Abstract: The present invention relates to a compound comprising a protected or deprotected thiol moiety S enabling Native Chemical Ligation, a photolabile linker moiety L and at least one hydrophilic polymer moiety HP. Furthermore, the present invention relates to a method for production of a peptide conjugate P12 comprising a peptide moiety P1 and a second moiety P2 covalently bound with another via an amide bond, wherein the compound of the present invention employed for selectively forming the amide bond between P1 and P2.
A compound for Preparing Polypeptides

The present invention relates to a compound comprising a protected or deprotected thiol moiety S enabling Native Chemical Ligation, a photolabile linker moiety L and at least one hydrophilic polymer moiety HP. Furthermore, the present invention relates to a method for production of a peptide conjugate P12 comprising a peptide moiety P1 and a second moiety P2 covalently bound with another via an amide bond, wherein the compound of the present invention employed for selectively forming the amide bond between P1 and P2.

Today, in peptide chemistry, it is frequently intended to provide longer polypeptide chains and to conjugate peptides to non-proteinogenic carrier structures. Concomitantly, it is often desired to obtain peptides that are selectively modified by enzymatic means such as, e.g., glycosylated peptides. In particular glycosylation is known to play a key role in numerous biological and biochemical functions such as, e.g., protein folding, cellular differentiation, cell-cell communication, cell-matrix interaction and viral invasion. Post-translational modifications such as glycosylation may play a crucial role in different types of biochemical recognition processes responsible for growth, development, infection, immune response, cell adhesion, signal transduction processes, formation and manifestation of neoplasia, cancer, metastases and autoimmune diseases. Today, many peptides subjected to post-translational modifications such as glycopeptides are of great interest for pharmaceutical purposes. Herein, such peptide may exemplarily serve as antibiotic, hormone and cytokine (e.g., juvenile human growth hormone, CD4 or tissue plasminogen activator), vaccine, artificial extracellular matrix, artificial glycocalyx or a coating for implants. Further, numerous of such peptides are of great interest for research purposes.

In general two strategies for obtaining peptides are known in the art: genetically-based methods (e.g., heterologous expression) and synthetical approaches (e.g., solid phase peptide synthesis (SPPS)). In comparison to genetically-based methods, synthetical approaches enable full flexibility with respect to the incorporation of modifications and may provide higher purities. Therefore, often synthetical approaches are preferred or even required to obtain the desired pattern of modifications.
However, in particular the preparation of longer polypeptide chains with specific modifications by means of synthetical peptide synthesis is still hampered by insufficiency of the compounds known in the art. As the synthetical approaches known in the art show significantly decreasing yields upon increasing chain length of the polypeptide product, improved methods enabling site-specific conjugation are required.

So far only few methods for site-specific conjugation of peptides enabling the omittance of orthogonal protecting groups (e.g., Boc- or Fmoc-based strategies) and non-peptidic moieties included into the polypeptide chain (e.g., maleimidy moieties) are known in the art. Further the conjugation in an aqueous environment is often preferred. Therefore, one of the most preferred reactions for site-specific conjugation is Native Chemical Ligation (NCL) – a method applicable for site-specific conjugation of unprotected peptides in an aqueous environment. Common NCL enables the conjugation of peptides bearing N-terminal cysteinyl (Cys) moieties with other peptides bearing thioesters at their C-termini. This however bears significant drawbacks. First, the rather rare amino acid moiety cysteinyl (Cys) is inevitably incorporated into the polypeptide product. This is often undesired when producing nature-like sequences because many natural peptide strands are free of cysteine or comprise cysteine in a position not suitable for NCL. Further, the presence of a cysteinyl moiety bears drawbacks such as, e.g., the possibilities of oxidation and of formations of undesired intra- and/or intermolecular disulfide bridges. Several methods have therefore been developed to subsequently remove the cysteinyl moiety, but this requires even a further procedural step. Purification of the polypeptide product is typically considerably difficult and laborious.

Merely few synthetical strategies have been developed to improve NCL in a way of avoiding the need of a cysteinyl moiety. One of these approaches is the use of photo-cleavable auxiliary compounds that enable the replacement of cysteinyl by another amino acid moiety (Marinzi et al., 2001; Kawakami et al., 2003; Marinzi et al., 2004; Chatterjee et al., 2007; Offer, 2010; Pellois and Muir, 2005). This bears the advantage that cysteinyl moieties can be avoided and the auxiliary compounds can be easily cleaved off in a mild manner by irradiation of light.

The NCL-based methods known in the art however merely partly solve the problems the experimenter is faced with. In fact, only peptides that are well-soluble in aqueous buffers can be efficiency used in such NCL-based methods. In contrast, peptides which are purely soluble in aqueous environments can hardly be conjugated by NCL because one or both of the educts and/or intermediate products (i.e., the peptide moieties and/or the conjugate of the two educt peptide moieties conjugated with the auxiliary compound) may
aggregate. In particular, enzymatic modification of the peptides (e.g., enzymatic glycosylation) however requires well-soluble peptides in a rather pure aqueous environment such as a buffer, still excluding many peptides from such methods known in the art. Summarized, the solubility of peptides to be subjected to an NCL reaction is still a considerable issue, in particular when enzymatic modifications are intended to be included.

For enabling enzymatic modifications, it is known that the conjugation of hydrophilic polymer moieties directly to the polypeptides to be subjected to enzymatic modifications can improve the yields of said enzymatic reactions WO2012/139777, Bello et al., 2014).

This however requires additional steps. In one additional step, the polypeptide has to be conjugated to the hydrophilic polymer. In a second additional step, the hydrophilic polymer has to be cleaved off the polypeptide. Further, additional purification steps may be required for removing the hydrophilic polymer and cleaving reagents from the polypeptide. Should obtaining a longer polypeptide product be desired, an additional step of a conjugation reaction such as an NCL reaction (typically performed in another environment) is required and potentially even further purification steps. Furthermore, these methods will include an undesired cysteinyl moiety into the polypeptide product.

In the view of the above, there is still and unmet need for improved means for preparing longer polypeptide products in a less laborious and time-consuming manner, in particular when the peptides are also intended to be enzymatically modified (e.g., glycosylated) and/or bear pure solubility in aqueous environments. Concomitantly, it is desired to avoid the need of cysteinyl moieties in the polypeptide products. Therefore, there is still an unmet need for improved compounds that enable the omittance of using cysteinyl residues in the conjugation and concomitantly improve solubility and purification steps.

Surprisingly, a compound that bears superior properties with respect to processability has been identified. Synthesis of a peptide, enzymatic modification thereof, conjugation to another moiety (such as another peptide) are particularly easily performable, concomitantly with less and improved purification steps and improved solubility in aqueous environments. Thereby, the incorporation of cysteinyl residues is evitable.

In a first aspect, the present invention relates to a compound comprising:

(A) a thiol moiety S comprising:
    (A1) at least one protected or deprotected thiol group, and
(A2) at least one primary or secondary amino group, wherein said thiol moiety S enables Native Chemical Ligation (NCL);

(B) a photolabile linker moiety L; and

(C) at least one hydrophilic polymer moiety HP;

wherein the photolabile linker moiety L is covalently bound to the at least one hydrophilic polymer moiety HP and to the thiol moiety S, and wherein HP does not comprise primary amino groups or hydroxyl groups when the thiol moiety S is not covalently bound to a peptide moiety P1.

Generally, it may be understood that also a salt of such compound may be embraced by the present invention. The compound may be an auxiliary compound for preparing polypeptide products. As used throughout the present invention, it will be well understood that primary amino groups may, depending on the chemical environment of the compound, will occur in the form of -NH₂ or -NH₃⁺.

The feature that the hydrophilic polymer HP does not comprise primary amino groups or hydroxyl groups when the thiol moiety S is not covalently bound to a peptide moiety P1 may be understand in the broadest sense.

Therefore, in case the compound of the present invention is no peptide conjugate, i.e., is not covalently bound to a peptide moiety P1 (therefore, most typically, the thiol moiety S comprises at least one primary amino group (-NH₂ or -NH₃⁺)), the hydrophilic polymer HP is free of primary amino groups (-NH₂ or -NH₃⁺) and hydroxyl groups (-OH). In contrast, in case the compound of the present invention is a peptide conjugate, i.e., covalently bound to a peptide moiety P1 (wherein the thiol moiety S comprises only one amino group (a secondary amino group) that forms part of an amide bond to the peptide moiety P1), the hydrophilic polymer HP may optionally comprise one or more primary amino groups and/or one or more hydroxyl groups.

Summarized, particularly preferred, the compound of the present invention not conjugated to a peptide moiety P1 comprises only one (a single) primary amino group and no hydroxyl group, wherein the one primary amino group most preferably forms part of the thiol moiety S.
Accordingly, in one embodiment, the only primary amino group of the compound is the one comprised in the thiol moiety S. Then, the compound for preparing polypeptide products or a salt thereof comprises:

(A) a thiol moiety S comprising:

(A1) one protected or deprotected thiol group, and
(A2) one primary amino group (-NH₂ or -NH₃⁺),
wherein said thiol moiety S enables Native Chemical Ligation (NCL);

(B) a photolabile linker moiety L; and

(C) at least one hydrophilic polymer moiety HP;

wherein the photolabile linker moiety L is covalently bound to the at least one hydrophilic polymer moiety HP and to the thiol moiety S, and wherein HP does not comprise primary amino groups or hydroxyl groups.

In a more preferred embodiment, the compound for preparing polypeptide products or a salt thereof consists of:

(A) a thiol moiety S comprising or (essentially) consisting of:

(A1) one protected or deprotected thiol group, and
(A2) one primary amino group (-NH₂ or -NH₃⁺),
wherein said thiol moiety S enables Native Chemical Ligation (NCL);

(B) a photolabile linker moiety L; and

(C) a hydrophilic polymer moiety HP;

wherein the photolabile linker moiety L is covalently bound to the at least one hydrophilic polymer moiety HP and to the thiol moiety S, and wherein HP does not comprise primary amino groups or hydroxyl groups.

Alternatively, in another embodiment, the compound for preparing polypeptide products or a salt thereof comprises:

(A) a thiol moiety S comprising:

(A1) at least one protected or deprotected thiol group, and
(A2) one secondary amino group (-NH-) forming part of an amino acid moiety, in particular a glycyl (Gly) moiety (-NH-CH₂-CO-),
wherein said thiol moiety S enables Native Chemical Ligation (NCL);

(B) a photolabile linker moiety L; and

(C) at least one hydrophilic polymer moiety HP;

wherein the photolabile linker moiety L is covalently bound to the at least one hydrophilic polymer moiety HP and to the thiol moiety S, and wherein HP may or may not comprise one or more primary amino group(s) and/or one or more hydroxyl group(s).
In a more preferred embodiment, the compound for preparing polypeptide products or a salt thereof consists of:

(A) a thiol moiety $S$ comprising or (essentially) consisting of:
   (A1) at least one protected or deprotected thiol group, and
   (A2) one secondary amino group (-NH-) forming part of an amino acid moiety, in particular a glycyl (Gly) moiety (-NH-CH$_2$-CO-), wherein said thiol moiety $S$ enables Native Chemical Ligation (NCL);

(B) a photolabile linker moiety $L$; and

(C) a hydrophilic polymer moiety $HP$; wherein the photolabile linker moiety $L$ is covalently bound to the at least one hydrophilic polymer moiety $HP$ and to the thiol moiety $S$, and wherein $HP$ may or may not comprise one or more primary amino group(s) and/or one or more hydroxyl group(s).

Alternatively, in another embodiment, the compound for preparing polypeptide products or a salt thereof comprises:

(A) a thiol moiety $S$ comprising:
   (A1) at least one protected or deprotected thiol group, and
   (A2) one secondary amino group (-NH-) forming part of a carbamate group (-NH-CO-0-), preferably conjugated with a protecting group, in particular an Fmoc protecting group,
   wherein said thiol moiety $S$ enables Native Chemical Ligation (NCL);

(B) a photolabile linker moiety $L$; and

(C) at least one hydrophilic polymer moiety $HP$; wherein the photolabile linker moiety $L$ is covalently bound to the at least one hydrophilic polymer moiety $HP$ and to the thiol moiety $S$, and wherein $HP$ may or may not comprise one or more primary amino group(s) and/or one or more hydroxyl group(s).

In a more preferred embodiment, the compound for preparing polypeptide products or a salt thereof consists of:

(A) a thiol moiety $S$ comprising or (essentially) consisting of:
   (A1) a protected or deprotected thiol group, and
   (A2) a secondary amino group (-NH-) forming part of a carbamate group (-NH-CO-0-), preferably conjugated with a protecting group, in particular an Fmoc protecting group,
   wherein said thiol moiety $S$ enables Native Chemical Ligation (NCL);
(B) a photolabile linker moiety L; and
(C) a hydrophilic polymer moiety HP;
wherein the photolabile linker moiety L is covalently bound to the at least one hydrophilic polymer moiety HP and to the thiol moiety S, and
wherein HP may or may not comprise one or more primary amino group(s) and/or one or more hydroxyl group(s).

As shown in the examples, it was found that a first peptide moiety P1 conjugated with such compound could be very well conjugated to a second moiety P2 by means of Native Chemical Ligation (NCL). In particular, it could be shown that the peptide moiety P1 conjugated with such compound could be modified enzymatically (as exemplified: glycosylated) prior to optionally be conjugated to a second moiety P2 by means of NCL. Herein, all intermediate products are well-soluble in aqueous environments and can be easily purified in merely few steps. Thereby the procedure is less laborious.

As used throughout the present invention, the terms "conjugated with", "conjugated to" or "bound to" may be understood interchangeably in the broadest sense as the covalent linkage of two or more molecular moieties with another.

The term "polypeptide product" as used in the context of the present invention as a modified or unmodified polypeptide strand that is the intermediate or final product of the present invention.

In general, as used throughout the present invention, the terms "polypeptide" and "peptide" may be understood interchangeably. A (poly)peptide moiety may be understood as a peptide strand that is covalently bound.

A (poly)peptide may be understood in the broadest sense as a molecular structure comprising two or more amino acids that are conjugated with another via an amide bond. Preferably, a (poly)peptide is mainly composed of amino acids. In the context of (poly)peptides, an amide bond may also be designated as "peptide bond" or "peptidic bond". A (poly)peptide may be a linear, cyclic or branched peptide. A (poly)peptide may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, more than 10, more than 15, more than 20, more than 25, more than 30, more than 40, more than 50, more than 75, more than 100, more than 200 or even more than 500 amino acids. A (poly)peptide that bears more than one secondary structure element and/or a tertiary structure element may also be designated as "protein domain" or "protein". Preferably, the (poly)peptides in the context of the present invention are small (poly)peptides of less than 100 amino acids.
As used herein, a (poly)peptide preferably consists of a single amino acid strand. Alternatively, a (poly)peptide may also comprise two or even more strands covalently conjugated with another by any means such as e.g., by one or more disulfide bond(s).

Preferably, the (poly)peptide (essentially) consists of natural L-amino acids. However, the (poly)peptide may also comprise one or more non-natural amino acid(s) such as, e.g., D-amino acid(s), beta amino acid(s), methylated amino acid(s) (e.g., N-methylated amino acid(s)). The (poly)peptide may even only consist of non-natural amino acids. The (poly)peptide may also be a retro-inverso peptide, thus a (poly)peptide mainly comprising D-amino acids and a reverse amino acid sequence compared to the corresponding naturally occurring peptide consisting of L-amino acids. As used in the context of the amino acid sequence, the term "naturally occurring" may be understood as a sequence that has at least 70 %, preferably at least 80 %, more preferably at least 90 %, even more preferably at least 95 % and most preferably 100 % sequence identity with the amino acid sequence found in nature.

The amino acid moieties may be positively charged, negatively charged or neutral. Also a whole (poly)peptide may be positively charged, negatively charged or neutral. A (poly)peptide may be hydrophilic or hydrophobic. A peptide may be well-water soluble or poorly water-soluble. As used herein, the term "poorly water-soluble" refers to a peptide which solubility is less than 5 mg/ml, less than 4 mg/ml, less than 3 mg/ml, less than 2 mg/ml, less than 1 mg/ml, less than 0.5 mg/ml or even less than 0.25 mg/ml peptide per water.

It will be understood by a person skilled in the art that a (poly)peptide may further bear any counter ion(s) known in the art, such as e.g., chloride ion(s), acetate ion(s), carbonate ion(s), hydrocarbonate ion(s), sodium ion(s), potassium ion(s), magnesium ion(s), and any ion(s) of the cleavage solution (e.g., TFA ion(s), bromide ion(s), perchlorate ion(s), ammonium ion(s)). Further, a (poly)peptide may be covalently or non-covalently associated to traces of one or more scavenger(s), such as, e.g., triisopropylsilane (TIS), dithiothreitol (DTT), anisole, thioanisole or 1,2-ethanedithiol.

Optionally, a (poly)peptide of the present invention may further bear one or more modification(s). A (poly)peptide may be amidated or capped at its C-terminus (e.g., by a non-amino acid moiety such as, e.g., a bead and/or a polymer, in particular a hydrophilic polymer). As used herein, the terms "C-terminus", "C-terminal end", "carboxy terminus", "carboxyterminus" and "carboxyterminal end" may be understood interchangeably.
Further, the N-terminus may be capped (e.g., by an acetyl moiety, a methyl moiety, a pyroglutamyl moiety, a bead and/or a polymer, in particular a hydrophilic polymer). As used herein, the terms "N-terminus", "N-terminal end", "amino terminus" and "amino terminal end" may be understood interchangeably. The capping and/or amidation at one terminus or both termini may render the peptide more stable against exopeptidases. As used herein, preferably, the N-terminus of a (poly)peptide is capped by a hydrophilic polymer.

Optionally, one or more amino acid residue(s) of the peptide may be lipidated, phosphorylated, sulfated, cyclized, oxidized, reduced, decarboxylated, acetylated, amidated, deamidated, biotinylated or bound to one or more other small molecule(s) and/or terpene(s). Further, a (poly)peptide may or may not form one or more intramolecular disulfide bond(s) or one or more intermolecular disulfide bond(s).

Optionally, a (poly)peptide of the present invention may be labeled radioactively (e.g., by \(^{3}\text{H}, {^{32}\text{P}, {^{35}\text{S}}, {^{14}\text{C}}, {^{99}\text{mTc}}\) or lanthanides (e.g., \(^{64}\text{Gd}\)) or may be labelled with a spin label, such as one or more heavy isotopes, e.g., \(^{13}\text{C}\), detectable by Nuclear Magnetic Resonance (NMR).

Optionally, a (poly)peptide of the present invention may be labeled by one or more small molecule dye(s) (e.g., Cy dye(s) (e.g., Cy3, Cy5, Cy5.5, Cy7), Alexa dye/s (e.g., Alexa Fluor 488, Alexa Fluor 546, Alexa Fluor 647, Alexa Fluor 680, Alexa Fluor 750), VisEn dye(s) (e.g. VivoTag680, VivoTag750), S dye(s) (e.g., S0387), DyLight fluorophore(s) (e.g., DyLight 750, DyLight 800), IRDye(s) (e.g., IRDye 680, IRDye 800), fluorescein dye(s) (e.g., fluorescein, carboxyfluorescein, fluorescein isothiocyanate (FITC)), rhodamine dye(s) (e.g., rhodamine, tetramethylrhodamine (TAMRA)) or HOECHST dye(s)) or one or more quantum dot(s).

The term "thiol moiety" as used herein may be understood in the broadest sense as any moiety comprising at least one protected or deprotected thiol group (-SA, wherein S represents sulfur and A represents hydrogen, a protecting group or a peptide moiety) and at least one primary or secondary amino group (-NH, wherein N represents nitrogen, H represents hydrogen and B represents hydrogen, a peptide moiety P1, an amino acid moiety, or a moiety activating the amino group). Herein, a primary amino group represents -\(\text{NH}_2\) or -\(\text{NH}_3^+\), respectively. Herein, a secondary amino group represents any bivalent -NH moiety, which may optionally also form part of an amide group (-NH-CO-), a carbamate (-NH-CO-0-), an urate (-NH-CO-NH-), a hydrazine (NH-NH-) etc.
Accordingly, one preferred embodiment refers to a compound wherein B is hydrogen. Then, the compound of the present invention comprises a primary amino group or a salt thereof (-NH$_2$ or -NH$_3^+$, respectively). Then, the compound of the present invention may be conjugated with an amino acid moiety, with more than one amino acid moieties consecutively conjugated with another or with a peptide moiety P1.

Accordingly, in an alternative preferred embodiment of the present, B is an amino acid moiety. Particularly preferable in this context is an amino acid moiety that bears no chiral carbon. Thus, B may particularly preferably be a glycine (Gly) moiety. Then, the whole compound may be designated as Aux-Gly. Alternatively, B may also be another natural occurring amino acid moiety (e.g., alanine (Ala), cysteine (Cys), asparaginic acid (Asp), glutamic acid/glutamate (Glu), phenylalanine (Phe), histidine (His), isoleucine (Iso), lysine (Lys), leucine (Leu), methionine (Met), asparagine (Asn), proline (pro), glutamine (Gin), arginine (Arg), serine (Ser), threonine (Thr), valine (Val), tryptophane (Trp), tyrosine (Tyr)). Alternatively, B may also be a non-natural occurring amino acid moiety such as e.g. a D-amino acid, a beta amino acid, a gamma amino acid, a delta amino acid, an epsilon amino acid, or a methylated amino acid (e.g., an N-methylated amino acid). Alternatively, B may also be a natural or non-natural occurring amino acid moiety conjugated in a non-natural way such as e.g. lysine conjugated via its epsilon amino group. The amino acid residue may be protected or deprotected. The amino acid residue may optionally also bear one or more dye moieties, e.g., one of those exemplified before. Pre-loading with an amino acid moiety as described herein may be subsequently easily extended by conjugating further amino acid moieties thereto. Then the compound of the present invention, wherein B is an amino acid residue, may act similar to a pre-loaded resin.

Accordingly, in an alternative preferred embodiment of the present, B is a consecutive sequence of amino acid moieties also designated as peptide moiety P1 as described in more detail herein. The person skilled in the art will notice that, throughout the present invention, P1 can also be an amino acid residue.

According to a highly preferred embodiment, the thiol moiety S comprises one protected or deprotected thiol group and one primary or secondary amino group.

Preferably, the at least one thiol group is protected and/or the at least one amino group is a primary amino group. More preferably, the at least one thiol group is protected and the at least one amino group is a primary amino group.
The thiol moiety S according to the present invention enables Native Chemical Ligation (NCL). There are preferably 2 to 5 consecutive atoms between the thiol group and the amino group, more preferably 2 to 4 consecutive atoms, more preferably 2 or 3 consecutive atoms, in particular 2 consecutive atoms. Highly preferably, such consecutive atoms are bound with another and with the thiol group and the amino group via single bonds and are preferably not incorporated into a cyclic structure. Most preferably, such consecutive atoms located between the thiol group and the amino group are carbon atoms, in particular between the thiol group and the amino group are one or more -CH₂- group(s) and/or a -CH(R)- group, wherein R is any residue or binding site to a linker Y or a photolabile linker moiety L. When Y represents a spacer, it is preferably a spacer of not more than 5 carbon atoms, preferably a spacer selected from the group consisting of methylene, ethylene, propylene, butylene and pentylene. Preferably, the thiol group and the amino group are therefore linked with another via a flexible spacer of consecutive carbon atoms.

Highly preferably, the thiol moiety S bears the following structure:

\[ \text{AS-}(\text{CR}^a \#)^\#\text{R}^-\text{NHB}, \]

wherein each \( R^a \) is independently from another H or a halogen (e.g. F, Cl, Br or I), in particular is H.

wherein A represents hydrogen or a protecting group;

wherein B represents nitrogen, H represents hydrogen and B represents hydrogen, a peptide moiety P₁, an amino acid moiety, or a moiety activating the amino group; and

wherein \( \#^1 \) represents the binding site to the photolabile linker moiety L or to a spacer Y binding to the photolabile linker moiety L; and

wherein \( x \) represents an integer of 1 to 5, preferably is 1 or 2, in particular is 1.

Particularly preferably, the thiol moiety S has the following structure:

\[ \text{AS-CH}_2\text{-CH}^\#\text{-NHB}, \]

wherein the residues are defined as defined herein.

Preferably, the compound or a salt thereof bears the following basic structure:

\[ \text{[thiol moiety S] - Y - [photolabile linker moiety L] - X - [hydrophilic polymer moiety HP]}, \]
wherein X and Y may independently from another be a bond or a spacer moiety.

In a preferred embodiment, the compound has the following structure of formula (I):

\[
\begin{array}{c}
\text{A} \xrightarrow{X} \text{NHB} \\
\text{L} \xrightarrow{Z} \text{HP}
\end{array}
\]

wherein A represents a protecting group residue or hydrogen, preferably a protecting group residue selected from the group consisting of a i.e./f-butylsulfanyl residue, a 3-nitro-2-pyridylsulfanyl moiety, a benzyl residue, a 4-methylbenzyl residue, a 4-methoxybenzyl residue, a 2,4,6-trimethoxybenzyl residue, a diphenylmethyl residue, a trityl residue, a tert-butyl residue, an acetamidomethyl residue, a trimethylacetamidomethyl residue, a 9-fluorenylmethyl residue, an allyloxy carbonylaminomethyl residue, and a 9H-xanthen-9-yl residue, in particular a tert-butylsulfanyl residue;

wherein X represents a spacer moiety or a bond, preferably a substituted or unsubstituted C1-5-alkylen spacer moiety selected or a bond, in particular a -CH₂- spacer;

wherein B represents hydrogen, a peptide moiety P₁, an amino acid moiety, or a moiety activating the amino group, in particular hydrogen;

wherein Y represents a bond or a spacer moiety, in particular a bond;

wherein L represents the photolabile linker moiety L;

wherein Z represents a spacer moiety or a bond, in particular a spacer moiety; and

wherein HP represents a hydrophilic polymer moiety.

It will be further understood that, as used throughout the present invention, "S" in a chemical formula represents sulfur, "N" in a chemical formula represents nitrogen, "H" in a chemical formula represents hydrogen. The subscript indices such as "₂" represent the number of atoms preceding such index.

In a preferred embodiment, the compound has the following structure of formula (la):

...
wherein A represents a protecting group residue, in particular a ie/f-butylsulfanyl residue; wherein X represents a substituted or unsubstituted C_{1,5}-alkylen spacer moiety, in particular -CH_{2}-; wherein L represents the photolabile linker moiety L; wherein Y represents a bond or a substituted or unsubstituted C_{1,5}-alkylen spacer moiety, in particular a bond; wherein Z represents a spacer moiety; and wherein HP represents a hydrophilic polymer moiety.

More preferably, the compound has the following structure of formula (lb):

wherein A represents a protecting group residue, in particular a ie/f-butylsulfanyl residue; wherein X represents a substituted or unsubstituted C_{1,5}-alkylen spacer moiety; wherein L represents the photolabile linker moiety L; wherein Z represents a spacer moiety; and wherein HP represents a hydrophilic polymer moiety.

Even more preferably, the compound has the following structure of formula (lc):

wherein A represents a protecting group residue, in particular a ie/f-butylsulfanyl residue; wherein L represents the photolabile linker moiety L; wherein Z represents a spacer moiety; and wherein HP represents a hydrophilic polymer moiety.
The term "photolabile linker moiety L" may be any linker known in the art that is cleavable by irradiation of light. Preferably, the linker is cleavable by irradiation light in the UV-A range. The UV-A range may be in the range of from 250-400 nm. Preferred structures of the photolabile linker moiety L are laid out below.

In a preferred embodiment, the photolabile linker moiety L has the following structure of formula (II):

wherein \( \# \) represents the binding site to the thiol moiety S via Y, wherein Y represents a bond or a spacer moiety, in particular a bond;

wherein \( \#' \) represents the binding site to the at least one hydrophilic polymer moiety HP via Z, wherein Z represents a spacer moiety or a direct bond, in particular a spacer moiety; and

wherein \( R^1 \) may be any residue of up to 40 carbon atoms, preferably a residue selected from the group consisting of \(-OR^{11}, \) hydrogen, halogen, \(-R^{11}, -0-C(0)-R^{11}, -NR^{11}R^{12}, -SiR^{11}R^{12}R^{13}, -Ge R^{11}R^{12}R^{13}, -SR^{11}, -SOR^{11}, \) and \(-SO_2R^{11}, \) wherein each of \( R^{11}, R^{12} \) and \( R^{13} \) represents a residue selected from the group consisting of a substituted or unsubstituted linear, branched or cyclic \( C_{1-20} \) alkyl residue, hydrogen, \(-OH, -NH_2, \) a halogen, a substituted or unsubstituted linear, branched or cyclic \( C_{1-20} \) heteroalkyl, \( C_{2-20} \)-alkenyl or \( C_{2-20} \)-heteroalkenyl residue and a substituted or unsubstituted \( C_{6-20} \) aryl residue, \( C_{5-20} \) heteroaryl residue, \( C_{1,3} \) arylalkyl residue, \( C_{4,3i} \) heteroarylalkyl residue, \( C_{4,3i} \) alkylarylalkyl residue, a \( C_{7-32} \) heteroarylalkylalkyl residue, and a dye, more preferably wherein \( R^1 \) is selected from the group consisting of \(-OCH_3, -OCH_2CH_3, -O(\text{CH}_2)_2\text{CH}_3, -OCH(\text{CH}_3)_2 \) or hydrogen, in particular \(-OCH_3\).

