Glycoprotein Synthesis: An Update

David P. Gamblin, Eoin M. Scanlan, and Benjamin G. Davis*

Department of Chemistry, University of Oxford, Chemistry Research Laboratory, Mansfield Road, Oxford OX1 3TA, United Kingdom

Received August 29, 2007

Contents

1. Introduction 131
   1.1. Aims 131
   1.2. Biological Importance of Glycopeptides/Glycoproteins 131
   1.3. Why Does Nature Employ Carbohydrates 132
   1.4. Potential Applications 133
   1.5. Need for Homogeneity/Single Glycoforms 134
   1.6. Retrosynthetic Analysis 135
2. Glycopeptide Assembly 135
   2.1. Strategies for Formation of the Glycosidic Linkage 135
      2.1.1. Linkage Variety 135
      2.1.2. Formation of Natural Linkages 136
      2.1.3. Synthetic Approaches to Unnatural Linkage Types 136
   2.2. Assembly Strategies 139
      2.2.1. Linear Assembly 139
      2.2.2. Convergent Assembly 139
      2.2.3. Elaborative and Mixed Assembly Strategies 139
      2.2.4. Native Ligation Assembly 140
3. Chemical Glycoprotein Synthesis 143
   3.1. Indiscriminate Convergent Glycosylation 143
   3.2. Chemoselective and Site-Specific Glycosylation 145
   3.3. Site-Selective Glycosylation 147
   3.4. Native Ligation Assembly 151
4. Enzymatic Glycoprotein and Glycopeptide Synthesis 151
   4.1. Glycan Extension 151
   4.2. Glycoprotein Remodeling (GPR) 152
   4.3. Enzymatic Formation of the Glycan−Protein/Peptide Link 153
   4.4. Glycopeptide Ligation 155
5. Molecular and Cell Biological Techniques 155
   5.1. Biosynthesis Augmentation 155
   5.2. Biosynthesis Regulation 156
   5.3. Codon Suppression Technology 157
   5.4. Expressed Protein Ligation (EPL) 157
6. References 159

1. Introduction

1.1. Aims

This review updates our original review on the synthesis of glycoproteins.[1] Its aims, therefore, are 2-fold: to highlight, first, key developments since 2002 and, second, provide, where relevant, pointers in the form of seminal examples to the unitiated. Thus, while we hope to produce a detailed update of methods for glycoprotein assembly, only select examples of the numerous potential applications are included. The field continues to expand apace, yet the conceptual categories chosen in 2002 still serve well as clear markers and so will also be used here. The excitement at this time is the emergence of genuinely viable strategies of direct relevance to novel probes of biological mechanism. These include approaches that now seem likely to yield synthetic proteins as clinical biopharmaceuticals in the near future.

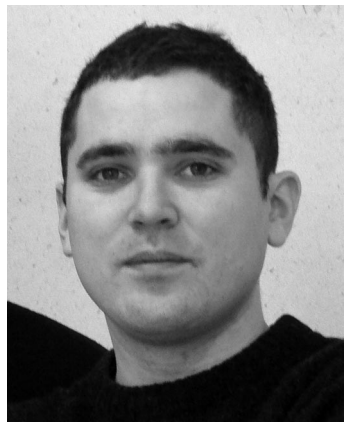
* To whom correspondence should be addressed. E-mail: Ben.Davis@chem.ox.ac.uk.

1.2. Biological Importance of Glycopeptides/Glycoproteins

Carbohydrates are ubiquitous in nature, and there is little doubt that the elaborate carbohydrate constructs observed on the surface of cells and proteins hold the key to understanding many complex biological processes. For example, glycoproteins have been reported to exhibit a pivotal role in processes as diverse as fertilization, neuronal development, hormone activities, immune surveillance, and inflammatory responses.[2−6] The carbohydrates of host cells are often employed by pathogens for cell entry and immunological evasion. In addition to their critical role in communication events, an interesting aspect of *N*-linked glycans is highlighted by their post-translational role in the 'quality control' of protein synthesis.[7−9] In the absence of correct glycosylation, many proteins fold incorrectly. It has been suggested that if proteins fail to fold correctly the glycans are incorrectly displayed and cannot be processed in trimming steps leading to expulsion of the protein via the endoplasmic reticulum-associated protein degradation (ERAD) pathway.[10] Therefore, the apparently superfluous glycan trimming steps seen in the endoplasmic reticulum may not simply be a means to glycan structure but steps along a "quality-controlled" protein production line and suggests a key role for added glycans as indicators of correct protein structure.[7,9,11,12] Moreover, glycans appear to stabilize tertiary structure[13,14] and also aid folding and transport by protecting proteins from proteolysis.[15] Enhanced proteolytic stability, as a result of glycosylation, has been observed, for example, through comparison of RNase A, an unglycosylated pancreatic ribonuclease, and RNase B, which bears a single high-mannose oligosaccharide at Asn34.[16−18]

Carbohydrates alter other physiochemical properties of proteins. In artic fish,[19] *O*-glycan-rich proteins act as an in vivo "antifreeze", preventing nucleation of ice, allowing them to survive temperatures of −2 °C. Glycopeptide analogues of these proteins have also demonstrated significant antifreeze properties in the cryopreservation of islet cells.[20,21]

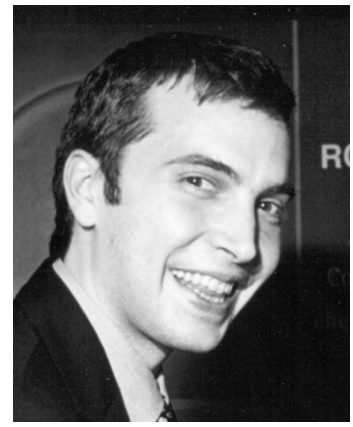
10.1021/cr078291i CCC: $71.50 © 2009 American Chemical Society
Published on Web 12/18/2008



Dr. David Gamblin was born in Taunton, England, in 1980. He was awarded his M.Chem. degree in 2002 and later obtained his D.Phil. degree in 2005 from the University of Oxford. His doctoral thesis focused on development and design of novel chemoselective reagents for site-selective protein modification, for which he was awarded the Europa medal in 2004. In 2005 he took a postdoctoral research fellowship working in collaboration with Professors Ben Davis and John Simons, FRS. David's research interests include precise protein modification, glycosylation chemistry, and glycan analysis using mass spectrometry. From 2005 until 2008 he was the Lecturer in Organic Chemistry at St. Edmund Hall, Oxford. In 2008 he left academic research and is currently is a school master at St. Pauls School, London.

Dr. Eoin Scanlan completed his degree in Chemistry and Applied Chemistry at the National University of Ireland, Galway, in 2000. He was awarded his Ph.D. degree in 2003 from the University of St. Andrews, U.K., where he worked under the supervision of Professor John Walton. His doctoral thesis focused on the development of oxime oxalate amides as tin-free precursors for radical-mediated synthesis of biologically active lactams. In January 2004 he took a position with Professor Philippe Renaud at the University of Bern, Switzerland, where he investigated the use of organoboranes as free-radical precursors. In July 2005 he moved to the University of Oxford, U.K., where he worked with Professor Ben Davis on the synthesis of complex plant cell wall oligosaccharides and regioselective glycosylation methodology. In March 2008 he was appointed as lecturer in Organic and Medicinal Chemistry at Trinity College Dublin. His group's research interests include development of novel synthetic methodology for construction of oligosaccharides and glycoconjugates and synthesis of carbohydrate-based therapeutics for treatment of cancer and diabetes.

This diverse role of protein function fine tuning suggests that post-translational modifications (PTMs) such as sugars may in part explain one of the most important current paradoxes in science. The number of genes found in higher animals has been much less than anticipated. At the same time the functional output is far more diverse than would have been expected from this comparatively restricted genetic



Professor Ben Davis is a Professor of Chemistry and Fellow and Tutor in Organic Chemistry, Pembroke College. After receiving his B.A. (1993) and D.Phil. (1996) degrees at Oxford he spent 2 years at Toronto, exploring protein design and biocatalysis. In 1998 he returned to the United Kingdom to take up a lecturership at the University of Durham. In 2001 he returned to Oxford and was made a Professor in 2006. His group's research centers on chemical biology with an emphasis on carbohydrates and proteins. In particular, the group's interests encompass synthesis and methodology, inhibitor design, protein engineering, drug delivery, molecular modeling, molecular biology, and glycoscience with the goal of medicinal application. He has published >100 papers and >20 patents and has given >120 invited lectures. This work has been funded by participation in >£10m of research income and has received the 1999 RSC Meldola medal and prize, the 2001 RSC Carbohydrate Award sponsored by Syngenta, an AstraZeneca Strategic Research Award, a DTI Smart Award, a Mitzutani Foundation for Glycoscience Award, the 2002 Philip Leverhulme Prize, the 2005 Royal Society Mullard Prize and Medal, the RSC 2005 Corday-Morgan Medal, and in 2006 the International Association for Protein Structure Analysis and Proteomics Young Investigator Award. He is cofounder of Glycoform, a small biotechnology company aimed at exploiting the therapeutic potential of glycoproteins. In 2003 he was named among the 100 top young innovators in the world by *Technology Review*, Massachusetts Institute of Technology's magazine of innovation. In 2008, he was awarded the U.K. Wain Medal for Chemical Biology and was the first U.K. recipient of the ACS Horace S. Isbell Award.

code.[22] Alterations of amino acid side chains such as glycosylation lead to higher structural and functional protein diversity and are therefore a prime candidate in explaining the seeming incongruity in gene number and biological function.[23]

1.3. Why Does Nature Employ Carbohydrates

The need for carbohydrates in biological events may be explained through the diverse range of constructs that they can form. Schreiber recently suggested[24] that carbohydrates in biology may best be compared simply with terpenes, yet this underestimates their structural diversity enormously. Carbohydrates are unparalleled in the number of structures they can adopt, and as a consequence, nature appears to exploit this structural diversity to convey information at a molecular level.[25] In terms of oligomerization, proteins and nucleic acids are effectively linear in structure (amide bonds in proteins and 3′−5′ linkage of phosphodiesters within DNA), where limited basis sets (4 for DNA, 20 for amino acids) give rise to limited variations. In addition to inherent configurational variation (gluco, manno, etc.), additional variety caused by ring size, branching, anomeric configuration, and modification (e.g., acylation, sulfation, and phosphorylation) gives carbohydrates strong potential for diversity. This inherent structural diversity parallels a wide range of functions within nature,[3,6,11,26,27] ranging from a source of energy and metabolic intermediates to the structural

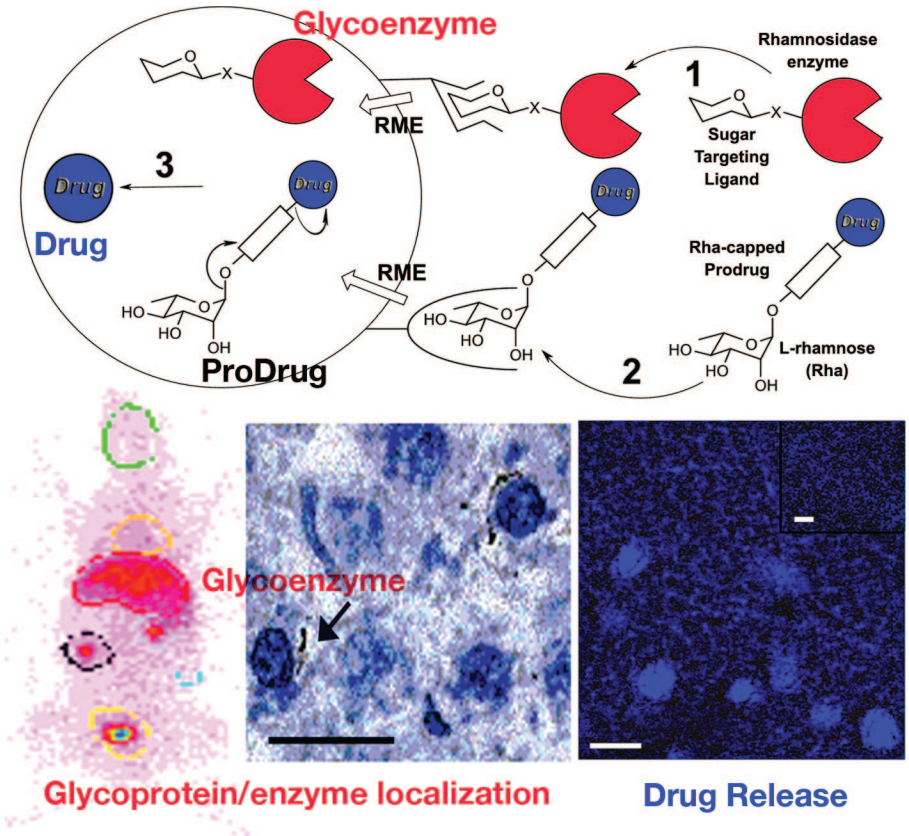


Figure 1. LEAPT: lectin-directed enzyme activated prodrug therapy.

components of plants (cellulose), animals (chitin), and nucleic acids (DNA, RNA). However, greater variation in oligosaccharide structure is exploited by nature through the combination of carbohydrates with proteins (glycopeptide/glycoprotein), the products of which have caused an explosion of interest within the scientific community. This structural diversity has been christened *glycocode,* a term that well represents the potential level of complex information that carbohydrate structures are able to convey.[1,28] This vast number of potential permutations represents a technological barrier and means that oligosaccharide portions of glycoproteins cannot be made simply on an iterative basis since there are far too many possible synthetic targets. It is therefore crucial that the design of new glycoproteins is guided by identification of the associated functions and activities of existing natural and resulting synthetic structures. In this regard, the number of possible glycans that we may need to target may not be as overwhelming as a statistical analysis of possible permutations might suggest. As we start to unpick the mechanisms by which the biosynthetic assembly of protein glycans is controlled we may start to unpick a greater sense of the 'rules' of these processes and their resulting functional effects in proteins; here again, precisely modified proteins will be crucial in the delineation of any such rules.

1.4. Potential Applications

Recent developments in glycopeptide/glycoprotein synthesis have been further honed by the application of products to an understanding of the vital role that they play in nature.[29−32] Although this is the principle reason for conducting this research,[33] the biological significance and application of both natural and synthesized glycopeptides/glycoproteins will only be discussed in passing here.[34] The reader is referred to our previous review for further examples of the biological and biotechnological applications of glycopeptides and glycoproteins.[33]

The specificity of the hepatic asialoglycoprotein receptor[35] has been widely exploited for liver targeting and is a standard for targeted drug delivery[36] and gene delivery[37,38] therapies.[39,40] A recent example termed lectin-directed enzyme-activated prodrug therapy (LEAPT) has been designed to exploit endogenous carbohydrate−lectin binding through the combination of biocatalysis with novel glycosylated enzymes and prodrugs.[41] In the first step of this bipartite strategy a glycosylated enzyme is targeted to specific cell types within the body that are determined by the selected carbohydrate ligand; next, a prodrug capped with a nonmammalian sugar is administered. Use of linkages in the prodrug that can only be cleaved by the activity of the glycosylated enzyme ensures that it is only released at the predetermined target site (Figure 1). The first system utilized a rhamnosidase, which was first deglycosylated following treatment with endoglycosidase-H (see section 3.1) and then chemically glycosylated using a lysine-glycosylation (IME, 2-iminomethoxymethyl thioglycoside) methodology.[42] When coadministered with model rhamnose-capped prodrugs, in vivo analysis showed a high level of drug in the target organ, the liver. Moreover, use of a prodrug of the anticancer drug doxorubicin in the system allowed promising treatment of an animal tumor model. The methodology shows possible adaptation to other disease targets by varying the sugar or to target other suitable receptors of medical relevance.

Such "sugar-targeting" approaches may also be used to target physiologically advantageous enzymes in so-called "enzyme replacement therapy" (ERT).[43] For example, the

presence of mannose-6-phosphate bearing high-mannose oligosaccharides on recombinant human acid α-glucosidase (rhGAA) has been shown to improve the clearance of glycogen in Pompe mice.[44] Infusion of glycosylated rhGAA into Pompe mice resulted in greater delivery of the enzyme to muscle tissues when compared to the unmodified enzyme with an impressive 5-fold increase in glycogen breakdown in affected tissues.

Recently, Vérez-Bencomo and co-workers developed a glycoconjugate vaccine composed of a fully synthetic capsular polysaccharide antigen of *Haemophilus influenzae* type b (Hib),[45] highlighting the fact that access to synthetic complex carbohydrate-based vaccines is feasible. Initial results from clinical trials demonstrated long-term protective antibody titers as compared to licensed commercial products. This Cuban work has set the international benchmark for synthetic vaccines and highlighted the need for further development of similar approaches to other human pathogens and perhaps greater innovation and courage in other countries.

