Extending Chemoselective Ligation to Sugar Chemistry: Convergent Assembly of Bioactive Neoglycoconjugates

Francesco Peri*

Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milano, Italy

Abstract: Glycoproteins and glycolipids play central roles in human health and disease, and their mimetics are primary candidates for drug development. Our understanding of the molecular level of phenomena based on molecular recognition of oligosaccharides by specific receptors has taken tremendous advantage from their chemical synthesis, which provides homogeneous material not attainable from biosynthetic systems. This review summarizes chemoselective approaches for the assembly of neoglycoconjugates. The objective of these methods is to make glycoconjugate synthesis accessible to a broader community, thus accelerating progress in biotechnology.

Keywords: Chemoselective ligation, neoglycoproteins, neoglycopeptides, sugar mimetics, drug design.

INTRODUCTION

Glycoproteins and gycolipids are the two classes of molecules localized on cell surfaces, displaying a rich mosaic of carbohydrate structures in the extracellular environment. The molecular recognition of protein- and lipid-linked oligosaccharides by specific protein receptors plays a key role in several physiological and pathological events, such as cell-cell recognition and adhesion, cell growth and differentiation, development, lymphocyte migration, oncogenesis and viral infection [1]. The possible changes in anomeric configuration, in the regiochemistry of points of attachment, in states of oxidation or reduction, and in substitution of monosaccharide units composing oligosaccharides, provide a wide range of topographies which serve as an extremely large and sophisticated vocabulary assisting in the development of a large number of molecular recognition events.

There is great demand, for reasons of both fundamental research and pharmaceutical therapeutics, to have access to homogeneous glycoproteins both O-linked (serine, threonine or tyrosine α-glycosides) and N-linked (asparagine β-glycosides). The technique usually employed for glycoprotein production, recombinant over-expression, produces heterogeneous and organism-dependent glycoforms of the same protein, the peptide backbone being associated with an array of oligosaccharide structures [2]. In this review several chemical and chemo-enzymatic methods will be described, allowing glycosylated proteins to be obtained without recombinant technology, in a homogeneous, chemically pure form.

Glycolipids represent a second class of molecules of interest as candidates for drug development. As an example, glycosphingolipids (G_{M1} and G_{b3}) provide an anchoring point on cell surfaces to bacterial cholera and shiga toxins, respectively [3]. Glycosphingolipids (GSP) of the "globo" and "ganglio" series are also specific tumor-associated antigens [4]. Galactosylceramide (GalCer), expressed in human neural and intestinal tissues, binds to human immunodeficiency virus HIV-1, allowing it to enter cells that lack the surface protein CD4 [5]. Synthetic analogs of lipid A, part of lipopolysaccharides (LPS) that allows the anchoring to gram-negative bacteria membrane, and of its precursors-lipid X and lipid IVa, have been synthesized in order to obtain nontoxic molecules with endotoxin antagonistic activity as potential therapeutic agents to combat septic shock [6].

The observation that over-expression of certain types of carbohydrate structures on glycoprotein is strictly correlated with specific tumors inspired the development of carbohydrate-protein or peptide conjugates as chemically homogeneous or partially homogenous synthetic vaccines [7].

The development of drugs based on glycopeptides, glycoproteins and glycolipids is mainly based on molecules containing carbohydrate mimetics, a term indicating any carbohydrate derivative or other compound that has multiple hydroxyl groups and looks somewhat like a sugar, maintaining structural and functional properties of the native sugar [8].

This review focuses mainly on mimetics in which the glycosidic linkage (α− or β−O-glycoside linkage of glycoproteins and glycolipids and β−N-glycoside linkage of glycoproteins) has been substituted with non-native linkages as showed in Fig. (**1**). These mimetics, called neoglycoconjugates, have a number of advantages over their native counterparts as therapeutic agents. They are more stable toward endogenous hydrolytic enzymes; are generally more resistant to the gastric acid environment; and have improved bioavailability and reduced clearance rates. The introduction of an "artificial" glycoside bond in synthetic glycoconjugates is advantageous for two reasons: the pharmacokinetic properties of the molecules are generally improved, and it is possible to use convergent chemoselective strategies for the synthesis of the molecules.

*Address correspondence to this author at the Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milano, Italy; Tel: (+39)02-6448-3453; Fax: (+39) 02-6448-3565; E-mail: francesco.peri@unimib.it

1389-5575/03 $41.00+.00 © 2003 Bentham Science Publishers Ltd.