A dye may be understood in the broadest sense as defined in the context of the polypeptide above.
Accordingly, more preferably, the photolabile linker moiety \( L \) has the following structure of formula (IIa):

\[
\begin{align*}
\text{CH}_2\text{O} & \quad \text{NO}_2 \\
# & \quad \text{\#'}
\end{align*}
\]

(IIa),

wherein \( \# \) represents the binding site to the thiol moiety \( S \); and wherein \( \#' \) represents the binding site to \( Z \).

Alternatively, the photolabile linker moiety \( L \) may be a linker based on a structure like \( 1\text{-}1\text{-}(2\text{-}nitrophenyl)ethyl, \) a \( 3\text{-}hydroxy\text{-}2\text{-}naphthalenemethanol \) ether, a \( 3\text{-}hydroxy\text{-}2\text{-}naphthalenemethanol \) ester or a \( 4,4\text{-}‘\text{dimethoxytrityl} \) ether.

The term "hydrophilic polymer" as used herein may be understood in the broadest sense as any polymeric molecule that has a hydrophilic surface that does (essentially) not comprise primary amino groups (-\( \text{NH}_2 \) or -\( \text{NH}_3^+ \)) or hydroxyl groups (-\( \text{OH} \)). A polymeric molecule comprises more than one monomer. The hydrophilic polymer may have a molecular mass of more than 500 Da, more than 1,000 Da, more than 1,500 Da, more than 2,000 Da, more than 5,000 Da or more than 10,000 Da. The hydrophilic polymer may be a soluble polymer, a gel-like polymeric matrix, a bead, a surface, a bead coating or the coating of a surface. Preferably, the hydrophilic polymer of the present invention is a soluble polymer.

In a preferred embodiment, the at least one hydrophilic polymer moiety \( HP \) is/are selected from the group consisting of

(a) polyethylene glycol (PEG) or derivatives thereof;
(b) polyethylene imine (PEI) or derivatives thereof;
(c) polyacrylic acid or derivatives thereof, preferably a polymer of methacrylic acid or derivative thereof, more preferably a polymer of hydroxypropyl methacrylate or hydroxyethyl methacrylate (HEMA) or derivatives thereof, in particular a polymer of N-(2-hydroxypropyl) methacrylamide (HPMA) or derivatives thereof;
(d) polysaccharide(s) or derivatives thereof,
(e) hydrophilic polypeptide(s)
(f) lipopolysaccharide(s); and/or
(g) conjugate(s) or blockpolymer(s) comprising two or more of the above,
In a more preferred embodiment, the hydrophilic polymer is PEG or derivatives thereof.

In a more preferred embodiment, the hydrophilic polymer is PEG or a derivative thereof of between 5 to 50 PEG units, even more preferably of between 10 to 40 PEG units, even more preferably of between 15 to 35 PEG units, in particular of between 20 to 30 PEG units.

Alternatively, the hydrophilic polymer is PEG or a derivative thereof of between 50 to 200 PEG units, preferably of 50 to 150 PEG units, in particular 50 to 100 PEG units.

In the context of the hydrophilic polymers, the term "derivative thereof" may be understood in the broadest sense as any substitution of said polymer. Exemplarily, the terminus of such hydrophilic polymer is capped by a protecting group by a dye. Exemplarily, the terminus of PEG may be capped by Fmoc, Boc or may be acylated (e.g., acetylated).

Alternatively, the hydrophilic polymer may be a bead having a hydrophilic surface or a bead having a coating comprising a hydrophilic polymer, in particular wherein the bead is a micro- or a nanobead. Alternatively, the hydrophilic polymer may be a solid material or a surface coating comprising a hydrophilic polymer. In this context, the hydrophilic polymer may optionally be the coating of a peptide microarray, the coating of a column matrix, in particular an affinity column matrix or the coating of an implant.

The hydrophilic polymer of the present invention may be conjugated with the peptide of the present invention via any functional group of the peptide. The hydrophilic polymer may be conjugated with the N-terminus of the peptide, an amino acid residue side chain, or the C-terminus of the peptide.

Depending on the polymer and the functional group of the peptide it is conjugated with, any conjugation strategy may be used, such as, e.g., formation of an ester bond, formation of an amide bond, formation of a thioester bond, formation of an ether bond, formation of a thioether bond or formation of a disulfide bond.

Preferably, the hydrophilic polymer is conjugated with the N- or the C-terminus of the peptide, more preferably the hydrophilic polymer is conjugated with the N- or the C-terminus of the peptide via the formation of an amide bond, more preferably the hydrophilic polymer is conjugated with the N-terminus of the peptide via the formation of an amide bond.
Such hydrophilic polymer may render the compound and its conjugates with peptide moieties more soluble in aqueous buffers. Further such conjugates may bear a lower tendency to interact with surfaces (e.g., plastic surfaces, glass surface and/or surfaces of column matrices) than the corresponding compound, peptide and/or conjugate not conjugated with a hydrophilic polymer. As used herein, interaction may be understood in the broadest sense as the stickiness and/or retention ability of the auxiliary compound, peptide and/or conjugate with or without being conjugated to at least one hydrophilic polymer when attached to a surface.

The hydrophilic polymer moieties are conjugated to the compound of the present invention by any means, in particular via the spacer Z. Such spacer Z may have any bivalent structure that does neither disturb the NCL not the photo-induced cleavage of the compound.

In a preferred embodiment, the spacer Z has the following structure of formula (III)

\[
\text{"-(CH}_2\text{)_nZ'-}" \quad \text{(III),}
\]

wherein " represents the binding site to the photolabile linker moiety L or to the hydrophilic polymer moiety HP, in particular to the photolabile linker moiety L;

wherein "" represents the binding site to the hydrophilic polymer moiety HP or to the photolabile linker moiety L, in particular to the hydrophilic polymer moiety HP;

wherein n represents an integer from 1 to 10, preferably wherein n is 1, 2 or 3, in particular 3;

wherein Z' represents a bivalent group or a bond, preferably wherein Z' is selected from the group consisting of -CO-NH-(CH\_2\text{)}\_m-NH-CO-, -NH-CO-, -NH-(CH\_2\text{)}\_m-NH-CO-, -NH-CO-(CH\_2\text{)}\_m-NH-CO-, -CO-NH-, -0-CO-, -CO-0-, -NH-CO-NH-, -0-CO-NH-, -NH-CO-0-, -0-CO-0-, -CO-NH-(CH\_2\text{)}\_m-NH-CO-, -NH-CO-, -NH-(CH\_2\text{)}\_m-0-CO-, -NH-(CH\_2\text{)}\_m-0-CO-, -0-(CH\_2\text{)}\_m-0-CO-, -0-(CH\_2\text{)}\_m-0-CO-, -O-(CH\_2\text{)}\_m-0-CO-, and -0-(CH\_2\text{)}\_m-0-CO-; and

wherein m represents an integer from 1 to 10, preferably wherein m is 1, 2 or 3, in particular 2.
In a more preferred embodiment, in the structure of formula (III), $Z'$ is selected from the group consisting of $-\text{CO-NH-}(\text{CH}_2)_m\text{-NH-CO-}$, $-\text{NH-CO-}$, and $\text{NH-(CH}_2)_m\text{-NH-CO-}$, and the other residues are selected as laid out above.

In an even more preferred embodiment, in the structure of formula (III), $Z'$ is $-\text{CO-NH-}(\text{CH}_2)_m\text{-NH-CO-}$ and the other residues are selected as laid out above.

Highly preferably, the spacer $Z$ has the following structure of formula (III),

wherein $#"$ represents the binding site to the photolabile linker moiety $L$;

wherein $#"'$ represents the binding site to the hydrophilic polymer moiety $\text{HP}$;

wherein $n$ is 1, 2 or 3, in particular 3;

wherein $Z'$ is selected from the group consisting of $-\text{CO-NH-}(\text{CH}_2)_m\text{-NH-CO-}$, $-\text{NH-CO-}$, and $\text{NH-(CH}_2)_m\text{-NH-CO-}$; and

wherein $m$ is 1, 2 or 3, in particular 2.

Particularly preferably, the spacer $Z$ has the following structure of formula (IIia)

$$#"-(\text{CH}_2)_3\text{-CO-NH-}(\text{CH}_2)_2\text{-NH-CO-}#"'$$ (IIia)

wherein $#"$ represents the binding site to the photolabile linker moiety $L$; and

wherein $#"'$ represents the binding site to the hydrophilic polymer moiety $\text{HP}$;

In total, the compound preferably has the following structure

![Chemical Structure](image)

wherein the residues $A$, $B$, $Y$, $Z$ and $R_1$ as well as $\text{HP}$ is defined as laid out above.

In a preferred embodiment, the compound has the following structure of formula (IV):
wherein \( R \) represents a linear or branched \( C_{1-20} \) alkoxy moiety or hydrogen, preferably \(-\text{OCH}_3\), hydrogen, \(-\text{OCH}_2\text{CH}_3\), \(-0(\text{CH}_2)_2\text{CH}_3\), or \(-\text{OCH}(\text{CH}_3)_2\), in particular \(-\text{OCH}_3\);

wherein \( Z \) represents a direct bond or a spacer moiety, in particular a spacer moiety as defined above;

wherein \( A \) represents a protecting group residue or hydrogen; and

wherein \( HP \) represents a hydrophilic polymer moiety, in particularly such as defined above.

In a more preferred embodiment, the compound bears one of the following structures (V)-(VII):

wherein \( R^1 \) represents a linear or branched \( C_{1-20} \) alkoxy moiety or hydrogen, preferably \(-\text{OCH}_3\), hydrogen, \(-\text{OCH}_2\text{CH}_3\), \(-0(\text{CH}_2)_2\text{CH}_3\), or \(-\text{OCH}(\text{CH}_3)_2\), in particular \(-\text{OCH}_3\);

wherein \( n \) represents an integer from 1 to 10, preferably wherein \( n \) is 1, 2 or 3, in particular 3; and

wherein \( m \) represents an integer from 1 to 10, preferably wherein \( m \) is 1, 2 or 3, in particular 2.

More preferably, the compound bears one of the following structures (Va)-(Vlla):
wherein PEG represents a polyethylene glycol moiety;
wherein y represents an integer from 5-50, preferably from 10-40, more preferably from 15-35, in particular 20-30;
wherein E represents a protecting group conjugated with the PEG via an amino or hydroxyl group, in particular Fmoc;
wherein R represents a linear or branched C12 alkoxo moiety or hydrogen, preferably -OCH3, hydrogen, -OCH2CH3, -O(CH2)2CH3, or -OCH(CH3)2, in particular -OCH3;
wherein n represents an integer from 1 to 10, preferably when n is 1, 2 or 3, in particular 3; and
wherein m represents an integer from 1 to 10, preferably when m is 1, 2 or 3, in particular 2.

As mentioned above, the amine moiety comprised in the thiol moiety S of the compound may also form (together with the amino group it is bound to) an amino acid moiety, optionally bound to a peptide moiety P1.

Particularly preferably, the compound of the present invention bears a structure of the following formula (XIV):
Alternatively, the compound of the present invention may bear a structure of the following formulae (XV) or (XVI):

\[
\text{H}_2\text{N} \begin{array}{c} \text{SS} \text{Bu} \\ \text{NO}_2 \\ \text{O} \end{array} \begin{array}{c} \text{CHR} \\ \text{PEG}_{27} \\ \text{N} \end{array} \text{(XV)}
\]

\[
\text{H}_2\text{N} \begin{array}{c} \text{SS} \text{Bu} \\ \text{NO}_2 \\ \text{O} \end{array} \begin{array}{c} \text{R}^{1} \\ \text{CH}_{n} \text{CHR}^{2} \text{CO} \text{CHR}^{2} \text{CO} \text{PEG}_{27} \\ \text{N} \end{array} \text{(XVI)},
\]

wherein \(R^{1}\) is defined as above, preferably is an alkoxy residue, in particular a methoxy residue; and
wherein \((\text{PEG})_{27}\) may have an unbound terminus or a terminus covered by a protecting group such as Fmoc.

Alternatively, the compound of the present invention may bear a structure of the following formulae (XVII) or (XIX):

\[
\text{H}_2\text{N} \begin{array}{c} \text{SS} \text{Bu} \\ \text{NO}_2 \\ \text{O} \end{array} \begin{array}{c} \text{R}^{1} \\ \text{CH}_{n} \text{CHR}^{2} \text{CO} \text{CHR}^{2} \text{CO} \text{PEG}_{27} \\ \text{N} \end{array} \text{(XVII)}
\]

\[
\text{H}_2\text{N} \begin{array}{c} \text{SS} \text{Bu} \\ \text{NO}_2 \\ \text{O} \end{array} \begin{array}{c} \text{R}^{1} \\ \text{CH}_{n} \text{CHR}^{2} \text{CO} \text{CHR}^{2} \text{CO} \text{PEG}_{27} \\ \text{N} \end{array} \text{(XVIII)}
\]

\[
\text{H}_2\text{N} \begin{array}{c} \text{SS} \text{Bu} \\ \text{NO}_2 \\ \text{O} \end{array} \begin{array}{c} \text{R}^{1} \\ \text{CH}_{n} \text{CHR}^{2} \text{CO} \text{CHR}^{2} \text{CO} \text{PEG}_{27} \\ \text{N} \end{array} \text{(XIX)},
\]

wherein \(R^{1}\), \(n\) and \(m\) are each independently from another is defined as above, preferably wherein \(R^{1}\) is an alkoxy residue, in particular a methoxy residue; and
wherein \((\text{PEG})_{27}\) may have an unbound terminus or a terminus covered by a protection group such as Fmoc.

In a preferred embodiment, in a compound of formula (I), \(B\) represents a peptide moiety \(P_{1}\), preferably a peptide moiety \(P_{1}\) characterized in that:

(a) it bears an N-terminal moiety of the structure of formula (VIII):

\[
\#^{\psi} \text{CHR}^{2} \text{CO} \#^{\psi} \text{(VIII)},
\]

wherein \#^{\psi}\ represents the binding site to the compound of the present invention;
wherein "#" represents the binding site to the remaining moieties of peptide moiety P1 (rest of P1); and

wherein R² represents hydrogen or a substituted or unsubstituted C₅₋₅-alkyl residue, preferably hydrogen, CH₃,-CH₂-OH or -CHOH-CH₃, in particular hydrogen

(b) it comprises less than 150 amino acid moieties, preferably less than 100 amino acid moieties, more preferably less than 50 amino acid moieties, even more preferably less than 40 amino acid moieties, even more preferably less than 35 amino acid moieties, more preferably less than 30 amino acid moieties, in particular less than 25 amino acid moieties; and/or

(c) it bears one or more non-peptidic modification(s), in particular wherein said peptide moiety P1 is glycosylated.

Preferably, the peptide moiety P1 bears a molecular mass of less than 10 kDa, more preferably less than 5 kDa, even more preferably less than 4 kDa, and most preferably less than 3 kDa.

Preferably, the N-terminal amino acid moiety of the peptide moiety P1 is a glycyl (Gly) moiety.

Exemplarily, the peptide moiety P1 bears the following sequence:

GVTSAPDTRPAPGSTAPPAH  (SEQ ID NO: 1)

Alternatively, also peptides moieties P1 of the peptide sequences disclosed in WO 2012/13977 to which an N-terminal G has been added may be used.

As used herein, the rest of P1 is peptide moiety P1 without its N-terminal amino acid moiety (which is preferably a glycyl (Gly) moiety).

More preferably, the compound has a structure of formula (lb):
wherein \( R^2 \) represents hydrogen or a substituted or unsubstituted C\(_{1-5}\)-alkyl residue, preferably hydrogen, CH\(_3\), -CH\(_2\)-OH or -CHOH-CH\(_3\), in particular hydrogen; and wherein the other residues A, L, Y, Z as well as P1 and HP are defined as above.

Accordingly, the compound may preferably have the structure of formula (lc):

\[
\begin{align*}
\text{A} & \overset{\text{NH-CHR}^2-\text{CO}}{\text{L}} - \text{Z} - \text{HP} \\
\end{align*}
\]

(lc)

wherein \( R^2 \) represents hydrogen or a substituted or unsubstituted C\(_{1-5}\)-alkyl residue, preferably hydrogen, CH\(_3\), -CH\(_2\)-OH or -CHOH-CH\(_3\), in particular hydrogen; and wherein the other residues A, L, Z as well as P1 and HP are defined as above.

Therefore, the structure may be exemplarily as follows:

\[
\begin{align*}
P2 & \overset{\text{SNH-CHR}^2-\text{CO}}{\text{Y}} - \text{L} - \text{Z} - \text{HP} \\
\end{align*}
\]

(ld) or

\[
\begin{align*}
P2 & \overset{\text{SNH-CHR}^2-\text{CO}}{\text{Y}} - \text{L} - \text{Z} - \text{HP} \\
\end{align*}
\]

(le).

Then, the molecule may intramolecularly rearrange, in particular when it is subjected to an aqueous environment:
In a more preferred embodiment, the compound has the following structure of formula (IX):

wherein \( R \) represents a linear or branched \( C_{1,3} \) alkoxy moiety or hydrogen, preferably \(-OCH_3\), hydrogen, \(-OCH_2CH_3\), \(-0(CH_2)_2CH_3\), or \(-OCH(CH_3)_2\), in particular \(-OCH_3\);

wherein \( Z \) represents a direct bond or a spacer moiety, in particular a spacer moiety as defined above;

wherein \( A \) represents a protecting group residue or hydrogen;

wherein \( HP \) represents a hydrophilic polymer moiety; and

wherein the rest of \( P1 \) represents the peptide moiety \( P1 \) without its N-terminal moiety \(-CHR_2-CO-\).

When \( A \) herein represents a peptide moiety, the structure may be as follows:
Then, the molecule may be restructured in an aqueous environment into:

Subsequently, the polypeptide conjugate P12 (i.e., P1-P2) may be cleaved off.

In a more preferred embodiment, the compound bears one of the following structures (X)-(XII):

wherein \( n \) represents an integer from 1 to 10, preferably wherein \( n \) is 1, 2 or 3, in particular 3;

wherein \( m \) represents an integer from 1 to 10, preferably wherein \( m \) is 1, 2 or 3, in particular 2;
and wherein the other residues A and R as well as HP and P₁ are defined as above.

When A is a peptide moiety, the structure may be as follows:

![Diagram](Xa) (Xla) (Xlla),

wherein the residues are defined as above.

Then, these molecules may be restructured in an aqueous environment into:

![Diagram](Xb) (Xlb) (Xllb),

wherein the residues are defined as above.

Subsequently, the polypeptide conjugate P₁₂ may be cleaved off. The structure is depicted above.

As mentioned above, the compound of the present invention may be used for the production of a polypeptide product that is a polypeptide conjugate P₁₂ comprising a peptide moiety P₁ and a second moiety P₂.

Therefore, in a further aspect, the present invention relates to a method for the production of a peptide conjugate P₁₂ comprising a peptide moiety P₁ and a second moiety P₂ covalently bound with another via an amide bond, said method comprising the steps of:

(i) providing a compound of the present invention, preferably wherein the thiol group of the thiol moiety S is protected and the amino group of the thiol moiety S is hydrogen;
(ii) adding a peptide moiety \( P_1 \) to the amino group covalently bound to the thiol moiety \( S \) of said compound of step (i) thereby forming a peptide conjugate \( PC_1 \), preferably wherein said peptide conjugate \( PC_1 \) bears the structure of formula (XIII):

\[
\begin{align*}
\text{AS}^X \text{NH} \text{P1} \\
\text{Y} \\
\text{L}^Z \text{HP}
\end{align*}
\]

wherein \( A, X, Y, L, Z, P_1 \) and \( HP \) are defined as above;

(iii) cleaving off the protecting group from the thiol group of the thiol moiety \( S \) the conjugate \( PC_1 \) so that said thiol moiety \( S \) bears an unprotected thiol group; and

(iv) conjugating a second moiety \( P_2 \) comprising a thioester with said peptide conjugate \( PC_1 \) obtained from step (iii) by means of Native Chemical Ligation, thereby forming a peptide conjugate \( PC_{12} \).

This method is further exemplified in the example section below. The outcome of step (iv) may be the a peptide conjugate \( PC_{12} \) having a structure according to formula (XIIIb):

\[
\begin{align*}
P_2 \text{S} \text{NH} \text{P1} \\
\text{Y} \\
\text{L}^Z \text{HP}
\end{align*}
\]

that may rearrange in an aqueous environment into a peptide conjugate \( PC_{12} \) having a structure according to formula (XIIIc):
Step (i) of providing a compound is shown in the example section below. Exemplarily, the compound may be produced and/or stored in a form wherein the amino group of the thiol moiety is protected by a protecting group. Then, this moiety is cleaved off in step (i). The compound may then optionally be purified and optionally be dissolved in a suitable solvent. Preferably, herein the compound has a structure of formula (I) or (Ia), (IV), (V), (VI), (VII), (Va), (Vb) or (VIIb), wherein B is H and A is a protecting group.

This means, in general terms, that the method preferably comprises the steps of:

(i) providing the compound of a structure of formula (Ia):

wherein A represents a protecting group residue, in particular a ie/f-butylsulfanyl residue;
wherein X represents a substituted or unsubstituted C_{1-5}-alkylen spacer moiety, in particular -CH$_2$-
wherein L represents the photolabile linker moiety L;
wherein Y represents a bond or a substituted or unsubstituted C_{1-5}-alkylen spacer moiety, in particular a bond;
wherein Z represents a spacer moiety; and
wherein HP represents a hydrophilic polymer moiety,
in particular wherein A, X, Y, L, Z, P1 and HP are defined as above;
(ii) adding a peptide moiety $P_1$ to the amino group covalently bound to the thiol moiety $S(-\text{NH}_2)$ of said compound of step (i) thereby forming a peptide conjugate $PC_1$, having the structure of formula (XIII):

$$\text{AS} \xrightarrow{X} \text{NH} \xrightarrow{Y} P_1 \xrightarrow{L-Z} \text{HP}$$

(XIII),

wherein $A, X, Y, L, Z, P_1$ and $\text{HP}$ are defined as in step (i);

(iii) cleaving off the protecting group from the thiol group of the thiol moiety $S$ the conjugate $PC_1$ so that said thiol moiety $S$ bears an unprotected thiol group $(-\text{SH})$ obtaining the compound of formula (XI$_{\text{Ma}}$):

$$\text{HS} \xrightarrow{X} \text{NH} \xrightarrow{Y} P_1 \xrightarrow{L-Z} \text{HP}$$

(XI$_{\text{Ma}}$);

(iv) conjugating a second moiety $P_2$ comprising an thioester with said peptide conjugate $PC_1$ obtained from step (iii) by means of Native Chemical Ligation, thereby forming a peptide conjugate $PC_{12}$ of formula (XI$_{\text{lie}}$):

$$\text{HS} \xrightarrow{Y} P_1 \xrightarrow{L-Z} \text{HP}$$

(XI$_{\text{lie}}$); and optionally

(v) irradiating the compound comprising the peptide conjugate $P_{12}$ obtained from step (iv) thereby cleaving said peptide conjugate $P_{12}$ off the photolabile linker $L$ and the hydrophilic polymer moiety $\text{HP}$ obtaining a polypeptide conjugate $P_{12}$ ($P_2-P_1$); and optionally

(vi) purifying the peptide conjugate $P_{12}$ obtained from step (iv) or (v).
Step (ii) of adding the peptide moiety may be performed by any means known in the art. This step is described further below.

Step (iii) of cleaving off the protecting group from the thiol group may be performed by any means known for that purpose in the art. The person skilled in the art will know how to adapt such cleaving step to the protecting group used. Exemplarily, when the thiol group of the thiol moiety S is protected by StBu (ie/f-butylsulfanyl), treatment with TCEP (tris(2-carboxyethyl)phosphine) may be used for deprotection. This is further exemplified in the example section below.

Step (iv) of conjugating a second moiety P2 to the peptide conjugate PC1 obtained from step (iii) is performed by Native Chemical Ligation. This is typically performed in an aqueous environment (e.g., in a buffer of pH 7-8). Exemplarily a NaPi buffer (sodium phosphate buffer) of pH 7.5 may be used at a temperature of 20-40°C. This exemplified further in the example section below.

The result us the method is the polypeptide conjugate P12.

As mentioned above, the peptide may also be modified enzymatically.

Therefore, in a preferred embodiment, the method further comprises the step (iiA) of modifying the peptide moiety P1 comprised in the compound obtained from step (ii) by means of one or more enzyme(s) prior to subjecting it to step (iii) or (iv).

As enzymes typically work best in an aqueous environment, the peptide conjugate PC1 may be provided in an aqueous buffer and admixed with a suitable amount of an enzyme and optionally cofactor(s) and educts and incubated in this mixture.

In a more preferred embodiment, the one or more enzyme(s) is/are glycosyltransferase(s) and said step of modifying the peptide moiety P1 is glycosylating.

As shown in the examples, the peptide conjugate PC1 may be further subjected to one or more glycosylating step(s) by means of one or more glycosyltransferase(s). Then a glycosylated peptide (glycopeptide) is obtained.

As used herein, the terms "glycosylated peptide" and "glycopeptide" may be understood interchangeably in the broadest sense as any peptide that comprises one or more
carbohydrate moiety/moieties. Preferably, the glycosylated peptide is a peptide that is glycosylated in a site-specific and/or stereospecific manner, more preferably in a site-specific and stereospecific manner.

As used herein, the terms "carbohydrate moiety", " saccharide moiety " and " sugar moiety " may be understood interchangeably in the broadest sense as a molecular residue of any carbohydrate known in the art. A carbohydrate moiety may comprise of one or more monosaccharide unit(s).

A monosaccharide unit may be any monosaccharide unit known in the art, such as, e.g., a triose (e.g., an aldotriose (e.g. glyceraldehyde) or a ketotriose (e.g., dihydroxyacetone)), a tetrose (e.g., an aldotetrose (e.g., erythrose, threose) or a ketotetrose (e.g., erythulose)), a pentose (e.g., an aldopentose (e.g., ribose, deoxyribose, arabinose, xylose, lyxose) or a ketopentose (e.g., ribulose, xylulose)), a hexose (e.g., an aldohexose (e.g., allose, altrose, glucose, mannose, gulose, idose, galactose, talose), other hexoses (e.g., glucuronic acid, galacturonic acid, N-acetylgalactosamine/N-acetylchitosamine, N-acetylgalactosamine/N-acetylchondrosamine, N-acetylgalactosamine/N-acetylchondrosamine, glucosamine/chitosamine, fucose, desoxygalactose, rhamnose, deoxymannose, chinovose, deoxyglucose) or a ketohexose (e.g., fructose)) or a higher monosaccharide unit such as, e.g., heptulose, sedoheptulose, desoxymanno-octulosonic acid/ketodesoxyoctonic acid, sialic acid/ N-acetylneuraminic acid. Each monosaccharide unit may be a D- or an L-saccharide and may form a linear or cyclic (e.g., pyranose or furanose) structure. Optionally, one or more of the functional group(s) of a monosaccharide unit may be replaced by other residue(s) such as, e.g., amino group(s), carboxylic group(s), hydroxyl group(s) carbonyl group(s), hydrogen atom(s), halogen atom(s), alkyl group(s) (e.g., methyl group(s), ethyl group(s), propyl group(s)), aminoalkyl group(s) (e.g., aminomethyl group(s), aminothethyl group(s), aminopropyl group(s)), hydroxyalkyl group(s) (e.g., hydroxymethyl group(s), hydroxyethyl group(s), hydroxypropyl group(s)), one or more amino acid(s), lipid(s), lipid acid(s) and/or terpene(s). Preferably, the employed glycosyltransferase(s) can still accept the carbohydrate moiety of which one or more of the functional group(s) is/are replaced by one or more other residues(s) as a substrate. Preferably, a monosaccharide unit is a pentose or a hexose, in particular a D-pentose or a D-hexose.

