abolishing Mg^{2+}-induced folding of the reconstituted nucleosome array. Remarkably, these results prove not only the functional equivalency between sLys(Ac) and Lys(Ac) but also that between sLys and Lys. (Note: The part on nucleosome array preparation and AUC experiments was done by Dr. Abdollah Allahverdi)

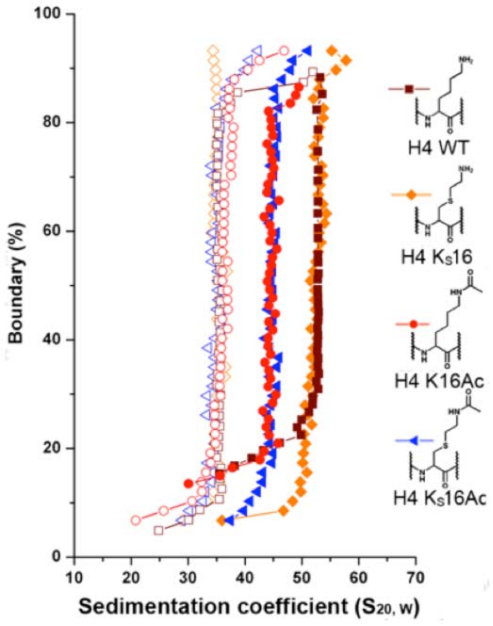


Figure 24. *Effects of H4 K16 acetylation on nucleosome array folding as seen from sedimentation distributions of the nucleosome arrays before (no Mg^{2+}, open symbols) and after Mg^{2+}-induced folding (with 1.0 mM $MgCl_2$, solid symbols).*

3.2.7 Site-specific installation of other acetyl-lysine analogues by thiol-ene coupling

In a similar way, peptide **1** was treated with vinyl acetate, N-allylacetamide, and N-methyl-N-vinylacetamide respectively to generate different acetyl-lysine analogues (Fig. 25, 26 and 27). Different from the reaction with NVA, no telomerization side products were detected.

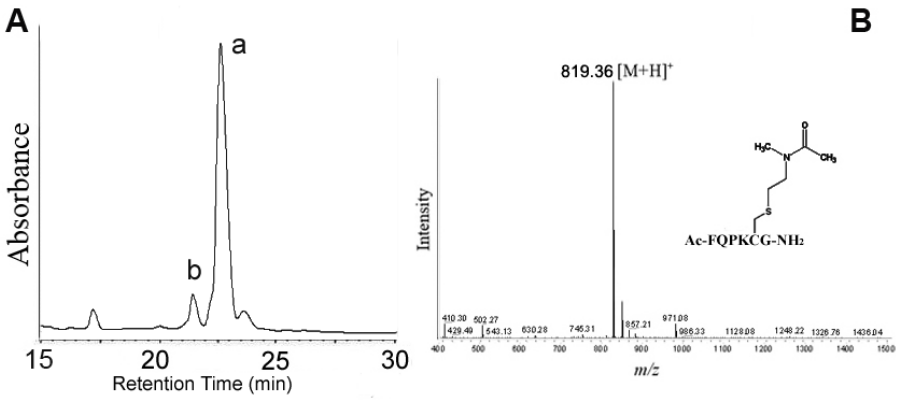


Figure 25. *Thiol-ene coupling reaction between peptide 1 and N-methyl-N-vinylacetamide. (A) C18 analytical HPLC profile of the reaction. Peak a: product; Peak b: peptide 1. (B) ESI-MS of peak a ($[M+H]^+$ found: 819.36, MW calcd: 818.40). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

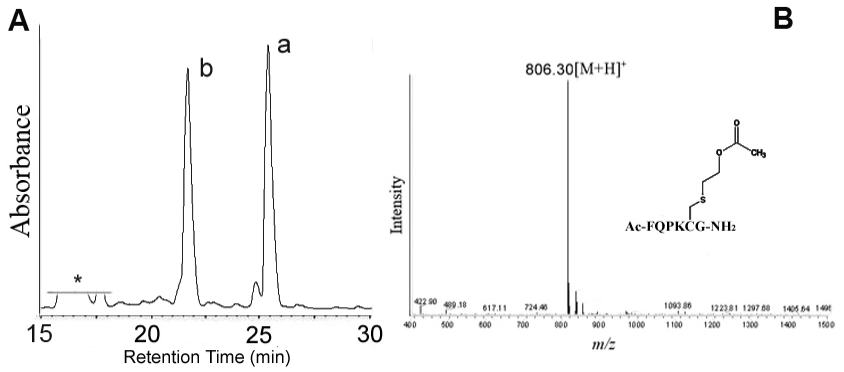


Figure 26. *Thiol-ene coupling reaction between peptide 1 and Vinyl acetate. (A) C18 analytical HPLC profile of the reaction. Peak a: product; Peak b: peptide 1; Peak *: unidentified peak. (B) ESI-MS of peak a ($[M+H]^+$ found: 806.30, MW calcd: 805.37). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

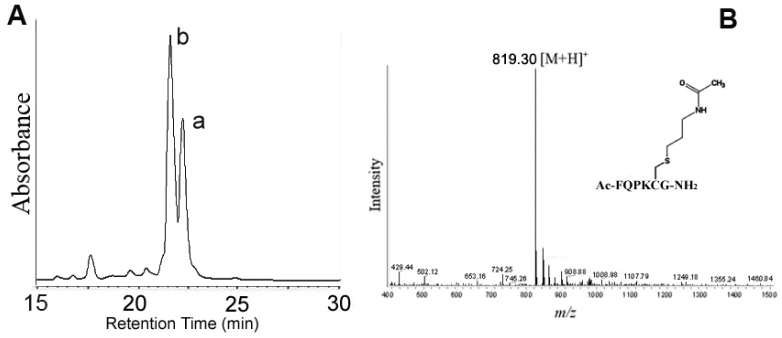


Figure 27. *Thiol-ene coupling reaction between peptide 1 and N-allylacetamide. (A) C18 analytical HPLC profile of the reaction. Peak a: product; Peak b: peptide 1. (B) ESI-MS of peak a ($[M+H]^+$ found: 819.30, MW calcd: 818.40). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

Although all the reagents contain double bonds, the reaction activity would be different due to their different structures. The NVA was the most active reagent and gave over 95 % conversion of the starting material for 1 h. When peptide **1** was treated with N-methyl-N-vinylacetamide, the yield was also quite good (78.7 %). For the other two reagents vinyl acetate and N-allyacetamide, the reaction furnished yields of 50.9 % and 38.3 %, respectively. As the free radical intermediates of the two reagents are less stable than that of NVA (Table 1), the yields are lower.

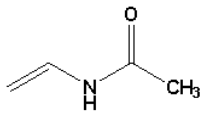
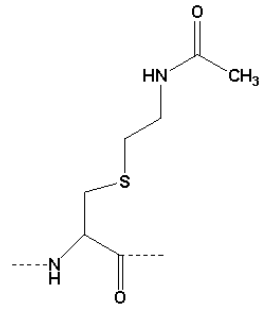
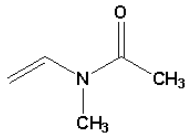
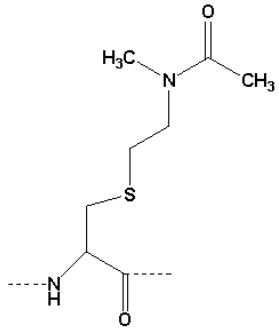
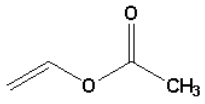
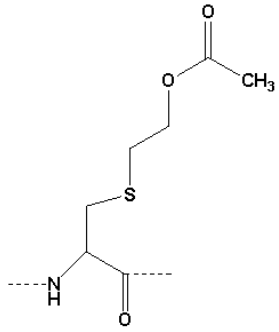
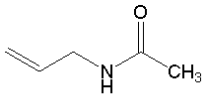
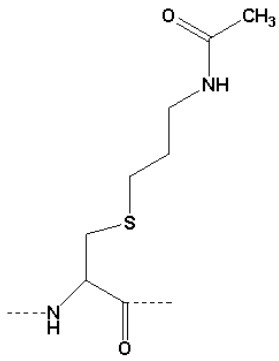
Thio-ene coupling reaction			
Peptide 1	Reagent	Product	Yield (%)
Ac-FQPKCG	 *N*-vinylacetamide		>95
Ac-FQPKCG	 *N*-methyl-*N*-vinylacetamide		78.7
Ac-FQPKCG	 vinyl acetate		50.9
Ac-FQPKCG	 *N*-allylacetamide		38.3

Table 1. *Yields of thio-ene coupling reaction between peptide 1 and different reagents for 1 h.*

3.3 Discussion and conclusion

In this research work, the thiol-ene coupling reaction was utilized to synthesize N-acetyl-thialysine which contains the desired acetylation group. The thiol-ene coupling reaction performed very well in model peptides. The inclusion of glutathione could successfully suppress the side reactions of radical telomerization and give an excellent yield in 1 h. The ubiquitin K48C was used in its native form and the reaction also gave a near quantitative conversion within 2 h. The H4 K_S16Ac and H3 K_S27Ac were also prepared by this method. Although a small amount of oxidated side product was found when H4 K_S16C was used as the substrate, it could be minimized by addition of dimethyl sulfide. The N-acetyl-thialysine could be recognized by the specific histone H4 K16Ac antibody in western-blot. This result indicates the N-acetyl-thialysine is structurally similar to the native acetyl-lysine. In the enzymatic assay, the acetyl group of N-acetyl-thialysine was successfully removed by the class III NAD-dependent histone deacetylase SIRT2. As the side chain of sLys is a little longer than that of native Lys, the efficiency of deacetylation from sLys was slightly slower than that from the native one. Moreover, in the analytical ultracentrifugation assay, the nucleosomal array containing the N-acetyl-thialysine shows the same behavior in Mg^{2+}-induced compaction assay as the one containing native K16(Ac) in H4. These results show that the sLys or the sLys(Ac) is functional equivalent to native Lys or the Lys(Ac) in the AUC assay.

Three other reagents, vinyl acetate, N-allylacetamide and N-methyl-N-vinylacetamide were also used to react with peptide **1** to generate different acetyl-lysine analogues. These analogues are structurally similar, however they are different one from another in the steric and electronic properties which would also translate into functional differences. My future work will continue to analyze the function and application of these analogues in histone proteins.

The N-acetyl-thialysine differs with natural acetyl-lysine only isosterically at the γ-position of the amino acid structure. Because of this, it is functionally equivalent or similar to the latter. The reaction system of thiol-ene coupling is robust and gives near quantitative yields of site-specifically acetylated proteins which can be purified in a simple chromatography or dialysis step. The ease of implementation of this method also makes it easily adoptable by researchers from the bioscience research community. As such, this radical reaction approach provides a convenient enabling tool for the study of lysine acetylation biology and will help to advance research in this important field.

3.4 Materials and methods

3.4.1 Model study of the thiol-ene coupling reaction between NVA and benzyl mercaptan

N-Vinyl acetamide (NVA) and benzyl mercaptan (BM) were mixed in the 0.2 M acetate buffer (pH 4.0, prepared from acetic acid and sodium acetate) at different ratios and concentrations, followed by addition of 5 mM of the initiator VA-044 in a 0.5 ml THIN WALL, CLEAR tube (Axygen). The tube containing the reaction mixture was placed in a Cole Parmer 9818-series darkroom UV light box at about 10 cm under the lamp (365 nm) at room temperature for 30 min. The reaction was monitored by C18 analytical HPLC and the product was purified by semi-prep HPLC. After lyophilization, the products were dissolved in $CDCl_3$ for NMR analysis.

3.4.2 Reaction of model peptides with NVA for the introduction of acetyl-lysine analog sLys (Ac)

The model peptides were synthesized as C-terminal carboxyamides using standard Fmoc solid phase peptide synthesis techniques. Rink-amide MBHA resin was utilized for the synthesis. Fmoc protection group was removed by treatment with 20% piperidine in DMF twice (2 min × 1, 20 min × 1). All the amino acids were used in 4 eq. of the resin, and the coupling was performed using 4 eq. of PyBOP and 8 eq. of DIEA for 2 h. For peptide 3 and 4, D-biotin was used at the last coupling step. After

sequence assembly, the final deprotection and cleavage was performed by using a cocktail of TFA : H2O : triisopropylsilane : 2-mercaptoethanol (94 : 2.5 : 2.5 : 1) for 3 h at room temperature. The peptide was then precipitated with diethyl ether and lyophilized. The crude peptides were purified by C18 semi-prep HPLC.

For the site-specific installation of the acetyl-lysine analog in the model peptides using, the thiol-ene coupling reaction, all the reactants were dissolved in the 0.2 M acetate buffer (pH 4.0) or phosphate buffer (pH 6.0) as indicated above and the final concentrations were as following:

Peptide: 5 mM (except for peptide 4, 1.25 mM).

NVA (vinyl acetate, N-allylacetamide, or N-methyl-N-vinylacetamide): 50 mM.

VA-044: 5 mM.

Glutathione (reduced form): 15 mM.

The reaction tube was irradiated under 365 nm UV for 1 h. The reaction products were analyzed by C18 analytical HPLC and confirmed by ESI or MALDI-TOF MS. MS analysis was done on either desalted samples (using a C18 zip-tip) or on HPLC-purified fractions.

