universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Synthesis of new generation PEGylated ligation auxiliaries for glycopeptide preparation"

verfasst von / submitted by

Maximilian Schrems BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2017 / Vienna 2017

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 863
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Biologische Chemie
Betreut von / Supervisor:	Univ.-Prof Dr. Christian Becker
Mitbetreut von / Co-Supervisor:	Dr. Claudia Bello

I Abstract

The combination of the inherent complexity of carbohydrates and the intricate cellular synthesis pathways so far only allowed for a modest understanding of the functions of protein glycosylation and also slowed down our ability to fully utilize the potential of these posttranslational modifications. Progress in the field of glycobiology may have a direct impact on various fields of research, such as drug development, immunology, general medicine and cell biology. This can only be achieved by expanding the spectrum of biological and chemical tools for the synthesis of glycoproteins.

A recently developed PEGylated ligation auxiliary exploits the advantages of ligation auxiliaries and adds functionality by introducing a polyethylene glycol (PEG) polymer in the scaffold that tunes solubility properties enabling controlled precipitation of peptide-auxiliary conjugates. Via this temporary modification of the peptide, elaborate purification procedures between steps of enzymatic peptide glycosylation can be reduced to simple precipitation and centrifugation. The possibility of combining native chemical ligation and mild removal conditions of a ligation auxiliary with the ability of fast, quantitative recovery of glycopeptides provides a valuable instrument for glycopeptide preparation.

Here, a synthesis route for a new generation ligation auxiliary was established. Its altered structure, compared to the original molecule, enables investigation of alternative coupling strategies, with the prospect of expansion of the auxiliary's versatility. Furthermore, a straightforward and short synthesis scheme improves its applicability and allows us to extend its usage in glycopeptide synthesis and other chemoenzymatic modification reactions of peptides.

II Zusammenfassung

Die inhärente chemische Komplexität der Kohlenhydrate in Kombination mit aufwendigen zellulären Synthesewegen erlaubt bis dato lediglich ein eingeschränktes Verständnis der Funktion von Proteinglykosylierungen und bremst dadurch unsere Möglichkeiten das Potential dieser posttranslationalen Modifikationen vollständig zu nutzen. Fortschritte im Bereich der Glykobiologie haben möglicherweise einen direkten Einfluss auf verschiedenste Forschungsgebiete, wie zum Beispiel auf die Arzneimittelentwicklung, Immunologie, Medizin und Zellbiologie. Dies kann jedoch nur durch die Erweiterung des Spektrums an biologischen und chemischen Werkzeugen zur Synthese von Glykoproteinen erreicht werden.

Ein kürzlich entwickeltes PEGyliertes Ligationsauxiliar nützt die Vorteile von Ligationsauxiliaren und fügt durch die Einführung eines Polyethylenglycol (PEG)-Polymers eine weitere Funktionalität zu dessen Grundstruktur hinzu, welche die Regulierung der Löslichkeitseigenschaften und somit die kontrollierte Fällung des Peptid-Auxiliar-Konjugats ermöglicht. Durch diese temporäre Modifikation eines Peptids können aufwändige Reinigungsschritte zwischen wiederholten Runden von enzymatischen Peptidglycosylierungen auf eine einfache Fällung und Zentrifugation reduziert werden. Die Möglichkeit der Kombination von nativer chemischer Ligation und milden Bedingungen zur Entschützung mit der anschließenden schnellen und quantitativen Gewinnung des Glykopeptids stellt ein wertvolles Instrument zur Herstellung von Glykopeptiden dar.

In dieser Arbeit wurde eine Syntheseroute zur Herstellung einer neuen Generation an PEGylierten Ligationsauxiliaren entwickelt. Seine veränderte Struktur, verglichen mit der des ursprünglichen Moleküls, ermöglicht die Erforschung alternativer Kopplungsstrategien, mit der Aussicht auf die Erweiterung der Einsatzfähigkeit des Auxiliars. Zusätzlich verbessert ein vereinfachtes und verkürztes Syntheseschema dessen Anwendbarkeit und ermöglicht die Herstellung von Glycopeptiden und anderen chemoenzymatisch modifizierten Peptiden.

III Table of contents

1 Introduction

1.1 Glycosylation

Current estimations concerning the number of protein encoding genes in the human genome range around 19.000-25.000,[1,2] substantially lower than previously thought.[1] Therefore, other processes that permit the complexity of life are necessarily involved. Looking at the human proteome, one can find that it consists of protein variants and modifications that are not or cannot be encoded in the genome. The estimated number of distinct protein species is still debated and an ongoing topic of research, but could range at around one million.[3] This expansion of protein characteristics can be attributed to alternative splicing, where messenger RNA is altered on a post-transcriptional level, to yield different amino acid (aa) sequences. The second path contributing to the diversification of the proteome are protein modifications after translation, hence, termed posttranslational modifications (PTMs).

Besides non enzymatic modifications,[4] the majority of PTMs are enzyme-catalysed attachments of functional groups to the side-chain of specific amino acids or, to a lesser extent, proteolytic cleavages. As reviewed by Walsh *et al*.,[5] the characteristics by which these PTMs can be sorted are numerous. They differ in regards to the modified amino acid side chain, nature of coupled moiety and their function.

The latter spans from signal transduction through phosphorylations, membrane anchoring via lipidation, to regulation of gene expression by histone modifications.

One of the most abundant PTMs in nature is the enzymatic linkage of a glycan to a protein, named glycosylation. These modified proteins are termed glycoproteins and cover almost the complete spectrum of protein functions. One way to classify protein glycosylations is by their linkage type to the amino acid residue, namely *O*-glycans, *N*-glycans, *C*-glycans and *S*-glycans.

In *O*-glycosylations, a carbohydrate is linked to the hydroxy group of serine or threonine and less usual of tyrosine, hydroxyproline and hydroxylysine.[6] *O*-linkage has been found to be the prevalent type in bacteria.[7] Synthesis occurs mainly in the Golgi apparatus of eukaryotic cells or the cytosol[6] and they are less complex than *N*-glycosylations.

N-glycans are connected to a protein via an amide bond to Asparagine, and known to be the most common glycosylations in eukaryotes[6] and archaea.[7] Initially assembled on a lipid carrier molecule on the luminal side of the endoplasmatic reticulum membrane, the complex glycan structure comprised of fourteen monosaccharides is transferred to the carboxamide side chain of Asn (consensus sequence, X $\neq$ Proline: N-X-S/T, rarely N-X-C, N-G and N-X-V[8]), by a membrane associated protein complex called oligosaccharyltransferase (OST).[5,9] After removal and addition of further sugars, the chaperones calreticulin and calnexin can recognize the glycan structure, facilitate proper folding of the protein, and flag it for degradation if not successful. Subsequent steps of trimming and elongation in the Golgi, with its numerous

glycosyltransferases and glycosidases, lead to the vast diversity of the *N*-glycans.[10]

Substantially less common is the mannosylation of tryptophan via formation of a C-C bond to the C2 carbon of the indole heterocycle (C-glycosylation). The questions regarding its exact biosynthesis and biological function remain yet to be answered.[11,12]

The fact that new and unusual variants of glycosidic linkage are still discovered, e.g. S-glycosylation of cysteine,[13–15] is exemplary for the important role that glycoproteins play in biochemistry and for the potential that still lies hidden.

The scientific research and knowledge around glycoproteins does not match its biological significance at the moment, especially in eukaryotic cells, where it is one of the most abundant PTMs. Reasons for this deficiency can be found in the nature of carbohydrate chemistry, biosynthesis and processing and challenges in the analysis of these structures.[16]

Amongst numerous other tasks, they are involved in the folding control, trafficking[17] and physical stability of proteins and cell membranes or act as ligands for disease causing bacteria or viruses.[18] While mutations in genes involved in the glycan/glycoprotein biosynthesis and processing can lead to severely affected phenotypes,[19] their discovery is often hindered by insufficient diagnostic methods.[20]

1.2 Glycopeptide/-protein synthesis

The combination of the aforementioned chemical properties of carbohydrates and the heterogenous processing lead to a number of different glycoforms present at a specific time in a cell. In contrast to other biopolymers, the synthesis of glycans is only dependent on expression and activity/specificity of the enzymes involved, their substrate availability and the spatial separation of those two. Moreover, the glycosylation pattern of a cell is also cell type dependent[21,22] and may vary during development/aging[20,23]and disease.[19,21,24] In order to investigate which role these different variants play in cells and organisms, homogenously glycosylated glycoproteins are needed.

In comparison to DNA and non-modified proteins, the use of the molecular machinery for the synthesis of complex glycopeptides cannot be exploited in a comprehensive manner, as its structural information is not stored in form of DNA/RNA. However, for certain applications where large amounts of glycoproteins are needed (e.g. therapeutic proteins), genetically engineered and relatively inexpensive production systems have been developed. With specifically designed prokaryotic, as well as eukaryotic expression systems, such as yeast, complex homogeneously glycosylated proteins are accessible.[22]

Although purification via HPLC from natural sources is possible, it is certainly laborious and inefficient for large scale preparation as it cannot eliminate micro-heterogeneity, therefore,

synthesis/semisynthesis is a viable option.

To achieve this, various approaches can be used, each presenting certain drawbacks and advantages. Here, only a brief overview of the available chemical and chemoenzymatic strategies will be outlined.

During convergent synthesis, a peptide is assembled via solid phase peptide synthesis (SPPS) and, before complete deprotection, chemoselectively glycosylated to form the natural glycan-peptide linkage. This approach is restricted to *N*-linked glycopeptides.[25,26] Additionally, a number of methods are available that enable the efficiently introduction of an unnatural glycan linkage, even under aqueous conditions. This includes various carbohydrate derivatives to yield S-linked glycans or Cu(I) catalysed azide-alkine cyloaddition to yield 1,4-disubstituted triazoles.[25]

Linear assembly uses preformed glycosylated amino acid building blocks during SPPS. Problems can arise during peptide elongation, as bigger glycans can hinder coupling after a certain peptide length.[25,27]

A variation of the latter is the use of monoglycosylated amino acid building blocks during SPPS and further enzymatic elongation of the glycan in solution (or, less efficiently, on resin). Advantages of this approach are: 1) coupling efficiency during SPPS is only minimally influenced by the glycan, 2) no solubility problems as possible with fully acetylated glycans, 3) various glycosyltransferases can be utilized to avoid chemical assembly of glycans and, with that, complex protection schemes.

These three strategies all share one major disadvantage: they are based on SPPS, which is restricted to peptides ranging from 40-50 aa (amino acids) or less. This limitation can be overcome by using native chemical ligation (NCL)[28]: full length glycoproteins can be

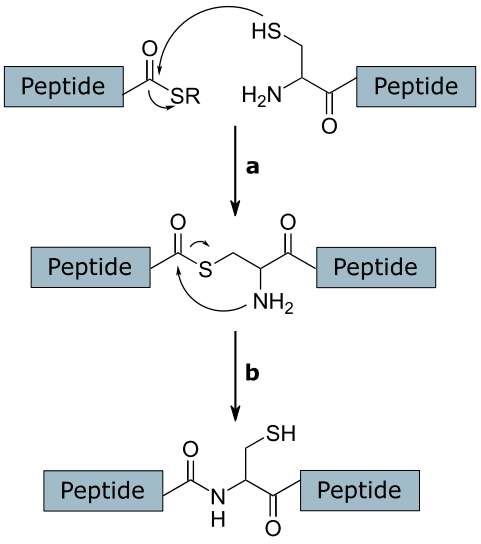


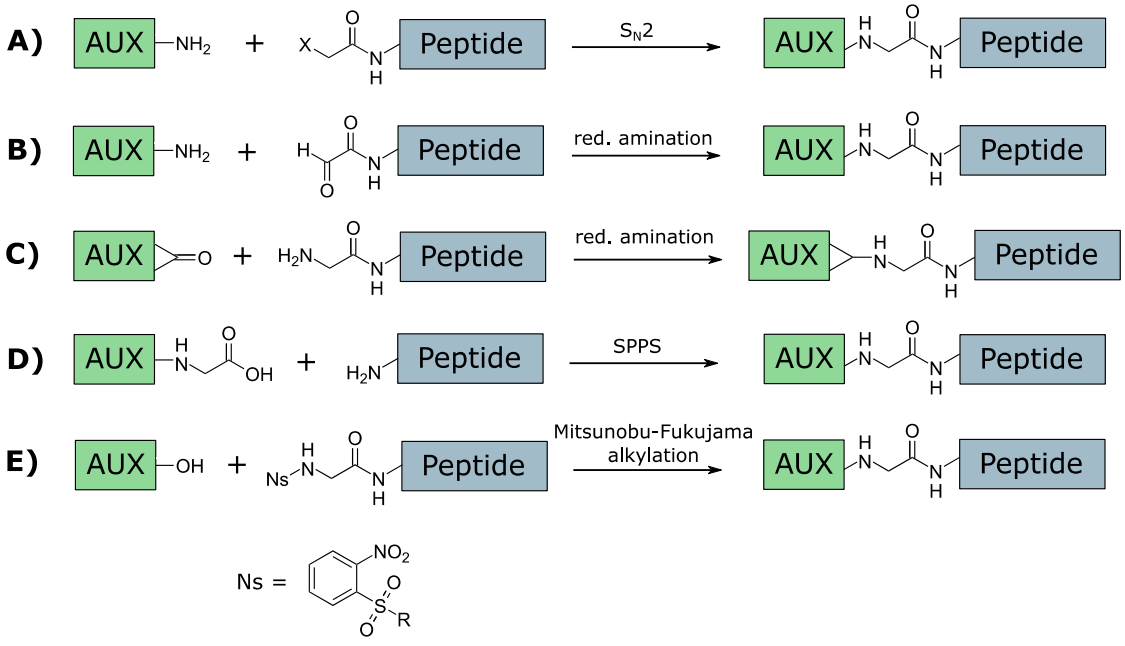
Figure 1 Reaction scheme native chemical ligation; a) Transthioesterification: C-terminal thioester undergoes nucleophilic attack by the N-terminal cysteine side chain and forms a reversible thioester intermediate; b) S→N acyl shift: irreversible rearrangement to a native peptide bond

assembled from peptide fragments, either from chemical sources or recombinantly expressed.

Besides a C-terminal thioester on one peptide, NCL requires a N-terminal cysteine on the second partner of the ligation (**Figure 1**). This latter prerequisite is often hard to meet, as cysteines are only found at a low frequency in natural proteins or are present at positions unfavourable for ligation. One way to overcome this limitation is mutation of the natural amino acid to cysteine. However, the exchange may interfere with the function of the protein and could lead to misfolding, caused by incorrect disulfide bridges.

Another method uses desulfurization of cysteine and other side chain thiolated amino acid derivatives, coupled during SPPS to restore a natural amino acid functionality. It has been shown that in combination with radical based desulfurization half of the proteogenic amino acids are accessible in this manner.[29,30] Synthesis of these building blocks is work intensive and therefore restricts its utilization.[31] For peptides containing multiple cysteine, selective desulfurisation is necessary. This can be accomplished by the protection of cystein side chains with an Acm (acetamidomethyl) protecting group. [32]

In addition to the temporary thiol functionality introduced on amino acid side chains, the peptide N-terminus can be modified with a cleavable moiety, termed auxiliary (N^{α}-auxiliary). Attachment is generally carried out via reductive amination, [31,33–36] S_N2 reaction of a modified peptide N-terminus[37–39] or by coupling of an auxiliary-conjugated amino acid during SPPS[40,41] (**Scheme 1**). The auxiliary mimics cysteine and therefore enables NCL, and its cleavage has to be compatible with SPPS. In general, two different conditions are used for removal after NCL: acid (or, less common, base[33]) mediated cleavage and photocleavage. Former methods are typically based on an ethanethiol/benzylamine scaffold, and acid lability is tuned through variation of the substituents on the aromatic ring, [34–36,42,43] with some exceptions. [31,33]



Scheme 1 **Substrates and reactions used for coupling ligation auxiliaries to peptides.**

Photocleavable auxiliaries[38,44–46] possess similar structures, but circumvent harsh removal condition, that might interfere with other modifications such as glycosylations, by using UV-light. Basis for this reaction is an o-nitrobenzyl scaffold (**Figure 2**).

A common problem that affects all auxiliaries is ligation efficiency. The influence of steric hindrance of the auxiliary at the junction during ligation is the smallest when both the N-terminal and C-terminal amino acid are glycine. A change of both or even only one of the amino acids at the junction dramatically influences the ligation, and poses a major limitation to the applicability of ligation auxiliaries.[33,37,41,42] Advances to overcome this constraint are dependent on multiple factors, such as the specific peptide sequence or its solubility. Optimization of ligation conditions may be key to increase ligation yields and therefore access junctions sterically more hindered.

1.3 PEGylated auxiliary

The advantages of photocleavable ligation auxiliaries can be exploited even further by the introduction of an additional functional group linked to its moiety. A polyethylene glycol (PEG) polymer permits precipitation of the auxiliary-peptide conjugate.[47] Attaching it to the auxiliary enables its clean removal at the end of the synthesis by UV radiation (**Figure 2**). The PEGylated auxiliary was designed to avoid a number of problems that otherwise occur during glycopeptide synthesis. Glycosylation of a peptide on resin can result in a non-quantitative reaction and therefore lead to a heterogenous mixture of glycoforms. Other PEGylation strategies need a protease recognition site to remove the polymer, possibly influencing protein function.[47] The temporary modification, based on the auxiliary's platform, enables quantitative separation of peptide and reaction mixture between glycosylation steps, allowing for straightforward ligation and easy cleavage.[38]

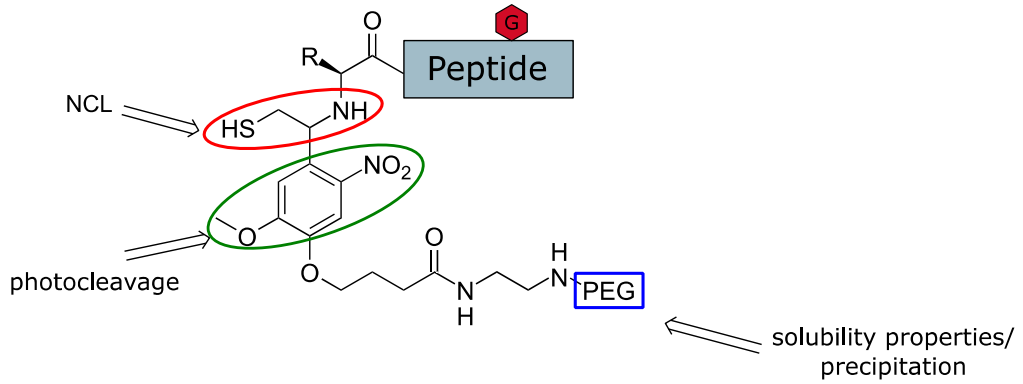
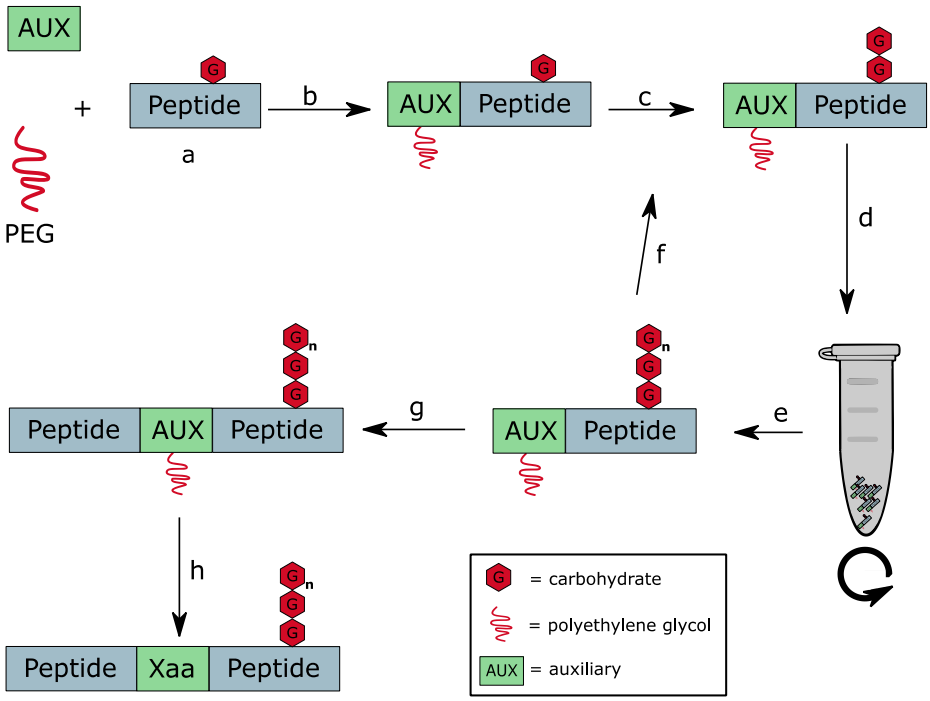


Figure 2 PEGylated auxiliary coupled to a peptide N-terminus (see also **Scheme 2**); functional moieties that introduce properties are highlighted; R = aa side chain

This separation step is crucial, as after enzymatic elongation of the glycan structure, unreacted activated sugar and glycosyltransferase have to be separated from the glycopeptide, to ensure

that the following elongation step is possible without side reactions. Exactly this purification (normally via HPLC) can be the cause of substantial losses in yield, [48] which is avoided by the PEG enabled precipitation. In **Scheme 2** the general steps for glycopeptide synthesis, using the auxiliary, are depicted.

The auxiliary is coupled to a monoglycosylated peptide, syntesized by SPPS (**Scheme 2a**). While still on resin, this auxiliary is modified with a polyethyleneglycol polymer (**b**). After removal from the solid support, the glycan is extended in solution with the help of an appropriate glycosyltransferase (**c**). The obtained glycopeptide is then precipitated taking advantage of the properties of PEG (**d**). This makes easy purification possible (**e**) and enables further clean glycan elongation (**f**). The fully glycosylated peptide can now be completed to its full size by NCL or ECL (**g**), expoiting the mercaptoethylene amino group of the auxiliary. At last, UV cleavage detaches the auxiliary, hence, also the PEG polymer, to leave behind nothing but a natural amino acid. In theory, the nature of this amino acid is dependent on the educts used in the reaction. Due to the auxiliary coupling reaction employed and steric reasons during ligation, until now only glycine was used with success. This approach was succesfully employed in the synthesis of variants of the epithelial tumor marker MUC1, carrying a mono-, di-, or trisaccharide.[38]



Scheme 2 General workflow of the PEGylated auxiliary assisted glycopeptide preparation; peptide obtained by SPPS, monoglycosylated; b) coupling of the auxiliary to peptide N-terminus and PEG attachment to auxiliary; c) enzymatic glycan elongation; d) precipitation and centrifugation; e) separation from supernatant; f) further elongation using a different enzyme; g) native chemical ligation; h) UV cleavage;

2 Objective

A new synthesis route will be established for a new generation PEGylated auxiliary for glycopeptide synthesis and subsequent ligation, which differs from the initially reported one by Bello *et al*.[38] in the substituent of its benzylic position. In this case, the amine (**Figure 3A**) is substituted by a hydroxy group (**Figure 3B**). In addition, the newly established synthesis should be more concise than the previous scheme.

