Chemoselective Modification Strategies for Peptides and Proteins in Aqueous Media

Dissertation zur Erlangung des akademischen Grades des

Doktors der Naturwissenschaften (Dr. rer. Nat.)

Eingereicht im Fachbereich Biologie, Chemie, Pharmazie

der Freien Universität Berlin

Vorgelegt von

Michaela Mühlberg

aus Berlin

2014

Die Arbeit wurde zwischen dem 01. Juli 2009 und dem 31. März 2014 unter der Leitung von Prof. Dr. Christian P. R. Hackenberger am Institut für Chemie und Biochemie der Freien Universität Berlin sowie am Leibniz-Institut für Molekulare Pharmakologie angefertigt.

1. Gutachter:	Prof. Dr. Christian P. R. Hackenberger
2. Gutachter:	Prof. Dr. Beate Koksch
Disputation am:	10. Juli 2014

Declaration

I herewith confirm that I have prepared this dissertation without the help of any impermissible resources. All citations are marked as such. The present thesis has neither been accepted in any previous doctorate degree procedure nor has it been evaluated as insufficient.

Berlin, 31th March 2014

Michaela Mühlberg

The work on this dissertation resulted so far in the following publications:

1. M. Mühlberg, D. M. M. Jaradat, R. Kleineweischede, I. Papp, D. Dechtrirat, S. Muth, M. Broncel, C. P. R. Hackenberger, Bioorganic & Medicinal Chemistry **2010**, *18*, 3679-3686.
   Acidic and basic deprotection strategies of borane-protected phosphinothioesters for the traceless Staudinger ligation
2. S. Sowa, M. Mühlberg, K. M. Pietrusiewicz, C. P. R. Hackenberger, Bioorganic & Medicinal Chemistry **2013**, *21*, 3465-3472.
   Traceless Staudinger Acetylation of Azides in Aqueous Buffers
3. M. Mühlberg, K. D. Siebertz, B. Schlegel, P. Schmieder, C. P. R. Hackenberger, Chemical Communications **2014**, DOI: 10.1039/c4cc00774c.
   Controlled Thioamide vs. Amide Formation in the Thioacid–Azide Reaction under Acidic Aqueous Conditions
4. M. Mühlberg, M. G. Hoesl, C. Kuehne, J. Dernedde, N. Budisa, C. P. R. Hackenberger, **2014**, submitted.
   Orthogonal Dual-Modification of Proteins by Oxime Ligation and Cu-Catalyzed Azide–Alkyne Cycloaddition

Acknowledgements

First, I would like to thank Prof. Dr. Christian Hackenberger for offering me the opportunity to work in his research group on such diverse and interesting projects and for the freedom he gave me to pursue different ideas. I would like to especially thank him also for the support that I got throughout my whole career.

I am grateful to Prof. Dr. Beate Koksch for being my second supervisor throughout my Ph.D., for helpful discussions and for being the second referee of my thesis.

Special thanks go to Ina Wilkening, Robert Vallée, Paul Majkut, Verena Böhrsch and Lukas Artner for fruitful discussions and their support in the lab, and many pleasant hours outside of the lab throughout all those years.

I also would like to thank Maria Glanz, Olaia Nieto-Garcia, Dominik Schumacher and Kristina Siebertz for all their support during the last year of my Ph.D. and their patience with me during my last months.

All current and former group members are thanked for the pleasant and friendly atmosphere, especially my lab mates Da'san Jaradat, Jordi Bertran, Simon Reiske and Divya Agrawal.

A thank you goes also to Nediljko Budisa, Michael Hösl and Nina Bohlke for their constant supply of proteins, fruitful discussions and the very pleasant collaboration on an interesting and challenging project.

I would also like to thank Prof. Pietrusiewicz and Sylwia Sowa for the productive collaboration and discussions.

I would like to thank Chris Weise for his time and support during my first MALDI-MS measurements and all the scientific troubles that came with it.

I would like to acknowledge Andreas Springer and the whole mass spectrometry team from the FU Berlin as well as Eberhard Krause and the whole mass spectrometry team from the FMP in Buch for their help with all my MS measurements over time and for helpful discussion on more complex problems.

I also thank Andreas Schäfer and the team of the NMR service at the FU Berlin as well as Peter Schmieder and the team of the NMR department at the FMP in Buch for their help and support with my NMR measurements with special thanks to Brigitte Schlegel.

Many thanks also go to Katrin Wittig, Marianne Dreißigacker and Katharina Tebel for their help and support with all administrative problems, especially Katta for many helpful and motivating conversations throughout my Ph.D.

For financial support, I would like to thank the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 765), the Fonds der chemischen Industrie, the Studienstiftung des deutschen Volkes, the Freie Universität Berlin, the Leibniz-Institute for Molecular Pharmacology and the Humboldt Universität Berlin.

Many thanks go to my friends, especially Silke Zobel, for all their support and patience with me during the last years.

Last but not least, I would like to thank Mark Archibald and my family, especially my mother, for their belief in me during all those years and their motivation and support, which helped me to get to the end.

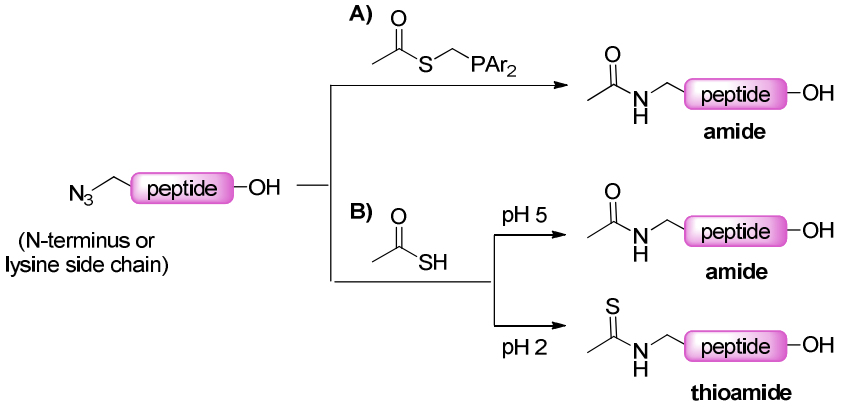
for my mother

Abstract

In nature, a protein's structure and function depends amongst other things on very precise posttranslational modifications such as phosphorylations, glycosylations and many others. Our ability to naturally and unnaturally modify a peptide or a protein in a chemoselective and thereby very distinct fashion enables us to examine biological phenomena in a very sophisticated approach.

During this thesis, we have focused on a variety of different ways of decorating peptides and proteins in a chemoselective manner. This work involves four different approaches that deal with different chemoselective labelling strategies to introduce either natural or unnatural modifications to peptides or proteins.

In the first project, the traceless Staudinger ligation and its application for selective acetylation of azido lysine peptides in aqueous media was probed (Scheme 1A). Different alkyl azides were reacted with two different phosphines: 1) commercially available hydrophobic (diphenylphosphino-)methanethiol acetate and 2) water-soluble bis(*p*-dimethylaminoethyl)-phosphinomethanethiol acetate. Initial studies showed no chemoselectivity of the reaction in the presence of basic amino acid side chains such as lysine. Consequently, the traceless Staudinger ligation was probed with an azido lysine without free amines present in the peptide. Thereby, only the water-soluble variant showed a good reactivity in aqueous buffered systems. The highest conversion of 61% was achieved with the water-soluble phosphine (10 eq.) in a phosphate buffer (0.4 M, pH 8)/DMF mixture (2:8).



Scheme 1: Selective acetylation of alkyl azido peptides by A) traceless Staudinger ligation or B) thioacid–azide reaction.

In the second project, the thioacid–azide reaction was probed for selective acetylation of alkyl azido peptides (Scheme 1B). As alkyl azides are more electron-rich and therefore less reactive than the commonly used sulfonyl azides, a side reaction of the thioacid with basic side chains

such as lysines was observed. Therefore, the reaction was probed at different pH values with an electron-rich azido butanoyl and a modestly electron-poor azido glycine peptide. The results showed a maximum conversion towards the desired acetylated peptides at pH 5 and conversion rates could be easily increased to > 99% at higher concentrations during the reaction. At lower pH, an unexpected and different main product was observed, which was not the desired amide but a thioamide (Scheme 1B). Thioamide formation could be increased at pH 2 to more than 90% for the modestly electron-poor azido peptide. It was shown that the ratio of thioamide and amide depends on the pH of the reaction mixture as well as on the nature of the azide, which allowed selective control over the desired product.

Within the third project, a new multiple-labelling approach for proteins was probed. Selective dual-modification of a thermophilic lipase (TTL) was achieved successfully by combination of two orthogonal functionalisation strategies: oxime ligation and CuAAC (Figure 1). To do so, the TTL was bearing several azidohomoalanine residues, which were incorporated by selective pressure induction as methionine analogues, and an *N*-terminal serine, which could be post-translationally reacted with periodate to yield an aldehyde as the second bioorthogonal tag. This strategy allowed modification of the protein with two different functional moieties: galactose for carbohydrate–protein binding studies and biotin for immobilisation of the protein. With these modifications, the protein could be applied in surface plasmon resonance measurements to study their different binding to a lectin and to determine lectin binding constants. Thereby, the biotin functionalisation allowed immobilisation of our constructs on a streptavidin-coated chip, which required significantly less amounts of our proteins for the measurements.

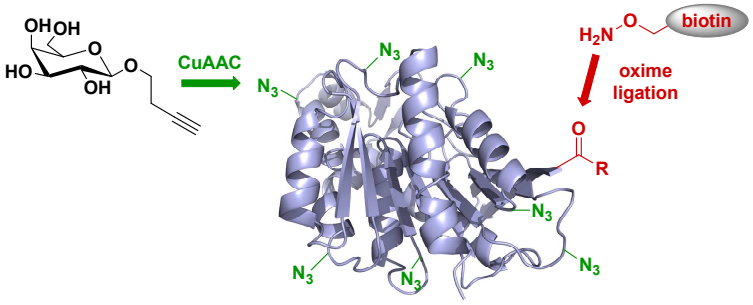


Figure 1: Dual-modification approach on TTL by oxime ligation and CuAAC.

In addition, dual-functionalisation a GFP variant was performed successfully by selective combination of thiazolidine formation on the chemically created *N*-terminal aldehyde and CuAAC on previously incorporated azidohomoalanine residues. It was shown that thiazolidine formation is more efficient than oxime ligation and yields a very stable product.

In the fourth project, a new way for selective post-translational aldehyde formation on a distinct position anywhere throughout a protein was planned. To do so, a couple of new pyrrolysine derivatives were designed, which are comprised of a lysine and an amino glycol unit attached to its side chain to allow subsequent periodate cleavage and selective formation of an aldehyde. These unnatural amino acid building blocks were synthesised successfully in good yields and probed by a cooperation partner for their incorporation during protein expression applying the pyrrolysine amber suppression technique. So far, initial attempts for the incorporation were not successful.

Kurzzusammenfassung

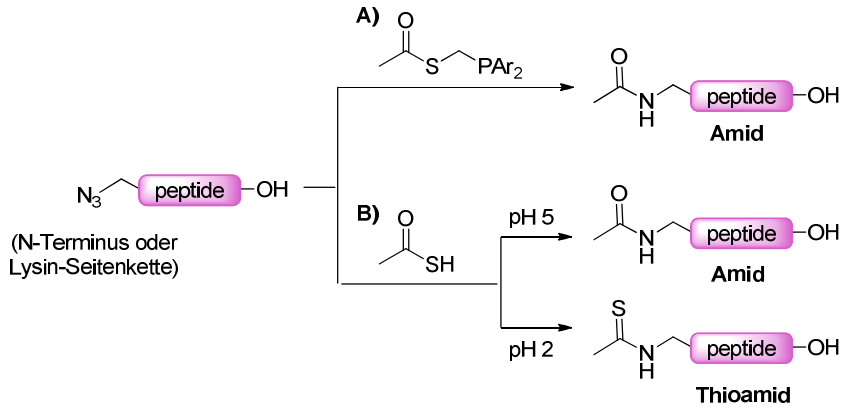
Die Struktur und Funktion eines Proteins hängen in der Natur unter anderem von spezifischen posttranslationalen Modifikationen ab wie z.B. Phosphorylierungen, Glykosylierungen und vielen weiteren. Unsere Fähigkeit, Peptide und Proteine auf natürliche und unnatürliche Weise chemoselektiv und damit sehr präzise zu modifizieren, erlaubt es uns, biologische Phänomene systematisch und differenziert zu untersuchen.

In dieser Doktorarbeit haben wir uns mit verschiedenen Möglichkeiten der chemoselektiven Funktionalisierung von Peptiden und Proteinen beschäftigt. Dabei wurden vier verschieden wissenschaftliche Ansätze ausgewählt, die auf unterschiedlichen chemoselektiven Funktionalisierungsstrategien zur natürlichen beziehungsweise unnatürlichen Modifikation von Peptiden und Proteinen beruhen.

Im ersten Projekt wurde die spurlose Staudinger Ligation bezüglich ihrer Anwedung für die selektive Acetylierung von Azidolysinpeptiden im wässrigen Medium untersucht (Schema 1A). Dabei wurden verschiedene Alkylazide mit zwei unterschiedlichen Phosphinen umgesetzt: 1) kommerziell erhältliches hydrophobes (Diphenylphosphin)methanthiolacetat und 2) wasserlösliches Bis(*p*-dimethylaminoethyl)phosphinmethanthiolacetat. Erste Studien zeigten, dass die Reaktion in Anwesenheit von basischen Aminosäureseitenketten wie Lysin nicht chemoselektiv abläuft. Aus diesem Grund wurde die spurlose Staudinger Ligation anschließend an einem Azidolysinpeptid ohne freie Amine getestet. Dabei zeigte nur das wasserlösliche Phosphin eine gute Reaktivität bezüglich des Azidolysinpeptids in wässrigen Puffersystemen. Der höchste Umsatz von 61% konnte mit dem wasserlöslichen Phosphin (10 eq.) in einem Phosphatpuffer (0.4 M, pH 8)/DMF-System erzielt werden.

Im zweiten Projekt wurde die Thiosäure–Azid Reaktion bezüglich ihrer Anwendbarkeit auf die selektive Acetylierung von Alkylazidpeptiden untersucht (Schema 1B). Da Alkylazide generell elektronenreicher und damit weniger reaktiv sind als die herkömmlich verwendeten Sulfonylazide, konnten unter pH-neutralen Reaktionsbedingungen eine Nebenreaktion mit basischen Aminosäureseitenketten wie Lysin beobachtet werden. Aus diesem Grund wurde die Reaktion bei unterschiedlichem pH-Wert mit zwei verschiedenen Aziden getestet: 1) einem elektronenreichen Azidobutanoylpeptid und 2) einem geringfügig elektronenärmeren Azidoglycinpeptid. Die Ergebnisse zeigten einen maximalen Umsatz zum gewünschten acetylierten Peptid bei pH 5, und die Umsatzraten konnten leicht durch Erhöhen der Konzentration der Reaktionslösung auf bis zu >99% angehoben werden. Bei niedrigeren pH-

Werten wurde ein unerwartetes neues Hauptproduckt beobachtet, bei welchem es sich nicht um das erwünschte Acetamid sondern um ein Thioacetamid handelte (Schema 1B). Die Bildung dieses Thioamids konnte bei der Reaktion mit dem elektronenärmeren Azidoglycinpeptid bei pH 2 auf mehr als 90% erhöht werden. Zusammenfassend haben unsere Ergebnisse gezeigt, dass das Verhältnis zwischen Thioamid und Amid sowohl vom pH-Wert der Reaktion als auch von den elektronischen Eigenschaften des Azides abhängt, was eine Kontrolle über das gewünschte Endprodukt erlaubt.



Schema 1: Selektive Acetylierung von Alkylazidpeptiden mittels A) spurloser Staudinger Ligation oder B) Thiosäure–Azid Reaktion.

Im dritten Projekt wurde ein neuer Multi-Funktionalisierungsansatz für Proteine untersucht. Durch die Kombination zweier orthogonaler Funktionalisierungsstrategien, der Oxim-Ligation und der CuAAC, konnte eine thermophile Lipase (TTL) erfolgreich zweifach modifiziert werden (Abbildung 1). Um das zu erreichen, besaß die TTL mehrere Azidohomoalanin-Reste, die mittels selektiver Druckinduktion als Methionin-Analoga eingebaut wurden, sowie ein *N*-terminales Serin, welches posttranslational mittels Periodat in ein Aldehyd als zweite bioorthogonale Gruppe überführt werden konnte. Diese Strategie ermöglichte die Funktionalisierung des Proteins mit zwei unterschiedlichen Molekülen: Galactose zur Erforschung von Kohlenhydrat-Protein Interaktionen und ein Biotin für eine zusätzliche Immobilisierung des Proteins. Diese beiden Modifikationen ermöglichten es schließlich, das Protein mittels Oberflächenplasmonenresonanzspektroskopie zu untersuchen und somit eventuelle Bindungen unserer Proteinkonjugate mit Lektinen sowie spezifischen Bindungskonstanten zu bestimmen. Dabei ermöglichte das eingeführte Biotin die Immobilisierung unserer Konstrukte auf Streptavidin-Chips, wodurch für die darauffolgenden Messungen deutlich weniger Material unserer Proteinkonjugate benötigt wurde.

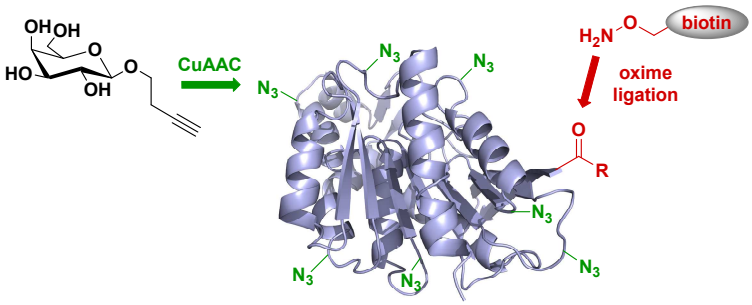


Abbildung 1: Dualer Modifikationsansatz an TTL mittels Oxim-Ligation und CuAAC.

Desweiteren konnte in einem zweiten Vorhaben erfolgreich ein eine modifizierte Variante von GFP dual-funktionalisiert werden. Dabei wurde die zuvor verwendete CuAAC an den Azidohomoalanin-Resten mit einer Thiazolidinbildung am synthetisch erzeugten Aldehyd kombiniert. Es konnte gezeigt werden, dass die Thiazoldinbildung im Vergleich zur Oximbildung effizienter ist und ein sehr stabiles Produkt liefert.

Im vierten Projekt wurde die selektive posttranslationale Einführung eines Aldehyds an einer beliebigen Stelle eines Proteins versucht. Dafür wurden einige neue Pyrrolysin-Derivate entworfen, welche auf einem Lysin basieren in dessen Seitenkette eine Aminoglycol-Einheit eingebaut wurde. Diese Einheit sollte im Anschluss an die Proteinexpression die selektive Bildung eines Aldehyds mittels Glykolspaltung ermöglichen. Die unnatürlichen Aminosäurebausteine konnten erfolgreich und in guten Ausbeuten synthetisiert werden. Anschließend wurde von einem Kooperationspartner versucht die Pyrrolysin-Derivate mittels der Pyrrolysin-Amber-Suppressionsmethode während der Proteintranslation in das gewuenscht Protein einzubauen. Erste Versuche waren bis jetzt jedoch erfolglos.

Abbreviations

aa	amino acid
aaRS	aminoacyl-tRNA synthetase
Ac	acetyl
Aha	azidohomoalanine
AIBN	2,2'-azobis(2-methyl-propionitrile)
Ald	aldehyde
All	allyl
Alloc	allyloxycarbonyl
Ar	aryl
ASGPR	asialoglycoprotein receptor
AU	absorbance units
Bn	benzyl
Boc	*tert*-butyloxycarbonyl
Bpa	*p*-benzoyl L-phenyl-alanine
Bu	butyl
C peptide	peptide at *C*-terminus
calcd.	calculated
CAT	chloramphenicol acetyl transferase
CBD	chitin binding domain
CCA	α-cyanocinnaminic acid
CuAAC	copper-catalyzed azide–alkyne cycloaddition
Cychex	cyclohexane
d	day
Da	Dalton
DABCO	1,4-diazabicyclo[2.2.2]octane
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCC	*N,N'*-dicyclohexylcarbodiimide
DFT	density functional theory
DHB	2,5-dihydroxybenzoic acid
DIAD	diisopropyl azodicarboxylate
DIC	*N,N'*-diisopropylcarbodiimide
DIPEA	*N,N*-diisopropylethylamine
Dmab	4-(*N*-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]amino)benzyl
DMF	*N,N*-dimethylformamide
DMS	dimethyl sulfide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
ECL	*Erythrina cristagalli* lectin
E. coli	*Escherichia coli*
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	ethylenediaminetetraacetic acid
em.	emission
eq.	equivalent
EPL	Expressed Protein Ligation
ESI	electrospray ionization
Et	ethyl
ex.	excitation
exper.	experimental
Fmoc	fluorenylmethyloxycarbonyl

FRET	fluorescence resonance energy transfer
GFP	green fluorescent protein
Gnd	guanidine
h	hour
HATU	1-[bis(dimethylamino)methylene)-1H-1,2,3-triazolo[4,5-b]pyridinium-3-oxide hexafluorophosphate
HBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HdeA	histone-dependent expression A
hex	hexane
HOBt	1-hydroxybenzotriazole
Hpg	homopropargylglycine
HPLC	high-performance liquid chromatography
HRMS	high resolution mass spectrometry
ivDde	1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl
KAHA	ketoacid hydroxylamine
LC-MS	liquid chromatography mass spectrometry
LDA	lithium diisopropyl amide
LRMS	low resolution mass spectrometry
MALDI	matrix-assisted laser desorption/ionisation
Me	methyl
MeCN	acetonitrile
min	minute
MMPP	magnesium monoperoxyphthalate
MPA	3-mercaptopropionic acid
MPAA	4-mercaptophenylacetic acid
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MS/MS	tandem mass spectrometry
n. d.	not determined
N peptide	*N*-terminal peptide
NCL	Native Chemical Ligation
Nle	norleucine
NMM	*N*-methylmorpholine
NMP	N-methyl pyrrolidone
NMR	nuclear magnetic resonance
OEG	oligoethylene glycol
o.n.	overnight
OTf	triflate
PBS	phosphate buffered saline
PEG	polyethylene glycol
PG	protecting group
Ph	phenyl
pKa	acid dissociation constant
ppm	parts per million
Pr	propyl
Pybop	benzotriazol-1-yloxy tripyrrolidinophosphonium hexafluorophosphate
Pyl	pyrrolysine

R	residue
Rhod.	Rhodamine B
RNA	ribonucleic acid
rt	room temperature
SA	sinapinic acid
SDS	sodium dodecyl sulfate
SEA	bis(2-sulfanylethyl)amido
Sec	selenocysteine
SeMet	selenomethionine
SPAAC	strain-promoted azide–alkyne cycloaddition
SPANC	strain-promoted alkyne–nitrone cycloaddition
SPI	selective pressure incorporation
SPPS	solid phase peptide synthesis
SPR	surface plasmon resonance
TCEP	tris(2-carboxyethyl)phosphine
TCFH	*N,N,N',N'*-tetramethylchloroformamidinium hexafluorophosphate
Tf	triflyl
TFA	trifluoroacetic acid
theor.	theoretical
THF	tetrahydrofuran
Thia	thiazolidine
THPTA	tris(3-hydroxypropyltriazolylmethyl)amine
TIC	total ion count
TIS	triisopropylsilane
TLC	thin layer chromatography
Tmob	2,4,6-trimethoxybenzyl
TMS	trimethylsilyl
ToF	time of flight
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
tRNA	transfer ribonucleic acid
Trt	trityl
Ts	tosyl
TSL	traceless Staudinger ligation
TTL	*Thermoanaerobacter thermohydrosulfuricus*
UV	ultraviolet
w/	with
w/o	without

Specific 3- and 1-Letter-Codes for amino acids were used upon the recommendation of the IUPAC-IUB commission for biochemical nomenclature.[1]

1. Introduction

Peptides and proteins are mostly known to people as an important part of nutrition. We take them up, digest them and use their building blocks – namely the essential amino acids that we cannot produce on our own – to form new proteins that take over different roles in our biological system, such as enzymes, hormones and specific receptors.[2] Decades of research has established that these small molecular machineries control the stability and function of the cell and once damage is done proteins take over repair functions or decide on cell degradation by several defined modes of interaction inside and outside of the cytoplasmic matrix.[2] *But how does our body know which proteins to synthesise for which purpose? And where do proteins get information on their specific structure and function?*

Peptides and proteins are primarily constituted of different amino acids that are partially synthesised in organisms or taken up by nutrition.[2] Their primary structure – also known as their specific amino acid sequence – is saved on the genome DNA located in the cellular nucleus. A cell is able to transcribe its DNA into the closely related RNA to transfer it out of the nucleus into the cytosol where it can then be read by the so-called ribosomes. Ribosomes are little machineries which can translate the RNA's information by reading it like a claviature putting amino acids together one after another to constitute the previously encoded protein.[2] However, the primary structure itself is not enough to give a protein its designated role of action. The three dimensional folded state of a protein is strongly related to its true and final role. Thereby, the secondary and tertiary structure of a protein is not solely dictated by the 23 proteinogenic amino acids but also by non-covalent interactions both intramolecular as well as with small molecules, such as co-factors, and by the introduction of additional covalent modifications on the protein's termini or side chains.[3] These so-called post-translational modifications, e.g. phosphorylation, acetylation, methylation and sulfation make the protein and its function unique and specific.[4] Selective control of these modifications can thereby activate or deactivate a protein's mode of action and trigger a whole different cascade of mechanism.

The field of chemical biology includes the synthesis of biological macromolecules, such as peptides and proteins, and their modification by elongating them, changing their primary structure,[5] incorporating non-natural amino acids[6] including those with altered chirality[7] and non-amide backbone bond,[8] and introducing non-proteinaceous moieties, such as sugars[9] or fatty acid chains[10] to name a few. Notably, the manipulation of natural or artificial proteins might provide scientists an insight into the protein's contribution in certain biological events by

influencing features, such as, fluorescent activity, solubility, immunostability and selective pharmacokinetic effects.[11]

To study a protein and its selective modifications and to get a better understanding of its detailed mode of action, (bio)chemists had to find a way to selectively synthesise these various proteins. There are two main approaches to achieve this goal. On one hand, scientists apply chemical synthetic strategies to attain the protein of choice.[12] During the last decades, scientists have worked on a variety of different improvements in this field, which will be discussed in more detail in chapter 1.1 – 1.3.

On the other hand, several techniques have been developed that allow the introduction of various modifications during biological protein expression. In a second reaction step, these modifications might then be further functionalised applying selective reaction strategies.[13] Chapter 1.4 and 1.5 will deal with this topic.

1.1. Amide Bond Formation by Capture–Rearrangement Strategies

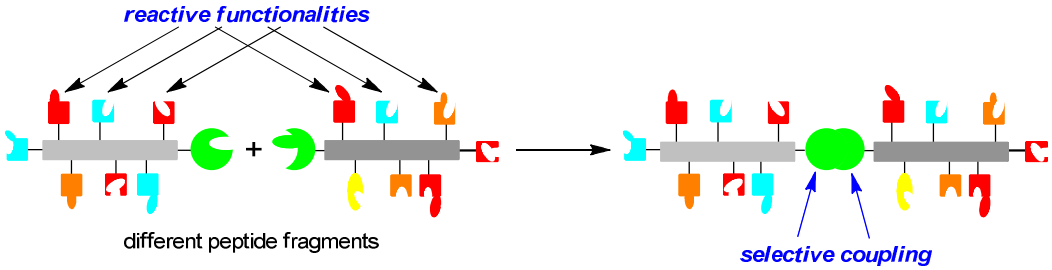
1.1.1. First Steps Towards Peptide and Protein Synthesis

The first attempt towards peptide synthesis reaches as far back as to the start of the 20th century. In 1901, Emil Fisher was the first to publish the successful synthesis of a dipeptide, glycylglycine, by hydrolysis of the diketopiperazine of glycine.[14] He was also the first to introduce the name *peptide* for short proteins – derived from the greek word *pepsis* meaning digestion. But back then the synthesis of peptides was very challenging and suitable methods for protection as well as amide bond formation were not established. Therefore, it took the scientific community almost another fifty years to develop appropriate methods to fully elucidate a peptide's structure and achieve its complete synthesis – namely cyclic polypeptide hormone oxytocin.[15] Published in 1953 by Vincent du Vigneaud, his ground-breaking work earned him a Nobel Prize only two years later in 1955. During the next decades, the introduction of proper ways to analyse peptides and proteins led to the discovery of a multitude of naturally occurring peptides. In addition, the development of new synthetic strategies enabled scientists to make and analyse a vast array of shorter peptides.[16]

In 1963, the introduction of Solid Phase Peptide Synthesis (SPPS) by Nobel Prize winner B. Merrifield gave an immense boost to the field of peptide and protein chemistry.[17] Scientists were now able to more easily synthesise longer peptides in moderate yields. Undesired byproducts and reagents could be easily removed by filtration after every reaction avoiding

cumbersome purification steps and the repetitive nature of peptide synthesis now allowed automation of the whole process, and only a few years later, its discovery led to the synthesis of a 124 amino acids long Ribonuclease A enzyme.[18] Although nowadays this method is well developed and routinely used, it is limited by certain factors, such as the length of the peptide and its amino acid composition. Especially, hydrophobic peptides tend to aggregate, which further limits the yield.[19] As a result, chemical biologists started to look out for new methods to synthesise longer and more demanding peptides and proteins. First approaches probed the ligation of protected peptide fragments,[19] but their poor solubility and steric hindrance during ligation gave rise to the development of protecting group free peptide ligation strategies. *But how to selectively form an amide bond between two unprotected peptide fragments with such a multitude of functional groups around?*

Based on the principle of chemoselectivity, scientists have come up with different methods to accomplish this goal.[20] The term chemoselectivity refers to a group of functional moieties that react solely with each other in the presence of a large variety of other functional units (Scheme 1.1). The difficulty is that of choosing appropriate functional groups on the subject of their reactivity towards others. Many known reactions take place between a nucleophile and an electrophile. To find groups that are inert towards this kind of reactivity – and therefore inert towards the majority of functionalities or in other words chemoselective – is crucial for the successful synthesis of long and unprotected peptides and proteins.[20] In addition, there must be a way to incorporate these chemoselective moieties into single peptide and protein fragments during synthesis.



Scheme 1.1: General idea of chemoselective ligation strategy.

1.1.2. Native Chemical Ligation (NCL)

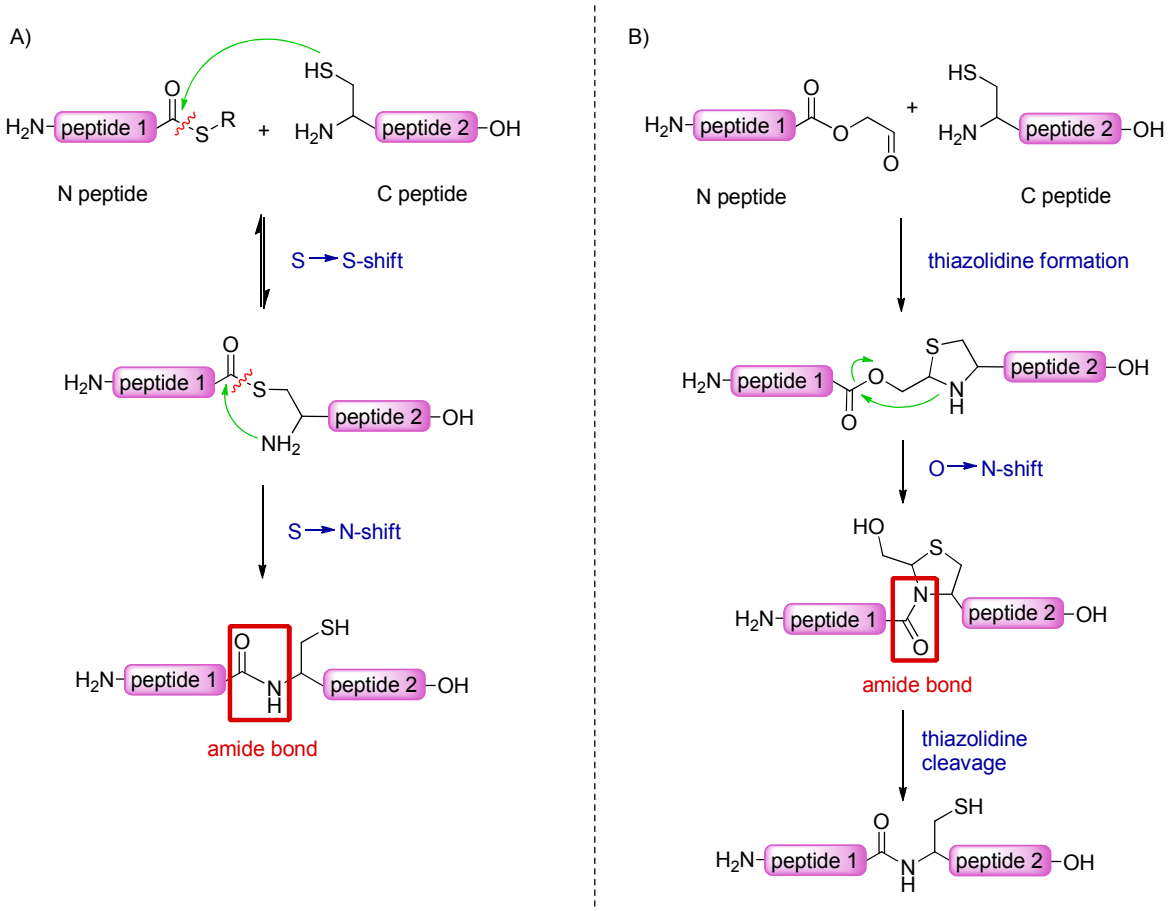
1.1.2.1. Initial Development and Concept of Native Chemical Ligation

In the early 1950s, Wieland and co-workers published the discovery of a very distinct and selective reaction behaviour of free cysteine.[21] The reaction of a valine thioester with cysteine

gave the dipeptide valine-cysteine as the sole product. They could show that the reaction proceeded over a transthioesterification with the thiol group of cysteine then leading *via* an intramolecular S→N-shift to the formation of the more stable valine-cysteine amide bond. These findings indicated a preferential reaction of N-terminally unprotected cysteines with C-terminal peptide thioesters.

It was only in 1994 that Dawson, Kent and co-workers published the first chemoselective application of the latter reaction towards the synthesis of human interleukin 8 (Scheme 1.2A).[22] They could show that this method could be easily applied to join unprotected peptide segments and achieve a moderate size without application of protecting groups. The convergent strategy requires an *N*-terminal Cys-residue of the C peptide – the one ending up in the *C*-terminal region of the ligation product – and a thioester moiety on the *C*-terminus of the N peptide – the one ending up in the *N*-terminal region of the newly formed peptide – to yield an amide bond (Scheme 1.2A). The mechanism follows a capture/rearrangement concept *via* a thiol-thioester exchange as described before by Wieland *et al*. with a final rapid intramolecular S→N-acyl shift that leads to the formation of a stable amide bond between the two peptide segments.[21]

This so-called Native Chemical Ligation (NCL) has many features. The reaction can be carried out in aqueous media under physiological conditions and is a powerful tool to join smaller peptides to generate small or medium-sized proteins under physiological conditions.[23] However, strong basic or acidic conditions could hamper the reaction. Under basic conditions the thioester moiety is prone to cleavage or attack by basic amino acid side chains such as lysine or arginine. In contrary, acidic conditions change the equilibrium of the reaction and could turn the irreversible amide-forming step into a reversible one. Hence, the reaction would be hindered and/or accompanied by the formation of several side products.[24] Kent and co-workers therefore explored the kinetics of NCL in solution and optimised the reaction conditions in aqueous buffer systems.[25] In addition, the usage of chaotropic reagents, such as urea or guanidinium hydrochloride, does not interfere with the chemoselective reaction, thus enabling the solubilisation and coupling of more hydrophobic peptides that would not be soluble under standard reaction conditions. This effect leads thus to an enhancement in peptide concentration and improved rates of the reaction.[22]



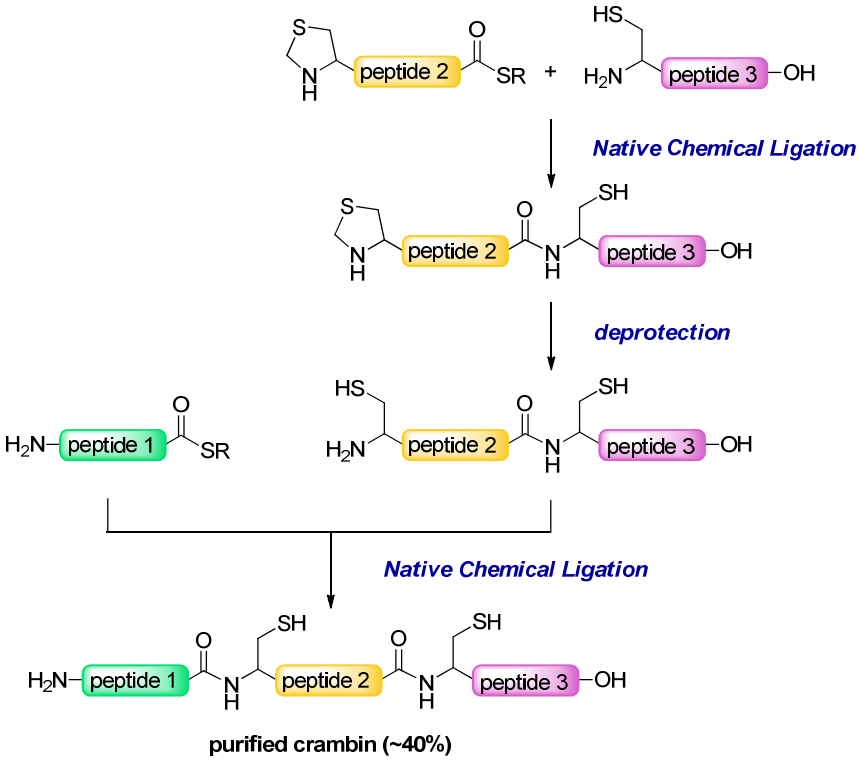
Scheme 1.2: General mechanism of A) Native Chemical Ligation and B) ligation by Tam *et al.*

In the same year, Tam *et al.* published a similar method (Scheme 1.2B).[26] The reaction is based on a more stable aldehyde ester N peptide that can react with a C peptide cysteine forming a thiazolidine. Bringing both peptides in close proximity, the amide bond between C and N peptide is formed subsequently by undergoing an O→N-shift yielding the desired peptide with a thiazolidine group in the backbone that would have to be cleaved afterwards (Scheme 1.2B). Although being based on a similar principle and applying a more stable ester group, this method did not gain as much attention as the Native Chemical Ligation.

1.1.2.2. Three-Segment Peptide Ligation Strategy by NCL (C-to-N)

In 2004, Kent and co-workers first published the total chemical synthesis of a protein, namely [V15A]crambin, in a three-segment strategy using NCL of unprotected peptides (Scheme 1.3).[27] The ligation of more than two fragments requires not just the synthesis of the known starting materials but also specific protection of at least one cysteine residue. The central segment does not only contain a thioester moiety to enable coupling to the C peptide but also a C-terminal cysteine for ligation towards the N peptide. The NCL reaction with the C peptide would lead to

partial fragment cyclisation in the absence of a cysteine protecting group. Kent *et al.* used an acetal-protecting strategy to avoid unwanted cyclisation. After the first ligation, the addition of methoxyamine hydrogen chloride leads to the unprotected ligation product. Subsequent readjustment of the pH to ~7 and addition of the second thioester peptide (in approximately 10% molar excess) resulted in quantitative ligation within 20 h. In 2008, Kent published another C-to-N sequential ligation regarding the Human insulin-like growth factor 1 protein.[28] Although this method still requires intermediary protection/deprotection steps of the non-reacting *N*-terminal cysteine residues, these recent results give an outlook on the versatile application of Native Chemical Ligation still undiscovered. This powerful method can be used to synthesise and manipulate many different (bio)molecules due to mild, aqueous and physiological reaction conditions and, nowadays, it is the most commonly used chemoselective ligation strategy for unprotected peptide fragments of up to ~150 amino acids.



Scheme 1.3: One-pot three-segment ligation strategy for protein synthesis by NCL.

1.1.2.3. Advancements in Peptide Thioester Synthesis

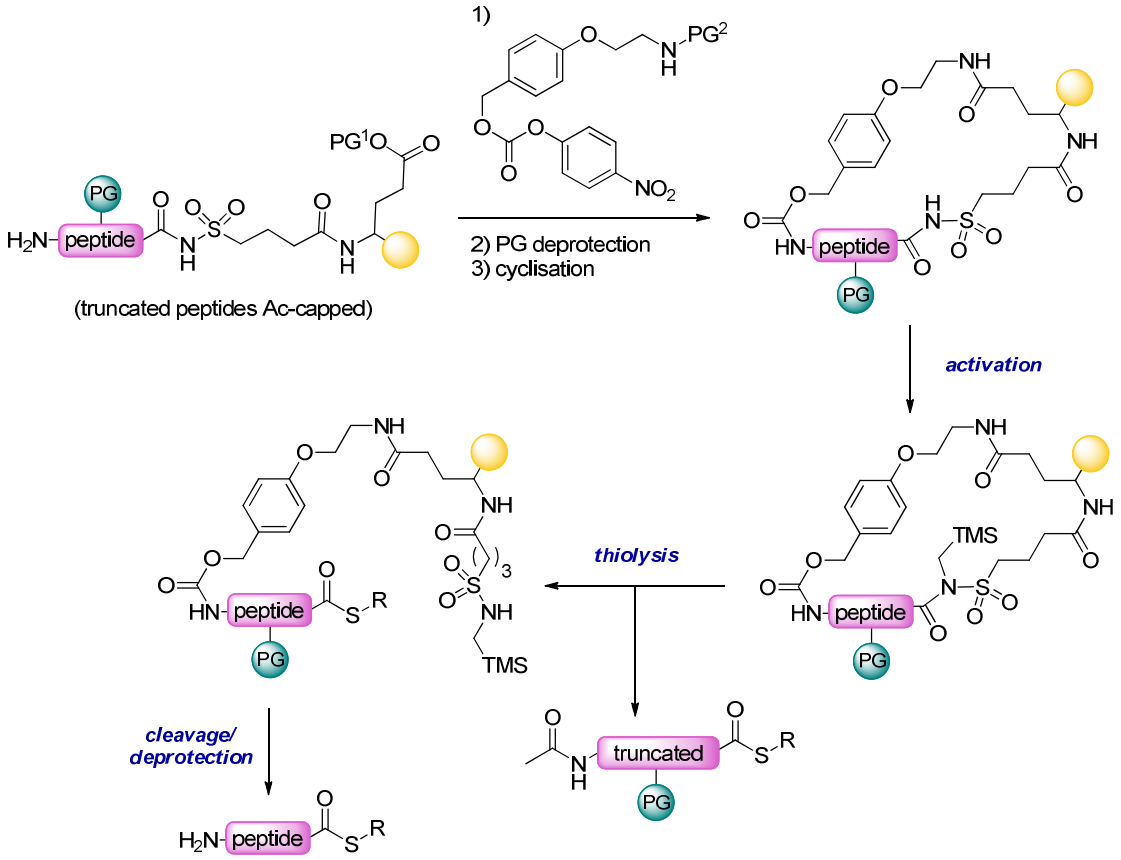
In contrast to the above summarised results, scientists have also encountered few drawbacks in the application of Native Chemical Ligation. One of the major drawbacks of NCL is the need for peptide thioesters, which are not only difficult to synthesise but are also prone to hydrolysis as well as oxidation. Especially in its early stages, thioester synthesis was limited. Few methods

were known for the selective synthesis of peptide thioesters without epimerisation or hydrolysis and they were all based on the application of Boc-chemistry in solid phase peptide synthesis.[29] In 1999, Dawson published the employment of a thiol resin which enabled the direct synthesis of any desired peptide thioester after full peptide cleavage and deprotection with HF.[30] But the toxic and corrosive nature of HF requires special equipment. In addition, repetitive TFA treatment hampers the synthesis of acid-labile structures, e.g., glycopeptides. This evoked the need for milder and more efficient Fmoc-based synthetic strategies for peptide thioesters.

To date, several alternative approaches have been published. One strategy involves the cleavage of protected peptide fragments from highly acid labile resin followed by thioester formation in solution.[31] Thereby, poor solubility due to high hydrophobicity of the protecting groups and their steric hindrance towards the reacting carboxylic acid call for solid support approaches. An alternative is the side chain anchoring of peptides, e.g. by Wong *et al.*, but this method demands an appropriate functional group in the side chain for resin anchoring.[32] Other approaches are based on the *in situ* synthesis of a thioester by an ester-thioester or amide-thioester rearrangement. The installation of a disulfide-protected thiol in β-position of the C-terminal amino acid leads after cleavage by addition of a free thiol to the deprotection of the disulfide bond and enables an O→S or N→S shift, respectively, to afford the desired thioester. Several different efficient strategies have been developed by Danishefsky[33] and Gaertner[34] and Meldal[35], respectively.

Another common strategy is the application of “safety catch” alkane sulfonamide resins that are stable towards base treatment, but upon addition of thiols the strategy yields the desired protected thioester peptide in solution.[36] An improvement of this strategy is the self-purifying “safety-catch” resin introduced by the group of Seitz (Scheme 1.4).[37]

Thereby, an N-to-C-cyclisation of the fully synthesised peptide retains upon thiol addition the desired peptide thioester on resin while all other peptidic material should be cleaved and washed off during that step. Afterwards, treatment with TFA furnishes the desired thioester peptide in a highly pure fashion. This strategy helps to avoid tedious purification steps that could lead to thioester hydrolysis and loss of material.



Scheme 1.4: Self-purifying "safety-catch" thioester synthesis.

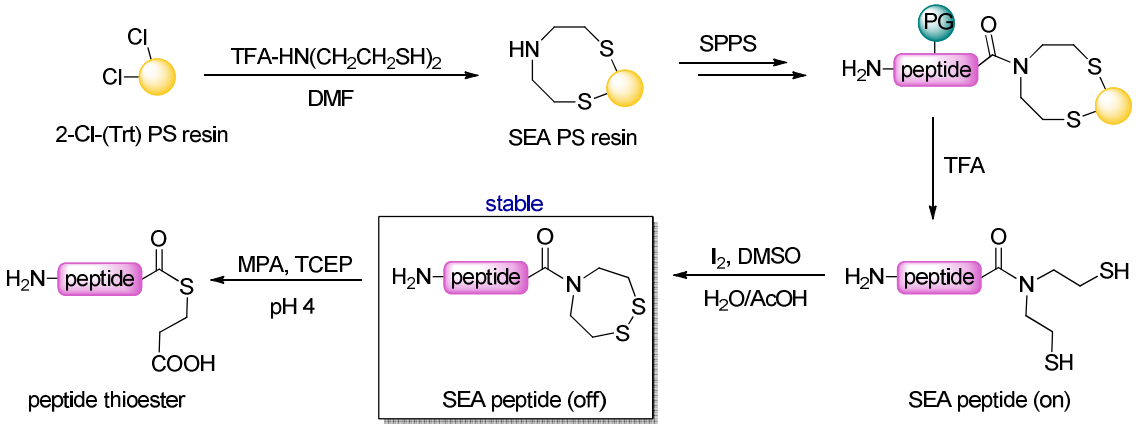
1.1.2.4. Kinetically Controlled Three-Segment Ligation by NCL (C-to-N)

The high reactivity of peptide thioesters does not only hamper their purification, but it also makes N-to-C sequential peptide ligation strategies more complicated due to the necessary thioester protection. Kent *et al.* partially solved this problem by introducing a kinetically controlled version of NCL making use of the inherent different reactivity of aryl- and alkylthiols.[25] They could show the efficacy of the method for the synthesis of several peptides and proteins constructed of up to six peptide fragments. However, this method does not completely avoid partial thioester hydrolysis.

1.1.2.5. Synthesis of Protected Peptide Thioester

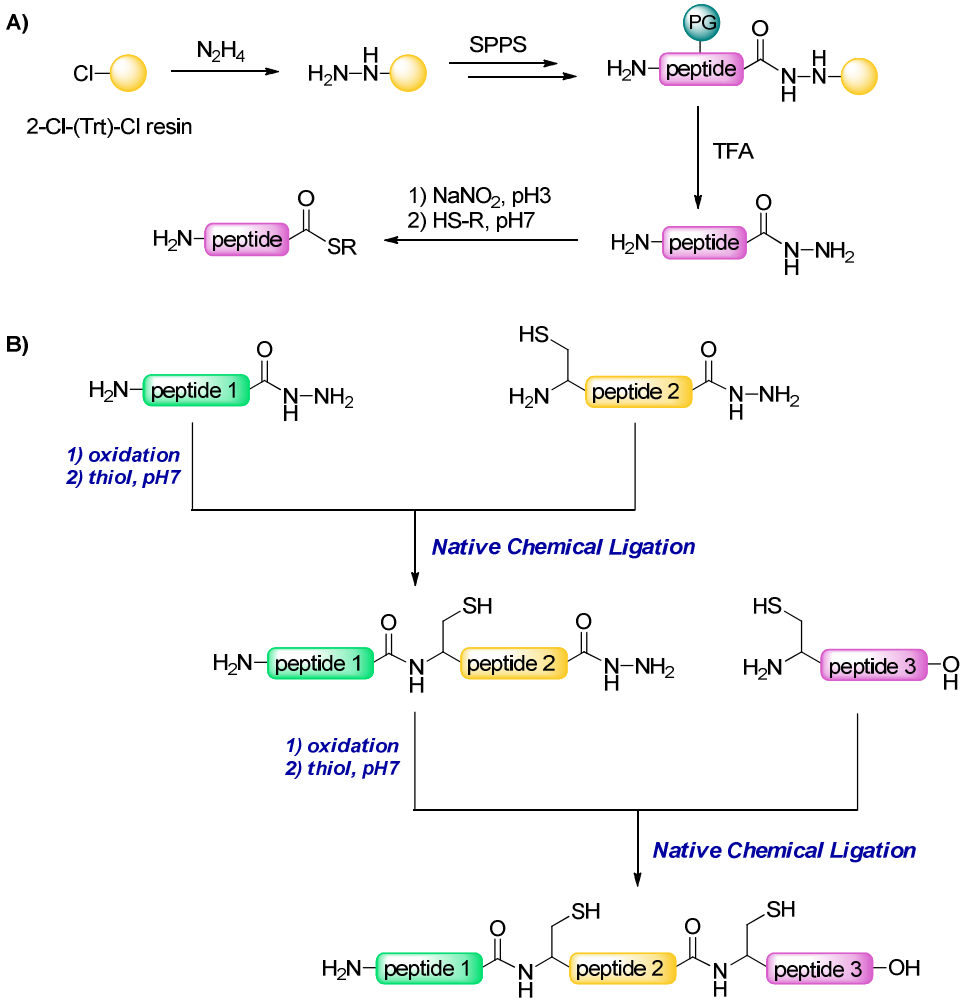
To further circumvent these problems, scientists are still on the lookout for new strategies that avoid excessive handling with unstable peptide thioesters. In 2010, the group of Melnyk published a new Fmoc solid-support synthesis strategy for peptide thioesters leading to the formation of alkylated amides as stable thioester precursors (Scheme 1.5).[35b] To do so, peptides

are synthesised by Fmoc-SPPS on a bis(2-sulfanylethyl)amido (SEA) polystyrene resin. Upon TFA cleavage, the peptide is released with a tertiary amide on its *C*-terminus with two free thiol groups attached. To protect the peptide from an undesired N→S shift, oxidation with iodine in acetic acid oxidises both thiols to form a stable disulfide bond. The peptide can now be purified and stored until further application. Upon thiol addition, these bis(2-sulfanylethyl)amido peptides can be easily transferred to their reactive thioester counterparts for direct application in NCL (Scheme 1.5) at pH 4 and 37 °C, though some more difficult amino acids such as proline might need temperatures of up to 65 °C for full conversion.



Scheme 1.5: Synthesis of SEA peptides by Melnyk (MPA: mercaptopropanoic acid, TCEP: (tris(2-carboxyethyl)phosphine)).

Only recently, the group of Lei Liu has published a new strategy that involves the direct and efficient synthesis of peptide hydrazides by Boc- and Fmoc-based solid phase peptide synthesis (Scheme 1.6A).[38] These peptide hydrazides are not only more stable during synthesis, purification and storage, but they can be readily converted into peptide thioesters when needed following a one-pot two-step protocol (Scheme 1.6A). Treatment of the hydrazide with $NaNO_2$ at pH 3–4 generates a peptide azide very efficiently within 20 min. Upon addition of 4-mercaptophenylacetic acid (MPAA) and readjusting the pH to 7.0, the desired peptide thioester is formed within 2 h in high yields. As the thioester formation takes place under NCL conditions it can directly be combined with the subsequent peptide ligation by NCL (Scheme 1.6B). This strategy does not only enable *in situ* generation of otherwise rather unstable thioesters but also offers a new approach for N-to-C sequential ligation strategies.[39] The hydrazides serve as protected thioester functionalities that can be converted *in situ* before addition of the next cysteine hydrazide peptide as shown in several publications with up to six peptide fragments (Scheme 1.6B).



Scheme 1.6: A) Synthesis of peptide thioester from hydrazides; B) Three-segment ligation with peptide hydrazides (N-to-C).

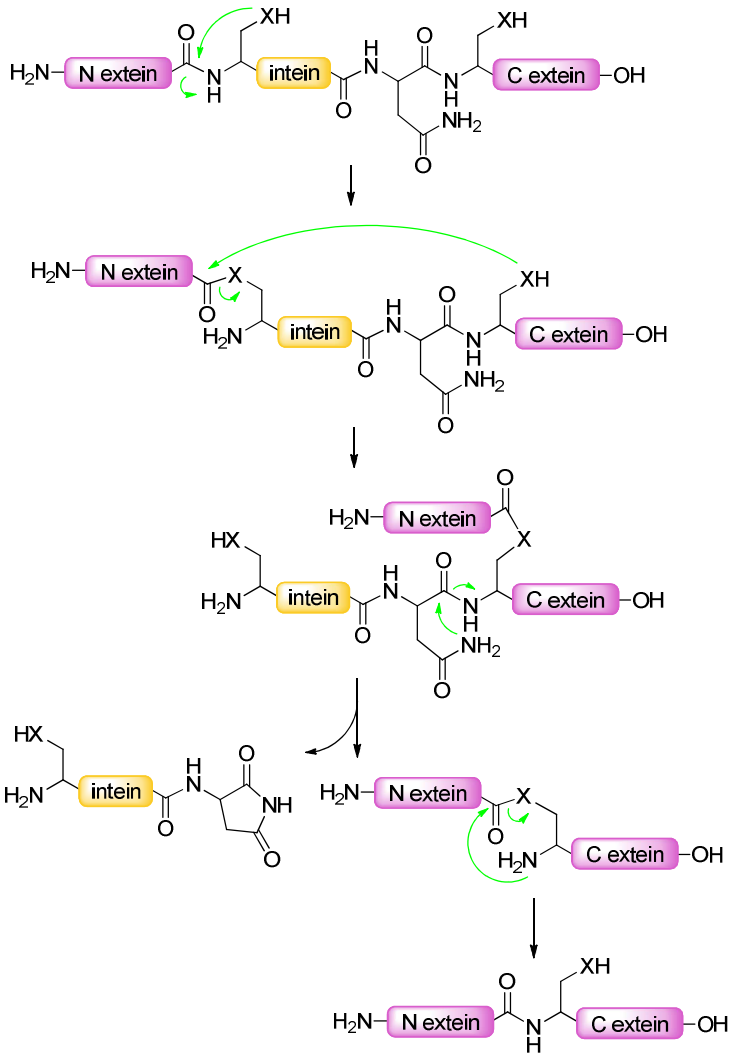
1.1.3. Protein Semi-Synthesis – Expressed Protein Ligation

One can envision that for larger proteins the concept of peptide synthesis and fragment ligation is a rather tedious procedure. Therefore, scientists have also applied the capture/rearrangement chemistry for the ligation of recombinant peptide and protein fragments with synthetic peptide segments.[40] On one hand, this strategy allows the incorporation of unnatural amino acids into a synthetic peptide thioester whereas the main part of the protein is biologically expressed bearing an *N*-terminal cysteine moiety, which can be achieved with different genetic methods such as site-directed mutagenesis and controlled cleavage with proteases.[41] NCL would then lead to the synthesis of a protein bearing an unnatural unit in close proximity to its *N*-terminus. The *vice versa* approach – Expressed Protein Ligation (EPL) – combines a recombinant thioester protein with a modified cysteine peptide and has also been proven to be a powerful tool.[42] Nevertheless, the necessity of a C-terminal protein thioester that is biologically expressed turned out to be more challenging than expected.

1.1.3.1. The Concept of Intein Splicing for Protein Thioester Synthesis

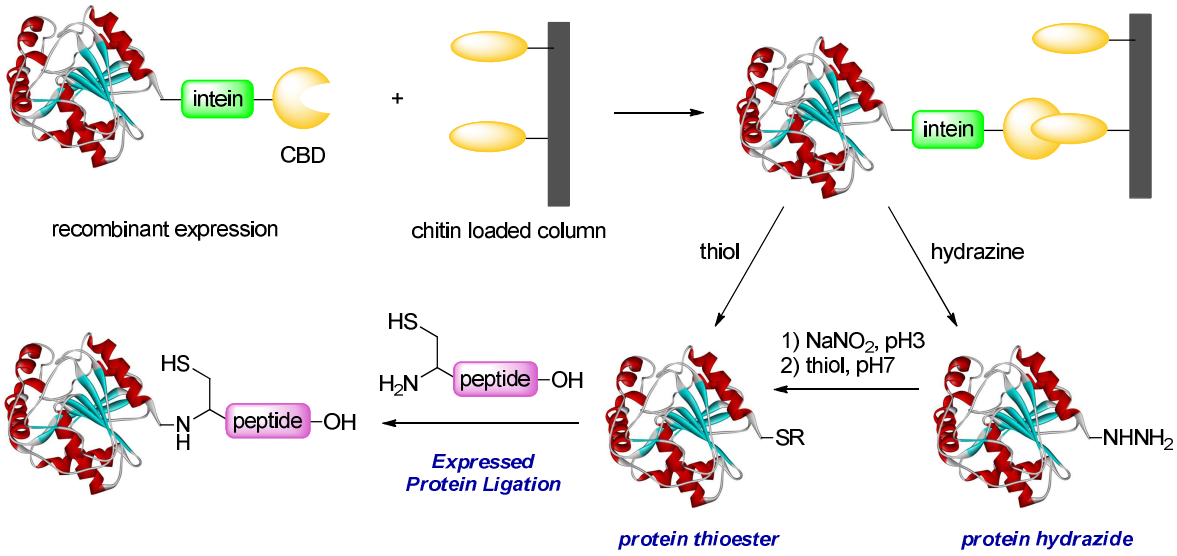
The concept was first published in 1998 by two individual groups and is based on preceding results in the field of intein *trans*-splicing.[43] In this biological process, two protein fragments – C and N exteins – are expressed separately each bearing an *N*- and *C*-terminal intein fragment, respectively. The individual intein pieces have no activity, but when combined to the full intein a cascade of rearrangement steps is triggered that combines the two external protein fragments and cuts out the intein to yield the active protein of choice.

The rearrangement mechanism, also termed as protein-splicing, is initiated by an N→S or N→O-acyl shift between the intein's *N*-terminal amino acid – threonine, serine or cysteine – and the *C*-terminal amino acid of the N extein that leads to the formation of a less stable (thio)ester bond (Scheme 1.7).[44] This usually leads to a conformational change that brings the ends of the intein and thereby the two extein fragments in close proximity. The *N*-terminal amino acid of the C extein can now undergo transesterification with the previously formed (thio)ester bond and combines the two exteins. The last amino acid of the intein is usually an asparagine. Its side chain amine can cleave the peptide bond between the intein and the C extein by formation of a cyclic imide and releases the desired protein that can eventually undergo an S→N- or O→N-acyl shift to yield a stable amide bond between the two exteins.[44] With this mechanism in mind, one can now envision the expression of a protein with a *C*-terminal intein which can be released after N→S-shift by the addition of excess free thiol to form a reactive thioester protein, which can then react with a synthetically provided cysteine peptide.



Scheme 1.7: Mechanism of protein splicing (X = S/O).

With concern towards the stability of protein thioesters and their storage, proteins are often expressed with a *C*-terminal intein–chitin binding domain (CBD, Scheme 1.8).[43b] The protein can then be fused to a chitin loaded column and separated from all other protein material. Upon addition of a thiol the thioester can be released from the column and directly reacted with a cysteine peptide *in situ*. Another recently published strategy is the application of split inteins for release and thioester formation upon intein reassembly on a column.[45]



Scheme 1.8: Protein Expression with intein-CBD, thioester/hydrazide formation and Expressed Protein Ligation.

1.1.3.2. Protein Hydrazides for Expressed Protein Ligation

A more efficient way to cleave the protein off the chitin binding column is the treatment with 8% hydrazine in phosphate buffer at pH 7.0 (Scheme 1.8).[46] This delivers protein hydrazide in high yields within a few hours. Protein hydrazides are equivalent to their peptide analogues more stable than thioesters and can be stored for longer as well as purified without loss of material due to hydrolysis. It enables further chemical modifications on the protein before activation by oxidation and formation of the more reactive thioester for subsequent NCL reactions. The only drawback is the rather harsh treatment of the protein due to several necessary pH changes and the reaction with sodium nitrite. In addition, Liu *et al.* could show on a model protein, the *E. coli* acid chaperone HdeA, that they could not only introduce an oxoester into the proteins backbone by using an α-hydroxy acid by genetic code expansion with a PylRS-$tRNA^{Pyl}_{CUA}$ pair (see chapter 1.5.2.2),[46] but they could also convert it to its corresponding α-hydrazide. The hydrazide was then converted to a more active thioester and applied in a NCL reaction with a cysteine peptide. The obtained protein had not only a correctly folded structure but also full bioactivity. This new approach allows for a more flexible protein semi-synthesis as it lacks the necessity of long intein fusion proteins. However, the many pH changes might be stressful to a protein and could result in undesired denaturation and/or aggregation.

To sum up, the field of protein synthesis is rapidly advancing. The need for chemically accessible peptides and proteins in large quantities has made the rather well developed Native Chemical Ligation an important tool in chemical biology. The constant development of new solutions in the field of the Fmoc-peptide thioester synthesis plus the newly advanced chemical systems that

can act as thioester protecting groups and be activated on demand for ligation have improved the field of peptide synthesis continuously and will do so in the upcoming years. In addition, scientists are constantly looking out for new methodologies to supplement the already know synthetic strategies and extent the field even further. Especially, the NCL's reliance on cysteine at the ligation site which is the second least common amino acid in proteins is still a major drawback. To circumvent this problem, many groups have recently worked on alternative strategies that employ β-mercapto amino acids, such as penicillamine or β-mercapto phenylalanine, to undergo NCL.[47] The thiol is then removed by reductive desulfurization methods, in which additional cysteines in the protein must be protected to avoid complete desulfurization. As coupling rates with more sterically hindered β-mercapto amino acids go down and radical desulfurization methods are still not that extensively established, the development of additional assembling strategies that rely on other functionalities than cysteine and thioesters and do not necessarily proceed *via* a capture/rearrangement mechanism is required.

An alternative is the Sec ligation, which relies on the ligation of seleno amino acids such as seleno cysteine and was first published in 2001 by the groups of Raines,[48] Hilvert[49] and van der Donk.[50] The fact that selenols possess a higher acidity than thiols allow fast peptide ligations also at slightly acidic pH, but the selenol's higher tendency to oxidise renders peptide ligations rather difficult.[51] Nevertheless, selenols exhibit the unique potential to be reduced to alanine in the presence of unprotected cysteine residues as published by the group of Dawson.[52] This strategy allows peptide ligations on alanine ligations sites without the need to protected potentially present cysteines. However, the difficult synthesis of selenol amino acid building block for peptide synthesis and their significantly lower redox potential does not render Sec ligations into a potential alternative for many peptide chemists.[53]

1.2. Amide Bond Formation Based on Azido Peptides

A very common approach towards chemoselective ligation of biomolecules is the introduction of non-native functional groups into biopolymers that react bio-orthogonally with another synthetically modified molecule without protection of distal functional groups.[54] Recent studies involve mostly the application of an azide group, which offers several advantages.[55] An azide possesses a high intrinsic energy distributed over a small volume. The group is not present in any naturally occurring compound and is bio-orthogonal towards almost all the other functional

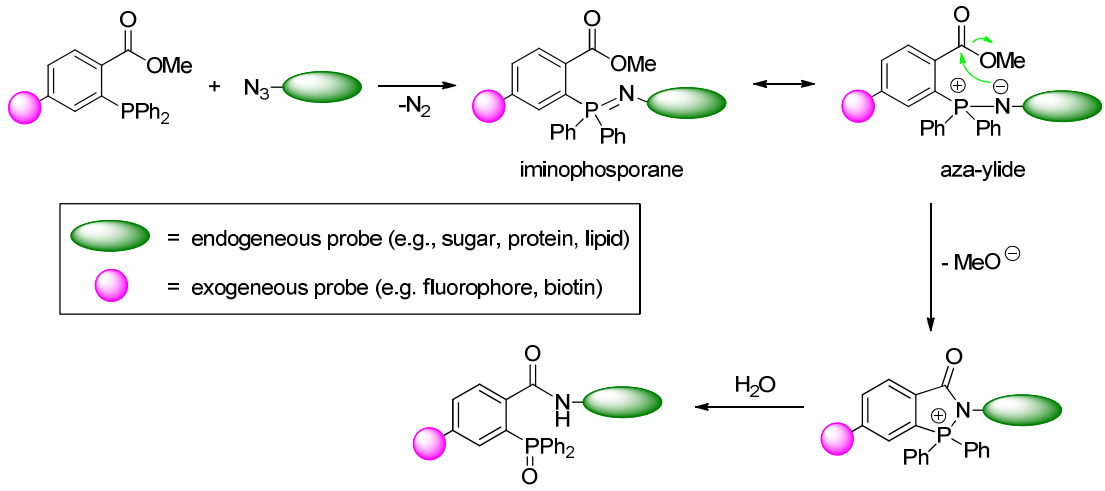
groups present in an organism (although reduction of aryl azides by free thiols is possible).[56] Furthermore, the azide exhibits a very specific reactivity towards some molecules.[57]

The incorporation of azides into peptides for ligation strategies is usually done by the introduction of azido amino acids. Depending on the subsequent transformation, the azide is situated either at the *N*-terminus or in the side chain of the amino acid. A multitude of synthesis protocols for a vast variety of azido amino acids has been published during the last decades.[58] There are two general ligation strategies known to convert azides into amide bonds which will be introduced in the upcoming chapters: the (traceless) Staudinger ligation and the thioacid–azide reaction.

1.2.1. Staudinger Ligation

All Staudinger reactions are based on the Staudinger reduction first presented by Hermann Staudinger in 1919.[59] It describes the reaction of a phosphine with an azide which leads to the formation of an iminophosphorane. Upon hydrolysis, the reaction yields a free amine and a phosphine oxide. To prevent the labile aza-ylide from hydrolysis, Bertozzi *et al.* introduced the Staudinger ligation in 2000 (Scheme 1.9), wherein an ester moiety was introduced as an electrophilic trap in *ortho* position of the phosphine on one of the aryl substituents.[60] After formation of the iminophosphorane, an intramolecular O→N shift is induced. The nitrogen of the iminophosphorane attacks the carbonyl group of the ester *via* formation of a five-membered cyclic intermediate. The alcohol of the ester is released and subsequent hydrolysis of the P-N bond yields an amide linkage and the phosphine oxide, thus the two products are still covalently bound to each other.

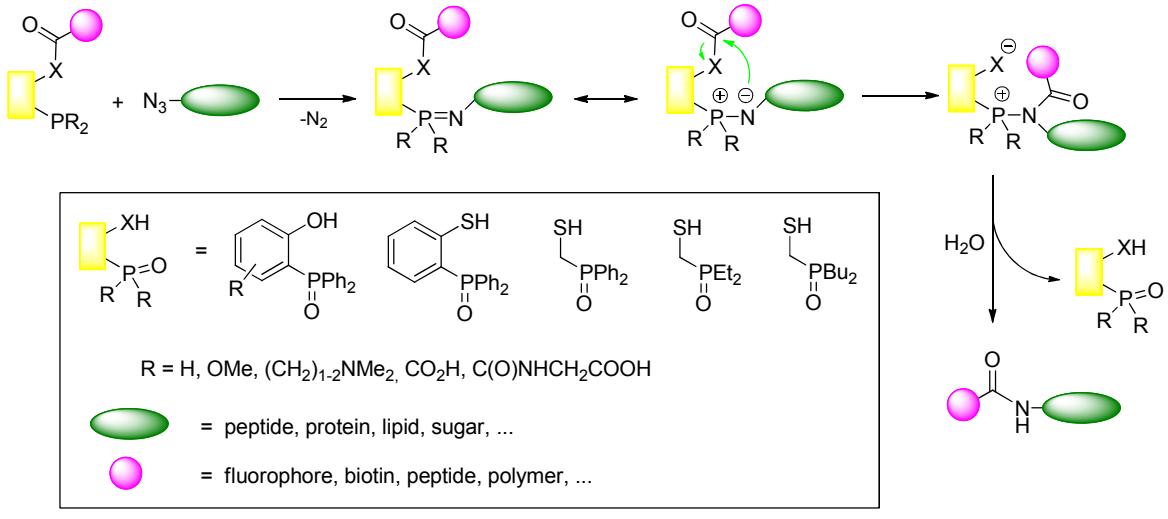
High selectivity and reaction rates of the intramolecular acylation step enable application of the reaction among biomolecules in an aqueous medium. Up to date, the Staudinger ligation has been applied in the synthesis of glycopeptides,[61] *in vitro*[55] and *in vivo*[62] biomolecular labelling, site-specific immobilisation of peptides and small molecules on a surface[63] and could furnish a fully active enzyme analogue.[64] Due to the residual bulk of the aryl phosphine oxide, the reaction has not been further employed in peptide and protein synthesis strategies.



Scheme 1.9: Mechanism of the Staudinger ligation.

1.2.2. Traceless Staudinger Ligation

To obtain a naturally elongated peptide, it would be necessary to get rid of the redundant aryl phosphine oxide, which is due to its size and hydrophobicity unacceptable in synthetic applications for natural peptides. The groups of Bertozzi and Raines simultaneously came up with the idea to switch the direction of the ester moiety on the aryl substituent (Scheme 1.10).[65] This way, the phosphine oxide is part of the leaving alcohol, which allows the attachment of the azide to any desirable carboxylic acid moiety. Several research groups came up with a variety of different phosphine derivatives (Scheme 1.10) bearing either an alkyl or arylthiol or -alcohol, allowing the rearrangement *via* a five- or six-membered intermediate.[66]



Scheme 1.10: Mechanism of the traceless Staudinger ligation.

During the last decade, further studies were performed on the electronic, steric as well as coulombic effects of these different aryl substituents on the phosphine (Scheme 1.10).[66a-c, 67]

They could show that bulky substituents slow down the reaction rate hampering the initial intermolecular reaction step between azide and phosphine. Electron-donating groups could abate the effect by enhancing the phosphines nucleophilicity as well as quicken the following S→N-shift. But enhancing the electron-donating character of the substituents, e.g. methoxy, was a tightrope walk as an increased electron density could also lead to faster protonation of the ylide nitrogen predominantly resulting in hydrolysis of the P-N bond yielding not the desired amide but the amine.[66b] The latter effect could be reduced by the introduction of less electron-donating alkyl chains carrying Lewis acid groups, such as dimethylamines, which were able to minimize protonation of the iminophosphorane nitrogen. These substituents did not only prove to be the best in yields but, in addition, showed much higher water-solubility enhancing the applicability of these phosphines in aqueous media.[66a]

During the last years, it has been shown that the traceless Staudinger ligation can not only be applied successfully in the cyclisation of several peptides of different lengths as published by Hackenberger,[68] but it can also be employed in the condensation of peptide fragments on solid support[64] as well as the site-specific immobilisation of azido proteins on a gold surface without perturbing the protein's structural and catalytic integrity.[69] In addition, the group of Raines could successfully generate a phosphinothioester with a genetically expressed protein.[66a] Unfortunately, it has not been put to a ligation, yet, but all the results give rise to the assumption that the traceless Staudinger ligation might be a promising candidate to support the field of peptide and protein synthesis.

One of the main features of the reaction is that not only no cysteine is needed at the ligation site, but also the reaction can proceed with any kind of azide, though studies have shown that electron-rich azides, e.g. any kind of alkyl amino acid azide, seem to further promote the reaction rate. However, the instability of the phosphine moiety towards oxygen makes it necessary to work with degassed solvents or with a high excess of the phosphine ester, which might lower the reaction's suitability for bigger protein ligations. In addition, the formation and purification of phosphinothioesters and -esters is still a scientific challenge.

1.2.3. Thioacid–Azide Reaction

1.2.3.1. Strategies for the Synthesis of Thioacid Peptides and Proteins

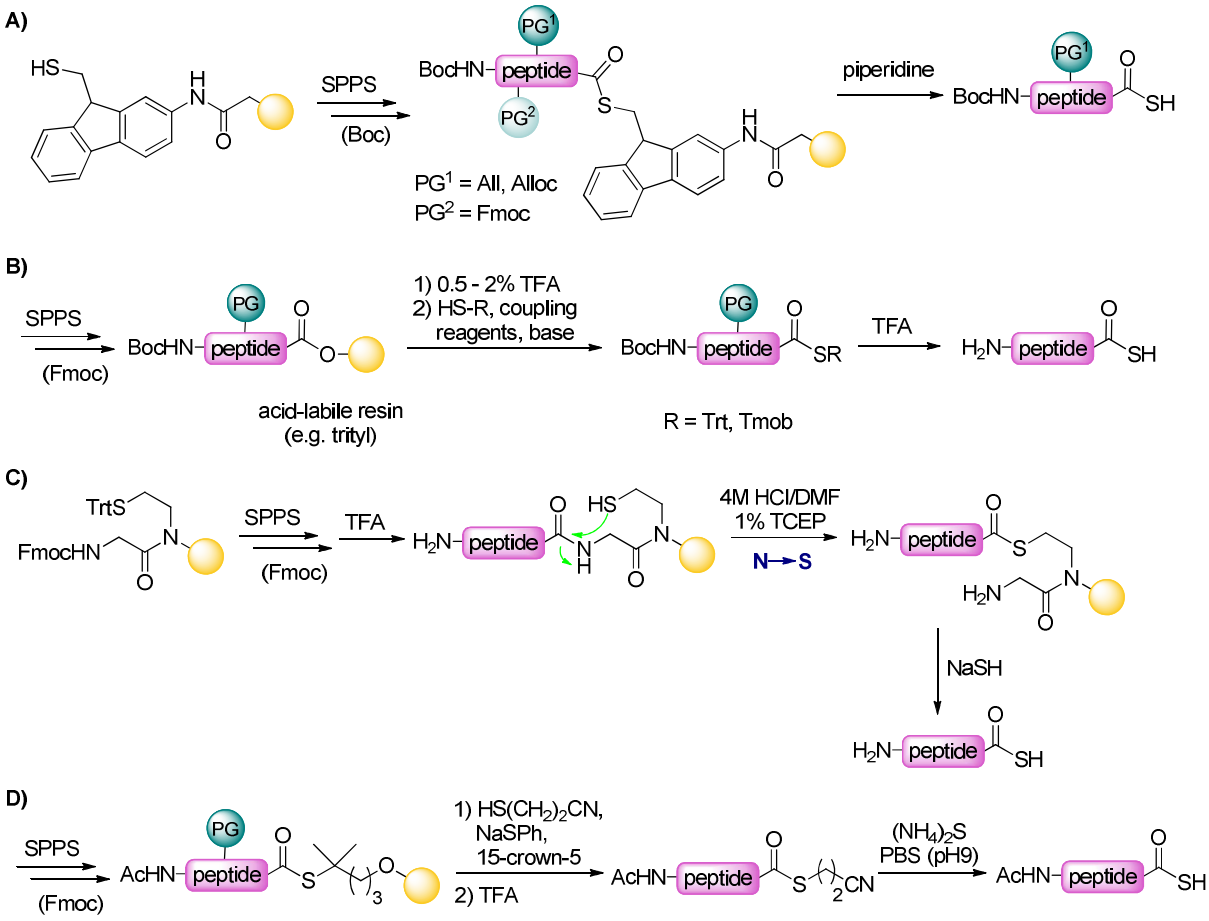
Another alternative to the well-known Native Chemical Ligation and the traceless Staudinger avoiding the synthesis of peptide (thio)esters might be the application of thioacids and their

selective reaction with azides.[70] This reductive amidation strategy has recently gained new attention in the field of bioorganic chemistry, but to date, few publications exist that describe efficient methods to successfully synthesise peptide and protein thioacids. First attempts are mostly based on solid phase Boc-chemistry but repetitive TFA treatment and final HF cleavage of the peptide from the resin can yield many side products excluding acid-labile peptides as mentioned above.[23, 71] In 2009, the group of Crich published a protocol that allowed solid phase Boc-synthesis of peptides on resin based on a 9-fluorenylmethyl thioester linker (Scheme 1.11A).[72] Final cleavage was therefore achieved by treatment with basic piperidine instead of the rather harsh HF conditions. The group of Liu published a protocol for the solid phase Boc-synthesis of resin-bound peptide thioesters.[73] Upon hydrothiolysis with hydrogen sulfide, the fully protected peptide thioacid could be released. With slightly lower yields, this approach can also be conveyed to Fmoc-protocols by application of a special Fmoc-deprotection cocktail to reduce possible aminolysis of the thioester.[74]

For peptides synthesised following the Fmoc-protocol, thioacid formation is often performed in solution phase. To do so, peptides are built up on an acid-labile resin, e.g. 2-chlorotrityl resin, and treatment with 1% TFA delivers a fully protected peptide in solution with a free terminal carboxy group. Afterwards, there are in general three different pathways to furnish a thioacid. On one side, the carboxylic acid can be activated, e.g. to a succinimide ester, and treatment with hydrosulphides delivers the desired thioacid.[75] On the other hand, 9-fluorenylmethyl, 2,4,6-trimethoxybenzyl and trityl thioesters can be formed in solution and converted into thioacids by base or acid treatment (Scheme 1.11B), respectively.[76] In a third strategy, Liu *et al*. published in 2008 an enzymatic approach converting peptide acids into thioesters that were subsequently cleaved by hydrothiolysis.[77]

Beside the first approach by Liu in 2008, merely two Fmoc-based SPPS protocols exist to date. In 2010, the group of Otaka in Japan published a strategy based on a (*N*-Fmoc-glycyl-*N*-sulfanyethyl)aminobenzoic acid linker (Scheme 1.11C).[78] During peptide synthesis on solid support, the thiol is trityl protected. Deprotection with TFA delivers the unprotected peptide in solution with a free thiol group on the linker attached. Treatment with 4N HCl/DMF in the presence of 1% TCEP induces an N→S shift towards a thioester. Upon addition of NaSH in phosphate buffer (1M, pH 9.2)/MeCN (85:15) an S→S shift leads to the formation of desired thioacid peptide. In 2012, Rademann *et al.* reported the application of a 4-mercapto-4-methylpentanol linker for peptide thioacid synthesis (Scheme 1.11D).[79] During peptide SPPS, the first amino acid is linked *via* a thioester that is stable towards standard Fmoc-deprotection

conditions but can be cleaved under mild basic conditions using sodium phenylthiolate. To prevent hydrothiolysis, 3-mercaptopropionitrile was added to capture the peptide as a thioester that can release the desired peptide thioacid upon treatment with ammonium sulfide under slightly basic conditions.



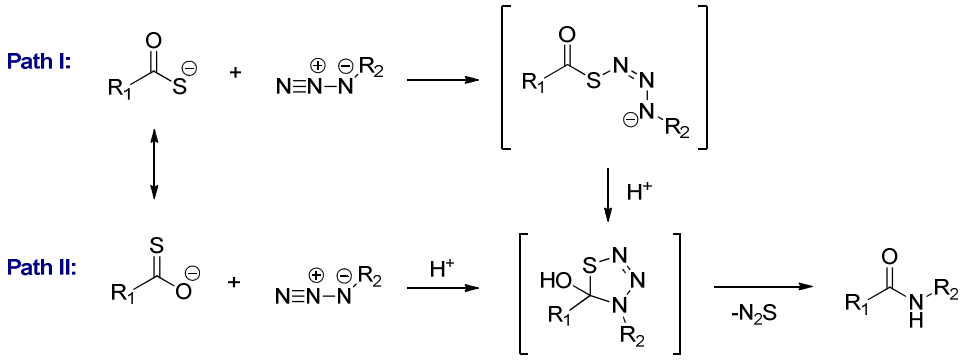
Scheme 1.11: Thioacid peptide synthesis: A) Boc-chemistry without HF, B) general in-solution approach, Fmoc-chemistry on resin by Otaka (C) and Rademann (D).

To bring this even one step further, the group of Begley has already published in 1998 a protocol to successfully express protein thioacids by taking advantage of the protein splicing mechanism.[80] An intein-CBD fusion protein was expressed and bound to chitin beads. Treatment with ammonium sulfide at 4 °C delivered the desired protein thioacid in solution. The recent advances in peptide and protein thioacid synthesis seem to promote the possible application of the thioacid–azide reaction in peptide and protein synthesis.

1.2.3.2. General Mechanism of the Thioacid–Azide Reaction

The general mechanism – as published by Williams *et al.* in 2006 – proceeds *via* formation of a thiatriazoline intermediate that undergoes subsequently a retro-[3+2]-cycloaddition to release

an amide and N_2S which quickly decomposes to nitrogen and sulfur – two non-hazardous side-products (Scheme 1.12).[81] Thereby, the nature of the precise mechanism is dependent on the electronic properties of the engaged azide. The route towards the thiatriazoline can proceed *via* two different pathways as shown by the group of Williams based on experimental results from Hammett correlation, kinetic data, chemical trapping experiments and ESI-MS/MS fragmentation studies as well as theoretical results by computational DFT calculations. Their findings encouraged previous speculations for electron-deficient and modestly electron-rich azides to undergo first of all an intermolecular reaction step between the electrophilic terminal nitrogen and the nucleophilic sulfur of the thiocarboxylate followed by the intramolecular formation of 5-membered cyclic thiatriazoline with the nucleophilic inner nitrogen adding to the electrophilic carbonyl group (Scheme 1.12, Path I). Thereby, it has to be mentioned that due to the highly polar character of electron-poor azides the formation of the N-S bond towards the linear intermediate seems to proceed rather fast and activates the azide additionally to undergo a more rapid thiatriazoline formation. Consequentially, electron-deficient azides, e.g., sulfonyl and glycosyl azides, afford higher yields in a short amount of time under relatively mild reaction conditions. The reaction proceeds comparably well in a variety of solvents ranging from organic to aqueous systems. In contrast, highly electron-rich azides, e.g., alkyl azides, can directly undergo formation of the thiatriazoline in a [2+3]-cycloaddition which seems to proceed much slower due to the rather unpolarized nature of the azide (Scheme 1.12, Path II). To generate satisfactory conversion yields, heating to elevated temperatures at higher concentrations is necessary.[81]



Scheme 1.12: General mechanism of thioacid–azide reaction.