The carbohydrate moiety may be composed of one, two, three, four, five or more monosaccharide unit(s). The carbohydrate moiety may be a linear, branched or cyclic carbohydrate moiety. It may also be branched more than one times or may bear a symmetric or asymmetric dendrimer-like structure.
When there are more than one monosaccharide units, the monosaccharide units may be conjugated with another by any bond known in the art such as, e.g., an ether bond (R^a-CO-NH-R^b), an amide bond (R^a-N=H-R^b), a secondary amine (R^a-NH-R^b), an imine (R^a=N-R^b), an ester bond (R^a-CO-0-R^b), an alkyl linker (R^a=[CH_2]_n-R^b), a carbonyl group (R^a-CO-R^b), an alkenyl group (e.g., R^a=CH-R^b, R^a=CH-[CH_2]_n-R^b or R^a-[CH_2]_n=CH-[CH_2]_n-R^b), an alkynyl group (e.g., R^a=C-R^b, R^a=C-[CH_2]_n-R^b or R^a-[CH_2]_n=CH-[CH_2]_n-R^b), a thioether bond (R^a-S-R^b), a thioester bond (R^a-CO-S-R^b) or a thionoester bond (R^a-CS-0-R^b). As used herein, the residues R^a and R^b, interchangeably, represent two carbohydrate moieties. It will be understood that the aforementioned bonds also comprise the charged and/or tautomeric forms thereof such as, e.g., R^a-N=C(OH)-R^b, R^a-NH_2^+-R^b, or R^a-NH^+=R^b.

Preferably, the monosaccharide units may be conjugated with another by an ether bond, a secondary amine, an ester bond or an amide bond.

A glycopeptide of the present invention may bear one or more glycosylation site(s). As used herein a glycosylation site may be understood in the broadest sense as an amino acid moiety to which one or more saccharides may be conjugated with. In particular, a glycosylation site may refer to an amino acid moiety in a sequential context of amino acids, typically, a sequential context of a consecutive amino acid strand, to which one or more carbohydrate moiety/moieties is/are conjugated in nature. The glycosylation pattern of the glycopeptide of the present invention may be equal to the naturally occurring glycosylation pattern or may be different.

As used herein, the term “glycosylation pattern” may be understood in the broadest sense as the location and number of glycosylated sites of the peptide, the type(s) of conjugated monosaccharide(s), the number of conjugated monosaccharide(s) and/or the sequential orientation of the conjugated monosaccharide(s) in the carbohydrate moiety/moieties. A carbohydrate moiety may be conjugated to an asparagyl or glutamyl residue via an amide bond or an ester bond, to a serinyl or threonyl residue via an ether bond, or to a cysteinyln residue via a thioether bond.

As used herein, the term “glycosylating” may be understood in the broadest sense as the conjugation of one or more carbohydrate moiety/moieties to the peptide. Preferably, in a first glycosylation step a glycosyltransferase specific for a particular carbohydrate moiety is used to conjugate said carbohydrate moiety with the peptide. After the glycosylation step, said glycosyltransferase may be removed. Optionally, one or more further
glycosylation step(s) are conducted. As used herein, a glycosylation step may be understood as a reaction cycle.

In each one glycosylation step, a further glycosyltransferase specific for the conjugation of a further carbohydrate moiety is used to conjugate said further carbohydrate moiety with the peptide moiety P1 or with a carbohydrate moiety already conjugated with the peptide moiety P1. More preferably, in a further glycosylation step, a further glycosyltransferase specific for the conjugation of a further carbohydrate moiety is used to conjugate said further carbohydrate moiety with the carbohydrate moiety already conjugated with the peptide moiety P1 and, thereby, elongates said carbohydrate moiety conjugated with the peptide moiety P1. Then, the carbohydrate moiety conjugated with the peptide moiety P1 grows with each glycosylation step.

Most preferably, in each one glycosylation step, a glycosyltransferase conjugates a further monosaccharide unit to the growing carbohydrate moiety conjugated with the peptide moiety P1.

As used in the context of the present invention, the terms "glycosyltransferase", "glycosylating enzyme" and "glycosylation enzyme" may be understood interchangeably in the broadest sense as any enzyme that can conjugate one or more carbohydrate moiety/moieties to the peptide moiety P1 and/or to the one or more carbohydrate moiety/moieties conjugated to the peptide moiety P1 of the present invention.

As mentioned below, P2 may optionally also be glycosylated. The glycosylation may, exemplarily, be further performed after NCL to obtain a peptide conjugate such as exemplarily MUC1(T)-Aux-MUC1(T). Therefore, notably, such glycosyltransferase may also or alternatively be able to conjugate one or more carbohydrate moiety/moieties to the peptide moiety P2 and/or to the one or more carbohydrate moiety/moieties conjugated to the peptide moiety P2 of the present invention.

Typically, a glycosyltransferase is an enzyme of enzyme classification (EC) class 2.4 that acts as a catalyst for the transfer of one or more monosaccharide(s), disaccharide(s), trisaccharide(s), oligosaccharide(s) or polysaccharide(s) from activated donor molecules to specific acceptor molecules (i.e., the peptide moiety P1). Glycosyltransferases using nucleotide diphospho-carbohydrates, nucleotide monophospho-carbohydrates and carbohydrate phosphates include hexosyltransferases (EC 2.4.1), pentosyltransferases (EC 2.4.2) and transferases of other glycosyl groups (EC 2.4.99). The glycosyltransferase of the present invention may be any glycosyltransferase known in the art such as
members of the families of glucosyltransferases, glucuronosyltransferases, galactosyltransferases, galacturonosyltransferases, fucosyltransferases, mannosyltransferases, N-acetylglucosaminyltransferases and N-acetylgalactosaminyltransferases (EC 2.4.1), xylosyltransferases, arabinosyltransferases and ribosyltransferases (EC 2.4.2), and sialyltransferases (EC 2.4.99).

Typically, a glycosyltransferase catalyses a site-specific glycosylation. The glycosyltransferase(s) of the present invention may be O-glycosyltransferase(s), N-glycosyltransferase(s) or S-glycosyltransferase(s), preferably O-glycosyltransferase(s) or N-glycosyltransferase(s), more preferably O-glycosyltransferase(s).

In a preferred embodiment, the method of the present invention further comprises the step of removing the glycosyltransferase(s) and/or reaction component(s) of the step of glycosylating the polymer-conjugated peptide moiety P1 by means of one or more glycosyltransferase(s). More preferably, the method of the present invention further comprises the step of removing the glycosyltransferase(s) and reaction component(s).

In a preferred embodiment, the step of removing the one or more enzyme(s) and/or reaction component(s) (step iib) is performed by chromatography and/or by precipitating the polymer-conjugated peptide, preferably by gel permeation chromatography (GPC) and/or precipitating the polymer-conjugated peptide, more preferably by a combination of GPC and precipitating the polymer-conjugated peptide (e.g. by a combination of microspin GPC and precipitating the polymer-conjugated peptide) or by precipitating as the only purification step, in particular by precipitating as the only purification step.

When precipitating as the only purification step, the step of removing the one or more enzyme(s) and/or reaction component(s) consists of precipitating the polymer-conjugated peptide.

As used herein, the term "removing" may be understood in the broadest sense as the decrease of the content of glycosyltransferase(s) and/or reaction component(s) by any means. Preferably, removing means that the content of glycosyltransferase(s) and/or reaction component(s) is decreased at least 10fold, more preferably at least 50fold, even more preferably at least 100fold, even more preferably at least 250fold, even more preferably at least 500fold and most preferably at least 1,000fold in comparison to the concentration of the reaction batch of the preceding the step of glycosylating the polymer-conjugated peptide moiety P1 by means of one or more glycosyltransferase(s) (glycosylation batch).
As used herein, the term "reaction component" may be understood in the broadest sense as any component that may be used for performing the step of glycosylating the polymer-conjugated peptide moiety P1. Said component(s) may be, e.g., one or more carbohydrate protecting group(s), one or more protected or unprotected carbohydrate moiety/moieties not conjugated with the peptide moiety P1, one or more cofactor(s) suitable for the employed glycosyltransferase(s), one or more side product(s) generated by the glycosyltransferase(s) and/or one or more precursor(s) usable by the employed glycosyltransferase(s).

As used herein, a cofactor suitable for the employed glycosyltransferase(s) may be, but may not be limited to a nucleotide triphosphate (e.g., adenosine triphosphate (ATP), guanosine triphosphate (GTP), uracil triphosphate (UTP)), lipids, vitamin(s) (e.g., biotin) and/or metal ions (e.g., iron ions, cobalt ions, magnesium ions).

As used herein, a side product generated by the glycosyltransferase(s) may be, but may not be limited to phosphate, pyrophosphate, uracil diphosphate (UDP), cytosine diphosphate (CDP), guanosine diphosphate (GDP), uracil monophosphate (UMP), cytosine monophosphate (CMP), uracil and/or cytosine.

As used herein, a precursor usable by the employed glycosyltransferase(s) may be, but may not be limited to one or more glycosyl moiety/moieties such as, e.g., one or more carbohydrate-phosphate(s), one or more carbohydrate-pyrophosphate(s), one or more carbohydrate-triphosphate(s), one or more nucleotide-diphosphate carbohydrate(s) (e.g., a uracil-diphosphate carbohydrate(s) (UDP-carbohydrate(s)), guanosin-diphosphate carbohydrate(s) (GDP-carbohydrate(s))), nucleotidiphosphate carbohydrate(s) (e.g., cytosine-monophosphate carbohydrate(s) (CMP-carbohydrate(s))), one or more lipid-linked carbohydrate donor(s) (e.g., terpenoid-linked carbohydrate donor(s) (e.g., dolichol- or polyisoprenol-linked carbohydrate donor(s))).

Further it may be understood that one or more buffer component(s) such as, e.g., salts, buffer substances (e.g., phosphate ions, hydrogenophosphate ions, TAPS (3-[[tris(hydroxymethyl)methyl]amino]propanesulfonic acid), Bicine (N,N-bis(2-hydroxyethyl)glycine), Tris(tris(hydroxymethyl)methylamino), Tricine (N-tris(hydroxymethyl)methylglycine), TAPSO ([N-Tris(hydroxymethyl)methylamino]-2-hydroxypropanesulfonic acid), HEPES (2-hydroxyethyl-1-piperazineethanesulfonic acid), TES (2-[[tris(hydroxymethyl)methyl]amino]ethanesulfonic acid), MOPS (3-(N-morpholino)propanesulfonic acid), PIPES (piperazine-N,N'-bis(2-ethanesulfonic acid),...
Cacodylate (dimethylarsinic acid), SSC (sodium citrate) and/or MES (2-(N-morpholino)ethanesulfonic acid)) may be removed.

Further it may be understood that one or more detergent(s) may be removed. As used herein, the term "detergent" may be understood interchangeably with "surfactant". A detergent may be every detergent used in the art such as, e.g., Triton (X-100, CF-54), Nonidet P-40, CHAPS (3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate), a polysorbate (glyceryl laurate, polyoxyethylene glycol sorbitan alkyl ester, e.g., polysorbate 20, polysorbate 21, polysorbate 40, polysorbate 60, polysorbate 61, polysorbate 65, polysorbate 80, polysorbate 81, polysorbate 85, polysorbate 120), a poloxamer (block copolymer of polyethylene glycol and polypropylene glycol), ADS (ammonium dodecyl sulfate), SLES (sodium lauryl ether sulfate), SMS (sodium myreth sulfate), DSS (dioctyl sodium sulfosuccinate), PFOS (perfluorooctanesulfonate), perfluorononanoate, PFOA (perfluorooctanoate), perfluorobutanesulfonate, sodium stearate, Sodium lauroyl sarcosinate, octenidine dihydrochloride, CTAB (cetyl trimethylammonium bromide), CTAC (cetyl trimethylammonium chloride), CPC (cetylpyridinium chloride), POEA (polyethoxylated tallow amine), BAC (benzalkonium chloride), BZT (benzethonium chloride), dimethyldioctadecylammonium chloride, DODAB (dioctadecyldimethylammonium bromide), cocamidopropyl hydroxysultaine, cocamidopropyl betaine, lecithin, a polyoxyethylene glycol alkyl ether, an octaethylene glycol monododecyl ether, a pentaethylene glycol monododecyl ether, a polyoxypropylene glycol alkyl ether, a decyl glucoside, lauryl glucoside, an octyl glucoside, a polyoxyethylene glycol octylphenol ether, a polyoxyethylene glycol alkylphenol ether, a polyoxyethylene glycol alkylphenol ether, nonoxynol-9, a glycerol alkyl ester or a sorbitan alkyl ester. Preferably, the one or more detergent(s) is/are Triton, Nonidet P-40, CHAPS, a polysorbate or a poloxamer.

Further, the reaction component(s) may also be residuals of the step of providing a peptide moiety P1 conjugated with one or more hydrophilic polymer(s) still present when performing the glycosylating step. Said component(s) may be, e.g., one or more amino acid protecting group(s), one or more protecting group(s) of the hydrophilic polymer(s), one or more coupling agent(s), one or more scavenger(s), one or more cleaving agent(s), one or more protected or deprotected amino acid(s) not conjugated with the peptide moiety P1, one or more protected or deprotected hydrophilic polymer(s) not conjugated with the peptide moiety P1, one or more cofactor(s) suitable for the employed of glycosyltransferase(s), one or more side product(s) generated by the glycosyltransferase(s) and/or one or more precursor(s) usable by the employed glycosyltransferase(s).
After the glycosylation, the glycosyltransferase(s) and/or reaction component(s) may be removed by any method(s) known in the art such as, e.g., one or more chromatographic method(s), one or more filtration method(s), one or more electrophoretic method(s), one or more precipitation-based method(s) and/or one or more dialysis method(s).

Preferably, the step of removing the glycosyltransferase(s) and/or reaction component(s) does not require any toxic agent(s) such as, e.g., acetonitrile and TFA. Therefore, purification may preferably be performed with an aqueous solution such as a buffer, more preferably a buffer suitable for the glycosyltransferase(s) of the next glycosylating step as shown in the examples.

Preferably, removing the glycosyltransferase(s) and/or reaction component(s) is performed by a combination of not more than three different methods, more preferably a combination not more than two different methods or even not more than a single method.

Preferably, removing the glycosyltransferase(s) and/or reaction component(s) is performed by a chromatographic method. As used throughout the present invention, a chromatographic method may be any chromatographic method known in the art such as, e.g., gel permeation chromatography (GPC), high performance liquid chromatography (HPLC), reversed phase HPLC (RP-HPLC), fast protein liquid chromatography (FPLC), Flash Chromatography (flash), Rapid Refluid Liquid Chromatography (RRLC), Rapid Separation Liquid Chromatography (RSLC), Ultra Fast Liquid Chromatography (UFLC), reversed phase UFLC (RT-UFLC), Ultra Performance Liquid Chromatography (UPLC) or reversed phase UPLC (RT-UPLC). The column may be a prepacked column and/or a spin column usable by the employment of a centrifugal force.

Most preferably, the step of removing the glycosyltransferase(s) and/or reaction component(s) after one glycosylating cycle is performed by a single chromatographic step and/or a single precipitating step. That means that only one column and/or one precipitation step is used. That may lead to higher yields of the glycosylated peptide moiety P1 in comparison to a multistep procedure in which more than one chromatographic steps and/or more than one precipitating steps are used. This leads to a particular improvement of yields when there are more than one glycosylation steps, thus, more than one glycosylating cycles.

When using a chromatographic method, the chromatographic matrix may preferably be a bead matrix. The beads may be spherical or may bear any other shape known in the art to
conduct chromatographic matrix. The beads may be of any material known in the art to be useful for preparing chromatographic material such as, e.g., silica, sugar-based bead material (e.g., agarose, sepharose), plastic bead material (e.g., polystyrene). The beads may bear a neutral, a positive or a negative zeta potential. The column may be a flow-through device or may be used batch-wise.

Preferably, removing the glycosyltransferase(s) and/or reaction component(s) is performed by a chromatographic method using an aqueous buffer as solvent. More preferably removing the glycosyltransferase(s) and/or reaction component(s) is performed by gel permeation chromatography (GPC).

As used herein, the terms "gel permeation chromatography", "molecular sieve chromatography", "gel filtration" and "size exclusion chromatography" may be understood interchangeably in the broadest sense as a chromatographic method for separating molecules via size. Typically the GPC matrix has a pore size in which several molecules that fit in (i.e., the polymer-conjugated peptide moiety P1) are retarded and that high-molecular weight components (i.e., glycosyltransferase(s)) pass by. It will be understood that in the context of the present invention the pore-size is selected accordingly. Preferably the GPC matrix has a hydrophilic surface such as, e.g., cross-linked polysaccharide material (e.g, a cross-linked dextran gel (e.g., Sephadex material such as, e.g., Sephadex G-10 or Sephadex G-15)).

Preferably, a chromatographic method is a fast chromatographic method such, e.g., a chromatographic method working with a pressure that is higher than ambient pressure. This pressure may be obtained by using a pump and/or by centrifuging the column. A person skilled in the art knows the higher pressure may typically lead to a higher flow rate. Consequently, the tailing of the elution profile is minimized and the peptide moiety P1 is obtainable faster, in a higher purity, a higher yield and higher concentrated. Undesired interaction with the column matrix and surface is minimized.

Most preferably, removing the glycosyltransferase(s) and/or reaction component(s) is performed by centrifuging a microspin GPC column as used in the examples.

In the context of the present invention, an aqueous buffer may be any aqueous buffer suitable for the glycosyltransferase(s) used in the context of the present invention such as, e.g., phosphate buffered saline (PBS), a HEPES buffer, a Tris buffer, a TAPS buffer, a Bicine buffer, a Tricine buffer, a TAPSO buffer, a TES buffer, a MOPS buffer, a PIPES buffer, a Cacodylate buffer, an SSC buffer, an MESS buffer, an acetate buffer or culture
medium. Typically, the aqueous buffer will bear a pH in a range of between 4 and 9, preferably in a range of between 5 and 9, more preferably in a range of between 6 and 8.

Alternatively or additionally, a filtration method may be used. As used throughout the present invention, a filtration method may be any filtration method known in the art, such as, e.g., dead-end filtration or cross-flow filtration. Filtration may be conducted batch-wise of in a continuous flow method. As used herein, the terms "cross-flow filtration", "crossflow filtration", "tangential flow filtration" or "tangential filtration" may be understood interchangeably. The filter may be of any material known in the context of filtration in the art, such as, e.g., plastic (e.g., nylon, polystyrene), metal, alloy, glass, ceramics, cellophane, cellulose, or composite material. The filter may be hydrophobic or hydrophilic. The surface of the filter may be neutral or positively charged or negatively charged. Preferably, the filtration method is a filtration method using an aqueous buffer as solvent.

Further, alternatively or additionally, an electrophoretic method may be used. As used throughout the invention, an electrophoretic method may be any electrophoretic method known in the art such as, e.g., gel electrophoresis or capillary (CE) electrophoresis. Preferably, the electrophoretic method is an electrophoretic method using an aqueous buffer as solvent.

Further, alternatively or additionally, a precipitation-based method may be used. As used throughout the present invention, a precipitation-based method may be any precipitation-based method known in the art such as, e.g., salting in, salting out or precipitation by adding one or more organic solvents (e.g., acetone, diethylether, ethanol, methanol, dichloromethane). Typically, precipitation is performed by adding a solvent in which only one or some of the component(s) is/are well-soluble and the other component(s) is/are not or poorly soluble and precipitate. The supernatant and/or the pellet may be used further. Exemplarily, precipitation may be diethyl ether precipitation.

Further, alternatively or additionally, a dialytical method (dialysis) may be used. Dialysis is based on diffusion and osmosis, respectively, and is well-known in the art. Dialysis is widely used in protein purification and is also used to provide an artificial replacement for lost kidney function in people with renal failure. The membrane may be of any material, such as, e.g., plastic (e.g., nylon, polystyrene), metal, alloy, glass, ceramics, cellophane, cellulose, or composite material. The membrane may be hydrophobic or hydrophilic. The surface of the membrane may be neutral or positively charged or negatively charged. Preferably, dialysis is dialysis based on an aqueous buffer as solvent.
As described above, most preferably, chromatography is a single chromatographic step, thus, the use of a single column only.

As used herein, the term "microspin GPC" may be understood in the broadest sense as any spin column packed with a GPC matrix that is usable in a centrifugal tube and has a maximal loading volume of less than 15 ml, less than 10 ml, less than 5 ml, less than 4 ml, less than 3 ml, less than 2 ml, less than 1 ml, less than 0.5 ml, or less than 0.25 ml. The microspin GPC column may be used in a centrifuge and subjected to a centrifugal force of at least 100 x g, at least 1,000 x g, at least 2,500 x g, at least 5,000 x g, at least 7,500 x g, at least 10,000 x g, at least 15,000 x g, at least 25,000 x g, or at least 50,000 x g, depending on the GPC matrix. In a microspin column, the glycosylated peptide moiety P1 is in contact with the column matrix for a comparably short time. The person skilled in the art will understand that this may increase yield and purity of the eluted glycosylated peptide moiety P1. A tailing in the elution profile is minimized.

The peptide conjugate obtained in step (ii) may be prepared by any means known in the art for adding a peptide to a structure like the compound.

In a preferred embodiment, the step (ii) of adding a peptide moiety P1 to the amino group covalently bound to the thiol moiety S of the compound is:

(iia) providing an unbound peptide P1 and conjugating said peptide P1 to the compound of step (i) that is optionally is attached to a solid phase; or

(ii b) synthesizing the peptide moiety P1 on the compound of step (i) by sequentially conjugating amino acid building blocks, preferably wherein said compound is attached to a solid phase, in particular by means of Fmoc- or Boc-based solid phase peptide synthesis (SPPS).

As used throughout the present invention, the term "providing an unbound peptide P1" may be understood in the broadest sense as obtaining a peptide from any source. The peptide P1 may be provided by chemical synthesis, obtained from a biotechnological method and/or extracted from a natural source. Preferably, the peptide P1 is provided by chemical synthesis. As used herein, the term "chemical synthesis" may refer to SPPS, liquid phase peptide synthesis or a combination of both. Here, the synthesis typically bases on the stepwise coupling of amino acids bearing protected side chains (orthogonal protecting groups). Typically, during synthesis, the peptide strand grows from the C-terminus to the N-terminus. However, there are alternative methods wherein the peptide
strand grows from the N-terminus to the C-terminus. Nowadays, the most common methods base on at least two different types of protecting groups that are cleavable under at least two different conditions, such as, e.g., the fluorenlyl-9-methoxycarbonyl/tert-butanyl- (Fmoc/tBu) protecting group scheme (Sheppard Tactics) or the tert-butoxycarbonyl/benzyl- (Boc/Bzl) protecting group scheme (Merrifield Tactics). Alternatively or additionally, the peptide P1 may be also provided by conjugating two or more peptide strand(s) with another by any conjugation method known in the art such as, e.g., Native Chemical Ligation (NCL), Click Chemistry, Maleimide-Thiol Conjugation, enzymatic conjugation, biochemical protein ligation and/or soluble handling conjugation. Alternatively, the peptide P1 may be obtained from a biotechnological method. Today, numerous biotechnological methods are well-known in the art such as, e.g., overexpression and/or heterologous expression, in particular heterologous expression based on cloning of one or more gene(s) in bacteria, insect cells, mammalian cells or yeast cells. The peptide may further be extracted by any means known in the art. Alternatively, the peptide may be extracted from a natural source by any means known in the art. Additionally, the peptide P1 may be purified by any means known in the art, such as, e.g., one or more chromatographic method(s), one or more filtration method(s), one or more electrophoretic method(s), one or more precipitation-based method(s), one or more dialysis method(s) or a combination of two or more thereof. The natural source may be any biological material such as, e.g., bacterial material, plant material, animal material or fungal material, such as e.g. tissue, liquids or secretion(s). The peptide obtained from a natural source may also be digested or partly digested by one or more protease(s). It will be understood by a person skilled in the art, that the aforementioned methods for providing a peptide may also be combined with another. In particular a peptide obtained from a biotechnological method or a natural source may further be purified and/or modified by chemical means known in the art.

As used herein, the terms "solid support" in the context of peptide synthesis may be understood interchangeably in the broadest sense as any solid matrix known for peptide synthesis in the art. Typically the solid support is a plastic bead. The solid support may be, but may not be limited to, chloromethyl resin (Merrifield resin), 4-benzyloxybenzyl alcohol resin (Wang resin), (2,4-dimethoxy)benzhydrylamine resin (Rink amide resin), 2,4-dialkoxybenzyl resin (super acid-sensitive resin, SASRIN®), 2-chlorotrityl resin, alphachlorotrityl chloride resin (Barlos resin), benzhydrylamine resin (BHA resin), chloromethyl resin, hydroxymethylbenzoic acid resin (HMBA resin), 4-hydroxyethyl-3-methoxyphenoxoetyric acid resin (HMPB resin), hydroxycrotonyl aminomethyl resin (HYCRAM resin), MBHA resin, oxime resin, 4-(hydroxymethyl)phenylacetamidomethyl
resin (PAM resin) and/or a resin with special cleavable linkers (e.g., photolabile linkers or safety-catch linkers).

Solid phase peptide synthesis (SPPS) methods typically base on the use of orthogonal protecting groups cleavable under special conditions. A protecting group may be any protecting group known in the art such as, e.g., an amino-protecting group of the urethane type (e.g., benzylxycarbonyl (Z), 4-methoxybenzyloxycarbonyl (Z(OMe)), 2-nitrobenzyloxycarbonyl (Z(2-N02)), 4-nitrobenzyloxycarbonyl (Z(N02)), chlorobenzyloxycarbonyl (Z(Cl), Z(2-Cl), Z(3-Cl), Z(2,4-Cl)), 3,5-dimethoxybenzyloxycarbonyl (Z(3,5-OMe)), alpha, alpha-dimethyl-3,5-dimethoxybenzyloxycarbonyl (Ddz), 6-nitroveratryloxycarbonyl (Nvoc), 4-(phenyldiazenyl)-benzyloxycarbonyl (Pz), a-methyl-2,4,5-trimethylbenzyloxycarbonyl (Bic), 2-(biphenyl-4-yl)-2-propoxycarbonyl (Bpoc), (4-phenylazophenyl)-isopropoxycarbonyl (Azoc), isonicotinylxoxycarbonyl (iNoc), ie/f-butoxycarbonyl (Boc), 2-cyano-ie/f-butoxycarbonyl (Cyoc), 2,2,2-trichloro-ie/f-butoxycarbonyl (Tcboc), adamantly-1-oxycarbonyl (Adoc), 1-(1-adamantyl)-1-methoxyxocarbonyl (Adpoc), isobornyloxycarbonyl (Iboc), fluorenyl-9-methoxycarbonyl (Fmoc), (2-nitrofluoren-9-yl)methoxycarbonyl (Fmoc(N02)), 2-(4-toluenesulfonyl)-ethoxycarbonyl (Tsoc), methylsulfonylethoxycarbonyl (Msc), 2-(4-nitrophenylsulfonyl)ethoxycarbonyl (Nsc), 2-(tert-butylsulfonyl)-2-propenyloxycarbonyl (Bspoc), 1,1-dioxbenzo[b]-thien-2-yimethoxyxocarbonyl (Bsmoc), 2-(methylsulfonyl)-3-phenyl-2-propenylxocarbonyl (Mspoc), allyloxycarbonyl (Aloe), 2-(trimethylsilyl)-ethoxycarbonyl (Teoc), triisopropylsilylethoxycarbonyl (Tipseoc), piperidinyloxycarbonyl (Pipoc), cyclopententlyloxycarbonyl (Poc), 3-nitro-1,5-dioxaaspiro[5.5]undec-3-ylmethoxycarbonyl (PTnm), 2-ethynyl-2-propyl-oxycarbonyl (Epoc)), a carboxy-protecting group of the ester type (e.g., methyl (Me), ethyl (Et), benzyl (Bzl), 4-nitrobenzyl (Nbz), 4-methoxybenzyl (Mob), 2,4-di methoxybenzyl (2,4-Dmb), o-chlorotrityl (Trt(2-Cl)), pyrimidyl-4-methyl(4-picolyl (Pic), 2-toluene-4-sulfonyl)-ethyl (Tse), phenacly (Pac), 4-methoxyphenacly (Pac(OMe), diphenylmethyl (Dpm), ier/butyl (Bu), cyclohexyl (Cy), 1-adamantyl (1-Ada), 2-adamantyl (2-Ada), dicyclopentylmethyl (Dcpm), 9-phenylfluoren-9-y1 (Pi), 9-fluorenylmethyl (Fm), 2-trimethylsilylthoxycarbonyl (TMSe), 2-phenyl- trimethylsilyl-ethyl (PTMSe), allyl (Al), 4-[N-[1-(4,4-dimethyl-2,6-dioxocyclohexanilidene)-3-methylbutyl]-amino]benzyl (Dmab)), a thiol-protecting group (e.g., benzyl (Bzl), 4-methylbenzyl (Bz(4-Me)), 4-methoxybenzyl (Mob), 2,4,6-trimethoxybenzyl (Tmb), diphenylmethyl (Dpm), trityl (Trt), ier/butyl (Bu), acetamidomethyl (Acm), trimethylacetamidomethyl (Tacm), 9-fluorophenylmethyl (Fm), ier/butylsulfanyl (StBu), 3-nitro-2-pyriddylsulfanyl (Npys), allyloxycarbanyaminomethyl (Alocam), 9H-xanthen-9-yl (Xan)), an imidazole protecting group (e.g., benzyl (Bzl), 2,4-dinitrophenyl (Dnp), benzoxycarbonyl (Bom), adamantly-1-oxycarbonyl (Adoc), triphenylmethyl (Trt), diphenylmethyl (Dpm), pyridydiphienylmethyl
(Pdpm), 4-toluenesulfonyl (Tosyl, Tos), 4-methoxybenzenesulfonyl (Mbs), tert-butoxymethyl (Bum), allyl (Al), allyloxymethyl (Alom)), or a hydroxyl-protecting group (e.g., benzyl (Bzl), 2,6-dichlorobenzyl (Deb), diphenylmethyl (Dbm), cyclohexyl (Cy), 2-bromobenzoxycarbonyl (Z(2-Br)), tert-butyl (tBu), 1-benzoxycarbonyl-amino-2,2,2-trifluoroethyl (Zte), methylthiomethyl (Mtm), allyl (Al), allylcarbonyl (Aloe)).