Auxiliaries are generally introduced by reductive amination, S_N2 reaction or as aa-auxiliary conjugate building blocks to the N-terminus of a peptide. Each of these methods possesses individual limitations, like the lack of chemical compatibility with a glycosidic linkage, difficulties in the preparation of N-terminally halogenated amino acids or the necessity of modification of the auxiliary with an amino acid for each application. However, this work tries to test the possibility and practicability of the newly generated auxiliary for coupling via alternative reactions, enabled by the introduced hydroxy moiety. One of these is the Mitsunobu-Fukuyama alkylation, whose conditions are mild and compatible with standard Fmoc peptide synthesis. In combination with the general strategy of a PEGylated photocleavable auxiliary, this revisited approach should enable further expansion of its application.

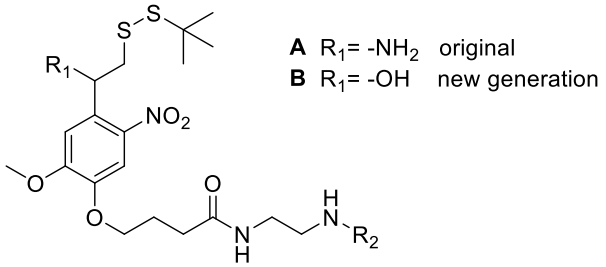


Figure 3 Photocleavable auxiliary **A**) original by Bello *et al.* **B**) synthesis aim of this work; R_2- site of later PEG attachment

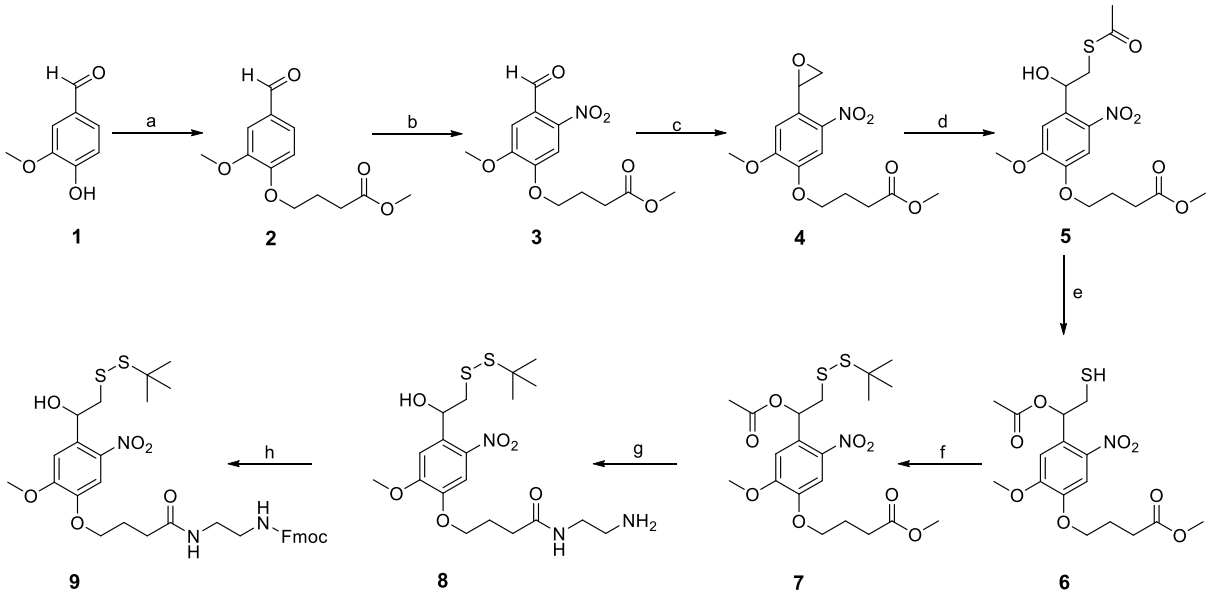
3 Results

3.1 Auxiliary synthesis

The synthesis pathway of the new generation auxiliary shares the starting material (vanillin) and the first two steps with the previously reported work.[38]

The question to answer at this stage is: What is the most feasible course of action to introduce an additional carbon at the aldehyde moiety, which could allow for further derivatisation to obtain the desired beta functionalized thiol motif? The solution by Bello *et. al.*[38] was the introduction of a methylene group at the aldehyde moiety via Wittig reaction, subsequent dihydroxylation and activation of the benzylic position as an electrophile by conversion of the vicinal diol into a cyclic sulfite ester.

In this work, rather than making the above-mentioned detour to form a C-C bond and create a suitable electrophile, the Corey-Chaykovsky reaction was employed. The key reaction of this synthesis directly yields an epoxide, which compared to the starting material, contains the needed extra methylene group. Using an appropriate nucleophile and the right reaction conditions the epoxides ring can be opened regioselectively.



Scheme 3 Synthesis of the new generation ligation auxiliary; a) K_2CO_3, phase transfer catalyst, Methyl 4-chlorobutyrate, ACN, 91%; b) $AcOH/HNO_3$, 70%; c) NaH, $(CH_3)S(I)O$, DMSO/THF, 69%; d) Pyridine, AcSH, 74%; e) SiO_2, petrol ether/ethyl acetate 1:1 (v/v), 92%; f) Activated tBuSH **16**, TEA, DMF/MeOH, 42%; g) Ethylenediamine, NaOMe cat., Toluene, 80%; h) DIEA, FmocCl, DCM, 90%;

3.1.1 Synthesis of Methyl 4-(4-formyl-2-methoxyphenoxy)butanoate 2

The synthesis (**Scheme 3**) starts with commercially available vanillin **1**, which was converted into its methyl butanoate ether in the first step, using the Williamson ether synthesis[38] (**Figure 4**). Here, a phenolate obtained by deprotonation of vanillin with potassium carbonate attacks methyl 4-chlorobutyrate in a nucleophilic substitution reaction. To ensure both proper transfer of potassium carbonate to the organic phase and catalysis of the substitution reaction, tetrabutylammonium iodide was added. The majority of product **2** was recrystallized from methanol. For complete recovery, the remaining crude mixture was subjected to flash column chromatography to give a combined yield of 91%.

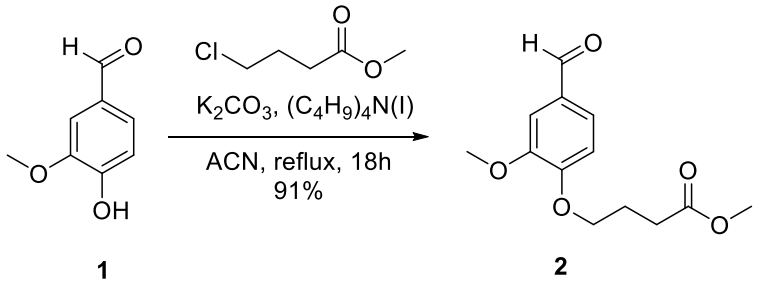


Figure 4 Williamson ether synthesis

3.1.2 Synthesis of Methyl 4-(4-formyl-2-methoxy-5-nitrophenoxy) butanoate 3

Nitration of compound **2** (**Figure 5**) was accomplished with fuming nitric acid and glacial acetic acid.[38] The highly electrophilic nitronium cation (NO_2^+), formed in situ, enables electrophilic aromatic substitution. Substituent effects in combination with steric factors resulted in the preferred formation of the C6 substituted product **3**. Analysis, however, showed an additional side product, presumably formed by ipso (C1) substitution of the aldehyde moiety (**Figure 6**, see discussion). Flash column chromatography proved to be challenging and produced a yield of 70%.

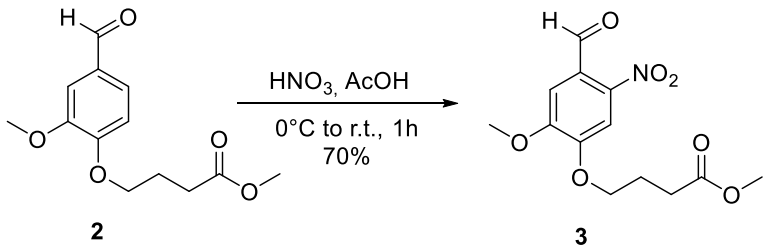


Figure 5 Aromatic nitration

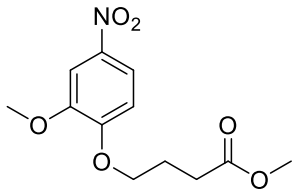


Figure 6 Ipso substituted side product

3.1.3 Synthesis of Methyl 4-(2-methoxy-5-nitro-4-(oxiran-2-yl)phenoxy)butanoate 4

Conversion of aldehyde **3** into epoxide **4** was achieved by using the Corey-Chaykovsky reaction (**Figure 7**).[49–51] During this step, a sulfur ylide is created *in situ* by reaction of trimethylsulfoxonium iodide with sodium hydride (**Figure 8**).

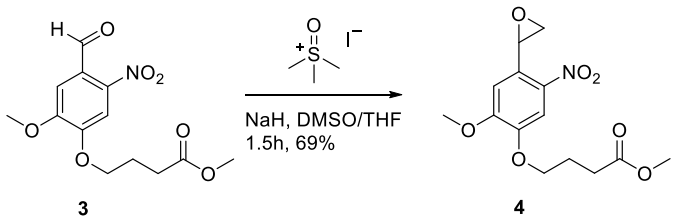


Figure 7 Corey-Chaykovsky epoxidation

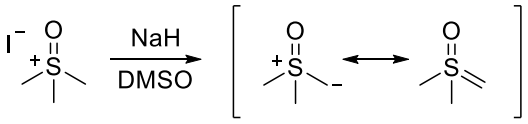


Figure 8 Formation of the sulfur ylide

With the addition of the aldehyde (**Figure 9**), the nucleophilic ylide adds to the carbonyl carbon. The oxygen anion of the newly formed betaine motif acts now as a nucleophile and is able to eliminate dimethylsulfoxide. Thus, in the course of one reaction, a C-C bond in the form of a methylene transfer is created together with a highly reactive epoxide. The use of a mixture of DMSO and THF as solvent allowed stirring of the reaction solution at temperatures below the melting point of DMSO. After column chromatography, the product was obtained in 69% yield.

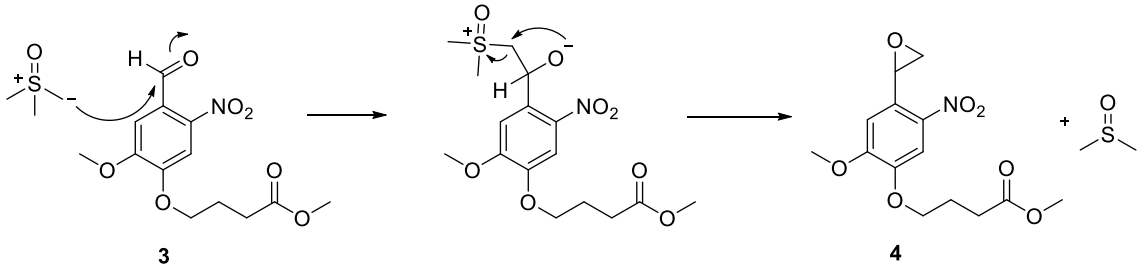


Figure 9 Nucleophilic attack of the ylide and elimination of dimethylsulfoxide through formation of the epoxide

3.1.4 Synthesis of Methyl 4-(4-(2-(acetylthio)-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate 5

To obtain the desired beta hydroxy mercaptan motif, epoxide **4** was opened with thioacetic acid in neat pyridine (**Figure 10**). Under basic conditions, complete regioselective opening of the epoxide at the less hindered primary carbon is achieved. Reaction monitoring via TLC showed the formation of a small amount of side product, which was later identified as the O acetylated species (**6**). To a smaller extent during the reaction and to a greater during column chromatography, the acetyl group migrated from the sulphur to the oxygen, to form the more stable ester. After purification via column chromatography, product **5** with a content of about 15% of **6** could be obtained in a combined yield of 74%. Because the following synthesis step consisted in migration of the acetyl group, no further effort was undertaken to separate the two isomers.

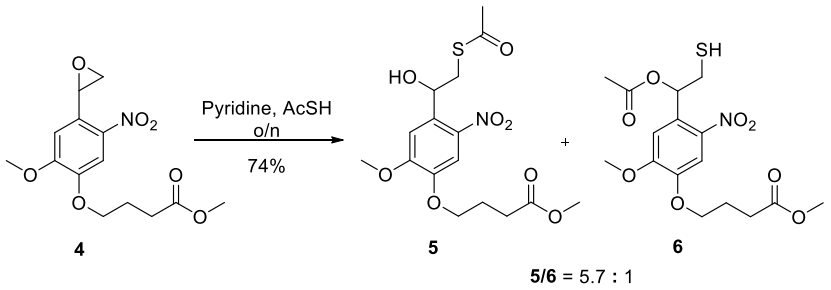
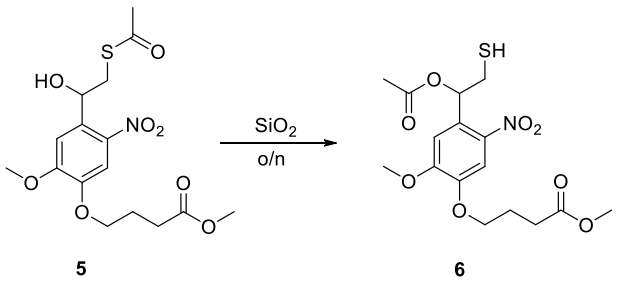


Figure 10 Opening of the epoxide with thioacetic acid in pyridine

3.1.5 Synthesis of Methyl 4-(4-(1-acetoxy-2-mercaptoethyl)-2-methoxy-5-nitrophenoxy)butanoate 6

To fully convert thioester **4** into ester **5**, migration of the acetyl group was promoted via the application of the mixture onto silica gel (**Figure 11**). As reported by Gao *et al.*,[52] silica gel catalyses this reaction under mild conditions. The proposed mechanism involves hydrogen bonding of surface hydroxy groups of the silica gel with the compounds hydroxy group. This polarizes the hydrogen oxygen bond and enables the nucleophilic attack of the alcohols oxygen on the carbonyl carbon of the acetyl group. At last, subsequent elimination gives the free thiol. Due to the higher stability of the ester bond compared to the thioester, this process is irreversible under the reaction conditions used.

To do so, compound **4** was mixed with the appropriate amount of silica gel and enough solvent (petroleum ether/ethyl acetate 1:1) to make proper stirring possible. After reaction overnight, the slurry was transferred onto a chromatography column and the product eluted with the previous solvent mixture, to afford the ester in 92% yield.

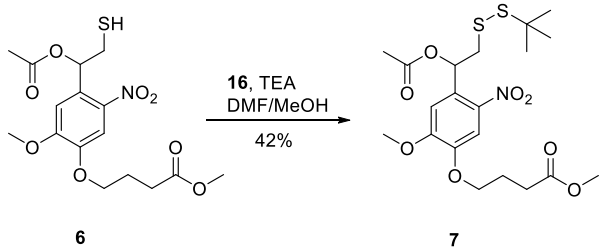


5 **6**

Figure 11 Silica gel catalysed S→O acetyl migration

3.1.6 Synthesis of Methyl 4-(4-(1-acetoxy-2-(tert-butyldisulfaneyl)ethyl)-2-methoxy-5-nitrophenoxy)butanoate 7

To prevent the thiol group to undergo unwanted side reactions e.g. homo-disulfide formation, a *tert*-butyl sulfanyl is introduced as a protecting group (**Figure 12**). This makes possible orthogonal protection and simple deprotection of the thiol in aqueous buffers, which is crucial for the auxiliaries use during native chemical ligation.



6 **7**

Figure 12 Disulfide protection of the thiol

Tert-butyl thiol (2-methylpropane-2-thiol) in its solid and odourless activated form as 1-(tert-Butylthio)-1,2-hydrazinecarboxmorpholide[38,53] (**16**) was used for this purpose. Reaction presumably proceeds in a S_N2 reaction, where first the thiol on the auxiliary is deprotonated by the base triethylamine. The nucleophilic thiolate is now able to attack the electrophilic sulfur and release an azodicarboxylate anion leaving group, which completes the process by deprotonating the triethylamine (**Figure 13**).

Analysis of the crude reaction mixture via TLC and NMR revealed the formation of a side product that was identified as the disulfide homodimer of starting material **6**. Additionally, small amounts of a styrene derivative, presumably produced by elimination reaction, could be detected. Column chromatography gave product **7** in a yield of 42%.

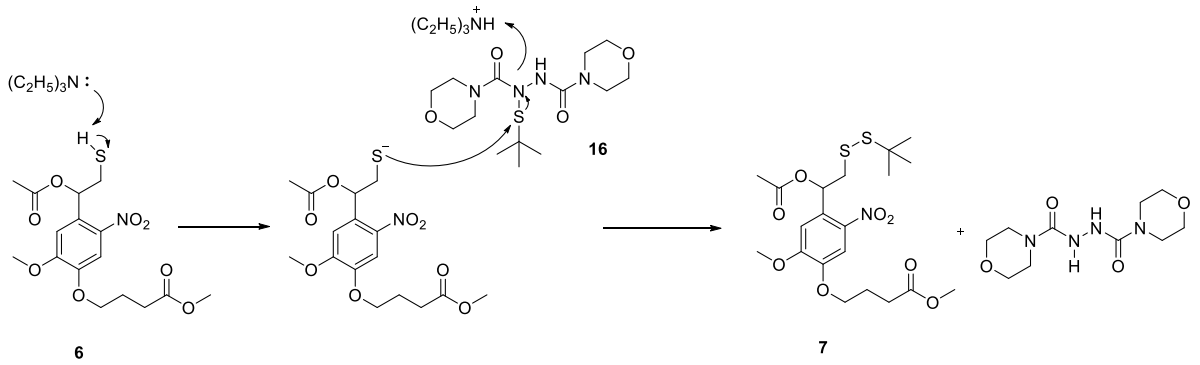
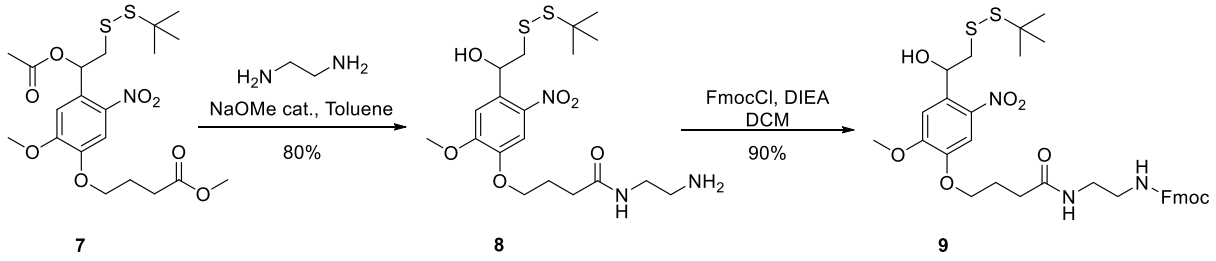


Figure 13 Possible reaction mechanism of the thiol protection with an activated tertiary butyl thiol (**16**)

3.1.7 Synthesis of (9H-fluoren -9-yl)methyl (2-(4-(4-(2-(tert-butyldisulfaneyl)-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy) butanamido)ethyl)carbamate 9

Compound **8** was synthesized via sodium methoxide catalysed amidation of ester **7** (**Figure 14**). In the procedure, slightly modified from Ohshima *et al.*[54], the starting material **7** is dissolved in toluene and an excess of ethylenediamine. Addition of a catalytic amount of sodium methoxide and heating to 50°C over night yielded 80% of **8** after column chromatography.

Fmoc (fluorenylmethoxycarbonyl) protection of amine **8** was straightforward and realized using Fmoc chloride in DCM with diisopropylethylamine. After column chromatography, a yield of 90% could be obtained.



7 **8** **9**

Figure 14 Ester-Amide exchange under sodium methoxide catalysis and Fmoc protection of the thereby created amine

3.2 Coupling of the auxiliary to the peptide

In solution, a number of reactions are available to convert primary into secondary amines. However, to modify a peptide N-terminus, the employed conditions of such a reaction have to be compatible with standard protecting groups and resins used in SPPS. One reaction that fulfils these requirements is the Fukuyama-Mitsunobu alkylation (**Figure 15**).[55] It has been shown that alkylations beyond a simple methylation[56] are possible in good yields,[57–59] though in these examples only primary alcohols were used.