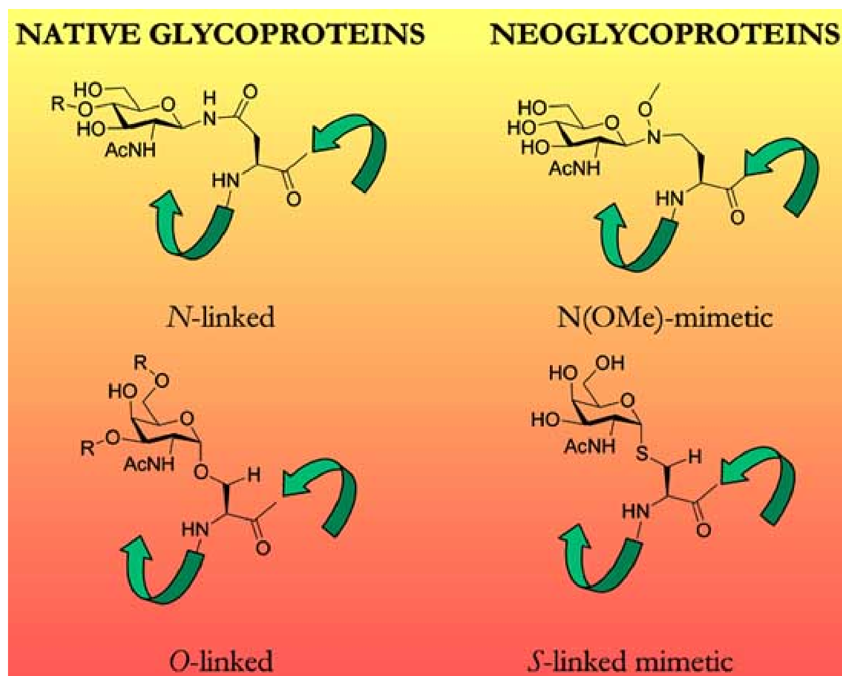

Fig. (1). *N*-linked and *O*-linked core saccharides found in mammalian glycoproteins and prototypical synthetic analogues (mimetics).

THE CHEMOSELECTIVE APPROACH TO NEOGLYCOCONJUGATE SYNTHESIS

Numerous methods exist for the synthesis of glycopeptides and glycoproteins: glycans are generally introduced into peptides during solid phase synthesis (SPPS) by means of glycosyl-amino acid building blocks [9]. This "linear" synthetic approach is generally laborious; requires extensive use of protecting groups; and allows the preparation of one glycoform per synthesis. The alternative "convergent" approach is based either on the enzymatic elongation of the oligosaccharide chain of a pre-synthesized

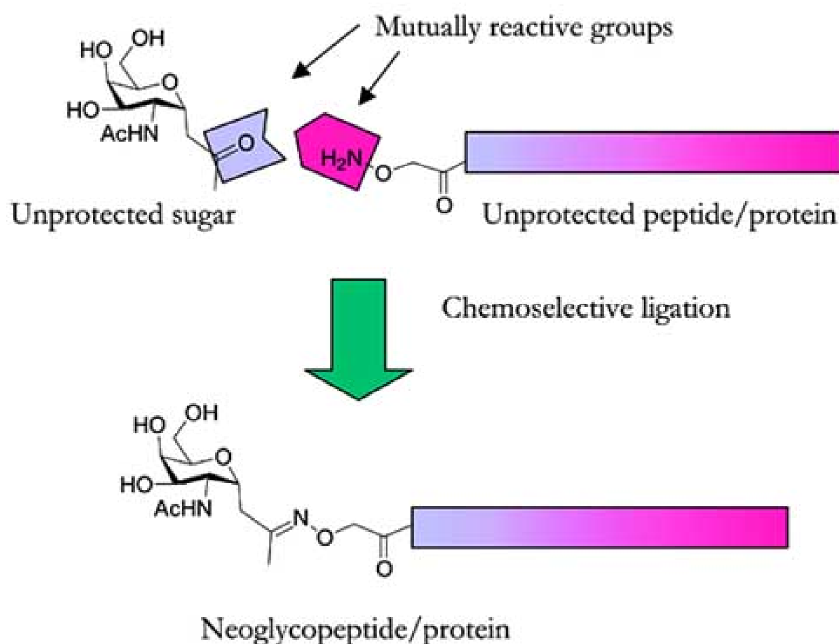

Fig. (2). Chemoseletive ligation between sugar and peptide in the case of oxime bond formation.

glycopeptide or glycosyl-amino acid [10,11], or on the conjugation of a fully elaborated, complex saccharide to short synthetic peptides [12].

Chemoselective ligation, first described by protein chemists as the coupling of two mutually and uniquely reactive groups in an aqueous environment, provides also access to complex neoglycoconjugates in an elegant and convergent way. In this technique, two uniquely reactive couples of functional groups (generally an electrophile and a nucleophile) are introduced in the peptide and sugar fragments, giving rise to selective covalent bond formation even in the presence of an array of other unprotected functional groups [Fig. (**2**)].

The chemoselective techniques so far developed for neoglycoconjugate synthesis fall into two broad categories: the first characterized by the reaction of a carbonyl group of a ketone or an aldehyde with strong nucleophiles, and the second by the addition of sulfur nucleophiles to a variety of electrophiles.

NEOGLYCOCONJUGATES BY CARBONYL-BASED CHEMOSELECTIVE CHEMISTRY

The absence of aldehydes and ketones on the side chains of the naturally occurring amino acids inspired a set of chemoselective reactions based on the condensation of the anomeric carbon of a reducing sugar (aldehyde group in the cyclic hemiacetal form) with a variety of non-natural nucleophile groups introduced into the peptide chain. Mutter and coworkers first demonstrated that the amino-oxyacetylated somatostatin analog RC-160 could be glycosylated in a chemoselective way after SPPS with a variety of unprotected mono- and oligosaccharides, obtaining oxime-linked neoglycopeptides with increased bioavailability [compound **1**, Fig. (**3**)] [13]. The main drawback of this method was the formation of the oxime linear form, quite different from the natural pyranose form, of the first attached sugar. In order to overcome this problem, an *N*-methylated amino-oxy group on the peptide moiety was developed, that allowed, after conjugation, the cyclic form of the first attached sugar to be maintained [compound **2**, Fig. (**3**)] [14,15]. The glycoside linkage of the attached sugar was obtained exclusively in the β-configuration, making these mimetics quite close structurally to natural *N*-linked glycoproteins. The same ligation strategy was employed to glycosylate with lactose the *N*-terminal segment of P-selectin glycoprotein ligand 1 (PSGL-1), in order to mimic the sialyl Lewis x-glycosylated threonine that is fundamental for efficient adhesion to P-selectin. The neoglycopeptide was further decorated by the chemoselective reaction of a cysteine side chain with iodoacetylated biotin [Fig. (**4**)] [16].