So far, most applications of the thioacid–azide reaction rely on employing electron-deficient azides for a vast variety of applications, e.g., sulfonamide resin loading,[82] protein biotinylation,[83] short glycopeptide synthesis,[76b] peptide ligation[79] and reactions with selenocarboxylates.[84]

One publication seems to outstand from the others as it describes a successful $RuCl_3$-promoted amide formation between thioacids and electron-poor glycosyl as well as electron-rich alkyl azides at room temperature in protic solvents: methanol and water.[85] In addition, a study by Gao *et al.* indicates a solvent dependent reactivity of different azides. Theoretical studies seem to show that electron-deficient azides prefer non-polar solvents for successful amide formation whereas electron-rich azides work equally fine in polar and non-polar solvents with a slight tendency towards polar solvents, such as, water and methanol.[81b] With these results in mind, one could envision the application of the thioacid–azide reaction in aqueous media for the formation of natural alkyl amides from electron-rich alkyl azides in peptide and protein synthesis.

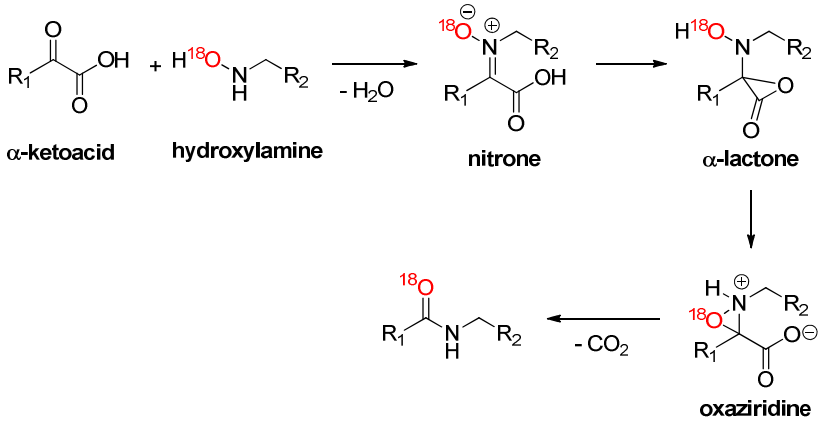
1.3. KAHA Ligation

Another method that relies on a chemoselective approach for amide bond formation but does not employ azides and has gained a lot of attention during the last years is the KAHA ligation (derived from its reagents: α-ketoacids and hydroxylamines).[86] Discovered and developed by the group of Bode and first published in 2006, the ligation strategy is based on the formation of nitrones from *N*-alkylhydroxylamines and aldehydes.[87] Nitrone formation seems to proceed fast with the equilibrium strongly leaning towards its product. With the sudden idea of employing not aldehydes or ketones but α-ketoacids in the reaction with *N*-alkylhydroxylamines a new chemoselective ligation strategy was born. The reaction proceeds very straightforward in DMF at 40 °C without any additional reagents producing carbon dioxide and water as the only side products. In addition, they could show that the stereochemistry of the α-keto amino acid is retained during the ligation with very high yields in amide formation.

A mechanistic study with isotopically labelled water $H_2{}^{18}O$ did reveal that the ligation is not as simple as expected and there is no ^{18}O introduced into the product at all during the reaction.[88] Thus, the carbonyl oxygen had to come from one of the two starting materials. In another experiment with ^{18}O-labelled hydroxylamine the product retained the full isotopic labelling (Scheme 1.13). Further investigations revealed the formation of an oxaziridine that then leads to decarboxylation to yield the desired amide.

In the successive years, the group not only managed to develop different synthesis pathways for α-ketoacid[89] and hydroxylamine amino acids and peptides[90] for a variety of successful peptide ligations,[91] but also developed different protection strategies for a consecutive three fragment peptide ligation.[92] Therefore, they introduced a non-canonical *N*-hydroxyamino acid –

5-oxaproline – that could be employed Fmoc-protected at the non-reacting *N*-terminus during sequential coupling without interfering with an ongoing ligation.[93] Upon Fmoc-cleavage the unprotected cyclic hydroxylamine could react with an α-ketoacid peptide leaving an unnatural homoserine behind at the ligation site. This approach also allowed the synthesis of an 82 amino acids long *C*-terminal variant of modifier protein UFM1 in a sequential three-fragment peptide ligation.[92]



Scheme 1.13: Mechanism of KAHA ligation with isotopically labelled hydroxylamine.

These results prove the potential of the KAHA ligation for peptide and protein synthesis with a few drawbacks, namely the oxidative formation of α-ketoacid peptides that requires oxone and is not compatible with methionine and cysteine residues, the introduction of homoserine as an essential amino acid for multiple fragment ligations and the not yet approached synthesis of expressed α-ketoacid proteins for a variation of Expressed Protein Ligation.

To sum up, the KAHA ligation seems very promising as an alternative to the widely spread NCL. It is highly chemoselective, proceeds well at sterically demanding ligation sites in good yields, does not require any additional reagents and leaves small innocuous molecules, such as H_2O and CO_2, as the only side products in the reaction.

1.4. Chemoselective Modification Strategies for Peptides and Proteins

All the above summarised ligation strategies seem to provide a way to selectively incorporate chemical handles and modifications into peptides and proteins close to the *N*- or *C*-terminus by the application of peptide and protein (semi)synthesis. *But how does one incorporate modifications into the middle of proteins without an elaborate multi-segment-ligation approach that is cumbersome and generally provides rather low yields? What can one do if the previously introduced modification hampers or complicates the ligation step or if the protein is misfolded due to just partial fragment expression? How can one introduce a modification on a readily*

expressed protein and investigate a modification's influence on the protein's structure and function?

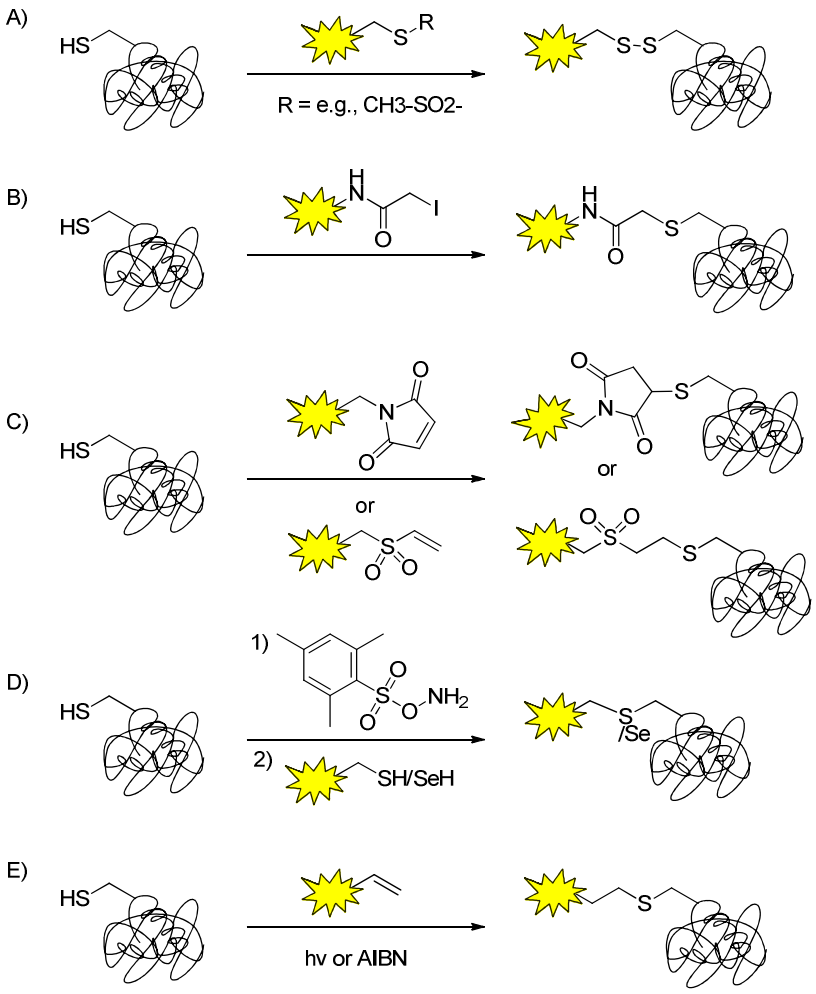
1.4.1. Canonical Amino Acids

There are several chemical modification techniques known to selectively address canonical amino acids, e.g. cysteine, lysine, tyrosine and tryptophane, allowing specific alterations at the *N*-terminus or at any desired type of amino acid in a protein.[54]

1.4.1.1. Cysteine Modification Strategies

Due to the versatile reactivity of thiols, cysteine is the most commonly used amino acid for chemical modification techniques on proteins.[94] Its high reactivity compared to other basic amino acids, such as, serine, tyrosine and threonine, allows very selective modifications. The first reaction that usually comes into mind when thinking of thiols and proteins is the formation of disulfide bonds – a common posttranslational modification to stabilise a protein's structure. Disulfides can be formed, among other methods, by using methane thiosulfonates, which can be nucleophilically attacked by a cysteine thiolate (Scheme 1.14A).[95] However, disulfides often exist in an equilibrium, which renders full conversion towards mixed disulfides rather difficult due to possible reduction and competition with symmetric disulfide formation.

Cysteines can also react with alkyl halides in a nucleophilic substitution (Scheme 1.14B)[96] and undergo addition reactions with maleimides, vinyl sulfones and electron-deficient alkynes (Scheme 1.14C).[97] Although these strategies are commonly used, they also bear certain drawbacks. It is known that – depending on pH and accessibility of the cysteine side chain – side reactions with slightly less reactive basic amino acids, such as lysines, can occur, which are sometimes difficult to control. A technique developed by Davis *et al.*, which is also applying the concept of Michael addition but in a *vice versa* approach, is based on a preceding formation of dehydroalanine from cysteine (Scheme 1.14D).[98] In this elimination reaction, the protein is treated with *O*-mesitylenesulfonylhydroxylamine rendering the cysteine's thiol into a good leaving group, which is eliminated under basic conditions to yield dehydroalanine. With such a Michael system in hand, different thiol or selenol reagents can now add to the formed double bond in a Michael addition to yield a stable thio- or selenoether linkage. Unfortunately, the elimination-addition strategy leads to the loss of stereo-information at C_α and, as a consequence, delivers two protein diastereomers, which might influence the protein's structure and function.



Scheme 1.14: Common bio-orthogonal modification techniques for cysteine.

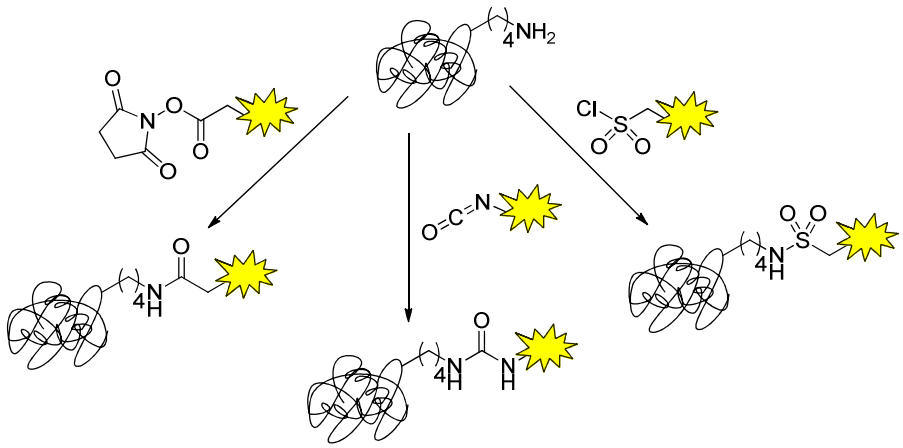
Another strategy to form thioether linkages is the thiol-ene chemistry, which has been used for a long time in polymer chemistry (Scheme 1.14E).[99] The approach exploits the propensity of thiols to react in the presence of a radical starter, e.g., hv or AIBN, with terminal alkenes. Although this modification technique seems to avoid side reactions with other basic amino acid side chains, thioether formation under radical conditions is reversible and thus can lead to incomplete conversion. Going one step further, scientists have also rediscovered the radical reaction of thiols with alkynes for protein modification, the thiol-yne reaction.[100] Alkynes are not only more reactive, but the formed vinyl sulfide can react again with a second thiol in a thiol-ene type reaction to introduce a second functionality.[101] Thus, the thiol-yne reaction has to be carefully handled to avoid cross-reactions between several cysteines in proteins. In addition, the group of Davis reported the opening of cystine bonds under the radical conditions applied in the thiol-ene/-yne ligation.[102] This protocol allows an additional functionalisation of former disulfide protected cysteines, but destroying disulfide bonds that are required to stabilise a protein's

structure and function might be crucial to the protein's activity thus highly limiting the reaction's applicability.

Overall, cysteine modification strategies feature efficient atom economy and, although cysteine is overall the least common amino acid, its low abundance often allows selective introduction to proteins by site-directed mutagenesis. However due to the cysteine's tendency to form disulfide bonds, modifications can be rather tricky sometimes and the different ligation reactions are often to a certain extent reversible.

1.4.1.2. Lysine Modification Strategies

A more prevalent amino acid that is used for protein modifications is lysine. Lysines can be modified with succinimide esters, isocyanates and sulfonyl chlorides (Scheme 1.15).[54] However, its high frequency in proteins usually leads to multiple additions on a protein molecule and kinetic control of site-selective lysine modification is rather challenging, cumbersome and not always possible.[103]

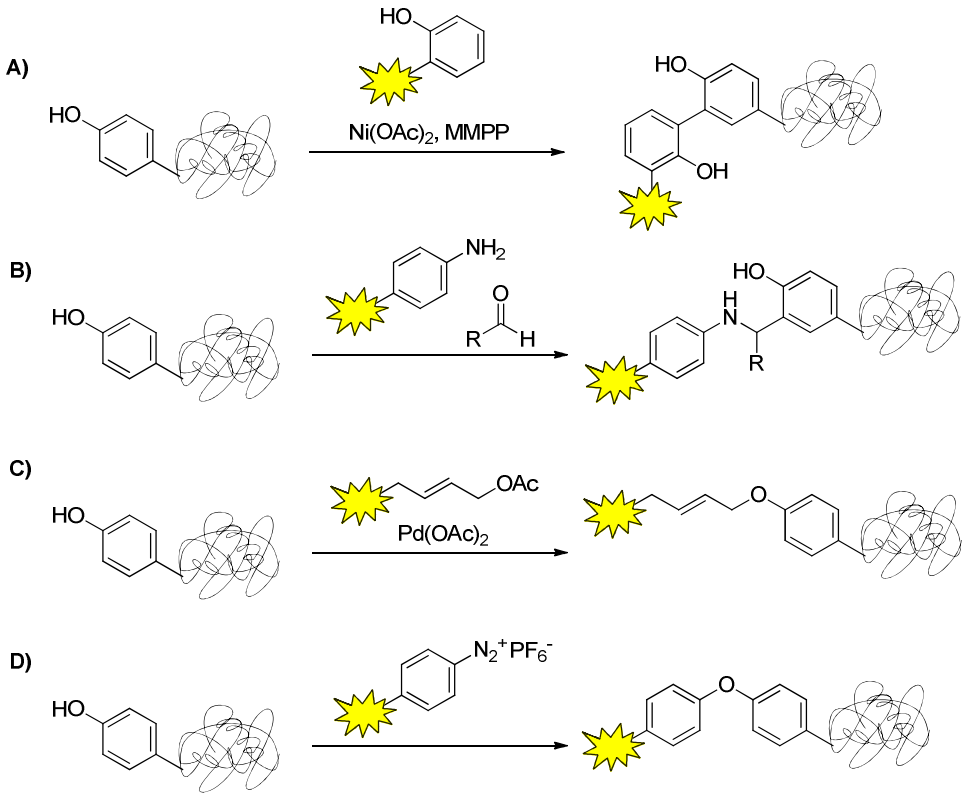


Scheme 1.15: Common bio-orthogonal lysine modifications on proteins with succinimide esters, isocyanates and sulfonyl chlorides (left to right).

1.4.1.3. Tyrosine and Tryptophane Modification Strategies

The amino acids tyrosine and tryptophane are relatively rare on a protein surface and allow more site-selective modifications.[104] However, there exist so far only few modification techniques for these amino acids. Tyrosine can be addressed in a transition-metal-based oxidative reaction, which crosslinks two phenol units (Scheme 1.16A).[105] First developed by the group of Kodadek for protein dimerisation, this strategy was further developed by Finn *et al.* to selectively introduce a variety of functionalities, e.g. biotin and alkynes, to a protein.[106] Another technique for controlled tyrosine modification was developed by Francis and co-workers and is

based on a three-component Mannich-type reaction with anilines and aldehydes (Scheme 1.16B).[104, 107] The same group also published a palladium-catalysed π-allylation (Scheme 1.16C).[108] Very recently, scientists have resumed studies on the electrophilic aromatic substitution of diazonium salts for the modification of tyrosine residues.[109] Its successful reintroduction relies greatly on the synthesis of more stable diazonium compounds for the application in aqueous media as shown by the group of Barbas III (Scheme 1.16D).[110]

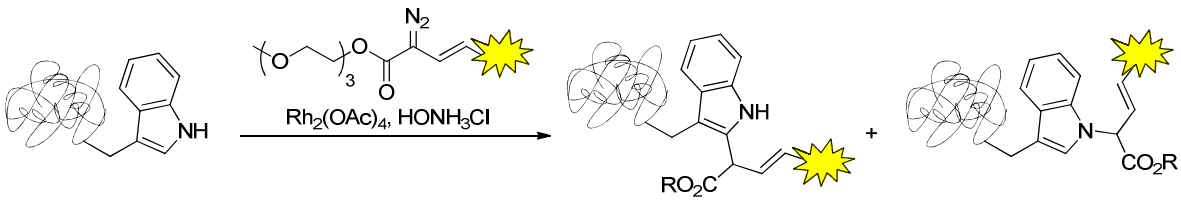


Scheme 1.16: Common bio-orthogonal tyrosine modification strategies.

Tryptophane as the rarest amino acid can react selectively with a rhodium carbenoid, which was generated *in situ* from a diazonium compound in the presence of $[Rh_2(OAc)_4]$ (Scheme 1.17).[111] However, the reaction, which was developed by Francis and Antos, requires very harsh conditions as it proceeds at pH 2 and might disturb the protein's overall structure.

As already mentioned, one of the main drawbacks of all of these methods is potential multi-functionalisation when there is more than one targeted amino acid side chain present in the overall protein. To overcome this problem, early works of Smith on site-directed mutagenesis allowed first changes of the amino acid sequence on DNA level to influence amino acid composition of a protein prior to expression.[112] The latter discovery was not only honoured with a Nobel Prize, but this method has also revolutionised the field of unnatural protein synthesis. However, changing a protein's primary structure at several positions can lead to

changes in fold and activity turning it useless for later applications. To introduce modifications into proteins more specifically without adding many changes to the natural amino acid sequence, new techniques were needed to achieve site-specificity as opposed to residue-specificity.



Scheme 1.17: Selective tryptophane modification developed by Francis and Antos.

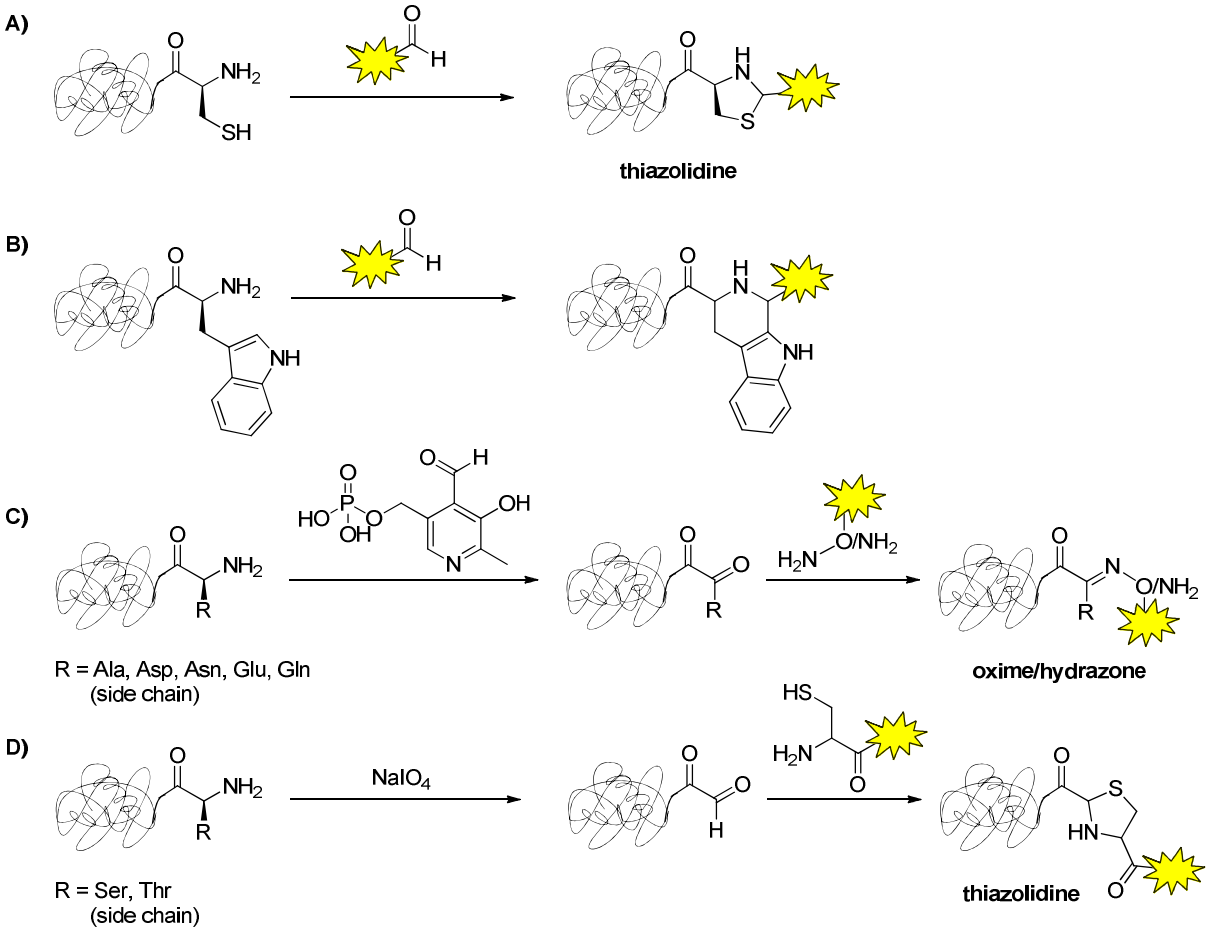
1.4.1.4. Modification of the *N*-Terminus

One possibility for a selective single modification on a protein is to address the unique reactivity of *N*-terminal amino acids leaving alone inner amino acid side chains. Depending on the amino acid present, there are a variety of more or less advanced techniques known ranging from chemical approaches to enzymatic ligation strategies.[54]

To start with *N*-terminal cysteine residues, they can be applied in amide bond ligation techniques, such as Native Chemical Ligation.[22] In addition, what is nowadays used as a common protection for *N*-terminal cysteines in these ligation strategies can also be applied for *N*-terminal cysteine modifications – namely the reaction with aldehydes and ketones to form thiazolidines (Scheme 1.18A).[27, 113] The speed and efficiency of thiazolidine formation thereby depends on the electrophilicity of the aldehyde or ketone with more electron-deficient carbonyl groups reacting faster and yielding a more stable product. The reaction of aldehydes and cysteines to form such cyclic structures also works with other *N*-terminal amino acids, e.g., serine, threonine, histidine and asparagine.[114] However, when tryptophane is positioned at the *N*-terminus it can also react with an aldehyde in a Pictet-Spengler reaction producing a stable C–C bond (Scheme 1.18B).[115]

There are also ways to selectively introduce aldehyde functionalities to a protein, which can be applied in a second modification step in a hydrazone or oxime ligation to achieve further functionalisation. To do so, one can theoretically treat any *N*-terminal amino acid side chain with pyridoxal-5-phosphate, which forms upon hydrolysis a pyruvamide that bears a ketone (aldehyde for Gly) and can be further modified (Scheme 1.18C).[116] Yields for this modification strategy depend very much on the *N*-terminal amino acid side chain leading to a variety of side-

reactions and conversion rates are not always satisfying.[117] A more straight-forward approach to full aldehyde formation on a protein's *N*-terminus is the oxidative cleavage of serine or threonine residues with sodium periodate (Scheme 1.18D).[118] Glycol cleavage thereby yields a highly reactive oxoacetamide group for subsequent modification with hydrazines, oxyamines or cysteines (Scheme 1.18C–D). However, treatment of a protein under oxidative conditions might risk oxidation of side chains, such as, methionine and cysteine. This effect can sometimes be toned down by careful pH control or addition of sulfur-containing compounds, but in most cases the application of periodate cleavage is limited to cysteine free proteins.



Scheme 1.18: Common bio-orthogonal modification strategies at the *N*-terminus of a protein.

As successful as some of these techniques are, they are limited to modifications at the *N* terminus of a protein. To introduce a functional moiety at a specific position away from the terminus and to avoid multiple functionalisation or excessive site-directed mutagenesis, a new concept is needed that is not solely based on canonical amino acids.

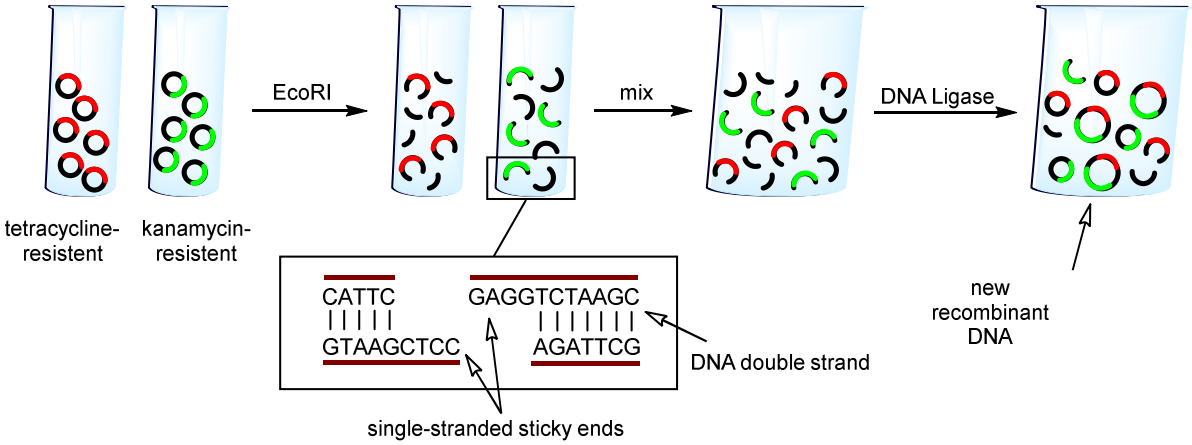
1.4.2. Biosynthetic Approach: *In Vivo* Introduction of Non-Canonical Amino Acids

1.4.2.1. Genetic Code Engineering

To be able to incorporate selective single modifications on a protein without having to rely on canonical amino acids, which might lead to multi-functionalisation or a variety of side reactions, scientists started to look out for a way to selectively incorporate non-canonical amino acids during protein expression. Ideally, such amino acid analogues would bear a bio-orthogonal moiety that could be modified later on in a chemoselective fashion. But the big question was: *How could one introduce a non-canonical amino acid into a protein other than using ligation techniques, such as Native Chemical Ligation?*

It was not before the early 1950s and 1960s that the group of Harold Tarver first proved the incorporation of an unnatural amino acid into proteins – namely ethionine as a methionine surrogate.[119] These experiments in mind, scientists started to target other amino acid analogues that could enter the bacterial amino acid pool the same way as natural amino acids and thereby take part in protein expression. But the incorporation of structurally and chemically similar amino acid analogues was always in competition with their naturally present counterparts. In 1957, a major breakthrough was achieved by Cowie and Cohen.[120] They employed a Met-auxotrophic *E. coli* strain for the introduction of seleno methionine (SeMet) during protein expression. Due to the lack of methionine in cells, cell growth is dependent on a continuous external supply of that amino acid. Replacing methionine with SeMet resulted in the synthesis of exclusively SeMet-bearing proteins all over the cell, which led to the formation of the first "unnatural" organism consisting of cells similar to their Met-counterparts, showing slower cell growth but nevertheless successful incorporation of a non-canonical amino acid. In addition to the latter discovery, the starvation of cells of a single amino acid was discovered as a supplementary tool to exclusively introduce unnatural amino acid analogues into proteins by an external amino acid supply.[121] In order to selectively synthesise just one desired unnatural protein in large amounts without interfering with the cell's proteome, it was only in 1973 that Cohen and Boyer revolutionised the field of protein expression developing new techniques to recombine DNA fragments and clone such genetically engineered material (Scheme 1.19).[122] To do so, two different gene plasmids bearing tetracycline and kanamycin drug-resistance, respectively, were isolated from bacterial *E.coli* strains. A gene restriction enzyme, EcoRI, was added to cut each plasmid off its restriction site producing short, single-stranded sequences called sticky ends. Upon mixing the cut plasmids, DNA ligase was added to randomly recombine

plasmid fragments – as the restriction enzyme produced DNA fragments with complementary ends – hopefully leading to plasmids that bear both drug-resistances.



Scheme 1.19: DNA fragmentation and recombination developed by Cohen and Boyle.

To replicate and isolate the desired plasmid, Cohen and Boyer mixed the plasmids with new bacteria and applied a “heat shock” to the cells. The “heat shock” technique was developed in 1970 by Mandel and Higa and is one of the most applied techniques for DNA delivery into cells, nowadays.[123] The strategy is based on a rapid change of lowered and raised temperature, which induces plasmid uptake also known as bacterial transformation. To conclude, the bacteria were grown in a culture containing both antibiotics, tetracycline and kanamycin. The set-up results in sole growth and survival of bacteria transformed with plasmids bearing both resistances on either one single or two separate genes, which was confirmed by subsequent analysis of the cut plasmids on agarose gel. Next, Cohen and Boyer could demonstrate the incorporation of genes from the South African clawed frog into *E.coli* cells – the first successful interspecies cloning.[124] With these results in hand, an efficient expression machinery was obtained, which allowed the production of large amounts of protein by inducing selective pressure on a translation system of a host cell. In 1977, Boyer and Itakura were first to succeed in the expression of a mammalian protein in bacteria – somastatin.[125] It could be shown that recombinant somastatin shows activities identical to the naturally occurring protein. From there on, the technology was taken on by many pharmaceutical companies to produce proteins, such as insulin, as biotechnological products for medical purposes.

As the newly developed strategy keeps a cell’s proteome unchanged targeting the production of a single protein, it was now possible to control the introduction of an unnatural amino acid solely into a protein of choice without affecting the host cell itself. Most importantly, the strategy is based on the principle to selectively impose an experimental pressure onto the cell

critically controlling factors such as fermentation parameters and amino acid supply.[126] The so-called SPI method – selective pressure incorporation – has been applied by many research groups to integrate a variety of non-canonical amino acids as amino acid surrogates during protein expression. Although many applications of this strategy have been published so far and there is a multitude of known surrogates to be introduced to a protein (Figure 1.1), there are also drawbacks to the auxotrophic expression system. Scientists are bound to introduce non-canonical amino acids that are structurally and chemically very similar to the substituted canonical amino acids and by substituting a specific amino acid, such as methionine, all methionine residues in a protein will be replaced at once without exception. Depending on the number of amino acids that are exchanged by a surrogate, an introduced bio-orthogonal handle might be present more than just once in a protein, which either requires selective site-directed mutagenesis on the DNA sequence or does not allow selective single-modifications anymore. In contrast, if one desires to incorporate a non-canonical amino acid more than just once throughout a protein sequence, selective pressure incorporation would be the method of choice.

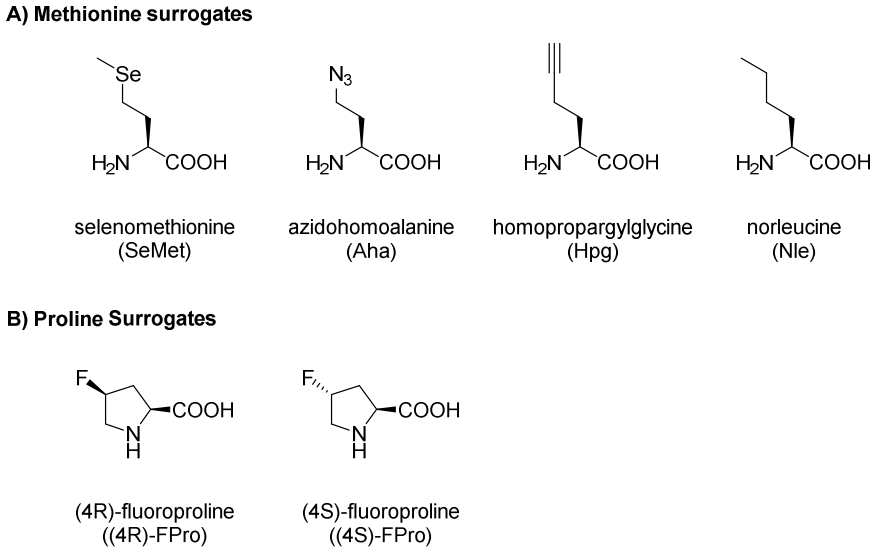


Figure 1.1: Non-canonical amino acid surrogates introduced by the SPI-method.

1.4.2.2. Genetic Code Expansion

Another possibility for selective incorporation of an unnatural amino acid is *via* the ribosomal translation system.[127] To do so, the non-canonical amino acid first needs to be attached to a tRNA in an aminoacylation step. The tRNA should then be recognized by its corresponding codon on the mRNA, which induces the transfer of the amino acid of choice to the protein by transacylation. To avoid cross-reactions, this specific codon must be orthogonal to the 61

degenerate triplet codons known for the incorporation of the 20 most common canonical amino acids.[2] This requirement in mind, there are only three known triplet codons left (Figure 1.2). These so-called stop or nonsense codons do not have a corresponding tRNA and therefore naturally terminate a protein synthesis.

	U		C		A		G		
U	UUU UUC UUA UUG	Phe Leu	UCU UCC UCA UCG	Ser	UAU UAC UAA UAG	Tyr Stop	UGU UGC UGA UGG	Cys Stop Trp	U C A G
C	CUU CUC CUA CUG	Leu	CCU CCC CCA CCG	Pro	CAU CAC CAA CAG	His Gln	CGU CGC CGA CGG	Arg	U C A G
A	AUU AUC AUA AUG	Ile Met	ACU ACC ACA ACG	Thr	AAU AAC AAA AAG	Asn Lys	AGU AGC AGA AGG	Ser Arg	U C A G
G	GUU GUC GUA GUG	Val	GCU GCC GCA GCG	Ala	GAU GAC GAA GAG	Asp Glu	GGU GGC GGA GGG	Gly	U C A G

Figure 1.2: Triplet codons and corresponding amino acids.

In 1989, the group of P. G. Schultz first published the application of a stop codon for the site-specific introduction of different unnatural amino acids in *E.coli* cells.[128] In this first approach, aminoacylation of the non-canonical amino acid to a corresponding tRNA was carried out chemically which proved to be very cumbersome and limited due to its stoichiometric necessity. Soon, the demand for *in vivo* acylation of unnatural amino acids to their desired tRNA increased. In organisms, aminoacylation is carried out by an enzyme called tRNA synthetase (aaRS). As every tRNA has its specific synthetase, it was necessary to develop a tRNA synthetase that selectively aminoacylates nonsense codon tRNA with the desired unnatural amino acid. Each component of the aaRS/tRNA base pair must act independently in the ribosomal aminoacylation machinery and not undergo any cross reactions with other base pairs. In addition, the amino acid of choice must be able to enter the cell and thereby be metabolically stable and nontoxic in its cellular environment.[129] First attempts were carried out in *E. coli* cells with endogenous aaRS/tRNA pairs but showed misaminoacylations with natural amino acids in the cell.[129] To find an orthogonal pair, *in vitro* studies were conducted locating aaRS/tRNA pairs from other organism than *E.coli* (*Ec*) that did not interfere with present *Ec*aaRS/*Ec*tRNA pairs and work efficiently enough in protein translations.

The first such pair was derived from methanogenic archaea applying engineered tyrosine RNA synthetase and a cognate nonsense suppressor tRNA (*Mj*TyrRS/*Mj*tRNA^{Tyr}), which was mutated to recognise the least deployed stop codon in *E. coli* (TAG) to interfere as little as possible with

the bacterial expression system.[130] To specifically incorporate unnatural amino acids, Schultz *et al.* developed a selection procedure that is based on a RNA synthetase library comprised of more than 10^8 active-site mutants of the *Mj*TyrRS. These were narrowed down in two steps.[131] First, the mutants' ability to suppress the nonsense codon was probed in the chloramphenicol acetyltransferase (CAT) gene expression. In the presence of an unnatural amino acid and different concentrations of chloramphenicol, the survival of the cell was probed. Successful aminoacylation of the suppressor *My*tRNA^{Tyr} with any amino acid and suppression of the amber stop codon leads to the synthesis of CAT that leads to cell survival. Effective mutants were then probed in a second round of screening. This time, expression of the toxic barnase gene carrying the amber stop codon at three permissive sites was performed in the absence of unnatural amino acids. This negative selection step could exclude all mutants that amino acylated endogenous amino acids and read over all nonsense codons resulting in the synthesis of the full length protein and cell death. Due to the double screening, an *Mj*TyrRS/*Mj*tRNA^{Tyr} pair was obtained that could incorporate >30 unnatural amino acids as efficiently and reliably as natural ribosomal aaRS/tRNA pairs.

During the last decade, the method of genetic code expansion was transferred to eukaryotic cells, such as *Saccharomyces cerevisiae* and different mammalian cell lines.[132] The orthogonal aaRS/tRNA pairs were thereby developed applying similar selection steps as described above. In addition, decoding pairs for TGA and 4-base codons have been developed for *E. coli* systems.[133] After the recognition of pyrrolysine (Pyl) as the 22nd genetically encoded amino acid, an orthogonal aaRS/tRNA pair was evolved for the introduction of pyrrolysine (Pyl) during protein expression.[134] These results were followed by the development of different mutants that enable selective incorporation of non-canonical amino acids by nonsense-suppression (for a selection, see Figure 1.3)[134b, 135] also allowing the introduction of two different unnatural amino acids (see chapter 1.6.2).

For the post-translational introduction of natural and unnatural modifications, a variety of bio-orthogonal handles has been introduced to proteins with the most common ones being azides, alkynes and ketones due to their small size and inert character in the presence of most functional groups at physiological pH (Figure 1.3).[20, 54, 136]

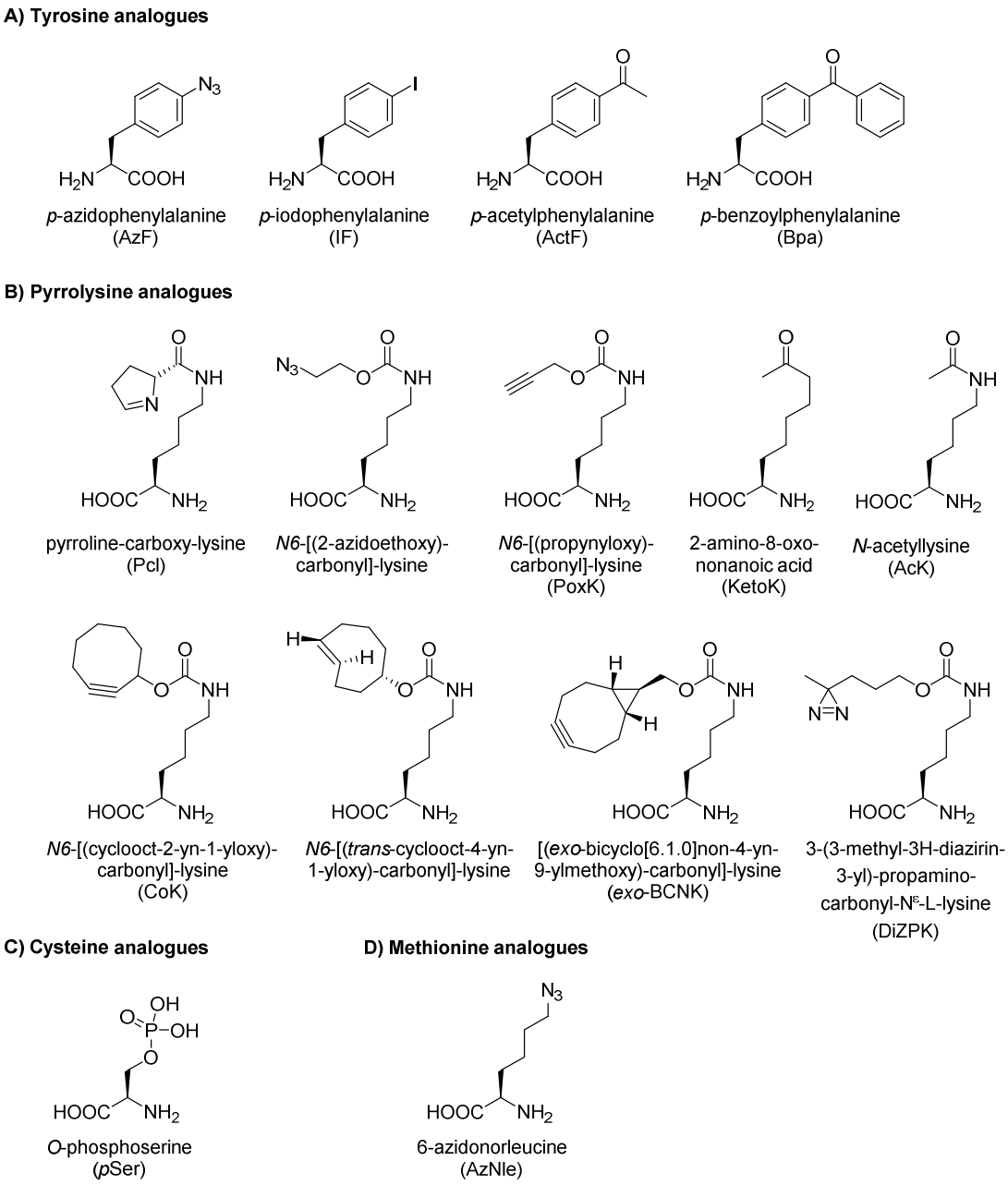


Figure 1.3: Selection of non-canonical amino acids for site-selective incorporation to proteins.

1.5. Chemoselective Modification Strategies

1.5.1. Reactions Based on Azides

The azide especially enjoys great popularity due to its many unique features as has been described before in chapter 1.2. During the last decades, a multitude of different chemoselective modifications have been developed and applied to azido biomolecules, such as the copper-catalyzed azide–alkyne cycloaddition (CuAAC) and the strain-promoted cycloaddition, different

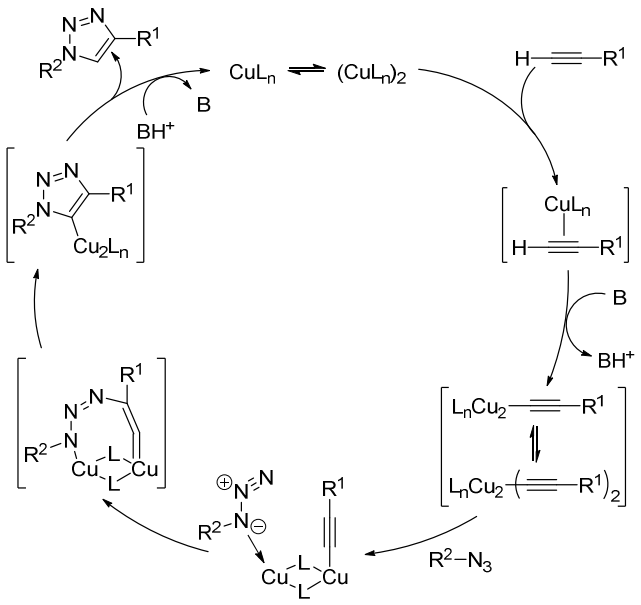
variations of the Staudinger reaction with phosphines, phosphites and phosphonites and the Sulfo-Click reaction between sulfonylazides and thioacids.

1.5.1.1. Copper-Catalysed Azide–Alkyne Cycloaddition (CuAAC)

The most frequently applied bioconjugation method with azides is still the copper-catalysed 1,3-dipolar cycloaddition (CuAAC) with terminal alkynes yielding 1,2,3-triazoles. In 1961, the reaction itself was first presented by Huisgen.[137] Without the application of any catalyst by sole heating, two possible products could be observed bearing the residues of alkyne and azide on the formed triazole either 1,2- or 1,4-oriented. In 2002, the groups of Meldal and Sharpless independently applied the azide–alkyne cycloaddition to peptide ligations and the synthesis of an inhibitor for the acetylcholine esterase, respectively.[138] Due to extremely low conversion rates, they added catalytic amounts of Cu(I)-salts to the reaction mixture. By doing so, they not only achieved a dramatic increase in conversion, but the reaction yielded solely the 1,4-disubstituted product.

The proposed mechanism suggests Cu(I) first to activate the alkyne by formation of a π-complex (Scheme 1.20).[139] As a consequence of the increased acidity of the alkyne, a copper acetylide is generated that can now form a bimetallic π-complex with an additional copper complex. The second copper coordinates to the azide's internal nitrogen and decreases its electron density. Thus activated, the terminal nitrogen of the azide is attacked in a nucleophilic fashion by the acetylide's β-carbon to form the first covalent C–N bond. Subsequent formation of the second covalent C–N bond leads to ring closure and to a triazole-copper complex, which undergoes a final copper-hydrogen exchange and yields the desired triazole product. This newly developed strategy enjoys ever since great popularity as its feasibility is rather quick and undemanding, and the reaction is easily applicable to a large variety of organic and bioorganic molecules in organic as well as aqueous media.[140]

Several variations have been developed, e.g., copper-stabilising ligands for organic and aqueous media[141] and ruthenium catalysis yielding exclusively 1,5-substituted triazole.[142] In contrast to the incorporation of azido amino acids, the *vice versa* approach that includes the incorporation of alkyne amino acids during protein expression followed by the modification of these with azido molecules *via* CuAAC has also shown great success.[143] However, the application of the CuAAC exhibits one major flaw in biological systems: copper is highly cytotoxic and leads quickly to cell death in living organisms.

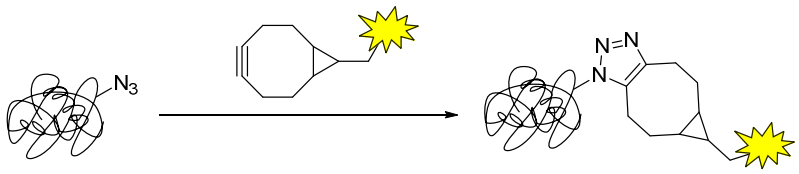


Scheme 1.20: Proposed mechanism for the copper-catalysed azide–alkyne cycloaddition (B = base).

1.5.1.2. Strain-Promoted Azide–Alkyne Cycloaddition

An advancement of the 1,3-dipolar cycloaddition that does not require the presence of Cu(I) exploits the highly energetic ring strain of cycloalkynes and was published by the group of Bertozzi in 2004 (Scheme 1.21).[144] Due to an angle deformation of 17° next to its triple bond, cyclooctyne – the smallest stable cycloalkyne – possesses a ring strain of not less than ~87 kJ/mol enhancing the molecules reactivity drastically. Therefore, triazole formation occurs without the necessity of any catalyst at room temperature. The so-called strain-promoted [3+2]-cycloaddition can even be accelerated in the presence of electron withdrawing groups, such as fluorine, next to the triple bond and a variety of cyclooctyne derivatives with higher reaction rates has been developed during the last decade. Merely the synthesis of such cyclooctynes is very complex and tedious and requires a multitude of reaction steps with moderate yields. In addition, Davis *et al.* could recently show that strained cyclooctynes have the propensity to react with free thiols in a thiol-yne type reaction but without the need of photoactivation.[102, 145] These results suggest partial limitations for the strain-promoted azide–alkyne cycloaddition in its application on thiol-containing biomolecules.

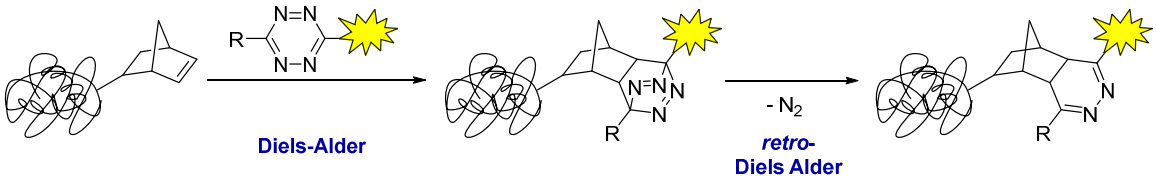
Nevertheless, several groups are currently investigating the applicability of the strain-promoted cycloaddition in biological systems, e.g., zebrafish and mice.[146] Very recently, few techniques have been published to introduce strained cycloalkynes, e.g., bicyclo[6.1.0]non-4-yn-9-ylmethanol, to proteins by genetic code expansion.[147]



Scheme 1.21: Strain-promoted azide–alkyne cycloaddition for protein modification.

1.5.1.3. Diels–Alder Reactions on Proteins

The incorporated cycloalkynes do not just allow the reaction with azido derivatives, but as reintroduced by the groups of Fox and Hilderbrand cyclooctynes, *trans*-cyclooctenes and norbonenes – all incorporated to proteins as pyrrolysine derivatives by amber suppression[147] – can also react with tetrazine in a rapid and efficient inverse electron-deficient Diels–Alder reaction just to undergo in a second step a *retro*-Diels–Alder reaction to release nitrogen and form a stable 4,5-dihydropyridazine (Scheme 1.22).[148] Though the presented labelling technique is still young, it has recently found a variety of applications in bioorthogonal protein functionalisations.[147, 149]



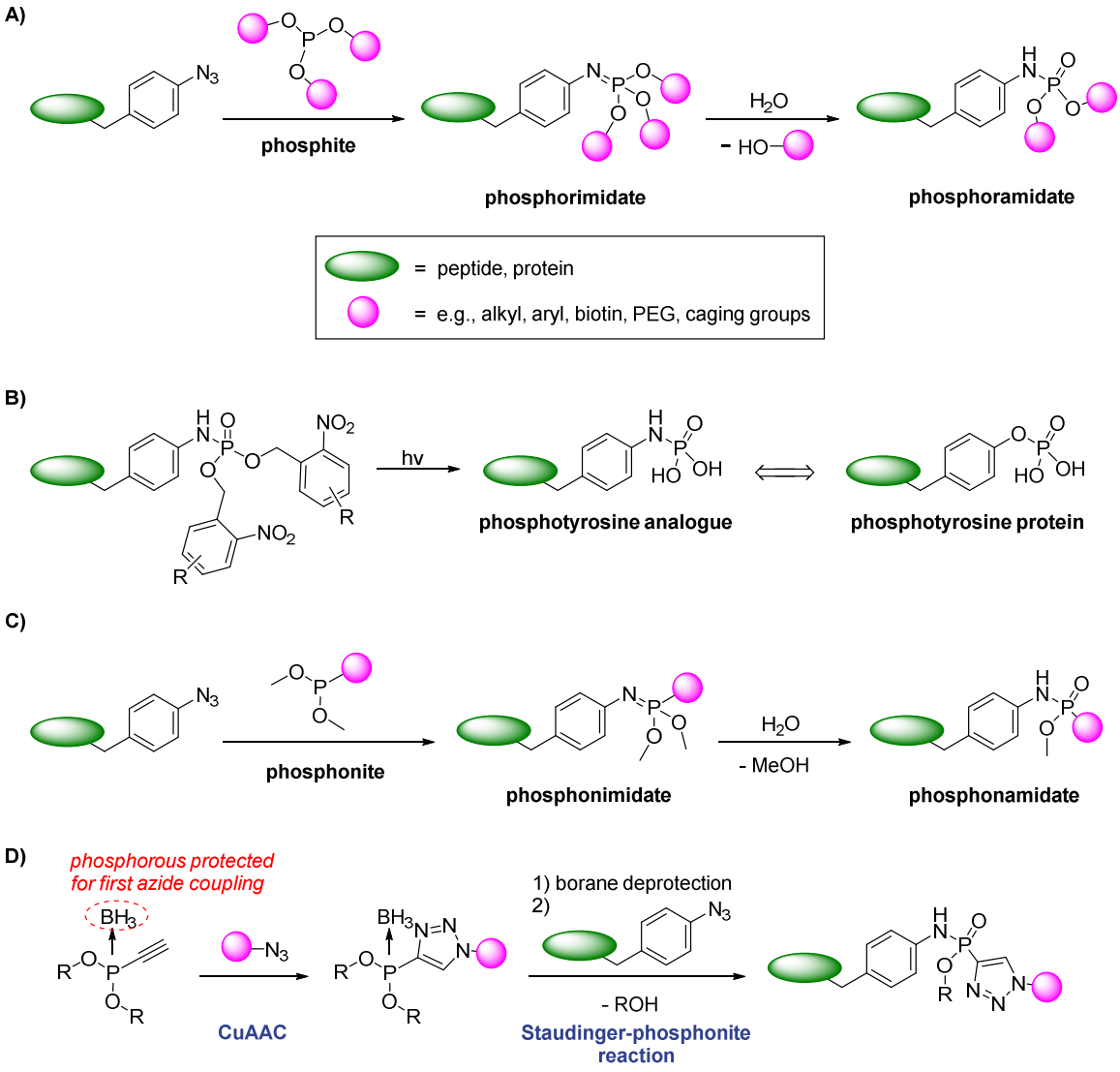
Scheme 1.22: Inverse electron-deficient Diels–Alder reaction.

1.5.1.4. Staudinger Ligations

Back to azide conversions, another bio-orthogonal ligation strategy that employs azides and can be applied in protein modifications is the previously mentioned Staudinger ligation and its traceless variant (see chapter 1.2.1 and 1.2.2). However as indicated before, one of the main drawbacks of this reaction is the employed phosphine's low solubility in aqueous media and its high tendency to oxidise, which requires a large excess during protein modification to still achieve high conversion rates. In addition, the iminophosphorane intermediate is prone to hydrolysis when the rearrangement is not fast enough limiting the range of the reactions application.

Hackenberger and co-workers have recently shown successful application of phosphites in selective modification of azido peptides and proteins (Scheme 1.23A).[150] Thereby, the azide reacts with a phosphite comparatively to the Staudinger ligation to first form a phosphorimidate.

Upon hydrolysis, not the amine but one of the alcohol groups attached to the phosphorous is released due to higher leaving group propensities and a phosphoramidate is formed, which is stable under physiological conditions. This strategy has not only been applied to introduce a variety of modifications, such as, PEG-chains, biotin and sugar moieties,[151] but with the introduction of light-cleavable caging groups to the phosphite a protein with an unprotected phosphoramidate could be generated – an analogue to natural phosphotyrosine (Scheme 1.23B).[150]



Scheme 1.23: A) General mechanism of Staudinger-phosphite reaction, B) formation of phosphotyrosine anaologues, C) general mechanism of Staudinger-phopshonite reaction, D) azide–azide ligation with alkyne phosphonite.

Due to difficulties to control leaving group propensities of different alcohols on phosphorimidates, mostly symmetrical phosphites bearing three identical alcohol functionalities were applied initially for peptide and protein modifications. For more precious, very hydrophobic and bulky substituents, e.g., biotin or oligosaccharides, a new approach was

developed that allowed the introduction of a single functionality, which could not be released during the Staudinger phosphite reaction. To do so, phosphonites bearing a stable P–C bond were successfully probed in the reaction with different azides (Scheme 1.23C).[152] However, the cumbersome synthesis of some of these phosphonites required a new synthetic strategy that allowed a more versatile approach towards functionalised phosphonites. This led to the development of a borane-protected alkyne phosphonite, which allowed direct combination of two azido molecules in a three-step procedure (Scheme 1.23D).[153] First, an azido molecule, e.g. azido oligosaccharide, can react selectively with the alkyne group by CuAAC as the phosphorous itself is borane-protected and not reactive enough. After purification and borane deprotection, e.g., by treatment with DABCO as a base, the functionally modified phosphonite can now react with a second azide, e.g. an azido protein, in a metal-free Staudinger phosphonite reaction avoiding the application of toxic copper on a biomolecule. First applications were shown for selective introduction of lactose moieties on azido polyglycerol in organic solvents, but also functionalisation of a protein with a biotin phosphonite could be achieved effectively.[153] However, beside many features of this ligation strategy, e.g., lack of toxic metals and benefit to connect two better available azido molecules, a problem is the instability and quick hydrolysis of phosphonites in aqueous media. For successful modification of biomolecules in buffer systems a high excess of phosphonite is required.

1.5.2. Ligations Based on Ketones

1.5.2.1. Oxime and Hydrazone Ligation as well as a Radical Approach

Beside the use of azido amino acids, scientists have also incorporated ketone amino acids into proteins (Figure 1.3).[132] Ketones have proven to be stable during protein expression and react sufficiently fast in oxime and hydrazone ligations. Although carbonyl groups react with alcohols and amines present in a protein, at physiological pH the equilibrium lies strongly on the side of the free ketone.[154] In contrast, oximes and hydrazones once formed are rather stable under physiological conditions and therefore represent an ideal way of protein modification.

The concept of ketone incorporation during protein expression has not been applied very often, yet. The group of Schultz has shown the incorporation of some ketone amino acids with *p*-acetylphenylalanine being the most successful one.[155] Subsequent reactions with hydroxylamine or hydrazines are usually executed at rather low pH of 3–5 to be efficient,[154] which might in some cases harm the protein's overall structure and function. To perform these

ligation strategies at physiological conditions, the group of Dawson has introduced the application of aniline as a catalyst that greatly accelerates the reaction at physiological pH and turns it into a rather mild protein modification.[156] As aniline has the propensity to denature proteins and shows rather low solubility in aqueous media, a variety of new biocompatible catalysts has been introduced recently, which show equal or higher efficiencies than aniline itself.[157]

Chin *et al.* have shown recently that ketones can not only be applied in selective oxime and hydrazone formations.[158] They have developed a way to introduce benzophenone as a ketone amino acid in good yields during protein expression (Figure 1.3). Its reactivity in oxime ligations is rather low. However, benzophenones are known as photocrosslinker and can be targeted as such also on proteins as shown by Chin. Careful handling of concentrations is necessary as this strategy could also lead upon UV irradiation to protein dimerisation.

1.5.2.2. Ketones vs. Aldehydes

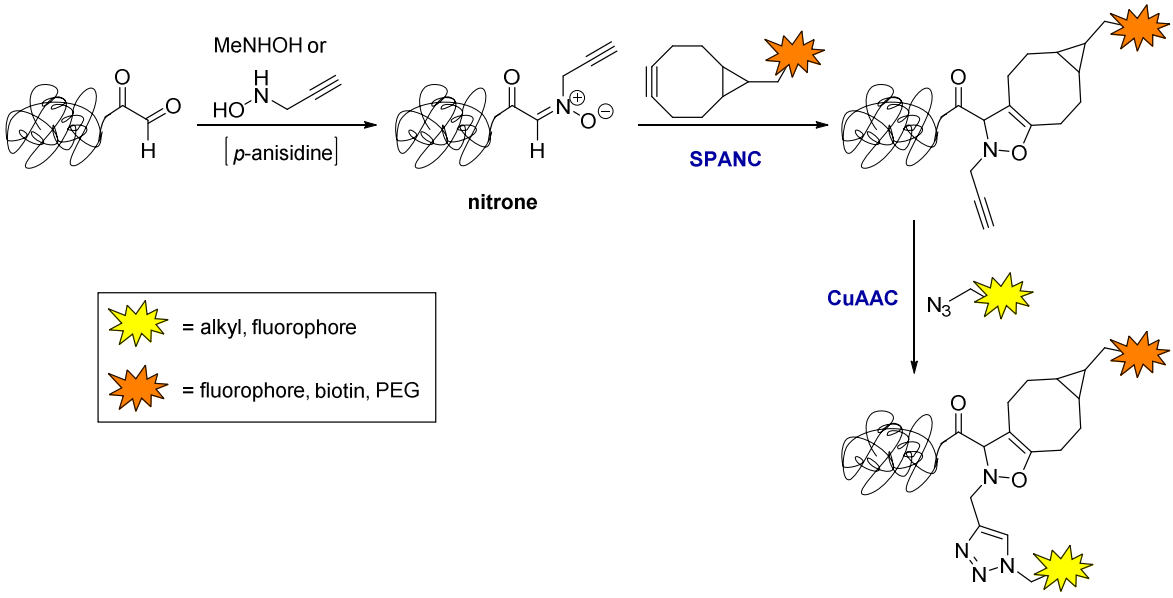
An alternative to the application of ketones are aldehydes. Due to the higher electrophilicity and less steric hindrance of aldehydes, they can react faster with hydroxylamines and hydrazines and thereby accelerate oxime and hydrazone formation.[159] However, the high reactivity renders their application in a protein expression system rather difficult and might be one of the main reasons why it has not been successfully addressed yet. Instead, as partially already discussed above common techniques to introduce aldehydes into proteins are solely posttranslational by chemical, e.g., periodate cleavage,[118] or enzymatic means, e.g. formylglycine synthesis by the group of Bertozzi using a cysteine containing hexapeptide tag.[160]

Once an aldehyde is introduced to a protein, it can not only be used in oxime and hydrazone ligations. As mentioned above, aldehydes and ketones can also react selectively with *N*-terminal cysteines to form thiazolidines, whereas aldehydes render the product faster and more stable (Scheme 1.18C-D).[161] This strategy was first applied for peptide and protein modifications by the group of Tam to introduce biotin to human parathyroid hormone and α_1-acid glycoprotein.

1.5.2.3. Strain-Promoted Azide–Alkyne Cycloaddition

As another selective aldehyde reaction, the group of van Delft has published a new modification strategy called strain-promoted alkyne–nitrone cycloaddition (SPANC, Scheme 1.24).[162] The nitrone is thereby formed by reaction of an aldehyde and *N*-alkyl hydroxylamine in the presence

of *p*-anisidine as catalyst. The obtained nitrone then reacts with a modified cyclooctyne in a [3+2] cycloaddition under physiological conditions in an almost quantitative fashion. Van Delft *et al.* also investigated the reaction kinetics and revealed that the reaction between nitrone and cyclooctyne is approximately 10–30 times faster than the strain-promoted azide–alkyne cycloaddition by Bertozzi.[162]



Scheme 1.24: Strain-promoted alkyne–nitrone cycloaddition (SPANC) for mono- and dual-functionalisation.

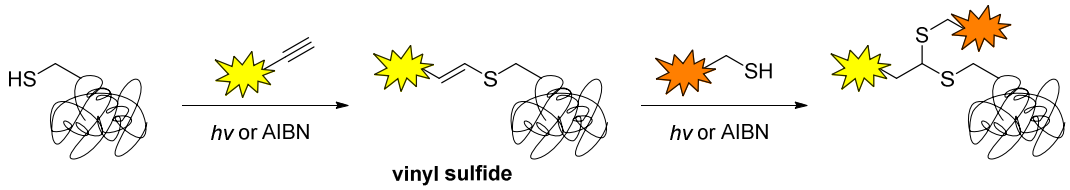
1.6. Dual-Functionalisation on Proteins

1.6.1. Introduction of Multiple Functionalisations on a Single Protein Site

In a very recent further development, the group of van Delft could apply their method in a selective dual-functionalisation of a protein (Scheme 1.24).[163] To do so, they used an *N*-terminal aldehyde created by periodate cleavage from serine and *N*-propargylhydroxylamine for the formation of the nitrone, which was reacted with a bicyclo[6.1.0]nonyne biotin. After SPANC, the terminal alkyne that was introduced with the hydroxylamine could be functionalised with any azide in a CuAAC. By doing so, they successfully introduced two distinct functional moieties to a protein's *N*-terminus.

Another strategy for dual-functionalisation on a single site in proteins is the thiol-yne reaction.[102] The work is based on a photoinduced thiol-yne reaction, in which the formed vinyl sulfide intermediate is trapped and reacted with a second but different thiol in a thiol-ene reaction to produce selective dual-functionalisation (Scheme 1.25). In doing so, the group of

Davis effectively introduced three carbohydrate moieties and three fluorophores consecutively to a protein in a quantitative fashion as judged by MALDI-MS.



Scheme 1.25: Dual-modification by thiol-yne chemistry.

Several groups have exploited the character of chemoselective ligation strategies to another extent by developing molecular handles that contain three different bio-orthogonal moieties, which can be applied in selective single-site dual-modification experiments. The group of Renard first published in 2010 a combination of thiol alkylation with iodoalkyl derivatives, oxime ligation and CuAAC[164] whereas Jones *et al.* combined ligation strategies, such as, SPAAC, thiol-Michael reaction and CuAAC, to successfully dual-functionalise biochips and proteins, respectively (Figure 1.4).[165]

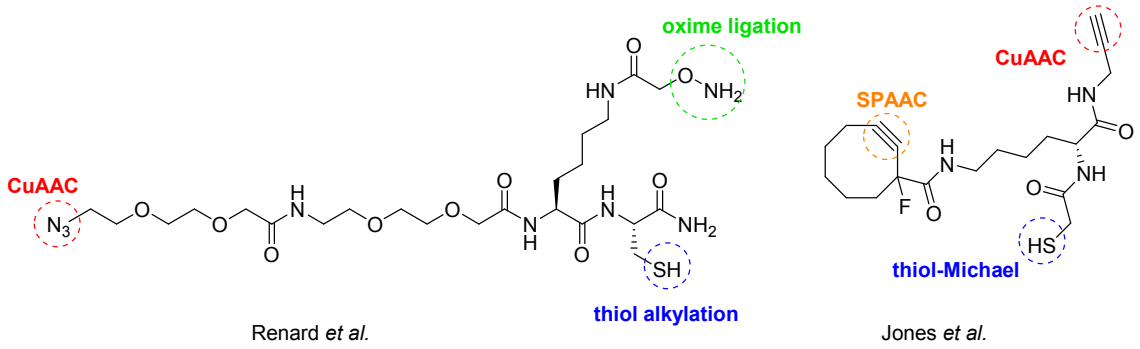


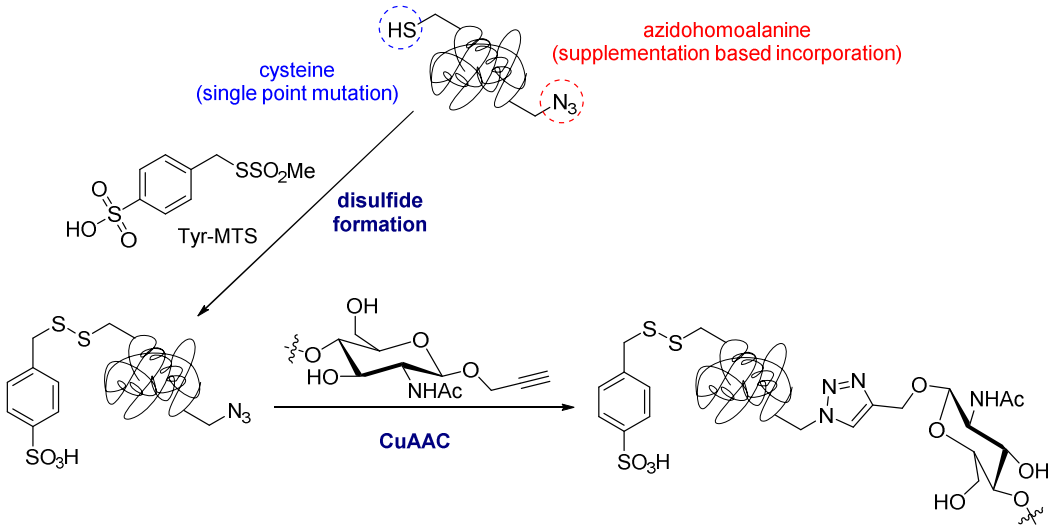
Figure 1.4: Bio-orthogonal handles combining three modifications strategies each.

1.6.2. Dual-Functionalisation of a Protein at Different Sites

Dual-functionalisation on a single protein site is definitely an advancement over mono-functionalisation, since it allows more versatility in protein modification. However, introducing two rather bulky moieties next to each other might not only hinder successful protein modification, but the close proximity could also have an influence on later applications, e.g., binding assays of carbohydrates or biotin and fluorescence measurements. Keeping this aspect in mind, new strategies that allow dual- or even triple-functionalisation at different sites of a protein would highly improve the field of protein modification.

During the last decade, several approaches have been published that combine different nonsense codon suppression techniques, such as, quadruplet codons and amber suppression,

summarised in a recent review by Budisa *et al*.[166] A major drawback of these strategies is the competition between the nonsense suppressor tRNA and present release factors, which are meant to stop the polymerase chain reaction cutting off the protein synthesis at the nonsense codon. Latest results by Sakamoto and coworkers to circumvent this problem by knocking out one of the major release factors seem promising.[167] However until this new method is not established further, the partial termination of translation can lead to much lower expression yields during protein synthesis. Especially for approaches, such as the combination of two suppression techniques (amber and quadruplet codon suppression) by Chin *et. al.*,[168] this factor plays an important role, which might render protein expression in some cases inefficient giving very little material to work with for further applications.



Scheme 1.26: Protein dual-modification at two distinct sites developed by Davis *et al*.

In 2007, the group of Davis published the first dual-functionalisation strategy that allows differential protein modification on two distinct sites combining canonical and non-canonical amino acids – namely cysteine and azidohomoalanine (Scheme 1.26).[169] With the overall help of site-directed gene mutagenesis, the latter amino acid was thereby incorporated by a supplementation based concept as methionine surrogate. Subsequent modifications with glycoconjugates as well as sulfo-tyrosine in a copper-catalyzed azide–alkyne cycloaddition and a disulfide bond formation, respectively, yielded a variety of modified protein scaffolds. Davis *et al.* could not only show that distinct combination of bio-orthogonal modification strategies can be applied for multiple and differential protein functionalisation, but they could further demonstrate the successful applicability of protein mimics in binding studies with human P-selectin.

In 2012, this concept was adopted by the group of Plückthun using the same amino acids for modification but varying the bio-orthogonal ligation strategies to SPAAC for azide functionalisation and thiol-Michael reaction to address the cysteine.[170]

1.7. Proteins as Scaffolds for Lectin Binding Studies

As the results of Davis *et al.* have shown, proteins have also gained importance as sheer scaffolds, which can be specifically functionalised and applied to study a variety of binding events. Especially in the field of glycoproteins, differently engineered protein scaffolds might give new insights into the mode of action of carbohydrate–carbohydrate and protein–carbohydrate interactions, e.g., lectin binding events.[143, 171] Lectins are complex proteins that can interact with different carbohydrate structures to trigger a variety of distinct functions in different parts of the biological system influencing, e.g., the immune system, communication and interaction as well as agglutination of cells, cell division and finally protein translation and folding.[172]

The binding mode of lectin proteins to carbohydrate complexes has not been fully elucidated yet, but it is assumed to be partially driven by the effect of multivalency (Figure 1.5).[173] Initial studies have shown that single lectin–glycan interactions seem rather weak (monovalent binding). If several glycans are connected on a cell membrane and a lectin protein bears multiple binding sites, a single binding event brings further binding partners into close proximity and drastically enhances the propensity for additional binding. Due to this effect – similar to chelating ligands – binding is more effective and the overall energy that is released is much higher than the actual sum of all binding events. This effect is called multivalency and can be found in a vast array of natural binding events, such as, the interaction of viruses and cells, the transport and uptake of vesicles and regulation of gene expression.

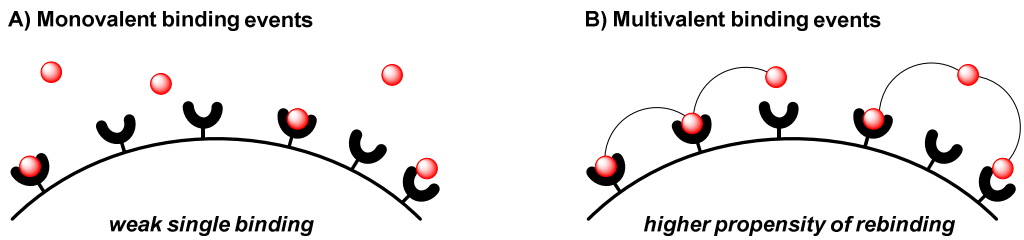


Figure 1.5: Comparison of monovalent and multivalent binding.

To study these important events in more detail, it would be desirable to have protein scaffolds that can be modified selectively with different functionalities. To do so, new ligation techniques and the meticulous combination of (bio)orthogonal ligation strategies for multiple protein

functionalisation could additionally help to understand these different binding events and elucidate possible multivalent interactions.

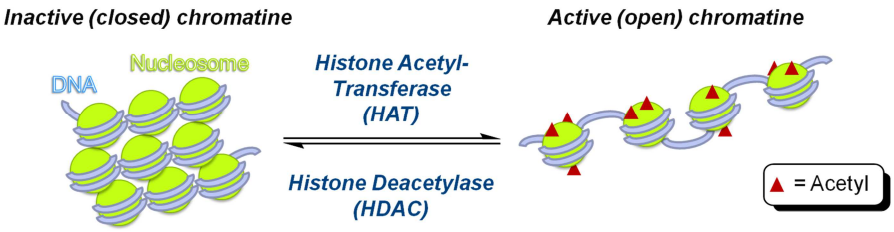
2. Objective

As already presented in the introduction part, the various structures and functions of peptides and proteins are not solely controlled by their amino acid composition, but also by the post-translational introduction of a multitude of additional functionalisations such as glycosylation, acetylation, fatty acylation, phosphorylation and sulfation. To study the influence of these functional moieties, this thesis focuses on the development and combination of chemoselective reaction strategies based on the selective incorporation of a variety of functionalisations onto a peptide or protein in aqueous media under ambient reaction conditions as well as on the development of new non-canonical amino acids for subsequent protein expression and functionalisation.

In that sense, the present thesis is comprised of three particular approaches.

2.1. Selective Peptide Acetylation in Aqueous Media

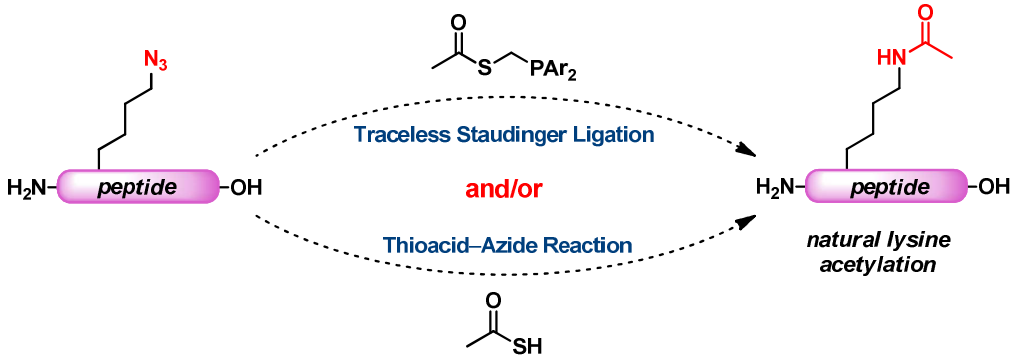
In the first part, we planned to develop selective acetylation strategies for peptides in aqueous media. Besides peptide elongation, selective amide bond forming reactions such as the traceless Staudinger ligation could also be applied in the important field of peptide and protein modification: selective amino acylations, such as fatty acylation, ubiquitination and acetylation.[174] The latter – protein acetylations – play a pivotal role in the newly emerging field of epigenetics, which describe the control of DNA transcription (Scheme 2.1).[174c, 175]. To be able to better understand and evaluate these biological processes on our nucleosome, scientists are looking for new tools to selectively introduce these modifications at certain positions of histone proteins applying a variety of enzymatic,[176] biotechnological[134b, 177] and bio-orthogonal strategies.[178]



Scheme 2.1: Protein acetylation in the field of epigenetics.

Acetyllysine residues can, for example, be introduced to proteins by amber suppression as pyrrolysine analogues or by peptide synthesis and subsequent peptide and protein ligation strategies. To further investigate their mode of action and especially the interplay of not only

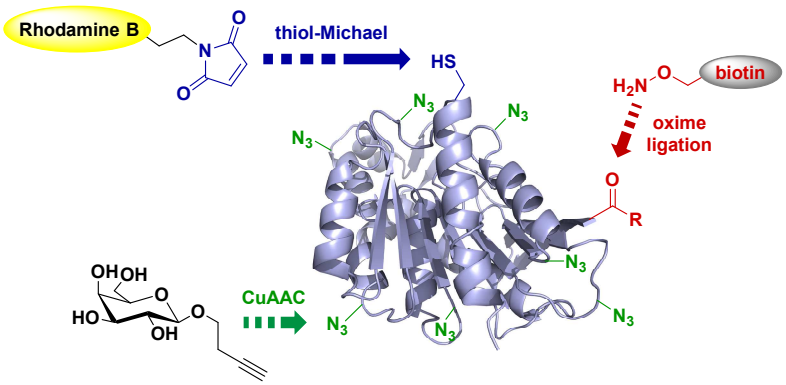
one but several of these modifications throughout a protein, a chemoselective modification strategy based on azido lysines would allow selective lysine acetylations at distinct positions as azido lysines can be introduced as methionine surrogates by supplementation based incorporation at multiple positions in a protein. To do so, we planned to synthesise a variety of *ε*-azido lysine peptides to react them either with phosphines *via* traceless Staudinger ligation or with thioacids in a thioacid–azide reaction and probe their chemoselectivity in the presence of basic side chains (Scheme 2.2). In the particular case of the traceless Staudinger ligation, we envisioned to improve the water-solubility of the employed phosphine to achieve more efficient selective peptide acetylations in aqueous media. Regarding the thioacid–azide reaction, we planned to address the applicability of more electron-rich alkyl azides such as azido lysine and azido glycine peptides avoiding known metal catalysts.



Scheme 2.2: Selective peptide acetylation of *ε*-azido lysine peptides by traceless Staudinger ligation or by thioacid–azide reaction.

2.2. Chemoselective Protein Modifications

In a second project, we envisioned to combine different known chemoselective ligation strategies to be able to introduce not only one but multiple different natural and unnatural functional moieties onto a protein, e.g. carbohydrates, fluorophores, FRET pairs and biotin, to achieve more versatility in biological studies (Scheme 2.3). That way, we planned to apply a thermophilic lipase as a protein scaffold to study carbohydrate–lectin binding events with the help of an additional biotin moiety for on-chip immobilisation. To do so, we had to combine three orthogonal modification techniques. We envisioned to apply the copper-catalysed azide–alkyne cycloaddition for the introduction of a variety of galactose moieties to azidohomoalanine residues. For further modification, we chose the oxime formation for the attachment of a biotin to a previously incorporated ketone or aldehyde and finally, as a possible third functionalisation strategy, we planned to use the thiol-Michael reaction for selective introduction of a fluorophore to a single cysteine residue.

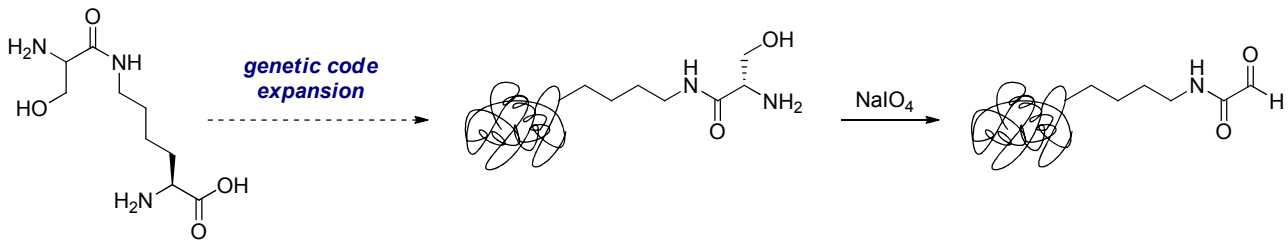


Scheme 2.3: Project outline for selective triple-modification of a protein.

2.3. Synthesis of Pyrrolysine Derivatives for Protein Synthesis Applying Amber Suppression Techniques

As a last objective of this work, to further enlarge the toolbox of unnatural amino acid incorporation to proteins using the amber suppression strategy, we envisioned the synthesis of different pyrrolysine derivatives. Initially, we planned to synthesise an already published azide derivative, which was first employed by the RINA GmbH to improve their PylRS/tRNA synthetase pair for the incorporation of pyrrolysine derivatives.

In addition, we envisioned the synthesis and incorporation of a new pyrrolysine derivative that would allow the post-translational formation of aldehydes on any desirable position of a protein for selective oxime ligation. The oxime ligation is a very reliable and clean modification technique, but the co-translational introduction of an aldehyde could not be shown, yet, to our best knowledge as aldehydes are too reactive. Few strategies have been developed so far that allow the post-translational formation of aldehydes but they are mostly limited to a protein's *N*-terminus. We therefore planned the synthesis and incorporation of a pyrrolysine derivative with a serine in its side chain, which could be used to post-translationally and selectively generate an aldehyde moiety by oxidative periodate cleavage (Scheme 2.4).



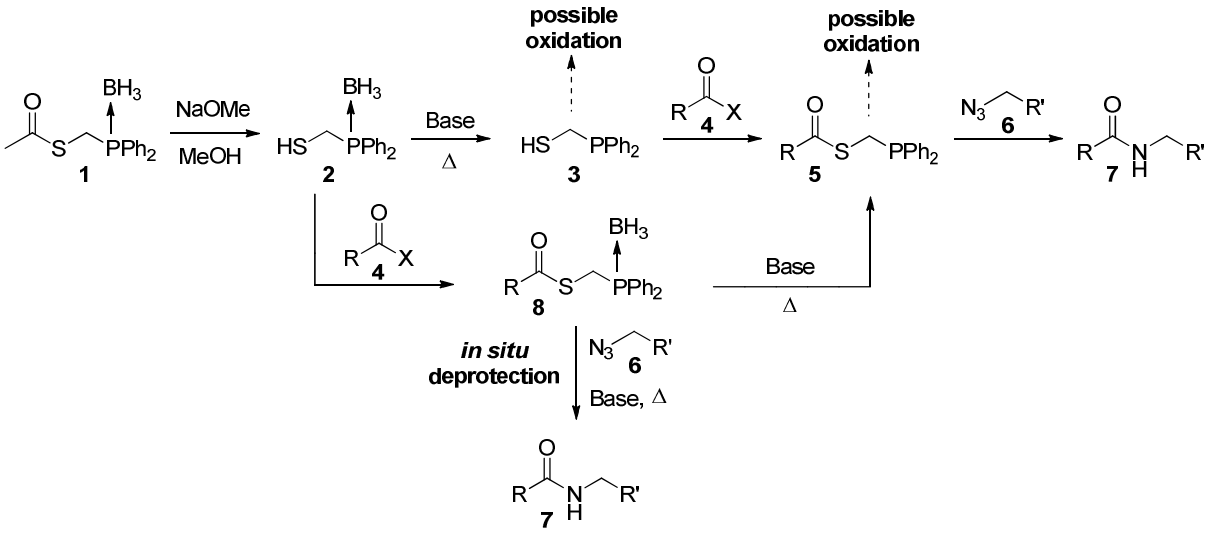
Scheme 2.4: Pyrrolysine derivative for potential posttranslational aldehyde formation in a protein.