In contrast to nucleophiles used during standard Mitsunobu reaction, e.g. carboxylic acids, amines are not a suitable substrate and have to be converted into nitrobenzenesulfonamide (nosyl amide). This necessity is explained by the fact that the betaine structure, formed during the reaction of triphenylphosphine with diisopropylazodicarboxylate, has to be protonated by the pronucleophile employed, which requires a pka of <11.[60] But the protection, in this case of the peptide N-terminus, serves two purposes. First, as mentioned, it acts as an activating group by increasing the acidity of the amide proton to a value that allows alkylation under Mitsunobu conditions. Secondly, it avoids further alkylation of the desired secondary to a tertiary amine. Furthermore, the conditions under which the nosyl group is cleaved, via the formation of a Meisenheimer complex, are relatively mild. [61]

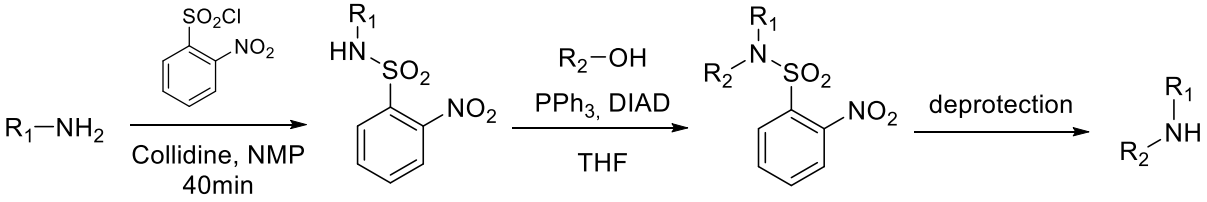
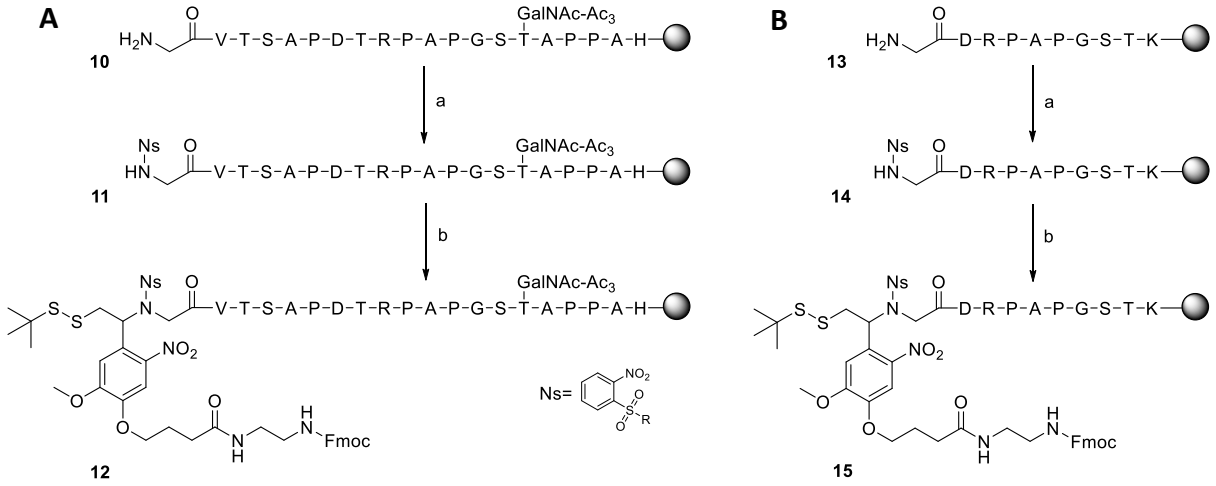


Figure 15 General reaction scheme Fukuyama-Mitsunobu alkylation; R1: peptide N-terminus, R2: Auxiliary

In order to find the appropriate conditions, the reaction was investigated using two different test peptides (**Scheme 4**). Peptides **10** and **13** (**Figure 16**) still bound to TentaGel resin were available from previous experiments, and differ in their length (20 vs. 10 aa) and in a peracetylated N-acetyl galactosamine residue. In both cases the N-terminal amino acid is glycine. First, a test cleavage of the peptides from the resin, still Fmoc protected on their N-termini, and a subsequent analysis via LC-MS were performed to evaluate the purity of the starting material (**Figure 17** and **Figure 18**).

Peptide **10**	Peptide **13**
H-GVTSA PDTRP APGST APPAH-**OH** GalNAc	**H**-GDRPA PGSTK-**OH**

Figure 16 Amino acid sequence of peptide **10** and peptide **13**



Scheme 4 Auxiliary coupling under Mitsunobu-Fukuyama conditions to **A**) peptide **10** and **B**) peptide **13** on resin; a) NsCl, Collidine, NMP; b) Auxiliary **9**, DIAD, PPh_3,THF;

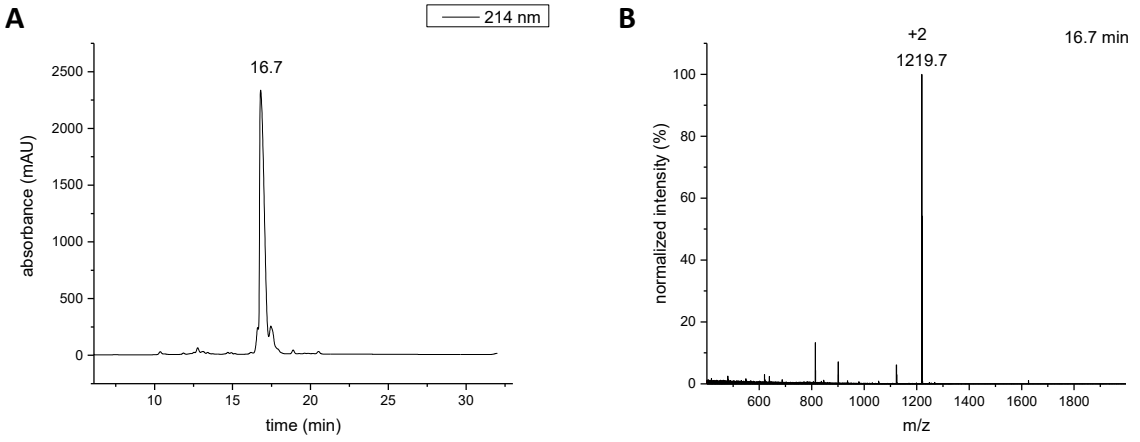


Figure 17 LC/MS analysis of peptide **10** +Fmoc after cleavage from resin; **A**) HPLC chromatogram of peptide **10** at 214nm; **B**) mass spectrum at 16.7min, calculated mass $C_{109}H_{156}N_{26}O_{38}$: 2438.6 Da, observed: 1219.7 $[M+2H]^{+2}$

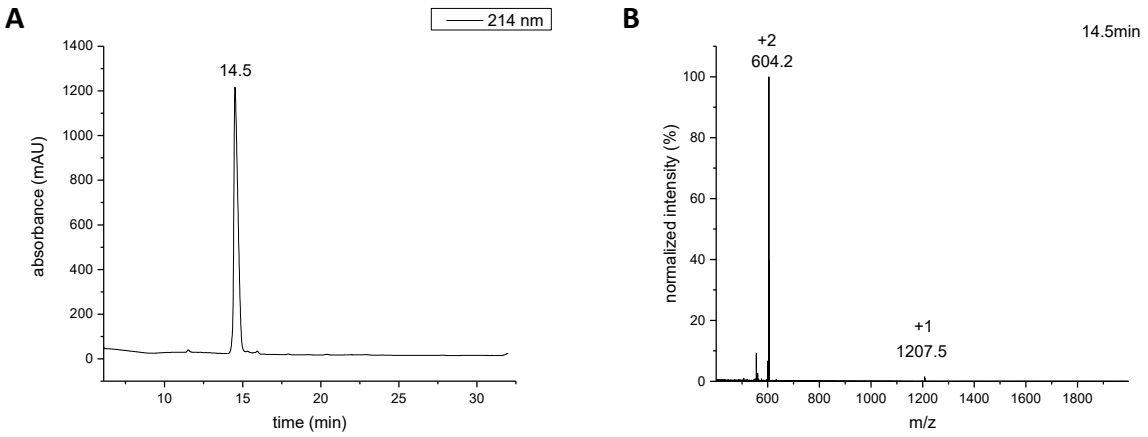


Figure 18 LC/MS analysis of peptide **13** +Fmoc after cleavage from resin; **A**) HPLC chromatogram of peptide **13** at 214nm; **B**) mass spectrum at 14.5min, calculated mass $C_{55}H_{78}N_{14}O_{17}$: 1206.6 Da, observed: 1207.5 $[M+H]^+$, 604.2 $[M+2H]^{2+}$

N-terminal protection of the glycine residue from peptides **10** and **13**, still on resin, as 2-nitrobenzenesulfonamides was achieved under the following conditions. The Fmoc deprotected peptides **10** and **13** on resin were treated with 2-nitrobenzenesulfonyl chloride and collidine in NMP. After 40min, LC/MS analysis shows complete protection (**Figure 20** and **Figure 21**). Coupling of auxiliary **9** was possible with standard Mitsunobu reagents, triphenylphosphine (PPh_3) and diisopropylazodicarboxylate (DIAD) in anhydrous THF. Initially, PPh_3 reacts with DIAD and forms a phosphonium intermediate. Simultaneously, the forming hydrazide anion (betaine structure) has to be protonated by the employed nucleophile, in this case the sulfonamide. This protonation step requires a sufficient acidity of the nucleophile and is the reason why unmodified amines cannot be used.

The alcohol of the auxiliary binds to the phosphonium intermediate in a substitution elimination reaction, and is now activated as a leaving group. At last, substitution of the deprotonated sulfonamide liberates triphenylphosphine oxide and gives the desired product (**Figure 19**).

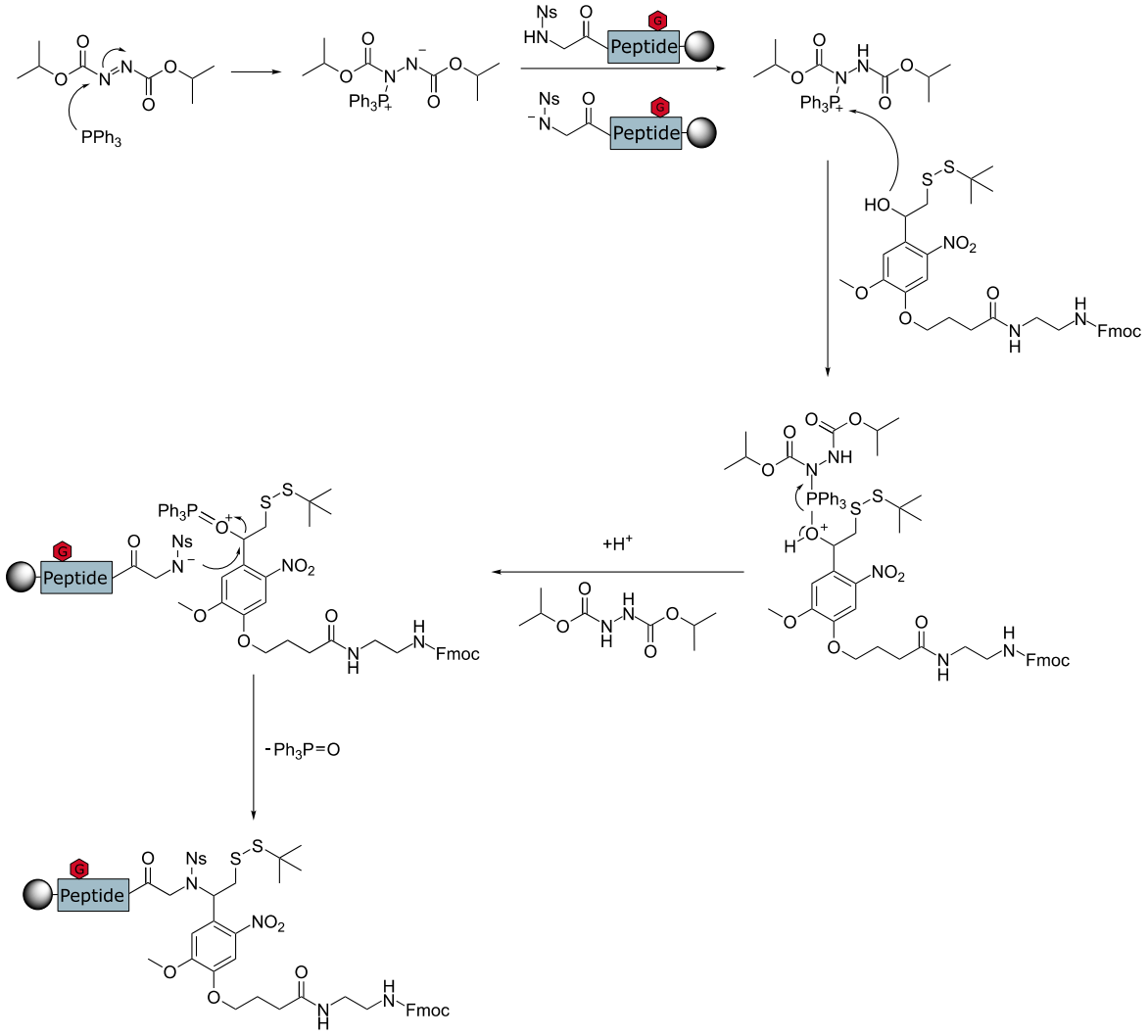


Figure 19 Detailed reaction mechanism of the Mitsunobu-Fukuyama alkylation

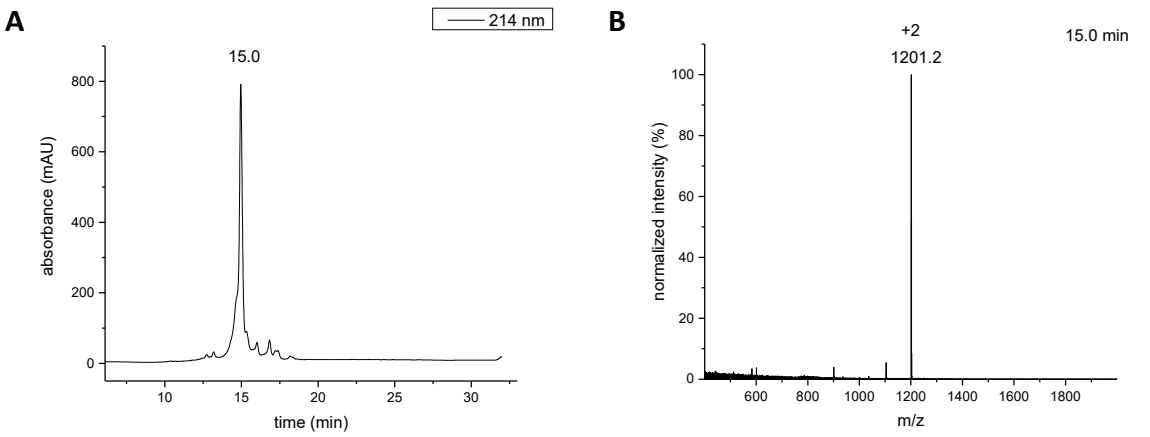


Figure 20 LC/MS analysis of peptide **11** after cleavage from the resin; **A**) HPLC chromatogram of peptide **11** at 214 nm; **B**) mass spectrum at 15.0 min, calculated mass $C_{100}H_{149}N_{27}O_{40}S$: 2401.5 Da, observed 1201.2 $[M+2H]^{2+}$

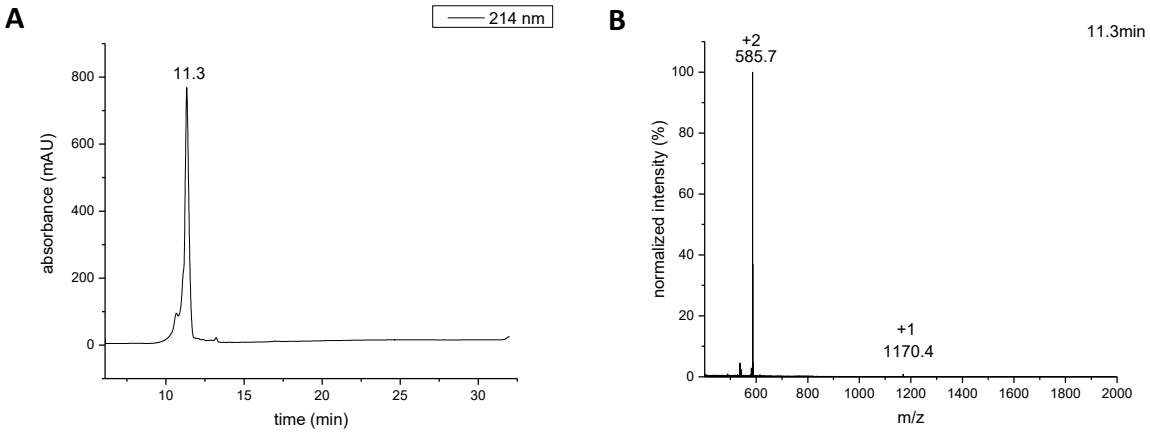


Figure 21 LC/MS analysis of peptide **14** after cleavage from the resin; **A**) HPLC chromatogram of peptide **14** at 214 nm; **B**) mass spectrum at 11.3 min, calculated mass $C_{46}H_{71}N_{15}O_{19}S$: 1169.48 Da, observed 1170.4 $[M+H]^{+}$, 585.7 $[M+2H]^{2+}$

Alkylation of peptide **11** led to a mixture of products (**Figure 22**), amongst them the peptide-auxiliary conjugate **12** (**Figure 22E**, calculated mass: 3066.24 Da, observed: 1534.89 $[M+2H]^{+2}$). Also still detectable is the starting material (**Figure 22B**, calculated mass: 2401.02 Da, observed: 1201 $[M+2H]^{+2}$) and two side products with a difference of -88 Da (**Figure 22C**, observed mass: 1490.26 $[M+2H]^{+2}$) and -16 Da (**Figure 22D**, observed mass: 1526.68 $[M+2H]^{+2}$, 1017.67 $[M+3H]^{+3}$).

The relatively high background signal in the mass spectra seen in **Figure 22B-E** is presumably caused by a too low concentration/injection volume of the sample or/and a structure-related reduced efficiency in ionisation.

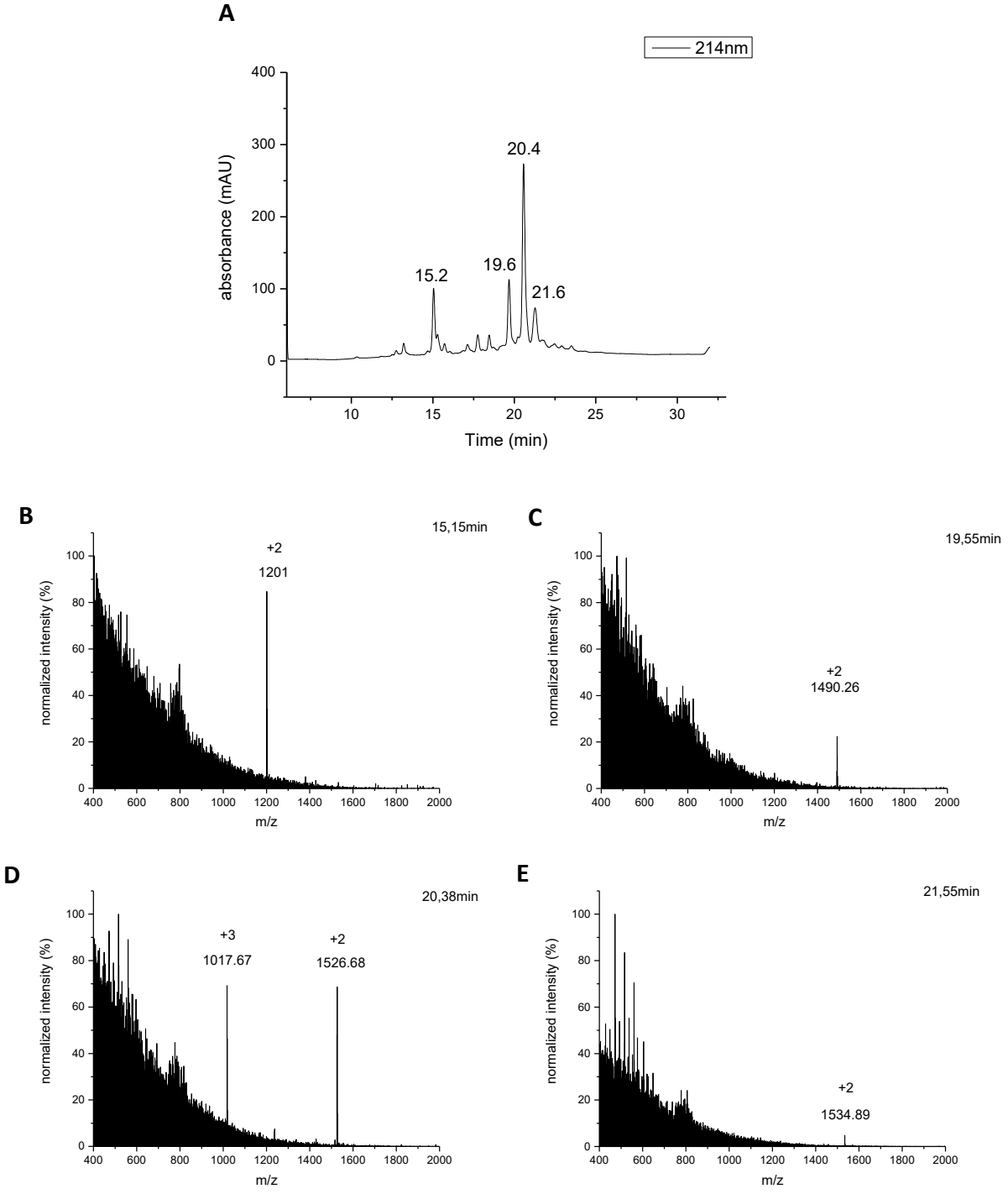


Figure 22 LC/MS analysis of Mitsunobu reaction of **11** after cleavage from the resin (**Scheme 4**); **A**) HPLC chromatogram at 214 nm; **B**) mass spectrum at 15.15 min, peptide **11**, calculated mass: 2401.02 Da, observed: 1201 $[M+2H]^{+2}$ **C**) mass spectrum at 19.56min, observed: 1490.26 $[M+2H]^{+2}$; **D**) mass spectrum at 20.4min, observed: 1526.68 $[M+2H]^{+2}$, 1017.67 $[M+3H]^{+3}$; **E**) mass spectrum at 21.6min, peptide **12**, calculated mass: 3066.24 Da, observed: 1534.89 $[M+2H]^{+2}$;

For peptide **13**, analysis shows, besides a side product and starting material, the expected signal (**Figure 23**, calculated mass: 1169.48 Da, observed: 918.55 $[M+2H]^{+2}$). This side product has a mass +72 Da compared to product **15** (**Figure 23D**, observed mass: 954.62 $[M+2H]^{+2}$).