The peptide P1 may be conjugated to the compound by any means known in the art. As mentioned above, herein a first amino acid of the formula

$$#^\text{v} \text{CH}_2 #^\text{w} \text{CO} #^\text{z}$$

(Vili),

wherein $#^\text{v}$ represents the binding site to the compound;

wherein $#^\text{w}$ represents the binding site to the remaining moieties of peptide moiety P1; and

wherein $R^2$ represents hydrogen or a substituted or unsubstituted C$_{1-5}$-alkyl residue, preferably hydrogen, CH$_3$-CH$_2$-OH or -CHOH-CH$_3$; in particular hydrogen, may preferably be formed.

As the compound use in the method of the present invention initially bears a primary amino group to which the peptide is conjugated, this amino group may be included in this first amino acid moiety. Therefore, conjugation is preferably performed by incubation with iodoacetic acid. This is exemplified further in the example section.

The final product obtained by the method may be the polypeptide conjugate P12 conjugated to the compound. In most cases, however, the indented final product will be the released unbound polypeptide conjugate P12.

Therefore, in a preferred embodiment, the method comprises the further step (v) of irradiating the compound comprising the peptide conjugate P12 obtained from step (iv) thereby cleaving said peptide conjugate P12 off the photolabile linker L and the hydrophilic polymer moiety HP.

This may be performed by any means known for that purpose. The wavelength and light intensity may be adapted to the individual photolabile linker L. Most typically, this will be performed in the UV-A range.

Optionally, the polypeptide conjugate P12 may also be purified.
Therefore, in a preferred embodiment, the method further comprises the step (vi) of purifying the peptide conjugate P12 obtained from step (iv) or (v).

In a more preferred embodiment, purification is performed by one or more chromatographic method(s) and/or one or more precipitation step(s), in particular gel permeation chromatography (GPC), fast protein liquid chromatography (FPLC), reversed phase high performance liquid chromatography (RP-HPLC) and/or diethyl ether precipitation.

The second moiety P2 used in the method may be any moiety bearing an alpha-thioester moiety.

In a preferred embodiment, the second moiety P2 is a peptide moiety, a peptide analogue moiety, a protein moiety, a lipid moiety, a polymeric scaffold moiety, a surface of a device, an oligonucleotide or a dye moiety.

In a particularly preferred embodiment, the second moiety P2 is a peptide moiety.

The surface of a device may exemplarily be a microarray coated with moieties comprising alpha-thioester moieties.

An oligonucleotide may be a natural oligonucleotide (e.g., consisting of a ribonucleic acid (RNA) strand or a deoxyribonucleic acid (DNA) strand) or may be an analogue of such natural oligonucleotide consisting of or comprising non-natural moieties such as, e.g., comprising xeno nucleic acid (XNA) (e.g., glycol nucleic acid (GNA), threose nucleic acid (TNA) and/or hexose nucleic acid (HNA)), locked nucleic acid (LNA), peptide nucleic acid (PNA) and/or morpholinos.

The second moiety P2 being a peptide moiety may be a peptide moiety P2 bearing the same sequence or a peptide moiety P2 bearing another sequence than the peptide moiety P1. The second peptide moiety P2 may be a glycopeptide or a non-glycosylated peptide. The second peptide moiety P2 may be a peptide moiety of less than 20 amino acids, more than 20 amino acids, more than 50 amino acids, more than 100 amino acids, more than 150 amino acids, more than 200 amino acids or even more than 250 amino acids.

Preferably, the peptide moieties P1 and P2 are conjugated with another by head-to-tail. Optionally, the second peptide moiety P2 may bear a sequence derived from the same naturally occurring peptide.
Alternatively, the peptide moiety P1 may be conjugated to another functional peptide such as, e.g., a fluorescent protein (e.g., (enhanced) green fluorescent protein (eGFP), red fluorescent protein (RFP), yellow fluorescent protein (YFP), cyan fluorescent protein (CFP) or mCherry), a cell-penetrating peptide (e.g., a polyarginine (e.g., R₇, Rs or R₉), a penetrating peptide, a Chariot peptide, a lactoferrin-derived peptide, a HIV Tat peptide, a SynB1 peptide, a Buforin peptide, a Magainin peptide, a HIV-gp41 peptide or a Kaposi FGF signal sequence peptide), or a reporter group (e.g., a histidine (HIS) tag., streptavidin or a peroxidase). Depending on the fusion partner, the obtainable fusion peptide may then bear novel functions such as, e.g., fluorescence, cell-permeability, or binding to novel binding partners.

The obtained polypeptide conjugate P12 may be purified by any means known in the art, e.g., by one or more chromatographic method(s), one or more filtration method(s), one or more electrophoretic method(s), one or more precipitation-based method(s) and/or one or more dialysis method(s). Preferably, purifying the polypeptide conjugate P12 is performed by GPC, FPLC, RP-HPLC and/or diethyl ether precipitation. More preferably, purifying the polypeptide conjugate P12 is performed by GPC, FPLC and/or RP-HPLC and diethyl ether precipitation.

As used herein, diethyl ether precipitation is preferably performed with cooled diethyl ether, such as, e.g., diethyl ether at 4°C, -20°C or -80°C. The cooled diethyl ether may be added to the broth containing the polypeptide conjugate P12. The polypeptide conjugate P12 may then be precipitated. Optionally, the reaction tube containing diethyl ether and polypeptide conjugate P12 is incubated at a low temperature such as, e.g., at 4°C, -20°C or -80°C. A polypeptide conjugate P12-containing pellet may be obtained by centrifugating the reaction tube. The diethyl ether solution may optionally further contain other organic solvents such as, e.g., ethanol. Optionally, ethanol is first added to the reaction containing the polypeptide conjugate P12, mixed, and diethyl ether is added before incubation at a low temperature such as, e.g., at 4°C, -20°C or -80°C. The diethyl ether solution may be discarded and/or evaporated. The obtained polypeptide conjugate P12-containing pellet may be dried or dissolved in a suitable solvent (e.g., an aqueous buffer or an organic, polar solvent (e.g. dimethyl sulfoxide (DMSO), dimethyl formamide (DMF), acetonitrile, ethanol and/or methanol)). The polypeptide conjugate P12 may first be purified by precipitation and subsequently purified by a chromatographic method or vice versa.

The polypeptide conjugate P12 may also be subjected to a further step of preserving the polypeptide conjugate P12.
As used herein, the term "polypeptide conjugate P12" may be understood in the broadest sense as any means for preparing a storable polypeptide conjugate P12 and elongating shelf-life of said polypeptide conjugate P12. Exemplarily, preserving the polypeptide conjugate P12 may be drying or freeze-drying. As used herein, the terms "freeze-drying", "lyophilization" and "cryodesiccation" may be understood interchangeable in the broadest sense as the removal of evaporable residuals from the polypeptide conjugate P12 in the frozen state. Typically, the polypeptide conjugate P12 is first dissolved in an aqueous solution, preferably in water, a water/acetonitrile mixture or a buffer with extensive volatile components. The sample is then frozen (e.g., preferably at -80° or liquid nitrogen) and, finally, the sample is dehydrated. The freeze-dried polypeptide conjugate P12 may be stored at any suitable temperature such as, e.g., at room temperature, at -20°, at -80°C or in liquid nitrogen. Before use, the polypeptide conjugate P12 may be dissolved in any suitable solvent (e.g., an aqueous buffer or an organic, polar solvent (e.g. DMSO, DMF, acetonitrile, ethanol, methanol or a mixture of two or more thereof)). Alternatively, the polypeptide conjugate P12 may be frozen such as, e.g., at -20°, -80°C or in liquid nitrogen. The polypeptide conjugate P12 may be frozen in the dry state or dissolved in a suitable solvent. Before use, the polypeptide conjugate P12 may be thawed. The polypeptide conjugate P12 may further be preserved by adding one or more preservative agent(s) such as, e.g., sodium azide (NaN₃) and/or benzoic acid. Further, the polypeptide conjugate P12 may be dissolved in an organic solvent such as, e.g., DMSO, DMF, acetonitrile, ethanol, methanol or a mixture of two or more thereof. Further, in order to prevent oxidation of one or more of the polypeptide conjugate P12 side chain(s) and/or the carbohydrate moiety/moieties, optionally, one or more reducing agent(s) may be added to the stored polypeptide conjugate P12 such as, e.g. traces of a thiol.

The present invention further relates to a polypeptide conjugate P12 obtained from the method of the present invention or salts thereof, in particular pharmaceutically acceptable salts thereof.

As the process of preparing the polypeptide conjugate P12 by means of the present invention, in particular when also treated by one or more enzymes (e.g., one or more glycosyltransferase(s)) is highly specific, the polypeptide conjugate P12 obtained from the method of the present invention is typically an extensively pure polypeptide. Therefore, said polypeptide conjugate P12 is particularly well usable for any means a particularly pure polypeptide conjugate P12 may be used for, in particular for one or more pharmaceutical purpose(s).
Exemplarily, said polypeptide conjugate P12 or a pharmaceutically acceptable salt thereof may be used for a pharmaceutical purpose such as, e.g., as antibiotic, hormone, cytokine, vaccine, artificial extracellular matrix, artificial glyocalyx and/or coating for an implant.

Preferably, said polypeptide conjugate P12 or a pharmaceutically acceptable salt thereof may be used as a vaccine, in particular a mixed vaccine, thus, a vaccine in that several polypeptide moieties of a polypeptide conjugate P12 of the present invention and at least a part of the one or more carbohydrate moiety/moieties conjugated with said polypeptide conjugate P12 in common serve as an antigen. Said antigen may be recognizable by an immune cell receptor or an antibody or fragment thereof. Said polypeptide conjugate P12 may also serve as a haptene.

Said polypeptide conjugate P12 or a pharmaceutically acceptable salt thereof may be used in a method for the treatment or prevention of a disease such as, e.g., cancer, thrombosis, myocardial infarction, viral or bacterial infection and/or stroke. In particular, the polypeptide conjugate P12 of the present invention may be used in a method for the prevention of a viral or bacterial infection or for the treatment of cancer by immune therapy.

Further, said polypeptide conjugate P12 or a salt thereof may be used for scientific research, in particular research dealing with cell-cell interaction, cell-matrix interaction, neoplastic cells, cancer, thrombosis, myocardial infarction, viral or bacterial infection and/or stroke. The peptide obtainable by the method of the present invention may also be used in a screening library or on a microarray for detecting novel interaction partners of enzymatically modified polypeptides(s) (e.g., glycopeptides(s)) or characterizing interaction pattern(s) of such polypeptides with cells or other peptides. The peptide may further be immobilized on an affinity chromatographic matrix and used for affinity chromatography.

Exemplarily, such polypeptide conjugate P12 or a salt thereof may be a short mucin-type O-glycopeptide. Preferable, said mucin-type O-glycopeptide is carrying the Tη-, T-, ST- or STn-antigen or combinations thereof, in particular Tη-, T- or STn-antigen O-glycans or combinations thereof usable as a cancer vaccine.

The invention is further explained by the following Examples and Figures, which are intended to illustrate, but not to limit the scope of the present invention.

**Brief Description of the Figures**
Scheme 1 shows a schematic representation of the auxiliary mediated chemoenzymatic approach, including all glycans installed on MUC1 peptides.

Scheme 2 shows the basic principle of using the compound of the present invention in a method of the present invention. It is demonstrated that an efficient approach for the preparation of site-specifically modified proteins has been established and applied to the synthesis of native homogeneously glycosylated MUC1 analogues. A photo-cleavable PEGylated auxiliary mediates the enzymatic site-specific O-glycosylation and subsequent ligation of MUC1 peptides and is eventually cleanly removed by UV irradiation.

Scheme 3 shows a schematic representation of the synthesis of an auxiliary-MUC1 peptide conjugate. In this case the MUC1 peptide (MUC1 (3Tn-3Ac)) carries three Tn antigens at specific positions. The synthesis of the conjugate does not involve iodoacetylation of the peptidyl resin and $\text{S}_n^2$ reaction, instead the auxiliary $\text{8}$ is first converted into an auxiliary-glycine conjugate $\text{10}$, which is in turn coupled to the N-terminus of the MUC1 (3Tn-3Ac) peptidyl resin via HATU-mediated coupling generating the conjugate Aux-MUC1 (3Tn-3Ac).

Figure 1 shows the synthesis of the photocleavable ligation auxiliary. a) Methyl 4-chloro butanoate, $\text{K}_2\text{CO}_3$, $\text{Bu}_4\text{N}$, $\text{CH}_3\text{CN}$, reflux, 80%; b) $\text{HN}_0\text{3,CH}_3\text{COOH}$, 0°C to rt, 93%; c) $\text{MePP}_3\text{Br}$, $\text{TMS}_2\text{NNa}$, THF, -10°C to 0°C, 65%; d) AD-mixa, $^1\text{BuOH}/\text{CH}_2\text{Cl}_2/\text{H}_2\text{O}$, 90%; e) $\text{SOCl}_2$, $\text{NEt}_3$, $\text{CH}_2\text{Cl}_2$, 0°C, f) $\text{NaN}_3$, DMF, 60°C; g) $\text{PP}_3$, $\text{H}_2\text{O}$, THF, 70°C; h) $\text{Boc}_2\text{O}$, DMAP, $\text{CH}_3\text{CN}$, 0°C, 63% (over four steps); i) $\text{PP}_3$, DIAD, $\text{CH}_3\text{COSH}$, THF, 0°C to rt, 84%; j) $\text{MeONa}$, 1-(butylthio)-1,2-hydrazine carboxmorpholide, MeOH, rt, 77%; k) 1,2-ethylenediamine, $\text{Zr(OBu)}_4$, HOBT, $\text{CH}_3\text{CN}$, 70%; l) 1. FmocPEG$_2\text{COOH}$, HATU, DIEA, $\text{CH}_3\text{CN}$, rt; 2. TFA, $\text{CH}_2\text{Cl}_2$, 68% (two steps); m) 1. FmocCl, DIEA, $\text{CH}_2\text{Cl}_2$, 0°C to rt; 2. TFA $\text{CH}_2\text{Cl}_2$, 92% (two steps).

Figure 2 shows the sequential glycosylation of peptide conjugate Aux-MUC1(Tn) gives peptide conjugates Aux-MUC1(\text{T}) and Aux-MUC1(\text{ST}). MUC1: Peptide consisting of the tandem repeat sequence of Mucin 1 $\text{VTSAPDTRPAGSTAPPAH}$ (SEQ ID NO: 2), the O-glycan on the side chain of Thr(14) are indicated in brackets. Tn: GalNaca, T: Gal31-3GalNaca, ST: Neu5Aca2-3Gal31-3GalNaca. Left panel: HPLC chromatogram and deconvoluted mass spectra of starting material Aux-MUC1(Tn) and of Aux-MUC1(T) and Aux-MUC1(ST) after precipitation. The peak at 2 min is the injection peak. Right panel: ESI-MS spectra of the conjugates. From bottom to top: Aux-MUC1(Tn): calculated mass: 4059 Da, found: 1354.0 Da ([M+3H]$^3^+$), 1015.7 Da ([M+4H]$^4^+$), 812.8 Da ([M+5H]$^5^+$), 677.8
Da ([M+6H]^{6+}), 580.9 Da ([M+7H]^{7+}); Aux-MUC1(T): calculated: 4222 Da, found: 1408.5 Da ([M+3H]^{3+}), 1056.7 Da ([M+4H]^{4+}), 845.4 Da ([M+5H]^{5+}), 704.6 Da ([M+6H]^{6+}); Aux-MUC1(ST): calculated: 4513 Da, found: 1505.2 Da ([M+3H]^{3+}), 1293.3 Da ([M+4H]^{4+}), 903.5 Da ([M+5H]^{5+}), 753.2 Da ([M+6H]^{6+}), 646.2 Da ([M+7H]^{7+}).

Figure 3 shows an HPLC chromatogram of the NCL reaction with non-glycosylated Aux-MUC1-OH and MUC1-SR and ESI-MS of the ligation product. MUC1: Peptide consisting of the tandem repeat sequence of Mucin 1 VTSAFDTRPAFPSTAPPAAH (SEQ ID NO: 2). Aux-MUC1-OH: auxiliary conjugate-MUC1 peptide. MUC1-SR: MUC1 peptide thioester.

MUC1-Aux-MUC1: ligation product, calculated mass: 5580 Da, found: 1861.0 Da ([M+3H]^{3+}), 1396.0 Da ([M+4H]^{4+}), 1116.9 Da ([M+5H]^{5+}), 931.0 Da ([M+6H]^{6+}), 798.2 Da ([M+7H]^{7+}), 698.6 Da ([M+8H]^{8+}), 621.1 Da ([M+9H]^{9+}). Asterisk indicates the ligation product (21.92 min), peak at 23.03 min: thiol deprotected Aux-MUC1-OH, peak at 24.48 min. Aux-MUC1-OH (with thiol protected as i.e/f-butylidisulfanyl) peak at 9.27 min: MUC1-SR, injection peak at 2 min.

Figure 4 shows HPLC chromatograms of the NCL reaction between Aux-MUC1(Tn) and MUC1-SR and ESI-MS of the ligation product. Calculated mass: 5783 Da, found: 1929.0 Da ([M+3H]^{3+}), 1447.0 Da ([M+4H]^{4+}), 1157.5 Da ([M+5H]^{5+}), 965.1 Da ([M+6H]^{6+}), 827.2 Da ([M+7H]^{7+}), 723.8 Da ([M+8H]^{8+}). Asterisk indicates the ligation product.

Figure 5 shows HPLC runs of the NCL reaction between Aux-MUC1(ST) and MUC1-SR and ESI-MS of the starting material and ligation product (from bottom to top): Aux-MUC1(ST) calculated mass: 4513 Da, found: 1505.2 Da ([M+3H]^{3+}), 1293.3 Da ([M+4H]^{4+}), 903.5 Da ([M+5H]^{5+}), 753.2 Da ([M+6H]^{6+}), 646.2 Da ([M+7H]^{7+}), unprotected Aux-MUC1(ST) calculated mass: 4425 Da, found: 1476.2 Da ([M+3H]^{3+}), 1070.0 Da ([M+4H]^{4+}), 886.1 Da ([M+5H]^{5+}), 738.5 Da ([M+6H]^{6+}), 633.3 Da ([M+7H]^{7+}), ligation product MUC1-Aux-MUC1(ST) calculated mass: 6237 Da, found: 1560.1 Da ([M+4H]^{4+}), 1248.2 Da ([M+5H]^{5+}), 1040.6 Da ([M+6H]^{6+}), 891.9 Da ([M+7H]^{7+}), 780.7 Da ([M+8H]^{8+}). Peaks at 2 min and 12 min in the chromatograms are the injection peak and MUC1-SR, respectively.

Figure 6 shows HPLC runs of the crude (A) and isolated (B) final product MUC1-G-MUC1(ST) and the corresponding ESI-MS (C). Calculated mass: 4355 Da, found: 1452.5 Da ([M+3H]^{3+}), 1089.8 Da ([M+4H]^{4+}), 871.8 Da ([M+5H]^{5+}), 727.0 Da ([M+6H]^{6+}), 623.4 Da ([M+7H]^{7+}). Asterisk indicates the unprotected ligation product.

Figure 7 shows HPLC runs and mass spectra of glycosylated Aux-MUC1-NHNH2, precursors of glycosylated peptide thioesters and the corresponding ESI-MS spectra (from
bottom to top): **Aux-MUC1(Tn)-NHNH$_2$** calculated mass: 4073 Da, found: 1358.9 Da ([M+3H]$^3+)$, 1019.3 Da ([M+4H]$^4+)$, 815.8 Da ([M+5H]$^5+)$, 680.0 Da ([M+6H]$^6+)$, 583.0 Da ([M+7H]$^7+)$, **Aux-MUC1(T)-NHNH$_2$** calculated mass: 4236 Da, found: 1413.1 Da ([M+3H]$^3+)$, 1060.1 Da ([M+4H]$^4+)$, 848.0 Da ([M+5H]$^5+)$, 706.8 Da ([M+6H]$^6+)$, **Aux-MUC1(ST)-NHNH$_2$** calculated mass: 4527 Da, found: 1510.6 Da ([M+3H]$^3+)$, 1132.6 Da ([M+4H]$^4+)$.

**Figure 8** shows an example for a protected compound of the present invention.

**Figure 9** shows an HPLC chromatogram and ESI-MS spectrum (major peak in the chromatogram) of Fmoc protected peptide MUC1. Calculated Mass for C$_{93}$H$_{134}$N$_{24}$O$_{26}$: 2051.0 Da, found: 1026.7 Da ([M+2H]$^2+)$, 685.0 Da ([M+3H]$^3+)$, 514.0 Da ([M+4H]$^4+)$.

**Figure 10** shows an HPLC chromatogram and ESI-MS spectrum of Fmoc protected MUC1 (Tn) peptide (GalNAc protected as Ac3). Calculated Mass for C$_{107}$H$_{153}$N$_{25}$O$_{37}$: 2380.1 Da, found: 1191.3 Da ([M+2H]$^2+)$, 794.7 Da ([M+3H]$^3+)$, 596.3 Da ([M+4H]$^4+)$.

**Figure 11** shows an HPLC chromatogram and ESI-MS spectrum of IAc-MUC1 peptide. Calculated mass for C$_{25}$H$_{36}$N$_{24}$C$_{18}$: 1996.8 Da, found: 1000.0 Da ([M+2H]$^2+)$, 666.9 Da ([M+3H]$^3+)$, 500.4 Da ([M+4H]$^4+)$.

**Figure 12** shows an HPLC chromatogram and ESI-MS spectrum (peak at 6.58 min) of IAc-MUC1 (Tn) peptide. Calculated Mass for C$_{194}$H$_{283}$N$_{29}$O$_{64}$S$_{2}$: 2325.9 Da, found: 1164.2 Da ([M+2H]$^2+)$, 776.6 Da ([M+3H]$^3+)$, 582.7 Da ([M+4H]$^4+)$.

**Figure 13** shows an HPLC chromatogram and ESI-MS spectrum (peak at 26.45 min, asterisk) of Aux(Fmoc)MUC1. Calculated Mass for C$_{157}$H$_{146}$N$_{28}$O$_{35}$S$_{2}$: 2551.2 Da, found: 1277.2 Da ([M+2H]$^2+)$, 851.7 Da ([M+3H]$^3+)$, 639.2 Da ([M+4H]$^4+)$.

**Figure 14** shows an HPLC chromatogram and ESI-MS spectrum (peak at 7.55 min, asterisk) of Aux(Fmoc)MUC1(Tn). Calculated Mass for C$_{158}$H$_{185}$N$_{29}$O$_{43}$S$_{2}$: 2880.3 Da, found: 1442.2 Da ([M+2H]$^2+)$, 961.5 Da ([M+3H]$^3+)$, 721.3 Da ([M+4H]$^4+)$.

**Figure 15** shows an HPLC chromatogram and ESI-MS spectrum of Aux-MUC1. Calculated Mass for C$_{173}$H$_{283}$N$_{29}$O$_{64}$S$_{2}$: 3854.9 Da, found: 1929.7 Da ([M+2H]$^2+)$, 1286.4 Da ([M+3H]$^3+)$, 965.2 Da ([M+4H]$^4+)$, 772.4 Da ([M+5H]$^5+)$, 643.8 Da ([M+6H]$^6+)$, 552.0 Da ([M+7H]$^7+)$.
Figure 16 shows an HPLC chromatogram and ESI-MS spectrum of Aux-MUC1(Tn). Calculated Mass for C173H283N29O64S2: 4059 Da, found: 1354.3 Da ([M+3H]3+), 1015.8 Da ([M+4H]4+), 812.9 Da ([M+5H]5+), 677.7 Da ([M+6H]6+), 580.9 Da ([M+7H]7+).

Figure 17 shows the results on an ESI-MS spectrum of MUC1-NHNH₂ peptide hydrazide. Calculated mass: 1842.9 Da, found: 1844.6 Da ([M+H]+), 922.6 Da ([M+2H]2+), 615.5 Da ([M+3H]3+), 461.9 Da ([M+4H]4+).

Figure 18 shows an HPLC chromatogram and ESI-MS spectrum of MUC1-SR peptide thioester. Calculated mass: 1952.8 Da, found: 977.8 Da ([M+2H]2+), 652.1 Da ([M+3H]3+), 489.6 Da ([M+4H]4+).

Figure 19 shows an HPLC chromatogram and ESI-MS spectrum of Aux-MUC1(Tn)-NHNH₂ (A) and of Aux(NH₂)-MUC1(Tn)-NHNH₂ (B). Aux-MUC1(Tn)-NHNH₂ calculated mass: 4073 Da, found: 1358.6 Da ([M+3H]3+), 1019.3 Da ([M+4H]4+), 815.8 Da ([M+5H]5+), 680.0 Da ([M+6H]6+), 583.0 Da ([M+7H]7+). Aux(NH₂)MUC1(Tr)-NHNH₂ calculated mass: 3852 Da, found: 1284.6 Da ([M+3H]3+), 964.0 Da ([M+4H]4+), 771.3 Da ([M+5H]5+), 642.9 Da ([M+6H]6+), 551.3 Da ([M+7H]7+).

Figure 20 shows the following:
A: time course (HPLC and ESI-MS) of the galactosylation reaction of Aux-MUC1(Tn) catalyzed by C1GalT1 enzyme. Aux-MUC1(Tn): calculated mass: 4059 Da, found: 1354.0 Da ([M+3H]3+), 1015.7 Da ([M+4H]4+), 812.8 Da ([M+5H]5+), 677.8 Da ([M+6H]6+), 580.9 Da ([M+7H]7+); Aux-MUC1(T): calculated: 4222 Da, found: 1408.5 Da ([M+3H]3+), 1056.7 Da ([M+4H]4+), 845.4 Da ([M+5H]5+), 704.6 Da ([M+6H]6+).