3.4.3 Introduction of sLys(Ac) into ubiquitin: Preparation of Ub K_S48Ac

As there is no cysteine in the ubiquitin protein, the gene code “AAA” for lysine 48 on

the ubiquitin gene was changed into "TGC" for cysteine using the QuikChangeTM site-directed mutagenesis kit (Stratagene) with the primers: Ubi K48C F: 5'-GTCTGATATTTGCCGGCTGTCAGCTGGAGGATGGCCG-3' and Ubi K48C R: 5'-CGGCCATCCTCCAGCTGACAGCCGGCAAATATCAGAC -3'. The plasmid containing ubiquitin K48C gene was transformed into BL21 (DE3) cells. The cells were grown in 1 L LB media containing ampicilin to OD_{600} of 0.6 and induced by a final concentration of 0.5 mM IPTG for 4 h at 37 ^{o}C. After centrifugation at 6000 rpm for 10 min at 4 ^{o}C, cells were resuspended in 50 ml lysis buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, pH 7.5) and broken by microfluidization (Microfluidics, Newton, USA). After centrifugation at 25,000 *g* for 30 min at 4 ^{o}C, 70 % perchloric acid was added to the supernatant at a ratio of 350 μl to 50 ml lysis buffer. The mixture was stirred for 10 min and centrifuged at 25,000 *g* for 30 min at 4 ^{o}C. The supernatant was filtered using 0.2 μm filters and dialyzed with 3.5 KDa cut-off dialysis tubing against 50 mM ammonium acetate buffer (pH 4.5 with 1 mM DTT). After filtration, the protein was purified with a HiTrap^{TM} SP FF 5 ml FPLC column (GE Healthcare Life Sciences). It was eluted with a linear gradient from 0% to 100% FPLC buffer B (50 mM ammonium acetate, pH 4.5, 1 mM DTT, 0.5 M NaCl) in buffer A (50 mM ammonium acetate, pH 4.5, 1 mM DTT) in 120 min at a flow rate of 0.5 ml / min. Usually the ubiquitin would be eluted at around 240 mM NaCl. After FPLC purification, the protein was dialyzed to ddH_2O followed by lyophilization.

The freeze-dried ubiquitin K48C was dissolved in the 0.2 M acetate buffer (pH 4.0).

The final concentrations of the reactants were as following:

Ubiquitin K48C: 0.5 mM.

NVA: 50 mM.

VA-044: 5 mM.

Glutathione: 15 mM.

The reaction tube was irradiated under 365 nm UV for 2 h. The reaction product was analyzed by C18 analytical HPLC and confirmed by MALDI-TOF MS. The protein remained soluble during the reaction period, indicating it stayed in its folded state.

3.4.4 Introduction of sLys(Ac) into histone H4 and H3: Preparation of H4 K_S16Ac and H3 K_S27Ac

K16C point mutation was introduced to H4 gene using the same mutagenesis kit with the primers: H4 K16C F: 5'-GGTAAAGGTGGTGCTTGCCGTCACCGTAAAGTT C-3' and H4 K16C R: 5'-GAACTTTACGGTGACGGCAAGCACCACCTTTACC-3'. The plasmid containing the H4 K16C gene was transformed into BL21 (DE3) pLysS cells. The cells were grown in 1 L LB media containing ampicillin and cholramphinicol to OD_{600} of 0.6 and induced by 0.4 mM IPTG for 3 h at 37 ^{o}C . After centrifugation at 6000 rpm for 10 min, cells were resuspended in 50 ml wash buffer (50 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, pH 7.5) and broken by microfluidization. The cell debris were removed by centrifugation at 20,000 *g* for 30 min at 4 ^{o}C. The pellet was washed twice with wash buffer containing 1% Triton

X-100 and once without Triton X-100. Then 1 ml of DMSO was added to the pellet and the pellet was stirred for 30 min. After that 10 ml unfolding buffer (6 M guanidinium HCl, 10 mM Tris-HCl, 10 mM DTT, pH 7.5) was added. After centrifugation, the supernatant was loaded on a 26/60 Sephacryl S-200 column and purified with gel filtration buffer (7 M de-ionized urea, 20 mM sodium acetate, 1 M sodium chloride, 5 mM beta-mercaptoethanol, 0.5 M EDTA, pH 5.2). After FPLC purification, the protein was purified again by C4 semi-prep HPLC followed by lyophilization.

The reaction was performed in the 0.2 M acetate buffer (pH 4.0) containing 6 M guanidinium HCl. Dimethyl sulfide was added to minimize oxidation of the thioester linkage in this product. The final concentrations of the reactants were as follows:

H4 K16C: 1 mM.

NVA: 50 mM.

VA-044: 5 mM.

Glutathione: 15 mM.

Dimethyl sulfide: 100 mM.

The reaction tube was irradiated under 365 nm UV for 2 h. The reaction products were analyzed by C4 semi-prep HPLC and by ESI or MALDI-TOF MS.

The C110A point mutation was introduced into the H3 gene using the mutagenesis kit with the primers:

H3 C110A F: 5’-GAGGACACCAACCTGGCCGCCATCCACGCCAAG-3’

H3 C110A R: 5’- CTTGGCGTGGATGGCGGCCAGGTTGGTGTCCTC-3’.

The K27C point mutation was introduced to the H3 C110A gene using the mutagenesis kit with the primers: H3 K27C F: 5’-AAGGCAGCCAGGTGCTCCGCTCCTGCTACC -3’ and H3 K27C R: 5’-AGCAGGAGCGGAGCACCTGGCTGCCTTGGTG-3’.

The expression of H3 K27C protein was the same as in the case of H4 K16C.

The reaction was performed in the 0.2 M acetate buffer (pH 4.0) containing 6 M guanidinium HCl. The final concentrations of the reactants were as follows:

H3 K27C: 1 mM.

NVA: 50 mM.

VA-044: 5 mM.

Glutathione: 15 mM.

The reaction tube was irradiated under 365 nm for 2 h UV. The product was purifid by C4 semi-prep HPLC and confirmed by ESI-MS.

3.4.5 Functional characterization of the N^{ε}-acetyl-lysine analog

3.4.5.1 Western-blot

3 μg of the control protein (H4 K16C) and H4 K_S16Ac were dissolved respectively in loading buffer and run on 15 % SDS-PAGE, then transferred onto a polyvinylidene

difluoride membrane. After blocked with 20 ml of TBST (50 mM Tris.HCl, 150 mM NaCl, 0.1% Tween 20, pH 7.4) containing 5% non-fat milk for 1 h, the membrane was incubated in 20 ml TBST with 5% non-fat milk containing histone H4 K16Ac antibody (1 : 1000 dilution) overnight at 4 ^{o}C. After washing with TBST 4 times, the membrane was incubated in 20 ml TBST with 5% non-fat milk containing the anti-rabbit IgG peroxidase conjugate (1 : 5000 dilution) for 1 h at room temperature. After washing with TBST 4 times, proteins were visualized by chemiluminescence.

3.4.5.2 SIRT2 HPLC assay

Peptide **1** with the acetyl-lysine analog, Ac-FQPKK$_S$(Ac)G-NH_2, and the native peptide substrate, Ac-FQPKK(Ac)G-NH_2, were used in this assay. The sequence of peptide **1** was based on residues 317-320 of p53 (Gln-Pro-Lys-Lys(Ac)) which is the best deacetylase substrate for the SIRT2 enzyme and the N-terminus was capped by an acetyl group to increase the hydrophobicity. 0.5 mM peptide and 0.1 mM NAD^+ were mixed in SIRT2 assay buffer (50 mM Tris, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$, 1 mg / ml BSA) and equilibrated at 37 ^{o}C for 10 min. The reaction was initiated by adding 0.1 mM SIRT2 enzyme (0.1 U / μl) at 37 ^{o}C and monitored by C18 analytical HPLC .

3.4.6 Nucleosomal Array Reconstitutions

The plasmid containing the 12-177-601 DNA was a gift from Prof. Richmond, ETH, and was transformed and amplified in HB101 cells. The plasmid was extracted as described previously. RNA and protein impurities were removed by gel filtration on a Sepharose 6 column with the use of TES2000 buffer. After excision with EcoRV, the 12-177-601 DNA was separated from short plasmid fragment by polyethylene glycol (PEG 6000). Finally 12-177-601 DNA was purified on a Sephacryl SF1000 column with TES100 buffer.

Wild type *Xenopus laevis* histones H2A, H2B, H3, and H4 were individually over-expressed in BL21 (DE3) pLysS cells in presence of ampicilin and cholramphinicol. Each histone was purified by gel filtration on a Sephacryl S-200 column and subsequently on a Resource S cation exchange column. H4 K_S16 was obtained by S-alkylation of H4 K16C with 2-bromoethylamine, and H4 K16Ac was prepared by chemical semi-synthesis. The histone octamer was formed using a molar ratio of 1:1:1.2:1.2 for H2A, H2B, H3, and H4. The histone octamer was purified on a Sephacryl S-200 gel filtration column.

The nucleosome array was reconstituted by step-wise dialysis using the histone octamer and 12-177-601 DNA as described previously. To prevent excessive binding of the histone to the DNA, 0.5 molecule of competitor Core Length DNA (150-bp) per one array was added to the reconstitution mixture. The exact amount of histone octamer to 12-177-601 DNA to get the stoichiometry of 12:1 was determined

empirically in small scale preparations using different ratios of histone ocatmer to the DNA template. The reconstituted arrays were purified as described previously. Purified array material and also the respective digests by ScaI were checked on 5% poly acrylamide gel electrophoresis (PAGE) to verify the quality of the nucleosomal arrays.

3.4.7 Mg^{2+}-induced compaction of nucleosome arrays

The sedimentation velocity experiment was carried out on a Beckman XL-I analytical ultracentrifuge equipped with an AN-50T rotor and monochrome scanner. The stock solution of array was diluted in TEK buffer to get $A^{259} = 0.8\ cm^{-1}$ (DNA concentration C_p =121 μM). The sample and reference (TEK buffer + salt at the same concentration as in the sample) were loaded into the 12 mm double channel cells and equilibrated under vacuum for 30 min at 3000rpm and 20 ^oC. Data measurement and analyses were carried out following earlier reported methods.

References

Allfrey, V. G., R. Faulkner, et al. (1964). "Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis." Proc Natl Acad Sci U S A 51: 786-794.

Barlev, N. A., L. Liu, et al. (2001). "Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases." Mol Cell 8(6): 1243-1254.

Chalker, J. M., G. J. Bernardes, et al. (2009). "Chemical modification of proteins at cysteine: opportunities in chemistry and biology." Chem Asian J 4(5): 630-640.

Choudhary, C., C. Kumar, et al. (2009). "Lysine acetylation targets protein complexes and co-regulates major cellular functions." Science 325(5942): 834-840.

Dondoni, A. (2008). "The emergence of thiol-ene coupling as a click process for materials and bioorganic chemistry." Angew Chem Int Ed Engl 47(47): 8995-8997.

Dorigo, B., T. Schalch, et al. (2003). "Chromatin fiber folding: requirement for the histone H4 N-terminal tail." J Mol Biol 327(1): 85-96.

Floyd, N., B. Vijayakrishnan, et al. (2009). "Thiyl glycosylation of olefinic proteins: S-linked glycoconjugate synthesis." Angew Chem Int Ed Engl 48(42): 7798-7802.

Griesbaum, K. (1970). "Problems and Possibilities of the Free-Radical Addition of Thiols to Unsaturated Compounds." Angew. Chem. Int. Ed. 9: 273-287.

Guan, K. L. and Y. Xiong (2011). "Regulation of intermediary metabolism by protein acetylation." Trends Biochem Sci 36(2): 108-116.

Guo, J., J. Wang, et al. (2008). "Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins." Angew Chem Int Ed Engl 47(34): 6399-6401.

Hake, S. B., A. Xiao, et al. (2004). "Linking the epigenetic 'language' of covalent histone modifications to cancer." Br J Cancer 90(4): 761-769.

Hoyle, C. E. and C. N. Bowman (2010). "Thiol-ene click chemistry." Angew Chem Int Ed Engl 49(9): 1540-1573.

Huang, R., M. A. Holbert, et al. (2010). "Site-specific introduction of an acetyl-lysine mimic into peptides and proteins by cysteine alkylation." J Am Chem Soc 132(29): 9986-9987.

Kim, S. C., R. Sprung, et al. (2006). "Substrate and functional diversity of lysine acetylation revealed by a proteomics survey." Mol Cell 23(4): 607-618.

Liu, L., D. M. Scolnick, et al. (1999). "p53 sites acetylated in vitro by PCAF and p300 are acetylated in vivo in response to DNA damage." Mol Cell Biol 19(2): 1202-1209.

Marks, P. A. and R. Breslow (2007). "Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug." Nat Biotechnol 25(1): 84-90.

Minucci, S. and P. G. Pelicci (2006). "Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer." Nat Rev Cancer 6(1): 38-51.

Neumann, H., S. Y. Peak-Chew, et al. (2008). "Genetically encoding N(epsilon)-acetyllysine in recombinant proteins." Nat Chem Biol 4(4): 232-234.

Norris, K. L., J. Y. Lee, et al. (2009). "Acetylation goes global: the emergence of acetylation biology." Sci Signal. 97: pe 76.

Raftery, M. A. and R. D. Cole (1966). "On the aminoethylation of proteins." J Biol Chem 241(15): 3457-3461.

Sakaguchi, K., J. E. Herrera, et al. (1998). "DNA damage activates p53 through a phosphorylation-acetylation cascade." Genes Dev 12(18): 2831-2841.

Shahbazian, M. D. and M. Grunstein (2007). "Functions of site-specific histone acetylation and deacetylation." Annu Rev Biochem 76: 75-100.

Shogren-Knaak, M., H. Ishii, et al. (2006). "Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions." Science 311: 844-847.

Shogren-Knaak, M., H. Ishii, et al. (2006). "Histone H4-K16 acetylation controls chromatin structure and protein interactions." Science 311(5762): 844-847.

Spange, S., T. Wagner, et al. (2009). "Acetylation of non-histone proteins modulates cellular signalling at multiple levels." Int J Biochem Cell Biol 41(1): 185-198.

van Holde, K. E. and W. O. Weischet (1978). "Boundary analysis of sedimentation-velocity experiments with monodisperse and polydisperse solutes." Biopolymers 17: 1387-1403.