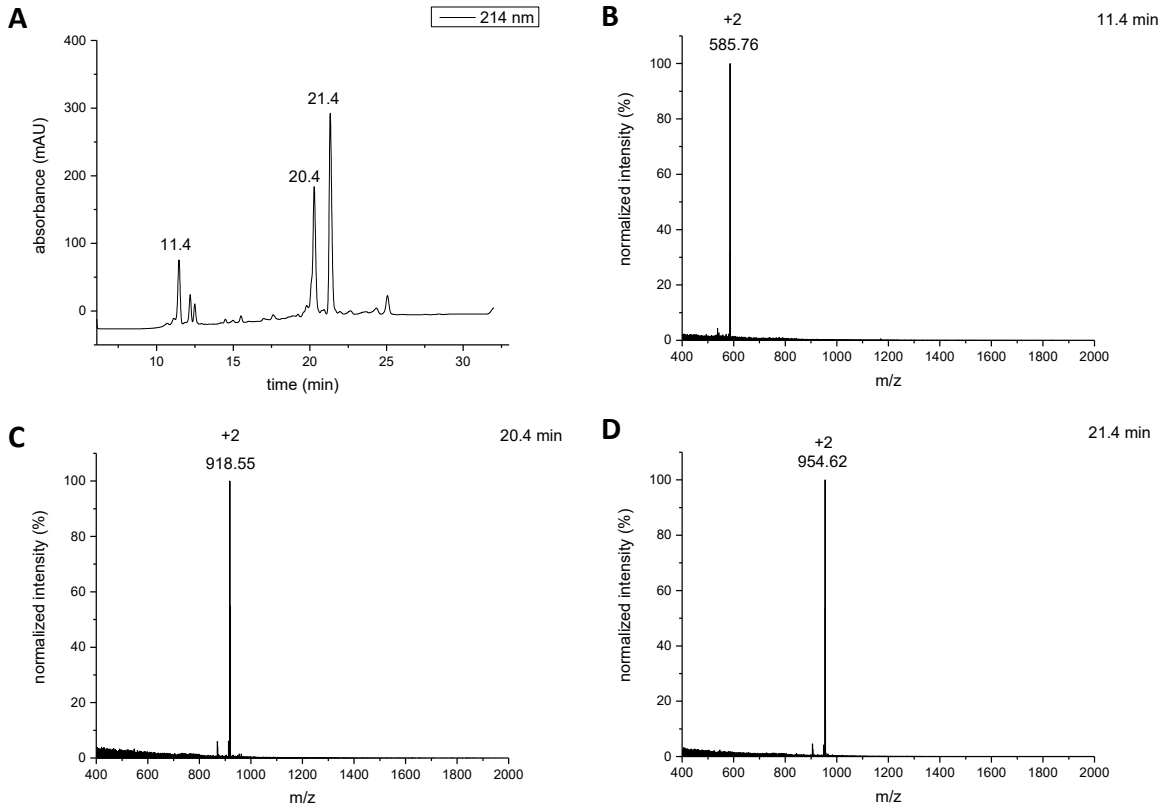


Figure 23 LC/MS analysis of Mitsunobu reaction of **14** after cleavage from the resin (**Scheme 4**); **A**) HPLC chromatogram at 214 nm; **B**) mass spectrum at 11.4 min, peptide **14**, calculated mass: 1169.48 Da, observed: 585.76 $[M+2H]^{+2}$; **C**) mass spectrum at 20.4 min, peptide **15**, calculated mass: 1834.7 Da, observed: 918.55 $[M+2H]^{+2}$; **D**) mass spectrum at 21.4 min, side product, observed mass: 954.62 $[M+2H]^{+2}$

4 Discussion

4.1 Auxiliary synthesis

In order to broaden the applicability of the PEGylated auxiliary,[38] a new synthetic route was established to yield a modified new generation PEGylated ligation auxiliary for glycopeptide synthesis. This enables the coupling of a peptide and the auxiliary via reactions not possible before and allows testing of different amino acids at the junction after cleavage, other than glycine.

Purification of the crude reaction mixture from the Williamson ether synthesis containing product **2** was done via recrystallization from methanol. It was observed that the boiling point of methanol is above the melting temperature of **2**, which led to the formation of an impure oil, rather than crystals. Therefore, lower temperatures and a higher volume of methanol had to be used. To then completely extract the product from the mother liquor after recrystallization, the remaining concentrated crude was subjected to a flash column chromatography.

In addition to **3**, aromatic nitration of **2** also gives a side product. Supported by NMR data (not shown), it is assumed that this is methyl 4-(2-methoxy-4-nitrophenoxy)butanoate, formed by ipso substitution of the substrates formyl moiety. The mixture of these two proved to be rather difficult to separate from one another, making two rounds of purification necessary. The yield of this nitration can certainly be improved further, by optimizing the purification step and regioselectivity of the reaction by tighter temperature control. Related work concerning the synthesis of ipso substituted arenes[62] and the mentioned intermediate structures involved suggest that an increased reaction time and continuous lowered temperatures might be necessary to achieve the desired selectivity.

During the Corey-Chaykovsky epoxidation a sulfur ylide reacts with imines or carbonyls, an aldehyde in this case, to yield epoxides or aziridines. During the classical conditions, dimethylsulfonium iodide or dimethylsulfoxonium iodide is treated with sodium hydride in DMSO to create the sulfur ylide. Subsequent addition of the carbonyl to the reaction mixture results in nucleophilic attack from the ylide and elimination of dimethyl sulfide or dimethylsulfoxide, dependent on the initial sulfur species used. Examination of both reagents showed that using dimethylsulfoxonium iodide gives the better yields, and reaction conditions are less temperature sensitive. The ylide formed using dimethylsulfonium iodide is unstable at room temperature and demands thorough cooling. This requires the addition of an appropriate amount of THF to prevent freezing. Even when an excess of sodium hydride and dimethylsulfonium iodide was used, starting material was detectable by TLC and NMR. Presumably because of the decomposition of the ylide.

Utilizing dimethylsulfoxonium iodide, the reaction could be performed at room temperature and no starting material was detectable after 75 min. reaction time. Interestingly, after

addition of aldehyde to the ylide, the solution turned dark blue and after quenching the reaction by addition of water the mixture took a brown color. No further analysis of the stereochemistry of the product of this reaction was performed, but as no chiral reactants or catalysts were used, it is assumed that a racemic mixture was obtained. Various approaches towards a stereoselective Corey-Chaykovsky epoxidation are known in the literature, [63–65] and may be an option, if the configuration of the auxiliary proves to be of importance in future experiments.

Opening of this epoxide to yield product **5** was tried with various sulfur nucleophiles, including sodium sulfide[66] and bis(trimethylsilyl)sulfide (HMDST)[67] with unsatisfactory results. Thioacetic acid in pyridine was chosen at the end, as these conditions gave acceptable yields. Under base catalysis, the ring opening is initiated with the attack of the thioacetate nucleophile at the sterically less hindered carbon. In contrast, acid catalysed reaction would lead to addition and opening at the higher substituted carbon. One side product which forms during this reaction was identified as compound **6**. To a small extend, also pyridine (and other bases)[68] catalyse the migration of the acetyl group from thioester **5** to the ester **6**. The amount of this side product was, after separation by column chromatography, more than the TLC would have initially indicated. It was at that point that the tendency of the substrate **5** to readily undergo acetyl migration became apparent, as also silica, and therefore column chromatography, catalyses this reaction[52]. Initial synthesis plans intended to deprotect the acetylated thiol **5** later with sodium methoxide, but throughout the work it was discovered that this distinctive feature can be used at our advantage. Instead of separating the two structural isomers, both were purified together from the crude reaction mixture and again subjected to silica gel, to facilitate acetyl migration, rather than to avoid it. Additionally, it was found that the suggested amount of silica gel per mmol thiol of 8.61 g by Gao et.al. [52] should not be reduced, as the reaction was not complete in experiments with a reduced amount.

Alternatively to the method described above, where silica gel and substrate are stirred together, a column can be packed as usual for column chromatography, and the educt mixture directly applied to it. As the former procedure allows better control for an even distribution of the substrate to the silica gel, stirring was chosen over direct application onto the packed column.

The free thiol (**6**) obtained by this procedure showed to be quite stable at air to oxidation as solid and in solution, as long as the environment was not basic. Initially, the thiol protection was pictured to work under similar conditions as used by Bello et. al.[38], by thiol deprotection/deprotonation using sodium methoxide and an activated tertiary butyl thiol (1-(tert-Butylthio)-1,2-hydrazinecarboxmorpholide **16**. Unfortunately, this treatment resulted not only in the formation of product **7**, but also in an elimination reaction, to give 30-55% of an vinyl derivative (methyl 4-(2-methoxy-5-nitro-4-vinylphenoxy)butanoate). This elimination is not restricted to ester **6** but also takes place if the isomeric thioester **5** is subjected to the same

treatment. Therefore other conditions had to be found.

Treatment with an excess of neat tBuSH (50equ.) and triethylamine resulted in the formation of compound **7** and various side products. Additionally, tBuSH has an disagreeable odour and is volatile, which makes oxidation by bubbling of air problematic. The use of a different base and an even higher excess of tBuSH are possible changes that may be tested in future experiments.

It was found that the aforementioned elimination poses a lesser issue when using triethylamine as a base. Therefore, activated tertiary butyl thiol (1-(tert-Butylthio)-1,2-hydrazinecarboxmorpholide) **16** in combination with triethylamine was examined in the protection of the thiol. Even under strictly oxygen free conditions, the reaction showed a disproportionally high formation of a homo dimer (disulfide) of **6**. Although not further verified, this behaviour could be explained by a side reaction of the activated thiol or its secondary product. By addition of methanol, a reduced amount of triethylamine and an increased excess of activated thiol, little side products/homo dimer could be observed on TLC and by NMR. After the improvement of this reaction in respect to the formation of a homodimer and elimination, it was even more unpleasant to see that, after column chromatography, only 42% yield could be obtained. Due to limited time, no further effort in optimization of the thiol protection reaction was made, but screening of other reaction conditions should lead to the desired improvement.

To obtain amide **8** from ester **7**, a modified procedure utilizing sodium methoxide as catalyst[54] was used. The substrate was treated with an excess of ethylenediamine and sodium methoxide in toluene at 50°C, not only to convert the ester into the amide, but also to simultaneously liberate the hydroxy group via aminolysis. This reaction poses a major progress in the experimental procedure compared to the original synthesis. There, a comparable substrate was converted to an amine using a zirconium (IV) catalyst, which is expensive and difficult to handle because of its very high water sensitivity.

Fmoc protection of amine **8** was straightforward and achieved with Fmoc-Cl in DCM.

4.2 Coupling

A major motivation for the synthesis of this new generation auxiliary was the prospect of its coupling to the N-terminus of the peptide via Mitsunobu-Fukuyama alkylation, independently of the nature of N-terminal amino acid of the peptide.

Nosylation of the peptides N-terminus was tested with nosyl chloride in combination with different bases and solvents, to find the best conditions. Analysis via LC/MS showed that, among collidine in DMF and DIEA (diisopropylethylamine) in DMF, the best result could be obtained with collidine in NMP, in regards of reaction completion and side product formation (data not shown).

The initial alkylation experiment, using a modified protocol by Demmer *et al*.[57], indicated that reaction does indeed occur, although the substrates are particularly sterically hindered. Analysis of alkylation of peptide **11** showed the desired product **12** (**Figure 22**) and a peak with a mass corresponding to its reduced form with -88 Da (≙ StBu). This can be explained with the ability of phosphines to reduce disulfides in the presence of water (contamination). However, the major product formed showed a mass of 3050 Da, 16 Da less than **12**. Possible explanations involve unwanted modification/oxidation of the disulfide motive, side reaction during cleavage or reduction of the nitro to a nitroso group, but neither of these theories could be fully confirmed.

To investigate these promising findings further, as well as to minimize side product formation and to reduce the complexity of the used peptide, the reaction was repeated under similar conditions and with several alterations in the procedure. Unfortunately, the results could not be reproduced, as no reaction was detectable at all. Also by changing numerous parameters, the initial findings could not be observed again.

The reason for these failures was identified in the THF used in these last experiments, as it was presumably contaminated with water. For the alkylation of peptide **14** (**Figure 23**), a freshly opened bottle of anhydrous THF was used. Interestingly, no disulfide reduction was observable in this case. Besides the expected compound **15**, the major product could be identified as +72 Da to the expected mass. As above, further investigation has to be done to identify the nature of this side product and find the right condition to avoid its formation.

The factor of water contamination is the most likely explanation for whether the reaction proceeds or not. It does not explain the difference in side products and the altered ratio of product to side product when alkylations of peptide **11** and **14** are compared (**Figure 22**/**Figure 23**). These findings could come from the different peptides that were used (**Scheme 4**), the different cleavage cocktails or the change in the order of addition of reagents. Additionally, the presence of water, which is necessary for phosphines to reduce disulfides, could have given rise to the different outcome of the two experiments. A clear explanation for this behaviour remains yet to be found and proven.

A factor that has also to be accounted for is the nature of the employed TentaGel solid phase. The combination of a polystyrene matrix modified with polyethyleneglycol gives good swelling properties with a broad spectrum of solvents, allows for homogenous reactions with sterically demanding reagents and enables the generation of a masked thioester via hydrazinolysis.[69] However, the hydrophilic nature of PEG also favours the retention of water, and therefore might influence the water sensitive alkylation reaction negatively. To account for this, the resin should be washed multiple times with anhydrous solvents and dried at high vacuum, until no further change in weight is observed.

It is clear that in both cases the reaction did not go to completion, as there is still starting material observable. This could be circumvented by addition of an even greater excess of reagents, or potentially better, a second round of alkylation.

As described previously by Dawson *et al.,*[45] the configuration of the stereocenter, introduced during Corey-Chaycovsky epoxidation on the benzylic position of the auxiliary, does not influence the rate of native chemical ligation. Therefore, no effort was undertaken to favour the formation of either enantiomer or to separate the racemic mixture. Additionally, during HPLC of the peptide auxiliary conjugate, no difference in the retention times of the two diastereoisomer conjugates was observed. Yet unanswered is the question on whether the configuration in the presence of the PEG polymer or chiral amino acids directly attached to the auxiliary is influencing ligation rates and efficiency.

Peptides **11** and **14** both carry an N-terminal glycine. Due to obvious steric reasons, this is the first amino acid to test coupling of the auxiliary and later the rate of ligation. Ligation auxiliaries in general and benzylamine derivatives in particular suffer from a substantial reduction in ligation efficiency if the amino acid at the junction (N-terminal at the auxiliary and C-terminal at the thioester) is not glycine. Using the new generation auxiliary in combination with the Mitsunobu alkylation, it will be possible to investigate conditions to improve auxiliary-mediated native chemical ligation in the presence of more sterically demanding amino acids than shown previously.[38]

Because of the limited time available, other conditions for the Mitsunobu reaction have not been investigated during this master thesis. After Mitsunobu, removal of the nosyl group on the N-terminal amine would be required. In general, mild cleavage of the nosyl group is performed by treatment with thiolate nucleophiles via formation of a meisenheimer complex.[61] Potentially problematic could be a possible reaction of the tert-butyl sulfanyl group of the auxiliary with the employed thiolate, e.g. thiol-disulfide exchange, giving the deprotected thiol. This may lead to unwanted side reactions. Or a different disulfide where the tBuS protecting group is substituted by the thiol used in deprotection, which presumably poses a lesser problem.

Another approach possible for the attachment of the auxiliary to the peptide N-terminus is the coupling of the auxiliary to an amino acid in solution. These building blocks can then be used in SPPS, to avoid direct coupling to the peptide, which may need to be optimized for each

individual experimental setup/peptide. Preliminary experiments (data not shown), converting the alcohol on the auxiliary into a mesylate, tosylate or halogen to enable S_N2 reaction with glycine (protected as *tert*-butyl ester) did not give satisfying results. Treatment of the auxiliary with methanesulfonyl chloride does not stop at the desired mesylate, but rather proceeds directly to the chloride. This problem may be avoided by using methanesulfonic anhydride. No formation in the case of the tosylate by reaction of the auxiliary with p-toluenesufonyl chloride was observed. Steric hindrance of the adjacent tertiary butyl group could be the reason for this lack of reactivity. In addition, also Mitsunobu Fukuyama alkylation in solution did not yield the desired product. The reason for this is presumably the aforementioned contamination of THF with water.

The tested conversion of the hydroxy group of the auxiliary into a suitable substrate for S_N2 reaction and the in solution Fukuyama alkylation strategy require further investigation, but could pose a valuable extension of the auxiliaries applicability.

5 Conclusion and outlook

During the course of this work, a novel synthetic route for a new generation UV cleavable PEGylated ligation auxiliary was developed. The new auxiliary has the potential to be coupled to the N-terminus of any peptide independent from the nature of the N-terminal amino acid. The key step of the synthesis, the Corey-Chaykovsky reaction, gave an epoxide, which was subsequently opened with a sulfur nucleophile, leading to a mercapto ethylenehydroxy group instead of the mercapto ethylene amino group of previously described auxiliaries. The combination of the reactivity of this moiety with the known capacity of the PEGylated auxiliary in order to drastically simplify peptide purification as well as to mediate native chemical ligation and to enable cleavage via UV light, may prove to be a valuable extension of the set of tools available for glycopeptide synthesis. The synthesis of the new generation auxiliary is straightforward, but an improvement of certain steps is still necessary, in order to obtain it in high overall yield.

The altered chemical structure of the novel auxiliary enables closer investigation of possible coupling reactions to the N-terminus to increase its applicability and versatility. First promising results could be obtained by applying the Mitsunobu-Fukuyama alkylation to the auxiliary coupling reaction. The mild conditions are highly compatible with standard Fmoc SPPS and allow its use in combination with glycosylated peptides.

Further research is needed to investigate the nosyl deprotection of the peptide and the auxiliary performance in coupling and ligation at junctions containing sterically hindered amino acids.

Synthesis of auxiliary-carrying amino acid building blocks for use in SPPS may even further simplify incorporation of the auxiliary.

In conclusion, a concise synthetic route towards a new generation PEGylated ligation auxiliary was successfully established during this work and first encouraging steps towards the novel strategy of coupling a ligation auxiliary via Mitsunobu-Fukuyama alkylation could be established.

6 Experimental

Chemicals and solvents were obtained in the highest purity available, and used without any further purification from Sigma-Aldrich (Taufkirchen, Germany), TCI-Europe (Zwijndrecht, Belgium), VWR (Darmstadt, Germany), J.T.Baker (Griesheim, Germany).

Reaction monitoring was done via thin layer chromatography (TLC sheets Alugram Xtra sil G/UV_{254} 0.2mm silical gel 60; Macherey-Nagel (Druen, Germany)) and detected with UV light at 254nm, $KMnO_4$ stain or Pancaldi reagent. Product purification was performed using flash column chromatography (Silica gel 60, 230-400 mesh, 0.04-0.063 mm, Macherey-Nagel (Druen, Germany)).

^{1}H NMR were recorded on a Bruker AV400 spectrometer at 400 MHz. The signal of the residual solvent was used as reference ($CDCl_3$: 7.27 ppm). Coupling constants are given in Hz. ^{13}C NMR were recorded on the same device stated above at 101 MHz. The signal of the residual solvent was used as reference ($CDCl_3$: 77 ppm).

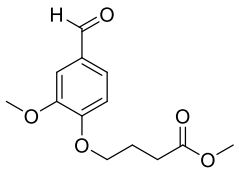
Infrared (IR) spectra were recorded on a FT-IR-Bruker Vertex 70-spectrometer.

High resolution mass spectra were obtained using a Bruker maXis ESI-Qq-oaRTOF system in positive ion mode.

Peptides and peptide auxiliary conjugates were analysed by liquid chromatography – mass spectrometry (LC-MS) on either a Waters AutoPurification HPLC/MS system with electrospray ionisation (ESI) in positive ion mode or on a Dionex Ultimate 3000 (Thermo Scientific) HPLC/MS instrument. The compounds were separated on a Kromasil 300-5-C4 column (4.6x150mm, 10μm particle size) with a flow rate of 1ml/min and applying a linear gradient from 5% to 65% of buffer B (ACN + 0.05 % TFA) in buffer A (ddH_2O + 0.05 % TFA) in 30min.

6.1 Auxiliary Synthesis

Methyl 4-(4-formyl-2-methoxyphenoxy)butanoate **2**



2

Vanillin **1** (10g, 66mmol) was dissolved in anhydrous acetonitrile (140 ml) under argon atmosphere. K_2CO_3 (18.2g, 131mmol, 2equ.), tetrabutylammonium iodide (4.85g, 13mmol, 0.2equ.) and methyl 4-chlorobuyrate (11.67g, 10.4ml, 85mmol, 1.3equ.) were added under stirring. The pink suspension was stirred under reflux for 18h, during which the colour changed to yellow. After cooling to room temperature, the solid was filtered off and washed thoroughly with ethyl acetate. The solvent of the combined filtrate and washings was evaporated under reduced pressure. The crude was dissolved in ethyl acetate and washed with H_2O (50ml) and brine (50ml). After drying of the organic phase over $MgSO_4$ the solvent was evaporated under reduced pressure.

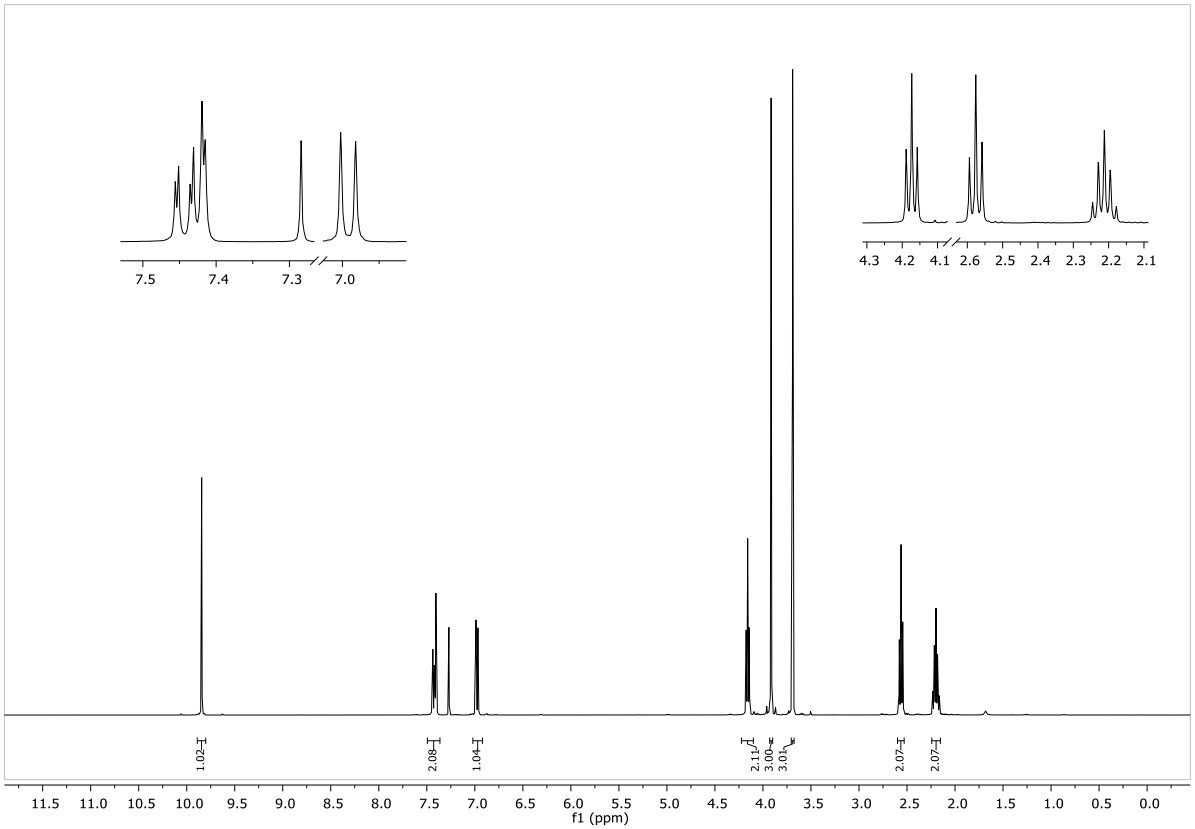
The crude was recrystaliyed from methanol. The supernatant was concentrated under reduced pressure and purified by flash column chromatography on silica gel (PE/EtAc 1:1) to recover the product that did not crystallize. The products of the two purifications (15.12g, 60mmol, 91%yield, white solid) gave identical analytical data.