Dendritic cells (DCs)-targeted synthetic vaccines have been synthesized using oxime-based ligation chemistry. DCs are able to activate helper-T cells by a mechanism involving internalization of exogenous antigen peptides by macropinocytosis or specific receptor-mediated uptake, followed by presentation of peptide segments by the major histocompatibility class II molecules. In order to induce a DC-mediated immunological response against the T_n epitope that is associated with a large group of human malignant

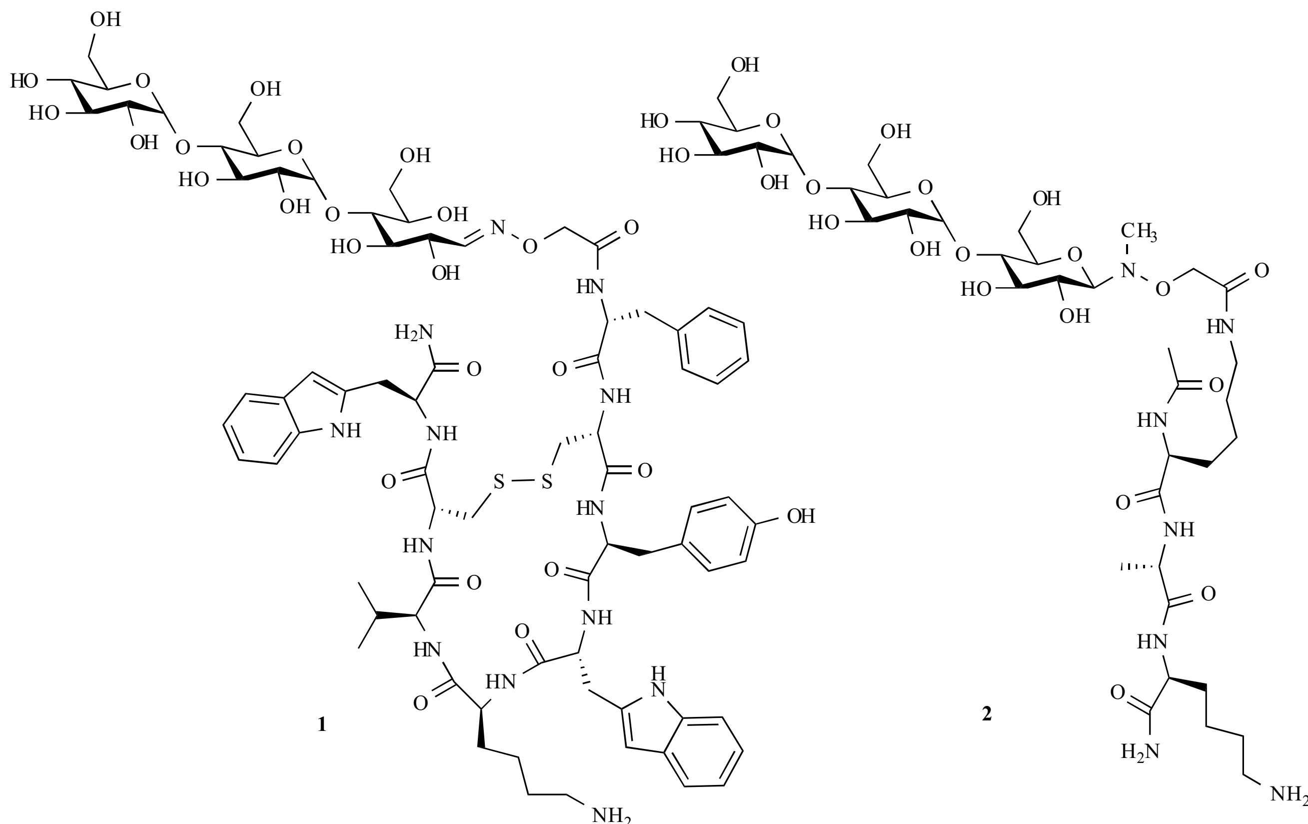

Fig. (3). Compound **1**: somatostatin analogue RC-160 linked to maltotriose through an oxime bond (with the first attached sugar in the open chain form); compound **2**: maltotriose has been conjugated to a tripeptide through an *N*-methyl amino-oxy linkage.

