Streamlining peptide synthesis with thiol-free native chemical ligation

2023

Ken Sakamoto

Abbreviations

Ac	acetyl
Ar	aryl
Boc	*tert*-butoxycarbonyl
t-Bu	*tert*-butyl
Dbz	3,4-diaminobenzene
DIPEA	*N,N*-diisopropylethylamine
DMF	*N,N*-dimethylformamide
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
EDT	1,2-ethanedithiol
ESI	electrospray ionization
EtOAc	ethyl acetate
Fmoc	9-fluorenylmethyloxycarbonyl
Gn·HCl	guanidine hydrochloride
GSH	reduced glutathione
GSSG	oxidized glutathione
HPLC	high-performance liquid chromatography
HRMS	high resolution mass spectrometry
MBom	*p*-methoxybenzyloxymethyl
MeDbz	*o*-amino(methyl)aniline
MeIm	1-methylimidazole
MeNbz	*N*-acyl-*N*'-methyl-benzimidazolinone
MESNa	sodium mercaptoethanesulfonate
MPA	3-mercaptopropionic acid
MPAA	4-mercaptophenylacetic acid
MS	mass spectrometry
MTG	methyl thioglycolate
NaH	sodium hydride
NCL	native chemical ligation
Nbz	*N*-acyl-benzimidazolinone
NMR	nuclear magnetic resonance
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
pUb	[Ser(PO_3H_2)^{65}]-ubiquitin
RP-HPLC	reversed-phase high-performance liquid chromatography

SEAlide	*N*-sulfanylethylanilide
SPPS	solid-phase peptide synthesis
TCEP	tris(2-carboxyethyl)phosphine
TFA	trifluoroacetic acid
TFET	2,2,2-trifluoroethanethiol
THF	tetrahydrofuran
TIS	triisopropylsilane
TLC	thin-layer chromatography
Trt	trityl
VA-044	2,2′-azobis[2-(2-imidazolin-2-yl)propane] dihydrochloride

The commonly used one- and three- abbreviations for amino acids:

A	(Ala)	Alanine
C	(Cys)	Cysteine
D	(Asp)	Aspartic acid
E	(Glu)	Glutamic acid
F	(Phe)	Phenylalanine
G	(Gly)	Glycine
H	(His)	Histidine
I	(Ile)	Isoleucine
K	(Lys)	Lysine
L	(Leu)	Leucine
M	(Met)	Methionine
N	(Asn)	Aspargine
P	(Pro)	Proline
Q	(Gln)	Glutamine
R	(Arg)	Arginine
S	(Ser)	Serine
T	(Thr)	Threonine
V	(Val)	Valine
W	(Trp)	Tryptophan
Y	(Tyr)	Tyrosine

Contents

Preface

Chemical synthesis of proteins and peptides offers a broad range of modification control, including specific modifications of side chains and backbone, introduction of non-natural amino acids, and post-translational modifications. Chemical synthesis enables homogeneous production of large quantities of proteins and has contributed significantly to the understanding of protein-related biological phenomena. Solid-phase peptide synthesis (SPPS) is the most common method of peptide synthesis, in which amino acids are coupled stepwise on a solid phase, but this method cannot be directly applied to the synthesis of peptides and proteins with more than 50 amino acids. Longer peptides and proteins are synthesized using ligation strategies between peptide segments prepared by SPPS. The current widely used ligation method in chemical synthesis is native chemical ligation (NCL), which involves peptide thioesters and N-terminal cysteine-containing peptides. NCL has achieved the synthesis of many proteins with practical applications in a wide range of research fields. However, NCL still faces challenges such as the synthesis of peptide thioesters and issues arising from thiol additives used as reaction accelerators.

In this study, I have addressed these challenges and proposed several solutions to the NCL method to make it more practical. In Chapter 1, I discuss the development of versatile synthetic methods for *N*-sulfanylethylanilide (SEAlide) peptides, which can be synthesized by Fmoc-SPPS and function as thioester equivalents, and their application to NCL. In Chapter 2, I investigated alternative reaction accelerators in NCL using peptide thioesters and found that imidazole was a suitable option. Imidazole promotes the NCL reaction as well as conventional additives, while avoiding the risk of additive contamination during high-performance liquid chromatography (HPLC) purification and enabling one-pot desulfurization reaction after NCL. In Chapter 3, I describe that 1,2,4-triazole is a practical additive to facilitate NCL with the thioester equivalent, peptide-MeNbz (MeNbz: *N*-acyl-*N*'-methyl-benzimidazolinone). 1,2,4-Triazole has similar advantages to imidazole in NCL and is more efficient in NCL with peptide-MeNbz.

I believe that the improved NCL methods achieved in this study will be widely used as practical peptide synthesis methods.

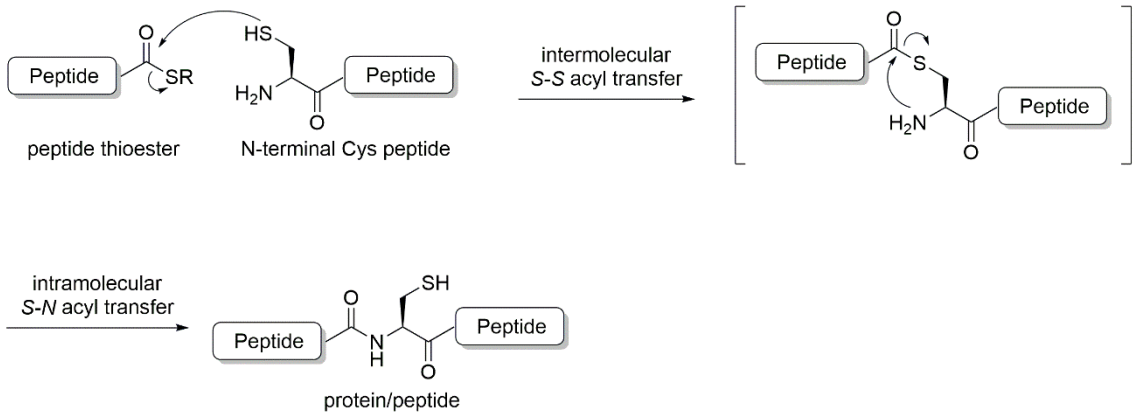
Chapter 1

Versatile synthetic methods of *N*-aminoacyl-*N*-sulfanylethylaniline linker for SEAlide peptide chemistry

Adapted with permission from J. Org. Chem. 2012, 77, 6948−6958. Copyright 2012 American Chemical Society.

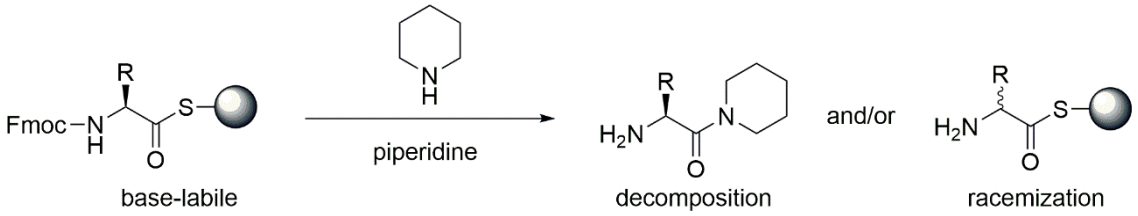
1.1 Introduction

Native chemical ligation (NCL) is a chemoselective reaction that involves the coupling of peptide thioesters with N-terminal cysteine-containing peptides.[1,2] This intermolecular *S–S* acyl transfer reaction and subsequent intramolecular *S–N* acyl transfer reaction (Scheme 1.1) enable the assembly of peptide segments, providing a robust method for the chemical synthesis of proteins. NCL facilitates the production of unprotected proteins under mild conditions in neutral aqueous solutions. To improve the efficiency of NCL, researchers have worked to develop effective techniques for the preparation of peptide thioesters. While Boc-based solid-phase peptide synthesis (Boc-SPPS) has been widely used in this regard, it is unsuitable for the preparation of peptide-thioester with acid-labile groups, and its experimental procedures are more complicated than those of Fmoc-based solid-phase peptide synthesis (Fmoc-SPPS).[3] Therefore, many groups, including ours, have investigated and developed peptide thioester synthesis using Fmoc-SPPS.[4]



Scheme 1.1. Mechanism of native chemical ligation.

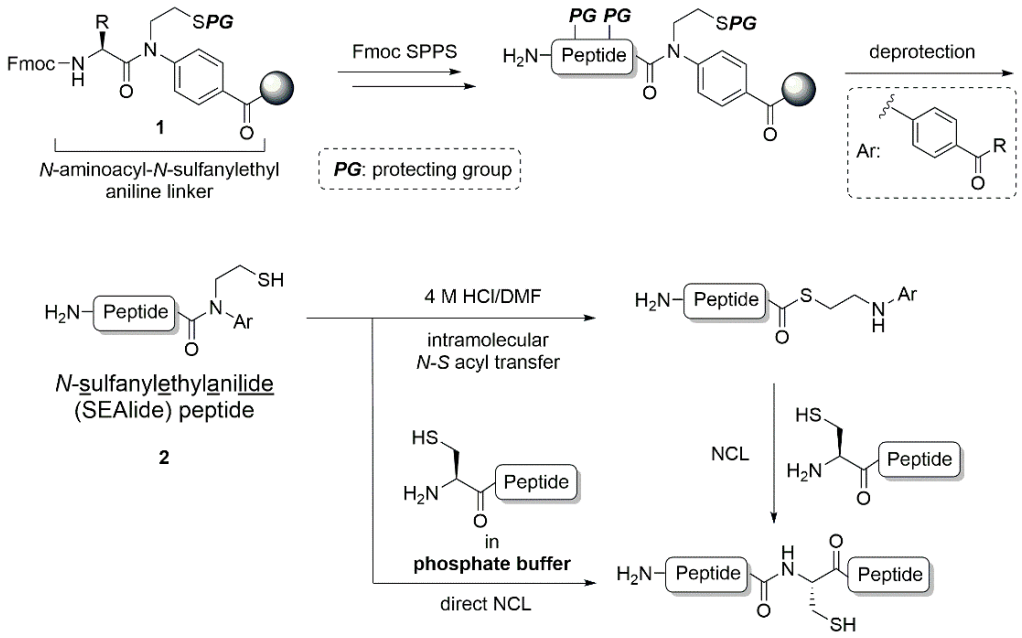
During the synthesis of peptide thioesters using Fmoc-SPPS, the thioester moiety is susceptible to racemization and degradation due to its labile nature under Fmoc deprotection conditions (Scheme 1.2). Various strategies have been explored to overcome this problem, such as the use of improved deprotection reagents, thioesterification of protected peptides in solution, and safety catch linkers that allow thioesterification prior to overall deprotection.[2,4] Of these approaches, peptide-*N*-acyl-benzimidazolinone (Nbz)[5] and peptide-hydrazides[6] have emerged as the most widely employed thioester precursors in recent years. Peptide-Nbz, synthesized by imidation of 3,4-diaminobenzene (Dbz) linker-based peptides prepared through Fmoc-SPPS, can be converted to peptide thioesters by the introduction of thiol additives. Peptide hydrazides, on the other hand, can be generated in Fmoc-SPPS and subsequently oxidized to peptide azides using sodium nitrite under aqueous acidic conditions. The resulting peptide azides can then be converted to peptide thioesters through the addition of thiol additives. An alternative method of accessing peptide thioesters is the *N–S* acyl transfer system, which has been the subject of extensive research. In this system, the peptide is synthesized in the amide form, which is resistant to Fmoc-SPPS, and then converted to the thioester form via intramolecular *N–S* acyl transfer. Several approaches based on *N–S* acyl transfer have been developed, including *N*-alkylcysteine peptides, cysteinyl prolyl ester peptides, and *N*-(2-sulfanylethyl)amide peptides.[4]



Scheme 1.2. Problems in the synthesis of peptide thioesters in Fmoc-SPPS.

Our group focused on the *N–S* acyl transfer reaction for thioester synthesis and developed an *N*-aminoacyl-*N*-sulfanylethylaniline linker **1** (Scheme 1.3).[7] The peptides with this linker, *N*-sulfanylethylanilide (SEAlide) peptides **2**, can be easily synthesized in Fmoc-SPPS and then converted to peptide thioesters by *N–S* acyl transfer under acidic conditions such as 4 M HCl in DMF. In addition, we discovered that SEAlide peptides exist in equilibrium with thioesters in the presence of phosphate salts and can therefore be used directly in the NCL reaction without the need for prior thioester

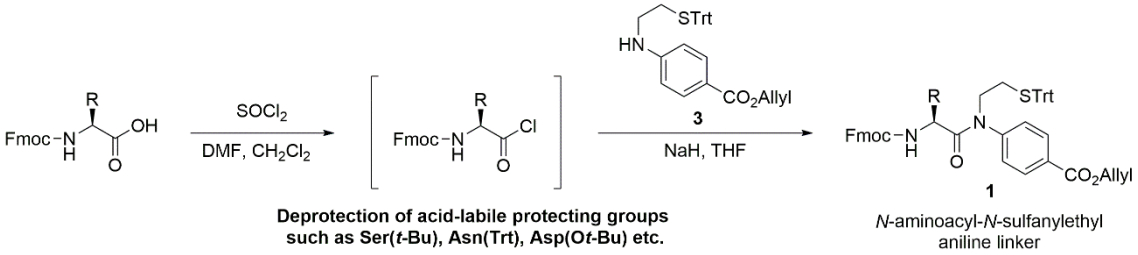
conversion.[8] We have successfully synthesized a 162-residue GM2 activator protein[9] and chemokine CXCL14[10] using SEAlide peptides and demonstrated their efficacy.



Scheme 1.3. SEAlide peptides as thioester precursors and equivalents.

1.2 Synthetic problems in the synthesis of N-aminoacyl-N-sulfanylethylaniline linker

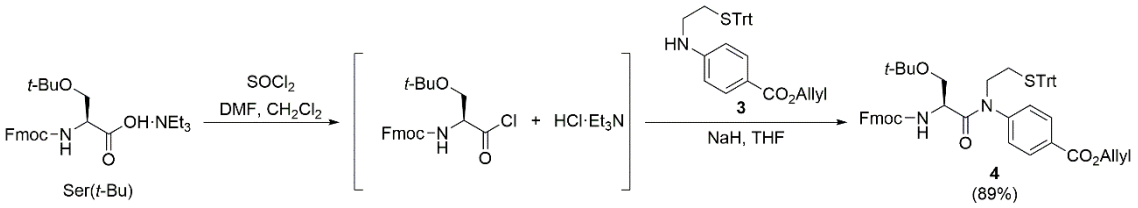
The synthesis of the *N*-aminoacyl-*N*-sulfanylethylaniline linker **1** is a crucial aspect of the early stages of SEAlide peptide preparation. When coupling amino acids to the aniline linker **3**, the aniline linker **3** is activated with sodium hydride (NaH) and then reacted with Fmoc amino acyl chlorides due to its low nucleophilicity (Scheme 1.4). The required acyl chlorides, such as Gly and Ala derivatives, are generated by thionyl chloride ($SOCl_2$), DMF in CH_2Cl_2 ($SOCl_2$-Method). However, this method is not expected to be applicable to amino acids with acid-labile protecting groups, such as *t*-Bu, Boc, or Trt groups (Scheme 1.4).[11] To further increase the versatility of SEAlide peptides, facile synthetic methods for the *N*-aminoacyl-*N*-sulfanylethylaniline linker **1** are crucial. In this investigation, I explored versatile coupling techniques for the aniline linker **3** that can be employed with 20 naturally occurring proteogenic amino acids without side reactions.



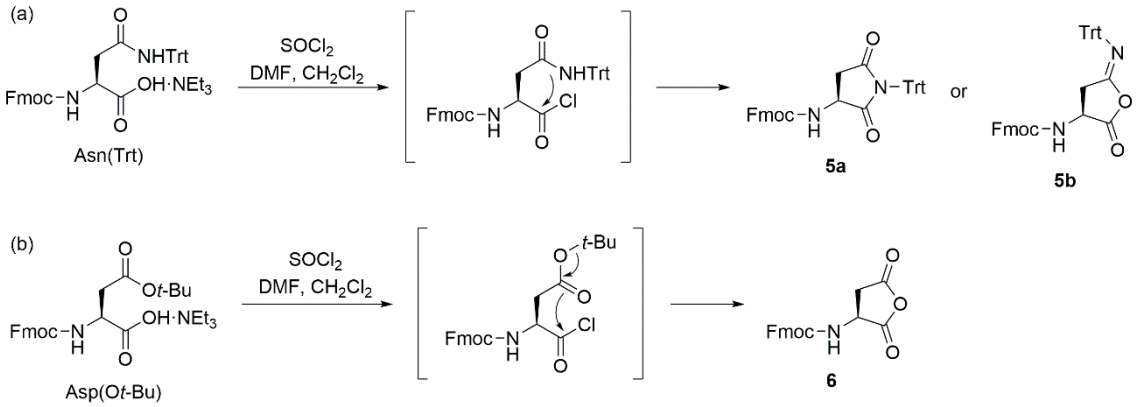
Scheme 1.4. Problems with the introduction of amino acids with acid-labile protecting groups into the aniline linker **3**.

1.3 Acylation reactions using amino acyl chlorides and their problems

In the $SOCl_2$-Method, the aniline linker **3** could be coupled in high yields to aliphatic amino acid derivatives such as Val, Ile, Leu, and Phe, as well as Gly and Ala derivatives (Gly: 94%, Ala: 80%, Val: 98%, Ile: 88%, Leu: 74%, Phe: 83%). However, the yield of Pro derivative was moderate (48%). To extend the applicability of the method to amino acids with acid-labile protecting groups, I developed a modified method in which $SOCl_2$ was used to treat the triethylammonium (Et_3N) salts of carboxylic acids (modified $SOCl_2$-Method).[12] This modification successfully afforded the desired product **4** from Fmoc-L-Ser(*t*-Bu)-OH in 89% yield without loss of the protecting group and with negligible racemization (Scheme 1.5). However, the reactions of Fmoc-L-Asn(Trt)-OH and Fmoc-L-Asp(O*t*-Bu)-OH failed to produce the desired products using this modified method. Activation of the Et_3N salt of Fmoc-L-Asn(Trt)-OH with $SOCl_2$ produced cyclic compound **5a** or **5b**, and that of Fmoc-L-Asp(O*t*-Bu)-OH produced anhydride **6** (Scheme 1.6). These undesirable products were probably due to intramolecular reactions of acid chlorides caused by functional groups of side chains.



Scheme 1.5. Condensation reaction of Ser(*t*-Bu) with aniline linker using Et_3N salt of Fmoc-Ser(*t*-Bu)-OH.

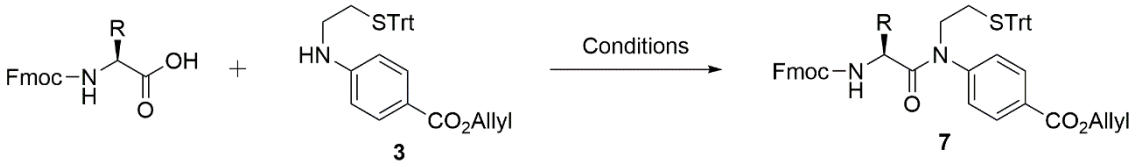


Scheme 1.6. Problems with the introduction of Asn(Trt) and Asp(O*t*-Bu) into the aniline linker **3**.

1.4 Examination of coupling conditions using $POCl_3$ and their applications to 20 naturally occurring proteogenic amino acids

The successful coupling of low nucleophilic aromatic amines with amino acid derivatives by Tesser *et al*. using phosphoryl chloride ($POCl_3$) in pyridine solvent[13] inspired me to apply $POCl_3$ to the synthesis of *N*-aminoacyl-*N*-sulfanylethylaniline linker **7** (Table 1.1). The desired product **7** was obtained in 48% isolated yield by reacting a 2-fold excess of Fmoc-L-Ala-OH and $POCl_3$ with the aniline linker **3** at 4 ºC in pyridine solvent. The use of a 5-fold excess of reagents further improved the reaction yield (75%) (entry 1). However, racemization of Fmoc-L-Ala-OH was confirmed during the coupling reaction (2.3%). This racemization problem was more severe for the coupling of Fmoc-L-Ser(*t*-Bu)-OH by this method (33%) (entry 2). The evaluation method for racemization is described below. To solve this problem, the excess amount of pyridine as a solvent was replaced with the equimolar amount of Et_3N to $POCl_3$ in CH_2Cl_2. The reaction of a 3-fold excess of Fmoc-L-Ser(*t*-Bu)-OH, $POCl_3$, and Et_3N with the aniline linker at 4 ºC for 1 hour in CH_2Cl_2 gave the desired product in moderate yield (51%) and significantly suppressed racemization (0.1%) (entry 3). Furthermore, application of this condition to Fmoc-L-Asn(Trt)-OH also produced the desired product in 88% isolated yield without the formation of the *N*-Trt cyclic imide compound **5** (entry 4). In an attempt to increase the coupling efficiency of Fmoc-L-Ser(*t*-Bu)-OH, I tried adding Fmoc-L-Ser(*t*-Bu)-OH, $POCl_3$, and Et_3N in 5-fold excess; however, even after 26 hours at room temperature, the reaction yield remained at 65% (entry 5).

Table 1.1. Investigation of coupling on the aniline linker **3** by $POCl_3$.

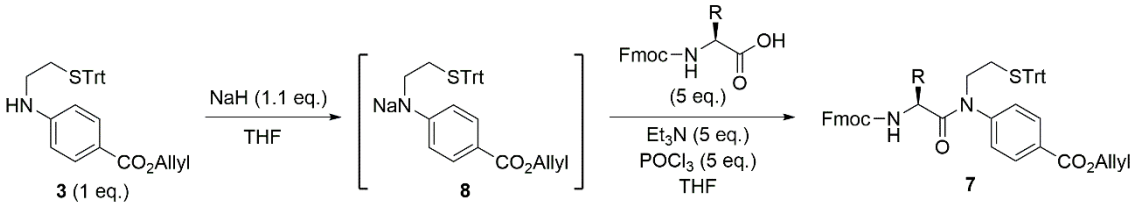


Entry	Fmoc amino acid	Conditions	Isolated yield (%) [Racemization (%)]
1	Ala	A	75 [2.3]
2	Ser(*t*-Bu)	A	48 [33]
3	Ser(*t*-Bu)	B	51 [0.1]
4	Asn(Trt)	C	88
5	Ser(*t*-Bu)	C	65

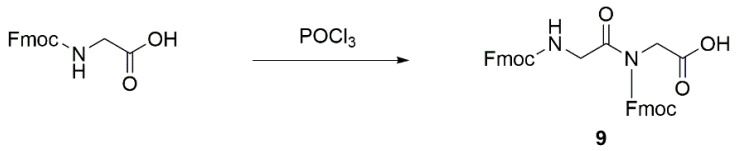
Condition A: **3** (1.0 eq.), Fmoc amino acid (5 eq.), $POCl_3$ (5 eq.), pyridine, 4 °C. **Condition B**: **3** (1.0 eq.), Fmoc amino acid (3 eq.), $POCl_3$ (3 eq.), Et_3N (3 eq.), CH_2Cl_2, 4 °C. **Condition C**: **3** (1.0 eq.), Fmoc amino acid (5 eq.), $POCl_3$ (5 eq.), Et_3N (5 eq.), CH_2Cl_2, room temperature.

With the aim of further improving the yield, NaH was utilized to increase the nucleophilicity of the aniline linker **3**. A reaction of a 5-fold excess of Fmoc-L-Ser(*t*-Bu)-OH, $POCl_3$, and Et_3N in THF was carried out with the activated aniline linker **8** using NaH at room temperature. The reaction proceeded efficiently within 24 hours, resulting in 90% isolated yield with a racemization rate of 1.8% (Table 1.2). Encouraged by these promising results, the same conditions were applied to all 20 naturally occurring proteogenic amino acids. With the exception of Fmoc-L-Pro-OH and Fmoc-L-His(Trt)-OH, coupling of the amino acid derivatives occurred efficiently with over 65% isolated yield as shown in Table 1.2. Although the coupling of Fmoc-L-Pro-OH resulted in low yield, further investigation was discontinued due to the unsuitability of C-terminal Pro-thioester peptide for the NCL reaction.[14] The coupling of Fmoc-L-His(Trt)-OH is described below. In addition, the coupling of Fmoc-Gly-OH and Fmoc-L-Ala-OH was less efficient than the method using $SOCl_2$. Investigation of the reaction with Fmoc-Gly-OH revealed the formation of the dipeptide **9**, which is formed by acylation of the carbamoyl nitrogen (Scheme 1.7). This compound was also detected on TLC analysis of the reaction by mixing Fmoc-Gly-OH, $POCl_3$, and Et_3N in CH_2Cl_2. These results suggest that Gly and Ala, which have a less sterically hindered carbamoyl nitrogen, may be more susceptible to this side reaction, resulting in low yields.[15] Sterically less-hindered Gly and Ala derivatives can be synthesized in high yields using the $SOCl_2$ method.

Table 1.2. Summary of coupling of Fmoc amino acids with aniline linker **3**.



Fmoc amino acid	Isolated yield (%) [Racemization (%)]	Fmoc amino acid	Isolated yield (%) [Racemization (%)]
Asp(O*t*-Bu)	66	Val	87
Asn(Trt)	89	Met	80
Thr(*t*-Bu)	84	Ile	87
Ser(*t*-Bu) (r.t.)	90 [1.8]	Leu	88
Ser(*t*-Bu) (4 ˚C)	91 [0.4]	Tyr(*t*-Bu)	86
Glu(O*t*-Bu)	72	Phe	88
Gln(Trt)	93	His(τ-Trt)	41 [23]
Pro	39	His(π-MBom)	88 [0.4]
Gly	65	Lys(Boc)	80
Ala	67	Arg(Pbf)	95
Cys(Trt)	72 [Not detected]	Trp	92

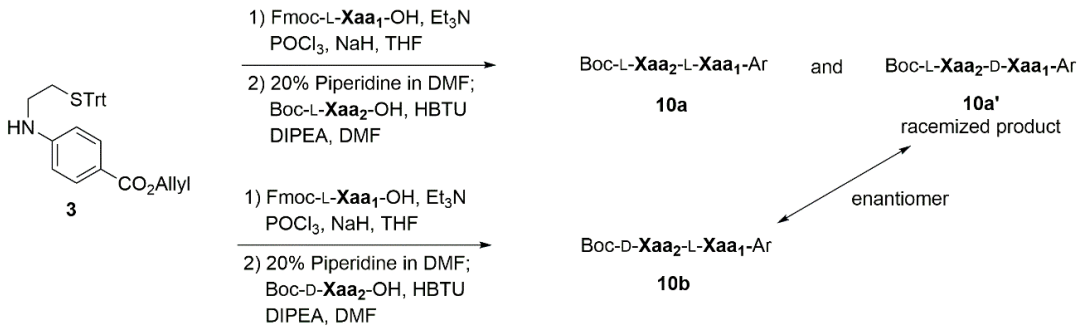


9

Scheme 1.7. Dipeptide **9** produced by acylation of the carbamoyl nitrogen in Fmoc-Gly-OH.

Given the susceptibility of Cys, Ser, and His to racemization in coupling reactions, we investigated the degree of racemization of Fmoc-L-Cys(Trt)-OH, Fmoc-L-Ser(*t*-Bu)-OH, and Fmoc-L-His(Trt)-OH during their coupling with the aniline linker (Table 1.2). To evaluate racemization, the resulting linkers were converted to the dipeptide linkers **10** by Fmoc deprotection followed by condensation with Boc-L-Leu-OH or Boc-D-Leu-OH. The Boc-D-Leu-L-Xaa derivative **10b**, the major component produced by coupling with Boc-D-Leu-OH, corresponds to the enantiomer of Boc-L-Leu-D-Xaa **10a'** derived from the racemized form. The HPLC analysis showed that Cys(Trt) did not undergo racemization (Table 1.2, Figure 1.1). Although the racemization of Ser(*t*-Bu) was 1.8% at room temperature, the reaction at 4 °C successfully suppressed it to 0.4% with a similar yield. However, severe racemization

(23%) occurred during the condensation of His(Trt). In recent years, Hibino *et al.* developed a *p*-methoxybenzyloxymethyl (MBom) group as a side chain protecting group of His to suppress racemization.[16] The MBom group of His(MBom) protects the π-position nitrogen of the imidazole ring, which promotes racemization, while the Trt group of His(Trt) protects the other nitrogen (see Experimental Section for structures). As expected, the coupling of Fmoc-L-His(MBom)-OH proceeded well, yielding the desired product in 88% isolation yield with negligible racemization (<0.4%) (Table 1.2, Figure 1.1).



Scheme 1.8. Preparation of diastereomeric mixtures for racemization study.