Fatalities resulting from malaria are caused by an inflammatory surge initiated by malarial toxin released from the parasite *Plasmodium falciparum*; glycosylphosphatidylinositols (GPI) are thought to be the primary toxin that underlies this pathology. In an attempt to produce a vaccine against malaria, Seeberger and co-workers chemically synthesized the GPI oligosaccharide and conjugated it to carrier proteins (see section 3.2).[46] Anti-GPI antibodies were obtained from immunized mice and shown to neutralize the pro-inflammatory activity of *P. falciparum* in vitro. More impressively, it was demonstrated that deaths from malarial parasites in animal models were greatly decreased, thereby establishing the idea that GPI conjugates could be used for antimalarial vaccine design. The rapid and high-yielding assembly of the GPI motif was possible through automated solid-phase glycosylation chemistry,[47] which could be adapted to include alternative building blocks for SAR studies.

Glycans may modulate serum half-life, and in an attempt to enhance the pharmacokinetics of synthetic insulin, Nishimura and co-workers combined mutagenesis with enzymatic synthesis in the production of a glycosylated insulin.[48] Standard insulin is rapidly degraded by the liver (within a few hours of administration), thus requiring frequent booster shots. Methods to increase in vivo activity of insulin have been investigated,[49] but these are limited by intricate administration regimes resulting in uncontrolled fluctuation in glucose blood levels caused by decreased water solubility.[50] Nishimura et al. tackled the water solubility and degradation problems by introducing sialic acid, a sugar that is known to increase glycoprotein half-life in vivo.[51,52] This required installation of accessible Gln residues through mutagenesis into the B chain of insulin and subsequent transglutaminase (TGase) catalyzed transamidiation with a lactosyl amine. The resulting glycoprotein was further modified through enzymatic sialylation with a Siaα2,6-transferase to make Sia2,6-Lac. The product insulin showed a longer-lasting in vivo activity than unmodified insulin. Interestingly, the same synthetic methods were also used in the creation of a dendrimeric display of sialic acid, leading to an insulin "glycodendriprotein" (see section 3.3).[53] Although the binding affinity of these modified insulins to receptor decreased as dendrimer size increased, an overall in vivo activity increase was observed due to the enhanced half-life caused by the higher presence of sialic acid.

Keratoconjunctivitis is a severe eye infection caused by adenoviruses; Kihlberg and co-workers developed multivalent displays of 3′-sialyllactose conjugated to HSA-carrier proteins ligated through squaric acid (see section 3.2).[54] 3′-Sialyllactose was chosen as a putative cellular marker for adenovirus serotype 37 infection, and the effectiveness of these glycoconjugates was demonstrated through strong inhibition of the binding of Ad37 to human corneal epithelial cells.

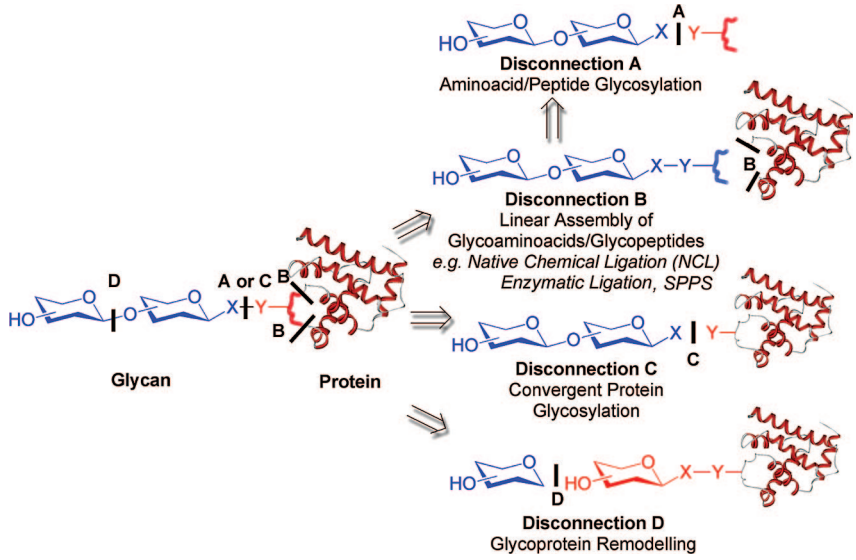
More recently, a 'dual-tag double-modification' strategy in combination with site-directed mutagenesis and orthogonal chemoselective ligation has enabled the precise site-selective attachment of different post-translational modifications to the surface of reporter proteins, thereby creating probes to survey protein/protein interactions (see section 3.3).[23] A combination of triazole formation and disulfide formation allowed for the position-specific attachment of a sLe^x tetrasaccharide and a sulfated tyrosine mimic, respectively, onto a reporter protein SSβG, a galactosidase which is capable of being visualized using the widely used X-Gal substrate. Furthermore, through the use of site-directed mutagenesis these modifications were located in a way that successfully mimicked the binding parameters of human protein ligand PSGL-1 to its protein receptor P-selectin. This synthetic protein was able to act as a potent reporter system for P-selectin and hence as a probe for in vivo inflammatory brain lesions caused by either acute or chronic interleukin-induced inflammation and within a *Plasmodium*-infected malarial disease model.

Single-site glycosylation through site-selective cysteine ligation has also been used to produce glycosylated fragments of IgG1,[55] shown to be capable of interacting with human FcγRI by inhibiting superoxide production in interferon-γ (IFNγ) stimulated U937 cells. This study thus impressively showed that the activity of a glycoprotein could be removed through deglycosylation of its natural glycan and then partially restored through the in vitro ligation of synthetic carbohydrates.

1.5. Need for Homogeneity/Single Glycoforms

As previously discussed, glycoproteins are fundamental to a variety of biological events. However, investigating these processes is complicated by the fact that naturally expressed glycoproteins often emerge as heterogeneous mixtures.[56] Biosynthesis of proteins is under direct genetic control from DNA-to-protein through transcription−translation. However, biosynthesis of the glycan portion is not template mediated; the glycan structure is subject to competition between various glycosyltransferases, glycosidases, and their substrates, the substrate specificity of these processing enzymes, and substrate/donor availability. This variable post-translational processing of glycans results in expression of an assortment of possible glycan structures called "glycoforms",[3] glycoproteins that possess the same protein backbone but differ in the oligosaccharide components and sites of glycosylation. Therefore, expression of glycoproteins in mammalian cell culture habitually leads to mixtures of glycoforms. The different properties demonstrated by each glycan within the resulting heterogeneous mixture present not only regulatory difficulties for therapeutic glycoproteins[57] but also problems in determining exact structure−activity relationships (SAR). There is therefore an urgent need to develop methodologies to provide access to alternative sources of homogeneous perhaps synthetic glycoproteins. Development of highly successful small-molecule carbohydrate-containing ligands

Scheme 1. Glycoprotein Retrosynthetic Analysis



has often involved careful SAR refinements.[58] If we are to achieve the same successes with glycoproteins as larger molecules using advanced synthetic organic chemistry then a new mode of synthetic biology must be developed with product homogeneity as one of its first key priorities.

1.6. Retrosynthetic Analysis

The aim of this review is to discuss and highlight the recent methodologies and synthetic strategies applied to preparation of glycoproteins and, where relevant, glycopeptides as model systems. These procedures encompass chemical, enzymatic, and molecular biological approaches used independently or in combination with one another. The net result is a formidable range of techniques that bridges the ever-decreasing boundaries between chemistry and biology. Scheme 1 displays a generalized pseudoretrosynthetic analysis for glycoproteins, and this review systematically discusses recent advances under the headings of each disconnection type.[1] The terms that have been used to describe disconnections A, B, C, and D are not necessarily systematic and are subjective ones that have been popularized by various workers in the field but nonetheless are helpful and shall be used where appropriate in this review (Scheme 1).

It is not the intention of this review to provide complete coverage of all the methods for construction of carbohydrate−amino acid conjugates but to instead focus on formation of glycosylated-polypeptide (glycoprotein) structures that have emerged since our last review. As a consequence, glycopeptide synthesis and assembly are only covered where it is of potential or direct relevance to glycoproteins. The reader is referred to recent reviews that have appeared on synthesizing shorter length glycopeptides.[59−63] Correspondingly, although they are a crucial requirement for glycoprotein synthesis, synthesis of oligosaccharides[64−72] or pure proteins[73−77] will not be covered unless of direct relevance. This review is a complete update of our first comprehensive review of 2002 on the synthesis of glycoproteins,[1] and the reader is also referred to other reviews.[22,63,78−85] All sources of glycoproteins have potential function (see section 1.4); therefore, there will be no distinction drawn between synthetic analogues, sometimes referred to as neoglycoproteins (neo simply means new and adds nothing to the nomenclature of this field), and those occurring naturally.

2. Glycopeptide Assembly

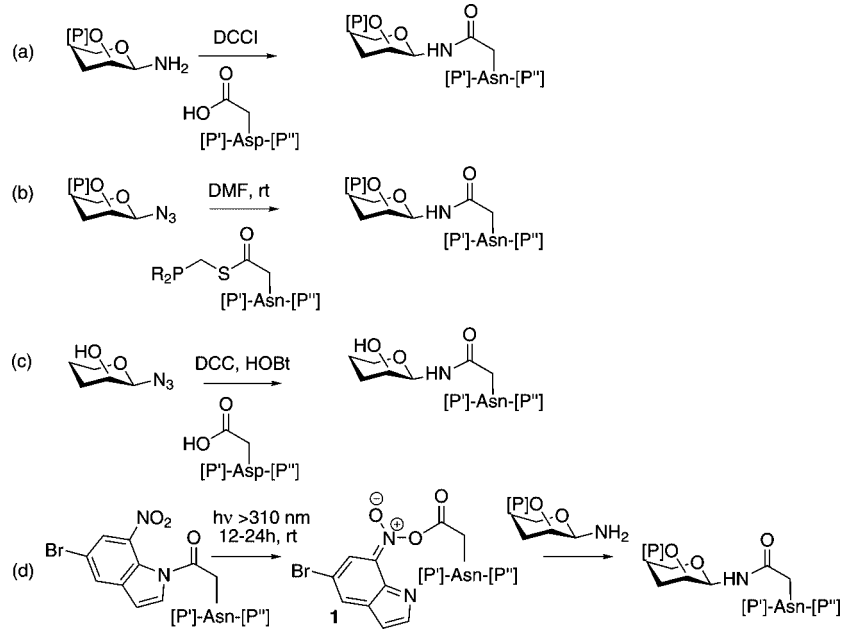
The urgent need for homogeneous glycopeptides and glycoproteins has resulted in a concerted effort to develop general and efficient methodologies for their assembly. Synthesis of glycopeptides requires a combination of synthetic methods from both carbohydrate and peptide chemistry. Central to any glycopeptide synthetic strategy is formation of the integral glycan−amino acid bond. Thus, the required carbohydrate structure is attached to an amino acid residue (typically asparagine for *N*-linked glycopeptides and serine or threonine for *O*-linked glycopeptides). Several reviews cover strategies for preparation of *N*-linked and *O*-linked glycoaminoacids, and our previous review covers all major strategies.[1,59,86,87] Any resulting glycosylated amino acid is then typically incorporated into standard solid-phase peptide synthesis (SPPS) techniques. Two factors generally limit this approach: first, the requirement not only for often extensive carbohydrate protection but also amino acid protection regimes and, second, the acid and base lability of glycosylated amino acid residues.[88] As a direct consequence of these limitations SPPS technology is typically limited to synthesis of peptide sequences of up to 50 residues. Necessary protection and deprotection regimes, use of particular supports, including introduction of specific linkers, and coupling methods have, in some examples, all been tailored to be compatible with the presence of carbohydrates. The following section outlines some of the advances since 2002 in formation of naturally or unnaturally linked glycopeptides, the products of which can or could be used in linear, convergent, or mixed strategies to give larger glycopeptides/glycoproteins. The reader is referred to several excellent reviews which cover these aspects in more detail.[86−98]

2.1. Strategies for Formation of the Glycosidic Linkage

2.1.1. Linkage Variety

Most natural glycan−protein linkages are *N*-(Asn) and *O*-(Ser/Thr) glycosides. The majority of motifs for *N*-linked are GlcNAcβ-Asn and for *O*-linked GlcNAcβ-Ser/Thr or GalNAcα-Ser/Thr, although other unusual linkage forms are also observed[99] including GlcNAc-α-Asn,[100] Glc-Asn,[101]

Scheme 2. Methods for Preparation of *N*-Linked Glycopeptides



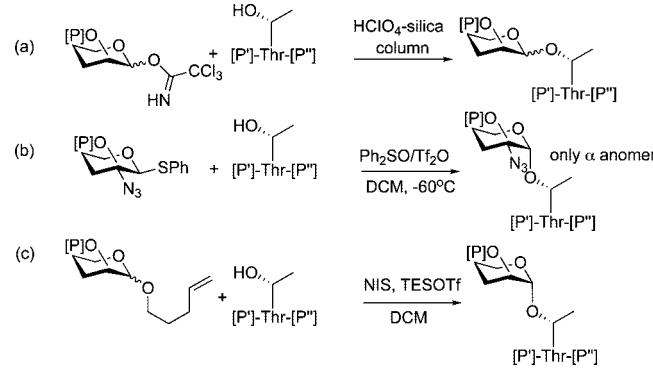
GalNAc-Asn,[102] Rha-Asn,[103] Man-α-Ser/Thr,[104] or even *C*-linked Manα-Trp.[105]

2.1.2. Formation of Natural Linkages

2.1.2.1. *N*-Linked Glycans. The most commonly employed route is reaction of a protected or deprotected glucosamine or equivalent with a suitably protected Asp derivative (Scheme 2a).[106] The Staudinger reaction of glycosyl azides with Asp side chain carboxylates or carboxyl derivatives in the presence of PPh_3,[107] also offers good yields to β-Asn *N*-glycosides (Scheme 2b).[108] Recently, Kiessling and co-workers described a Staudinger ligation for *N*-glycosylation using glycosyl azides and asaparagine-derived phosphinothioesters.[109] In several examples complete β selectivity was observed and the process showed a marked dependence on the type of phosphinothioester used. We recently expanded the Staudinger reaction to be fully compatible with deprotected sugars (Scheme 2c).[110] *N*-Glycosyl asparagines have also been prepared via mild photochemical coupling in which a photoreactive side chain amide of aspartate was condensed with a glycosyl amine (Scheme 2d).[111] Photolysis activates the γ-carboxylate through formation of intermediate **1**, rendering it susceptible to nucleophilic attack with glycosyl amines. Orthogonally protected photoreactive peptides have also been employed by Michael and co-workers as part of their convergent synthesis of *N*-glycopeptide fragments of the glycoprotein hormone hEPO (human erythropoietin).[112] This novel phototransamidation strategy furnished *N*-linked glycopeptides in good yield and avoided formation of aspartimide

2.1.2.2. *O*-Linked Glycans. The *O*-glycosidic bond in *O*-linked glycans is normally installed through glycosylation of the hydroxyl group of a suitably protected Ser or Thr using standard glycosyl donors.[113,114] Trichloroacetimidate (TCA) donors have been used previously, but a recent innovation by Field and co-workers describes an 'on-column' approach where Lewis-acid-catalyzed activation of the TCA occurs with perchloric acid immobilized on silica (Scheme 3a).[115] The procedure also allows in situ purification of products, and a range of glycosylated amino acids was prepared by this route. Activation of thioglycosides using a Ph_2SO/Tf_2O promoter system has been reported by Boons and co-workers in the stereospecific preparation of T_n and T_f antigen building blocks (Scheme 3b).[116] Use of *n*-pentenyl glycosides as glycosyl donors for *O*-glycopeptide synthesis has also been revisited (Scheme 3c).[117]

Scheme 3. Methods for Preparation of *O*-Linked Glycopeptides



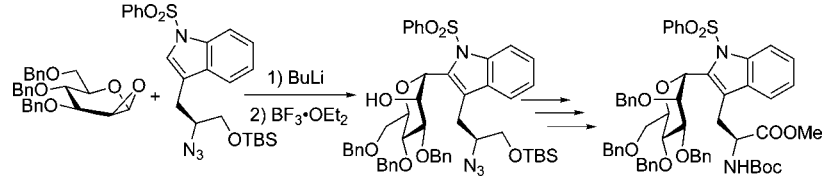
2.1.2.3. *C*-Linked Glycans. A novel variant of glycoproteins that incorporates a *C*-glycosylated amino acid has been identified. Ito and co-workers described the synthesis of one such *C*-linked glycosyl amino acid, namely, C-2-α-D-C-mannopyranosyl-L-tryptophan, via ring opening of a protected 1,2-anhydromannose by a lithiated indole derivative (Scheme 4).[118] Functional group conversion and deprotection furnished a glycoamino acid suitable for peptide elongation.