Wang, Q., Y. Zhang, et al. (2010). "Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux." Science 327(5968): 1004-1007.

Yang, X. J. (2004). "The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases." Nucleic Acids Res 32(3): 959-976.

Chapter 4: PNA-templated Disulfide Bond Formation

4.1 Introduction

Watson-Crick base pairing in nucleic acids provides a mechanism for the faithful transfer of genetic information in biological systems. It guides all the reactions in the fundamental processes of DNA replication, repair and recombination, transcription of DNA into RNA and translation of mRNA into proteins whereby component building blocks are recruited and assembled alongside a nucleic acid template. Assisted by other macromolecular partners, the template plays the key role of holding the respective reactants in close proximity and correct geometry and thus increasing their effective molarity at the reactive sites to ensure highly efficient and specific reactions.

In a way mimicking nature's use of effective molarity to control chemical reactivity, chemists have exploited this principle of complementary base pairing in developing the concept of nucleic acid templated synthesis (Li and Liu 2004). The first attempt was the formation of a phosphodiester bond through the hybridization of DNA or RNA (Fig. 1). An oligomer complementary to the template was successfully formed by the reaction between the activated 5' phosphate and the 3' hydroxy group of a nucleotide monomer (Naylor and Gilham 1966; Orgel 1995).

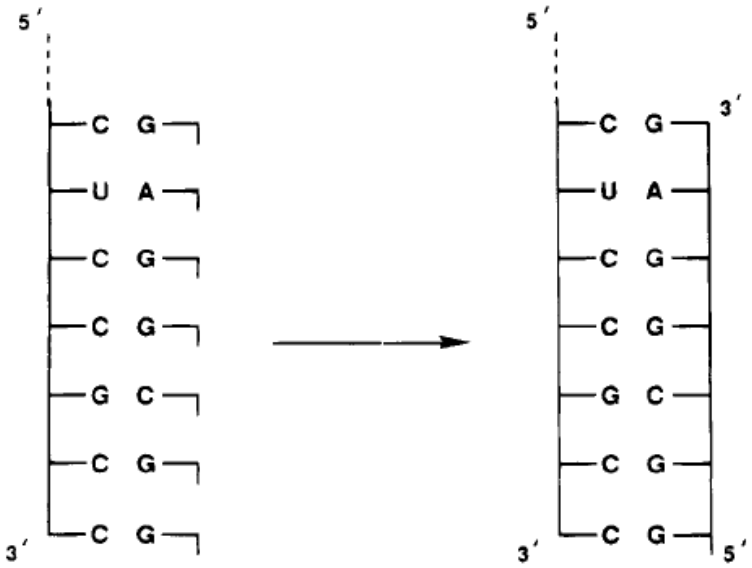


Figure 1. *The formation of phosphodiester bond through nucleic acid templated synthesis (adopted from* Orgel 1995*).*

4.1.1 DNA-templated organic synthesis (DTS)

Recent efforts are more focused on using DNA-templated organic synthesis (DTS) as an enabling technique for organic synthesis. When attached to oligonucleotide adapters, two small molecule substrates can be brought together through DNA hybridization, thereby enhancing the reaction rate between their reactive functional groups via the same phenomenon of effective molarity (Fig. 2) (Gartner and Liu 2001).

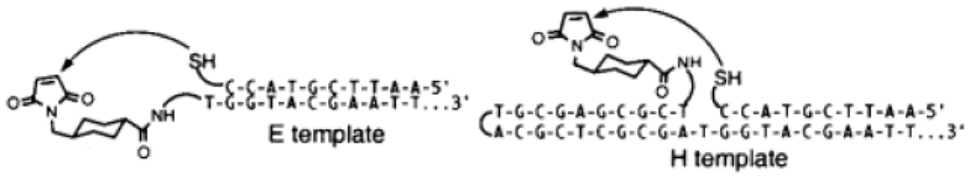


Figure 2. *DNA-templated synthesis by hairpin (H) and end-of-helix (E) DNA templates (adopted from* Gartner and Liu 2001*).*

Research in recent years has shown the surprising versatility of DTS, with its applications demonstrated in many settings including the construction of complex organic molecules and polymer products (Fig. 3) (Calderone and Liu 2004), creation of evolvable and encoded compound libraries (Gartner and Liu 2001; Doyon, Snyder et al. 2003; Gartner, Tse et al. 2004) and the discovery of novel organic synthesis reactions (Kanan, Rozenman et al. 2004; Momiyama, Kanan et al. 2007; Rozenman, Kanan et al. 2007).

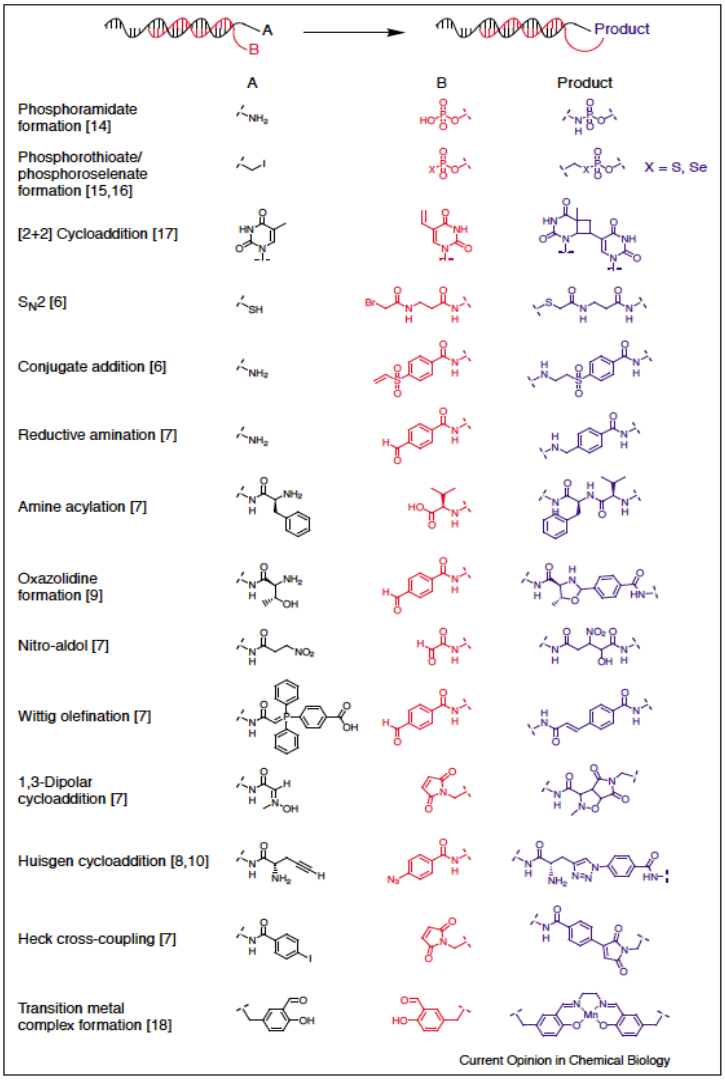


Figure 3. *Reactions directed by DTS (adopted from* Calderone and Liu 2004*)*

4.1.2 Strategies of our study: PNA-templated organic synthsis (PTS)

Inspired by these results, we sought to use PNA in lieu of DNA for similar PNA-templated reactions. PNA is a synthetic mimic of DNA and RNA which can bind to complementary DNA, RNA or another PNA with high affinity and sequence specificity according to Watson-Crick base pairing rules (Watson and Crick 1953). In PNA, the nucleobases are attached via methylenecarbonyl likers to a *poly-N*-(2-aminoethyl)glycine (aeg) backbone (Nielsen, Egholm et al. 1991; Egholm, Buchardt et al. 1993; Nielsen and Haaima 1997; Braasch and Corey 2002). The high binding affinity of PNA to DNA or RNA is due partly to the lack of negative charge on its backbone and presumably to the proper interbase distances, the high flexibility of the aminoethyl linkers, and eventually intramolecular hydrogen bonding (Fig. 4) (Hanvey, Peffer et al. 1992; Hyrup, Egholm et al. 1994; Wittung, Nielsen et al. 1994; Kushon, Jordan et al. 2001). PNA can be easily synthesized by the widely used solid phase peptide synthesis method due to the amide linkage in its backbone.

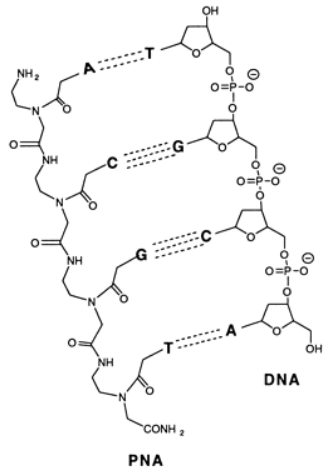


Figure 4. *The structure of peptide nucleic acid (PNA) and its binding with DNA.*

Peptide nucleic acid (PNA) shows several better properties than DNA. First of all, the PNA monomers are like amino acids, as a result PNA can be easily synthesized by the widely used Fmoc solid phase peptide synthesis. As such, PNA is also synthetically compatible with many types of organic compounds in direct solid phase synthesis. Second, with its peptide-like features, it is very easy to synthesize peptide-PNA conjugates by the chemical ligation methods. Third, the thermal stability of the PNA/PNA duplex is considerably higher than that of the PNA/DNA or DNA/DNA duplex, suggesting that shorter PNAs can be used as templates for templated reactions. Fourth, the complementary PNA/PNA duplex can be formed in the parallel mode, which can simplify the synthesis of peptide-PNA conjugates for templated peptide dimerization. Last, PNA is stable to nucleases and proteases, and can be stored for a long time without degradation. All these features would facilitate the preparation of substrate-PNA conjugates and make PNA-templated synthesis an easy-to-use approach. In this study we show that Watson-Crick base pairing between parallel complementary PNA strands can direct peptide heterodimerization via disulfide bond formation.

Disulfide bonds play an important role in stabilizing the bioactive conformation of many peptides and proteins (Thornton 1981; Gilbert 1990; Betz 1993; Annis, Hargittai et al. 1994; Sevier and Kaiser 2002; Witt 2008). Selective disulfide formation is a challenging synthetic problem in peptide chemistry, especially in the case of multimeric or polythiol peptide. For example, to prepare a simple peptide

heterodimer linked via a single disulfide bridge, random oxidative disulfide formation between the sulfhydryl groups on the peptides would produce three dimeric products: two homodimers and the desired heterodimer. If there are two disulfide bonds in the peptide heterodimer, the formation of disulfide bridges would be more complicated. Eight different products can be produced through random oxidation.

We reasoned that such problems could be tackled effectively in DNA- or PNA-templated reaction systems. In such a system, two peptides are linked to a pair of complementary PNA strands and PNA hybridization would bring the reactive Cys sulfhydryls into proximity, thus ensuring selective hetero-disulfide bond formation via air oxidation. Figure 5 shows the outline of the reaction steps involved in the preparation of a heterodimer linked via two disulfide bonds. First, the two peptide-PNA conjugates **1** and **2** are prepared by thioester-mediated chemical ligation between a peptide thioester and a cysteinyl-PNA, respectively. Second, the two conjugates are mixed together for the annealing between the two PNAs on the basis of the Watson-Crick base pairing rules. Third, the mixture is allowed to stand at room temperature for 24 h to allow the first disulfide bond formation between the two free Cys residues via air oxidation to form the conjugate **3**. The reaction would be specific due to the close proximity of the two Cys residues in the two peptide-PNA conjugates. Finally, the iodine is used for the oxidation and concomitant removal of Acm on the other two Cys residues near the N-termini to form the second disulfide specifically (Kamber, Hartmann et al. 1980).

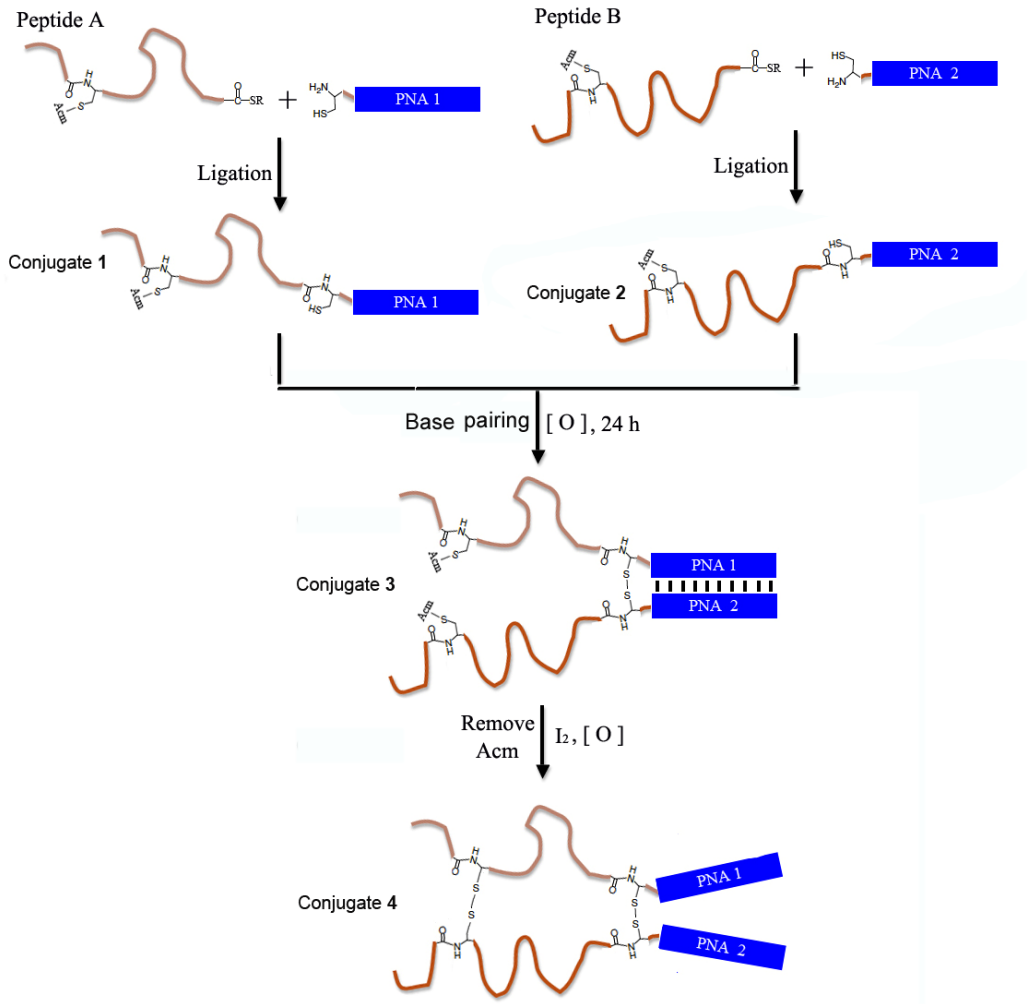


Figure 5. *PNA-templated disulfide bond formation for peptide heterodimerization. Step1: conjugate* ***1*** *and conjugate* ***2*** *synthesized by chemical ligation; Step 2: complementary PNA pairing between two conjugates and formation of the first disulfide bond; Step 3: formation of the second disulfide bond through iodine oxidation with removal of Acm.*

Our results show that this reaction system is highly selective with the desired heterodimer formed exclusively. Since PNA synthesis is fully compatible with peptide synthesis, this method has an advantage over and can be more useful than the DNA-templated reaction systems.