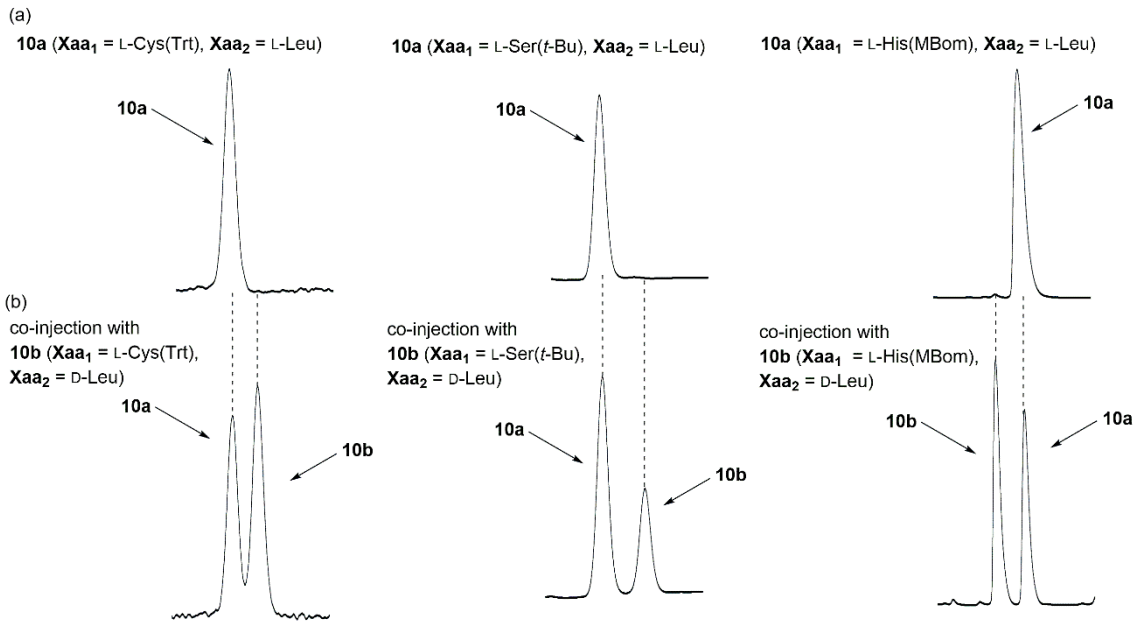
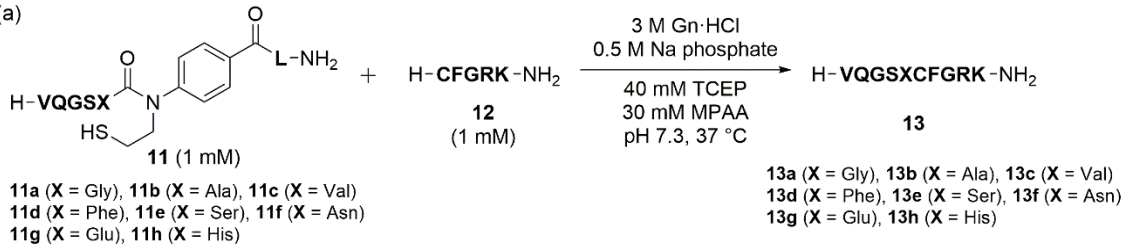


Figure 1.1. HPLC data of racemization study of C-terminal amino acids in the coupling process. (a) **10a** (**Xaa1** = L-Cys(Trt), L-Ser(*t*-Bu) or L-His(MBom), **Xaa2** = L-Leu); (b) co-injection of **10a** and **10b** (**Xaa1** = L-Cys(Trt), L-Ser(*t*-Bu) or L-His(MBom), **Xaa2** = D-Leu).

1.5 NCL reaction using SEAlide peptides with C-terminal diversity

In this study, I evaluated the efficacy in NCLs of SEAlide peptides derived from diverse *N*-aminoacyl-*N*-sulfanylethylaniline linkers, whose synthetic approaches were established in the previous Chapter. Dawson *et al.* investigated the effect of C-terminal amino acids on NCL reaction rates and ranked their reactivity as follows: Gly, Cys, His > Phe, Met, Tyr, Ala, Trp > Asn, Ser, Asp, Gln, Glu, Lys, Arg > Leu, Thr > Val, Ile > Pro.[17] Based on this ranking, I examined the NCLs of SEAlide peptides **11**, containing Gly, His, Phe, Ala, Asn, Ser, Glu, and Val at the C-terminus (Figure 1.2). SEAlide peptides **11** and the N-terminal cysteine-containing peptide **12** were incubated at 37 °C in a 3.0 M guanidine(Gn)·HCl-0.5 M sodium phosphate buffer (pH 7.3) with 40 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and 30 mM 4-mercaptophenylacetic acid (MPAA)[18]. As a result, the ligation products **13** were obtained with high purity within 24-36 hours, except for the Glu and Val derivatives (Figure 1.2, 1.3). Although the Glu and Val reactions were incomplete within 24 hours, the ligation products were obtained in high yields by extending the reaction time to 48 hours for the Glu derivative and 72 hours for the Val derivative. As previously reported, the reaction rate varied depending on the type of C-terminal amino acid. Next, I confirmed the epimerizations of the SEAlide peptides containing Ala, Ser, and His at the C-terminus. The authentic epimerization products, H-**VQS**-D-**X-CFGRK**-NH_2, were synthesized separately using Fmoc-SPPS. The degree of epimerization of the SEAlide peptides was estimated from the peak area of the HPLC analysis: 1.0% for Ala, 2.1% for His, and 6.7% for Ser (Figure 1.2). The susceptibility of Ser-sites to epimerization is also observed in other NCL reactions, such as cysteinyl-prolyl-ester ligation.[19] Reaction conditions that can suppress Ser-site epimerization have not yet been found.

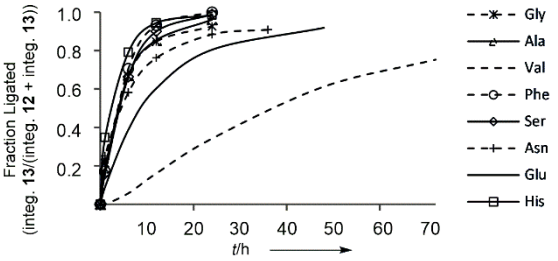
(a)



11a (**X** = Gly), **11b** (**X** = Ala), **11c** (**X** = Val)
11d (**X** = Phe), **11e** (**X** = Ser), **11f** (**X** = Asn)
11g (**X** = Glu), **11h** (**X** = His)

13a (**X** = Gly), **13b** (**X** = Ala), **13c** (**X** = Val)
13d (**X** = Phe), **13e** (**X** = Ser), **13f** (**X** = Asn)
13g (**X** = Glu), **13h** (**X** = His)

(b)



(c)

X	Reaction time (h)	Isolated yield (%)	Epimerization (%)
13a (**X** = Gly)	24	68	-
13b (**X** = Ala)	24	61	1.0
13c (**X** = Val)	72	60	n.d.
13d (**X** = Phe)	24	77	n.d.
13e (**X** = Ser)	24	70	6.7
13f (**X** = Asn)	24	67	n.d.
13g (**X** = Glu)	48	70	n.d.
13h (**X** = His)	24	59	2.1

Figure 1.2. Investigation of the NCL reactions of various SEAlide peptides **11** with N-terminal cysteine-containing peptide **12**. (a) NCL reaction scheme and conditions. (b) Kinetic analysis of NCL reactions: the fraction ligated was determined by HPLC separation and integration of ligated peptides **13** (integ. **13**) detected at 220 nm as a fraction of the sum of the unreacted N-terminal cysteinyl peptide (integ. **12**) and integ. **13**. (c) Table summarizing the results of NCL reactions: n.d. = not determined.

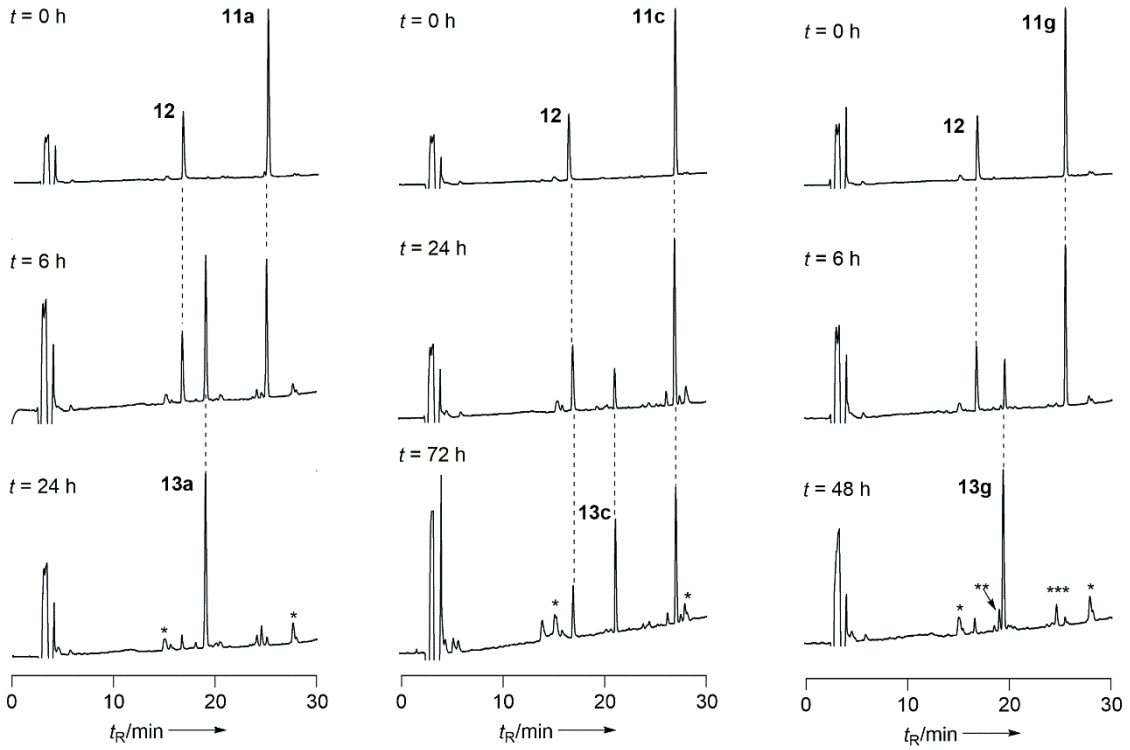


Figure 1.3. Representative HPLC data of NCL using SEAlide peptides **11** (**X** = Gly **11a**, L-Val **11c** or L-Glu **11g**). *The detected impurity was non-peptidic. **The structure of the impurity was deduced to be an isomeric form at Glu, as its observed mass was the same as that of **13g**.[20] ***Desulfurized derivative of SEAlide peptide **11g**.

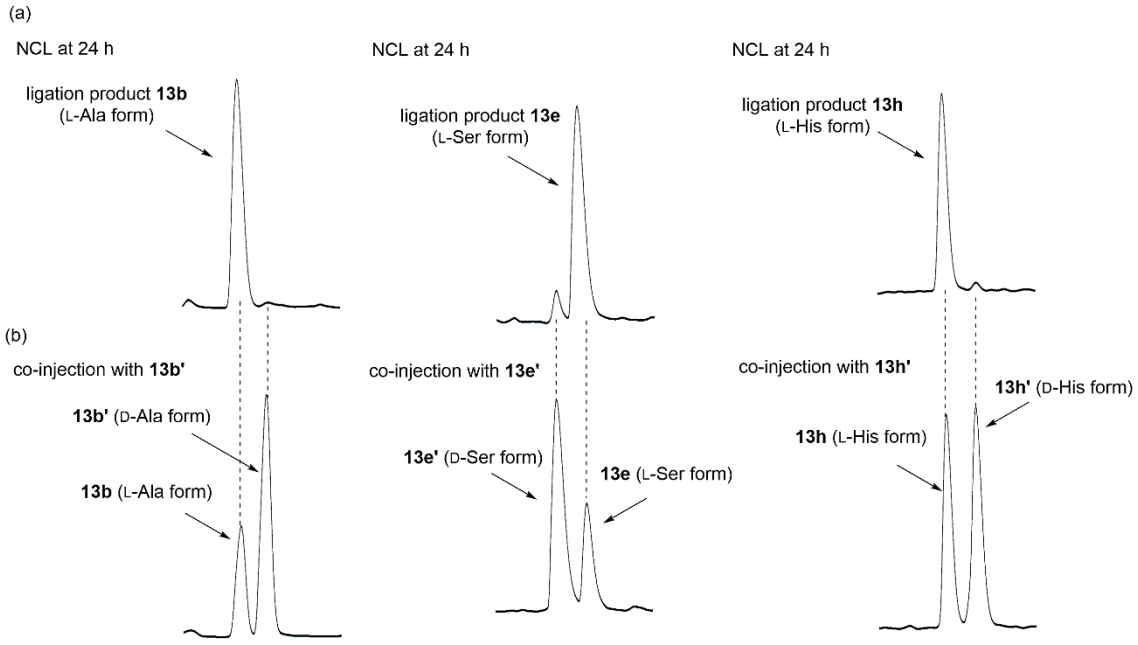


Figure 1.4. HPLC data of epimerization study of C-terminal amino acids of SEAlide peptides **11**. (a) NCL at 24 hours with C-terminal Ala, Ser or His-SEAlide; (b) co-injection of **13** (H-**VQGSX-CFGRK**-NH_2: **X** = L-Ala **13b**, L-Ser **13e** or L-His **13h**), and **13'** (H-**VQGS**-D-**X-CFGRK**-NH_2).

1.6 Problems with C-terminal Asp-containing SEAlide peptide

Deprotection of the C-terminal Asp-containing SEAlide (Asp-SEAlide) peptide resin with TFA produced two compounds with the same mass as the target compound **11i** (**X** = Asp), which eluted separately during HPLC analysis (Figure 1.5, Scheme 1.9). The later eluted compound was eliminated by treatment with sodium mercaptoethanesulfonate (MESNa) under neutral conditions, which resulted in the formation of the aniline linker **15** as a degradation product (Scheme 1.9). Conversely, the earlier eluted compound was not affected by MESNa treatment. This later eluted compound is presumed to be the thioester-type SEAlide peptide **14**, which disappeared due to thiolysis or anhydride formation under MESNa treatment conditions. To isolate the earlier eluted compound, amide-type SEAlide peptide **11i**, I used reversed-phase HPLC (RP-HPLC) with 0.1% TFA, and the purity of the fractionated solution was confirmed by analytical HPLC to contain a single amide-type SEAlide peptide **11i**. However, after lyophilization, a mixture of amide-type **11i** and thioester-type **14** SEAlide peptides was identified. It is noteworthy that I did not observe any *N–S* acyl transfer in other amino acid-containing SEAlide peptides during deprotection and

lyophilization. The results suggest that Asp-SEAlide peptide **11i** is readily converted to thioester-type **14** under acidic conditions. I hypothesized that the specific *N*–*S* acyl transfer is facilitated by acid-base catalysis through the 7-membered ring formation of Asp-carboxyl group, which remains undissociated under acidic conditions (Scheme 1.10). Several experiments confirm this hypothesis by demonstrating cleavage of the Asp-Pro bond under acidic conditions.[21] Therefore, I predicted that HPLC purification under neutral conditions, where the carboxyl group of Asp exists in its ionic state, would yield a single amide-type SEAlide peptide **11i**. By purifying a mixture of amide-type SEAlide peptide **11i** and thioester-type SEAlide peptide **14** in a 10 mM ammonium acetate (pH 6.1)-acetonitrile system and lyophilizing it, I successfully obtained pure amide-type SEAlide peptide **11i**.

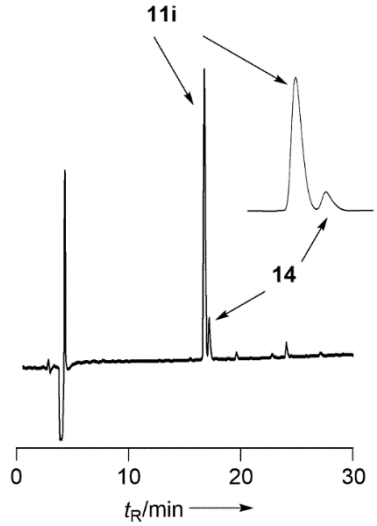
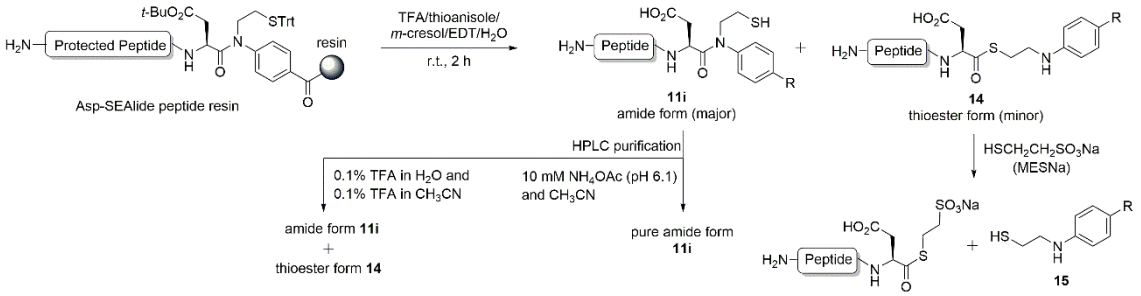
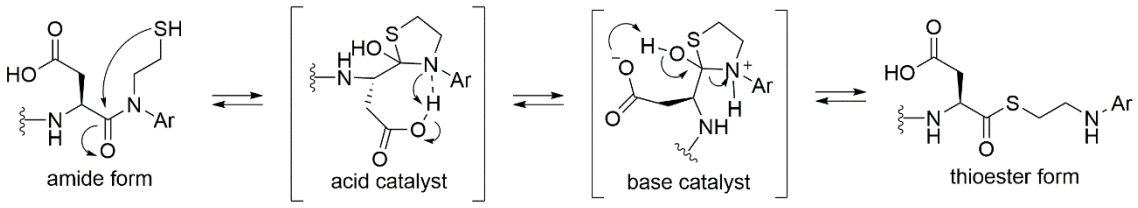


Figure 1.5. HPLC data after deprotection of the C-terminal Asp-containing SEAlide.



Scheme 1.9. Overview of *N*–*S* acyl transfer of Asp-SEAlide peptide proceeding under acidic conditions.



Scheme 1.10. Proposal of a feasible reaction mechanism for thioester formation from Asp-SEAlide peptide **11i** in acidic environment.

Application of the resulting C-terminal Asp-containing SEAlide peptide **11i** to the NCL reaction produced a mixture of α-peptide and β-peptide (Figure 1.6). A similar result has been documented in NCL reactions using C-terminal Asp-containing peptide thioesters (Scheme 1.11).[20] However, the NCL reaction of the Asp-Cys moiety remains an unsettled matter.

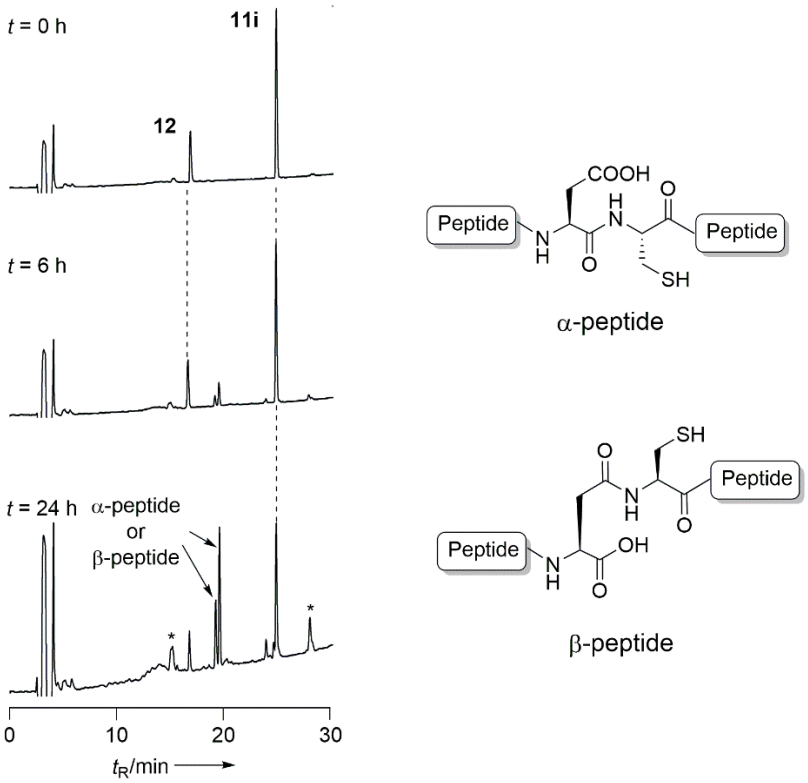
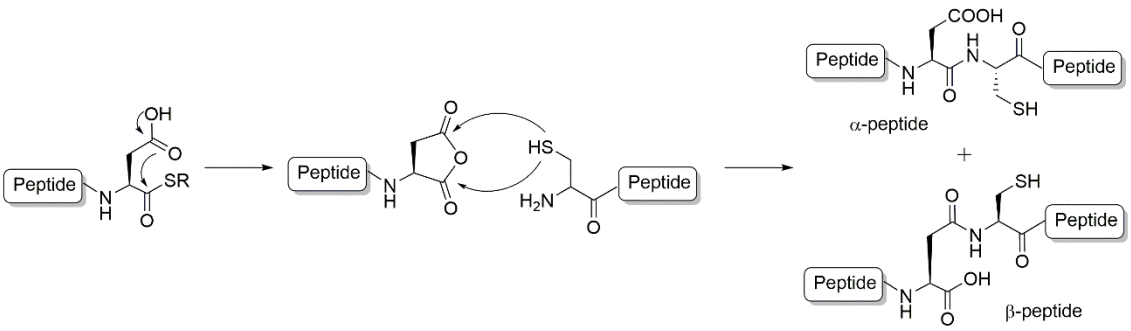


Figure 1.6. HPLC data of NCL with C-terminal Asp-containing SEAlide peptide **11i**. *Non-peptidic impurity.



Scheme 1.11. Reaction mechanism to produce a mixture of α-peptide and β-peptide from C-terminal Asp-thioester.

1.7 Conclusion

In this Chapter, I have developed the practical synthetic techniques for *N*-aminoacyl-*N*-sulfanylethylaniline linkers. The effectiveness of these methods has been demonstrated by the successful coupling of 20 naturally occurring proteogenic amino acids to the aniline linker. In addition, SEAlide peptides with various C-terminal amino acids have been synthesized using these techniques and were effectively used as peptide thioester equivalents in the NCL. Otaka *et al.* have demonstrated the usefulness of SEAlide peptides in NCL chemistry by performing one-pot, multi-fragment condensation under kinetic conditions. Therefore, the development of this technology provides significant opportunities for the use of SEAlide peptides. For example, a 162-residue glycoprotein was chemically synthesized using SEAlide peptides synthesized by the methods established in this study.[9] I hope that the continued use of SEAlide peptides in protein synthesis will contribute to the advancement of protein research in the future.

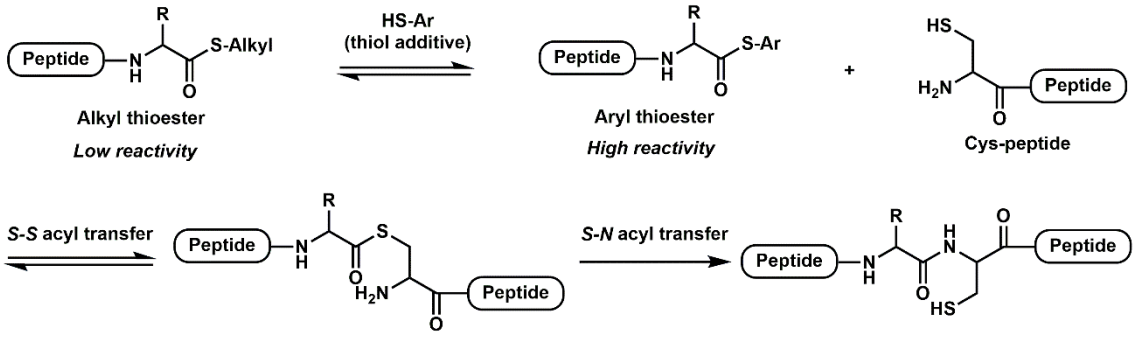
Chapter 2

Streamlining peptide synthesis with imidazole facilitating native chemical ligation

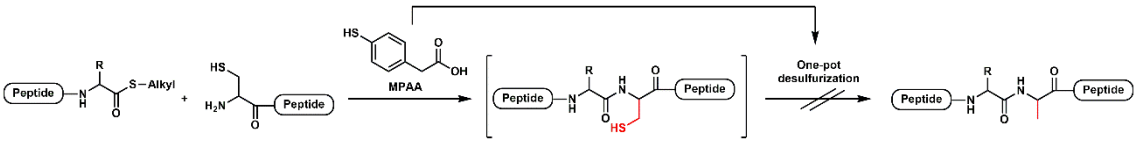
2.1 Importance of additive selection in NCL

NCL reactions proceed through intermolecular *S*–*S* acyl transfer between thioesters and the side chains of Cys residues, followed by intramolecular *S*–*N* acyl transfer. Alkyl thioesters are commonly used in NCLs due to their easy preparation and long-term stability, but their low reactivity makes intermolecular *S*–*S* acyl transfer the rate-limiting step.[2] Thiol additives are necessary to accelerate the reaction by forming highly reactive thioesters in situ, with 4-mercaptophenylacetic acid[18] (MPAA) being a popular choice due to its low odor, high solubility, and strong accelerating effect (Scheme 2.1). However, MPAA has similar elution properties with peptides in reversed-phase HPLC (RP-HPLC), which means that MPAA has the potential to co-elute with the target peptides.[22] This co-elution presents a significant challenge, especially in purification processes involving large amounts of MPAA. The advent of the radical desulfurization reaction, which is capable of converting Cys to Ala, has broadened the scope of NCL beyond its traditional Cys-containing ligation sites.[23] This procedure eliminates the need for Cys at the ligation site, thereby enabling ligation at diverse sites, including Ala, Phe, and many other amino acids. The combination of NCL and the desulfurization reaction has proven to be a popular and effective approach, leading to the synthesis of numerous bioactive peptides. The development of a one-pot NCL-desulfurization technique would be an immensely valuable advancement due to the time-consuming purification process. However, the use of aryl thiol additives, such as MPAA, in NCL can impede radical desulfurization reactions, making it unsuitable for such one-pot reactions (Scheme 2.2).[24] There have been reports of techniques to remove MPAA after NCL or "thiol-additive-free NCL" using specifically designed peptide-thioesters; however, these procedures are not simple.[25] In addition, direct production of MPAA-thioesters via SPPS (solid-phase peptide synthesis) remains challenging.[26] As a result, researchers worldwide have explored alternative additives to MPAA, and 2,2,2-trifluoroethanethiol (TFET)[27] and methyl thioglycolate (MTG)[28] have emerged as promising options (Table 2.1). Despite their ability to enhance the NCL reaction without impeding desulfurization, they are burdened with issues such as

unpleasant odors and the high volatility of TFET. In this study, I aimed to identify additives that could address these concerns, and my results show that imidazole[29] is a better choice.



Scheme 2.1. Reaction mechanism of NCL promoted by aryl thiols.



Scheme 2.2. Inhibition of desulfurization reaction by MPAA.

Table 2.1. Comparison of additives for NCL.

Substrate	pK_a	Desulfurization	Odor
MPAA	6.6	unamenable	odorless
2,2,2-trifluoroethanethiol (TFET)	7.3	amenable	malodor
methyl thioglycolate	7.9	amenable	malodor
imidazole	7.0*	amenable	odorless

* pK_a of the conjugate acid

2.2 Imidazole as a non-thiol additive for NCL

The alkyl thioesters, such as MPA (3-mercaptopropionic acid) thioesters (pK_a = 10.6),[30] which have low activity, can be converted to highly active aryl thioesters using the frequently used additive MPAA (pK_a = 6.6),[18] as shown in Scheme 2.1. The reactivity of the active ester increases with the acidity of the leaving group, but this also increases the likelihood of side reactions such as hydrolysis. As an effective additive, MPAA is likely to have the appropriate pK_a to produce moderate stability and reactivity in its esters. Therefore, additives with a pK_a similar to that of MPAA would promote the NCL reaction in the same manner as MPAA, and imidazole (pK_a of the conjugate acid =

7.0)[31] would be expected to react with the alkyl thioester to form acyl imidazole, followed by the coexisting acyl imidazolinium reacting similarly with the MPAA ester. Furthermore, as a non-thiol additive, imidazole was not expected to inhibit the desulfurization reaction. In addition, the high polarity of imidazole would avoid inhibiting the RP-HPLC purification of the peptide.

To test this hypothesis, imidazole was compared with MPAA, TFET, and MTG using the model peptides peptide thioester **16a** and cysteine peptide **17** (Figure 2.1A). The Imidazole-NCL conditions (2.5 M imidazole, 6.0 Gn·HCl, 30 mM TCEP, pH 7.1, rt) successfully reacted at a rate comparable to 50 mM MPAA. Although the hydrolysis of the alkyl thioester was slightly faster than the MPAA method, no other significant by-products were observed. Furthermore, the reactivity of Imidazole-NCL was more efficient than that of TFET and MTG, the utility of which has been previously reported (Figure 2.1A). The reaction rate was also found to be dependent on the imidazole concentration. Increasing the imidazole concentration enhanced the reaction rate (Figure 2.1B). Since the reaction rate was similar at 2.5 M and 5 M imidazole, 2.5 M imidazole was used as the standard condition. The accelerating effect of pyrazole, an azole compound with a different pK_a (pK_a = 14.2, pK_a of the conjugate acid = 2.48), was found to be poor, as expected, and almost identical to that of no additive (Figure 2.1C). These results indicate that imidazole can efficiently promote the NCL reaction due to its appropriate pK_a.