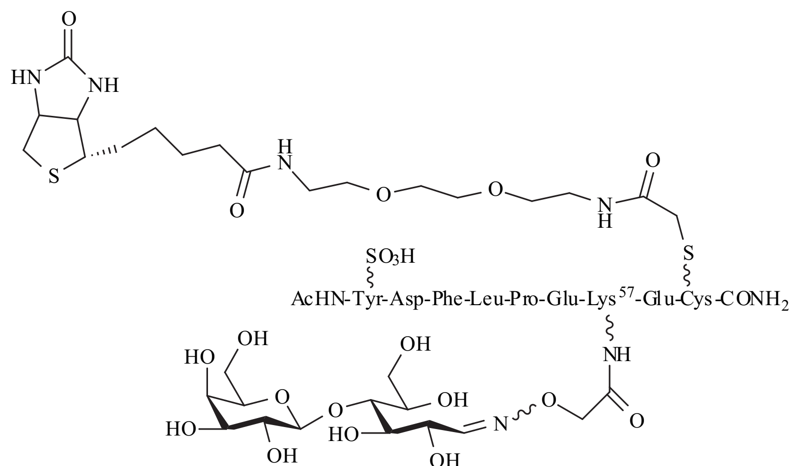

Fig. (4). Mimetic of the *N*-terminal segment of PSGL-1: the Thr^{57}(sialyl Lewis x) has been replaced by an oxime-linked lactosyl amino-oxyacetyl lysine.

tumors, one or two copies of the C-glycoside analog of *N*-acetyl-D-galactosamine (D-GalNAc), the sugar present in the T_n epitope, were conjugated by means of oxime bonds to the immunogenic peptide OVA $^{(327\text{-}339)}$, obtaining compound **3** [Fig. (**5**)] [17]. In this molecule, the C-glycoside mimics the α-anomeric configuration of the natural T_n epitope Thr/Ser(D-GalNAc). It was demonstrated *in vitro* that compound **3** was more efficiently internalized by DCs and consequently presented to T cells than its analog containing only one sugar unit, indicating that the C-saccharide epitope GalNAc either acted as an internalization agent by binding specific DCs surface sugar receptors, or promoted receptor clustering essential for DCs activation [18]. *In vivo* assays on C57BL/6 mice indicated that specific T and B cell responses against the tumor were induced by immunization with molecule **3**, and antibody production was higher than that obtained using the mono-glycosylated analog.

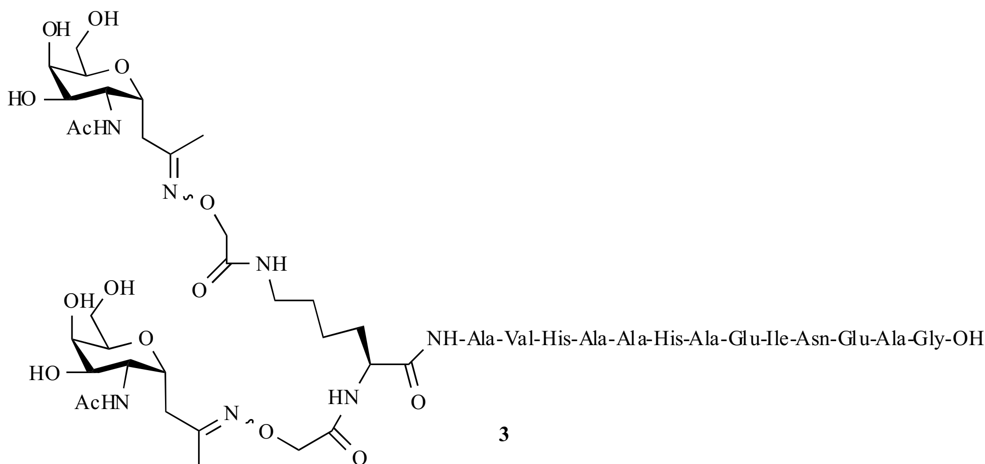

Fig. (5). Synthetic vaccine assembled by chemoselective synthesis: two copies of the C-saccharide analogue of the T_n epitope, were conjugated by means of oxime bonds to the immunogenic peptide OVA $^{(327\text{-}339)}$.

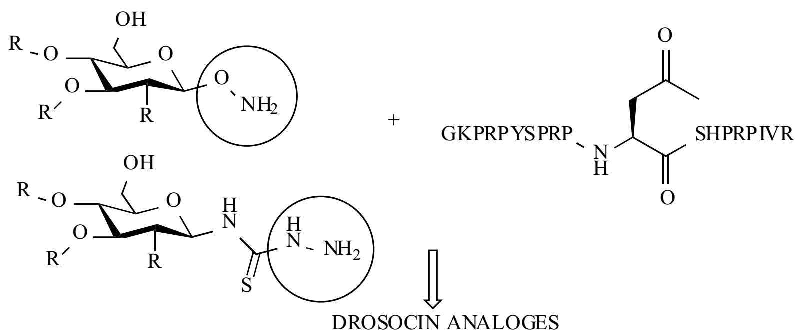

Fig. (6). β-oxyamino and β-thiohydrazide glycosides reacted selectively with the peptide bearing a ketone group giving drosocin analogues.