3. Results & Discussion

3.1. Traceless Staudinger Ligation for Selective Amide Bond Formations on Peptides

3.1.1. General Overview

During the last decades, amide bond forming strategies have become important tools in the field of peptide and protein synthesis.[12] Beside the many strategies known to date, there is still a high demand for new approaches that allow higher versatility and a possible combination with already developed techniques. One such ligation method is the traceless Staudinger ligation (TSL) between a phosphine and an azide, which has been exploited in many directions during the last decade and has been applied so far in, e.g., peptide immobilisations on surfaces, coupling reactions of short protected (glyco)peptide fragments and cyclisations of unprotected peptides.[179] Diphenylphosphinemethanethiol (**3**) developed by Raines *et al*. has thereby proven to be the most versatile and efficient phosphine linker in amide bond formations by the traceless Staudinger ligation (Scheme 3.1), but the phosphine's poor water-solubility and its high propensity for oxidation limits its application for efficient peptide couplings with regard to longer coupling times and aqueous solvent systems.[66a]

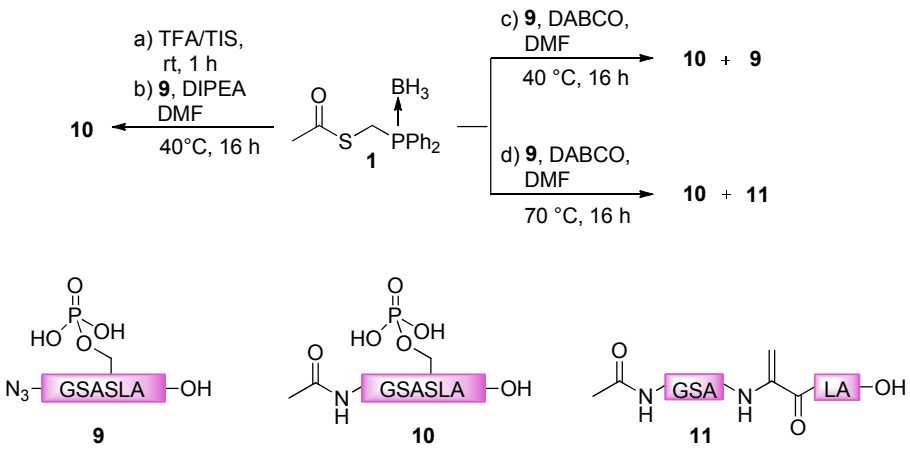


Scheme 3.1: Practicability of borane-protecting group for thioester synthesis with diphenylmethanethiol (3) and basic borane-deprotection for subsequent TSL reaction starting from commercially available borane-protected diphenylphosphinomethanethiol acetate (1).

In terms of applicability in the presence of oxygen, a borane-protecting group has been introduced by the group of van Maarseveen to allow efficient thioester bond formation without

undesired phosphine oxidation.[180] Alongside, a basic borane-deprotection strategy was published to regain reactive free phosphine for subsequent Staudinger ligations, which could also be applied in an *in situ* protocol (Scheme 3.1). In 2008, the group of Hackenberger published a complementary acidic deprotection strategy that would allow simultaneous side chain deprotection of peptide phosphino thioester fragments for chemoselective investigations of the traceless Staudinger ligation in peptide cyclisation reactions in DMF.[68]

To show propensities and limitations of the basic and the acidic deprotection strategies, we investigated their applicability in different ligation reactions including small organic molecules, carbohydrates and peptides, taking in account several factors, such as solvents, reaction temperature and functional group compatibility (see chapter 3.1.2). Based on phosphine linker **3**, we could show that the basic approach can be applied in a highly efficient one-pot strategy for traceless Staudinger ligations in organic solvents, such as toluene and DMF. However, deprotection with bases, e.g., DABCO, requires elevated temperatures which might interfere with the reactions applicability in biological systems as exemplified on azido phosphoserine peptide **9** (Scheme 3.2). Traceless Staudinger acetylation with borane-protected diphenylphosphinomethanethiol acetate **1** did not only yield the desired acetamide phosphoserine peptide **10** but also dehydroalanine variant **11** upon dephosphorylation (Scheme 3.2, path d). In addition, subsequent transformations of peptides in aqueous media would involve removal of poorly volatile organic solvents, such as DMF. In contrast, the acidic deprotection in neat TFA allows afterwards full removal by simple concentration under vacuum. Beside the fact that protected peptide phosphinothioesters can be fully unprotected during borane removal, which allows investigations on the reactions chemoselectivity, another benefit of the acidic approach is the initial formation of a protonated phosphonium salt, which cannot react with azides in a traceless Staudinger ligation unless deprotonation occurs, which can be selectively triggered by the addition of a base (see chapter 3.1.2). However, the addition of an excess of base, such as DIPEA, might lead to side-reactions with basic amino acid side chains, such as lysines or the *N*-terminus. This limitation under basic conditions might confine the reactions overall applicability in chemoselective transformations.



Scheme 3.2: Comparison study of acidic and basic borane deprotection on phosphorylated peptide 9 followed by TSL.

To address the rather poor solubility of phosphine linker **3** in aqueous media, Raines and co-workers as well as the group of Marx have synthesised a variety of new more water-soluble phosphine ligands based on thorough studies on coulombic effects and the influence of possible charges present in the phenyl substituents.[66] To our knowledge, these phosphines have not been used in any ligation strategy of unprotected peptide fragments. Therefore, we envisioned to synthesise water-soluble phosphine linker **12** that would allow us to study the propensity of the traceless Staudinger ligation on unprotected peptides in buffer aqueous systems (Figure 3.1, see chapter 3.1.3). First attempts towards the synthesis of a new water-soluble phosphine linker **12** were unsuccessful, which might be due to the strong electron-donating effect of our chosen phenyl substituent and the resulting high reactivity of the phosphorous.

As a consequence, we decided to work with the previously published water-soluble phosphine **13**,[66a] which was synthesised for us by the group of Prof. Pietrusiewicz at the university of Lublin (Figure 3.1, see chapter 3.1.4). They developed an improved synthetic route towards bis(4-[(*N,N*-dimethylamino)ethyl]-phenyl)(acetylthiomethyl)phosphine trisborane (**13**) (see chapter 3.1.5), which was first published by the group of Raines. With this phosphine in hand, initial studies on selective acetylation reactions of small molecule azides were carried out to directly compare hydrophobic phosphine **1** and more water-soluble phosphine **13** in different solvent mixtures of DMF and phosphate buffer (see chapter 3.1.4 and 3.1.5). To our delight, the initial results presented higher conversion rates for water-soluble phosphine **13** with an increasing water content. Best results were achieved in a 1:4 DMF/phosphate buffer system (0.1 M, pH 8) with 88% amide formation judged by LC-UV.

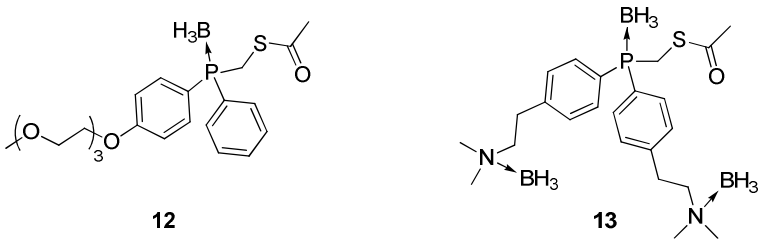
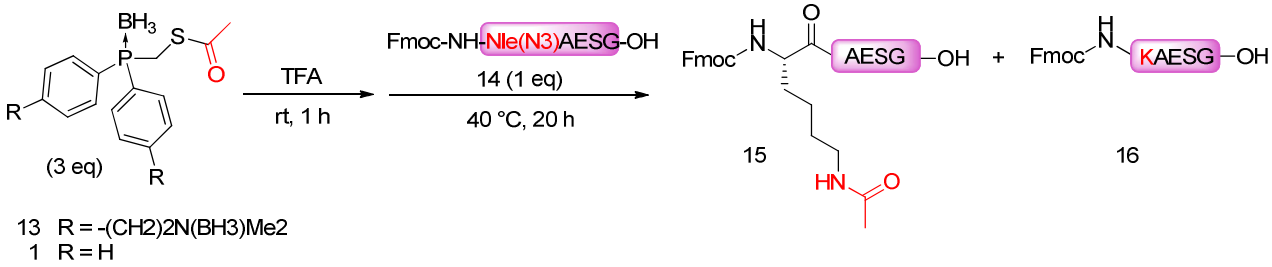


Figure 3.1: Water-soluble phosphine derivatives 12 and 13.

To go one step further, traceless Staudinger acetylations were then probed with both phosphines, **1** and **13**, on unprotected azido norleucine peptide **14** with an *N*-terminal Fmoc-group for UV-detection (Scheme 3.3, see chapter 3.1.5). Due to the Fmoc's hydrophobicity the ratio of DMF/phosphate buffer had to be increased to achieve better peptide solubility. Conversions were checked by LC-UV and the results seemed very promising. Solely, water-soluble phosphine variant **13** showed relevant product formation of amide **15**. Besides, we could observe formation of amine **16** due to partial iminophosphorane hydrolysis. However, this effect could possibly be a result of steric hindrance on the peptide conjugate itself due to the close proximity of bulky phosphine **13** and the *N*-terminal Fmoc-group. These results indicate the importance and necessity of more water-soluble phosphine derivatives for effective transformations in buffered systems (see chapter 3.1.5).



Scheme 3.3: Traceless Staudinger ligation on azido norleucine peptide 14 with phosphine 1 and 13.

Responsibility assignment for publications:

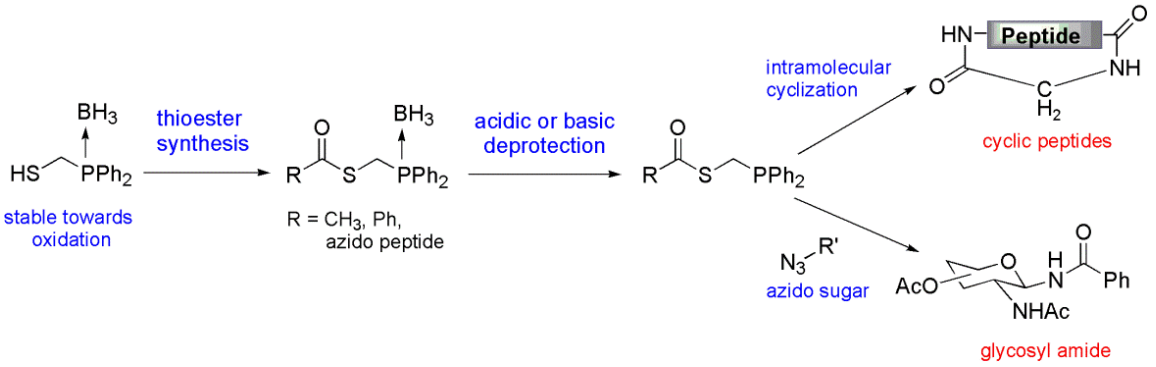
Chapter 3.1.2: The concept of research was provided by Professor C. P. R. Hackenberger. Decha Dechtrirat performed the stability studies of borane-protected phosphines. Optimisation of basic borane-deprotection conditions was executed by Ilona Papp. Sylvia Muth synthesised the borane-protected peptide thioester with fully unprotected side chains. Rolf Kleineweischede performed the peptide cyclisation studies. Da'san Jaradat was responsible for the synthesis of the azidosugar and subsequent traceless Staudinger ligations. The phosphoserine peptide **9** was synthesised by Malgorzata Broncel. Synthesis and traceless Staudinger ligation of azidoglycine, stability studies on unprotected phosphine **3**, and all reactions on the phosphorylated peptide **9**

were executed by the author. Furthermore, the publication was planned and written by the author under supervision of Professor C. P. R. Hackenberger.

Chapter 3.1.5: The concept of research was provided by Professor C. P. R. Hackenberger. The phosphine synthesis of **13** was developed and executed by Sylwia Sowa. Benzylazide synthesis and traceless Staudinger ligations were planned and supervised by the author and executed by Sylwia Sowa and the author. Azidonorleucine and the corresponding peptide were synthesised by the author. Furthermore, the publication was planned and written by the author under supervision of Professor C. P. R. Hackenberger.

3.1.2. Acidic and Basic Deprotection Strategies of Borane-Protected Phosphinothioesters for the Traceless Staudinger Ligation

Michaela Mühlberg, Da'san M. M. Jaradat, Rolf Kleineweischede, Ilona Papp, Decha Dechtrirat, Silvia Muth, Malgorzata Broncel, Christian P. R. Hackenberger



Scheme 3.4: General overview for applications of borane-protected phosphinothioesters in traceless Staudinger ligations.

Abstract:

The traceless Staudinger ligation has recently found various applications in the field of peptide synthesis and modification, including immobilisation and cyclisation strategies. In this report, we utilize the traceless Staudinger ligation in the formation of amide bonds, which allows the acquisition of acylated aminosugars and peptides as well as the cyclisation of peptides. A key element in these synthetic procedures is the use of a borane-protected phosphinomethanethiol, which is demonstrated to be prone towards oxidation in its unprotected form, during the synthesis of phosphinothioesters. In combination with acidic and basic deprotection strategies for the borane-protected phosphinothioesters, amide bonds can be formed in the presence of azides in moderate to good overall yields.

This chapter was published in the following journal:

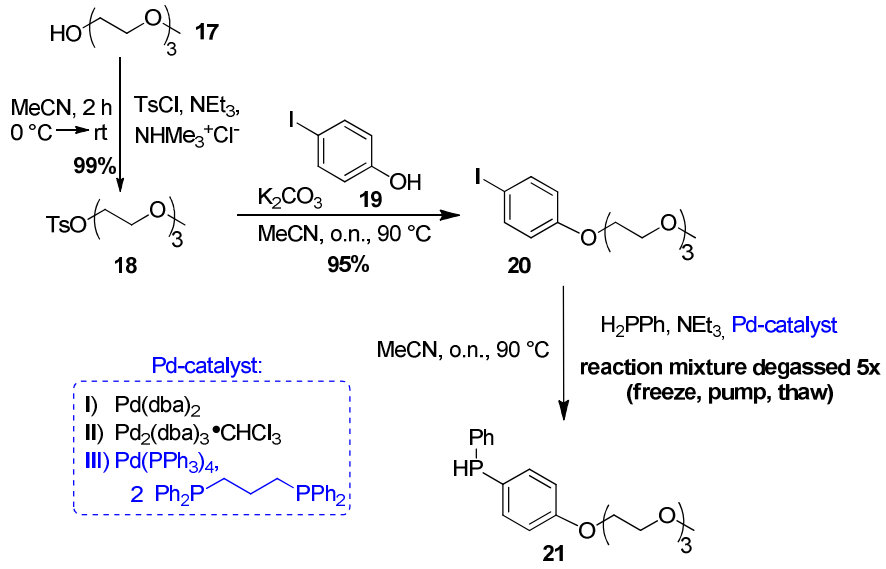
Michaela Mühlberg,[+] Da'san M. M. Jaradat,[+] Rolf Kleineweischede,[+] Ilona Papp, Decha Dechtrirat, Silvia Muth, Malgorzata Broncel, Christian P. R. Hackenberger
([+] These authors contributed equally to this work.)
Bioorganic & Medicinal Chemistry **2010**, *18*, 3679-3686.
Received: 18 January 2010, Accepted: 6 April 2010
First published on the web: 9 April 2010

The original article is available at:

http://dx.doi.org/10.1016/j.bmc.2010.04.015

3.1.3. Synthesis of a new Water-Soluble Phosphine

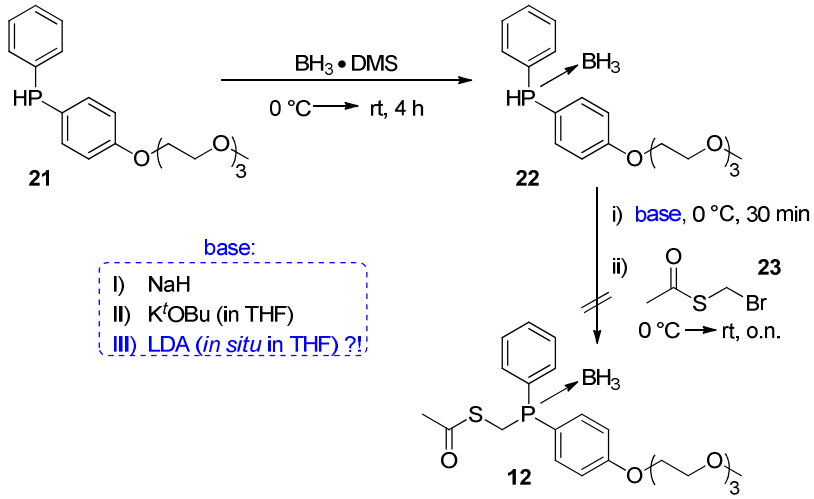
We envisioned to apply the traceless Staudinger ligation for amide bond forming reactions on peptides in aqueous media. Previous studies with rather hydrophobic phosphine **1** have shown to be inefficient due to poor solubility and competing iminophosphorane hydrolysis.[181] To achieve better water-solubility, triethylene glycol unit **17** should be introduced on one of the phenyl rings of phosphine linker **19** (Scheme 3.5). In addition, introduction in *para* position should enhance nucleophilicity of the phosphorous and could therefore circumvent iminophosphorane hydrolysis by increasing the velocity of the intramolecular rearrangement, the second step of the traceless Staudinger ligation.



Scheme 3.5: Synthesis of phosphine 21 by introduction of pegylated phenyl substituent 20.

For attachment of triethylene glycol monomethyl ether (**17**), the free hydroxyl group was tosylated and introduced to *p*-iodophenol (**19**) by nucleophilic substitution to form **20**. Coupling of aryl substituent **20** to phenylphosphine in a C–C-cross coupling reaction was probed with three different Pd-catalysts: I) $Pd(dba)_2$, II) $Pd(dba)_3{\cdot}CHCl_3$ and III) $Pd(PPh_3)_4$ with 1,3-bis(diphenylphosphino)propane (Scheme 3.5).[182] To avoid undesired phosphine oxidation, the reaction mixture was degassed in five freeze–pump–thaw cycles before refluxing it overnight at 90 °C. Conversions were checked by ^{31}P-NMR and presented the third catalyst as the best and quickest option to form disubstituted phosphine **21**. To further investigate the reaction, different solvents were probed, e.g., acetonitrile, tetrahydrofurane, dimethylformamide and dimethylsulfoxide, with DMF and MeCN showing best conversions. Due to the high instability of **21**, borane-protection was probed in situ to decrease possible phosphine oxidation (Scheme 3.6). Therefore, phosphine **21** was treated with a 1 M solution of $BH_3{\cdot}DMS$ or $BH_3{\cdot}THF$.

After 4 h, the reaction was again checked by ^{31}P-NMR with BH_3·DMS showing better conversion rates to borane-complex **22**.



Scheme 3.6: Borane protection and 2nd substitution on phosphine 21.

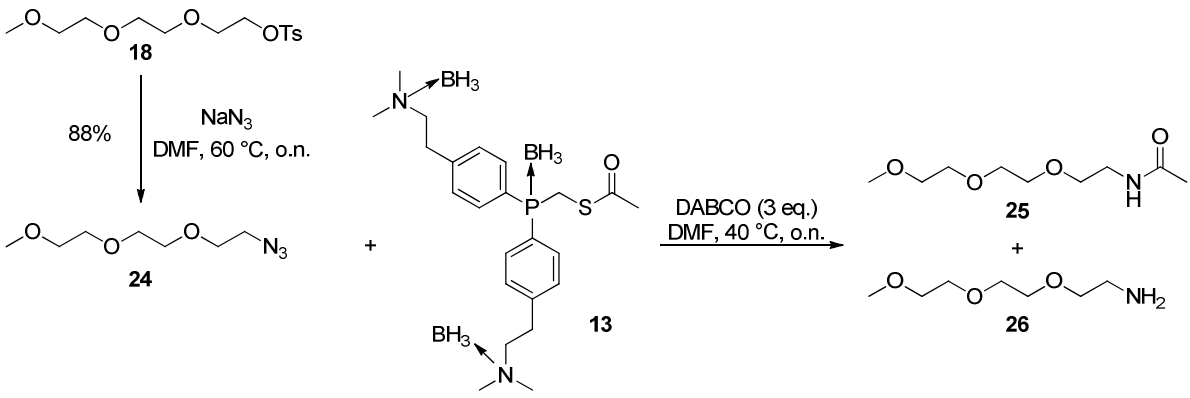
For the introduction of the third phosphine substituent S-(bromomethyl) ethanethioate (**23**), three different bases – sodium hydride, potassium *tert*-butoxide and lithium diisopropylamine – were probed for effective deprotonation of phosphine **22**. ^{31}P-NMR spectra showed a wide range of new signals mostly around 30 ppm. Compared to published NMR data by the group of Raines, borane protected phosphine **12** should give a signal around 17 ppm.[67] Higher ppm values seem to indicate the formation of secondary and tertiary phosphine oxides due to undesired oxidation despite previous borane-protection. One reason for this could be the higher nucleophilicity of phosphorous, which comes from the *para*-substituted triethylene glycol chain. Due to the high instability of phosphine **22**, the desired phosphine **12** could therefore not be successfully synthesised. As a consequence, we considered to switch to less electron-donating alkyl chains on the phenyl substituents bearing tertiary amines to enable higher water-solubility as previously described by Raines *et al*.[66a]

3.1.4. First Acetylation Studies with Different Phosphines

For further studies on the traceless Staudinger ligation in aqueous media, we planned to use phosphine **13**, which was synthesised for us by the group of Prof. Pietrusiewicz from the University of Lublin (Scheme 3.7). Phosphine **13** was initially published by Raines *et al.* and promised to be more reactive in aqueous media due to its higher water-solubility.[66a] First test reactions with a very simple water-soluble azide were planned to probe the reaction in buffered systems at different pH values. Therefore, 1-azido-2-(2-(2-methoxyethoxy)-ethoxy)ethane (**24**)

was synthesised from 2-(2-(2-methoxyethoxy)-ethoxy)ethyl 4-methyl-benzenesulfonate (**18**) in a nucleophilic substitution reaction with sodium azide (Scheme 3.7).

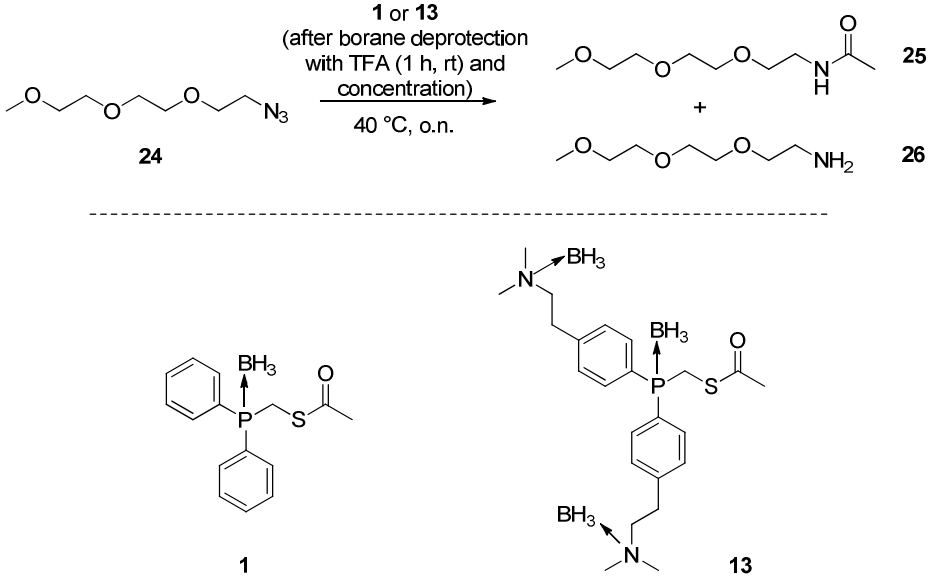
In a first attempt, the traceless Staudinger ligation was probed in dry DMF with 1 eq. of phosphine **13** and 1 eq. of azide **24**. DABCO was added for *in situ* borane deprotection and subsequent traceless Staudinger ligation in a one-pot approach (Scheme 3.7, Table 3.1, entry 1). After stirring at 40 °C overnight, LCMS measurements showed full conversion of azide **24**. The product, acetamide **25**, was detected in high quantities and amine **26**, which is a side product due to hydrolysis of intermediary formed iminophosphorane, could hardly be detected.



Scheme 3.7: Traceless Staudinger ligation with azide 24 and phosphine 13 in DMF and synthesis of azide 24.

For possible applications on peptides and proteins, the reaction should be probed in aqueous media. Two different solvent systems were employed – a DMF/phosphate buffer system (4:1; 0.7 M, pH 7) and pure phosphate buffer (0.7 M, pH 7) without solvent additive (Table 3.1, entry 2–3). In the first case, phosphine **13** was deprotected with DABCO in dry DMF for 4 h at 40 °C. Afterwards, azide **24** was predissolved in the buffer and added to the reaction mixture, which was stirred overnight at 40 °C. For the reaction in pure buffer, deprotection with DABCO seemed difficult as it would require full removal of DMF afterwards. Therefore, the deprotection strategy was changed to TFA (see also chapter 3.1.1). Phosphine **13** was treated with TFA for one hour, concentrated on high vacuum and redissolved in phosphate buffer. After addition of azide **24**, the pH was adjusted to 7.0 and the reaction was stirred overnight at 40°C. In both cases, conversion was again checked by LCMS. To evaluate the different reaction conditions, we decided to superficially judge the different reaction outcomes by comparing the total ion count signals (TIC) in the measured LCMS spectra of azide **24** (starting material), amine **26** (side product) and amide **25** (product). One can state that in both cases conversion of azide **24** seems almost complete, but compared to the reaction in dry DMF the formation of undesired

amine **26** increased with a decrease in DMF, which could be assigned to faster iminophosphorane hydrolysis in aqueous media.



Scheme 3.8: Traceless Staudinger ligation of 1-azido-2-(2-(2-methoxyethoxy)ethoxy)ethane (24) with phosphine 1 and 13, respectively, in different buffered solvent systems (for conditions see Table 3.3, entry 1–3).

Based on results published by the group of Raines,[66a] which indicate higher product formation as well as lower amine formation at a higher pH, the traceless Staudinger reaction with azide **24** was probed at pH 8 and 9 (Table 3.1, entry 4–9). In addition, we decided to compare the efficiency of water-soluble phosphine **13** towards the previously published, more commonly used and commercially available borane protected *S*-(diphenylphosphino)methyl ethanethioate (**1**) in buffered aqueous media (Scheme 3.8). Phosphine **1** is expected to produce lower yields than the variant **13** in aqueous media due to its poor water-solubility and lower reactivity (see above). The two phosphines were now applied under different reaction conditions in the conversion of 1-azido-2-(2-methoxyethoxy)ethoxy)ethane (**24**) (Table 3.1). In all cases, deprotection was achieved with TFA to allow better comparison of the results. For reactions in dry DMF, DIPEA was added to the reaction mixture to suppress residual TFA. Reactions in buffer were set up and the pH was adjusted immediately. After stirring overnight at room temperature – as stated in the latest publications by Raines *et al.* – conversions were checked by LCMS. In accordance to our expectations, conversion of azide **24** was approximately the same for all solvent systems. It can therefore be assumed that the reaction between phosphine and azide is rather fast and the rate limiting step is the rearrangement of the intermediary iminophosphorane towards the desired product **25**. The highest product formation was detected for the reaction in dry DMF (Table 3.1, entry 4–5). Both phosphines **13** and **1** gave the same results under these conditions. In contrast to previous results published by Raines

et al., formation of amine **26** increased with higher pH (Table 3.1, entry 8–9). Formation of amide **25** seemed to be favoured at pH 8 (Table 3.1, entry 6–7). Non-substituted *S*-(diphenylphosphino)methyl ethanethioate (**1**) thereby yielded higher amide formation compared to the more reactive phosphine **13**. These results indicate a stronger competition between hydrolysis and rearrangement of the iminophosphorane in the case of phosphine **13** leading to more amine formation.

Table 3.1: Traceless Staudinger ligation of 1-azido-2-(2-(2-methoxyethoxy)ethoxy)ethane (24) with phosphines 1 and 13 in different media at 40 °C (*ratios estimated by TIC (only for comparison of different conditions, no yields)).

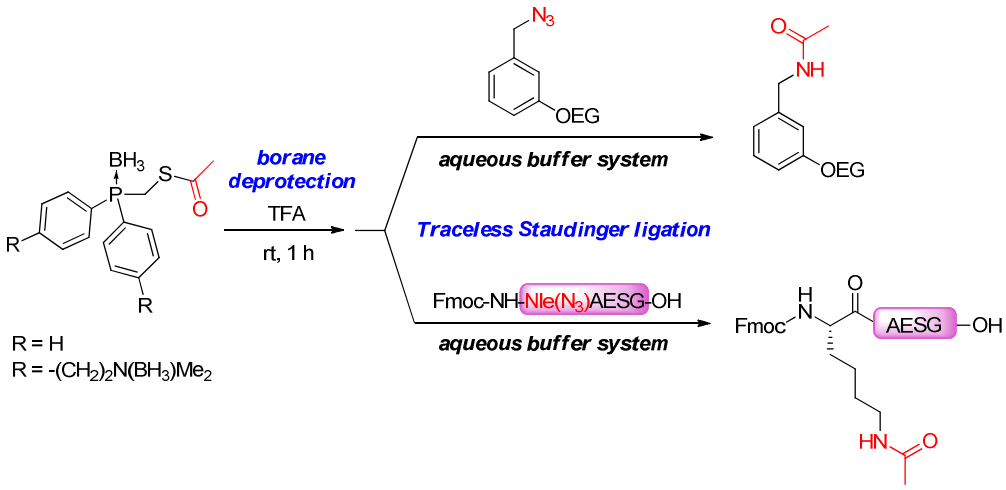
entry	phosphine	borane-deprotection	solvent system	azide 24*	amide 25*	amine 26*
1	**13**	DABCO	DMF (dry)	–	100%	–
2	**13**	DABCO	phosphate buffer (0.7 M, pH 7)/DMF (1:4)	5%	90%	5%
3	**13**	TFA	phosphate buffer (0.7 M, pH 7)	5%	48%	47%
4	**1**	TFA	DMF (dry), DIPEA (12 eq.)	10%	90%	–
5	**13**	TFA	DMF (dry) , DIPEA (12 eq.)	10%	90%	–
6	**1**	TFA	phosphate buffer (0.4 M, pH 8)	10%	70%	20%
7	**13**	TFA	phosphate buffer (0.4 M, pH 8)	5%	55%	40%
8	**1**	TFA	phosphate buffer (0.4 M, pH 9)	10%	45%	45%
9	**13**	TFA	phosphate buffer (0.4 M, pH 9)	10%	40%	50%

With these initial results in hand, the reaction should now be probed with new azido compounds including azido peptides in a variety of solvent mixtures of DMF and phosphate buffer (pH 8) as preliminary results have shown a considerable decline in iminophosphorane hydrolysis (Table 3.1, entry 2).

3.1.5. Traceless Staudinger Acetylation of Azides in Aqueous Buffers

Sylwia Sowa,[+] Michaela Mühlberg,[+] K. Michal Pietrusiewicz,

Christian P. R. Hackenberger



Scheme 3.9: Selective traceless Staudinger acetylations in aqueous buffer systems (OEG = oligoethylene glycol).

Abstract:

In this paper, we demonstrate the applicability of water-soluble *p*-dimethylaminoethyl substituted phosphinomethanethiol in acetyl transfer reactions by the traceless Staudinger Ligation with unprotected ε-azido lysine containing peptides in aqueous buffer systems. Additionally, we present an improved synthesis pathway for the water-soluble phosphinothiol linker requiring less reaction steps in a comparable overall yield in comparison to previously published protocols.

This chapter was published in the following journal:

Sylwia Sowa,[+] Michaela Mühlberg,[+] K. Michal Pietrusiewicz, Christian P. R. Hackenberger

([+] These authors contributed equally to this work.)

Bioorganic & Medicinal Chemistry **2013**, *21*, 3465-3472.

DOI: 10.1016/j.bmc.2013.02.045

First published on the web: 7 March 2013

The original article is available at:

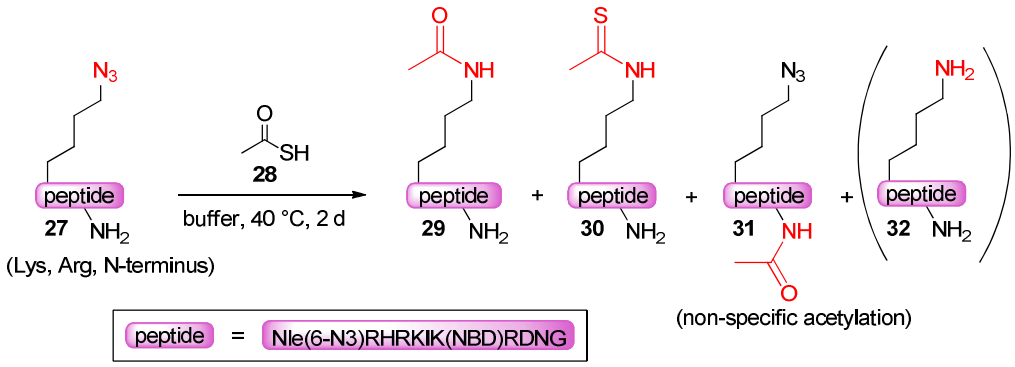
http://dx.doi.org/10.1016/j.bmc.2013.02.045

3.2. The Thioacid–Azide Reaction with Alkyl Azides

3.2.1. General Overview

In the previous chapter, we could show that amide bond forming reactions such as the traceless Staudinger ligation can be applied in selective lysine side chain acetylations in peptides. Another suitable reaction for this kind of approach could be the thioacid–azide reaction.[70a, 70b] The reaction is generally applied with electron-poor azides, e.g. sulfonyl- and glycosylazides, due to their higher reactivity towards thioacids.[70c] Nevertheless, successful employment of alkyl azides might allow the application of the thioacid–azide reaction for natural amide bond formations in peptides, e.g. selective acetylation or peptide bond formation.

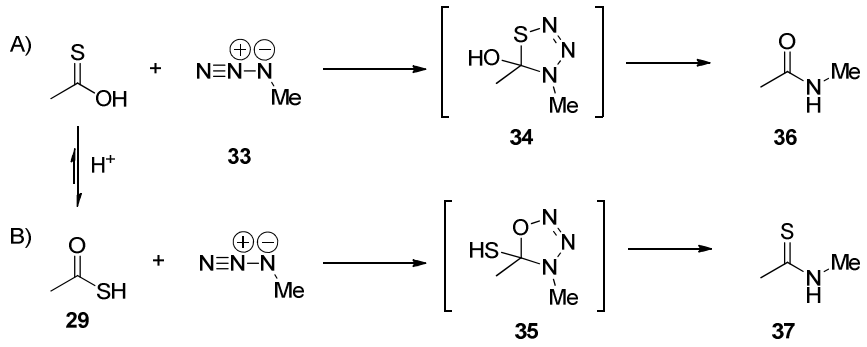
To do so, thioacetic acid (**28**) was reacted with azidonorleucine peptide **27** derived from histone protein H4[183] to yield acetyllysine peptide **29** under aqueous buffered conditions (Scheme 3.10, see chapter 3.2.2). Unfortunately, we could not observe any product formation under physiological or slightly basic conditions. Instead, the results showed that thioacetic acid (**28**) seems to react under these conditions preferentially with basic amines from amino acid side chains and the *N*-terminus to yield unselectively acetylated peptides **31**. As a consequence, the reaction was probed at lower pH values to suppress undesired reactions with lysine side chains. Thereby, we observed the formation of a new product, thioacetamide **30**, which increased with decreasing pH. These findings suggested a new mechanism for the thioacid–azide reaction at lower pH (Scheme 3.11).



Scheme 3.10: Reaction of ε-azido lysine peptide 27 with thioacetic acid (28) in buffered systems of different pH (NBD = *ε*-nitrobenzoxadiazole, Nle(6-N_3) = 6-azido norleucine).

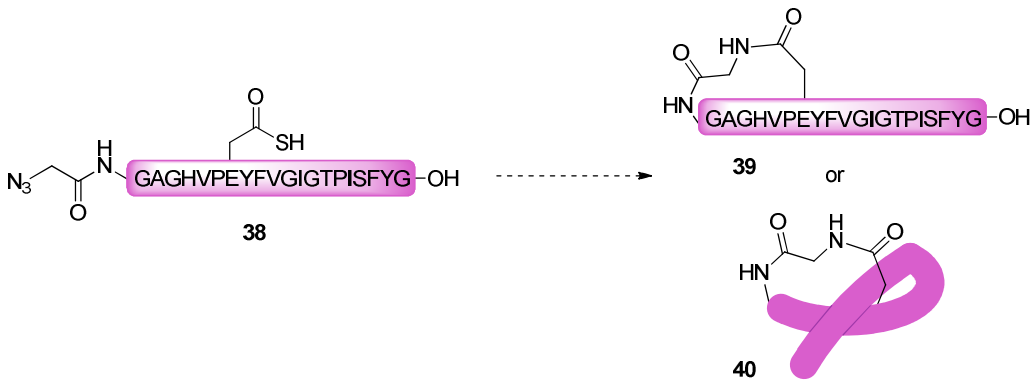
In addition, ^{1}H- and ^{13}C-NMR studies were performed on two azido peptides with an electron-rich and a modestly electron-poor azide, respectively (see chapter 3.2.2). Different solvent systems, pH values and concentrations were probed to better understand thioamide formation. Basic conditions could fully suppress thioamide formation, but as expected conversion rates for

electron-rich azides were very low. The results confirmed an increase of thioamide formation towards lower pH values (< 3). In addition, we could show that the conversion for the modestly electron-poor azido peptide could be easily increased at higher peptide concentrations, whereas the ratio of thioamide to amide formation could be controlled by pH with the highest amide formation at slightly acidic pH (ammonium acetate (100 mM, pH 4)/DMF, 7:3) and the highest thioamide formation at pH 2 (100 mM HCl/KCl buffer) with 88% amide and 92% thioamide, respectively, as judged by NMR (see chapter 3.2.2).



Scheme 3.11: Proposed mechanism for thioacid–azide reaction under acidic conditions.

It was also envisioned to use the reaction between thioacid and azide for peptide ligations (see chapter 3.2.3). Microcin *J25* an antibiotic peptide that inhibits nucleotide transfer during RNA synthesis of bacteria and thereby bacterial transcription, features in nature a *lariat protoknot* structure.[184] To yield this peptide with its natural structure in a synthetic approach, cyclisation between the *N*-terminus and the side chain glutamic acid should occur after the tail finds its natural fold, which might be supported in solution under more or less physiological conditions. Therefore, linear variant **38** of Microcin *J25* was synthesised and probed to cyclise *via* thioacid–azide reaction in solution (Scheme 3.12). Unfortunately, the chosen peptide model seemed problematic as selective thioacid formation on the glutamic acid side chain either proofed challenging with bulkier thiols yielding very low conversion rates or led to glutarimide formation under basic reaction conditions.



Scheme 3.12: Proposed cyclisation of Microcin *J25* by thioacid–azide reaction.

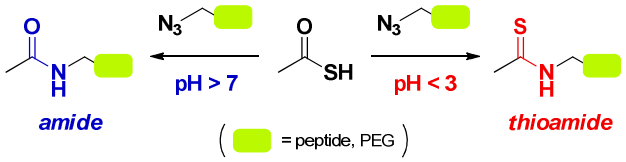
Responsibility assignment for publication (chapter 3.2.2):

The concept of research was provided by Professor C. P. R. Hackenberger. Brigitte Schlegel performed the NMR measurements. NMR spectra of *γ*-azido butanoic acid and azido glycine peptides were first analysed by Peter Schmieder. Further analysis to determine conversions and thioamide/amide ratios were performed by the author. Kristina Siebertz conducted final experiments with oligoethylene glycol azides under supervision of the author. Everything else was done by the author. Furthermore, the publication was planned and written by the author under supervision of Professor C. P. R. Hackenberger.

3.2.2. Controlled Thioamide *vs.* Amide Formation in the Thioacid–Azide Reaction under Acidic Aqueous Conditions

Michaela Mühlberg, Kristina D. Siebertz, Brigitte Schlegel, Peter Schmieder,

Christian P. R. Hackenberger



Scheme 3.13: Thioamide *vs.* amide formation in the thioacid–azide reaction on alkyl azido peptides.

Abstract:

The thioacid–azide reaction and its chemoselectivity were probed with alkyl azides for a potential application to form amide bonds in aqueous solvents. Our results reveal that under acidic conditions thioamides were formed as major reaction products suggesting a competing mechanism, whereas reactions to amides predominated at slightly higher pH values.

This chapter was published in the following journal:

Michaela Mühlberg, Kristina D. Siebertz, Brigitte Schlegel, Peter Schmieder, Christian P. R. Hackenberger

Chemical Communications **2014**, *50*, 4603-4606

Received: 28th January 2014, Accepted: 12th March 2014

First published on the web: 12th March 2014

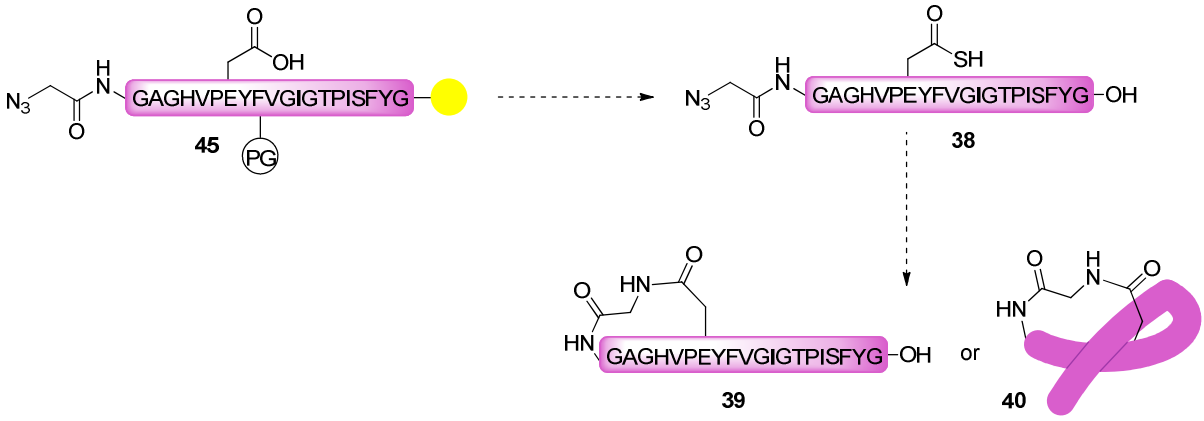
The original article is available at:

http://dx.doi.org/10.1039/C4CC00774C

3.2.3. The Thioacid–Azide Reaction – Towards Peptide Cyclisation

In a next step, we wanted to transfer the concept of selective amidation to the cyclisation of the bacterial peptide Microcin *J25*.[184] The bacterial antibiotic peptide is found in nature in a *lariat knot* structure where the *C*-terminal peptide tail goes through a loop formed by a head-to-side chain cyclisation (Figure 3.2A). The cycle is comprised out of eight amino acids connecting the *N*-terminus with a side chain glutamic acid. As the peptide is difficult to biologically express, there are few published attempts to synthesise Microcin *J25* by chemical methods either on the solid support or in solution by Native Chemical ligation upon introduction of an unnatural cysteine.[185] But to possibly yield the peptide's natural amino acid sequence and its typical lasso structure, we pictured that cyclisation should preferably proceed in solution without side chain protection to enable preformation of the peptide towards it three-dimensional structure. In addition, the rather slow amidation rates by thioacid–azide reaction could even support preformation of the natural lasso structure.

To selectively induce head-to-side chain cyclisation of our peptide by thioacid–azide reaction, we had to first convert the side chain of glutamic acid from an acid to a thioacid and introduce an azide to the peptide's *N*-terminus (Scheme 3.14). Upon cleavage and deprotection, peptide **38** should then be allowed to cyclise in solution in a chemoselective fashion.



Scheme 3.14: Planned synthesis of azido thioacid peptide 38 and subsequent head-to-side chain cyclisation by thioacid–azide reaction.

To do so, a linear precursor **41** of Microcin *J25* was synthesised by SPPS with N-terminal azido glycine and an orthogonally protecting group for the glutamic acid side chain (Figure 3.2B), which upon selective deprotection is supposed to be transformed into a thioacid before TFA deprotection and cleavage from the resin. We probed three commercially available protecting groups – allyl,[186] N-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl]-amino benzyl (Dmab)[187] and 1,1-dimethyl-1-phenylmethyl (^{i}PrPh)[188] – for the glutamic acid side chain of

peptide **41** and tested their deprotection on solid support (Figure 3.2, Scheme 3.15). For subsequent analysis, the peptide was fully deprotected and cleaved off the resin with trifluoroacetic acid to be checked qualitatively by MS.

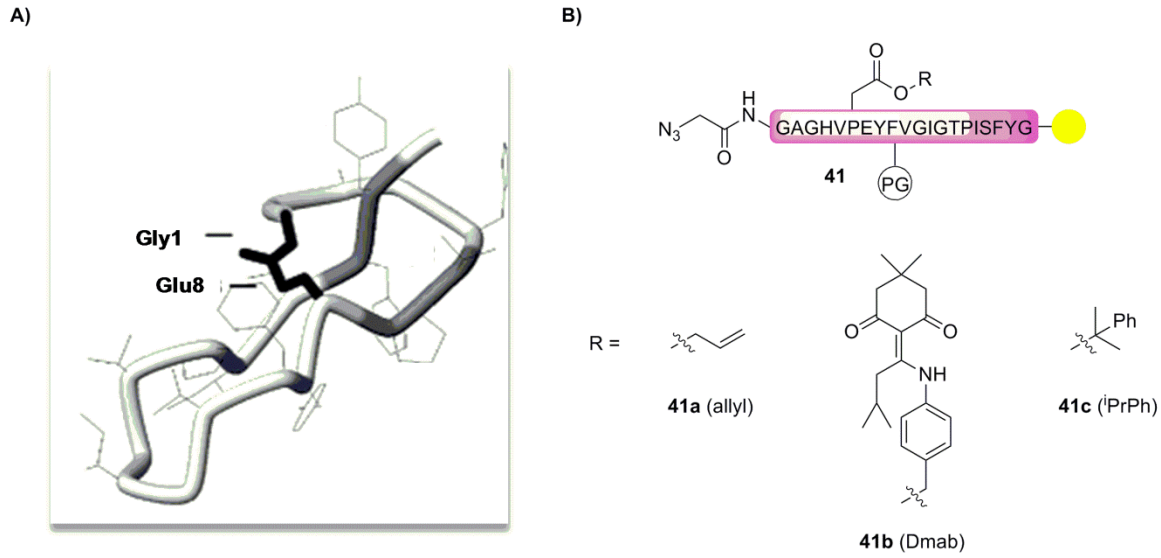


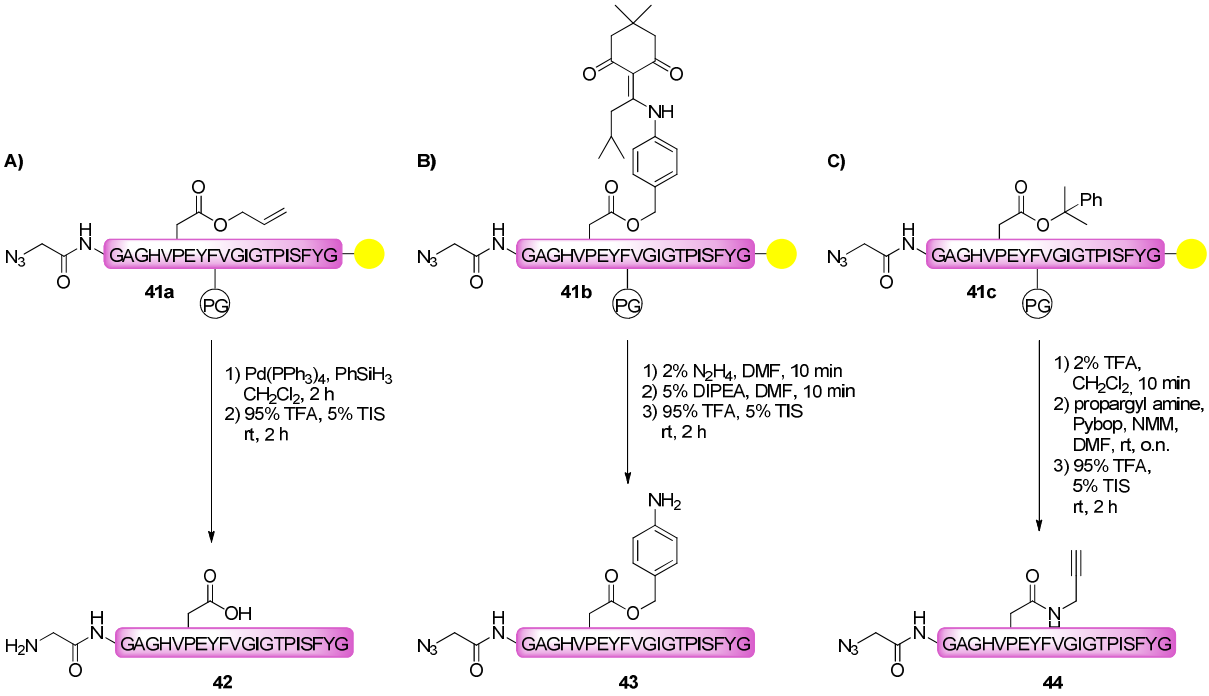
Figure 3.2: A) Lasso structure of Microcin *J25*, B) azido peptide 41 as a linear synthetic precursor of Microcin *J25* on resin with orthogonal glutamic acid side chain protecting groups.

First, the commonly used allyl protecting group was probed, but it turned out that common rather mild deprotection conditions ($Pd(PPh_3)_4/CHCl_3$/AcOH/NMM)[189] were not strong enough to achieve full removal and harsher conditions with $Pd(PPh_3)_4$ and phenylsilane led to an expected reduction of the azide towards peptide **42** (Scheme 3.15A).

The Dmab group gave better results, but although the general deprotection of Dmab with 2% hydrazine[190] to peptide **43** was successful, the subsequent base treatment[191] to cleave the intermediary *p*-aminobenzylester **43** for release of the free acid **46** was incomplete (Scheme 3.15B). Standard conditions of 5% DIPEA in DMF for 10 min showed almost no ester cleavage. Harsher conditions of 50 mM NaOH for three hours led to approximately 50% ester cleavage.

The last protecting group was *iso*-propylphenyl (^{i}PrPh), an acid labile ester, which should be cleaved with 2% TFA in dichloromethane in 10 min (Scheme 3.15C).[188] As we had to subsequently cleave the peptide from the resin with high amounts of TFA to check conversion by MS, it would have been impossible to judge on the efficiency of the preceding deprotection with only 2% TFA. As a consequence, we had to develop a new strategy to proof successful orthogonal deprotection of the glutamic acid side chain. After selective deprotection with 2%

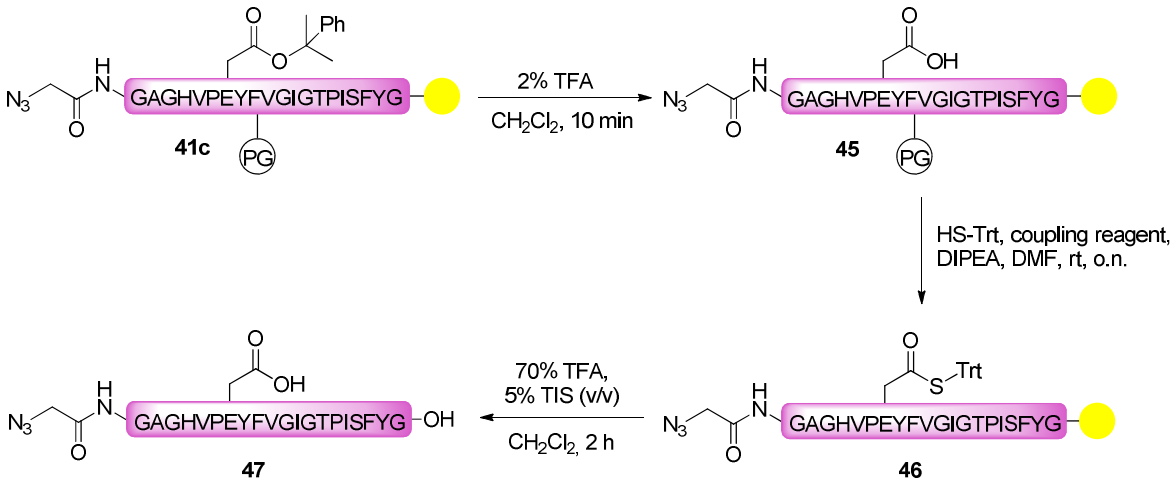
TFA, we therefore reacted the free glutamic acid side chain on resin with propargyl amine, using Pybop as a common amide coupling reagent and *N*-methylmorpholine (NMM) as a base in DMF overnight (Scheme 3.15C). Afterwards, the peptide was washed thoroughly and treated with 95% TFA to be fully deprotected and cleaved from the resin. MS measurements showed complete formation of amide **43** on the glutamic acid side chain, which proofs high efficiency of ^{i}PrPh as orthogonal glutamic acid side chain protecting group for our peptide model.



Scheme 3.15: Different orthogonal protecting groups and selected deprotection conditions for the glutamic acid side chain of peptide 41.

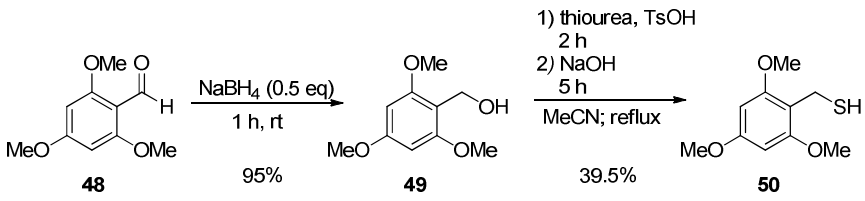
With these promising results in hand, we went on to probe selective thioacid peptide formation applying different reaction strategies known in literature (see also chapter 1.2.3.1). In general, there are two common ways to achieve selective thioacid formation usually performed on a peptide C-terminus – either by activation of the free acid and hydrothiolysis with a hydrogen sulfide, e.g., NaSH,[75] or by thioester formation with thiols and conversion to the desired thioacid by TFA treatment.[76] We decided against the first approach as it is known that longer treatment of azides with hydrogen sulfides can lead to azide reduction. Therefore, we turned towards the second strategy for thioacid peptide formation. A good leaving group upon TFA treatment to release sulfur compounds is trityl,[76b] which is therefore also a common acid-labile protecting group for cysteines. First reactions towards thioesters on peptide **45** with tritylthiol were probed under dry reaction conditions to avoid hydrolysis of trityl thioester **46** and applied different coupling reagents such as HATU, DIC, DCC and Pybop to determine which one is best

for thioester formations (Scheme 3.16). After TFA treatment and LCMS analysis, none of the performed reactions seemed to yield desired thioacid peptide **38** (see Scheme 3.14), but only free glutamic acid peptide **47** could be detected.



Scheme 3.16: First attempts to address selective thioacid formation from peptide 41c with tritylthiol (coupling reagent: Pybop, DCC, DIC or HATU).

One reason for this could be the thiol. As tritylthiol is rather bulky, we decided to switch to the smaller 2,4,6-trimethoxybenzyl thiol (**50**) (Tmob-SH, Scheme 3.17).[76a] It was reported by Vetter that its thioester is rather stable and can be easily converted to a thioacid during side chain deprotection and cleavage from the resin by treatment with 70% TFA (5% TIPS; 25% CH_2Cl_2) for 2 h.[76a] To do so, we first had to synthesise thiol **50** following a protocol from the literature.[192] The synthesis started with 2,4,6-trimethoxybenzaldehyde (**48**). Reduction with $NaBH_4$ yielded alcohol **49** that was then transformed into the corresponding thiol **50** by treatment with urea and *p*-toluenesulfonic acid and subsequent sodium hydroxide addition (Scheme 3.17).

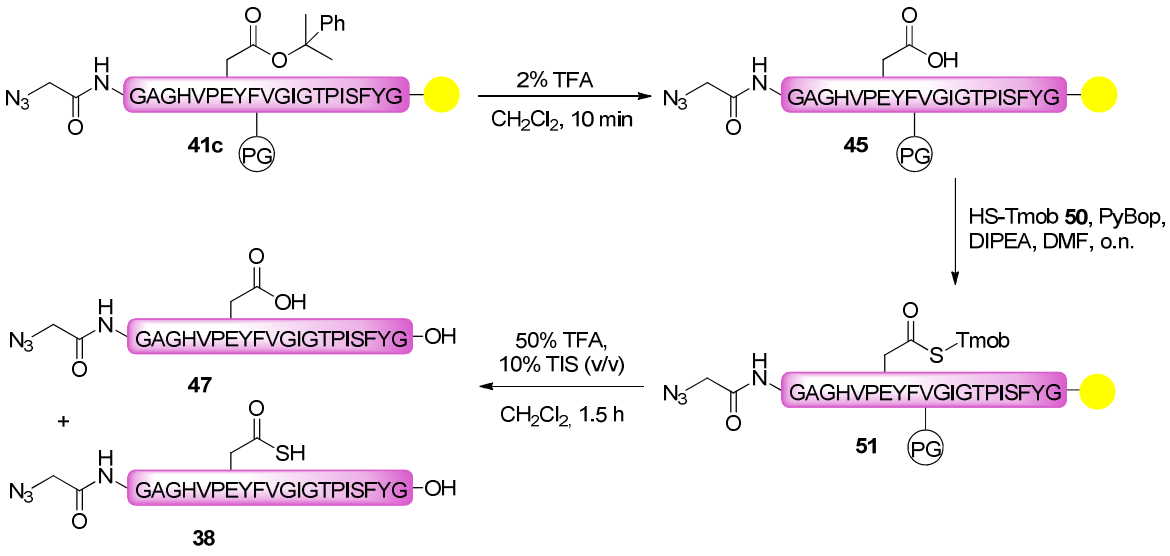


Scheme 3.17: Synthesis of 2,4,6-trimethoxybenzyl thiol (50).

With the new thiol in hand, we could now again focus on the thioacid peptide synthesis. After selective deprotection of peptide **41c** on resin under acidic conditions, we performed selective thioacid formation with thiol **50**, DIPEA as a base and different coupling reagents such as Pybop, HATU and DIC were again probed to yield thioester **51**. Subsequent TFA cleavage and deprotection with 70% TFA (as described above) yielded the desired thioacid peptide **38** and glutamic acid peptide **47** as well as TFA ester or amides of both peptides. Thereby, HATU and DIC

showed much lower conversion rates towards the desired thioacid peptide **38** than Pybop. We therefore decided to concentrate on Pybop as coupling reagent for subsequent thioester formations (Scheme 3.18). To solve the problem of the high formation of TFA ester peptides during the deprotection/cleavage step, we first probed the hydrolysis of the formed ester after the reaction was finished. Therefore, the reaction mixture was treated with different solutions overnight, such as 1 M NaOAc-solution, NaOMe (30% in MeOH) and 1 M NaOH. None of these led to even partial hydrolysis. Another possibility was to change the cleavage cocktail. Again, different conditions – a) HCl in dioxane, b) HCl in CH_2Cl_2, c) TFA in dioxane, d) TFA in CH_2Cl_2, 10% triisopropylsilane – were probed on the starting azido peptide **41c** on resin, but only method d) showed successful deprotection and in addition no TFA ester formation. A higher amount of TIS seemed to be essential for the prevention of TFA ester formation.

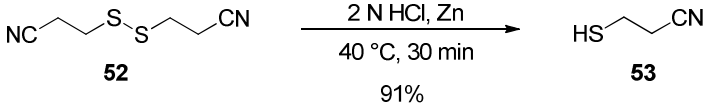
In a next approach, thioester formation on resin was probed with 50eq. of thiol **50**, Pybop and DIPEA in DMF overnight (Scheme 3.18). The amount of thioacid peptide **38** detected by LCMS was thereby as expected increased in comparison to the previously employed reaction with DIC, but the main peak was still the fully unprotected azido peptide **47** with the free glutamic acid side chain. Low thioacid formation might be rationalised to steric hindrance at the glutamic acid side chain as ester cleavage after deprotection of Dmab (Scheme 3.15) and thioester formation with more bulky tritylthiol (Scheme 3.16) also showed unexpected problems.



Scheme 3.18: Synthesis of thioacid peptide 38 *via* 2,4,6-trimethoxybenzyl (Tmob) thioester 51 with optimised conditions (PG = protecting groups).

Following these results, we decided to probe another even smaller thiol for thioester formation to achieve higher conversion of starting peptide **45** during on-resin thioester synthesis. The

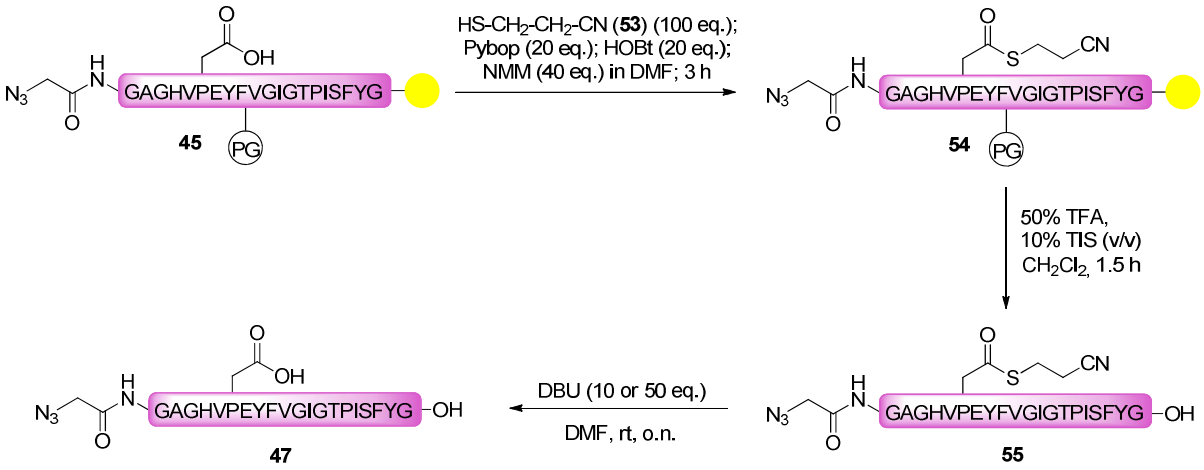
group of Rademann published in 2012 a protocol for a *C*-terminal peptide thioacid formation with 3-mercaptopropionitrile **53** that could be released upon β-elimination under basic conditions to yield the desired peptide thioacid.[79] 3-Mercaptopropionitrile **53** was synthesised *via* reduction of the corresponding commercially available disulfide **52** with zinc and hydrochloric acid (Scheme 3.19). Another feature of this thiol is that its cleavage proceeds under basic conditions, which enables actual detection of the formed thioester after TFA treatment.



Scheme 3.19: Synthesis of 3-mercaptopropionitrile 53.[79]

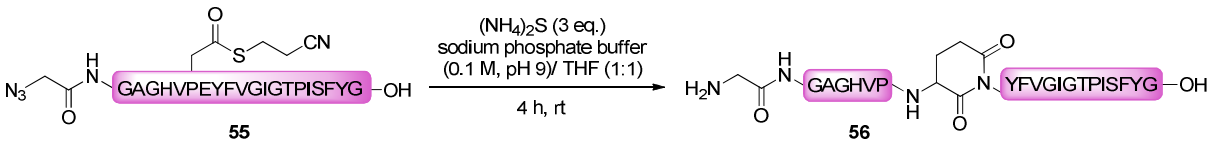
For peptide thioester formation, the orthogonally deprotected peptide **45** with the free glutamic acid side chain was reacted with thiol **53**, Pybop, HOBt and NMM in DMF (Scheme 3.20). After TFA treatment, peptide conversion was checked by MS indicating much better conversion towards thioester **54** than previous results with more bulky trityl- and 2,4,6-trimethoxybenzylthiol. With this result, we decided to directly go on and probe the subsequent formation of thioacid peptide **38** in solution without further optimisation (Scheme 3.20). Following the same protocol by the group of Rademann, the cleaved and fully deprotected thioester peptide **55** was concentrated under reduced vacuum and afterwards redissolved in dry DMF with 10 or 50 equivalents of 1,8-diazabicycloundec-7-ene (DBU) as a base.[79] The reaction was checked by MS and showed no thioacid formation but solely thioester hydrolysis back to the free glutamic acid side chain of peptide **47**. As a consequence, we decided to probe a different way of deprotection also published by Rademann *et al*.[79] To do so, the peptide thioester was redissolved in a mixture of phosphate buffer (100 mM, pH 9) and THF (1:1) due to its high hydrophobicity and after addition of ammonium sulfide the reaction mixture was shaken at room temperature for 45 minutes. We expected the azide to be at least partially reduced in the presence of ammonium sulfide, but we thought that by cautious control of the reaction time and the conditions we might be able to suppress azide reduction to a minimum. Surprisingly, LCMS measurements showed as a main product not thioacid peptide **38** but a peptide with the mass of the desired cyclisation product **39** indicating a possible undefined azide reduction and loss of water, which might correspond to a our intended peptide cyclisation by thioacid–azide reaction. In addition, we also observed as expected uncyclised peptide **42** with the azide being reduced to the amine (see also Scheme 3.20). Therefore, we probed different peptide (0.5–4 mM) and ammonium sulfide concentrations (1–3 eq.) to suppress formation of undesired side product **42** and increase the formation of our reduction–dehydration product to allow purification and

further analysis to find out if the desired cyclisation product **39** was formed or if any other side chains might have reacted instead. Best results could be achieved with a rather low peptide concentration of 0.5 mM and three equivalents of $(NH_4)_2S$. Lower amounts of ammonium sulfide seemed to reduce the formation of the apparent dehydration product and led to thioester hydrolysis to yield **47**.



Scheme 3.20: Synthesis of 2-cyanoethane thioester 55 and first attempts towards thioacid peptide 38.

With these results in hand, we could finally purify the reaction product by preparative HPLC to analyse it further by MS/MS. Surprisingly, analysis by MS/MS revealed that the isolated peptide is not the desired cyclisation product **39**, but uncyclised peptide **56** with a glutarimide formed within the peptide backbone and the azide being reduced to an amine (Scheme 3.21). It seems that basic conditions for thioacid formation induce glutarimide formation when the glutamic acid side chain is activated as a thioester. In this way, we can say that the thioacid formation strategy published by Rademann and co-workers is not applicable to our approach of an intramolecular peptide cyclisation by the thioacid–azide reaction due to the reduction of azides. In addition, this strategy should not be applied to thioacid formations on glutamic acid and probably also aspartic acid side chains due to possible glutarimide/aspartimide formation.



Scheme 3.21: Glutarimide formation from thioester peptide 55.

In contrast, the previously employed acidic approach might still be an option as successful formation of the thioacid peptide could be shown by MS and, in addition, glutarimide formation should not occur under acidic deprotection conditions. Although peptide thioester formation for

our model peptide was very low, we decided to probe cyclisation reactions *via* Tmob-thioester peptide **50**. To perform peptide cyclisation in solution (Scheme 3.18), the glutamic acid side chain of peptide **41c** was selectively deprotected on resin with 2% TFA and converted with Tmob-SH **49**, Pybop and DIPEA in DMF to the desired thioester peptide **50** (Scheme 3.18). After full cleavage/deprotection, the peptide mixture was concentrated under vacuum and re-dissolved in different solvent systems to probe peptide cyclisation towards peptide **39** (Table 2). The solutions were stirred overnight at room temperature and then checked by LCMS. The amidation reaction took 1–3 days until no thioacid peptide **38** could be detected anymore. The desired reduction–dehydration mass that might correspond to peptide cyclisation could only be detected for two of the probed solvent systems (Table 3.2, entry 4 and 7). However, to draw any further conclusions from our results we had to find a way to determine the conversion of the reaction.

Table 3.2: Peptide cyclisation by thioacid–azide for Microcin *J25* synthesis.

entry	solvent system, c(peptide) = 1 mM	T [°C]	yield[a]
1	H_2O (6 M Gnd-HCl, 3 mM lutidine	30	n.d
2	sodium phosphate buffer (0.1 M, pH 8)	30	n.d.
3	Tris-HCl buffer (50 mM, pH 8, 200 mM NaCl)	30	n.d.
4	sodium citrate buffer (0.1 M, pH 2.5)/DMF (1:1)	30	6.0
5	sodium citrate buffer (0.1 M, pH 2.5, 200 mM NaCl)/DMF (1:1)	30	n.d.
6	sodium citrate buffer (0.1 M, pH 2.5, 200 mM LiCl)/DMF (1:1)	30	n.d.
7	DMF, 3 mM lutidine	30	9.0

[a] Determined by MS (comparison with deuterated product **60**).

Unfortunately, it was not possible to determine the conversion of the reaction from the LC-UV chromatogram. Due to the inherent impurity of the peptide on resin after SPPS and the formation of several unassigned side products during the following reaction steps, there is a high peak overlap in the region of our desired peptide (Figure 3.3). Considering these facts together with the low conversion yield towards the reduction–dehydration product, conversion could not be determined properly and purification seemed also rather difficult. In comparison to the previously formed glutarimide peptide **56**, the now synthesised reduction–dehydration product has a slightly shifted retention time, which is even closer to the mostly formed free glutamic acid azido peptide **47**.

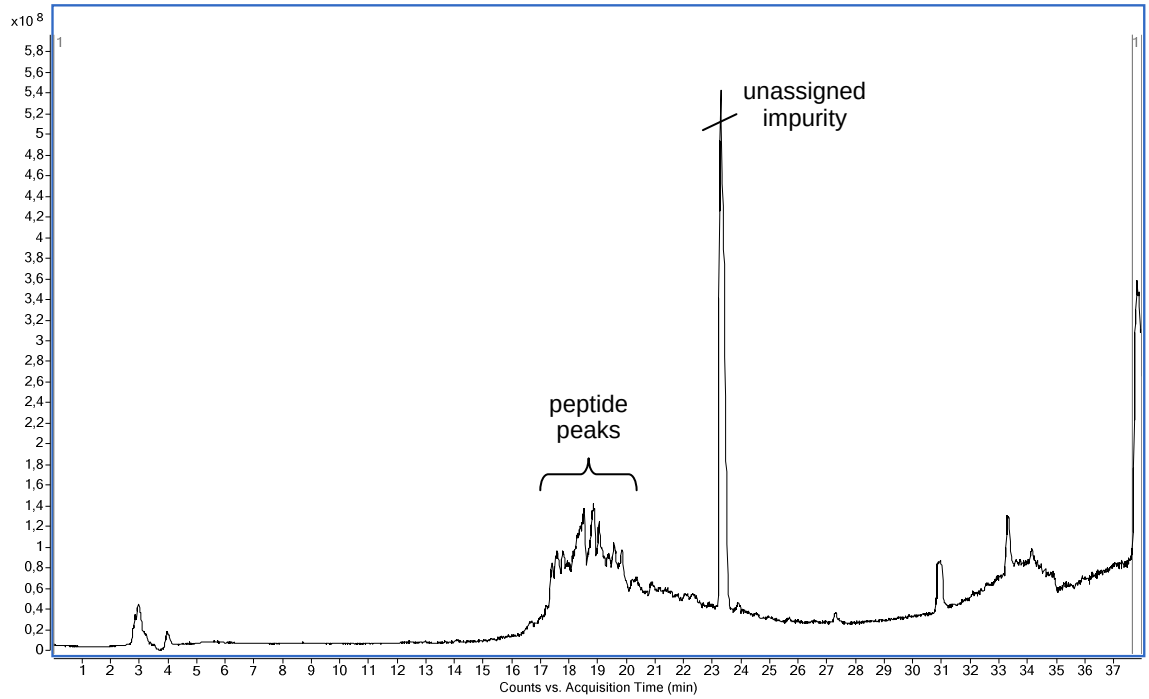
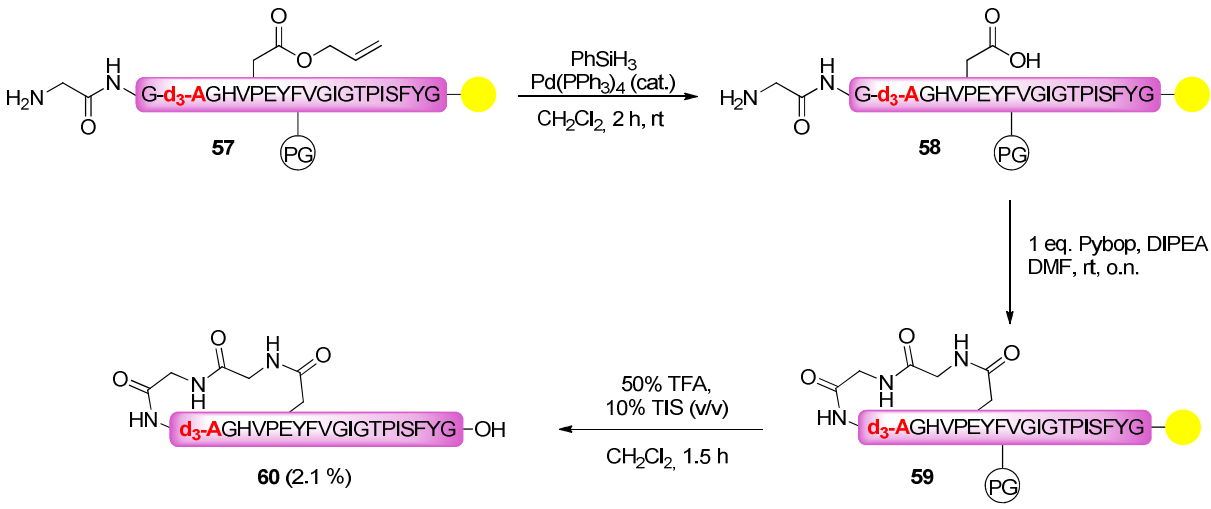


Figure 3.3: LC-UV chromatogram at 220 nm of probed peptide cyclisation (Table 3.2, entry 7).

To gain further information on the detected peptide including retention time and conversion, a standard for **39** was needed. Therefore, we synthesised deuterated variant **60** of cyclised peptide **39** by normal SPPS (Scheme 3.22).In this standard, alanine was substituted by d_3-alanine to further allow quantification of the formed product. It is commonly accepted that MS intensities of isotopic isomers are seen as equivalent. Both compounds are chemically identical and therefore have the same ionisation ability. By addition of a known amount of the deuterated peptide to our reaction mixture we can now quantify the formation of the nondeuterated peptide by comparison of the total ion count signals.[193]

To synthesise deuterated peptide **60**, the cyclisation reaction on resin was first probed with a non-deuterated peptide (Scheme 3.22). The peptide was synthesised with an allyl protecting group for orthogonal deprotection of the glutamic acid side chain to allow selective on resin cyclisation with the free *N*-terminal amine. The allyl protecting group was removed by treatment with phenylsilane and $Pd(PPh_3)_4$ in CH_2Cl_2 for two hours. Afterwards different coupling reagents, such as DIC, Pybop and HATU, were probed for final peptide cyclisation overnight. The two latter reagents gave very good results judged by LCMS. Afterwards deuterated peptide **60** was synthesised by Fmoc-SPPS, allyl group deprotection, on-resin peptide cyclisation and TFA deprotection and cleavage. Final purification by preparative HPLC yielded deuterated cyclic peptide **60** in overall 2.1% (Scheme 3.22).



Scheme 3.22: Synthesis of deuterated cyclic peptide 60.

To quantify conversion rates by LCMS, a known amount of deuterated product **60** was added to the reaction mixture before each LCMS measurement. The corresponding TIC (total ion count) signals for both peptides **39** and **59**, which eluted with the same retention time, were integrated and compared to calculate the amount of the formed dehydration product. Due to the absence of any arginine or lysine side chains in the peptide and therefore any undesired side reactions with basic amino acid side chains (see also chapter 3.2.2), DMF with 2,6-lutidine as a base (Table 3.2, entry 7)[70c] seemed to show the best conversion of 9% towards the reduction–dehydration product as judged by LCMS. Furthermore, as expected no thioamide formation could be detected due to the rather basic cyclisation conditions (see also chapter 3.2.2). Cyclisation in a 1:1 mix of sodium citrate buffer (0.1 M, pH 2.5) and DMF yielded overall 6% of the reduction–dehydration product (Table 3.2, entry 4). As already mentioned, the product could not be isolated by preparative HPLC due to its small amount and its overlap with the more dominant glutamic acid azido peptide **47** and several other unassigned peptide peaks. We could therefore not further analyse the apparent cyclisation product, yet. Further evaluation of its identity and structure is therefore still outstanding and necessary to claim successful cyclisation.

To further investigate the applicability of the thioacid–azide reaction for intramolecular peptide cyclisations, a new peptide target should be chosen that allows thioacid formation in higher amounts either on resin or in solution. In addition, one should also further investigate the thioacid–azide reaction in peptide fragment ligations between C-terminal peptide thioacids and N-terminal azido peptides as we could show that α-azido amino acid building blocks such as azido glycine show high conversion rates towards amides at slightly acidic pH values.

3.3. Proteins as Scaffolds for Carbohydrate-Binding Studies – Potential Dual- and Triple-Modification of Proteins

3.3.1. General Idea and First Protein Model TTL(D221Bpa,S261C)[Aha]

Proteins are not only highly important as molecular machineries that control a vast variety of cellular processes but, in addition, chemical biologists apply them as synthetically modified scaffolds to study the diversity of biological mechanisms in more detail. Therefore, scientists have developed only a small number of new (bio)chemical techniques, e.g., expansion of the genetic code by amber suppression or auxotroph supplementation, and chemical modifications of natural amino acids, to very selectively decorate a complex protein with different modifications and functionalities (see chapter 1.4). However, as research goals get more demanding, (bio)chemists started to search for techniques that not only allow the introduction of one but multiple different functionalisations in parallel. To do so, the careful combination of already established methods is necessary (see also chapter 1.6). To date, only a few scientists have succeeded in this task. Most strategies published involve the introduction of two different functionalities on a single modification site.[102, 163-164, 165] To selectively introduce two different functional moieties on two distinct sites of a protein, first attempts have been made, which employ the introduction of one or two non-canonical amino acids by nonsense and/or quadruplet codon suppression.[168, 194] Although recombinant expression strains have been engineered to improve incorporation efficiency,[167, 195] double labelling approaches suffer from low protein yields due to competition of NCAA incorporation with translational frame shifting or termination, so-called context effects, and low catalytic efficiency of engineered aminoacyl-tRNA synthetases.[166] To avoid genetic code expansion, the group of Davis published in 2007 the incorporation of an azidohomoalanine (Aha) residue by supplementation based incorporation (SPI) in combination with the modification of a single natural cysteine.[169] Although a variety of modification strategies for solvent-exposed cysteines are widely used nowadays,[196] there are also some drawbacks to the employment of cysteines such as residue-specificity in natural amino acid modifications,[58a] the cysteine's high tendency for disulfide bond formation and side-reactions with basic side chains, e.g., lysines.[197] Therefore, a demand has risen for the application of alternative more site-specific bio-orthogonal labelling strategies for new multiple-labelling approaches of proteins.

We envisioned the development of a selective triple-modification strategy for a single protein on three distinct sites *via* three orthogonal modification strategies: copper-catalysed azide–alkyne

cycloaddition (CuAAC), oxime ligation and thiol-Michael reaction. We planned to carry out the experiments with a model protein, TTL(D221Bpa,S261C)[Aha] (Figure 3.4), which is known to be readily tolerant to high temperature, solvents and other additives, and an enzymatic assay is available as a control for retained protein integrity and catalytic function.[198] The protein was expressed by the group of Budisa using, on one hand, auxotroph supplementation for unnatural protein translation. This incorporation strategy allows for the introduction of the methionine analogue azido homoalanine in several positions throughout the whole protein for a later introduction of galactose moieties by CuAAC. On the other hand, a single cysteine was incorporated by site-directed mutagenesis to introduce a fluorophore by thiol-Michael reaction for fluorescent quantification. Finally, as a last modification strategy, a benzophenone moiety (Bpa) was introduced by amber suppression to allow selective biotin labelling of the protein by oxime formation for later immobilisation on a streptavidin coated surface.

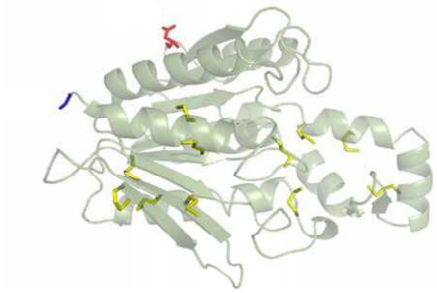
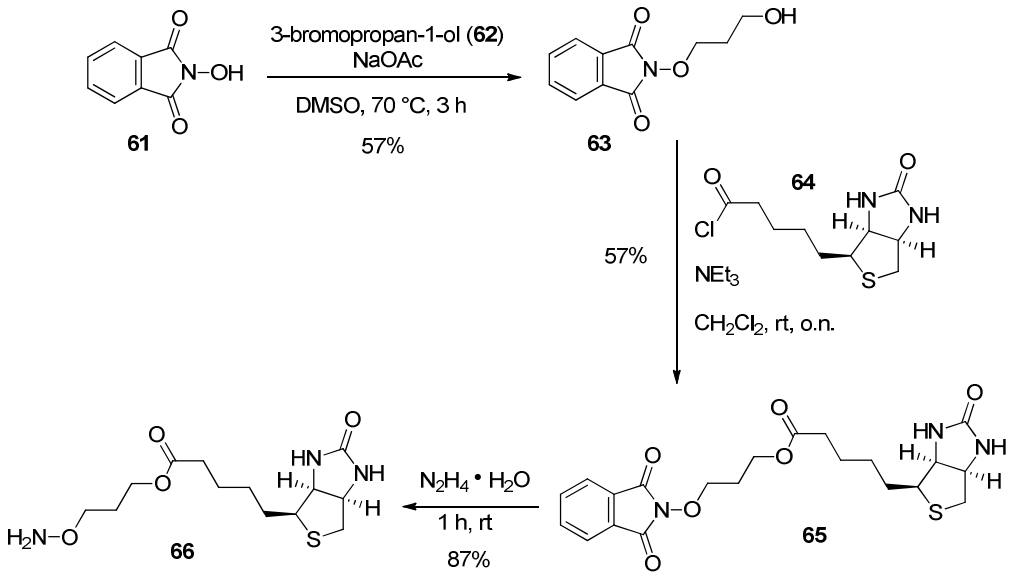


Figure 3.4: Modified protein TTL(D221Bpa,S261C)[Aha] (red: benzophenone, blue: cysteine, yellow: azides).

This modified thermophile bacterial lipase should enable selective multi-functionalisation by three different reactions. However, we decided to first probe each modification reaction individually on the protein level to check the suitability of our model protein.

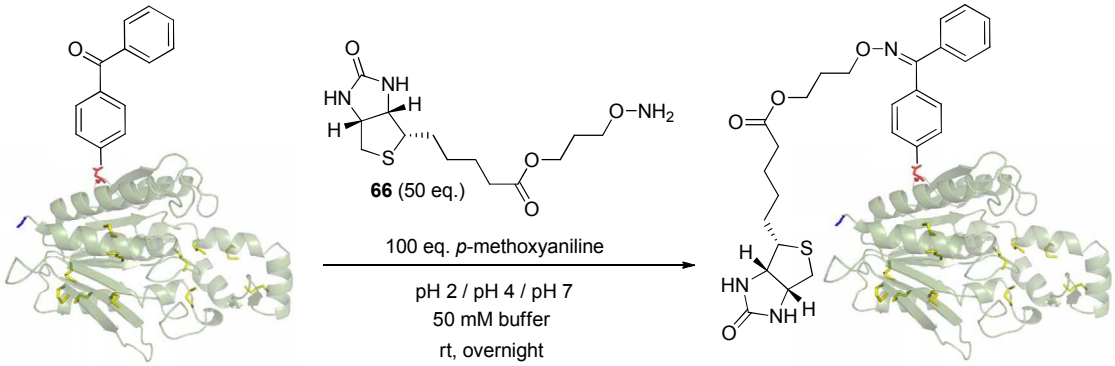
3.3.1.1. Oxime Ligation

To probe the oxime ligation on our protein of choice, we first focused on the synthesis of an appropriate biotin hydroxylamine derivative **66** (Scheme 3.23). To do so, *N*-hydroxy phthalimide (**61**) was coupled to a propanol linker and subsequently bound to biotin by ester formation. Biotin was first reacted with thionyl chloride to yield acyl chloride **64**, which was then added to a solution of alcohol **63**, which was synthesised from 3-bromopropan-1-ol (**62**) and *N*-hydroxy phthalimide (**61**) in 57% yield. Triethylamine was applied as a base and the desired protected hydroxylamine **65** was obtained in 57% yield. Phthalimide protected *O*-alkyl hydroxylamine **65** was deprotected in a solution of hydrazine over 1 h to achieve the final biotin hydroxylamine derivative **66** in 87% yield.