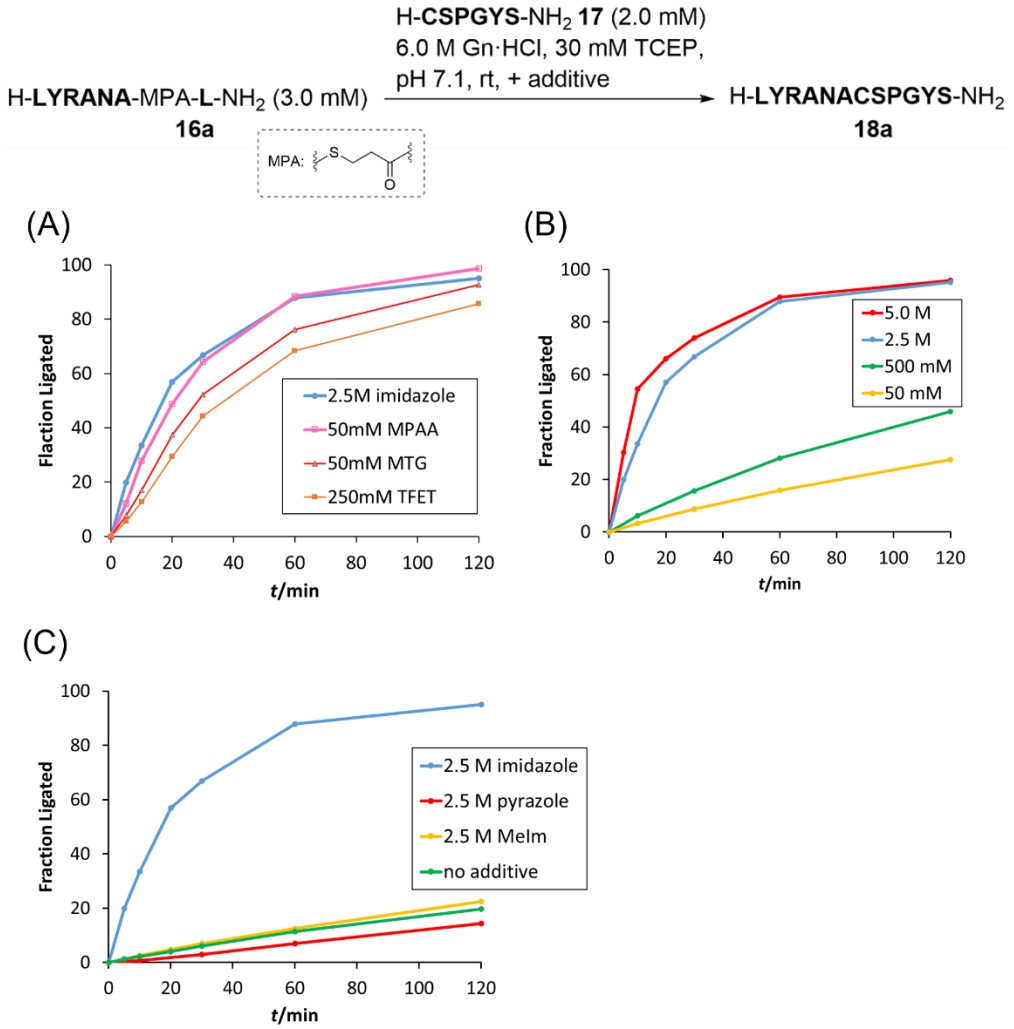
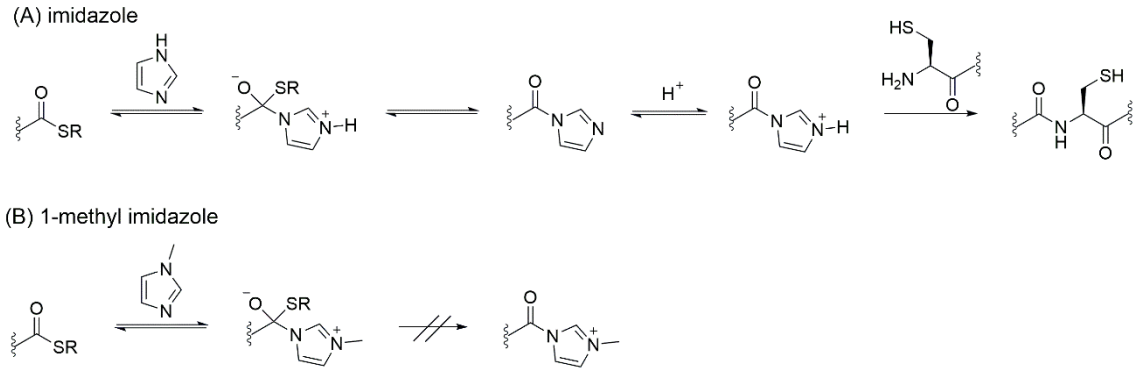


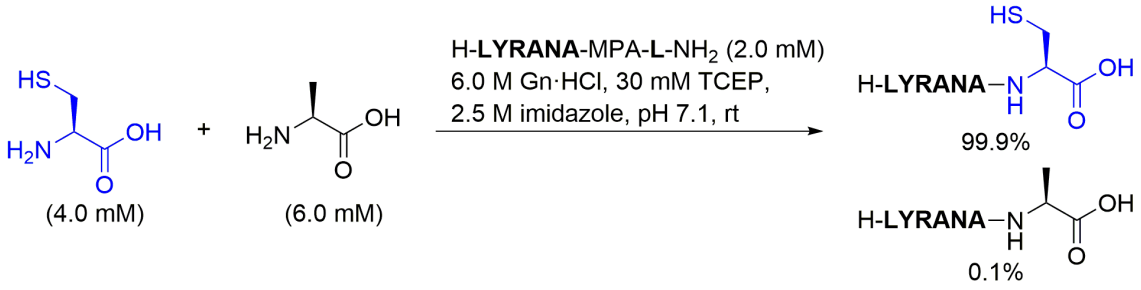
Figure 2.1. Kinetic analysis of NCL between H-**LYRANA**-MPA-**L**-NH_2 (**16a**) and H-**CSPGYS**-NH_2 (**17**): (A) comparison of imidazole and reported NCL additives (the concentrations recommended in the report were used), (B) comparison of the concentration dependence of imidazole, and (C) comparison with various azole additives. The fraction ligated was determined using the HPLC peak areas at λ=280 nm.

1-Methylimidazole (MeIm) is an azole compound with pK_a similar to imidazole, but with a Me group on the nitrogen atom (pK_a of the conjugate acid = 7.1). However, as shown in Figure 2.1C, MeIm failed to promote the NCL reaction. This is probably because MeIm cannot form an acylimidazole intermediate due to the presence of the methyl group. Therfore, this result strongly supports the notion that the Imidazole-NCL is facilitated by the involvement of the acylimidazole form (Scheme 2.3).



Scheme 2.3. Possible mechanism of NCL using imidazole (A) and 1-methylimidazole (B).

To investigate the chemoselectivity of Imidazole-NCL, a mixture of cysteine and alanine was subjected to the reaction with a thioester peptide. As a result, 99.9% of the reaction products were obtained with Cys, confirming the Cys-selectivey of Imidazole-NCL. (Scheme 2.4).



Scheme 2.4. Chemoselectivity of Imidazole-NCL with a mixture of cysteine and alanine.

2.3 Application of Imidazole-NCL to peptide thioesters with various C-terminal amino acids

With these results, I proceeded to assess the performance of Imidazole-NCL using peptide thioesters featuring diverse C-terminal amino acids. Peptide thioesters with Gly, Phe, Leu, and Thr at the C-terminus enabled the reaction to proceed in a few hours without the formation of by-products (Figure 2.2). The reaction kinetics were dependent on the type of C-terminal amino acid, a phenomenon that had also been observed in NCL with thiol additives.[17] In particular, Val-thioesters exhibited sluggish reaction rates, necessitating the use of 5.0 M imidazole to accelerate the process to a

degree comparable to that achieved with the widely used MPAA (100 mM) (Figure 2.2). Furthermore, the racemization of Imidazole-NCL was Ala: 2% (2 h), Phe: 2% (1 h), Val: 3% (5 h) (Figure 2.3).

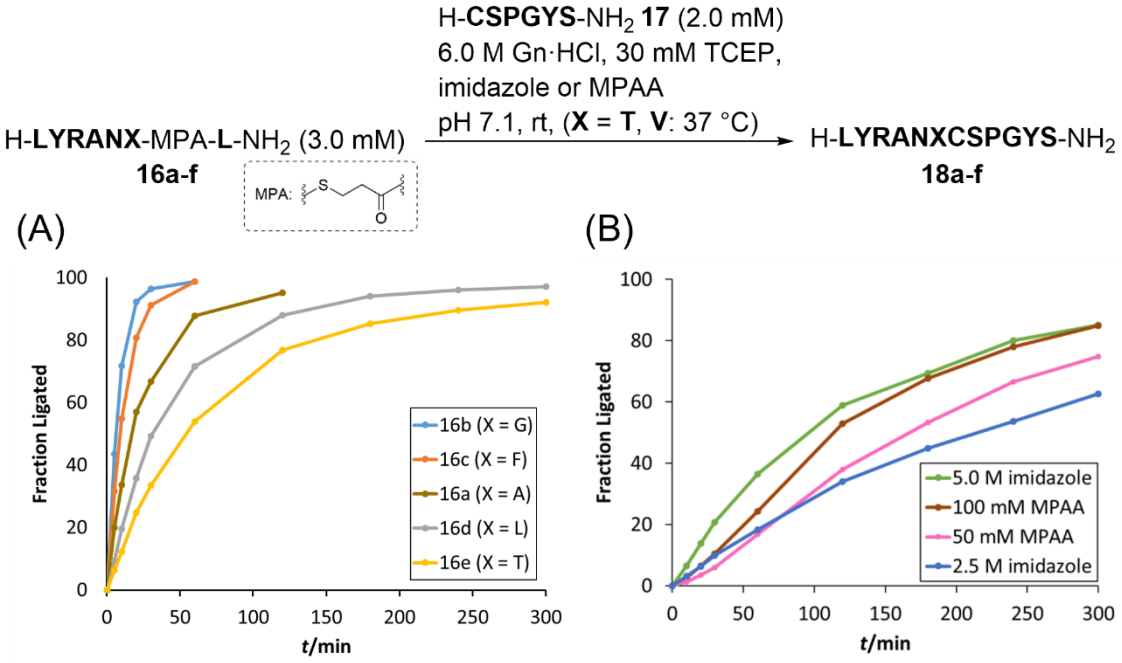


Figure 2.2. Kinetic analysis of (A) NCL with 2.5 M imidazole using peptide thioesters (H-**LYRANX**-MPA-**L**-NH_2: **16a**, **X=A**; **16b**, **X=G**; **16c**, **X=F**; **16d**, **X=L**; **16e**, **X=T**), and (B) comparison of imidazole and MPAA in the case of thioester **16f** (**X=V**). The fraction ligated was determined using the HPLC peak areas at λ=280 nm.

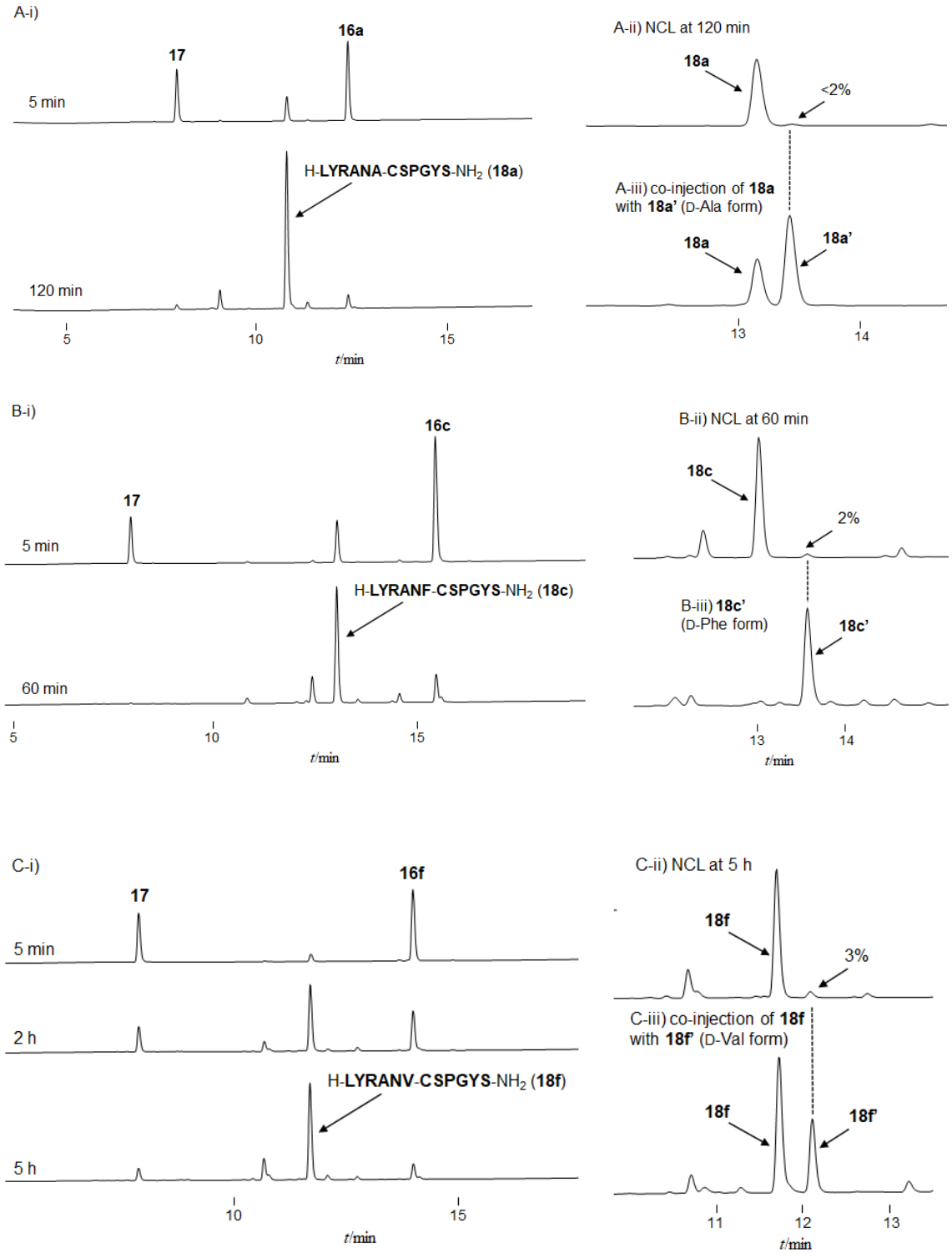


Figure 2.3. The HPLC traces show the representative Im-NCL between peptide thioesters (H-**LYRANX**-MPA-**L**-NH_2: **16a**, **X** = **A** (for A); **16c**, **X** = **F** (for B); **16f**, **X** = **V** (for C)) and Cys-peptide (H-**CSPGYS**-NH_2, **17**). The rate of epimerization was confirmed by comparing with the authentic corresponding epimer (ii/iii).

2.4 Synthesis of adiponectin using Imidazole-NCL

Having found that Imidazole-NCL proceeded efficiently with the model sequence, I investigated the synthesis of adiponectin, an adipocytokine known for its ability to enhance insulin sensitivity and characterized by four distinct domains.[32] I attempted to synthesize adiponectin(19-107), which contains a variable region with the collagen component (Adiponectin(19-107): H-**ET**^{20} **TTQGPGVLLP**^{30} **LPKGACTGWM**^{40} **AGIO*GHO*GHN**^{50} **GAO*GRDGRDG**^{60} **TPGEKGEKGD**^{70} **PGLIGPKGDI**^{80} **GETGVPGAEG**^{90} **O*RGFPGIQGR**^{100} **KGEPGEG**-OH, **O*** = (2S, 4R)-4-hydroxyproline). Using SPPS, I synthesized adiponectin(19-35)-thioester **19** and cysteine peptide(36-107) **20** as fragments for Imidazole-NCL (Figure 2.4). The Imidazole-NCL reaction was completed within 3 hours and yielded no significant by-products other than the hydrolyzed peptide of thioester **19**. The crude solution containing a large amount of imidazole was then purified by RP-HPLC, and the desired product **21** was successfully isolated in 85% yield. The results clearly demonstrated that imidazole does not interfere with RP-HPLC purification, as was initially hypothesized.

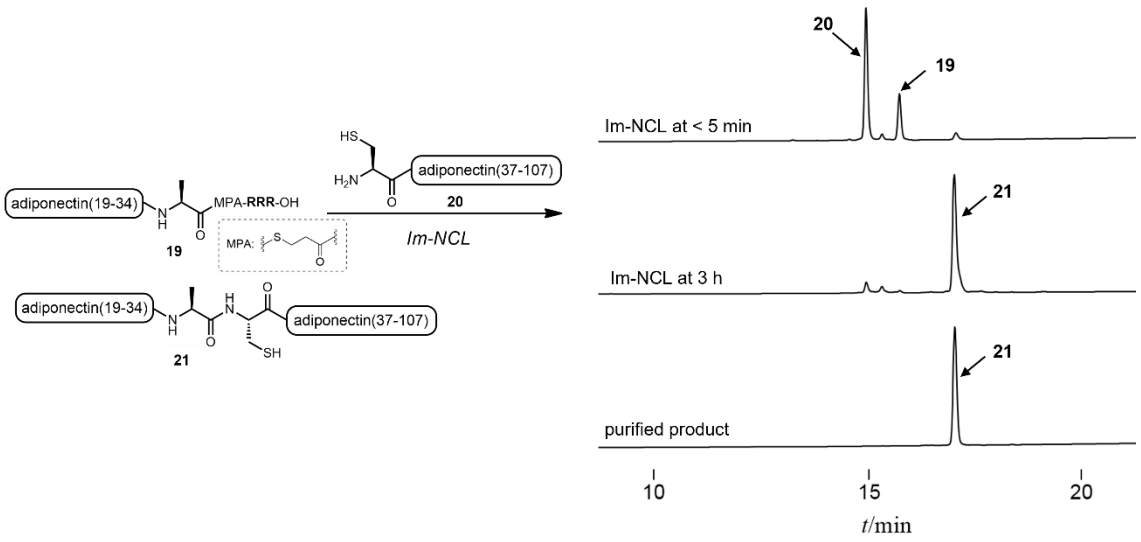


Figure 2.4. Synthesis of adiponectin(19-107) **21** by Imidazole-NCL.

2.5 One-pot NCL and desulfurization using Imidazole-NCL

Finally, I aimed to synthesize [Ser(PO_3H_2)^{65}]-ubiquitin (pUb) using a one-pot Imidazole-NCL and desulfurization method. The pUb is a modified form of ubiquitin that is phosphorylated and has been implicated in Parkinson's disease.[33] I hypothesized

that using imidazole instead of MPAA would not inhibit the desulfurization reaction, making it possible to synthesize phosphorylated ubiquitin in one step. In a previous synthesis using MPAA, the desired product peak was buried in the MPAA peak during RP-HPLC purification (Figure 2.5). Imidazole-NCL would solve this problem and simplify the purification process. I synthesized thioester of pUb(1-45) **22** and [Cys^{46}]-pUb(46-76) **23** using standard Boc and Fmoc-SPPS methods. The resulting thioester peptide **22** and cysteine peptide **23** were then subjected to Imidazole-NCL as shown in Figure 2.6. The reaction was completed within 3 hours and the thioester peptide **22** disappeared. The desulfurization reaction was then carried out by adding VA-044 and TCEP. The reaction proceeded efficiently overnight at 40 °C, and the one-pot reaction was achieved without any major by-products except for the hydrolyzed peptide derived from the thioester. The epimerization of Phe at the ligation site was only 1% (Figure 2.7). During the RP-HPLC purification after the one-pot reaction, imidazole did not affect the purification process, and the desired product **25** was successfully obtained in a yield of 69%. The obtained results clearly demonstrate that imidazole can serve as a highly efficient NCL catalyst in terms of productivity. Moreover, it enables a simple and direct desulfurization process, which cannot be achieved by the conventional MPAA-assisted NCL.

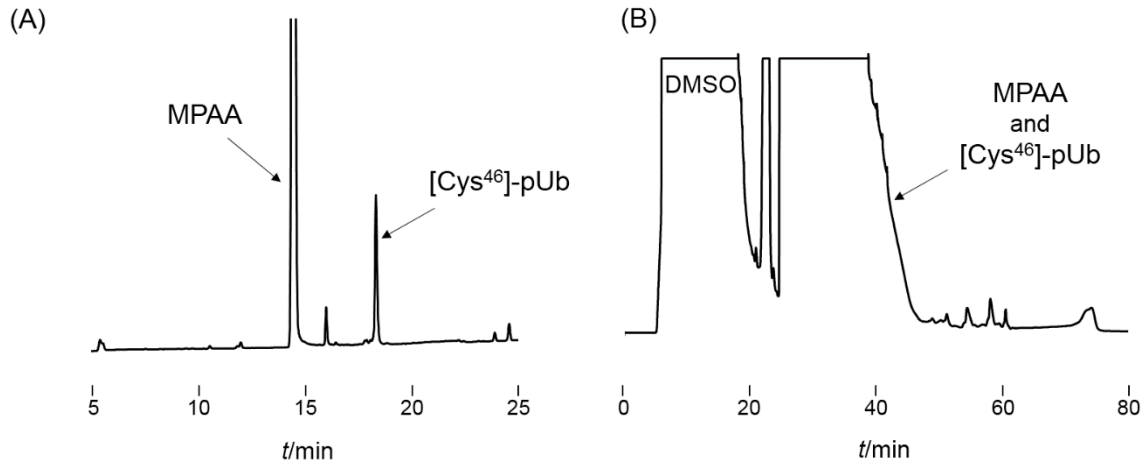


Figure 2.5. Co-elution of the ligation product "[Cys^{46}, Ser(PO$_3$H$_2$)^{65}]-ubiquitin ([Cys^{46}]-pUb)" with MPAA was observed after NCL using MPAA: (A) on an analytical-scale RP-HPLC and (B) on a preparative-scale RP-HPLC. To prevent the production of insoluble precipitates, DMSO was added before preparative-scale RP-HPLC.

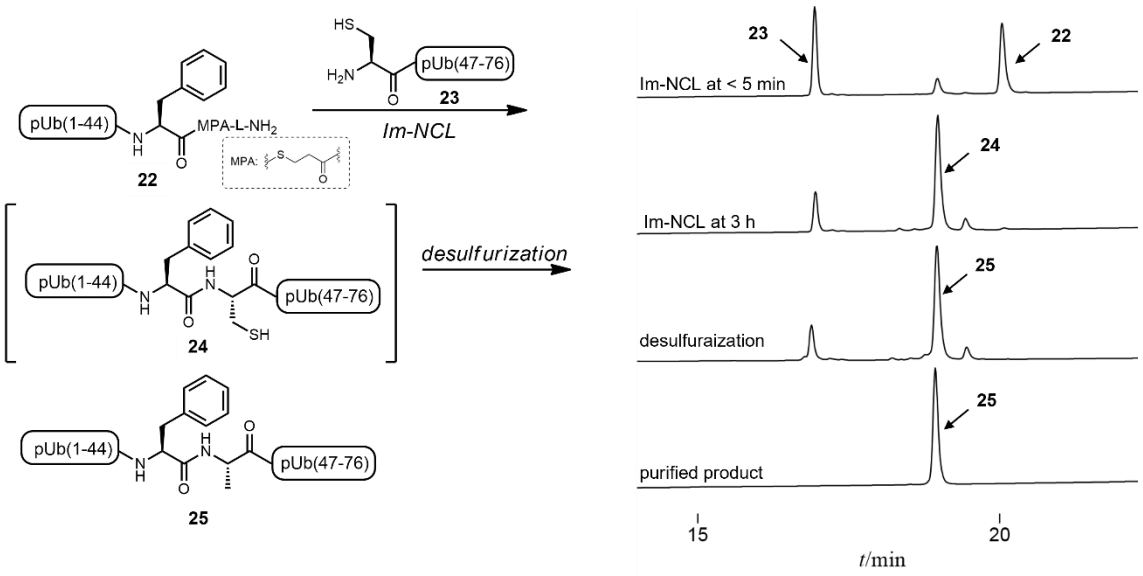


Figure 2.6. Synthesis of pUb **25** by one-pot Imidazole-NCL and desulfurization reaction.

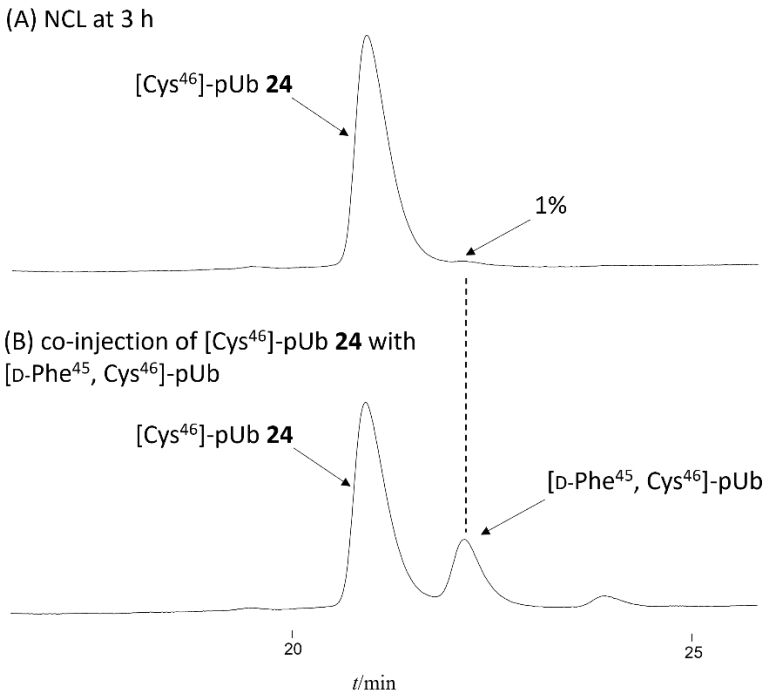


Figure 2.7. Epimerization at the ligation site (Phe^{45}) during Im-NCL, resulting in [Cys^{46}]-pUb **24** was evaluated by RP-HPLC using the authentic epimer [D-Phe^{45}, Cys^{46}]-pUb: (A) reaction mixture after 3 hours and (B) result of co-injection of the authentic epimer into the reaction mixture.

2.6 Conclusion

In this study, I have successfully demonstrated the efficacy of using imidazole instead of MPAA for efficient NCL using peptide-alkylthioester. The Imidazole-NCL method showed similar efficiency to the traditional MPAA-assisted NCL, and I was able to successfully synthesize adiponectin(19-107) and pUb using this method. The occurrence of side reactions, such as epimerization, was insignificant. In addition, one-pot desulfurization after NCL can be achieved since imidazole does not interfere with the radical reaction. Furthermore, during the final RP-HPLC purification, the highly hydrophilic nature of imidazole allowed for easy removal from the NCL product. Based on these results, I believe that Imidazole-NCL has the potential to become a standard method for the efficient protein synthesis in the future.

Chapter 3

Facile peptide synthesis via 1,2,4-triazole-assisted native chemical ligation with peptide-*N*-acyl-*N'*-methyl-benzimidazolinone

3.1 Problem of imidazole using peptide-N-acyl-N'-methyl-benzimidazolinone in NCL

Peptide thioesters play an important role in native chemical ligation (NCL). However, their application in Fmoc-SPPS is challenging due to their instability under Fmoc deprotection conditions. While various methods have been developed to synthesize peptide thioesters by Fmoc-SPPS, one of the most commonly used approaches is based on the use of 3,4-diaminobenzene (Dbz) and *o*-amino(methyl)aniline (MeDbz) linkers (Figure 3.1).[5] Peptide-Nbz (peptide-*N*-acyl-benzimidazolinone) and peptide-MeNbz (peptide-*N*-acyl-*N*'-methyl-benzimidazolinone) can be prepared by conventional Fmoc-SPPS using the Dbz and MeDbz linkers. After synthesis of peptide-Dbz/MeDbz by general Fmoc-SPPS, peptide-Nbz/MeNbz are prepared by treatment with *p*-nitrophenyl chloroformate. After deprotection with TFA, peptide-Nbz/MeNbz are reacted with thiols to form peptide thioesters. Alternatively, peptide-Nbz/MeNbz can be used directly as thioester equivalents in NCL. In this process, peptide-Nbz/MeNbz are converted in situ to peptide thioesters using aryl thiol additives such as 4-mercaptophenylacetic acid (MPAA). In Chapter 2, I demonstrated the efficacy of imidazole and expected that the use of imidazole with peptide-Nbz/MeNbz would enable practical NCL with the advantage of non-thiol additives. However, the NCL reaction between peptide-MeNbz and Cys-peptide using imidazole showed unexpectedly low efficiency due to hydrolysis of peptide-MeNbz as shown below. Therefore, in this Chapter, I explored alternative non-thiol additives for NCL with peptide-MeNbz.

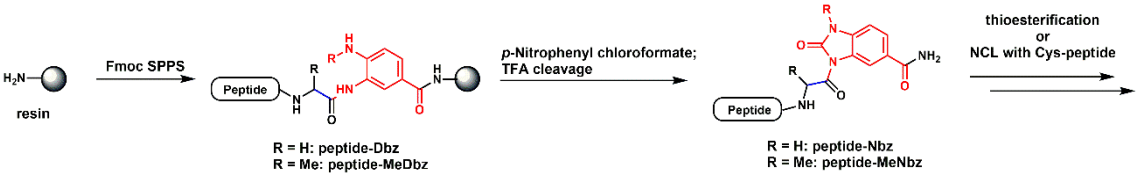


Figure 3.1. Summary of peptide-Nbz/MeNbz chemistry.