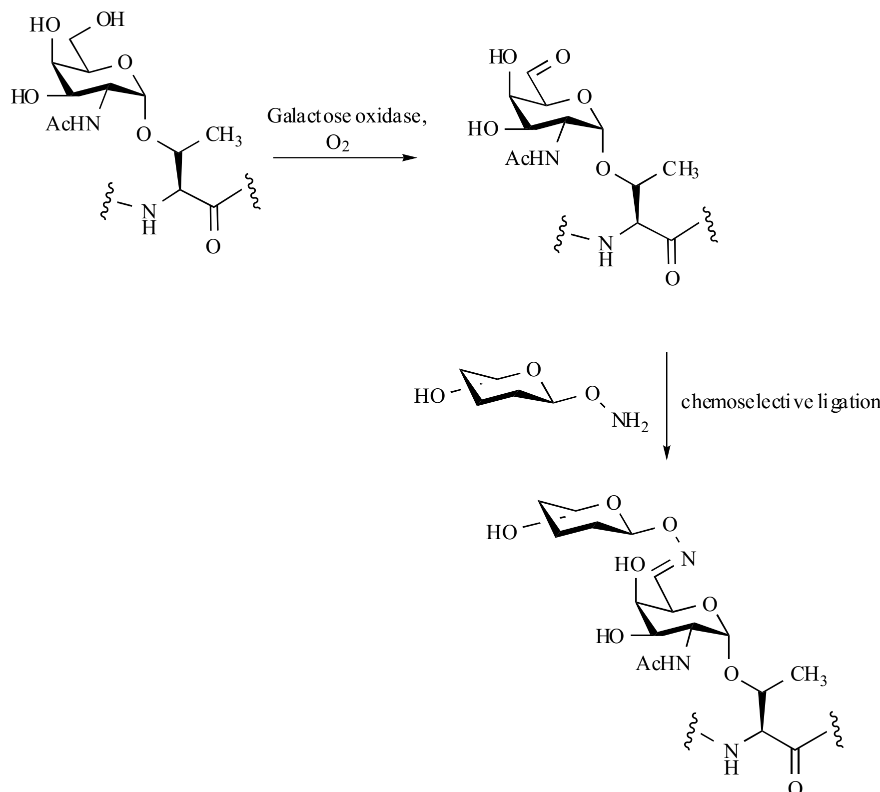

Fig. (7). Synthesis of oxime-linked drosocin analogues by chemoselective extension of the saccharide part.

In a similar chemoselective approach, Bertozzi and coworkers used the ketone group of the unnatural amino acid, (2*S*)-aminolevulinic acid, as an electrophile for chemoselective reaction with sugars presenting different nucleophile groups as appendages at the anomeric position, thus obtaining analogs of the *O*-linked glycopeptide drosocin [Fig. (**6**)] [19]. It was also proved that, in the case of oxime-linked derivative, the neoglycopeptide possessed activity as an inhibitor of bacterial growth similar to that of the native glycopeptide [20].

The chemoselective strategy was extended to the formation of sugar-sugar linkages in the synthesis of drosocin analogs: the glycosyl amino acid Thr/Ser(Ac_3-D-GalNAc) was incorporated into the peptide by solid phase synthesis, then the sugar was deprotected and selectively oxidized to the corresponding C-6 aldehyde by treatment with the enzyme galactose oxidase. Finally, the glycopeptide was conjugated through reactions with mono- and disaccharides bearing amino-oxy groups at the anomeric position [Fig. (**7**)] [21].

The ultimate extension of this chemistry has been shown by decorating glycoproteins anchored to the cell membrane by chemoselective covalent reactions in the cellular environment. Through the addition of chemically modified metabolic precursors of sialic acid, *N*-levulinoylmannosammine and *N*-azidoacetylmannosammine, (cell surface glycoproteins) can be made to display ketones and azides, which may be extended by selective covalent reactions. In the case of ketones, complementary nucleophiles as hydrazides and amino-oxy groups react forming stable adducts. This technology has been applied to the remodeling of cells with alternative glycoforms [22]; to the selective targeting of toxins and diagnostic agents to tumor cells [23]; and to modulate the cell surface immunoreactivity [24]. The azide group on the other hand has been shown to undergo a Staudinger reaction with a triphenylphosphine, and a prototypical decoration of Jurkat cells with water-soluble biotinylated phosphine has been shown [Fig. (**8**)] [25].

NEOGLYCOCONJUGATES BY SULFUR NUCLEOPHILE-BASED CHEMOSELECTIVE CHEMISTRY

The superior nucleophilic character of thiols inspired the use of these functional groups in combination with a series of electrophiles for chemoselective formation of neoglycoconjugates.

The natural amino acid cysteine is a chemoselective point of attachment for site-selective reactions. Site-directed mutagenesis combined with chemical modification provides, in principle, a general method allowing both regio- and sugar-specific glycosylation of protein. The serine protease subtilisin *Bacillus lentus* (SBL), which is not naturally glycosylated, was mutated in order to introduce cysteine residues in precise sites. Mutants were then reacted with protected and unprotected monosaccharides bearing a methanethiosulfonate or an ethyl-tethered methanethiosulfonate as selective thiol electrophiles in anomeric position, thus generating disulfide linkages [26].

The glycoprotein hormone erythropoietin (EPO) is naturally glycosylated, and *N*-linked glycans at asparagines 24, 38, and 83 are essential for *in vitro* activity. Flitsch and