B and C: HPLC traces (B) and ESI-MS spectra (C) of sequential glycosylation of Aux-MUC1(Tn) and precipitation. To calculate the yield of the precipitation, the dry precipitate was dissolved in ddH₂O (same volume as the glycosylation reaction volume) and analyzed by HPLC in the same conditions used for monitoring the glycosylation. The ratio between the areas of the peaks (precipitated conjugate/conjugate in reaction mixture) gave the percentage of the recovery. Aux-MUC1(Tn): calculated mass: 4059 Da, found: 1354.0 Da ([M+3H]3+), 1015.7 Da ([M+4H]4+), 812.8 Da ([M+5H]5+), 677.8 Da ([M+6H]6+), 580.9 Da ([M+7H]7+); Aux-MUC1(T): calculated: 4222 Da, found: 1408.5 Da ([M+3H]3+), 1056.7 Da ([M+4H]4+), 845.4 Da ([M+5H]5+), 704.6 Da ([M+6H]6+); Aux-MUC1(ST): calculated: 4513 Da, found: 1505.2 Da ([M+3H]3+), 129.3 Da ([M+4H]4+), 903.5 Da ([M+5H]5+), 753.2 Da ([M+6H]6+), 646.2 Da ([M+7H]7+).
Figure 21 shows HPLC traces and ESI-MS spectra of sequential glycosylation of Aux(NH$_2$)-MUC1(Tn)-NHNH$_2$ and precipitation. Aux(NH$_2$)-MUC1(Tn)-NHNH$_2$:\ calculated mass: 3852 Da, found: 1284.6 Da ([M+3H]$^3+$), 964.0 Da ([M+4H]$^4+$), 771.3 Da ([M+5H]$^5+$), 642.9 Da ([M+6H]$^6+$), 551.3 Da ([M+7H]$^7+$); Aux(NH$_2$)$_2$-MUC1(T)-NHNH$_2$:\ calculated mass: 4013 Da, found: 1339.2 Da ([M+3H]$^3+$), 1004.5 Da ([M+4H]$^4+$), 803.5 Da ([M+5H]$^5+$), 669.9 Da ([M+6H]$^6+$), 574.5 Da ([M+7H]$^7+$); Aux$_{nNH_2}$-MUC1(T)-NHNH$_2$:\ calculated mass: 4304 Da, found: 1077.1 Da ([M+4H]$^4+$), 862.0 Da ([M+5H]$^5+$), 718.5 Da ([M+6H]$^6+$), 616.0 Da ([M+7H]$^7+$).

Figure 22 shows HPLC chromatograms of crude Aux-MUC1(ST)-NHNH$_2$ and ESI-MS of purified product. Calculated mass: 4637 Da, found: 1546.7 Da ([M+3H]$^3+$), 1160.4 Da ([M+4H]$^4+$), 928.3 Da ([M+5H]$^5+$), 773.7 Da ([M+6H]$^6+$), 663.4 Da ([M+7H]$^7+$).

Figure 23 shows HPLC chromatograms of the deprotection reaction of Aux-MUC1(ST) at different time points and ESI-MS of the starting material (Aux-MUC1(ST) (A), calculated mass: 4513 Da, found: 1505.2 Da ([M+3H]$^3+$), 129.3 Da ([M+4H]$^4+$), 903.5 Da ([M+5H]$^5+$), 753.2 Da ([M+6H]$^6+$)), of the desired Aux-MUC1(ST) with unprotected thiol (B), calculated mass: 4425 Da, found: 1476.2 Da ([M+3H]$^3+$), 1107.0 Da ([M+4H]$^4+$), 886.1 Da ([M+5H]$^5+$), 738.5 Da ([M+6H]$^6+$)) and of the dimer MUC1-Aux(SS)Aux-MUC1 (C), calculated mass: 8848 Da, found: 1770.5 Da ([M+5H]$^5+$), 1475.9 Da ([M+6H]$^6+$), 1265.2 Da ([M+7H]$^7+$), 1106.8 Da ([M+8H]$^8+$), 984.1 Da ([M+9H]$^9+$).

Figure 24 shows the photorelease of MUC1-G-MUC1 peptide. HPLC of the crude after irradiation and of the purified unprotected product (indicated by asterisk) and ESI-MS of pure MUC1-G-MUC1. calculated mass: 3697 Da, found: 1233.4 Da ([M+3H]$^3+$), 925.7 Da ([M+4H]$^4+$), 740.9 Da ([M+5H]$^5+$), 617.4 Da ([M+6H]$^6+$), 529.2 Da ([M+7H]$^7+$).

Figure 25 shows the photorelease of MUC1-G-MUC1(Tn) peptide. HPLC of the crude after irradiation and of the purified unprotected product (indicated by asterisk, purification and final analysis were done using two different HPLC columns) and ESI-MS of pure MUC1-G-MUC1(Tn). calculated mass: 3901 Da, found: 1302.0 Da ([M+3H]$^3+$), 976.4 Da ([M+4H]$^4+$), 781.3 Da ([M+5H]$^5+$), 651.3 Da ([M+6H]$^6+$), 558.5 Da ([M+7H]$^7+$).

Figure 26 shows an HPLC chromatogram and an ESI-MS spectrum of MUC1(Tn$_2$)-NHNH$_2$ peptide hydrazide. Calculated mass: 2047 Da, found: 1024.1 Da ([M+H]$^+$), 683.2 Da ([M+2H]$^2+$), 512.7 Da ([M+3H]$^3+$). Asterisk indicates the desired peptide hydrazide.
Figure 27 shows an HPLC chromatogram and an ESI-MS spectrum of MUC1(Tn\textsubscript{7})-SR peptide thioester. Calculated mass: 2157 Da, found: 1079.0 Da ([M+2H]\textsuperscript{2+}), 719.9 Da ([M+3H]\textsuperscript{3+}), 540.2 Da ([M+4H]\textsuperscript{4+}). Some hydrolyzed peptide thioester is also detectable (probably the little shoulder at 12.0 min).

Figure 28 shows HPLC chromatograms (t=0 and t=24h) of the NCL reaction between Aux-MUC1(Tn) (peaks at 25.15 min and 25.48 min, same mass as unprotected HS-Aux-MUC1(Tn) 3970 Da) and MUC1(Tn\textsubscript{7})-SR (peak at 11.78 min, the shoulder that increase over time is hydrolyzed thioester) and ESI-MS (direct injection) of isolated product MUC1(Tn\textsubscript{7})-Aux-MUC1(Tn): calculated mass: 5987 Da, found: 1498.2 Da ([M+4H]\textsuperscript{4+}), 1198.1 Da ([M+5H]\textsuperscript{5+}), 998.9 Da ([M+6H]\textsuperscript{6+}), 856.4 Da ([M+7H]\textsuperscript{7+}), 749.4 Da ([M+8H]\textsuperscript{8+}).

Figure 29 shows HPLC chromatograms and ESI-MS spectra of MUC1(Tn\textsubscript{7})-Aux-MUC1(Tn) (t = 0 galactosylation reaction, left) and of MUC1(T\textsubscript{7})-Aux-MUC1(T) (after precipitation, right): calculated mass for MUC1(Tn\textsubscript{7})-Aux-MUC1(Tn):

Figure 29A: 5987 Da, found: 1498.2 Da ([M+4H]\textsuperscript{4+}), 1198.1 Da ([M+5H]\textsuperscript{5+}), 998.9 Da ([M+6H]\textsuperscript{6+}), 856.4 Da ([M+7H]\textsuperscript{7+}), 749.4 Da ([M+8H]\textsuperscript{8+}), calculated mass for MUC1(T\textsubscript{7})-Aux-MUC1(T): 6312 Da, found: 1579.1 Da ([M+4H]\textsuperscript{4+}), 1263.3 Da ([M+5H]\textsuperscript{5+}), 1053.0 Da ([M+6H]\textsuperscript{6+}), 902.9 Da ([M+7H]\textsuperscript{7+}), 702.2 Da ([M+9H]\textsuperscript{9+}). Asterisk indicate compound MUC1(Tn\textsubscript{7})-Aux-MUC1(T) (left) or compound MUC1(T\textsubscript{7})-Aux-MUC1(T) (right).

Figure 29B: 5765 Da, found: 1442.2 Da ([M+4H]\textsuperscript{4+}), 1154.0 Da ([M+5H]\textsuperscript{5+}), 961.8 Da ([M+6H]\textsuperscript{6+}), 824.6 Da ([M+7H]\textsuperscript{7+}), 721.7 Da ([M+8H]\textsuperscript{8+}), 641.4 Da ([M+9H]\textsuperscript{9+}), calculated mass for MUC1(T\textsubscript{7})-Aux-MUC1(T) : 6089Da, found: 1523.3 Da ([M+4H]\textsuperscript{4+}), 1218.9 Da ([M+5H]\textsuperscript{5+}), 1015.8 Da ([M+6H]\textsuperscript{6+}), 870.9 Da ([M+7H]\textsuperscript{7+}), 762.2 Da ([M+8H]\textsuperscript{8+}), 677.6 Da([M+9H]\textsuperscript{9+}).

Figure 30 shows an a total ion current (TIC) chromatogram (A), HPLC chromatogram (B) and mass spectrum (C, 27.2-29.3 min of the TIC) of the mixture of conjugate MUC1(ST\textsubscript{7})-Aux-MUC1(ST) and MUC1(T\textsubscript{7})-Aux-MUC1(ST) / MUC1(ST\textsubscript{7})-Aux-MUC1(T). Calculated mass for MUC1(ST\textsubscript{7})-Aux-MUC1(ST): 6693 Da, found: 1379.6642 Da ([M+5H]\textsuperscript{5+}), 1149.8871 Da ([M+6H]\textsuperscript{6+}), 985.7612 Da ([M+7H]\textsuperscript{7+}), 862.7917 Da ([M+8H]\textsuperscript{8+}). Calculated mass for MUC1(T\textsubscript{7})-Aux-MUC1(ST) / MUC1(ST\textsubscript{7})-Aux-MUC1(T): 6603 Da, found: 1321.4431 Da ([M+5H]\textsuperscript{5+}), 1101.5366 Da ([M+6H]\textsuperscript{6+}), 944.1739 Da ([M+7H]\textsuperscript{7+}). Also detectable the mass of MUC1(ST\textsubscript{7})-Aux-MUC1(ST) - H\textsubscript{2}O (28 min) and of MUC1(ST\textsubscript{7})-Aux-MUC1(ST) + H\textsubscript{2}O (28.9nm); and an HPLC chromatogram (D) (t = 0: beginning of sialylation reaction, after precipitation: sialylated conjugate after precipitation procedure) and mass spectrum of conjugate MUC1(sT\textsubscript{7})-Aux-MUC1(sT). Calculated mass for MUC1(sT\textsubscript{7})-Aux-MUC1(sT): 6671 Da, found: 1669.3 Da ([M+4H]\textsuperscript{4+}), 1334.9 Da...
Figure 31 shows a mass spectrum (28.0-28.2 min of the TIC (Figure 30A)) of the mixture of conjugate MUC1(T$_7$)-Aux-MUC1(ST) / MUC1(ST$_7$)-Aux-MUC1(T) Calculated mass for MUC1(T$_7$)-Aux-MUC1(ST) / MUC1(ST$_7$)-Aux-MUC1(T): 6603 Da, found: 1321.4431 Da ($[M+5H]^{5+}$), 1101.5366 Da ($[M+6H]^{6+}$), 944.1739 Da ($[M+7H]^{7+}$). The second series visible in the spectrum account for the mass of the product (MUC1(ST$_7$)-Aux-MUC1(ST))$\cdot$H$_2$O.

Figure 32 shows HPLC chromatograms and ESI-MS spectra of the final purified products MUC1(T$_7$)-G-MUC1(T) and of MUC1(sT$_7$)-G-MUC1(sT): calculated mass for MUC1(T$_7$)-G-MUC1(T): 4429.1 Da, found: 1477.4 Da ($[M+3H]^{3+}$), 1108.3 Da ($[M+4H]^{4+}$), 886.8 Da ($[M+5H]^{5+}$), 739.2 Da ($[M+6H]^{6+}$), calculated mass for MUC1(sT$_7$)-G-MUC1(sT): 501.3 Da, found: 1671.3 Da ($[M+3H]^{3+}$), 1253.8 Da ($[M+4H]^{4+}$), 1003.1 Da ($[M+5H]^{5+}$), 836.9 Da ($[M+6H]^{6+}$).

Figure 33 shows an HPLC chromatogram of the crude product Aux-MUC1(3Tn-3Ac) obtained after coupling of the Aux-Gly to the N-terminus of the peptide containing three Tn antigens still protected as acetates. The MS spectrum in the figure correspond to the desired product indicated by the asterisk. The peak at 11.1 min is the PEGylated peptide (a small amount of MUC1(3Tn-3Ac) was not coupled to the auxiliary, as a consequence the activated PEG then reacted with the free N-terminus of the peptide). The peak at 8.9 min corresponds to the peptide modified with a tetramethyl guanidinium group (generated from excess HATU used as activating agent for the coupling).
Examples

Materials

9-Fluorenylmethoxycarbonyl (Fmoc)-protected amino acids and 2-(1 H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Merck (Darmstadt, Germany), Iris Biotech (Marktredwitz, Germany), ORPEGEN Peptide Chemicals (Heidelberg, Germany) and Sigma-Aldrich (Taufkirchen, Germany). TentaGel® R PHB-His(Trt)-Fmoc was purchased from Rapp polyme (Tuebingen, Germany). Fmoc-Thr(GalNAc-Ac₃)-OH was purchased from Sussex Research (Ottawa, Canada), UDP-Gal and CMP-Neu5Ac from Merck (Darmstadt, Germany). /V,V-Dimethylformamide (DMF), Methanol, dichloromethane (DCM), acetonitrile (ACN) and trifluoroacetic acid (TFA) were obtained from Biosolve (Valkenswaard, The Netherlands). Fmoc-PEG₉-COOH was from Polypure (Oslo, Norway). All other chemicals were obtained from Sigma-Aldrich at the highest purity available and used without further purification.

Synthesis and Use of the Compounds

Here, the synthesis and application of compounds comprising a photolabile moiety is demonstrated to allow attachment of the soluble polymer PEG for efficient enzymatic glycosylation, provides the functional groups to link two peptide moieties via NCL, and that the photolabile moiety can be cleanly removed by UV irradiation leaving a native glycine at the ligation site (Schemel ).

We based the first part of the auxiliary synthesis on the procedure developed by Dawson and coworkers (Marinzi et al., 2004). Starting from vanillin, the desired methyl butanoate, nitro and methylene groups were efficiently introduced (Figure 1). The planned subsequent Sharpless aminohydroxylation was not as efficient, despite numerous attempts under different conditions. Therefore, a stepwise strategy was adopted, starting with the Sharpless dihydroxylation to obtain the diol, which was then converted into the corresponding amino alcohol in three (respectively four) steps and 63% overall yield.

The Mitsunobu reaction, followed by protecting group exchange (acetyl to tert-butylsulfanyl), gave the protected mercaptoaminoethyl group essential for the NCL reaction.
Catalytic ester-amide exchange provided the attachment point for the PEG polymer and potentially for many other modifications. Final PEGylation of the free primary amine, removal of the Boc group and purification gave the desired auxiliary in 68% yield over the last two steps.

With the desired auxiliary in hands we investigated the conditions for its incorporation into a MUC1 peptide replacing a native N-terminal glycine residue (Scheme 1). The MUC1 tandem repeat peptide was synthesized by Fmoc-based solid phase peptide synthesis (SPPS) and the N-terminus of peptidyl resin MUC1 was modified with iodoacetic acid (Figure 11, ESI) via HBTU-mediated coupling and then submitted to an S_N2 reaction with auxiliary 7. Unfortunately, all tested conditions gave poor conversion even after several days of reaction. To improve this critical step, the synthesis of the MUC1 peptide was repeated on TentaGel resin, with lower peptide loading and better swelling properties when compared to the previously used Wang PS resin. To reduce any potential steric hindrance by the long PEG chain during the substitution reaction, auxiliary 8, missing the PEG moiety, was installed first and subsequently PEGylated on the resin. The synthesis of the auxiliary was modified accordingly by introducing an Fmoc protecting group on the free primary amine instead of the PEG (Figure 1, m).

Fortunately the S_N2 reaction of auxiliary 8 with the iodoacetylated MUC1 peptidyl (TentaGel) resin proceeded with very high conversion (Figure 13, ESI) and the PEGylation on resin gave complete conversion. After cleavage from the resin and purification, the pure MUC1-auxiliary conjugate Aux-MUC1 (Figure 15, ESI) was obtained in 34% yield.

This strategy was applied for the preparation of an auxiliary carrying MUC1 peptide including the Tn antigen (GalNacα1-0-Thr) at position 14 (Aux-MUC1(Tn)). The use of Fmoc-Thr(GalNac-Ac_2)-OH ensures full control over the O-glycan attachment site. Here iodoacetylation, attachment of the auxiliary, cleavage from the resin and subsequent hydrazinolysis of the acetyl protecting groups on the glycan gave Aux-MUC1(Tn) (Figure 2 and 16, ESI) in 4% overall yield (based on synthesis scale).

In a first attempt to prepare Aux-MUC1-Tn the GalNAc deprotection was performed on resin by overnight incubation with 5% (v/v) hydrazine monohydrate in methanol. Surprisingly, this not only led to acetyl removal but to complete release of the peptide from TentaGel resin. Further investigation demonstrated that hydrazine quantitatively cleaved the otherwise protected peptide from TentaGel resin by nucleophilic attack on the linker, forming a peptide hydrazide. This initially appeared as a nuisance on route to Aux-MUC1(Tn), however since peptide hydrazides can be efficiently converted into the
corresponding peptide thioesters, we decided to take advantage of this intermediate to generate larger MUC1 segments by sequential NCL reactions with auxiliary and hydrazide modified MUC1 peptides. In order to fully exploit the bi-functional MUC1 peptides we needed to prove that MUC1-NHNH$_2$ can be easily converted into the corresponding thioester and later on that chemoenzymatic glycosylation works on Aux-MUC1-NHNH$_2$.

We knew from previous attempts that peptide thioesters hydrolyze under glycosylation conditions, therefore a peptide hydrazide could be an ideal masked peptide thioester sufficiently stable to allow enzymatic glycosylation and subsequent conversion into the desired glycosylated peptide thioester. Moreover, the same peptidyl resin provided access both to the peptide acid and to the peptide hydrazide.

To test conversion of MUC1-NHNH$_2$ into the corresponding MUC1-thioester MUC1-SR, MUC1-peptidyl resin was incubated with 5% (v/v) hydrazine monohydrate in methanol overnight. Protected MUC1-NHNH$_2$ was obtained in 85% yield and deprotection of the side chains in solution by treatment with a mixture of TFA/TIS/H$_2$O (92.5:5:2.5) gave the desired MUC1-NHNH$_2$ (Figure 17, ESI). Treatment with NaN$_3$ followed by addition of MesNa and a RP-HPLC purification gave the MesNa thioester MUC1-SR (Figure 18, ESI) in 14% overall yield.

The next step was to prove that the PEGylated auxiliary was indeed able to facilitate the fast sequential site specific enzymatic glycosylation of the MUC1 peptides previously described for N-terminally PEGylated MUC1 (Bello et al., 2014). Glycosylation of Aux-MUC1(Tn) with recombinant human C1GalT1 produced Aux-MUC1(T) with a single T antigen (Gal31-3GalNAca disaccharide attached to Thr(14)) in excellent conversion (Figure 2 and 20A and B, ESI). The incubation with a mixture of ethanol and diethyl ether at -80° C induced the precipitation of Aux-MUC1(T), which was then collected in 95% yield by centrifugation (Figure 20B and C, ESI) and used in the next glycosylation step without further purification. Incubation of Aux-MUC1(T) with CMP-Neu5Ac in the presence of recombinant ST3Gal1 allowed efficient elongation of the disaccharide to the Neu5Aca2-3Gal31-3GalNAca- trisaccharide and to obtain Aux-MUC1(ST) containing the sialyl-T antigen in 90% yield (Figure 2 and 20B and C, ESI). By using similar conditions for glycosylation reactions but by increasing incubation times with ethanol and ether at -80°C (12 h instead of 4 h) we were able to efficiently and site-specifically O-glycosylate the auxiliary-modified MUC1 peptides and to further simplify the sequential glycosylation procedure (Bello et al., 2014) by omitting the GPC spin column step.

Next, the glycosylated peptide conjugates were linked to MUC1-SR via NCL to demonstrate all advantages of the auxiliary. Non-glycosylated Aux-MUC1 was used to
establish and optimize the ligation conditions for which the buffer composition was essential. First the thiol group of the auxiliary was deprotected via incubation with TCEP at 24°C for 6 h before addition of MUC1-SR.

Optimized NCL conditions for MUC1 peptides that give 65% conversion after two days of incubation are NaPi buffer (pH 7.5) at 30°C with Aux-MUC1 at 8 mM concentration and a 2.5-fold excess of MUC1-SR (Figure 3).

The optimized ligation conditions were applied for the synthesis of glycosylated MUC1-Aux-MUC1(Tn), which was obtained from MUC1-SR and Aux-MUC1(Tn) in 78% conversion after 36 h incubation at 30°C (Figure 4).

After demonstrating the individual functions of the auxiliary in MUC1 peptides we coupled glycosylation and NCL reaction by performing sequential enzymatic glycosylation of Aux-MUC1(Tn) to the sialylated core 1-containing conjugate Aux-MUC1(ST) (Figure 2), which was recovered by precipitation and directly used in a ligation reaction with MUC1-SR (Figure 5).

In this case consecutive additions of TCEP aliquots were needed to efficiently remove the ie/f-butylsulfanyl group from the auxiliary (Figure 23, ESI). Addition of MUC1-SR and MesNa to the ligation mixture led to the desired product in one day and in 70% conversion (Figure 5).

Finally, light induced removal of the PEGylated auxiliary was demonstrated for all three ligation products. Luckily this was easily accomplished by UV irradiation of the crude ligation mixtures in water or in a water/acetonitrile mixture. In all cases no starting material was detectable after 30 minutes of irradiation with an UV-A lamp and simultaneously a new peak containing the desired product was formed and easily isolated by HPLC purification (Figure 6 and Figures 24 and 25, ESI).

As described above, peptide hydrazides are potentially useful thioester precursors that remain unaffected in glycosylation reactions in which peptide thiosteres quickly hydrolyze. Hydrazine-induced cleavage of Aux-MUC1(Tn) peptidyl resin gave protected Aux-MUC1(Tn)-NHNH₂ (Figure 19A, ESI) that, after acidic deprotection in solution, was successfully used in sequential glycosylation reactions giving the corresponding Aux-MUC1(ST)-NHNH₂ hydrazide with similar yields as found for Aux-MUC1(ST)-OH (Figure 7).
Treatment with NaN0₂ followed by ascorbic acid (to avoid nitrosamine formation, a side reaction we observed in initial reactions) and MesNa gave access to the desired thioester Aux-MUC1(ST)-SR in 63% conversion.

A MUC1 peptide carrying the Tn antigen at a different position (Thr(7) instead than Thr(14)) was also synthesized via Fmoc-based SPPS and used to prepare a different peptide thioester. Hydrazine cleavage, deprotection and treatment with NaN0₂ and MesNa smoothly led to the desired MUC1(Tn₇)-SR in 43% yield (Figure 27). This peptide thioester was used in NCL with Aux-MUC1(Tn) to access MUC1(Tn₇)-Aux-MUC1(Tn) (Figures 28 and 29-left). This conjugate was efficiently used in the glycosylation-precipitation approach leading to a peptide (MUC1(Tn₇)-Aux-MUC1(T), Figure 29) consisting of two MUC1 tandem repeats and containing each a T antigen at different positions. In a next step sialylated MUC1(ST₇)-Aux-MUC1(ST) was obtained (Figure 30 and 31). Similarly MUC1(T)-Aux-MUC1 (T) and MUC1(ST)-Aux-MUC1 (ST) where obtained (with the T and ST antigen at position 14 in both MUC1 segments) and the auxiliary was fully converted into a glycine residue (G) by irradiation with 365 nm light generating MUC1(T)-G-MUC1(T) and MUC1 (ST)-G-MUC1 (ST) (Figure 32). These results indicate the power of this approach for the chemoenzymatic synthesis of glycosylated peptides and its combination with NCL to obtain larger polypeptides with different but specific glycosylation patterns.

The conjugation of the auxiliary to the peptide via SN2 reaction is the limiting step in the method. Indeed, while giving conversions that are higher that 50% with a non-modified MUC1 peptide or a peptide containing modifications far away from the N-terminus, the yields drop in the case of peptides carrying the modifications close to the N-terminus. In order to improve the efficiency of this step, we thought to first conjugate the auxiliary to a glycine and then to couple the Aux-Gly conjugate to the free N-terminus (non-iodoacetylated) peptide via HATU-mediated coupling (Scheme 3).

For this purpose i-butylglyoxalate was prepared via periodate oxidation of di-i-butyl-L-tartrate and directly used in a reductive amination reaction with auxiliary 8. Aux-Gly(tBu) 9 was obtained in 52% yield after purification via flash column chromatography on silica gel (ca 20% of the starting material can be recovered unreacted after the reductive amination). Removal of the tBu group by acidic treatment, followed by purification and freeze-drying, gave the desired product Aux-Gly 10 in 91% yield (scheme 3). This compound was used to prepare a new Aux-MUC1 conjugate in which the MUC1 peptide contains three Tn antigens (MUC1(3Tn)). MUC1 (3Tn-3Ac), in which the hydroxyl groups on the sugars are protected as acetates, was synthesized by Fmoc-based SPPS. After removal of the terminal Fmoc protecting group, the peptidyl resin was incubated with a solution of Aux-Gly 10, HATU and DIEA in DMF. After 2h incubation the resin was
washed and a test cleavage confirmed that the coupling reaction proceeded with a very high conversion (>90%). The Fmoc group on the auxiliary side chain was removed and then the Aux-MUC1 (3Tn-3Ac) conjugate was submitted to PEGylation. A new test cleavage was performed showing that the PEGylation was complete and that the desired product could be obtained (Figure 33). This new procedure could be also used, instead than with Auxiliary 8, with PEGylated auxiliary 7 (Figure 1), which could be modified with glycine via reductive amination. The obtained Aux(PEG)-Gly could be then coupled to the peptidyl resin via HATU-mediated coupling. In this way it would be possible to simplify the overall procedure, by avoiding the iodoacetylation of each peptidyl resin and the final PEGylation of each Aux-peptide conjugate, as both would be already included in the auxiliary scaffold.

In summary, we have developed a new PEGylated ligation auxiliary that efficiently supports the quantitative enzymatic glycosylation of peptides in solution and can mediate NCL reactions. Addition of a PEG polymer to the previously described ligation auxiliary turns the auxiliary-peptide conjugates into excellent substrates for the site-specific enzymatic glycosylation approach we previously described (Bello et al., 2014) and allows further simplification of the intermediate purification step. The auxiliary-modified (and glycosylated) peptides can be used in NCL reactions with other MUC1 tandem repeat peptides carrying a C-terminal othioester. All peptide α-thioesters used here are smoothly obtained via hydrazinolysis followed by oxidation of the corresponding hydrazide. Subsequent NCL reactions carried out in this study give conversions of more than 65% and the auxiliary is cleanly removed from the ligation products. The new PEGylated auxiliary in combination with generating peptide α-thioesters via direct hydrazinolysis allows preparation of site-selectively O-glycosylated peptide α-thioesters, opening the way to the synthesis of polypeptides comprising two or more MUC1 tandem repeats with different glycosylation patterns. Controlling thioester generation and deprotection of the thiol group within the auxiliary also allows the controlled extension of each building block in C- and N-terminal direction. We will now use this approach to create a library of site-selectively O-glycosylated MUC1 variants with different glycosylation pattern for a detailed study of the role of these patterns in MUC1 function.

However, this approach is not limited to the synthesis of glycosylated MUC1 analogues, but it is in principle applicable to the synthesis of many larger, posttranslationally modified proteins, only limited by the availability of the suitable chemistry or enzymes that introduce the desired PTMs.

Further Details Regarding the Synthesis and Use of the Compounds
Reaction monitoring, product purification and analysis by RP-HPLC/MS

For analytical liquid chromatography-mass spectrometry (LC-MS) a Waters AutoPurification HPLC/MS system (3100 Mass Detector, 2545 Binary Gradient Module, 2767 Sample Manager and 2489 UVA/visible Detector) was used. Mass spectra were acquired by electrospray ionization (ESI-MS) operating in positive ion mode. Separation was achieved with a Kromasil 300-5-C4 or 300-5-C18 column (50 µm, 5 µm particle size) at a flow rate of 1 mL/min running a linear gradient from 5% to 65% of (ACN + 0.05 % TFA) in (ddH2O + 0.05 % TFA) in 10 min.

Product analysis and reaction monitoring were performed by analytical RP-HPLC on a Dionex Ultimate 3000 instrument (Columns: C4 Kromasil 300-5-C4 150 µm, 5 µm particle size and BioBasic-4 150 µm, 5 µm particle size, Thermo Fisher) at a flow rate of 1 mL/min with a linear gradient from 5 to 65% buffer B in buffer A over 30 min (buffer A: 0.1% (v/v) TFA in ddH2O, buffer B: 0.08% (v/v) TFA in ACN).