3.2 1,2,4-Triazole as an alternative additive in NCL using peptide-N-acyl-N'-methyl-benzimidazolinone

This research began by applying imidazole conditions to NCL using peptide-MeNbz. Peptide-MeNbz **26** (H-**LYRANX**-MeNbz, **X** = Ala, Phe, Asn, or Leu) was synthesized using Fmoc-SPPS according to established protocols. Initial attempts to promote the NCL reaction between peptide-MeNbz **26a** (**X** = Ala) and Cys-peptide **27** using imidazole (2.5 M) proved unsatisfactory, as the reaction was accompanied by hydrolysis of peptide-MeNbz **26a**, resulting in a lower yield (Figure 3.2, 3.3). It was postulated that the progression of the excess hydrolysis reaction was due to the acid-base catalysis effect mediated by imidazole. The higher reactivity of peptide-MeNbz compared to peptide thioesters may have resulted in a greater degree of hydrolysis induced by imidazole. Therefore, I decided to change the reaction accelerator from imidazole (pK_a = 7 for conjugated acids)[31] to triazoles[34] (pK_a = 9-10)[31] to fine-tune the reactivity of the intermediate. Both 2.5 M 1,2,4-triazole and 1,2,3-triazole facilitated the reaction with peptide-MeNbz at rates comparable to 50 mM MPAA while suppressing the hydrolysis reaction (Figure 3.2, 3.3). However, 1,2,3-triazole yielded an unidentified by-product with a mass value of +28 relative to the target peptide, possibly due to impurities in the reagent or degradation of 1,2,3-triazole. Therefore, I focused on 1,2,4-triazole and performed further investigations. Despite the fact that 1,2,4-triazole required an excess amount relative to MPAA, I anticipated that it would not be retained and would be easily removed by RP-HPLC. In addition, the non-thiol nature of 1,2,4-triazole, like imidazole, would offer several advantages, including the ability to perform one-pot desulfurization reactions.

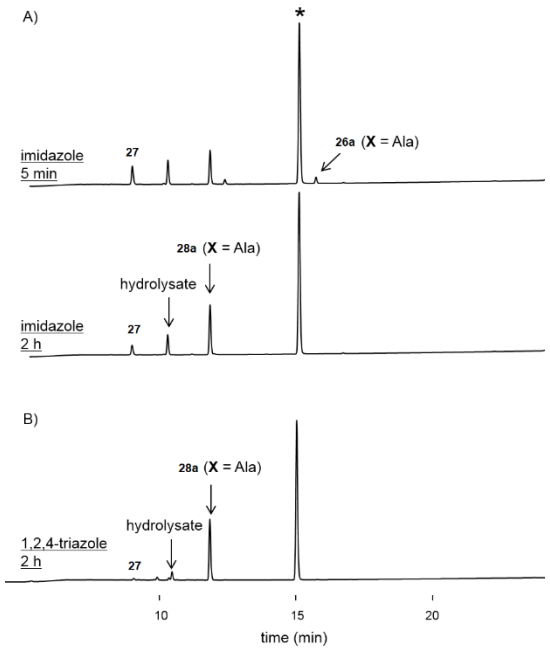


Figure 3.2. Efficiency of imidazole and 1,2,4-triazole for NCL using peptide-MeNbz was compared. The NCL reaction was carried out between H-**LYRANA**-MeNbz-**L**-NH_2 (**26a**) and H-**CSPGYS**-NH_2 (**27**). When the reaction was stopped, a higher amount of hydrolysate of peptide-MeNbz **26a** was observed in the case of NCL using imidazole (A) compared to NCL using 1,2,4-triazole (B). *MeNbz-**L**-NH_2.

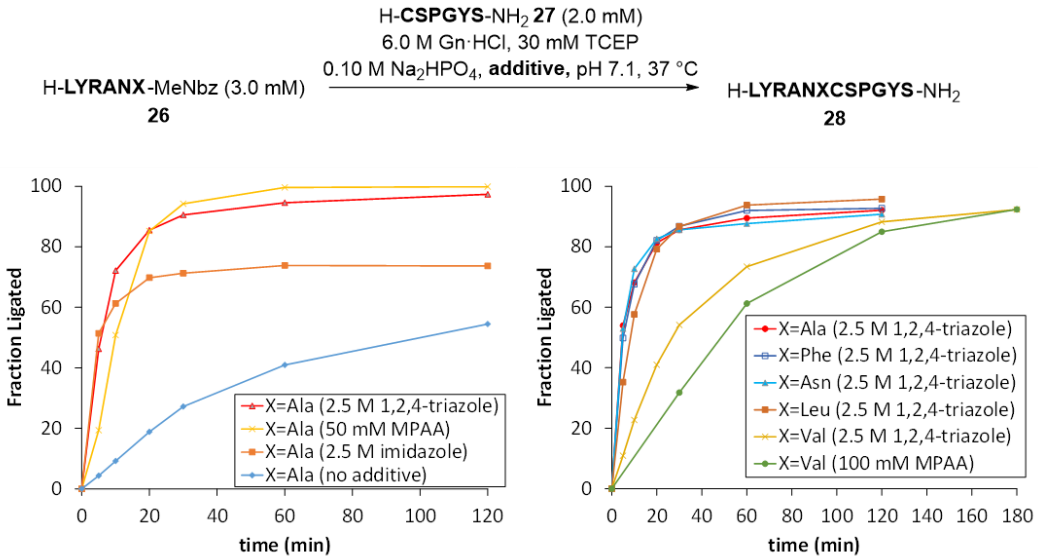
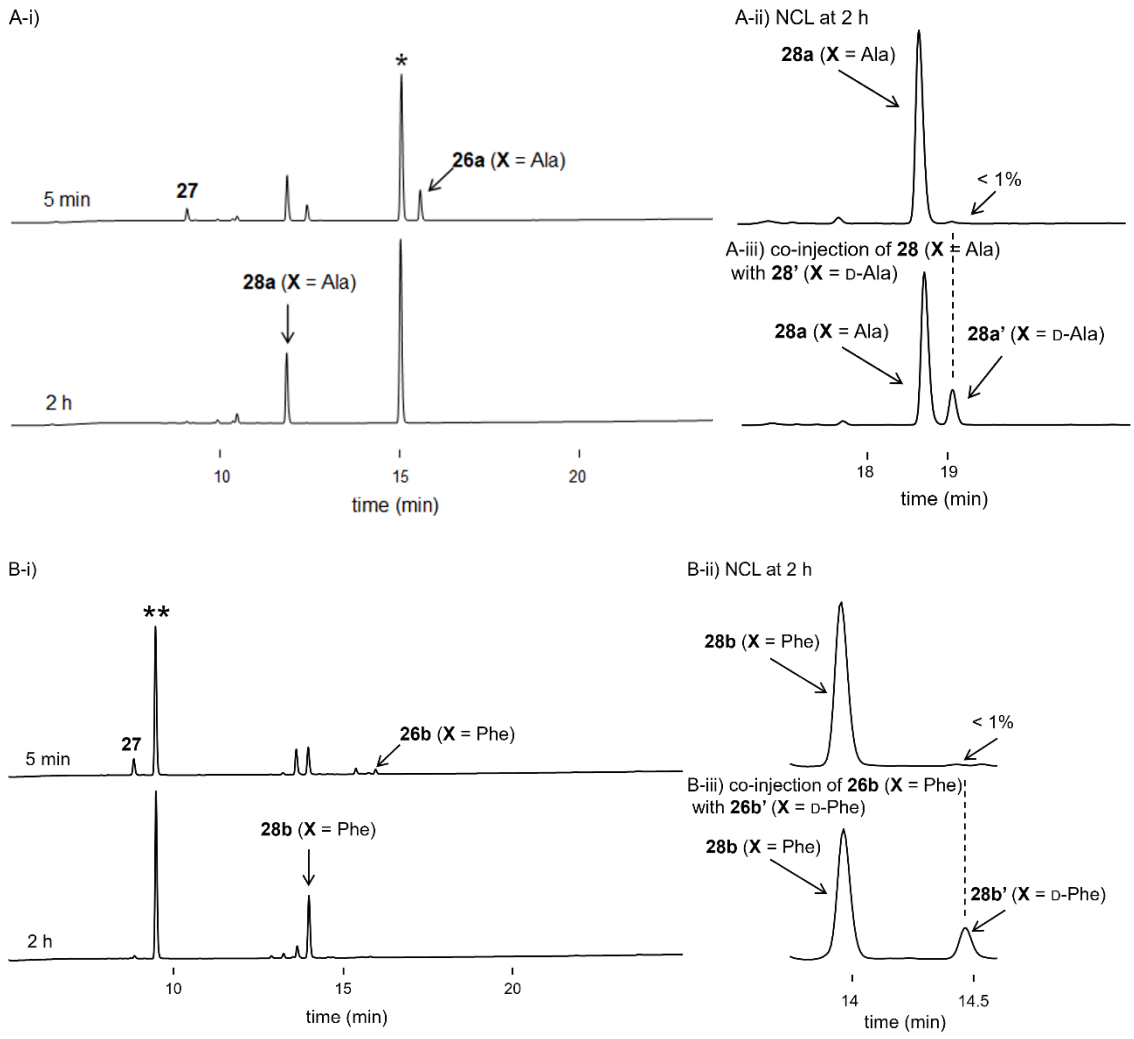


Figure 3.3. Kinetic analysis for the NCL between H-**LYRANX**-MeNbz **26** and H-**CSPGYS**-NH_2 **27**. The fraction ligated was determined using the HPLC peak areas at λ=280 nm.

3.3 Effect of 1,2,4-triazole on NCL using MeNbz peptides with various C-terminus

With the knowledge that 1,2,4-triazole is capable of accelerating ligation with MeNbz peptide, I proceeded to investigate the potential of 1,2,4-triazole-assisted NCL with peptide-MeNbz containing various C-terminal residues (Figure 3.3). The reaction proceeded rapidly (within 2 hours) at sites with lower steric hindrance (**X** = Ala, Phe, Asn, or Leu). The Val site, which is relatively bulky and therefore has a slower reaction rate, required 3 hours to complete. Notably, the reaction at the Val site with 1,2,4-triazole (2.5 M) was slightly faster than with MPAA (100 mM). Furthermore, the degree of epimerization observed during the reactions was less than 0.1% (**X** = Ala, Phe, Val) (Figure 3.4).



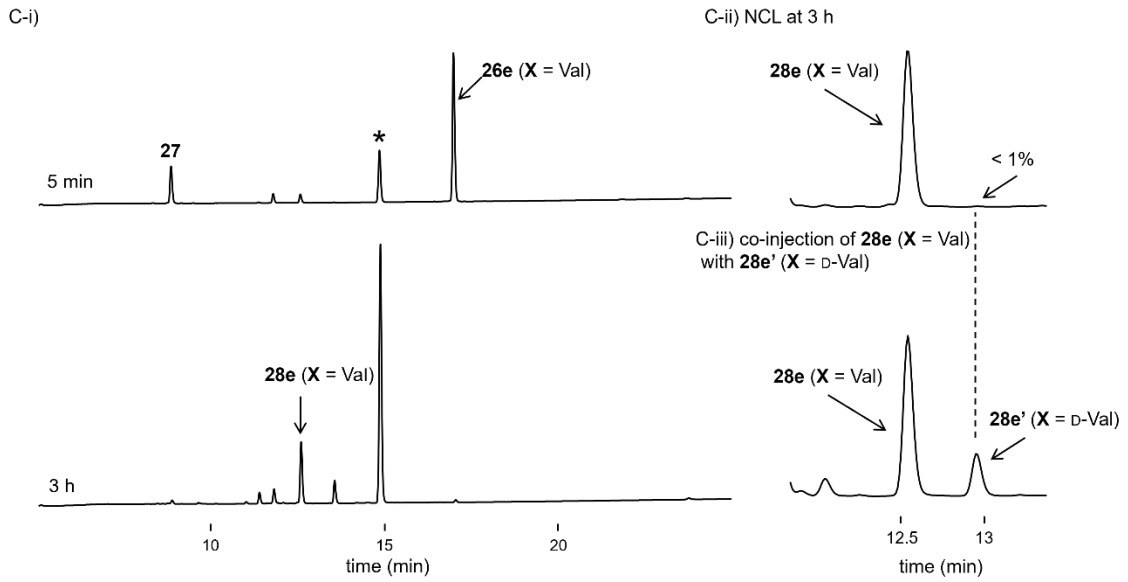


Figure 3.4. A, B, C-i) The HPLC traces show representative 1,2,4-triazole-aided NCL between peptide-MeNbz [H-**LYRANX**-MeNbz, A) **26a** (**X** = Ala); B) **26b** (**X** = Phe); C) **26e** (**X** = Val)] and Cys-peptide (H-**CSPGYS**-NH_2 (**27**). A, B, C-ii, iii) The rate of epimerization was confirmed by comparing with the authentic corresponding epimer [ii) reaction mixture after 2-3 hours, iii) result of co-injection of the authentic epimer into the reaction mixture]. * MeNbz-Leu-NH_2. **MeNbz-NH_2.

3.4 One-pot NCL and desulfurization reaction with 1,2,4-triazole

In this study, we evaluated the feasibility of a one-pot desulfurization reaction with 1,2,4-triazole-assisted NCL between peptide-MeNbz **26a** and Cys-peptide **27** (Figure 3.5). Upon completion of the NCL reaction with 2.5 M 1,2,4-triazole, the desulfurization reaction was initiated through the addition of desulfurization reagents including VA-044 and TCEP. Notably, the desulfurization reaction proceeded efficiently, comparable to the Imidazole-NCL (Chapter 2). It was also found that 1,2,4-triazole, like imidazole, is not retained by RP-HPLC and can be easily analyzed and purified.

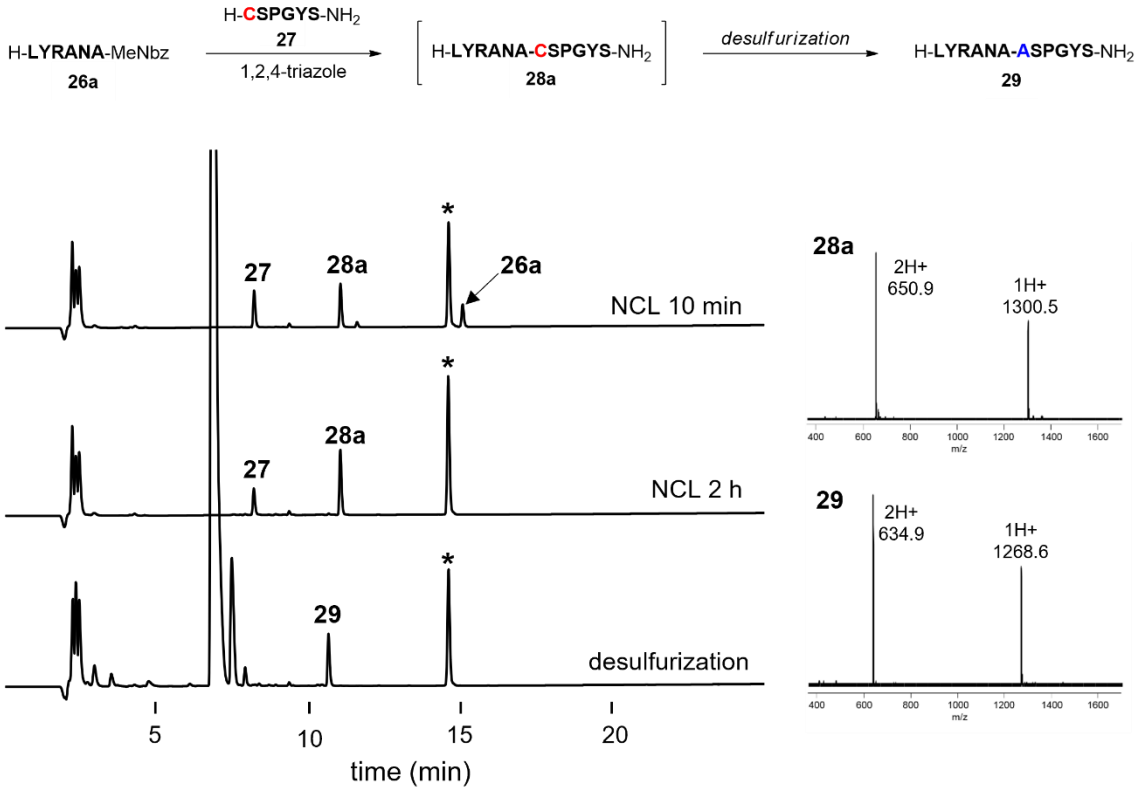


Figure 3.5. One-pot desulfurization after thiol-additive-free NCL using 1,2,4-triazole. *MeNbz-**L**-NH_2.

3.5 One-pot NCL and disulfide bond formation reaction with 1,2,4-triazole

To further demonstrate the usefulness of 1,2,4-triazole with its thiol-free nature, I next attempted a one-pot NCL and disulfide bond formation for the synthesis of human brain natriuretic peptide (BNP) (1-32),[35] which contains a single disulfide bond (Figure 3.6). It is important to note that the 1,2,4-triazole-NCL, which involves the use of peptide-MeNbz, is completely thiol-free, as the reaction system contains no thiol compounds other than the Cys-peptide. Conversely, the Imidazole-NCL using peptide thioesters is not completely thiol-free because alkyl thiols released from the peptide thioesters are present. I hypothesized that the thiol-free nature of the Triazole-NCL would inhibit inappropriate disulfide bond formation and facilitate one-pot disulfide bond formation. To evaluate this hypothesis, I first performed the NCL of peptide-MeNbz **30** and Cys-peptide **31** in the presence of 1,2,4-triazole (2.5 M) and the reaction was completed within 4 hours. The resulting reduced BNP(1-32) **32** was then

subjected to a disulfide bond formation reaction by adding iodine. The reaction proceeded rapidly without any side reactions and successfully produced BNP(1-32) **33**. Although this one-pot reaction contained TCEP, a reducing reagent for NCL, the remaining TCEP was consumed during the oxidation reaction and did not impede disulfide bond formation. In contrast, I found that MPAA was unsuitable for one-pot NCL and disulfide bond formation (Figure 3.7). As expected, the disulfide formation reaction with MPAA produced large amounts of MPAA adduct **34**. These results demonstrated the superiority of 1,2,4-triazole over MPAA.

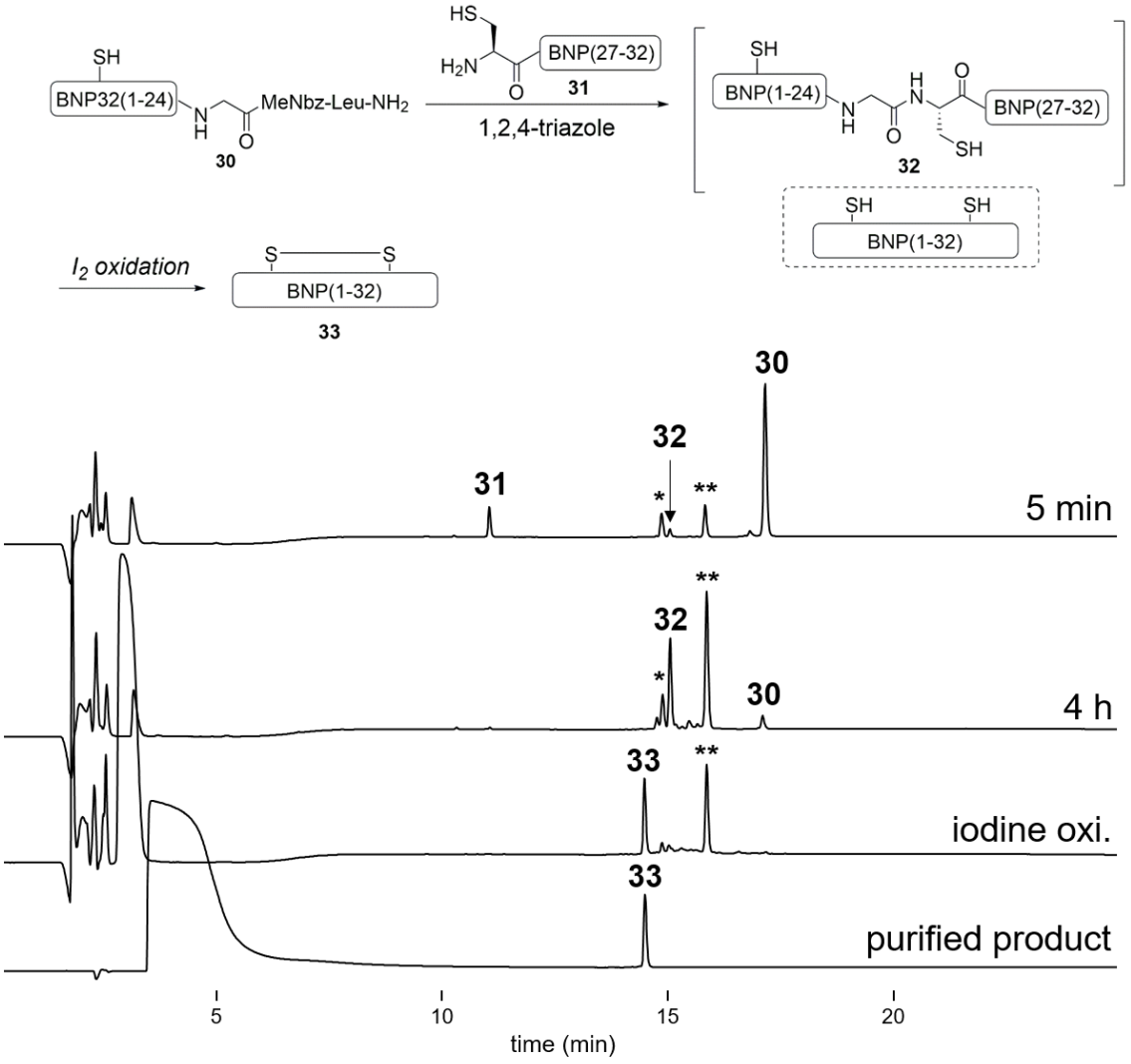


Figure 3.6. Synthesis of BNP(1–32) **33** using one-pot NCL and desulfurization reaction with 1,2,4-triazole. *Hydrolysate and intramolecular thiolactone derived from **30**. **MeNbz-**L**-NH_2.

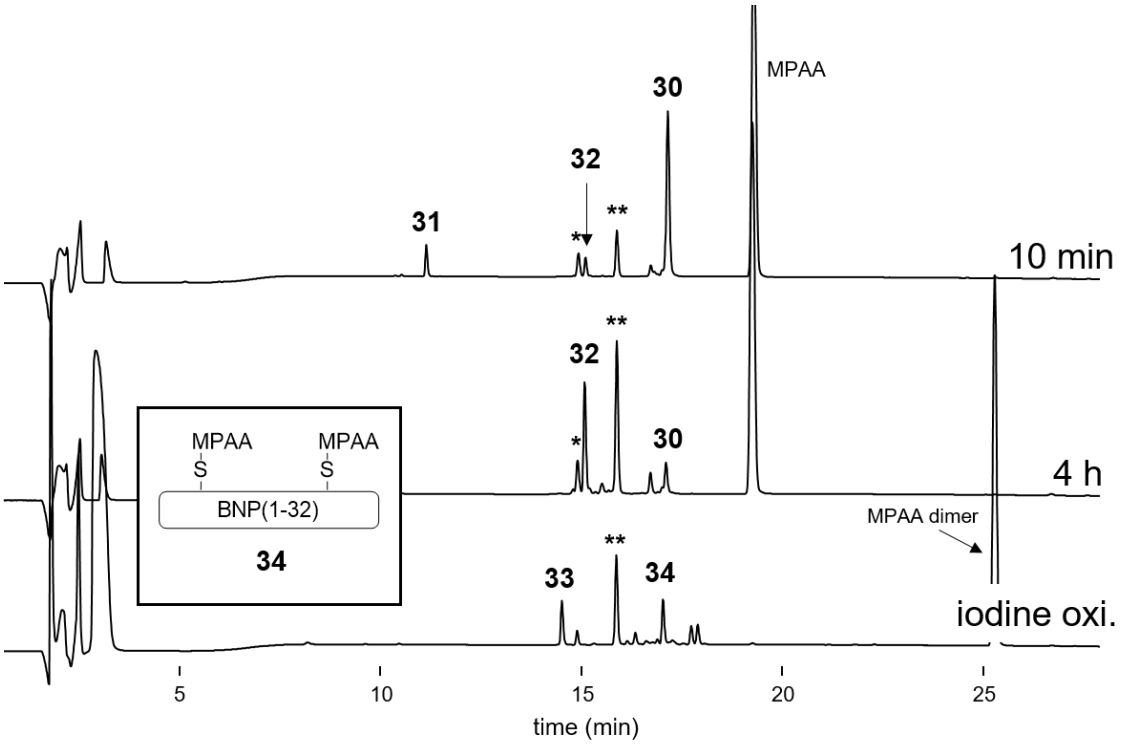


Figure 3.7. Synthesis of BNP(1–32) **33** using one-pot NCL and desulfurization reaction with MPAA. *Hydrolysate and intramolecular thiolactone derived from **30**. **MeNbz-**L**-NH_2.

3.6 One-pot NCL and resin purification with 1,2,4-triazole

In this study, I investigated the use of activated Thiopropyl Sepharose resin[36] to purify the desired peptides from the NCL mixture, taking advantage of the thiol-free property of 1,2,4-triazole. Commercially available activated Thiopropyl Sepharose resin contains reactive disulfide bonds that form disulfides with the free thiol group of the desired peptide. After washing the resin to remove impurities, the desired peptide can be easily recovered through treatment with reducing agents. This purification system was used by Maroney *et al.* after NCL with MPAA, but required pre-purification to remove MPAA, which competes with disulfide bond formation.[36e] I expected that NCL with 1,2,4-triazole would allow for one-pot purification using the resin. Furthermore, it was hypothesized that an excess of peptide-MeNbz, which contains no Cys residues, would result in the production of a highly pure desired product through thiol resin purification, since theoretically only the desired product should have thiol groups at the end of the

reaction. To evaluate this concept, I synthesized model peptide **37** (Figure 3.8). Peptide-MeNbz **35** and Cys-peptide **36** were ligated using 1,2,4-triazole-assisted NCL, and excess TCEP, the reducing agent for NCL, was quenched by the addition of 4-azidobenzoic acid.[37] The resulting solution was then treated with activated Thiopropyl Sepharose resin for 2 hours. After washing the peptide-bound resin, the desired peptide was released from the resin using dithiothreitol (DTT). The eluted solution was desalted by size exclusion chromatography and the desired product was successfully produced with more than 90% purity without HPLC purification.

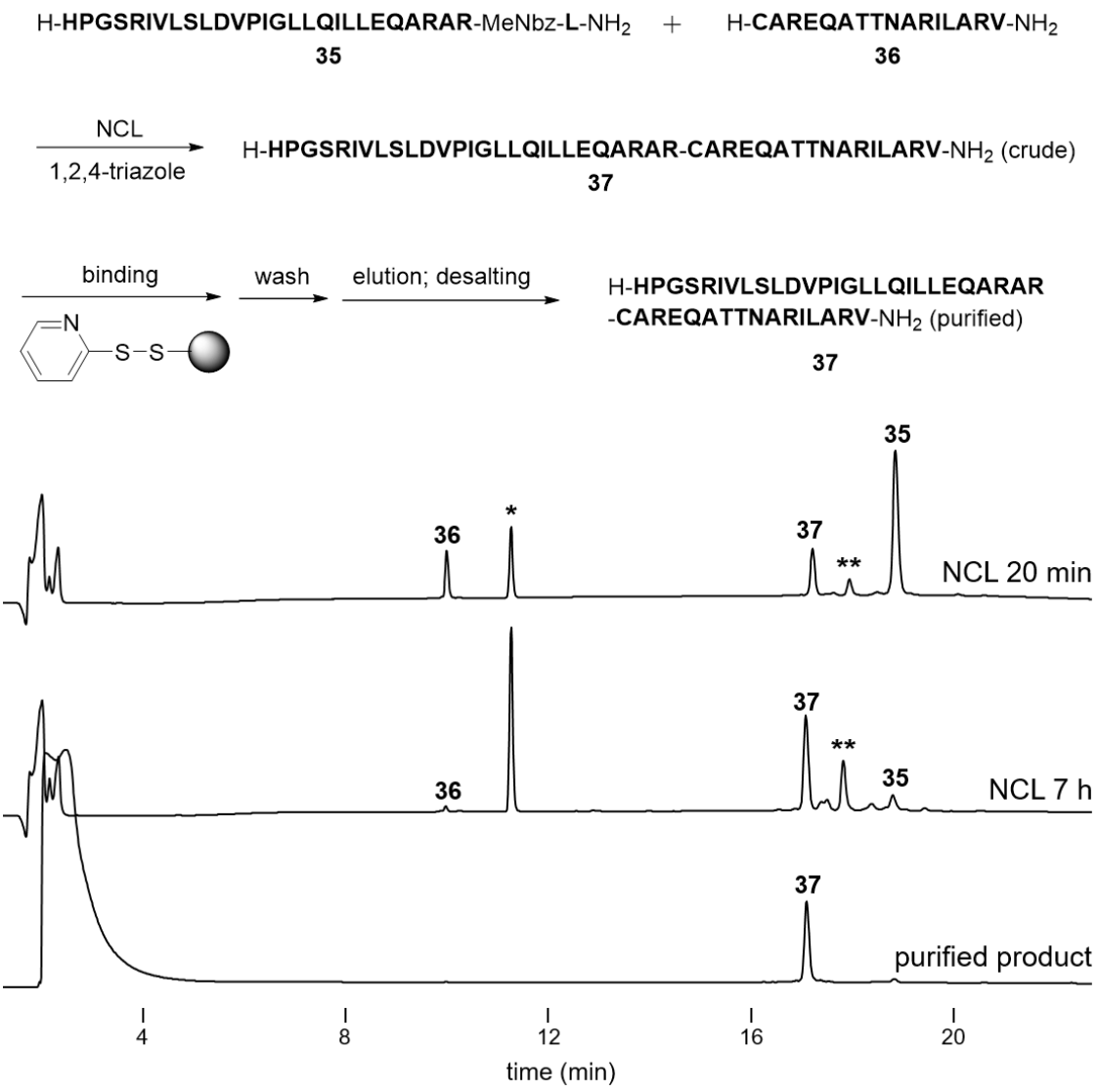


Figure 3.8. Synthesis of model peptide **37** by one-pot NCL and resin purification with 1,2,4-triazole. *MeNbz-Leu-NH_2. **Hydrolysate of **35**.