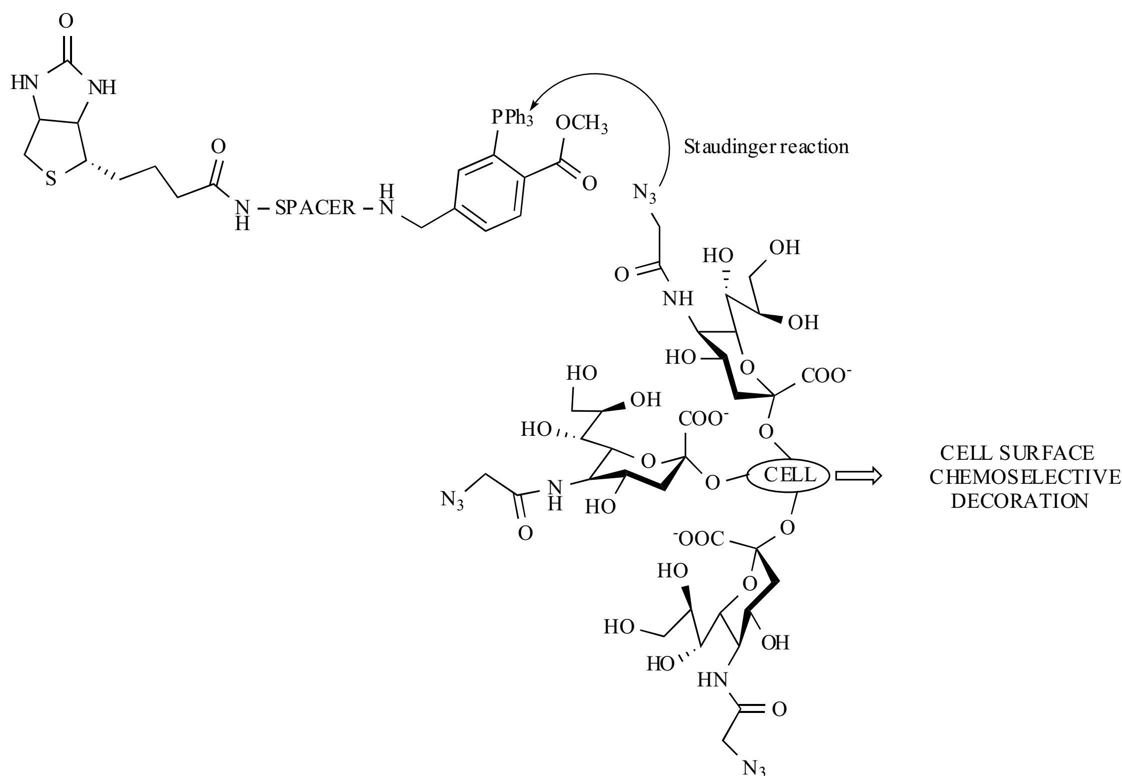

Fig. (8). Cell surface engineering by chemoselective biotinylation.

coworkers undertook a systematic and elegant study on this protein through Asn→Cys mutation at the natural *N*-linked glycosylation sites [27]. The three mutants N24C, N38C, and N83C were refolded in the native form, and then reacted with an excess of glycosyl iodoacetamide, affording neoglycoprotein selectively glycosylated on the desired regions. It was also demonstrated that, despite the non-natural linker between sugar and protein, the stability of neoglycoproteins was similar to that of wild-type EPO.

In model studies with short peptides, cysteines incorporated during SPPS were reacted chemoselectively with sugars bearing maleimide [28], disulfide [29] and 2-bromoethyl [30] electrophiles. Reversing the position of the electrophile, tripeptides containing dehydroalanine, derived from selenocysteine by oxidative elimination, were reacted with thioglycosides, thus obtaining *S*-linked neoglycopeptides through chemoselective Michael addition [31].

The thiol-based ligation was used as a glycosylation technique in combination with aldehyde/hydrazide chemistry to build up complex dendrimeric synthetic vaccines targeted to DCs.

It has been reported that mannose receptors on the DC surface can selectively bind and internalize molecules and micro-organisms coated with D-mannose, L-fucose or *N*-acetyl-D-glucosamine residues by means of simultaneous interactions with several carbohydrate recognition domains. Synthetic vaccines were obtained by linking several copies of (2-thioethyl)-α-D-mannopyranoside to an *N*-chloroacetyl-L-lysinyl branched core bearing a hydrazine function (or a glyoxylyl aldehyde) that was successively reacted with the glyoxylyl aldehyde group (or the hydrazide group) of the antigen peptides. An array of synthetic vaccines was generated, presenting different epitopes linked to the poly-mannose cluster [Fig.(**9**)] [32,33]. This synthetic approach, based on the combination of orthogonal hydrazone and thioether methods, is one elegant and rare example of double ligation strategy, one reaction being used for sugar-peptide, the other for peptide-peptide conjugation.

Two orthogonal ligation techniques have been used for the chemoselective elaboration of *O*-linked glycopeptide mimetics. In this case, two reactions respectively based on oxime and thioether bonds formation, have been applied in different syntheses [34].

The well-known "native ligation" developed in parallel by Kent for peptide fragment assembly and based on the reaction of an *N*-terminal cysteine with a *C*-terminal thioester followed by S-to-N acyl migration and subsequent amide bond formation, allows the synthesis of proteins of over 100 residues in length. This approach has been extended to the total synthesis of the homogeneous glycosylated lymphotactin (Lptn) where the fragment with *N*-terminal cysteine contains several *O*-linked glycosides incorporated in a classical way during SPPS [35]. In this context, the "protein expressed ligation" allowed the union of recombinant proteins bearing a *C*-terminal thioester with synthetic peptides, through native ligation. With this technique, it has been possible to accomplish the semi-synthesis of homogeneous *O*-linked mucin-type glycoproteins [36].