4. 2 Results

4.2.1 Formation of two disulfide bonds between PNA-peptide conjugates

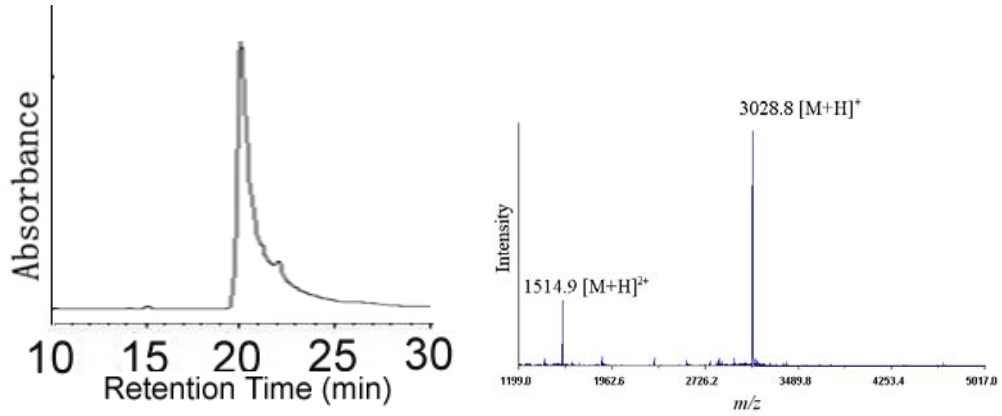
As mentioned above, we first designed two peptide-PNA conjugates to test this idea of PNA-directed disulfide formation. Each conjugate was prepared using chemical ligation between a C-ter thioester peptide and a PNA containing a Cys at the amino end. Therefore, two relatively large peptide components, each comprising 17 and 14 amino acid residues respectively, are linked to the amino termini of two PNA strands. Therefore, the Cys residue, to be engaged in the disulfide bond, is found near the C-ter end of the peptide component in each conjugate. The two PNA strands with a 10-nucleobase sequence were designed for hybridization in the parallel orientation and the Tm of such a parallel PNA-PNA duplex is about 48 ºC. The PNA sequences were designed as following:

PNA 1: Cys-Ala-Ala-GTAGATCACT-Gly-$CONH_2$

PNA 2: Cys-Ala-Ala-CATCTAGTGA-Gly-$CONH_2$

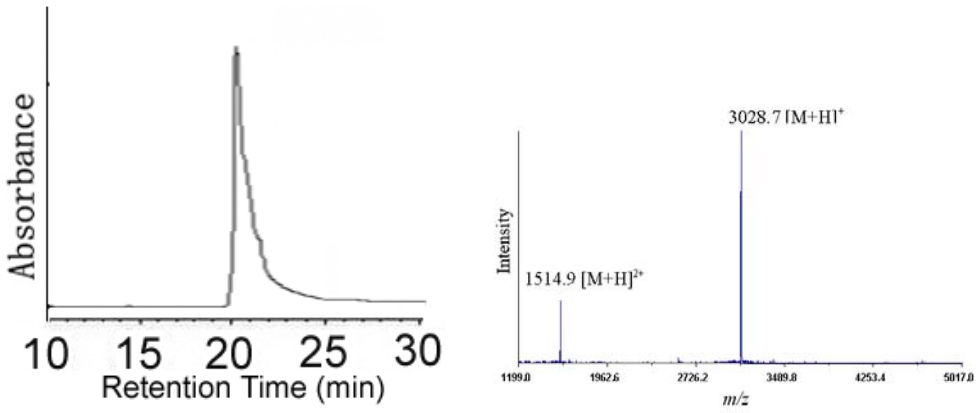
4.2.1.1 Synthesis and purification of PNAs

The PNAs were synthesized by the standard Fmoc solid phase peptide synthesis (SPPS) method with minor revisions. For the synthesis of PNA, the Fomc deprotection was performed three times for 1 min, 5 min and 10 min each to avoid side reactions. And in the coupling step, 4 eq. of PyBOP activated amino acid was used to the resin and the reaction conducted at room temperature for 1 h. And 2 eq. of PyBOP activated PNA monomer was used to the resin for coupling for 5 h. After TFA cleavage and HPLC purification, PNA 1 and PNA 2 were obtained (Fig. 6 and 7).



A **B**

Figure 6. *Characterization of PNA 1. (A) C18 analytical HPLC profile of PNA1. (B) Mass spectrum of the PNA determined by MALDI-TOF MS. ($[M+H]^+$ found: 3028.8, MW calcd: 3027.4). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*



A **B**

Figure 7. *Characterization of PNA 2. (A) C18 analytical HPLC profile of PNA2. (B) Mass spectrum of the PNA determined by MALDI-TOF MS. ($[M+H]^+$ found: 3028.7, MW calcd: 3027.4). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

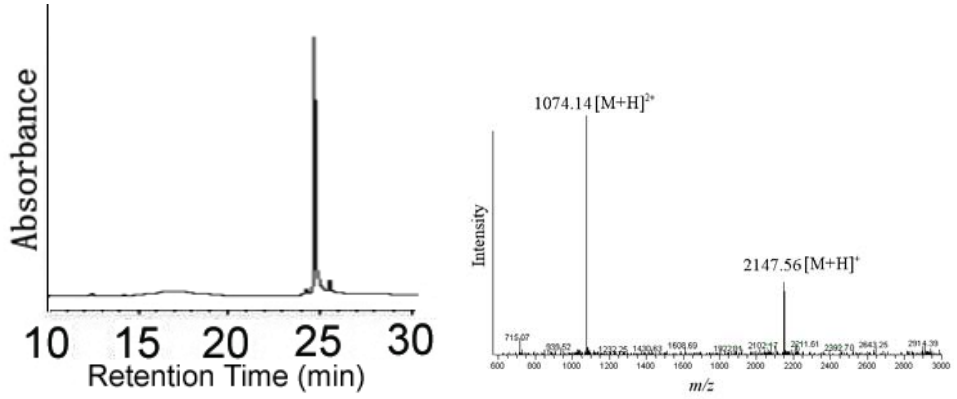
4.2.1.2 Synthesis and purification of peptide thioesters

For the two peptide thioesters, the sequences were designed from the insulin A chain and B chain. There were two cysteines in each peptides and the one which was near the N-terminus was protected by Acm. After the formation of the first disulfide bond between the cysteins near the C-termini, the Acm can be removed by iodine oxidation. The sequences of the peptide thioesters were designed as follows:

Peptide A: Lys-Arg-Gln-Ala-Cys(Acm)-Thr-Ser-Ile-Ala-Ser-Leu-Tyr-Gln-Leu-Glu-Asn-Tyr-$SCH_2CH_2CONH_2$

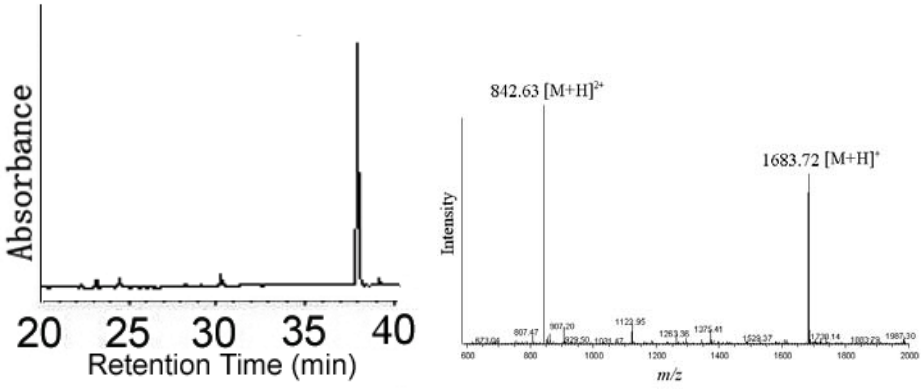
peptide B: His-Leu-Cys(Acm)-Gly-Leu-Ser-His-Leu-Val-Glu-Ala-Leu-Tyr-Ala-$SCH_2CH_2CONH_2$

The peptide thioesters were synthesized by the Boc solid phase peptide synthesis (SPPS) method. Trt-SCH_2CH_2COOH was loaded first onto the MBHA resin in order to synthesize the peptide with a thioester at the C-terminus. After TFMSA cleavage and HPLC purification, peptide A and peptide B were obtained (Figs. 8 and 9).



A **B**

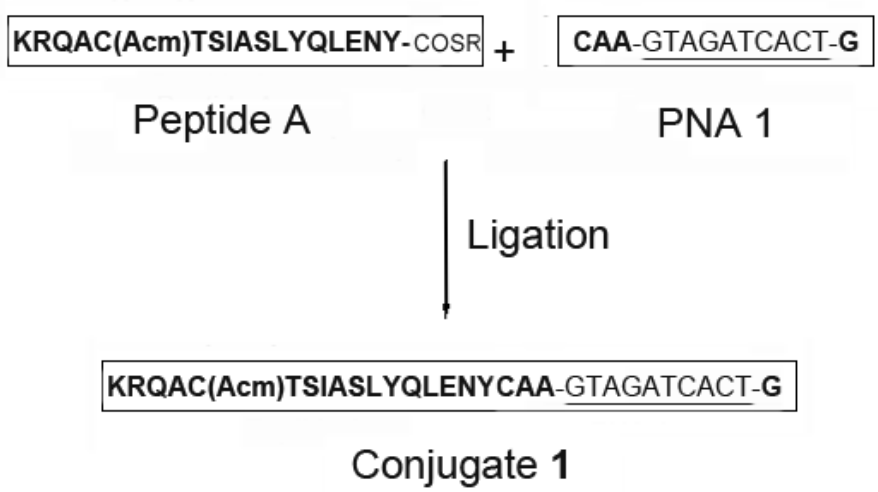
Figure 8. *Characterization of peptide A. (A) C18 analytical HPLC profile of peptide 1. (B) Mass spectrum of the peptide determined by ESI-MS. ($[M+H]^+$ found: 2147.56, MW calcd: 2146.27). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*



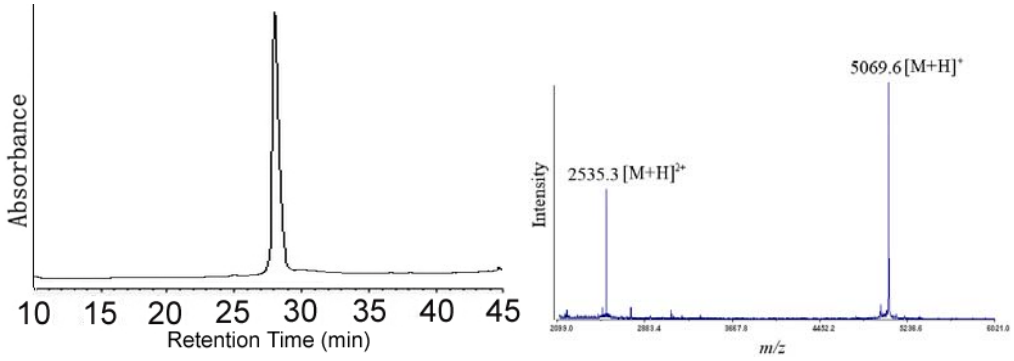
A **B**

Figure 9. *Characterization of peptide B. (A) C18 analytical HPLC profile of peptide 2. (B) Mass spectrum of the peptide determined by ESI-MS. ($[M+H]^+$ found: 1683.72, MW calcd: 1682.64). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

4.2.1.3 Ligation of peptide A and PNA 1



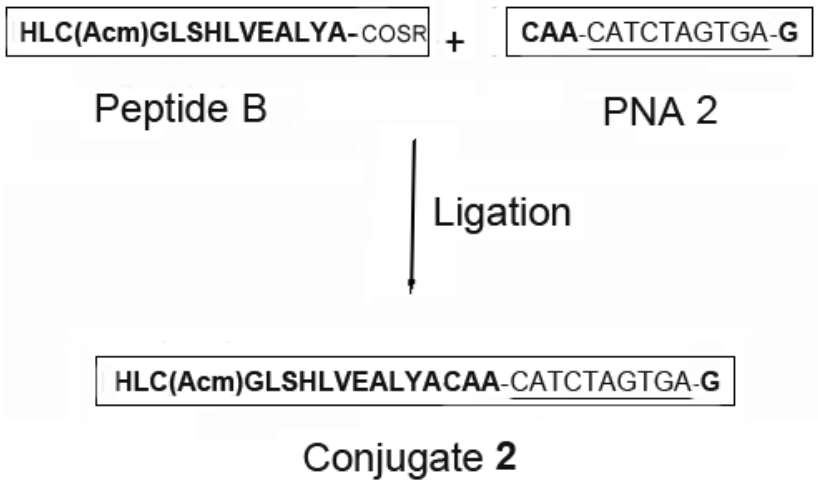
The reaction of thioester-mediated chemical ligation between peptide A and PNA 1 was monitored by analytical RP-HPLC and the reaction was complete after 6 hrs. The purity of conjugate **1** was checked by HPLC and was confirmed by MALDI-TOF MS (Fig. 10).