2.1.3. Synthetic Approaches to Unnatural Linkage Types

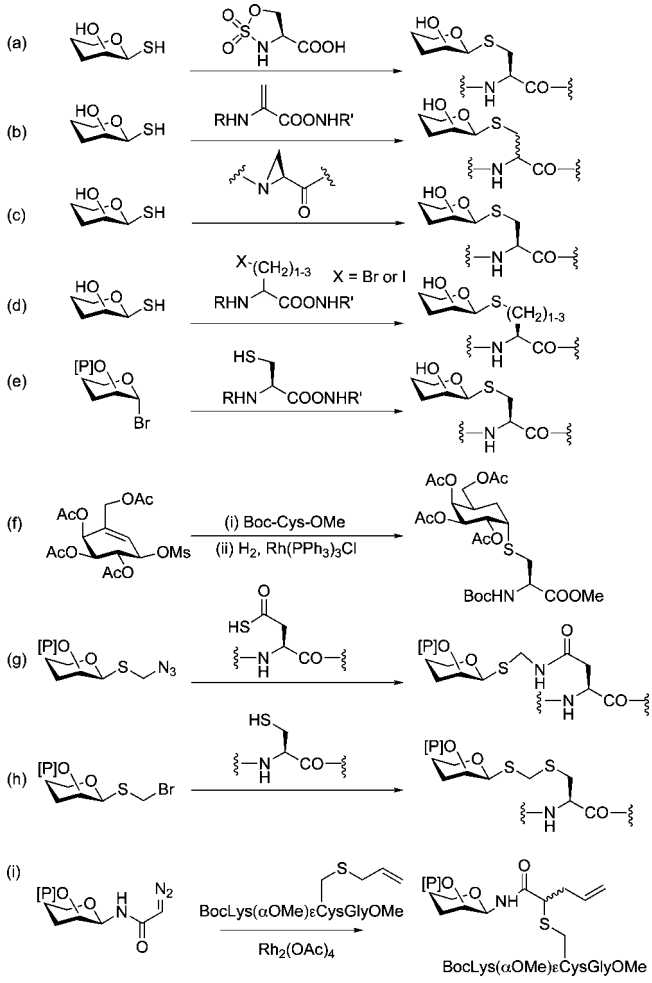
A variety of methods have been developed for preparation of unnatural glycan−peptide linkages, and most of those which follow have generally been applied to, or indeed are restricted strategically to, peptide systems. However, some may also applicable to larger protein systems (Scheme 5a−h). As a primary linkage type, thioethers or variants have featured heavily (Scheme 5).

Use of a serine-derived sulfamidate (Scheme 5a), based on previous methods,[119] allows reaction with thiohexoses to create *S*-linked glycoaminoacids in water. This is a method

Scheme 4. Ito's Synthesis of *C*-Mannosyl Tryptophan[118]



Scheme 5. Methods for Preparation of Thioether-Linked Glycopeptides



that could potentially be adapted to convert *N*-terminal serines to glycosylcysteines in peptides, although some amino acid epimerization was noted which may partly limit its applicability.[120] Michael addition of various protected and unprotected thiosugars to dehydroalanine (Dha) in small peptides (Scheme 5b),[121] introduced through oxidative elimination of phenylselenocysteine, has produced a variety of glycopeptide mimics. To this end, thio-isosteres of the four tumor-associated carbohydrate antigens, T_n, T, ST_n, and 2,6-ST, were prepared as pairs of diastereoisomers due to uncontrolled stereoselective reprotonation of the addition intermediate.[122] Recently, several methods for introduction of Dha into proteins have opened up the possibility of translating this glycoconjugation into full-length proteins.[123−125] In order to better control stereoselectivity, van der Donk and Gin and co-workers employed aziridine-containing peptides (or Azy-peptides) as reagents for preparation of glycopeptides (Scheme 5c).[126,127] Azy-containing peptides can be prepared by standard SPPS using FmocAzyOH as a novel amino acid monomer. Unfortunately, attempts to extend Azy-containing peptides by sequential coupling of single Fmoc-amino acids were unsuccessful due to intramolecular aziridine *N*-deacetylation by the liberated amine upon Fmoc removal at the [Azy+2] position. This problem was overcome through use of a dipeptide building block that allowed simultaneous incorporation of the [Azy+2] and [Azy+3] residues, thereby avoiding aziridine deacetylation. The resulting Azy-peptides were treated with 1-thio-sugars to furnish glycopeptides in 43% yield over 10 steps in a convergent manner.[127] Solid-supported unprotected thioglucose can be used to displace side-chain amino acid iodides to furnish *S*-linked glycoaminoacids.[128] Similarly, iodide displacement of an iodoserine or Mitsunobu coupling using GlcNAc-thiol gave Fmoc-Cys[GlcNAc], although in this study racemization was observed during iodide displacement due to in situ elimination to Dha and subsequent conjugate addition.[129] Peptides containing β-bromoalanine and γ-bromohomoalanine were prepared by direct bromination of peptides or incorporation of a bromoamino acid into the peptide synthesis. These were then coupled with protected 1-thio-sugars to furnish *S*-linked glycopeptides in good yield (Scheme 5d).[130] One major advantage of this methodology is that, due to the mild reaction conditions, epimerization at the α-carbon of the amino acid was not observed. The methodology could also be applied to deprotected 1-thio-sugars, resulting in formation of partially protected glycopeptides.[130,131] More recently, Wong and co-workers adapted this procedure for a one-pot strategy for *S*-linked glycopeptide synthesis with the in situ generation of glycosyl thiolates in the presence of bromoalanine.[132] This method was used in the synthesis of an *S*-linked cycloglycopeptide mimic of tyrocidine, where it should be noted that the *S*-linked analogue displayed a greater inhibitory activity against *B. subtilis* than the natural antibiotic. In a reverse mode alkylation, Schmidt and co-workers employed glycosyl bromides with thiol-containing peptides (Scheme 5e).[133] Homocysteine-containing glycopeptides, prepared by native chemical ligation (see section 2.2.4), were glycosylated in the presence of Na_2CO_3 with glycosyl halides to furnish *S*-linked glycopeptides in good yield. In this fashion, synthesis of an *S*-linked glycopeptide analogue carrying two sugar residues derived from Tamm−Horsfall glycoprotein, the most abundant glycoprotein present in human urine, was achieved (Figure 2).[134]

Use of mesylates to alkylate cysteine residues has also been applied to an isostere of a galactosylated peptide (Scheme 5f).[135] In this approach, a 'carbagalactose' derivative in which the ring oxygen of galactose was replaced with a methylene unit (synthesized using ring-closing metathesis

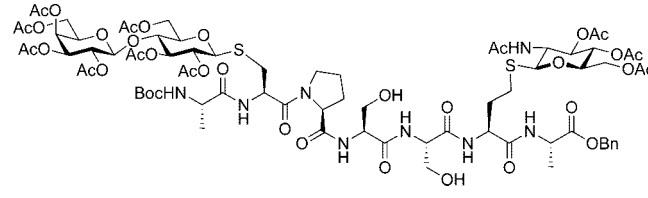
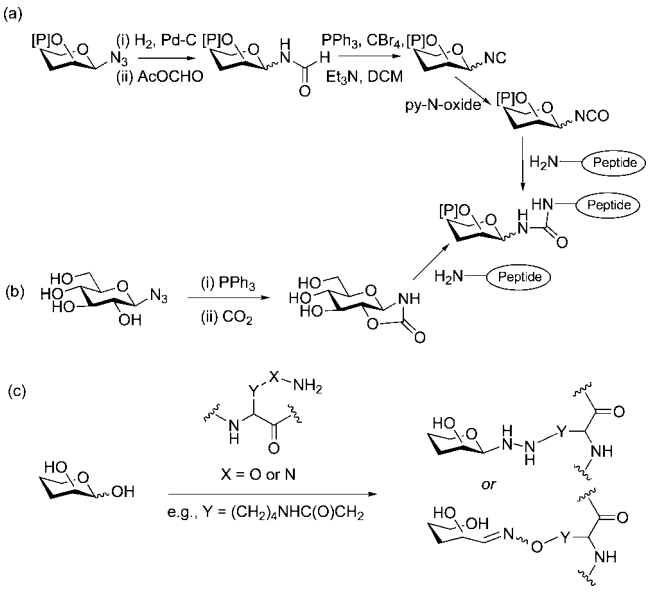
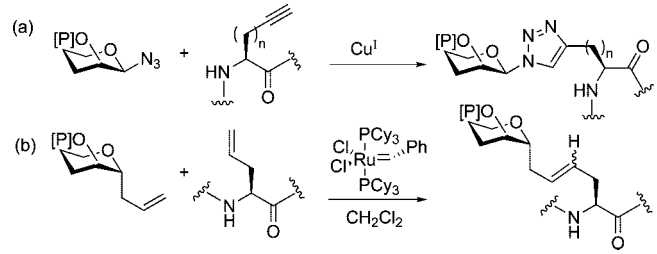


Figure 2. Tamm−Horsfall glycopeptide mimic prepared by cysteine alkylation.[134]

Scheme 6. Unnatural *N*-Linked Glycopeptide Examples



Scheme 7. Metal-Mediated Glycoconjugation Strategies for Unnatural Glycopeptide Synthesis



followed by mesylation) was reacted with a Boc-protected cysteine-containing tripeptide. Reaction of glycosylthiomethyl azides directly with peptide-derived thioacids has resulted in formation of a variety of glycopeptidomimetics (Scheme 5g). In an analogous fashion, a glycosylthiomethyl bromide was coupled to a peptide through the side-chain thiol of cysteine under basic phase-transfer conditions (Scheme 5h).[136] The synthetic pathways and utilities of thioglycopeptides have been comprehensively reviewed elsewhere.[98] Recently, Crich and co-workers reported a modified Kirmse−Doyle reaction for ligation of assorted functionality onto cysteine residues (Scheme 5i).[137] In this approach a 2,3-sigmatropic rearrangement of allylic sulfur ylids, generated from the rhodium-catalyzed addition of a diazoalkane (using stabilized and unstabilized carbenoids) to an allylic sulfide, was used to ligate carbohydrates to a dipeptide.

Recently, Dondoni and co-workers described a Hantzsch-type three-component approach for preparation of *C*-linked glycosyl amino acids.[138] The *C*-glycosylmethyl pyridylalanine products were prepared via a thermally induced cyclocondensation reaction using an aldehyde−ketoester−enamino ester system. Both α and β configurations at the anomeric center could be accessed, and the positions of the carbohydrate and amino acid sectors in the pyridine ring could be controlled.

Glycosylisocyanates prepared by oxidation of isocyanides, in turn prepared from anomeric formamides, react with amines to give urea linkages, and this has allowed one example of a glycosyl−amino acid conjugate to be synthesized (Scheme 6a).[139,140] Steyermark's glucopyranosyl oxazolidinone has been successfully applied to ligation of glucose to amino acids in water (Scheme 6b).[141] The oxazolidinone was prepared from trapping of an anomeric glycosyl aza-ylid with CO_2; addition of either lysine or cysteine results in urea-tethered glycopeptides.

Ligation approaches based on C=N bond formation have also been successfully applied to unnaturally linked glycopeptide synthesis.[142] Functionalization of both α- and ϵ-amino groups of lysine with aminooxyacetyl groups, before condensation of the newly introduced free amine with reducing sugars, generated the desired oxime linkages (Scheme 6c). In a similar fashion, reaction of a 12-aa peptide containing an *N*-terminal aminooxyacetyl function has also been described.[143] Aminooxyacetyl Lys capping and oxime formation with the reducing end of lactose has also been applied to a sulfopeptide.[144] In addition, Imperiali and co-workers developed unnatural tripeptides containing either β-hydroxylamine or β-hydrazide alanine (Scheme 6c) and reacted them with the reducing terminus of *N*-acetylglucosamine.[145]

The Cu(I)-catalyzed [3 + 2] addition between organic azides and acetylenes as reported by the groups of Meldal[146] and Sharpless[147] generally results in formation of the corresponding 1,4-disubstituted 1,2,3-triazoles in high yields. Application of this reaction to azidoglycosides and acetylenic amino acids has allowed preparation of triazole-linked glycopeptides (Scheme 7a).[148−150] Peptide model systems containing azido and acetylenic amino acids were used to evaluate reaction conditions for eventual protein glycoconjugation (see section 3.3) using triazole formation.[23] This methodology has been employed by Danishefsky and co-workers for preparation of complex multiglycan peptides.[151] This chemistry has also been used in the synthesis of a library of carbohydrate-modified cyclic 13-mers related to tyrocidine.[152] The resulting triazole-linked glycopeptides were screened in antibacterial and hemolytic assays and in general demonstrated improved activity as compared to the natural motif. Thorson et al. developed a procedure to rapidly access variant libraries of the cyclic peptide vancomycin based upon an efficient chemoenzymatic route.[153,154] The vancomycin biosynthetic enzyme, GtfE, was used with a variety of natural and unnatural NDP sugar donors to create a glycosylated vancomycin library which contained a unique azide 'handle'. Subsequent ligation of this azido group with other carbohydrate moieties gave a second-stage library that showed increased activity with respect to unmodified vancomycin. More recently, Wang and co-workers employed the Cu(I)-catalyzed [3 + 2] addition to conjugate four units of the D1 arm tetrasaccharide[155] (Manα1,2Manα1,2Manα1,3Manα-) of the high-mannose *N*-glycan onto a cyclic decapeptide. The scaffold showed a high affinity toward antibody 2G12 compared to the monomeric D1 arm, highlighting a significant clustering effect in 2G12 recognition. Recently, Ziegler and co-workers employed Cu(I)-catalyzed [3 + 2] additions for synthesis of a combinatorial library of glycosylated amino acids; a number of orthogonally protected L-asparaginic and (*S*)-3-aminobutyric acids bearing a glycosidic moiety linked through the triazole spacer were prepared.[156]

Olefin cross metathesis (OM) has been employed as a strategy for both co- and post-translational synthetic strategies of *C*-linked glycopeptides (Scheme 7b) in organic solvents.[157−159] Tripeptides incorporating the allyl functionality were prepared and successfully ligated with *C*-allyl glycosides to furnish the *C*-linked glycopeptides in yields of up to 68%. OM, using the Grubbs second-generation catalyst, has been

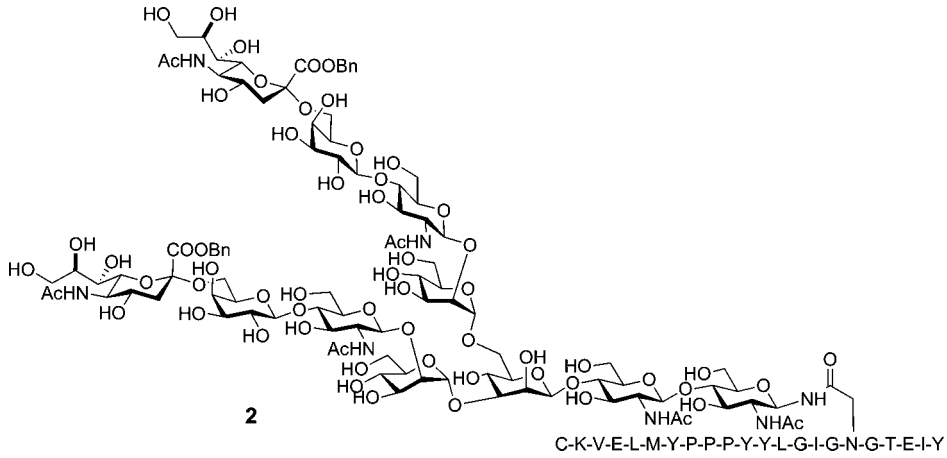


Figure 3. Benzyl-protected intermediate used during the synthesis of CTLA-4.[165]

successfully employed to link α- and *β-C*-allyl glycosides with protected vinyl glycines to furnish products in 57−94% yields; subsequent palladium-catalyzed hydrogenation gave saturated *C*-glycosyl amino acids in good yield.[160] A similar approach has also been employed to produce glycomimetics of globo-H.[161] Interestingly, Nolen and co-workers report increased efficiency with *O*-allyl glycosides than with the corresponding *C*-allyl derivatives in some cases.[162] Discovery of allyl sulfides as privileged partners in aqueous OM has recently allowed the first examples of OM on proteins including glycoconjugation reactions with allyl glucosides.[163]

2.2. Assembly Strategies

2.2.1. Linear Assembly

These techniques rely heavily on SPPS methodology. A noteworthy example is preparation of a cyclic 47-aa HIV-1 V3 domain containing two GlcNAc residues.[164] More recently, Kajihara and co-workers employed linear assembly in their synthesis of the sialylglycopeptide CTLA-4, **2**.[165] The sialyl oligosaccharide, isolated from egg yolk, was converted into an Asn-linked building block which was subsequently elongated by SPPS. Impressively, only benzyl protection of the sialic acid C-1 carboxylates residues was required during the synthesis and allowed use of lightly protected intermediates such as **2** (Figure 3).