3.7 Conclusion

In this Chapter, I have shown that the utility of 1,2,4-triazole as a catalyst for facilitating NCL with peptide-MeNbz. The promoting efficiency of NCL using 2.5 M 1,2,4-triazole is comparable to that of the conventional NCL using MPAA. Importantly, 1,2,4-triazole does not affect reaction monitoring and product isolation by RP-HPLC. Furthermore, this method eliminates the need for thiol additives, allowing for one-pot Cys modification after NCL such as desulfurization. The utility of this method was further supported by its successful application in one-pot disulfide formation and purification using activated Thiopropyl Sepharose resin. The thiol-free NCL methodology presented in Chapters 2 and 3 has been widely adopted by various research groups, and its value is indisputable.[38] I trust that it will continue to be a valuable tool for protein preparation and contribute significantly to the advancement of the life sciences.

Chapter 4

Conclusions

1. I have developed the practical synthetic techniques for *N*-aminoacyl-*N*-sulfanylethylaniline linkers. The effectiveness of these methods has been demonstrated by the successful coupling of 20 naturally occurring proteogenic amino acids to the aniline linker. In addition, SEAlide peptides with various C-terminal amino acids have been synthesized using these techniques and were effectively used as peptide thioester equivalents in the NCL. Otaka *et al.* have demonstrated the usefulness of SEAlide peptides in NCL chemistry by performing one-pot, multi-fragment condensation under kinetic conditions. Therefore, the development of this technology provides significant opportunities for the use of SEAlide peptides. For example, a 162-residue glycoprotein was chemically synthesized using SEAlide peptides synthesized by the methods established in this study. I hope that the continued use of SEAlide peptides in protein synthesis will contribute to the advancement of protein research in the future.

2. I have successfully demonstrated the efficacy of using imidazole instead of MPAA for efficient NCL using peptide-alkylthioester. The Imidazole-NCL method showed similar efficiency to the traditional MPAA-assisted NCL, and we were able to successfully synthesize adiponectin(19-107) and pUb using this method. The occurrence of side reactions, such as epimerization, was insignificant. In addition, one-pot desulfurization after NCL can be achieved since imidazole does not interfere with the radical reaction. Furthermore, during the final RP-HPLC purification, the highly hydrophilic nature of imidazole allowed for easy removal from the NCL product. Based on these results, I believe that Imidazole-NCL has the potential to become a standard method for the efficient protein synthesis in the future.

3. I have shown that the utility of 1,2,4-triazole as a catalyst for facilitating NCL with peptide-MeNbz. The promoting efficiency of NCL using 2.5 M 1,2,4-triazole is comparable to that of the conventional NCL using MPAA. Importantly, 1,2,4-triazole does not affect reaction monitoring and product isolation by RP-HPLC. Furthermore, this method eliminates the need for thiol additives, allowing for

one-pot Cys modification after NCL such as desulfurization. The utility of this method was further supported by its successful application in one-pot disulfide formation and purification using activated Thiopropyl Sepharose resin. The thiol-free NCL methodology presented in Chapters 2 and 3 has been widely adopted by various research groups, and its value is indisputable.[38] I trust that it will continue to be a valuable tool for protein preparation and contribute significantly to the advancement of the life sciences.

The use of alternative reaction accelerators in NCL represents a promising avenue for practical protein synthesis. In particular, recent research has demonstrated the feasibility of chemical synthesis of proteins using non-thiol additives.[38] This innovation has the potential to expand the scope and versatility of protein synthesis techniques. I also believe that the synthetic proteins produced by these innovative methods will be valuable tools for elucidating complex biological phenomena.

Experimental section

General Methods

All reagents and solvents were purchased from Peptide Institute, Inc. (Osaka, Japan), FUJIFILM Wako Pure Chemical Corporation (Osaka, Japan), Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan), Nacalai Tesque, Inc. (Kyoto, Japan), Watanabe Chemical Industries, Ltd. (Hiroshima, Japan), Sigma-Aldrich Co. LLC. (St. Louis, MO), and Merck KGaA (Darmstadt, Germany). Silica gel (KANTO KAGAKU N-60) was used for column chromatography.

In Chapter 1, mass spectra were recorded on Waters MICROMASS® LCT PREMIER™ or Bruker Esquire200T. NMR spectra were recorded on a JEOL GSX400 spectrometer operating at 400 MHz frequency for ^{1}H and 100 MHz frequency for ^{13}C in $CDCl_3$, with TMS or residual $CHCl_3$ signal used as internal standards. Optical rotations were measured using a JASCO P-2200 polarimeter. For HPLC separations, a Cosmosil 5C_{18}-AR-II analytical column (Nacalai Tesque, 4.6 × 250 mm, flow rate 1.0 mL/min), a YMC-Pack Pro C18 RS column (4.6 × 150 mm, flow rate 1.0 mL/min), a 5C_{18}-AR-II semi-preparative column (Nacalai Tesque, 10 × 250 mm, flow rate 3.0 mL/min), or a 5C_{18}-AR-II preparative column (Nacalai Tesque, 20 × 250 mm, flow rate 10 mL/min) was employed. A solvent system consisting of 0.1% TFA in H_2O and 0.1% TFA in CH_3CN was used at a flow rate of 1 mL/min (room temperature) with detection at 220 nm. Absorbance was measured with BECKMAN SPECTROPHOTOMETER DU-650. In Chapters 2 and 3, mass spectra were obtained using an Agilent G1956B LC/MSD detector on an Agilent 1100 series HPLC system. Analytical HPLC was performed on a YMC-ODS AA12S05-1546WT column (4.6 x 150 mm). A solvent system consisting of 0.1% TFA in H_2O and 0.1% TFA in CH_3CN was used at a flow rate of 1 mL/min (40 °C) with detection at 220 nm unless otherwise stated. Preparative HPLC was performed on a YMC-ODS AA12S05-2530WT column (30 x 250 mm). A solvent system consisting of 0.1% TFA in H_2O and 0.1% TFA in CH_3CN was used at a flow rate of 20 mL/min (room temperature) with detection at 220 nm.

Automated peptide synthesis by Boc-SPPS was performed on an ABI 433A peptide synthesizer. The peptide chain was elongated using *in situ* neutralization protocols[39] with Boc-amino acid/HCTU/6-Cl-HOBt/DIEA (4/4/4/6 equiv.) as coupling reagents in NMP for 2 hours. The following side-chain-protected amino acids were

employed: Arg(Tos), Asp(OcHex), Cys(MeBzl), Glu(OcHex), His(Bom), Lys(ClZ), Ser(Bzl), Thr(Bzl), Tyr(BrZ). Manual peptide synthesis by Fmoc-SPPS was performed on a KMS-3 (Kokusan Chemical Co., Ltd, Tokyo, Japan) peptide synthesizer. The peptide chain was elongated using the coupling protocols with Fmoc-amino acid/DIC/HOBt·H_2O (3/3/3 equiv.) in DMF for 2 hours. Automated peptide synthesis by Fmoc-SPPS was performed on an ABI 433A peptide synthesizer (Applied Biosystems, MA). The peptide chain was elongated using the coupling protocols with Fmoc amino acid/DIC/OxymaPure (4/4/4 equiv.). For the -Asp-Gly- moiety, Fmoc-Asp(O*t*-Bu)-(Dmb)Gly-OH was adopted. The following side-chain-protected amino acids were used: Arg(Pbf), Asn(Trt), Asp(O*t*-Bu), Cys(Trt), Glu(O*t*-Bu), Gln(Trt), His(Trt) for adiponectin(36-107), His(MBom)[16] for [Ser(PO_3H_2)[65]]-ubiquitin, Hyp(*t*-Bu), Lys(Boc), Ser[PO(OBzl)OH], Ser(*t*-Bu), Trp(Boc), Thr(*t*-Bu), Tyr(*t*-Bu). In addition, the following protected amino acids were used as N-terminal amino acids: Boc-Leu-OH (for peptide **26**), Boc-Ser(*t*-Bu)-OH (for BNP(1-25)-MeNbz-Leu-NH_2 (**30**)), Boc-His(Trt)-OH (for peptide **35**).

Chapter 1

Typical experimental protocol utilizing Fmoc amino acyl chloride ($SOCl_2$-Method)

The aniline linker **3** (100 mg, 0.21 mmol) was dissolved in THF (2.0 mL) and stirred at 4 °C with 55% NaH (suspended in mineral oil) (10 mg, 0.23 mmol) for 30 minutes at room temperature. Separately, Fmoc-Leu-OH (148 mg, 0.42 mmol) was dissolved in CH_2Cl_2 (1.5 mL) and treated with $SOCl_2$ (0.31 mL, 4.2 mmol) and a drop of DMF for 3 hours at room temperature. The reaction mixture was then concentrated and dried under reduced pressure, and the resulting residue was mixed with the previously prepared sodium anilide solution at 4 °C. The mixture was allowed to react for 15 hours at room temperature and quenched with saturated $NaHCO_3$ solution (sat. $NaHCO_3$ aq.). After extraction with EtOAc, the organic phase was washed with $NaHCO_3$ aq., 10% citric acid aq., and brine. After drying over $MgSO_4$ and removal of the solvent under reduced pressure, purification was performed by silica gel column chromatography using EtOAc-hexane. As a result, the desired Fmoc-Leu-incorporating linker (allyl ester) was obtained in an isolated yield of 74% (126 mg).

Typical experimental protocol utilizing Fmoc amino acyl chloride prepared from triethylammonium salts of Fmoc amino acid (modified $SOCl_2$-Method)

Fmoc-Ser(*t*-Bu)-OH (480 mg, 1.3 mmol) and Et_3N (174 μL, 1.4 mmol) were dissolved in CH_2Cl_2 (12 mL), followed by the addition of $SOCl_2$ (137 μL, 2.0 mmol) and a drop of DMF, and the reaction was refluxed for 2 hours. After removing the solvent under reduced pressure, the residue was dissolved in THF (3 mL), and the formed salts were removed by filtration. The resulting filtrate was mixed with a solution of sodium anilide (0.63 mmol) in THF (6 mL) prepared as in $SOCl_2$-Method and stirred for 20 hours at room temperature. The reaction was quenched with sat. $NaHCO_3$ aq.. After extraction with EtOAc, the organic phase was washed sequentially with sat. $NaHCO_3$ aq., 10% citric acid aq., and brine, and dried over $MgSO_4$. The solvent was then removed under reduced pressure, and the crude product was purified by silica gel column chromatography using EtOAc-hexane to give the desired Fmoc-Ser(*t*-Bu)-incorporated linker (allyl ester) in an isolated yield of 89% (469 mg).

Typical experimental protocol for Condition A in Table 1.1 (Entry 2)

Fmoc-Ser(*t*-Bu)-OH (399 mg, 1.0 mmol) and the aniline linker **3** (100 mg, 0.21 mmol) were dissolved in pyridine (2.0 mL), followed by dropwise addition of $POCl_3$ (97

mL, 1.0 mmol) at −15 °C. The reaction was stirred for 3 hours at 4 °C. The reaction was then quenched with sat. $NaHCO_3$ aq. and the resulting mixture was extracted with EtOAc. The organic layer was sequentially washed with sat. $NaHCO_3$ aq., 10% citric acid, and brine, and then dried over $MgSO_4$. The solvent was removed under reduced pressure and the crude product was purified by column chromatography on silica gel using EtOAc-hexane. The desired compound was obtained in an isolated yield of 48% (85 mg).

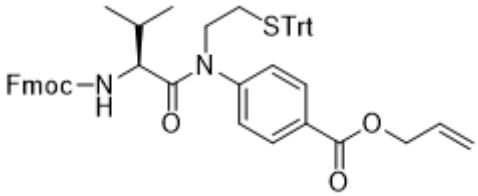
Typical experimental protocol for Condition C in Table 1.1 (Entry 5)

Fmoc-Ser(*t*-Bu)-OH (2.0 g, 5.2 mmol), the aniline linker **3** (500 mg, 1.0 mmol) and Et_3N (0.73 mL, 5.2 mmol) were dissolved in CH_2Cl_2 (17 mL), followed by dropwise addition of $POCl_3$ (0.49 mL, 5.2 mmol) at 4 °C. The reaction was stirred for 26 hours at room temperature. The reaction was then quenched with sat. $NaHCO_3$ aq. and the resulting mixture was extracted with EtOAc. The organic layer was sequentially washed with sat. $NaHCO_3$ aq., 10% citric acid, and brine, and then dried over $MgSO_4$. The solvent was removed under reduced pressure and the crude product was purified by column chromatography on silica gel using EtOAc-hexane. The desired compound was obtained in an isolated yield of 65% (571 mg).

Typical experimental protocol in Table 1.2

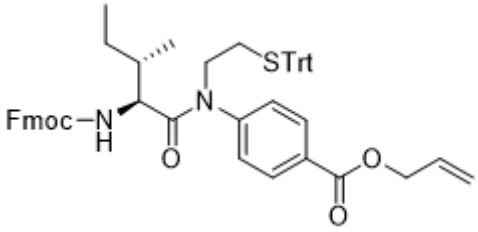
Fmoc-Ser(*t*-Bu)-OH (400 mg, 1.05 mmol) was dissolved in THF (2.4 mL) and treated sequentially with Et_3N (146 μL, 1.05 mmol) and $POCl_3$ (98 μL, 1.05 mmol) at 4 °C. The resulting mixture was then immediately treated with a solution of sodium anilide (0.21 mmol, 1.1 mL) in THF prepared as in $SOCl_2$-Method. The reaction was allowed to proceed for 24 hours at room temperature. The reaction was then quenched with sat. $NaHCO_3$ aq. and the resulting mixture was extracted with EtOAc. The organic layer was sequentially washed with sat. $NaHCO_3$ aq., 10% citric acid, and brine, and then dried over $MgSO_4$. The solvent was removed under reduced pressure and the crude product was purified by column chromatography on silica gel using EtOAc-hexane. The desired compound was obtained in an isolated yield of 90% (159 mg).

Allyl 4-[(Fmoc-L-Val-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 87% (831 mg); $[\alpha]^{21}_{D}$ 100.5 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 0.68 (3H, d, *J* = 6.5 Hz), 0.76 (3H, d, *J* = 6.5 Hz), 1.75 (1H, oct, *J* = 6.5 Hz), 2.28 (1H, ddd, *J* = 12.8, 9.0 and 5.5 Hz), 2.55 (1H, ddd, *J* = 12.8, 8.4 and 5.4 Hz), 3.34 (1H, ddd, *J* = 14.2, 8.4 and 5.5 Hz), 3.58 (1H, ddd, *J* = 14.2, 9.0 and 5.4 Hz), 4.05 (1H, dd, *J* = 9.0 and 6.5 Hz), 4.20 (1H, t, *J* = 7.0 Hz), 4.30 (1H, dd, *J* = 10.4 and 7.0 Hz), 4.38 (1H, dd, *J* = 10.4 and 7.0 Hz), 4.85 (2H, d, *J* = 5.6 Hz), 5.33 (1H, d, *J* = 10.4 Hz), 5.34 (1H, d, *J* = 9.0 Hz), 5.44 (1H, d, *J* = 17.2 Hz), 6.06 (1H, ddt, *J* = 17.2, 10.4 and 5.6 Hz), 7.06 (2H, d, *J* = 8.2 Hz), 7.09-7.36 (17H, m), 7.39 (2H, t, *J* = 7.3 Hz), 7.58 (2H, dd, *J* = 7.4 and 2.8 Hz), 7.76 (2H, d, *J* = 7.3 Hz), 8.05 (2H, d, *J* = 8.2 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 17.2, 19.5, 29.2, 31.5, 47.2, 49.0, 56.4, 65.9, 66.9, 66.9, 118.7, 120.0, 125.1, 125.1, 126.7, 127.0, 127.7, 127.8, 128.4, 129.5, 129.8, 131.1, 131.9, 141.3, 143.8, 144.5, 145.0, 155.9, 165.3, 171.6; HRMS (ESI-TOF) *m*/*z* calcd for $C_{51}H_{49}N_2O_5S$ $[M + H]^+$: 801.3362, found: 801.3358.

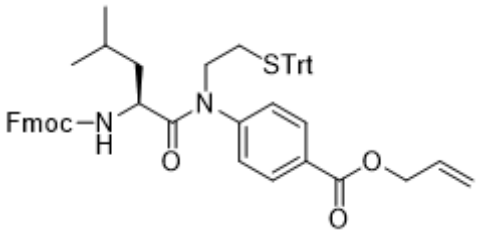
Allyl 4-[(Fmoc-L-Ile-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 87% (848 mg); $[\alpha]^{21}_{D}$ 114.9 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 0.67 (3H, d, *J* = 6.6 Hz), 0.73 (3H, t, *J* = 7.2 Hz), 0.84-0.98 (1H, m), 1.31-1.45 (1H, m), 1.45-1.56 (1H, m), 2.28 (1H, ddd, *J* = 13.3, 9.0 and 5.4 Hz), 2.55 (1H, ddd, *J* = 13.3, 8.3 and 5.1 Hz), 3.37 (1H, ddd, *J* = 14.3, 8.3 and 5.4 Hz), 3.60 (1H, ddd, *J* = 14.3, 9.0 and 5.1 Hz), 4.08 (1H, dd, *J* = 8.8 and 7.5 Hz), 4.20 (1H, t, *J* = 7.0 Hz), 4.31 (1H, dd, *J* = 10.5 and 7.0 Hz), 4.38 (1H, dd, *J* = 10.5 and 7.0 Hz), 4.85 (2H, d, *J* = 5.6 Hz), 5.30 (1H, d, *J* = 8.8 Hz), 5.33 (1H, d, *J* = 10.5 Hz), 5.44 (1H, d, *J* = 17.4 Hz), 6.06 (1H, ddt, *J* = 17.4, 10.5 and 5.6 Hz), 7.06 (2H, d, *J* = 8.0 Hz), 7.10-7.36 (17H, m), 7.40 (2H, t, *J* = 7.4 Hz), 7.59 (2H, dd, *J* = 7.2 and 2.5 Hz), 7.76 (2H, d, *J* = 7.5 Hz), 8.04 (2H, d, *J* = 8.0 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 11.2, 15.6, 23.9, 29.2, 38.2, 47.2, 49.0,

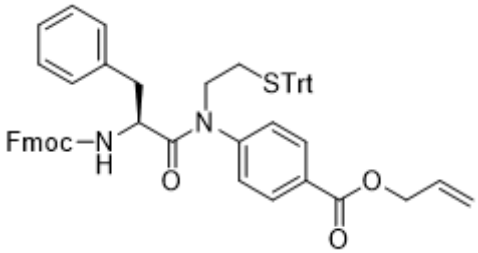
55.8, 65.9, 66.8, 66.9, 118.6, 119.9, 125.1, 125.1, 126.6, 127.0, 127.7, 127.8, 128.4, 129.5, 129.6, 129.8, 131.1, 132.0, 141.3, 143.8, 144.5, 145.0, 155.8, 165.3, 171.7; HRMS (ESI-TOF) *m*/*z* calcd for $C_{52}H_{51}N_2O_5S$ $[M + H]^+$: 815.3519, found: 815.3549.

Allyl 4-[(Fmoc-L-Leu-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 88% (75 mg); $[\alpha]^{18}_D$ 97.4 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 0.37 (3H, d, *J* = 6.0 Hz), 0.72 (3H, d, *J* = 6.2 Hz), 1.08-1.20 (1H, m), 1.32-1.49 (2H, m), 2.30 (1H, ddd, *J* = 12.9, 9.3 and 5.4 Hz), 2.53 (1H, ddd, *J* = 12.9, 8.3 and 5.3 Hz), 3.28 (1H, ddd, *J* = 13.8, 8.3 and 5.4 Hz), 3.61 (1H, ddd, *J* = 13.8, 9.3 and 5.3 Hz), 4.15-4.40 (4H, m), 4.86 (2H, d, *J* = 5.6 Hz), 5.27 (1H, d, *J* = 9.2 Hz), 5.33 (1H, d, *J* = 10.5 Hz), 5.43 (1H, d, *J* = 17.1 Hz), 6.06 (1H, ddt, *J* = 17.1, 10.5 and 5.6 Hz), 7.05-7.35 (19H, m), 7.39 (2H, t, *J* = 7.4 Hz), 7.58 (2H, dd, *J* = 7.2 and 4.8 Hz), 7.76 (2H, d, *J* = 7.4 Hz), 8.06 (2H, d, *J* = 8.1 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 20.8, 23.2, 24.3, 29.2, 42.3, 47.1, 48.9, 50.2, 65.9, 66.9, 66.9, 118.6, 119.9, 119.9, 125.1, 126.6, 127.0, 127.7, 127.8, 128.4, 129.5, 129.8, 131.1, 131.9, 141.3, 143.8, 143.9, 144.5, 144.9, 156.0, 165.2, 172.5; HRMS (ESI-TOF) *m*/*z* calcd for $C_{52}H_{50}N_2NaO_5S$ $[M + Na]^+$: 837.3338, found: 837.3304.

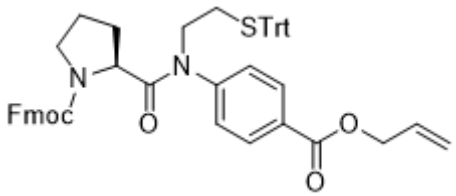
Allyl 4-[(Fmoc-L-Phe-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 88% (883 mg); $[\alpha]^{19}_D$ 63.8 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 2.21-2.41 (2H, m), 2.71 (1H, dd, *J* = 12.9 and 6.1 Hz), 2.85 (1H, dd, *J* = 12.9 and 8.8 Hz), 3.25-3.45 (2H, m), 4.15 (1H, t, *J* = 7.1 Hz), 4.21-4.40 (1H, m), 4.32 (2H, d, *J* = 7.1 Hz), 4.85 (2H, d, *J* = 5.5 Hz), 5.33 (1H, d, *J* = 10.4 Hz), 5.38 (1H, d, *J* = 8.9 Hz), 5.43 (1H, d, *J* = 17.3 Hz), 6.05 (1H, ddt, 17.3, 10.4 and 5.5 Hz), 6.88 (2H, d, *J* = 7.2 Hz), 7.09-7.34 (22H, m), 7.39 (2H, t, *J* = 7.4 Hz), 7.55 (2H, dd, *J* = 7.2 and 4.4 Hz),

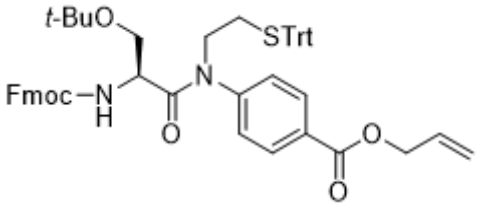
7.76 (2H, d, J = 7.4 Hz), 7.88 (2H, d, J = 8.1 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 29.0, 39.8, 47.1, 48.9, 52.9, 65.9, 66.9, 66.9, 118.6, 119.9, 125.1, 125.1, 126.7, 127.0, 127.7, 127.8, 128.0, 128.5, 129.5, 129.7, 130.9, 132.0, 132.0, 135.8, 141.3, 143.7, 143.8, 144.5, 155.2, 165.3, 170.9; HRMS (ESI-TOF) *m*/*z* calcd for $C_{55}H_{49}N_2O_5S$ $[M + H]^+$: 849.3362, found: 849.3356.

Allyl 4-[(Fmoc-L-Pro-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 39% (66 mg); $[\alpha]^{19}_D$ 62.0 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.56-2.07 (4.3H, m), 2.13-2.34 (1.2H, m), 2.42-2.54 (0.5H, m), 3.12-3.34 (1H, m), 3.34-3.51 (2H, m), 3.51-3.72 (1H, m), 3.87-4.05 (1H, m), 4.10-4.40 (3H, m), 4.67-4.81 (2H, m), 5.20-5.30 (1H, m), 5.31-5.41 (1H, m), 5.91-6.05 (1H, m) 6.50-6.78 (1H, m), 6.95-7.57 (22H, m), 7.64-7.74 (2H, m), 7.76-7.88 (1H, m), 7.91-8.10 (1H, m); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 23.3, 24.4, 29.3, 30.2, 31.4, 46.9, 47.1, 47.3, 47.5, 48.5, 48.9, 57.0, 57.5, 65.8, 65.9, 66.8, 67.0, 67.3, 118.5, 118.6, 119.9, 125.0, 125.1, 125.3, 126.6, 126.9, 127.0, 127.0, 127.1, 127.6, 127.6, 127.8, 127.9, 128.4, 128.6, 129.4, 129.5, 129.6, 131.0, 131.9, 131.9, 141.2, 143.8, 144.0, 144.1, 144.4, 144.6, 145.2, 145.7, 154.2, 154.7, 165.2, 165.4, 171.5, 171.9; HRMS (ESI-TOF) *m*/*z* calcd for $C_{51}H_{46}N_2NaO_5S$ $[M + Na]^+$: 821.3025, found: 821.3010.

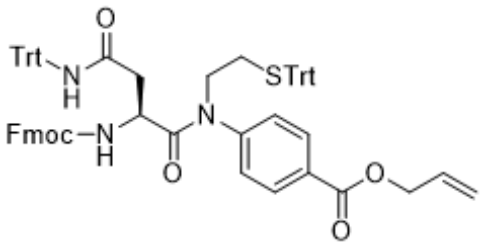
Allyl 4-[(Fmoc-L-Ser(*t*-Bu)-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 90% (159 mg); $[\alpha]^{20}_D$ 52.1 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.06 (9H, s), 2.35-2.54 (2H, m), 3.18-3.28 (1H, br m), 3.32 (1H, t, J = 7.8 Hz), 3.42 (1H, ddd, J = 13.5, 8.3 and 5.2 Hz), 3.54 (1H, ddd, J = 13.5, 8.3 and 5.8 Hz), 4.17 (1H, t, J = 7.0 Hz), 4.30 (2H, d, J = 7.0 Hz) 4.25-4.38 (1H, m), 4.85 (2H, d, J = 5.4 Hz), 5.32 (1H, d, J = 10.4 Hz), 5.36-5.48 (1H, m), 5.43 (1H, d, J = 17.2 Hz), 6.04 (1H, ddt, J = 17.2, 10.4 and 5.4 Hz), 7.05-7.43 (21H, m), 7.57 (2H, dd, J = 7.2 and 3.6 Hz),

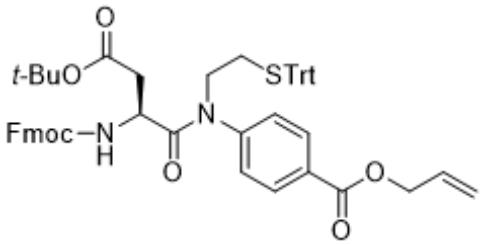
7.75 (2H, d, J = 7.4 Hz), 8.03 (2H, d, *J* = 8.3 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 27.2, 29.2, 47.1, 49.2, 51.6, 62.5, 65.8, 67.0, 73.5, 118.5, 119.9, 125.1, 126.6, 127.0, 127.6, 127.8, 128.4, 129.5, 129.7, 130.9, 132.0, 141.2, 143.8, 143.8, 144.5, 145.1, 155.5, 165.3, 170.3; HRMS (ESI-TOF) *m*/*z* calcd for $C_{53}H_{53}N_2O_6S$ $[M + H]^+$: 845.3624, found: 845.3643.

Allyl 4-[(Fmoc-L-Asn(Trr)-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 89% (980 mg); $[\alpha]^{19}_D$ 2.4 (*c* 0.40, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 2.37 (2H, t, *J* = 7.4 Hz), 2.50 (1H, br d, *J* = 15.2 Hz), 2.60 (1H, dd, *J* = 15.2 and 5.2Hz), 3.35 (2H, t, *J* = 7.4 Hz), 4.05 (1H, t, *J* = 6.3 Hz), 4.17 (2H, d, *J* = 6.3 Hz), 4.38-4.55 (1H, m), 4.76 (2H, d, *J* = 5.4 Hz), 5.26 (1H, d, *J* = 10.4 Hz), 5.37 (1H, d, *J* = 17.4 Hz), 5.43-5.56 (1H, m), 5.97 (1H, ddt, *J* = 17.4, 10.4 and 5.4 Hz), 6.70 (1H, s), 6.91 (2H, d, *J* = 7.9 Hz), 7.06-7.34 (32H, m), 7.38 (2H, t, *J* = 7.3 Hz), 7.49 (2H, d, *J* = 6.7 Hz), 7.75 (2H, d, *J* = 7.2 Hz), 7.89 (2H, d, *J* = 7.9 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 28.9, 36.0, 39.4, 46.9, 49.3, 49.4, 50.2, 65.6, 66.9, 67.1, 70.4, 74.0, 118.4, 119.8, 119.9, 124.8, 124.9, 124.9, 125.0, 126.5, 126.8, 126.9, 127.0, 127.4, 127.6, 127.8, 127.9, 128.4, 128.6, 129.3, 129.5, 130.9, 131.8, 141.1, 141.1, 141.8, 143.4, 143.5, 143.6, 143.6, 144.3, 144.4, 155.0, 155.6, 165.1, 168.2, 170.2, 172.9, 174.8; HRMS (ESI-TOF) *m*/*z* calcd for $C_{69}H_{59}N_3NaO_6S$ $[M + Na]^+$: 1080.4022, found: 1080.4039.