Scheme 3.23: Synthesis of biotin hydroxylamine 66.

To probe oxime formation on the protein level, test reactions were set up wherein hydroxylamine **66** and *p*-methoxy aniline[156a] (catalyst) were added to a solution of TTL(D221Bpa,S261C)[Aha] (0.6 mg/ml) in Tris/HCl-buffer (20 mM) at different pH values, namely pH 2.0, 4.0 and 7.5 and shaken overnight. Analysis by MALDI-MS showed mostly starting material and ~10% product formation for the reactions at pH 2 and pH 4 (Scheme 3.24). The reaction at pH 7.5 turned completely black overnight and showed no protein material in the MALDI-MS spectrum.



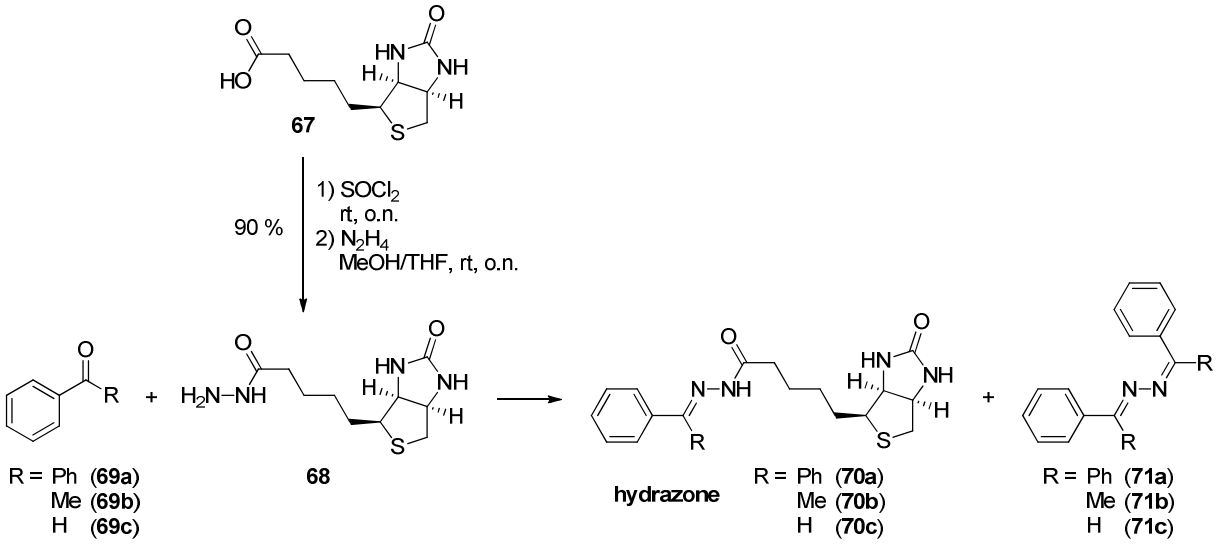
Scheme 3.24: Oxime formation with biotin hydroxylamine 66 on benzophenone protein TTL(D221Bpa,S261C)[Aha].

We simultaneously synthesised and probed a second biotin compound for selective labelling, namely biotin hydrazide **68**,[199] which is meant to form a hydrazone. The synthesis of biotin hydrazide **68** proceeded rather smoothly in 90% yield (Scheme 3.25). Hydrazone formation was then probed with the TTL under different reaction conditions at pH 4 and pH 7 with different aniline derivatives as catalyst,[156a, 157a, 159] but none of the applied reaction conditions showed

any product formation. In comparison to biotin hydroxylamine **66**, these results are not surprising as hydrazones are known to be less stable than oximes und thus more difficult to form under similar reaction conditions.

Due to the poor conversion rates, we decided to address the benzophenone as a possible cause for our ligation problems. To do so, we compared different carbonyl compounds for their reactivity in hydrazone formation with biotin hydrazide **68** – namely benzophenone amino acid (**69a**), acetophenone (**69b**) and benzaldehyde (**69c**) (Scheme 3.25). As aldehydes are known to be more reactive in hydrazone and oxime formations, we expected an increase in conversion from **69a** to **69c**. We decided to probe the influence of different solvents, organic and aqueous, at different pH values and also two different known catalysts (Table 3.3). In this sense, we chose aniline and the in terms of protein toxicity more benign *p*-amino phenylalanine as catalysts, both already known for their ability to improve hydrazone and oxime ligations at physiological pH.[156a, 157a]

The reaction of 4-benzoyl-L-phenylalanine (**69a**) with biotin hydrazide **68** delivered no product **70a** under all probed reaction conditions (Table 3.3, entry 1–3), which was probably due to its low solubility and its rather low reactivity. As expected, acetophenone (**69b**) gave better results with hydrazide **68**. The reaction in ethanol showed no desired product formation (Table 3.3, entry 1). However, hydrazone formations in a buffer-acetonitrile-system at pH 4.5 yielded after 3 h about 25% of the product **70b** as judged by LC-UV (Table 3.3, entry 5–6). At neutral pH, hydrazone formation seems to proceed worse (Table 3.3, entry 5 and 9), which is not surprising as Schiff base formations are known to proceed faster at slightly acidic pH.[200] In addition, the results indicate no prominent temperature effect for hydrazone formations on acetophenone (**69b**) (Table 3.3, entry 5–6) suggesting good applicability for proteins also at lower temperatures. Surprisingly, LC-UV data showed an unexpected negative effect of the applied catalysts as product formation with acetophenone (**69b**) could not be observed in the presence of aniline and *p*-amino phenylalanine at the same pH (Table 3.3, entry 7–8). In addition, the reaction with acetophenone (**69b**) and biotin hydrazide **68** seemed to yield an unexpected second hydrazone **71b**, which does not contain biotin but two acetophenone moieties connected by a hydrazine, in small portions (Scheme 3.25). LCMS analysis suggests cleavage of the N–C(O)-bond between hydrazone and biotin in the formed product. To exclude cleavage on the LCMS under acidic conditions, the LCMS was performed without acid. The detected result was the same supporting our first assumption that hydrazone **71b** is formed in the reaction itself.



Scheme 3.25: Synthesis of biotin hydrazide 68 and hydrazone formation with carbonyl compounds 69a–b.

As already expected, benzaldehyde (**69c**) gave the best yields in hydrazone formations towards **70c** with hydrazide **68**. Using ethanol as a solvent, the reaction yielded 50% of the desired product as judged by LCMS. The yield did not drastically change for the other solvent systems at pH 4.5 and pH 7 without the addition of catalyst. For this case, we could observe a small increase in product formation at slightly acidic pH when the catalyst was added to the reaction mixture. Unfortunately, we detected a negative influence of the catalyst for the reaction with benzaldehyde (**69c**) at pH 7.0 as formation of the desired product **70c** proceeded best without catalyst (Table 3.3, entry 17). The yield was slightly decreased by aniline and drastically dropped in the case of *p*-amino phenylalanine (Table 3.3, entry 18–19). This could be due to the high concentration of the catalyst which shifts the equilibrium between the desired product and the catalyst's imine towards the latter. Again, undesired cleavage of the biotin from the formed hydrazone towards **71c** could be detected by LCMS this time to a much higher extent. These results might be another possible reason why no protein modification with biotin hydrazide **68** could be observed.

In conclusion, best results were obtained with benzaldehyde at pH 4.5 in the presence of aniline. For this reason, we decided to switch from a benzophenone with rather low reactivity to a protein bearing an aldehyde moiety, which should allow faster and more efficient oxime formation on our protein of choice and will be discussed in chapter 3.3.2. In addition, we abandoned biotin hydrazide **68** for further experiments.

Table 3.3: Hydrazone formations probed between biotin hydrazide 68 and carbonyl compounds 69a–c (t = 3 h).

entry	reagent	solvent	T [°C]	catalyst (100 mM)	conversion to hydrazone 70a–c*
1	**69a**	EtOH	40	–	0%
2		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	40	–	0%
3		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	40	–	0%
4	**69b**	EtOH	25	–	0%
5		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	25	–	25%
6		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	40	–	25%
7		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	40	aniline	0%
8		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	40	*p*-amino Phe	0%
9		phosphate buffer (0.1 M, pH 7.0)	40	–	0%
10		phosphate buffer (0.1 M, pH 7.0)	40	aniline	0%
11		phosphate buffer (0.1 M, pH 7.0)	40	*p*-amino Phe	0%
12	**69c**	EtOH	25	–	50%
13		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	25	–	50%
14		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	40	–	50%
15		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	40	aniline	60%
16		NH_4OAc (30 mM, pH 4.5, 20% MeCN)	40	*p*-amino Phe	60%
17		phosphate buffer (0.1 M, pH 7.0)	40	–	50 – 60%
18		phosphate buffer (0.1 M, pH 7.0)	40	aniline	40 – 50%
19		phosphate buffer (0.1 M, pH 7.0)	40	*p*-amino Phe	20%

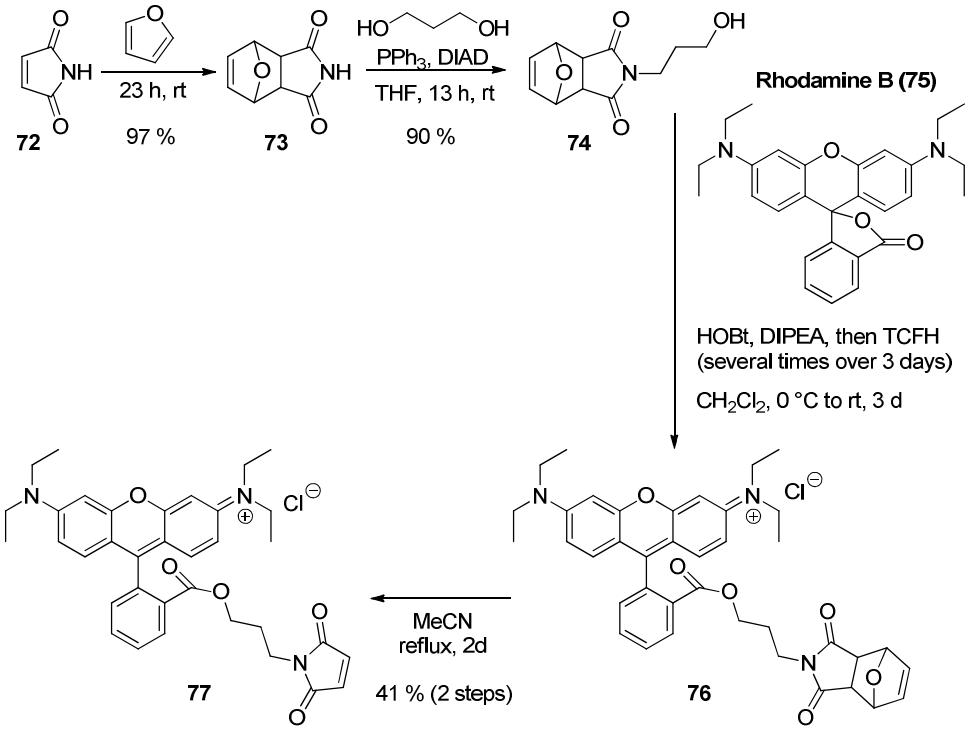
* Determined by LC-UV (λ = 280 nm).

3.3.1.2. Thiol-Michael Reaction

As the next modification strategy, we decided to probe fluorescent labelling of cysteine by thiol-Michael reaction on our protein model.[97a] To do so, we first had to synthesise an appropriate maleimide fluorophore derivative and therefore planned to use Rhodamine B for fluorescent labelling as it is rather cheap and known to cause less photobleaching than other common dyes, e.g. fluorescein.[201]

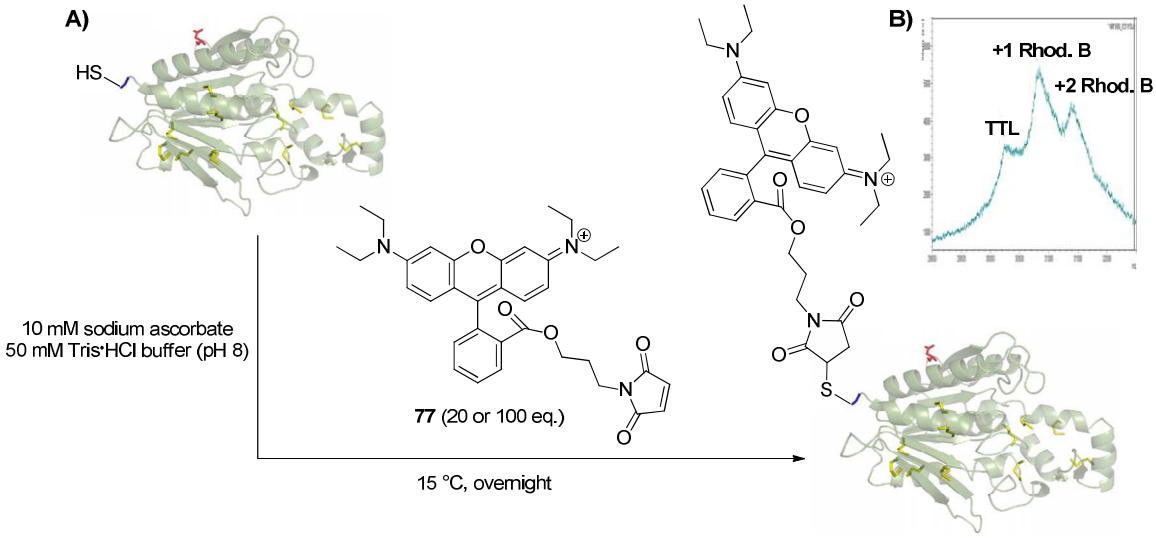
In a first step, maleimide **72** was reacted with furan in a [3+2]-cycloaddition to yield a protected maleimide **73** (Scheme 3.26).[202] Afterwards a propanol linker was introduced to yield compound **74**, which should then be coupled to Rhodamine B (**75**) by ester formation. Several approaches were performed to yield a stable linkage between maleimide alcohol **77** and Rhodamine B (**75**). Unfortunately, all our attempts to promote the coupling under a variety of conditions: acyl chloride formation, different coupling reagents, different bases and coupling with catalytic amounts of sulphuric acid, failed.[203] Finally, however using N,N,N',N'-tetramethylchloroformamidinium hexafluorophosphate (TCFH)[204] and HOBt with DIPEA as a base, ester bond formation could be achieved under very mild conditions and in good conversion to yield the desired ester **76** (Scheme 3.26). The crude reaction mixture of **76** was

then concentrated and without further purification re-dissolved in acetonitrile and refluxed for 2 d to achieve maleimide deprotection in a *retro*-Diels–Alder cycloaddition, which yielded the desired Rhodamine B maleimide **77** in 41% over two days.



Scheme 3.26: Synthesis of Rhodamine B maleimide 77.

Rhodamine maleimide **77** was probed in first test reactions with the TTL from the Budisa group (Scheme 3.27A, Table 3.4, entry 1–2). Maleimide **77** (20 or 100 eq.) was added to a solution of the lipase in Tris buffer (0.6 mg/ml, 20 mM buffer, pH8).[205] To reduce potentially formed disulfide bridges and improve reactivity, sodium ascorbate was added. The mixture was shaken overnight at 15 °C and purified by membrane ultracentrifugation. MALDI-MS showed in both cases a very similar result with a main peak, which corresponds to the single modified protein, the desired product (Scheme 3.27B). Besides, a low amount of starting material and a double reacted protein could be detected as discussed further below.



Scheme 3.27: Thiol-Michael reaction on TTL with Rhodamine maleimide 80 at pH 8: A) reaction scheme, B) MALDI-MS (Rhod. B = Rhodamine maleimide 77).

The thiol-Michael reaction is known to proceed faster under basic conditions, but other basic side chain groups, such as lysines, can interfere under these conditions and lead to undesired double coupling.[206] Therefore, the reaction was furthermore probed under different conditions at slightly lower pH 7.5 (Table 3.4, entry 3–6).

Table 3.4: First probed conditions for Thiol-Michael reaction with 77 on TTL overnight (20 mM Tris·HCl buffer).

entry	pH	T [°C]	Rhodamine maleimide 77 [eq.]
1	8.0	15	20
2	8.0	15	100
3	7.5	40	1
4	7.5	40	2
5	7.5	40	5
6	7.5	40	20

The reactions applying just 1 equivalent of Rhodamine B derivative **77** showed mostly starting material after one day and a rather small product shoulder (Table 3.4, entry 3). Subsequently, it was probed how an increased amount of Rhodamine maleimide **77** influences the reaction at 40 °C (Figure 3.5, Table 3.4, entry 4–6). With two equivalents of **77**, conversion was almost negligible as shown by MALDI-MS analysis (Figure 3.5A, Table 3.4, entry 4). With increase of maleimide **77**, product formation increased as well. Five equivalents yielded in 50% conversion towards the product (Figure 3.5B, Table 3.4, entry 5) and 20 equivalents yielded the same result as the reaction before at pH 8 with a main peak of the product two big shoulders for the starting material and the double-reacted protein (Figure 3.5C, Table 3.4, entry 6).

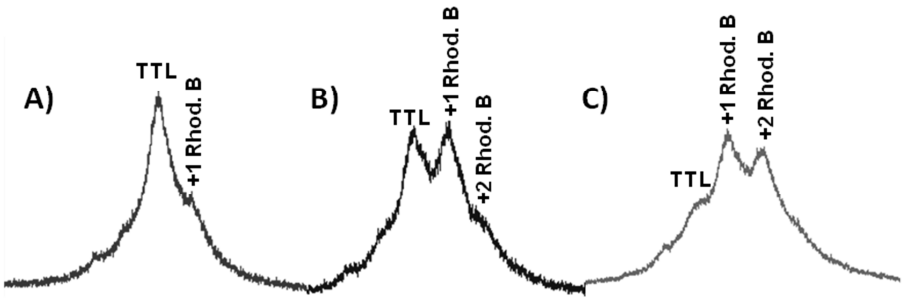


Figure 3.5: MALDI-MS spectra of thiol-Michael reaction on TTL at pH 7.5: A) 2 eq., B) 5 eq., C) 20 eq. of 77 (Rhod. B = Rhodamine maleimide 77).

Due to the still high amount of double addition, we decided to perform the thiol-Michael reaction at pH 6.5 as basic side chains such as lysine should react considerably slower than cysteines at that pH.[206] The reaction was probed at 40 °C adding 10 and 50 equivalents of Rhodamine maleimide **77**, respectively. Furthermore, the reaction with 50 eq. was probed with an 80% lower concentration. MALDI-MS measurements (Figure 3.6) showed as expected an increased product formation when more maleimide is present (Figure 3.6A and B). But as seen before at higher pH values, the reaction with 50 eq. of maleimide **77** showed a small shoulder, which indicates double addition of maleimide **77** on the protein. Decreasing the concentration of the reaction shows a slower product formation and no obvious shoulder after 24 h reaction time similar to the reaction at 40 °C with 10 eq. Rhodamine maleimide **77** (Figure 3.6C).

Comparing the reactivity of a thiol and, e.g., a lysine chain at pH 6.5, one might expect the thiol to react faster with the added maleimide moiety. The formation of a double reacted protein shoulder indicates the thiol might not be able to react fast enough. This could be due to steric hindrance of the addressed cysteine side chain. Though it was expected that – based on the protein's crystal structure in combination with *PyMOL* as a prediction program (open-source molecular visualisation system)[207] – the thiol should be exposed to the solvent, it is possible that in reality the thiol group is situated differently. Therefore, two different chaotropic agents and a detergent were probed to disrupt the tertiary structure of the protein and enable better and more selective thiol maleimide coupling: guanidinium hydrochloride (Gnd·HCl), aminoguanidine hydrochloride and sodium dodecyl sulfate (SDS).[208] The different additives were applied to the reaction mixture at pH 6.5, 40 °C with 10 equivalents maleimide **77**. In the first trial, the concentration of the chaotropic agents and the detergent was set to 100 mM to avoid full disruption of the tertiary structure. SDS seemed to form micelles in the reaction mixture and enclose the Rhodamine maleimide. It was not possible to wash the Rhodamine off by membrane ultracentrifugation and therefore to detect any protein mass by MALDI-MS. Thus, we decided to neglect the detergent for our further studies.

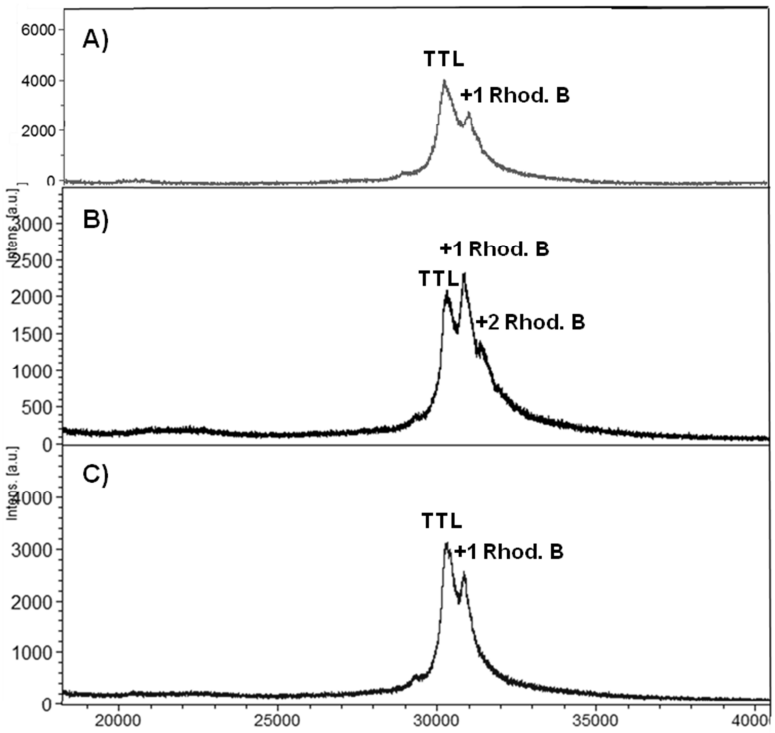


Figure 3.6: MALDI-MS spectra of thiol-Michael reaction at pH 6.5, 40 °C: A) 10 eq. 77, B) 50 eq. 77, C) 50 eq. 77 (80% lower concentration) (Rhod. B = Rhodamine maleimide 77).

Thiol-Michael reaction in the presence of Gnd·HCl and aminoguanidine·HCl showed no visible product formation in the MALDI-MS spectra, which indicates a disturbance of the reaction by chaotropic agents. At higher concentrations of Gnd·HCl and aminoguanidine·HCl (1 M) and twice as much maleimide **77** (20 eq.), MALDI-MS measurements revealed oligomer formation of the protein rather than product formation.

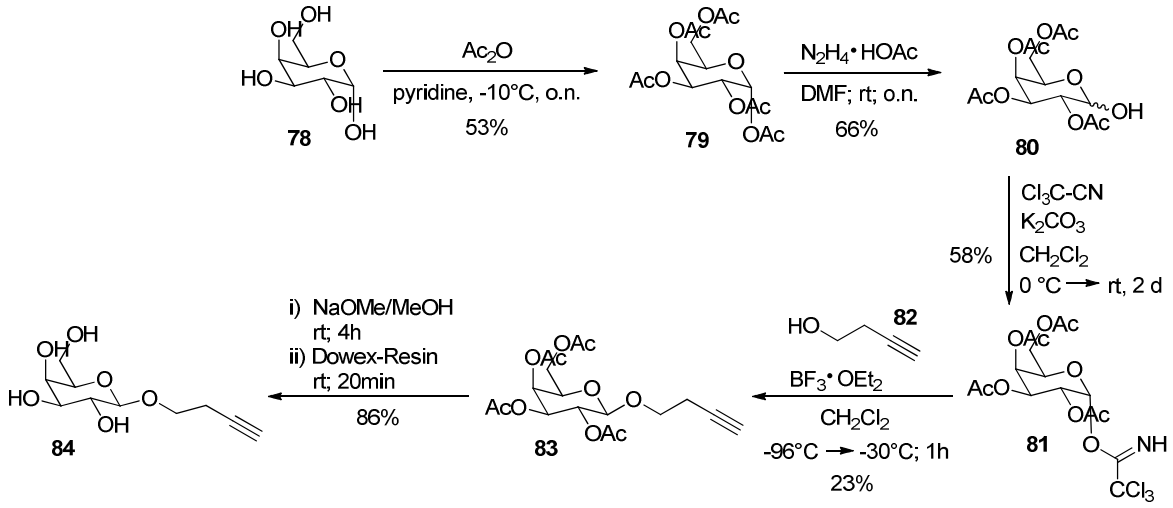
With the best conditions leading to ~ 50% product formation and ~10% double addition so far, we decided to evaluate the reactivity of the cysteine moiety. Therefore, the TTL was reacted with phenylvinyl sulfone. This reagent is known to be more reactive towards thiols than maleimides at increased pH.[209] The reaction with the TTL lipase was probed at 27°C overnight at pH 7.5 which should enable good and fast conversion. The sulfone was applied to the reaction mixture with 3.5 and 8.5 equivalents, respectively, with and without addition of either sodium ascorbate or TCEP as reducing agent. All reactions showed partial product formation but no full conversion despite the good reactivity of phenylvinyl sulfone. These results clearly indicate that the cysteine is not reactive enough, which might be due to a less solvent exposed position than expected.

We therefore decided to probe a new position for the cysteine in the next protein construct, which will be discussed further in chapter 3.3.3.1.

3.3.1.3. CuAAC

Finally, the third modification strategy was probed to selectively functionalise the azidohomoalanine residues on the TTL by CuAAC. We planned to synthesise β-butinyle galactose **84** (Scheme 3.28), which could also be used subsequently in carbohydrate–protein binding studies.

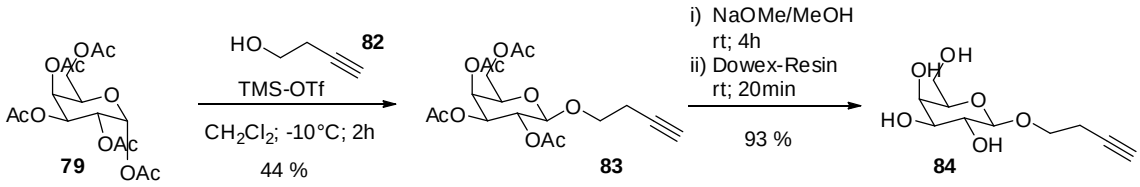
For that purpose, the synthesis of β-butinyle galactose **84** was carried out *via* two different reaction strategies starting from peracetylated galactose **79**, which was obtained from galactose **78** in the reaction with acetic anhydride in pyridine[210] in 53% yield (Scheme 3.28). Alkyne functionalisation of the anomeric position in **79** was tested by different methods. We could not observe any product formation by either direct substitution of alcohol **82** in presence of $BF_3{\cdot}OEt_2$ as a Lewis base[211] or by the synthesis of the anomeric bromide with HBr/AcOH followed by substitution with **82** in the presence of a silver salt.[212] However, when we performed selective deprotection of the anomeric hydroxyl group towards **80** and subsequent conversion to the more reactive trichloroacetimidate *α*-galactose **81** (Scheme 3.28) we achieved formation of the desired acetal **83** upon addition of $BF_3{\cdot}OEt_2$ and 3-butyn-1-ol (**82**).[211] In a last step, but-3-ynyl 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside (**83**) was fully deprotected with a NaOMe/MeOH-solution followed by Na/H-exchange with a DOWEX-resin, which yielded the pure but-3-ynyl β-D-galactopyranoside (**84**) in an overall yield of 8%.



Scheme 3.28: Synthesis of β-butinyle galactose 84 *via* trichloroacetamide 81.

Simultaneously, we probed the synthesis of but-3-ynyl β-D-galactopyranoside (**84**) following another reaction strategy, which yielded the desired product **84** in only two steps by direct

substitution with butyn-3-ol (**82**) on peracetylated galactose **79** with the use of trimethylsilyl triflate as Lewis acid[213] in an overall yield of 41% (Scheme 3.29).



Scheme 3.29: Synthesis of β-butinyle galactose 84 by Lewis acid catalysis with TMS-OTf.

First CuAACs were probed with butinyle galactose **84** and TTL(D221Bpa,S261C)[Aha]. The lipase has eleven azidohomoalanine residue incorporated into its sequence with four azido groups being more or less situated on the surface of the protein (determined by *PyMOL*).[207] First coupling reactions were performed with a 100-fold excess of alkyne galactose **84** per azide with $CuSO_4$ as catalyst and sodium ascorbate as reducing agent. In addition, a water-soluble Cu-ligand (THPTA: tris(hydroxylpropyl) triazolylmethyl amine) was added to accelerate the reaction and protect the protein from oxidative side-reactions and amino guanidine was used as chaotropic agent.[214] The reaction was left overnight at 25 °C. After purification, MALDI-MS measurements showed no starting material but a broad peak slightly upshifted (Figure 3.7). The mass corresponds to a 4-fold clicked protein which leads to the assumption that under these rather mild reaction conditions only the homoazidoalanines on the outer sphere of the protein reacted.

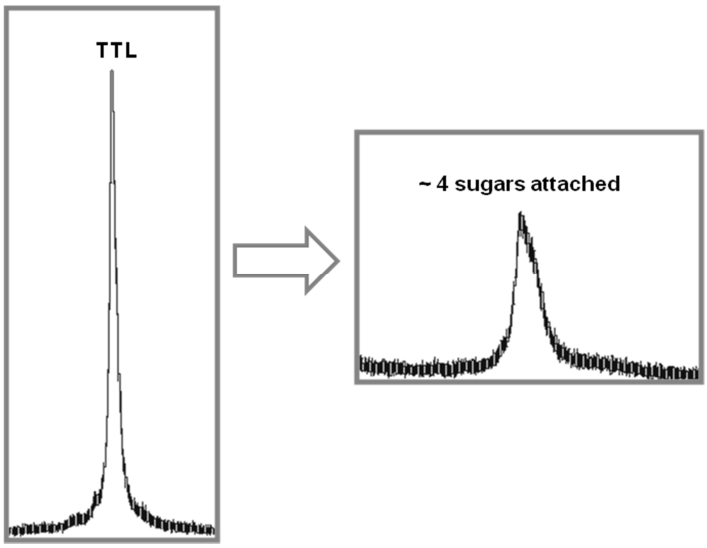


Figure 3.7: MALDI-MS of TTL before (left) & after (right) CuAAC with β-butynyl-galactose 84.

Different reaction conditions such as Cu-concentration, protein concentration, pH and the addition of aminoguanidine as a chaotropic agent were probed and conversions were checked by MALDI-MS (see experimental section, Table 5.11). As the MALDI peaks were usually rather broad and therefore difficult to analyse, different matrices were applied for protein detection

after the reactions: A) sinapinic acid (SA, as before), B) DHB (2,5-dihydroxybenzoic acid; known for complex glycoproteins) and C) “Super-DHB” (DHB, 10% 2-hydroxy-5-methoxybenzoic acid).[215] For some reactions the MALDI spectra looked very much alike, but especially broader peaks in SA with a big shift to higher masses gave more distinct peaks with “Super-DHB” that were shifted back to lower masses. The sharper peaks seem to indicate that the newly chosen matrix can lower the degree of salt adducts for MS-detection.

In summary, we could show that previously employed benzophenone is not reactive enough for an efficient oxime ligation. In addition, conversions with the cysteine required large amounts of maleimide **77**, which indicated that the cysteine is not accessible enough for a selective modification. Therefore, we decided to use a new protein for subsequent attempts on a triple protein modification, which, as discussed before, will bear an *N*-terminal serine instead of a benzophenone and the position of the cysteine should be shifted to a different hopefully more solvent-exposed site.

3.3.2. Selective Dual-Modification on Newly Engineered S-TTL[Aha]

3.3.2.1. General Overview

As a next step, the dual-modification of the protein should be probed by combination of the oxime ligation and the CuAAC. Taking our results from the mono-functionalisation into account, we used a protein, expressed by the group of Budisa, with the eleven azides and an *N*-terminal serine. As described before, the *N*-terminal serine allows the selective oxidation with periodate to the aldehyde,[118] which should be much more reactive than the previously used benzphenone moiety. At this point, we wanted to concentrate first on the oxime ligation. It was reported before that oxidative periodate treatment of a protein might lead to cysteine oxidation. Therefore, the group of Budisa expressed a cysteine-free TTL bearing an *N*-terminal serine, S TTL[Aha], for first attempts on glycol cleavage and oxime ligation on the protein level. In addition, one methionine residue was mutated back to its original amino acid due to a previously observed instability of the protein.[216] Protein expression usually relies on the application of methionine as a start codon, which is cleaved after successful expression by a methionine aminopeptidase to yield the protein of choice. Unfortunately, the efficiency of the Met-aminopeptidase cleavage decreases when the *N*-terminal methionine is exchanged for an unnatural amino acid such as azidohomoalanine.[217] We therefore obtained a protein mixture (Aha)S-TTL[Aha] containing approximately 50% AhaS-TTL[Aha] and 50% S-TTL[Aha] (judged by MS, see also Scheme 3.33), which was nevertheless applied for subsequent dual-modification experiments. However, this *N*-terminal heterogeneity did not hamper our subsequent application, since only biotinylated protein could bind to the chip for SPR studies (see chapter 3.3.2.3).

As a first step, aldehyde formation on an *N*-terminal serine and subsequent oxime formation were initially probed under different reaction conditions on peptide level and a new biotin hydroxylamine derivative was synthesised (see chapter 3.3.2.2).

Afterwards, protein dual-modification was successfully probed on the cysteine-free (Aha)S-TTL[Aha]. After glycol cleavage and oxime formation with biotin hydroxylamine **66**, butinyle galactose **84** was effectively coupled to different degrees to the azidohomoalanine residues on the TTL by CuAAC. To our knowledge, this shows the first combination of auxotroph supplementation and chemical modification by glycol cleavage for the successful dual-modification of a protein (see chapter 3.3.2.3). Three differently modified proteins (**Gal-0**: zero galactose units attached, **Gal-1**: one to two galactose units attached, **Gal-3**: three to five

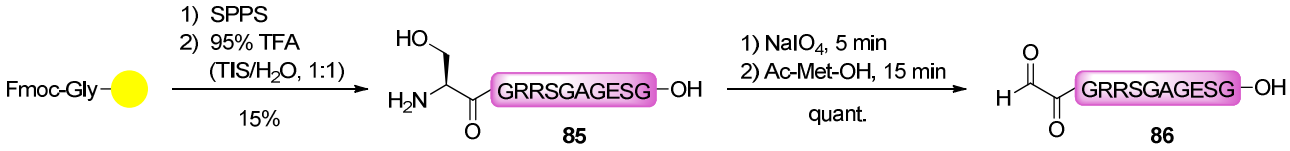
galactose units attached) were probed in lectin binding studies with ECL. The proteins were immobilised on a streptavidin-coated chip by its biotin linker and different concentrations of ECL were probed for binding to the galactose-TTL models in Surface Plasmon Resonance (SPR) measurements showing a slight increase in binding efficiency from **Gal-1** to **Gal-3** (see chapter 3.3.2.3).

In addition, we probed thiazolidine formation as an alternative modification reaction for aldehydes with the potential for selective cleavage due to their lower stability compared to oximes (see chapter 3.3.2.4).

3.3.2.2. Initial Studies on Glycol Cleavage and Oxime Formation

Before focussing on the modification of our new TTL, we decided to probe selective glycol cleavage with periodate and oxime formation on a model peptide. To do so, a random peptide sequence with several serine and threonine residues was chosen (Scheme 3.30). Peptide **85** was synthesised by SPPS and purified by semi-preparative HPLC with a total yield of 15%.

Glycol cleavage was probed with sodium periodate for 5-15 min at rt.[218] This yielded in all cases quantitative conversion to the desired aldehyde peptide **86** (Scheme 3.30). For subsequent reactions with hydrazines or hydroxylamines, residual sodium periodate had to be quenched. First attempts with sodium nitrite[218] yielded a side product based on unspecific addition of SO_2 to the reaction product **86**. However, with another reducing agent, *N*-acetylmethionine,[218] the formation of peptide **86** proceeded quantitatively without any by-product formation (Scheme 3.30). *N*-Acetylmethionine was therefore applied for all further studies.

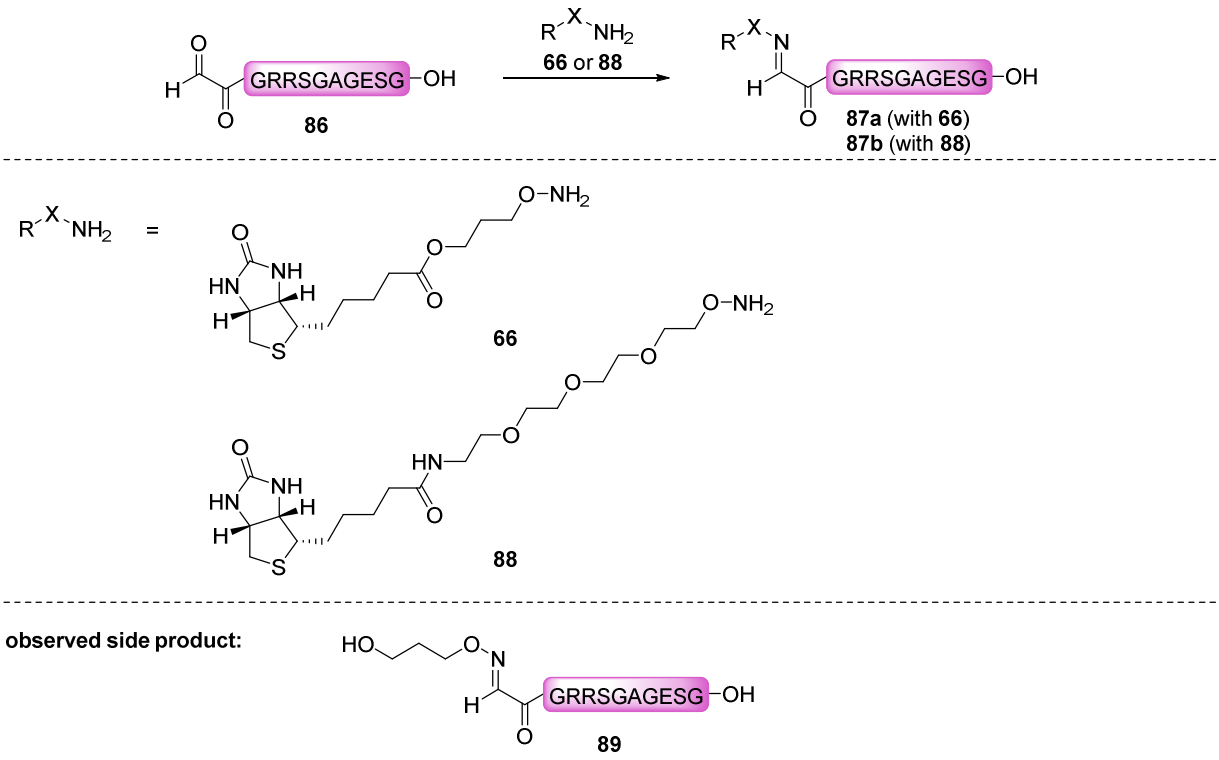


Scheme 3.30: Synthesis of *N*-terminal serine peptide 85 and probed glycol cleavage towards aldehyde peptide 86.

Once we established the glycol cleavage conditions, first oxime ligations were performed with two different reagents: biotin hydroxylamine **66** and commercially available more water-soluble biotin hydroxylamine **88** (Scheme 3.31). Different reaction conditions were varied, such as buffer and pH, temperature, catalyst and equivalents.

First studies with hydroxylamine **66** were conducted without catalyst. At pH 7 no product formation to **87a** could be detected after 24 h at 30 °C probably due to its low water-solubility.

As it is known that oxime ligations proceed much faster at slightly acidic pH,[200] we lowered the pH to 4.5 and added 20% acetonitrile for better solubility, which yielded more than 80% conversion of aldehyde peptide **86** (judged by UV). Upon addition of aniline as catalyst,[156a] we could accelerate the reaction immensely. After 3 h hardly any aldehyde **86** could be detected but mostly peptide oxime **87a** and some peptide anilinium imine. Unfortunately, biotin ester **66** did not seem stable enough and was partially hydrolysed during the reaction to oxime **89** (Scheme 3.31).

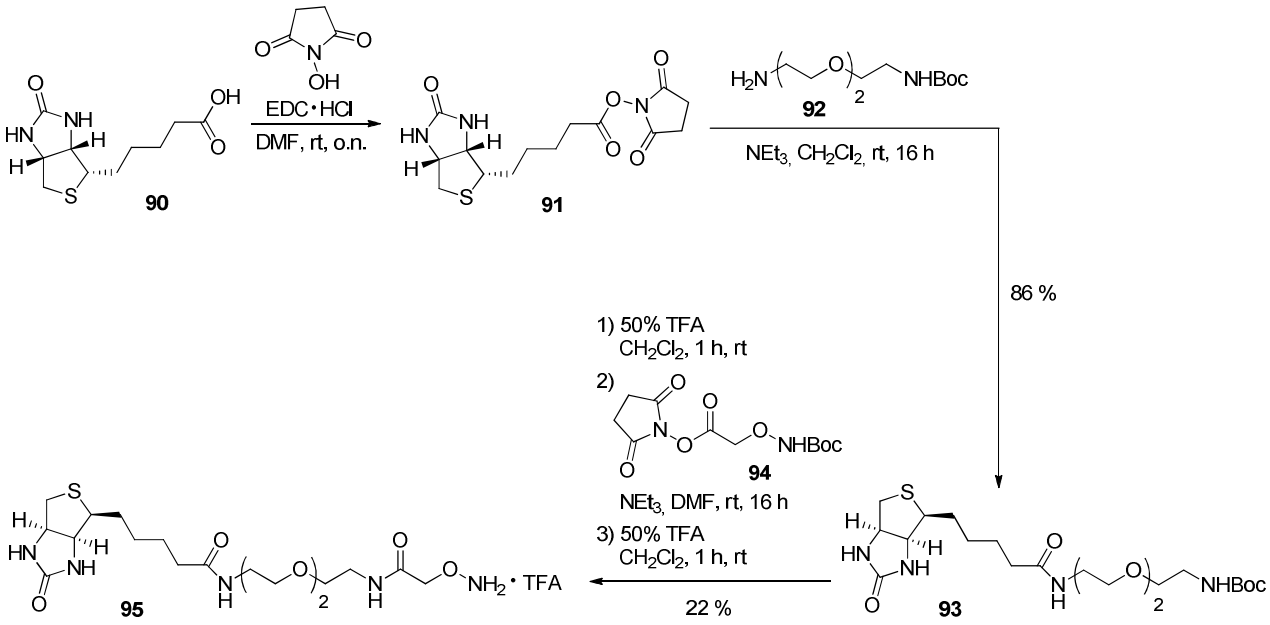


Scheme 3.31: Oxime formation on aldehyde peptide 86 with hydroxylamines 66 and 88.

In comparison to hydroxylamine **66**, biotin hydroxylamine **88** showed better conversion rates at pH 7. Two previously applied catalysts with higher water-solubility and lower protein-toxicity than aniline were probed to allow milder reaction conditions for later protein studies, *p*-methoxyaniline and *p*-amino phenylalanine. Hydroxylamine **88** yielded almost full conversion towards the desired oxime biotin peptide **87b** after 3 h with methoxyaniline[156a] at 30 °C and after 16 h with *p*-amino phenylalanine[157a] at 10 °C with the latter result showing promising for a good applicability of the reaction for oxime formation on our protein of choice.

Due to the good results for hydroxylamine **88** in the probed oxime formations but also due to its high price, we decided to synthesise our own biotin derivative **95** with stable amide bonds and a short oligo ethylene glycol chain to enhance water-solubility as biotin hydroxylamine was rather

expensive and we had only little amounts left of it. For the synthesis of biotin hydroxylamine **95**, biotin succinimide (**91**), which was obtained from D-biotin (**90**), was first coupled with a short ethylene glycol linker **92** to enhance water-solubility and yield compound **93**. After Boc-deprotection, attachment of a succinimide *N*-Boc α-aminoxyacetate **94** moiety followed by a second Boc-deprotection afforded the desired biotin hydroxylamine derivative **95** in five steps with an overall yield of 19% (Scheme 3.32). Hydroxylamine **95** was applied in all further oxime formations on protein level as further described in chapter 3.3.2.3 and 3.3.4.3.

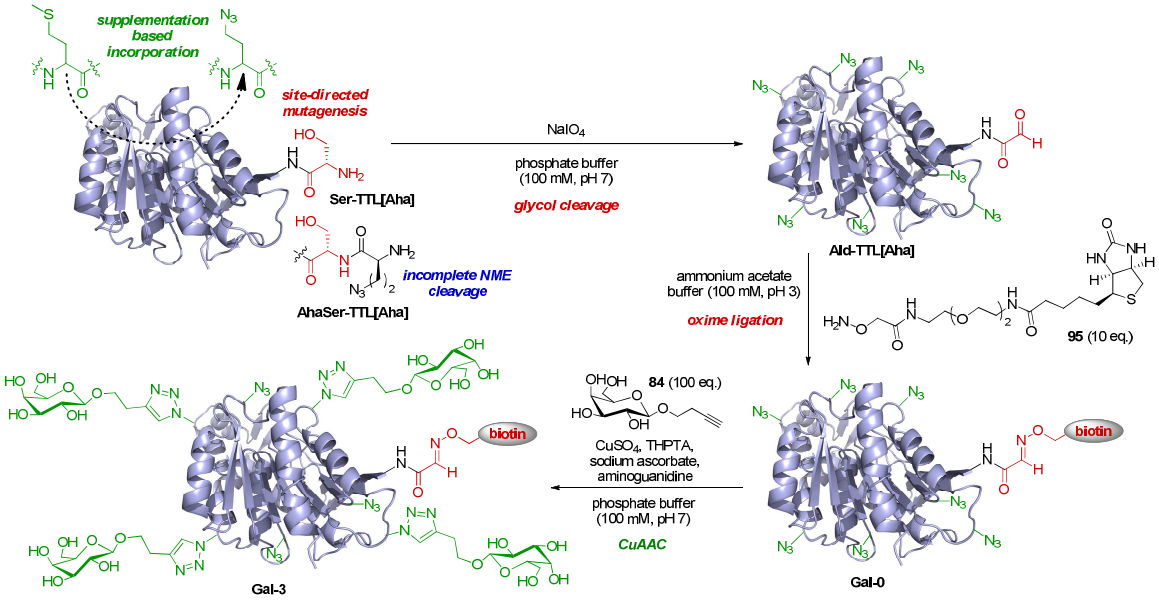


Scheme 3.32: Synthesis of biotin hydroxylamine 95.

3.3.2.3. Orthogonal Dual-Modification of Proteins by Oxime Ligation and Cu-Catalyzed Azide–Alkyne Cycloaddition

After successful glycol cleavage and oxime ligation studies on the peptide level, we now decided to probe our dual-modification approach by oxime ligation and CuAAC on our model protein mixture (Aha)S-TTL[Aha]. Oxime ligation and CuAAC have been reported previously to be orthogonal to each other in DNA model systems and proteins.[164a, 194b, 219] Nevertheless, since glycol cleavage is needed to generate the *N*-terminal aldehyde,[118] we initialised our synthetic route with $NaIO_4$ treatment since the galactose units installed by CuAAC would be efficient targets for a glycol cleavage, as shown previously.[161] Based on our optimization experiments for the periodate treatment of *N*-terminal Ser peptides (see above), the TTL was treated with sodium periodate in a phosphate buffer at pH 7 and 15 °C for 1 h and quenched with *N*-acetyl methionine to quantitatively form the aldehyde Ald-TTL[Aha] in a mixture with the unreactive

AhaSer-TTL[Aha] (Scheme 3.33, Figure 3.8A). For the oxime ligation with the newly synthesised biotin hydroxylamine derivative **95**, several reaction conditions were screened to achieve full conversion based on MALDI-MS analysis for the Ald-TTL[Aha], in which the unreactive AhaSer-TTL[Aha] served as a reference point (Figure 3.8A). With rather mild reaction conditions at pH 7 with *p*-anisidine as a catalyst only 10 % product was formed,[156a] We therefore decided to lower the pH and increasing the amount of hydroxylamine **95** to promote the desired Schiff's base formation (Table 3.5). Finally, full conversion to **Gal-0** could be achieved in an ammonium acetate buffer (100 mM, pH 3.0) with 20 eq. hydroxylamine **95**. Oxime formation at higher pH required much larger amounts of hydroxylamine **95** with only 50% conversion with 50 eq. of **95**.



Scheme 3.33: Protein design & dual-functionalisation of TTL: periodate cleavage, oxime ligation and CuAAC.

Table 3.5: Conditions for oxime ligation with hydroxylamine 95 on Ald-TTL[Aha].

entry	hydroxylamine 95	pH	catalyst (10 eq.)	conversion*
1	30 eq.	7.4	*p*-anisidine	10%
2	50 eq.	5.0	–	50%
3	30 eq.	5.0	–	30%
4	30 eq.	4.5	–	60%
5	30 eq.	4.0	–	65%
6	30 eq.	3.5	–	80%
7	30 eq.	3.0	–	100%
8	20 eq.	3.0	–	100%
9	10 eq.	3.0	–	80%
10	5 eq.	3.0	–	55%

† 50 µM, 15 °C, overnight. * Estimated by MALDI-MS in comparison to unreacted AhaS-TTL[Aha]takes.

Successful biotinylation towards **Gal-0** could also be shown by SDS-PAGE and Western Blot analysis (Figure 3.8B–C, lane 3), where we can observe a small band shift on the SDS gel as well as a strong signal for **Gal-0** on the Western Blot, which is the result from a specific non-covalent biotin–streptavidin binding with streptavidin being attached to fluorescent luminol.

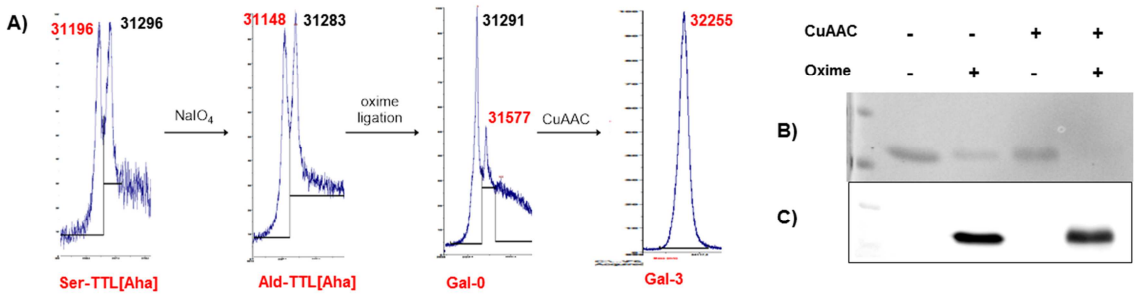


Figure 3.8: Dual-functionalisation of TTL: A) MALDI-MS spectra (red: modified protein, black: reference AhaS-TTL[Aha]), B) SDS-PAGE of TTL protein conjugates (Coomassie stain), C) Western Blot (anti-biotin) (lane 1: protein ladder, lane 2: (Aha)S-TTL[Aha], lane 3: Gal-0 + AhaSer-TTL[Aha], lane 4: CuAAC on (Aha)S-TTL[Aha], lane 5: Gal-3).

In the next step, we decided to directly probe CuAAC on our biotinylated protein **Gal-0** with previously synthesised β-butynyl galactose **84** (Scheme 3.33). Corresponding to our former attempts on CuAAC on TTL(D221Bpa,S261C)[Aha], the reaction was performed in phosphate buffer (100 mM, 100 mM NaCl, pH 7) with varying amounts of $CuSO_4$ to potentially obtain different degrees of glycosylation. After purification of our protein samples, conversions were checked by MALDI-MS, SDS-PAGE and Western Blot (Figure 3.8). Depending on the different Cu^{2+} concentrations, we successfully achieved dual-functionalisation of our TTL with different amounts of galactose units per protein (Table 3.6).

Table 3.6: Conditions for CuAAC on biotinylated TTL (Gal-0) (5 eq. THPTA to Cu^{2+}, 8mM aminoguanidine, 100 eq. alkyne 84 per azide, 50 eq. sodium ascorbate/Cu^{2+}, 15 °C, overnight).

entry	c [µM] (protein)	$CuSO_4$ (to alkyne 84)	product*
1	10	5 mol%	3-5 clicked sugars
2	10	10 mol%	3-4 clicked sugars (**Gal-3**)
3	10	30 mol%	1-2 clicked sugars (**Gal-1**)
4	10	50 mol%	1-2 clicked sugars

* Ratios judged by MALDI-MS in comparison to biotinylated TTL.

Surprisingly, lower concentrations of $CuSO_4$ led to higher degrees of functionalisation with 3–5 galactose units as the highest carbohydrate labelling (5 mol% Cu^{2+}), whereas higher amounts of Cu^{2+} led to lower degrees of functionalisation with only 1–2 clicked sugars (30 mol %, **Gal-1**). It was not possible to achieve a higher galactose functionalisation, which might be due to the inaccessibility of some of the azidohomoalanines. In addition, it should be noted that higher concentration of Cu^{2+} also led to precipitation and loss of protein material, which indicates strong binding of Cu^{2+} to our protein of choice, which might also result in labelling inhibition.

Final purification of our protein material by centrifuge membrane filtration yielded approximately 20–35% of the initially employed (Aha)Ser-TTL[Aha] as judged by UV.

To ensure the stability of TTL throughout the dual-labelling process, we performed a lipase activity assay to demonstrate that the enzymatic activity could be retained (Figure 3.9C).[198b] Therefore, different protein samples – (Ald)Ser-TTL[Aha], **Gal-0**, **Gal-1**, and **Gal-3** (see also Table 3.6) – were incubated with *p*-nitrophenyl palmitate for 1 h at 50 °C. Upon ester cleavage, the formation of *p*-nitrophenol can be quantified by UV due to a specific absorption at 410 nm. We could thereby show that the enzymatic activity and therefore the protein's integrity did not interfere with our dual-modification approach.

In a final experiment, our protein scaffold was planned to be immobilised on a streptavidin gold chip to monitor carbohydrate–protein binding interactions by surface plasmon resonance (SPR) and to measure lectin binding constants (Figure 3.9A and B). This immobilisation strategy should allow easy handling and reproducible orientation, which are notable improvements over the alternative active ester immobilisation. In addition, our approach required considerably less amount of the inhibiting glycoconjugate in comparison to the reverse approach, which involves immobilisation of lectin and titration of the binder.

The SPR studies were conducted by the group of Dr. Dernedde at the Charité Berlin. First, the qualitative binding of *Erythrina cristagalli* lectin (ECL) to proteins **Gal-1** and **Gal-3** as well as **Gal-0** as a negative control was probed. Although **Gal-3** was not the protein sample with the highest galactose functionalisation (3–4 moieties instead of 3–5), we chose it for further binding studies due to its more narrow range of functionalisation. The three protein samples were each immobilised on a streptavidin-coated chip (Figure 3.9A). Then, ECL was passed over the chip at different concentrations to determine the relative binding affinity for the immobilised glycosyl-TTL coated surface. At a concentration of 10 µM ECL, the results showed significant binding of both glycosylated protein samples towards the lectin. In contrast, the non-glycosylated lipase showed no binding at all. To further characterise the binding efficiency, K_D-values were determined by SPR measurements (Figure 3.9B). Again, **Gal-0** showed no binding. Both glycosylated proteins, **Gal-1** and **Gal-3**, presented very similar and rather low K_D-values (60 and 70 µM, respectively) with a slight tendency for stronger binding for the higher glycosylated protein **Gal-3**. As the two Gal-binding sites of ECL are localised on opposite sides,[220] our rather short butynyl linker might not be able to fully bend around the protein to achieve a multivalent effect.[173, 221]

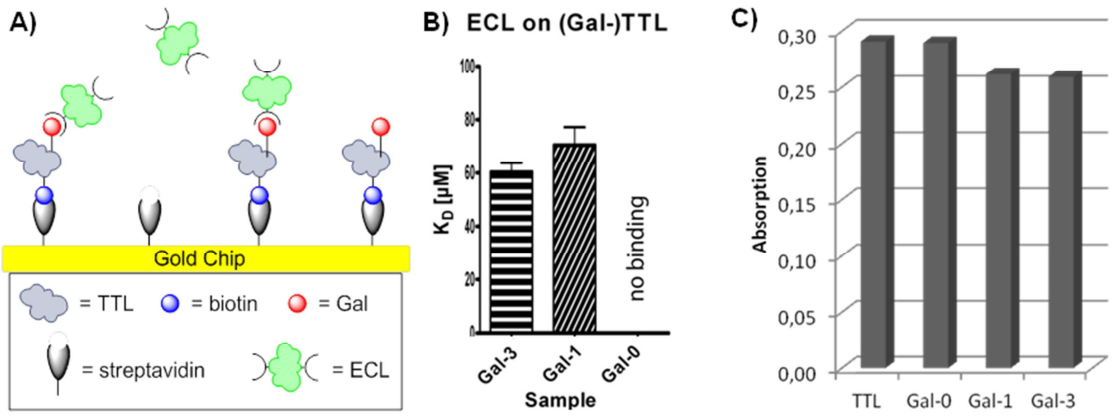
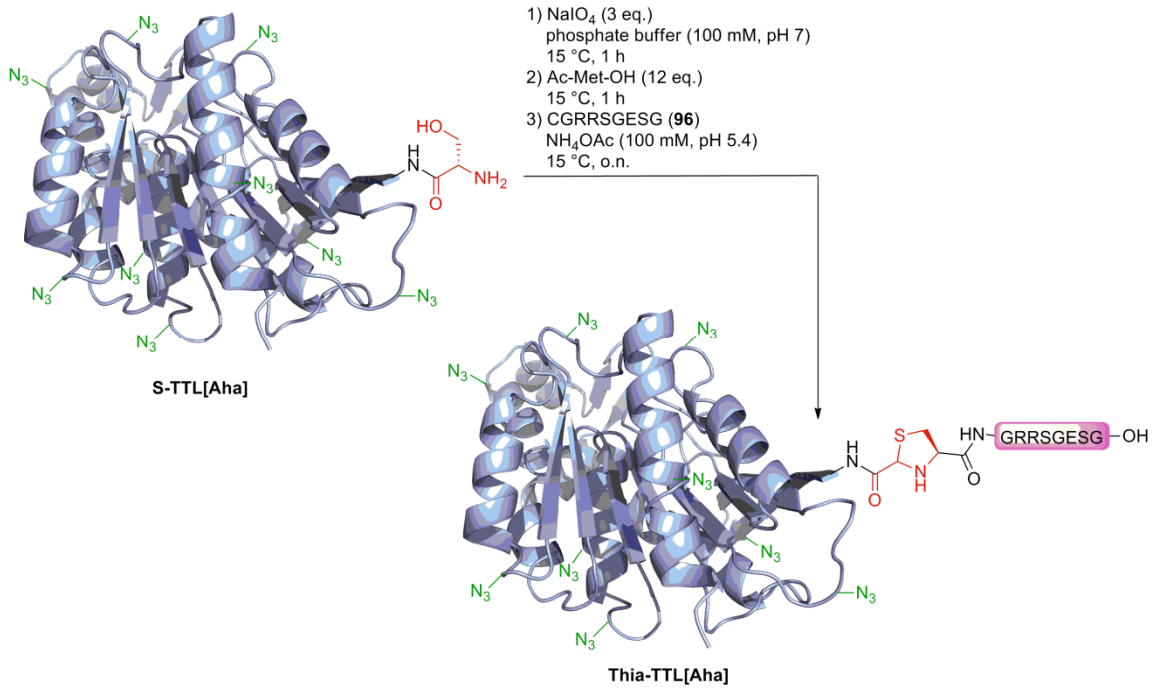


Figure 3.9: SPR measurements performed by the group of Dr. Dernedde: A) set-up and B) K_D-values. C) Results of lipase activity assay.

3.3.2.4. Thiazolidine Formation on S-TTL[Aha] and Potential Thiazolidine Cleavage

After our dual-modification of proteins by oxime ligation and CuAAC (chapter 3.3.2.3), we decided to probe thiazolidine formation as a second modification technique for aldehydes, which is less commonly applied.[161] Thiazolidine groups are generally used as protecting groups for pseudoproline derivatives in peptide synthesis and for *N*-terminal cysteines in sequential peptide ligations by NCL and are known to be cleaved at more acidic pH in the presence of thiols or methoxyamine.[27, 222] The thiazolidine formation has been applied in only few ligation strategies during the last decades, e.g. in the synthesis of peptide-peptide,[223] peptide-oligonucleotide,[224] peptide/protein-lipid[113a, 223] and protein-peptoid conjugates[113b] and in peptide immobilisation on a sensor chip[225] with all methods generally employing the N-terminal cysteine of a protein or peptide and a synthetic aldehyde building block. To our knowledge, the only reported application of a thiazolidine formation on an aldehyde protein was reported in 1996 by the group of Tam wherein an aldehyde was generated either from an N-terminal serine or glycan units by periodate treatment and subsequently reacted with a cysteine biotin in a thiazolidine formation.[161]

In addition, we envisioned that this might allow us to introduce a biotinylated tag for purification, which can be cleaved off afterwards (as applied in NCL for sequential ligation) and exchanged for another functional hydroxylamine or cysteine tag.



Scheme 3.34: Thiazolidine formation with cysteine peptide 96 on S-TTL[Aha].

To do so, we synthesised a trivial water-soluble cysteine peptide **96** to probe thiazolidine formation on our (Aha)S-TTL[Aha] protein mixture (see also chapter 3.3.2.3), which lacks cysteine and was previously applied in our dual-modification experiments (Scheme 3.34). Aldehyde formation was performed with three equivalents $NaIO_4$ in phosphate buffer (100 mM, pH 7) followed by periodate quenching with *N*-acetyl methionine.[118b] Thiazolidine formation was performed in an ammonium acetate buffer (100 mM, pH 5.4) with 10, 20 and 30 equivalents of cysteine peptide **96**, respectively.[26, 113b] After purification, the samples were analysed by MALDI-MS, which showed full conversion to Thia-TTL[Aha] with 30 equivalents of cysteine peptide **96**.

As a next step, we decided to probe thiazolidine cleavage on our purified protein sample – a mixture of Thia-TTL[Aha] and unreactive AhaS-TTL[Aha] (see chapter 1.3.2.2). Different conditions were probed varying pH (2–8), temperature (15–37 °C), reaction time (3–24 h) and different known additives such as methoxyamine hydrochloride, sodium 3-mercaptopropane 1-sulfonate and cysteine, for successful thiazolidine cleavage.[27, 222] None of the probed conditions yielded the desired aldehyde protein or protein adduct (outcome depending on applied additive).

As a conclusion, we can say that the thiazolidine formation seems a good and milder alternative to the previously applied oxime ligation with a very stable product due to the electron deficient aldehyde.

3.3.3. Further Attempts Towards Triple-Modification of Proteins

As we were still interested to also develop a selective triple-modification of a protein, the group of Budisa expressed a new TTL protein variant with the cysteine being newly positioned to another hopefully more solvent-exposed position (see also chapter 3.3.1.2). In addition, a serine was placed at second position from the *N*-terminus by site-directed mutagenesis as shown before in our dual-modification approach. The starting codon for the *N*-terminal amino acid in protein expression is usually a methionine, which in our case is replaced by azidohomoalanine by selective pressure incorporation. *N*-terminal methionine can subsequently be cleaved off by an enzyme called methionine (Met) aminopeptidase to yield an *N*-terminal serine for subsequent glycol cleavage and oxime formation. Unfortunately, Met-aminopeptidase cleavage did not seem to proceed as efficiently on our *N*-terminal azidohomoalanine again leaving us with a protein mixture with and without *N*-terminal azidohomoalanine (see also chapter 3.3.2). Following MALDI analysis, we knew that only about half of our protein material possessed an *N*-terminal serine for efficient modification and the two proteins bore ten and eleven azidohomoalanine residues, respectively. However, as the introduction of a biotin by oxime ligation would eventually enable us to separate those two proteins by purification over an avidin or streptavidin column, we decided to proceed with our triple modification strategy on the protein mixture at hand: (Aha)S-TTL(D221C)[Aha].

3.3.3.1. Thiol-Michael Reaction

We first decided to probe thiol-Michael reaction with our previously synthesised Rhodamine maleimide **77** (see also chapter 3.3.1.2). Initial MS data indicated partial dimer formation of the newly synthesised lipase, which might be due to disulfide formation. Indeed, first attempts in thiol-Michael reaction with 10 eq. maleimide **77** in a phosphate buffer (100 mM, pH 7) were not successful, but MALDI-MS showed mostly the protein dimer. We therefore decided to treat the protein with 5 mM dithiothreitol (DTT) as reducing agent prior to thiol-Michael reaction.[226] Analysis by MALDI-MS showed much higher product formation, which supported our initial assumption of partial disulfide formation. Unfortunately, we could not only observe product formation as the main product but also left-over starting material, undesired double-addition and possible DTT adducts (Figure 3.10A). To address the formation of possible DTT adducts, we next probe different concentrations of DTT (20 µM to 5 mM) as well as two other reducing agents: *β*-mercaptoethanol and glutathione (both 1 mM).[2, 227] Best conversion could thereby be achieved with DTT applied in a 200 µM concentration showing much better conversion of the starting cysteine TTL (Figure 3.10B). However, these results led to much higher undesired

multiple addition of maleimide **77**. As discussed before, this is possibly due to solvent-exposed lysine residues that are known to undergo maleimide addition at higher pH. To successfully decrease the lysines' reactivity, we decided to probe the reaction at slightly lower pH decreasing it from 7.0 to 6.5. Therefore, the TTL was reacted with two and five equivalents of Rhodamine maleimide **77** at pH 6.5 and checked by MALDI-MS (Figure 3.10C). With only two equivalents of **77**, we could almost suppress double addition. However, this led to an overall decrease in protein conversion. Low product formation could also be due to partial reformation of disulfide bonds during the reaction. We therefore decided to probe TCEP as an *in situ* reducing agent[228] as it should not interfere with the thiol-Michael reaction. Two different TCEP concentrations, 200 µM and 1 mM, were tested with best results at lower concentration leading to a surprisingly complete suppression of double addition but no visible increase in overall conversion (Figure 3.10D). Even an increase in maleimide **77** from 2 to 13 eq. in the reaction mixture did not change anything about the ratio of starting protein and product as detected by MALDI-MS. Therefore, we decided to first go along with the other modification strategies.

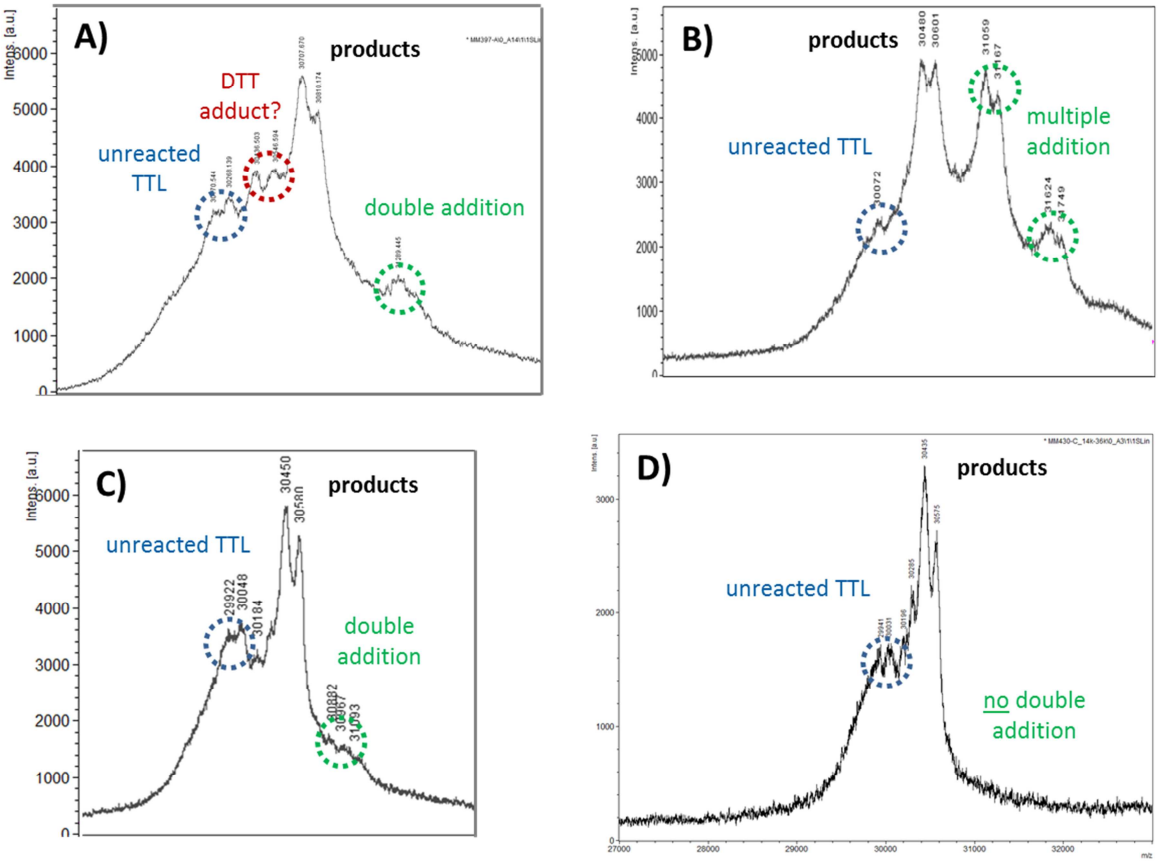


Figure 3.10: MALDI-MS spectra of thiol-Michael reaction with 77 on TTL A) 5 mM DTT, 5 eq. 77, B) 200 µM DTT, 5 eq. 77, C) 200 µM DTT, 2 eq. 77, D) 200 µM TCEP, 5 eq. 77.

3.3.3.2. Copper-Catalysed Azide–Alkyne Cycloaddition

To probe CuAAC as the second modification technique on our new TTL model with butinyle galactose **84**, several reaction conditions were applied varying the amount of Cu^{2+} as well as protein concentration in the reaction mixture (Table 3.7). Results were checked by MALDI-MS, whereby the reactions at 5 µM protein concentration showed no conversion at all. When enhancing the concentration to 10 µM, best conversion in CuAAC could be achieved with 35 mol% Cu^{2+}. An increased concentration of Cu^{2+} led to lower conversion rates again as well as partial precipitation of the protein. Surprisingly, MALDI-MS spectra showed not only starting TTL and product peaks but also two peaks around 16 and 17 kDa. It seemed that the protein partially degraded. To make sure if degradation occurs due to our reaction conditions, we rechecked the starting protein by MS and SDS-PAGE and could detect that about 50% of the protein seemed broken apart. Nevertheless, we decided to further analyse our modified protein by trypsin digestion and MALDI-MS to get an indication which azide residues react predominantly with butinyle galactose **84**. Unfortunately, we could not detect any galactose peptides in our protein mixture. Instead, our results indicated a high instability of our protein in its middle region showing several peptide fragments resulting from inaccurate protein cleavage. Therefore, we decided to temporarily stop working with the present protein but changed over to the better known and more stable GFP (green fluorescent protein), which will be discussed in the following chapter.

Table 3.7: First reaction conditions for CuAAC on newly expressed TTL with alkyne 84 (15 °C, overnight, 100 eq. alkyne 84 (per azide), 5 eq. THPTA (per Cu^{2+}), 1.2 mM aminoguanidine).

entry	c [µM] (protein)	Cu_2SO_4 (to alkyne 84)	sodium ascorbate (to alkyne 84)	product*
1	10	10 mol%	50 eq. (sequential, 4 parts)	2–4 clicked sugars
2	10	35 mol%	50 eq. (at once)	1–2 clicked sugars
3	10	35 mol%	50 eq. (sequential, 4 parts)	4–11 clicked sugars
4	10	50 mol%	50 eq. (sequential, 4 parts)	3–4 clicked sugars
5	5	10 mol%	50 eq. (sequential, 4 parts)	no reaction
6	5	35 mol%	50 eq. (sequential, 4 parts)	no reaction

* estimated by MALDI-MS.

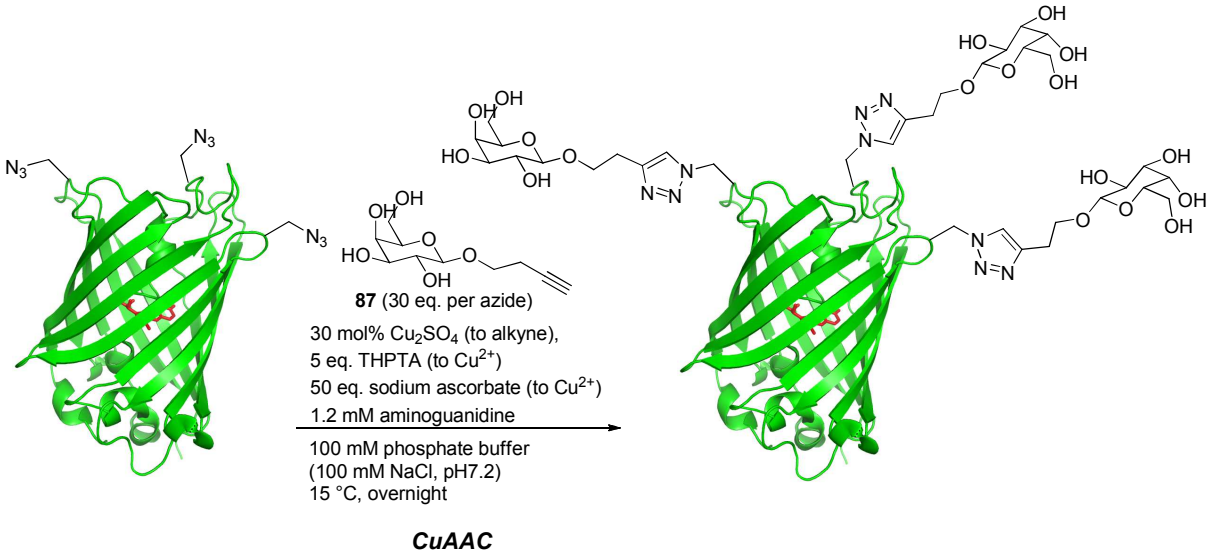
3.3.4. Dual-Modification of Green Fluorescent Protein

Green fluorescent protein (GFP), a protein first described by Osamu Shimomura,[229] emits green fluorescent light upon irradiation with blue or ultraviolet light. During the last decades, GFP has gained much importance in the field of biology as it could be attached to a variety of proteins and used to observe different biological events.[230] In its natural structure, the protein exhibits a

barrel-like conformation formed of eleven *β*-sheets. In its central region, GFP is covalently bound to a chromophore, 4-(*p*-hydroxybenzylidene)imidazolidin-5-one, which is responsible for the protein's fluorescence. On the open sides of the barrel, several short peptide loops are presented, which seem predestined for the introduction of unnatural modifications.[231]

We envisioned that dual-functionalisation of GFP could similar to our previous study with the lipase yield a glycoprotein-scaffold for carbohydrate–protein interaction studies, which can be immobilised on a surface and bears additionally an intrinsic fluorescence to better determine its density on a Streptavidin chip. To do so, we asked the group of Budisa to engineer a GFP with three methionines, one in each of the three external loops of GFP, which were later to be replaced by azidohomoalanines, and with an *N*-terminal serine residue for subsequent glycol cleavage and biotin attachment.

3.3.4.1. Copper-Catalysed Azide–Alkyne Cycloaddition



Scheme 3.35: Copper-catalysed azide–alkyne cycloaddition on S-GFP[Aha] with alkyne galactose 87.

Before we received our desired S-GFP[Aha] model protein for dual-labelling, first CuAACs were probed on a model protein in hand without *N*-terminal serine, GFP[Aha]-His_6 (Scheme 3.35). We decided to start with a rather low amount of of Cu^{2+} (10 mol% per alkyne) to avoid excessive binding to the protein and varied the amount of alkyne sugar **84** applied in the cycloaddition reaction (10, 30 and 50 eq. per azide). All reactions were checked by MALDI-MS. We achieved an addition of up to two galactose moieties with 50 eq. alkyne. By slightly increasing the amount of Cu^{2+} to 30, 50 and 100 mol% while keeping the amount of alkyne at 30 eq., full conversion could be achieved with only 30 mol% Cu^{2+} and 30 equivalents of butinyle galactose **84**.

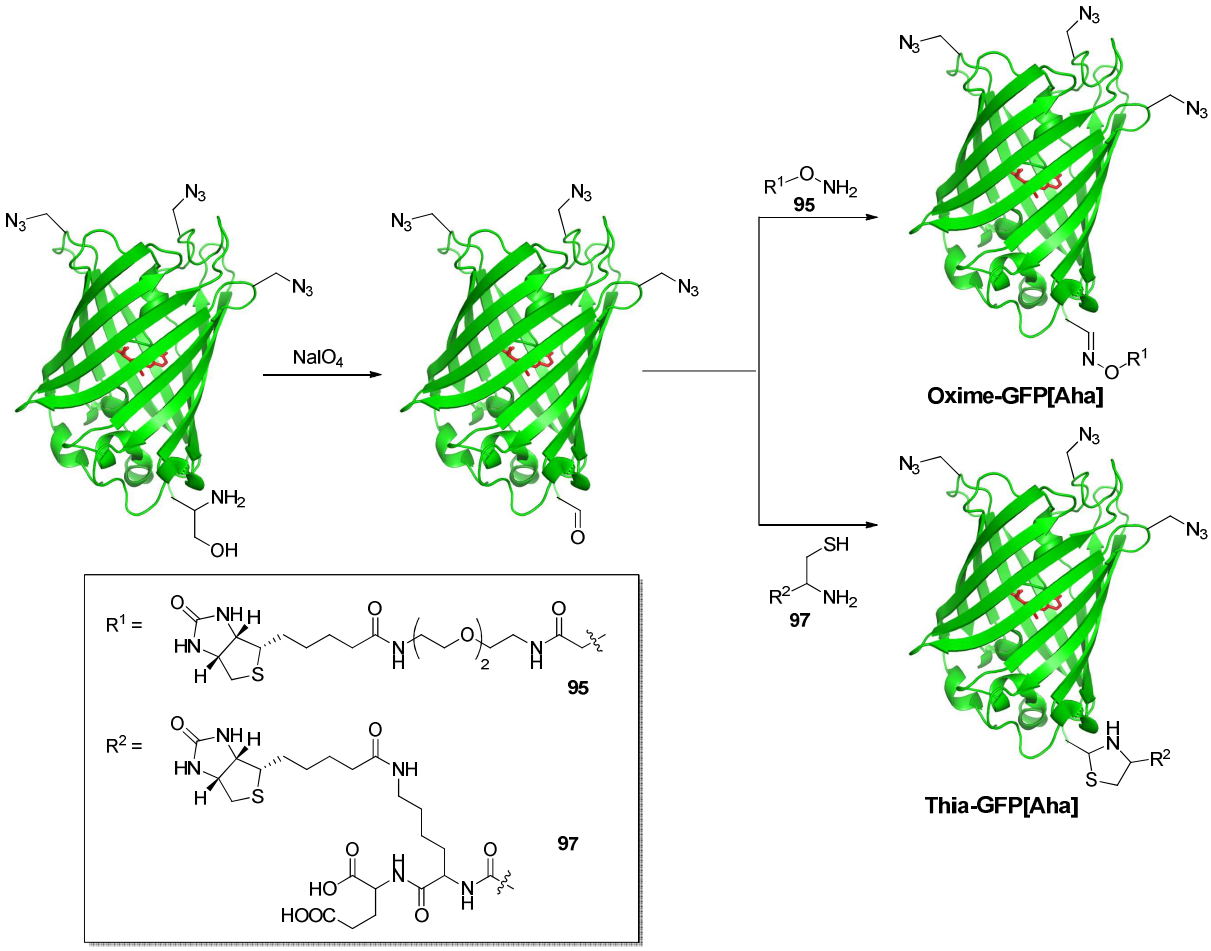
3.3.4.2. Oxime *vs.* Thiazolidine Formation

Once we got our optimised conditions for full conversion in the CuAAC with alkyne sugar **84**, we decided to focus on the attachment of a biotin moiety to the *N*-terminus of GFP. Therefore, we received our new model protein with an *N*-terminal serine, S-GFP[Aha], which we oculd now apply in our double modification approach. As a first step, we decided to probe different conditions for the glycol cleavage as GFP possesses two cysteines in its amino acid sequence and potential oxidation of these could be a limitation to our strategy – especially as one of the cysteine residues is in close proximity and possibly place a crucial role with the fluorophore.

We first probed our standard glycol cleavage conditions, as shown before with our TTL (see chapter 3.3.2.3), with three equivalents sodium periodate in a phosphate buffer (100 mM, pH 7.2) at 15 °C for 30 min.[118b] Immediately after periodate addition, we could observe a complete loss of the typical yellow colour of the protein solution. Purification of the protein mixture and subsequent MALDI-MS analysis showed a slight increase in the protein mass, which might correspond to oxidation of the two cysteine moieties and would explain the loss of colour, which could not be regained. To suppress undesired oxidation, we decided to switch the pH to a value above the cysteines reported pK_a of 8.3.[3] To do so, we performed glycol cleavage at pH 9.0 and to our delight we could observe no change of colour. In addition, the detected mass by MALDI-MS indicated sole aldehyde formation.

In the next step, we decided to probe further modification of our chemically generated aldehyde moiety. To so do, we envisioned to compare two different chemoselective strategies to each other: oxime ligation with our biotin hydroxylamine **95** and thiazolidine formation with newly synthesised biotin cysteine peptide **97** (Scheme 3.36).

Regarding the thiazolidine formation, we planned to carry out the reaction at pH 5.4 as it is known that the reaction proceeds best under these conditions.[26] However, we also probed more neutral conditions with and without the help of aniline as a catalyst according to previous studies on improved oxime formations in the presence of a catalyst (Table 3.8).[156a] By MALDI-MS analysis, we could observe that product formation was 80–90% with 30 eq. of cysteine biotin peptide **97** at pH5.4 (Table 3.8, entry 3). As expected at neutral pH, thiazolidine formation seemed to drop a little (Table 3.8, entry 6). However upon addition of 10 mM aniline as a catalyst, conversion towards the desired product, Thia-GFP[Aha], could be increased to the same results as detected at pH 5.4 (Table 3.8, entry 8). Higher conversions could be achieved with an increased amount of cysteine peptide **97** (Table 3.8, entry 5).



Scheme 3.36: Aldehyde formation on S-GFP[Aha] and subsequent oxime or thiazolidine formation with biotin hydroxylamine 95 and biotin cysteine peptide 97, respectively.