Another linear assembly strategy reported by Guo and co-workers describes the preparation of CD-52,[166] a GPI-anchored glycopeptide expressed on virtually all human lymphocytes and sperm cells (Scheme 8). The oligosaccharide amine **3** was linked to a side-chain carboxy Asp residue side chain through DCC/HOBT coupling to form the *N*-glycan. The sensitive α-fucoside linkage remained intact despite treatment with 18% TFA necessary for deprotection of the glycopeptides. The oligosaccharide remained unprotected for the entirety of the glycopeptide synthesis.

Kunz demonstrated an impressive SPPS of an 18-residue fragment of the CD62P ligand PSGL-1 containing an *O*-linked hexasaccharide.[167] A more recent example of linear glycopeptide synthesis has been reported by Wong and co-workers. This method employs traceless Staudinger ligation methodology (see section 5.2) in the coupling of an azide containing peptide/glycopeptide to a phosphinothioester peptide/glycopeptide. This allowed production of glycopeptides of up to 10 residues in length, displaying up to two acetylated monosaccharides.[168]

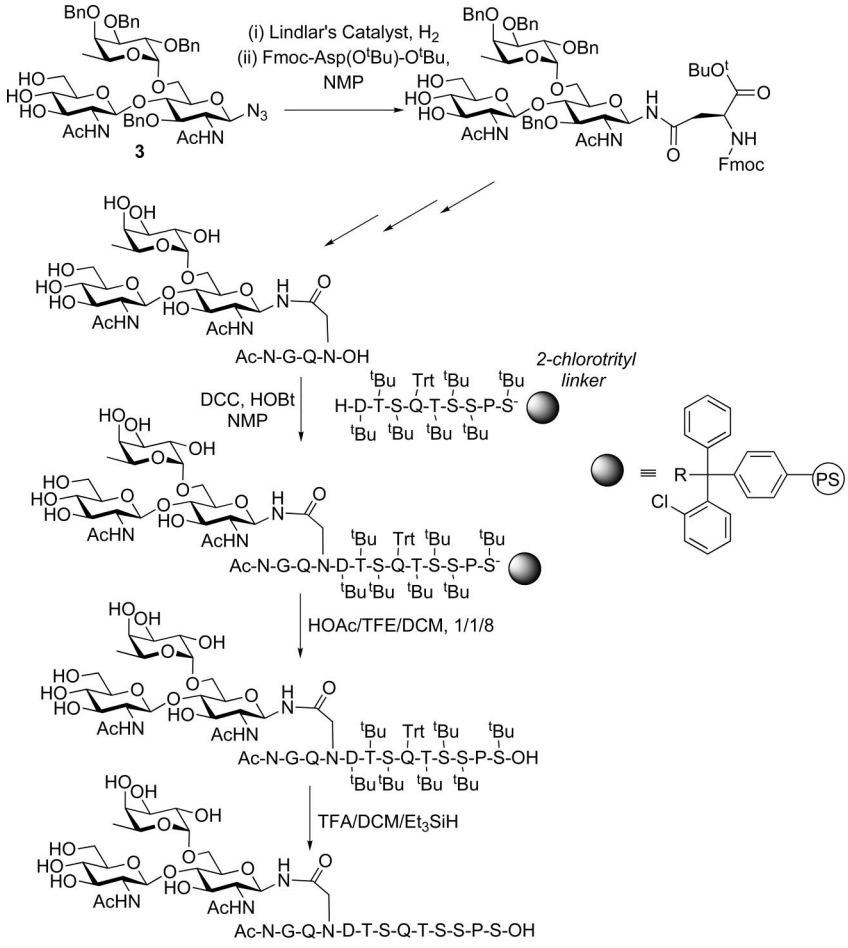
2.2.2. Convergent Assembly

As described above, a linear strategy in glycopeptide synthesis is more usual since direct convergent peptide glycosylation often suffers from low yield. Despite this limitation, some examples of convergent glycopeptide synthesis have been reported. For example, Lansbury and co-workers applied glycosylamine coupling with carboxylates to a prescient convergent strategy for glycopeptide synthesis. HBTU-mediated coupling of GlcNAc glucosamine[169] or more complex glycans, e.g., $Man_5(GlcNAc)_2$,[170] with the aspartate carboxylate side chain in a pentapeptide allowed formation of an Asn-linked *N*-acetylglucosaminyl-containing glycopeptide. As part of their investigation into preparation of the prostate-specific antigen (PSA) glycopeptides, Danishefsky and co-workers developed a 'universal' strategy for preparation of complex *N*-linked glycopeptides from a common glycal precursor (Scheme 9).[171,172] It was found that a sequence consisting of Kochetkov amination[173] of an oligosaccharide bearing a free reducing terminus followed by application of Lansbury aspartylation[170] and then native chemical ligation (NCL) provided a route to complex *N*-linked polypeptides. Evaluation of conditions for optimal glycosyl asparagine formation revealed use of 2 equiv of peptide and 5 equiv of HATU, giving 58−70% yield with concomitant aspartimide and nonglycosylated asparagine formation.[171] This method was applied initially to incorporation of an epimer of the core pentasaccharide into glycopeptides using a glycosylated peptide (CANAS) cassette[171] and then to the core pentasaccharide itself in a subsequent similar study[172] with the same peptide coupling partner. It should also be noted that Kotchetkov amination can be acclerated using microwave irradiation.[174]

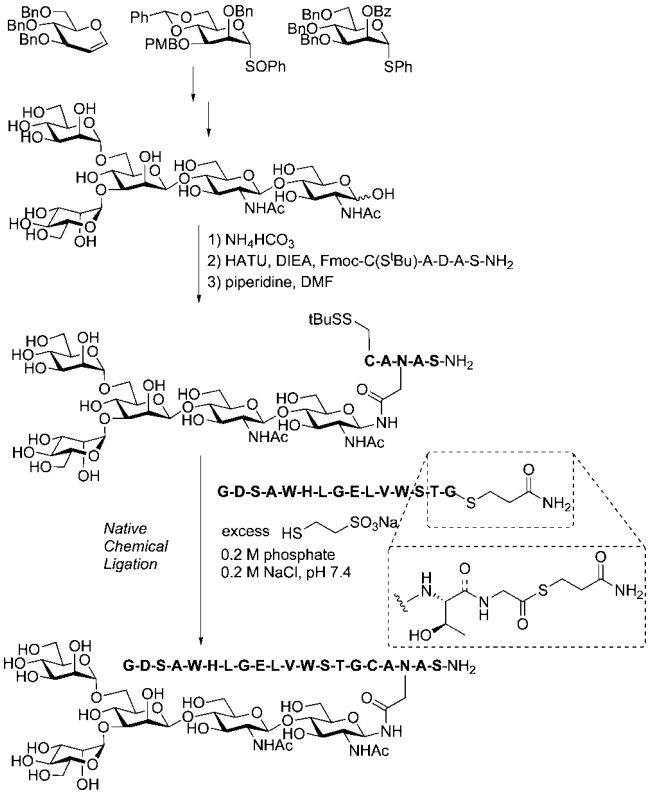
2.2.3. Elaborative and Mixed Assembly Strategies

A mixed strategy which combines both linear assembly coupled with convergent elaboration may also be successful. For example, combination of convergent synthesis and polycondensation has been used to make mucin-like glycoprotein mimics of antifreeze glycoprotein (AFGP).[21,175] AFGP are glycopeptides consisting of a glycosylated tripeptide (Ala-Thr-Ala) repeat. They depress the freezing point of water and are critical to survival of polar and deep sea fish. A variety of glycosyl fluorides, including an *O*-6-sulfated GalNAc, were coupled to a protected Ala-Thr-Ala tripeptide using Cp_2ZrCl_2 activation (Scheme 10).[176] Poly-

Scheme 8. Synthesis of a CD-52 Glycopeptide by Guo et al.[166]



Scheme 9. Danishefsky's 'Universal' Strategy for Preparation of Complex *N*-Linked Glycopeptides[172]

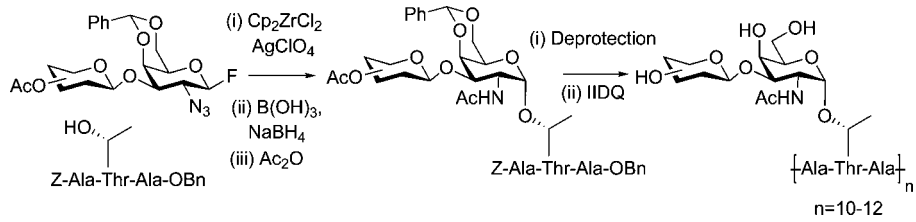


merizations of these units were conducted with 1-isobutoxycarbonyl-2-isobuthoxy-1,2-dihydroquinoline (IIDQ) or 4-(4,6-dimethoxy-1,3,5-trianzin-2-yl)-4-methylmorpholinium chloride (DMT-MM), thereby resulting in 10- to 12-mer repeats (Scheme 10). Impressively, it was demonstrated that such synthetic AFGPs were able to prevent ice nucleation and inhibit cryoinjury in islets cells following cryopreservation and thawing.[20]

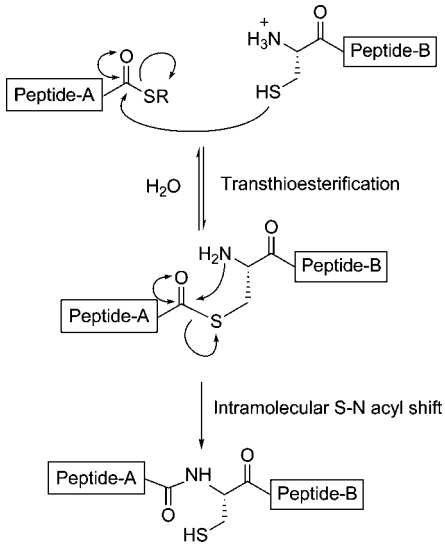
2.2.4. Native Ligation Assembly

Chemical synthesis of glycopeptides has been heavily investigated, and the wide range of techniques employed to install the peptide−glycan link has been comprehensively reviewed.[1,59] Routine SPPS is typically limited to peptides containing <50 amino acid residues; in terms of glycoprotein synthesis this technology is insufficient. However, the linear coupling of such glycopeptide fragments allows chain lengths approaching or even matching those of naturally occurring glycoproteins to be achieved; native chemical ligation (NCL) provides a powerful tool for such glycopeptide coupling.[177−181] This process, which has been excellently reviewed elsewhere,[182−185] involves chemoselective reaction of the *N*-terminal cysteine residue on one peptide with the *C*-terminal thioester of another (Scheme 11). Ligation essentially proceeds via transthioesterfication followed by a spontaneous and irreversible $S \rightarrow N$ acyl shift to give a native peptide bond. This methodology was first introduced for protein synthesis by Kent and co-workers in the 1990s[178] based on observations by Wieland[186] in the 1950s and has since been variously refined to enhance its utility.[185]

Scheme 10. Nishimura's Synthetic Antifreeze Glycoprotein Mimic[176]



Scheme 11. Mechanism of Native Chemical Ligation (NCL)



The fact that NCL can be carried out in aqueous media in the absence of protecting groups has seen its growing application to glycoprotein synthesis, and some impressive examples have been reported (Schemes 12−14). For example, Danishefsky and co-workers described incorporation of the core pentasaacchride[172] and an epimer[171] into the same 20-residue glycopeptide (see section 2.2.2) and synthesis of a highly mannosylated 20-residue fragment of the HIV viral protein gp120 (Figure 4).[187,188]

More recently, Hojo and co-workers described the synthesis of a MUC2 domain composed of 141 amino acids with 42 GalNAc residues (Scheme 12).[189] A smaller MUC2 thioester unit was prepared and sequentially condensed through activation of the thioester with silver chloride. This represents one the largest glycoproteins so far constructed solely employing chemical synthesis. Kajihara and co-workers described the synthesis of the sialylglycopeptide CTLA-4 (113−150) which incorporates two sialyloligosaccharides (Figure 5). This impressive synthesis represented the first application of NCL to construction of complex glycopeptides containing fully unprotected sialyloligsaccharides.[165] More recently, in collaboration with Dawson, they incorporated this glycan into a 76-residue protein, MCP-3.[190]

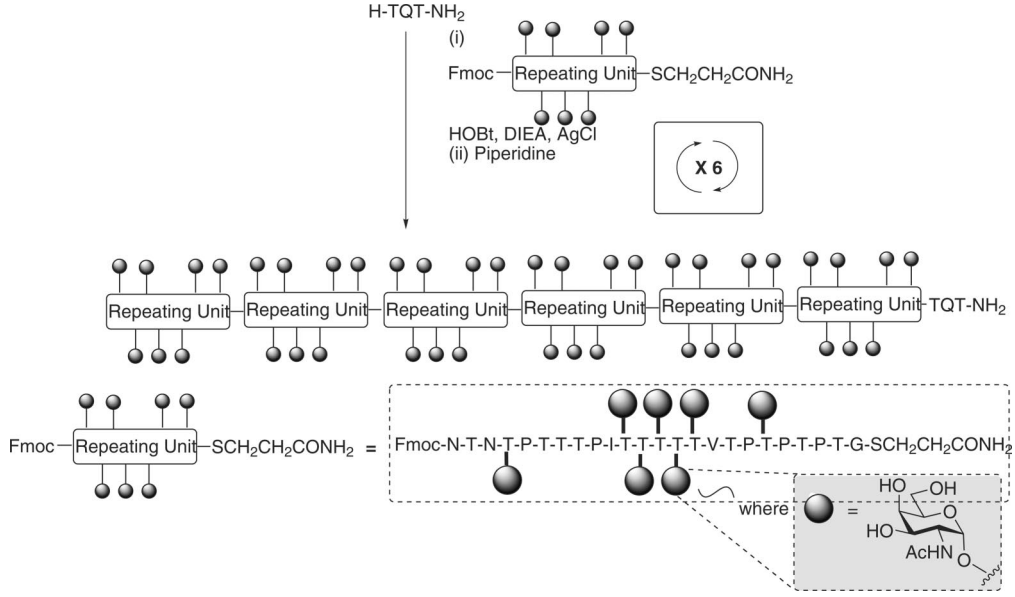
Despite the advances and refinements of NCL technology, certain limiting factors still prohibit its widespread application. One obvious limitation of NCL is the requirement for an *N*-terminal cysteine residue. Cysteine residues are not commonly found in natural peptide sequences and often have to be artificially introduced. This problem has in part been addressed with noncysteine-containing peptides through selective desulfurization of cysteine to alanine[191−193] or β-(SSEt)-Phe derivatives to phenylalanine[194] after ligation. Danishefsky and co-workers also employed a thiol auxiliary (the 4,5,6-trimethoxy-2-mercaptobenzoyl (Tmb) thiol auxiliary, originally developed by Dawson and co-workers[195]) as part of their preparation of a 'trifunctional' (three different glycosylation types) glycopeptide (Scheme 13).[196] Three glycopeptide fragments were coupled sequentially to furnish the desired tripeptide. It was demonstrated that the auxiliary used in the first ligation could be selectively removed chemically or retained during the final ligation (which involved an unmasked cysteine) without affecting the yield. This synthesis of a trifunctional glycopeptide and prior synthesis of a bifunctional glycopeptide[197] also made use of *ortho*-disulfide phenolic peptide esters as latent or 'masked' acyl donors for the NCL coupling.

A glycosylated fragment of RNase has been successfully prepared by Unverzagt and co-workers using SPPS and NCL methodology, thus demonstrating the first example of the synthesis of a complex-type *N*-linked glycopeptide using NCL.[198] An Fmoc-protected asparagine, glycosylated with a complex unprotected biantennary heptasaccharide, was introduced using PyBOP in the presence of DIPEA onto a pentapeptide attached to solid support via a Rink-amide-'safety catch' linker.[199] Activation of the glycosyl asparagine in situ gave the highest coupling yields, and free hydroxyls could be capped without activating the safety catch linker. The resulting glycopeptide was further extended by SPPS and released from the safety catch linker by sodium thiophenolate (Scheme 14). The resulting thioester was coupled to protein fragment RNase41−68 through NCL. This novel linker construct facilitated rapid analysis by LC-MS through acidic cleavage of the Rink-amide linker, thereby avoiding the standard two-step cleavage reaction often necessitated by SPPS.