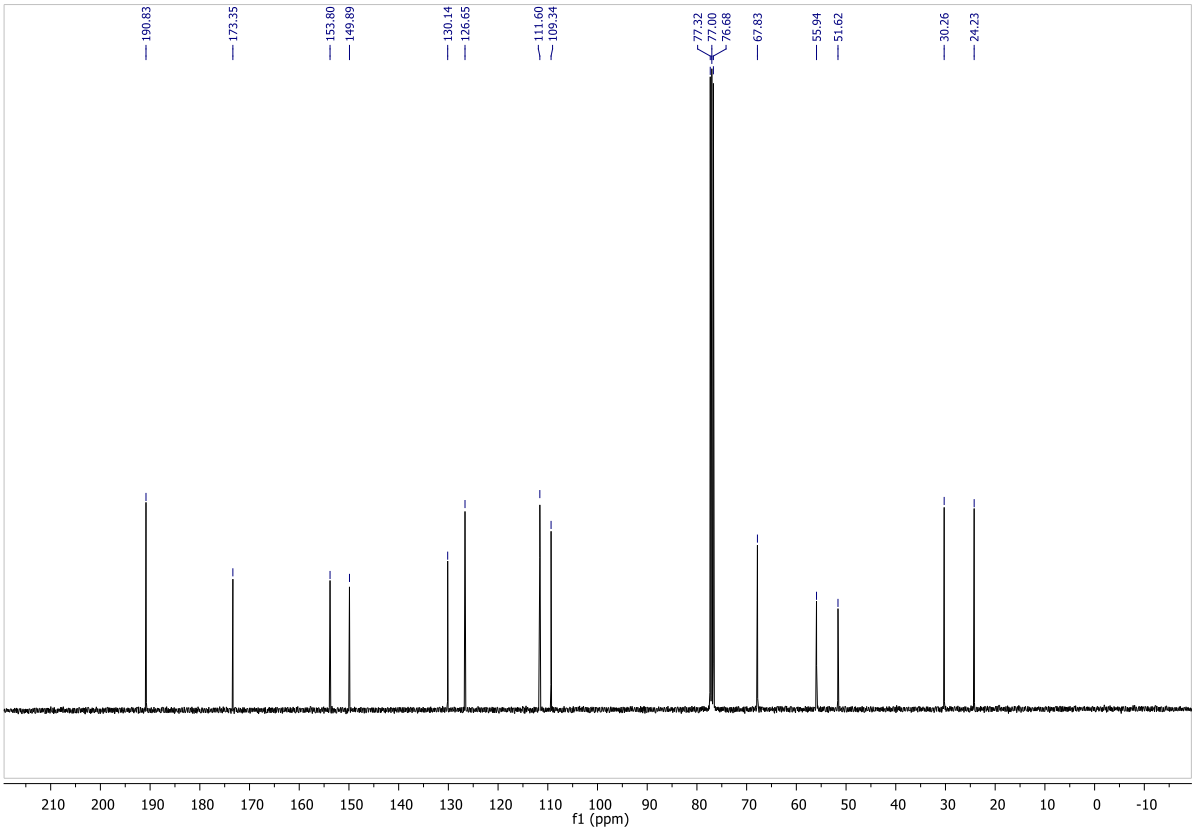
^{1}H NMR (400 MHz, $CDCl_3$) δ: 9.84 (s, 1H, C**H**O), 7.47 – 7.37 (m, 2H, $\mathbf{H}_{Ar}$), 6.98 (d, J=8.1, 1H, $\mathbf{H}_{Ar}$), 4.16 (t, J=6.3, 2H, OC$\mathbf{H}_2$CH$_2$), 3.92 (s, 3H, OC$\mathbf{H}_3$), 3.69 (s, 3H, COOC$\mathbf{H}_3$), 2.56 (t, J=7.2, 2H, CH$_2$C$\mathbf{H}_2$CO), 2.26 – 2.13 (m, 2H, CH$_2$C$\mathbf{H}_2$CH$_2$).

^{13}C NMR (101 MHz, $CDCl_3$) δ: 190.81 (**C**HO), 173.34 (**C**OOCH$_3$), 153.79 ($\mathbf{C}_{Ar}$), 149.88 ($\mathbf{C}_{Ar}$), 130.13 ($\mathbf{C}_{Ar}$), 126.65 ($\mathbf{C}_{Ar}$), 111.60 ($\mathbf{C}_{Ar}$), 109.34 ($\mathbf{C}_{Ar}$), 67.83 (O**C**H$_2$CH$_2$), 55.95 (O**C**H$_3$), 51.63 (COO**C**H$_3$), 30.27 (**C**H$_2$CO), 24.24 (CH$_2$**C**H$_2$CH$_2$).

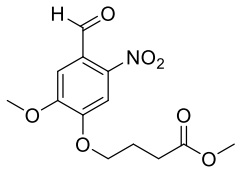
HRMS (ESI): mass calculated for $C_{13}H_{16}O_5$: 252.0998 g/mol; found 253.1071 $[M+H]^+$, 275.0892 $[M+Na]^+$, 291.0630 $[M+K]^+$

IR (ATR): cm^{-1} 2950, 2835, 2727, 2359, 1731, 1679, 1584, 1508, 1466, 1423, 1397, 1339, 1262, 1237, 1195, 1160, 1133, 1028, 945, 866, 809, 781, 730, 654, 590


7.5
7.4
7.3
7.0
4.3
4.2
4.1
2.6
2.5
2.4
2.3
2.2
2.1
1.02
2.08
1.04
2.11
3.00
3.01
2.07
2.07
11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)


190.83
173.35
153.80
149.89
130.14
126.65
111.60
109.34
77.32
77.00
76.68
67.83
55.94
51.62
30.26
24.23
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

Methyl 4-(4-formyl-2-methoxy-5-nitrophenoxy)butanoate **3**



3

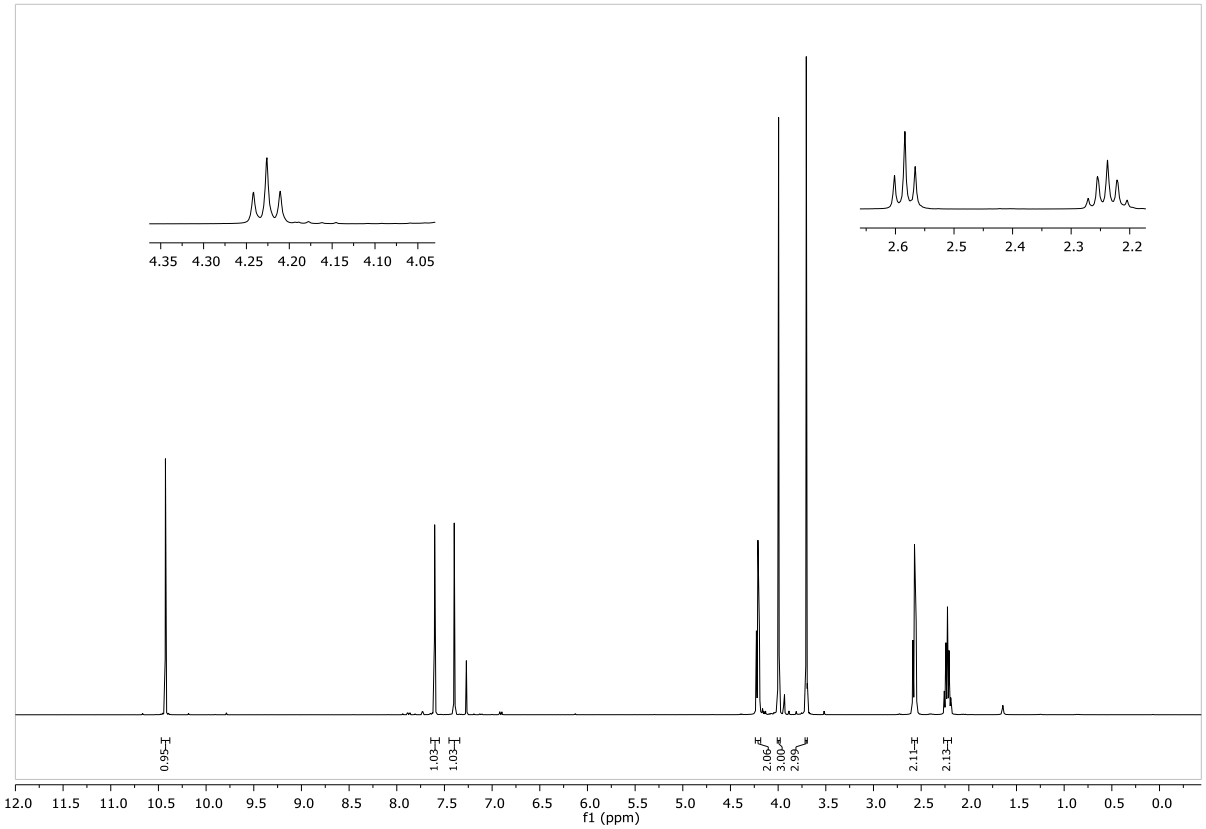
An ice cold mixture of glacial acetic acid (126ml) and fuming nitric acid (34ml) was added dropwise to Methyl 4-(4-formyl-2-methoxyphenoxy)butanoate **2** (14.5g, 57.5mmol) under cooling on ice over the course of 25min. After consumption of the starting material (1h, monitored by TLC), the reaction was quenched by pouring it on ice. The mixture was extracted with DCM (4x 50ml) and the combined organic phases washed with water (4x 50ml) and sat. aq $NaHCO_3$ (2x 50ml) until neutral pH. The crude was subjected to flash column chromatography (DCM/petrolether/ethyl acetate 8.75:1:0.25 to DCM/ethyl acetate 4:1). Non pure fractions were again purified (DCM/petrol ether/ethyl acetate 8.9:1:0.1), giving the pure product as yellow solid in 70% total yield (40.27mmol, 11.96g) after evaporation of the solvent under reduced pressure.

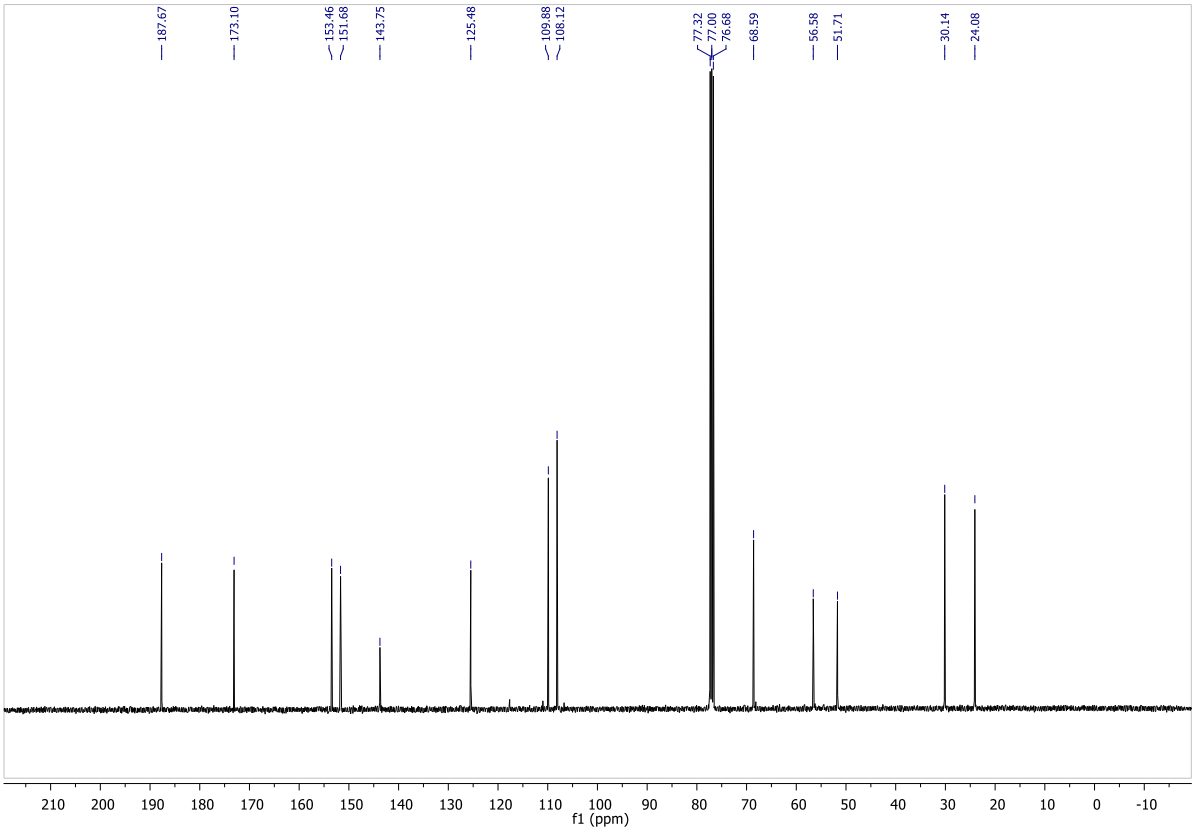
^{1}H NMR (400 MHz, $CDCl_3$) δ: 10.44 (s, 1H, C**H**O), 7.62 (s, 1H, $\mathbf{H}_{Ar}$), 7.41 (s, 1H, $\mathbf{H}_{Ar}$), 4.23 (t, J=6.2, 2H, OC$\mathbf{H}_2$CH$_2$), 4.01 (s, 3H, OC$\mathbf{H}_3$), 3.72 (s, 3H, COO$\mathbf{H}_3$), 2.58 (t, J=7.1, 2H, C$\mathbf{H}_2$CO), 2.30 – 2.16 (m, 2H, CH$_2$C$\mathbf{H}_2$CH$_2$).

^{13}C NMR (101 MHz, $CDCl_3$) δ: 187.69 (**C**HO), 173.12 (**C**OOCH$_3$), 153.48 ($\mathbf{C}_{Ar}$), 151.70 ($\mathbf{C}_{Ar}$), 143.78 ($\mathbf{C}_{Ar}$), 125.51 ($\mathbf{C}_{Ar}$), 109.90 ($\mathbf{C}_{Ar}$), 108.15 ($\mathbf{C}_{Ar}$), 68.62 (O**C**H$_2$CH$_2$), 56.62 (O**C**H$_3$), 51.75 (COO**C**H$_3$), 30.18 (**C**H$_2$CO), 24.12 (CH$_2$**C**H$_2$CH$_2$).

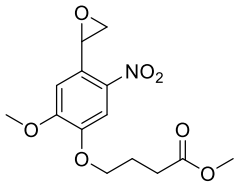
HRMS (ESI): mass calculated for $C_{13}H_{15}NO_7$: 297.0849 g/mol; found 298.0918 $[M+H]^+$, 320.0739 $[M+Na]^+$

IR (ATR): cm^{-1} 2952, 2360, 1734, 1687, 1602, 1572, 1519, 1467, 1439, 1404, 1331, 1283, 1223, 1169, 1060, 1003, 944, 881, 818, 738,


4.35 4.30 4.25 4.20 4.15 4.10 4.05
2.6 2.5 2.4 2.3 2.2
0.95
1.03
1.03
2.06
3.00
2.99
2.11
2.13
12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)


187.67
173.10
153.46
151.68
143.75
125.48
109.88
108.12
77.32
77.00
76.68
68.59
56.58
51.71
30.14
24.08
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

Methyl 4-(2-methoxy-5-nitro-4-(oxiran-2-yl)phenoxy)butanoate **4**



4

Sodium hydride (1.97g, 60% in mineral oil, 8.9mmol, 1,5eq.) was washed under argon atmosphere with anhydrous THF. Trimethylsulfoxonium iodide (360mg, 8.9mmol, 1.5equ.) in 17ml anhydrous DMSO was added under argon and stirring at 0°C and the solution was further stirred at room temperature. After gas evolution stopped (40min.), a solution of **3** (1.77g, 6mmol) in THF (7ml) was added dropwise at 0°C. The dark blue solution was lead to warm up to room temperature under stirring. The reaction was quenched by addition of water after 75min, upon which the mixture turned dark brown. Diethyl ether was added and the two phases were separated. The aqueous phase was washed with diethyl ether (2x) and the combined organic phases washed with 5% LiCl (2x) and water, dried over $MgSO_4$ and the solvent was evaporated under reduced pressure.

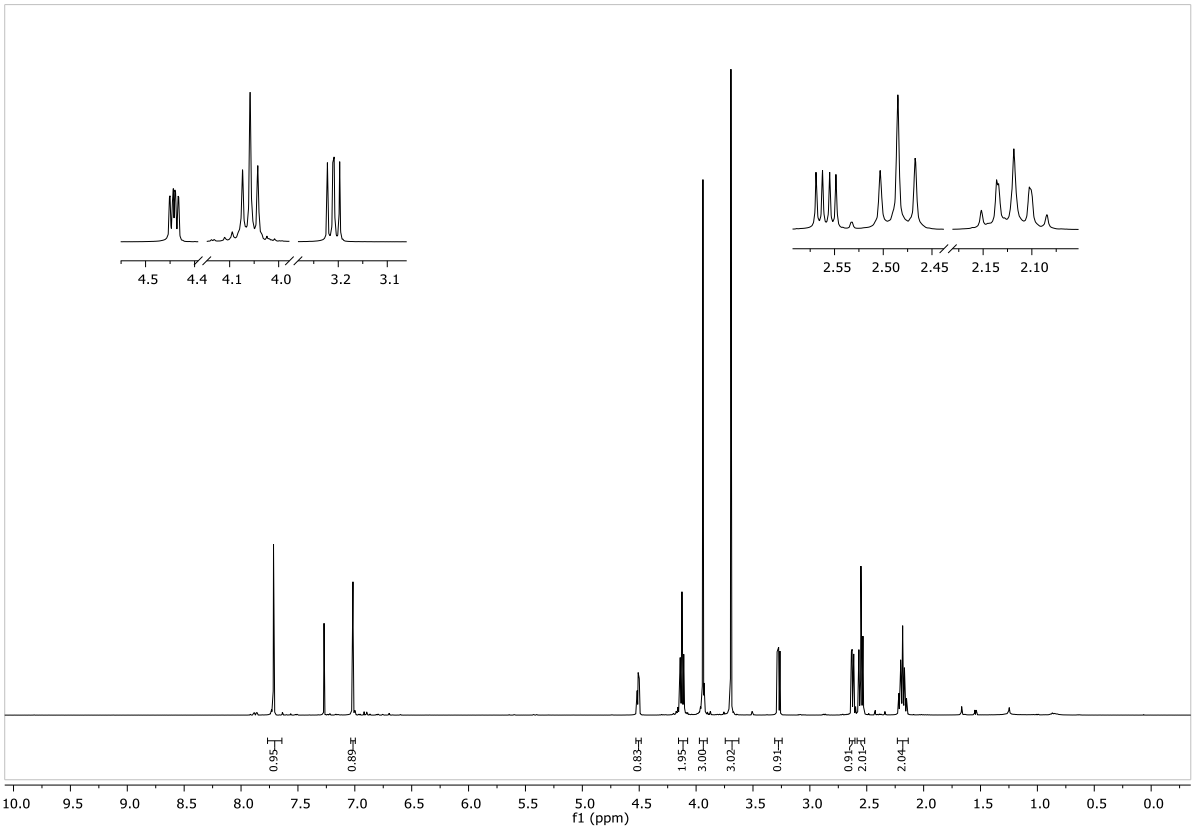
Flash column chromatography (petrol ether/ethyl acetate 1:1, crude dry loaded) gave epoxide **4** in 69% yield (4.11mmol, 1.28g) as a yellow solid.

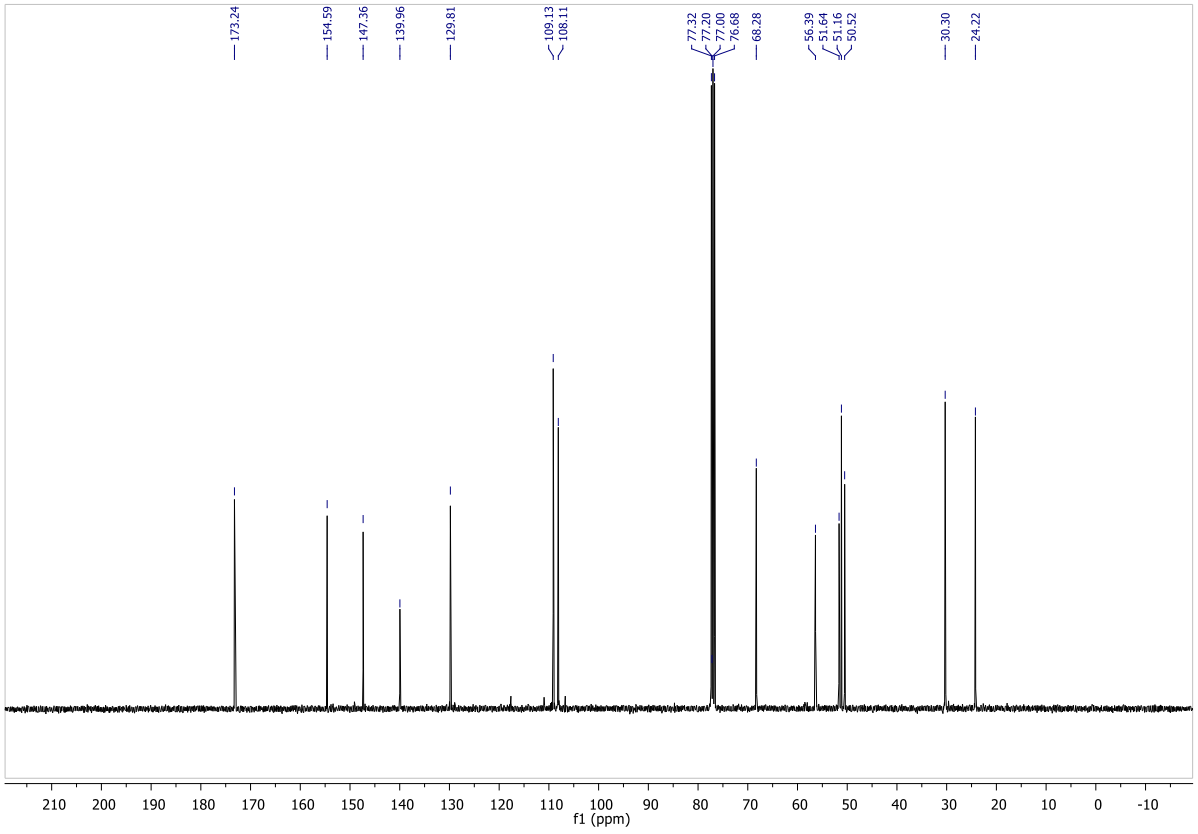
^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.65 (s, 1H, $\mathbf{H}_{Ar}$), 6.95 (s, 1H, $\mathbf{H}_{Ar}$), 4.44 (ddd, J=4.4, 2.6, 0.6, 1H, C**H**OCH_2), 4.06 (t, J=6.3, 2H, OC$\mathbf{H}_2$$CH_2$), 3.87 (s, 3H, OC$\mathbf{H}_3$), 3.63 (s, 3H, COOC$\mathbf{H}_3$), 3.21 (dd, J=5.5, 4.4, 1H, CHOC$\mathbf{H}_2$), 2.56 (dd, J=5.5, 2.6, 1H, CHOC$\mathbf{H}_2$), 2.49 (t, J=7.2, 2H, C$\mathbf{H}_2$CO), 2.12 (tt, J=7.2, 6.2, 2H, CH_2C$\mathbf{H}_2$$CH_2$).