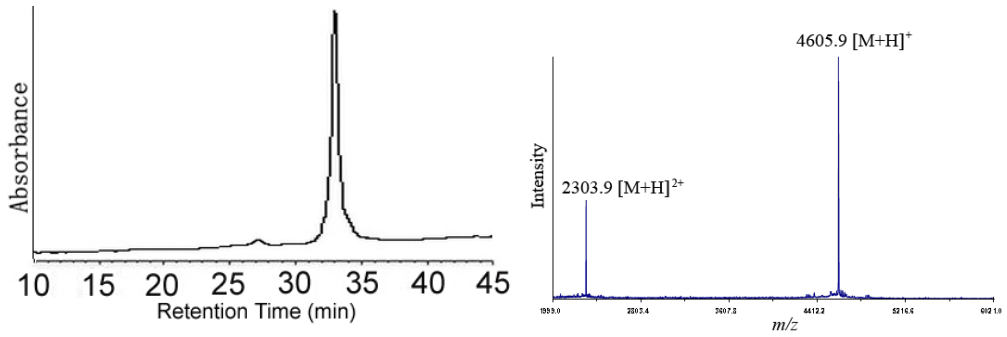
A **B**

Figure 10. *Characterization of the purified conjugate* ***1****. (A) C18 analytical HPLC profile of purified conjugate* ***1****. (B) Mass spectrum of the conjugate* ***1*** *determined by MALDI-TOF MS. ($[M+H]^+$ found: 5069.6, MW calcd: 5068.7). HPLC condition: 0% to 50% of buffer B in buffer A in 50 min.*

4.2.1.4 Ligation of peptide B and PNA 2



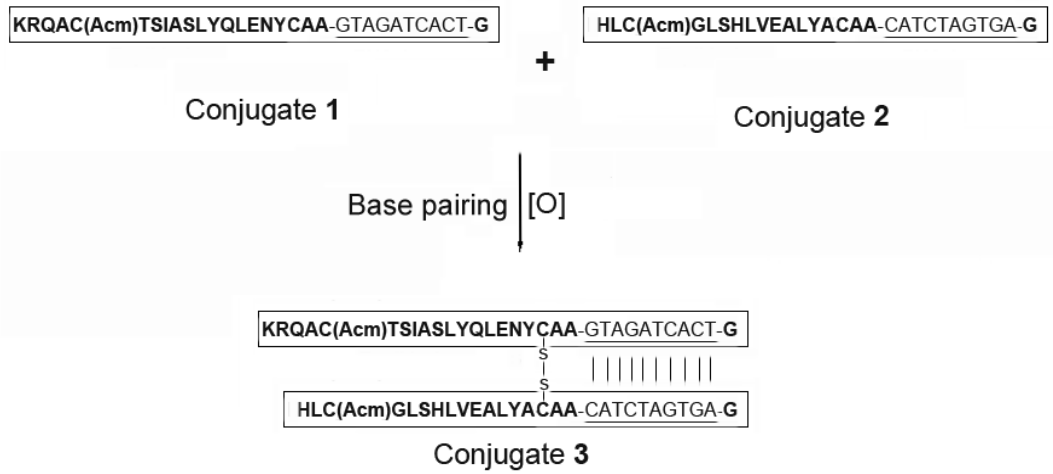
The ligation reaction between peptide B and PNA 2 was performed under the same conditions. The purified conjugate **2** was checked by HPLC and confirmed by MALDI-TOF MS (Fig. 11).



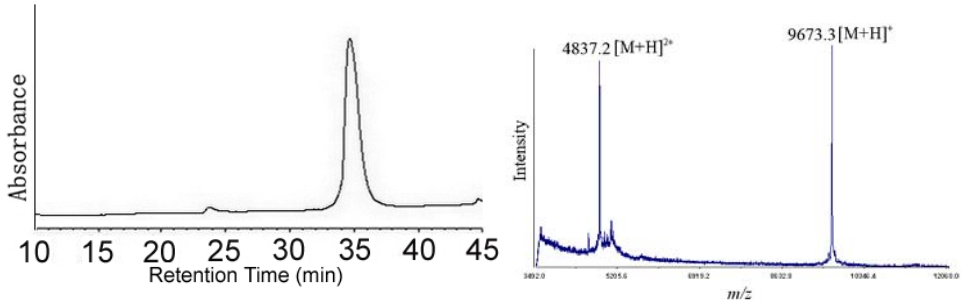
A **B**

Figure 11. *Characterization of purified conjugate* ***2****. (A) C18 analytical HPLC profile of purified conjugate* ***2****. (B) Mass spectrum of the conjugate* ***2*** *determined by MALDI-TOF MS. ([M+H]$^+$ found: 4605.9, MW calcd: 4604.1). HPLC condition: 0% to 50% of buffer B in buffer A in 50 min.*

4.2.1.5 Formation of first disulfide bond



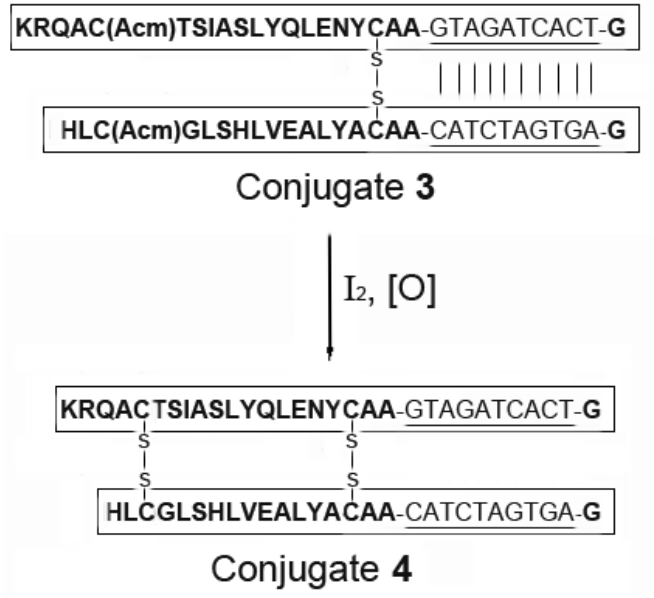
The above two conjugates (10 μM) were then dissolved in 10 mM sodium phosphate buffer (pH 7.0), and the mixture solution was placed in a 70 ^{o}C water bath for 10 min and cooled slowly to room temperature. During this process, annealing between the two PNA occurred according to the Watson-Crick base pairing. The mixture was allowed to stand at room temperature for 24 h to allow the formation of first disulfide bond between the two free Cys residues via air oxidation to form conjugate **3**. As shown in HPLC and MS analyses, the reaction was specific due to the close proximity of the two Cys residues and the product was formed cleanly with the expected -2 Da MW for the formation of disulfide bond (Fig. 12). No homodimer of conjugate **1** or conjugate **2** was detected.



A **B**

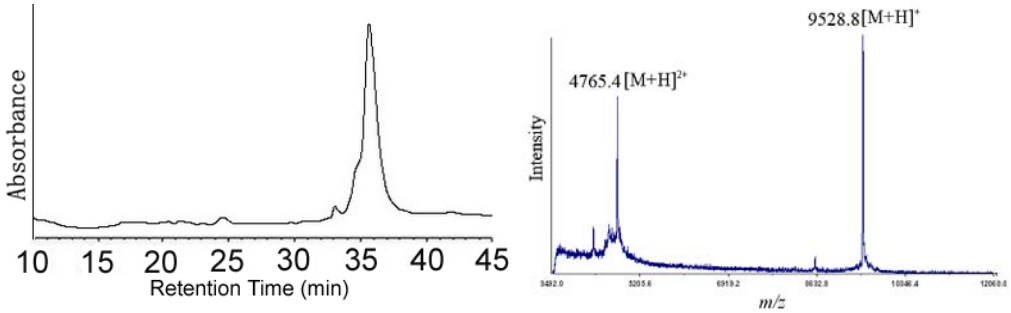
Figure 12. *Characterization of conjugate* ***3****. (A) C18 analytical HPLC profile of conjugate* ***3****. (B) Mass spectrum of the conjugate* ***3*** *determined by MALDI-TOF MS. ([M+H]*$^{+}$ *found: 9673.3, MW calcd: 9672.8). HPLC condition: 0% to 50% of buffer B in buffer A in 50 min.*

4.2.1.6 Formation of second disulfide bond



The formation of a second disulfide bond in the above product was subsequently achieved through iodine oxidation with concomitant removal of Acm on the other two Cys residues near the N-termini of the peptide components. Briefly, the Cys (Acm) peptide-PNA conjugate **3** was dissolved in 50% MeOH to a final concentration of

1mg/ml. 1.5 eq. of HCl was added, followed by 5-10 eq. of 0.1 M I_2 (in MeOH). After 30 min of reaction, excess iodine was quenched by sodium thiosulfate. The final conjugate **4** with two disulfide bonds was purified by HPLC, and the purified product was characterized by MALDI-MS with the expected -144 Da MW for the removal of Acm and formation of the second disulfide bond (Fig. 13). Therefore, the presence of the first disulfide made the formation of the second disulfide an intramolecular reaction which was effective and also very specific.



A **B**

Figure 13. *Characterization of conjugate* ***4****. (A) C18 analytical HPLC profile of conjugate* ***4****. (B) Mass spectrum of the conjugate* ***4*** *determined by MALDI-TOF MS. ($[M+H]^+$ found: 9528.8, MW calcd: 9527.6). HPLC condition: 0% to 50% of buffer B in buffer A in 50 min.*

4.2.2 Trypsin digestion and release of peptide heterodimer from PNA

To further validate this PNA-templated reaction system, a second experiment was conducted in which the S-S linked peptide dimeric product was released from the PNA components and isolated. In order to do this, two new peptide-PNA conjugates were designed, each having a Lys residue between the peptide and PNA components.

No other Lys or Arg residues were present in the peptides. This would allow specific tryptic cleavage at the position after the Lys residue (Fig. 14). The results show that the disulfide bond was selectively formed and the peptide conjugate was successfully released by the treatment of trypsin. This technology would greatly broaden the application of PNA-templated synthesis.

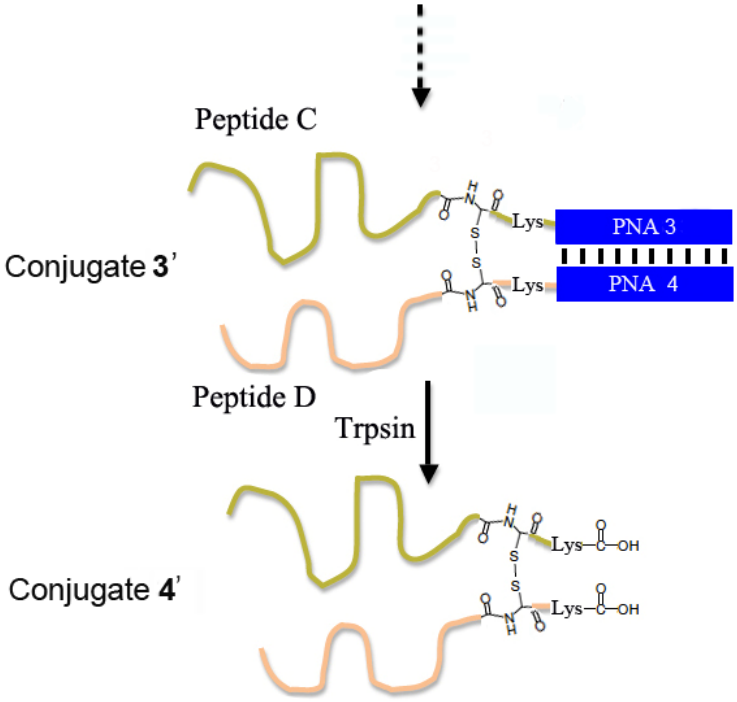
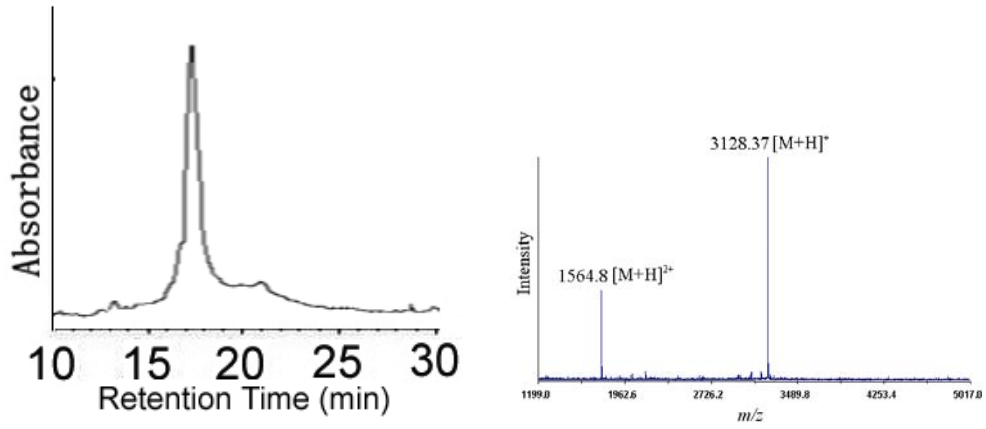


Figure 14. *Trypsin digestion and release of peptide conjugate* ***4'***.