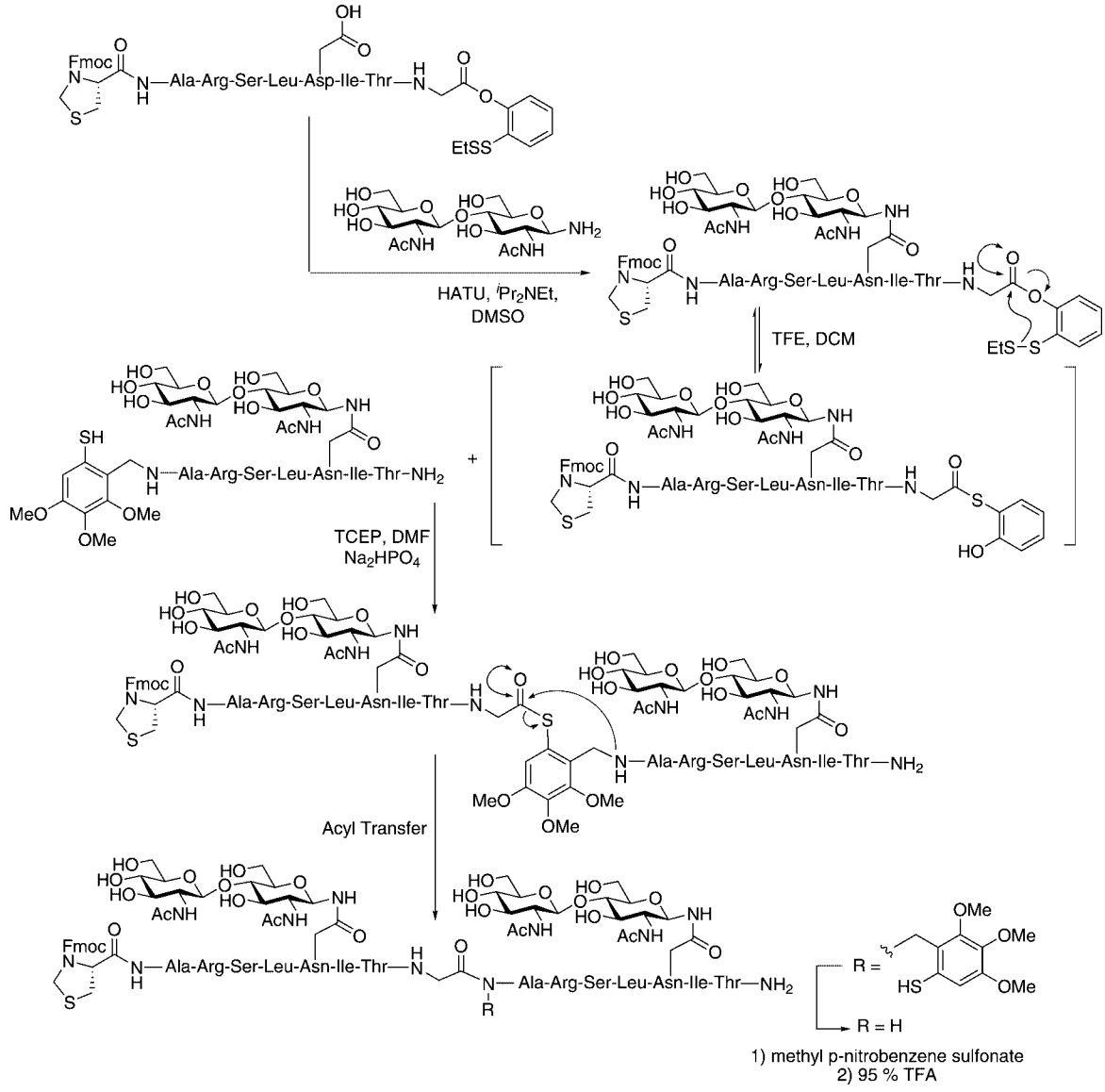
Most auxiliaries used in NCL proceed through five- or six-membered ring transition states. However, Wong and co-workers recently reported a sugar-assisted glycopeptide ligation which proceeds via a surprising 14-membered ring transition state (Scheme 15).[200,201] The sugar moiety of a glycopeptide, modified with a thiol functionality at the C-2 position, was used to mimic the cysteine function at the ligation site. The authors proposed that the sugar moiety controls the proximity of the *N*-terminal amine to the thioester, and this allows $S \rightarrow N$ acyl transfer to occur despite the larger ring size. When coupled with desulfurization using Dawson's conditions this methodology represents a traceless ligation.[202] A disadvantage of this methodology is the necessity to desulfurize the thiol handle in order to regenerate the native carbohydrate *N*-acetyl group as this is not compatible with any side-chain thiols of cysteine residues that may be present in the target peptide sequence. Recently, this disadvantage has recently been circumvented through a second-generation approach that utilizes a thiol side chain in a base-labile *O*-3-thioacetate in the sugar that mediates ligation.[203]

It has been demonstrated by Boons and co-workers that incorporating the glycopeptidic reactants of NCL reactions into liposomes can allow assembly of peptide segments that are normally poorly soluble in aqueous solution.[204] These

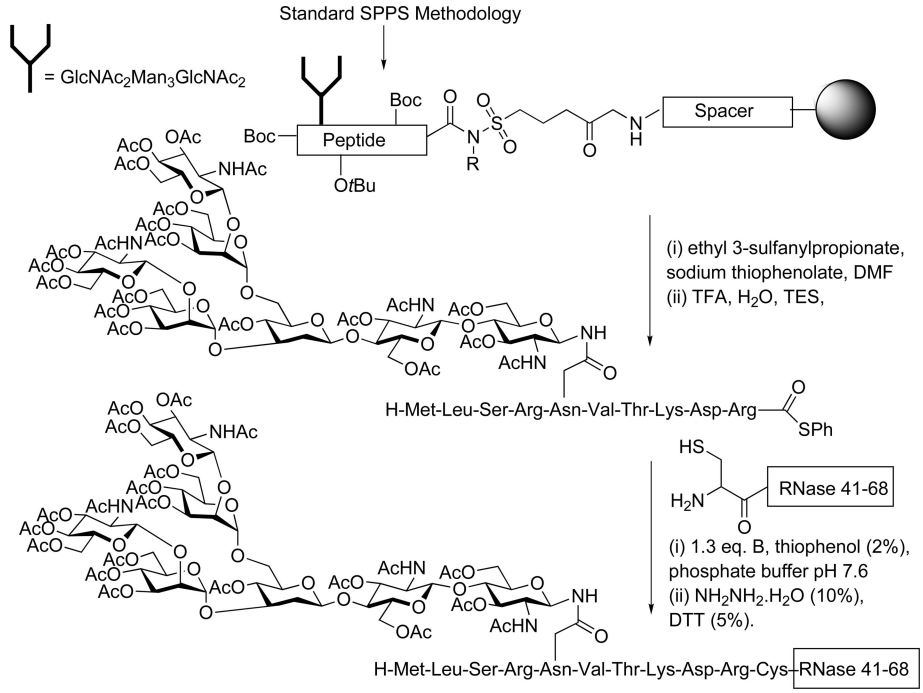
Scheme 12. Hojo's Synthesis of a MUC2 Repeat Peptide Composed of 141 Amino Acids with 42 GalNAc Residues[189]



Scheme 13. Use of Thiol Auxiliary by Danishefsky and Co-Workers in Preparation of a 'Trifunctional' Glycopeptide[196]



Scheme 14. Unverzagt and Co-Workers Demonstrated the First Example of the Synthesis of a Complex-Type *N*-Linked Glycopeptide Using SPPS and NCL[198]



liposome-mediated NCL reactions were found to deliver higher reaction rates and improved yields over existing methods, and the methodology was applied to the synthesis of lipophilic glycopeptides.[204,205]

Interestingly, Danishefksy and co-workers[206] and Wong and co-workers[207] have both demonstrated that native ligation may not even require a cysteinyl or auxiliary thiol nucleophile in the acyl acceptor moiety; direct aminolysis of thioesters in 4:1 *N*-methyl-pyrrolidinone:buffer[207] or disulfide phenolic esters activated by Ag^{+} [206] or phosphine,[206] can also yield glycopeptides. This raises the question as to whether prior ligations thought to be proceeding via NCL may, in fact, have also been accomplished without the intermediacy of a thioester and subsequent $S \rightarrow N$ intramolecular acylation. Indeed, unusually large cyclic transition states have been proposed in some prior studies to accommodate continued concepts of intramolecular transfer.[203,208]

3. Chemical Glycoprotein Synthesis

Typically syntheses of glycoproteins adopt one of two strategies. The first is formation of the putative glycan−protein link early to form glycopeptide building blocks that may then be assembled (linear assembly Scheme 1, mode B). The second is construction of the link late on in a synthesis once the protein scaffold for its presentation is in place (convergent assembly Scheme 1, mode C). Given the requirements for protection that need to be considered in the use of glycosylated building blocks and the instability that may be associated with the link,[88] it is clear why the latter approach is an attractive option. However, although the convergent glycosylation of oligopeptides may be successful, it is limited by a lack of chemo- and/or regioselectivity when applied to proteins, and this remains a key challenge.[88] Convergent protein glycosylation methods (Scheme 1, mode C) have been subdivided here into the categories of chemoselective, site specific, and site selective. This distinction is subjectively based on the overall glycosylation strategy and does not necessarily reflect the inherent specificity or selectivity of a given ligation. For example, use of the same chemoselective (e.g., thiol reactive) reaction will give rise to (i) indiscriminate glycosylation with a variety of available cysteine thiols, (ii) site specificity when used to glycosylate a single, naturally occurring, cysteine thiol, and (iii) site selectivity when used to modify a single cysteine thiol that is artificially introduced to a preselected position.

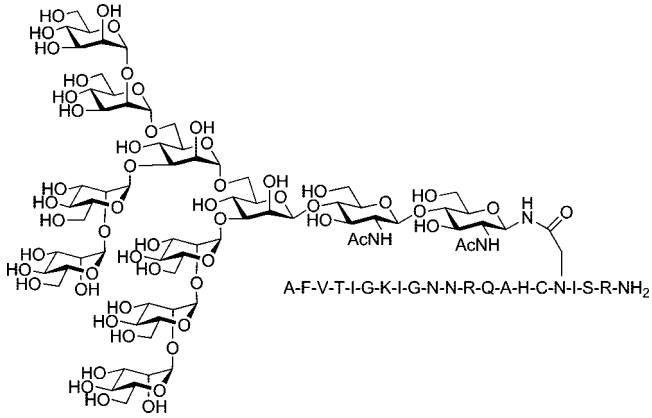


Figure 4. HIV viral protein gp120 fragment synthesized by Danishefsky and co-workers.[187,188]

3.1. Indiscriminate Convergent Glycosylation

Although first introduced some 30 years ago, indiscriminate protein glycosylation is still dominated by two methodologies: 2-iminomethoxymethyl thioglycosides (IME)[42] and reductive amination, both of which rely upon the high abundance of surface lysine residues (Scheme 16a). IME methodology, pioneered by Lee in 1976, involves conversion of the nitrile in cyanomethyl thioglycosides into a methyl imidate group following exposure to sodium methoxide or methanolic HCl. These IME−thioglycosides can be readily reacted with simple amines or protein scaffolds. A recent example highlighting the mild nature of IME reagents

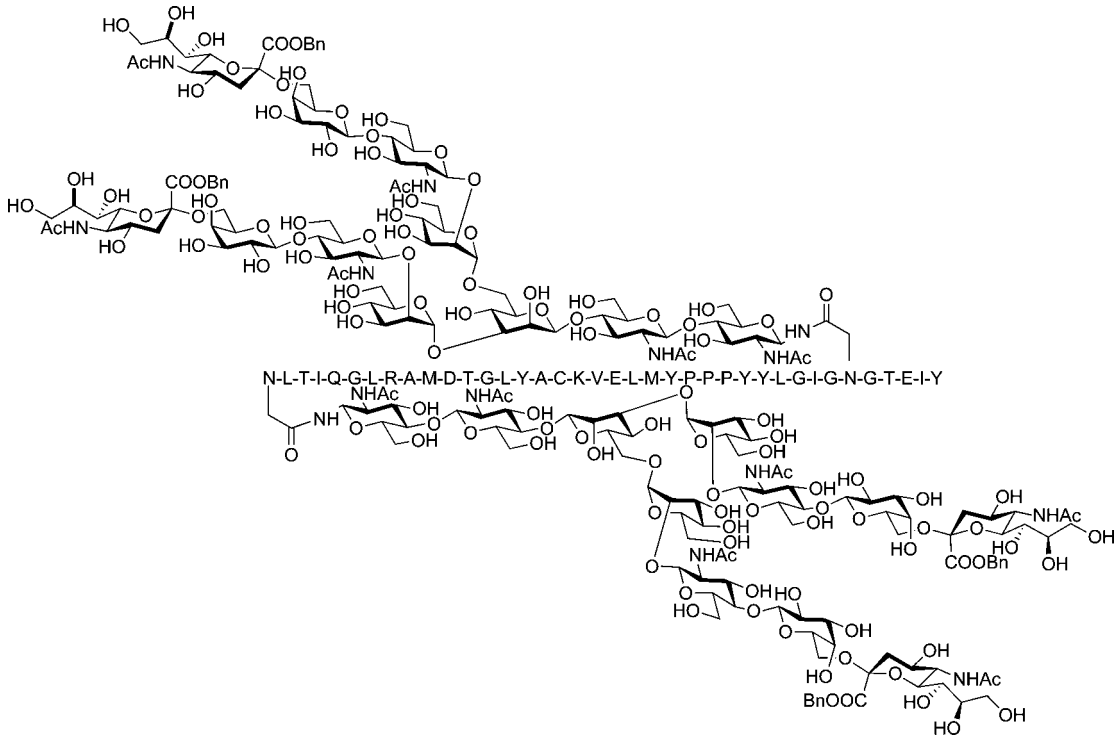
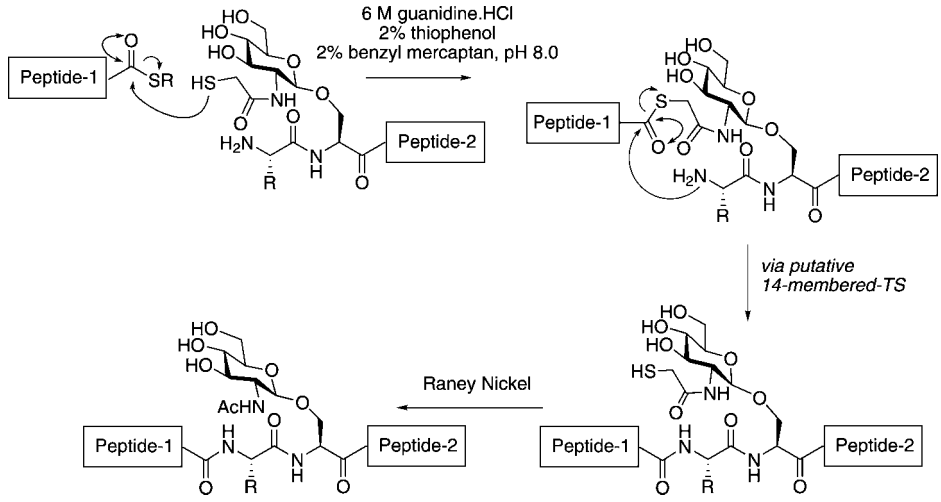


Figure 5. Sialylglycopeptide CTLA-4 (113−150).