For oxime formation, we also probed the reaction at pH 5.4 and at pH 7 with aniline as a catalyst. In addition, oxime formation was performed at pH 3.0, which should allow higher conversion to our desired oxime protein Oxime-GFP[Aha] (Table 3.8).[200] After purification, all reactions were checked by MALDI-MS. It seemed that the protein itself is not stable at very low pH as the solution turned immediately colourless and no protein material could be detected after conducting the reaction at pH 3.0 (Table 3.8, entry 1). At pH 5.4, oxime formation seemed to be drastically lower than thiazolidine formation (Table 3.8, entry 2–5) with only 30–40% product formation in the presence of 50 eq. biotin hydroxylamine **95**. At neutral pH in the presence of 10 mM aniline (Table 3.8, entry 7), oxime formation seemed to slightly increase to 50%. These results indicate that thiazolidine formation might be a good alternative to the better known oxime formation especially with proteins that are instable at lower pH values.

Table 3.8: Reaction conditions probed for oxime and thiazolidine formation (15 °C, overnight).

entry	hydroxylamine 95	cysteine peptide 97	buffer / catalyst	product formation*
1	30 eq.	–	NH_4OAc (100 mM, pH 3.0)	–
2	30 eq.	–	NH_4OAc (100 mM, pH 5.4)	–
3	–	30 eq.	NH_4OAc (100 mM, pH 5.4)	80–90%
4	50 eq.	–	NH_4OAc (100 mM, pH 5.4)	30–40%
5	–	50 eq.	NH_4OAc (100 mM, pH 5.4)	> 90%
6	–	30 eq.	phosphate buffer (100 mM, 100 mM NaCl, pH 7)	60–70%
7	30 eq.	–	phosphate buffer (100 mM, 100 mM NaCl, pH 7) / 10 mM aniline	50%
8	–	30 eq.	phosphate buffer (100 mM, 100 mM NaCl, pH 7) / 10 mM aniline	80–90%

* Roughly estimated by MALDI-MS in comparison to non-reacted protein. No yields! Only for comparison.

To be absolutely sure that periodate cleavage and thiazolidine formation do not interfere with the protein's fluorescence, we checked our reaction mixture after purification (Table 3.8, entry 5) as well as a sample of unreacted S-GFP[Aha] by UV (250–650 nm) and fluorescence spectroscopy (ex. 491 nm, em. 511 nm) and to our delight could observe no visible loss in fluorescence (Figure 3.11).

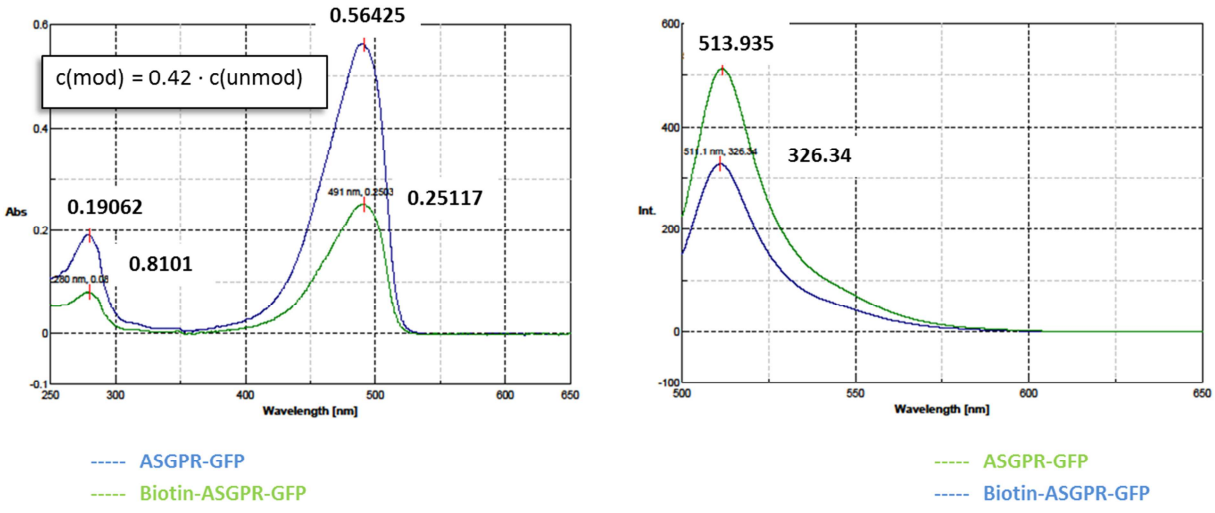


Figure 3.11: UV (left) and fluorescence spectra (right, ex. 491 nm, em. 511 nm) of Thia-GFP[Aha] (Table 3.8, entry 5).

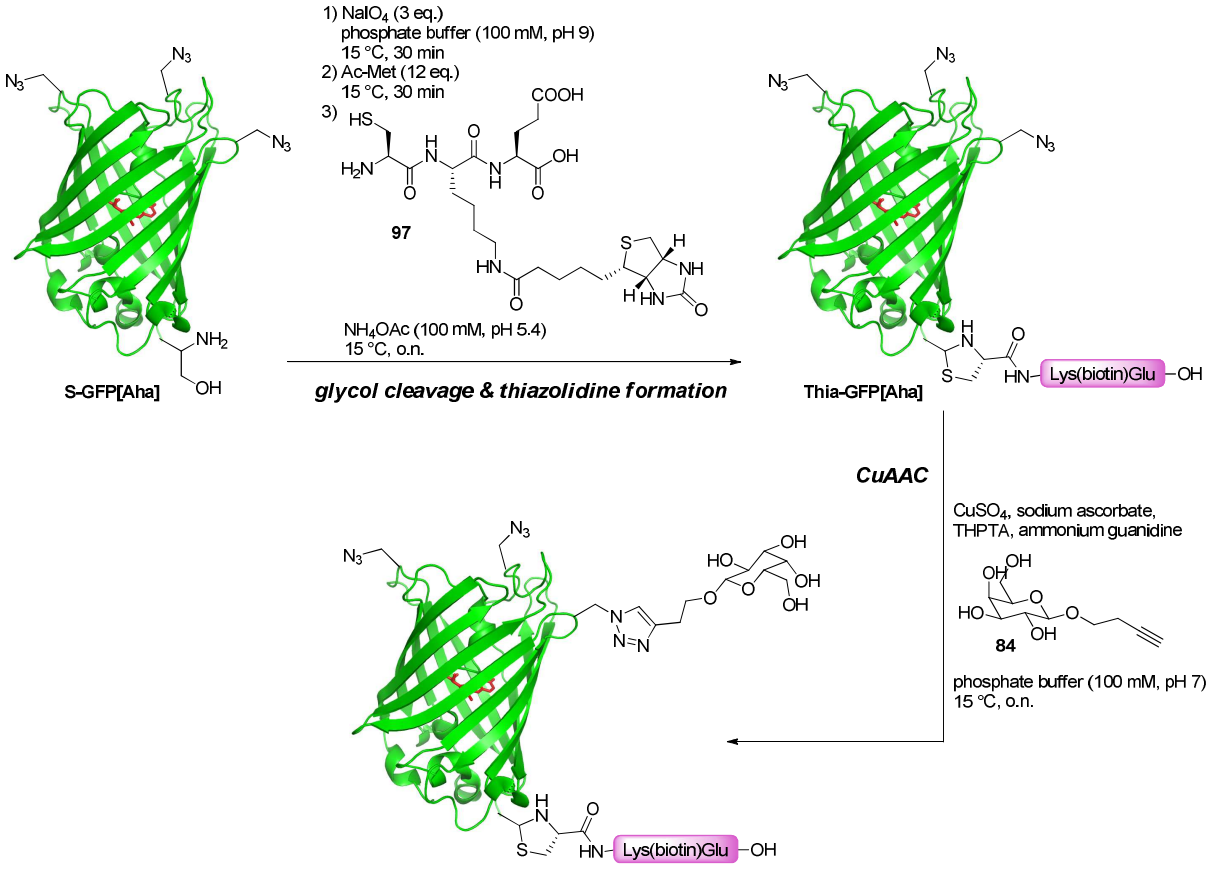
3.3.4.3. Sequential Dual-Modification of GFP

Once we demonstrated that both functionalisation strategies are working individually, we decided to develop a selective dual-functionalisation for our model protein S-GFP[Aha] by subsequent thiazolidine formation and CuAAC.

Therefore, the protein was first reacted with sodium periodate at pH 9 (as described in the previous chapter), quenched with *N*-acetyl methionine and, after buffer exchange to ammonium

acetate (100 mM, pH 5.4), the protein was reacted with biotin cysteine peptide **97** to achieve complete thiazolidine formation to Thia-GFP[Aha] (Scheme 3.37).

For subsequent modification with our butinyle galactose **84** by CuAAC, we decided to first probe the reaction conditions that yielded full conversion with the previous GFP[Aha]-His_6 that did not contain an *N*-terminal serine (see chapter 3.3.4.1). We therefore reacted Thia-GFP[Aha] with 30 eq. of alkyne galactose **84** (per azide) in the presence of 30 mol% Cu^{2+} (per alkyne). Both proteins have three Aha-residues placed in the three loops of GFP though they were slightly repositioned from our first construct GFP[Aha]-His_6 to the newer protein S-GFP[Aha]. This might be one of the reasons that the reaction under these conditions did not yield as expected full functionalisation with three galactose moieties per protein but only one alkyne sugar could be successfully attached to the protein. Another reason could be the additional biotin moiety that we attached to the N-terminus so we decided to again probe different amounts of Cu^{2+} and butinyle galactose **84**. Surprisingly, all reactions yielded only one galactose unit attached to our protein as shown by MALDI-MS analysis.



Scheme 3.37: Sequential dual-functionalisation of S-GFP[Aha] with biotin cysteine 97 in a thiazolidine formation and butinyle galactose 84 in CuAAC.

After dual-modification and purification, we also checked the protein's fluorescent activity (ex. 491 nm, em. 511 nm). All samples showed a strong yellow colour. However, none of the measured fluorescent spectra showed any residual fluorescence in all four probes, which could be due to the multiple buffer exchange or the applied CuAAC conditions. This result shows how important it is to properly check the fluorescent activity and not only rely on the visible colour of the protein sample.

Although, fluorescence could not be retained during initial attempts CuAAC, we successfully achieved in our selective dual-functionalisation of S-Gal[Aha] with one galactose unit attached.

3.4. Synthesis of Pyrrolysine Derivatives for Protein Incorporation by Amber Suppression

Finally in the last part of this thesis, we wanted to synthesise a variety of pyrrolysine derivatives to be applied in the expression of unnatural proteins by amber suppression techniques (Figure 3.12B). Pyrrolysine was first discovered in 2002 as part of methane-producing archaea and is known as the 22nd proteinogenic amino acid (Figure 3.12A).[232] As this naturally occurring amino acid is encoded by the amber stop codon, pyrrolysine's RS/tRNA base pair can be either directly employed or modified to introduce pyrrolysine derivatives *via* the UAG stop codon.[233] In collaboration with the RINA GmbH, the introduction of our unnatural amino acid building blocks during protein synthesis by this technique was probed to achieve the selective incorporation of chemoselective functional moieties.

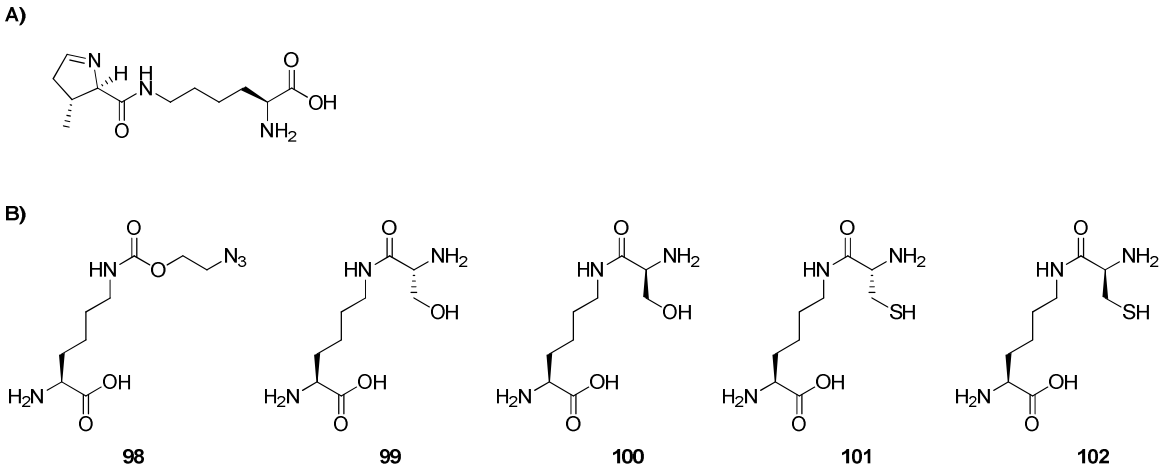
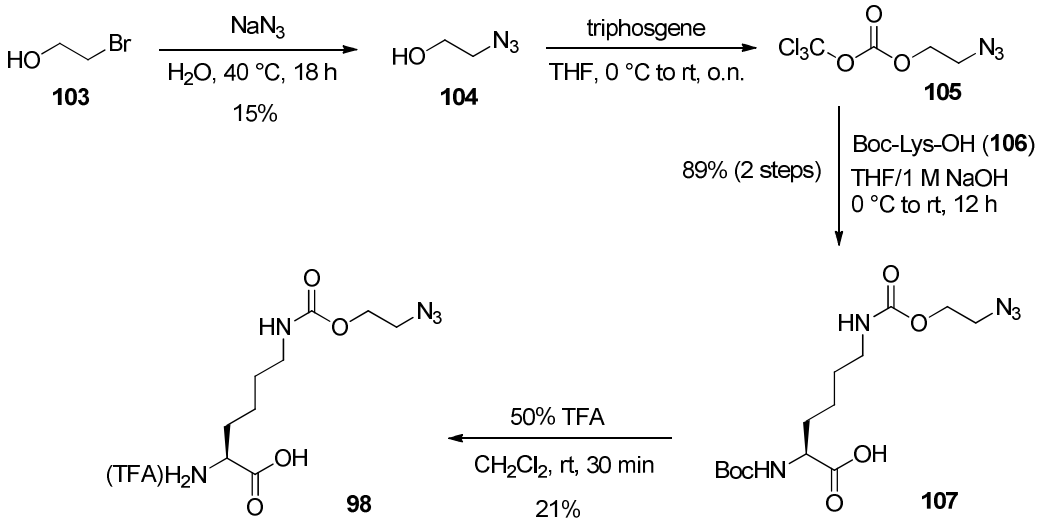


Figure 3.12: A) Structure of Pyrrolysine (Pyl), B) planned pyrrolysine derivatives 98–102.

3.4.1. Azide-Containing Derivative

First, we decided to synthesise azide-containing lysine carbamate **98** (Figure 3.12B),[135e] which has been successfully incorporated into proteins before and was applied by the RINA GmbH to improve their PylRS/tRNA pair. We synthesised *S*-2-Amino-6-((2-azidoethoxy)carbonylamino)-hexanoic acid **98** following the previously published protocol by the group of Chin (Scheme 3.38).[135e] Therefore, 2-azidoethanol (**104**) was prepared in 15% yield from 2-bromoethanol (**103**) and afterwards reacted with triphosgene to yield 2-azidoethyl trichloromethyl carbonate (**105**). After concentration, carbonate **105** was directly reacted with Boc-L-Lys-OH (**106**) to yield the desired Boc-protected product **107** in 89% yield over two steps. After Boc-deprotection, compound **98** was purified by reversed phase HPLC yielding a white powder in 21%.

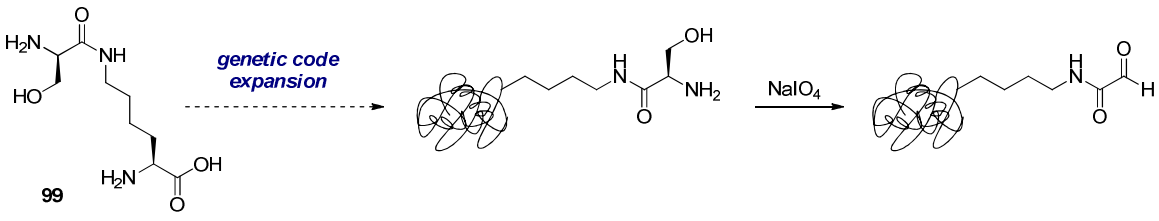


Scheme 3.38: Synthesis of H_2N-L-Lys(C(O)$OCH_2CH_2N_3$)-OH (98).

First experiments on the incorporation of H_2N-L-Lys(C(O)$OCH_2CH_2N_3$)-OH (**98**) by amber suppression during protein expression of Calmodulin with the method developed by the RINA company showed very promising results and the azide could be tagged and visualised with a phosphine attached to the fluorophore Alexa 650 in a Staudinger ligation.

3.4.2. "Pre-Aldehyde" Amino Acid Building Blocks

In the next step, we envisioned that the implementation of the serine derivative **99** would allow – wherever placed – selective aldehyde formation in a protein by glycol cleavage with sodium periodate as shown before in our dual-modification strategy of TTL with *N*-terminal serine (chapter 3.3.2.3).[118b] Direct incorporation of selective aldehyde tags has not been shown to date as the aldehyde itself is too reactive during protein expression and post-translational aldehyde formations on proteins are usually restricted to the *N*-terminus.[116, 118a, 160b, 234] Our strategy would be a major improvement as it would allow the post-translational introduction of aldehydes on any desired position of a protein (Scheme 3.39).



Scheme 3.39: New Pyrrolysine derivative 99 for potential posttranslational aldehyde formation in a protein.

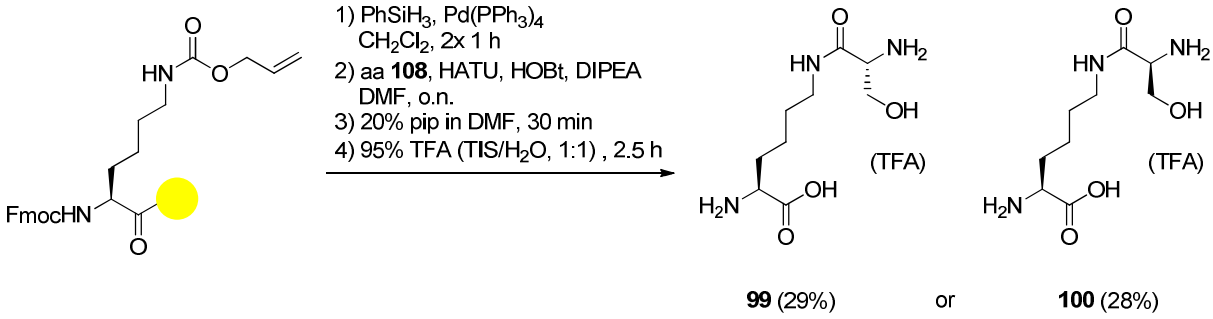
Therefore, we intended to synthesise H_2N-L-Lys(D-Ser)-OH (**99**) as a potential pyrrolysine surrogate following previously published results reporting the successful incorporation of

cysteine derivative **101**,[235] which we decided to synthesise as a potential positive control. Taking into account that the stereocenter on the serine might have an influence on the incorporation efficiency, we planned to also synthesise diastereomers H_2N-L-Lys(L-Ser)-OH (**100**) and H_2N-L-Lys(L-Cys)-OH (**102**).

For the synthesis of the four dipeptides **99–102** an in-solution approach was chosen in the beginning. For the synthesis of H_2N-L-Lys(D-Ser)-OH (**99**), Boc-L-Lys-OMe was reacted with Fmoc-D-Ser(O^{t}Bu)-OH (**D-108**) using EDC·HCl and HOBt as coupling reagents. After stirring the reaction mixture overnight, the formed product seems to have lost the Fmoc protecting group and racemised at C_α of either serine or lysine. As it was not possible to assign the position of the racemised stereocenter by comparison of ^{1}H- and ^{13}C-NMR spectra, we abandoned the in-solution approach for our dipeptides.

For further dipeptide synthesis, a solid phase approach was chosen. To enable selective deprotection of the lysine side chain and coupling of Fmoc-L-Ser(O^{t}Bu)-OH (**L-108**), we decided to probe two different protecting groups: Alloc (allylcarbonyl) and ivDde (1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)-3-methylbutyl).[236] For the first path, a Wang resin (0.7 mmol/g) was functionalised with Fmoc-L-Lys(Alloc)-OH (Scheme 3.40).

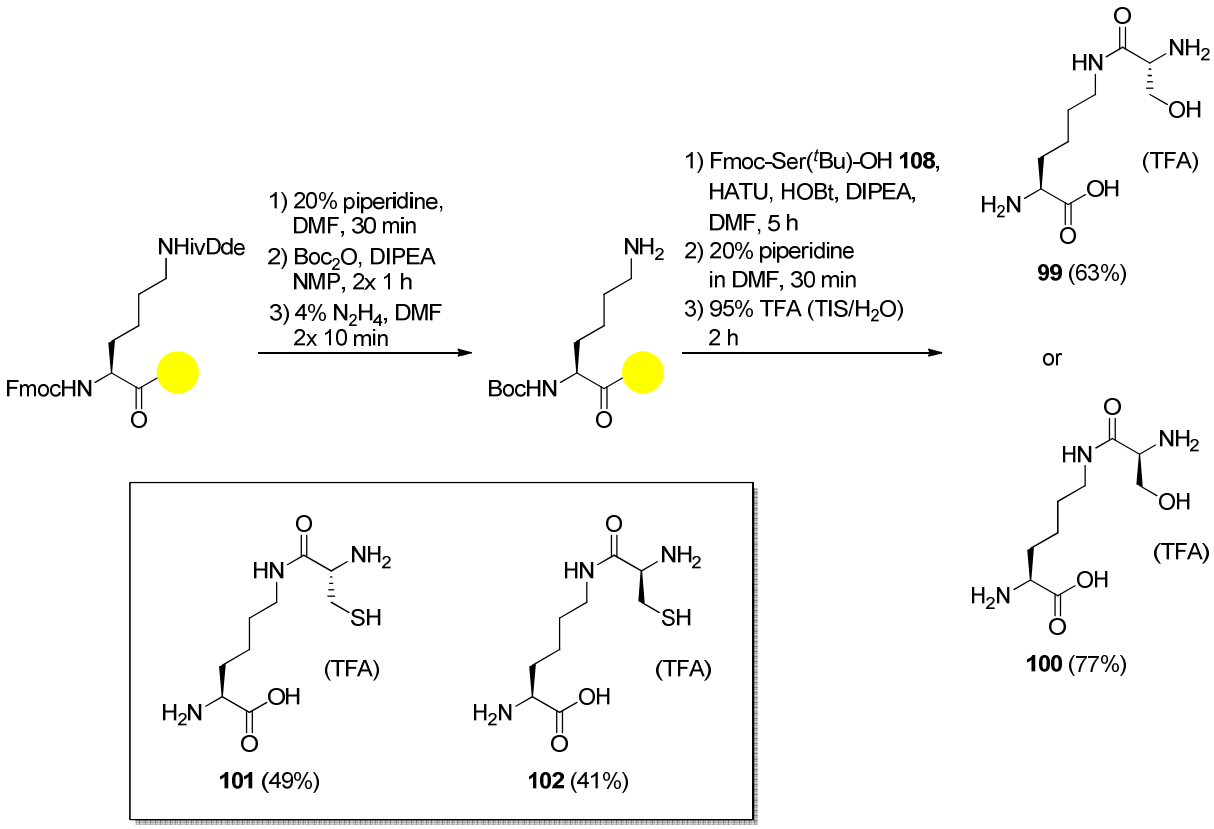
After selective deprotection of Alloc under standard conditions,[236a] **L-108** was coupled to the free lysine side chain. Final deprotection and cleavage from the resin afforded dipeptide **100** in an overall yield of 28%. In the same manner, diastereomeric dipeptide **99** was synthesised in 29% yield.



Scheme 3.40: Solid phase dipeptide synthesis of 99 and 100 with orthogonal Alloc deprotection.

As an alternative protecting group for the lysine side chain, we simultaneously probed ivDde (Scheme 3.41). We started from preloaded Fmoc-Lys(ivDde)-Wang resin. As a first step, the Fmoc-protecting group was exchanged to a Boc-group as Fmoc itself is not stable to ivDde deprotection conditions. Afterwards, ivDde was cleaved off with 4% hydrazine[236b] in DMF and **D-108** or **L-108** was coupled to the free lysine side chain. After Fmoc-deprotection with piperidine,

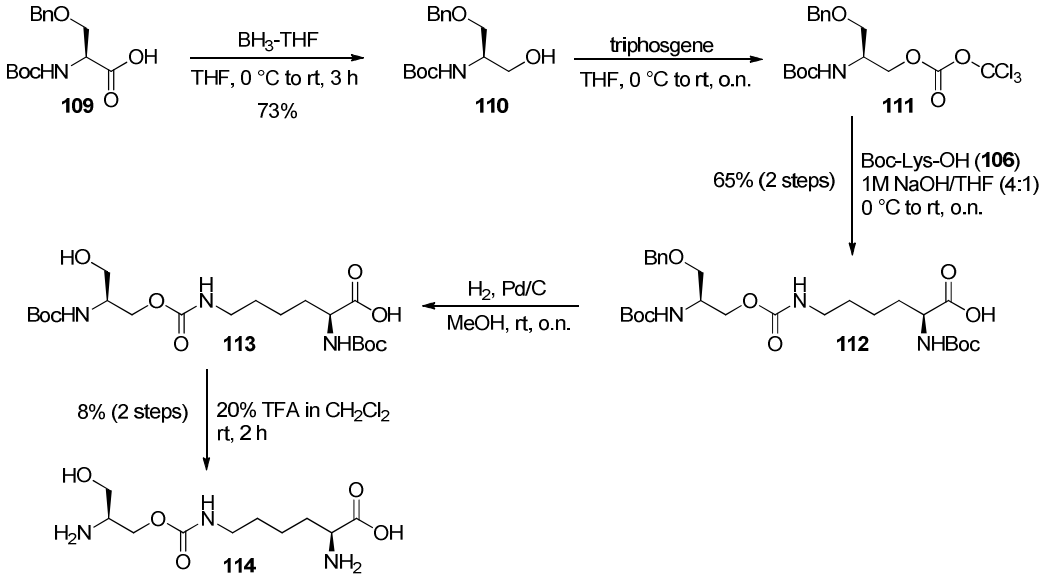
the peptide was cleaved off the resin and purified by HPLC to yield our desired dipeptides **99** and **100** in 63% and 77% yield, respectively. Due to the much higher yields with ivDde as a protecting group, we decided to also synthesise our cysteine containing dipeptides **101** and **102** by the second route yielding 49% and 41% of each, respectively (Scheme 3.41).



Scheme 3.41: Synthesis of dipeptides 99–102 on solid support starting from Fmoc-L-Lys(ivDde)-Wang resin.

All four substrates were probed in protein expressions by pyrrolysine amber suppression by the RINA GmbH, but unfortunately no successful incorporation could be detected so far under the probed conditions.

One possibility for the poor incorporation could be a lack in proton donors next to the lysine side chain functionality as compared to lysine carbamate **99**. We therefore decided to synthesise lysine carbamate **114** (Scheme 3.42), which might allow better incorporation with the potential for subsequent periodate cleavage and aldehyde formation on protein level. The synthesis was started with Boc-L-Ser(Bn)-OH (**109**). We first reduced the carboxylic acid in **109** to an alcohol **110**,[237] which was then reacted with triphosgene to yield carbonate **111**. The free side chain amine of Boc-L-Lys-OH (**106**) was then coupled to carbonate **111** to form the desired protected carbamate **112**. Final benzyl-deprotection proofed to be rather inefficient, which led after Boc-deprotection to only 8% of lysine carbamate **114** as a TFA salt over two steps.



Scheme 3.42: Synthesis of lysine carbamate 114.

Although initial test by the RINA GmbH on amino acid incorporation during protein expression were not successful, possible improvements might be achieved by changing the tRNA/RNA base pair or improve the building block's suitability by, e.g., oxazolidine formation on the serine side chain to reduce the hydrophilicity of our unnatural amino acid.

4. Summary and Outlook

In this thesis, we were engaged in the development of different modification strategies, which were applied for the selective functionalisation of peptides and proteins in aqueous media. The various methods were probed for their chemoselectivity, reagent scope and applicability in buffered aqueous systems. Thereby, the presented work can be sub-divided into three main projects: the selective acetylation of azido peptides (chapter 3.1 and 3.2), the chemoselective dual-functionalisation of proteins (chapter 3.3) and the synthesis of new pyrrolysine derivatives for protein incorporation (chapter 3.4).

4.1. Project 1 – Selective Peptide Acetylation in Aqueous Media

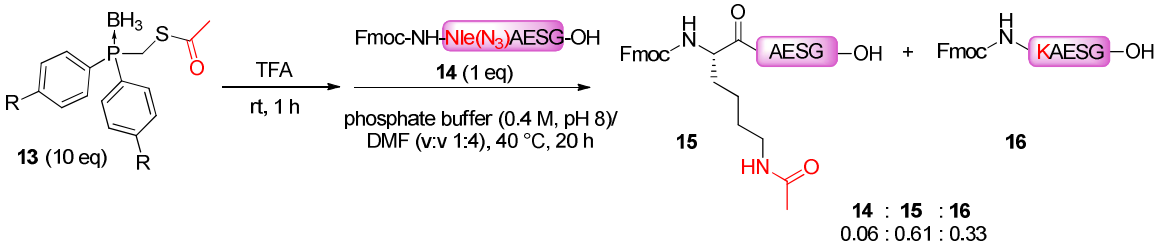
Within the first project, we focused on the development of selective acetylation techniques for alkyl azido peptides in aqueous media. The precise and efficient introduction of acetylations on lysine side chains at distinct positions throughout a peptide or protein would allow more thorough studies of the influence and effect of those posttranslational modifications, which play an important role in the field of epigenetics. We therefore applied two individual ligation strategies – either the traceless Staudinger ligation or the thioacid–azide reaction – on alkyl azido peptides to study the scope of the individual reactions in aqueous media.

4.1.1. Traceless Staudinger Ligation

In terms of the traceless Staudinger ligation, we first conducted a study on the commercially available borane-protected diphenylphosphinomethanethiol (**3**) as an appropriate phosphine linker for the traceless Staudinger ligation. Upon borane-protection, the phosphine preserves a high stability towards oxidation. The Borane group can be selectively removed to trigger the traceless Staudinger ligation with the free phosphine and different azido molecules such as peptides and carbohydrates. For borane-deprotection, we compared two previously published strategies, acidic and basic, with each other to demonstrate their individual advantages and disadvantages. It was shown that the basic deprotection with DABCO proceeds best in organic solvents at elevated temperatures and can be applied in an *in situ* protocol in the presence of the subsequently reacting azide to give good yields in the traceless Staudinger ligation. However, these deprotection conditions are too harsh for the application in biological systems under physiological conditions. In contrast, the second deprotection approach under acidic conditions with neat TFA proceeds well at room temperature not only delivering the free phosphine. In addition, the TFA treatment also removes all acid labile side chain protecting groups on

potentially employed peptides to further explore the reaction's chemoselectivity on biomolecules. In contrast to the typically applied organic solvents in the basic deprotection approach such as DMF, toluene or benzene, TFA can be easily removed under vacuum though residual TFA has to be quenched afterwards by addition of a strong buffer or several equivalents of a base to avoid the competing Staudinger reduction. In subsequent studies, we unfortunately observed that the traceless Staudinger ligation is not fully chemoselective under physiological conditions. Instead, we noticed that basic amino acid side chains such as Lys can attack the activated phopshinothioester under these conditions as shown in cyclisation studies with small peptides as well as in first reactions with fully unprotected *ε*-lysine peptides.

Nevertheless, we decided to further investigate the scope of the traceless Staudinger ligation. We probed the two different phosphines – commercially available phosphine **1** in comparison with more water-soluble phosphine **13** first published by Raines – in selective acetylation studies with water-soluble azides in aqueous buffered systems in the absence of free amine groups. Two different azido compounds were employed, a water-soluble benzyl azide and more complex *ε*-azido norleucine containing peptide **14**, and we could show that though phopshine **1** exhibited superior reactivity under oxygen-free conditions in a buffer/DMF mixture (1:1) with increasing amounts of buffer only the water-soluble phosphine **13** could yield reasonable amounts of product formation exceeding even previous results with hydrophobic phosphine **1**. This effect proofed even stronger in the reaction with azido lysine peptide **13**, where only water-soluble phosphine **13** was able to deliver selective lysine side chain acetylation (Scheme 4.1). These results support a possibly broad applicability of water-soluble phosphine boranes as a ligation reagent in selective peptide transformations in aqueous media with the limitation of a necessary oxygen-free environment.



Scheme 4.1: Best solvent system and reaction conditions for traceless Staudinger ligation with water-soluble phosphine 13 and azido lysine peptide 14.

To improve the applicability of phosphines for the traceless Staudinger ligation in aqueous solvent systems, the development of new water-soluble phosphine linkers with an even better water-solubility and higher stability towards air might allow their application in not only peptide

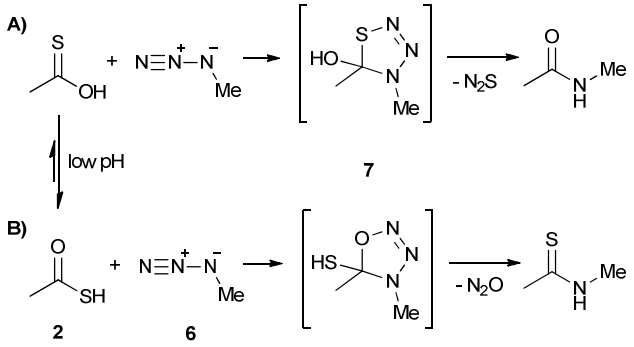
but also selective protein modifications/ligations. In addition, one might be able to improve the chemoselectivity of the reaction by applying an oxoester instead of the more reactive thioester. Though the oxoester might slow down the intramolecular rearrangement from the aza-ylide, its lower reactivity might additionally restrict unselective reactions with basic amino acid side chains.

4.1.2. Thioacid–Azide Reaction

In this approach, we planned to apply the thioacid–azide reaction for the selective acetylation of alkyl azido peptides. In this reaction, a thioacid reacts with an azide to form a stable amide bond upon release of N_2S. The thioacid–azide reaction is usually applied with very reactive electron-poor azides such as sulfonyl azides whereupon chemoselectivity can be achieved by kinetic control due to the rather fast reactivity of the employed azide. To transfer this reaction for the formation of native amide bonds on peptides instead of the otherwise formed sulfonamides, we decided to perform the reaction with less reactive electron-rich alkyl azides and probe its applicability for selective amide bond formations in aqueous media.

We performed first test reactions on the thiaocid–azide reaction with alkyl azides and could show that the reaction is not chemoselective under basic or neutral reaction conditions, which were reported to promote the reaction when sulfonyl azides are employed. It seemed that under these conditions free amino groups can directly react with the thioacid present and thereby lead to unselective amide bond formations. Therefore, we decided to probe different buffered solvent systems and pH values to suppress this undesired side reaction and enhance the probability for selective reaction with the azide. Two different alkyl azides were employed in this study – an electron-rich γ-azido butanoic acid peptide and a modestly electron-poor α-azido glycine peptide. During this work we observed the unexpected formation of thioamides as a side-product instead of amides, which increased with decreasing pH. We could show that it is possible to control the ratio of thioamide and amide formation especially with electron poor azides by pH. Here, we observed almost full conversion towards thioamides at pH < 3 and towards amides at pH > 5. In addition, we could show that the reaction could be driven towards very high conversion rates by simply enhancing the concentration of the reaction while the thioamide/amide ratio did not change. In addition, based on previous results of the group of Prof. Rademann on thioamide formation by the metal catalysed thioacid–azide reaction, we proposed a mechanism for our thioamide formation at low pH, which is solely dependent on the protonation state of the thioacid (Scheme 4.2). When the pH is lower than the pKa of the applied thioacid, the thioacid exists predominantly with a stable C=O instead of a C=S bond, which then

reacts with the azide in a [3+2]-cycloaddition towards an oxatriazolidine and yields the thioamide upon loss of N_2O.



Scheme 4.2: Proposed mechanism for thioacid–azide reaction with alkyl azides under acidic conditions: A) by Williams, B) by our group.

As a next step, it would be interesting to further investigate the influence of the electronic nature of the employed azide towards electron-poor azides but also of the thioacid in the thioamide formation to further maximise the formation of only one of the two products. In addition, one should probe if at higher concentrations where the reaction seems to proceed faster and yield higher conversion rates the thioacid–azide reaction might be possibly applied in a chemoselective fashion again when only little thioacid is applied to suppress undesired reactions with free amines.

4.2. Project 2 – Chemoselective Protein Modifications

Within the second project, we envisioned to combine different orthogonal ligation strategies for the selective dual- or triple modification of two proteins, a thermostable lipase and GFP. This strategy should enable us to synthesise carbohydrate–protein scaffolds with an additional biotin moiety for immobilisation on a streptavidin-coated chip for subsequent SPR studies.

We were able to develop a very effective strategy to introduce two different unnatural functional moieties to a thermostable lipase (TTL). Together with the group of Prof. Budisa, we combined co- and posttranslational methods to introduce two orthogonal handles, azides and an aldehyde, by auxotroph supplementation based incorporation and by glycol cleavage of an *N*-terminal serine, respectively. This strategy enabled us to then selectively apply two orthogonal ligation techniques for specific protein labelling: the CuAAC for azide functionalisation with alkyne galactose moieties and the oxime ligation for aldehyde modification with a biotin hydroxylamine (Figure 4.1). In a final step, three biotinylated proteins, which were functionalised to different degrees with galactose moieties, were immobilised on streptavidin coated chips.

Subsequent lectin binding studies by surface plasmon resonance measurements with ECL could show a increased binding propensities for our protein conjugates to the applied lectin and K_D values could be determined.

To achieve better or maybe even bivalent binding to ECL of our different protein scaffolds, the introduction of a longer and more flexible galactose linker might be of help. In addition, repositioning of some of the azidohomoalanine residues might allow yet a higher degree of galactose functionalisation on our protein model.

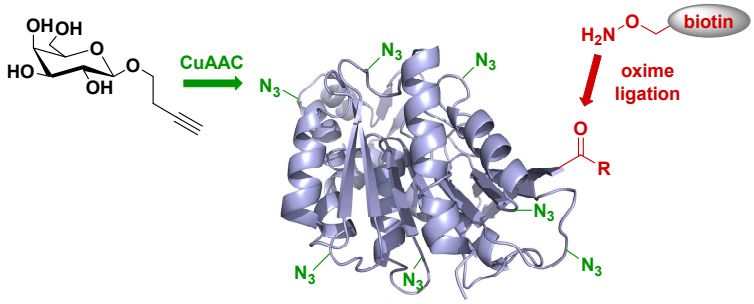


Figure 4.1: Dual-functionalisation approach by CuAAC and oxime ligation on our TTL model protein.

In a side project, we compared our applied oxime ligation with a second modification technique, namely the thiazolidine formation and could observe that thiazolidine formations work at less acidic pH and therefore under milder reaction conditions for many proteins more efficiently than oxime ligations. In addition, we envisioned that the previously formed thiazolidine could be cleaved in analogy to its application in sequential NCL experiments. That way, we could re-engineer the aldehyde after possible purification *via* our biotin and modify it again with something else. Unfortunately, we were not able to cleave the thiazolidine again as previously envisioned. We suspect that the thiazolidine is too stable, which is probably due to the high electrophilicity of our employed aldehyde. Nevertheless, we could show that thiazolidine formation in general is more efficient than oxime formation at higher pH and therefore allows milder protein modification resulting in a very stable functionalisation.

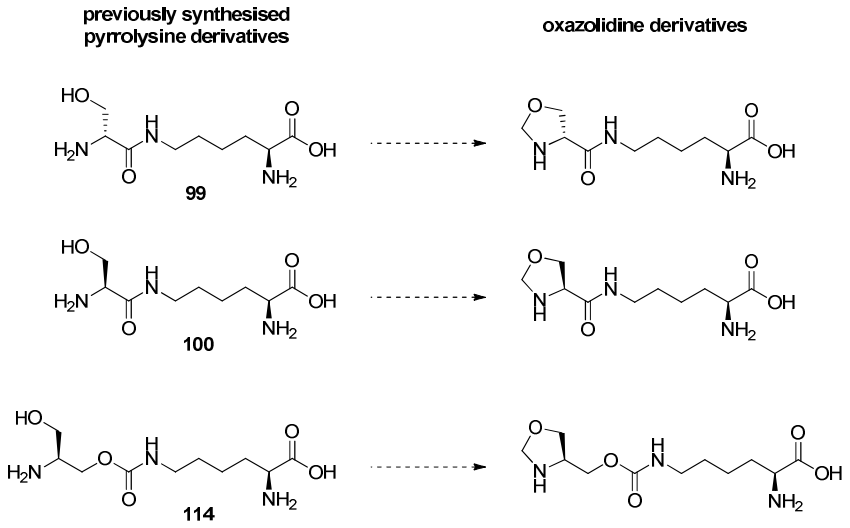
In a second approach, we were able to apply our dual-modification strategy to a second protein, GFP, which was expressed by the group of Prof. Budisa, with an N-terminal serine and three different azidohomoalanine residues present in the three outer loops of the protein. We successfully employed the GFP protein, in a glycol cleavage followed by thiazolidine formation on the *N*-terminus with a biotin cysteine building block and could show that the inherent fluorescence was still preserved after successful biotinylation. In the final step, we probed the functionalisation of the biotinylated GFP by CuAAC with alkyne galactose but were so far only

able to introduce one out of three galactose moieties. In addition, fluorescence measurements revealed a loss of fluorescence of the GFP after CuAAC.

To go on with this project, it is necessary to scan more reaction conditions for the CuAAC to achieve full functionalisation and to figure out the responsible factors for the detected loss of fluorescence. The dual-functionalisation approach on a protein scaffold with intrinsic fluorescence would allow an additional fluorescence read-out in possible biological applications of our model proteins.

4.3. Project 3 – Synthesis of Pyrrolysine Derivatives for Protein Incorporation Applying Amber Suppression Techniques

Within the third project, we decided to synthesise new pyrrolysine derivatives, which would allow the posttranslational formation of aldehydes on a distinct position of a protein by glycol cleavage. Therefore, an amino glycol unit had to be introduced into the side chain of a lysine.



Scheme 4.3: During this thesis synthesised pyrrolysine derivatives 99, 100 and 114 for post-translational aldehyde formation (left) and their corresponding oxazolidine forms (right).

We could successfully synthesise two lysine derivatives **99** and **100** with either a D- or an L-serine attached to its side chain by amide formation applying solid phase peptide synthesis and different orthogonal protection strategies (Scheme 4.3). In addition, we synthesised pyrrolysine derivative **114** with a carbamate linkage between a glycerol-like building block and a lysine side chain (Scheme 4.3). Incorporation of these amino acid building blocks in an *in vitro* expression system was not successful to date. It is possible that they are compared to the natural pyrrolysine too hydrophilic due to their free amino and hydroxy groups and thereby hamper successful incorporation. The protection as oxazolidine moieties with simple formaldehyde

might help to enhance their propensity for effective introduction during protein expression (Scheme 4.3). In addition, incorporation might be more efficient with a mutated PylRS/tRNA base pair.

5. Experimental

5.1. Materials and Methods

Materials: Solvents such as ethyl acetate, hexane and CH_2Cl_2 were purchased as p.a. grade and distilled once prior to use. Reagents – including deuterated as well as dry solvents and acetonitrile – are commercially available as reagent grade and did not require further purification. The resins as well as Fmoc-protected natural L-amino acids were purchased from Novabiochem or Iris Biotech. Thin-layer chromatography (TLC) was performed with precoated silica gel plates and visualized by UV light (λ = 254 nm) or $KMnO_4$ solution. The reaction mixtures were purified by column chromatography over silica gel (60−240 mesh).

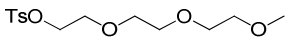
Methods:

NMR spectra were recorded on Bruker (300 MHz, 600 MHz) or on *Jeol ECX-400* or *-500* (400 MHz, 500 MHz) spectrometers at room temperature. Chemical shifts (δ) are reported in ppm relative to residual solvent peak ($CDCl_3$: 7.26 (^{1}H) and 77.16 (^{13}C), D_2O: 4.79 (^{1}H), d_6-DMSO: 2.50 (^{1}H) and 39.52 (^{13}C), d_4-MeOD 3.31 (^{1}H) and 49.00 (^{13}C)). The analytical HPLC-MS was applied from WatersTM with a 717plus autosampler, a 600S controller, 2 pumps 616 and a 2489 UV/Visible detector connected to a 3100 mass detector (low resolution, all Waters Corporation, Milford, Massachusetts, USA). The RP-HPLC-column was a Kromasil C18 (5µm, 250x4.6mm with a flow rate of 0.8 mL/min, MeCN/H_2O (0.1% TFA)). For fluorescence LC-UV a Jasco FP-2020Plus fluorescence detector (JASCO Inc., Easton, Maryland,USA) was attached used with the analytical HPLC. Data was processed by Empower Pro software (Waters Corporation). Semi-preparative HPLC purification of the peptides was performed on a JASCO LC-2000 Plus system using a reversed phase C18 column (5 lm, 25 x 250 mm, constant flow of 16.0 mL/min), consisting of a Smartline Manager 5000 with interface module, two Smartline Pump 1000 HPLC pumps, a 6-port-3-channel injection valve with 2.5 mL loop, a UV detector (UV-2077) and a high pressure gradient mixer. Preparative HPLC purification was performed on a Gilson® PLC 2020 Personal Purification System using a Nucleodur VP 250/32 C18 HTec 5 µm column from Macherey-Nagel at a flow rate of 35 mL/min. High-resolution mass spectra (HRMS) were collected with an Agilent 6210 ToF LC/MS system (Agilent Technologies, Santa Clara, California, USA) using as an eluent water and acetonitrile in a 1:1 mixture (with 0.1% TFA) at a flow rate of 0.2 mL/min or for LCMS analysis H_2O and MeCN (both including 1% acetic acid) were used as eluent (0.5 mL).

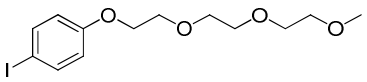
Peptide Synthesis:

Peptides were synthesised with an ABI 433A Peptide Synthesiser (Applied Biosystems, Inc., Foster City, California, USA) *via* standard Fmoc-based conditions (Fast-moc protocol with HOBt/HBTU conditions) on a preloaded Fmoc-Gly-Wang resin (0.79 mmol/g). Non-canonical amino acids (2 eq.) were coupled manually with HATU (2 eq.), HOBt (2 eq.) and DIPEA (2 eq.) overnight.

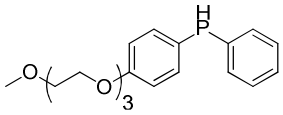
5.2. Synthesis of Phosphine Derivative 5



2-(2-(2-Methoxyethoxy)ethoxy)ethyl-4-methylbenzenesulfonate (18). Tosylchloride (67.0 mmol, 12.8 g), triethylene glycol monomethyl ether (**17**) (67.0 mmol, 11.0 g) and triethylamine (134 mmol, 13.6 g) were dissolved in 200 mL acetonitrile. Trimethylamine hydrochloride (67.0 mmol, 6.40 g) was added dropwise at 0 °C. The reaction mixture was allowed to warm to room temperature and stirred for 2 h. After filtration and concentration the residue was extracted with 100 mL 1 N HCl and 100 mL toluene in a separation funnel. The organic phase was washed with 50 mL 1 N HCl and afterwards the aqueous phases were combined and washed with 100 mL toluene. The combined organic phases were dried over $MgSO_4$ and concentrated. The desired product **18** was obtained in 99% yield (21.1 g, 66.3 mmol) as pale yellow oil: R_f 0.20 (3:1 hex:EtOAc); $\delta_H(CDCl_3$; 250 MHz): δ = 7.79 (d, *J = 8.3 Hz*, 2H), 7.33 (d, *J = 8.3 Hz*, 2H), 4.15 (t, *J = 4.5 Hz*, 2H), 3.70-3.51 (m, 10H), 3.36 (s, 3H), 2.44 ppm (s, 3H). The spectroscopic data was consistent with the literature.[238]



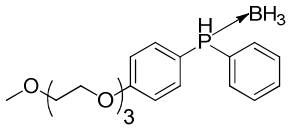
1-(2-(2-(2-Methoxyethoxy)ethoxy)ethoxy)-4-iodobenzene (20). *p*-Iodophenol (**19**) (7.50 mmol, 1.66 g) and K_2CO_3 (14.4 mmol, 2.00 g) were mixed in a flask and dried under vacuum. 150 ml acetonitrile and tosylate **18** (6.30 mmol, 2.00 g) were added and the reaction mixture refluxed at 90 °C and stirred overnight. The suspension was allowed to cool down to room temperature, filtered and concentrated to yield the crude product. Purification with flash column chromatography (cyclohexene : ethylacetate 4:1; $R_f \approx 0.25$) gave the pure product **20** in 95% yield (5.99 mmol, 2.19 g) as a yellowish oil: R_f 0.25 (3:1 hex:EtOAc); $\delta_H(CDCl_3$; 250 MHz): δ = 7.54 (d, *J = 9.0 Hz*, 2H), 6.69 (d, *J = 9.0 Hz*, 2H), 4.09 (bt, 2H), 3.84 (bt, 2H), 3.75-3.50 (m, 8H), 3.37 ppm (s, 3H); $\delta_C(CDCl_3$; 100 MHz): δ = 158.97, 138.44, 117.35, 83.20, 72.20, 71.13, 70.93, 70.85, 69.89, 67.82, 59.33 ppm. The spectroscopic data was consistent with the literature.[239]



4-[1-(2-(2-(2-Methoxyethoxy)ethoxy)ethoxy]-phenyl phenylphosphine (21). Pegylated iodophenol **20** (0.682 mmol, 250 mg) was dissolved in 5 ml solvent (Table 6.1). After addition of triethylamine (1.37 mmol, 139 mg), $Pd(PPh_3)_4$ (2 mol%, 15.8 mg), 1,3-bis(diphenylphosphino)propane (4 mol%, 11.3 mg) and phenyl-phosphine (0.682 mmol, 75.2 mg), the reaction mixture was degassed (freeze-pump-thaw). The reaction was heated up to a certain temperature (Table 5.1) and stirred overnight. Conversion was checked by ^{31}P-NMR: δ_P(162 MHz): δ = -123 (phenylphosphine), -42 (product **21**), 20 ppm (oxidised product). m/z: $[M+H]^+$ (monoisotopic) 365.1512; m/z: $[M+H]^+$ (ESI-ToF) 365.1468.

Table 5.1: Different reaction conditions for preparation of phosphine 21.

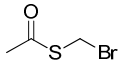
#	Solvent	Temperature [°C]	δ_P (400 MHz) -42 ppm	20 ppm
1	MeCN	90	×	(×)
2	THF	70	–	×
3	DMF	120	×	(×)
4	DMSO	120	–	×



4-[1-(2-(2-(2-Methoxyethoxy)ethoxy)ethoxy]-phenyl phenylphosphine borane (22). To the reaction mixture of **21** was directly added a 1 M solution of a borane complex (6.82 mmol; Table 5.2). A change of colour towards orange-red could be observed. After 4 h conversion was controlled with ^{31}P-NMR by direct measurement of the crude reaction mixture: ^{31}P-NMR (162 MHz): δ = -1.82 ppm (d, *J* = 55.8 Hz).

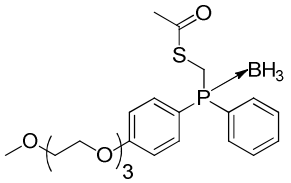
Table 5.2: Different reaction conditions for preparation of phosphine borane complex 22.

#	solvent	borane-complex (1M) BH_3·THF	BH_3·DMS
1	MeCN	no product	product
2	DMF	n.d.	product



***S*-Bromomethyl ethanethioate (23).** Thioacetic acid (330 mmol, 25 g) was added to paraformaldehyde (330 mmol, 10 g) and refluxed for 3 h. Vacuum distillation (2 mbar, 70 °C)

gave the pure *S*-hydroxymethyl ethanethioate as a yellow oil (24.8 g, 71%): δ_H($CDCl_3$; 250 MHz): δ = 5.04 ppm (s, 2H), 2.39 (s, 3H). The spectroscopic data was consistent with the literature.[240] Tribromophosphine was added dropwise for 30 min at 0 °C to *S*-hydroxymethyl ethanethioate. After further stirring for 30 min, the reaction mixture was allowed to warm to room temperature, stirred for 1 h, poured into 30 ml ice/water and stirred for 1 h. After extraction with diethylether (3 × 15 ml), the combined organic phases were dried ($MgSO_4$) and concentrated. Vacuum distillation (200 mTorr, 32 °C) yielded the pure product **23** as colourless oil (12.0 g, 47%): ^{1}H-NMR ($CDCl_3$; 250 MHz): δ = 4.71 (s, 2H), 2.41 ppm (s, 3H). The spectroscopic data was consistent with the literature.[240]

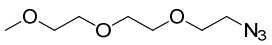


***S*-((4-(2-Methoxyethoxy)phenyl)(phenyl)phosphino)methyl ethanethioate borane (12).** To the crude mixture of **22** a base (0.682 mmol, Table 5.3) was added at 0 °C and stirred for 30 min. *S*-bromomethyl ethanethioate (**23**) (0.682 mmol, 115 mg) was added and the reaction mixture was stirred over night. Conversion was checked by ^{31}P-NMR (162 MHz) (see Table 5.3).

Table 5.3: Different reaction conditions for preparation of phosphine borane complex 12 (*traces).

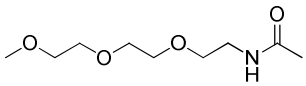
solvent	base	^{31}P-NMR [ppm]
MeCN	NaH	(27.2)*; (32.9)*; 60.6
MeCN	K^{t}OBu in THF	27.6; 32.0
DMF	LDA (*in situ* in THF)	(-27.8)*; (-8.7)*; 27.0; 29.9

5.3. First Acetylation Studies by Traceless Staudinger Ligation with Different Phosphines



1-Azido-2-(2-(2-methoxyethoxy)ethoxy)ethane (24). 2-(2-(2-Methoxyethoxy)ethoxy)-ethyl 4-methylbenzenesulfonate **18** (1.00 g, 2.14 mmol) and sodium azide (0.82 mg, 12.6 mmol) were dissolved in 10 ml DMF and stirred overnight at 60°C. The reaction was cooled to rt and diluted with water. The aqueous phase was extracted with diethylether. The organic layers were combined, washed with brine, dried and concentrated to yield a colourless clear oil (559 mg, 88%): ^{1}H-NMR ($CDCl_3$, 400 MHz) δ = 3.86-3.83 (m, 8H), 3.57-3.53 (m, 2H), 3.39-3.37 (m, 5H). The spectroscopic data was consistent with the literature.[241]

General protocols for traceless Staudinger ligation to yield N-(2-(2-(2-methoxyethoxy)-ethoxy)ethyl)acetamide (25).



A) Phosphine **13** (5.00 mg, 10.9 µmol) was dissolved in 0.5 ml dry DMF with 3eq. DABCO (3.70 mg, 32.8 µmol) and azide **24** (2.07 mg, 10.9 µmol). The mixture was heated to 40°C and stirred overnight. Conversion was checked by LC-LRMS (Table 5.4, entry 1).

B) Phosphine **13** (5.00 mg, 10.9 µmol) was dissolved in 0.4 ml dry DMF with 3eq. DABCO (3.70 mg, 32.8 µmol) and stirred for 4 h at 40°C. Azide **24** (2.07 mg, 10.9 µmol) was added in ~0.1 ml phosphate buffer (0.1M, pH 7) and the reaction was stirred overnight at 40°C. Conversion was checked by LC-LRMS (Table 5.4, entry 2).

C) The phosphine (**1:** 3.14 mg, 10.9 µmol; **13:** 5.00 mg, 10.9 µmol) was dissolved in 0.5 ml neat TFA and stirred for 1 h at rt. Azide **24** (2.07 mg, 10.9 µmol) was added in ~0.5 ml phosphate buffer (see Table 5.4), the pH was adjusted and the reaction was stirred overnight at 40 °C. Conversion was checked by LC-LRMS (Table 5.4, entry 3-9).

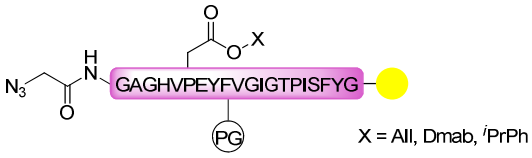
LRMS: m/z: $[M+H]^+$ (monoisotopic) 190.1 (azide **24**), 206.1 (amide **25**), 164.1 (amine **26**); m/z: $[M+H]^+$ (ESI-ToF) 190.1 (azide **24**), 206.1 (amide **25**), 164.1 (amine **26**).

Table 5.4: Traceless Staudinger ligation of 1-azido-2-(2-(2-methoxyethoxy)ethoxy)ethane (24) with phosphines 1 and 13 in different media at 40 °C (*ratios estimated by TIC (only for comparison of different conditions, no yields)).

entry	phosphine	borane-deprotection	solvent system	azide 24*	amide 25*	amine 26*
1	**13**	DABCO	DMF (dry)	–	100%	–
2	**13**	DABCO	phosphate buffer (0.7 M, pH 7)/DMF (1:4)	5%	90%	5%
3	**13**	TFA	phosphate buffer (0.7 M, pH 7)	5%	48%	47%
4	**1**	TFA	DMF (dry), DIPEA (12 eq.)	10%	90%	–
5	**13**	TFA	DMF (dry) , DIPEA (12 eq.)	10%	90%	–
6	**1**	TFA	phosphate buffer (0.4 M, pH 8)	10%	70%	20%
7	**13**	TFA	phosphate buffer (0.4 M, pH 8)	5%	55%	40%
8	**1**	TFA	phosphate buffer (0.4 M, pH 9)	10%	45%	45%
9	**13**	TFA	phosphate buffer (0.4 M, pH 9)	10%	40%	50%

5.4. Synthesis of Microcin *J25* Derivatives and Cyclisation Studies

5.4.1. Glutamic Acid Side Chain Deprotection Strategies



All peptides **41** were synthesised as described above (see chapter 5.1). After TFA cleavage the peptides were checked by LRMS:

41a: Allyl: m/z (monoisotopic): $[M+2H]^{2+}$ 1096.0; m/z (exp.): $[M+2H]^{2+}$ (ESI-ToF) 1096.0.

41b: Dmab: m/z (monoisotopic): $[M+2H]^{2+}$ 1231.6; m/z (exp.): $[M+2H]^{2+}$ (ESI-ToF) 1231.7.

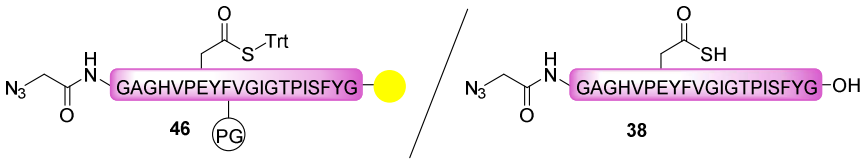
41c: ^{i}PrPh: m/z (monoisotopic): $[M+2H]^{2+}$ 1076.0; m/z (exp.): $[M+2H]^{2+}$ (ESI-ToF) 1076.1.

Allyl deprotection: The synthesised peptide **41a** on resin was treated with $Pd(PPh_3)_4/CHCl_3/AcOH/NMM$,[189] washed thoroughly with DMF and CH_2Cl_2, treated with 95% TFA (2.5% each TIS/H_2O) for 2 h, concentrated and checked by LRMS: m/z (monoisotopic): $[M+2H]^{2+}$ 1096.0 (allyl ester **41a**), 1063.0 (amino peptide **42**); m/z (exp.): $[M+2H]^{2+}$ (ESI-ToF) 1096.0 (allyl ester **41a**), 1062.9 (amino peptide **42**).

Dmab deprotection: Peptide **41b** on resin (5 mg, 1 µmol) was treated with 1 mL 2% $N_2H_4 \cdot H_2O$ in DMF for 10 min[190] and washed thoroughly with DMF and CH_2Cl_2. After addition of a basic mixture (a) 5% DIPEA in DMF, b) 5 mM NaOH, c) 50 mM NaOH, d) 250 mM NaOH in dioxane), the resin was shaken for: a) 10 min, b) 2 h, c) 1-3 h or 15 min at rt (a,b,c) or at 0 °C (d) and washed thoroughly with DMF and CH_2Cl_2. After TFA cleavage/deprotection with 95% TFA (2.5% TIS, 2.5% H_2O) for 2 h, the cleavage mixture was concentrated and checked by LRMS: m/z (monoisotopic): $[M+2H]^{2+}$ 1076.0 (peptide **47**), 1128.5 (*p*-aminobenzylester **43**); m/z (exp.): $[M+2H]^{2+}$ (ESI-ToF) 1076.1 (peptide **47**), 1128.8 (*p*-aminobenzylester **43**).

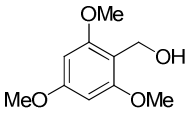
^{i}PrPh deprotection: Peptide **41c** on resin (5 mg, 1 µmol) was treated with 2% TFA in dry CH_2Cl_2 (0.5 mL) and shaken for 10 min at rt.[188] After thorough washing with CH_2Cl_2, propargylamine (1.3 µL, 20 µmol), Pybop (10.4 mg, 20 µmol) and NMM (4.4 µL, 20 eq.) were added with 2 mL dry DMF to the peptide on resin and shaken overnight at rt. The resin was washed thoroughly with dry DMF and CH_2Cl_2 and treated with 95% TFA (2.5% TIS, 2.5% H_2O) for 2 h at rt. The cleavage mixture was concentrated and checked by LRMS: m/z (monoisotopic): $[M+2H]^{2+}$ 1094.5 (propargylamide **44**); m/z (exp.): $[M+2H]^{2+}$ (ESI-ToF) 1094.6 (propargylamide **44**).

5.4.2. Different Approaches for the Thioacid Formation

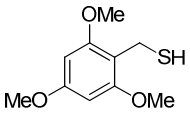


Probed formation of trityl thioester 46 / thioacid 38:

The peptide on resin **41c** (~0.01 mmol) was treated with 2% TFA in CH_2Cl_2 for 10–15 min and then washed thoroughly with CH_2Cl_2. For thioesterification of the free acid, a cocktail of tritylthiol, Pybop/HATU/DIC/DCC (one of them) and DIPEA (0.5 mmol each) in dry DMF was added to the resin and it was shaken overnight at room temperature. The resin was washed again with DMF and CH_2Cl_2 and treated with 70% TFA (5% TIS in dry CH_2Cl_2) for 2 h. The cleavage solution was concentrated under vacuum and checked by MS: LRMS: m/z (monoisotopic): $[M+2H]^{2+}$ 1076.0 (hydrolysed azido peptide **47**); m/z (exp.): $[M+2H]^{2+}$ (ESI-ToF) 1076.0 (hydrolysed azido peptide **47**).

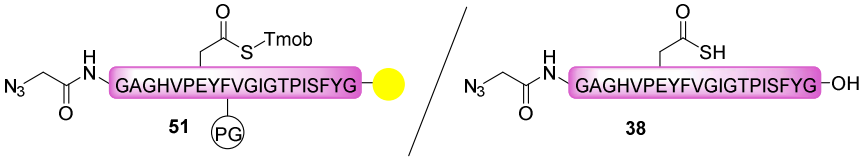


2,4,6-Trimethoxybenzyl alcohol (49). To a suspension of 2,4,6-trimethoxy benzaldehyde **48** (15.3 mmol, 3.00 g) in ethanol (~40 mL) was slowly added $NaBH_4$ (15.3 mmol, 0.289 g). The reaction turned into a clear solution. It was left stirring at rt and checked by TLC. After 1 h the reaction was quenched with H_2O (~50 mL) and the aqueous phase was extracted with Et_2O (3 x 40 mL). The combined organic phases were dried and concentrated to give a yellow viscous oil. Column chromatography (EtOAc/Cychex; 2:1) yielded the desired product **49** as a viscous oil (7.58 mmol, 3.03 g, quant.): R_f (EtOAc/Cychex; 1:1) = 0.4; ^{1}H-NMR (d_6-DMSO, 400 MHz): δ(ppm) = 6.18 (s, 2H), 4.38 (d, J = 5.4 Hz, 2H), 4.22 (t, J = 5.4 Hz, OH), 3.76 (s, 3H), 3.74 (s, 6H); ^{13}C-NMR (d_6-DMSO, 400 MHz): δ(ppm) = 160.7, 159.3, 110.1, 90.7, 55.8, 55.4, 51.5; HRMS: m/z (calculated): $[M+Na]^+$ (monoisotopic) 221.0784; m/z (exp.): $[M+Na]^+$ (ESI-ToF) 221.0795. The spectroscopic data was in accordance with the literature.[192]



2,4,6-Trimethoxybenzylthiol (50). Alcohol **49** (7.30 mmol, 1.45 g), thiourea (8.32 mmol, 0.634 g) and toluenesulfonic acid hydrate (7.57 mmol, 1.44 g) were dissolved in acetonitrile (~150 mL) and refluxed for 1 h. The reaction was quenched with NaOH (10% aq. solution, ~50 mL) and refluxed for 4 h. After acidification with 2.5 M HCl to pH 1, the mixture was extracted with Et_2O.

The combined organic phases were washed with brine, dried and concentrated to give a yellow liquid. Column chromatography (EtOAc/Cychex; 1:25) yielded the desired product **50** as a colourless oil (2.42 mmol; 33%): R_f (EtOAc/Cychex; 1:25) = 0.3; ^{1}H-NMR (d_6-DMSO, 400 MHz): δ(ppm) = 6.20 (s, 2H), 3.77 (s, 3H), 3.74 (s, 6H), 3.57 (d, *J* = 7.6 Hz, 2H), 2.29 (t, *J* = 7.6 Hz, 1H); ^{13}C-NMR (d_6-DMSO, 400 MHz): δ(ppm) = 160.2, 160.0, 109.5, 90.9, 55.9, 55.4, 16.2. HRMS: m/z (calculated): $[M+Na]^+$ (monoisotopic) 221.0784; $[2M-H_2+Na]^+$ (monoisotopic) 449.1063; m/z (exp.): $[2M-H_2+Na]^+$ (ESI-ToF) 449.1081. The spectroscopic data was in accordance with the literature.[192]



Tmob thioester 51 / thioacid 38:

The peptide on resin (~0.01 mmol) was treated with 2% TFA in CH_2Cl_2 for 10–15 min and then washed thoroughly with CH_2Cl_2. For thioesterification of the free acid, a cocktail of **50**, Pybop and DIPEA (0.5 mmol each) in dry DMF was added to the resin and it was shaken overnight at room temperature. The resin was washed again with DMF and CH_2Cl_2 and treated with 50% TFA (10% TIPS in dry CH_2Cl_2) for 1.5 h. The cleavage solution was concentrated under vacuum and checked by LRMS: m/z (monoisotopic): $[M+2H]^{2+}$ 1076.0 (hydrolysed azido peptide **47**), 1084.0 (thioacid **38**), 1124.0 (TFA ester of **47**), 1132.0 (TFA ester of **38**); m/z (experimental): $[M+2H]^{2+}$ (ESI-ToF) 1076.2 (hydrolysed azido peptide **47**), 1084.4 (thioacid **38**), 1124.2 (TFA ester of **47**), 1132.1 (TFA ester of **38**).

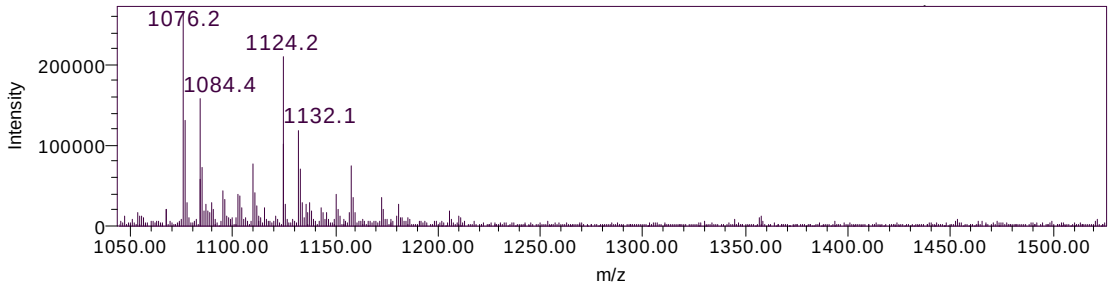
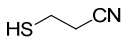
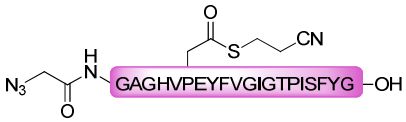


Figure 5.1: LRMS of crude reaction mixture from thioacid 38 formation via Tmob thioester 51.



Mercaptopropionitrile (53). 3,3'-Dithiobis(propionitrile) (**52**) (1.72 g, 10.0 mmol) was suspended in 25 mL 2 N HCl and heated up to 40 °C. Zn powder (1.31 g, 20.0 mmol) was added slowly. The reaction was stirred for 30 min at 40 °C and cooled to rt. The reaction mixture was extracted

with CH_2Cl_2 (4× 10 mL). The organic layers were combined, dried over $MgSO_4$ and concentrated to yield the product as a colourless liquid (1.58 mg, 9.05 mmol, 91%): ^{1}H-NMR ($CDCl_3$, 300 MHz): δ = 2.79-2.72 (m, 2H), 2.68-2.64 (m, 2H), 1.78 ppm (t, J = 8.5 Hz, SH); ^{13}C-NMR ($CDCl_3$, 75 MHz): δ = 117.97, 22.78, 20.51 ppm. The analytical data was in accordance with the literature.[79]



3-Cyano ethyl thioester peptide 55. The peptide on resin (~0.01 mmol) was treated with 2% TFA in CH_2Cl_2 for 10–15 min and then washed thoroughly with CH_2Cl_2. For thioesterification of the free acid, a cocktail of **53**, Pybop and DIPEA (0.5 mmol each) in dry DMF was added to the resin and it was shaken overnight at room temperature. The resin was washed again with DMF and CH_2Cl_2 and treated with 50% TFA 10% TIPS, 40% dry CH_2Cl_2) for 1.5 h. The cleavage solution was concentrated under vacuum and checked by LRMS: m/z (monoisotopic): $[M+2H]^{2+}$ 1076.0 (hydrolysed azido peptide **47**), 1110.5 (thioester **55**), m/z (experimental): $[M+2H]^{2+}$ (ESI-ToF) 1075.8 (hydrolysed azido peptide **47**), 1110.6 (thioester **55**).

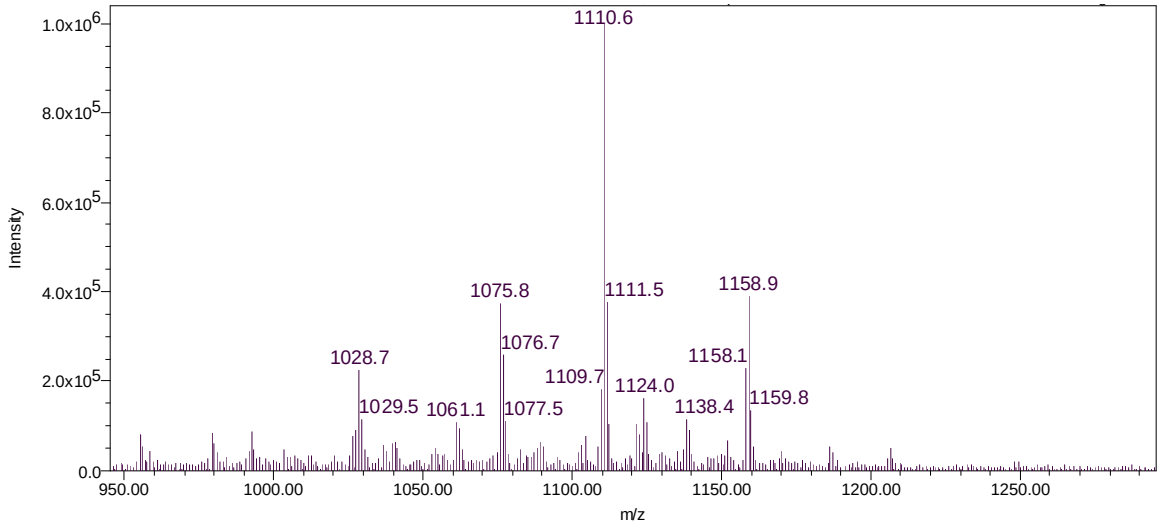
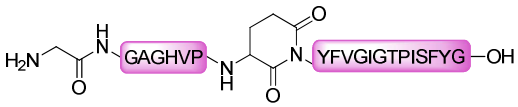


Figure 5.2: LRMS spectrum of crude thioester 55.

5.4.3. Thioacid–Azide Cyclisation



Peptide cyclisation by thioacid–azide reaction from 3-cyano ethyl thioester peptide 55. The peptide on resin (~0.01 mmol) was treated with 2% TFA in CH_2Cl_2 for 10–15 min and then washed thoroughly with CH_2Cl_2. For thioesterification of the free acid, a cocktail of **53**, Pybop

and DIPEA (0.5 mmol each) in dry DMF was added to the resin and it was shaken overnight at room temperature. The resin was washed again with DMF and CH_2Cl_2 and treated with 50% TFA (10% TIPS, 40% dry CH_2Cl_2) for 1.5 h. The cleavage solution was concentrated under vacuum, redissolved in sodium phosphate buffer (0.1 M, pH 9)/THF (1:1) with 3 eq. $(NH_4)_2S$ and stirred for 2 h. Purification by preparative HPLC (35 mL/min, 16 to 64% MeCN in 70 min, t_r = 32.5 min) yielded 6.2 mg (92% purity) of glutarimide peptide **56**: MALDI-MS: m/z (calculated): $[M+H]^+$ 2107.0284; m/z (MALDI-MS(CCA)): $[M+H]^+$ 2107.2700.

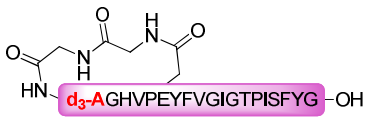
MS/MS fragmentation (MALDI-MS (CCA)):

	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}	b_{13}	b_{14}
theor.	243.11	380.17	479.24	576.29	–	850.39	997.45	1096.52	1153.54	1266.63	1323.65
exper.	243.15	380.25	479.32	576.38	–	850.54	997.64	1096.72	1153.76	1266.86	1323.90

	b_{15}	b_{16}	b_{17}	b_{18}	b_{19}	b_{20}
theor.	1424.70	1521.75	1634.83	1721.87	1868.93	2032.00
exper.	1424.95	1522.01	1635.10	1722.34	1869.21	2032.32

	y_2	y_6	y_7	y_8	y_9	y_{10}	y_{11}	y_{12}	y_{13}
theor.	239.10	683.34	784.39	841.41	954.49	1011.52	1110.58	1257.65	1420.71
exper.	239.14	683.46	784.52	841.53	954.65	1011.68	1110.79	1257.87	–

	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	y_{20}
theor.	–	1628.80	1727.87	1864.93	1921.95	1992.99	2050.01
exper.	–	1629.07	1728.13	1865.21	1922.24	1993.29	2050.34



Deuterated cyclised Microcin variant 60. Peptide synthesis (0.05 mmol) was carried out on a peptide synthesiser. The Wang resin was preloaded with glycine (0.6 mmol). Fmoc-Glu(OAll)-OH (5 eq.) and all subsequent amino acids until azido glycine were coupled manually (10 eq. each) following the Fastmoc protocol. After cleavage of the last Fmoc-group, the glutamic acid side chain was deprotected. The allyl group was cleaved under mild reductive conditions using tetrakis-(triphenylphosphine)palladium(0) (25.0 mg, 22.0 µmol) and phenyl silane (123 µL,

1 mmol) in dry CH_2Cl_2 (2 mL). The resin was shaken for 1 h and the procedure was repeated. Afterwards the resin was washed thoroughly with CH_2Cl_2 (5×3 mL), DMF (3×3 mL) and CH_2Cl_2 (3×3 mL). For cyclisation Pybop (260 mg, 0.05 mmol), DIPEA (8.5 µL, 0.05 mmol) and dry DMF (~4 mL) were added to the resin and it was shaken overnight at rt. The resin was washed thoroughly with CH_2Cl_2 and DMF. Cleavage and final deprotection was performed with 95% TFA (H_2O/TIS 2.5:2.5 v/v). Precipitation in cold diethylether yielded the crude peptide that was further purified by semi-prep. HPLC (MeCN from 50% to 80% in 40 min; t_r = 5 min) to yield 2.6% of the pure peptide as a white solid: HRMS: m/z (calculated): $[M+H]^+$ 2110.0472, $[M+2H]^{2+}$ 1055.5272 (product); m/z (experimental): $[M+H]^+$ (ESI-ToF) 2110.0552, $[M+2H]^{2+}$ (ESI-ToF) 1055.5329 (product).

Peptide cyclisation by thioacid–azide reaction on Tmob-protected peptide 51. The peptide on resin (~0.01 mmol) was treated with 2% TFA in CH_2Cl_2 for 10–15 min and then washed thoroughly with CH_2Cl_2. For thioesterification of the free acid, a cocktail of **50**, Pybop and DIPEA (0.5 mmol each) in dry DMF was added to the resin and it was shaken overnight at room temperature. The resin was washed again with DMF and CH_2Cl_2 and treated with 50% TFA 10% TIPS, 40% dry CH_2Cl_2) for 1.5 h. The cleavage solution was concentrated under vacuum, redissolved (see Table 5.5) and stirred overnight at room temperature. After lyophilization of the samples and redissolving them in H_2O/water, conversion was checked by LC-HRMS (assigned masses): m/z (calculated): $[M+2H]^{2+}$ 1076.0183 (hydrolysed azido peptide **47**), 1054.0178 ("reduction–dehydration product"); m/z (experimental): $[M+2H]^{2+}$ (ESI-ToF) 1076.0233 (hydrolysed azido peptide **47**), 1054.0226 ("reduction–dehydration product").

Table 5.5: Probed peptide cyclisation by thioacid–azide reaction starting from Tmob-thioester peptide 51.

entry	solvent system, c(peptide) = 1 mM	T [°C]	yield[a]
1	H_2O (6 M Gnd-HCl, 3 mM lutidine	30	n.d
2	sodium phosphate buffer (0.1 M, pH 8)	30	n.d.
3	Tris-HCl buffer (50 mM, pH 8, 200 mM NaCl)	30	n.d.
4	sodium citrate buffer (0.1 M, pH 2.5)/DMF (1:1)	30	6.0
5	sodium citrate buffer (0.1 M, pH 2.5, 200 mM NaCl)/DMF (1:1)	30	n.d.
6	sodium citrate buffer (0.1 M, pH 2.5, 200 mM LiCl)/DMF (1:1)	30	n.d.
7	DMF, 3 mM lutidine	30	9.0

[a] Determined by MS (comparison with deuterated product **60**).

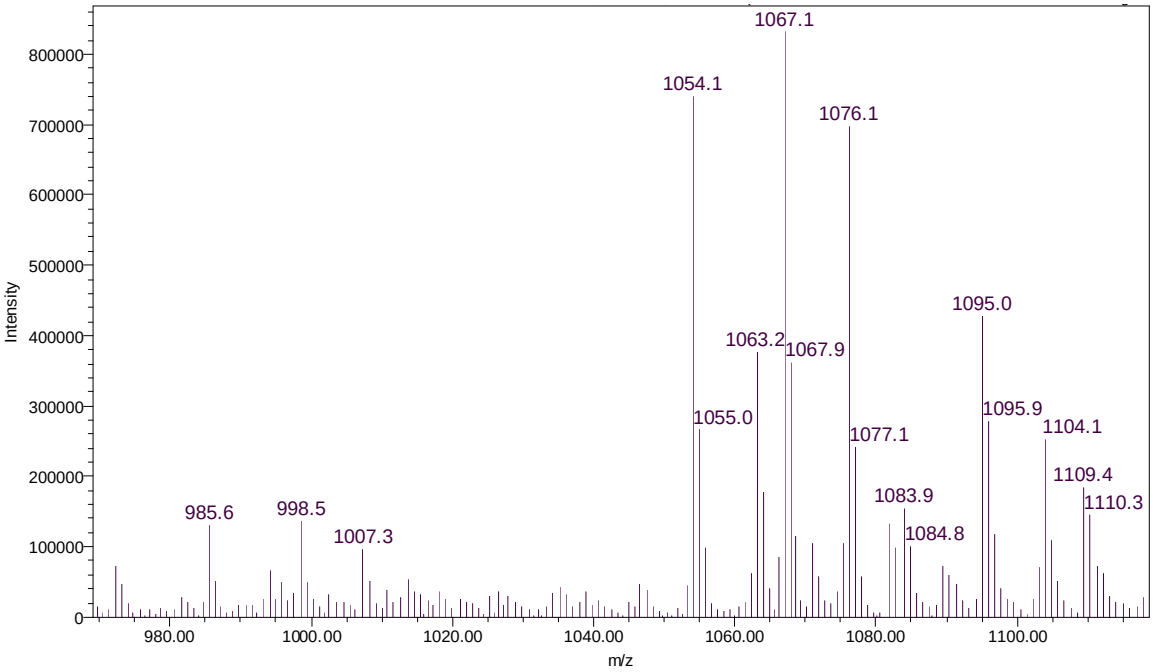
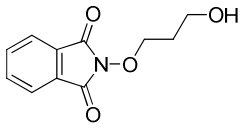


Figure 5.3: LRMS spectrum of crude reaction mixture of attempted peptide cyclisation from thioester 51.

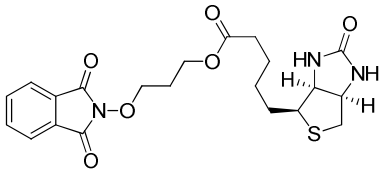
5.5. Protein Modifications on TTL(D221Bpa,S261C)[Aha]

5.5.1. Oxime Ligation

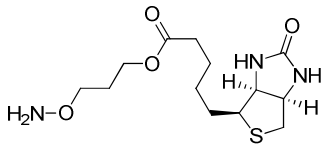
5.5.1.1. General Synthesis of Biotin Derivatives **66** and **68**



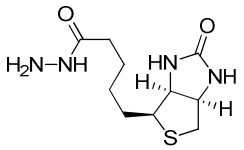
2-(3-Hydroxypropoxy)benzo[c]azolidine-1,3-dione (63). A solution of sodium acetate (37.4 mmol, 3.06 g), N-hydroxyphtalimide (**61**) (12.7 mmol, 2.00 g) and 3-bromoethanol (**62**) (33.1 mmol, 4.60 g) in 20 mL DMSO was stirred for 3 h at 70 °C. After it was cooled down to room temperature, 20 mL water was added to the reaction mixture and the whole was extracted with CH_2Cl_2 (3 × 12.5 mL). The combined organic phases were washed with water (2 × 12.5 mL), 3N HCl (10 mL), and brine, dried over $MgSO_4$, and concentrated. Recrystallization from ethanol yielded the product in 57% (7.23 mmol, 1.48 g) as yellowish platelets: ^{1}H-NMR ($CDCl_3$, 400 MHz): δ(ppm) = 7.83-7.80 (m, 2H), 7.75-7.73 (m, 2H), 4.38-4.34 (m, 2H), 3.93-3.90 (m, 2H), 2.64 (bs, OH), 2.01-1.95 (m, 2H); ^{13}C-NMR ($CDCl_3$, 400 MHz): δ(ppm) = 163.9, 134.8, 128.9, 123.8, 76.1, 59.1, 30.9. The analytical data is in accordance with the literature.[242]



3-(1,3-dioxoisoindolin-2-yloxy)propyl-5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate (65). Thionyl chloride (13.8 mmol, 1 mL) was added slowly to biotin (1.03 mmol, 250 mg) and the solution was stirred for 1 h at room temperature. After removal of thionyl chloride under vacuum, the crude acyl chloride **64** was dissolved in dry CH_2Cl_2, and a solution of **63** (2.06 mmol, 423 mg) in dry CH_2Cl_2 and NET_3 (3.09 mmol, 0.43 mL) were added and stirred overnight. Filtration, concentration and flash column chromatography (EtOAc:hex; 1:2) gave the product, which was used directly for the next step: R_f = 0.3; ^{1}H-NMR ($CDCl_3$, 400 MHz): δ(ppm) = 7.78-7.69 (m, 4H), 6.82 (bs, NH), 4.86 (bt, 1H), 4.56-4.44 (m, 2H), 4.36 (bdd, 1H), 3.20 (bq, 1H), 3.11 (d, *J* = 13.6 Hz, 1H), 2.96 (dd, *J* = 13.6, 5.4 Hz, 1H), 2.30 (t, *J* = 7.3 Hz, 2H), 2.12 (q, *J* = 6.2 Hz, 2H), 2.04 (q, *J* = 6.2 Hz, 2H), 1.73 (m, 1H), 1.62 (m, 3H), 1.44 (m, 2H).