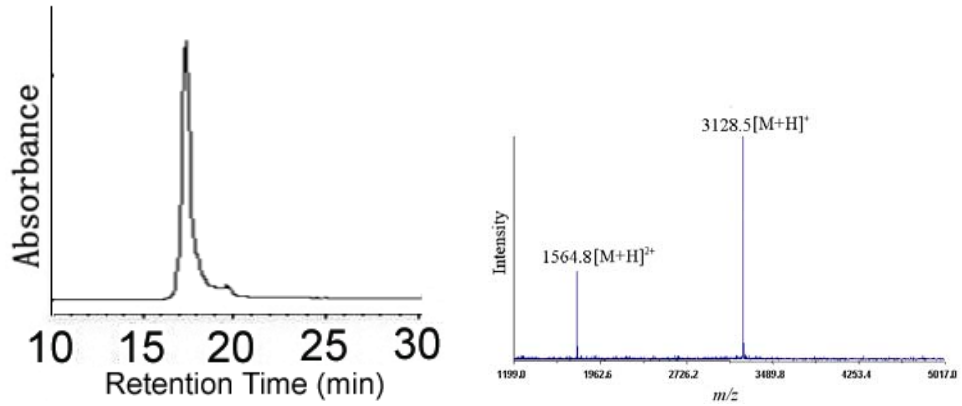
4.2.2.1 Synthesis and purification of PNAs

PNA 3 (Cys-Gly-Lys-Gly-GAACATTTGC-Gly-$CONH_2$) and PNA 4 (Cys-Gly-Lys-Gly-CTTGTAAACG-Gly-$CONH_2$) were also synthesized by the Fmoc solid phase peptide synthesis (SPPS) method. After TFA cleavage and HPLC purification, PNA 3 and PNA 4 were obtained (Fig. 15 and 16).



A **B**

Figure 15. *Characterization of PNA 3. (A) C18 analytical HPLC profile of PNA 3. (B) Mass spectrum of the PNA determined by MALDI-TOF MS. ($[M+H]^+$ found: 3128.4, MW calcd: 3127.5). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*



A **B**

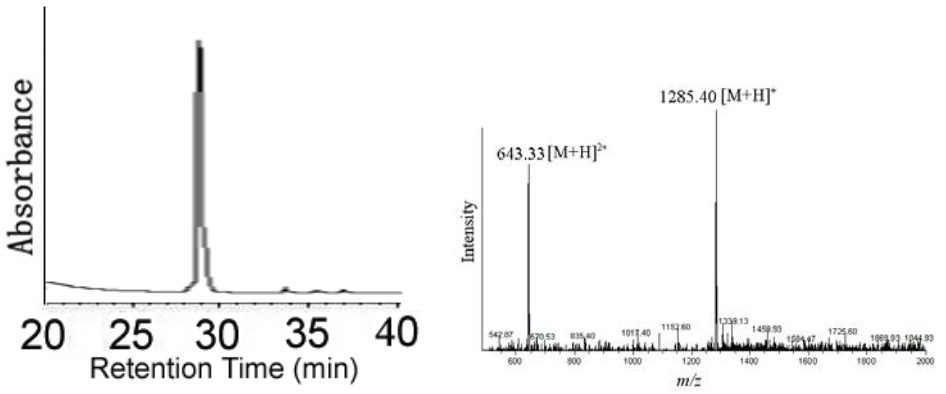
Figure 16. *Characterization of PNA 4. (A) C18 analytical HPLC profile of PNA 4. (B) Mass spectrum of the PNA determined by MALDI-TOF MS. ($[M+H]^+$ found: 3128.5, MW calcd: 3127.5). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

4.2.2.2 Synthesis and purification of peptide thioesters

peptide C (Thz-Gln-Gly-Ser-Gly-Ala-Thr-Glu-Ser-Ala-Leu-Tyr-$SCH_2CH_2CONH_2$)

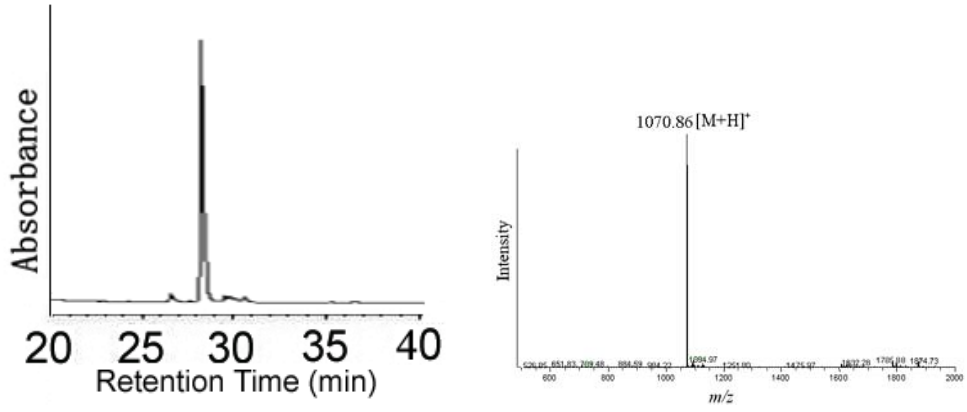
and peptide D (Thr-Ser-Gly-Glu-Ala-Tyr-Val-Ser-Leu-Gly-$SCH_2CH_2CONH_2$) were also

synthesized by the Boc solid phase peptide synthesis (SPPS) method. After TFMSA cleavage and HPLC purification, peptide C and peptide D were obtained and confirmed by ESI-MS (Fig. 17 and 18).



A **B**

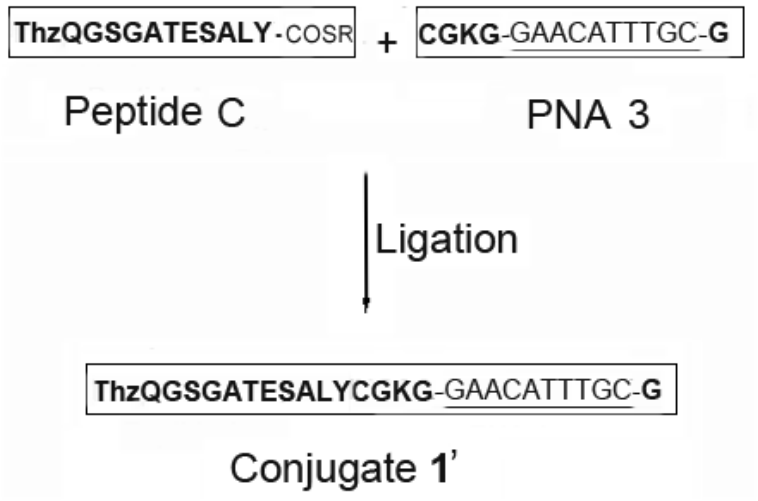
Figure 17. *Characterization of peptide C. (A) C18 analytical HPLC profile of peptide 3. (B) Mass spectrum of the peptide determined by ESI-MS. ($[M+H]^+$ found: 1285.40, MW calcd: 1285.28). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*



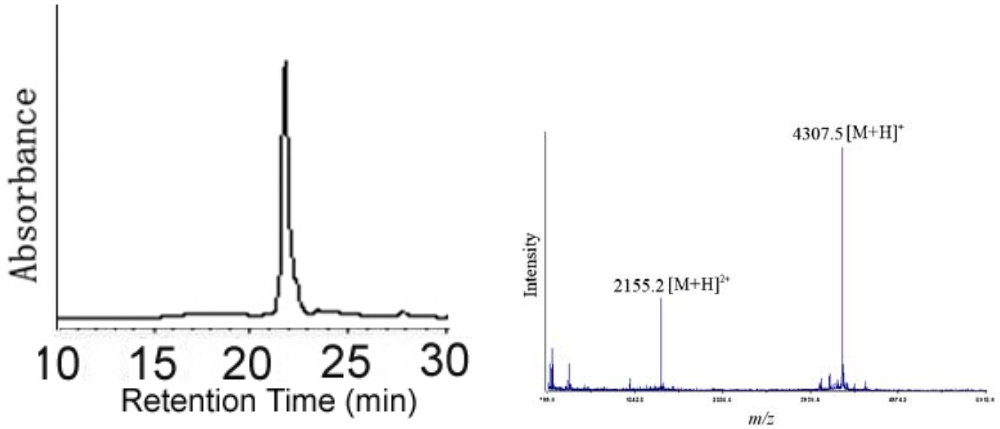
A **B**

Figure 18. *Characterization of peptide D. (A) C18 analytical HPLC profile of peptide 4. (B) Mass spectrum of the peptide determined by ESI-MS. ($[M+H]^+$ found: 1070.86, MW calcd: 1069.46). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

4.2.2.3 Ligation of peptide C and PNA 3



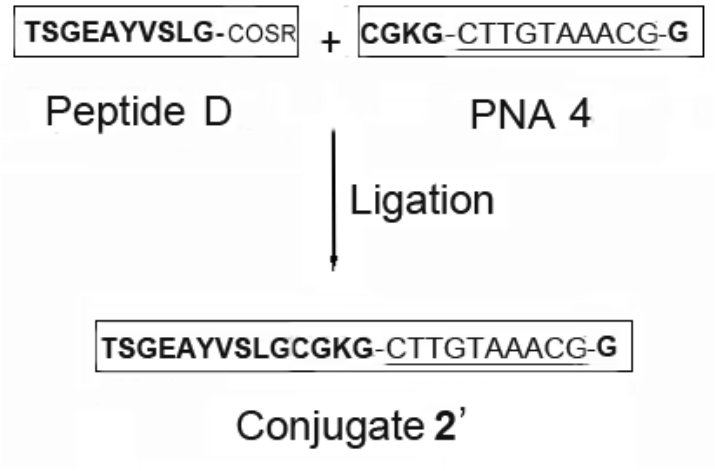
The freeze-dried PNA 3 and peptide C were dissolved in the ligation buffer (100 mM NaH_2PO_4, 50 mM TCEP, pH 8.0) and the mixture was allowed to react for 6 h. The purified conjugate **1'** was checked by HPLC and confirmed by MALDI-TOF MS (Fig. 19).



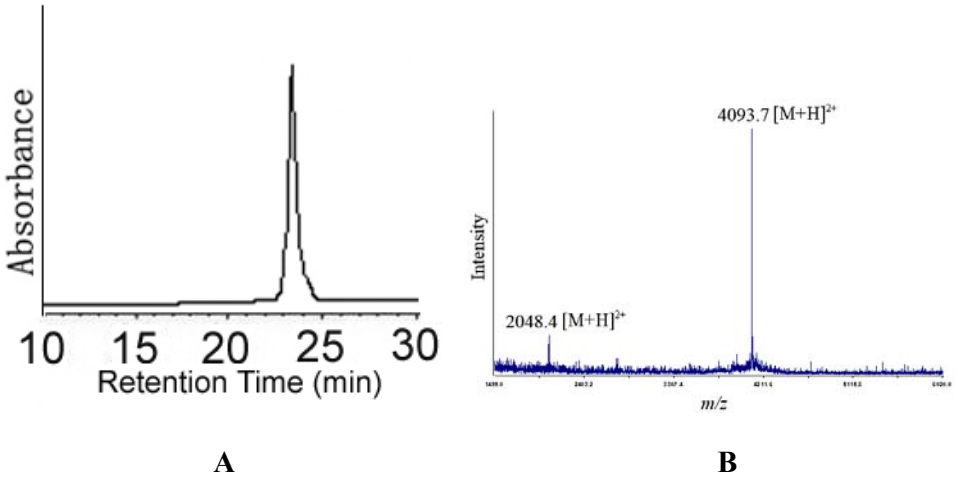
A **B**

Figure 19. *Characterization of purified conjugate* ***1'****. (A) C18 analytical HPLC profile of purified conjugate* ***1'****. (B) Mass spectrum of the conjugate* ***1'*** *determined by MALDI-TOF MS. ($[M+H]^+$ found: 4307.5, MW calcd: 4307.5). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

4.2.2.4 Ligation of peptide D and PNA 4



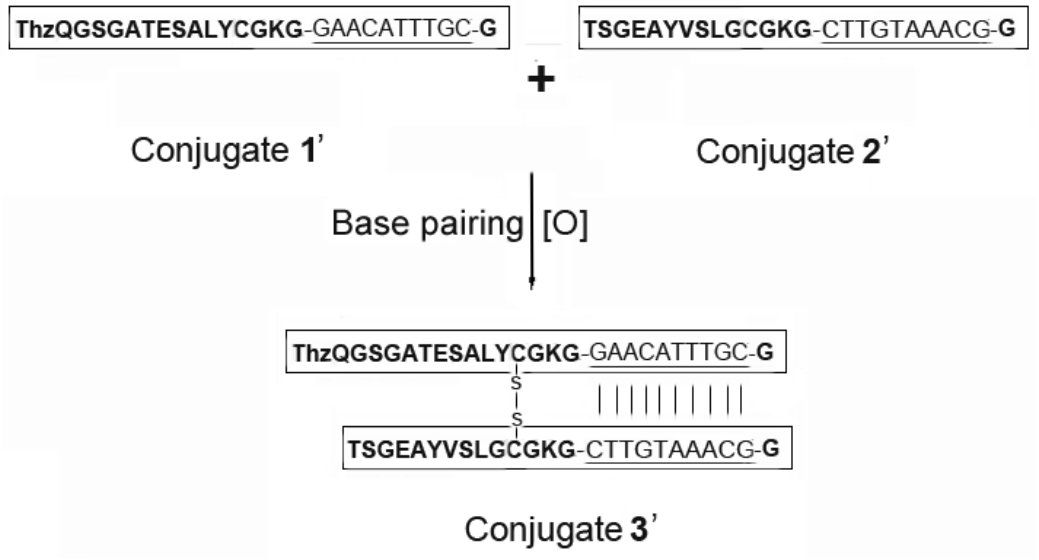
The ligation was performed in the same condition. The purified conjugate **2'** was checked by HPLC and confirmed by MALDI-TOF MS (Fig. 20).