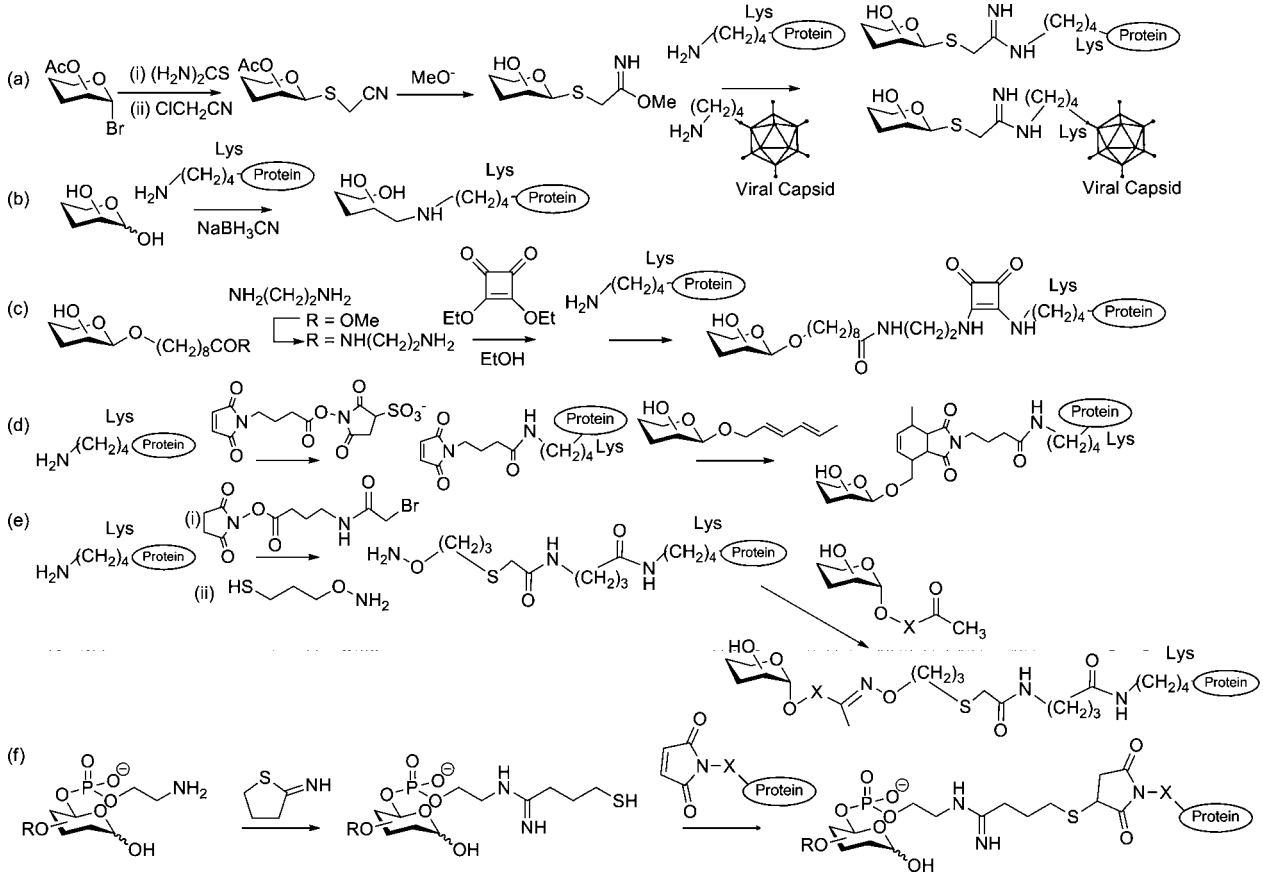
Scheme 15. Strategy for Sugar-Assisted Chemical Synthesis of Glycopeptides[200]



involved incorporation of ~20 000 glycans onto the surface of adenovirus.[209] Use of two different IME's allowed for galactosylation or mannosylation, and careful control of conditions allowed three different levels of modification. Other recent applications of this methodology include LEAPT, an in vivo bipartite drug delivery strategy,[41] which required use of IME reagents to galactosylate the surface of a rhamnosidase enzyme. Reductive amination has been used to modify albumin with the reducing terminus of lactose (Scheme 16b) and has been applied to aldehyde functionality generated by ozonolysis of unsaturated spacer arms.[210] For example, reduction of glycosyl azides and coupling to 4-pentenoic acid followed by subsequent ozonolysis and reductive amination to KLH produced glycoconjugates capable of causing an immunological response in C57BL/6 mice.[211,212] Generally, reductive amination is normally limited by low protein loading levels due to steric factors that are a consequence of short spacer arms. This can be improved by use of a second spacer arm, for example, a hydrazide spacer; the result was a 5-fold increase in the loading of sialyl-GalNAc to KLH,[213] for example.

Squaric acid diesters can also be used as coupling reagents for low molecular weight amines. The first step involves formation of a monoamide link between the squarate and the ligand with subsequent attachment of a second amine via formation of the 1,2-bisamide (Scheme 16c).[214,215] Saksena et al. conjugated the terminal hexasaccharide of the *O*-specific polysaccharide of *V. cholerae* O:1 serotype via three variable alkyl chain linkers to BSA through squaric acid.[216] The authors report a 9:11 molar ratio of carbohydrate to protein after 14 days of reaction. The effect of linker length in this strategy has also been investigated in production of vibriocidal and protective antibodies.[217] A similar carbohydrate has also been ligated to BSA in this fashion.[218] More recently, this approach has been used to create multiply modified human serum albumin carrying up to 19 siaLacNAc residues.[54] These conjugates showed effective inhibition in the binding of adenovirus Ad37 to human corneal epithelial

Scheme 16. Indiscriminate Convergent Protein Glycosylation



(HCE) cells; it should also be noted that the levels of inhibition were dependent upon the average copy number of glycans in the glycoconjugates used. Squarate has also been used to conjugate thio-oligosaccharides to either BSA or tetanus toxoid (TT).[219] The resulting glycoconjugates demonstrated that carbohydrate epitopes of conjugate vaccines could be altered to contain *S*-linked residues that produce similar results to the native *O*-linked epitope. Bundle suggested the use of an adipic acid-based linker system derived from the corresponding *p*-nitrophenyl diester; reaction of aminoalkyl glycosides in dry DMF gave the corresponding amide half-esters in good yields, and these were then used to modify BSA at pH 7.5.[220]

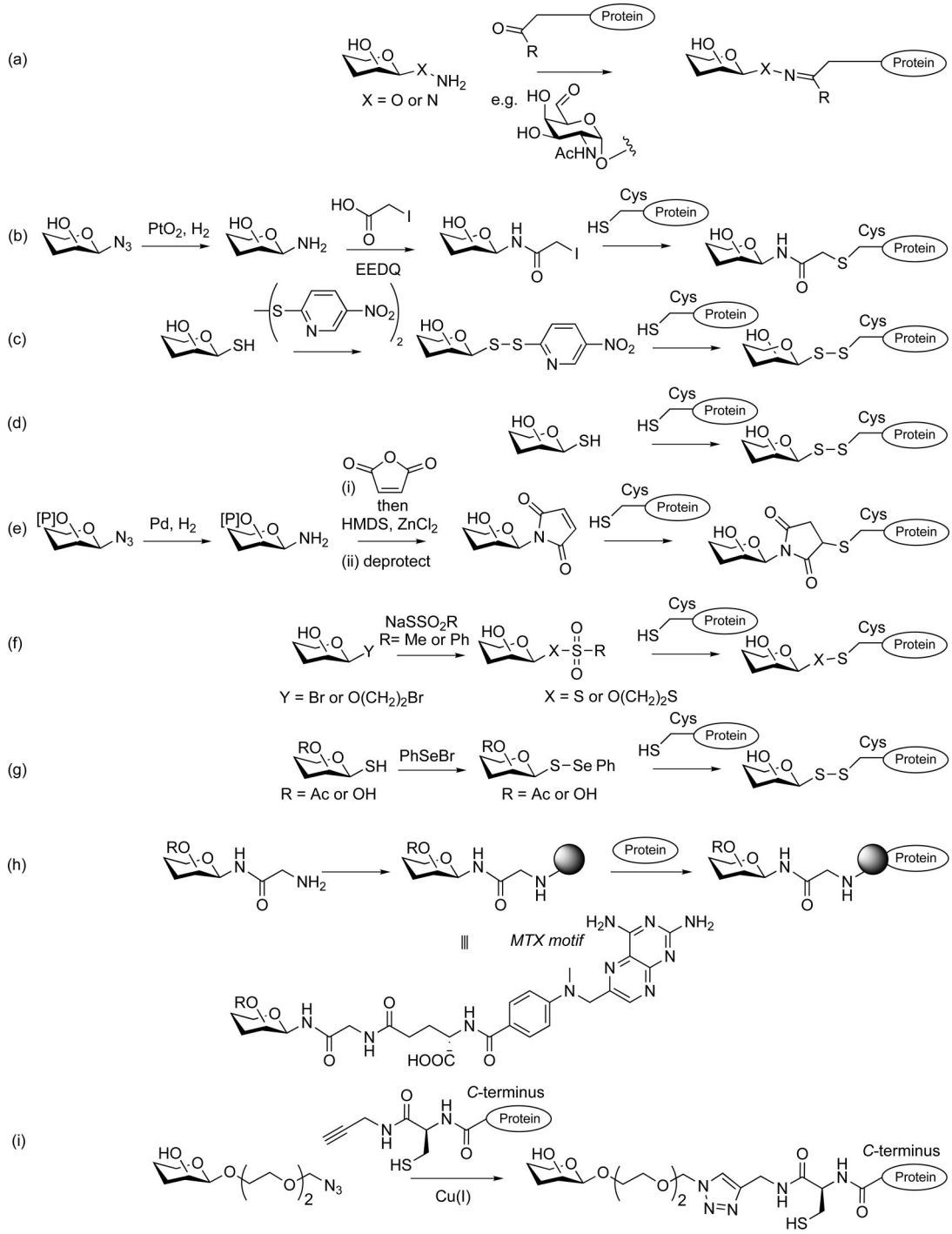
Pozsgay and co-workers recently described the use of Diels—Alder chemistry under mild conditions to glycosylate human serum albumin (HSA) with a series of glycosides, modified to incorporate a conjugated diene (Scheme 16d).[221,222] This methodology involved a sequential process of ligating electron-deficient double bonds, installed through 3-sulfosuccinimidyl-4-maleimidobutyrate linkers, to the protein prior to cycloaddition. The average incorporation after 8 h was 13 hexasaccharides per HSA molecule. The uncoupled saccharide could be removed and recycled by diafiltration in pure form. More recently, this approach has be used to conjugate capsular polysaccharide fragments of the Gram-negative bacterium *Neisseria meningitidis* to human serum albumin with up to 26 saccharide units per albumin.[222] Pozsgay and co-workers also reported indiscriminate glycosylation through an aminooxy-thiol heterobifunctional linker based upon observations of Bertozzi (Scheme 16e).[223] Initially the protein is bromoacetylated through 3-(bromoacetamido)propionate. The bromide is then displaced by the thiol of the heterofunctional linker to a leave an aminooxy handle on the protein. The final step involved chemoselective coupling of the modified protein to keto-derivatized carbohydrates, exemplified through ligation of mono- and tetrasaccharides and a negatively charged ribitol-phosphate to BSA. Sequential acylation (using NHS esters) of lysines with alkyne- and azide-containing modifications followed by Cu(I)-catalyzed [3 + 2] reaction with carbohydrate azides and acetylenes, respectively, has also been used to indiscriminately conjugate the surface of cowpea mosaic virus.[224]

Other linkers have been employed by Seeberger and co-workers in construction of potential vaccines with the aim of generating antibodies against synthetic glycosylphosphatidylinositol (GPI) anchors (see also section 1.4):[46] synthetic GPIs were treated with Traut's reagent[225] to convert primary amine into thiol, which was subsequently conjugated to maleimide-activated KLH or OVA (Scheme 16).

3.2. Chemoselective and Site-Specific Glycosylation

In order to increase the selectivity and predictability of protein glycosylation various approaches have been developed, all of which exploit the chemoselectivities of various enzymatic and chemical methods. Bertozzi and co-workers introduced an aldehyde tag selectively onto the C-6 of a GalNAc residue in the antimicrobial 19-residue peptide drosocin following the specific action of galactose oxidase (Scheme 17a).[226] This aldehyde was next chemoselectivity conjugated with aminooxy glycosides to introduce further oligosaccharides via formation of an oxime in an analogous manner to the use of other spacer-arm hydrazides in reactions with cell surface aldehydes.[227] This

Scheme 17. Site-Specific Protein Glycosylation Methods



use of galactose oxidase and aminooxy chemistry has also been adopted by Shimaoka et al.[228] In a similar approach, Dumy and co-workers employed aminooxy-derivatized T_n-antigens in so-called regioselectively addressable functionalized templates (RAFTs).[229] These cyclic peptide motifs were used to present both T_n antigens and a $CD4^+$ helper T-cell peptide motif (KLFVWKITYKDT) from type-I polio virus. Both modifications were attached to the template in a sequential manner with the aldehyde tag being introduced to the RAFT platform via the action of $NaIO_4$. It should be noted that in vitro and in vivo studies clearly demonstrated recognition of the carbohydrate by T_n-specific monoclonal antibodies and the nonimmunogenic nature of the template itself, thereby highlighting such RAFTs as potential synthetic vaccine scaffolds. $NaIO_4$ has been used to treat recombinant human acid α-glucosidase with the aim of oxidizing sialic acid diols to aldehydes.[44] These aldehydes were then condensed with glycosyl hydrazines to remodel the glycan to contain mannose 6-phosphate motifs. The resulting constructs were intended for use in enzyme replacement therapy to treat Pompe's disease (see section 1.4). Similar chemoselective C═N ligations have also been previously applied to various glycopeptide syntheses.[143−145,230,231] It should be noted that these strategies can offer the potential for site-selective glycosylation when coupled with other methods for introduction of a ketone tag (see below and section 3.3).

Adoption of Flitsch's iodoacetamide methodology[232] (see below and Scheme 17b) by Bertozzi and co-workers enabled chemoselective ligation of Glc- and Gal-iodoacetamides to

a C-3 thiol in a GalNAcα-Thr peptide.[233] The required thiol-protected glycopeptide was constructed via SPPS to give a glycosylated 17-aa sequence corresponding to the *N*-terminal region of P-selectin glycoprotein (PSGL-1). Subsequent thiol deprotection and exposure to Glc- and Gal-iodoacetamides allowed the synthesis of a Glc/Galβ(1,3)GalNAcα-Thr mimetic. It should be noted that no selectivity was observed if a cysteine residue was present in the peptide backbone. In a similar fashion, alkylation of C-2-modified terminal mannoses of the *N*-linked core pentasaccharide with bromoacetamido trisaccharides allowed preparation of a biantennary *N*-linked glycopeptide of CD52.[234]

Although chemoselective, some of the above approaches have circumvented the heart of glycoprotein/peptide synthesis, formation of the putative carbohydrate−protein link, through the requirement for a pre-existing glycan marker in the peptide/protein structure. Similarly, many enzymatic methods also necessitate that a protein−*N*-glycan link be present from the start (described in section 4.2).[164,235−241] As these latter methods alter one glycan structure for another they are therefore better termed glycoprotein remodeling (GPR) and offer the glycoscientist no choice over the site of glycosylation. Several methods have been proposed that tackle this central issue through exploitation of specific functionality present in the protein or peptide sequence. The increased nucleophilicity and comparative rarity of thiols, found in cysteine residues, make them good nucleophiles to react with soft, thiol-specific electrophiles. Among the first approaches was that of Flitsch and co-workers, who reacted the α-iodoacetamide of *N*-acetyl-D-glucosamine with bovine serum albumin (BSA) to modify the single free cysteine present (Scheme 17b).[232] The technique is also compatible with SPPS, as shown by the in situ glycosylation of a 12-aa peptide sequence, which was further extended via SPPS to produce a glycosylated 29-residue section of EPO.[242] The resulting glycopeptide α-thioesters also offer possible access to NCL strategies. A more vigorous test of the use of carbohydrate haloacetamides, beyond simple monosaccharides, saw introduction of chitotriose and a heptasaccharide, stripped from the surface of horseradish peroxidase, to BSA.[243] Furthermore, a sialylated biantennary complex *N*-glycan has been introduced to an 11-mer peptide,[244] and a glucosylated $Man_9GlcNAc_2$ motif has been conjugated to Cys152 of dihydrofolate reductase (DHFR) or Cys212 of carbonic anhydrase using this method.[245] Macmillan and co-workers recently used copper-catalyzed Huisgen coupling reactions[147] to link GlcNAc via a triazole linkage to iodoacetamide reagents.[150] These were then used to alkylate the Cys in a 12-aa peptide. The resulting peptide was then coupled to a human EPO fragment (1−19) via NCL. Impressively, Ito and co-workers employed iodoacetamides to act as covalent irreversible inhibitors of yeast Png1, a deglycosylating enzyme that releases *N*-linked glycans and relies upon a catalytic triad of Cys191, His, and Asp for its mode of action.[246] In this study, using a chitobiose iodoacetamide, only Cys191 was found to be modified among the 14 available Cys residues in the protein (as determined by tryptic digest), suggesting that the high affinity of Png1 toward the chitobiose core of the iodoacetamide allowed for this unusual specificity. This result was also confirmed by the observation that BSA, proteins from *E. coli* extract, and EPO were all inefficiently glycosylated unless a very high concentration of the carbohydrate was used.

Cysteine modification has also be exploited by Boons and co-workers using either dithiopyridyl methodology to make disulfide-linked BSA-*N*-acetyl-D-glucosamine constructs (Scheme 17c)[247] or aerial oxidation (see section 3.3) with disulfide exchange to form a variety of glycopeptides/proteins (Scheme 17d).[55,248] It should be noted that development of a disulfide exchange approach was required due to the incompatibility of glycans beyond the size of monosaccharides with dithiopyridyl methodology. Cysteine modification can also be achieved with glycosylmaleimides; recent direct functionalization of a cysteinyl-containing 11aa peptide and BSA (Scheme 17e)[249] is similar in concept to earlier spacer-arm maleimide-terminated reagents.[213] In a more stringent test of this chemistry Wang and co-workers synthesized an array of mono- and oligosaccharides, which included $Man_9GlcNAc_2$, containing maleimide-bearing spacer arms.[250] Ligation of maleimide-activated mannose and high mannose to T20, a 36-mer of HIV-1 gp41, proceeded smoothly in good yields. More recently, thiolated KLH or a universal T-helper epitope cysteine-containing peptide has been modified with a spacer-linked malemide functionalized with a novel dendritic display of oligomannose.[251] These carbohydrate clusters have been shown to bind to the HIV-neutralizing antibody 2G12,[252] and the conjugates were used in an attempt to raise antiglycan neutralizing antibodies. However, rabbit immunization studies revealed that only a small fraction of raised antibodies were directed instead toward the glycan antigen with the majority being directed to the conjugate linker; as a result, the antisera showed minimal cross reactivity to HIV-1 gp120.