Crude peptides or peptide-auxiliary conjugates were purified by reverse phase-high performance liquid chromatography (RP-HPLC) on Waters AutoPurification HPLC/MS system. According to the amount and hydrophobicity of the peptides to be purified, different columns were used: Kromasil 300-10-C4 column (250 x 2.12 mm, 10 µm particle size), Kromasil 300-10-C4 column (250 x 10 mm, 10 µm particle size). Detection for all the chromatographic methods occurred at 214 nm and 280 nm wavelengths.

**MUC1(ST7)-Aux-MUC1(ST) conjugate** was analyzed by RP-HPLC/HR-ESI-TOF on a Dionex Ultimate 3000 (Thermo Scientific) capillary HPLC coupled with a maXis HD ESI-QTOF (Bruker) instrument. An Agilent Technologies Zorbax 300SB C18 column (150x 0.3 mm, 3.5 µm particle size) was used for optimal separation with a flow rate of 6 µL/min and a gradient of 10% to 60% buffer B in buffer A in 40 min (in this case buffer A: ddH2O with 0.1% v/v formic acid, buffer B: ACN with 0.1% v/v formic acid).

Reaction monitoring and analysis for the auxiliary synthesis

Reactions were monitored by thin layer chromatography (Merck silica gel 60F254 plates). Detection and staining were carried out by UV light (254 nm) and KMn04 [KMn04 (3g), K2CO3 (20g), AcOH (0.25 ml) and H2O (300 ml)] or Pancaldi reagent [NH4]3MoO4 (21 g), Ce(S0)2 (1 g), H2S04 (31 ml) and H2O (470 ml)]. Purification of products was performed by flash column chromatography (silica gel 60, 230-400 mesh, 0.04-0.063 mm, Macherey-Nagel (Dijren, Germany)). 1HNMR were recorded on Bruker AC 250 spectrometer at 250 MHz, on Bruker AV-400 at 400 MHz and Bruker AV-500 at 500 MHz. The signals of residual solvents were used as reference (CDCl3: 7.27 ppm, CD3OD: 3.32 ppm). Coupling constants are given in Hz. 13CNMR spectra were recorded on Bruker AV-400 at 101 MHz
and Bruker AV-500 at 125MHz. The signals of residual solvents were used as reference (CDCl₃: 77 ppm, CD₃OD: 49 ppm). IR spectra were recorded on a Perkin Elmer Spectrum 2000 spectrometer. High resolution Mass spectra were recorded on a maXis HD ESI-QTOF Instant Expertise mass spectrometer. The synthesis scheme of the photocleavable ligation auxiliaries is depicted in Fig. 1.

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

![Methyl 4-(4-formyl-2-methoxyphenoxy)butanoate](image)

Vanillin (10.0 g, 66mmol) was dissolved in anhydrous acetonitrile (140 ml) under argon. Anhydrous K₂CO₃ (18.4 g, 133mmol, 2 eq) was added, followed by tetrabutylammonium iodide (4.4 g, 13 mmol, 0.2 eq) and methyl 4-chlorobutyrate (9.5 ml, 11.3 g, 83 mmol, 1.3 eq). The pink suspension was heated at reflux giving a yellow solution that was stirred at reflux for 18h. The carbonate was filtered off and the solvent was evaporated under reduced pressure. The crude product was re-crystalized from methanol, giving pure methyl 4-(4-formyl-2-methoxyphenoxy)butanoate (13.4g, 53mmol, 80% yield). ¹H NMR (500 MHz, CDCl₃) δ 9.87 (s, 1H, CHO), 7.49 - 7.41 (m, 2H, HAr), 7.01 (d, J = 8.2 Hz, ¹H, HAr), 4.19 (t, J = 6.3 Hz, 2H, OCH₂), 3.94 (s, 3H, OCH₃), 3.72 (s, 3H, COOCH₃), 2.59 (t, J = 7.2 Hz, 2H, CH₂CO), 2.27 - 2.18 (m, 2H, CH₂). ESI-MS: m/z calculated for C₁₃H₁₆O₅: 252.1 ; found 253.0 ([M+H]+). Analytical and spectroscopic data are in accordance with those already reported.

**Methyl 4-(4-formyl-5-nitro-2-methoxyphenoxy)butanoate**

![Methyl 4-(4-formyl-5-nitro-2-methoxyphenoxy)butanoate](image)

A cold solution of fuming nitric acid (24 ml) and glacial acetic acid (90 ml) was added dropwise to solid methyl 4-(4-formyl-2-methoxyphenoxy)butanoate (10.1 g, 41.1 mmol) at 0°C under stirring. After all the acid was added, the solution was stirred at room temperature for 3h then poured into ice and extracted with diethyl ether and ethyl acetate.
The combined organic phases were neutralized with solid sodium bicarbonate (pH controlled with pH-paper), then water was added and the two layers were separated. The organic layer was washed with brine, dried over MgSO$_4$ and the solvent was evaporated under reduced pressure. The obtained solid (1.14 g, 38.2 mmol, 93% yield) was dried under high vacuum and then used in the next step without any further purification. $^1$H NMR (250 MHz, CDCl$_3$): $\delta$ 10.48 (s, 1H, CHO), 7.64 (s, 1H, H$_\text{A}'$), 7.44 (s, 1H, H$_\text{A}''$), 4.25 (t, $J = 6.2$ Hz, 2H, OCH$_2$), 4.03 (s, 3H, OCH$_3$), 3.74 (s, 3H, COOCH$_3$), 2.61 (t, $J = 7.1$ Hz, 2H, CH$_2$CO), 2.32 - 2.18 (m, 2H, CH$_2$). Analytical and spectroscopic data are in accordance with those already reported.

Methyl 4-(2-methoxy-5-nitro-4-vinylphenoxy)butanoate (2)

Dry methylphosphonium bromide (18.1 g, 50 mmol, 1.3 eq) was suspended in anhydrous THF (75 ml) in a three-neck round-bottom flask and the suspension was stirred under argon at -5°C. Sodium bis-(trimethylsilyl)amide (1 M in THF, 52 ml, 50 mmol, 1.3 eq) was added dropwise and the obtained yellow solution was stirred for 1h at -5°C. Methyl 4-(formyl-5-nitro-2-methoxyphenoxy)butanoate (11.3 g, 40 mmol) was dissolved in dry THF (105 ml) in a two-neck round-bottom flask. The obtained solution was added dropwise to the solution of ylide at 0°C. The purple solution was stirred at 0°C for 5 min then at room temperature for 18h. After consumption of the starting material (TLC pentane/ethyl acetate 3:2), the solvent was evaporated under reduced pressure. Water and chloroform were added to the crude and the two phases were separated. The aqueous phase was extracted with chloroform. The collected organic layers were washed with brine, dried over MgSO$_4$ and the solvent was evaporated under reduced pressure. Pure methyl 4-(2-methoxy-5-nitro-4-vinylphenoxy)butanoate 2 (7.8 g, 26 mmol, 65% yield) was obtained after flash column chromatography on silica gel (pentane/ethyl acetate from 4:1 to 3:1).

$^1$H NMR (250 MHz, CDCl$_3$): $\delta$ 7.59 (s, 1H, HAr), 7.40 - 7.21 (m, 1H, CH=CH$_2$), 6.96 (s, 1H, HAr), 5.63 (dd, $J = 17.3$, 0.9 Hz, 1H, CH=CHH H$_{1a}$), 5.43 (dd, $J = 10.9$, 0.9 Hz, 1H, CH=CHH H$_{1a}$), 4.13 (t, $J = 6.2$ Hz, 2H, OCH$_2$), 3.97 (s, 2H, OCH$_3$), 3.71 (s, 3H, COOCH$_3$), 2.57 (t, $J = 7.2$ Hz, 2H, CH$_2$CO), 2.26 - 2.13 (m, 2H, CH$_2$). ESI-MS: m/z calculated for C$_{14}$H$_{17}$O$_6$: 295.1; found 296.1 ([M+H]$^+$). Analytical and spectroscopic data are in accordance with those already reported.
AD-mix a (10.2 g) was dissolved in BuOH/CH₂Cl₂/H₂O (1:1:2, 68 ml) at room temperature. The mixture was stirred at room temperature until the reagent was completely dissolved, then the orange solution was cooled down to 0°C. Methyl 4-(2-methoxy-5-nitro-4-vinylphenoxy)butanoate (2.1 g, 7.1 mmol) was added and the solution was stirred at room temperature for 18 h. After consumption of the starting material (TLC ethyl acetate/pentane 4:1) sodium sulfite (10.2 g) was added and the suspension was stirred at room temperature for 1 h. Water was added, followed by ethyl acetate. The two phases were separated and the aqueous phase was extracted with ethyl acetate. The combined organic layers were washed with brine, dried over magnesium sulfate and the solvent was evaporated under reduced pressure. The pure product 3 (2.1 g, 6.4 mmol, 90% yield) was obtained after filtration through silica gel. 

\[ ^1H\text{ NMR (400 MHz, CDCl}_3\text{)} \delta 7.54 \text{ (s, 1H, } H_A) , \]
\[ 7.25 \text{ (s, 1H, } H_A') , \]
\[ 5.45 \text{ (dd, } J = 7.3, 3.1 \text{ Hz, 1H, CHO), } 4.09 - 3.99 \text{ (m, 3H, OCH}_2\text{ +AcOEt) , } 3.95 \text{ (dd, } J = 11.2, 3.2 \text{ Hz, 1H, CHHOH) , } 3.89 \text{ (s, 3H, OCH}_3\text{) , } 3.62 \text{ (s, 3H, COOCH}_3\text{) , } 3.50 \text{ (dd, } J = 11.2, 7.4 \text{ Hz, 1H, CHHOH) , } 2.48 \text{ (t, } J = 7.2 \text{ Hz, 2H CH}_2\text{CO) , } 2.15 - 2.06 \text{ (m, 2H, CH}_2\text{). } ^{13}C\text{ NMR (101 MHz, CDCl}_3\text{)} \delta 173.79, 154.53, 147.67, 140.24, 132.11, 110.20, 109.69, 70.75, 68.66, 67.44, 56.78, 52.13, 30.76, 24.65. \]

HR-ESI-MS: m/z calculated for C₁₄H₁₹NNaO₈⁺: 352.1003; found: 352.1003 ([M+Na⁺])

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

\[ \text{Methyl 4-(4-(1,2-dihydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate (3).} \]

\[ \text{Methyl 4-(4-(2-azido-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate} \]
Diol 3 (2.97 g, 9.0 mmol) was dissolved in dichloromethane (29 ml) at 0°C. Triethylamine (1.82 ml, 1.32 g, 13 mmol, 1.5 eq) was added, followed by thionyl chloride (2.6 ml, 36 mmol, 4 eq). The solution was stirred at 0°C. After 30 min the reaction was finished (TLC, ethyl acetate/pentane 4:1). The solution was diluted with diethyl ether and water, the two phases were separated, the organic layer was washed with brine, dried over MgSO₄ and the solvent was evaporated under reduced pressure. Methyl 4-(4-(1,2-dihydroxyethyl cyclic sulfite)-2-methoxy-5-nitrophenoxy)butanoate was obtained as yellow solid (2.96 g, 7.9 mmol, 88% yield) and used in the following reaction without any further purification. The crude (2.96 g, 7.9 mmol) was dissolved in DMF, sodium azide (1.28 g, 19.7 mmol, 2.5 eq) was added and the solution was stirred at 60°C. After consumption of the starting material (TLC ethyl acetate/pentane 4:1) the DMF was partially removed under reduced pressure. Water was added and the solution was extracted with diethyl ether. The combined organic layers were washed with brine and dried over MgSO₄. The solvent was evaporated under reduced pressure and the crude was dried under high vacuum overnight. The desired methyl 4-(4-(2-azido-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate was obtained as an orange oil (2.83 g, (still containing 20 mol% DMF (NMR)) 6 mmol, 76% yield) and used in the next step without any further purification.

1H NMR (500 MHz, CDCl₃) δ 7.65 (s, 1H, H/sub N), 7.13 (s, 1H, H/sub /sub A), 5.60 (dd, J = 7.3, 3.6 Hz, 1H, CHN₃), 4.15 (tt, J = 6.3, 3.3 Hz, 2H, OCH₂), 4.02 (dd, J = 8.4, 3.2 Hz, 1H, CHHOH), 3.73 (dd, J = 11.5, 7.3 Hz, 1H, CHHOH) 3.72 (s, 3H, COOCH₃), 2.58 (tt, J = 7.1, 3.4 Hz, 2H, CH₂CO), 2.26 - 2.17 (m, 2H, CH₂).

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

Methyl 4-(4-(2-azido-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate (955.7 mg, 2.7 mmol) was dissolved in THF/H₂O 3:1 (9.4 ml), triphenylphosphate (813.1 mg, 3.1 mmol, 1.2 eq) was added and the reaction mixture was stirred at 60°C for 1h (monitored by TLC: ethyl acetate/petroleum ether 4:1). The solution was cooled-down to room temperature and then the solvent was evaporated under reduced pressure. Pure methyl 4-(4-(2-amino-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate (765.6 mg, 2.33 mmol, 87% yield) was obtained after flash column chromatography on silica gel (CH₂Cl₂/MeOH 93:7 +1%
Methyl 4-((N-te/f-butyloxycarbonyl) (2-amino-1-hydroxyethyl)-2-methoxy-5-nitrophenox)butanoate (4)

Methyl 4-((2-amino-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate 7 (765.6 mg, 2.33 mmol) was dissolved in ethanol (4.6 ml) at 0°C. Boc₂O (606.4 mg, 2.77 mmol, 1.2 eq) was added portionwise followed by triethylamine (0.3 ml, 2.16 mmol, 0.9 eq) and the solution was stirred at 0°C. Dichloromethane (1 ml) was added to dissolve the yellow precipitate that formed. The reaction was stirred at 0°C for 1 h (followed by TLC CH₂Cl₂/MeOH 95:5 + 1% NH₄OHaq) and then the solvent was evaporated under reduced pressure. The crude was dissolved in CH₂Cl₂, water was added and the two phases were separated. The aqueous phase was extracted with CH₂Cl₂, the collected organic phases were washed with brine, dried over MgSO₄ and the solvent was evaporated under reduced pressure. The pure methyl 4-((N-te/f-butyloxycarbonyl) (2-amino-1-hydroxyethyl)-2-methoxy-5-nitrophenoxy)butanoate 8 (794.5 mg, 1.85 mmol, 80% yield) was obtained after purification by flash column chromatography on silica gel (CH₂Cl₂/MeOH 95:5 + 1% NH₄OHaq). ¹H NMR (400 MHz, CDCl₃) δ 7.65 (s, 1H, H₁A), 7.03 (s, 1H, H₂A), 5.62 (bs, 1H, NHBoc), 5.48 (bs, 1H, CHNHBoc), 4.14 (t, J = 6.2 Hz, 2H, OCH₂), 4.07 (dd, J = 11.3, 4.1 Hz, 1H, CHHO), 3.96 (s, 3H, OCH₃), 3.89 (dd, J = 10.5, 5.5 Hz, 1H, CHHO), 3.73 (s, 3H, COOCH₃), 2.58 (t, J = 7.2 Hz, 2H, CH₂CO), 2.26 - 2.15 (m, 2H, CH₂), 1.43 (bs, 9H, (CH₃)₂C). ESI-MS: m/z calculated for C₁₉H₂₈N₂O₉: 428.4; found: 428.6 ([M+H]⁺) and 451.1 ([M+Na]⁺).
Triphenylphosphine (661.5 mg; 2.52 mmol, 2.1 eq) was dissolved in anhydrous THF (19.5 ml) at 0°C under argon atmosphere. Diisopropyl azodicarboxylate (0.55 ml, 2.79 mmol, 2.4 eq) was added and the solution was stirred at 0°C for 30 min. A white precipitate formed. Methyl 4-(4-N-[(f-butoxycarbonyl)amino-2-(f-butyl disulfanyl)ethyl]-2-methoxy-5-nitrophenox)butanoate (8) (505.3 mg, 1.17 mmol) was dissolved in anhydrous THF (24 ml) and the obtained solution was added dropwise to the solution of triphenylphosphine at 0°C. Neat thioacetic acid (0.18 ml, 2.55 mmol, 2.2 eq) was then added dropwise and the dark red solution was stirred at 0°C for 5 min, then at room temperature for 16 h. After consumption of the starting material (TLC CH₂Cl₂/MeOH 95:5 + 1% NH₄OHaq and pentane/ethyl acetate 2:1) the solvent was evaporated under reduced pressure. The residue was dissolved in 60 ml MeOH and the solution incubated at 0°C for 24 h. The white precipitate (triphenylphosphine oxide) was filtered off and washed with cold methanol. The solvent was evaporated and pure methyl 4-(4-(2-(acetylsulfanyl)-1-((f-butoxycarbonyl)amino)ethyl)-2-methoxy-5-nitrophenox)butanoate (5) (486.5 mg, 1.01 mmol, 86% yield) was obtained after purification of the crude by flash column chromatography on silica gel (pentane/ethyl acetate 2:1). ¹H NMR (360 MHz, CDCl₃) δ 7.64 (s, 1H, Hₐ), 7.04 (s, 1H, Hₐ'), 5.70 (bs, 1H, NHBoc), 5.40 (bs, 1H, CHNHBoc), 4.13 (t, J = 6.1 Hz, 2H, OCH₂), 3.98 (s, 3H, OCH₃), 3.72 (s, 3H, COOCH₃), 3.37-3.28 (m, 2H, CH₂S), 2.58 (t, J = 7.2 Hz, 3H, CH₂CO), 2.42 (s, 3H, CH₃CO), 2.28 - 2.13 (m, 2H, CH₂), 1.41 (bs, 9H, (CH₃)₃C). ESI-MS: m/z calculated for C₂₁H₂₅N₂O₉S₂: 486.2; found: 486.6 ([M+H]⁺) and 509.1 ([M+Na]⁺).

Methyl 4-(4-{1-[(f-butoxycarbonyl)amino-2-(f-butyl disulfanyl)]ethyl}-2-methoxy-5-nitrophenox)butanoate (5)
Methyl 4-(4-(2-(acetylsulfanyl)-1-(isopropoxycarbonyl)amino)ethyl)-2-methoxy-5-nitrophenoxy)butanoate (467 mg, 0.96 mmol) was dissolved in anhydrous methanol (28 ml) under argon atmosphere, sodium methoxide (1.9 ml, 0.5 M solution in MeOH) was added dropwise and the solution was stirred at room temperature for 15 min. 2-Methylpropane-2-thiol (3.3 ml, 30 eq) was added dropwise. The solution was stirred at room temperature and air was bubbled through it for 18 h. The reaction was followed by TLC (pentane/ethyl acetate 2:1) and extra 2-methylpropane-2-thiol (2 x 0.5 ml) was added. After 24 h the starting material was almost completely consumed. After solvent evaporation and purification by flash column chromatography on silica gel (pentane/ethyl acetate 3:1) the pure product was obtained in moderate yield (236 mg, 0.44 mmol, 46% yield).

Alternative synthesis:
Methyl 4-(4-(2-(acetylsulfanyl)-1-(isopropoxycarbonyl)amino)ethyl)-2-methoxy-5-nitrophenoxy)butanoate (101 mg, 0.21 mmol) was dried under high vacuum for 2 h and then dissolved in anhydrous THF (825 µl) and MeOH (225 µl), previously degassed. Sodium methoxide (500 µl, 0.5 M solution in MeOH) was added dropwise and the solution was stirred at room temperature for 10 min. Freshly prepared 1-(isopropylthio)-1,2-hydrazinecarboxmorpholide3 (92 mg, 0.27 mmol, 1.3 eq) was added and the mixture was stirred at room temperature for 3 h. The reaction mixture was washed with citric acid (0.1 M aqueous solution) until pH 6 and with brine and then dried over MgSO4. The solvent was evaporated and the pure product was obtained (67 mg, 0.13 mmol, 62% yield) after purification by flash chromatography on silica gel (petroleum ether/ethyl acetate 2:1 (a second fraction (22 mg) containing both the desired product and the product with hydrolyzed methyl ester was recovered as well).

1H NMR (250 MHz, CDCl3) δ 7.63 (s, 1H, Hα), 7.03 (s, 1H, Hα), 5.68 (bs, 1H, NHBoc), 5.55 (m, 1H CHNHBoc), 4.14 (t, J = 6.4 Hz, 2H, OCH2), 3.98 (s, 3H, OCH3), 3.73 (s, 3H, COOCH3), 3.34 (dd, J = 13.7, 4.8 Hz, 1H, CHHS), 3.09 (m, 1H, CHHS), 2.58 (t, J = 7.2 Hz, 2H, CH2CO), 2.32 - 2.12 (m, 2H, CH2), 1.43 (s, 9H, (CH3)3COC(0)), 1.36 (s, 9H,
(CH$_3$)$_3$CS). ESI-MS: m/z calculated for C$_{23}$H$_{36}$N$_2$O$_8$S$_2$: 532.2; found: 532.6 ([M+H]$^+$) and 555.2 ([M+Na]$^+$).

**Tert-butyl (1-(4-(4-((2-aminoethyl)amino)-4-oxobutoxy)-5-methoxy-2-nitrophenyl)-2-(te/f-butyldisulfanvDethvDcarbamate (6)**

Ester 5 (288 mg, 0.54 mmol) and HOBr were dissolved in extra-dry acetonitrile in a heat-dried two-neck round-bottom flask under argon atmosphere. Ethylenediamine (1 ml, 14.9 mmol, 27 eq) and zirconium (IV) ie/f-butoxide (230 µl, 0.59 mmol, 1.1 eq) have been added under water and oxygen-free atmosphere. The solution was stirred at room temperature for 18 h. The solvent and the amine were evaporated under reduced pressure and pure ierf-butyl (1-(4-(4-((2-aminoethyl)amino)-4-oxobutoxy)-5-methoxy-2-nitrophenyl)-2-(ie/f-butyldisulfanyel)ethyl)carba-mate (6) (277 mg, 0.49 mmol, 91% yield) was obtained after purification of the crude by flash column chromatography on silica gel (CHCl$_3$/MeOH 9:1 + 1% NEt$_3$). $^1$H NMR (500 MHz, MeOD, mixture of rotamers) $\delta$ 7.60 (s, 1H, H$_A$), 7.17 (s, 1H, H$_A$), 5.63 - 5.52 (m, 1H, CHNHBoc), 4.11 (t, J = 6.2 Hz, 2H, OCH$_2$), 3.95 (s, 3H, OCH$_3$), 3.40 (m), 3.29 (t, J = 6.3 Hz, 2H, CONHCH$_2$), 3.23 - 3.17 (m, 2H, CHHS), 2.92 (dd, J = 24.2, 12.1 Hz, 2H, CHHS), 2.77 (t, J = 6.3 Hz, 2H,CH$_2$NH$_2$), 2.45 (t, J = 7.8 Hz, 2H, CH$_2$CO ), 2.20 - 2.06 (m, 2H, CH$_2$), 1.44 (s, 9H, (CH$_3$)$_3$COC(0)), 1.36 (s, 9H, (CH$_3$)$_3$CS). $^{13}$C NMR (101 MHz, MeOD) $\delta$ 177.31, 177.31, 156.81, 156.81, 150.11, 135.45, 112.63, 111.94, 71.26, 70.95, 58.38, 53.29, 51.10, 50.89, 50.68, 50.46, 50.25, 50.04, 49.83, 43.64, 43.24, 34.88, 31.73, 27.73. HR-ESI-MS: m/z calculated for C$_{24}$H$_{41}$N$_{40}$.S$_2$: 561.241 1; found: 561.2402 ([M+H]$^+$) and 583.2221 ([M+Na]$^+$).

(9H-fluoren-9-yl)methyl (95-(4-(1-amino-2-(te/f-butylsulfinothioyl)ethyl)-2-methoxy-5-nitrophenoxy)-87,92-dioxo-3,6,9,12,15,18,21,24,27,30,33,36,39,42,45,48,51,54,57,60,63,66,69,72,75,78,81,84-octacosaox-88,91-diaza-pentananonacontyQcarbamate (7)
Fmoc-PEG$_2$-OH (476 mg, 0.31 mmol, 1.5 eq) was dissolved in a solution of HATU (0.5 M in acetonitrile, 420 µL). DIEA (72.4 µL, 2 eq) was added and the solution was diluted with acetonitrile (0.7 mL) and mixed with a solution of tert-butyl (1-((4-(1-amino-2-(tert-butyldisulfanyl)ethyl)-2-methoxy-5-nitrophenoxyl)butanamido)ethyl)carbamate (6) (118 mg, 0.21 mmol) and mixed in dry acetonitrile (1.5 mL). The reaction mixture was stirred at room temperature for 18 h. After consumption of the starting material (TLC: CH$_2$Cl$_2$/MeOH 9:1 +1% NEt$_3$) the solvent was evaporated under reduced pressure. The crude product was dissolved in CH$_2$Cl$_2$/TFA 1:1 (1.5 mL) and the solution was stirred at room temperature for 1 h. The solvent was evaporated under reduced pressure and pure 7 was obtained (283 mg, 0.14 mmol, 68% yield) after flash column chromatography on silica gel (CH$_2$Cl$_2$/MeOH 93:7 +1% NEt$_3$). IR (KBr): cm$^{-1}$ 3331, 2868, 1715, 1658, 1518, 1454, 1348, 1274, 1250, 1094, 946, 841, 763, 742.

$^1$H NMR (500 MHz, MeOD) δ 7.83 (d, J = 7.5 Hz, 2H, H$_A$), 7.68 (d, J = 7.4 Hz, 2H, H$_B$), 7.61 (s, 1H, H$_C$), 7.42 (t, J = 7.4 Hz, 3H, H$_D$), 7.34 (t, J = 7.4 Hz, 2H, H$_E$), 7.29 (s, 1H, H$_F$), 5.03 (dd, J = 8.1, 5.4 Hz, 1H, CHNH$_2$), 4.39 (d, J = 6.8 Hz, 2H, CH$_2$(Fmoc)), 4.23 (t, J = 6.8 Hz, 1H, CH(Fmoc)), 4.12 (t, J = 6.2 Hz, 2H, OCH$_2$), 3.99 (s, 3H, OCH$_3$), 3.81-3.45 (m, 116H, CH$_2$(PEG)), 3.36 - 3.21 (m, CH$_2$ (from HN(CH$_2$)$_2$NH) and CHHS + MeOH (residual solvent)), 3.07 (dd, J = 13.5, 8.2 Hz, 1H, CHHS), 2.53 - 2.37 (m, 4H, CH$_2$ (from HN(CH$_2$)$_2$NH and CH$_2$CO), 2.20 - 2.07 (m, 2H, CH$_2$), 1.36 (s, 9H, (CH$_3$)$_3$CS).

$^{13}$C NMR (126 MHz, MeOD) δ 174.13, 172.94, 157.44, 153.68, 147.65, 143.95, 141.42, 141.22, 127.42, 126.80, 124.80, 119.59, 109.91, 109.02, 70.18, 70.10, 70.01, 69.90, 69.87, 69.54, 68.36, 66.79, 66.18, 55.66, 49.28, 40.39, 38.71, 38.49, 36.28, 32.05, 28.82, 24.84. HR-ESI-MS: m/z calculated for C$_{93}$H$_{169}$N$_9$S$_{36}$O$_{36}$S$_{2}$+: 1987.0279, found: 1987.0022 ([M+H$^+$]). 994.0184 ([M+2H$^{2+}$]) and 663.0132 ([M+3H]$^{3+}$).