^{13}C NMR (101 MHz, $CDCl_3$) δ = 173.27 (**C**OOCH_3), 154.62 ($\mathbf{C}_{Ar}$), 147.39 ($\mathbf{C}_{Ar}$), 140.00 ($\mathbf{C}_{Ar}$), 129.85 ($\mathbf{C}_{Ar}$), 109.16 ($\mathbf{C}_{Ar}$), 108.14 ($\mathbf{C}_{Ar}$), 68.32 (O**C**$H_2$$CH_2$), 56.43 (O**C**$H_3$), 51.68 (COO**C**$H_3$), 51.21 (**C**HO$CH_2$), 50.56 (CHO**C**$H_2$), 30.34 (**C**$H_2$CO), 24.26 ($CH_2$**C**$H_2$$CH_2$).

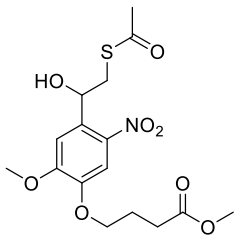
HRMS (ESI): mass calculated for $C_{14}H_{17}NO_7$: 311.1005; found 312.1078 $[M+H]^+$, 334.0898 $[M+Na]^+$

IR (ATR): cm^{-1} 2952, 2360, 1735, 1617, 1578, 1519, 1469, 1439, 1405, 1334, 1280, 1257, 1217, 1173, 1062, 1008, 937, 878, 844, 814, 760, 705


4.5
4.4
4.1
4.0
3.2
3.1
2.55
2.50
2.45
2.15
2.10
0.95
0.89
0.83
1.95
3.00
3.02
0.91
0.91
2.01
2.04
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


173.24
154.59
147.36
139.96
129.81
109.13
108.11
77.32
77.20
77.00
76.68
68.28
56.39
51.64
51.16
50.52
30.30
24.22
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

Methyl 4-(4-(2-(acetylthio)-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate **5**



5

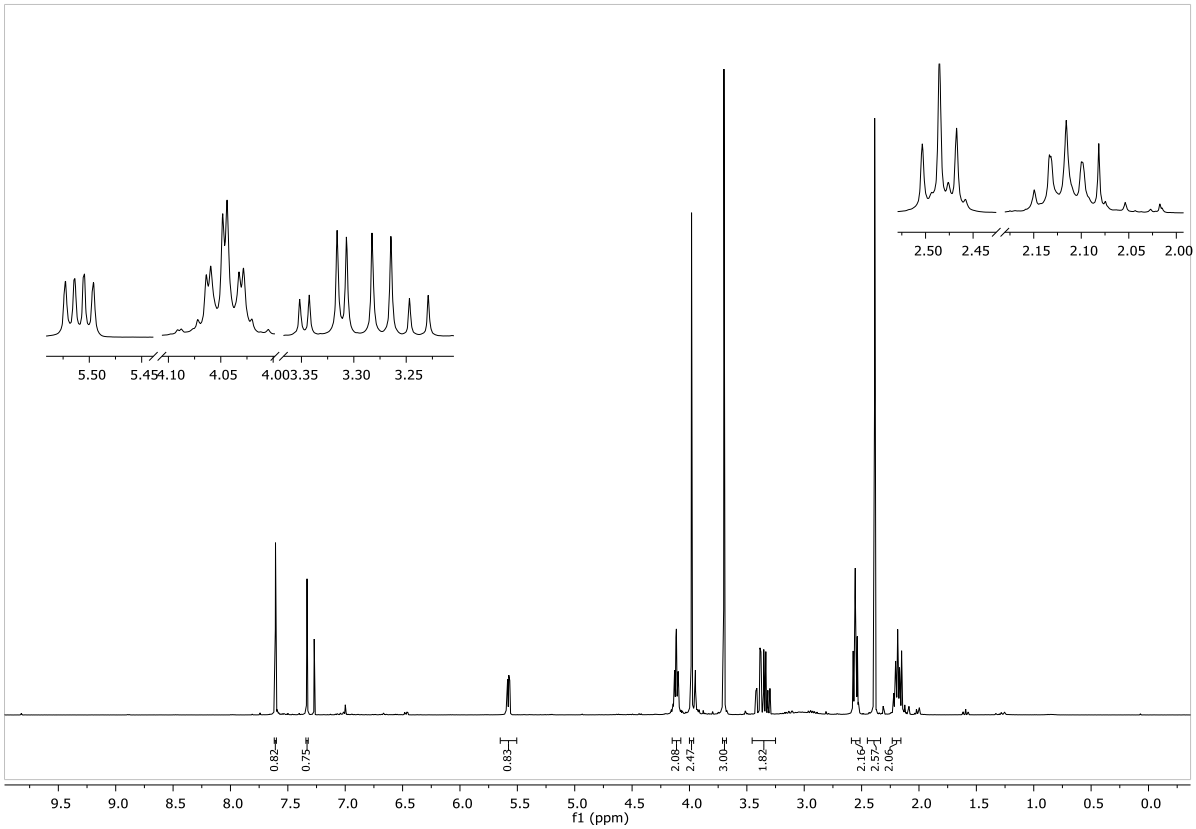
Thioacetic acid (2.86g, 2.6ml, 37.6mmol, 1.6equ) and pyridine (12ml) were added to **3** (7.27g, 23.34mmol) at 0°C under stirring and the solution was heated to 25°C. After stirring for 16h, the reaction was quenched by addition of DCM and a saturated aqueous solution of ammonium chloride. The organic phase was washed with a saturated aqueous solution of ammonium chloride (3x) and water (1x), dried over $MgSO_4$, and the solvent evaporated. To eliminate residual pyridine, the crude was dissolved in toluene and concentrated under vacuum. After flash column chromatography (petrolether/ethyl acetate 2:1 later changed to 1:1), S acetylated (**5**) and O acetylated (**6**) products could be obtained as yellow, oily liquid in a combined yield of 74% (6.7g, 17.3mmol).

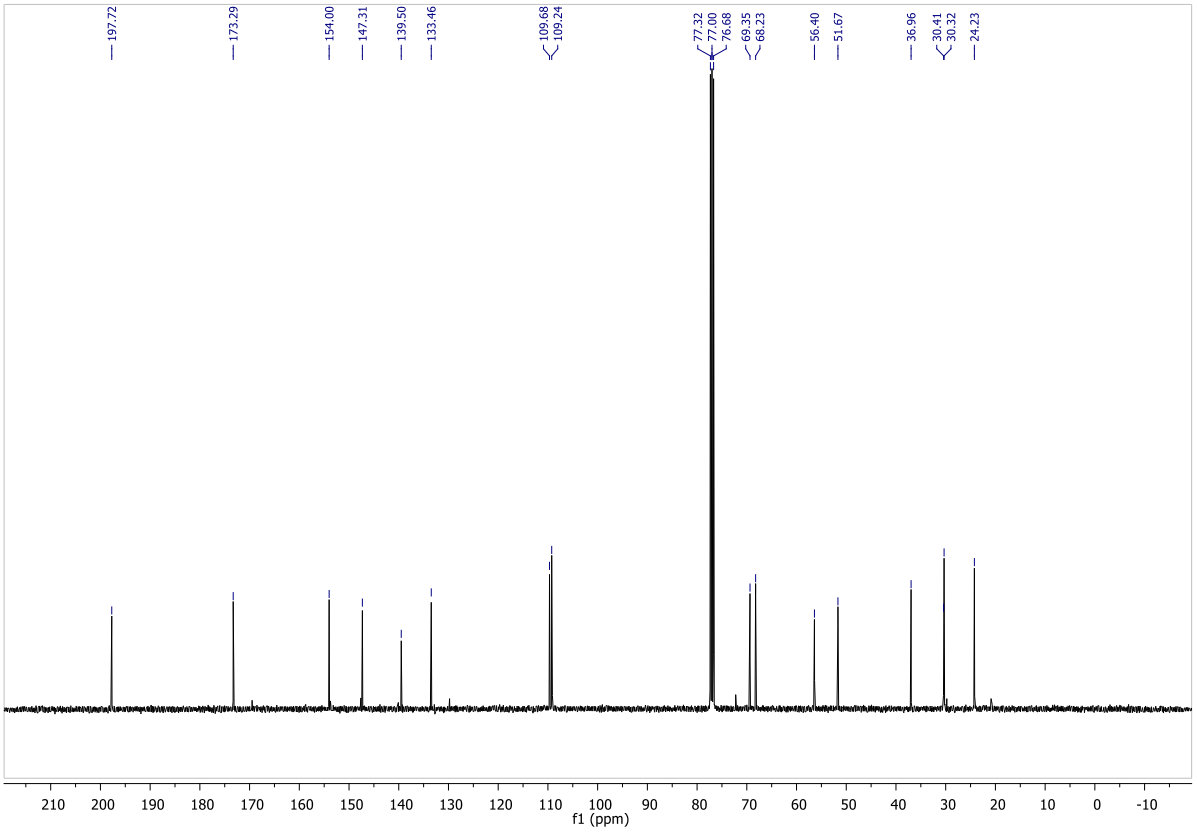
^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.60 (s, 1H, $\mathbf{H}_{Ar}$), 7.32 (s, 1H, $\mathbf{H}_{Ar}$), 5.57 (dd, *J*=7.2, 3.5, 1H, C**H**OH), 4.22 – 4.03 (m, 2H, OC$\mathbf{H}_2$CH$_2$), 3.97 (s, 3H, OC$\mathbf{H}_3$), 3.69 (s, 3H, COO$\mathbf{H}_3$), 3.43 – 3.25 (m, 2H, C$\mathbf{H}_2$S), 2.55 (t, *J*=7.2, 2H, C$\mathbf{H}_2$CO), 2.38 (s, 3H, SCOC$\mathbf{H}_3$), 2.23 – 2.11 (m, 2H, CH$_2$C$\mathbf{H}_2$CH$_2$).

^{13}C NMR (101 MHz, $CDCl_3$) δ: 197.74 (S**C**OCH$_3$), 173.32 (**C**OOCH$_3$), 154.03 ($\mathbf{C}_{Ar}$), 147.34 ($\mathbf{C}_{Ar}$), 139.52 ($\mathbf{C}_{Ar}$), 133.49 ($\mathbf{C}_{Ar}$), 109.71 ($\mathbf{C}_{Ar}$), 109.27 ($\mathbf{C}_{Ar}$), 69.39 (**C**OH), 68.26 (O**C**H$_2$CH$_2$), 56.43(O**C**H$_3$), 51.71 (COO**C**H$_3$), 37.00 (**C**H$_2$S), 30.45 (**C**H$_2$CO), 30.36 (SCO**C**H$_3$), 24.27 (CH$_2$**C**H$_2$CH$_2$).

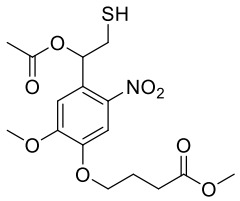
HRMS (ESI): mass calculated for $C_{16}H_{21}NO_8S$: 387.0988 g/mol; found 410.0884 $[M+Na]^+$, 426.0625 $[M+K]^+$

IR (ATR): cm^{-1} 3507, 2951, 1734, 1693, 1613, 1578, 1518, 1468, 1439, 1400, 1330, 1272, 1214, 1175, 1134, 1062, 953, 877, 818, 754, 627


5.50
5.45
4.10
4.05
4.00
3.35
3.30
3.25
2.50
2.45
2.15
2.10
2.05
2.00
0.82
0.75
0.83
2.08
2.47
3.00
1.82
2.16
2.57
2.06
f1 (ppm)


197.72
173.29
154.00
147.31
139.50
133.46
109.68
109.24
77.32
77.00
76.68
69.35
68.23
56.40
51.67
36.96
30.41
30.32
24.23
f1 (ppm)

Methyl 4-(4-(1-acetoxy-2-mercaptoethyl)-2-methoxy-5-nitrophenoxy)butanoate **6**



6

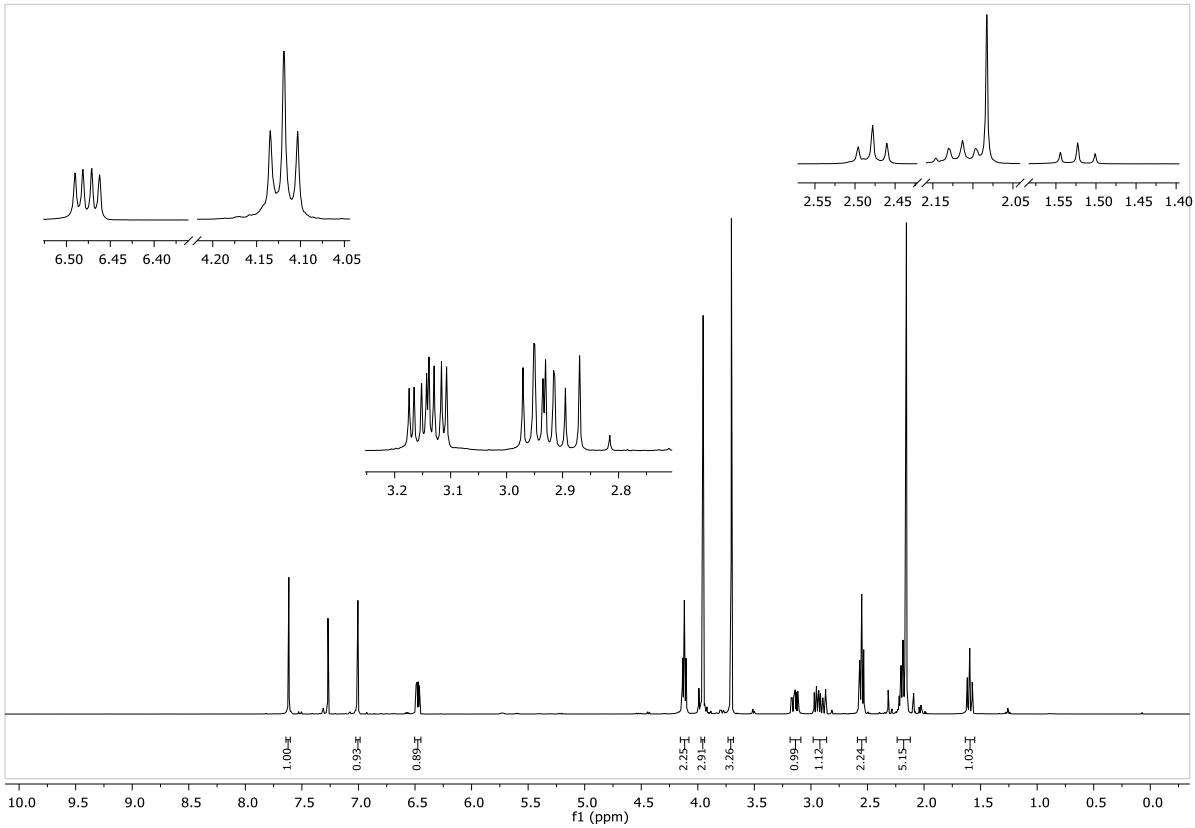
Silica gel (8.5g/mmol) was added to the mixture of **5** and **6** (6.7 g, 17 mmol) and enough solvent (approximately 300ml, petrolether/ethyl acetate 1: 1) to enable stirring. After 17 h the mixture was poured into a chromatography column, and the compound eluted from the silica gel gave the pure product as a yellow solid in 92% yield (15.6mmol, 6.05g), after evaporation of the solvent under vacuum.

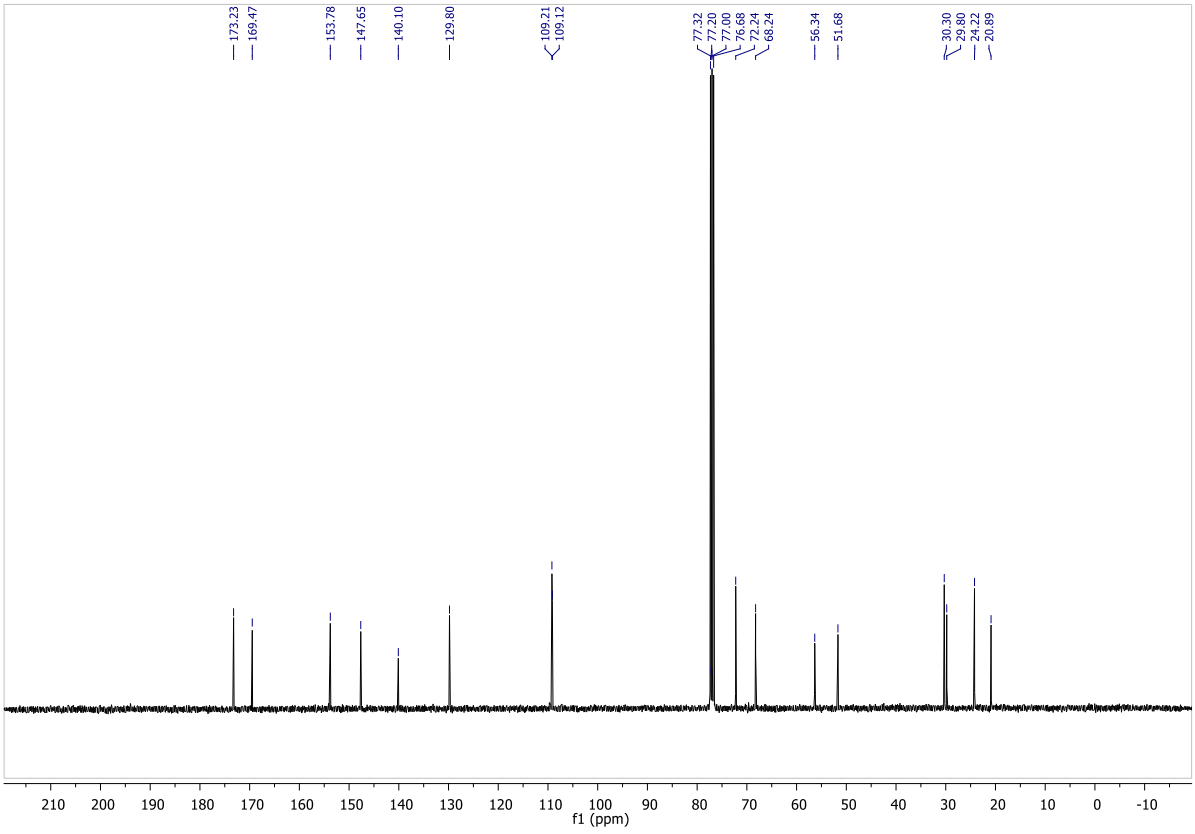
^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.61 (s, 1H, **H**$_{Ar}$), 7.00 (s, 1H, **H**$_{Ar}$), 6.47 (dd, *J*=7.6, 3.6, 1H, C**H**OHCH$_2$), 4.11 (t, *J*=6.2, 2H, OC**H**$_2$CH$_2$), 3.94 (s, 3H, OC**H**$_3$), 3.69 (s, 3H, COOC**H**$_3$), 3.13 (ddd, *J*=14.2, 8.9, 3.5, 1H, C**H**$_2$SH), 2.99 – 2.83 (m, 1H, C**H**$_2$SH), 2.54 (t, *J*=7.2, 2H, C**H**$_2$CO), 2.22 – 2.15 (m, 2H, CH$_2$C**H**$_2$CH$_2$), 2.15 (s, 3H, OCOC**H**$_3$), 1.59 (t, *J*=8.7, 1H, S**H**).

^{13}C NMR (101 MHz, $CDCl_3$) δ: 173.24 (**C**OOCH$_3$), 169.49 (O**C**OCH$_3$), 153.79 (**C**$_{Ar}$), 147.67 (**C**$_{Ar}$), 140.11 (**C**$_{Ar}$), 129.82 (**C**$_{Ar}$), 109.22 (**C**$_{Ar}$), 109.14 (**C**$_{Ar}$), 72.27 (C**H**CH$_2$), 68.27 (O**C**H$_2$CH$_2$), 56.37 (O**C**H$_3$), 51.71 (COO**C**H$_3$), 30.33 (**C**H$_2$CO), 29.83 (**C**H$_2$SH), 24.24 (CH$_2$**C**H$_2$CH$_2$), 20.92 (OCO**C**H$_3$).