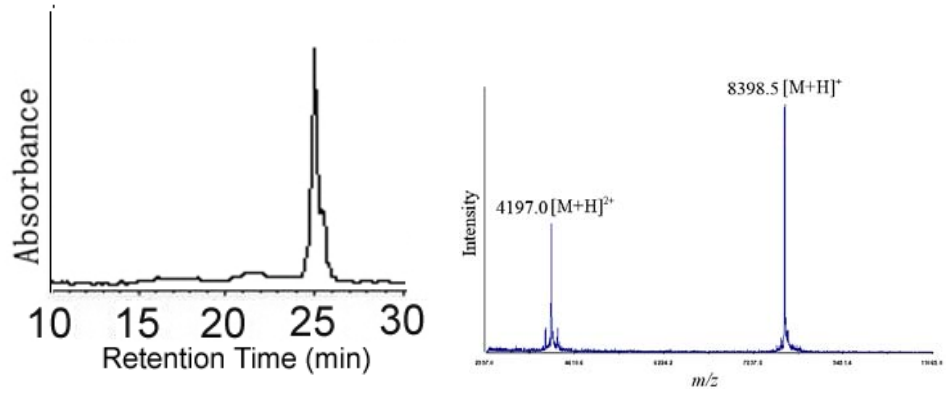
A **B**

Figure 20. *Characterization of purified conjugate* ***2'****. (A) C18 analytical HPLC profile of purified conjugate* ***2'****. (B) Mass spectrum of the conjugate* ***2'*** *determined by MALDI-TOF MS. ($[M+H]^+$ found: 4093.7, MW calcd: 4092.0). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

4.2.2.5 Formation of disulfide bond



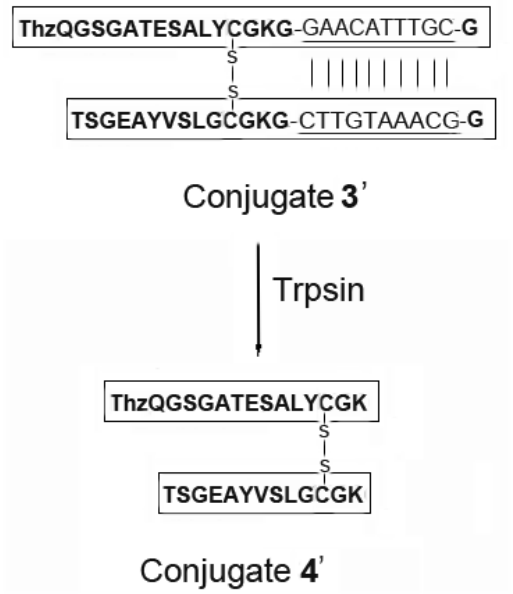
After PNA hybridization and air oxidation under the same conditions as for the preparation of conjugate **3**, the disulfide bond was formed between the two peptide and the conjugate **3'** was purified by HPLC and its identity confirmed by MALDI-MS analysis (Fig. 21).



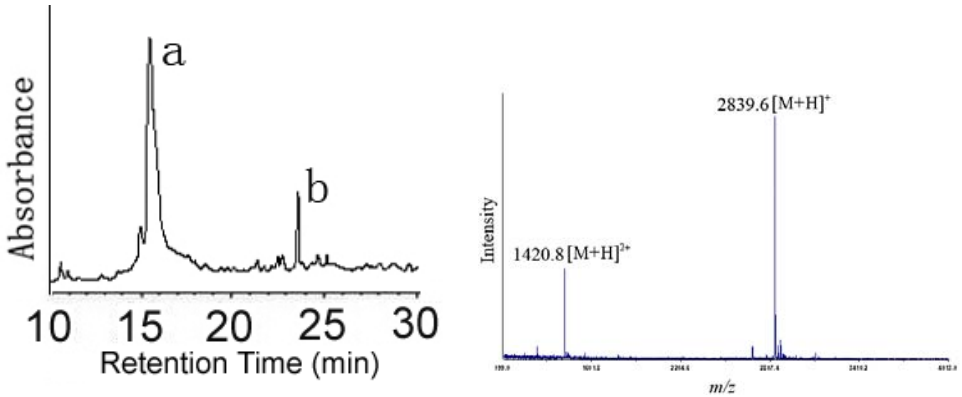
A **B**

Figure 21. *Characterization of conjugate* ***3'****. (A) C18 analytical HPLC profile of conjugate* ***3'****. (B) Mass spectrum of the conjugate determined* ***3'*** *by MALDI-TOF MS. ([M+H]$^+$ found: 8398.5, MW calcd: 8397.5). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min.*

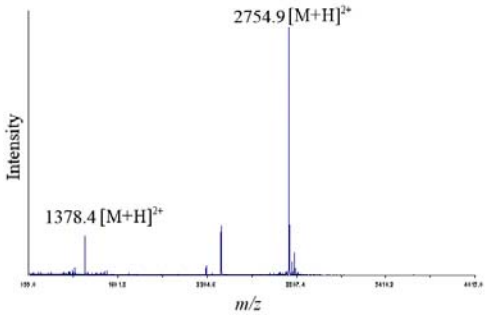
4.2.2.6 Trypsin digestion



The freeze-dried conjugate was dissolved in 0.1 M $NaHCO_3$ buffer (pH 8.3), and trypsin was added at a ratio of 1 : 20 (w/w) to the conjugate. The reaction mixture was allowed to stand at room temperature overnight. The trypsin treated mixture was subjected to HPLC analysis. MS analysis clearly confirmed the expected mass of the heterodimeric peptide (peak b) as well as that of the PNAs (peak a) separated after trypsin digestion (Fig. 22).



A **B**



C

Figure 22. *Characterization of trypsin digestion. (A) C18 analytical HPLC profile of trypsin digestion. Peak a: PNAs; Peak b: peptide conjugate* ***4'****. (B) Mass spectrum of the peak a determined by MALDI-TOF MS ($[M+H]^+$ found: 2839.6, MW calcd: 2839.2). Mass spectrum of the peak b determined by MALDI-TOF MS ($[M+H]^+$ found: 2754.9, MW calcd: 2754.5). HPLC condition: 0% to 40% of buffer B in buffer A in 40 min. (Note: the two PNAs have identical base and amino acid compostion and therefore the same MW.)*

4.3 Discussion and conclusion

A seemingly simple reaction, oxidation of sulfhydryls to form disulfide bonds can be very complex in its outcome owing to problems of random cross-linking between thiols of virtual identical reactivity. In many natural proteins containing S-S bonds, correct pairing of Cys residues is helped by the protein's tendency to fold into the most stable conformation in which the respective sulfhydryl groups are brought together for selective bridging. This is yet another example of nature's use of effective molarity in controlling the selectivity of –SH oxidation, as first demonstrated in the

landmark study of the oxidative refolding of reduced ribonuclease S (Anfinsen, Haber et al. 1961; Anfinsen 1973). However, for cases where such intrinsic property of effective molarity is absent, such as the heterodimeric systems mentioned at the beginning of this chapter, selective formation of an interchain asymmetric disulfide is not possible. This problem can be overcome by installing an artificial proximity-driving device such as PNA hybridization.

The above data show that intermolecular disulfide bond formation can be achieved very specifically and efficiently under the direction of Watson-Crick base pairing. This system is quite useful since, compared to DNA-templated organic synthesis (DTS), PTS is a better substitute for the following reasons. Firstly, as mentioned before, PNA synthesis is mostly like peptide synthesis using commercially available PNA monomers by the widely used Boc or Fmoc SPPS. Secondly, with its peptide-like features, PNA can be ligated to another peptide by chemical ligation methods in place of the complicated biotin-streptavidin coupling system that is often used for conjugation between the DNA and protein. Thirdly, the complementary PNA/PNA can be bound in parallel mode, which can simplify the procedures to prepare PNA-peptide conjugates. Above all, PNA has a relatively simple structure and is virtually nondegradable in biological systems. These features make PNA an attractive platform for many chemical applications like the one discussed in this chapter. More experiments are underway to further demonstrate its use in more complex conjugate systems and other chemical reactions.

4.4 Materials and methods

4.4.1 Materials

All amino acids and coupling reagents were purchased from GL Biochem (Shanghai, China) and Novabiochem (Germany). Rink amide resin was purchased from Merck. All the PNA monomers were purchased from ASM research chemicals. The trypsin was from Sigma company. All other chemical reagents were purchased from commercial suppliers.

4.4.2 Methods

4.4.2.1 Synthesis of PNAs

The Rink Amide PEGA resin was swelled in DMF/DCM (5:1) solvent for 3 hrs. The Fmoc deprotection was performed with 20% piperidine in DMF twice (2 min × 1, 20 min × 1). After washing the resin with DMF (2 ×), DCM (2 ×) and DMF (2 ×), Fmoc-Gly-OH was coupled onto the resin. The procedure of coupling the monomers was almost the same as in Fmoc SPPS with a slight modification. Fmoc deprotection was carried out with 20% piperidine in DMF three times (1 min × 1, 5 min × 1, 10 min × 1). The monomer and PyBOP were mixed together in two-fold excess over the resin substitution rather than the four-fold excess used in standard Fmoc-SPPS, and the coupling time was prolonged to 5 hrs. The ninhydrin test was used throughout the synthesis to monitor the coupling and deprotection steps. The Fmoc deprotection and

monomer coupling cycles were continued until the last residue. After the final Fmoc deprotection, the resin was washed with DMF (2 ×), DCM (2 ×) and dried in vacuum. The resin was cleaved by 95% TFA, 1% H2O, 2.5% TIS and 1.5% 2-mercaptoethanal for 1 h. The PNA was precipitated by ethyl ether, and then purified by C18 semi-HPLC.

The side chain of cysteine was protected by Trt, and the free amine groups on the bases of the PNA monomers (A, C, G) were protected by Bhoc.

4.4.2.2 Synthesis of peptide thioesters

The peptide thioesters were manually synthesized employing standard t-Boc chemistry. First, Trt-SCH2CH2COOH was coupled onto MBHA resin. The trityl group was removed by treatment with a cocktail containing TFA/TIS/b-mercaptoethanol/DCM (5:2.5:2.5:90). For the coupling of amino acids, Boc-amino acid (4 eq.) and benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) (4 eq.) were dissolved in DCM. DIEA (12 eq.) was added in the solution. After 2 min of activation, the mixture was mixed with resin. The reaction was undertaken for 1.5 h. The coupling reaction was monitored with the Kaiser test. The Boc group was removed by treatment with 30% TFA in DCM for 10 min, followed by 15 min. After sequence assembly, peptide thioesters were cleaved from the resin with a cocktail consisting of TFMSA/TFA/p-cresol/methyl phenyl

sulfide (1:7:1:1) for 1 h. The crude peptides were purified with C18 semi-preparative HPLC.

4.4.2.3 Thioester-mediated chemical ligation between PNA and peptide thioester

The freeze-dried peptide and PNA were dissolved in degassed buffer (100 mM NaH_2PO_4, 50 mM TCEP, pH 8.0), in a ratio of 1 eq of the peptide thioester to 1.5 eq of the cysteinyl-PNA. The reaction components were mixed gently and the reaction tube was incubated at 37 ^oC for 6 h. Progress of the ligation reaction was monitored by C18 analytical HPLC.

4.4.2.4 PNA pairing

The two conjugates (10 μM) were then dissolved in 10 mM sodium phosphate buffer (pH 7.0), and the mixture solution was placed in a 70 ^oC water bath and cooled slowly to room temperature. During this process, annealing between the two PNA occurred on the basis of the Watson-Crick base pairing rules.

4.4.2.5 Oxidation

The mixture was allowed to stand at room temperature for 24 h to allow disulfide bond formation between the two free Cysteine residues via air oxidation. After

desalination by C4 ZipTip, the oxidation products were detected by MALDI-TOF MS.

4.4.2.6 Removal of Acm

The peptide-PNA conjugate **3** was dissolved in 50% MeOH to a final concentration of 1mg/ml. 1.5 eq. of 1M HCl was added, followed by 5-10 eq. of 0.1 M I_2 (in MeOH). After 30 min reaction, excess iodine was quenched by sodium thiosulfate. The final conjugate **4** with two disulfide bonds was purified by HPLC.

4.4.2.7 Trypsin digestion

The freeze-dried conjugate was dissolved in 0.1 M $NaHCO_3$ buffer (pH 8.3), the trypsin was added in a ratio of 1 : 20 to the conjugate. The reaction mixture was allowed to stand at room temperature overnight and was analyzed by HPLC.

References

Anfinsen, C. B. (1973). "Principles that govern the folding of protein chains." Science 181(96): 223-230.

Anfinsen, C. B., E. Haber, et al. (1961). "The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain." Proc Natl Acad Sci U S A 47: 1309-1314.

Annis, I., B. Hargittai, et al. (1994). "Disulfide Bond Formation in peptides." Methods in Enzymology 289: 198-221.

Betz, S. F. (1993). "Disulfide bonds and the stability of globular proteins." Protein Science 2: 1551-1558.

Braasch, D. A. and D. R. Corey (2002). "Novel antisense and peptide nucleic acid strategies for controlling gene expression." Biochemistry 41(14): 4503-4510.