(9H-fluoren-9-yl)methyl (2-(4-(4-(1-amino-2-(tert-butyldisulfanyl)ethyl)-2-methoxy-5-nitrophenoxyl)butanamido)ethyl)carbamate (8).
Ter-butyl (1-(4-(4-((2-aminoethyl)amino)-4-oxobutoxy)-5-methoxy-2-nitrophenyl)-2-(i.e/f-butyldisulfanyl)ethyl)carbamate (6) (260 mg, 0.46 mmol) was dissolved in anhydrous dichloromethane (1.2 ml) and the solution was cooled down to 0°C. DIEA (111 µl, 1.4 eq) was added, followed by dropwise addition of a solution of FmocCl (165 mg, 0.64 mmol, 1.4 eq) in anhydrous dichloromethane (0.8 ml). The reaction mixture was stirred at 0°C for 30 min, dichloromethane (1 ml) was added and the solution was stirred at room temperature for 18 h. After consumption of the starting material (TLC, CHCl₃/MeOH 9:1 + 1% NEt₃) the solvent was evaporated under reduced pressure and the crude was dissolved in CH₂Cl₂/TFA (3 ml, 1:2 mixture) at 0°C and stirred at 0°C for 1 h. The solvent was evaporated under reduced pressure and the pure product (290 mg, 0.42 mmol, 92% yield) was obtained after purification by recrystallization from MeOH/Et₂O. IR (pure): cm⁻¹ 3430, 2631, 1663, 1523, 1445, 1336, 1280, 1193, 1101, 1062, 995, 729. ¹H NMR (400 MHz, MeOD) δ 7.78 (d, J = 7.5 Hz, 1H, H₃), 7.68 (s, 1H, H₅), 7.63 (d, J = 7.4 Hz, 1H, H₄), 7.38 (t, J = 7.5 Hz, 1H, H₆), 7.30 (t, J = 7.4 Hz, 1H, H₇), 7.14 (s, 1H, H₈), 6.36 (dd, J = 8.4, 5.9 Hz, 1H, CHₐ), 4.43 (d, J = 6.8 Hz, 2H, CH₂(Fmoc)), 4.18 (t, J = 6.8 Hz, 3H, CH(Fmoc)), 3.94 (s, 3H, OCH₃), 3.37 (dd, J = 14.2, 5.9 Hz, 6H, CHHS), 3.29 - 3.16 (m, 5H, CHHS and NH(CH₂)₂NH), 2.39 (t, J = 7.3 Hz, 2H, CH₂CO), 2.10 (p, J = 6.7 Hz, 2H, CH₂), 1.35 (s, 9H, (CH₃)₂CS). ¹³C NMR (101 MHz, MeOD) δ 175.60, 159.06, 155.44, 150.18, 145.34, 143.05, 142.61, 128.79, 128.16, 126.18, 125.29, 120.94, 111.07, 110.94, 69.89, 67.80, 57.21, 50.40, 48.46, 43.28, 41.31, 40.45, 33.38, 30.17, 26.24. HR-ESI-MS: m/z calculated for C₂₄H₃₄N₄O₇S₂⁺: 683.2568; found: 683.2553 ([M+H]⁺) and 705.2375 ([M+Na]⁺).

Ter-butyl glvoxalate
Di-tert-butyldtartrate (152 mg, 0.58 mmol) was dissolved in MeOH/H₂O 4:1 at 0°C. Sodium periodate (303 mg, 1.42 mmol, 2.4 eq) was added in portions and the reaction was stirred at 0°C for 4 h (monitoring by TLC CH₂Cl₂/AcOEt 95:5). The suspension was poured in water and extracted with Et₂O. The organic layers were dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure. The product was obtained as a colorless oil (147 mg, 97% yield) and used in the next step without further purification.
Tert-butyl (1-(4-(2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)ethyl)amino)-2-oxoethoxy)-5-methoxy-2-nitrophenyl)-2-(tert-butyldisulfanyl)ethyl)glvcinate (9)

Tert-butylglyoxalate (30 mg, 0.23 mmol, 1.2 eq) was dissolved in 1,2-dichloroethane (1.2 ml), (9H-fluoren-9-yl)methyl (2-(4-(4-(1-amino-2-(tert-butyldisulfanyl)ethyl)-2-methoxy-5-nitro-phenoxy)butanamido)ethyl)carbamate (8) (151 mg, 0.21 mmol) was added and the reaction was stirred at 45°C for 2h. Sodium triacetoxyborohydride (66 mg, 0.31 mmol, 1.5 eq) was added and the reaction was stirred at 45°C for 18h (monitoring by TLC CHpClp/MeOH 97:3 + 0.1% NNOH). The solvent was evaporated and the product was recovered pure as a light yellow solid (82 mg, 0.11 mmol, 52% yield) after purification by flash column chromatography on silica gel (CHpCH2 gradient from 1% to 3% MeOH + 0.1% NH4OH).n

1H NMR (400 MHz, MeOD) δ 7.78 (d, J = 7.5 Hz, 2H), 7.63 (d, J = 7.4 Hz, 2H), 7.48 (s, 1H), 7.38 (t, J = 7.4 Hz, 2H), 7.31 (dd, J = 13.2, 6.0 Hz, 3H), 4.67 (dd, J = 8.2, 4.8 Hz, 1H), 4.33 (d, J = 6.8 Hz, 2H), 4.18 (t, J = 6.7 Hz, 1H), 4.06 (t, J = 6.2 Hz, 2H), 3.89 (s, 3H), 3.70-3.29 (m, 2H), 3.27 - 3.21 (m, 1H), 3.21 - 3.11 (two overlapping signals: d, J = 16.9 Hz, 1H and dd, J = 4.8,13.2 1H), 3.02 (d, J = 16.9 Hz, 1H), 2.88 (dd, J = 13.2, 8.3 Hz, 1H), 2.40 (t, J = 7.3 Hz, 2H), 2.21 - 2.02 (m, 2H), 1.45 (s, 9H), 1.33 (s, 9H).

13C NMR (101 MHz, MeOD) δ 175.71 , 172.69, 159.01 , 155.07 , 148.81 , 145.33 , 144.08, 142.58, 131.82, 128.75, 128.13, 126.15, 120.91 , 111.53, 110.29, 82.60, 69.76, 67.75, 56.85, 56.72, 50.34, 48.70, 48.45, 48.00, 41.31 , 40.43, 33.52, 30.22, 28.33, 26.35.

HR-ESI-MS: m/z calculated for C49H53N4O9S2+: 797.3248, found: 797.3238 ([M+H]+) and 819.3055 ([M+Na]+).

(1-(4-(2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)ethyl)amino)-2-oxoethoxy)-5-methoxy-2-nitrophenyl)-2-(tert-butyldisulfanyl)ethyl)glvcine (10)
Tert-butyl (1-((2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)ethyl)(amino)-2-oxoethoxy)-5-methoxy-2-nitrophenyl)-2-((tert-butylsulfonyl)ethyl)glycinate (9) (82 mg, 0.11 mmol) was dissolved in TFA (2 ml) containing 1% of TIS and 0.5% of H$_2$O as scavengers and the solution was stirred at room temperature for 2h (monitoring via TLC: CH$_2$Cl$_2$/MeOH 93:7 +0.1% NH$_4$OH). The solvent was evaporated under reduced pressure and the residue was dissolved in a mixture of acetonitrile and water 7:3 and freeze-dried. The crude product was obtained as a light yellow powder, which was purified via flash column chromatography on silica gel (CH$_2$Cl$_2$/MeOH + 0.1% AcOH, gradient from 95:5 to 92:8 v/v) giving pure (1-((2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)ethyl)(amino)-2-oxoethoxy)-5-methoxy-2-nitrophenyl)-2-((tert-butylsulfonyl)ethyl)glycine (10) (73 mg, 0.10 mmol, 91% yield) as white foam (freeze-dried from ACN/H$_2$O).

$^1$H NMR (400 MHz, MeOD) $\delta$ 7.67 (d, $J = 7.5$ Hz, 2H), 7.52 (m, 3H), 7.27 (t, $J = 7.4$ Hz, 2H), 7.22 - 7.13 (m, 3H), 5.18 (t, $J = 7.2$ Hz, 1H), 4.23 (d, $J = 6.9$ Hz, 2H), 4.08 (t, $J = 6.8$ Hz, 1H), 4.00 (t, $J = 6.2$ Hz, 2H), 3.82 (s, 3H), 3.62 - 3.52 (m, 2H), 3.29 - 3.24 (m, 2H), 3.20 - 3.15 (m, 2H), 3.15 - 3.08 (m, 2H), 2.29 (t, $J = 7.3$ Hz, 2H), 2.05 - 1.94 (m, 2H), 1.21 (s, 9H).

$^{13}$C NMR (101 MHz, MeOD) $\delta$ 175.61, 170.20, 159.04, 155.33, 150.32, 145.33, 144.58, 142.59, 128.76, 128.13, 126.15, 123.84, 120.92, 112.16, 110.86, 69.88, 67.78, 57.59, 57.17, 48.88, 48.46, 48.05, 43.50, 41.30, 40.44, 33.41, 30.13, 26.25.

HR-ESI-MS: m/z calculated for C$_6$H$_{4}$N$_{4}$O$_{9}$S$_2$: 741.2629, found: 741.2618 ([M+H]$^+$) and 763.2437 ([M+Na]$^+$).

**Solid phase peptide synthesis**

Mucin peptide MUC1 (VTSAPDTPAPGSTAPPAH (SEQ ID NO: 2)) was synthesized on a solid support using the fluorenylmethoxycarbonyl (Fmoc) strategy. Preloaded TentaGel® R PHB-His(Trt)-Fmoc resin (TentaGel® R PHB-His(Trt)-Fm resin) (1g) was used as the solid support for the synthesis of the peptides on a 0.2 mmol scale. The resin was loaded into the reaction vessel of an automated peptide synthesizer (PTI Tribute), and swollen in DMF. The amino acids used in the synthesis were carrying the following orthogonal side-chain protecting groups: Thr(O'Bu), Ser(O'Bu), Arg(Pbf), Asp(O' Bu). Double couplings were performed after coupling of each proline residue. Mucin peptide MUC1(Tn) (VTSAPDTPAPGST(GalNAc-Ac3)APPAH) was synthesized manually until
Ser(13) on a solid support using the fluorenylmethoxycarbonyl (Fmoc) strategy. Preloaded TentaGel® R PHB-His(Trt)-Fmoc, resin (1g) was used as the solid support for the synthesis of the peptide on a 0.2 mmol scale. N-terminal Fmoc deprotection was performed by incubating the resin twice with a piperidine solution (20% v/v in DMF) for 3 and 7 min, respectively. After washing with DMF, the resin was incubated for 30 min with the pre-activated amino acid (2.5 eq, activated with 2.4 eq HBTU (0.5M, DMF) and 5.0 eq DIEA, 3 min) and washed again with DMF before starting a new deprotection step. Coupling and deprotection reaction of Ala(15) were monitored by ninhydrine test. Fmoc-Thr(GalNAc-Ac3)-OH (160 mg, 0.24 mmol, 1.2 eq) was dissolved in DMF (477 µl) and activated with HATU (91 mg, 0.24 mmol, 1.2 eq) and DIEA (87 µl, 0.5 mmol, 2.5 eq) for 2 min. The resin was incubated with the solution of activated modified amino acid for 2h and then washed with DMF. Coupling and deprotection of Fmoc-Thr(GalNAc-Ac3)-OH were monitored by ninhydrine test, as well as the coupling of the following serine (Ser(13)). The resin was washed with DMF, MeOH and DCM, dried under vacuum and transferred into the reaction vessel of the Fmoc-synthesizer (PTI Tribute) and the synthesis was continued automatically.

Mucin peptide MUC1(Tn7) (VTSAPDT(GalNAc-Ac3)RPAPGSTAPPAH) was synthesized manually on a solid support using the fluorenylmethoxycarbonyl (Fmoc) strategy. Preloaded TentaGel® R PHB-His(Trt)-Fmoc, resin (0.5g) was used as the solid support for the synthesis of the peptide on a 0.1 mmol scale. N-terminal Fmoc deprotection was performed by incubating the resin twice with a piperidine solution (20% v/v in DMF) for 3 and 7 min, respectively. After washing with DMF, the resin was incubated for 30 min with the pre-activated amino acid (2.5 eq, activated with 2.4 eq HBTU (0.5M, DMF) and 5.0 eq DIEA, 3 min) and washed again with DMF before starting a new deprotection step. Coupling and deprotection reaction of Arg(8) were monitored by ninhydrine test. Fmoc-Thr(GalNAc-Ac3)-OH (80 mg, 0.12 mmol, 1.2 eq) was dissolved in DMF (238.5 µl) and activated with HATU (45.5 mg, 0.12 mmol, 1.2 eq) and DIEA (43.5 µl, 0.25 mmol, 2.5 eq) for 2 min. The resin was incubated with the solution of activated modified amino acid for 2h and then washed with DMF. Coupling and deprotection of Fmoc-Thr(GalNAc-Ac3)-OH were monitored by ninhydrine test, as well as the coupling of the following aspartic acid (Asp(6)). The sequence was completed by coupling the remaining amino acids following the standard coupling and deprotection procedure. The peptidyl-resin was then washed, dried and directly used for the preparation of the corresponding peptide hydrazide (page 38 and following).

For peptides MUC1 and MUC1(Tn), after the last coupling, the resin was washed with DCM and dried under vacuum. The peptidyl resins were used in the subsequent reactions.
after performing a test cleavage to verify the efficiency of the synthesis: 10 mg of peptidyl-resin were incubated with a mixture of TFA/TIS/H₂O (92.5:5:2.5) for 2 h at room temperature. The free peptide was precipitated by addition of cold diethyl ether and centrifugation. Finally, the supernatant was removed and the peptide washed with cold diethyl ether, dissolved in a mixture of H₂O/CH₃CN 1:1 and analyzed by RP-HPLC and ESI-MS.

Iodoacetylation of MUC1 peptidyl resins

ICH₂C(Q)-VTSAPDTRPAPGSTAPPAGH-OH

The resin (0.4566 g) was swollen in DMF for 1 h, then the Fmoc protecting group was removed by incubating the resin twice with a solution of piperidine (20% v/v in DMF) for 3 and 7 min respectively. After extensive washing with DMF, the resin was incubated with a solution of iodoacetic acid (47 mg, 0.23 mmol) and DIC (35.7 µl, 0.23 mmol) in DMF (992 µl) for 30 min. After removal of the supernatant and washing with DMF, the coupling was repeated for 20 min. The peptide resin was washed with DMF, DCM and MeOH, dried under vacuum and a test cleavage was performed: 10 mg of the resin were incubated with a mixture of TFA/TIS/H₂O (92.5:5:2.5) for 2 h at room temperature. The free peptide was precipitated by addition of cold diethyl ether and centrifugation. Finally, the supernatant was removed and the peptide washed with cold diethyl ether, dissolved in a mixture of H₂O/CH₃CN 1:1 and analyzed by RP-HPLC and ESI-MS.

ICH₂C(Q)-VTSAPDTRPAPGST(GalNAc-Ac3)APPAH-OH

The resin (0.450 g) was swollen in DMF for 1 h, then the Fmoc protecting group was removed by incubating the resin twice with a solution of piperidine (20% v/v in DMF) for 3 and 7 min respectively. After extensive washing with DMF, the resin was incubated with a solution of iodoacetic acid (37 mg, 0.18 mmol) and DIC (28 µl, 0.18 mmol) in DMF (792 µl) for 30 min. After removal of the supernatant and washing with DMF, the coupling was repeated for 20 min. The peptide resin was washed with DMF, DCM and MeOH, dried under vacuum and a test cleavage was performed: 10 mg of the resin were incubated with a mixture of TFA/TIS/H₂O (92.5:5:2.5) for 2 h at room temperature. The free peptide was precipitated by addition of cold diethyl ether and centrifugation. Finally, the supernatant was removed and the peptide washed with cold diethyl ether, dissolved in a mixture of H₂O/CH₃CN 1:1 and analyzed by RP-HPLC and ESI-MS.

Conjugation of auxiliary (8) to MUC1 peptide

The iodoacetylated peptidyl MUC1 resin (247 mg) was swollen in CH₃CN (600 µl) for 2 h. DIEA (17 µl, 0.1 mmol) was added followed by the auxiliary (8) (59.0 mg, 0.1 mmol in 300 µl
CH$_3$CN). The suspension was stirred at 28°C for 18h, then the supernatant was removed (the unreacted auxiliary was recovered by recrystallization) and the resin washed with DMF, DCM and MeOH. A test cleavage was performed to check the efficiency of the S$_2$N$_2$ reaction: 10 mg of the resin were incubated with a mixture of TFA/TIS/H$_2$O (92.5:5:2.5) for 2h at room temperature. The free peptide was precipitated by addition of cold diethyl ether and centrifugation. Finally, the supernatant was removed and the peptide washed with cold diethyl ether, dissolved in a mixture of H$_2$O/CH$_3$CN 1:1 and analyzed by RP-HPLC and ESI-MS: the major peak in the chromatogram correspond to the desired product, some unreacted iodoacetylated peptide could still be detected (peak at 14.5 min), together with a few side products.

After test cleavage the remaining Aux-peptidyl resin conjugate was directly used in the following PEGylation step.

Conjugation of auxiliary (8) to MUC1 (Tn) peptide

The iodoacetylated peptidyl MUC1 (Tn) resin (360 mg) was swollen in CH$_3$CN (750 µl) for 2h. DIEA (9 µl, 0.06 mmol) was added followed by the auxiliary (8) (39.7 mg, 0.06 mmol, in 300 µl CH$_3$CN). The suspension was stirred at 28°C for 18h, then the supernatant was removed (the unreacted auxiliary was recovered by recrystallization) and the resin washed with DMF, DCM and MeOH. A test cleavage was performed to check the efficiency of the S$_2$N$_2$ reaction: 10 mg of the resin were incubated with a mixture of TFA/TIS/H$_2$O (92.5:5:2.5) for 2h at room temperature. The free peptide was precipitated by addition of cold diethyl ether and centrifugation. Finally, the supernatant was removed and the peptide washed with cold diethyl ether, dissolved in a mixture of H$_2$O/CH$_3$CN 1:1 and analyzed by RP-HPLC and ESI-MS: the major peak in the chromatogram correspond to the desired product, some unreacted iodoacetylated peptide could still be detected, together with a few side products. After test cleavage, the remaining Aux-peptidyl resin conjugate was directly used in the following PEGylation step.

PEGylation of Aux(Fmoc)MUC1 peptidyl resin and cleavage from the solid support

Aux(Fmoc)MUC1 peptidyl resin was swollen in DMF for 1h. The Fmoc protecting group on the auxiliary was removed by incubating the peptidyl resin twice with a solution of piperidine (20% v/v in DMF) for 3 min and 7 min respectively. In the meanwhile Fmoc-PEG$_2$-$\mu$COOH (130 mg, 83.8 µmol, 1.37 eq) was dissolved in a solution of HATU (149.5 µl (0.5M in DMF/CH$_3$CN 3:2), 71.9 µmol, 1.19 eq) and then DIEA (26 µl, 149.6 µmol, 2.5eq) was added. The deprotected peptidyl resin was washed with DMF and then incubated with the solution of activated PEG for 20h at room temperature under shaking. The supernatant was removed and the resin was washed extensively with DMF, DCM and
MeOH and dried under vacuum. The reaction was checked by performing a test cleavage. The PEGylated auxiliary-MUC1 conjugate was cleaved from the resin: the peptidyl resin was incubated with a mixture of TFA/TIS/H₂O (1.5 ml, 92.5:5:2.5) for 3 h at room temperature. The free conjugate was precipitated by addition of cold diethyl ether and recovered by centrifugation. After washing with diethyl ether and centrifugation, the conjugate was dissolved in CH₃CN/H₂O 1:1 (1% TFA) and freeze-dried. Pure Aux-MUC1 (19 mg, 4.9 μmol, 34% yield) was obtained after purification of the crude product by RP-HPLC (C4 column, gradient: 5 to 25% buffer B in buffer A in 5 min, then 25% to 55% buffer B in buffer A in 45 min, flow-rate: 20 ml/min, buffer A: ddH₂O + 0.05% TFA, buffer B: CH₃CN + 0.05% TFA).

PEGylation of Aux(Fmoc)MUC1(Tn) peptidyl resin, cleavage from the solid support and glycan deprotection in solution

Aux(Fmoc)MUC1(Tn) peptidyl resin was swollen in DMF for 1 h. The Fmoc protecting group on the auxiliary was removed by incubating the peptidyl resin twice with a solution of piperidine (20% v/v in DMF) for 3 min and 7 min respectively. In the meanwhile Fmoc-PEG₂₋COOH (146 mg, 94.8 μmol, 1.37 eq) was dissolved in a solution of HATU (160 μl (0.5M in DMF/CH₃CN 3:2), 80 μmol, 1.19 eq) and then DIEA (29.7 μl, 171 μmol, 2.5eq) was added. The deprotected peptidyl resin was washed with DMF and the incubated with the solution of activated PEG for 20 h at room temperature under shaking. The supernatant was removed and the resin was washed extensively with DMF, DCM and MeOH and dried under vacuum. The reaction was monitored by performing a test cleavage. The PEGylated auxiliary-MUC1 conjugate was cleaved from the resin: the peptidyl resin was incubated with a mixture of TFA/TIS/H₂O (2 ml, 92.5:5:2.5) for 3 h at room temperature. The free conjugate was precipitated by addition of cold diethyl ether and collected by centrifugation. After washing with diethyl ether and centrifugation, the conjugate was dissolved in CH₃CN/H₂O 1:1 (1% TFA) and freeze-dried. The crude product was dissolve in a 8% v/v solution of H₂NNH₂·H₂O in CH₃CN (3 ml), stirred at room temperature for 18 h and then freeze dried (monitoring by HPLC-MS: the glycan was completely deprotected, a side product could be detected accounting for the Aux-MUC1 (Tn) conjugate without Fmoc on the terminal amine of the PEG chain). The desired pure Aux-MUC1 (Tn) was obtained (9.2 mg, 2.3 μmol, 4% yield, based on peptide synthesis scale) after purification of the crude by RP-HPLC (C4 column, gradient: 5-45% buffer B in buffer A in 45 min, flow-rate: 20 ml/min, buffer A: ddH₂O + 0.05% TFA, buffer B: CH₃CN + 0.05% TFA).

Synthesis of VTSAPDTRPAPGSTAPPAH-SR, MUC1-SR peptide thioester.
MUC1 peptidyl resin (388mg, ca 78 µmol peptide) was swollen in DMF for 1h. N-terminal Fmoc protecting group was removed by incubating the resin twice with a 20% v/v solution of piperidine in DMF (3 and 7 min respectively). The resin was washed with DMF, DCM and MeOH and dried, then it was incubated with a 5% (v/v) \( \text{H}_2\text{NNH}_2\cdot\text{H}_2\text{O} \) in \( \text{CH}_3\text{CN/H}_2\text{O} \) 1:1 (3ml) for 18h under shaking. The supernatant was removed and the resin washed with \( \text{CH}_3\text{CN/H}_2\text{O} \) 1:1. The collected solutions containing the desired fully protected peptide hydrazide (analysis by HPLC-MS) were neutralized with TFA and freeze-dried. The crude reaction was dissolved in TFA/TIS/H\(_2\)O (2 ml, 92.5:5:2.5) and shaken at room temperature for 2h. The unprotected peptide hydrazide was precipitated with cold diethyl ether and recovered by centrifugation. After washing with diethyl ether and centrifugation, the crude was dissolved in \( \text{H}_2\text{O/CH}_3\text{CN} \) 1:1 and freeze-dried.

The crude peptide hydrazide was dissolved in 5.5 ml of buffer 1 (6M Gnd-HCl, 0.2M Na\(_2\)HP0\(_4\), pH 3) at -10°C. A 0.2M aqueous solution of NaN\(_3\) (1.2ml), was added dropwise and the solution was stirred at -10°C for 30 min. A solution of MesNa (0.45M in buffer 2 (6M Gnd-HCl, 0.2M Na\(_2\)HP0\(_4\), pH 7.5), 440 µl) was added and the reaction mixture was stirred at -10°C for 10 min and then at room temperature for 20 min. The pure thioester (21 mg, 11 µmol, 14% yield,) was obtained after purification by rp-HPLC (the reaction mixture was directly used for injection without any dilution. C4 column, gradient 5-30% buffer B in buffer A in 45 min, flow: 20 ml/min. 6.7 mg of hydrolyzed peptide thioester were recovered as well).

**Synthesis of Aux-MUC1 (Tn)-NNH\(_2\)**

Aux-MUC1 (Tn) peptidyl resin (50mg) was swollen in 5% (v/v) \( \text{H}_2\text{NNH}_2\cdot\text{H}_2\text{O} \) in MeOH (1ml) and shaken at room temperature for 18h. The supernatant was collected, the resin was washed with MeOH (0.2 ml) and the washing solution was added to the supernatant. The solution was neutralized with acetic acid and then the solvent was evaporated under reduced pressure. The crude was dissolved in buffer A (ddH\(_2\)O + 0.1% TFA, 0.4 ml) and freeze-dried. The crude lyophilized product was dissolved in TFA/TIS/H\(_2\)O (0.5 ml, 92.5:5:2.5) and rotated at room temperature for two hours. The deprotected Aux-MUC1 (Tn)-NNH\(_2\) was precipitated by addition of cold diethyl ether and recovered by centrifugation. After washing with cold diethyl ether and recovery by centrifugation, the crude product was dissolved in \( \text{H}_2\text{O/CH}_3\text{CN} \) 1:1 and freeze-dried. Purification by RP-HPLC allowed recovery of pure AuxMUC1(Tn)-NNH\(_2\) (0.5 mg, 4% yield) and of pure Aux(NH\(_2\))-MUC1 (Tn)NNH\(_2\) (free amine on the PEG due to partial removal of Fmoc by hydrazine, 0.9 mg, 7% yield). Both auxiliary-peptide hydrazide conjugates were used efficiently in glycosylation reaction and converted to the corresponding peptide thioesters (see description below).
Recombinant expression and purification of human C1GALT1 and ST3GAL1

Human C1GALT1 (Uniprot Q9NS00, residues 34-363), C1GALT1C1 (Uniprot Q96EU7, residues 29-318), and ST3GAL1 (Uniprot Q11201, residues 51-340) were generated as soluble secreted fusion proteins in HEK293 suspension cultures by procedures similar to those described for ST6GAL1. Briefly, the respective protein coding regions were amplified from human Mammalian Gene Collection clones to truncate NH2-terminal transmembrane sequences and append an NH2-terminal fusion peptide of a tobacco etch virus (TEV) protease cleavage site and flanking Gateway attB recombination site sequences to facilitate Gateway cloning using a two-step adapter PCR method. PCR products were gel purified and cloned into the pDONR221 plasmid vector using Gateway BP Clonase II (Invitrogen) and confirmed by DNA sequencing. Expression constructs were generated using the Gateway LR Clonase recombination with the pDONR221 entry clones a Gateway adapted version of the pGen2 mammalian expression vector (pGen2-DEST). The resulting expression constructs encode fusion proteins comprised of an NH2-terminal signal sequence, an 8xHis tag, an AviTag recognition site, the "superfolder" GFP coding region, the 7 amino acid recognition sequence of the tobacco etch virus (TEV) protease, followed by the corresponding enzyme catalytic domain. Recombinant enzyme expression was accomplished by transient transfection of HEK293f cells (FreeStyle 293-F cells, Life Technologies, Grand Island, NY) as previously described. In the case of C1GALT1, secreted expression of the enzyme was facilitated by co-expression with the selective chaperone C1GALT1C1. The cultures were harvested, clarified by centrifugation, and the recombinant product was purified by Ni-NTA Superflow (Qiagen, Valencia, CA) chromatography as described. The eluted recombinant enzyme preparation was concentrated to ~1 mg/ml by ultrafiltration with a 10-kDa molecular mass cutoff membrane (Millipore, Billerica, MA) and used directly for enzymatic modification of glycopeptide substrates.

General procedure for the galactosylation of Aux-peptide conjugates with C1GalT1

The conjugate (Aux-MUC1(Tn)) or Aux-MUC1(Tn)-NHNH₂ (with or without Fmoc), reaction scale 50-300C°g, final concentration (fc): 0.24 mM) was dissolved in an aqueous solution of TRIS buffer (fc: 75mM, pH7.5), Triton X-100 (fc: 0.06%), MnCl₂ (fc: 10 mM), UDP-Gal (fc: 2mM) at room temperature. The enzyme, human C1GalT1, was added and the reaction was shaken at 37°C for 6h (reaction monitored by HPLC-MS). The reaction mixture (50µl) was added to 450µl of EtOH. Et₂O (1ml) was added and the solution was incubated for 18h at -80°C. The conjugate carrying a T antigen was collected by centrifugation (15000g x 30min), dried under the fume-hood (30 min at room temperature) and used in the next glycosylation step without any further purification (Recovery: 95%
(Aux-MUC1(T), Figure 20B and C), 81% (Aux-MUC1(Tn)-NHNH$_2$), and 78% (Aux(NH$_2$)-MUC1(Tn)-NHNH$_2$) (Figure 21).