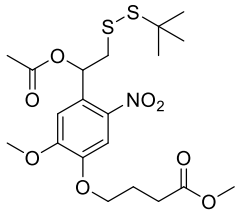
HRMS (ESI): mass calculated for $C_{16}H_{21}NO_8S$: 387.0988 g/mol; found 410.0884 $[M+Na]^+$, 426.0625 $[M+K]^+$

IR (ATR): cm^{-1} 2949, 1734, 1579, 15181477, 1438, 1402, 1370, 1332, 1271, 1217, 1174, 1060, 1033, 942, 873, 818, 755, 722, 613


6.50 6.45 6.40
4.20 4.15 4.10 4.05
2.55 2.50 2.45 2.15
2.05 1.55 1.50 1.45 1.40
3.2 3.1 3.0 2.9 2.8
1.00 0.93 0.89 2.25 2.91 3.26 0.99 1.12 2.24 5.15 1.03
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)


173.23
169.47
153.78
147.65
140.10
129.80
109.21
109.12
77.32
77.20
77.00
76.68
72.24
68.24
56.34
51.68
30.30
29.80
24.22
20.89
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

Methyl 4-(4-(1-acetoxy-2-(tert-butyldisulfaneyl)ethyl)-2-methoxy-5-nitrophenoxy)butanoate **7**



7

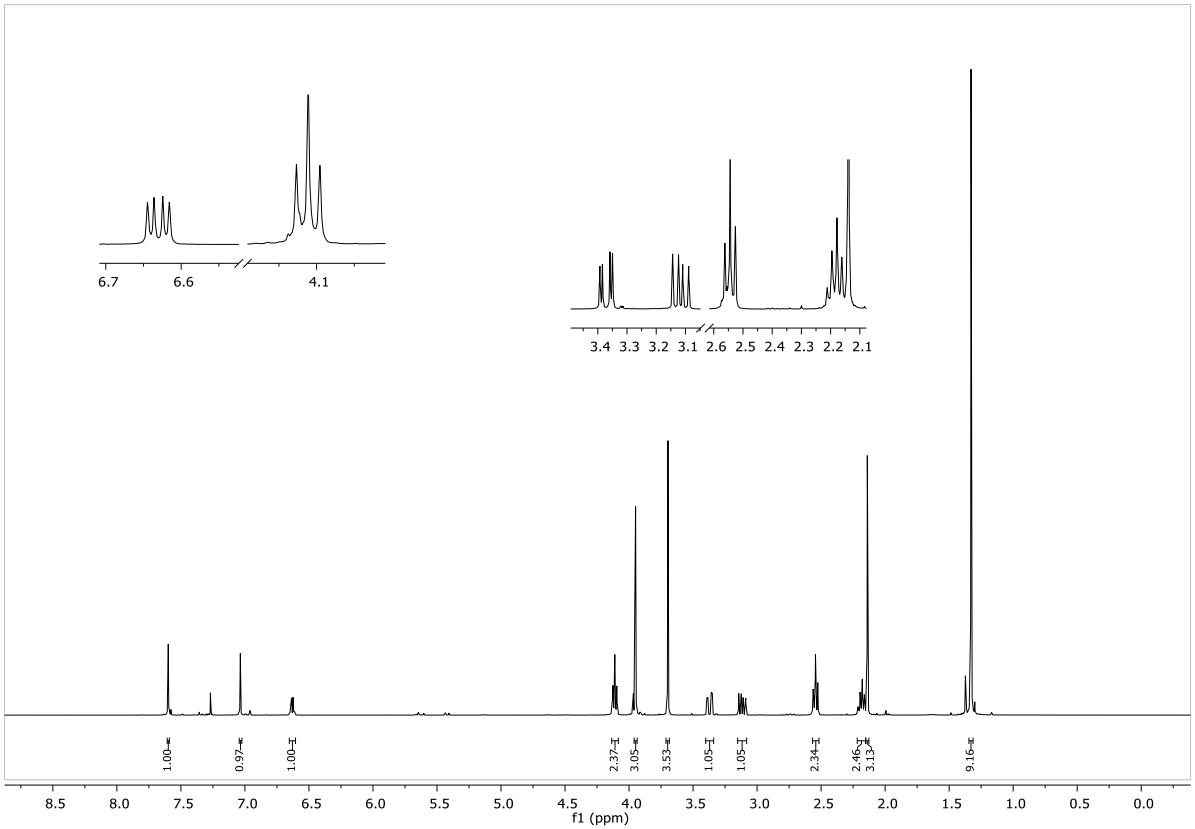
Methyl 4-(4-(1-acetoxy-2-mercaptoethyl)-2-methoxy-5-nitrophenoxy)butanoate **6** (128mg, 0.33mmol) was dissolved in previously degassed DMF (2ml) under argon atmosphere. 1-(tert-Butylthio)-1,2-hydrazinecarboxmorpholide[38,53] (**16**, 344mg, 1mmol, 3equ.) was dissolved in MeOH (1,7ml) and this solution was added to the solution of **6** dropwise at 0°C, followed by triethylamine (22mg, 31µl, 0.22mmol 0.67 equ.) added dropwise. After consumption of the starting material, as confirmed via TLC (petrolether/ethylacetate 1:1), the reaction was quenched via concentration of the mixture under vacuum and addition of diethylether. The organic phase was washed with 5% LiCl (2x 20ml), water (1x 20ml) and brine (1x 20ml), dried over $MgSO_4$ and the solvent evaporated under reduced pressure. Product **7** could be obtained after flash column chromatography (petrolether/ethyl acetate 1:1), in 42% yield (66mg, 0.139mmol) as a yellow oil.

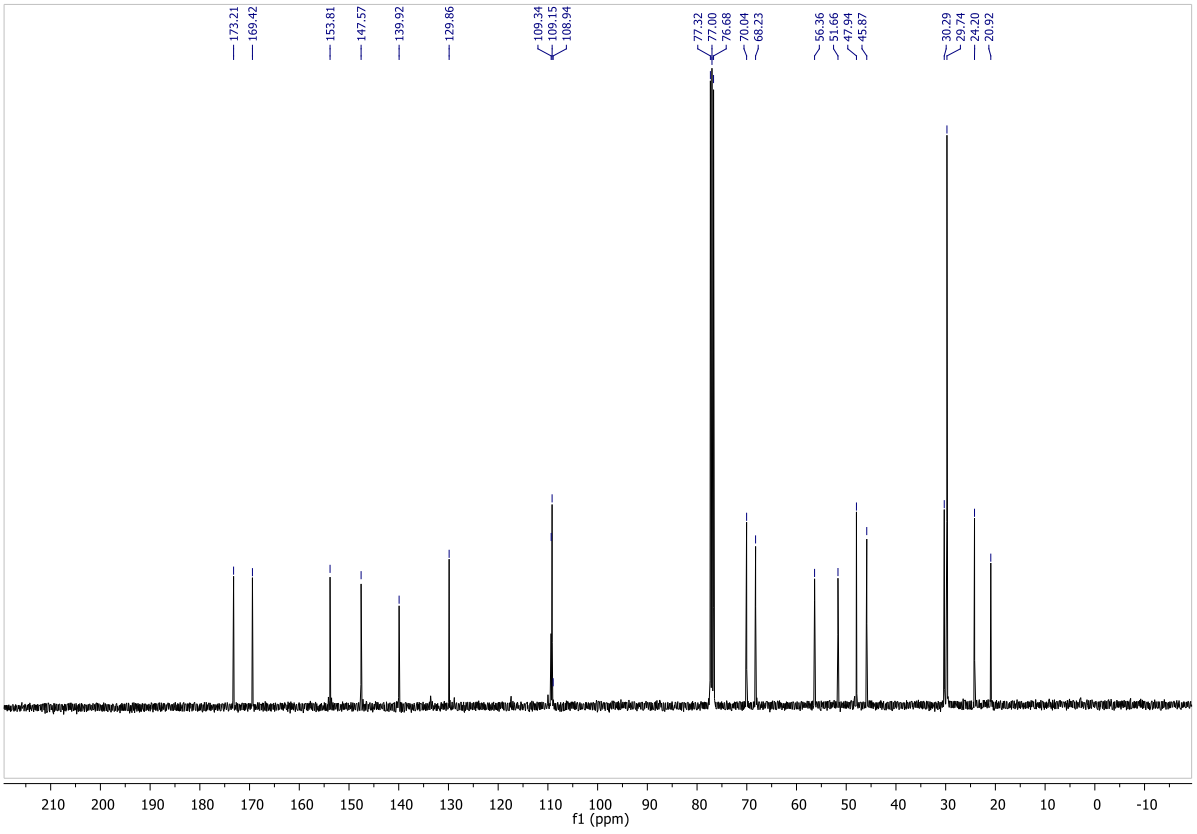
^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.61 (s, 1H, **C**$_{Ar}$), 7.05 (s, 1H, **C**$_{Ar}$), 6.65 (dd, *J*=8.1, 3.5, 1H, C**H**OCOCH$_3$), 4.13 (t, *J*=6.2, 2H, OC**H**$_2$CH$_2$), 3.96 (s, 3H, OC**H**$_3$), 3.71 (s, 3H, COOC**H**$_3$), 3.39 (dd, *J*=13.8, 3.5, 1H, C**H**$_2$SS), 3.13 (dd, *J*=13.8, 8.1, 1H, C**H**$_2$SS), 2.56 (t, *J*=7.2, 2H, C**H**$_2$CO), 2.20 (q, *J*=6.4, 2H, CH$_2$C**H**$_2$CH$_2$), 2.15 (s, 3H, OCOC**H**$_3$), 1.34 (s, 9H, SC(C**H**$_3$)$_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ: 173.24 (**C**OOCH$_3$), 169.45 (O**C**OCH$_3$), 153.84 (**C**$_{Ar}$), 147.60 (**C**$_{Ar}$), 139.96 (**C**$_{Ar}$), 129.90 (**C**$_{Ar}$), 109.38 (**C**$_{Ar}$), 109.19 (**C**$_{Ar}$), 70.08 (**C**HCH$_2$), 68.27 (OC**H**$_2$CH$_2$), 56.40 (O**C**H$_3$), 51.70 (COO**C**H$_3$), 47.99 (CH**C**H$_2$), 45.91 (**C**(CH$_3$)$_3$), 30.33 (**C**H$_2$CO), 29.79 (C(**C**H$_3$)$_3$), 24.24 (CH$_2$**C**H$_2$CH$_2$), 20.97 (OCO**C**H$_3$).

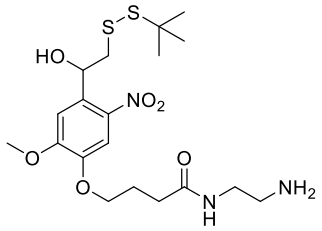
HRMS (ESI): mass calculated for $C_{20}H_{29}NO_8S_2$: 475.1335 g/mol; found 498.1231 $[M+Na]^+$

IR (ATR): cm^{-1} 2957, 2361, 1735, 1613, 1579, 1519, 1470, 1438, 1401, 1363, 1331, 1273, 1214, 1169, 1060, 1022, 930, 872, 817, 733, 702, 612


6.7
6.6
4.1
3.4 3.3 3.2 3.1 2.6 2.5 2.4 2.3 2.2 2.1
1.00
0.97
1.00
2.37
3.05
3.53
1.05
1.05
2.34
2.46
3.13
9.16
8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)


173.21
169.42
153.81
147.57
139.92
129.86
109.34
109.15
108.94
77.32
77.00
76.68
70.04
68.23
56.36
51.66
47.94
45.87
30.29
29.74
24.20
20.92
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

N-(2-aminoethyl)-4-(4-(2-(tert-butyldisulfanyl)-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanamide **8**



8

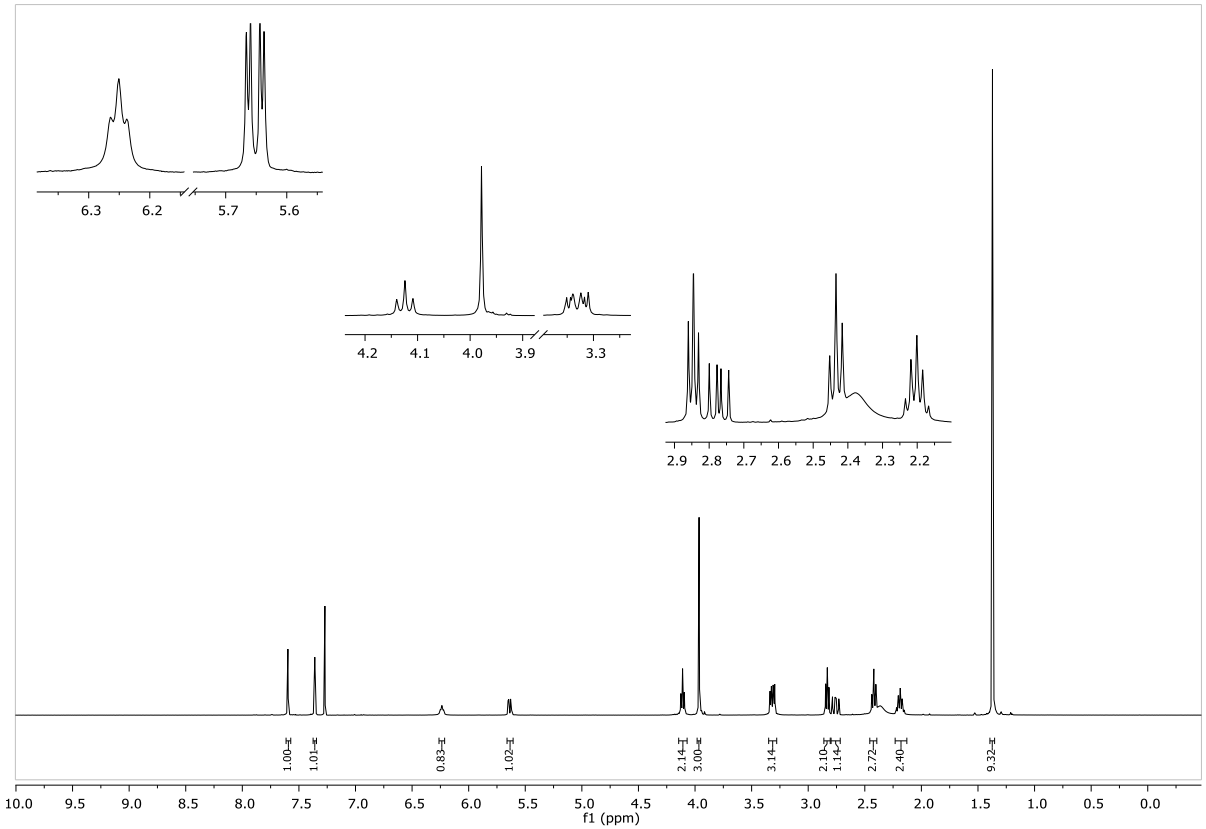
Methyl 4-(4-(1-acetoxy-2-(tert-butyldisulfaneyl)ethyl)-2-methoxy-5-nitrophenoxy)butanoate **7** (320mg, 0.67mmol) was dissolved in toluene (0.75ml). After addition of ethylenediamine (750µl) and of a catalytic amount of sodium methoxide, the solution was heated to 50°C and stirred at that temperature for 18h. The solvent was evaporated under vacuum and the residual ethylenediamine coevaporated three times with toluene. Product **8** was obtained pure as a yellow oil in 80% yield (250mg, 0.54mmol) after column chromatography on silica gel (DCM/MeOH 9:1 +NH_4OH)

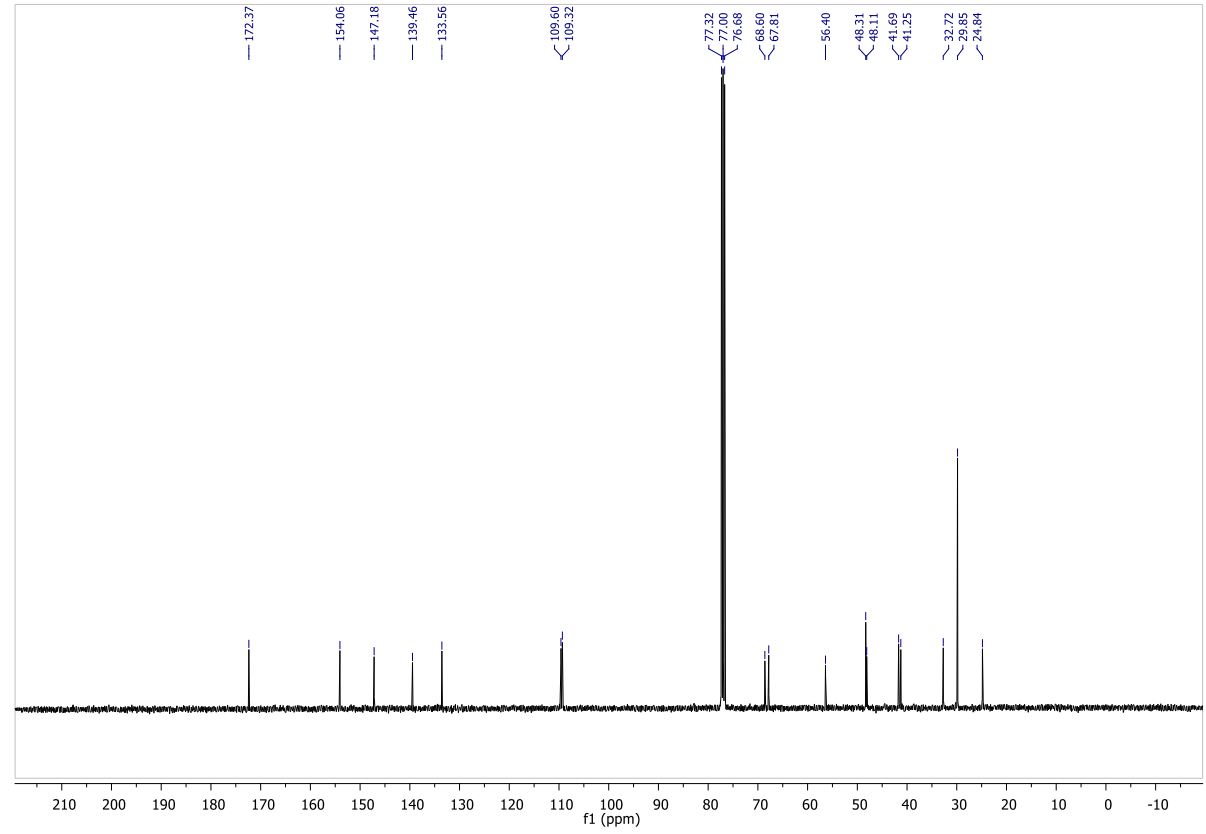
^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.61 (s, 1H, **H**$_{Ar}$), 7.37 (s, 1H, **H**$_{Ar}$), 6.25 (t, *J*=5.7, 1H, CON**H**), 5.65 (dd, *J*=8.9, 2.7, 1H, C**H**OHCH$_2$), 4.12 (t, *J*=6.2, 2H, OC**H**$_2$CH$_2$), 3.98 (s, 3H, OC**H**$_3$), 3.38 – 3.24 (m, 3H, CHOHC**H**$_2$, NHC**H**$_2$CH$_2$), 2.85 (t, *J*=5.9, 2H, CH$_2$C**H**$_2$NH$_2$), 2.77 (dd, *J*=13.4, 9.0, 1H, CHOHC**H**$_2$), 2.43 (t, *J*=7.2, 2H, C**H**$_2$CO), 2.20 (p, *J*=6.7, 2H, CH$_2$C**H**$_2$CH$_2$), 1.39 (s, 9H, C(C**H**$_3$)$_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ: 172.38 (**C**ONH), 154.07 (**C**$_{Ar}$), 147.19 (**C**$_{Ar}$), 139.47 (**C**$_{Ar}$), 133.57 (**C**$_{Ar}$), 109.62 (**C**$_{Ar}$), 109.33 (**C**$_{Ar}$), 68.62 (O**C**H$_2$CH$_2$), 67.84 (**C**HOHCH$_2$), 56.42 (O**C**H$_3$), 48.33 (**C**H$_2$SS), 48.13 (S**C**(CH$_3$)$_3$), 41.72 (CH$_2$CONH**C**H$_2$CH$_2$), 41.28 (CH$_2$**C**H$_2$NH$_2$COO), 32.74 (**C**H$_2$CO), 29.87 (C(**C**H$_3$)$_3$), 24.86 (CH$_2$**C**H$_2$CH$_2$).

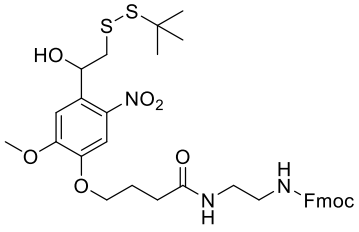
HRMS (ESI): mass calculated for $C_{19}H_{31}N_3O_6S_2$: 461.1654 g/mol; found 426.1727 $[M+H]^+$, 484.1546 $[M+Na]^+$, 500.1284 $[M+K]^+$

IR (ATR): cm^{-1} 3296, 2960, 1648, 1576, 1518, 1471, 1362, 1330, 1271, 1214, 1168, 1062, 879, 731


6.3
6.2
5.7
5.6
4.2
4.1
4.0
3.9
3.3
2.9
2.8
2.7
2.6
2.5
2.4
2.3
2.2
1.00
1.01
0.83
1.02
2.14
3.00
3.14
2.10
1.14
2.72
2.40
9.32
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


172.37
154.06
147.18
139.46
133.56
109.60
109.32
77.32
77.00
76.68
68.60
67.81
56.40
48.31
48.11
41.69
41.25
32.72
29.85
24.84
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

(9H-fluoren-9-yl)methyl (2-(4-(4-(2-(tert-butyldisulfaneyl)-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanamido)ethyl)carbamate **9**



9

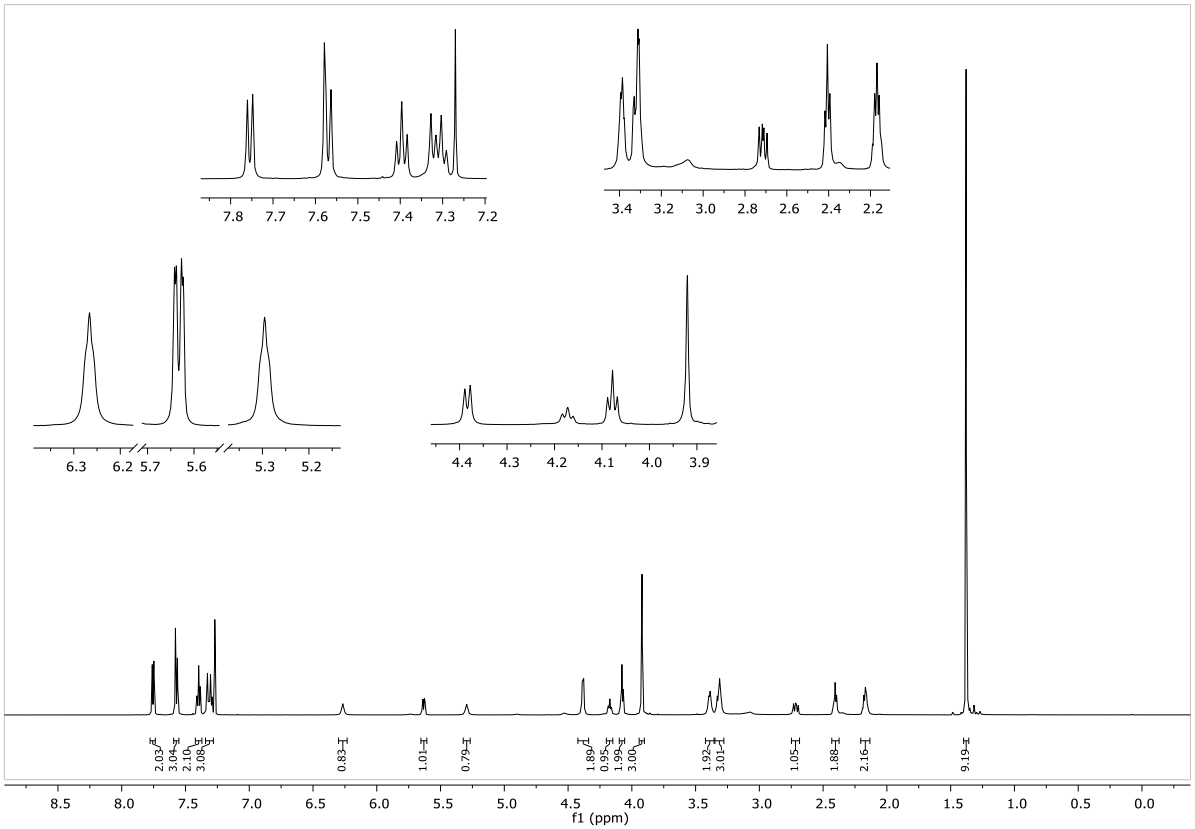
N-(2-aminoethyl)-4-(4-(2-(tert-butyldisulfanyl)-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy) butanamide **8** (446mg, 0.97mmol) was dissolved in DCM (3ml) at 0°C, then N,N-diisopropylethylamine (149mg, 202µl, 1.2equ.) was added dropwise. Whilst still stirring on ice, Fmoc chloride (299mg, 1.164mmol, 1.2equ.) in DCM (1.5ml) was added. After consumption of the starting material (35min, monitored via TLC), DCM was added to the reaction mixture, and the solution was extracted with 0.1M HCl. The organic phase was dried over $MgSO_4$, the solid was filtered off and then the solvent was evaporated under vacuum. Product **9** was obtained pure (595mg, 0.87mmol, 90%) as a yellow solid after flash column chromatography on silica gel (DCM+2.5% MeOH to 3% MeOH).