3-(aminooxy)propyl 5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate (66). 3-(1,3-dioxoisoindolin-2-yloxy)propyl 5-((3aS,4S,6aR)-2-oxohexa-hydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate (**65**) (0.59 mmol, 265 mg) was mixed with hydrazine hydrate (2.36 mmol, 115 µL) and stirred for 1 h at room temperature. Filtration and concentration yielded the desired product **66** as a viscous yellowish oil in 87% (0.51 mmol, 163 mg): ^{1}H-NMR ($(CD_3)_2CO$, 300 MHz): δ(ppm) = 6.23 (bs, NH), 6.01 (bs, NH), 4.55 – 4.44 (m, 1H), 4.40 – 4.26 (m, 1H), 4.12 (t, *J* = 6.5 Hz, 2H), 4.04 (t, *J* = 6.1 Hz, 2H), 3.70 – 3.46 (m, 1H), 3.22 (dd, *J* = 11.9, 7.2 Hz, 1H), 2.70 (d, *J* = 12.6 Hz, 1H), 2.33 (t, *J* = 7.4 Hz, 2H), 1.93 (p, *J* = 6.4 Hz, 2H), 1,80 (bs, NH_2), 1.75 – 1.39 (m, 6H); ^{13}C-NMR ($(CD_3)_2CO$, 75 MHz): δ(ppm) = 173.7, 154.6, 70.1, 62.5, 61.8, 60.8, 56.4, 41.0, 34.4, 29.2, 25.6, 21.7, 15.3; HRMS (ESI-ToF): m/z = 318.1482; $[M+H]^+$ (calc: m/z = 318.1482).



5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanehydrazide[199] **(68).** Biotin (300 mg, 1.23 mmol) was suspended in 3 mL dry methanol. Thionyl chloride (0.30 mL,

4.00 mmol) was added dropwise to the suspension. A clear solution was formed and it was stirred overnight at rt. The reaction mixture was concentrated under reduced pressure to yield a white powder. It was redissolved in 2.5 mL dry methanol and hydrazine (10 mL, 10 mmol, 1 M in THF) was added dropwise. The solution was stirred overnight and concentrated under reduced pressure to give a white solid. The crude was diluted with 70 mL H_2O and washed with CH_2Cl_2 (3 x 30 mL). The aqueous phase was concentrated to yield the desired product as a white solid (1.11 mmol, 90%): ^{1}H-NMR (d_6-DMSO/D_2O, 400 MHz): δ(ppm) = 4.62 (dd, *J* = 7.9, 4.8 Hz, 1H), 4.43 (dd, *J* = 7.8, 4.6 Hz, 1H), 3.40-3.27 (m, 1H), 3.02 (dd, *J* = 13.0, 5.0 Hz, 1H), 2.80 (d, *J* = 13.0 Hz, 1H), 2.26 (t, *J* = 7.4 Hz, 2H), 1.86-1.54 (m, 4H), 1.53-1.37 (m, 2H); HRMS (ESI-ToF): m/z = 259.1243; $[M+H]^+$ (calc: m/z = 259.1223). The analytical data was in accordance with the literature.[199]

5.5.1.2. Hydrazone Formation on Small Molecules

Biotin hydrazide **68** was mixed 1:1 with 4-benzoyl-L-phenylalanine (**69a**), acetophenone (**69b**) and benzaldehyde (**69c**), respectively, and suspended/dissolved (see Table 5.6). It was shaken for 3h and conversion was checked by LCUV-LRMS:

4-benzoyl-L-phenylalanine (**69a**): LRMS: m/z (calculated): $[M+H]^+$ 510.2 (product **70a**), 259.1 (biotin hydrazide **68**), m/z (experimental): $[M+H]^+$ 259.1 (biotin hydrazide **68**);

acetophenone (**69b**): LRMS: m/z (calculated): $[M+H]^+$ 361.2 (product **70b**), 259.1 (biotin hydrazide **68**), 237.1 (di(1-phenylethylidene)hydrazone **71b**) m/z (experimental): $[M+H]^+$ 361.1 (product **70b**), 259.1 (biotin hydrazide **68**), 237.1 (di(1-phenylethylidene)hydrazone **71b**);

benzaldehyde (**69c**): LRMS: m/z (calculated): $[M+H]^+$ 347.2 (product **70c**), 259.1 (biotin hydrazide **68**), 209.1 (di(benzylidene)hydrazone **71c**); m/z (experimental): $[M+H]^+$ 347.0 (product **70c**), 259.1 (biotin hydrazide **68**), 209.1 (di(benzylidene)hydrazone **71c**).

Table 5.6: Hydrazone formations probed between biotin hydrazide 68 and carbonyl compounds 69a–c (t = 3 h).

entry		T [°C]	solvent	catalyst (100 mM)	conversion to hydrazone 70a–c*
1	**69a**	25	EtOH	–	0%
2		25	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	–	0%
3		40	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	–	0%
4	**69b**	25	EtOH	–	0%
5		25	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	–	25%
6		40	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	–	25%
7		40	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	aniline	0%
8		40	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	*p*-amino Phe	0%
9		40	phosphate buffer (0.1 M, pH 7.0)	–	0%
10		40	phosphate buffer (0.1 M, pH 7.0)	aniline	0%
11		40	phosphate buffer (0.1 M, pH 7.0)	*p*-amino Phe	0%
12	**69c**	25	EtOH	–	50%
13		25	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	–	50%
14		40	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	–	50%
15		40	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	aniline	60%
16		40	NH_4OAc (30 mM, pH 4.5, 20% MeCN)	*p*-amino Phe	60%
17		40	phosphate buffer (0.1 M, pH 7.0)	–	50 – 60%
18		40	phosphate buffer (0.1 M, pH 7.0)	aniline	40 – 50%
19		40	phosphate buffer (0.1 M, pH 7.0)	*p*-amino Phe	20%

5.5.1.3. Protocol for Oxime & Hydrazone Ligation on TTL(D221Bpa,S261C)[Aha]

Oxime ligation. Three solutions of the TTL(D221Bpa, S261C)[Aha] (0.5 ml, 0.6 mg/mL, 20 mM Tris/HCl, pH 8) were mixed with 3.41 µl biotin hydroxylamine **66** (0.145 M in acetone) and 0.1 mg *p*-methoxy aniline. The pH was set to 2.0, 4.0 and 7.5, respectively, and the reactions were shaken for 2 d at 15 °C. The solutions were centrifuge filtered and washed seven times with Millipore water. MALDI-MS measurements showed the following results for pH 2 and 4: m/z: $[M+H]^+$ = 30213 Da (MALDI-MS (SA)); 30227 Da (calculated; starting material); 30540 Da (broad; experimental); 30618 Da (calculated; product). The reaction at pH 7.5 turned completely black and no protein could be detected.

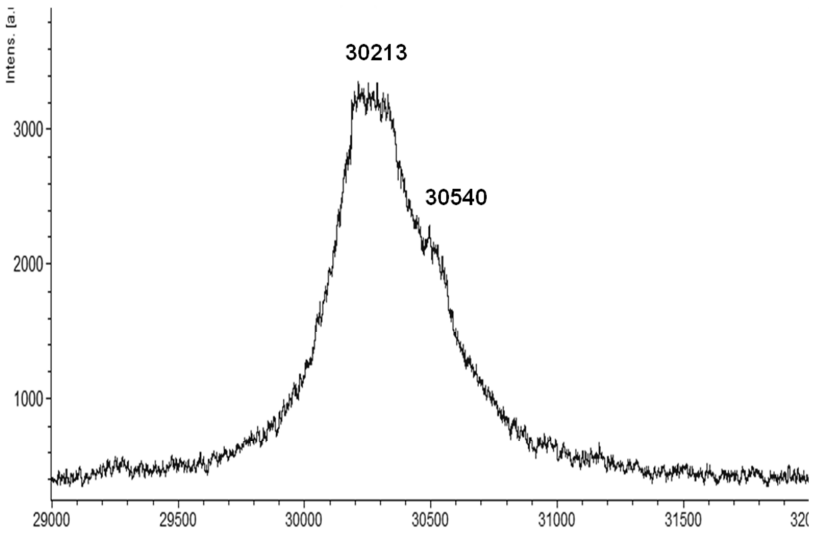


Figure 5.4: MALDI-MS spectrum of oxime ligation with TTL(D221Bpa, S261C)[Aha] and hydroxylamine 66 at pH 2.

Hydrazone ligation. A solution of TTL(D221Bpa, S261C)[Aha] (conditions see Table 5.7) was mixed with biotin hydrazide (10eq; stock solution in Millipore water) and shaken overnight at 25 °C.

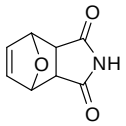
Table 5.7: Conditions for hydrazone formation with TTL Lipase and biotin hydrazide 68.

entry	buffer system	additive
1	anilinium acetate buffer (100 mM, pH 4.6)	–
2	phosphate buffer (20 mM, 100 mM NaCl, pH 7.4)	*p*-methoxyaniline (100 mM)
3	phosphate buffer (20 mM, 100 mM NaCl, pH 7.4)	*p*-aminophenylalanine (10 mM)
4	ammonium acetate buffer (100 mM, pH 4.9)	aniline (100 mM)

The solutions were purified by membrane ultracentrifuge filtration (3x w/ phosphate buffer (100 mM, pH7) and 5x w/ Millipore water) and concentrated to ~50 µl. Conversion was checked by MALDI-MS (SA): m/z (calculated): $[M+H]^+$ 30227 Da (starting material), 30467 Da (hydrazone product), m/z (MALDI-MS (SA)): $[M+H]^+$ 30307 (entry 1), 30321 (entry 3).

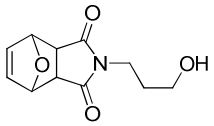
5.5.2. Thiol-Michael Reaction

5.5.2.1. General Synthesis of Rhodamine Maleimide **77**

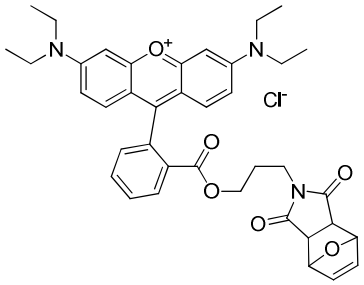


Furan-protected maleimide (73). Maleimide (**72**) (61.8 mmol, 6.00 g) and furan (618 mmol, 44.8 ml) were mixed in a round bottom flask and stirred at room temperature over night.

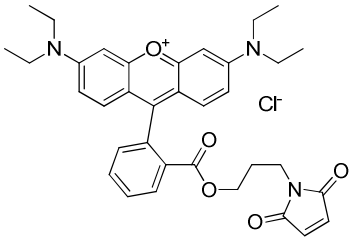
Concentration under high vacuum yielded 97% (60.0 mmol, 9.90 g) of the desired product **73**: ^{1}H-NMR ($CDCl_3$, 400 MHz): δ(ppm) = 6.50-6.47 (m, 2H), 5.32-5.29 (m, 2H), 3.58-3.52 (m, 1H), 2.88-2.87 (m, 1H); ^{13}C-NMR ($CDCl_3$-d_4-MeOD, 400 MHz): δ(ppm) = 136.2, 134.2, 80.7, 79.0, 47.2. The analytical data is in accordance with the literature.[202]



Furan-protected *N*-(3-Hydroxypropyl)-maleimide (74). Furan-protected maleimide **73** (6.06 mmol, 1.00 g), 1,3-propandiol (17.0 mmol, 1.23 mL), PPh_3 (16.7 mmol, 4.37 g) and neopentyl alcohol (10.3 mmol, 1.12 mL) were dissolved in 150 mL THF at -78 °C and the solution was stirred for 5 min. After addition of DIAD in 50 mL THF, stirring went on for 15 min before the reaction was allowed to warm to room temperature and stirred for 14 h. The yellow solution was concentrated and flash column chromatography (R_f = 0.25; EtOAc:hex; 3:1) yielded a mixture of the product and $P(O)Ph_3$ containing 1.21 g (5.4 mmol, 90%) of **77**. It was used without further purification. The analytical data is in accordance with the literature.[243]



Furan-protected 1-(3-hydroxypropyl)-1H-pyrrole-2,5-dione Rhodamine B ester (76). A solution of Rhodamine B (**75**) (0.59 mmol, 276 mg), furan-protected 1-(3-hydroxypropyl)-1H-pyrrole-2,5-dione (**74**) (1.73 mmol, 386 mg), HOBt (1.15 mmol, 156 mg) and DIPEA (2.30 mmol, 0.39 mL) in CH_2Cl_2 was cooled down to 0 °C and TCFH (1.15 mmol, 156 mg) was added portionwise. The reaction was left to warm to room temperature and stirred overnight. The same amount of TCFH and DIPEA was added again and the reaction mixture left stirring for 2 d. The reaction was checked by LCMS and concentrated when all Rhodamine B was consumed. The crude product **76** was used without further purification in the next step. HRMS: m/z (calculated): $[M+H]^+$ 648.307; m/z (experimental): $[M+H]^+$ (ESI-ToF) 648.302.



1-(3-hydroxypropyl)-1H-pyrrole-2,5-dione Rhodamine B ester (77). The crude ester **76** was dissolved in acetonitrile and refluxed for 2 d. Purification of 1/3 of the reaction mixture by HPLC (5min: 0 to 5% MeCN, 60 min: 5 to 65% MeCN; t_r = 41 min) yielded the pure product (50 mg, 81 µmol, 41%): ^{1}H-NMR (400 MHz, D_2O/d_3-MeCN 2:1) δ(ppm) = 8.87 (dd, *J* = 7.7, 1.1 Hz, 1H), 8.55-8.32 (m, 2H), 8.06-7.87 (m, 1H), 7.76-7.62 (m, 2H), 7.55 (d, *J* = 9.5 Hz, 2H), 7.46 (d, *J* = 2.2 Hz, 2H), 7.35 (t, *J* = 1.5 Hz, 2H), 4.53 (t, *J* = 6.1 Hz, 2H), 4.21 (dd, *J* = 13.7, 6.6 Hz, 8H), 3.88 (t, *J* = 6.6 Hz, 2H), 2.30-2.15 (m, 2H), 1.85 (t, *J* = 7.0 Hz, 12H); ^{13}C-NMR (400 MHz, D_2O/d_3-MeCN 2:1) δ(ppm) = 172.6, 166.7, 159.1, 158.6, 156.6, 156.6, 135.4, 135.3, 134.2, 134.0, 132.0, 132.0, 131.3, 131.2, 130.1, 115.3, 115.3, 114.2, 96.9, 96.9, 64.1, 46.6, 35.2, 27.8, 12.8; HRMS: m/z (calculated): $[M+H]^+$ 580.281; m/z (experimental): $[M+H]^+$ (ESI-ToF) 580.2813.

5.5.2.2. Protocol for Thiol-Michael Reaction on TTL(D221Bpa,S261C)[Aha]

Thiol-Michael approach. A solution of TTL(D221Bpa, S261C)[Aha] (0.5 mL, 0.6 mg/mL, for buffer see Table 5.8) was mixed with Rhodamine B maleimide **77** (4.9 g/L in Millipore water/acetonitrile), 10 µL sodium ascorbate (0.5 M in Millipore water) and in some cases a chaotropic agent; and the mixture was shaken for 1 d (for conditions see Table 5.8 and Table 5.9). The solution was centrifuge filtered and washed with 20% MeCN (in Millipore water, 3×) and with Millipore water (4×). MALDI-MS measurements showed the following results: m/z (calculated): $[M+H]^+$ = 30227 Da (TTL(D221Bpa, S261C)[Aha]); 30807 Da (product – one fold reacted); 31388 Da (two fold reacted); m/z (MALDI-MS (SA)): $[M+H]^+$ = 30226 Da (TTL(D221Bpa, S261C)[Aha]), 30808 Da (product – one fold reacted), 31379 Da (two fold reacted).

Table 5.8: First probed reaction conditions for Thiol-Michael reaction with 77 on TTL overnight.

entry	T [°C]	pH	Rhodamine maleimide 77 [eq.]
1	15	8.0	20
2	15	8.0	100
5	25	7.5	1
6	40	7.5	2
7	40	7.5	5
8	40	7.5	20

Table 5.9: Thiol-Michael reaction with 77 on TTL at pH 6.5 overnight at 40 °C (100 mM phosphate buffer, pH 6.5).

entry	c [µM] (protein)	maleimide 77 [eq.]	additive (c [mM])
1	23.16	10	–
2	23.16	50	–
3	4.63	50	–
4	23.16	10	Gnd·HCl (100)
5	23.16	10	aminoguanidine·HCl (100)
6	23.16	10	SDS (100)
7	23.16	20	Gnd·HCl (1000)
8	23.16	20	aminoguanidine·HCl (1000)

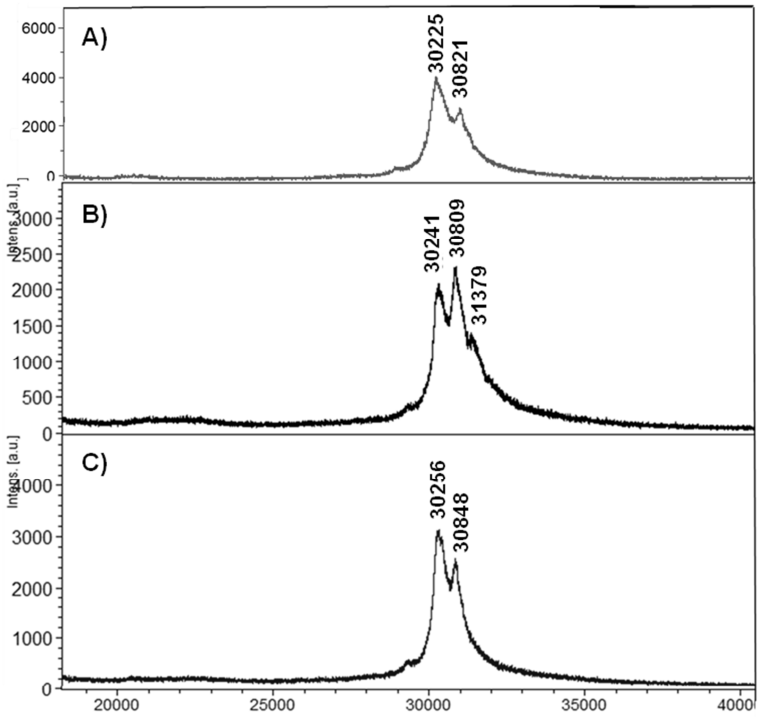


Figure 5.5: MALDI-MS spectra of thiol-Michael reaction at pH 6.5, 40 °C: A) 10 eq. of 77, B) 50 eq. of 77, C) 50 eq. of 77 (80% lower concentration).

Reaction with phenyl vinyl sulfone. A solution of TTL(D221Bpa, S261C)[Aha] (0.7 mg/mL; 20 mM phosphate buffer, 100 mM NaCl, pH7.4) was mixed with phenylvinyl sulfone and a possible additive (TCEP or sodium ascorbate) and shaken overnight at 27°C (for conditions see Table 5.10). The solutions were purified by membrane ultracentrifuge filtration (4x w/ NH_4HCO_3 (30 mM) and 5x w/ Millipore water) and concentrated to ~50 µL. Conversion was checked by MALDI-MS (SA): m/z (calculated): $[M+H]^+$ 30227 Da (starting material), 30395 Da (product), 30563 (two-fold addition); m/z (MALDI-MS (SA)): see Table 5.10.

Table 5.10: Reaction of TTL(D221Bpa, S261C) with phenyl vinyl sulfone (27 °C, o.n.).

entry	phenyl vinyl sulfone [µl]	additive	MALDI-MS (SA): $[M+H]^+$ [Da]
1	2 (3.4 eq.)	–	30241, 30321, 30483
2	2 (3.4 eq.)	sodium ascorbate (0.5 M)	30256, 30346, 30542
3	2 (3.4 eq.)	TCEP (48 mM)	30237, 30414
4	5 (8.5 eq.)	–	30202, 30326, 30506

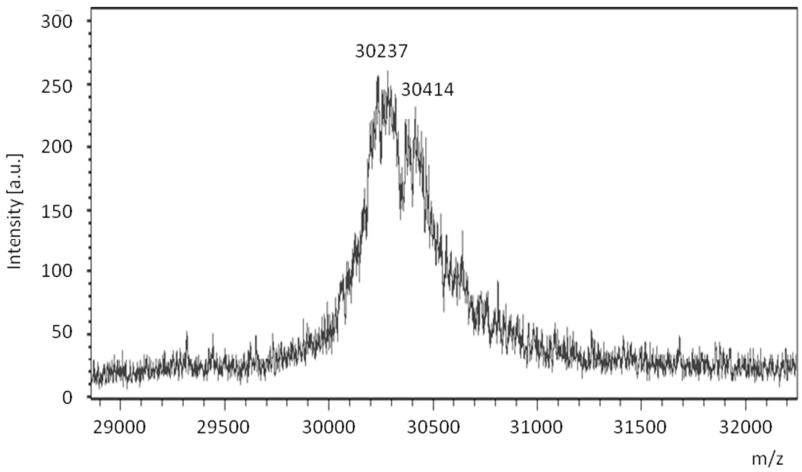
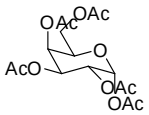


Figure 5.6: MALDI-MS after reaction of TTL with phenyl vinyl sulfone (Table 5.10, entry 3).

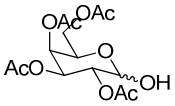
5.5.3. CuAAC

5.5.3.1. General Synthesis of Butinyle Galactose **84**

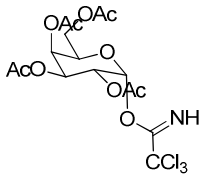
Path I towards 83:



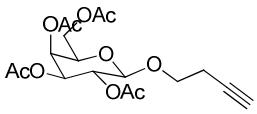
Peracetylated *α*-Galactose (79). Galactose (**78**) (11.3 mmol, 2.00 g) was dissolved in 200 mL pyridine and cooled down to -10 °C. Ac_2O (86.7 mmol, 8.00 mL) was added slowly and stirring was continued overnight. After addition of ice a white solid was filtered off and dried under vacuum to yield the desired product **79** in 53% (5.87 mmol, 2.29 g) yield: ^{1}H-NMR ($CDCl_3$, 400 MHz): δ(ppm) = 6.38 (d, *J = 4.0 Hz*, 1H), 5.50 (d, *J = 4.0 Hz*, 1H), 5.33 (s, 2H), 4.34 (bt, 1H), 4.09 (dt, *J = 4.0, 8.0 Hz*, 2H), 2.16 (s, 3H), 2.15 (s, 1H), 2.04 (s, 3H), 2.02 (s, 3H), 2.00 (s, 3H). The analytical data is in accordance with the literature.[210]



2,3,4,6-Tetra-*O*-acetyl galactosepyranose (80). A solution of **79** (1.41 mmol, 0.55 g) and freshly prepared hydrazinium acetate[244] (1.55 mmol, 0.183 g) in 5 mL DMF was stirred for overnight at room temperature. The reaction mixture was poured into EtOAc (30 mL) and washed with aq. sat. $NaHCO_3$ solution (3 times, 20 mL). The organic layer was dried over $MgSO_4$ and concentrated. Column chromatography (EtOAc:Cychex; 1:6 → 1:2) yielded the desired product **80** as a colourless liquid (0.93 mmol, 324.1 mg, 66%): R_f(EtOAc:Cychex, 1:1) 0.3; ^{1}H-NMR ($CDCl_3$, 400 MHz): δ(ppm) = 5.52 (d, *J* = 3.5 Hz, 1H), 5.47 (dd, *J* = 3.3, 1.2 Hz, 1H), 5.41 (dd, *J* = 10.8, 3.3 Hz, 1H), 5.16 (dd, *J* = 10.8, 3.6 Hz, 1H), 5.07 (dd, *J* = 4.6, 1.6 Hz, 0.65H), 4.47 (td, *J* = 6.5, 0.8 Hz, 1H), 4.09 (dd, *J* = 4.7, 1.7 Hz, 1H), 3.95 (td, *J* = 6.6, 1.2 Hz, 0.35H), 3.24 (bs, OH), 2.15 (s, 1H), 2.14 (s, 2H), 2.10 (s, 1H), 2.09 (s, 2H), 2.05 (s, 2H), 2.04 (s, 1H), 1.99 (s, 1H), 1.98 (s, 2H). The analytical data is in accordance with the literature.[245]



2,3,4,6-Tetra-*O*-acetyl-α-galactopyranosyl trichloroacetimidate (81). A solution of **80** (0.93 mmol, 324 mg) and K_2CO_3 (0.93 mmol, 129 mg) in CH_2Cl_2 was cooled down to 0 °C and trichloroacetonitrile (2.79 mmol, 0.28 mL) was added dropwise. The reaction mixture was warmed to room temperature and stirred for 2 d. Filtration, concentration and column chromatography (EtAOc:Cychex; 1:6 → 1:4) gave the desired product **81** as a colourless liquid in 58% yield (0.54 mmol, 266 mg): R_f(EtOAc:Cychex; 1:1) 0.65; ^{1}H-NMR ($CDCl_3$, 400 MHz): δ(ppm) = 8.67 (s, NH), 6.60 (d, *J* = 3.5 Hz, 1H), 5.56 (dd, *J* = 3.0, 1.2 Hz, 1H), 5.43 (dd, *J* = 10.9, 3.1 Hz, 1H), 5.36 (dd, *J* = 10.9, 3.5 Hz, 1H), 4.44 (td, *J* = 6.6, 0.8 Hz, 1H), 4.17 (dd, *J* = 10.4, 5.7 Hz, 1H), 4.08 (dd, *J* = 11.3, 6.7 Hz, 1H), 2.17 (s, 3H), 2.03 (s, 3H), 2.02 (s, 3H), 2.01 (s, 3H). The analytical data is in accordance with the literature.[246]



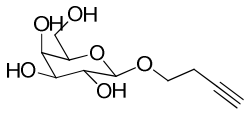
2,3,4,6-Tetra-*O*-acetyl-1-*O*-but-3-ynyl-β-galactopyranoside (83). A solution of **81** (0.54 mmol, 266 mg) and 3-butyn-1-ol (**82**) (1.08 mmol, 82.0 µL) in dry CH_2Cl_2 was cooled down to -90 °C. $BF_3{\cdot}OEt_2$ (3.24 mmol, 0.41 mL) was added dropwise. The reaction mixture was stirred for 50 min and poured into aq. sat. $NaHCO_3$ solution. Extraction with EtOAc (3 times), drying with $MgSO_4$

and concentration yielded the crude product. Column chromatography (EtOAc:Cychex; 1:6 → 1:3) gave the desired product as a colourless liquid in 23% yield (0.12 mmol, 49.8 mg): R_f(EtOAc:Cychex; 1:1) 0.55; ^{1}H-NMR ($CDCl_3$, 400 MHz): δ(ppm) = 5.36 (d, *J* = Hz, 1H), 5.18 (dd, *J* = Hz, 1H), 4.99 (dd, *J* = Hz, 1H), 4.51 (d, *J* = Hz, 1H), 4.18-4.07 (m, 2H), 3.94-3.89 (m, 2H), 3.65 (dt, *J* = Hz, 1H), 2.45 (dt, *J* = Hz, 2H), 2.11 (s, 3H), 2.09 (s, 1H), 2.03 (s, 3H), 2.02 (s, 3H), 1.95 (s, 3H), 1.22 (bs, OH); ^{13}C-NMR ($CDCl_3$, 100 MHz): δ(ppm) = 170.3, 170.2, 170.1, 169.4, 101.3, 80.6, 70.8, 70.7, 69.6, 68.6, 67.9, 67.0, 61.3, 20.8, 20.6, 20.6, 20.5, 19.8; HRMS: (ESI-ToF): m/z (calculated): $[M+Na]^+$ = 255.0839 ($C_{10}H_{16}O_6Na^+$), m/z (experimental): $[M+Na]^+$ = 255.0835. The analytical data is in accordance with the literature.[247]

Path II towards 83:

2,3,4,6-Tetra-*O*-acetyl-1-*O*-but-3-ynyl-β-galactopyranoside (83). A solution of per-acetylated galactose (**79**) (0.76 mmol, 300 mg) and 3-butyn-1-ol (**82**) (1.54 mmol, 120 µL) in dry CH_2Cl_2 was cooled down to -10 °C. TMS-OTf (3.84 mmol, 0.70 mL) was added dropwise. The reaction mixture was stirred for 2 h and poured into aq. sat. $NaHCO_3$-solution. Extraction with EtOAc (3 times), drying with $MgSO_4$ and concentration yielded the crude product. Column chromatography (EtOAc:Cyc; 1:4 → 1:2) gave the desired product **83** as a viscous oil in 44% yield (0.34 mmol, 136.7 mg).

Deprotection of 83 towards 84:



1-*O*-But-3-ynyl-β-galactopyranoside (84). To a solution of **83** (0.42 mmol, 170 mg) in 10 mL methanol was added NaOMe (32% in MeOH, 0.40 mL) and the reaction was stirred for 6 h at room temperature. Dowex-Exchange-Resin (3g) was added and stirring proceeded for 50 min. Filtration and concentration afforded the desired product **84** as a white solid in 93% yield (0.39 mmol, 91.1 mg): ^{1}H-NMR (CD_3OD, 400 MHz): δ(ppm) = 4.25 (d, *J* = 7.4 Hz, 1H), 3.96 (dt, *J* = 9.7, 7.3 Hz, 1H), 3.82 (dd, *J* = 3.2, 1.1 Hz, 1H), 3.79-3.66 (m, 3H), 3.54-3.48 (m, 2H), 3.46 (dd, *J* = 9.7, 3.2 Hz, 1H), 2.51 (td, *J* = 7.3, 2.7 Hz, 2H), 2.26 (t, *J* = 2.7 Hz, 1H); ^{13}C-NMR (CD_3OD, 100 MHz): δ(ppm) = 105.0, 81.8, 76.7, 74.9, 72.4, 70.6, 70.3, 69.0, 62.5, 20.6; HRMS: (ESI-ToF): m/z (calculated): $[M+Na]^+$ = 255.0839 ($C_{10}H_{16}O_6Na^+$), m/z (experimental): $[M+Na]^+$ = 255.0835.

5.5.3.2. Protocol for CuAAC on TTL(D221Bpa,S261C)[Aha]

CuAAC approach. A solution of the lipase was mixed with $CuSO_4$/THPTA (premixed), sodium ascorbate, aminoguanidine and 1-*O*-but-3-ynyl-β-galactopyranoside (**84**) and shaken overnight at 25 °C (for further conditions see Table 5.11). The solutions were centrifuge-filtered and washed three times with buffer/EDTA-solution (50 mM Tris/HCl; 200 mM NaCl; 5 mM EDTA; pH 8) and 4 times with Millipore water. For MALDI-MS measurements, different matrices were probed: A) SA (sinapinic acid), B) DHB (2,5-dihydroxybenzoic acid) and C) "Super-DHB" (DHB, 10% 2-hydroxy-5-methoxybenzoic acid).

Table 5.11: Conditions for CuAAC on TTL with butynyl galactose 84 in phosphate buffer (# Tris buffer,* pH 7.5).

entry	c [µM] (protein)	c [mM] (alkyne 84)	Cu^{2+} catalyst (to alkyne 84)	Na ascorbate (to alkyne 84)	THPTA (to alkyne 84)	aminoguanidine c [mM]
1#	100.0	110.00	3 mol%	4.5 mol%	–	1.2
2#	100.0	110.00	3 mol%	4.5 mol%	15 mol%	1.2
3	23.16	25.48	5 mol%	15 mol%	5 mol%	–
4*	23.16	25.48	5 mol%	15 mol%	5 mol%	–
5	23.16	25.48	10 mol%	30 mol%	50 mol%	–
6*	23.16	25.48	10 mol%	30 mol%	50 mol%	–
7	23.16	25.48	10 mol%	15 mol%	–	–
8	10.00	11.00	1 mol%	1.5 mol%	5 mol%	5.0

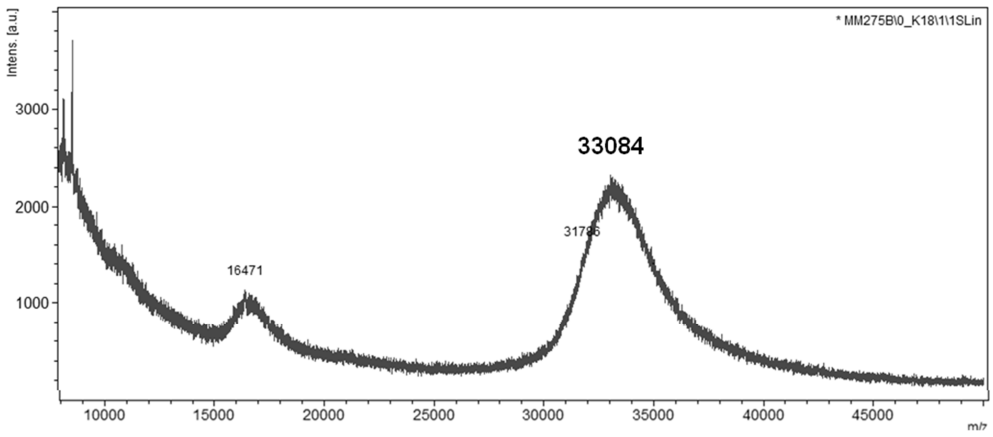


Figure 5.7: MALDI-MS after CuAAC with sinapinic acid as matrix (Table 5.11, entry 4).

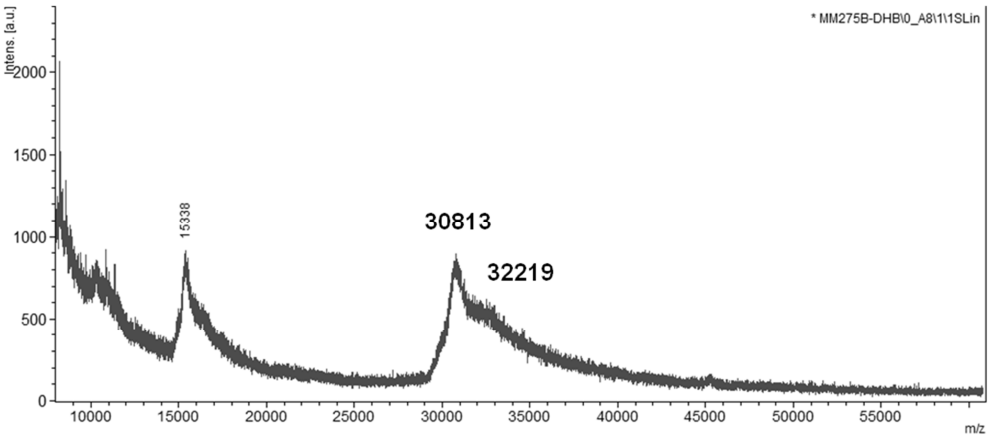


Figure 5.8: MALDI-MS after CuAAC with DHB as matrix (Table 5.11, entry 4).

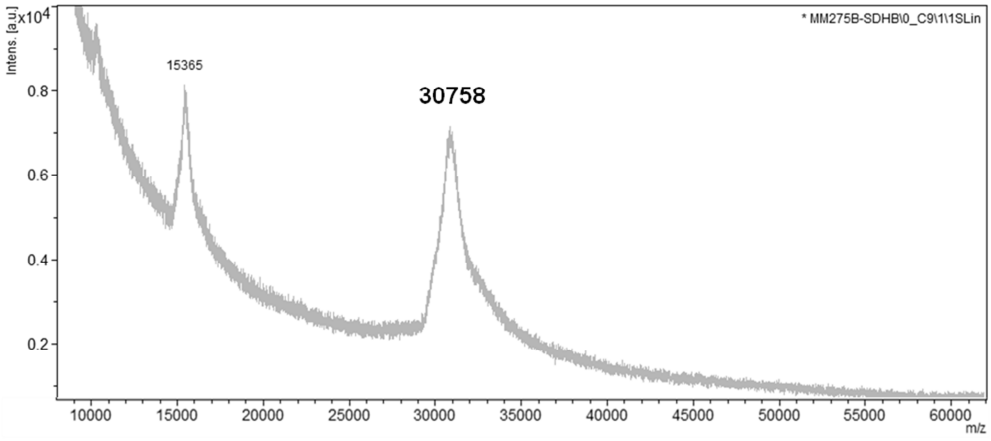
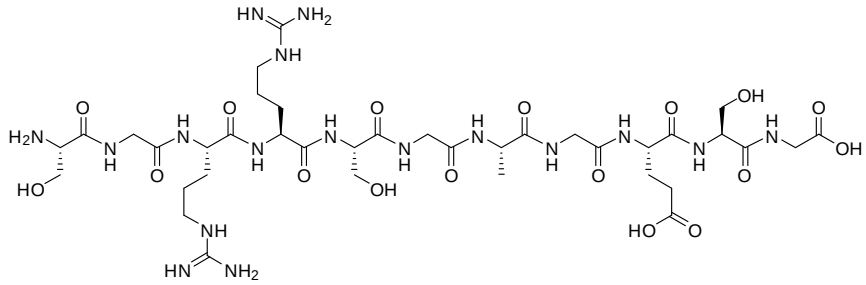


Figure 5.9: MALDI-MS after CuAAC with "Super-DHB" as matrix (Table 5.11, entry 4).

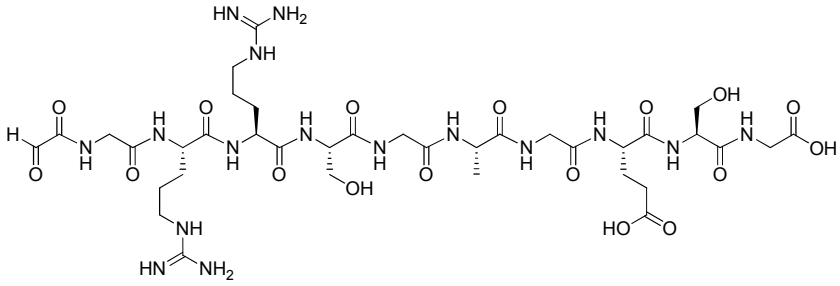
5.6. Protein Dual-Functionalisation on Ser-TTL[Aha]

5.6.1. Initial Experiments on Glycol Cleavage and Oxime Formation



Peptide synthesis. The *N*-terminal serine peptide **85** (H_2N-SGRRSGAGESG-OH) was synthesised by SPPS on a preloaded Fmoc-Gly-Wang resin (0.8 mmol/g) following the standard Fastmoc-protocol. It was cleaved and deprotected with 95% TFA (TIS/H_2O, 1:1) for 2 h, precipitated in cold ether and purified by semi-prep HPLC (14 mL/min, 5 min at 5% MeCN, in 45 min to 50% MeCN, t_r = 10 min). Pure fractions were lyophilised and combined to yield 15% (20.7 mg,

15.2 mmol) of the desired peptide: HRMS: (ESI-ToF): m/z (calculated): $[M+H]^+$ = 1020.4817 ($C_{37}H_{66}O_{17}N_{17}^+$), $[M+2H]^{2+}$ = 510.7445; m/z (experimental): $[M+2H]^{2+}$ = 510.7505.



Glycol cleavage to yield aldehyde peptide 86. Peptide **85** (0.8 mg, mmol) was dissolved in PBS buffer (0.1 M, pH 7) to a concentration of 2 mg/mL. $NaIO_4$ (4 µL, 0.25 M in H_2O) was added to the peptide solution. It was shaken for 5 min at rt. Reducing agent (5 µL, 1 M in H_2O) was added and the solution was shaken for 15 min at rt. Conversion was checked by MS:

A) $NaNO_2$ as reducing agent: LRMS: m/z (calculated): $[M+H]^+$ 1007.5 (hydrate (**86**+H_2O)), 1071.4 (hydrate (**86**+H_2O) +SO_2); m/z (experimental): $[M+H]^+$ (ESI-ToF) 1007.2 (hydrate (**86**+H_2O)), 1071.0 (hydrate (**86**+H_2O) +SO_2);

B) Ac-Met-OH as reducing agent: HRMS: m/z (calculated): $[M+H]^+$ 1007.5 (hydrate (**86**+H_2O)), m/z (experimental): $[M+H]^+$ (ESI-ToF) 1007.2 (hydrate (**86**+H_2O)); HRMS: (ESI-ToF): m/z (calculated): $[M+H]^+$ = 989.4395, $[M+2H]^{2+}$ = 504.2287 ($C_{36}H_{61}O_{17}N_{16}^+$), m/z (experimental): $[M+2H]^{2+}$ = 504.2285.

Oxime formation.

A) PBS: After glycol cleavage the peptide solution was mixed with **66** or **88** and the catalyst (if applied) and shaken overnight (see Table 5.12, discussion). Conversion was checked by LC-LRMS.

B) NH_4OAc: After glycol cleavage the solution was purified by analytical LCMS and lyophilised. The isolated peptide **86** was redissolved in the desired solvent system, mixed with reagent **66** or **88** and catalyst (if applied) and shaken overnight (see Table 6, discussion). Conversion was checked by LC-LRMS.

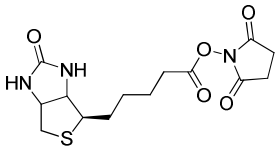
LRMS: **66**: m/z (calculated): $[M+H]^+$ 989.4 (aldehyde **86**), 1288.6 (product **87a**), 1062.5 (oxime w/o biotin **89**), $[M+2H]^{2+}$ 532.7 (imine w/ aniline); m/z (experimental): $[M+H]^+$ (ESI-ToF) $[M+H]^+$ 990.2 (aldehyde **86**), $[2M+H]^+$ 644.4 (product **87a**), 531.5 (oxime w/o biotin **89**), 532.5 (imine w/ aniline);

LRMS: **88**: m/z (calculated): $[M+H]^+$ 435.2 (hydroxylamine **88**), $[M+2H]^{2+}$ 504.2 (hydrate (**86**+H_2O)), 703.3 (product **87b**), m/z (experimental): $[M+H]^+$ (ESI-ToF) 435.1 (hydroxylamine **88**), $[M+2H]^{2+}$ 504.4 (hydrate **86**), 703.2 (product **87b**).

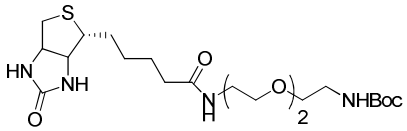
Table 5.12: Reaction conditions for oxime formation on peptide 86.

entry	reagent (eq)	solvent system	catalyst	T [°C]
1	**66** (3.0)	phosphate buffer (0.1 M, pH 7)	–	40
2	**66** (13)	NH_4OAc (0.1 M, pH 4.5), 20% MeCN	–	30
3	**66** (13)	NH_4OAc (0.1 M, pH 4.5), 20% MeCN	aniline (100 mM)	30
4	**88** (1.5)	phosphate buffer (0.1 M, pH 7)	–	40
5	**88** (10)	phosphate buffer (0.1 M, pH 7)	methoxyaniline (100 mM)	30
6	**88** (10)	phosphate buffer (0.1 M, pH 7)	*p*-amino Phe (10 mM)	10

5.6.2. Synthesis of Biotin Hydroxylamine 95

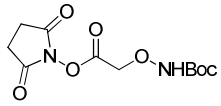


(+)-Biotin *N*-hydroxysuccinimide ester (91). To a solution of (+)-biotin (**90**) (0.82 mmol, 200 mg) and *N*-hydroxy succinimide (0.89 mmol, 102 mg) was added EDC·HCl (0.96 mmol, 184 mg). The reaction mixture was stirred overnight at room temperature and concentrated to give a white solid. The crude solid was mixed with isopropanol, heated up to 70 °C and cooled down. The pure product was filtered off as white powder in 91% yield (0.75 mmol, 255 mg): ^{1}H-NMR (DMSO-D_6, 400 MHz): δ(ppm) = 6.43 (s, 1H), 6.37 (s, 1H), 4.35-4.27 (m, 1H), 4.19-4.08 (m, 1H), 3.10 (dd, *J* = 11.8, 7.2 Hz, 1H), 2.92-2.78 (m, 5H), 2.67 (t, *J* = 7.4 Hz, 2H), 2.58 (d, *J* = 12.5 Hz, 1H), 1.71-1.34 (m, 6H); ^{13}C-NMR (DMSO-D_6, 100 MHz): δ(ppm) = 170.3, 169.0, 162.7, 61.0, 59.2, 58.9, 55.3, 30.0, 27.8, 27.6, 25.5, 24.3. The analytical data is in accordance with the literature.[248]

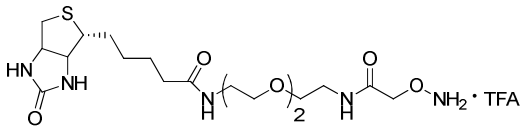


***tert*-Butyl-(2-(2-(2-(5-((3aS,4S,6aR)-2-oxohexahydro-1*H*-thieno[3,4-d]imidazol-4-yl)pentanamido)ethoxy)-ethoxy)ethyl)carbamate (93).** Biotin derivative **91** (0.29 mmol, 100 mg), Boc-1-amino-3,6-dioxa-8-octanamine (**92**) (0.44 mmol, 109 mg) and triethylamine (0.59 mmol, 81.0 µL) were dissolved in dry DMF (5 mL). The reaction mixture was stirred overnight at room temperature. DMF was removed, the residue taken up with CH_2Cl_2 and washed with water. The organic layer was dried with $MgSO_4$ and concentrated to yield the crude product. Column chromatography (CH_2Cl_2 + MeOH (slowly increasing from 0% to 7.5%)) yielded the desired product **3** in 86% (0.25 mmol, 118 mg) as a white sticky solid: R_f(CH_2Cl_2:MeOH; 9:1)

0.4; ^{1}H-NMR ($CDCl_3$, 400 MHz): δ(ppm) = 6.74 (s, 1H), 5.94 (s, 1H), 4.53-4.39 (m, 1H), 4.34-4.20 (m, 1H), 3.66-3.47 (m, 8H), 3.40 (dd, *J* = 10.1, 5.1 Hz, 2H), 3.33-3.22 (m, 2H), 3.11 (dd, *J* = 11.8, 7.1 Hz, 1H), 2.86 (dd, *J* = 12.8, 4.8 Hz, 1H), 2.71 (d, *J* = 12.8 Hz, 1H), 2.20 (t, *J* = 7.5 Hz, 2H), 1.79-1.54 (m, 4H), 1.49-1.32 (m, 10H), 1.27-1.17 (m, 1H); ^{13}C-NMR ($CDCl_3$, 100 MHz): δ(ppm) = 173.5, 164.4, 156.1, 79.4, 70.1, 61.9, 60.3, 55.8, 40.6, 40.4, 39.2, 36.1, 28.5, 28.4, 28.2, 25.7. The analytical data is in accordance with the literature.[249]



2,5-Dioxopyrrolidin-1-yl 2-(((tert-butoxycarbonyl)amino)oxy)acetate (94). 2-(((tert-butoxycarbonyl)amino)-oxy)acetic acid (1.50 mmol, 287 mg) and *N*-hydroxysuccinimide (1.65 mmol, 190 mg) were dissolved in dry DMF. EDC·HCl (1.80 mmol, 227 mg) was added and the reaction was stirred overnight at room temperature. The mixture was diluted with water and extracted with ethylacetate. The organic layer was dried and concentrated to yield the product as a yellow liquid with some DMF impurities (~59%, 255 mg). The product was applied in the next reaction step without further purification.



N-(2-(2-(2-(2-(Aminooxy)acetamido)ethoxy)ethoxy)ethyl)-5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (95). Biotin derivative **94** (0.24 mmol, 115 mg) was dissolved in CH_2Cl_2/TFA (4:1) and stirred for 1 h. The mixture was concentrated and dried under vacuum. The residue was redissolved in dry DMF with compound **94** (0.48 mmol, 1.40 mg) and triethylamine (0.58 mmol, 58.9 mg) and stirred at room temperature overnight. Concentration and purification by column chromatography (CH_2Cl_2 + MeOH (slowly increasing from 0% to 7%); R_f(CH_2Cl_2:MeOH; 9:1) 0.3) yielded the Boc-protected product. The compound was then redissolved in CH_2Cl_2/TFA (4:1) and stirred for 1 h. The mixture was concentrated and dried under vaccum to yield the desired product **95** as TFA salt (1:1.16) in 22% (53.0 µmol, 30.7 mg) as a white sticky solid: ^{1}H-NMR (D_2O, 400 MHz): δ(ppm) = 4.73-4.59 (m, 3H), 4.45 (bdd, *J* = 7.6, 4.6 Hz, 1H), 3.76-3.60 (m, 8H), 3.50 (t, *J* = 5.1 Hz, 2H), 3.45-3.31 (m, 3H), 3.02 (dd, *J* = 13.1, 4.8 Hz, 1H), 2.80 (d, *J* = 13.0 Hz, 1H), 2.30 (t, *J* = 7.1 Hz, 2H), 1.83-1.37 (m, 6H).; ^{13}C-NMR (D_2O, 100 MHz): δ(ppm) = 177.0, 168.8, 165.4, 162.9 (q, *J* = 35.3 Hz), 116.4 (q, *J* = 291.7 Hz), 71.8, 69.4, 68.9, 68.6, 39.7, 38.9, 38.7, 35.5, 27.9, 27.7, 25.1; ^{19}F-NMR (D_2O, 376 MHz): δ(ppm) = −75.53; HRMS:

(ESI-ToF): m/z (calculated): $[M+H]^+$ = 448.2224 ($C_{18}H_{34}N_5O_6S^+$), m/z (experimental): $[M+H]^+$ = 448.2224.

5.6.3. Dual-Modification on (Aha)S-TTL[Aha]

General protocol for glycol cleavage & oxime ligation on TTL: A solution of the TTL (12 µM; 100 mM phosphate buffer, 100 mM NaCl, pH 7.2) was mixed with $NaIO_4$ (3 eq.) and shaken for 1 h at 15 °C. *N*-Acetyl methionine (12 eq.) was added to the mixture and shaken for 1 h at 15 °C. The buffer was exchanged by centrifuge membrane filtration (14000 r/min). For the different buffers and catalysts see Table 5.13. Biotin hydroxylamine **95** was added to the protein solution and the mixture was shaken overnight at 15 °C. For MALDI-MS analysis, the solutions were centrifuge-filtered (14000 r/min) and washed 4 times with ammonium acetate solution (100 mM, pH 7) and 4 times with ultrapure water. The proteins were analyzed by MALDI-MS measurements (Figure 5.10) and by SDS-PAGE (Coomassie stain) and western blotting (streptavidin–peroxidase antibody, 1:1000) using a Mini-Protean Tetra cell system (BioRad): m/z (calculated): $[M+H]^+$ = 31517 Da (**Gal-0**) , 31245 Da (AhaS-TTL[Aha]); m/z (MALDI-MS ("Super-DHB")): $[M+H]^+$ = 31537 Da (**Gal-0**), 31252 Da (AhaS-TTL[Aha]).

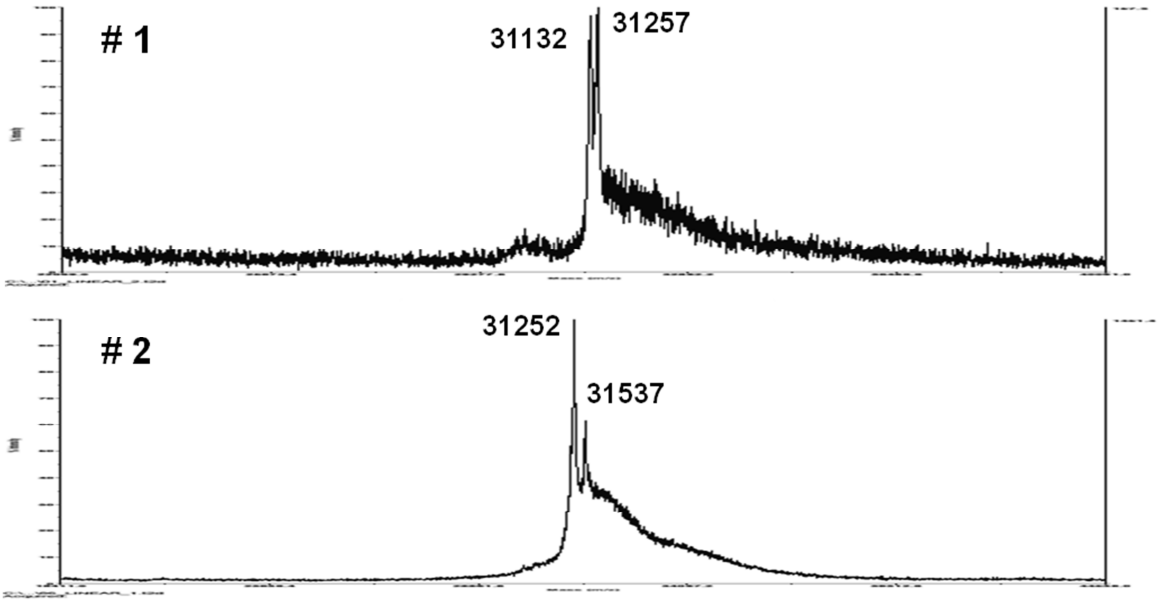


Figure 5.10: MALDI-MS spectra of (Aha)S-TTL[Aha] before (#1) & after (#2) glycol cleavage & oxime ligation with 95.

Table 5.13: Conditions for oxime ligation with hydroxylamine 95 on Ald-TTL[Aha].

entry	hydroxylamine 95	pH	catalyst (10 eq.)	conversion*
1	30 eq.	7.4	*p*-anisidine	10 %
2	50 eq.	5.0	–	50 %
3	30 eq.	5.0	–	30 %
4	30 eq.	4.5	–	60 %
5	30 eq.	4.0	–	65 %
6	30 eq.	3.5	–	80 %
7	30 eq.	3.0	–	100 %
8	20 eq.	3.0	–	100 %
9	10 eq.	3.0	–	80 %
10	5 eq.	3.0	–	55 %

† 50 µM, 15 °C, overnight. * Estimated by MALDI-MS in comparison to unreacted AhaS-TTL[Aha].

For subsequent dual-functionalisation, the samples were centrifuge-filtered with Dulbecco's PBS buffer (100 mM, pH 7) after oxime ligation and directly applied in the CuAAC.

General protocol for CuAAC on TTL: A solution of the TTL (10µM; 100 mM phosphate buffer, 100 mM NaCl, pH 7) was mixed with $CuSO_4$ (1 M in 100 mM phosphate buffer, 100 mM NaCl, pH 7), sodium ascorbate (50 eq. to Cu^{2+}) and 1-*O*-but-3-ynyl-β-galactopyranoside (**7**) (1100 eq. to protein), 80 µL THPTA (5 eq. to Cu^{2+}), and aminoguanidine (8 mM) and shaken overnight at 15 °C. For the different $CuSO_4$ concentrations see Table 5.14. The solutions were centrifuge-filtered (14000 r/min) and washed 3 times with buffer/EDTA-solution (100 mM phosphate buffer, 100 mM NaCl, 5mM EDTA, pH 7) and 4 times with ultrapure water. The proteins were analyzed by MALDI-MS measurements (Table 5.15 and Table 5.16) and by SDS-PAGE (Coomassie stain) and western blotting (streptavidin–peroxidase antibody, 1:1000) using a Mini-Protean Tetra cell system (BioRad) (see Figure 1 in manuscript). Protein concentrations were checked by UV (λ = 280 nm).

Table 5.14: Conditions for CuAAC on biotinylated TTL (Gal-0) (5 eq. THPTA to Cu^{2+}, 8mM aminoguanidine, 100 eq. alkyne 84 per azide, 50 eq. sodium ascorbate/Cu^{2+}, 15 °C, overnight).

entry	c [µM] (protein)	$CuSO_4$ (to alkyne 84)	product*
1	10	5 mol%	3-5 clicked sugars
2	10	10 mol%	3-4 clicked sugars (**Gal-3**)
3	10	30 mol%	1-2 clicked sugars (**Gal-1**)
4	10	50 mol%	1-2 clicked sugars

* Ratios judged by MALDI-MS in comparison to biotinylated TTL.

Table 5.15: Theoretical MALDI-ToF-MS data for CuAAC on Gal-0.

protein sample	$[M+H]^+$ (theoretical, average)
TTL lipase	31119 + (31245) Da
Biotin TTL (**Gal-0**)	(31245) + 31517 Da
Biotin TTL + 1 Gal	(31477) + 31749 Da
Biotin TTL + 2 Gal	(31709) + 31981 Da
Biotin TTL + 3 Gal	(31942) + 32214 Da
Biotin TTL + 4 Gal	(32174) + 32446 Da
Biotin TTL + 5 Gal	(32406) + 32678 Da

Table 5.16: Experimental MALDI-ToF-MS data for CuAAC on Gal-0.

entry	protein sample	$[M+H]^+$ (experimental, average)
# 1	TTL	31132 + (31257) Da
# 2	Biotin TTL (**Gal-0**)	(31252) + 31537 Da
# 3	Biotin (Gal-)TTL (Table 5.14, entry 1)	peak max.: ~ 32255
# 4	Biotin (Gal-)TTL (Table 5.14, entry 2) (**Gal-3**)	peak max.: ~ 32234
# 5	Biotin (Gal-)TTL (Table 5.14, entry 3) (**Gal-1**)	peak max.: ~ 31944
# 6	Biotin (Gal-)TTL (Table 5.14, entry 4)	peak max.: ~ 31875

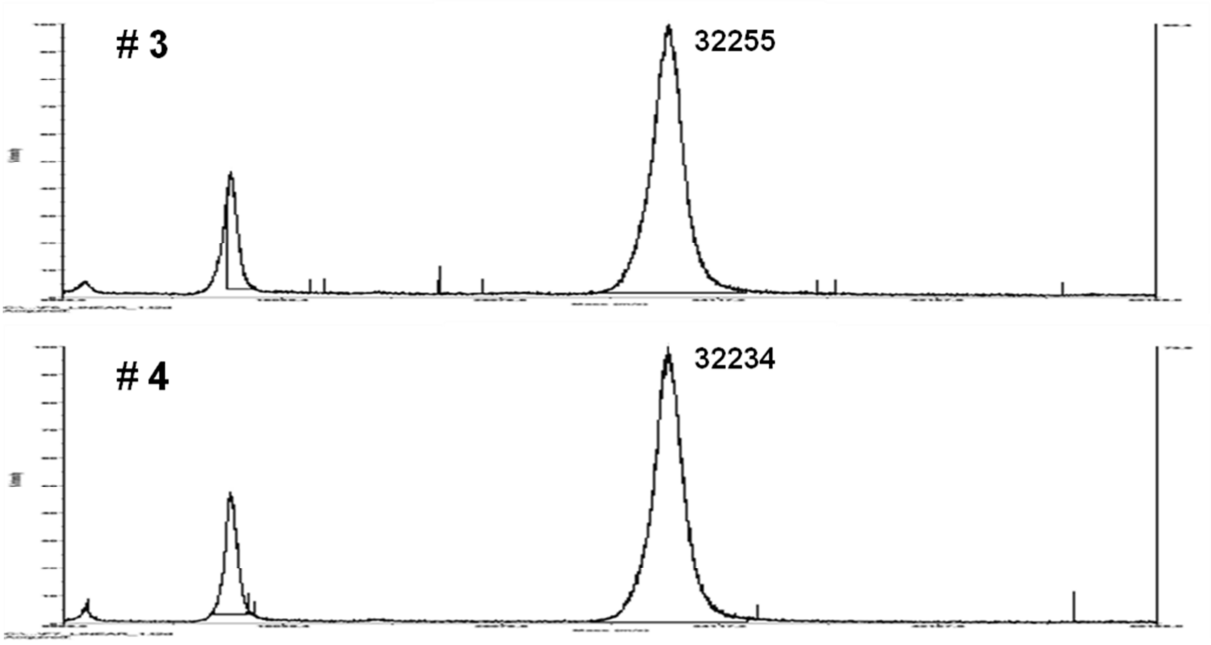


Figure 5.11: MALDI-MS spectra (see Table 5.16, entry 1 and 2).

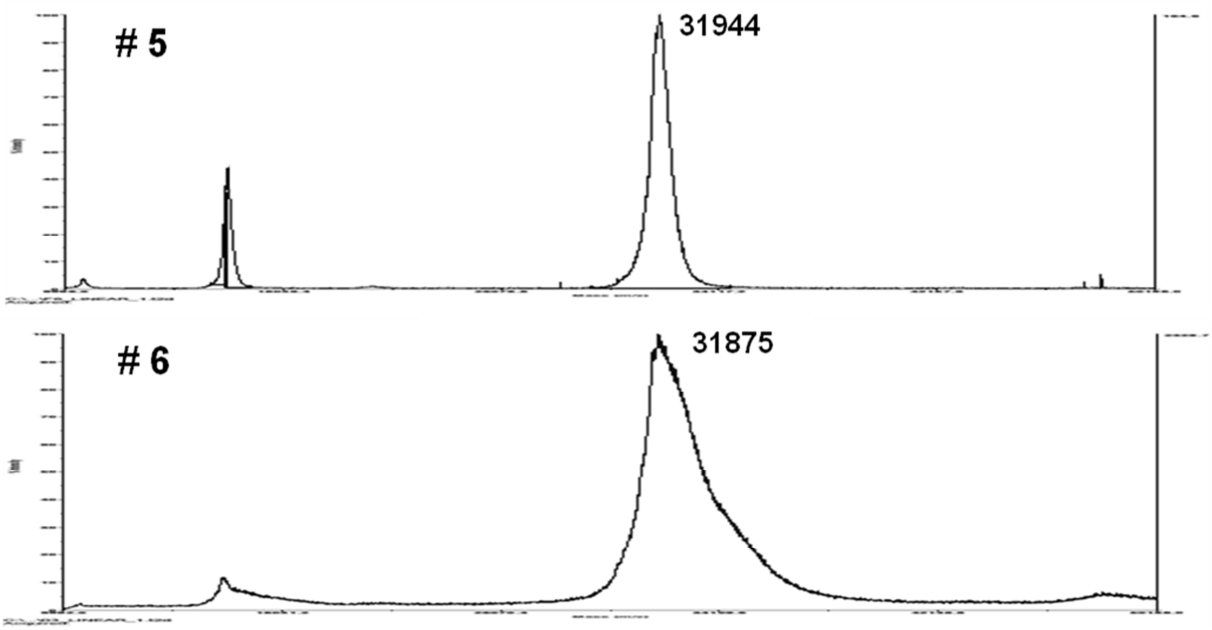
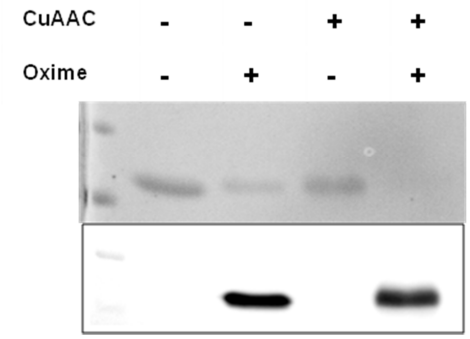


Figure 5.12: MALDI-MS spectra (see Table 5.16, entry 3 and 4).

SDS-PAGE (Coomassie stain) + Western Blot (streptavidin-peroxidase antibody):



lane 1: protein ladder

lane 2: Ser-TTL[Aha] & AhaSer-TTL[Aha] = (Aha)Ser-TTL[Aha]

lane 3: (Aha)Ser-TTL[Aha] + $NaIO_4$ + oxime ligation (with biotin hydroxylamine 2)

= **Gal-0** + AhaSer-TTL[Aha]

lane 4: (Aha)Ser-TTL[Aha] + CuAAC

lane 5: **Gal-0** + AhaSer-TTL[Aha] + CuAAC = **Gal-3**

Lipase activity test.[198b] Lipase activity was determined by measuring the hydrolysis of *p*-nitrophenyl palmitate (pNPP; Sigma). Solution A (10 mM *p*-nitrophenyl palmitate in 10 mL ethanol) and solution B (100 mg gummi arabicum in 90 mL Tris-HCl buffer (50 mM, pH 8) were mixed 1:9 and centrifuged (ultraturrax, 3 min, 20000 min^{-1}) to get solution C. For each

measurement, 450 µL solution C were mixed with 50 µL enzyme solution (0.13 nmol protein). The contribution of autohydrolysis was assessed by including a blank that contained the same volume of 50 mM Tris·HCl pH 8.0 instead of enzyme (background measurement). The samples were shaken at 50 °C for 1 h and measured by UV. Absorbance of released *p*-nitrophenol was measured at λ = 410 nm (Figure 5.13).

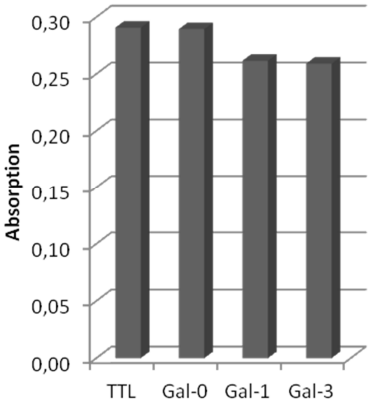
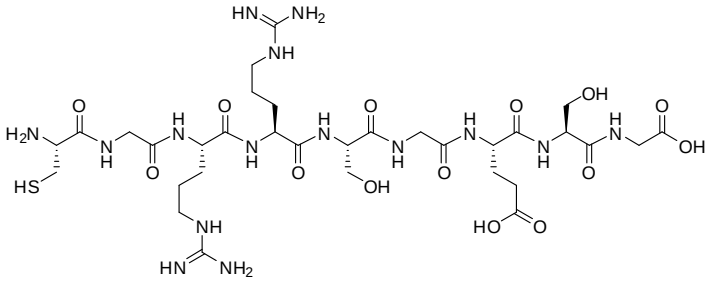


Figure 5.13: Lipase activity probed by the ester cleavage of 4-nitrophenyl palmitate.

5.6.4. Thiazolidine Formation on Ser-TTL[Aha] and Potential Cleavage



Peptide synthesis. Cysteine peptide **96** with the sequence H_2N-CGRRSGESG-OH was synthesised by SPPS on a preloaded Fmoc-Gly-Wang resin (0.7 mmol/g, 0.1 mmol) following the standard Fastmoc-protocol. It was cleaved/deprotected with 70% TFA (TIS/H_2O, 1:1) for 2 h, precipitated in cold ether and purified by preparative HPLC (35 mL/min, 15 min at 0% MeCN, in 35 min to 100% MeCN, t_r = 23.5 min). Pure fractions were lyophilised and combined to yield 9% (10.7 mg, 8.56 mmol) of the desired peptide: HRMS: (ESI-ToF): m/z (calculated): $[M+H]^+$ = 908.4003 ($C_{32}H_{58}O_{14}N_{15}^+$), m/z (experimental): $[M+H]^+$ = 908.4001.

Thiazolidine formation on Ser-TTL[Aha]:

A solution of (Aha)S-TTL[Aha] (0.42 mg/mL, 100 mM phosphate buffer, pH 7) was treated with $NaIO_4$ (3 eq.) and shaken for 1 h at 15 °C. *N*-Acetyl methionine (12 eq.) was added and the solution was shaken for 1 h at 15 °C. Afterwards, the buffer was exchanged by centrifuge

filtration to an ammonium acetate buffer (100 mM, pH 5.4, 3×). Cysteine peptide **96** (10, 20 or 30 eq.) was added to the protein solution (2.7 µM) and the mixture was shaken overnight at 15 °C. The solution was centrifuge-filtered and washed with ammonium acetate solution (100 mM, pH 7, 3×) and with Millipore water (4×). MALDI-MS measurements showed the following results: m/z (calculated): $[M+H]^+$ = 31088 Da (Ald-TTL[Aha]) , 31245 Da (AhaS-TTL[Aha]), 31977 Da (Thia-TTL[Aha]); m/z (MALDI-MS ("Super-DHB")): $[M+H]^+$ = 31076 Da (Ald-TTL[Aha]), 31291 Da (AhaS-TTL[Aha]), 32016 Da (Thia-TTL[Aha]).

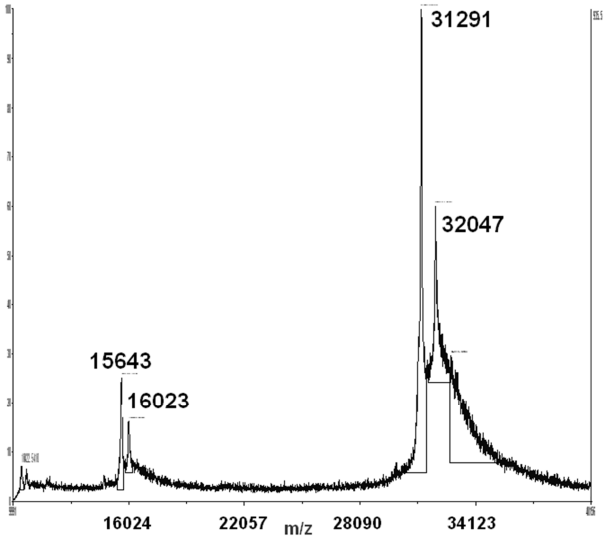


Figure 5.14: MALDI-MS spectrum after thiazolidine formation with 30 eq. of 96.

Thiazolidine cleavage on Ser-TTL[Aha]:

Table 5.17: Conditions probed for thiazolidine cleavage of Thia-TTL[Aha].

entry	buffer	pH	additive	T [°C]	t [h]
1	ammonium citrate (100 mM)	3	–	27	5
2	ammonium citrate (100 mM)	3	–	37	24
3	phosphate buffer (100 mM)	2	methoxyamine·HCl (0.2 M)	15	3
	phosphate buffer (100 mM)	4	methoxyamine·HCl (0.2 M)	25	3
4	phosphate buffer (100 mM)	2	methoxyamine·HCl (0.2 M)	37	4
5	phosphate buffer (100 mM)	2	methoxyamine·HCl (0.2 M)	37	16
6	phosphate buffer (100 mM)	2	methoxyamine·HCl (0.5 M)	37	4
7	phosphate buffer (100 mM)	2	methoxyamine·HCl (1.0 M)	37	4
8	sodium citrate (100 mM)	3	paraformaldehyde (0.8 M)	37	4
9	phosphate buffer (100 mM)	8	thiol* (0.2 M)	25	3
10	phosphate buffer (100 mM)	8	thiol* (0.4 M)	25	16
11	ammonium acetate (100 mM)	5	cysteine (0.2 M)	15	16
12	ammonium acetate (100 mM)	5	cysteine (0.5 M)	37	3

* thiol = sodium 3-mercaptopropane 1-sulfonate

A solution of the thiazolidine protein was centrifuge filtered to change the buffer if needed. Possible reagents were added and the mixture was shaken at a given temperature for a given

time (for conditions see Table 5.17). The solution was centrifuge filtered and washed with its buffer (3×) and with Millipore water (4×). Conversion was checked by MALDI-MS (“Super-DHB”): m/z (calculated): $[M+H]^+$ 31977 Da (Thia-TTL[Aha]); m/z (MALDI-MS (“Super-DHB”)): $[M+H]^+$ 32016 Da (Thia-TTL[Aha]).

5.7. Further Attempts Towards Triple-Modification of Proteins

5.7.1. Thiol-Michael Reaction on (Aha)S-TTL(D221C)[Aha]

Reducing agents except TCEP. A solution of (Aha)S-TTL(D221C)[Aha] (209 µL, 0.36 mg/mL) in 100 mM phosphate buffer (pH 7) was mixed with reducing agent (see Table 5.18, entry 1–9) and shaken for 3 h at 15 °C. The reducing agent was removed by centrifuge filtration with 100 mM phosphate buffer (for pH see Table 5.18). The protein solution was mixed with Rhodamine B maleimide **77** (4.9 g/L in Millipore water/acetonitrile, see Table 5.18); and the mixture was shaken overnight at 15 °C. The solution was centrifuge filtered and washed with 20% MeCN (in Millipore water, 3×) and with Millipore water (4×). MALDI-MS measurements showed the following results: m/z (calculated): $[M+H]^+$ = 30139 Da (S-TTL(D221C)[Aha]), 30718 Da (product – one fold reacted), 31297 Da (two fold reacted); m/z (MALDI-MS (“Super-DHB”)): $[M+H]^+$ = 29922 Da (TTL(D221C)[Aha]), 30460 Da (product – one fold reacted), 31010 Da (two fold reacted).

TCEP as reducing agent. A solution of (Aha)S-TTL(D221C)[Aha] (209 µL, 0.36 mg/mL) in 100 mM phosphate buffer (pH 7) was mixed with TCEP (see Table 5.18, entry 10–13) and Rhodamine B maleimide **77** (4.9 g/L in Millipore water/acetonitrile, see Table 5.18); and the mixture was shaken overnight at 15 °C. The solution was centrifuge filtered and washed with 20% MeCN (in Millipore water, 3×) and with Millipore water (4×). MALDI-MS measurements showed the following results: m/z (calculated): $[M+H]^+$ = 30139 Da (S-TTL(D221C)[Aha]), 30718 Da (product – one fold reacted); m/z (MALDI-MS (“Super-DHB”)): $[M+H]^+$ = 29941 Da (TTL(D221Bpa, S261C)[Aha]), 30435 Da (product – one fold reacted).

Table 5.18: Reaction conditions probed for thiol-Michael reaction of the TTL and maleimide 77 (15 °C, overnight).

entry	c [µM] (protein)	maleimide 77 [eq.]	pH	reducing agent [mM]
1	10	10	7	5.0 (DTT)
2	10	5	7	5.0 (DTT)
3	10	5	7	1.0 (DTT)
4	10	5	7	0.2 (DTT)
5	10	5	7	0.02 (DTT)
6	10	5	7	1.0 (*β*-mercaptoethanol)
7	10	5	7	1.0 (glutathione)
8	10	5	6.5	0.2 (DTT)
9	10	2	6.5	0.2 (DTT)
10	10	2	6.5	1.0 (TCEP, in situ)
11	10	2	6.5	0.2 (TCEP, in situ)
12	10	5	6.5	0.2 (TCEP, in situ)
13	10	13	6.5	0.2 (TCEP, in situ)

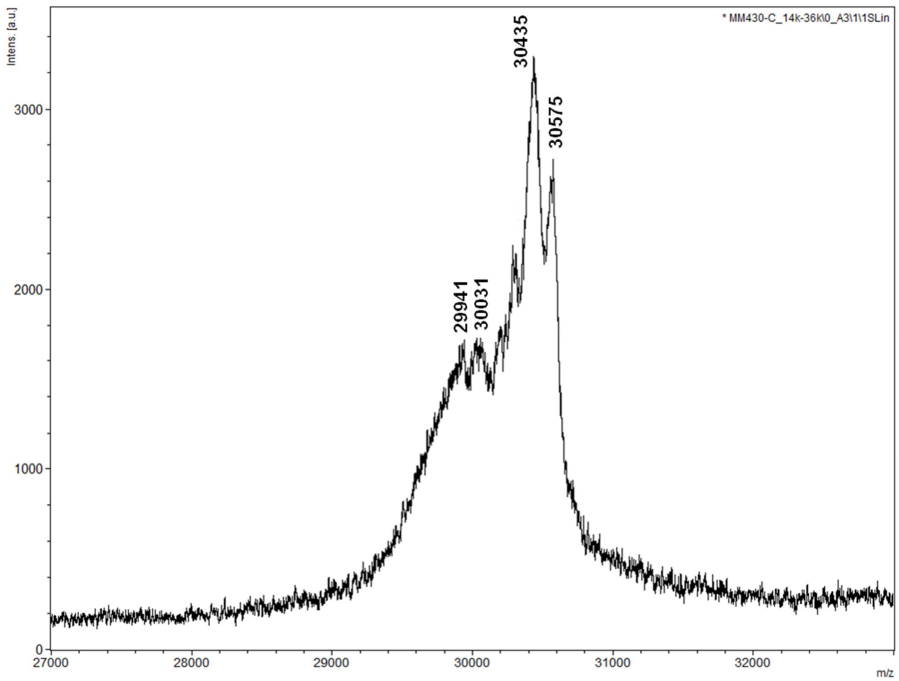


Figure 5.15: MALDI-MS spectra of thiol-Michael reaction with 77 on TTL (200 µM TCEP, 5 eq. of 77, Table 5.18, entry 12).

5.7.2. CuAAC on (Aha)S-TTL(D221C)[Aha]

A solution of the TTL was mixed with $CuSO_4$/THPTA (premixed), sodium ascorbate, aminoguanidine and 1-*O*-but-3-ynyl-β-galactopyranoside (**84**) and shaken overnight at 15 °C (for further conditions see Table 5.19). The solutions were centrifuge-filtered and washed three times with buffer/EDTA-solution (50 mM Tris/HCl; 200 mM NaCl; 5 mM EDTA; pH 8) and 4 times with Millipore water. MALDI-MS measurements showed the following results: m/z (calculated): $[M+H]^+$ = 30202 Da ((Aha)S-TTL(D221C)[Aha]), 30371 Da (1× sugar), 30603 Da (2× sugar),

30835 Da (3× sugar), 31067 Da (4× sugar), 31300 Da (5× sugar), 31532 Da (6× sugar), 31764 Da (7× sugar), 31996 Da (8× sugar), 32228 Da (9× sugar), 32460 Da (10× sugar), 32692 Da (11× sugar); m/z (MALDI-MS ("Super-DHB")): $[M+H]^+$ = 30214 Da ((Aha)S-TTL(D221C)[Aha]) (for further results see Table 3.7).

Table 5.19: Reaction conditions probed for CuAAC on (Aha)S-TTL(D221C)[Aha] (15 °C, 100 eq. alkyne 84 (per azide), 5 eq. THPTA (per Cu^{2+}), 1.2 mM aminoguanidine, overnight).

entry	c [µM] (protein)	Cu_2SO_4 (to alkyne)	sodium ascorbate (to alkyne)	m/z $[M+H]^+$ (MALDI-MS)	product*
1	10	10 mol%	50 eq. (sequential, 4 parts)	30710–31193 Da	2-4 clicked sugars
2	10	35 mol%	50 eq. (at once)	30453/30776 Da	1-2 clicked sugars
3	10	35 mol%	50 eq. (sequential, 4 parts)	30902–32700 Da (max. 31245 Da)	4-11 clicked sugars
4	10	50 mol%	50 eq. (sequential, 4 parts)	30918/31242 Da	3-4 clicked sugars
5	5	10 mol%	50 eq. (sequential, 4 parts)	30102 Da (-His_6)	no reaction
6	5	35 mol%	50 eq. (sequential, 4 parts)	30187 Da	no reaction

* estimated by MALDI-MS.

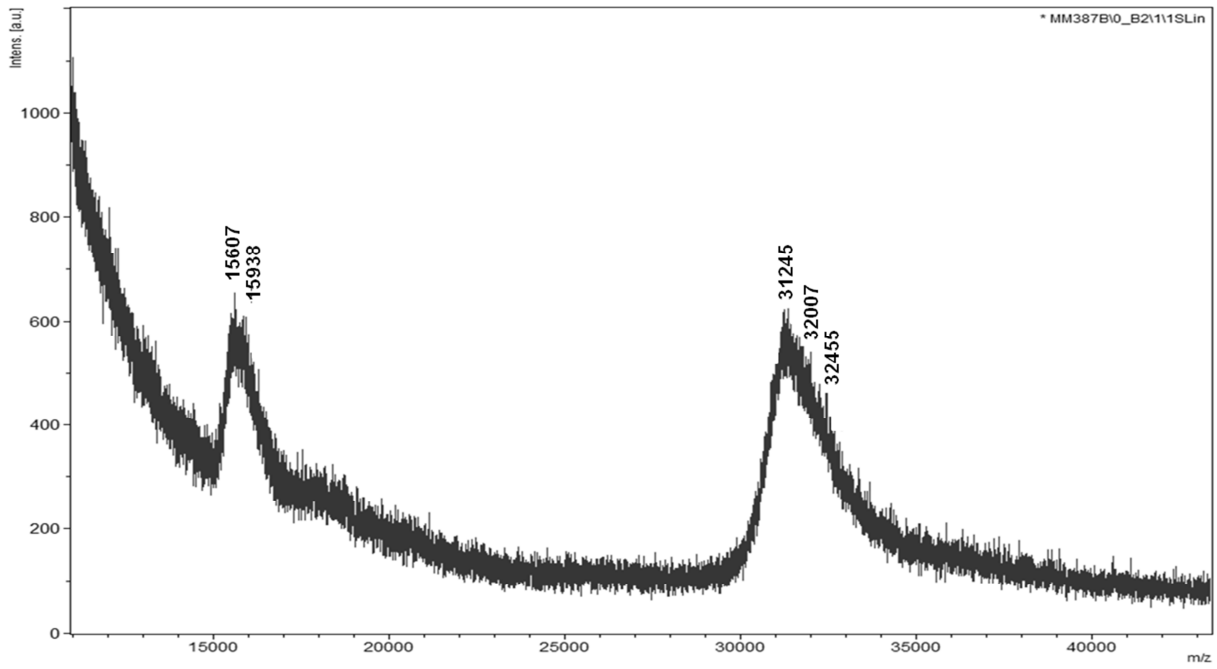


Figure 5.16: MALDI-MS spectrum after CuAAC on (Aha)S-TTL(D221C)[Aha] with alkyne galactose 84 (Table 5.19, entry 3).

5.7.3. Dual-Modification on Green Fluorescent Protein

5.7.3.1. CuAAC on GFP[Aha]-His_6

A solution of the GFP (0.25 mg/mL, 9.12 µM) was mixed with $CuSO_4$/THPTA (premixed), sodium ascorbate, aminoguanidine and 1-*O*-but-3-ynyl-β-galactopyranoside (**84**) and shaken overnight at 15 °C (for further conditions see Table 5.20). The solutions were centrifuge-filtered and

washed three times with buffer/EDTA-solution (50 mM Tris/HCl; 200 mM NaCl; 5 mM EDTA; pH 8) and 4 times with Millipore water. MALDI-MS measurements showed the following results: m/z (calculated): $[M+H]^+$ = 27410 Da (S-GFP[Aha]-His_6), 27642 Da (1× sugar), 27874 Da (2× sugar), 28106 Da (3× sugar); m/z (MALDI-MS ("Super-DHB")): $[M+H]^+$ = 27457 Da (S-GFP[Aha]-His_6) (for further results see Table 5.20).