General procedure for the sialylation of Aux-peptide conjugates with C1GalT1

To an aqueous solution of the recovered conjugate (Aux-MUC1(T) or Aux-MUC1(T)-NHNH$_2$ (with or without Fmoc), reaction scale 50-300 µg, final concentration (fc): 0.24 mM) MgCl$_2$ (0.2M) and CMP-Neu5Ac (fc 2 mM) were added, followed by the enzyme, human ST3Gal1, and the reaction mixture was incubated for 4h at 37°C (reaction monitored by HPLC-MS) under gentle shaking. The reaction mixture (50 µl) was added to 450 µl of EtOH. Et$_2$O (1 ml) was added and the solution was incubated for 18h at -80°C. The conjugate carrying a sialyl-T antigen was recovered by centrifugation (15000g x 30min), dried under the fume-hood (30 min at room temperature) and, in the case of Aux-MUC1(ST), directly used in the ligation reaction. Recovery: 90% (Aux-MUC1(ST), Figure 20B and C), 93% (Aux-MUC1(ST)-NHNH$_2$) and 80% (Aux(NH$_2$)-MUC1(ST)-NHNH$_2$) (Figure 21).

Synthesis of Aux-MUC1 (ST)-SR

The conjugate Aux-MUC1 (ST)-NHNH$_2$ recovered by precipitation and centrifugation after glycosylation was dissolved in buffer 1 (6M Gnd-HCl, 0.2M Na$_2$HPO$_4$, pH 3, 2.8 µl) at -15°C. A 0.05M NaN02aq solution (0.4 µl, 1.3 eq) was added and the solution incubated at -15°C for 40 min. A 0.32 M solution of ascorbic acid (0.25 µl, 5 eq) was added followed by 1 µl of a 0.06M solution of MesNa in buffer 2 (6M Gnd-HCl, 0.2M Na$_2$HPO$_4$, pH 7.5) and stirred at room temperature for 20 min. The solution has been diluted with 60 µl of ddH$_2$O and directly used for RP-HPLC analysis and purification (63% conversion). The collected material was used for ESI analysis (direct injection).

General procedure for the Native Chemical Ligation of auxiliarp-peptide conjugates

The auxiliary-peptide conjugate (0.08-0.25 µmol) was dissolved in degassed ligation buffer (0.2M NaPi buffer pH 8.5, 35 mM TCEP, final pH 7.5. Concentration of conjugate ranging from 5 mM to 8.5 mM) and the solution was shaken at 24°C for 5h to allow deprotection of the thiol group (in the case of Aux-MUC1(ST) 36h incubation at 30°C and further addition of TCEP (20 eq in total) were needed to achieve deprotection of the tert-butylidisulfanyl group (Figure 23)). The solid peptide thioester (2.5-3.5 eq, eventually addition of small aliquot if hydrolysis occurred) was added and the reaction mixture was incubated at 30°C for 1-2 days (reaction monitored by RP-HPLC-MS) under gentle shaking. The efficiency of the reaction was estimated by integration of the peaks area of the HPLC chromatograms. The ligated product was submitted to photocleavage without further purification. In one of the assays done with Aux-MUC1 the product was purified by
HPLC (Aux-MUC1 : 0.23 μmol, obtained 0.11 μmol MUC1-Aux-MUC1, 48% isolated yield).

General procedure of UV photorelease of the auxiliary after ligation

The crude ligation solution was diluted with degassed H$_2$O or H$_2$O/CH$_3$CN 7:3 (final concentration: 0.5-1.0 mM) containing TCEP (1 eq) in a glass tube under argon and irradiated with a UV-A lamp (UV reactor) for 30 min. The unprotected ligation product was obtained in all cases in high conversion and recovered pure after HPLC purification.

Synthesis of VTSAP DT(GalNAc)RPAPGSTAP PAH-SR, MUC1(Tn$_2$)-SR

MUC1(Tn$_2$) peptidyl resin (108mg, ca 55 μmolpeptide) was swollen in DMF for 1h. N-terminal Fmoc protecting group was removed by incubating the resin twice with a 20% v/v solution of piperidine in DMF (3 and 7 min respectively). The resin was washed with DMF, DCM and MeOH and dried, then it was incubated with a 5% (v/v) H$_2$N NH$_2$H$_2$O in CH$_3$CN/H$_2$O 1:1 (2.5ml) for 18h under shaking. The supernatant was removed and the resin washed with CH$_3$CN/H$_2$O 1:1. The collected solutions containing the desired fully protected peptide hydrazide (analysis by HPLC-MS) were neutralized with TFA and freeze-dried. The crude was dissoloned in TFA/TIS/H$_2$O (2 ml, 92.5:5:2.5) and shaken at room temperature for 2h. The unprotected peptide hydrazide was precipitated with cold diethyl ether and recovered by centrifugation. After washing with diethyl ether and centrifugation, the crude was dissolved in H$_2$O/CH$_3$CN 1:1 and freeze-dried. The crude peptide hydrazide (6mg, 2.3 μmol) was dissolved in 113 μl of buffer 1 (6M GnHCl, 0.2M Na$_2$HP0$_4$·pH3) at -10°C. A 0.2M aqueous solution of NaNO$_3$ (22 l), was added dropwise and the solution was stirred at -10°C for 20min. A solution of MesNa (0.45M in buffer 2 (6M GnHCl, 0.2M Na$_2$HP0$_4$·pH7.5), 66μl) was added and the reaction mixture was stirred at -10°C for 5 min. A solution of ascorbic acid (2.2mM, 13 μl) was added and the solution was stirred at room temperature for 10 min. The pure thioester (2.1 mg, 1 μmol, 43% yield,) was obtained after purification by RP-HPLC (the reaction mixture was diluted with ddH$_2$O (total volume 3.5ml) and directly used for injection. C18 column, gradient 5-45% buffer B in buffer A in 45min, flow: 10ml/min).

Synthesis of peptide-conjugate MUCKTnYVAux-MUCKTn via Native Chemical Ligation

The reaction was carried out following the general procedure for NCL (pag 35). Aux-MUC1(Tn): 0.39 mg, 96 nmol; MUC1(Tn$_2$)-SR: 0.47 mg, 218 nmol, 2.3 eq. Total reaction volume: 12μl. Aux-MUC1(Tn) was first incubated only with TCEP solution. After 2h the starting material was completely deprotected and started forming a dimer (disulfide, peak at 27.74 min, Figure 28) that was reduced by addition of TCEP during ligation. The reaction was monitored by RP-HPLC and after 24h it did not proceed further. The pure
ligation product was obtained (0.14 mg, 23nmol, 24% isolated yield) after purification of the crude reaction mixture by RP-HPLC.

**Galactosylation of MUC1(Tn$_7$)-Aux-MUC1(Tn) with C1GalT1**

The isolated ligation product MUC1(Tn$_7$)-Aux-MUC1(Tn) was directly used in the galactosylation reaction following the general procedure for galactosylation described above. The only modification to the procedure was the use of a higher amount of UDP-Gal (4mM final concentration) because of the different stoichiometry of the reaction. TCEP (1 $\mu$L of a 70mM aqueous solution) was added to reduce the disulfide that the free thiol on the auxiliary was forming. After 4h the starting material was converted into the desired product (Figure 29, left). Ethanol (393 $\mu$L) was added and then, after vortexing, diethyl ether (870 $\mu$L). The solution was incubated at -80°C overnight. After centrifugation (30min x 15000g, 5°C) the supernatant was removed and the solid product dried at room temperature under the fume hood. The recovered MUC1(T$_7$)-Aux-MUC1(T) was dissolved in ddH$_2$O (19$\mu$L + 1$\mu$L of 70mM TCEP solution) and analyzed by RP-HPLC (Figure 29, right). Recovery after precipitation: 49%.

**Sialylation of MUC1(T$_2$)-Aux-MUC1(T) with ST3Gal1**

To an aqueous solution of the recovered freeze-dried conjugate MUC1(T$_7$)-Aux-MUC1(T) (2$\mu$L, final concentration (fc): 0.35 mM), Tris HCl (pH 7.5, fc: 50mM), MgCl$_2$ (fc: 20mM), CMP-Neu5Ac (fc: 4mM) and DTT (fc: 40mM) were added, followed by the enzyme, human ST3Gal1 (final reaction volume: 20 $\mu$L), and the reaction mixture was incubated for 18h at 37°C (reaction monitored by HPLC-MS) under gentle shaking. Ethanol (225 $\mu$L) was added to the reaction mixture followed by diethyl ether (500 $\mu$L). The solution was incubated at -80°C overnight. After centrifugation (30min x 15000g, 5°C) the supernatant was removed and the solid product dried at room temperature under the fume hood. The recovered MUC1(ST7)-Aux-MUC1(ST) was dissolved in ddH$_2$O and analyzed by RP-HPLC (Figure 30B) and HR-ESI-TOF-MS (Figure 30A and C). The mass analysis demonstrated that the starting material was efficiently converted into the desired product. The mass of a conjugate carrying one sialyl-T and one T antigen (partial sialylation) can be also detected (Figure 31C and 32). Considering that the difference between the product and the partially sialylated conjugate consist in one sugar unit and that the sugar in these conjugate has a minimal influence on the ionization of the compound (as it can be seen in Figure 20A, UV trace and ionization intensity correlate very well), it can be estimated that the partially sialylated conjugate represent less than 10% of the mixture. The masses of the conjugates MUC1(T$_7$)-Aux-MUC1(T) and MUC1(Tn$_7$)-Aux-MUC1(Tn) are barely detectable indicating that the glycosylation reactions are almost quantitative. Recovery after precipitation: 38%.
References

WO2012/1 39777,


J. Offer, Native Chemical Ligation with Na Acyl Transfer Auxiliaries; Biopolymers 2010, 94, 530-541.

Claims

1. A compound comprising:

(A) a thiol moiety \( S \) comprising:
   (A1) at least one protected or deprotected thiol group, and
   (A2) at least one primary or secondary amino group,

wherein said thiol moiety \( S \) enables Native Chemical Ligation (NCL);

(B) a photolabile linker moiety \( L \); and

(C) at least one hydrophilic polymer moiety \( HP \),

wherein the photolabile linker moiety \( L \) is covalently bound to the at least one hydrophilic polymer moiety \( HP \) and to the thiol moiety \( S \), and wherein \( HP \) does not comprise primary amino groups or hydroxyl groups when the thiol moiety \( S \) is not covalently bound to a peptide moiety \( P_1 \).

2. The compound of claim 1, wherein said compound has the following structure of formula (I):

\[
\begin{align*}
\text{A} & \text{X} \text{Y} \text{Z} \text{L} \text{HP} \\
\text{AS} & \text{NHB} \\
\end{align*}
\]

wherein \( A \) represents a protecting group residue or hydrogen, preferably a protecting group residue selected from the group consisting of a \( \text{ie/f-butylsulfanyl} \) residue, a \( 3\text{-nitro-2-pyridylsulfanyl} \) moiety, a benzyl residue, a \( 4\text{-methylbenzyl} \) residue, a \( 4\text{-methoxybenzyl} \) residue, a \( 2,4,6\text{-trimethoxybenzyl} \) residue, a diphenylmethyl residue, a trityl residue, a \( \text{ie/tert-butyl} \) residue, an acetamidomethyl residue, a trimethylacetamidomethyl residue, a \( 9\text{-fluorenylmethyl} \) residue, an allyloxycarbonylaminomethyl residue, and a \( 9\text{H-xanthen-9-yl} \) residue, in particular a \( \text{tert-butylsulfanyl} \) residue;
wherein X represents a spacer moiety or a bond, preferably a substituted or unsubstituted C\(_{1-5}\)-alkylen spacer moiety selected or a bond, in particular a -CH\(_2\)- spacer;

wherein B represents hydrogen, a peptide moiety P1, an amino acid moiety, or a moiety activating the amino group, in particular hydrogen;

wherein Y represents a bond or a spacer moiety, in particular a bond;

wherein L represents the photolabile linker moiety L;

wherein Z represents a spacer moiety or a bond, in particular a spacer moiety; and

wherein HP represents a hydrophilic polymer moiety.

3. The compound of any of claims 1 or 2, wherein the photolabile linker moiety L has the following structure of formula (II):

\[
\begin{array}{c}
\text{NO}_2 \\
\text{R} \\
\text{O} \\
\text{#} \\
\text{#'}
\end{array}
\]

wherein # represents the binding site to the thiol moiety S via Y, wherein Y represents a bond or a spacer moiety, in particular a bond;

wherein #' represents the binding site to the at least one hydrophilic polymer moiety HP via Z, wherein Z represents a spacer moiety or a direct bond, in particular a spacer moiety; and

wherein R\(^1\) may be any residue of up to 40 carbon atoms, preferably a residue selected from the group consisting of -OR\(^{11}\), hydrogen, halogen, -R\(^{11}\), -0-C(0)-R\(^{11}\), -NR\(^{11}\)R\(^{12}\), -SiR\(^{11}\)R\(^{12}\)R\(^{13}\), -Ge R\(^{11}\)R\(^{12}\)R\(^{13}\), -SR\(^{11}\), -SOR\(^{11}\), and -SO\(_{2}\)R\(^{11}\), wherein each of R\(^{11}\), R\(^{12}\) and R\(^{13}\) represents a residue selected from the group consisting of a substituted or unsubstituted linear, branched or cyclic C\(_{1-20}\) alkyl residue, hydrogen, -OH, -NH\(_2\), a halogen, a substituted or unsubstituted linear, branched or cyclic C\(_{1-20}\) heteroalkyl, C\(_{2-20}\)-alkenyl or C\(_{2-20}\)-heteroalkenyl residue and a substituted or unsubstituted C\(_{6-20}\) aryl residue, C\(_{6-20}\) heteroaryl residue, C\(_{7-32}\) arylalkyl residue, C\(_{6-20}\)
31 heteroarylalkyl residue, C₈-3₃ alkylarylalkyl residue, a C₇-3₂ heteroalkylarylalkyl residue, and a dye, more preferably wherein R¹ is selected from the group consisting of -OCH₃, -OCH₂CH₃, -0(CH₂)₂CH₃, -OCH(CH₃)₂ or hydrogen, in particular -OCH₃.

4. The compound of any of claims 1 to 3, wherein the at least one hydrophilic polymer moiety HP is/are selected from the group consisting of
(a) polyethylene glycol (PEG) or derivatives thereof;
(b) polyethylene imine (PEI) or derivatives thereof;
(c) polyacrylic acid or derivatives thereof, preferably a polymer of methacrylic acid or derivative thereof, more preferably a polymer of hydroxypropyl methacrylate or hydroxyethyl methacrylate (HEMA) or derivatives thereof, in particular a polymer of N-(2-hydroxypropyl) methacrylamide (HPMA) or derivatives thereof;
(d) polysaccharide(s) or derivatives thereof,
(e) hydrophilic polypeptide(s)
(f) lipopolysaccharide(s); and/or
(g) conjugate(s) or blockpolymer(s) comprising two or more of the above,
preferably the hydrophilic polymer moiety is PEG or derivatives thereof, more preferably PEG or a derivative thereof of between 5 to 50 PEG units, even more preferably of between 10 to 40 PEG units, even more preferably of between 15 to 35 PEG units, in particular of between 20 to 30 PEG units.

5. The compound of any of claims 2 to 4, wherein the spacer Z has the following structure of formula (III)

#"-(CH₂)ₙ-Z'#"  (III),

wherein #" represents the binding site to the photolabile linker moiety L or to the hydrophilic polymer moiety HP, in particular to the photolabile linker moiety L;

wherein #" represents the binding site to the hydrophilic polymer moiety HP or to the photolabile linker moiety L, in particular to the hydrophilic polymer moiety HP;

wherein n represents an integer from 1 to 10, preferably wherein n is 1, 2 or 3, in particular 3;
wherein $Z'$ represents a bivalent group or a bond, preferably wherein $Z'$ is selected from the group consisting of -CO-NH-(CH$_2$)$_m$-NH-CO-, -NH-CO-, -NH-(CH$_2$)$_m$-NH-CO-, -CO-NH-(CH$_2$)$_m$-NH-CO-, -NH-CO-(CH$_2$)$_m$-NH-CO-, -CO-NH-, -0-CO-, -0-(CH$_2$)$_m$-CO-, -0-(CH$_2$)$_m$-CO-, -0-(CH$_2$)$_m$-CO-, -0-(CH$_2$)$_m$-CO-, -0-(CH$_2$)$_m$-CO-, -0-(CH$_2$)$_m$-CO-, and -0-(CH$_2$)$_m$-CO-; and

wherein $m$ represents an integer from 1 to 10, preferably wherein $m$ is 1, 2 or 3, in particular 2,

preferably wherein $Z'$ is selected from the group consisting of -CO-NH-(CH$_2$)$_m$-NH-CO-, -NH-CO-, and NH-(CH$_2$)$_m$-NH-CO-, in particular wherein $Z'$ is -CO-NH-(CH$_2$)$_m$-NH-CO-.

6. The compound of any of claims 1 to 5, wherein said compound has the following structure of formula (IV):

![Structure (IV)](image)

wherein $R^1$ represents a linear or branched C$_{1-20}$ alkoxy moiety or hydrogen, preferably -OCH$_3$, hydrogen, -OCH$_2$CH$_3$, -0(CH$_2$)$_2$CH$_3$, or -0CH(CH$_3$)$_2$, in particular -OCH$_3$;

wherein $Z$ represents a direct bond or a spacer moiety, in particular a spacer moiety as defined in claim 5;

wherein $A$ represents a protecting group residue or hydrogen; and

wherein HP represents a hydrophilic polymer moiety,

in particular wherein said compound bears one of the following structures (V)-(VII):
wherein n represents an integer from 1 to 10, preferably wherein n is 1, 2 or 3, in particular 3; and

wherein m represents an integer from 1 to 10, preferably wherein m is 1, 2 or 3, in particular 2.

7. The compound of any of claims 2 to 5, wherein B represents a peptide moiety P1, preferably a peptide moiety P1 characterized in that:

(a) it bears an N-terminal moiety of the structure of formula (VI I):

\[ #^v \text{-CH}_2\text{R}^2\text{-CO-}#^w \]  

(VII)

wherein \(#^v\) represents the binding site to the compound of any of the preceding claims;

wherein \(#^w\) represents the binding site to the remaining moieties of peptide moiety P1; and

wherein R^2 represents hydrogen or a substituted or unsubstituted C_{1-5}-alkyl residue, preferably hydrogen, -CH_3, -CH_2-OH or -CHOH-CH_3, in particular hydrogen

(b) it comprises less than 150 amino acid moieties, preferably less than 100 amino acid moieties, more preferably less than 50 amino acid moieties, even more preferably less than 40 amino acid moieties, even more preferably less than 35 amino acid moieties, more preferably less than 30 amino acid moieties, in particular less than 25 amino acid moieties; and/or
(c) it bears one or more non-peptidic modification(s), in particular wherein said peptide moiety P1 is glycosylated.

8. The compound of claim 7, wherein said compound has the following structure of formula (IX):

\[
\text{rest of } P1 \quad \text{NO}_2 \\
\text{HP}
\]

wherein R represents a linear or branched C_{1-20} alkoxy moiety or hydrogen, preferably -OCH₃, hydrogen, -OCH₂CH₃, -O(CH₂)₂CH₃, or -OCH(CH₃)₂, in particular -OCH₃;

wherein Z represents a direct bond or a spacer moiety, in particular a spacer moiety as defined in claim 5;

wherein A represents a protecting group residue or hydrogen;

wherein HP represents a hydrophilic polymer moiety; and

wherein the rest of P1 represents the peptide moiety P1 without its N-terminal moiety -CHR₂-CO-,

in particular wherein said compound bears one of the following structures (X)-(XII):

\[
\text{rest of } P1 \quad \text{rest of } P1 \quad \text{rest of } P1 \\
\text{HP} \quad \text{HP} \quad \text{HP}
\]

wherein n represents an integer from 1 to 10, preferably wherein n is 1, 2 or 3, in particular 3; and
wherein \( m \) represents an integer from 1 to 10, preferably wherein \( m \) is 1, 2 or 3, in particular 2.

9. A method for the production of a peptide conjugate P12 comprising a peptide moiety P1 and a second moiety P2 covalently bound with another via an amide bond, said method comprising the steps of:

(i) providing a compound of any of claims 1 to 5, preferably wherein the thiol group of the thiol moiety S is protected and the amino group of the thiol moiety S is hydrogen;

(ii) adding a peptide moiety P1 to the amino group covalently bound to the thiol moiety S of said compound of step (i) thereby forming a peptide conjugate PC1, preferably wherein said peptide conjugate PC1 bears the structure of formula (XIII):

\[
\begin{align*}
\text{AS}^X \text{Y} & \text{NH} \text{P1} \\
\text{L} \text{Z} & \text{HP}
\end{align*}
\]

wherein A, X, Y, L, Z, P1 and HP are defined as in any of claims 1 to 8, in particular wherein the structure is further defined as in any of claims 7 or 8;

(iii) cleaving off the protecting group from the thiol group of the thiol moiety S the conjugate PC1 so that said thiol moiety S bears an unprotected thiol group; and

(iv) conjugating a second moiety P2 comprising an thioester with said peptide conjugate PC1 obtained from step (iii) by means of Native Chemical Ligation, thereby forming a peptide conjugate PC12.

10. The method of claim 9, further comprising the step (iiA) of modifying the peptide moiety P1 comprised in the compound obtained from step (ii) by means of one or more enzyme(s) prior to subjecting it to step (iii) or (iv), in particular wherein said one or more enzyme(s) is/are glycosyltransferase(s) and said step of modifying the peptide moiety P1 is glycosylating.
11. The method of claim 10, further comprising the step of removing the one or more enzyme(s) and/or reaction component(s) (step (ii)), in particular wherein the step of removing the one or more enzyme(s) and/or reaction component(s) is performed by chromatography and/or by precipitating the polymer-conjugated peptide, preferably by gel permeation chromatography (GPC) and/or precipitating the polymer-conjugated peptide, more preferably by a combination of GPC and precipitating the polymer-conjugated peptide or by precipitating as the only purification step, in particular by precipitating as the only purification step.

12. The method of at least one of claims 9 to 11, wherein the step (ii) of adding a peptide moiety P1 to the amino group covalently bound to the thiol moiety S of the compound is:

(iia) providing an unbound peptide P1 and conjugating said peptide P1 to the compound of step (i) that is optionally is attached to a solid phase; or

(iiib) synthesizing the peptide moiety P1 on the compound of step (i) by sequentially conjugating amino acid building blocks, preferably wherein said compound is attached to a solid phase, in particular by means of Fmoc- or Boc-based solid phase peptide synthesis (SPPS).

13. The method of at least one of claims 9 to 12, wherein said method comprises the further step (v) of irradiating the compound comprising the peptide conjugate P12 obtained from step (iv) thereby cleaving said peptide conjugate P12 off the photolabile linker L and the hydrophilic polymer moiety HP.

14. The method of at least one of claims 9 to 13, further comprising the step (vi) of purifying the peptide conjugate P12 obtained from step (iv) or (v), preferably by one or more chromatographic method(s) and/or one or more precipitation step(s), in particular gel permeation chromatography (GPC), fast protein liquid chromatography (FPLC), reversed phase high performance liquid chromatography (RP-HPLC) and/or diethyl ether precipitation.

15. The method of at least one of claims 9 to 14, wherein the second moiety P2 is a peptide moiety, a peptide analogue moiety, a protein moiety, a lipid moiety, a polymeric scaffold moiety, a surface of a device, an oligonucleotide or a dye moiety, in particular is a peptide moiety.
Scheme 3
Fig. 2

Absorbance (mAU)

800
700
600
500
400
300
200
100
0

Aux-MUC1(ST) 26.26
Aux-MUC1(T) 26.36
Aux-MUC1(Tn) 26.50

Aux-MUC1(ST) calculated mass: 4513 Da
Aux-MUC1(T) calculated mass: 4222 Da
Aux-MUC1(Tn) calculated mass: 4059 Da
Fig. 4

Absorbance (mAU)

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Absorbance Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>200</td>
</tr>
<tr>
<td>10</td>
<td>400</td>
</tr>
<tr>
<td>15</td>
<td>600</td>
</tr>
<tr>
<td>20</td>
<td>800</td>
</tr>
<tr>
<td>25</td>
<td>1000</td>
</tr>
<tr>
<td>30</td>
<td>1200</td>
</tr>
<tr>
<td>35</td>
<td>1400</td>
</tr>
</tbody>
</table>

- t = 0
- t = 24h
- t = 36h

Mass spectra:

- Aux-MUC1-OH
- Aux-MUC1(Tn)
- MUC1-SR
- MUC1-OH
- MUC1-Aux-MUC1(Tn)
Fig. 7

Absorbance (mAU)

Aux-MUC1(Tn)-NHNH₂ after precipitation

Aux-MUC1(ST)-NHNH₂

Aux-MUC1(T)-NHNH₂ after precipitation

Aux-MUC1(T)-NHNH₂

Aux-MUC1(Tn)-NHNH₂

Aux-MUC1(ST)-NHNH₂
Fig. 20C

Aux-MUC1(ST)
calculated mass: 4513 Da

Aux-MUC1(T)
calculated mass: 4222 Da

Aux-MUC1(Tn)
calculated mass: 4059 Da
Fig. 21

Absorbance (mAU)

| Time (min) | MUC1 Compound
|------------|----------------|
| 0-5        | Aux_{nmsc}MUC1(ST)-NHNH$_2$
| 5-10       | Aux_{nmsc}MUC1(ST)-NHNH$_2$
| 10-15      | Aux_{nmsc}MUC1(T)-NHNH$_2$
| 15-20      | Aux_{nmsc}MUC1(T)-NHNH$_2$
| 20-25      | C1GalT1 t=6h
| 25-30      | Aux_{nmsc}MUC1(Tn)-NHNH$_2$
| 30-35      | C1GalT1 t=0

After precipitation:

- ST3Gal1 t=4h

M/z values for each compound indicate the mass-to-charge ratio for each species of interest.
Fig. 29A
Fig. 30D

Calcd: 6671 Da
Found: 6671 Da

Absorbance (mAU)

after precipitation

Calcd: 6671 Da
Found: 6671 Da

Intensity (%)
Fig. 32A

MUC1 = VTSAPDTRPAPGSTAPPAH

Calcd: 4429.0956 Da
Found: 4429.1176 Da

1. NaNO₂
2. MesNa

1. C1GalT1
2. EtOH/Et₂O

1. ST3Gal1
2. EtOH/Et₂O

hv
A. CLASSIFICATION OF SUBJECT MATTER

INV. C07K1/02 C07K7/08

According to International Patent Classification (IPC), the following national classification and IPC codes are used:

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols):

C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched:

Electronic database consulted during the international search (name of database and, where practicable, search terms used):

1. EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>KAWAKAMI T ET AL: &quot;A photoremoveable ligating auxiliary for use in poly peptides synthesis&quot;, TETRAHEDRON LETTERS, PEGAMON, GB, vol. 44, no. 32, 4 August 2003 (2003-08-04), pages 6059-6061, XP004437524, ISSN: 0040-4039, DOI: 10.1016/50040-4039 (03)01463-1</td>
<td>1,4, 9, 12, 14, 15</td>
</tr>
<tr>
<td>A</td>
<td>Schemes 1, 2;, compounds 1, 6-10</td>
<td>2, 3, 5-8, 13</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:
  *A* document defining the general state of the art which is not considered to be of particular relevance,
  *E* earlier application or patent but published on or after the international filing date,
  *L* later document which may throw doubts on priority claim(s) one of which is cited to establish the publication date of another citation or other special reason (as specified),
  *O* document referring to an oral disclosure, use, exhibition or other means,
  *P* document published prior to the international filing date but later than the priority date claimed.

Date of the actual completion of the international search: 13 April 2016

Name and mailing address of the ISA: European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016

Date of mailing of the international search report: 20/04/2016

Authorized officer: Jetter, Sonya
<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>ROBERT K. MCGINTY ET AL: &quot;Chemical ly ubi qui ty lated hi stone H2B stimul ates hDotIL-medi ated intranucl eosomal methyl ati on&quot;, NATURE, vol. 453, no. 7196, 30 April 2008 (2008-04-30), pages 812-816, X055264580, United Kingdom ISSN: 0028-0836, DOI: 10.1038/nature06906 figure 1; compounds 1, 3, 4</td>
<td>1, 4, 9-12, 14, 15</td>
</tr>
<tr>
<td>X</td>
<td>CHIARA MARINZI ET AL: &quot;An o-nitrobenzyl scaffold for pepti de ligation: syntheses and applicati ons&quot;, BIOORGANIC &amp; MEDICINAL CHEMISTRY, vol. 12, no. 10, 1 May 2004 (2004-05-01), pages 2749-2757, XP055264293, GB ISSN: 0968-0896, DOI: 10.1016/j.bmc.2004.02.039 Schemes 1, 4;; page 2753, col umn 1, line 1 - line 29; figure 1; table 1</td>
<td>1, 4</td>
</tr>
</tbody>
</table>