FUTURE DIRECTIONS

The chemoselective ligation techniques developed so far allow mainly the conjugation of sugars to peptides and proteins. The major limitation for neoglycoprotein synthesis

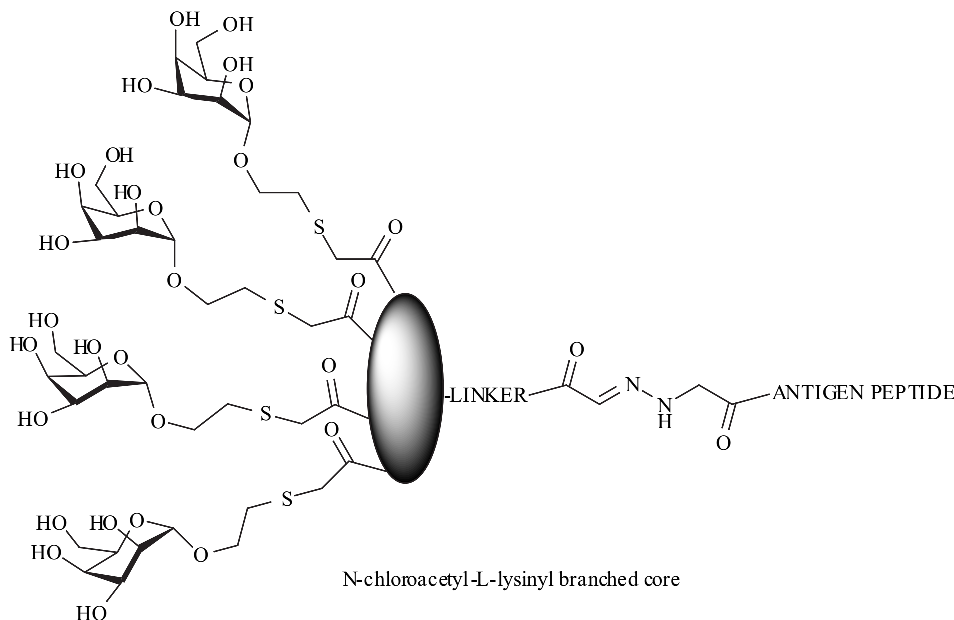

Fig. (9). Clustered mannoside-antigen conjugate obtained by one-pot, orthogonal, chemoselective ligations.

is the need to insert the natural amino acid cysteine or artificial amino acids at the desired site of conjugation. This can be accomplished by incorporating the residue during SPPS in the case of unnatural amino acids or by site-directed mutagenesis for the introduction of cysteine. The development of highly productive techniques for incorporating unnatural amino acids into proteins *in vivo* [37] would provide an extremely powerful and general method to install desired reacting groups at precise sites for chemoselective glycosylation.

Another important target for the future would be the extension of the ligation chemistry to the synthesis of neoglycolipids; even though the extreme lipophilicity of these molecules discourages the use of aqueous media for their preparation, convergent strategies could greatly enhance the production of diversity in drug development processes.

Ligation chemistry has been very recently extended to sugar-sugar bond formation. The reaction of a methoxyamino group on C-6 position of one sugar unit with the anomeric carbon of an unprotected reducing sugar, allowed the preparation of di- and trisaccharides with the non-natural *N*(*O*Me) glycoside linkage both in solution and in solid phase [Fig. (**10**)] [38]. This technique can be extended to the construction of mimetics of oligosaccharides with β(1-6) glycoside bond, with the following synthetic advantages: the ligation reaction is stereoselective leading exclusively to the β-glycoside bond, and the formed *N*(*O*Me) linkage is isosteric to the natural one.

The extension of this methodology to the formation of glycoside linkages with different regiochemistry and anomeric configuration would enlarge the repertoire of reactions available for the preparation of complex branched oligosaccharide mimetics.

REFERENCES

[1] Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. Essentials of Glycobiology, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, **1999**.

[2] Cummings, D. A. *Glycobiology* **1991**, *1*, 115-130.

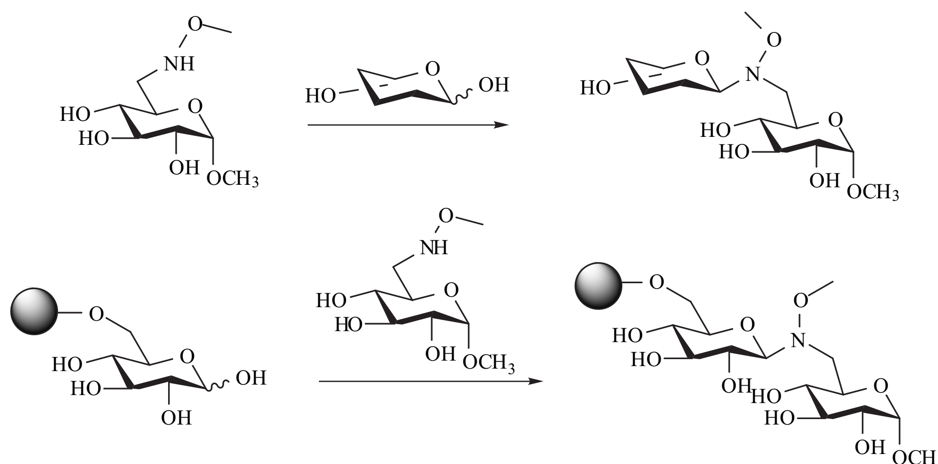

Fig. (10). General scheme of the chemoselective assembly of β(1-6)-linked *N*(*O*Me) oligosaccaride analogues in solution and in solid phase.