Table 5.20: Reaction conditions probed for CuAAC on GFP[Aha]-His_6 with butinyle galactose 84 (9.1 µM protein, 50 eq. sodium ascorbate (to Cu^{2+}), 5 eq. THPTA (to Cu^{2+}), 1.2 mM aminoguanidine, 15 °C, overnight).

entry	alkyne 84 (per azide)	Cu_2SO_4 (to alkyne 84)	m/z: $[M+H]^+$ (MALDI-MS)
1	10 eq.	10 mol%	27738 Da
2	30 eq.	10 mol%	27907 Da
3	50 eq.	10 mol%	27985 Da
4	30 eq.	30 mol%	28213 Da
5	30 eq.	50 mol%	27951 Da
6	30 eq.	100 mol%	27957 Da

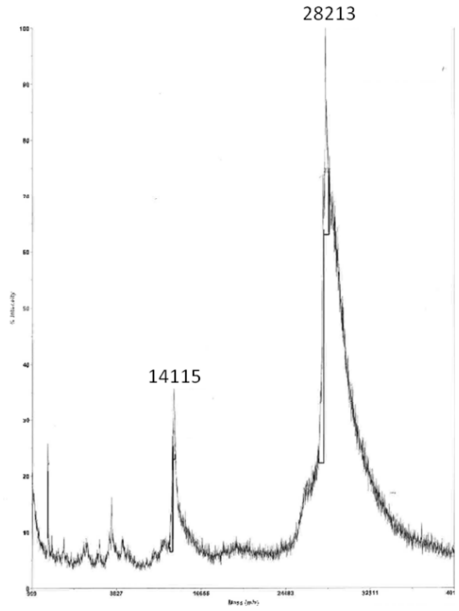
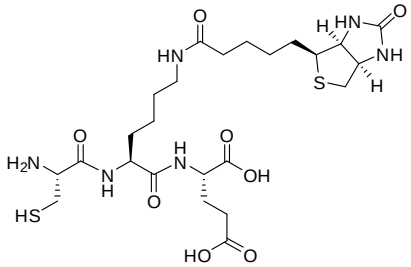


Figure 5.17: MALDI-MS spectrum of CuAAC on GFP[Aha]-His_6 (Table 5.20, entry 4).

5.7.3.2. Thiazolidine Formation on S-GFP[Aha]



Peptide synthesis. Cysteine biotin peptide **97** with the sequence H_2N-CK(Biotin)E-OH was synthesised by SPPS on a preloaded Fmoc-Glu(^{t}Bu)-Wang resin (0.58 mmol/g) following the standard Fastmoc-protocol. It was cleaved/deprotected with 95% TFA (TIS/H_2O, 1:1) for 3 h, precipitated in cold ether and purified by preparative HPLC (35 mL/min, 10 min at 0% MeCN, in 50 min to 100% MeCN t_r = 21.3 min) . Pure fractions were lyophilised and combined to yield 34% (24.1 mg, 33.53 µmol, TFA salt) of the desired peptide: HRMS: (ESI-ToF): m/z (calculated): $[M+H]^+$ = 605.2422 ($C_{24}H_{41}O_8N_6S_2^+$), m/z (experimental): $[M+H]^+$ = 605.2419.

Protein modification. A solution of S-GFP[Aha] (0.25 mg/mL, 100 mM phosphate buffer, pH 9) was treated with $NaIO_4$ (3 eq.) and shaken for 1 h at 15 °C. *N*-Acetyl methionine (12 eq.) was added and the solution was shaken for 1 h at 15 °C. Afterwards, the buffer was exchanged to the reaction buffer by centrifuge filtration to an (3×). Cysteine biotin peptide **97** was added to the protein solution (60 µM) and the mixture was shaken overnight at 15 °C (for further details see Table 5.21). The solution was centrifuge-filtered and washed with ammonium acetate solution (100 mM, pH 7, 3×) and with Millipore water (4×). MALDI-MS measurements showed the following results: m/z (calculated): $[M+H]^+$ = 26703 Da (S-GFP[Aha]), 27258 Da (Thia-GFP[Aha]); m/z (;ALDI-MS ("Super-DHB")): $[M+H]^+$ = 26598 Da (S-GFP[Aha]), 27169 Da (Thia-GFP[Aha]).

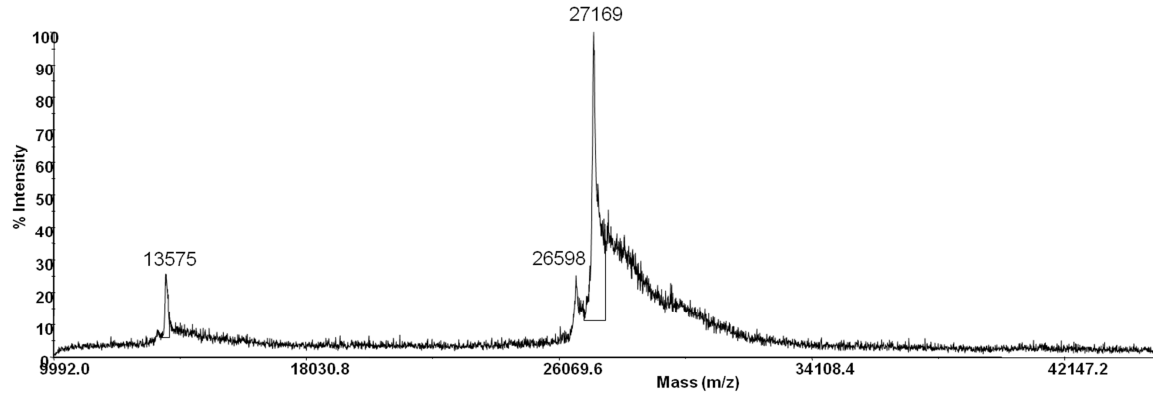


Figure 5.18: MALDI-MS spectrum of Thia-GFP[Aha] (Table 5.21, entry 5).

5.7.3.3. Oxime Formation on S-GFP[Aha]

A solution of S-GFP[Aha] (0.25 mg/mL, 100 mM phosphate buffer, pH 9) was treated with $NaIO_4$ (3 eq.) and shaken for 1 h at 15 °C. *N*-Acetyl methionine (12 eq.) was added and the solution was shaken for 1 h at 15 °C. Afterwards, the buffer was exchanged to the reaction buffer by centrifuge filtration to an (3×). Biotin hydroxylamine **95** was added to the protein solution (60 µM) and the mixture was shaken overnight at 15 °C (for further details see Table 5.21). The solution was centrifuge-filtered and washed with ammonium acetate solution (100 mM, pH 7, 3×) and with Millipore water (4×). MALDI-MS measurements showed the following results: m/z (calculated): $[M+H]^+$ = 26703 Da (S-GFP[Aha]), 27101 Da (Oxime-GFP[Aha]); m/z (MALDI-MS ("Super-DHB")): $[M+H]^+$ = 26608 Da (S-GFP[Aha]), 27003 Da (Oxime-GFP[Aha]).

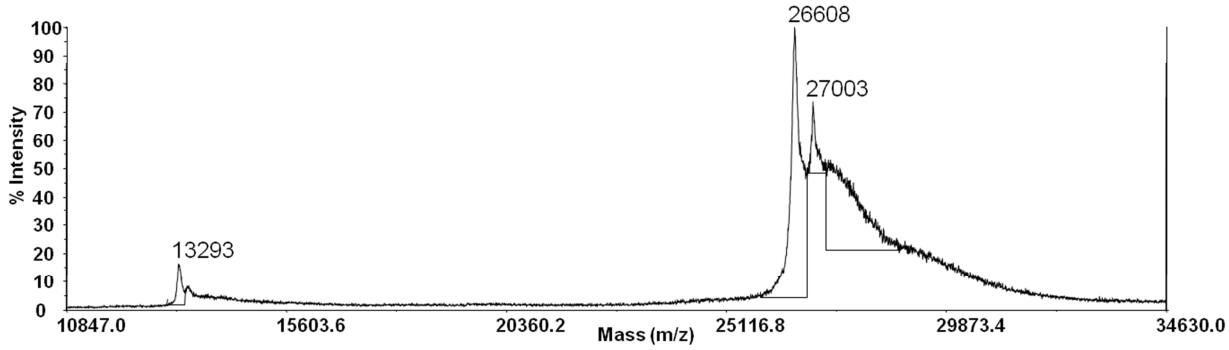


Figure 5.19: MALDI-MS spectrum of Oxime-GFP[Aha] (Table 5.21, entry 7).

Table 5.21: Reaction conditions probed for oxime and thiazolidine formation on S-GFP[Aha] (15 °C, overnight).

entry	hydroxylamine 95	cysteine peptide 97	buffer / catalyst	product formation*
1	30 eq.	–	NH_4OAc (100 mM, pH 3.0)	–
2	30 eq.	–	NH_4OAc (100 mM, pH 5.4)	–
3	50 eq.	–	NH_4OAc (100 mM, pH 5.0)	30–40%
4	–	30 eq.	NH_4OAc (100 mM, pH 5.4)	80–90%
5	–	50 eq.	NH_4OAc (100 mM, pH 5.4)	> 90%
6	–	30 eq.	phosphate buffer (100 mM, 100 mM NaCl, pH 7)	60–70%
7	30 eq.	–	phosphate buffer (100 mM, 100 mM NaCl, pH 7) / 10 mM aniline	50%
8	–	30 eq.	phosphate buffer (100 mM, 100 mM NaCl, pH 7) / 10 mM aniline	80–90%

* Product formation estimated by MALDI-MS for qualitative comparison.

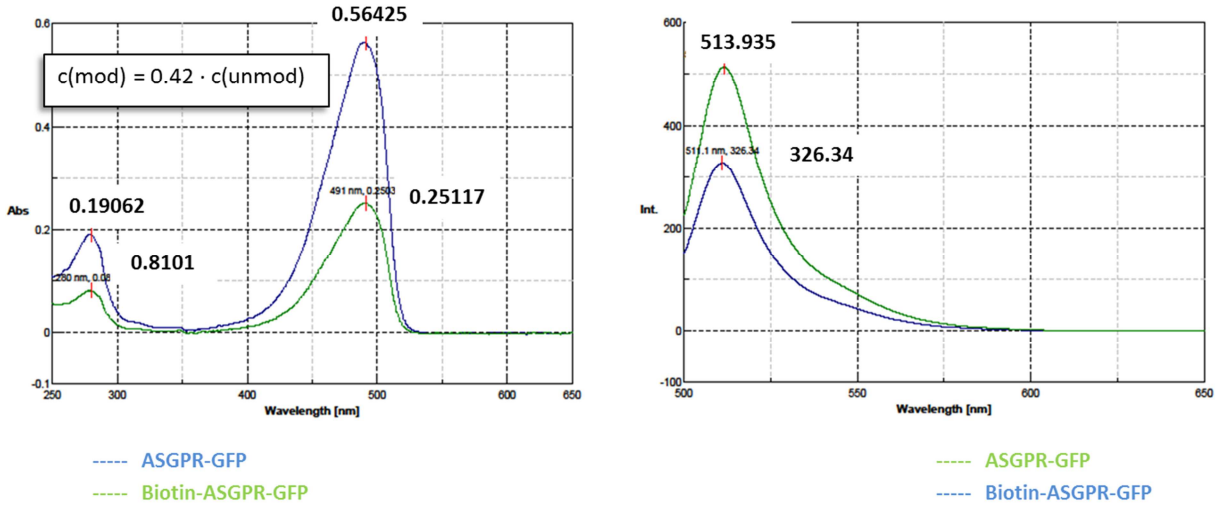


Figure 5.20: UV (left) and fluorescence spectra (right, ex. 491 nm, em. 511 nm) of Thia-GFP[Aha] (Table 6.16, entry 5).

5.7.3.4. Dual-Modification on S-GFP[Aha]

A solution of S-GFP[Aha] (0.58 mg/mL, 100 mM phosphate buffer, pH 9) was treated with $NaIO_4$ (3 eq.) and shaken for 1 h at 15 °C. *N*-Acetyl methionine (12 eq.) was added and the solution was shaken for 1 h at 15 °C. Afterwards, the buffer was exchanged to ammonium acetate buffer (100 mM, pH 5.4) by centrifuge filtration (3×). Cysteine biotin peptide **97** (100 eq.) was added to the protein solution (90 µM) and the mixture was shaken overnight at 15 °C. The buffer was exchanged to 100 mM phosphate buffer (pH 7) by centrifuge filtration. The protein solution was mixed with $CuSO_4$/THPTA (premixed), sodium ascorbate, aminoguanidine and 1-*O*-but-3-ynyl-β-galactopyranoside (**84**) and shaken overnight at 15 °C (for further conditions see Table 5.22). The solutions were centrifuge-filtered and washed three times with buffer/EDTA-solution (50 mM Tris/HCl; 200 mM NaCl; 5 mM EDTA; pH 8) and 4 times with Millipore water. MALDI-MS measurements showed the following results: m/z (calculated): $[M+H]^+$ = 27258 Da (Thia-GFP[Aha]), 27490 Da (Thia-GFP[Aha] +1× sugar). For experimental results see Table 5.22.

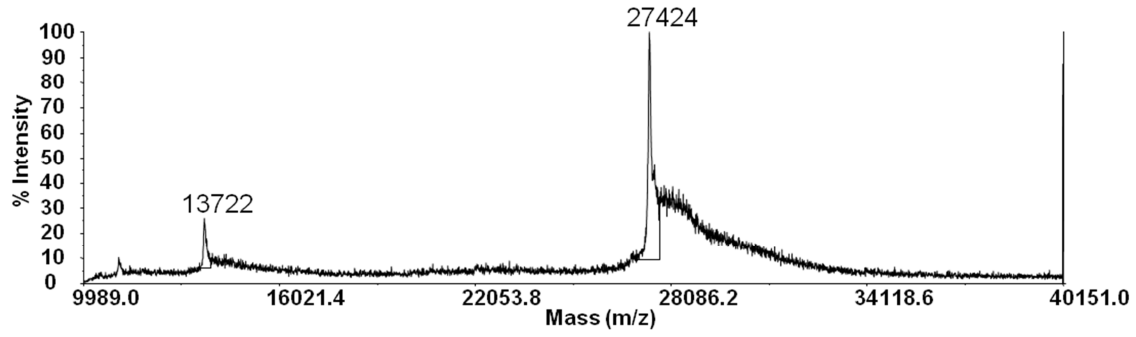


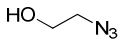
Figure 5.21: MALDI-MS after thiazolidine formation and CuAAC of S-GFP[Aha] (Table 5.22, entry 1).

Table 5.22: Conditions probed for CuAAC in dual-functionalisation of S-GFP[Aha] (11.5 µM protein, 50 eq. ascorbate (to Cu^{2+}), 5 eq. THPTA (to Cu^{2+}), 0.6 mM aminoguanidine, 15 °C, overnight).

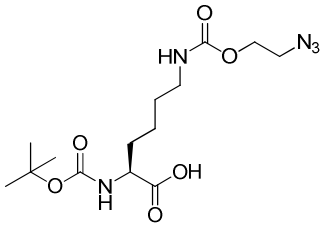
entry	alkyne 84 (to azide)	$CuSO_4$ (to alkyne)	m/z: $[M+H]^+$ (MALDI-MS ("Super-DHB"))
1	10 eq.	10 mol%	27424 Da
2	20 eq.	5 mol%	27419 Da
3	30 eq.	30 mol%	27423 Da
4	50 eq.	15 mol%	27416 Da

5.8. Synthesis of Pyrrolysine Derivatives

5.8.1. *S*-2-Amino-6-((2-azidoethoxy)carbonylamino)hexanoic acid (98)

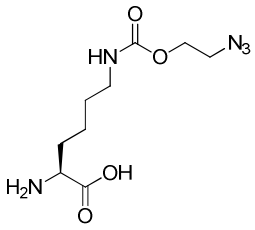


2-Azidoethanol (104). 2-Bromoethanol (**103**) (6.0 g, 48 mmol) was dissolved in 10 mL H_2O. Sodium azide (6.3 g, 96 mmol) was added and the mixture was stirred for 18 h at 40°C. The reaction was cooled down to rt and diluted with 10 mL H_2O. The aqueous phase was extracted with CH_2Cl_2 (3 x 10 mL). The organic layers were combined, dried over $MgSO_4$, filtered and concentrated. Purification by column chromatography (Cychex:EtOAc, 5:1, R_f = 0.1) yielded the product as a clear liquid (617 mg, 15%): ^{1}H-NMR ($CDCl_3$, 400 MHz) δ = 3.77 (t, *J* = 4.8 Hz, 2H), 3.43 (t, *J* = 4.8 Hz, 2H), 2.26 ppm (bs, OH). The analytical data was in agreement with the literature.[250]



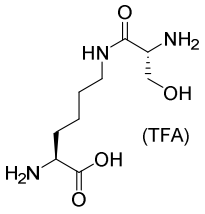
***S*-15-Azido-2,2-dimethyl-4,12-dioxo-3,13-dioxa-5,11-diazapentadecane-6-carbox-ylic acid (107).** 2-Azidoethanol (**104**) (500 mg, 5.74 mmol) was dissolved in dry THF and cooled to 0 °C. Triphosgene (0.17 g, 5.74 mmol) was added. The reaction was allowed to warm to rt and stirred overnight. It was concentrated to dryness and redissolved in 2.5 mL dry THF. The solution was added dropwise to a second solution of Boc-L-Lys-OH **106** in 20 mL NaOH (1 M) and 5 mL THF at 0 °C. After stirring overnight at rt, The reaction was again cooled to 0 °C and the pH was adjusted to 2 with 1 M HCl. The aqueous phase was extracted with EtOAc. The combined organic phases were washed with brine, dried and concentrated to yield the clean product as a white solid (1.83 g, 89%): ^{1}H-NMR ($CDCl_3$, 400 MHz) δ = 5.34-5.28 (m, 1H), 4.30-4.27 (m, 1H), 4.23-4.21 (m,

1H), 3.46-3.43 (m, 2H), 3.21-3.15 (m, 2H), 1.83-1.43 ppm (m, 15H). The analytical data was in agreement with the literature.[135e]



***S*-2-Amino-6-((2-azidoethoxy)carbonylamino)hexanoic acid (TFA salt) (98).** Boc-L-Lys(C(O)OCH$_2$CH$_2$N$_3$) (**107**) (1.83 g, 5.09 mmol) was dissolved in 50% TFA in CH_2Cl_2 and stirred for 30 min at rt. The solution was concentrated under vacuum. The residue was re-dissolved in 10 mL MeOH and poured into 100 mL Et_2O. White precipitate was formed, filtered and dried under vacuum. Purification on a preparative HPLC (35 mL/min; 5% to 100% MeCN in 55 min, t_r = 15 min) and lyophilization yielded the desired product as a white foam (403.8 mg, 21%): ^{1}H-NMR (D_2O, 400 MHz) δ = 4.17-4.15 (m, 2H), 3.95 (t, *J* = 6.3 Hz, 1H), 3.46-3.43 (m, 2H), 3.09-3.06 (m, 2H), 2.00-1.80 (m, 2H), 1.52-1.31 ppm (m, 4H); ^{13}C-NMR (D_2O, 400 MHz) δ = 172.4, 162.9, 162.6, 158.1, 117.8, 114.9, 63.8, 53.1, 49.9, 39.9, 29.5, 28.4, 21.5 ppm; HRMS: m/z (calculated): $[M+H]^+$ 260.1353, m/z (ESI-ToF): $[M+H]^+$ 260.1385. The analytical data was in agreement with the literature.[135e]

5.8.2. Synthesis of L-Lys(D-Ser) (99)

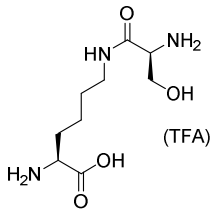


On-resin approach with Alloc-deprotection. Chlorotrityl resin (1.4 mmol/g, 0.1 mmol) was reacted with Fmoc-L-Lys(Alloc)-OH (4 eq., 0.4 mmol) and DIPEA (4 eq.) in CH_2Cl_2 for 2 h at room temperature. The resin was washed and treated with a solution of MeOH/CH_2Cl_2/DIPEA (5:80:15) for 15 min two times. After thorough washing, the resin was mixed with $Pd(PPh_3)_4$ (cat) and 100 µL $PhSiH_3$ in CH_2Cl_2 and shaken for 1 h. After thorough washing with CH_2Cl_2 and DMF, the resin was treated with a solution of Fmoc-D-Ser(O^{t}Bu)-OH (**D-108**) (10 eq.), HBTU (10 eq.), HOBt (10 eq.) and DIPEA (10 eq.) in DMF and shaken for 2 h. After thorough washing and Fmoc-deprotection with 20% piperidine in DMF for 30 min, the resin was treated with 95% TFA (TIS:H_2O, 2.5% each) for 2.5 h. The crude product was precipitated in cold ether, filtered, dried

and purified by preparative HPLC (35 mL/min, 0% MeCN for 15 min, t_r = 7 min) to yield the desired product as a white powder (4.9 mg, 13.2 µmol, 29%, TFA salt).

On-resin approach with ivDde-deprotection. Preloaded Fmoc-L-Lys(ivDde)-Wang resin (0.54 mmol/g, 0.1 mmol) was swollen in DMF for 30 min. The resin was treated with 20% piperidine in DMF (2× 15 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was mixed with a solution of Boc_2O (1.5 eq.) and DIPEA (1.5 eq.) in NMP and shaken for 1 h (2×). The resin was washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 4% $N_2H_4 \cdot H_2O$ in DMF (2× 10 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was mixed with a solution of Fmoc-D-Ser(^{t}Bu)-OH (**D-108**) (10 eq.), HATU (10 eq.), HOBt (10 eq.) and DIPEA (10 eq.) in DMF and shaken for 5 h. The resin was washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 20% piperidine in DMF (2× 15 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 95% TFA (TIS:H_2O, 2.5% each) for 2.5 h. The crude product was precipitated in cold ether, filtered, dried and purified by preparative HPLC (35 mL/min, 0% MeCN for 20 min, t_r = 7 min(MM524)) to yield the desired product as a white powder (14.7 mg, 63.0 µmol, 63%, 115.8 µmol TFA): ^{1}H-NMR (D_2O, 300 MHz) δ = 4.11-3.79 (m, 4H), 3.36-3.10 (m, 2H), 2.06-1.80 (m, 2H), 1.67-1.29 ppm (m, 4H); ^{13}C-NMR (D_2O, 75 MHz) δ = 172.4, 167.4, 162.9 (q, *J* = 35.2 Hz), 116.3 (q, *J* = 291.3 Hz), 60.2, 54.5, 53.0, 39.0, 29.3, 27.7, 21.5 ppm (s); ^{19}F-NMR (D_2O, 282 MHz) δ = -75.64 ppm; HRMS: m/z (calculated): $[M+H]^+$ 234.1448, m/z (ESI-ToF): $[M+H]^+$ 234.1416.

5.8.3. Synthesis of L-Lys(L-Ser) (100)

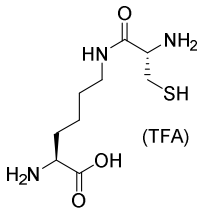


On-resin approach with Alloc-deprotection. Chlorotrityl resin (1.4 mmol/g, 0.1 mmol) was reacted with Fmoc-L-Lys(Alloc)-OH (4 eq., 0.4 mmol) and DIPEA (4 eq.) in CH_2Cl_2 for 2 h at room temperature. The resin was washed and treated with a solution of MeOH/CH_2Cl_2/DIPEA (5:80:15) for 15 min two times. After thorough washing, the resin was mixed with $Pd(PPh_3)_4$ (cat) and 100 µL $PhSiH_3$ in CH_2Cl_2 and shaken for 1 h. After thorough washing with CH_2Cl_2 and DMF, the resin was treated with a solution of Fmoc-L-Ser(O^{t}Bu)-OH (**L-108**) (10 eq.), HBTU (10 eq.), HOBt (10 eq.) and DIPEA (10 eq.) in DMF and shaken for 2 h. After thorough washing and Fmoc-deprotection with 20% piperidine in DMF for 30 min, the resin was treated with 95% TFA

(TIS:H_2O, 2.5% each) for 2.5 h. The crude product was precipitated in cold ether, filtered, dried and purified by preparative HPLC (35 mL/min, 0% MeCN for 15 min, t_r = 7 min) to yield the desired product as a white powder (6.5 mg, 27.9 µmol, 28%).

On-resin approach with ivDde-deprotection. Preloaded Fmoc-L-Lys(ivDde)-Wang resin (0.54 mmol/g, 0.2 mmol) was swollen in DMF for 30 min. The resin was treated with 20% piperidine in DMF (2× 15 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was mixed with a solution of Boc_2O (1.5 eq.) and DIPEA (1.5 eq.) in NMP and shaken for 1 h (2×). The resin was washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 4% $N_2H_4 \cdot H_2O$ in DMF (2× 10 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was mixed with a solution of Fmoc-L-Ser(^{t}Bu)-OH (**L-108**) (10 eq.), HATU (10 eq.), HOBt (10 eq.) and DIPEA (10 eq.) in DMF and shaken for 5 h. The resin was washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 20% piperidine in DMF (2× 15 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 95% TFA (TIS:H_2O, 2.5% each) for 2.5 h. The crude product was precipitated in cold ether, filtered, dried and purified by preparative HPLC (35 mL/min, 0% MeCN for 15 min, t_r = 7 min) to yield the desired product as a white powder (36.1 mg, 154.8 µmol, 77%, 173.8 µmol TFA): ^{1}H-NMR (D_2O, 300 MHz) δ (ppm) = 4.15-3.86 (m, 4H), 3.29 (td, *J* = 6.8, 2.6 Hz, 2H), 1.97 (bp, 2H), 1.69-1.37 (m, 4H); ^{13}C-NMR (D_2O, 75 MHz) δ (ppm) = 172.6, 167.4, 162.9 (q, *J* = 35.5 Hz), 116.4 (q, *J* = 291.2 Hz), 60.3, 54.6, 53.2, 39.1, 29.4, 27.8, 21.6; ^{19}F-NMR (D_2O, 565 MHz) δ (ppm) = -75.54; HRMS: m/z (calculated): $[M+H]^+$ 234.1448, m/z (ESI-ToF): $[M+H]^+$ 234.1452.

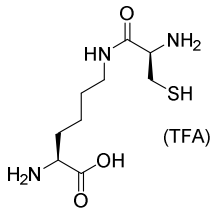
5.8.4. Synthesis of L-Lys(D-Cys) (101)



On-resin approach with ivDde-deprotection. Preloaded Fmoc-L-Lys(ivDde)-Wang resin (0.54 mmol/g, 0.1 mmol) was swollen in DMF for 30 min. The resin was treated with 20% piperidine in DMF (2× 15 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was mixed with a solution of Boc_2O (1.5 eq.) and DIPEA (1.5 eq.) in NMP and shaken for 1 h (2×). The resin was washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 4% $N_2H_4 \cdot H_2O$ in DMF (2× 10 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was mixed with a solution of Fmoc-D-Cys(^{t}Bu)-OH (10 eq.), HATU (10 eq.), HOBt (10 eq.) and DIPEA (10 eq.) in DMF

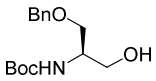
and shaken for 5 h. The resin was washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 20% piperidine in DMF (2× 15 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 95% TFA (TIS:H_2O, 2.5% each) for 2.5 h. The crude product was precipitated in cold ether, filtered, dried and purified by preparative HPLC (35 mL/min, 0% MeCN for 20 min, t_r = 11.5 min) to yield the desired product as a white powder (36.1 mg, 154.8 µmol, 77%, 173.8 µmol TFA): ^{1}H-NMR (D_2O, 300 MHz) δ = 4.02 (t, *J* = 5.9 Hz, 1H), 3.90 (t, *J* = 6.2 Hz, 1H), 3.30-2.85 (m, 5H), 1.94-1.24 (m, 8H); ^{13}C-NMR (D_2O, 75 MHz) δ = 172.40 (s), 167.58 (s), 118.10 (s), 116.16 (d, *J* = 292.1 Hz), 54.36 (s), 52.90 (s), 38.94 (s), 29.28 (s), 27.63 (s), 24.67 (s), 22.03 (s), 21.43 ppm (s); ^{19}F-NMR (D_2O, 282 MHz) -75.59 ppm. The analytical data is in accordance with the literature.[235b]

5.8.5. Synthesis of L-Lys(L-Cys) (102)

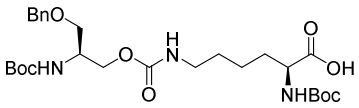


On-resin approach with ivDde-deprotection. Preloaded Fmoc-L-Lys(ivDde)-Wang resin (0.54 mmol/g, 0.1 mmol) was swollen in DMF for 30 min. The resin was treated with 20% piperidine in DMF (2× 15 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was mixed with a solution of Boc_2O (1.5 eq.) and DIPEA (1.5 eq.) in NMP and shaken for 1 h (2×). The resin was washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 4% $N_2H_4 \cdot H_2O$ in DMF (2× 10 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was mixed with a solution of Fmoc-L-Cys(^{t}Bu)-OH (10 eq.), HATU (10 eq.), HOBt (10 eq.) and DIPEA (10 eq.) in DMF and shaken for 5 h. The resin was washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 20% piperidine in DMF (2× 15 min) and washed thoroughly with DMF and CH_2Cl_2. The resin was treated with 95% TFA (TIS:H_2O, 2.5% each) for 2.5 h. The crude product was precipitated in cold ether, filtered, dried and purified by preparative HPLC (35 mL/min, 0% MeCN for 20 min, t_r = 12.5 min) to yield the desired product as a white powder (10.2 mg, 40.9 µmol, 41%, 39.3 µmol TFA): ^{1}H-NMR (D_2O, 400 MHz) δ = 4.17 (t, *J* = 5.8 Hz, 1H), 3.99 (t, *J* = 6.1 Hz, 1H), 3.45-3.02 (m, 5H), 2.05-1.37 (m, 8H); ^{13}C-NMR (D_2O, 300 MHz) δ = 172.95 (s), 167.72 (s), 162.97 (q, *J* = 35.4 Hz), 116.35 (q, *J* = 291.7 Hz), 73.47 (s), 54.55 (s), 53.43 (s), 44.57 (s), 39.13 (s), 29.56 (s), 27.82 (s), 24.84 (s), 22.22 (s), 21.64 (s); ^{19}F-NMR (D_2O, 282 MHz) -75.59 ppm. The analytical data is in accordance with the literature.[235b]

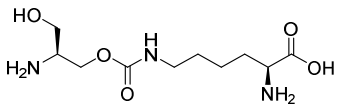
5.8.6. Synthesis of Lysine Carbamate 114



***tert*-Butyl (R)-(1-(benzyloxy)-3-hydroxypropan-2-yl)carbamate (110).** *O*-benzyl-*N*-(*tert*-butoxycarbonyl)-L-serine (**109**) (1.00 g, 3.39 mmol) was dissolved in 6 mL dry THF and the solution was cooled to 0 °C. A solution of BH_3·THF in THF (1 M, 3.40 mmol) was added. The reaction was stirred for 1 h at 0 °C and for 3 h at rt. The reaction mixture was cooled to 0 °C, quenched with methanol and concentrated by vacuum. Purification was column chromatography (EtOAc:hex, 1:1) gave the desired product in 73% yield (693 mg, 2.46 mmol): R_f = 0.5; ^{1}H-NMR ($CDCl_3$, 300 MHz) δ = 7.47-7.20 (m, 5H), 5.16 (bs, 1H, NH), 4.52 (s, 2H), 3.95-3.51 (m, 5H), 1.44 (s, 9H); ^{13}C-NMR ($CDCl_3$, 75 MHz) δ = 156.08, 137.64, 128.54, 127.93, 127.69, 79.72, 73.53, 70.80, 64.11, 51.64, 28.39 ppm. The analytical data was in agreement with the literature.[237]



(6S,15S)-6-((benzyloxy)methyl)-15-((*tert*-butoxycarbonyl)amino)-2,2-dimethyl-4,9-dioxo-3,8-dioxa-5,10-diazahexadecan-16-oic acid (112). Alcohol **110** (658 mg, 2.34 mmol) was dissolved in 6 mL dry THF and cooled to 0 °C. Triphosgene (692 mg, 2.34 mmol) was added, the reaction was stirred overnight at rt. The mixture was concentrated and redissolved in 2.5 mL THF (solution A). Boc-L-Lys-OH (**106**) (693 mg, 2.80 mmol) was dissolved in 10 mL NaOH (1 M) and 2.5 mL THF and slowly cooled to 0 °C. Solution A was slowly added and the reaction mixture was stirred overnight at rt. The mixture was concentrated by vacuum and purification by preparative HPLC (35 mL/min, 15 min at 0% MeCN, t_r = 6.3 min) gave the desired product in 65% yield (841 mg, 1.52 mmol): ^{1}H-NMR ($CDCl_3$, 300 MHz) δ = 7.42-7.24 (m, 5H), 5.39-4.87 (m, 2H), 4.51 (s, 2H), 4.40-4.08 (m, 3H), 3.65 – 3.40 (m, 2H), 3.28-3.98 (m, 2H), 1.87-1.67 (m, 2H), 1.59-1.31 (m, 21H).



***N*^{6}-(((S)-2-amino-3-hydroxypropoxy)carbonyl)-L-lysine (114).** Compound **112** (885 mg, 1.60 mmol) was dissolved in 20 mL dry methanol. Pd/$BaSO_4$ (10 mol%) was added. The reaction was flushed several times with H_2 and stirred under a H_2 atmosphere overnight at rt. The reaction mixture was filtered over Celite, concentrated under vacuum and redissolved in 20 mL CH_2Cl_2. After the addition of 5 mL of TFA, the reaction was stirred for 2 h at rt. The solution was

concentrated under high vacuum and purification by preparative HPLC (35 mL/min, 10 min at 0% MeCN, in 50 min to 100% MeCN, t_r = 23.0 min) gave the desired product in 8% yield (7.75 mg, 29.4 µmol, 29.4 µmol TFA): ^{1}H-NMR (D_2O, 300 MHz) δ = 4.30 (dd, *J* = 12.1, 3.9 Hz, 1H), 4.20 (dd, *J* = 12.1, 6.4 Hz, 1H), 3.92 (t, *J* = 5.9 Hz, 1H), 3.83 (dd, *J* = 12.2, 4.6 Hz, 1H), 3.73 (dd, *J* = 12.2, 6.4 Hz, 1H), 3.66 – 3.53 (m, 1H), 3.12 (t, *J* = 6.6 Hz, 2H), 2.00 – 1.80 (m, 2H), 1.60 – 1.29 (m, 4H); ^{13}C-NMR (D_2O, 75 MHz) δ = 173.1, 163.0 (q, *J* = 35.3 Hz), 157.5, 116.4 (q, *J* = 309.5 Hz), 61.8, 58.4, 53.4, 51.8, 40.0, 29.5, 28.2, 21.4 ppm; ^{19}F-NMR (D_2O, 282 MHz) δ = -75.65 ppm.

6. References

[1] A. Cornishbowden, *Eur. J. Biochem.* **1984**, *138*, 9-37.

[2] J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 6. ed., Freeman, New York u.a., **2007**.

[3] D. Voet, J. G. Voet, C. W. Pratt, *Principles of biochemistry*, 4. ed., Wiley, Hoboken, NJ u.a., **2013**.

[4] a) C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, *Angew. Chem. Int. Ed.* **2005**, *44*, 7342-7372; b) S. Prabakaran, G. Lippens, H. Steen, J. Gunawardena, *Wiley Interdisciplinary Reviews-Systems Biology and Medicine* **2012**, *4*, 565-583.

[5] F. Sherman, J. W. Stewart, J. H. Parker, E. Inhaber, N. A. Shipman, Putterma.Gj, R. L. Gardisky, Margolia.E, *J. Biol. Chem.* **1968**, *243*, 5446-5456.

[6] T. Hohsaka, M. Sisido, *Curr. Opin. Chem. Biol.* **2002**, *6*, 809-815.

[7] D. H. Appella, L. A. Christianson, D. A. Klein, D. R. Powell, X. L. Huang, J. J. Barchi, S. H. Gellman, *Nature* **1997**, *387*, 381-384.

[8] P. E. Nielsen, *Pseudo-peptides in drug discovery*, Wiley-VCH, Weinheim, **2004**.

[9] V. Wittmann, S. J. Danishefsky, *Glycopeptides and glycoproteins: synthesis, structure, and application*, Springer, Berlin u.a., **2007**.

[10] L. Brunsveld, J. Kuhlmann, K. Alexandrov, A. Wittinghofer, R. S. Goody, H. Waldmann, *Angew. Chem. Int. Ed.* **2006**, *45*, 6622-6646.

[11] C. Walsh, *Posttranslational modification of proteins : expanding nature's inventory*, Roberts and Co. Publishers, Englewood, Colo., **2006**.

[12] V. R. Pattabiraman, J. W. Bode, *Nature* **2011**, *480*, 471-479.

[13] L. Davis, J. W. Chin, *Nature Reviews Molecular Cell Biology* **2012**, *13*, 168-182.

[14] E. Fischer, E. Fourneau, *Über einige Derivate des Glykocolls*, Schade, Berlin, **1901**.

[15] a) V. Duvigneaud, C. Ressler, J. M. Swan, C. W. Roberts, P. G. Katsoyannis, S. Gordon, *J. Am. Chem. Soc.* **1953**, *75*, 4879-4880; b) V. Duvigneaud, C. Ressler, S. Trippett, *J. Biol. Chem.* **1953**, *205*, 949-957.

[16] a) M. Bergmann, L. Zervas, *Berichte der deutschen chemischen Gesellschaft (A and B Series)* **1932**, *65*, 1192-1201; b) L. A. Carpino, *J. Am. Chem. Soc.* **1957**, *79*, 4427-4431; c) H. G. Khorana, *Chem Ind-London* **1955**, 1087-1088; d) F. C. Mckay, N. F. Albertson, *J. Am. Chem. Soc.* **1957**, *79*, 4686-4690; e) J. C. Sheehan, G. P. Hess, *J. Am. Chem. Soc.* **1955**, *77*, 1067-1068; f) M. Goodman, W. J. Mcgahren, *Tetrahedron* **1967**, *23*, 2031-2050.

[17] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.

[18] B. Gutte, R. B. Merrifield, *J. Biol. Chem.* **1971**, *246*, 1922-1941.

[19] S. B. H. Kent, *Annu. Rev. Biochem* **1988**, *57*, 957-989.

[20] C. P. R. Hackenberger, D. Schwarzer, *Angew. Chem. Int. Ed.* **2008**, *47*, 10030-10074.

[21] T. Wieland, E. Bokelmann, L. Bauer, H. U. Lang, H. Lau, *Justus Liebigs Annalen der Chemie* **1953**, *583*, 129-149.

[22] P. Dawson, T. Muir, I. Clark-Lewis, S. Kent, *Science* **1994**, *266*, 776-779.

[23] M. Schnolzer, S. B. H. Kent, *Science* **1992**, *256*, 221-225.

[24] C. Haase, O. Seitz, *Angew. Chem. Int. Ed.* **2008**, *47*, 1553-1556.

[25] D. Bang, B. L. Pentelute, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2006**, *45*, 3985-3988.

[26] C. F. Liu, J. P. Tam, *P Natl Acad Sci USA* **1994**, *91*, 6584-6588.

[27] D. Bang, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2004**, *43*, 2534-2538.

[28] Y. Sohma, B. L. Pentelute, J. Whittaker, Q. X. Hua, L. J. Whittaker, M. A. Weiss, S. B. H. Kent, *Angew. Chem. Int. Ed.* **2008**, *47*, 1102-1106.

[29] a) H. Hojo, S. Aimoto, *Bull. Chem. Soc. Jpn.* **1991**, *64*, 111-117; b) H. Hojo, Y. Kwon, Y. Kakuta, S. Tsuda, I. Tanaka, K. Hikichi, S. Aimoto, *Bull. Chem. Soc. Jpn.* **1993**, *66*, 2700-2706.

[30] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *P Natl Acad Sci USA* **1999**, *96*, 10068-10073.

[31] a) R. R. Flavell, M. Huse, M. Goger, M. Trester-Zedlitz, J. Kuriyan, T. W. Muir, *Org. Lett.* **2002**, *4*, 165-168; b) S. Flemer, *J. Pept. Sci.* **2009**, *15*, 693-696; c) Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman, C. R. Bertozzi, *J. Am. Chem. Soc.* **1999**, *121*, 11684-11689.

[32] a) S. Ficht, R. J. Payne, R. T. Guy, C. H. Wong, *Chem-Eur J* **2008**, *14*, 3620-3629; b) A. Ishii, H. Hojo, Y. Nakahara, Y. Ito, Y. Nakaharai, *Biosci., Biotechnol., Biochem.* **2002**, *66*, 225-232.

[33] J. D. Warren, J. S. Miller, S. J. Keding, S. J. Danishefsky, *J. Am. Chem. Soc.* **2004**, *126*, 6576-6578.

[34] P. Botti, M. Villain, S. Manganiello, H. Gaertner, *Org. Lett.* **2004**, *6*, 4861-4864.

[35] a) W. Hou, X. H. Zhang, F. P. Li, C. F. Liu, *Org. Lett.* **2011**, *13*, 386-389; b) N. Ollivier, J. Dheur, R. Mhidia, A. Blanpain, O. Melnyk, *Org. Lett.* **2010**, *12*, 5238-5241.

[36] a) R. Ingenito, E. Bianchi, D. Fattori, A. Pessi, *J. Am. Chem. Soc.* **1999**, *121*, 11369-11374; b) G. W. Kenner, Mcdermot.Jr, R. C. Sheppard, *Journal of the Chemical Society D-Chemical Communications* **1971**, 636-637.

[37] F. Mende, O. Seitz, *Angew. Chem. Int. Ed.* **2007**, *46*, 4577-4580.

[38] G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. K. Cui, L. Liu, *Angew. Chem. Int. Ed.* **2011**, *50*, 7645-7649.

[39] a) G. M. Fang, J. X. Wang, L. Liu, *Angew. Chem. Int. Ed.* **2012**, *51*, 10347-10350; b) J. Liang, G. M. Fang, X. L. Huang, Z. Q. Mei, J. Li, C. L. Tian, L. Liu, *Sci China Chem* **2013**, *56*, 1301-1306.

[40] R. M. Hofmann, T. W. Muir, *Curr. Opin. Biotechnol.* **2002**, *13*, 297-303.

[41] G. K. Busch, E. W. Tate, P. R. J. Gaffney, E. Rosivatz, R. Woscholski, R. J. Leatherbarrow, *Chem. Commun.* **2008**, 3369-3371.

[42] T. W. Muir, *Annu. Rev. Biochem* **2003**, *72*, 249-289.

[43] a) T. C. Evans, J. Benner, M.-Q. Xu, *Protein Sci.* **1998**, *7*, 2256-2264; b) T. W. Muir, D. Sondhi, P. A. Cole, *P Natl Acad Sci USA* **1998**, *95*, 6705-6710; c) K. Severinov, T. W. Muir, *J. Biol. Chem.* **1998**, *273*, 16205-16209.

[44] a) C. J. Noren, J. M. Wang, F. B. Perler, *Angew. Chem. Int. Ed.* **2000**, *39*, 450-466; b) H. Paulus, *Annu. Rev. Biochem* **2000**, *69*, 447-496.

[45] M. Vila-Perelló, Z. Liu, N. H. Shah, J. A. Willis, J. Idoyaga, T. W. Muir, *J. Am. Chem. Soc.* **2012**, *135*, 286-292.

[46] Y. M. Li, M. Y. Yang, Y. C. Huang, Y. T. Li, P. R. Chen, L. Liu, *Acs Chem Biol* **2012**, *7*, 1015-1022.

[47] Q. Q. He, G. M. Fang, L. Liu, *Chin. Chem. Lett.* **2013**, *24*, 265-269.

[48] R. J. Hondal, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2001**, *123*, 5140-5141.

[49] R. Quaderer, A. Sewing, D. Hilvert, *Helv. Chim. Acta* **2001**, *84*, 1197-1206.

[50] M. D. Gieselman, L. L. Xie, W. A. van der Donk, *Org. Lett.* **2001**, *3*, 1331-1334.

[51] a) D. Besse, F. Siedler, T. Diercks, H. Kessler, L. Moroder, *Angewandte Chemie-International Edition in English* **1997**, *36*, 883-885; b) R. E. Huber, R. S. Criddle, *Arch. Biochem. Biophys.* **1967**, *122*, 164-&.

[52] N. Metanis, E. Keinan, P. E. Dawson, *Angew. Chem. Int. Ed.* **2010**, *49*, 7049-7053.

[53] L. R. Malins, K. M. Cergol, R. J. Payne, *Chembiochem* **2013**, *14*, 559-563.

[54] E. M. Sletten, C. R. Bertozzi, *Angew. Chem. Int. Ed.* **2009**, *48*, 6974-6998.

[55] N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo, C. R. Bertozzi, *Acs Chem Biol* **2006**, *1*, 644-648.

[56] J. V. Staros, H. Bayley, D. N. Standring, J. R. Knowles, *Biochem. Biophys. Res. Commun.* **1978**, *80*, 568-572.

[57] R. Breinbauer, M. Kohn, *Chembiochem* **2003**, *4*, 1147-1149.

[58] a) A. B. Hughes, *Amino acids, peptides and proteins in organic chemistry*, Wiley-VCH, Weinheim, **2009**; b) A. L. Isaad, F. Barbetti, P. Rovero, A. M. D'Ursi, M. Chelli, M. Chorev,

A. M. Papini, *Eur. J. Org. Chem.* **2008**, 5308-5314; c) Y. H. Lau, D. R. Spring, *Synlett* **2011**, 1917-1919; d) J. T. Lundquist, J. C. Pelletier, *Org. Lett.* **2001**, *3*, 781-783; e) H. Johansson, D. S. Pedersen, *Eur. J. Org. Chem.* **2012**, 4267-4281.

[59] H. Staudinger, J. Meyer, *Helv. Chim. Acta* **1919**, *2*, 635-646.

[60] E. Saxon, C. R. Bertozzi, *Science* **2000**, *287*, 2007-2010.

[61] Y. He, R. J. Hinklin, J. Y. Chang, L. L. Kiessling, *Org. Lett.* **2004**, *6*, 4479-4482.

[62] P. F. van Swieten, M. A. Leeuwenburgh, B. M. Kessler, H. S. Overkleeft, *Org Biomol Chem* **2005**, *3*, 20-27.

[63] M. Kohn, R. Wacker, C. Peters, H. Schroder, L. Soulere, R. Breinbauer, C. M. Niemeyer, H. Waldmann, *Angew. Chem. Int. Ed.* **2003**, *42*, 5830-5834.

[64] B. L. Nilsson, R. J. Hondal, M. B. Soellner, R. T. Raines, *J. Am. Chem. Soc.* **2003**, *125*, 5268-5269.

[65] a) B. L. Nilsson, L. L. Kiessling, R. T. Raines, *Org. Lett.* **2000**, *2*, 1939-1941; b) E. Saxon, J. I. Armstrong, C. R. Bertozzi, *Org. Lett.* **2000**, *2*, 2141-2143.

[66] a) A. Tam, M. B. Soellner, R. T. Raines, *J. Am. Chem. Soc.* **2007**, *129*, 11421-11430; b) A. Tam, M. B. Soellner, R. T. Raines, *Org Biomol Chem* **2008**, *6*, 1173-1175; c) A. Tam, R. T. Raines, *Biorg. Med. Chem.* **2009**, *17*, 1055-1063; d) S. H. Weisbrod, A. Marx, *Synlett* **2010**, 787-789.

[67] M. B. Soellner, A. Tam, R. T. Raines, *J. Org. Chem.* **2006**, *71*, 9824-9830.

[68] R. Kleineweischede, C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2008**, *47*, 5984-5988.

[69] J. Kalia, N. L. Abbott, R. T. Raines, *Bioconjugate Chem.* **2007**, *18*, 1064-1069.

[70] a) G. H. Hakimelahi, G. Just, *Tetrahedron Lett.* **1980**, *21*, 2119-2122; b) T. Rosen, I. M. Lico, D. T. W. Chu, *J. Org. Chem.* **1988**, *53*, 1580-1582; c) N. Shangguan, S. Katukojvala, R. Greenburg, L. J. Williams, *J. Am. Chem. Soc.* **2003**, *125*, 7754-7755.

[71] J. Blake, C. H. Li, *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1981**, *78*, 4055-4058.

[72] D. Crich, K. Sana, *J. Org. Chem.* **2009**, *74*, 7383-7388.

[73] X. H. Zhang, X. W. Lu, C. F. Liu, *Tetrahedron Lett.* **2008**, *49*, 6122-6125.

[74] X. Q. Li, T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **1998**, *39*, 8669-8672.

[75] A. S. Goldstein, M. H. Gelb, *Tetrahedron Lett.* **2000**, *41*, 2797-2800.

[76] a) S. Vetter, *Synth. Commun.* **1998**, *28*, 3219-3223; b) V. W. Wittmann, O. Keiper, J. Mannuthodikayil, *J. Pept. Sci.* **2010**, *16*, 91; c) D. Crich, K. Sana, S. P. Guo, *Org. Lett.* **2007**, *9*, 4423-4426.

[77] X. H. Tan, R. L. Yang, A. Wirjo, C. F. Liu, *Tetrahedron Lett.* **2008**, *49*, 2891-2894.

[78] A. Shigenaga, Y. Sumikawa, S. Tsuda, K. Sato, A. Otaka, *Tetrahedron* **2010**, *66*, 3290-3296.

[79] R. Raz, J. Rademann, *Org. Lett.* **2012**, *14*, 5038-5041.

[80] C. Kinsland, S. V. Taylor, N. L. Kelleher, F. W. McLafferty, T. P. Begley, *Protein Sci.* **1998**, *7*, 1839-1842.

[81] a) R. V. Kolakowski, N. Shangguan, R. R. Sauers, L. J. Williams, *J. Am. Chem. Soc.* **2006**, *128*, 5695-5702; b) Y. L. Gao, *Central European Journal of Chemistry* **2010**, *8*, 308-319.

[82] R. Merkx, M. J. van Haren, D. T. S. Rijkers, R. M. J. Liskamp, *J. Org. Chem.* **2007**, *72*, 4574-4577.

[83] X. H. Zhang, F. P. Li, X. W. Lu, C. F. Liu, *Bioconjugate Chem.* **2009**, *20*, 197-200.

[84] a) X. Wu, L. Hu, *The Journal of Organic Chemistry* **2007**, *72*, 765-774; b) X. H. Wu, L. Q. Hu, *Tetrahedron Lett.* **2005**, *46*, 8401-8405.

[85] F. Fazio, C. H. Wong, *Tetrahedron Lett.* **2003**, *44*, 9083-9085.

[86] J. W. Bode, *Abstr Pap Am Chem S* **2008**, *236*.

[87] J. W. Bode, R. M. Fox, K. D. Baucom, *Angew. Chem. Int. Ed.* **2006**, *45*, 1248-1252.

[88] I. Pusterla, J. W. Bode, *Angew. Chem. Int. Ed.* **2012**, *51*, 513-516.

[89] a) L. Ju, A. R. Lippert, J. W. Bode, *J. Am. Chem. Soc.* **2008**, *130*, 4253-4255; b) L. Ju, J. W. Bode, *Org Biomol Chem* **2009**, *7*, 2259-2264; c) M. A. Flores, J. W. Bode, *Org. Lett.* **2010**, *12*, 1924-1927.
[90] T. Fukuzumi, J. W. Bode, *J. Am. Chem. Soc.* **2009**, *131*, 3864-3865.
[91] a) J. Wu, J. Ruiz-Rodriguez, J. M. Comstock, J. Z. Dong, J. W. Bode, *Chem Sci* **2011**, *2*, 1976-1979; b) T. Fukuzumi, L. Ju, J. W. Bode, *Org Biomol Chem* **2012**, *10*, 5837-5844.
[92] A. O. Ogunkoya, V. R. Pattabiraman, J. W. Bode, *Angew. Chem. Int. Ed.* **2012**, *51*, 9693-9697.
[93] V. R. Pattabiraman, A. O. Ogunkoya, J. W. Bode, *Angew. Chem. Int. Ed.* **2012**, *51*, 5114-5118.
[94] J. M. Chalker, G. J. L. Bernardes, B. G. Davis, *Acc. Chem. Res.* **2011**, *44*, 730-741.
[95] a) B. G. Davis, R. C. Lloyd, J. B. Jones, *J. Org. Chem.* **1998**, *63*, 9614-9615; b) B. G. Davis, M. A. T. Maughan, M. P. Green, A. Ullman, J. B. Jones, *Tetrahedron-Asymmetr* **2000**, *11*, 245-262.
[96] E. Calce, M. Leone, L. Monfregola, S. De Luca, *Org. Lett.* **2013**, *15*, 5354-5357.
[97] a) G. T. Hermanson, *Bioconjugate techniques*, 2nd ed., Academic Press, Amsterdam ; Boston, **2008**; b) H. Y. Shiu, T. C. Chan, C. M. Ho, Y. Liu, M. K. Wong, C. M. Che, *Chem-Eur J* **2009**, *15*, 3839-3850.
[98] G. J. L. Bernardes, J. M. Chalker, J. C. Errey, B. G. Davis, *J. Am. Chem. Soc.* **2008**, *130*, 5052-5053.
[99] a) A. Dondoni, *Angew. Chem. Int. Ed.* **2008**, *47*, 8995-8997; b) G. Triola, L. Brunsveld, H. Waldmann, *J. Org. Chem.* **2008**, *73*, 3646-3649.
[100] R. Hoogenboom, *Angew. Chem. Int. Ed.* **2010**, *49*, 3415-3417.
[101] M. Lo Conte, S. Pacifico, A. Chambery, A. Marra, A. Dondoni, *J. Org. Chem.* **2010**, *75*, 4644-4647.
[102] M. Lo Conte, S. Staderini, A. Marra, M. Sanchez-Navarro, B. G. Davis, A. Dondoni, *Chem. Commun.* **2011**, *47*, 11086-11088.
[103] a) X. Chen, K. Muthoosamy, A. Pfisterer, B. Neumann, T. Weil, *Bioconjugate Chem.* **2012**, *23*, 500-508; b) X. Chen, L. Henschke, Q. Z. Wu, K. Muthoosamy, B. Neumann, T. Weil, *Org Biomol Chem* **2013**, *11*, 353-361.
[104] N. S. Joshi, L. R. Whitaker, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 15942-15943.
[105] T. Kodadek, I. Duroux-Richard, J. C. Bonnafous, *Trends in Pharmacological Sciences* **2005**, *26*, 210-217.
[106] S. Meunier, E. Strable, M. G. Finn, *Chem Biol* **2004**, *11*, 319-326.
[107] a) J. M. McFarland, N. S. Joshi, M. B. Francis, *J. Am. Chem. Soc.* **2008**, *130*, 7639-7644; b) D. W. Romanini, M. B. Francis, *Bioconjugate Chem.* **2008**, *19*, 153-157.
[108] S. D. Tilley, M. B. Francis, *J. Am. Chem. Soc.* **2006**, *128*, 1080-1081.
[109] H. Ban, J. Gavrilyuk, C. F. Barbas, *J. Am. Chem. Soc.* **2010**, *132*, 1523-+.
[110] J. Gavrilyuk, H. Ban, M. Nagano, W. Hakamata, C. F. Barbas, *Bioconjugate Chem.* **2012**, *23*, 2321-2328.
[111] J. M. Antos, J. M. McFarland, A. T. Iavarone, M. B. Francis, *J. Am. Chem. Soc.* **2009**, *131*, 6301-6308.
[112] C. A. Hutchison, S. Phillips, M. H. Edgell, S. Gillam, P. Jahnke, M. Smith, *J. Biol. Chem.* **1978**, *253*, 6551-6560.
[113] a) H. Gaertner, R. Offord, P. Botti, G. Kuenzi, O. Hartley, *Bioconjugate Chem.* **2008**, *19*, 480-489; b) B. C. Lee, R. N. Zuckermann, *Acs Chem Biol* **2011**, *6*, 1367-1374.
[114] J. P. Tam, Q. T. Yu, Z. W. Miao, *Biopolymers* **1999**, *51*, 311-332.
[115] X. F. Li, L. S. Zhang, S. E. Hall, J. P. Tam, *Tetrahedron Lett.* **2000**, *41*, 4069-4073.
[116] J. M. Gilmore, R. A. Scheck, A. P. Esser-Kahn, N. S. Joshi, M. B. Francis, *Angew. Chem. Int. Ed.* **2006**, *45*, 5307-5311.

[117] R. A. Scheck, M. T. Dedeo, A. T. Lavarone, M. B. Francis, *J. Am. Chem. Soc.* **2008**, *130*, 11762-11770.
[118] a) H. B. F. Dixon, L. R. Weitkamp, *Biochem. J* **1962**, *84*, 462-&; b) K. F. Geoghegan, J. G. Stroh, *Bioconjugate Chem.* **1992**, *3*, 138-146.
[119] M. Levine, H. Tarver, *J. Biol. Chem.* **1951**, *192*, 835-850.
[120] D. B. Cowie, G. N. Cohen, *Biochim. Biophys. Acta* **1957**, *26*, 252-261.
[121] a) G. Hortin, I. Boime, *Methods Enzymol.* **1983**, *96*, 777-784; b) B. D. Sykes, Weingart.Hi, Schlesin.Mj, *P Natl Acad Sci USA* **1974**, *71*, 469-473.
[122] S. N. Cohen, A. C. Y. Chang, H. W. Boyer, R. B. Helling, *P Natl Acad Sci USA* **1973**, *70*, 3240-3244.
[123] M. Mandel, A. Higa, *J. Mol. Biol.* **1970**, *53*, 159-&.
[124] J. F. Morrow, S. N. Cohen, A. C. Y. Chang, H. W. Boyer, H. M. Goodman, R. B. Helling, *P Natl Acad Sci USA* **1974**, *71*, 1743-1747.
[125] R. H. Scheller, R. E. Dickerson, H. W. Boyer, A. D. Riggs, K. Itakura, *Science* **1977**, *196*, 179-180.
[126] a) N. Budisa, C. Minks, S. Alefelder, W. Wenger, F. M. Dong, L. Moroder, R. Huber, *FASEB J.* **1999**, *13*, 41-51; b) N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann, R. Huber, *Eur. J. Biochem.* **1995**, *230*, 788-796.
[127] N. Budisa, *Angew. Chem. Int. Ed.* **2004**, *43*, 6426-6463.
[128] C. J. Noren, S. J. Anthonycahill, M. C. Griffith, P. G. Schultz, *Science* **1989**, *244*, 182-188.
[129] D. R. Liu, P. G. Schultz, *P Natl Acad Sci USA* **1999**, *96*, 4780-4785.
[130] L. Wang, T. J. Magliery, D. R. Liu, P. G. Schultz, *J. Am. Chem. Soc.* **2000**, *122*, 5010-5011.
[131] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498-500.
[132] L. Wang, J. Xie, P. G. Schultz, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225-249.
[133] a) B. Moore, B. C. Persson, C. C. Nelson, R. F. Gesteland, J. F. Atkins, *J. Mol. Biol.* **2000**, *298*, 195-209; b) J. C. Anderson, N. Wu, S. W. Santoro, V. Lakshman, D. S. King, P. G. Schultz, *P Natl Acad Sci USA* **2004**, *101*, 7566-7571; c) T. S. Young, P. G. Schultz, *J. Biol. Chem.* **2010**, *285*, 11039-11044.
[134] a) A. Ambrogelly, S. Gundllapalli, S. Herring, C. Polycarpo, C. Frauer, D. Soll, *P Natl Acad Sci USA* **2007**, *104*, 3141-3146; b) H. Neumann, S. Y. Peak-Chew, J. W. Chin, *Nature Chemical Biology* **2008**, *4*, 232-234.
[135] a) T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto, S. Yokoyama, *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818-822; b) T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, S. Yokoyama, *Chem Biol* **2008**, *15*, 1187-1197; c) P. R. Chen, D. Groff, J. T. Guo, W. J. Ou, S. Cellitti, B. H. Geierstanger, P. G. Schultz, *Angew. Chem. Int. Ed.* **2009**, *48*, 4052-4055; d) D. P. Nguyen, M. M. G. Alai, P. B. Kapadnis, H. Neumann, J. W. Chin, *J. Am. Chem. Soc.* **2009**, *131*, 14194-+; e) D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters, J. W. Chin, *J. Am. Chem. Soc.* **2009**, *131*, 8720-+; f) T. Fekner, M. K. Chan, *Curr. Opin. Chem. Biol.* **2011**, *15*, 387-391; g) Y. S. Wang, X. Q. Fang, A. L. Wallace, B. Wu, W. S. R. Liu, *J. Am. Chem. Soc.* **2012**, *134*, 2950-2953.
[136] R. K. V. Lim, Q. Lin, *Chem. Commun.* **2010**, *46*, 1589-1600.
[137] R. Huisgen, *P Chem Soc London* **1961**, 357.
[138] a) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596-+; b) C. W. Tornoe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057-3064.
[139] a) F. Himo, T. Lovell, R. Hilgraf, V. V. Rostovtsev, L. Noodleman, K. B. Sharpless, V. V. Fokin, *J. Am. Chem. Soc.* **2005**, *127*, 210-216; b) V. D. Bock, H. Hiemstra, J. H. van Maarseveen, *Eur. J. Org. Chem.* **2006**, 51-68.
[140] M. Meldal, C. W. Tornoe, *Chem. Rev.* **2008**, *108*, 2952-3015.

[141] a) V. O. Rodionov, S. I. Presolski, D. D. Diaz, V. V. Fokin, M. G. Finn, *J. Am. Chem. Soc.* **2007**, *129*, 12705-12712; b) V. Hong, N. F. Steinmetz, M. Manchester, M. G. Finn, *Bioconjugate Chem.* **2010**, *21*, 1912-1916.

[142] a) L. Zhang, X. G. Chen, P. Xue, H. H. Y. Sun, I. D. Williams, K. B. Sharpless, V. V. Fokin, G. C. Jia, *J. Am. Chem. Soc.* **2005**, *127*, 15998-15999; b) B. C. Boren, S. Narayan, L. K. Rasmussen, L. Zhang, H. T. Zhao, Z. Y. Lin, G. C. Jia, V. V. Fokin, *J. Am. Chem. Soc.* **2008**, *130*, 14900-14900.

[143] L. M. Artner, L. Merkel, N. Bohlke, F. Beceren-Braun, C. Weise, J. Dernedde, N. Budisa, C. P. R. Hackenberger, *Chem. Commun.* **2012**, *48*, 522-524.

[144] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046-15047.

[145] R. van Geel, G. J. M. Pruijn, F. L. van Delft, W. C. Boelens, *Bioconjugate Chem.* **2012**, *23*, 392-398.

[146] K. E. Beatty, J. D. Fisk, B. P. Smart, Y. Y. Lu, J. Szychowski, M. J. Hangauer, J. M. Baskin, C. R. Bertozzi, D. A. Tirrell, *Chembiochem* **2010**, *11*, 2092-2095.

[147] K. Lang, L. Davis, S. Wallace, M. Mahesh, D. J. Cox, M. L. Blackman, J. M. Fox, J. W. Chin, *J. Am. Chem. Soc.* **2012**, *134*, 10317-10320.

[148] a) M. L. Blackman, M. Royzen, J. M. Fox, *J. Am. Chem. Soc.* **2008**, *130*, 13518-+; b) N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297-2299; c) N. K. Devaraj, R. Weissleder, *Acc. Chem. Res.* **2011**, *44*, 816-827; d) M. T. Taylor, M. L. Blackman, O. Dmitrenko, J. M. Fox, *J. Am. Chem. Soc.* **2011**, *133*, 9646-9649.

[149] K. Lang, L. Davis, J. Torres-Kolbus, C. J. Chou, A. Deiters, J. W. Chin, *Nat Chem* **2012**, *4*, 298-304.

[150] R. Serwa, I. Wilkening, G. Del Signore, M. Muhlberg, I. Claussnitzer, C. Weise, M. Gerrits, C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2009**, *48*, 8234-8239.

[151] a) V. Bohrsch, R. Serwa, P. Majkut, E. Krause, C. P. R. Hackenberger, *Chem. Commun.* **2010**, *46*, 3176-3178; b) D. M. M. Jaradat, H. Hamouda, C. P. R. Hackenberger, *Eur. J. Org. Chem.* **2010**, 5004-5009; c) R. Serwa, P. Majkut, B. Horstmann, J. M. Swiecicki, M. Gerrits, E. Krause, C. P. R. Hackenberger, *Chem Sci* **2010**, *1*, 596-602; d) R. A. Serwa, J. M. Swiecicki, D. Homann, C. P. R. Hackenberger, *J. Pept. Sci.* **2010**, *16*, 563-567; e) V. Bohrsch, T. Mathew, M. Zieringer, M. R. J. Vallee, L. M. Artner, J. Dernedde, R. Haag, C. P. R. Hackenberger, *Org Biomol Chem* **2012**, *10*, 6211-6216; f) N. Nischan, A. Chakrabarti, R. A. Serwa, P. H. M. Bovee-Geurts, R. Brock, C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2013**, *52*, 11920-11924.

[152] M. R. J. Vallee, P. Majkut, I. Wilkening, C. Weise, G. Muller, C. P. R. Hackenberger, *Org. Lett.* **2011**, *13*, 5440-5443.

[153] M. R. J. Vallee, L. M. Artner, J. Dernedde, C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2013**, *52*, 9504-9508.

[154] W. P. Jencks, *J. Am. Chem. Soc.* **1959**, *81*, 475-481.

[155] a) V. W. Cornish, K. M. Hahn, P. G. Schultz, *J. Am. Chem. Soc.* **1996**, *118*, 8150-8151; b) D. J. Oshannessy, M. J. Dobersen, R. H. Quarles, *Immunology Letters* **1984**, *8*, 273-277; c) Z. W. Zhang, B. A. C. Smith, L. Wang, A. Brock, C. Cho, P. G. Schultz, *Biochemistry-Us* **2003**, *42*, 6735-6746.

[156] a) A. Dirksen, T. M. Hackeng, P. E. Dawson, *Angew. Chem. Int. Ed.* **2006**, *45*, 7581-7584; b) Y. Zeng, T. N. C. Ramya, A. Dirksen, P. E. Dawson, J. C. Paulson, *Nature Methods* **2009**, *6*, 207-209.

[157] a) A. R. Blanden, K. Mukherjee, O. Dilek, M. Loew, S. L. Bane, *Bioconjugate Chem.* **2011**, *22*, 1954-1961; b) M. Rashidian, M. M. Mahmoodi, R. Shah, J. K. Dozier, C. R. Wagner, M. D. Distefano, *Bioconjugate Chem.* **2013**, *24*, 333-342.

[158] J. W. Chin, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *P Natl Acad Sci USA* **2002**, *99*, 11020-11024.

[159] A. Dirksen, P. E. Dawson, *Bioconjugate Chem.* **2008**, *19*, 2543-2548.

[160] a) I. S. Carrico, B. L. Carlson, C. R. Bertozzi, *Nature Chemical Biology* **2007**, *3*, 321-322; b) T. Dierks, A. Dickmanns, A. Preusser-Kunze, B. Schmidt, M. Mariappan, K. von Figura, R. Ficner, M. G. Rudolph, *Cell* **2005**, *121*, 541-552.
[161] L. S. Zhang, J. P. Tam, *Anal. Biochem.* **1996**, *233*, 87-93.
[162] X. H. Ning, R. P. Temming, J. Dommerholt, J. Guo, D. B. Ania, M. F. Debets, M. A. Wolfert, G. J. Boons, F. L. van Delft, *Angew. Chem. Int. Ed.* **2010**, *49*, 3065-3068.
[163] R. P. Temming, L. Eggermont, M. B. van Eldijk, J. C. M. van Hest, F. L. van Delft, *Org Biomol Chem* **2013**, *11*, 2772-2779.
[164] a) G. Clave, H. Volland, M. Flaender, D. Gasparutto, A. Romieu, P. Y. Renard, *Org Biomol Chem* **2010**, *8*, 4329-4345; b) G. Viault, S. Dautrey, N. Maindron, J. Hardouin, P. Y. Renard, A. Romieu, *Org Biomol Chem* **2013**, *11*, 2693-2705.
[165] D. M. Beal, V. E. Albrow, G. Burslem, L. Hitchen, C. Fernandes, C. Lapthorn, L. R. Roberts, M. D. Selby, L. H. Jones, *Org Biomol Chem* **2012**, *10*, 548-554.
[166] M. G. Hoesl, N. Budisa, *Angew. Chem. Int. Ed.* **2011**, *50*, 2896-2902.
[167] T. Mukai, A. Hayashi, F. Iraha, A. Sato, K. Ohtake, S. Yokoyama, K. Sakamoto, *Nucleic Acids Res.* **2010**, *38*, 8188-8195.
[168] H. Neumann, K. H. Wang, L. Davis, M. Garcia-Alai, J. W. Chin, *Nature* **2010**, *464*, 441-444.
[169] S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105-1109.
[170] M. Simon, U. Zangemeister-Wittke, A. Pluckthun, *Bioconjugate Chem.* **2012**, *23*, 279-286.
[171] B. G. Davis, *Chem. Rev.* **2002**, *102*, 579-601.
[172] H. Ghazarian, B. Idoni, S. B. Oppenheimer, *Acta Histochemica* **2011**, *113*, 236-247.
[173] C. Fasting, C. A. Schalley, M. Weber, O. Seitz, S. Hecht, B. Koksch, J. Dernedde, C. Graf, E. W. Knapp, R. Haag, *Angew. Chem. Int. Ed.* **2012**, *51*, 10472-10498.
[174] a) M. D. Resh, *Biochimica Et Biophysica Acta-Molecular Cell Research* **1999**, *1451*, 1-16; b) O. Kerscher, R. Felberbaum, M. Hochstrasser, *Annu. Rev. Cell Dev. Biol.* **2006**, *22*, 159-180; c) K. Sadoul, J. Wang, B. Diagouraga, S. Khochbin, *Journal of Biomedicine and Biotechnology* **2011**, 1.
[175] J. Patel, R. R. Pathak, S. Mujtaba, *Nutrition & Metabolism* **2011**, *8*, 12.
[176] P. J. J. Robinson, W. An, A. Routh, F. Martino, L. Chapman, R. G. Roeder, D. Rhodes, *J. Mol. Biol.* **2008**, *381*, 816-825.
[177] H. Neumann, S. M. Hancock, R. Buning, A. Routh, L. Chapman, J. Somers, T. Owen-Hughes, J. van Noort, D. Rhodes, J. W. Chin, *Mol Cell* **2009**, *36*, 153-163.
[178] a) M. Shogren-Knaak, H. Ishii, J. M. Sun, M. J. Pazin, J. R. Davie, C. L. Peterson, *Science* **2006**, *311*, 844-847; b) M. Manohar, A. M. Mooney, J. A. North, R. J. Nakkula, J. W. Picking, A. Edon, R. Fishel, M. G. Poirier, J. J. Ottesen, *J. Biol. Chem.* **2009**, *284*, 23312-23321; c) F. P. Li, A. Allahverdi, R. L. Yang, G. B. J. Lua, X. H. Zhang, Y. Cao, N. Korolev, L. Nordenskiold, C. F. Liu, *Angew. Chem. Int. Ed.* **2011**, *50*, 9611-9614; d) M. Simon, J. A. North, J. C. Shimko, R. A. Forties, M. B. Ferdinand, M. Manohar, M. Zhang, R. Fishel, J. J. Ottesen, M. G. Poirier, *P Natl Acad Sci USA* **2011**, *108*, 12711-12716.
[179] C. I. Schilling, N. Jung, M. Biskup, U. Schepers, S. Brase, *Chem. Soc. Rev.* **2011**, *40*, 4840-4871.
[180] R. J. Detz, S. A. Heras, R. de Gelder, P. W. N. M. van Leeuwen, H. Hiemstra, J. N. H. Reek, J. H. van Maarseveen, *Org. Lett.* **2006**, *8*, 3227-3230.
[181] M. Muehlberg, One-pot three-fragment coupling of unprotected peptides via chemoselective ligation strategies, Master Thesis, Freie Universität Berlin **2009**.
[182] a) O. Herd, A. Hessler, M. Hingst, P. Machnitzki, M. Tepper, O. Stelzer, *Catal. Today* **1998**, *42*, 413-420; b) J. Tsuji, Wiley, Chichester, **2004**.
[183] D. L. Hartl, L. A. Snyder, D. Freifelder, *Basic genetics*, Jones and Bartlett Publishers, Boston, **1988**.

[184] a) J. O. Solbiati, M. Ciaccio, R. N. Farias, J. E. Gonzalez-Pastor, F. Moreno, R. A. Salomon, *J. Bacteriol.* **1999**, *181*, 2659-2662; b) M. J. Bayro, J. Mukhopadhyay, G. V. T. Swapna, J. Y. Huang, L. C. Ma, E. Sineva, P. E. Dawson, G. T. Montelione, R. H. Ebright, *J. Am. Chem. Soc.* **2003**, *125*, 12382-12383; c) K. A. Wilson, M. Kalkum, J. Ottesen, J. Yuzenkova, B. T. Chait, R. Landick, T. Muir, K. Severinov, S. A. Darst, *J. Am. Chem. Soc.* **2003**, *125*, 12475-12483.

[185] L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526-533.

[186] S. Friedrichbochnitschek, H. Waldmann, H. Kunz, *J. Org. Chem.* **1989**, *54*, 751-756.

[187] T. Conroy, K. A. Jolliffe, R. J. Payne, *Org Biomol Chem* **2009**, *7*, 2255-2258.

[188] C. W. Yue, J. Thierry, P. Potier, *Tetrahedron Lett.* **1993**, *34*, 323-326.

[189] S. A. Kates, B. G. Delatorre, R. Eritja, F. Albericio, *Tetrahedron Lett.* **1994**, *35*, 1033-1034.

[190] W. C. Chan, B. W. Bycroft, D. J. Evans, P. D. White, *J Chem Soc Chem Comm* **1995**, 2209-2210.

[191] K. F. Medzihradszky, N. P. Ambulos, A. Khatri, G. Osapay, H. A. Remmer, A. Somogyi, S. A. Kates, *Lett. Pept. Sci.* **2001**, *8*, 1-12.

[192] M. Matoba, T. Kajimoto, M. Node, *Synlett* **2007**, 1930-1934.

[193] a) J. R. Barr, V. L. Maggio, D. G. Patterson, G. R. Cooper, L. O. Henderson, W. E. Turner, S. J. Smith, W. H. Hannon, L. L. Needham, E. J. Sampson, *Clin. Chem.* **1996**, *42*, 1676-1682; b) S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol Cell Proteomics* **2002**, *1*, 376-386; c) H. Zhang, X. J. Li, D. B. Martin, R. Aebersold, *Nat. Biotechnol.* **2003**, *21*, 660-666; d) J. M. Asara, X. Zhang, B. Zheng, L. A. Maroney, H. R. Christofk, N. Wu, L. C. Cantley, *Nat Protoc* **2006**, *1*, 46-51; e) V. Mayya, D. K. Han, *Expert Rev Proteomic* **2006**, *3*, 597-610; f) K. G. Kline, M. R. Sussman, *Annu Rev Biophys* **2010**, *39*, 291-308; g) F. Xie, T. Liu, W. J. Qian, V. A. Petyuk, R. D. Smith, *J. Biol. Chem.* **2011**, *286*, 25443-25449; h) M. L. Steinhauser, A. P. Bailey, S. E. Senyo, C. Guillermier, T. S. Perlstein, A. P. Gould, R. T. Lee, C. P. Lechene, *Nature* **2012**, *481*, 516-U131; i) R. Martello, A. Mangerich, S. Sass, P. C. Dedon, A. Burkle, *Acs Chem Biol* **2013**, *8*, 1567-1575.

[194] a) W. Wan, Y. Huang, Z. Y. Wang, W. K. Russell, P. J. Pai, D. H. Russell, W. R. Liu, *Angew. Chem. Int. Ed.* **2010**, *49*, 3211-3214; b) B. Wu, Z. Y. Wang, Y. Huang, W. S. R. Liu, *Chembiochem* **2012**, *13*, 1405-1408; c) I. Nikic, T. Plass, O. Schraidt, J. Szymanski, J. A. G. Briggs, C. Schultz, E. A. Lemke, *Angew. Chem. Int. Ed.* **2014**, *53*, 2245-2249.

[195] a) D. B. F. Johnson, J. F. Xu, Z. X. Shen, J. K. Takimoto, M. D. Schultz, R. J. Schmitz, Z. Xiang, J. R. Ecker, S. P. Briggs, L. Wang, *Nature Chemical Biology* **2011**, *7*, 779-786; b) M. J. Lajoie, A. J. Rovner, D. B. Goodman, H. R. Aerni, A. D. Haimovich, G. Kuznetsov, J. A. Mercer, H. H. Wang, P. A. Carr, J. A. Mosberg, N. Rohland, P. G. Schultz, J. M. Jacobson, J. Rinehart, G. M. Church, F. J. Isaacs, *Science* **2013**, *342*, 357-360.

[196] J. M. Chalker, G. J. L. Bernardes, Y. A. Lin, B. G. Davis, *Chem-Asian J* **2009**, *4*, 630-640.

[197] D. Agrawal, C. P. R. Hackenberger, *Indian J Chem A* **2013**, *52*, 973-991.

[198] a) M. Royter, M. Schmidt, C. Elend, H. Hobenreich, T. Schafer, U. T. Bornscheuer, G. Antranikian, *Extremophiles* **2009**, *13*, 769-783; b) U. K. Winkler, M. Stuckmann, *J. Bacteriol.* **1979**, *138*, 663-670.

[199] S. Y. Chen, X. R. Zhao, J. Y. Chen, J. Chen, L. Kuznetsova, S. S. Wong, I. Ojima, *Bioconjugate Chem.* **2010**, *21*, 979-987.

[200] T. P. King, S. W. Zhao, T. Lam, *Biochemistry-Us* **1986**, *25*, 5774-5779.

[201] H. Giloh, J. W. Sedat, *Science* **1982**, *217*, 1252-1255.

[202] S. Biswas, K. Kinbara, N. Oya, N. Ishii, H. Taguchi, T. Aida, *J. Am. Chem. Soc.* **2009**, *131*, 7556-+.

[203] R. Bruckner, M. Harmata, *Organic mechanisms : reactions, stereochemistry and synthesis*, Springer, Berlin, **2010**.

[204] J. Tulla-Puche, A. Torres, P. Calvo, M. Royo, F. Albericio, *Bioconjugate Chem.* **2008**, *19*, 1968-1971.
[205] Y. L. Zhang, V. S. Bhatt, G. Y. Sun, P. G. Wang, A. F. Palmer, *Bioconjugate Chem.* **2008**, *19*, 2221-2230.
[206] M. S. Masri, M. Friedman, *J. Protein Chem.* **1988**, *7*, 49-54.
[207] W. Lyford DeLano, PyMOL - molecular visualization system, www.pymol.org.
[208] C. A. Smith, E. J. Wood, *Biological molecules*, 1st ed., Chapman & Hall, London ; New York, **1991**.
[209] M. J. Roberts, M. D. Bentley, J. M. Harris, *Adv Drug Deliver Rev* **2002**, *54*, 459-476.
[210] C. S. Chao, M. C. Chen, S. C. Lin, K. K. T. Mong, *Carbohydr. Res.* **2008**, *343*, 957-964.
[211] F. S. Ekholm, M. Polakova, A. J. Pawlowicz, R. Leino, *Synthesis-Stuttgart* **2009**, 567-576.
[212] a) E. Fischer, E. F. Armstrong, *Berichte der deutschen chemischen Gesellschaft* **1901**, *34*, 2885-2900; b) W. Koenigs, E. Knorr, *Berichte der deutschen chemischen Gesellschaft* **1901**, *34*, 957-981.
[213] P. B. Maloisel J.-L., Tiden A.-K.; Novel Compounds. EU Patent WO9816537, April 23, 1998.
[214] S. I. Presolski, V. Hong, S. H. Cho, M. G. Finn, *J. Am. Chem. Soc.* **2010**, *132*, 14570-14576.
[215] B. D. GmbH, Matrix Guide to Sample Preparation, http://www.chemistry.jhu.edu/Spectrometry/MatrixSelection.pdf.
[216] L. Merkel, Y. Cheburkin, B. Wiltschi, N. Budisa, *Chembiochem* **2007**, *8*, 2227-2232.
[217] a) M. G. Hoesl, C. G. Acevedo-Rocha, S. Nehring, M. Royter, C. Wolschner, B. Wiltschi, N. Budisa, G. Antranikian, *Chemcatchem* **2011**, *3*, 213-221; b) C. G. Acevedo-Rocha, M. G. Hoesl, S. Nehring, M. Royter, C. Wolschner, B. Wiltschi, G. Antranikian, N. Budisa, *Catal Sci Technol* **2013**, *3*, 1198-1201.
[218] G. T. Hermanson, *Bioconjugate techniques*, 2. ed., Elsevier Acad. Press, Amsterdam [u.a.], **2008**.
[219] C. H. Wong, S. C. Zimmerman, *Chem. Commun.* **2013**, *49*, 1679-1695.
[220] K. Turton, R. Natesh, N. Thiyagarajan, J. A. Chaddock, K. R. Acharya, *Glycobiology* **2004**, *14*, 923-929.
[221] L. L. Kiessling, J. E. Gestwicki, L. E. Strong, *Angew. Chem. Int. Ed.* **2006**, *45*, 2348-2368.
[222] a) D. Forget, D. Boturyn, E. Defrancq, J. Lhomme, P. Dumy, *Chem-Eur J* **2001**, *7*, 3976-3984; b) C. T. Pool, J. G. Boyd, J. P. Tam, *J Pept Res* **2004**, *63*, 223-234.
[223] D. Bonnet, K. Thiam, E. Loing, O. Melnyk, H. Gras-Masse, *J. Med. Chem.* **2001**, *44*, 468-471.
[224] M. Villien, E. Defrancq, P. Dumy, *Nucleos Nucleot Nucl* **2004**, *23*, 1657-1666.
[225] J. D. Wade, K. Hojo, K. Kawasaki, T. G. Johns, B. Catimel, J. Rothacker, E. C. Nice, *Anal. Biochem.* **2006**, *348*, 315-317.
[226] W. W. Cleland, *Biochemistry-Us* **1964**, *3*, 480-&.
[227] C. Hwang, A. J. Sinskey, H. F. Lodish, *Science* **1992**, *257*, 1496-1502.
[228] U. T. Rüegg, J. Rudinger, in *Methods in Enzymology, Vol. Volume 47* (Eds.: C. H. W. Hirs, N. T. Serge), Academic Press, **1977**, pp. 111-116.
[229] O. Shimomura, F. H. Johnson, Y. Saiga, *Journal of Cellular and Comparative Physiology* **1962**, *59*, 223-&.
[230] M. Chalfie, S. Kain, *Green fluorescent protein : properties, applications, and protocols*, 2nd ed., Wiley-Interscience, Hoboken, N.J., **2006**.
[231] F. Yang, L. G. Moss, G. N. Phillips, *Nat. Biotechnol.* **1996**, *14*, 1246-1251.
[232] a) B. Hao, W. M. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, M. K. Chan, *Science* **2002**, *296*, 1462-1466; b) G. Srinivasan, C. M. James, J. A. Krzycki, *Science* **2002**, *296*, 1459-1462.
[233] W. T. Li, A. Mahapatra, D. G. Longstaff, J. Bechtel, G. Zhao, P. T. Kang, M. K. Chan, J. A. Krzycki, *J. Mol. Biol.* **2009**, *385*, 1156-1164.