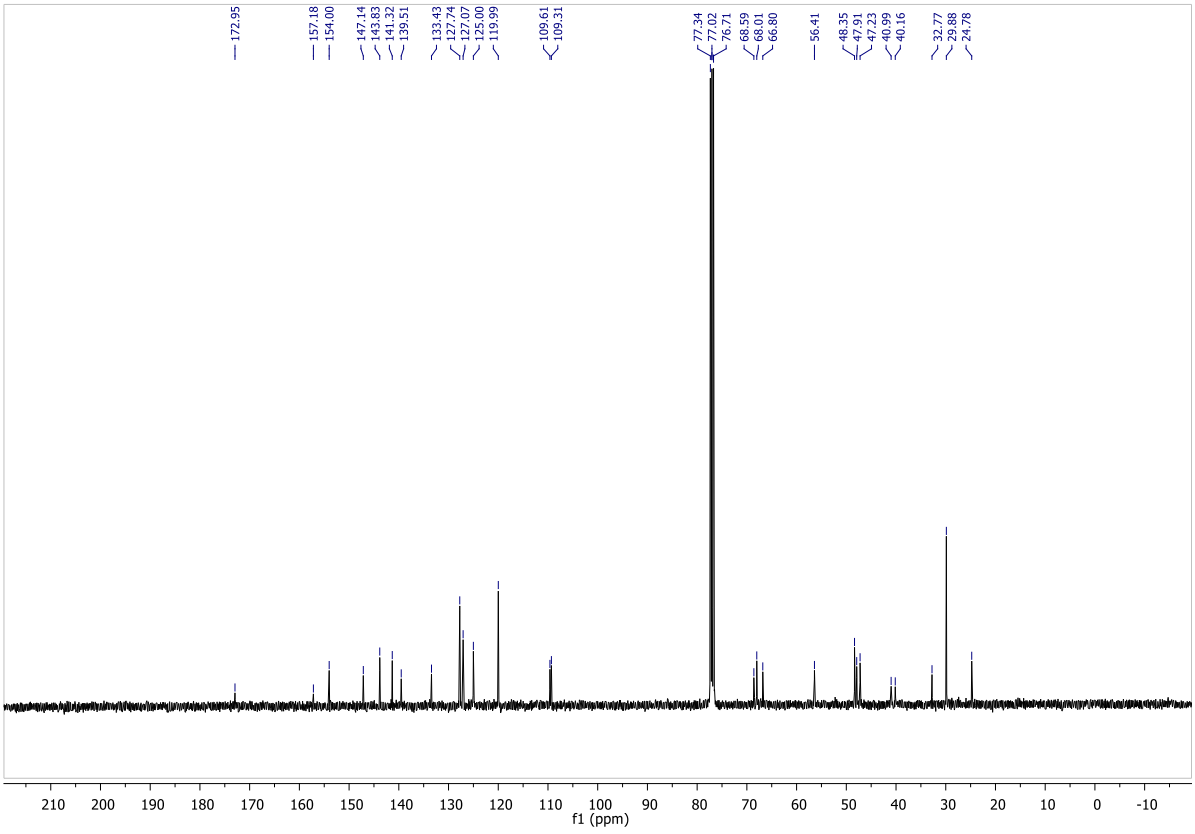
^{1}H NMR (600 MHz, $CDCl_3$) δ: 7.74 (d, J=7.5, 2H, H_{Ar}), 7.59 – 7.54 (m, 3H, , H_{Ar}), 7.39 (t, J=7.5, 2H, H_{Ar}), 7.34 – 7.27 (m, 3H, H_{Ar}), 6.26 (t, J=5.7, 1H, $CH_2CON\mathbf{H}CH_2$), 5.62 (dd, J=9.0, 2.7, 1H, $C\mathbf{H}OHCH_2$), 5.29 (t, J=6.0, 1H, $N\mathbf{H}COOCH_2$), 4.37 (d, J=6.9, 2H, CH_2(Fmoc)), 4.16 (t, J=7.0, 1H, CH(Fmoc)), 4.07 (t, J=6.1, 2H, OCH_2CH_2), 3.91 (s, 3H, OCH_3), 3.41 – 3.35 (m, 2H, $CH_2CONHC\mathbf{H}_2CH_2$), 3.34 – 3.27 (m, 3H, CH_2SS, $CH_2C\mathbf{H}_2NHCOO$), 2.70 (dd, J=13.5, 9.0, 1H, CH_2SS), 2.40 (t, J=7.2, 2H, $OCH_2C\mathbf{H}_2CONH$), 2.16 (p, J=6.6, 2H, $OCH_2C\mathbf{H}_2CH_2$), 1.37 (s, 9H, $SC(CH_3)_3$).

^{13}C NMR (101 MHz, $CDCl_3$) δ: 172.76 ($CH_2\mathbf{C}ONH$), 156.99 ($NH\mathbf{C}OO$), 153.81 (C_{Ar}), 146.95 (C_{Ar}), 143.64 (C_{Ar}), 141.13 (C_{Ar}), 139.32 (C_{Ar}), 133.24 (C_{Ar}), 127.55 (C_{Ar}), 126.88 (C_{Ar}), 124.82 (C_{Ar}), 119.81 (C_{Ar}), 109.42 (C_{Ar}), 109.13 (C_{Ar}), 68.41 ($O\mathbf{C}H_2CH_2$), 67.83 ($\mathbf{C}HOHCH_2$), 66.62 (CH_2(Fmoc)), 56.23 (OCH_3), 48.17 ($\mathbf{C}H_2SS$), 47.73 (CH(Fmoc)), 47.05 ($S\mathbf{C}(CH_3)_3$), 40.81 ($\mathbf{C}H_2NHCOO$), 39.97 ($CH_2CONH\mathbf{C}H_2$), 32.59 ($\mathbf{C}H_2CONH$), 29.70 ($SC(\mathbf{C}H_3)_3$), 24.60 ($OCH_2\mathbf{C}H_2CH_2$).

HRMS (ESI): mass calculated for $C_{34}H_{41}N_3O_8S_2$: 683.2335 g/mol; 684.2408 $[M+H]^+$, 706.2231 $[M+Na]^+$,722.1979 $[M+K]^+$

IR (ATR): cm^{-1} 3318, 3067, 2940, 2248, 1702, 1649, 1576, 1515, 1449, 1401, 1361, 1328, 1267, 1213, 1164, 1060, 907, 876, 815, 759, 726, 646


7.8 7.7 7.6 7.5 7.4 7.3 7.2
3.4 3.2 3.0 2.8 2.6 2.4 2.2
6.3 6.2 5.7 5.6 5.3 5.2
4.4 4.3 4.2 4.1 4.0 3.9
2.03 3.04 2.10 3.08 0.83 1.01 0.79 1.89 0.95 1.99 3.00 1.92 3.01 1.05 1.88 2.16 9.19
8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)


172.95
157.18 154.00
147.14 143.83 141.32 139.51
133.43 127.74 127.07 125.00 119.99
109.61 109.31
77.34 77.02 76.71
68.59 68.01 66.80
56.41
48.35 47.91 47.23 40.99 40.16
32.77 29.88 24.78
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

6.2 General procedure for peptide nosylation

Peptidyl TentaGel® R resins (initial resin loading 0.18-0.2mmol/g) **10** or **13** were available from previous synthesis (fluorenylmethoxycarbonyl (Fmoc) SPPS strategy). Test cleavage from the resin and analysis via LC-MS shows the expected peptide masses (**Figure 17,** calculated: 2438.6 Da; observed: 1219.7 $[M+2H]^{+2}$ for peptide **10** and **Figure 18,** calculated: 1206.6 Da; observed: 1207.5 $[M+H]^{+}$, 604.2 $[M+2H]^{+2}$ for peptide **13**).

Prior to use, the resin was swollen in NMP for 2h. Fmoc deprotection of the N-terminus was done by incubation of the resin with 20% piperidine in NMP (v/v) consecutively for 3 and 7 minutes and washing with NMP (4x).

The resin was then incubated for 40 min with a 0.19 mM solution o-Nosyl chloride (5equ.) and collidine (10equ.) in NMP. Afterwards, the resin was washed with NMP four times. Test cleavage from the resin and analysis via LC-MS shows the expected masses (**Figure 20** of calculated: 2401.5 Da; observed: 1201.2 $[M+2H]^{+2}$ for peptide **11** and **Figure 21,** observed: 1169.5 Da; observed: 1170.4 $[M+H]^{+}$, 585.7 $[M+2H]^{+2}$ for peptide **14**).

6.3 On resin Fukuyama-Mitsunobu alkylation

For alkylation of the nosylated peptides **12** and **15**, two slightly different approaches in regards of order of reagent addition and cleavage were used.

Peptide **12**:

The solvent of peptidyl resin **12** (aprox. 25mg ≙ 3.2μmol) was changed from NMP to anhydrous THF, by washing 4 times with anhydrous THF. PPh_3 (5eq, 16μmol, 4.2mg) was dissolved in anhydrous THF (37.5 mM) in a round-bottomed flask. DIAD (5 eq, 16μmol, 3.14μl, 3.2mg) was added at 0°C and the solution was stirred for 10min at that temperature. After addition of the auxiliary (**9**, 5eq, 16μmol, 11mg) to this mixture, it was then added to the resin. After 1h of constant shaking, the supernatant was removed and the resin washed with THF (4x) and NMP (4x).

Peptide **15**:

Thoroughly dried peptidyl resin **15** (aprox. 25mg ≙ 3.2μmol) (washed sequentially with anhydrous MeOH, THF and DCM and dried at high vacuum for 6 days) was swollen in anhydrous THF for 18h. After washing with anhydrous THF, a solution of PPh_3 (5eq, 16μmol, 4.2mg) and the auxiliary (**9**, 5eq, 16μmol, 11mg) in anhydrous THF (37.5 mM) was added.

Solvation of the reagents was ensured by shaking of the reaction vessel, diisopropyl azo dicarboxylate (DIAD, 5 eq, 16μmol, 3.14μl, 3.2mg) was added, and the reaction solution mixed vigorously. After 1.5h of constant shaking, the reaction was quenched by washing the resin with THF (4x) and NMP (4x).

Cleavage was done using TFA/TIS/H_2O/DMS (90:2.5:2.5:5) for peptide **12** and TFA/TIS/H_2O (92.5:5:2.5) for peptide **15** at r.t. for 2 hours. Addition of cold Et_2O led to precipitation of the product. By centrifugation, removal of the supernatant and washing with cold Et_2O, the peptide was isolated, dried and afterwards dissolved in ACN/H_2O 50:50 for further analysis by LC/MS.

7 Abbreviations

aa	amino acid
Acm	acetamidomethyl
ACN	Acetonitrile
ATR	attenuated total reflection
AUX	Auxiliary
DCM	dichloromethane
DIAD	diisopropyl azodicarboxylate
DIEA	N,N-diisopropylethylamine
DMF	N,N-dimehtylformamide
DMS	dimethylsulfide
DMSO	dimethylsulfoxide
ECL	expressed chemical ligation
Fmoc	Fluorenylmethyloxycarbonyl
GalNAc	N-acetylgalactosamine
HMDST	bis(trimethylsilyl)sulfide
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrum
LC-MS	liquid chromatography-mass spectrometry
NCL	native chemical ligation
NMP	N-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
Ns/nosyl	here: 2-/o-nitrobenzenesulfonyl
OST	Oligosaccharyltransferase
PEG	polyethylene glycol
PTM	Posttranslational modification
SPPS	solid phase peptide synthesis
TCEP	tris-(2-carboxyethyl) phosphine
TEA	triethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIS	triisopropylsilane
TLC	thin layer chromatography

8 Bibliography

[1] I. Human Genome Sequencing Consortium, *Nature* **2004**, *431*, 931–945.

[2] I. Ezkurdia, D. Juan, J. M. Rodriguez, A. Frankish, M. Diekhans, J. Harrow, J. Vazquez, A. Valencia, M. L. Tress, *Hum. Mol. Genet.* **2014**, *23*, 5866–5878.

[3] A. Bairoch, R. Apweiler, *Nucleic Acids Res.* **2000**, *28*, 45–48.

[4] P. A. C. Cloos, S. Christgau, *Matrix Biol.* **2002**, *21*, 39–52.

[5] C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, *Angew. Chem., Int. Ed.* **2005**, *44*, 7342–7372.

[6] R. G. Spiro, *Glycobiology* **2002**, *12*, 43R–56R.

[7] R. K. Upreti, M. Kumar, V. Shankar, *Proteomics* **2003**, *3*, 363–379.

[8] D. F. Zielinska, F. Gnad, J. R. Wiśniewski, M. Mann, *Cell* **2010**, *141*, 897–907.

[9] D. J. Kelleher, R. Gilmore, *Glycobiology* **2006**, *16*, 47–62.

[10] J. Roth, *Chem. Rev.* **2002**, *102*, 285–303.

[11] A. Furmanek, J. Hofsteenge, *Acta Biochim. Pol.* **2000**, *47*, 781–789.

[12] Y. Ihara, Y. Inai, M. Ikezaki, I. L. Matsui, S. Manabe, Y. Ito, in *Glycoscience: Biology and Medicine*, Springer Japan, **2015**, pp. 1–8.

[13] J. Stepper, S. Shastri, T. S. Loo, J. C. Preston, P. Novak, P. Man, C. H. Moore, V. Havlicek, M. L. Patchett, G. E. Norris, *FEBS Lett.* **2011**, *585*, 645–650.

[14] T. J. Oman, J. M. Boettcher, H. Wang, X. N. Okalibe, W. A. van der Donk, *Nat. Chem. Biol.* **2011**, *7*, 78–80.

[15] P. Lafite, R. Daniellou, *Nat. Prod. Rep.* **2012**, *29*, 729–738.

[16] K. Mariño, J. Bones, J. J. Kattla, P. M. Rudd, *Nat. Chem. Biol.* **2010**, *6*, 713–723.

[17] K. W. Moremen, M. Tiemeyer, A. V Nairn, *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 448–462.

[18] A. Varki, *Glycobiology* **1993**, *3*, 97–130.

[19] H. H. Freeze, M. Aebi, *Curr. Opin. Struct. Biol.* **2005**, *15*, 490–498.

[20] K. Ohtsubo, J. D. Marth, *Cell* **2006**, *126*, 855–867.

[21] R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720.

[22] G. Walsh, R. Jefferis, *Nat. Biotechnol.* **2006**, *24*, 1241–1252.

[23] R. Haltiwanger, J. Lowe, *Annu. Rev. Biochem.* **2004**, *73*, 491–537.

[24] D. H. Dube, C. R. Bertozzi, *Nat. Rev. Drug Discovery* **2005**, *4*, 477–488.

[25] D. P. Gamblin, E. M. Scanlan, B. G. Davis, *Chem. Rev.* **2009**, *109*, 131–163.

[26] Z. Guo, N. Shao, *Med. Res. Rev.* **2005**, *25*, 655–678.

[27] C. Unverzagt, Y. Kajihara, *Chem. Soc. Rev.* **2013**, *42*, 4408–4420.

[28] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.

[29] L. R. Malins, R. J. Payne, *Aust. J. Chem.* **2015**, *68*, 521–537.

[30] Q. Wan, S. J. Danishefsky, *Angew. Chem., Int. Ed.* **2007**, *46*, 9248–9252.

[31] S. F. Loibl, Z. Harpaz, O. Seitz, *Angew. Chem., Int. Ed.* **2015**, *54*, 15055–15059.

[32] B. L. Pentelute, S. B. H. Kent, *Peptides* **2007**, 8805–8807.

[33] Z. Harpaz, S. Loibl, O. Seitz, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 1434–1437.

[34] P. Botti, M. R. Carrasco, S. B. H. Kent, *Tetrahedron Lett.* **2001**, *42*, 1831–1833.

[35] T. Kawakami, K. Akaji, S. Aimoto, *Org. Lett.* **2001**, *3*, 1403–1405.

[36] J. Offer, P. E. Dawson, *Org. Lett.* **1999**, *2*, 23–26.

[37] J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646.

[38] C. Bello, S. Wang, L. Meng, K. W. Moremen, C. F. W. Becker, *Angew. Chem., Int. Ed.* **2015**, *54*, 1–6.

[39] C. Chatterjee, R. K. McGinty, J. P. Pellois, T. W. Muir, *Angew. Chem., Int. Ed.* **2007**, *46*, 2814–2818.

[40] S. Tchertchian, O. Hartley, P. Botti, *J. Org. Chem.* **2004**, *69*, 9208–9214.

[41] D. Macmillan, D. W. Anderson, *Org. Lett.* **2004**, *6*, 4659–4662.

[42] C. Marinzi, S. J. Bark, J. Offer, P. E. Dawson, *Bioorg. Med. Chem.* **2001**, *9*, 2323–2328.

[43] B. Wu, J. Chen, J. D. Warren, G. Chen, Z. Hua, S. J. Danishefsky, *Angew. Chem., Int. Ed.* **2006**, *45*, 4116–4125.

[44] T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **2003**, *44*, 6059–6061.

[45] C. Marinzi, J. Offer, R. Longhi, P. E. Dawson, *Bioorg. Med. Chem.* **2004**, *12*, 2749–2757.

[46] C. Nadler, A. Nadler, C. Hansen, U. Diederichsen, *Eur. J. Org. Chem.* **2015**, *2015*, 3095–3102.

[47] C. Bello, K. Farbiarz, J. F. Möller, C. F. W. Becker, T. Schwientek, *Chem. Sci.* **2014**, *5*, 1634.

[48] L. Liu, C. S. Bennett, C.-H. Wong, *Chem. Commun.* **2006**, 21–33.

[49] D. K. Nielsen, A. G. Doyle, *Angew. Chem., Int. Ed.* **2011**, *50*, 6056–6059.

[50] R. D. Aher, B. S. Kumar, A. Sudalai, *Synlett* **2014**, *25*, 97–101.

[51] E. J. Corey, M. Chaykovsky, *J. Am. Chem. Soc.* **1965**, *87*, 1353–1364.

[52] Y. Jin, J. Li, L. Peng, C. Gao, *Chem. Commun.* **2015**, *51*, 15390–15393.

[53] E. Wünsch, L. Moroder, S. Romani, *Hoppe-Seyler's Z. Physiol. Chem.* **1982**, *363*, 1461–1464.

[54] T. Ohshima, Y. Hayashi, K. Agura, Y. Fujii, A. Yoshiyama, K. Mashima, *Chem. Commun.* **2012**, *48*, 5434.

[55] T. Fukuyama, C. K. Jow, M. Cheung, *Tetrahedron Lett.* **1995**, *36*, 6373–6374.

[56] L. Yang, K. Chiu, *Tetrahedron Lett.* **1997**, *38*, 7307–7310.

[57] O. Demmer, I. Dijkgraaf, M. Schottelius, H. J. Wester, H. Kessler, *Org. Lett.* **2008**, *10*, 2015–2018.

[58] R. D. Viirre, R. H. E. Hudson, *Org. Lett.* **2001**, *3*, 3931–3934.

[59] Y. Rew, M. Goodman, *J. Org. Chem.* **2002**, *67*, 8820–8826.

[60] S. Fletcher, *Org. Chem. Front.* **2015**, *2*, 739–752.

[61] T. Kan, T. Fukuyama, *Chem. Commun.* **2004**, 353–359.

[62] C. E. Barnes, M. W. Johnson, H. W. H. Lee, P. C. Myhre, *J. Org. Chem.* **1979**, *44*, 3925–3930.

[63] V. K. Aggarwal, J. Richardson, *Chem. Commun.* **2003**, 2644–2651.

[64] V. K. Aggarwal, C. L. Winn, *Acc. Chem. Res.* **2004**, *37*, 611–620.

[65] A. Piccinini, S. A. Kavanagh, P. B. Connon, S. J. Connon, *Org. Lett.* **2010**, *12*, 608–611.

[66] F. Charmantray, P. Dellis, V. Hélaine, S. Samreth, L. Hecquet, *Eur. J. Org. Chem.* **2006**, *2006*, 5526–5532.

[67] A. Degl'Innocenti, A. Capperucci, A. Cerreti, S. Pollicino, S. Scapecchi, I. Malesci, G. Castagnoli, *Synlett* **2005**, *20*, 3063–3066.

[68] G. Wang, L. Peng, Y. Zheng, Y. Gao, X. Wu, T. Ren, C. Gao, J. Han, *RSC Adv.* **2015**, *5*, 5674–5679.

[69] C. Bello, F. Kikul, C. F. W. Becker, *J. Pept. Sci.* **2015**, *21*, 201–207.