[3] Merritt, E. A.; Hol, W. G. *J. Curr. Opin. Struct. Biol.* **1995**, *5*, 165-171.
[4] Hakomori, S.-I.; Zhang, Y. *Chem. Biol.* **1997**, *4*, 97-104.
[5] Villard, R.; Hammache, D.; Delapierre, G.; Fotiadu, F.; Buono, G.; Fantini, J. *Chem. Bio. Chem.* **2002**, *3*, 517-525.
[6] Chaby, R. *DDT* **1999**, *4*, 209-221.
[7] Allen, J. R.; Danishefsky, S. J. *Angew. Chem. Int. Ed. Engl.* **2000**, *39*, 836-863.
[8] Sears, P.; Wang, C.-H. *Angew. Chem. Int. Ed.* **1999**, *38*, 2300-2324.
[9] Seitz, O. *Chem. Bio. Chem.* **2000**, *1*, 214-246.
[10] Blixt, O.; Allin, K.; Pereira, L.; Datta, A.; Paulson, J. C. *J. Am. Chem. Soc.* **2002**, *124*, 5739-5746.
[11] Ramos, D.; Rollin, P.; Klaffke, W. *Angew. Chem. Int. Ed.* **2000**, *39*, 396-398.
[12] Meinjohanns, E.; Meldal, M.; Paulsen, H.; Dwek, R. A.; Bock, K. *J. Chem. Soc. Perkin Trans. 1* **1998**, 549-560.
[13] Cervigni, S. E.; Dumy, P.; Mutter, M. *Angew. Chem. Int. Ed.* **1996**, *35*, 1230-1232.
[14] Peri, F.; Dumy, P.; Mutter, M. *Tetrahedron* **1998**, *54*, 12269-12278.
[15] Carrasco, M. R.; Nguyen, M. J.; Burnell, D. R.; MacLaren, M. D., Hengel, S. M. *Tetrahedron Lett.* **2002**, *43*, 5727-5729.
[16] Durieux, P.; Fernandez-Carneado, J.; Tuchscherer, G. *Tetrahedron Lett.* **2001**, *42*, 2297-2299.
[17] [0] Peri, F.; Cipolla, L.; Rescigno, M.; La Ferla, B.; Nicotra, F. *Bioconj. Chem.* **2001**, *12*, 325-328.
[18] Cipolla, L.; Rescigno, M.Leone, A.; Peri, F.; La Ferla, B.; Nicotra, F. *Bioorg. Med. Chem.* **2002**, *10*, 1639-1646.
[19] Rodriguez, E. C.; Marcaurelle, L. A.; Bertozzi, C. *J. Org. Chem.* **1998**, *63*, 7134-7135.
[20] Marcaurelle, L. A.; Rodriguez, E. C.; Bertozzi, C. R. *Tetrahedron Lett.* **1998**, *39*, 8417-8420.
[21] Rodrigues, E. C.; Winans, K. A.; King, D. S.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1997**, *119*, 9905-9906.
[22] Yarema, K. J.; Mahal, L. K.; Bruhel, R. E.; Rodriguez, E. C.; Bertozzi, C. R. *J. Biol. Chem.* **1998**, *273*, 31168-31179.
[23] Lemieux, G. A.; Yarema, K. J.; Jacobs, C. L.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1999**, *121*, 4278-4279.
[24] Lemieux, G. A.; Bertozzi, C. R. *Chem. Biol.* **2001**, *8*, 265-275.
[25] Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007-2010.
[26] Davis, B. C.; Lloyd, R. C.; Bryan Jones , J. *J. Org. Chem.* **1998**, *63*, 9614-9615.
[27] Macmillan, D.; Bill, R. M.; Sage, K. A.; Fern, D.; Flitsch, S. *Chem. Biol.* **2001**, *8*, 133-145.
[28] Shin, I.; Jung, H.-J.; Lee, M.-R. *Tetrahedron Lett.* **2001**, *42*, 1325-1328.
[29] Macindoe, W. M.; vanOijen, A. H.; Boons, G.-J. *Chem. Commun.* **1998**, 847-848.
[30] Bengtsson, M.; Broddefalk, J.; Dahmén, J.; Henriksson, K.; Kihlberg, J.; Lonn, H.; Srinvasa, B. R.; Stenval, K. *Glycoconj. J.* 1998, *15*, 223-231.
[31] Gieselman, M. D.; Zhu, Y.; Zhou, H.; Galonic, D.; van der Donk, W. A. *Chem. Bio. Chem.* **2002**, *3*, 709-716.
[32] Grandjean, C.; Rommens, C.; Gras-Masse, H.; Melnyk, O. *Angew. Chem. Int. Ed.* **2000**, *39*, 1068-1071.
[33] Grandjean, C.; Gras-Masse, H.; Melnyk, O. *Chem. Eur. J.* **2001**, *7*, 230-239.
[34] Marcaurelle, L. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* 2001, *123*, 1587-1595.
[35] Marcaurelle, L. A.; Mizoue, L.; Wilken, J.; Oldham, L.; Kent, S. B. H.; Handel, T. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2001**, *7*, 1129-1132.
[36] Tolbert, T. J.; Wong, C.-H. *J. Am. Chem. Soc.* **2000**, *122*, 5421-5428.
[37] Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498-500.
[38] Peri, F.; Deutman, A.; La Ferla, B.; Nicotra, F. *Chem. Commun.* **2002**, 1505-1505.

Copyright © 2003 EBSCO Publishing