[234] P. Wu, W. Q. Shui, B. L. Carlson, N. Hu, D. Rabuka, J. Lee, C. R. Bertozzi, *P Natl Acad Sci USA* **2009**, *106*, 3000-3005.
[235] a) X. Li, T. Fekner, J. J. Ottesen, M. K. Chan, *Angew. Chem. Int. Ed.* **2009**, *48*, 9184-9187; b) D. P. Nguyen, T. Elliott, M. Holt, T. W. Muir, J. W. Chin, *J. Am. Chem. Soc.* **2011**, *133*, 11418-11421.
[236] a) H. Kunz, C. Unverzagt, *Angewandte Chemie-International Edition in English* **1984**, *23*, 436-437; b) S. R. Chhabra, B. Hothi, D. J. Evans, P. D. White, B. W. Bycroft, W. C. Chan, *Tetrahedron Lett.* **1998**, *39*, 1603-1606.
[237] M. E. Jung, S. W. Yi, *Tetrahedron Lett.* **2012**, *53*, 4216-4220.
[238] R. Heathcote, J. A. S. Howell, N. Jennings, D. Cartlidge, L. Cobden, S. Coles, M. Hursthouse, *Dalton T* **2007**, 1309-1315.
[239] G. Fernandez, F. Garcia, L. Sanchez, *Chem. Commun.* **2008**, 6567-6569.
[240] G. K. Farrington, A. Kumar, F. C. Wedler, *Org. Prep. Proced. Int.* **1989**, *21*, 390-392.
[241] A. Chakraborty, S. Dey, S. Sawoo, N. N. Adarsh, A. Sarkar, *Organometallics* **2010**, *29*, 6619-6622.
[242] S. J. Wey, M. E. Augustyniak, E. D. Cochran, J. L. Ellis, X. Q. Fang, D. S. Garvey, D. R. Janero, L. G. Letts, A. M. Martino, T. L. Melim, M. G. Murty, S. K. Richardson, J. D. Schroeder, W. M. Selig, A. M. Trocha, R. S. Wexler, D. V. Young, I. S. Zemtseva, B. M. Zifcak, *J. Med. Chem.* **2007**, *50*, 6367-6382.
[243] F. He, Y. L. Tang, M. H. Yu, S. Wang, Y. L. Li, D. B. Zhu, *Adv. Funct. Mater.* **2007**, *17*, 996-1002.
[244] K. C. Patil, J. P. Vittal, C. C. Patel, *P Indian as-Chem Sci* **1980**, *89*, 87-93.
[245] T. N. Van, S. Claessens, P. Habonimana, K. A. Tehrani, L. Van Puyvelde, N. De Kimpe, *Synlett* **2006**, 2469-2471.
[246] H. Cheng, X. H. Cao, M. Xian, L. Y. Fang, T. B. Cai, J. J. Ji, J. B. Tunac, D. X. Sun, P. G. Wang, *J. Med. Chem.* **2005**, *48*, 645-652.
[247] X. R. Li, J. Guo, J. Asong, M. A. Wolfert, G. J. Boons, *J. Am. Chem. Soc.* **2011**, *133*, 11147-11153.
[248] T. P. Wang, Y. J. Chiou, Y. Chen, E. C. Wang, L. C. Hwang, B. H. Chen, Y. F. Chen, C. H. Ko, *Bioconjugate Chem.* **2010**, *21*, 1642-1655.
[249] C. Gnaccarini, W. Ben-Tahar, W. D. Lubell, J. N. Pelletier, J. W. Keillor, *Biorg. Med. Chem.* **2009**, *17*, 6354-6359.
[250] J. del Barrio, L. Oriol, R. Alcala, C. Sanchez, *Macromolecules* **2009**, *42*, 5752-5760.

7. Appendix

7.1. NMR spectra

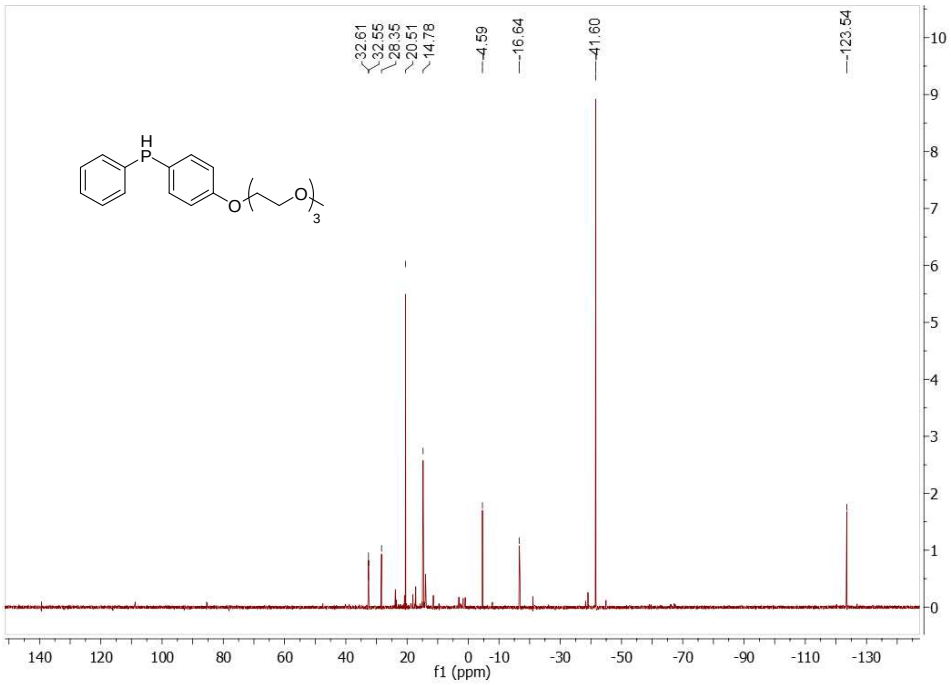


Figure 7.1: ^{31}P-NMR of crude 20 before borane protection.

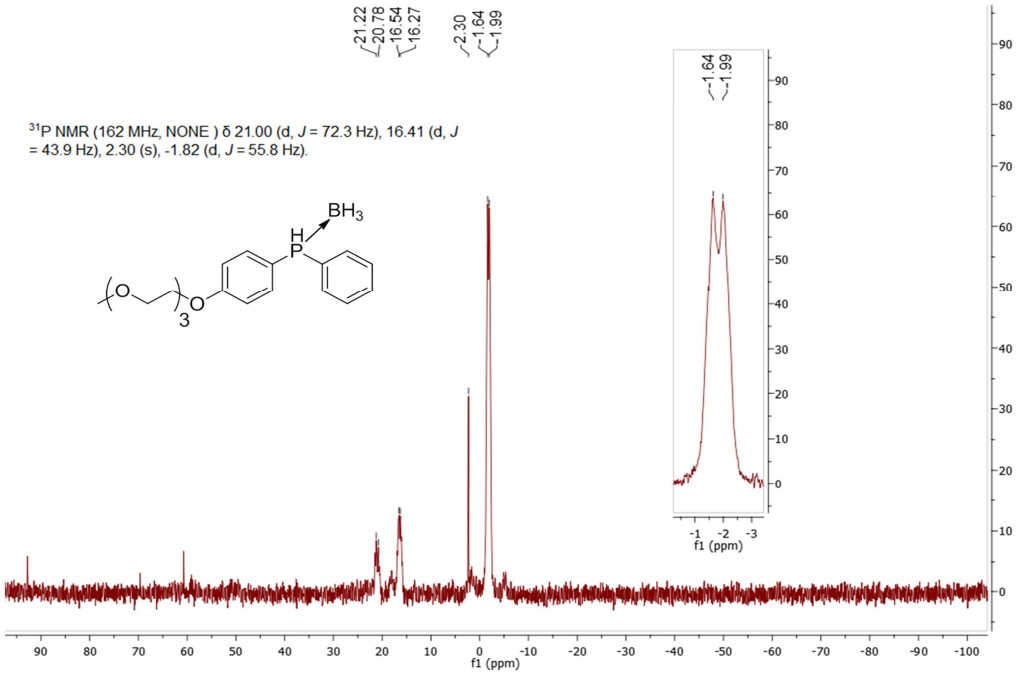


Figure 7.2: ^{31}P-NMR of compound 22.

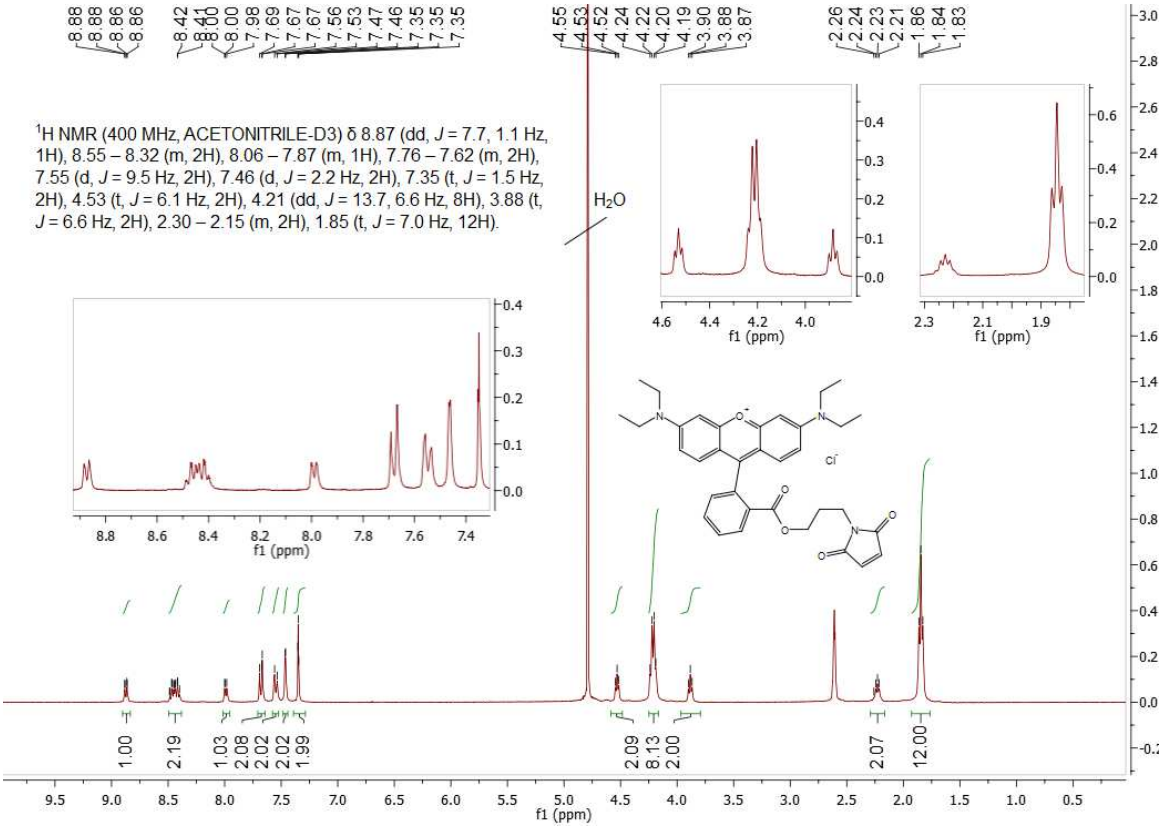


Figure 7.3: ^{1}H-NMR of compound 77.

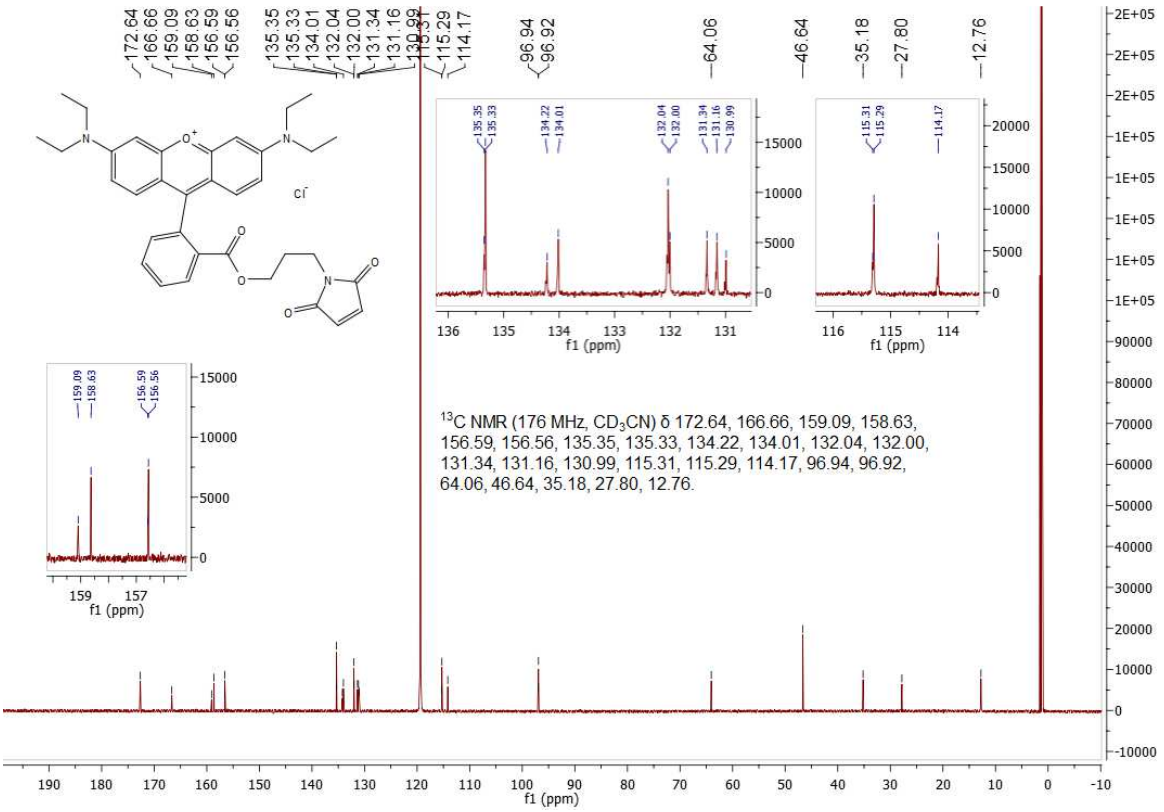


Figure 7.4: ^{13}C-NMR of compound 77.

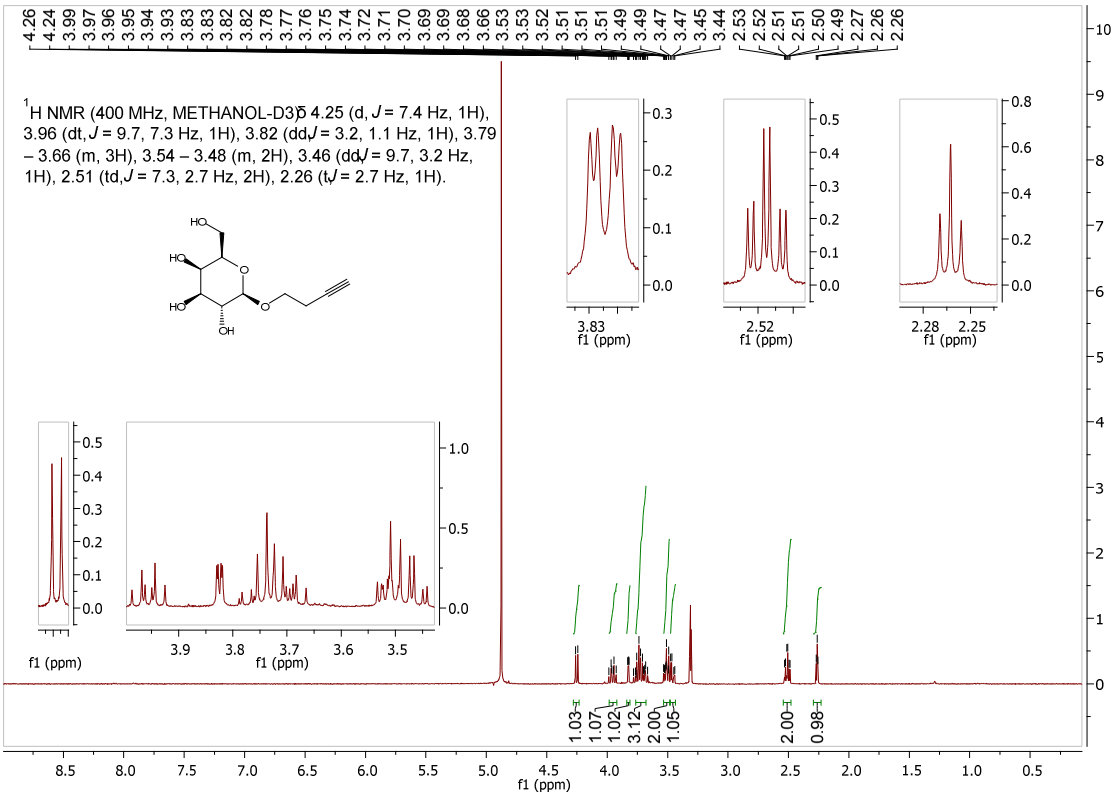


Figure 7.5: ^{1}H-NMR (CD_3OD, 400 MHz) of compound 84.

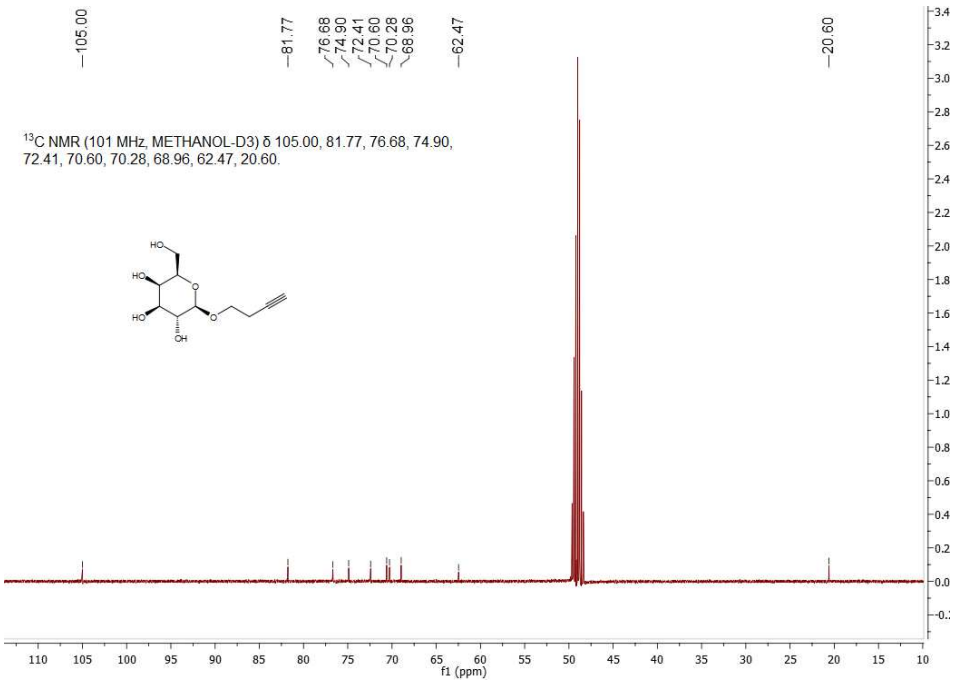


Figure 7.6: ^{13}C-NMR (CD_3OD, 100 MHz) of compound 84.

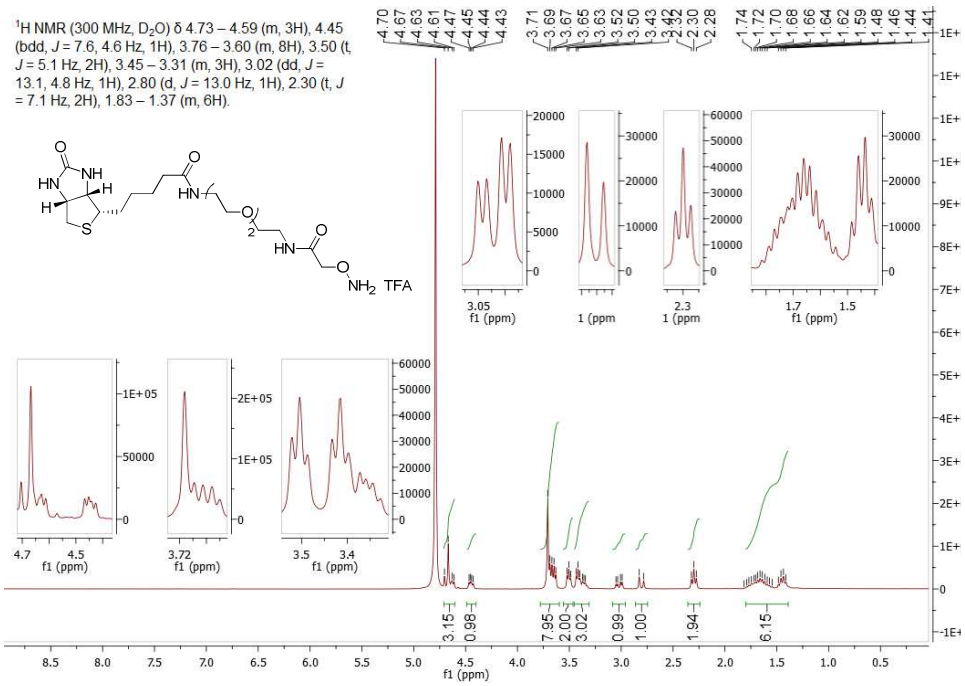


Figure 7.7: ^{1}H-NMR (D_2O, 400 MHz) of compound 95.

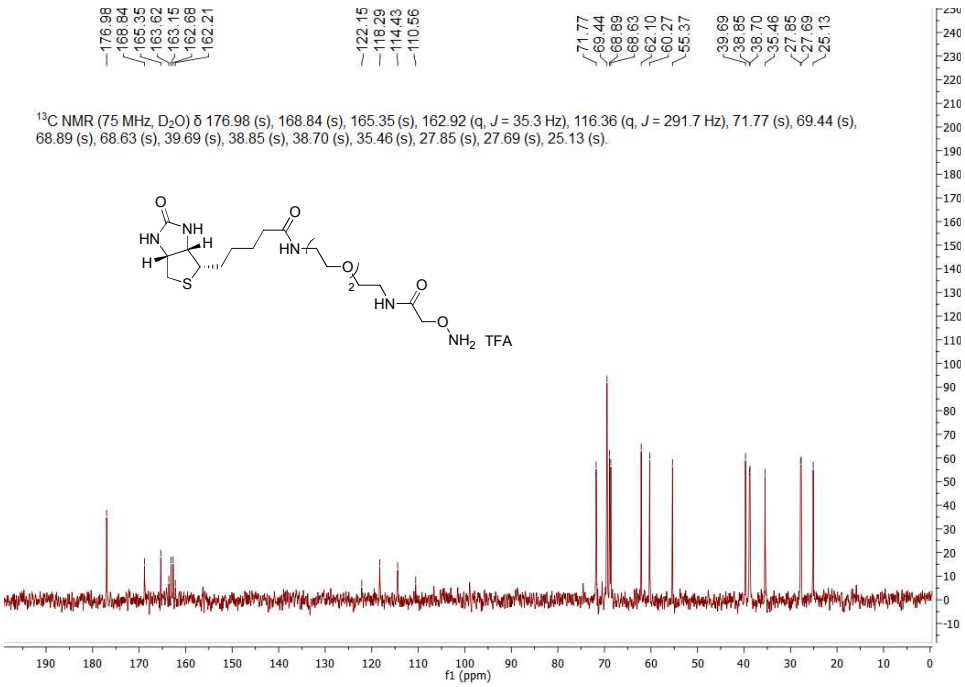


Figure 7.8: ^{13}C-NMR (D_2O, 100 MHz) of compound 95.

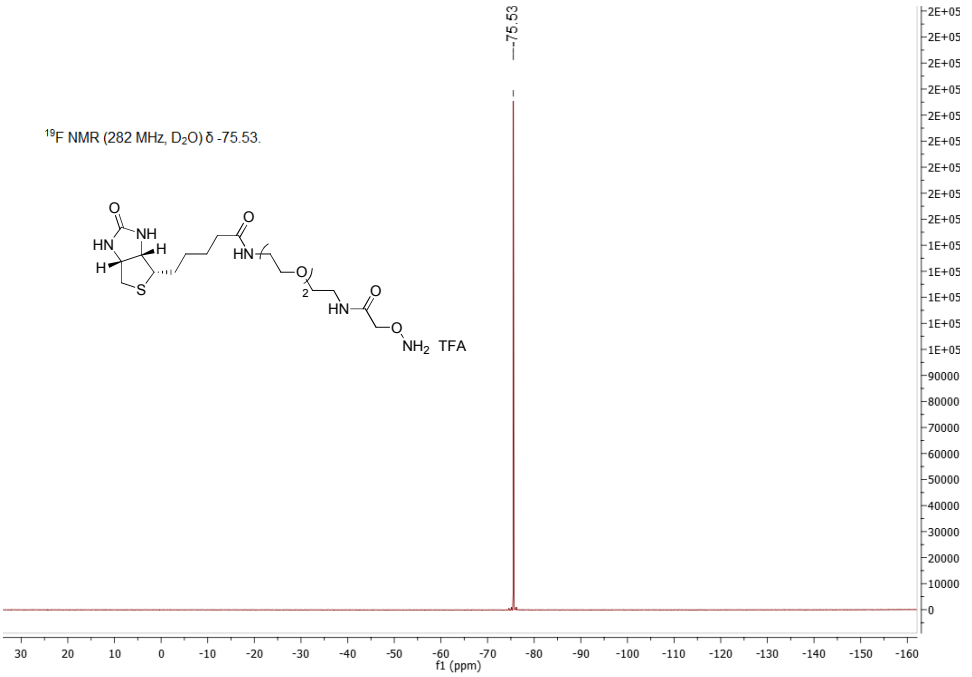


Figure 7.9: ^{19}F-NMR (D_2O, 376 MHz) of compound 95.

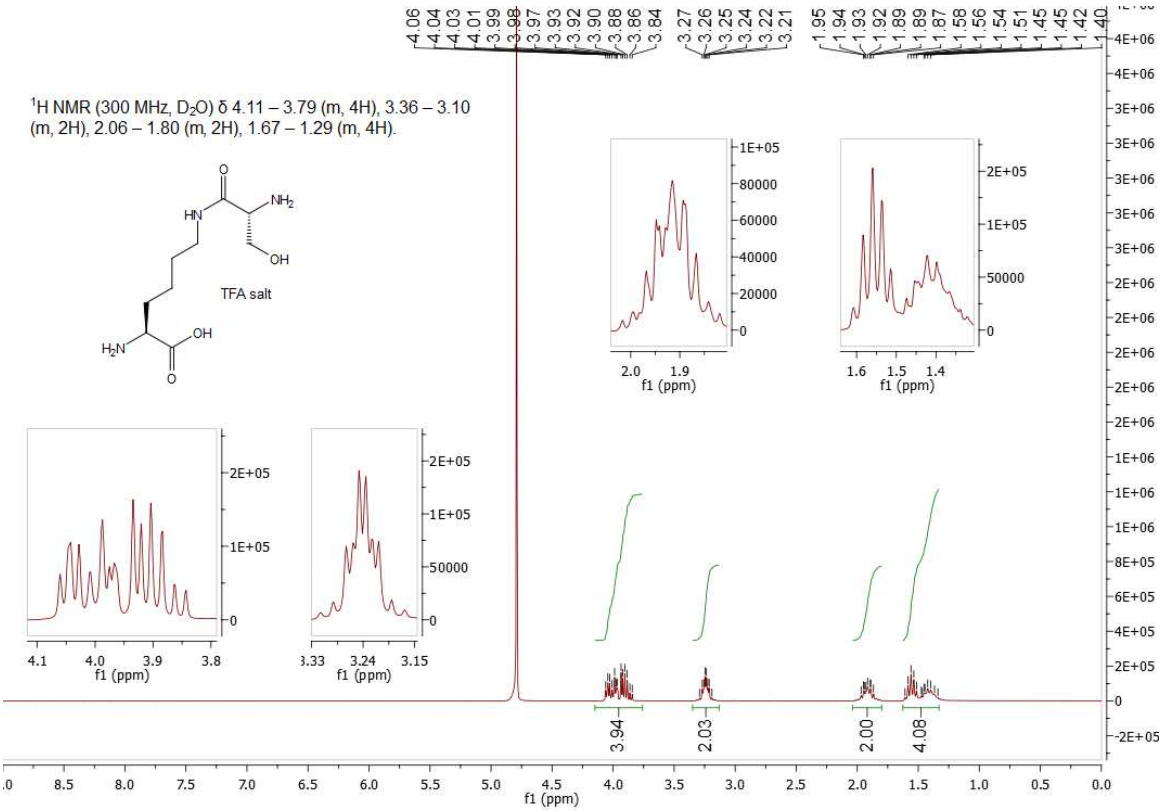


Figure 7.10: ^{1}H-NMR of compound 99.

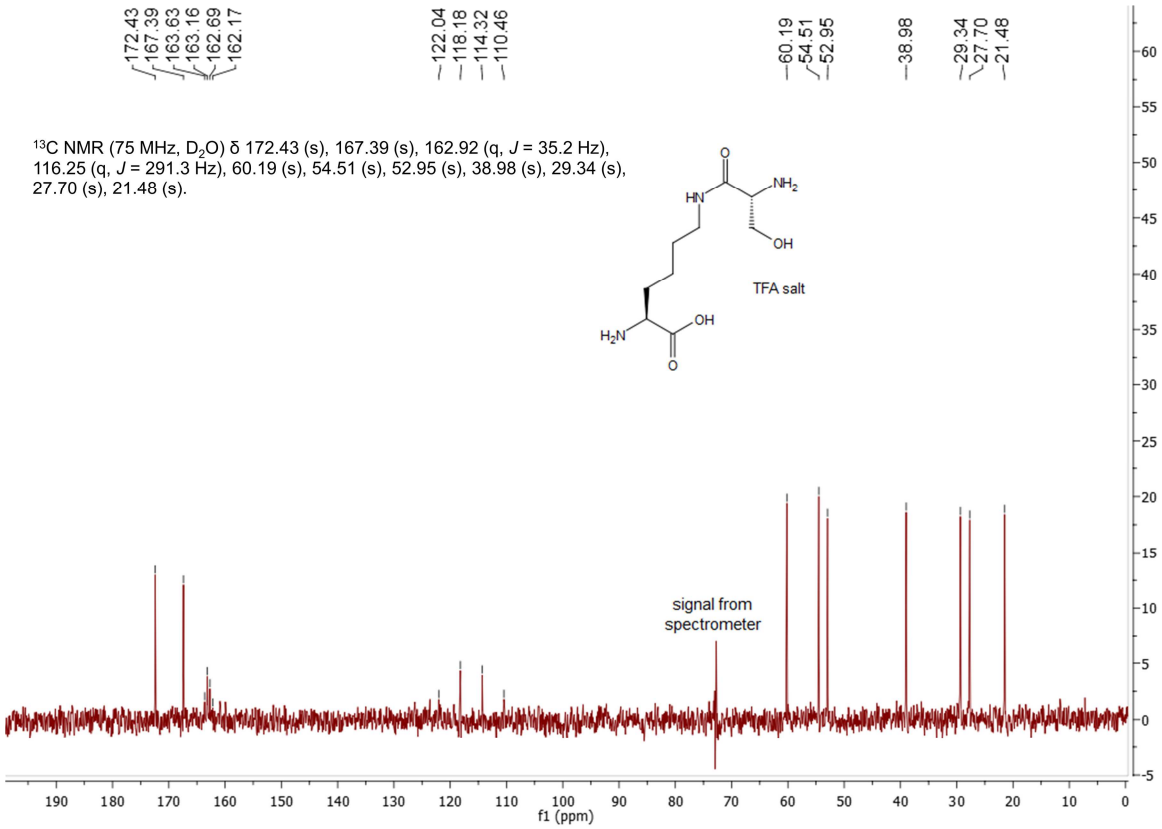


Figure 7.11: ^{13}C-NMR of compound 99.

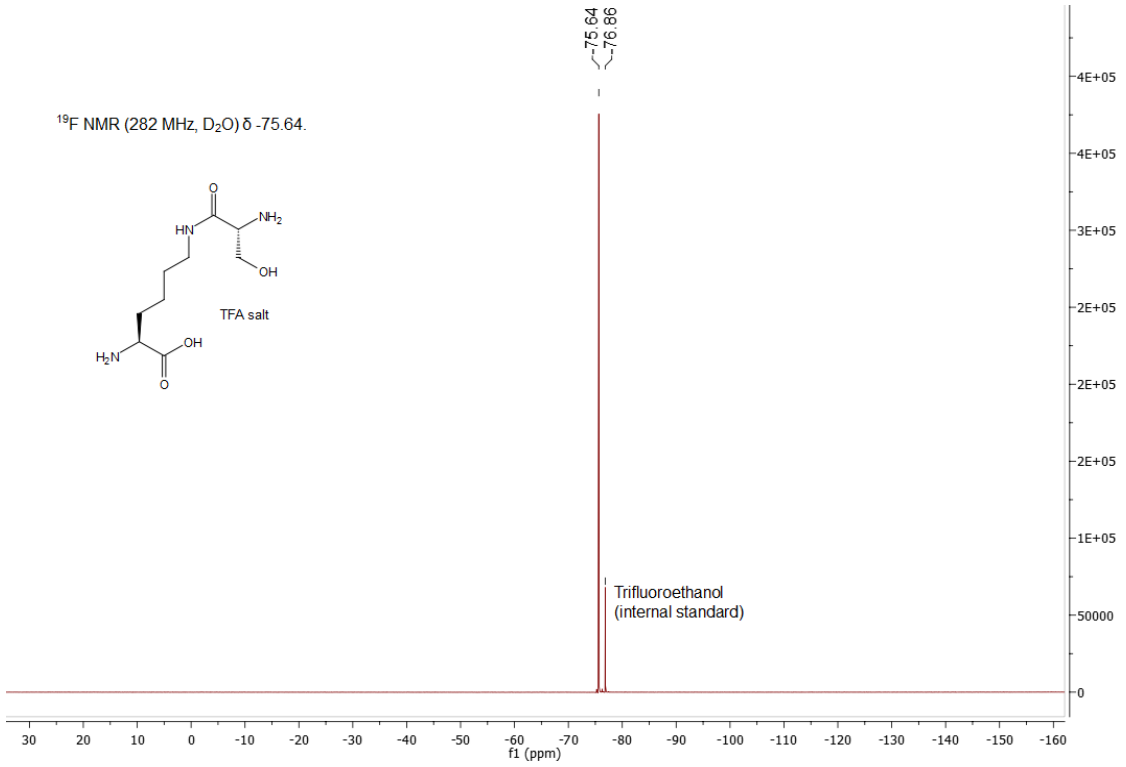


Figure 7.12: ^{19}F-NMR of compound 99.

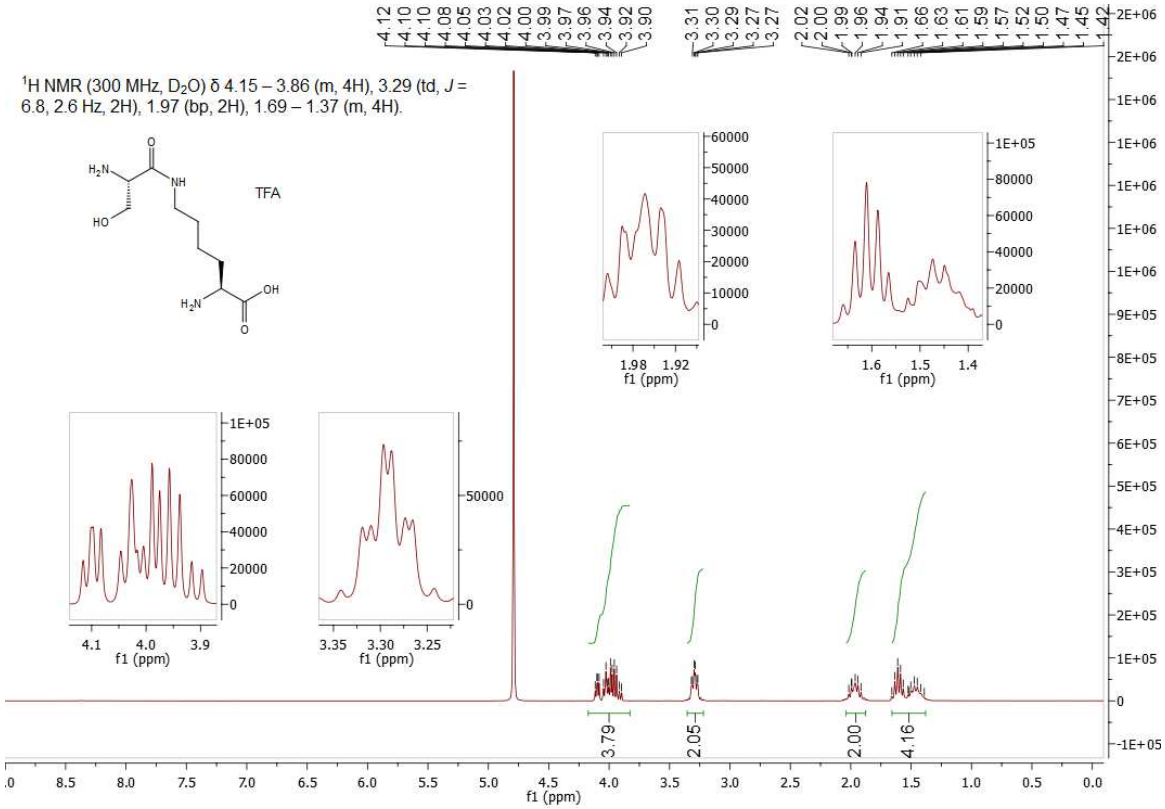


Figure 7.13: ^{1}H-NMR of compound 100.

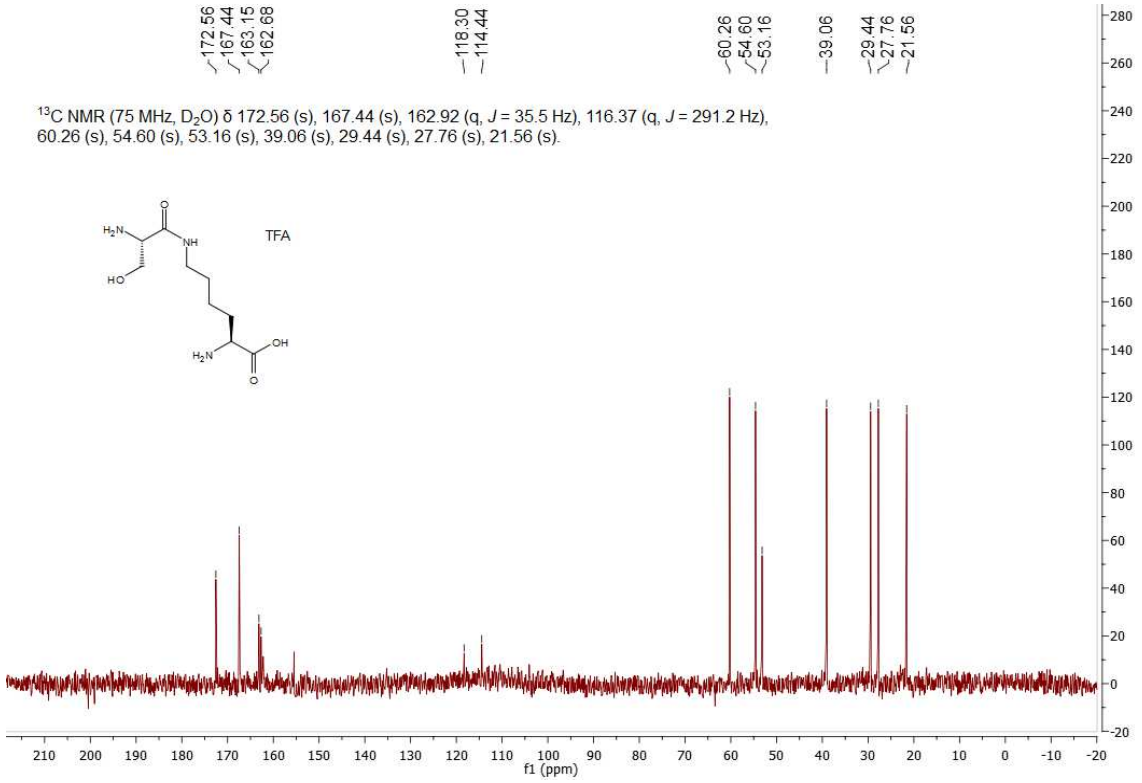


Figure 7.14: ^{13}C-NMR of compound 100.

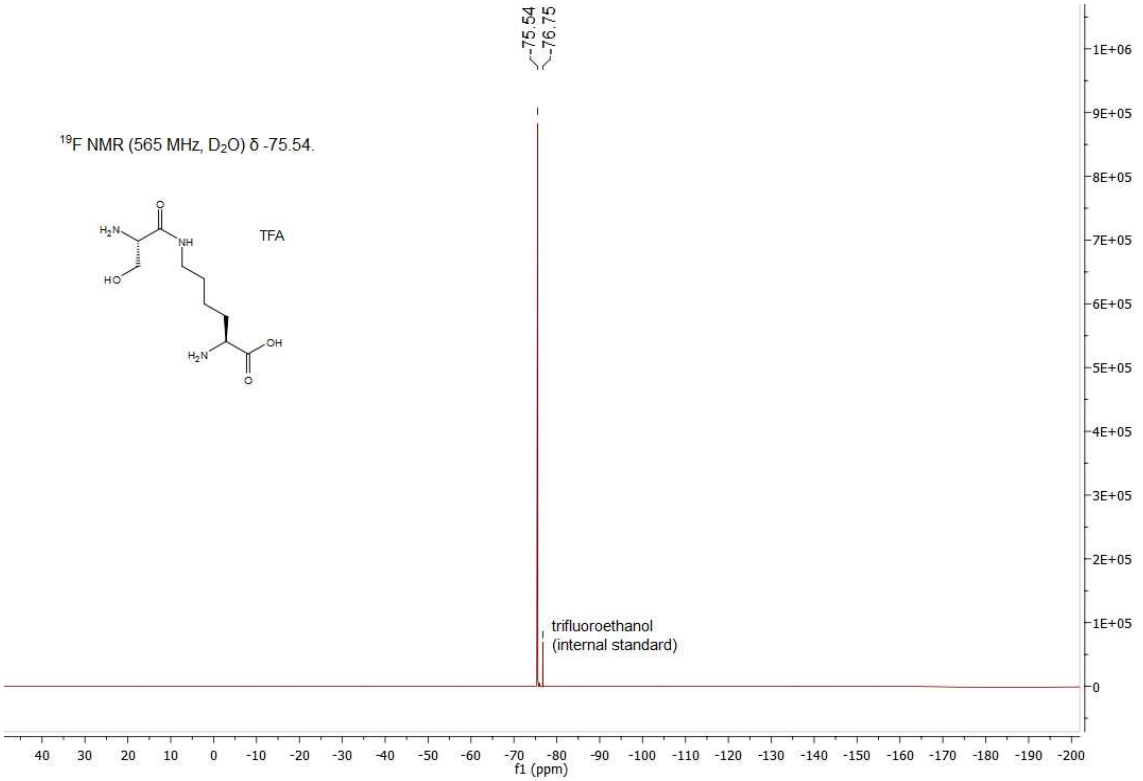


Figure 7.15: ^{19}F-NMR of compound 100.

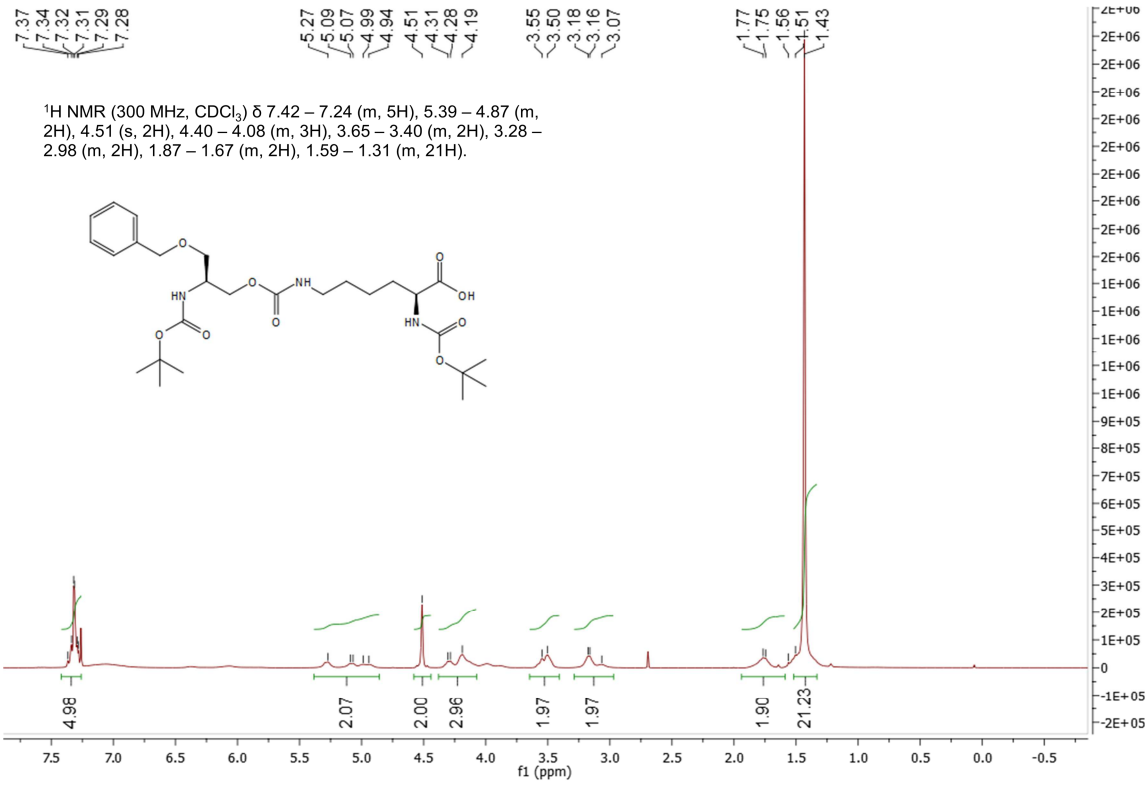


Figure 7.16: ^{1}H-NMR of compound 113.

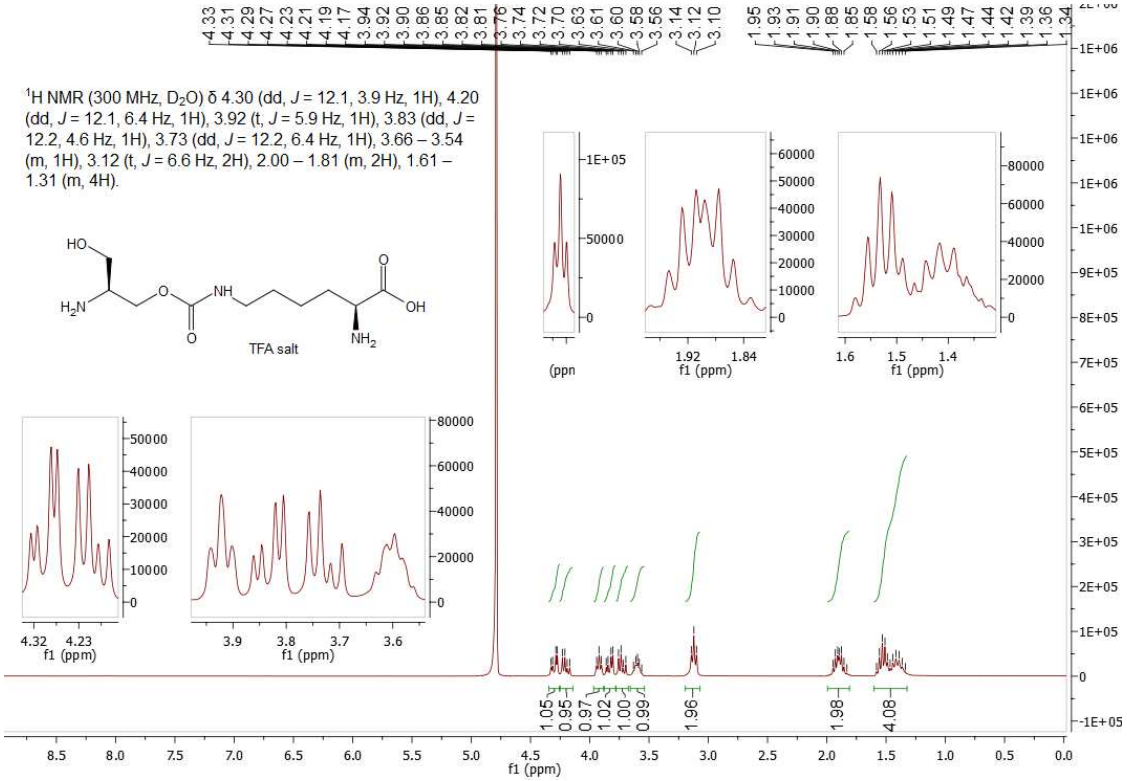


Figure 7.17: ^{1}H-NMR of compound 114.

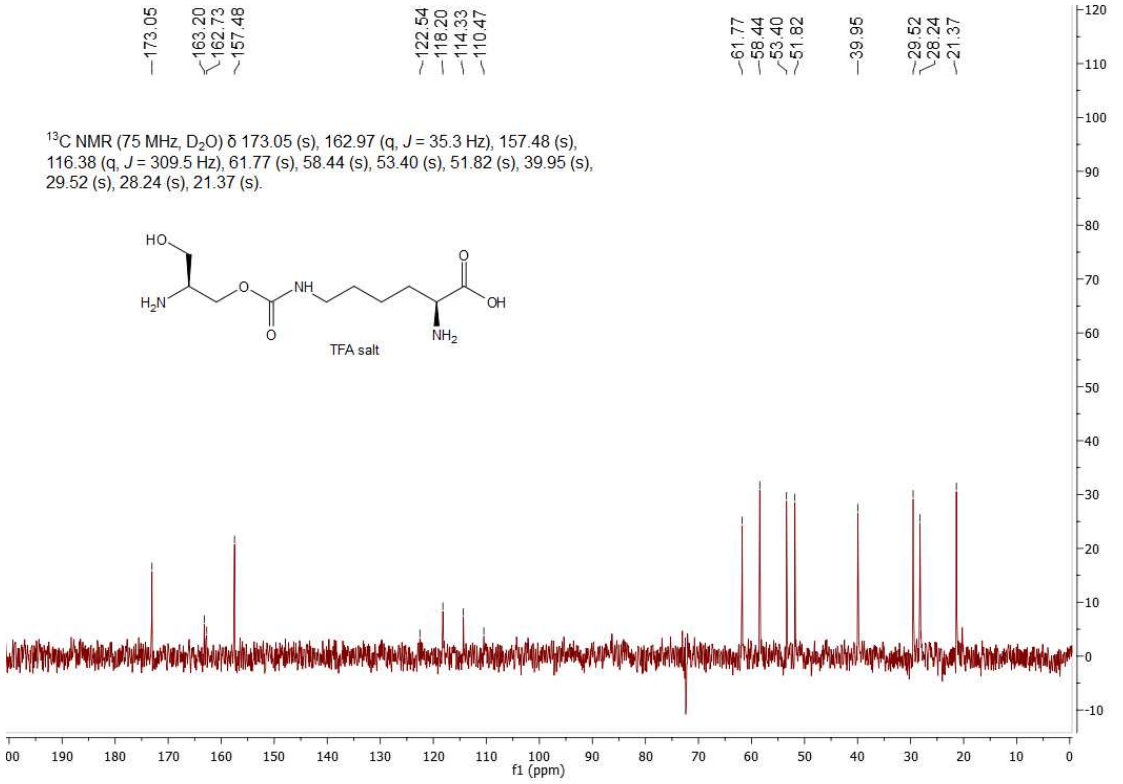


Figure 7.18: ^{13}C-NMR of compound 114.

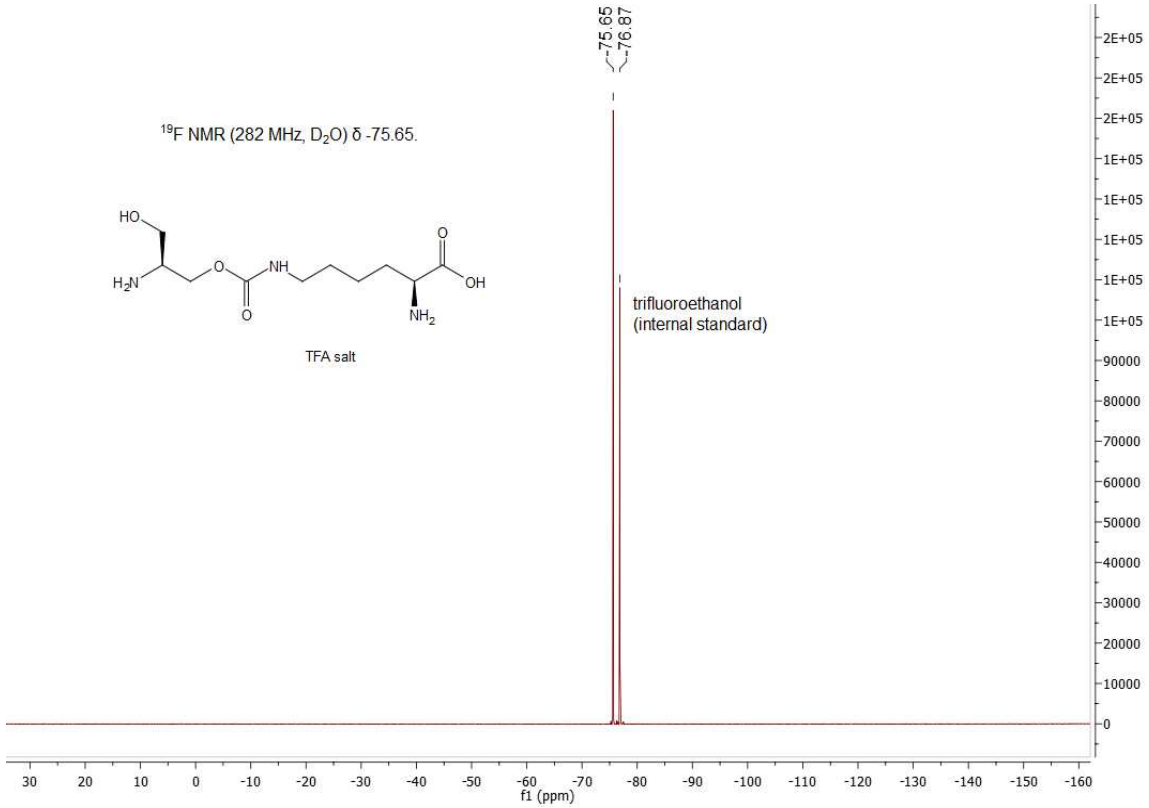


Figure 7.19: ^{19F}-NMR of compound 114.

7.2. HRMS Data

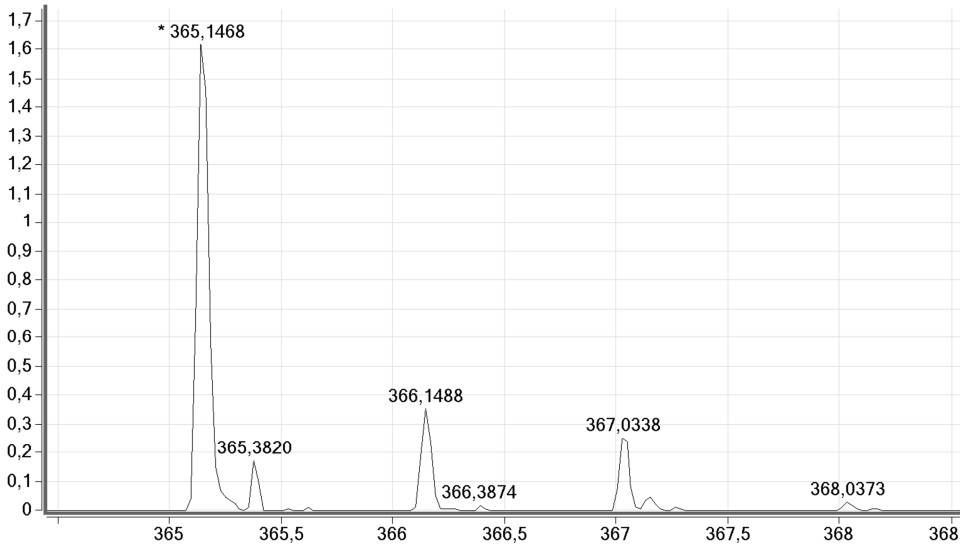


Figure 7.20: HRMS (ESI-ToF) of compound 21.

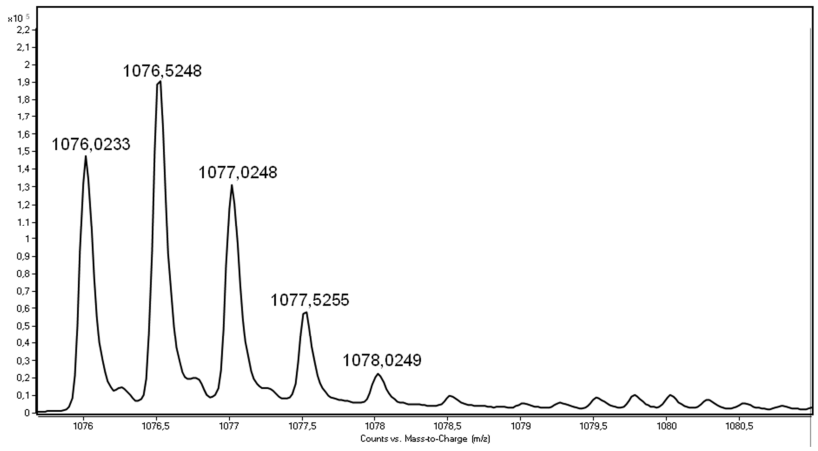


Figure 7.21: HRMS (ESI-ToF) of compound 47.

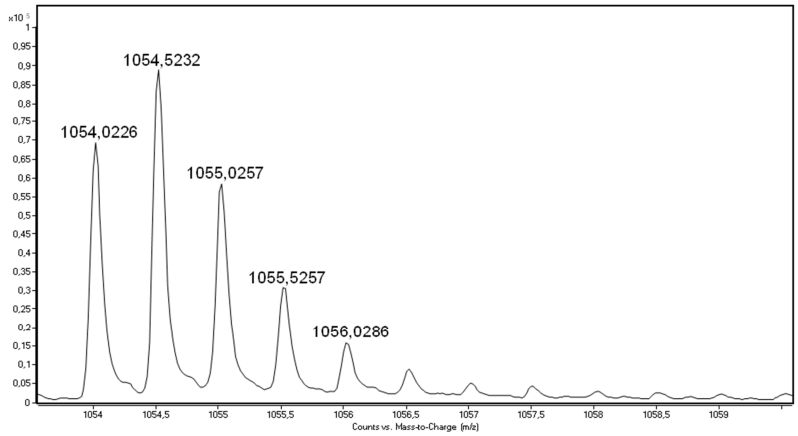


Figure 7.22: HRMS (ESI-ToF) of detected "reduction–dehydration product" with Tmob peptide thioester 51.

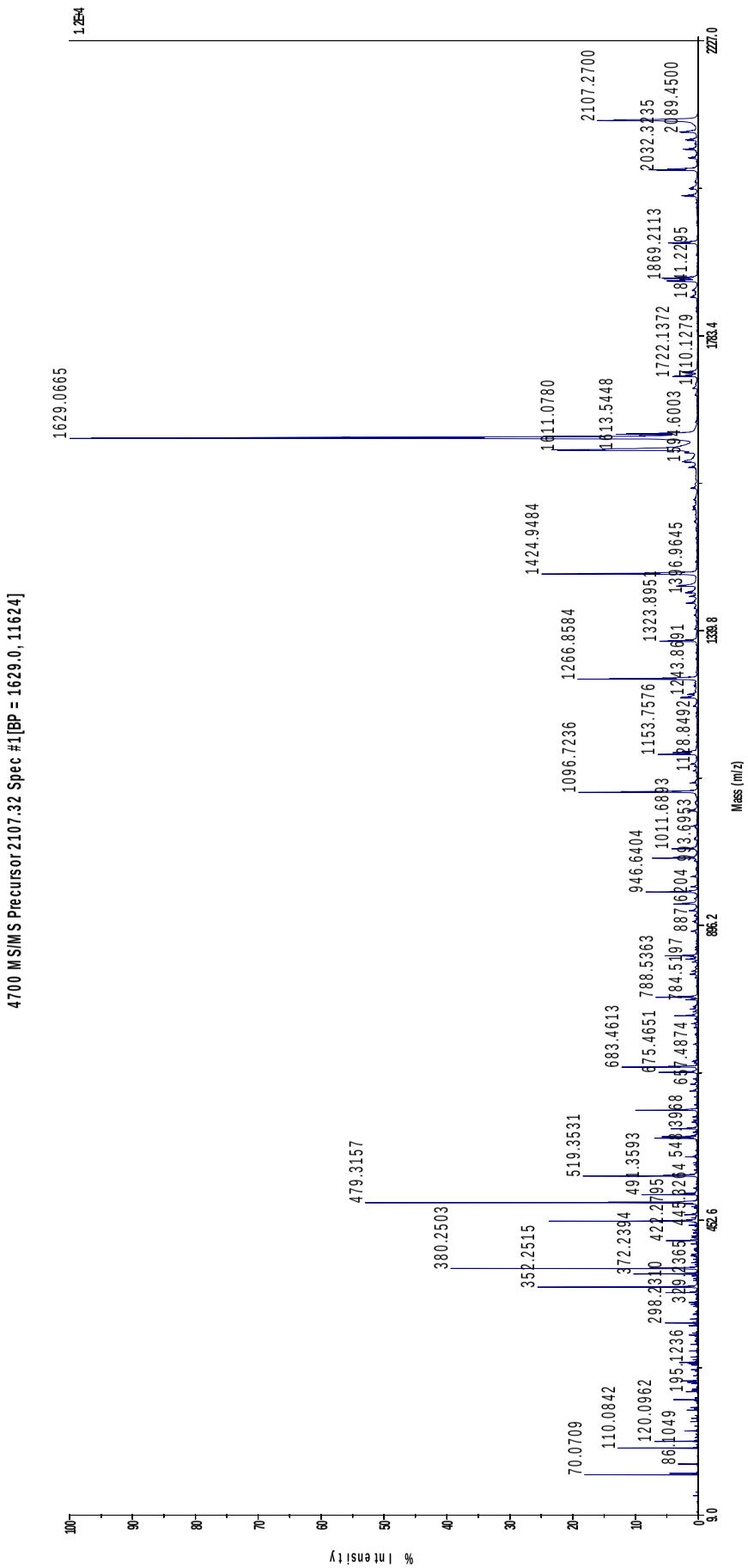


Figure 7.23: MALDI MS/MS spectrum of glutarimide peptide 56.

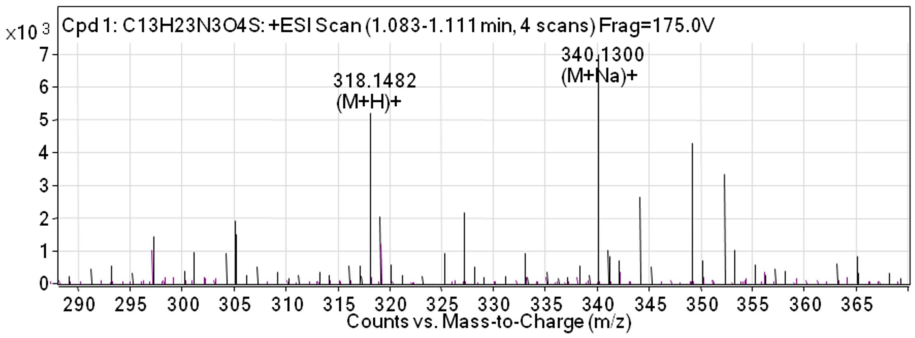


Figure 7.24: HRMS (ESI-ToF) of compound 66.

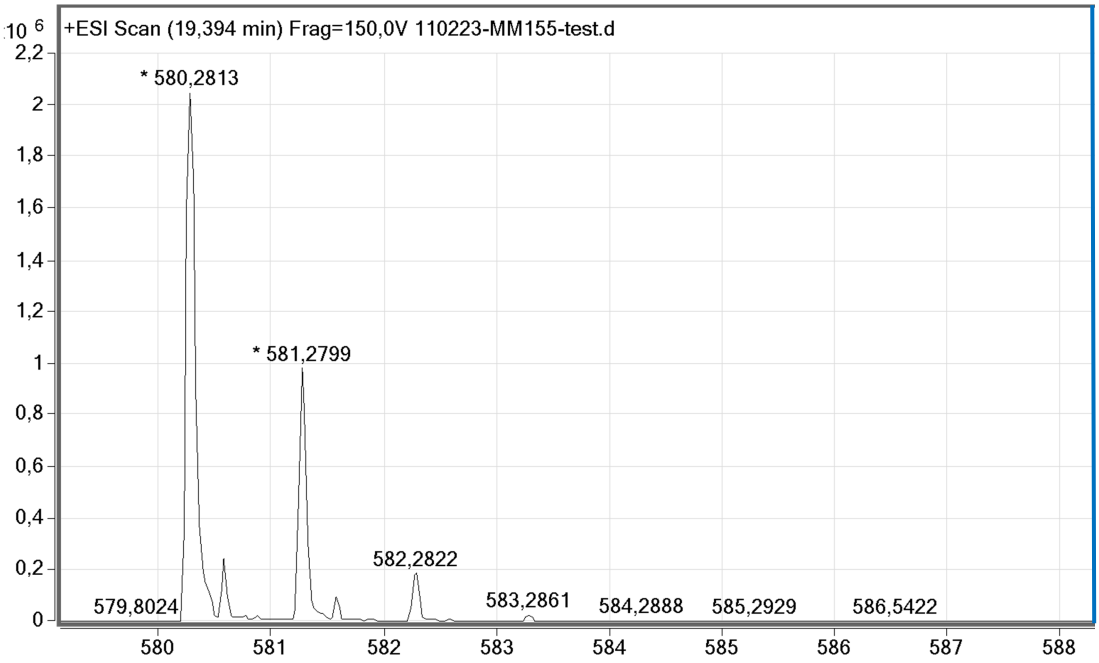


Figure 7.25: HRMS (ESI-ToF) of compound 77.

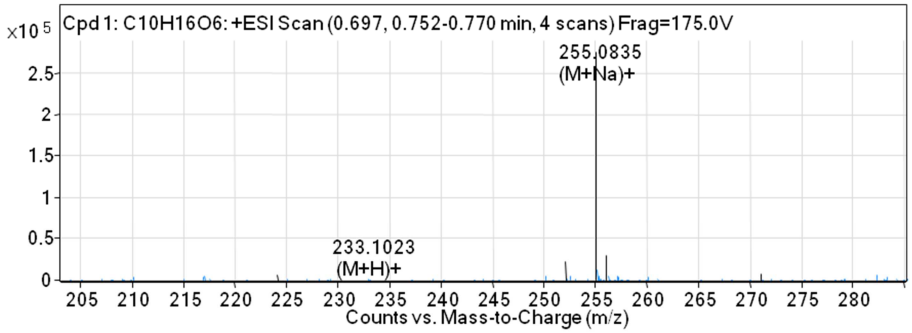


Figure 7.26: HRMS (ESI-ToF) of compound 84.

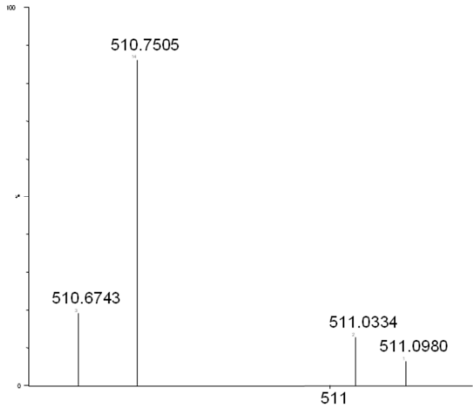


Figure 7.27: HRMS (ESI-ToF) of compound 85.

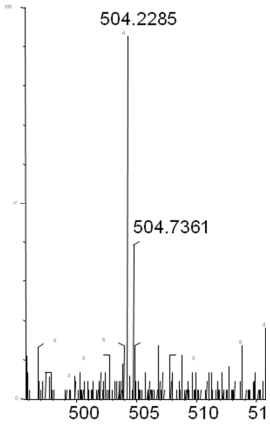


Figure 7.28: HRMS (ESI-ToF) of compound 86.

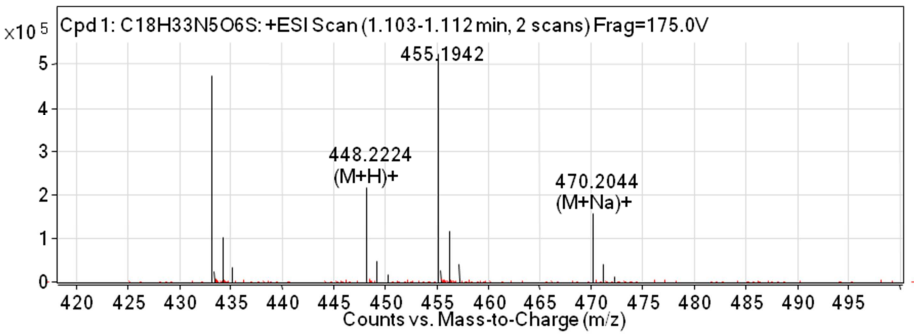


Figure 7.29: HRMS (ESI-ToF) of compound 95.

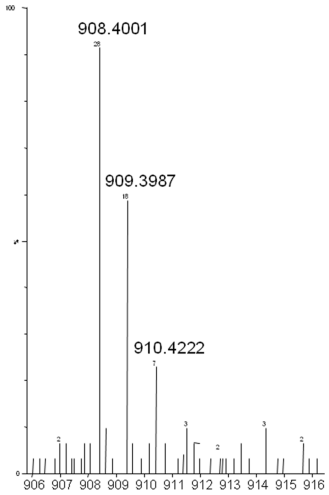


Figure 7.30: HRMS (ESI-ToF) of compound 96.

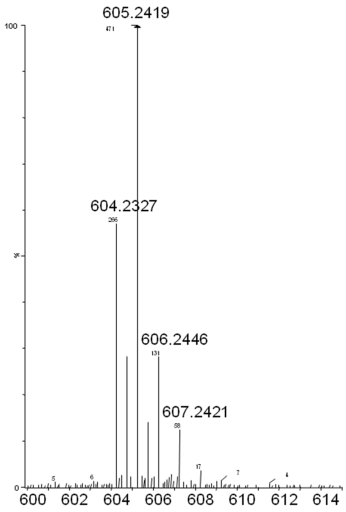


Figure 7.31: HRMS (ESI-ToF) of compound 97.

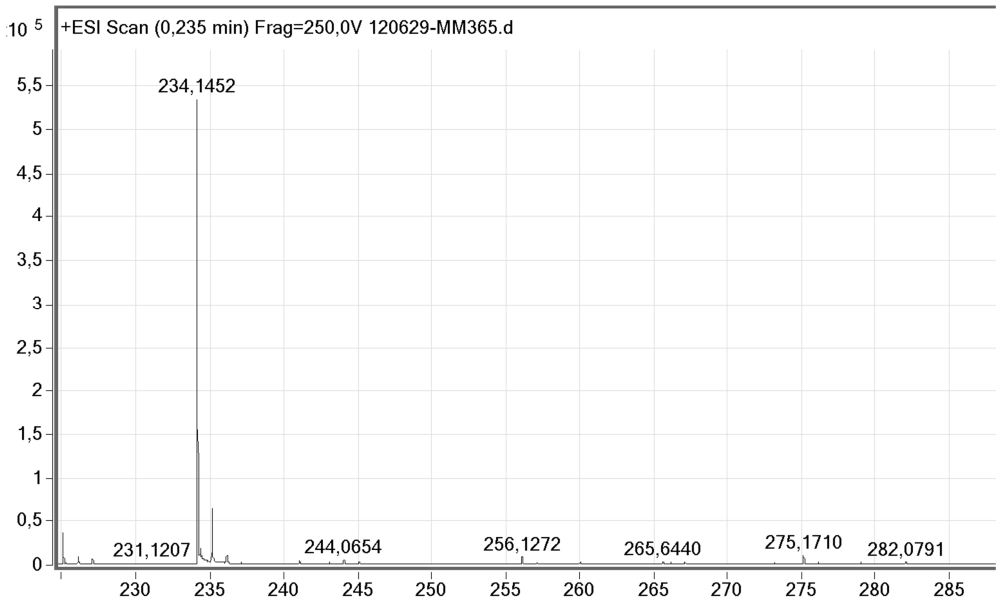


Figure 7.32: HRMS (ESI-ToF) of compound 203.

7.3. Hydrazone Formation on Small Molecules (Selected UV-Traces)

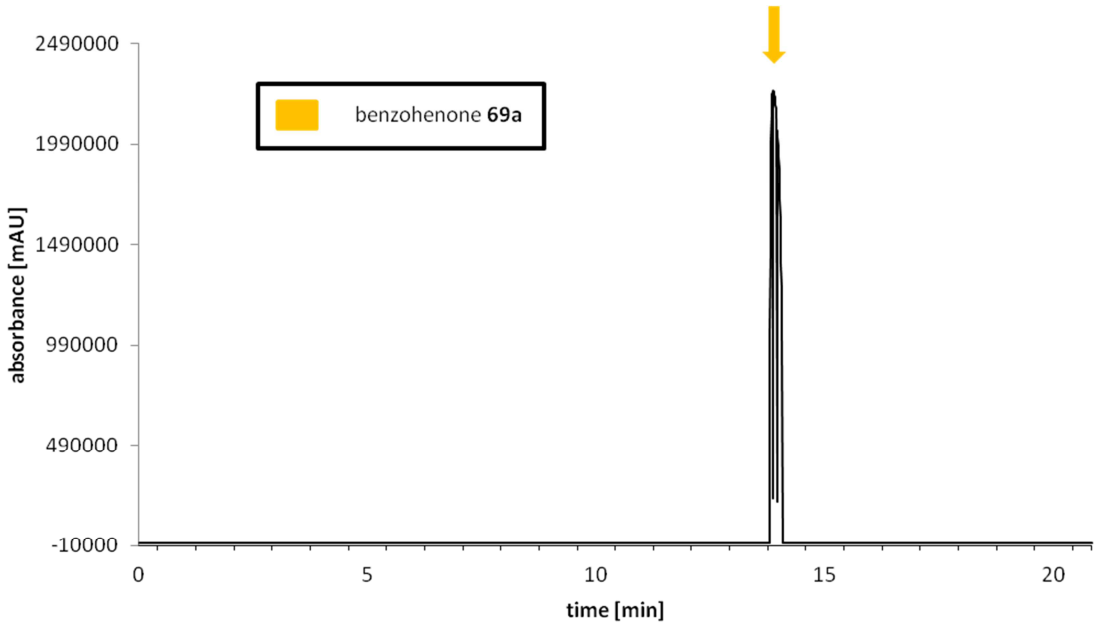


Figure 7.33: LCUV chromatogram (280 nm) of hydrazone formation with benzophenone 69a and biotin hydrazide 68 (Table 5.6, entry 3).

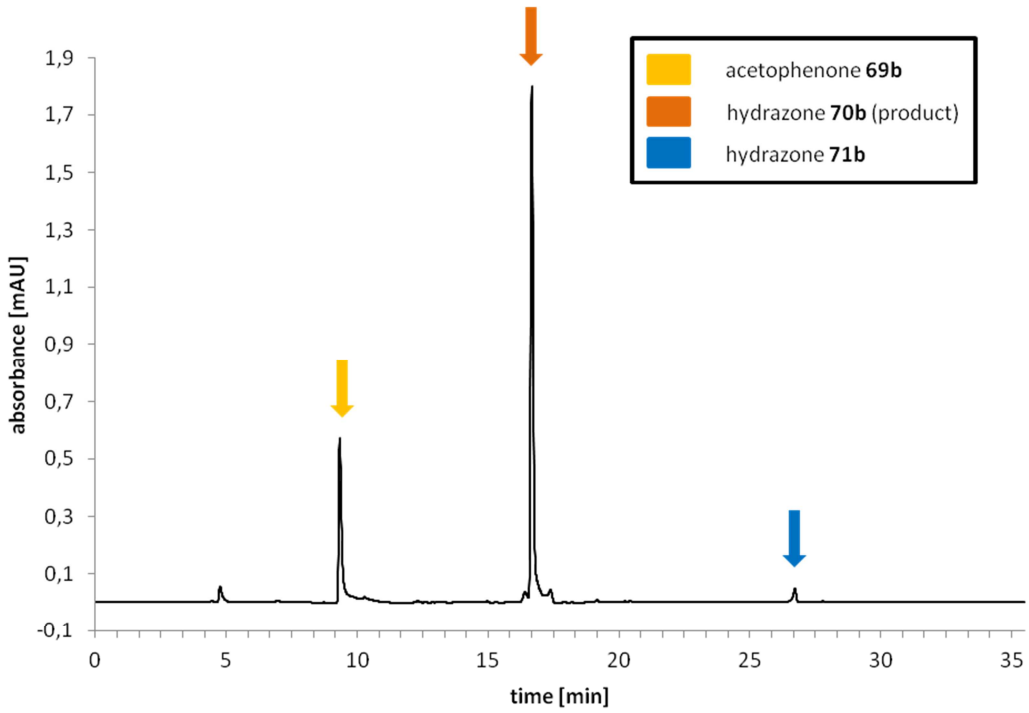


Figure 7.34: LCUV chromatogram (280 nm) of hydrazone formation with acetophenone 69b and biotin hydrazide 68 (Table 5.6, entry 6).

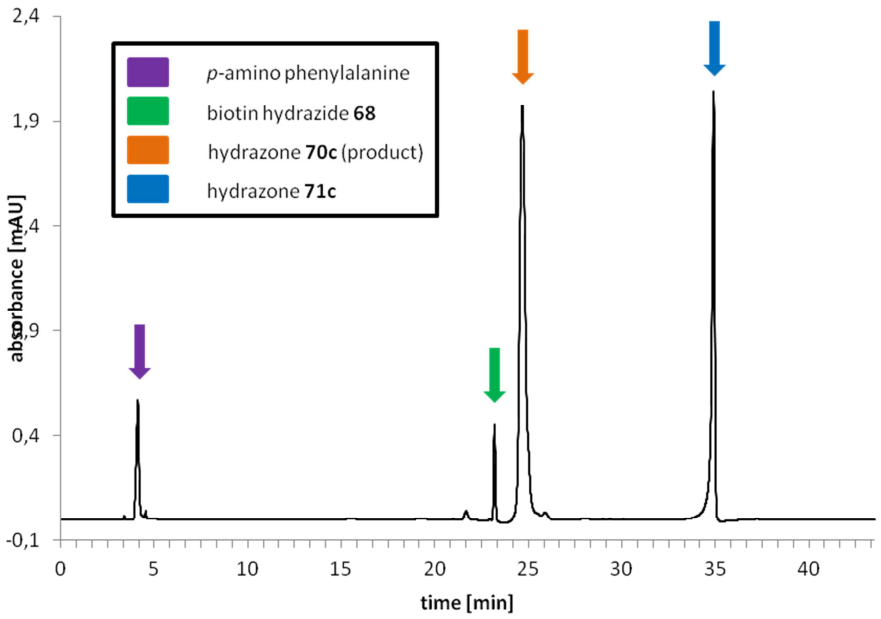


Figure 7.35: LCUV chromatogram (280 nm) of hydrazone formation with benzaldehyde 69c and biotin hydrazide 68 (Table 5.6, entry 19).

7.4. Oxime Formation on Aldehyde Peptide 86 (Selected UV-Traces)

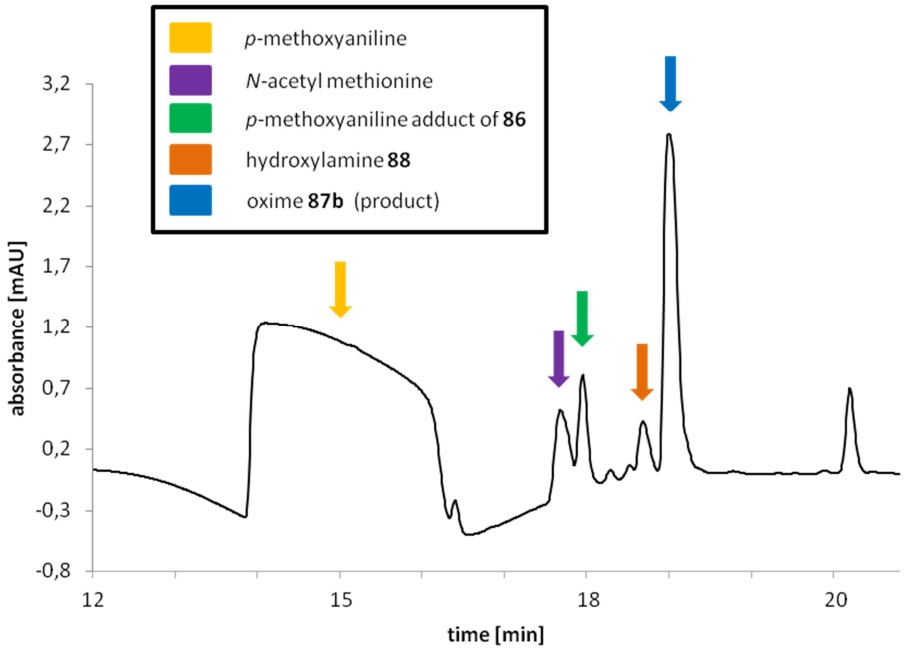


Figure 7.36: LCUV chromatogram (220 nm) of oxime formation with hydroxylamine 88 on peptide 86 (Table 5.12, entry 5).

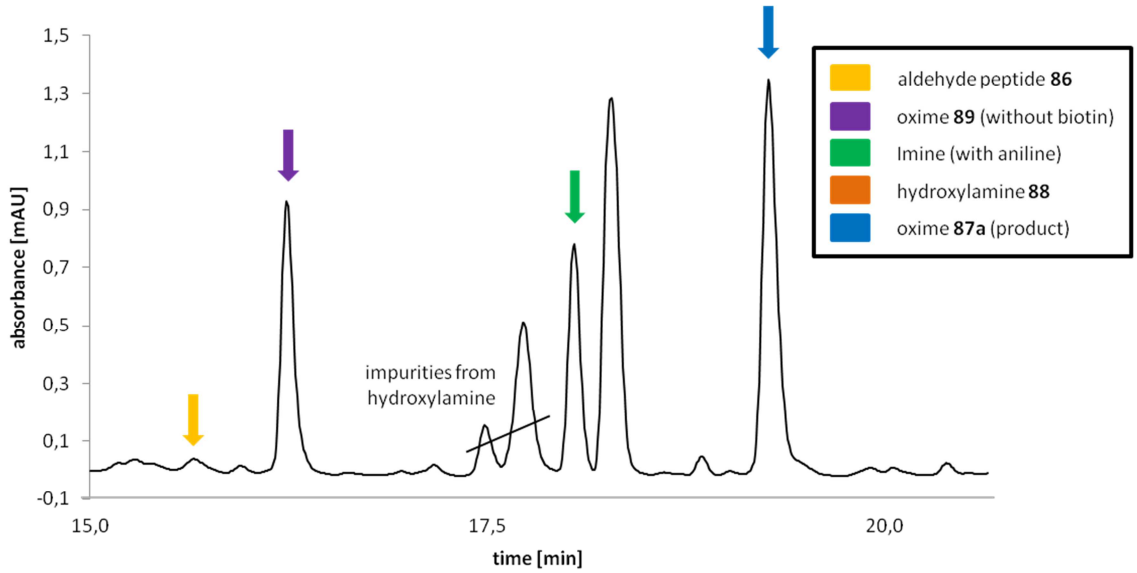


Figure 7.37: LCUV chromatogram (220 nm) of oxime formation with hydroxylamine 66 on peptide 86 (Table 5.12, entry 2).