Calderone, C. T. and D. R. Liu (2004). "Nucleic-acid-templated synthesis as a model system for ancient translation." Current Opinion in Chemical Biology 8: 645-653.

Doyon, J. B., T. M. Snyder, et al. (2003). "Highly sensitive in vitro selections for DNA-linked synthetic small molecules with protein binding affinity and specificity." J Am Chem Soc 125(41): 12372-12373.

Egholm, M., O. Buchardt, et al. (1993). "PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogenbonding rules." Nature 365: 566-568.

Gartner, Z. J. and D. R. Liu (2001). "The Generality of DNA-Templated Synthesis as a Basis for Evolving Non-Natural Small Molecules." J. Am. Chem. Soc 123: 6961-6963.

Gartner, Z. J., B. N. Tse, et al. (2004). "DNA-templated organic synthesis and selection of a library of macrocycles." Science 305(5690): 1601-1605.

Gilbert, H. F. (1990). "Molecular and Cellular Aspects of Thiol-Disulfide Exchange." Advances in Enzymology 63: 69-172.

Hanvey, J. C., N. J. Peffer, et al. (1992). "Antisense and Antigene Properties of Peptide Nucleic Acids." Science 258: 1481-1485.

Hyrup, B., M. Egholm, et al. (1994). "Structure-Activity Studies of the Binding of Modified Peptide Nucleic Acids (PNAs) to DNA." J. Am. Chem. Soc 116: 7964-7970.

Kamber, B., A. Hartmann, et al. (1980). "The Synthesis of Cystine Peptides by Iodine Oxidation of S-Trityl-cysteine and S-Acetamidomethyl-cysteine Peptides." Helvetica Chimica Acta 63:

899-915.

Kanan, M. W., M. M. Rozenman, et al. (2004). "Reaction discovery enabled by DNA-templated synthesis and in vitro selection." Nature 431(7008): 545-549.

Kushon, S. A., J. P. Jordan, et al. (2001). "Effect of secondary structure on the thermodynamics and kinetics of PNA hybridization to DNA hairpins." J Am Chem Soc 123(44): 10805-10813.

Li, X. Y. and D. R. Liu (2004). "DNA-Templated Organic Synthesis: Nature's Strategy for Controlling Chemical Reactivity Applied to Synthetic Molecules." Angew. Chem. Int. Ed. 43: 4848-4870.

Momiyama, N., M. W. Kanan, et al. (2007). "Synthesis of acyclic alpha,beta-unsaturated ketones via Pd(II)-catalyzed intermolecular reaction of alkynamides and alkenes." J Am Chem Soc 129(8): 2230-2231.

Naylor, R. and P. T. Gilham (1966). "Studies on Some Interactions and Reactions of Oligonucleotides in Aqueous Solution." Biochemistry 5: 2722-2728.

Nielsen, P. E., M. Egholm, et al. (1991). "Sequence-Selective Recognition of DNA by Strand Displacement wiht a Thymine-Substituted Polyamide." Science 254: 1497-1500.

Nielsen, P. E. and G. Haaima (1997). "Peptide nucleic acid (PNA). A DNA mimic with a pseudopeptide backbone " Chem. Soc. Rev. 26: 73-78.

Orgel, L. E. (1995). "Unnatural Selection in Chemical Systems." Acc. Chem. Res 28: 109-118.

Rozenman, M. M., M. W. Kanan, et al. (2007). "Development and initial application of a hybridization-independent, DNA-encoded reaction discovery system compatible with organic solvents." J Am Chem Soc 129(48): 14933-14938.

Sevier, C. S. and C. A. Kaiser (2002). "Formation and transfer of disulphide bonds in living cells." Nature Reviews Molecular and Cellular Biology 3: 836-847.

Thornton, J. M. (1981). "Disulphide Bridges in Globular Proteins." Journal of Molecular Biology 151: 261-287.

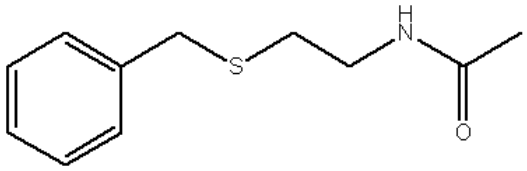
Watson, J. D. and F. H. C. Crick (1953). "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid." Nature 171: 737-738.

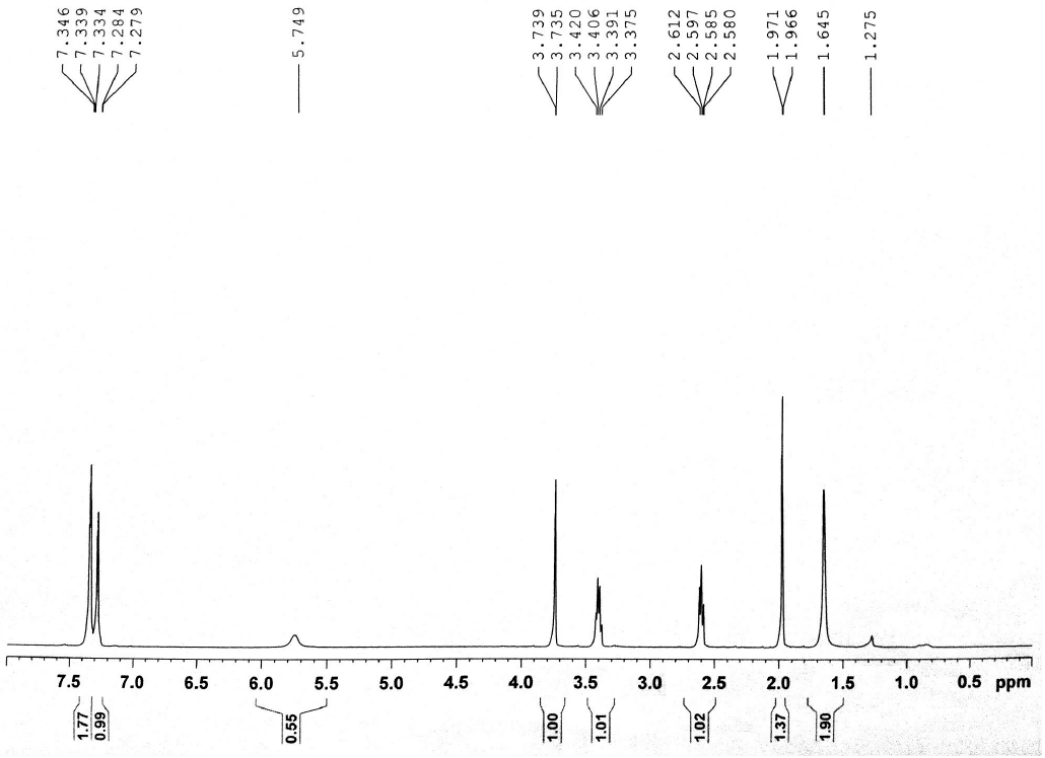
Witt, D. (2008). "Recent developments in disulfide bond formation." Synthesis: 2491-2509.

Wittung, P., P. E. Nielsen, et al. (1994). "DNA-like double helix formed by peptide nucleic acid." Nature 368(6471): 561-563.

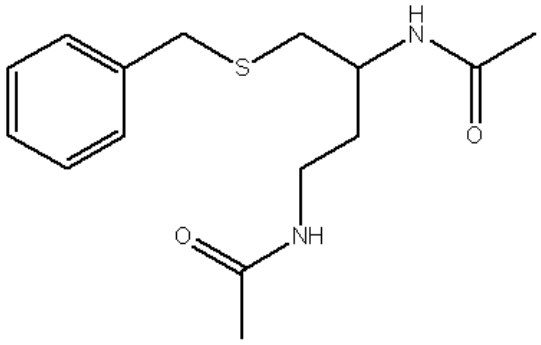
Appendix: NMR spectrum of the products prepared by thiol-ene coupling between benzyl mercaptan and N-vinylacetamide in this study (Chapter 3, pp 63)

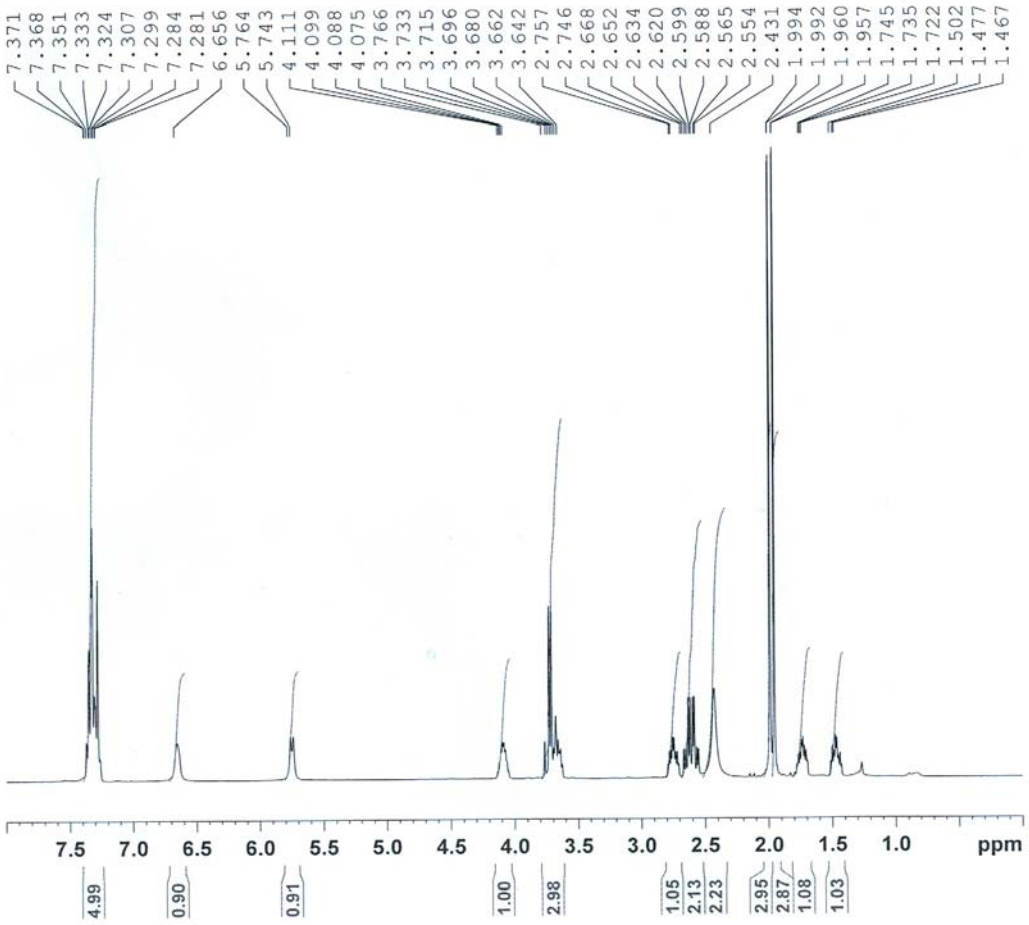
Appendix1. ^{1}H NMR spectrum of the product.





Appendix 2. ^{1}H NMR spectrum of the side product, two NVA were added to the benzyl mercaptan.





Publications

1. **Li, F.**; Alahverdi, A.; Yang, R.; Lua, G.; Zhang, X.; Cao, Y.; Korolev, N.; Nordenskiöld, L.; Liu, C. F., A direct method for site-specific protein acetylation. *Angew. Chem. Int. Ed.* **2011**, 50, 9611-9614.

2. Yi, C. H.; Pan, H.; Seebacher, J.; Jang, I.; Hyberts, S. G.; Heffron, G. J.; Heiden, M. G. V.; Yang, R.; **Li, F**.; Locasale, J. W.; Sharfi, H.; Zhai, B.; Rodriguez-Mias, R.; Luithardt, H.; Cantley, L. C.; Daley, G. Q.; Asara, J. M.; Gygi, S. P.; Wagner, G.; Liu, C. F.; Yuan, J., Metabolic Regulation of Protein N-Alpha-Acetylation by Bcl-xL Promotes Cell Survival. *Cell* **2011**, 146, 607-620.

3. Hou, W.; Zhang, X.; **Li, F.**; Liu, C. F., Peptidyl N,N-bis(2-mercaptoethyl)-amides as thioester precursors for native chemical ligation. *Org Lett* **2011**, 13, (3), 386-9.

4. Zhang, X.; **Li, F**.; Liu, C. F., Synthesis of histone H3 proteins by a thioacid capture ligation strategy. *Chem Commun* **2011**, 47, (6), 1746-8.

5. **Li, F.**; Zhang, X.; Hou, W.; Liu, C. F., Disulfide bond formation directed by Watson-Crick base pairing. *Proceedings of the 11th Chinese International Peptide Symposium*, **2011**, p.116-119.

6. Yang, R.; Pasunooti, K. K.; **Li, F**.; Liu, X. W.; Liu, C. F., Synthesis of K48-linked diubiquitin using dual native chemical ligation at lysine. *Chem Commun* **2010**, 46, (38), 7199-201.

7. **Li, F.**; Liu, C. F., Thioacid capture ligation at valine. *Proceedings of the 21th American Peptide Symposium*, **2009**, p.31-32

9. Yang, R.; Pasunooti, K. K.; **Li, F.**; Liu, X. W.; Liu, C. F., Dual native chemical ligation at lysine. *J Am Chem Soc* **2009**, 131, (38), 13592-3.

10. Zhang, X.; **Li, F**.; Lu, X. W.; Liu, C. F., Protein C-terminal modification through thioacid/azide amidation. *Bioconjug Chem* **2009**, 20, (2), 197-200.

Patent:

Application No.	Title	Date
SG 201103480-8	Site-specific installation of acetyl-lysine analogs into target molecules.	2011