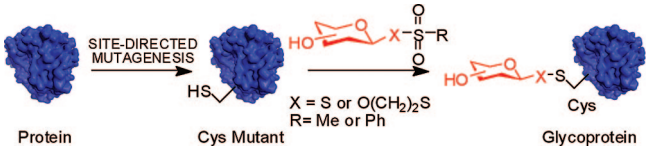
In an unusual strategy Ito and co-workers created noncovalently bound glycoproteins through use of dehydrofolate reductase (DHFR) and a linker moiety, methotrexate (MTX), its high-affinity ligand (K_D < 1 nM) (Scheme 17h).[253−255] A library of MTX−carbohydrate conjugates, prepared from combination of glycosyl amine, a glycine spacer, and MTX, all formed 1:1 complexes with DHFR. The tight binding in these complexes also facilitated facile isolation and purification of the conjugates through lectin affinity chromatography.

The Cu(I)-catalyzed [3 + 2] addition between organic azides and acetylenes has also been used to ligate GlcNAc to mannose binding protein (MBP) through a PEG spacer (Scheme 17i).[256] MBP was expressed as an intein fusion protein in *E. coli* using the IMPACT system (see section 5.4), and an alkyne tag was incorporated with alkynated cysteine via NCL into the C terminus of MBP. Treatment of alkynated MBP with azide-containing glycoconjugates in the presence of Cu(I) and a tri(triazolyl)amine ligand produced the desired glycoprotein in 6 h.

3.3. Site-Selective Glycosylation

To determine precise structure−activity relationships complete control of glycan and the site of glycosylation are needed. This problem was first addressed through a combination of site-directed mutagenesis and chemical modification.[1,31,32,257,258] This dual-step strategy provides the chemist with a flexible method that allows for both regio- and glycan-specific glycosylation of proteins (Scheme 18). The procedure involves introduction of, for example, cysteine, via site-directed mutagenesis, as a chemoselective thiol tag to preselected positions within a given protein. The free thiol is subsequently modified with thiol-specific carbohydrate reagents (Scheme 17f), such as glycosyl methanethiosulfonates, GlycoMTS,[257,259] or sec-

Scheme 18. Example of 'Tag−Modify' Approach: Site-Directed Mutagenesis in Combination with Chemoselective Ligation Allows for Site-Selective Protein Modification



ond-generation glycosyl phenylthiosulfonates, GlycoPTS.[260] The latter have recently been employed in ^{18}F labeling of glycopeptides.[261]

Initial work employing this strategy chose four representative sites on the serine protease subtilisin *Bacillus lentus* (SBL), which does not naturally contain cysteine, in order to probe varying thiol environmental characteristics and thereby provide a broad test of the methodology. Wide applicability with respect to the sugar moiety was evaluated using a representative library of protected or deprotected, tethered or untethered, mono- and disaccharide methanethiosulfonates. The homogeneous glycoproteins formed allowed the first systematic determinations of the properties of synthetic glycoforms, thus providing detailed glycan structure−hydrolytic activity relationships for a library of 48 glycosylated forms of SBL to be determined.[262] Furthermore, these glycosylated enzymes displayed an enhanced catalytic activity in peptide synthesis and, impressively, syntheses of D-amino acids dipeptides that were not possible using the native protein.[263−265]

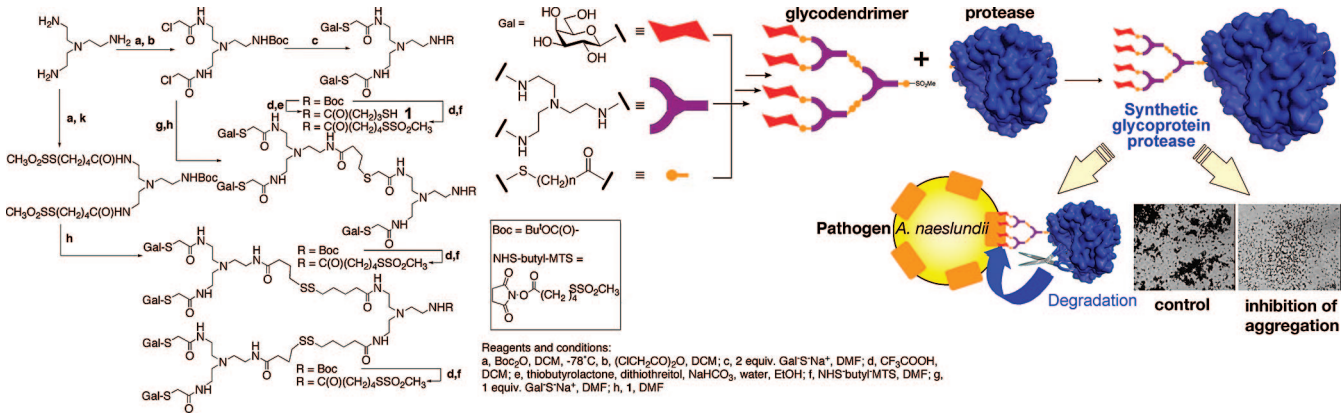
Following on from these promising results, MTS reagents have been employed to tether a variety of biological ligands, including carbohydrates, onto proteins. These ligands can be used to successfully target given proteins to receptor proteins, demonstrating increased selectivity in some cases by over 350-fold.[266] When the targeted protein is a protease this tactic can be used for targeted receptor protein degradation. Thus, initially the appropriate ligand acts as a targeting mechanism to bind the receptor protein, which is then catalytically degraded by the attached serine protease. More recently, the glycoMTS method has also allowed the synthesis of the first examples of a homogeneous protein bearing symmetrically branched multivalent glycans in which both the site of glycosylation and the structure of the glycan introduced have been predetermined (Scheme 19).[267] This new class of glycoconjugate, the glycodendriprotein, allowed for increased protease localization, demonstrated by a high level of coaggregation inhibition between the pathogen *A. naeslundii* and copathogen *Streptococcus oralis*.[268]

A new generation of cysteine site-selective protein glycosylation strategies relies upon selenenylsulfide-mediated conjugation (GlycoSeS; Scheme 17g).[269] In this approach site-selective glycoconjugation of cysteine-containing proteins and glycosyl thiols through phenyl selenenylsulfide intermediates is possible. Preactivation of either cysteine mutant protein or glycosyl thiol occurs following exposure to PhSeBr (Scheme 20). This procedure was demonstrated on simple cysteine-containing peptides and also shown to be successful on a variety of different proteins using as little as 1 equiv of glycosyl thiol in some cases, thereby highlighting the compatibility of selenenylsulfide-mediated glycosylation with a variety of protein thiol environments. This high-yielding procedure also allowed the first example of multi-site-selective glycosylation, coupling of a heptasaccharide (the largest to date), and first enzymatic extension of a disulfide-linked glycoprotein. More recently, a novel direct thionation reaction of both protected and unprotected reducing sugars readily yields glycosyl thiols,[270] the products of which, when combined with Glyco-SeS, allow for a direct one-pot protein glycosylation method.[269] These disulfide-linked glycoproteins may also be desulfurized to yield thioether-linked analogues.[271]

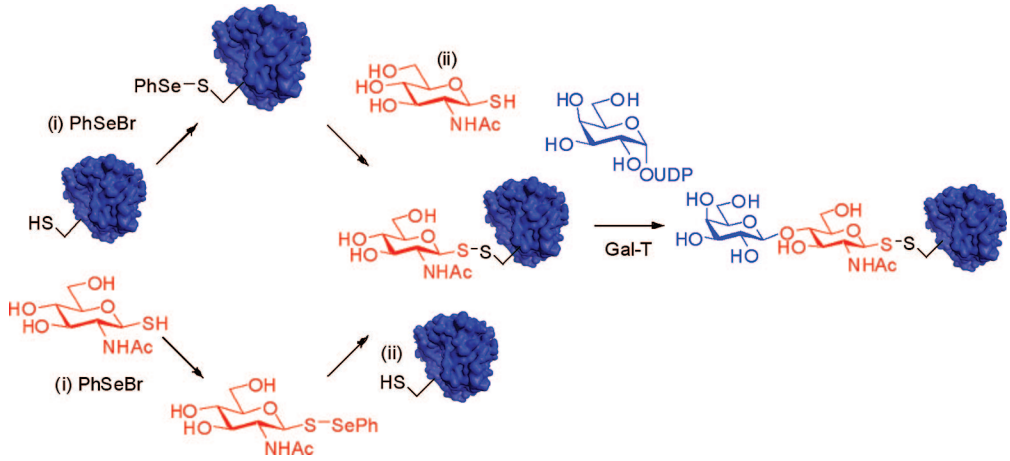
An extension by Flitsch and co-workers of the 'tag−modify' mutagenesis-modification approach has allowed the glycosylation of erythropoietin (EPO) using glycosyl iodoacetamides (Scheme 17b).[272] Several cysteine mutations were introduced at natural *N*-glycan glycosylation sites (N24C, N38C, and N83C) in EPO, where it is known that carbohydrate motifs are critical to the function of the mammalian hormone. Although requiring forcing conditions, 500 equiv of sugar reagent, to obtain partial glycosylation (>60% on N38C, 30% on N24C and N83C as determined by proteolytic digest), the glycosylated EPO was purified using lectin affinity column chromatography to give single, homogeneous glycoprotein. Under these forcing conditions additional nonspecific glycosylation of histidine residues, probably in the His_{10}-tag, was also observed, thereby indicating that such glycosyliodoacetamides may not be absolutely selective; this lack of selectivity was circumvented by carrying out the modification in the presence of excess imidazole. It was possible to confirm glycosylation of N83C unambiguously by proteolytic digest, and it should be noted that, as in the case of BSA, the three disulfide bonds in these EPO mutants were untouched by glycosylation of the single free cysteine. Glycosyl iodoacetamides have also been applied to dihydofolate reductase (DHFR), where monoglycosylated forms of DHFR showed increased thermolytic stability.[273] Again, here, use of glycosyl iodoacetamides at pH 9 resulted in uncontrolled glycosylation of lysine residues in addition to the desired cysteines. Fortunately, this problem could be minimized at pH 7, giving a mixture of monoglycosylated and unmodified protein. Purification of this mixture was accomplished using biotinylated thiosulfonates[274] to scavenge the unmodified DHFR cysteine mutant followed by avidin-based affinity chromatography to yield pure glycoproteins using a procedure similar to that employed previously.[266] More recently, Withers and co-workers employed glycosyl iodoacetamides in the site-selective modification of an *endo*xylanase from *Bacillus circulans* (Bcx). Ingeniously, when a protected thiol in the conjugated sugar was chemically exposed it could be enzymatically extended using Withers' 'thioglycoligase' strategy catalyzed by mutant enzyme Abg E170G.[275]

Boons and coauthors obtained homogeneous disulfide-linked glycoproteins by introducing a cysteine mutation into the Fc region of immunoglobulin G (IgG1). Distinct glycoforms of IgG1-Fc were prepared through disulfide exchange between thiol-containing sugars and oxidized protein in yields of up to 60% (Scheme 21).[55] Unfortunately, Boons' previous dithiopyridyl methodology (Scheme 17c) was not compatible with disaccharides and larger carbohydrates.[276] It should be noted that the procedure did not disrupt internal cystine disulfide bridges. These glycosylated antibodies were subsequently screened for their ability to inhibit the JY cell superoxide burst achieved when FcγRI receptors, found on U937 leukocytes, interact with Fc receptors on JY cells. As expected, the natural fully glycosylated Fc fragment completely inhibited this process, while the aglycosylated fragment showed minimal activity, reiterating the importance of

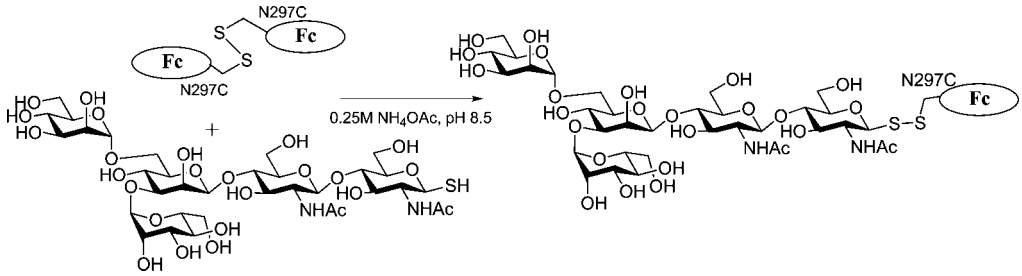
Scheme 19. Construction of a Glycodendriprotein Allowed Targeted Destruction of Pathogen Surface Proteins and Inhibition of Pathogen Binding[268]



Scheme 20. Glyco-SeS[269]



Scheme 21. Boons' Disulfide Exchange Methodology Produces Well-Defined IgG Glycoproteins[55]

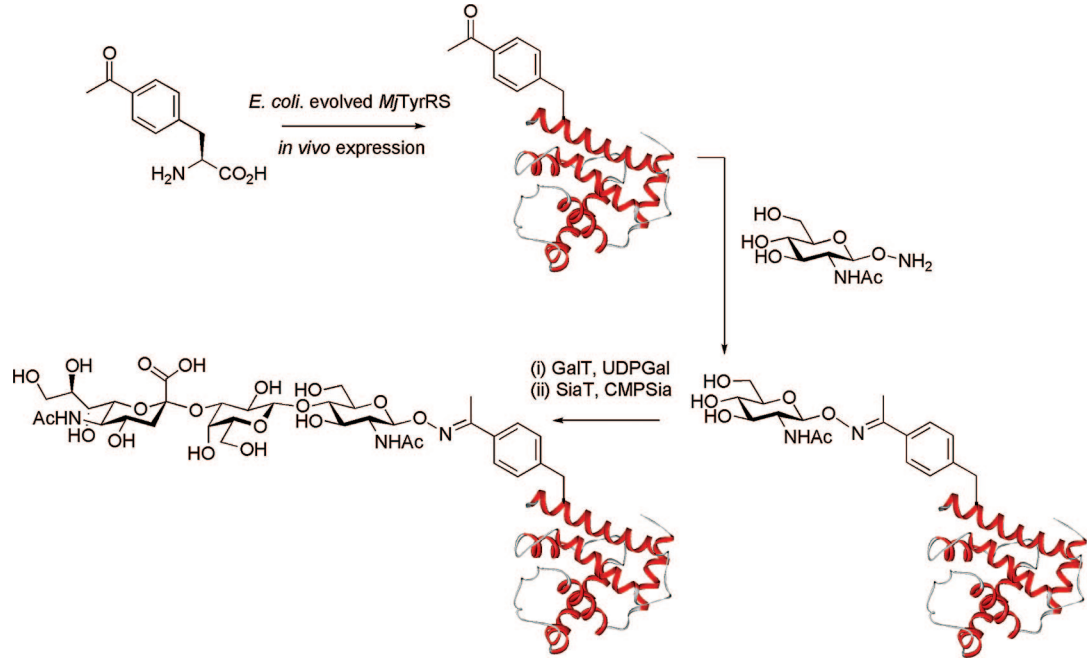


the carbohydrate motif. Pleasingly, the artificial glycoconjugates synthesized in this study displayed partial superoxide inhibition, indicating that protein activity can be restored or more importantly generated by the in vitro attachment of oligosaccharides.