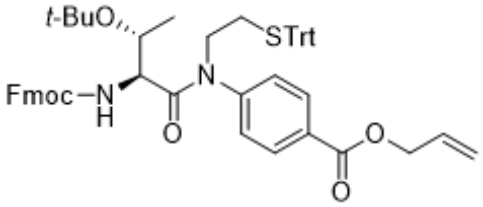
Allyl 4-[(Fmoc-L-Asp(O*t*-Bu)-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 68% (62 mg); $[\alpha]^{19}_D$ 47.1 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.36 (9H, s), 2.27 (1H, br dd, *J* = 14.8 and 6.0 Hz), 2.33-2.54 (3H, m), 3.38 (1H, ddd, *J* = 14.0, 8.4 and 5.4 Hz), 3.52 (1H, ddd, *J* = 14.0, 9.2 and 4.8 Hz), 4.16

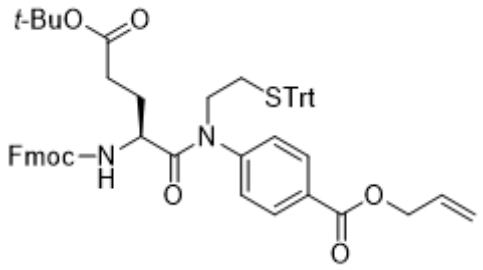
(1H, t, J = 7.0 Hz), 4.21-4.35 (2H, m), 4.40-4.54 (1H, br m), 4.83 (2H, d, J = 5.6 Hz), 5.30 (1H, d, J = 10.4 Hz), 5.41 (1H, d, J = 17.1 Hz), 5.52 (1H, d, J = 9.0 Hz), 6.02 (1H, ddt, J = 17.1, 10.4 and 5.6 Hz), 7.06-7.36 (19H, m), 7.39 (2H, t, J = 7.4 Hz), 7.56 (2H, dd, J = 6.4 and 4.8 Hz), 7.76 (2H, d, J = 7.4 Hz), 8.02 (2H, d, J = 8.2 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 27.9, 29.1, 38.2, 47.0, 49.0, 49.2, 65.8, 67.0, 67.0, 81.3, 118.5, 119.9, 125.1, 126.6, 127.0, 127.7, 127.8, 128.1, 129.5, 129.6, 129.9, 131.2, 131.9, 141.2, 143.7, 143.8, 144.5, 144.7, 155.2, 165.2, 169.0, 169.9; HRMS (ESI-TOF) m/z calcd for $C_{54}H_{52}N_2NaO_7S$ $[M + Na]^+$: 895.3393, found: 895.3373.

Allyl 4-[(Fmoc-L-Thr(*t*-Bu)-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 84% (698 mg); $[\alpha]^{23}_D$ 122.0 (c 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 0.89 (3H, d, J = 5.6 Hz), 1.03 (9H, s), 2.20-2.36 (1H, br m), 2.45-2.66 (1H, br m), 3.10-3.34 (1H, br m), 3.51-3.66 (1H, br m), 3.66-3.82 (1H, br m), 4.15-4.30 (1H, m), 4.22 (1H, t, J = 7.1 Hz), 4.35 (2H, d, J = 7.1 Hz), 4.86 (2H, d, J = 5.6 Hz), 5.33 (1H, d, J = 10.4 Hz), 5.44 (1H, d, J = 17.4 Hz), 5.54 (1H, d, J = 9.0 Hz), 6.06 (1H, ddt, J = 17.4, 10.4 and 5.6 Hz), 7.05-7.35 (19H, m), 7.40 (2H, t, J = 7.4 Hz), 7.61 (2H, dd, J = 7.4 and 7.2 Hz), 7.77 (2H, d, J = 7.5 Hz), 8.06 (2H, d, J = 8.3 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 20.1, 28.1, 29.1, 47.2, 49.4, 56.5, 65.9, 67.0, 67.7, 74.2, 118.6, 119.9, 125.2, 126.6, 127.0, 127.6, 127.7, 127.8, 127.9, 128.2, 129.5, 129.6, 131.2, 131.9, 141.3, 143.8, 144.0, 144.5, 145.4, 156.1, 165.3, 170.0; HRMS (ESI-TOF) m/z calcd for $C_{54}H_{54}N_2NaO_6S$ $[M + Na]^+$: 881.3600, found: 881.3580.

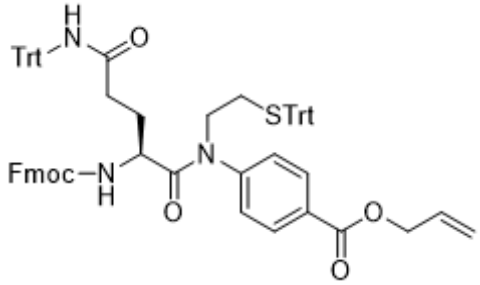
Allyl 4-[(Fmoc-L-Glu(O*t*-Bu)-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 72% (665 mg); $[\alpha]^{19}_D$ 81.5 (c 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.36 (9H, s), 1.65-1.83 (2H, m), 1.96-2.19 (2H, m), 2.23-2.37 (1H,

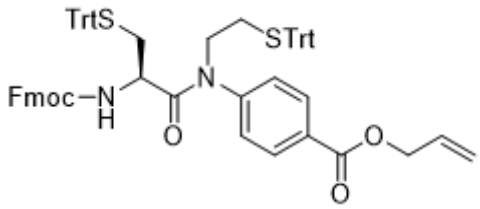
m), 2.44-2.59 (1H, m), 3.33 (1H, ddd, J = 13.4, 8.8 and 5.9 Hz), 3.57 (1H, ddd, J = 13.4, 9.0 and 6.1 Hz), 4.08-4.24 (1H, m), 4.18 (1H, t, J = 7.2 Hz), 4.32 (2H, d, J = 7.2 Hz), 4.86 (2H, ddd, J = 5.6, 1.5 and 1.3 Hz), 5.33 (1H, dtd, J = 10.4, 1.3 and 1.2 Hz), 5.44 (1H, dtd, J = 17.2, 1.5 and 1.2 Hz), 5.53 (1H, d, J = 8.5 Hz), 6.06 (1H, ddt, J = 17.2, 10.4 and 5.6 Hz), 7.07-7.36 (19H, m), 7.40 (2H, t, J = 7.4 Hz), 7.58 (2H, dd, J = 6.8 and 5.0 Hz), 7.76 (2H, d, J = 7.4 Hz), 8.05 (2H, d, J = 8.2 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 27.5, 28.0, 29.1, 30.9, 47.1, 48.9, 51.1, 65.8, 66.9, 80.6, 118.6, 119.9, 125.1, 126.3, 127.0, 127.8, 128.3, 129.5, 129.6, 130.0, 131.3, 131.9, 141.2, 143.7, 143.8, 144.5, 144.6, 155.7, 165.2, 171.1, 171.8; HRMS (ESI-TOF) m/z calcd for $C_{55}H_{54}N_2NaO_7S$ $[M + Na]^+$: 909.3549, found: 909.3569.

Allyl 4-[(Fmoc-L-Gln(Trt)-2-tritylsulfanylethyl)amino]benzoate



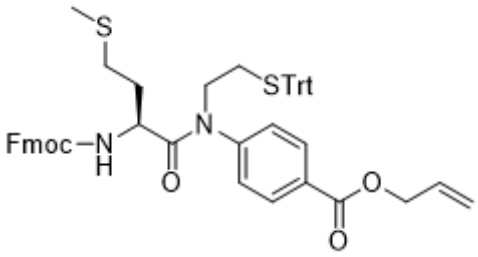
White amorphous solid, yield: 93% (1032 mg); $[\alpha]^{19}_D$ 48.3 (c 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.64-1.88 (2H, m), 2.00-2.22 (2H, m), 2.28 (1H, ddd, J = 13.5, 9.0 and 5.3 Hz), 2.46 (1H, ddd, J = 13.5, 8.0 and 5.3 Hz), 3.29 (1H, ddd, J = 14.1, 8.0 and 5.3 Hz), 3.48 (1H, ddd, J = 14.1, 9.0 and 5.3 Hz), 4.06-4.23 (2H, m), 4.24-4.41 (2H, m), 4.84 (2H, d, J = 5.5 Hz), 5.31 (1H, d, J = 10.4 Hz), 5.42 (1H, d, J = 17.2 Hz), 5.62 (1H, d, J = 8.2 Hz), 6.04 (1H, ddt, J = 17.2, 10.4 and 5.5 Hz), 6.47 (1H, s), 7.00-7.43 (36H, m), 7.56 (2H, d, J = 6.6 Hz), 7.74 (2H, dd, J = 7.6 and 2.8 Hz), 7.98 (2H, d, J = 8.2 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 27.8, 29.2, 32.7, 47.1, 49.0, 51.3, 65.8, 66.8, 67.0, 70.5, 118.6, 119.9, 125.1, 126.6, 127.0, 127.0, 127.7, 127.8, 127.9, 128.3, 128.6, 129.5, 129.6, 130.0, 131.3, 131.9, 141.2, 141.3, 143.7, 143.8, 144.5, 155.8, 165.2, 170.4, 171.0; HRMS (ESI-TOF) m/z calcd for $C_{70}H_{62}N_3O_6S$ $[M + H]^+$: 1072.4359, found: 1072.4370.

Allyl 4-[(Fmoc-L-Cys(Tr)-2-tritylsulfanylethyl)amino]benzoate



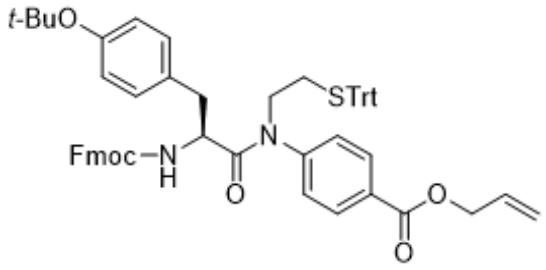
White amorphous solid, yield: 72% (78 mg); $[\alpha]^{18}_{D}$ 16.8 (*c* 1.00, $CHCl_3$); ^1H NMR ($CDCl_3$, 400 MHz) δ = 2.06 (1H, dd, *J* = 12.4 and 6.8 Hz), 2.20-2.34 (2H, m), 2.38-2.51 (1H, m), 3.26-3.50 (2H, m), 4.19 (1H, t, *J* = 7.3 Hz), 4.22-4.36 (3H, m), 4.86 (2H, d, *J* = 5.6 Hz). 5.34 (1H, d, *J* = 10.4 Hz), 5.45 (1H, d, *J* = 17.1 Hz), 5.38-5.50 (1H, m), 6.06 (1H, ddt, *J* = 17.1, 10.4 and 5.6 Hz), 6.86 (2H, d, *J* = 7.4 Hz), 7.08-7.34 (32H, m), 7.39 (2H, t, *J* = 7.4 Hz), 7.59 (2H, d, *J* = 7.2 Hz), 7.76 (2H, d, *J* = 7.4 Hz), 7.95 (2H, d, *J* = 8.3 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 29.1, 34.0, 47.0, 49.2, 50.6, 65.9, 66.5, 66.9, 67.1, 118.6, 119.9, 125.2, 126.6, 126.7, 127.1, 127.7, 127.8, 127.9, 128.1, 129.4, 129.5, 129.9, 131.2, 131.9, 141.2, 143.7, 143.8, 144.2, 144.4, 144.5, 155.3, 165.1, 169.6; HRMS (ESI-TOF) *m*/*z* calcd for $C_{68}H_{58}N_2NaO_5S_2$ $[M + Na]^+$: 1069.3685, found: 1069.3671.

Allyl 4-[(Fmoc-L-Met-2-tritylsulfanylethyl)amino]benzoate



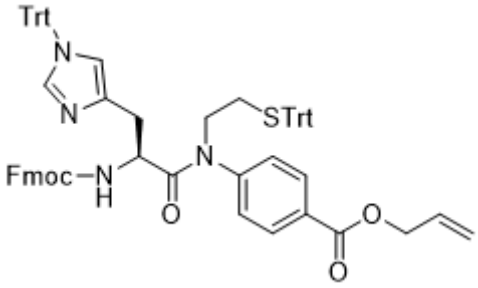
White amorphous solid, yield: 80% (690 mg); $[\alpha]^{18}_{D}$ 68.3 (*c* 1.00, $CHCl_3$); ^1H NMR ($CDCl_3$, 400 MHz) δ = 1.56-1.78 (2H, m), 1.88 (3H, s), 2.14-2.25 (1H, m), 2.25-2.38 (2H, m), 2.42-2.60 (1H, m), 3.38 (1H, ddd, *J* = 13.0, 8.1 and 6.5 Hz), 3.55 (1H, ddd, *J* = 13.0, 8.2 and 6.5 Hz), 4.18 (1H, t, *J* = 6.9 Hz), 4.25-4.39 (3H, m), 4.85 (2H, d, *J* = 5.4 Hz), 5.32 (1H, d, *J* = 10.4 Hz), 5.43 (1H, d, *J* = 17.4 Hz), 5.48 (1H, d, *J* = 8.9 Hz), 6.05 (1H, ddt, *J* = 17.4, 10.4 and 5.4 Hz), 7.03-7.36 (19H, m), 7.39 (2H, t, *J* = 7.4 Hz), 7.57 (2H, dd, *J* = 7.6 and 5.6 Hz), 7.76 (2H, d, *J* = 7.4 Hz), 8.06 (2H, d, *J* = 7.9 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 15.4, 29.2, 29.7, 32.5, 47.1, 48.9, 51.0, 65.9, 66.9, 66.9, 118.6, 119.9, 125.1, 126.7, 127.0, 127.7, 127.8, 128.2, 129.5, 130.1, 131.3, 131.9, 141.3, 143.7, 143.8, 144.4, 144.6, 155.8, 165.2, 171.3; HRMS (ESI-TOF) *m*/*z* calcd for $C_{51}H_{49}N_2O_5S_2$ $[M + H]^+$: 833.3083, found: 833.3077.

Allyl 4-[(Fmoc-L-Tyr(*t*-Bu)-2-tritylsulfanylethyl)amino]benzoate



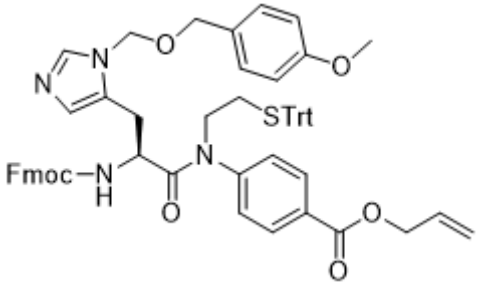
White amorphous solid, yield: 86% (773 mg); $[\alpha]^{22}_{D}$ 25.9 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.31 (9H, s), 2.23-2.42 (2H, m), 2.67 (1H, dd, *J* = 13.0 and 5.9 Hz), 2.83 (1H, dd, *J* = 13.0 and 9.0 Hz), 3.24-3.43 (2H, m), 4.16 (1H, t, *J* = 7.0 Hz), 4.22-4.43 (3H, m), 4.85 (2H, d, *J* = 5.5 Hz), 5.33 (1H, d, *J* = 10.5 Hz), 5.36 (1H, d, *J* = 8.9 Hz), 5.44 (1H, d, *J* = 17.2 Hz), 6.06 (1H, ddt, *J* = 17.2, 10.5 and 5.5 Hz), 6.79 (4H, s), 7.09-7.36 (19H, m), 7.40 (2H, t, *J* = 7.4 Hz), 7.56 (2H, dd, *J* = 6.8 and 2.4 Hz), 7.76 (2H, d, *J* = 7.5 Hz), 7.90 (2H, d, *J* = 8.0 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 28.8, 29.0, 39.3, 47.1, 48.9, 52.9, 65.8, 66.9, 67.0, 78.4, 118.6, 119.9, 124.1, 125.1, 125.1,126.7, 127.0, 127.7, 127.7, 127.8, 127.9, 128.1, 129.5, 129.7, 129.8, 130.0, 130.6, 130.8, 132.0, 141.3, 143.8, 143.8, 144.5, 144.5, 154.5, 155.2, 165.3, 171.0; HRMS (ESI-TOF) *m*/*z* calcd for $C_{59}H_{57}N_2O_6S$ $[M + H]^+$: 921.3937, found: 921.3939.

Allyl 4-[(Fmoc-L-His(Trt)-2-tritylsulfanylethyl)amino]benzoate



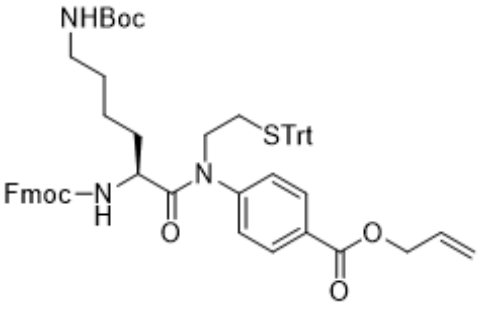
White amorphous solid, yield: 41% (745 mg); $[\alpha]^{20}_{D}$ 32.9 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 2.15-2.38 (1H, br m), 2.52-2.71 (2H, br m), 2.71-2.87 (1H, br m), 3.17-3.37 (1H, br m), 3.40-3.61 (1H, m), 4.18 (1H, t, *J* = 7.3 Hz), 4.20-4.35 (2H, m), 4.37-4.49 (1H, m), 4.87 (2H, d, *J* = 5.4 Hz), 5.34 (1H, d, *J* = 10.4 Hz), 5.45 (1H, d, *J* = 17.3 Hz), 6.07 (1H, ddt, *J* = 17.3, 10.4 and 5.4 Hz), 6.13 (1H, d, *J* = 8.0 Hz), 6.50 (1H, s), 7.02-7.45 (37H, m), 7.59 (2H, dd, *J* = 7.6 and 2.8 Hz), 7.76 (2H, d, *J* = 7.4 Hz), 8.02 (2H, d, *J* = 8.2 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 29.1, 30.4, 47.0, 49.2, 51.9, 65.7, 66.9, 67.0, 75.3, 118.5, 119.6, 119.8, 125.2, 125.3, 126.5, 127.0, 127.0, 127.1, 127.5, 127.7, 127.9, 128.0, 128.3, 129.4, 129.7, 131.0, 131.9, 135.6, 138.2, 141.1, 142.1, 143.8, 143.9, 144.5, 145.1, 155.7, 165.2, 170.6; HRMS (ESI-TOF) *m*/*z* calcd for $C_{71}H_{60}N_4NaO_5S$ $[M + Na]^+$: 1103.4182, found: 1103.4187.

Allyl 4-[(Fmoc-L-His(MBom)-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 88% (91 mg); $[\alpha]^{20}_{D}$ 45.4 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 2.25 (1H, ddd, *J* = 12.8, 9.2 and 6.0 Hz), 2.44 (1H, ddd, *J* = 12.8, 8.8 and 6.0 Hz), 2.77 (1H, dd, *J* = 14.7 and 6.7 Hz), 2.93 (1H, dd, *J* = 14.7 and 7.9 Hz), 3.20 (1H, ddd, *J* = 14.0, 8.8 and 6.0 Hz), 3.60 (1H, ddd, *J* = 14.0, 9.2 and 6.0 Hz), 3.78 (3H, s), 4.07 (1H, d, *J* = 11.6 Hz), 4.12 (1H, d, *J* = 11.6 Hz), 4.15 (1H, t, *J* = 6.9 Hz), 4.24-4.41 (3H, m), 4.77 (1H, d, *J* = 11.2 Hz), 4.81 (1H, d, *J* = 11.2 Hz), 4.83 (2H, d, *J* = 5.5, 1.7 and 1.5 Hz), 5.30 (1H, dtd, *J* = 10.4, 1.5 and 1.3 Hz), 5.42 (1H, dtd, *J* = 17.2, 1.7 and 1.3 Hz), 5.50 (1H, d, *J* = 8.8 Hz), 6.03 (1H, ddt, *J* = 17.2, 10.4 and 5.5 Hz), 6.70 (1H, s), 6.72-6.81 (1H, br m), 6.84 (2H, d, *J* = 8.5 Hz), 7.05-7.35 (21H, m), 7.38 (2H, t, *J* = 7.3 Hz), 7.55 (2H, dd, *J* = 7.2 and 3.2 Hz), 7.74 (2H, d, *J* = 7.5 Hz), 7.94 (2H, d, *J* = 8.5 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 27.5, 29.0, 47.0, 48.8, 51.2, 55.2, 65.9, 66.9, 69.0, 72.3, 114.0, 118.7, 120.0, 125.0, 126.1, 126.6, 127.0, 127.7, 127.8, 127.9, 128.1, 129.4, 129.6, 129.7, 129.9, 131.0, 131.9, 138.2, 141.3, 143.7, 144.3, 155.3, 159.6, 165.1, 170.6; HRMS (ESI-TOF) *m*/*z* calcd for $C_{61}H_{56}N_4NaO_7S$ $[M + Na]^+$: 1011.3767, found: 1011.3791.

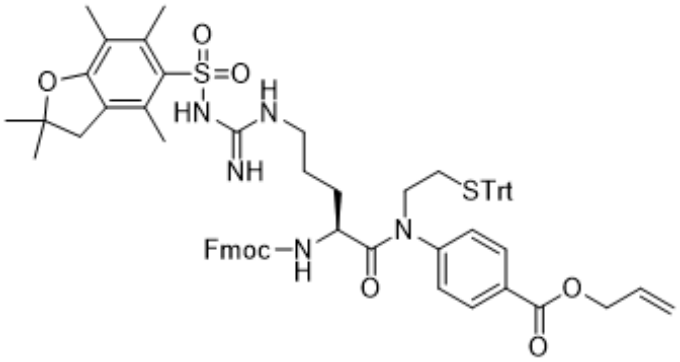
Allyl 4-[(Fmoc-L-Lys(Boc)-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 80% (779 mg); $[\alpha]^{21}_{D}$ 69.1 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.01-1.30 (4H, br m), 1.31-1.52 (2H, m) 1.41 (9H, s), 2.29 (1H, ddd, *J* = 12.8, 9.4 and 5.4 Hz), 2.53 (1H, ddd, *J* = 12.8, 9.2 and 5.1 Hz), 2.81-3.05 (2H, br s), 3.31 (1H, ddd, *J* = 14.5, 9.2 and 5.4 Hz), 3.60 (1H, ddd, *J* = 14.5, 9.4 and 5.1 Hz), 4.10-4.26 (1H, m), 4.19 (1H, t, *J* = 7.2 Hz), 4.33 (2H, d, *J* = 7.2 Hz), 4.37 (1H, m), 4.86

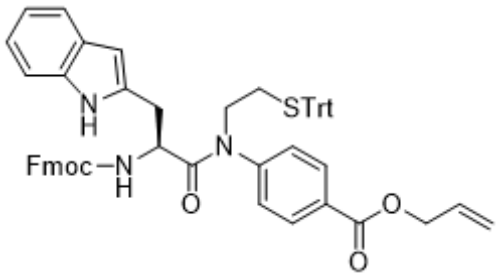
(2H, ddd, J = 6.0, 2.0 and 1.6 Hz), 5.33 (1H, dtd, J = 11.2, 2.0 and 1.2 Hz), 5.38-5.48 (1H, m), 5.44 (1H, dtd, J = 16.8, 1.6 and 1.2 Hz), 6.06 (1H, ddt, J = 16.8, 11.2 and 6.0 Hz), 7.04-7.45 (19H, m), 7.40 (2H, t, J = 7.4 Hz), 7.59 (2H, dd, J = 7.2 and 4.8 Hz), 7.76 (2H, d, J = 7.5 Hz), 8.06 (2H, d, J = 8.2 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 22.1, 28.4, 29.2, 29.2, 32.5, 40.1, 47.1, 48.9, 51.3, 65.9, 66.9, 66.9, 79.1, 118.6, 119.9, 125.1, 126.6, 127.0, 127.6, 127.8, 128.3, 129.4, 129.6, 130.0, 131.2, 131.9, 141.2, 143.7, 143.8, 144.4, 144.7, 155.8, 155.9, 165.1, 171.7; HRMS (ESI-TOF) m/z calcd for $C_{57}H_{60}N_3O_7S$ $[M + H]^+$: 930.4152, found: 930.4143.

Allyl 4-[(Fmoc-L-Arg(Pbf)-2-tritylsulfanylethyl)amino]benzoate



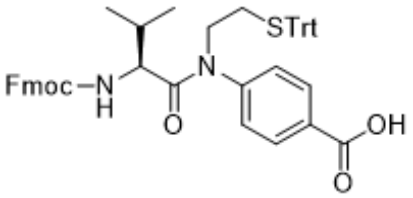
White amorphous solid, yield: 95% (1101 mg); $[\alpha]^{22}_D$ 46.3 (c 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.16-1.34 (2H, m), 1.41 (6H, s), 1.43-1.55 (2H, m), 2.06 (3H, s), 2.19-2.33 (1H, m), 2.43-2.51 (1H, m), 2.42- 2.51 (1H, m), 2.48 (3H, s), 2.54 (3H, s), 2.66-2.96 (2H, br m), 2.89 (2H, s), 3.23 (1H, ddd, J = 13.6, 9.0 and 5.4 Hz), 3.63 (1H, ddd, J = 13.6, 9.2 and 5.5 Hz), 4.06-4.20 (1H, m), 4.13 (1H, t, J = 6.4 Hz), 4.21-4.37 (2H, m), 4.85 (2H, d, J = 5.5 Hz), 5.32 (1H, d, J = 10.4 Hz), 5.43 (1H, d, J = 17.2 Hz), 5.92 (1H, d, J = 8.1 Hz), 5.98-6.16 (3H, m), 7.04-7.33 (19H, m), 7.37 (2H, t, J = 7.4 Hz), 7.54 (2H, t, J = 7.4 Hz), 7.74 (2H, d, J = 7.5 Hz), 8.04 (2H, d, J = 8.0 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 12.4, 17.9, 19.3, 24.6, 28.5, 29.1, 30.2, 40.5, 43.1, 47.0, 48.9, 50.9, 66.0, 66.9, 67.3, 86.4, 117.5, 118.7, 119.9, 124.7, 125.0, 125.1, 126.7, 127.0, 127.2, 127.7, 127.7, 127.8, 127.9, 128.3, 129.4, 130.2, 131.3, 132.5, 138.5, 141.2, 143.6, 144.4, 144.4, 144.4, 155.7, 156.3, 158.9, 165.1, 171.4; HRMS (ESI-TOF) m/z calcd for $C_{65}H_{67}KN_5O_8S_2$ $[M + K]^+$: 1148.4068, found: 1148.4044.

Allyl 4-[(Fmoc-L-Trp-2-tritylsulfanylethyl)amino]benzoate



White amorphous solid, yield: 92% (852 mg); $[\alpha]^{23}_{D}$ 82.1 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 2.16 (1H, ddd, *J* = 12.4, 8.4 and 7.2 Hz), 2.29 (1H, ddd, *J* = 12.4, 7.4 and 6.9 Hz), 2.93 (1H, dd, *J* = 13.8 and 5.4 Hz), 3.06 (1H, dd, *J* = 13.8 and 9.2 Hz), 3.27-3.47 (2H, m), 4.19 (1H, t, *J* = 7.0 Hz), 4.32 (1H, dd, *J* = 18.0 and 7.0 Hz), 4.35 (1H, dd, *J* = 18.0 and 7.0 Hz), 4.51 (1H, ddd, *J* = 9.2, 8.8 and 5.4 Hz), 4.83 (2H, d, *J* = 5.5 Hz), 5.34 (1H, d, *J* = 10.4 Hz), 5.44 (1H, d, *J* = 17.3 Hz), 5.49 (1H, d, *J* = 8.8 Hz), 6.06 (1H, ddt, *J* = 17.3, 10.4 and 5.5 Hz), 6.18-6.52 (1H, br s), 6.76 (1H, d, *J* = 1.7 Hz), 6.93 (1H, t, *J* = 7.3 Hz), 7.09-7.37 (21H, m), 7.40 (2H, t, *J* = 7.4 Hz), 7.58 (2H, dd, *J* = 7.6 and 3.6 Hz), 7.66 (2H, d, *J* = 6.9 Hz), 7.71-7.82 (1H, m), 7.77 (3H, d, *J* = 7.6 Hz) ; ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 29.1, 29.7, 47.1, 48.5, 52.0, 65.7, 66.7, 66.9, 110.0, 110.9, 118.4, 118.6, 119.6, 119.9, 122.2, 123.0, 125.1, 125.1, 126.6, 127.0, 127.4, 127.6, 127.7, 127.8, 129.3, 129.5, 129.6, 130.5, 132.0, 135.9, 141.2, 143.8, 144.3, 144.5, 155.4, 165.3, 171.7; HRMS (ESI-TOF) *m*/*z* calcd for $C_{57}H_{50}N_3O_5S$ $[M + H]^+$: 888.3471, found: 888.3444.

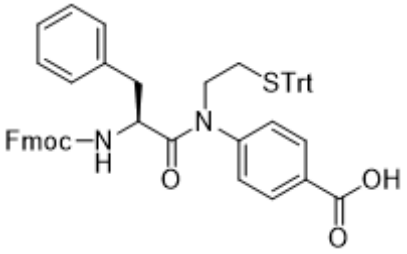
4-[(Fmoc-L-Val-2-tritylsulfanylethyl)amino]benzoic acid



White amorphous solid, yield: 80% (76 mg); $[\alpha]^{24}_{D}$ 82.8 (*c* 0.40, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 0.75 (3H, d, *J* = 6.5 Hz), 0.81 (3H, d, *J* = 6.5 Hz), 1.75-1.94 (1H, m), 2.23-2.38 (1H, m), 2.48-2.65 (1H, m), 3.39 (1H, ddd, *J* = 13.9, 8.5 and 5.4 Hz), 3.60 (1H, ddd, *J* = 13.9, 8.8 and 5.2 Hz), 4.11 (1H, dd, *J* = 9.5 and 7.2 Hz), 4.22 (1H, t, *J* = 7.0 Hz), 4.34 (1H, dd, *J* = 10.4 and 7.0 Hz), 4.43 (1H, dd, *J* = 10.4 and 7.0 Hz), 5.80 (1H, d, *J* = 9.5 Hz), 7.08 (2H, d, *J* = 8.1 Hz), 7.10-7.44 (19H, m), 7.63 (2H, dd, *J* = 7.6 and 2.8 Hz), 7.76 (2H, d, *J* = 7.2 Hz), 8.09 (2H, d, *J* = 8.1 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 17.3, 19.4, 29.1, 31.7, 47.2, 49.1, 56.5, 67.0, 67.0, 120.0, 125.1, 125.2, 126.7, 127.1, 127.3, 127.7, 127.8, 127.9, 128.4, 129.5, 131.7, 141.3, 143.8, 144.5, 145.3, 146.8, 156.1, 168.4, 171.7; HRMS (ESI-TOF) *m*/*z* calcd for $C_{48}H_{45}N_2O_5S$ $[M + H]^+$: 761.3049, found:

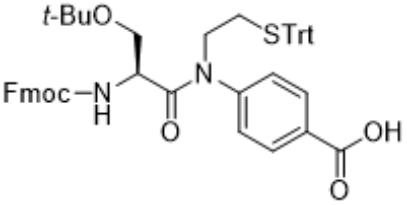
761.3057.

4-[(Fmoc-L-Phe-2-tritylsulfanylethyl)amino]benzoic acid



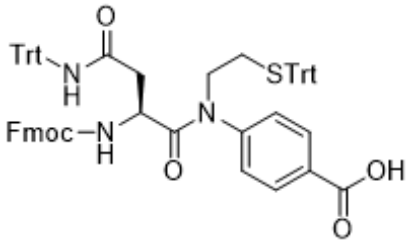
White amorphous solid, yield: 93% (176 mg); $[\alpha]^{19}_{D}$ 53.5 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 2.23-2.43 (2H, m), 2.78 (1H, dd, *J* = 12.9 and 5.9 Hz), 2.91 (1H, dd, *J* = 12.9 and 9.2 Hz), 3.28-3.45 (2H, m), 4.18 (1H, t, *J* = 7.0 Hz), 4.25-4.45 (3H, m), 5.86 (1H, d, *J* = 8.9 Hz), 6.95 (2H, d, *J* = 7.2 Hz), 7.10-7.35 (22H, m), 7.39 (2H, t, *J* = 7.2 Hz), 7.59 (2H, dd, *J* = 7.2 and 1.2 Hz), 7.76 (2H, d, *J* = 7.2 Hz), 7.93 (2H, d, *J* = 8.0 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 28.9, 39.8, 47.1, 49.0, 53.0, 67.0, 67.1, 119.9, 125.1, 125.2, 126.7, 127.0, 127.7, 127.8, 127.9, 128.1, 128.5, 128.7, 129.5, 129.6, 129.6, 131.4, 132.1, 132.2, 132.2, 135.8, 141.3, 143.7, 143.8, 144.5, 144.6, 155.5, 168.8, 171.3; HRMS (ESI-TOF) *m*/*z* calcd for $C_{52}H_{45}N_2O_5S$ $[M + H]^+$: 809.3049, found: 809.3044.

4-[(Fmoc-L-Ser(*t*-Bu)-2-tritylsulfanylethyl)amino]benzoic acid



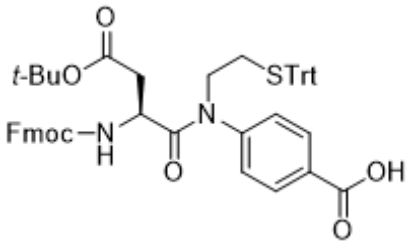
White amorphous solid, yield: quant. (187 mg); $[\alpha]^{20}_{D}$ 45.4 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.08 (9H, s), 2.35-2.55 (2H, m), 3.20-3.34 (1H, m), 3.34-3.50 (2H, m), 3.50-3.65 (1H, m), 4.19 (1H, t, *J* = 7.1 Hz), 4.32 (2H, d, *J* = 7.1 Hz), 4.35-4.46 (1H, m), 5.68 (1H, d, *J* = 8.8 Hz), 7.10-7.35 (19H, m), 7.39 (2H, t, *J* = 7.3 Hz), 7.59 (2H, dd, *J* = 6.8 and 4.0 Hz), 7.75 (2H, d, *J* = 7.3 Hz), 8.06 (2H, d, *J* = 8.3 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 27.2, 29.1, 47.0, 49.3, 51.6, 62.5, 67.0, 67.1, 73.7, 119.9, 125.2, 126.7, 127.0, 127.7, 127.8, 127.9, 128.5, 129.1, 129.5, 131.5, 141.3, 143.7, 143.8, 144.5, 145.4, 155.8, 169.3, 169.3; HRMS (ESI-TOF) *m*/*z* calcd for $C_{50}H_{49}N_2O_6S$ $[M + H]^+$: 805.3311, found: 805.3332.

4-[(Fmoc-L-Asn(Trt)-2-tritylsulfanylethyl)amino]benzoic acid



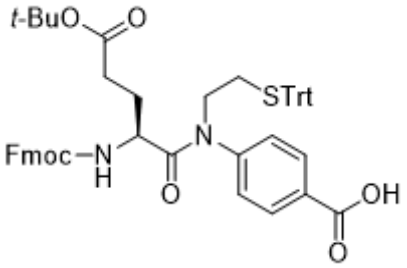
White amorphous solid, yield: 83% (160 mg); $[\alpha]^{24}_{D}$ −17.0 (*c* 0.40, $CHCl_3$); ^1H NMR ($CDCl_3$, 400 MHz) δ = 2.28-2.47 (2H, t, *J* = 6.7 Hz), 2.63-2.86 (2H, m), 3.15-3.46 (2H, m), 4.11 (1H, t, *J* = 6.7 Hz), 4.23 (1H, dd, *J* = 10.3 and 6.7 Hz), 4.30 (1H, dd, *J* = 10.3 and 6.7 Hz), 4.39-4.69 (1H, m), 6.34 (1H, d, *J* = 7.6 Hz), 6.77 (1H, s), 6.92 (2H, d, *J* = 7.4 Hz), 7.04-7.34 (32H, m), 7.36 (2H, t, *J* = 7.4 Hz), 7.55 (2H, dd, *J* = 13.6 and 7.2 Hz), 7.74 (2H, d, *J* = 7.4 Hz), 7.82 (2H, d, *J* = 7.4 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 28.9, 39.3, 47.0, 49.5, 49.5, 67.0, 67.0, 70.7, 119.9, 125.2, 125.2, 126.6, 127.0, 127.0, 127.2, 127.7, 127.8, 127.9, 128.7, 129.4, 129.6, 131.5, 141.3, 143.5, 143.8, 144.2, 144.5, 144.7, 155.6, 168.2, 169.0, 170.3; HRMS (ESI-TOF) *m/z* calcd for $C_{66}H_{56}N_3O_6S$ $[M + H]^+$: 1018.3890, found: 1018.3877.

4-[(Fmoc-L-Asp(O*t*-Bu)-2-tritylsulfanylethyl)amino]benzoic acid



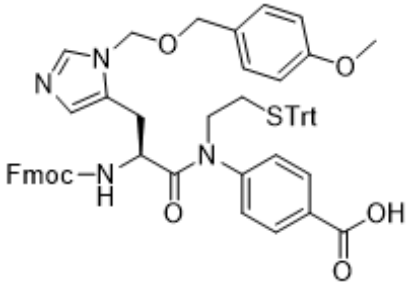
White amorphous solid, yield: 98% (188 mg); $[\alpha]^{23}_{D}$ 80.3 (*c* 1.00, $CHCl_3$); ^1H NMR ($CDCl_3$, 400 MHz) δ = 1.37 (9H, s), 2.23-2.44 (2H, m), 2.44-2.60 (2H, m), 3.29-3.46 (1H, m), 3.46-3.61 (1H, m), 4.17 (1H, t, *J* = 7.1 Hz), 4.29 (2H, d, *J* = 7.1 Hz), 4.45-4.61 (1H, br m), 5.77 (1H, d, *J* = 9.1 Hz), 7.07-7.35 (19H, m), 7.38 (2H, t, *J* = 7.4 Hz), 7.58 (2H, dd, *J* = 7.6 and 2.0 Hz), 7.75 (2H, d, *J* = 7.4 Hz), 8.06 (2H, d, *J* = 8.2 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 28.0, 29.1, 38.3, 47.0, 49.0, 49.3, 67.0, 67.2, 81.5, 120.0, 125.2, 126.7, 127.0, 127.7, 127.8, 128.2, 128.5, 128.6, 129.5, 131.7, 132.1, 132.2, 141.3, 143.7, 143.8, 144.5, 145.2, 155.4, 169.1, 169.4, 170.1; HRMS (ESI-TOF) *m/z* calcd for $C_{51}H_{49}N_2O_7S$ $[M + H]^+$: 833.3260, found: 833.3247.

4-[(Fmoc-L-Glu(O*t*-Bu)-2-tritylsulfanylethyl)amino]benzoic acid



White amorphous solid, yield: 86% (164 mg); $[\alpha]^{24}_{D}$ 57.0 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 1.38 (9H, s), 1.66-1.87 (2H, m), 2.00-2.24 (2H, m), 2.25-2.40 (1H, m), 2.45-2.60 (1H, m), 3.30-3.45 (1H, m), 3.58 (1H, ddd, *J* = 13.4, 8.8 and 5.9 Hz), 4.20 (1H, t, *J* = 7.2 Hz), 4.16-4.28 (1H, m), 4.35 (2H, d, *J* = 7.2 Hz), 5.92 (1H, d, *J* = 8.3 Hz), 7.06-7.37 (19H, m), 7.39 (2H, t, *J* = 7.4 Hz), 7.61 (2H, dd, *J* = 7.6 and 2.4 Hz), 7.76 (2H, d, *J* = 7.4 Hz), 8.10 (2H, d, *J* = 8.0 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 25.6, 28.0, 29.1, 30.8, 47.1, 49.0, 51.1, 67.1, 68.0, 80.7, 119.9, 125.2, 126.7, 127.0, 127.2, 127.7, 127.8, 127.9, 128.3, 128.6, 129.5, 131.8, 132.1, 141.3, 143.7, 143.8, 144.5, 144.9, 155.9, 168.6, 171.3, 172.0; HRMS (ESI-TOF) *m*/*z* calcd for $C_{52}H_{51}N_2O_7S$ $[M + H]^+$: 847.3417, found: 847.3391.

4-[(Fmoc-L-His(MBom)-2-tritylsulfanylethyl)amino]benzoic acid



White amorphous solid, yield: 90% (52 mg); $[\alpha]^{20}_{D}$ 27.8 (*c* 1.00, $CHCl_3$); ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 2.16-2.30 (1H, m), 2.38-2.54 (1H, m), 2.81 (1H, dd, *J* = 14.8 and 6.6 Hz), 2.97 (1H, d, *J* = 14.8 and 7.7 Hz), 3.24-3.38 (1H, m), 3.52-3.67 (1H, m), 3.73 (3H, s), 4.08-4.26 (1H, m), 4.13 (1H, d, *J* = 11.6 Hz), 4.20 (1H, d, *J* = 11.6 Hz), 4.28-4.48 (3H, m), 4.85 (1H, d, *J* = 10.7 Hz), 4.96 (1H, d, *J* = 10.7 Hz), 5.63 (1H, d, *J* = 8.5 Hz), 6.65-6.92 (1H, m), 6.71 (1H, s), 6.83 (2H, d, *J* = 8.4 Hz), 7.06-7.44 (22H, m), 7.50-7.64 (3H, m), 7.66-7.79 (2H, m), 8.07 (2H, d, *J* = 8.2 Hz); ^{13}C NMR ($CDCl_3$, 75 MHz) δ = 27.4, 29.1, 47.1, 48.7, 51.6, 55.2, 66.9, 69.5, 72.9, 114.1, 120.0, 125.1, 126.7, 126.8, 127.1, 127.6, 127.7, 127.8, 127.9, 129.5, 129.7, 131.4, 131.9, 137.7, 141.3, 143.5, 143.7, 144.4, 155.4, 159.6, 168.3, 170.7; HRMS (ESI-TOF) *m*/*z* calcd for $C_{58}H_{53}N_4O_7S$ $[M + H]^+$: 949.3635, found: 949.3661.

Typical experimental protocol for the check of racemization during coupling reaction of Fmoc amino acid to the aniline linker

Fmoc-Ser(*t*-Bu)-linked aniline linker **7** (50 mg, 0.059 mmol) was treated with 20% piperidine in DMF (0.5 mL) for 10 min at room temperature. The resulting residue was then evaporated under reduced pressure. A separately prepared solution of Boc-L/D-Leu-OH (16 mg each, 0.065 mmol), HBTU (22 mg, 0.059 mmol) and DIPEA (10 mL, 0.059 mmol) dissolved in DMF (0.5 mL) and stirred at room temperature for 10 min was added to the residue and allowed to react for 30 min. The reaction was then quenched by adding sat. $NaHCO_3$ aq., and the organic layer was extracted with EtOAc. The organic layer was washed with sat. $NaHCO_3$ aq., 10% citric acid aq., and brine, and dried over $MgSO_4$. The resulting samples obtained by coupling with Boc-L-Leu-OH or Boc-D-Leu-OH were analyzed by RP-HPLC. The major coupling product **10b** (D-Leu-L-Ser dipeptide) obtained from Boc-D-Leu coupling corresponds to the epimer of the racemization product **10a'** (L-Leu-D-Leu dipeptide). Therefore, in the HPLC analysis of the Boc-L-Leu-coupling sample, the minor peak corresponding to the D-Leu-L-Ser dipeptide **10b** was used as the peak of the racemic product **10a'** to determine the degree of racemization.

Typical experimental protocol for the synthesis of SEAlide peptides

NovaSyn® TGR resin (Rink amide type: 0.25 mmol amine/g, 0.40 g, 0.10 mmol) was added to a reaction vessel. Fmoc-Leu-OH (100 mg, 0.30 mmol) and DIPCDI (48 μL, 0.30 mmol) and HOBt·H_2O (48 mg, 0.30 mol) in DMF were added. The reaction mixture was shaken for 2 hours at room temperature. After washing with DMF, Fmoc deprotection was performed with 20% piperidine in DMF to obtain the Leu-incorporated resin (H-Leu-NovaSyn® TGR resin). Subsequently, the resin was treated with a mixture of an acid-type *N*-Fmoc amino acyl-*N*-sulfanylethylaniline linker (Fmoc-**X**-*NAr*-OH: 0.20 mmol), HATU (76 mg, 0.20 mmol), and DIPEA (36 mL, 0.20 mmol) for 3 hours at room temperature to obtain an Fmoc-**X**-*NAr*-incorporated resin. The resin was subjected to standard Fmoc-SPPS as described in General Methods. The resulting protected peptide resin for model SEAlide peptide was exposed to TFA-based cocktail with TFA–thioanisole–*m*-cresol–EDT–H_2O (*v/v*, 80:5:5:5:5 or 90:2.5:2.5:2.5:2.5) for 90–120 min at room temperature. The resulting product was filtered, treated with cold Et_2O, and the precipitate was collected by centrifugation, washed with Et_2O, and model SEAlide peptide **11** (H-**VQGSX**-*NAr*-**L**-NH_2: **X** = **G** (**11a**), **A** (**11b**), **V** (**11c**), **F** (**11d**), **S** (**11e**), **N** (**11f**), **D**, **E** (**11g**), and **H** (**11h**)) was obtained. The crude materials, except for **D** (**11i**),

were purified using RP-HPLC with a linear gradient of a solvent system consisting of 0.1% TFA in H_2O and 0.1% TFA in CH_3CN. For Asp-containing peptide **11i**, liner gradient of CH_3CN in 10 mM NH_4OAc buffer was used.

11a (**X** = **G**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 18.4 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{32}H_{52}N_9O_9S$ $[M + H]^+$: 738.4, found 738.1.

11b (**X** = **A**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 18.7 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{33}H_{54}N_9O_9S$ $[M + H]^+$: 752.4, found 752.2.

11c (**X** = **V**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 20.8 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{35}H_{58}N_9O_9S$ $[M + H]^+$: 780.4, found 780.5.

11d (**X** = **F**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 23.4 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{39}H_{58}N_9O_9S$ $[M + H]^+$: 828.4, found 828.6.

11e (**X** = **S**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 17.5 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{33}H_{54}N_9O_{10}S$ $[M + H]^+$: 768.4, found 768.1.

11f (**X** = **N**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 17.1 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{34}H_{55}N_{10}O_{10}S$ $[M + H]^+$: 795.4, found 795.2.

11g (**X** = **E**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 18.9 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{35}H_{56}N_9O_{11}S$ $[M + H]^+$: 810.4, found 810.2.

11h (**X** = **H**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 10 to 40% over 30 min, retention time = 15.0 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{36}H_{56}N_{11}O_9S$ $[M + H]^+$: 818.4, found 818.4.

11i (early eluted): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 19.0 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{45}H_{75}N_{16}O_{14}S$ $[M + H]^+$ 1095.5, found 1095.5.

14 (later eluted): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 19.4 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{45}H_{75}N_{16}O_{14}S$ $[M + H]^+$ 1095.5, found 1095.6.

Synthesis of N–terminal Cys-containing peptide 12

N-Terminal Cys-containing peptide **12** was prepared by standard Fmoc-SPPS

protocol on NovaSyn® TGR resin (Rink amide type: 0.25 mmol amine/g, 0.40 g, 0.10 mmol).

Typical experimental protocol for NCLs of SEAlide Peptides 11 with N–Terminal Cys-containing peptide 12

The NCL reaction between SEAlide peptide **11** (1.0 mM) and N-terminal Cys-containing peptide **12** (1.0 mM) was performed at 37 °C in 3.0 M Gn·HCl-0.5 M sodium phosphate buffer, pH 7.3, with the addition of 30 mM MPAA and 40 mM TCEP. Analysis of each reaction was performed by collecting 5 μL of the reaction mixture at each time point and subjecting it to reverse-phase HPLC analysis using a Cosmosil 5C_{18}-AR-II analytical column (4.6×250 mm; detection at 220 nm) with a concave gradient (curve 7 of the Waters 600E) of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min. The progress of ligation was quantified by integration of the ligated products **13** as a fraction of the sum of the unreacted N–terminal Cys-containing peptide and the ligated products. After 24-72 hours, each reaction was purified by preparative reverse-phase HPLC using a Cosmosil 5C18-AR-II preparative column (20×250 mm) to obtain the desired ligation peptide. Epimerization during the coupling of **13b**, **13e**, and **13h** was analyzed by comparing the HPLC results of crude ligation products with those of D-amino acid-containing reference peptides synthesized by standard Fmoc-SPPS.

13a (**X** = **G**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 12.0 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{43}H_{73}N_{16}O_{12}S$ $[M + H]^+$: 1037.5, found 1037.4.

13b (**X** = **A**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 12.3 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{44}H_{75}N_{16}O_{12}S$ $[M + H]^+$: 1051.5, found 1051.6.

13c (**X** = **V**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 10 to 40% over 30 min, retention time = 12.3 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{46}H_{79}N_{16}O_{12}S$ $[M + H]^+$: 1079.6, found 1079.8.

13d (**X** = **F**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 35% over 30 min, retention time = 18.7 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{50}H_{79}N_{16}O_{12}S$ $[M + H]^+$: 1127.6, found 1128.0.

13e (**X** = **S**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 12.0 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{44}H_{75}N_{16}O_{13}S$ $[M + H]^+$: 1067.5, found 1067.8.

13f (**X** = **N**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 12.2 min, LRMS (ESI-Ion Trap) *m*/*z*

calcd for $C_{45}H_{76}N_{17}O_{13}S$ $[M + H]^+$: 1094.6, found 1094.2.

13g (**X** = **E**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 10 to 25% over 30 min, retention time = 11.9 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{46}H_{77}N_{16}O_{14}S$ $[M + H]^+$: 1109.6, found 1109.2.

13h (**X** = **H**): Analytical HPLC conditions: linear gradient of 0.1% TFA/CH_3CN in 0.1% TFA/H_2O, 5 to 45% over 30 min, retention time = 16.7 min, LRMS (ESI-Ion Trap) *m*/*z* calcd for $C_{47}H_{77}N_{18}O_{12}S$ $[M + H]^+$: 1117.6, found 1117.6.

Chapter 2

Sequences of the target proteins

Adiponectin(19-107): H-**ET**^{20} **TTQGPGVLLP**^{30} **LPKGACTGWM**^{40} **AGIO*GHO*GHN**^{50} **GAO*GRDGRDG**^{60} **TPGEKGEKGD**^{70} **PGLIGPKGDI**^{80} **GETGVPGAEG**^{90} **O*RGFPGIQGR**^{100} **KGEPGEG**-OH, **O*** = (2*S*, 4*R*)-4-hydroxyproline

[Ser(PO_3H_2)^{65}]-ubiquitin: H-**MQIFVKTLTG**^{10} **KTITLEVEPS**^{20} **DTIENVKAKI**^{30} **QDKEGIPPDQ**^{40} **QRLIFAGKQL**^{50} **EDGRTLSDYN**^{60} **IQKES*TLHLV**^{70} **LRLRGG-OH**, **S*** = Ser(PO_3H_2)

Syntheses of fragments for model peptides

*H-**LYRANA**-MPA-**L**-NH$_2$* (**16a**)

The peptide was synthesized using Boc-Ala-MPA-Leu-MBHA resin (0.15 mmol) by automated Boc-SPPS procedure as described in the General Methods. After treatment with HF/*p*-cresol (v/v, 90/10) at −2 °C to −5 °C for 1 hour, the peptide resin was purified by preparative HPLC. The desired compound **16a** was obtained in an isolated yield of 16 % (22 mg). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was greater than 98% (t_R = 12.5 min); LRMS (ESI+) calcd $[M+H]^+$ 907.5, found 907.5.

*H-**LYRANG**-MPA-**L**-NH$_2$* (**16b**)

The peptide **16b** was synthesized in the same manner as described for **16a** (51 mg, 38%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was greater than 98% (t_R = 12.0 min); LRMS (ESI+) calcd $[M+H]^+$ 893.5, found 893.5.

*H-**LYRANF**-MPA-**L**-NH$_2$* (**16c**)

The peptide **16c** was synthesized in the same manner as described for **16a** (40 mg, 27%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was greater than 98% (t_R = 15.5 min); LRMS (ESI+) calcd $[M+H]^+$ 983.5, found 983.5.

*H-**LYRANL**-MPA-**L**-NH$_2$* (**16d**)

The peptide **16d** was synthesized in the same manner as described for **16a** (19 mg, 13%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was greater than 98% (t_R = 15.1 min); LRMS (ESI+) calcd $[M+H]^+$ 949.5, found 949.5.

*H-**LYRANT**-MPA-**L**-NH_2* (**16e**)

The peptide **16e** was synthesized in the same manner as described for **16a** (82 mg, 35%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was greater than 98% (t_R = 12.2 min); LRMS (ESI+) calcd $[M+H]^+$ 937.5, found 937.4.

*H-**LYRANV**-MPA-**L**-NH_2* (**16f**)

The peptide **16f** was synthesized in the same manner as described for **16a** (25 mg, 18%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was greater than 98% (t_R = 14.0 min); LRMS (ESI+) calcd $[M+H]^+$ 935.5, found 935.5.

*H-**CSPGYS**-NH_2* (**17**)

The peptide **17** was synthesized by Fmoc-SPPS as described in General Methods. (110 mg, 72%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was greater than than 98% (t_R = 8.1 min); LRMS (ESI+) calcd $[M+H]^+$ 612.2, found 612.2.

Typical procedure for kinetics studies of Im-NCL

A ligation buffer (6.0 M Gn·HCl, 30 mM TCEP, pH 7.1) containing various additives was added to peptide thioesters **16a-f** (3.0 mM) and Cys-peptide **17** (2.0 mM). A 5.0 M imidazole buffer with 3.0 M Gn·HCl was used instead of 6.0 M Gn·HCl due to the solubility problem. The NCL reaction without additives was performed using a ligation buffer (100 mM Na_2HPO_4, 6.0 M Gn·HCl, 30 mM TCEP, pH 7.1). The ligation solution was incubated at room temperature, and the reaction progress was monitored by taking aliquots at different time points. Evaluation of the reaction progress was performed by quenching the reaction with a 5% TFA solution in water and a 0.5 M TCEP solution from Sigma-Aldrich, followed by quantification of the unreacted Cys-peptide **17** and the desired ligation products (**18a-f**) at λ=280 nm using analytical HPLC.

Syntheses of fragments of adiponectin

Adiponectin(19-35)-MPA-Arg-Arg-Arg (**19**)

The peptide **19** was synthesized using Boc-Ala-MPA-Arg-Arg-Arg-PAM resin (0.3 mmol) by automated Boc-SPPS procedure as described in the General Methods. After treatment with HF/*p*-cresol (v/v, 80/20) at −6 °C for 1 hour, the peptide resin was purified by preparative HPLC. The desired compound **19** was obtained in an isolated yield of 60% (400 mg). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was greater than 98% (t_R = 15.8 min); LRMS (ESI+) calcd 2235, found 2235.

Adiponectin(36-107) (**20**)

The peptide **20** was synthesized using Fmoc-Gly-Wang-PEG-resin (0.2 mmol) by automated Fmoc-SPPS procedure as described in the General Methods. The resulting peptide resin was deprotected by TFA/H_2O/TIS/thiophenol (v/v, 95/5/5) at room temperature for 2 hours and then purified by preparative HPLC (71 mg, 8.8%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 90% (t_R = 15.0 min); LRMS (ESI+) calcd 7059, found 7060.

Adiponectin(19–107) (**21**)

Adiponectin(19-35)-MPA-Arg-Arg-Arg **19** (5.1 mg, 1.9 μmol) and adiponectin(36-107) **20** (10 mg, 1.3 μmol) were dissolved in a ligation buffer (6.0 M Gn·HCl, 2.5 M imidazole, 30 mM TCEP, pH 7.1) to a final volume of 0.80 mL, followed by incubation at 37 °C for 3 hours. Subsequently, dithiothreitol (1.9 mg, 13 μmol) was added and the reaction mixture was incubated for 30 minutes. The resulting solution was acidified to pH 1–2 using TFA and subjected to preparative HPLC purification using a 0.1% aqueous TFA-CH_3CN system. After lyophilization, the desired protein was obtained as a white amorphous powder (11 mg, 85%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was higher than 98% (t_R = 17.0 min), LRMS (ESI+) m/z calcd 8719; found: 8720.

Syntheses of fragments of [Ser(PO_3H_2)^{65}]-ubiquitin

*[Ser(PO_3H_2)^{65}]-ubiquitin (1-45)-MPA-**L**-NH_2* (**22**)

The peptide **22** was synthesized using automated Boc-SPPS protocol on MBHA resin, as described in the General Methods. The resulting peptide resin was deprotected by HF/*p*-cresol (v/v, 85/15) containing 5 equivalents of thiophenol at −4°C for 1 hour and then purified via preparative HPLC (110 mg, 14%). Analytical HPLC (1–60%

CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 94% (t_R = 20.1 min); LRMS (ESI+) calcd 5313, found 5313.

[Cys^{46}, Ser(PO$_3$H$_2$)^{65}]-ubiquitin(46-76) (**23**)

The peptide **23** was synthesized using an automated Fmoc-SPPS protocol on a Wang-PEG resin, as described in the General Methods. The resulting peptide resin was deprotected by TFA/H_2O/TIS/thiophenol (v/v, 92.5/2.5/2.5/2.5) with 20 eq. $MeONH_2 \cdot HCl$[40] at rt for 2 hours and then purified by preparative HPLC (140 mg, 15%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 96% (t_R = 16.9 min); LRMS (ESI+) calcd 3580, found 3581.

Synthesis of [Ser(PO$_3$H2)$_{65}$]-ubiquitin (**25**)

[Ser(PO$_3$H$_2$)^{65}]-ubiquitin (1-45)-MPA-**L**-NH_2 **22** (4.9 mg, 0.80 μmol) and [Cys^{46}, Ser(PO$_3$H$_2$)^{65}]-ubiquitin(46-76) **23** (5.0 mg, 1.1 μmol) were dissolved in 0.8 mL of a ligation buffer (6.0 M Gn·HCl, 2.5 M imidazole, 30 mM TCEP, pH 7.1). The resulting solution was incubated at room temperature for 3 hours. The reaction was complete with approximately 1% epimerization. This was confirmed by comparison with an independently prepared epimer (Figure S2 in the Experimental sections). The ligation mixture was then combined with a desulfurization buffer (0.80 mL; 6.0 M Gn·HCl, 0.1 M Na_2HPO_4, 450 mM TCEP, and 80 mM glutathione, pH 4.4) and an aqueous solution (90 mL) of VA-044 (30 mg, 93 mM). The resulting solution (final pH 6.4) was incubated at 37 °C and monitored by LC-MS until the conversion was completed (15 hours). After completion, the mixture was acidified to pH 1–2 with TFA, and subjected to preparative HPLC using a 0.1% aqueous TFA-CH_3CN system. The desired fractions were collected, lyophilized, and yielded the desired protein as a white amorphous powder (5.6 mg 69%). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): purity was higher than 98% (t_R =18.9 min); LRMS (ESI+) calcd 8643; found: 8644.

Chapter 3

Sequences of the target proteins

BNP(1-32): H-**SPKMVQGSGC**[10] **FGRKMDRISS**[20] **SSGLGCKVLR**[30] **RH**-OH,
Model peptide **11**: H-**HPGSRIVLSL**[10] **DVPIGLLQIL**[20] **LEQARARCAR**[30] **EQATTNARIL**[40] **ARV**-NH_2

Syntheses of fragments for model peptides

*H-**LYRANA**-MeNbz-**L**-NH_2* (**26a**)

The peptide was synthesized using automated Fmoc-SPPS protocol on a Fmoc-Ala-MeNbz-Leu-Rink amide resin (0.10 mmol) as described in the General Methods. Boc-Leu was introduced for the N-terminal amino acid. Subsequently, the resulting peptide on the resin was activated using a solution of *p*-nitrophenyl chloroformate solution (1 mmol) in CH_2Cl_2 for 30 minutes. The resin was then washed with CH_2Cl_2 and treated with DIEA 0.5 M in DMF for 15 minutes. The resulting resin was deprotected using TFA/TIS/H_2O (v/v, 95/2.5/2.5) for 2 hours. Subsequently, the crude product was purified by preparative HPLC to yield **26a** [70 mg, 0.057 mmol, 57% (calculated as **26a**·2TFA, M_{calcd}: 1221)]. Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): purity was 97% (t_R = 15.5 min); LRMS (ESI+) calcd $[M+H]^+$ 993.5, found 993.5.

*H-**LYRANF**-MeNbz-NH_2* (**26b**)

The peptide **26b** was synthesized in the same manner as described for **26a** (33 mg). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 92% (t_R = 16.1 min); LRMS (ESI+) calcd $[M+H]^+$ 956.5, found 956.4.

*H-**LYRANN**-MeNbz-**L**-NH_2* (**26c**)

The peptide **26c** was synthesized in the same manner as described for **26a** (26 mg). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 91% (t_R = 14.4 min); LRMS (ESI+) calcd $[M+H]^+$ 1037, found 1037.

*H-**LYRANL**-MeNbz-NH_2* (**26d**)

The peptide **26d** was synthesized in the same manner as described for **26a** (30

mg). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 96% (t_R = 15.6 min); LRMS (ESI+) calcd $[M+H]^+$ 922.5, found 922.5.

*H-**LYRANV**-MeNbz-**L**-NH_2* (**26e**)

The peptide **26e** was synthesized in the same manner as described for **26a** (69 mg). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 98% (t_R = 17.1 min); LRMS (ESI+) calcd $[M+H]^+$ 1022, found 1022.

*H-**CSPGYS**-NH_2* (**27**)

The peptide **27** was synthesized by Fmoc-SPPS as described in General Methods (110 mg). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 97% (t_R = 9.0 min); LRMS (ESI+) calcd $[M+H]^+$ 612.2, found 612.2.

General procedure of kinetics investigation of triazole-aided NCL

Peptide-MeNbz **26** (3.0 mM) and Cys-peptide **27** (2.0 mM) were dissolved in ligation buffer (6.0 M Gn·HCl, 30 mM TCEP, 0.10 M Na_2HPO_4, pH 7.1) with various additives. For NCL with imidazole, ligation buffer (pH 7.1) containing 6.0 M Gn·HCl, 30 mM TCEP was used. The reaction mixture was incubated at 37 °C. For analysis of each reaction, an aliquot of 5.0 μL was taken from the reaction mixture at each time point and quenched with 45 μL of 5% TFA aq. and 5.0 μL of neutral TCEP solution (0.5 M, Sigma-Aldrich), and analyzed by analytical HPLC. The degree of ligation was quantified by detecting the unreacted Cys-peptide **27** and the desired ligated product **28** with detection at 280 nm.

Syntheses of fragments of BNP(1-32)

BNP(1-25)-MeNbz-Leu-NH_2 (**30**)

The peptide **30** was synthesized in the same manner as described for **26a** (28 mg). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 96% (t_R = 17.1 min); LRMS (ESI+) calcd 2858, found 2858.

BNP(26-32) (**31**)

The peptide **31** was synthesized by Fmoc-SPPS as described in General Methods (71 mg). Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 97% (t_R = 11.0 min); LRMS (ESI+) calcd $[M+H]^+$ 911.5, found 911.5.

BNP(1-32) (**33**)

A solution of BNP(1-25)-MeNbz-Leu-NH_2 (**30**) (1.7 μmol) and BNP(26-32) (**31**) (0.93 μmol) in 0.70 mL ligation buffer (6.0 M Gn·HCl, 2.5 M 1,2,4-triazole, 30 mM TCEP, 0.10 M Na_2HPO_4, pH 7.1) was incubated at room temperature for 4 hours. The reaction mixture was then diluted with 50% aq. AcOH (6.3 mL) and treated with 0.45 mL of a solution of 0.10 M iodine in MeOH. The reaction was quenched by adding 1.0 M ascorbate aq., and the resulting product was purified by preparative HPLC using a 0.1% aqueous TFA–CH_3CN system. The resulting product was then lyophilized to give a white amorphous powder **33** [3.0 mg, 0.69 μmol, 73% (calculated as 78TFA, M_{calcd}: 4376)]. Analytical HPLC (1–60% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): purity was 98% (t_R = 14.5 min); LRMS (ESI+) m/z calcd: 3463; found: 3463.

Syntheses of fragments of Model peptide 37

*H-**HPGSRIVLSL**^{10} **DVPIGLLQIL**^{20} **LEQARAR**-MeNbz-**L**-NH_2* (**35**)

The peptide **35** was synthesized in the same manner as described in ref. 4 (47 mg). Analytical HPLC (10–80% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 95% (t_R = 18.8 min); LRMS (ESI+) calcd 3251, found 3251.

*H-**CAR**^{30} **EQATTNARIL**^{40} **ARV**-NH_2* (**36**)

The peptide **36** was synthesized in the same manner as described in ref. 4 (168 mg). Analytical HPLC (10–80% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): HPLC purity was 92% (t_R = 10.0 min); LRMS (ESI+) calcd 1771, found 1771.

*H-**HPGSRIVLSL**^{10} **DVPIGLLQIL**^{20} **LEQARARCAR**^{30} **EQATTNARIL**^{40} **ARV**-NH_2* (**37**)

Peptide-MeNbz **35** (2.6 μmol) and Cys-peptide **36** (1.6 μmol) were dissolved in 1.0 mL of a ligation buffer (6.0 M Gn·HCl, 2.5 M 1,2,4-triazole, 5.0 mM TCEP, 0.10 M Na_2HPO_4, pH 7.1). The resulting solution was incubated at room temperature for 8 hours, and TCEP was quenched by addition of 4-azidobenzoic acid (25 mmol). Thiopropyl–Sepharose 6B (1.0 mL, 25 mmol/mL) was then added to the solution, and the mixture was stirred at room temperature for 1 hours. The Sepharose resin was sequentially washed with H_2O, 0.1% TFA aq, 50% AcOH aq, and H_2O. The bound peptide was released from the resin using a reductive buffer (10 mM DTT, 10 mM AcONH4, pH 7.3). The resulting solution was desalted with Sephadex G-10 using 0.1% TFA in H_2O. After lyophilization, the desired compound was obtained as a white amorphous powder [7.2 mg, 1.1 μmol, 68% (corrected by the results of amino acid

analysis)]. Analytical HPLC (10–80% CH_3CN/0.1% TFA for 25 min, detected at 220 nm): purity was 92% (t_R = 17.1 min); LRMS (ESI+) calcd: 4719; found: 4719. Amino acid analysis: Asp 2.00 (2), Thr 1.90 (2), Ser 1.79 (2), Glu 4.96 (5), Pro 1.94 (2), Cys 0.64 (1), Gly 2.05 (2), Ala 5.95 (6), Val 2.66 (3), Ile 3.57 (4), Leu 6.79 (7), NH3 8.29 (5), His 0.95 (1), Arg 5.90 (6).

References

1. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776−779.
2. For reviews of NCL: (a) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923−960. (b) Kent, S. B. H. *Chem. Soc. Rev.* **2009**, *38*, 338−351. (c) Otaka, A.; Sato, K.; Ding, H.; Shigenaga, A. *Chem. Rec.* **2012**, *12*, 479−490. (d) Wilson, R. M.; Dong, S.; Wang, P.; Danishefsky, S. J. *Angew. Chem. Int. Ed*. **2013**, *52*, 7646−7665. (e) Hojo, H. *Curr. Opin. Struct. Biol*. **2014**, *26*, 16−23. (f) Malins, L. R.; Payne, R. J. *Curr. Opin. Chem. Biol.* **2014**, *22*, 70−78. (g) Rohrbacher, F.; Wucherpfennig, T. G.; Bode, J. W. *Top. Curr. Chem.* **2014**, *363*, 1−31. (h) Bondalapati, S.; Jbara, M.; Brik, A. *Nat. Chem.* **2016**, *8*, 407−418. (i) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J.-C. M.; Melnyk, O. *Chem. Rev*. **2019**, *119*, *12*, 7328–7443.
3. (a) Hojo, H.; Aimoto, S. Bull. *Chem. Soc. Jpn.* **1991**, *64*, 111–117. (b) Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180–193. (c) Aimoto, S. *Peptide Sci.* **1999**, *51*, 247–265. (d) Bang, D.; Pentelute, B. L.; Gates, Z. P.; Kent, S. B. *Org. Lett.* **2006**, *8*, 1049–1052.
4. For a review of the synthesis of peptide thioesters by Fmoc-SPPS: (a) Li, X.; Kawakami, T.; Aimoto, S. *Tetrahedron Lett.* **1998**, *19*, 8669–8672. (b) Macmillan, D.; Adams, A.; Remdjee, B. *Isr. J. Chem.* **2011**, *51*, 885–899. (c) Mende, F.; Seitz, O. *Angew. Chem. Int. Ed.* **2011**, *50*, 1232–1240. (d) Zheng, J. S.; Tang, S.; Huang, Y. C.; Liu, L. *Acc. Chem. Res.* **2013**, *46*, 2475–2484. (e) Yan, B.; Shi, W.; Ye, L.; Liu, L. *Curr. Opin. Chem. Biol.* **2018**, *46*, 33–40.
5. (a) Blanco-Canosa, J. B.; Dawson, P. E. *Angew. Chem. Int. Ed.* **2008**, *47*, 6851–6855. (Nbz) (b) Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. *J. Am. Chem. Soc.* **2015**, *137*, 7197–7209. (MeNbz)
6. Fang, G. M.; Li, Y. M.; Shen, F.; Huang, Y. C.; Li, J. B.; Lin, Y.; Cui, H. K.; Liu, L. *Angew. Chem. Int. Ed.* **2011**, *50*, 7645–7649.
7. Tsuda, S.; Shigenaga, A.; Bando, K.; Otaka, A. *Org. Lett.* **2009**, *11*, 823−826.
8. (a) Sato, K.; Shigenaga, A.; Tsuji, K.; Tsuda, S.; Sumikawa, Y.; Sakamoto, K.; Otaka, A. *ChemBioChem* **2011**, *12*, 1840−1844. (b) Ding, H.; Shigenaga, A.; Sato, K.; Morishita, K.; Otaka, A. *Org. Lett.* **2011**, *13*, 5588−5591. (c) Otaka, A.; Sato, K.; Ding, H.; Shigenaga, A. *Chem. Rec.* **2012**, *12*, 479−490.

9. Sato, K.; Shigenaga, A.; Kitakaze, K.; Sakamoto, K.; Tsuji, D.; Itoh, K.; Otaka, A. *Angew. Chem. Int. Ed.* **2013**, *52*, 7855–7859.
10. Tsuji K, Tanegashima K, Sato K, Sakamoto K, Shigenaga A, Inokuma T, Hara T, Otaka A, *Bioorg. Med. Chem.* **2015**, *23*, 5909–5914.
11. (a) Carpino, L. A.; Sadat-Aalaee, D.; Chao, H. G.; DeSelms, R. H. *J. Am. Chem. Soc.* **1990**, *112*, 9651–9652. (b) Carpino, L. A.; Beyermann, M.; Wenschuh, H.; Bienert, M. *Acc. Chem. Res.* **1996**, *29*, 268–274.
12. Brain, E. G.; Doyle, F. P.; Mehta, M. D.; Miller, D.; Nayler, J. H. C.; Stove, E. R. *J. Chem. Soc.* **1963**, 491–497.
13. (a) Rijkers, D. T. S.; Adams, H. P. H. M.; Hemker, H. C.; Tesser, G. I. *Tetrahedron* **1995**, *51*, 11235–11250. (b) Hojo, K.; Maeda, M.; Iguchi, S.; Smith, T.; Okamoto, H.; Kawasaki, K. *Chem. Pharm. Bull.* **2000**, *48*, 1740–1744. (c) Ginn, J. D.; Bosanac, T.; Chen, R.; Cywin, C.; Hickey, E.; Kashem, M.; Kerr, S.; Kugler, S.; Li, X.; Prokopowicz Iii, A.; Schlyer, S.; Smith, J. D.; Turner, M. R.; Wu, F.; Young, E. R. R. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 5153–5156.
14. (a) Pollock, S. B.; Kent, S. B. H. *Chem. Commun.* **2011**, *47*, 2342–2344. (b) Durek, T.; Alewood, P. F. *Angew. Chem., Int. Ed.* **2011**, *50*, 12042–12045.
15. Wieland, T.; Heinke, B. *Justus Liebigs Ann. Chem.* **1956**, *599*, 70–80.
16. Hibino, H.; Nishiuchi, Y. *Tetrahedron Lett.* **2011**, *52*, 4947–4949.
17. Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068–10073.
18. Johnson, E. C. B.; Kent, S. B. H. *J. Am. Chem. Soc.* **2006**, *128*, 6640–6646.
19. Kawakami, T.; Aimoto, S. *Tetrahedron* **2009**, *65*, 3871–3877.
20. (a) Villain, M.; Gaertner, H.; Botti, P. *Eur. J. Org. Chem.* **2003**, *2003*, 3267–3272. (b) Dang, B., Kubota, K., Mandal, K., Bezanilla, F., Kent, S. B. H. *J. Am. Chem. Soc.* **2013**, *135*, 11911−1191.
21. (a) Piszkiewicz, D.; Landon, M.; Smith, E. L. *Biochem. Biophys. Res. Commun.* **1970**, *40*, 1173–1178. (b) Fairchild, T. A.; Fulton, D.; Fontana, J. T.; Gratton, J.-P.; McCabe, T. J.; Sessa, W. C. *J. Biol. Chem.* **2001**, *276*, 26674–26679. (c) Catak, S.; Monard, G.; Aviyente, V.; Ruiz-López, M. F. *J. Phys. Chem. A* **2008**, *112*, 8752–8761. (d) Li, N.; Fort, F.; Kessler, K.; Wang, W. *J. Pharm. Biomed. Anal.* **2009**, *50*, 73–78.
22. Cowper, B.; Sze, T. M.; Premdjee, B.; White, A. B.; Hacking, A.; Macmillan, D. *Chem. Commun.* **2015**, *51*, 3208–3210.
23. (a) Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526–533. (b) Wan, Q.; Danishefsky, S. J. *Angew. Chem. Int. Ed.* **2007**, *46*, 9248–9252. (c) Rohde, H.;

Seitz, O. *Biopolymers* **2010**, *94*, 551–559. (d) Tsuda, S.; Mochizuki, M.; Nishio, H.; Yoshiya, T. *ChemBioChem* **2016**, *17*, 2133–2136.

24. Rohde, H.; Schmalisch, J.; Harpaz, Z.; Diezmann, F.; Seitz, O. *ChemBioChem* **2011**, *12*, 1396–1400.
25. (a) Van de Langemheen, H.; Brouwer, A. J.; Kemmink, J.; Kruijtzer, J. A.; Liskamp, R. M. *J. Org. Chem.* **2012**, *77*, 10058–10064. (b) Moyal, T.; Hemantha, H. P.; Siman, P.; Refua, M.; Brik, A. *Chem. Sci.* **2013**, *4*, 2496–2501. (c) Cergol, K. M.; Thompson, R. E.; Malins, L. R.; Turner, P.; Payne, R. *J. Org. Lett.* **2014**, *16*, 290–293. (d) Schmalisch, J.; Seitz, O. *Chem. Commun.* **2015,** *51*, 7554–7557. (e) Boll, E.; Drobecq, H.; Ollivier, N.; Blanpain, A.; Raibaut, L.; Desmet, R.; Vicogne, J.; Melnyk, O. *Nat. Protoc.* **2015**, 10, 269–292. (f) Tsuda, S.; Yoshiya, T.; Mochizuki, M.; Nishiuchi, Y. *Org. Lett.* **2015**, *17*, 1806–1809. (g) Tsuda, S.; Mochizuki, M.; Nishio, H.; Yoshiya, T.; Nishiuchi, Y. *Biopolymers* **2016**, *106*, 503–509. (h) Wang, Y.; Han, L.; Yuan, N.; Wang, H.; Li, H.; Liu, J.; Chen, H.; Zhang, Q.; Dong, S. *Chem. Sci.* **2018**, *9*, 1940–1946.
26. Bang, D.; Pentelute, B. L.; Gates, Z. P.; Kent, S. B. H. *Org. Lett.* **2006**, *8*, 1049–1052.
27. Thompson, R. E.; Liu, X.; Alonso-Garcia, N.; Pereira, P. J.; Jolliffe, K. A.; Payne, R. J. *J. Am. Chem. Soc.* **2014**, *136*, 8161–8164.
28. Huang, Y. C.; Chen, C. C.; Gao, S.; Wang, Y. H.; Xiao, H.; Wang, F.; Tian, C. L.; Li, Y. M. *Chem. Eur. J.* **2016,** *22*, 7623–7628.
29. While imidazole has been utilized as a catalyst for acylation in peptide chemistry as documented in various papers, this is the initial report that highlights the ability of imidazole to accelerate the NCL reaction involving alkylthioesters: (a) Bender, M. L.; Turnqest, B. W. *J. Am. Chem. Soc.* **1957**, *79*, 1652–1655. (b) Bruice, T. C.; Schmir, G. L. *J. Am. Chem. Soc.* **1957**, *79*, 1663–1667. (c) Bruice, T. C.; Schmir, G. L. *J. Am. Chem. Soc.* **1958**, *80*, 148–156. (d) Bruice, T. C.; Schmir, G. L. *J. Am. Chem. Soc.* **1958**, *80*, 2265–2267. (e) Kirsch, J. F.; Jencks, W. P. *J. Am. Chem. Soc.* **1964**, *86*, 833–837. (f) Yamada, H.; Kuroki, R.; Hirata, M.; Imoto, T. *Biochemistry* **1983**, *22*, 4551–4556. (g) Zhang, L.; Tam, J. P. *Tetrahedron Lett.* **1997**, *38*, 3–6. (h) Payne, R. J.; Ficht, S.; Greenberg, W. A.; Wong, C. H. *Angew. Chem. Int. Ed.* **2008**, *47*, 4411–4415. (i) Li, Y.; Yongye, A.; Giulianotti, M.; Martinez-Mayorga, K.; Yu, Y.; Houghten, R. A. *J. Comb. Chem.* **2009**, *11*, 1066–1072. (j) Li, Y.; Giulionatti, M.; Houghten, R. A. *Org. Lett.* **2010**, *12*, 2250–2253. (k) Fang, G. M.; Cui, H. K.; Zheng, J. S.; Liu, L. *ChemBioChem* **2010**, *11*, 1061–1065. (l) Suchy, M.; Elmehriki, A. A. H.; Hudson, R. H. E. *Org. Lett.* **2011**, *13*, 3952–3955. (m) Weissenborn, M.

J.; Castangia, R.; Wehner, J. W.; Sardzik, R.; Lindhorst, T. K.; Flitsch, S. L. *Chem. Commun.* **2012**, *48*, 4444–4446. (n) Byler, K. G.; Li, Y.; Houghtena, R. A.; Martinez-Mayorga, K. *Org. Biomol. Chem.* **2013**, *11*, 2979–2987.

30. Szajewski, R. P.; Whitesides, G. M. *J. Am. Chem. Soc.* **1980**, *102*, 2011–2026.
31. Catalan, J.; Elguero, *J. Adv. Heterocycl. Chem.* **1987**, *41*, 187–274.
32. (a) Mochizuki, M.; Taichi, M.; Hibino, H.; Takuwa, A.; Yoshida, T.; Ohkubo, T.; Nishiuchi, Y. *Tetrahedron Lett.* **2014**, *55*, 3073–3076. (b) Takuwa, A.; Yoshida, T.; Maruno, T.; Kawahara, K.; Mochizuki, M.; Nishiuchi, Y.; Kobayashi, Y.; Ohkubo, T. *FEBS Lett.* **2016**, *590*, 195–201.
33. (a) Koyano, F.; Okatsu, K.; Kosako, H.; Tamura, Y.; Go, E.; Kimura, M.; Kimura, Y.; Tsuchiya, H.; Yoshihara, H.; Hirokawa, T.; Endo, T.; Fon, E. A.; Trempe, J. F.; Saeki, Y.; Tanaka, K.; Matsuda, N. *Nature* **2014**, *510*, 162–166. (b) Han, C.; Pao, K. C.; Kazlauskaite, A.; Muqit, M. M. K.; Virdee, S. *ChemBioChem* **2015**, *16*, 1574–1579.
34. (a) Beyerman, H. C.; Maassen van Den Brink, W. *Recueil des Travaux Chimiques des Pays-Bas* **1961**, *80*, 1372–1375. (b) Jager, J.; Engberts, J. B. F. N. *J. Org. Chem.* **1985**, *50*, 1474–1478. (c) Haak, J. R.; Engberts, J. B. F. N.; Blandamer, M. J. *J. Am. Chem. Soc.* **1985**, *107*, 6031–6035. (d) Nusselder, J. J. H.; Engberts, J. B. F. N. *J. Org. Chem.* **1987**, *52*, 3159–3162.
35. (a) Sudoh, T.; Maekawa, K.; Kojima, M.; Minamino, N.; Kangawa, K.; Matsuo, M. *Biochem. Biophys. Res. Commun.* **1989**, *159*, 1427–1434. (b) Kambayashi, Y.; Nakao, K.; Mukoyama, M.; Saito, Y.; Ogawa, Y.; Shiono, S.; Inouye, K.; Yoshida, N.; Imura, H. *FEBS Lett.* **1990**, *259*, 341–345.
36. (a) Egorov, T. A.; Svenson, A.; Ryden, L.; Carlsson, J. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 3029–3033. (b) Ji, Y.; Majumder, S.; Millard, M.; Borra, R.; Bi, T.; Elnagar, A. Y.; Neamati, N.; Shekhtman, A.; Camarero, J. A. *J. Am. Chem. Soc.* **2013**, *135*, 11623–11633. (c) Guo, J.; Gaffrey, M. J.; Su, D.; Liu, T.; Camp II, D. G.; Smith, R. D.; Qian, W.-J. *Nat. Protoc.* **2014**, *9*, 64–75. (d) Aboye, T.; Kuang, Y.; Neamati, N.; Camarero, J. A. *ChemBioChem* **2015**, *16*, 827–833. (e) Campecino, J. O.; Dudycz, L. W.; Tumelty, D.; Berg, V.; Cabelli, D. E.; Maroney, M. J. *J. Am. Chem. Soc.* **2015**, *137*, 9044–9052.
37. Henkel, M.; Rockendorf, N.; Frey, A. *Bioconjugate Chem.* **2016**, *27*, 2260–2265.
38. (a) Tsuda, S.; Mochizuki, M.; Ishiba, H.; Yoshizawa-Kumagaye, K.; Nishio, H.; Oishi, S.; Yoshiya, T. *Angew. Chem. Int. Ed*. **2018**, *57*, 2105–2109. (b) Chisholm, T. S.; Clayton, D.; Dowman, L. J.; Sayers, J.; Payne, R. J. *J. Am. Chem. Soc.* **2018**, *140*, 9020–9024. (c) Sato, K.; Tanaka, S.; Yamamoto, K.; Tashiro, Y.; Narumi, T.;

Mase, N. *Chem. Commun.* **2018**, *54*, 9127–9130. (d) Gless, B. H.; Olsen, C. A. *J. Org. Chem.* **2018**, *83*, 10525–10534. (e) Yoshiya, T.; Tsuda, S.; Masuda, S. *ChemBioChem* **2019**, *20*, 1906–1913. (f) De Rosa, L.; Di Stasi, R.; D'Andrea, L. D. *Tetrahedron* **2019**, *75*, 894–905. (g) Mann, G.; Satish, G.; Meledin, R.; Vamisetti, G. B.; Brik, A. *Angew. Chem. Int. Ed.* **2019**, *58*, 13540–13549. (h) Shimamoto, S.; Mitsuoka, N.; Takahashi, S.; Kawakami, T.; Hidaka, Y. *The Protein Journal* **2020**, *39*, 711–716. (i) Vamisetti, G. B.; Satish, G.; Sulkshane, P.; Mann, G.; Glickman, M. H.; Brik, A. *J. Am. Chem. Soc.* **2020**, *142*, 19558–19569. (j) Hawala, I.; De Rosa, L.; Aime, S.; D'Andrea, L. D. *Chem. Commun.* **2020**, *56*, 3500–3503. (k) Wu, Y.; Li, C.; Fan, S.; Zhao, Y.; Wu, C. *Bioconjugate Chem.* **2021**, *32*, 2065–2072. (l) Naudin, E. A.; McEwen, A. G.; Tan, S. K.; Poussin-Courmontagne, P.; Schmitt, J.-L.; Birck, C.; DeGrado, W. F.; Torbeev, V. *J. Am. Chem. Soc.* **2021**, *143*, 3330–3339. (m) Mino, T.; Sakamoto, S.; Hamachi, I. *Biosci. Biotechnol. Biochem.* **2021**, *85*, 53–60. (n) Liao, P.; He, C. *Front. Chem.* **2021**, *9*. (o) De Rosa, L.; Hawala, I.; Di Stasi, R.; Stefania, R.; Capozza, M.; Nava, D.; D'Andrea, L. D. *J. Org. Chem.* **2023**, *88*, 4546–4553.

39. Schnoelzer, M., Alewood, P., Jones, A., Alewood, D., Kent, S. B. H. *Int. J. Pept. Protein. Res.* **1992**, *40*, 180–193.
40. Taichi, M., Kimura, T., Nishiuchi, Y. *Int. J. Pept. Res. Ther.* **2009**, *15*, 247–253.

Acknowledgement

I express my deepest gratitude and sincere, wholehearted appreciation to Prof. Akira Otaka (Tokushima University) for his kind guidance, constructive support, and heartfelt encouragement. In addition, I feel honored to have been given the opportunity of being the one to study peptide science from the beginning.

I wish to express my sincere and heartfelt gratitude to Dr. Shugo Tsuda, Dr. Masayoshi Mochizuki, and Dr. Taku Yoshiya (Peptide Institute, Inc.) for their kind support, constructive discussions and their careful perusing of my original manuscript.

I also wish to express my gratitude to Prof. Akira Shigenaga, and Dr. Kohei Sato for their generous encouragement and constructive discussion.

I am grateful to Prof. Tsubasa Inokuma, Dr. Kohei Tsuji and all other colleagues (Tokushima University) for further applications of SEAlide unit.

Finally, I am most grateful to my wife for various constant supports. I am also grateful to my parents for their constant encouragement.

List of publications

This study was published in the following papers.

1. Synthetic Procedure for *N*-Fmoc Amino Acyl-*N*-Sulfanylethylaniline Linker as Crypto-Peptide Thioester Precursor with Application to Native Chemical Ligation
   Ken Sakamoto, Kohei Sato, Akira Shigenaga, Kohei Tsuji, Shugo Tsuda, Hajime Hibino, Yuji Nishiuchi, and Akira Otaka
   J. Org. Chem. **2012**, *77*, 6948−6958.

2. Imidazole-Aided Native Chemical Ligation: Imidazole as a One-Pot Desulfurization-Amenable Non-Thiol-Type Alternative to 4-Mercaptophenylacetic Acid
   Ken Sakamoto, Shugo Tsuda, Masayoshi Mochizuki, Yukie Nohara, Hideki Nishio, and Taku Yoshiya
   Chem. Eur. J. **2016**, *22*, 17940–17944.

3. 1,2,4-Triazole-aided native chemical ligation between peptide-*N*-acyl-*N'*-methyl-benzimidazolinone and cysteinyl peptide
   Ken Sakamoto, Shugo Tsuda, Hideki Nishio and Taku Yoshiya
   Chem. Commun. **2017**, *53*, 12236–12239.