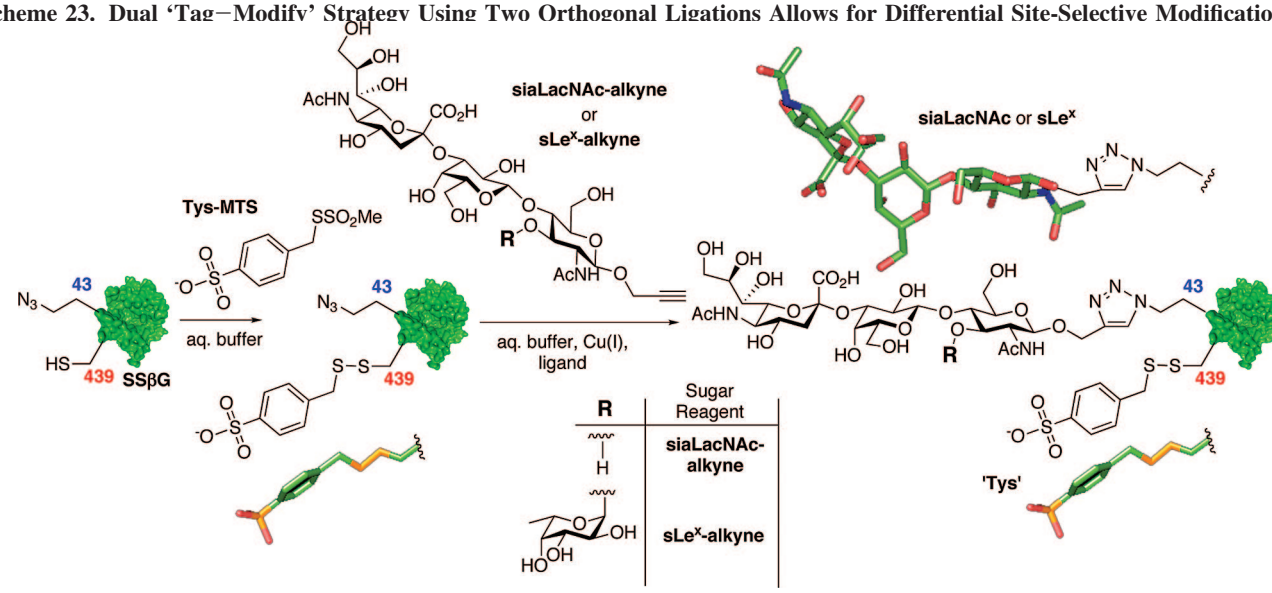
While the above strategies focused on the site-selective glycosylation of naturally occurring amino acids, Schultz and co-workers employed amber triplet codon suppression methodology to introduce unnatural ketone "handles" (*p*-acetyl-L-phenylalanine) into the Z domain of staphylococcal protein A.[277,278] Successful in vivo incorporation of the ketone handle required development of a tRNA synthetase−tRNA pair from *Methanococcus jannaschii*, capable of inserting unnatural amino acids at suppressed amber codon sites. This site-selectively incorporated unnnatural electrophilic amino acid was then reacted with an *N*-acetyl glucosamine aminooxy saccharide[279] for modification using the method developed by Bertozzi and co-workers (Scheme 22, see section 3.2).[226] Furthermore, the resulting glycoconjugate was subsequently extended with galactosyl- and sialyl-transferase to form sialylLacNAc, showing that the unnatural oxime linkage tolerated enzymatic extension conditions.

It has recently been demonstrated that a second mutually compatible chemical protein group, a second chemical "tag", can allow for dual and differential modification.[23,280] This allowed construction of a synthetic protein probe capable of detecting mammalian brain inflammation and disease (Scheme 23). In this second-generation approach three tags were incorporated: thiols in natural cysteine residues and azides and alkynes in the unnatural amino acids azidohomoalanine (Aha) and homopropargylglycine (Hpg). These chemical tags for site-selective conjugation allowed variable modification at multiple predetermined sites in the protein backbone. The first modification was based on cysteine modification using the well-established MTS reagents.[262] The second chemoselective ligation was accomplished using the copper(I)-catalyzed Huisgen cycloaddition (CCHC).[146,147] This reaction has been demonstrated in a variety of aqueous peptide

Scheme 22. Unnatural Amino Acid Incorporation Using Amber Codon Suppression and Chemoselective Ligation Allows for Site-Selective Protein Glycoconjugation[279]



Scheme 23. Dual 'Tag−Modify' Strategy Using Two Orthogonal Ligations Allows for Differential Site-Selective Modification[a]



[a] This approach was used here to synthesize a functional mimic of the glycoprotein ligand of P-selectin, PSGL-1.[23]

systems (see section 2.1.3), but these were the first examples of quantitative site-selective protein modification. Initial optimization studies of CCHC reactions were conducted on a fully competent enzyme substrate SsβG-Aha43, a 10-point mutant of the galactosidase SSβG ($(Met)_{10}(Cys)_1 \rightarrow (Met43)_1(Ile)_9(Ser)_1$) expressed in the Met(−) auxotrophic strain *E. coli* B834(DE3) with Met analogue Aha.[281,282] Coupling between a variety of alkyne glycosides as reagents and the azide in SsβG-Aha43 highlighted a requirement for highly pure (99.999%) Cu(I) in order to achieve >95% protein glycosylation. The CCHC reaction was also shown to be possible (in the 'reverse sense') with alkyne-containing proteins such as SsβG-Hpg1 using azido-derivatized sugars as reagents. Interestingly, SsβG-Hpg1 is a product of the retention of Hpg at *N*-terminal Met site, thereby allowing *N*-terminal modification; a very similar observation has also been made in more recent studies.[283]

Using these two ligations together, first examples of multisite differential modification were demonstrated on SsβG-Aha43-Cys439 using first a glucose MTS reagent to modify thiol and then a galactosyl alkyne to modify azide.[280] This protocol was also successfully applied to a Tamm−Horsfall glycoprotein fragment ($Thp_{295-306}$ Aha298Cys303). In all cases the native functions of proteins were maintained after modification, and in addition, all modified proteins demonstrated additional sugar-specific binding to lectins. Furthermore, this differential chemical, post-translational modification (PTM) of proteins was used to mimic biological processes that are dependent upon PTMs to mediate protein−protein interfaces. Investigations into the P-selectin-glycoprotein ligand-1 (PSGL-1) identified two critical PTMs required for binding to P-selectin in the primary rolling/adhesion phases of the inflammatory response; a sulfate attached to Tyr48 and either of the *O*-glycans siaLacNAc

or sLe^x attached to Thr57.[284,285] As a more vigorous test of the chemical differential ligation strategy, a novel sulfotyrosine mimic (Tys) was installed onto position 439 of the lacZ-type reporter enzyme SsβG and a siaLacNAc-alkyne or sLe^x-alkyne mimic at 43, thereby reconstructing the structural parameters of the natural interaction between PSGL-1 and P-selectin (Scheme 23).

It should be noted that the convergent nature of this approach allowed for incorporation and systematic analysis of truncated and altered glycans. Excitingly, synergistic binding was noted between SsβG(Tys)(siaLacNAc) and P-selectin, which was further enhanced upon fucosylation to SsβG(Tys)(sLe^x), highlighting the fundamental importance of both glycan and sulfation motifs in binding. As a consequence, the resulting differentially modified protein, when used in combination with its galactosidase activity in the presence of X-Gal (a LacZ-type reporter substrate),[286] was used to detect in vivo inflammatory brain lesions resulting from acute[287] and chronic[288] interleukin-1β (IL-1β)-adenovirus-induced inflammation. The LacZ-type reporter mimic was also successfully used to detect cerebral malaria. One of the pathological hallmarks of cerebral malaria is P-selectin-mediated platelet aggregation,[289] and pleasingly, at a 6 day time point the same reporter protein also was able to readily detect infection in a *Plasmodium* disease model.

Dehydroalanine (Dha) has been used for a number of years as a tag in peptidic systems for modification via conjugate addition (Scheme 5b and section 2.1.3) but until recently has not been used for protein modification due to a lack of ready methods for its incorporation. Szostak[123] and Schultz[124] have proposed the use of selenocysteine derivatives that can be installed through reconstituted translation systems or amber-codon suppression methodology, respectively, and then converted to Dha through oxidative selenoxide elimination (along with concomitant Met sulfoxidation). Dha may also be formed directly from Cys without oxidation of Met or alteration of any other amino acids using a novel sulfonylhydroxylamine-induced sulfilimine/sulfenamide elimination.[125] Incorporation of Dha in these proteins allows further protein conjugations including glycoconjugation.[124,125] The sulfilimine/sulfenamide elimination strategy also allows subsequent re-elimination back to Dha, in essence providing a form of reversible, chemically switched conjugation.[125] This latter system has also recently been used to introduce *S*-allyl cysteine into proteins; the allyl sulfide motif turns out to be a privileged motif for aqueous olefin cross metathesis, thereby allowing the first examples of olefin cross metathesis on proteins, including in glycoconjugation.[163]

3.4. Native Ligation Assembly

Expansion in the methods that are available for accessing and linking glycopeptide through NCL-type strategies (see section 2.2.4) has recently seen purely chemical construction of small intact glycoproteins via Mode B NCL assembly for the first time. Examples include incorporation of a complete human Complex-type sialyloligosaccharide into the 76-residue glycoprotein MCP-3 based on conventional coupling of thioesters and Cys-terminated peptides and the 82-residue glycoprotein/peptide diptericin ε using so-called sugar-assisted ligation methods.[201] It should be noted that these sequences are still somewhat short of typical glycoproteins (and the description of protein versus peptide perhaps somewhat a question of semantics) and still in length regions potentially accessible by standard SPPS; indeed, glycosylated diptericins have been synthesized previously using SPPS.[290] Nonetheless, these impressive assemblies highlight the growing power and potential of these strategies in even larger and more complex glycoproteins. The complementary use of expressed fragments also offers a powerful approach (see section 5.4).

4. Enzymatic Glycoprotein and Glycopeptide Synthesis

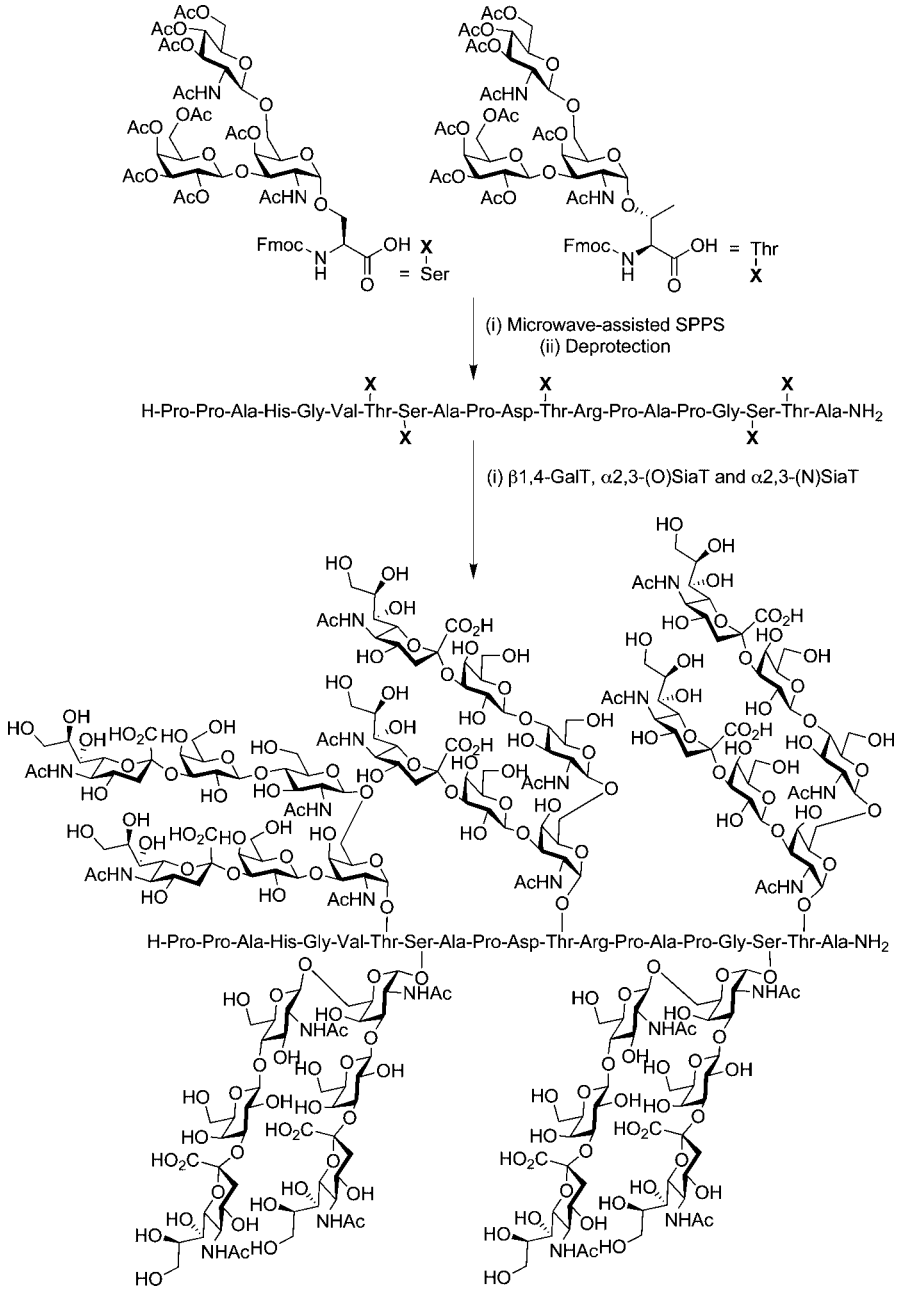
The strategic approaches outlined in the previous sections are equally applicable to enzyme-catalyzed techniques: mode D methods that manipulate glycans include glycan extension (section 4.1) and remodeling (section 4.2), mode C methods for convergent modification (section 4.3), and mode A/B methods for peptide assembly (section 4.4).

4.1. Glycan Extension

Enzymatic methods to develop existing carbohydrate displays on peptide/protein scaffolds have proved particularly successful. In early examples, Paulson and co-workers employed a sialyltransferase and CMP-*N*-Ac-neuraminic acid to restore 95% of the sialic acids to a fully desialylalted protein.[291] More recently, Ito and co-workers employed and explored the function of UDP-Glc:glycoprotein glucosyltransferase (UGGT) in the monoglucosylation of a synthetic $Man_9GlcNAc_2$ MTX complex (see section 3.2), an enzyme/substrate system that plays a critical role in the calnexin/calreticulin (CNX/CRT) protein folding "quality control" cycle.[292] Glycosyltransferases, due to their inherent specificity, are often used to elaborate glycopeptide structures either prior to or after linear assembly (see section 2.2.1). An impressive example from Nishimura and co-workers demonstrated the combination of β1,3-*N*-acetylglucosaminyltransferase (LgtA) and β-1,4-galactosyltransferase in the formation of a library of *O*-linked lactosaminoglycans.[293] Furthermore, this work was extended to include two different α-2,3-sialyltransferases, which when used in combination with β1,4-GalT incorporated up to five sialylated hexasaccharides into a MUC1 glycopeptide (Scheme 24).[294] Much of this work has strong resonance with the use of glycosyltransferases in oligosaccharide synthesis.[48,66,69,295−298]

Glycosyltransferase-mediated elaboration is a chemoselective technique which is particularly powerful for assembly of challenging structures rich in functionality such as sulfopeptides,[299] although strategic considerations may be important with regard to their frequently stringent substrate specificity. It should also be noted that enzyme availability is sometimes a stumbling block in such techniques; in particular, the number of readily available branching GlcNAc-transferases is particularly limited, and commercial supplies of glycosyltransferases continue to be variable. In an excellent display of the power of glycosyltransferases in glycopeptide elaboration, a short sulfated *N*-terminal domain of PSGL-1 has been synthesized on a small scale and characterized by HPLC and MS.[300] Enzymatic methods are also fully applicable on the solid phase with an early example involving the synthesis of dipeptide sLe^x-Asn-Phe on a glycine-linked aminopropyl silica support.[301] Enzymatic extensions have also been performed on water-soluble polymer producing sLe^x-modified 12- and 11-aa peptides.[302,303] Here, the glycopeptide was constructed through standard SPPS to include a keto-containing "transporter" group via a

Scheme 24. Use of Glycosyltransferases in the Synthesis of a Glycosylated Fragment of MUC1[294]

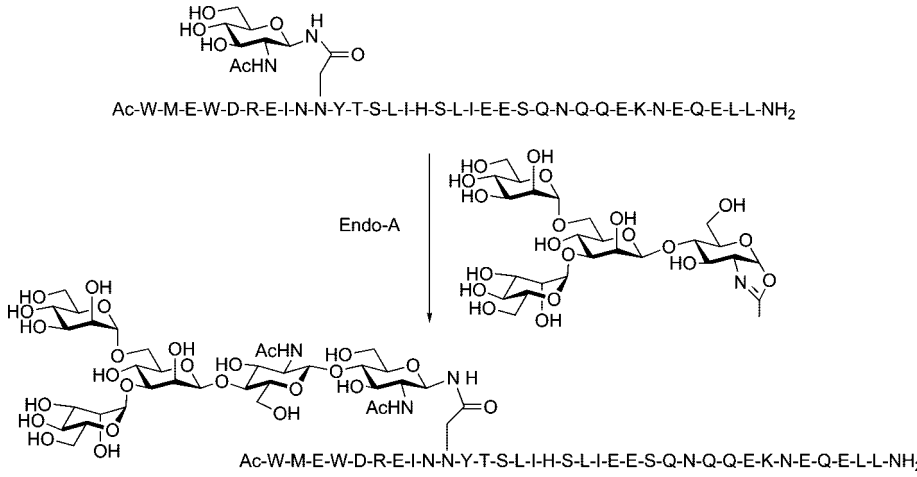


modified *C*-terminus proline residue. The glycopeptide was then released from the SPPS resin and trapped onto a water-soluble polyacrylamide derivative through chemoselective ligation. Subsequent enzymatic extension followed by photolytic cleavage of the transporter moiety furnished the glycopeptide in 12% overall yield requiring minimal purification.

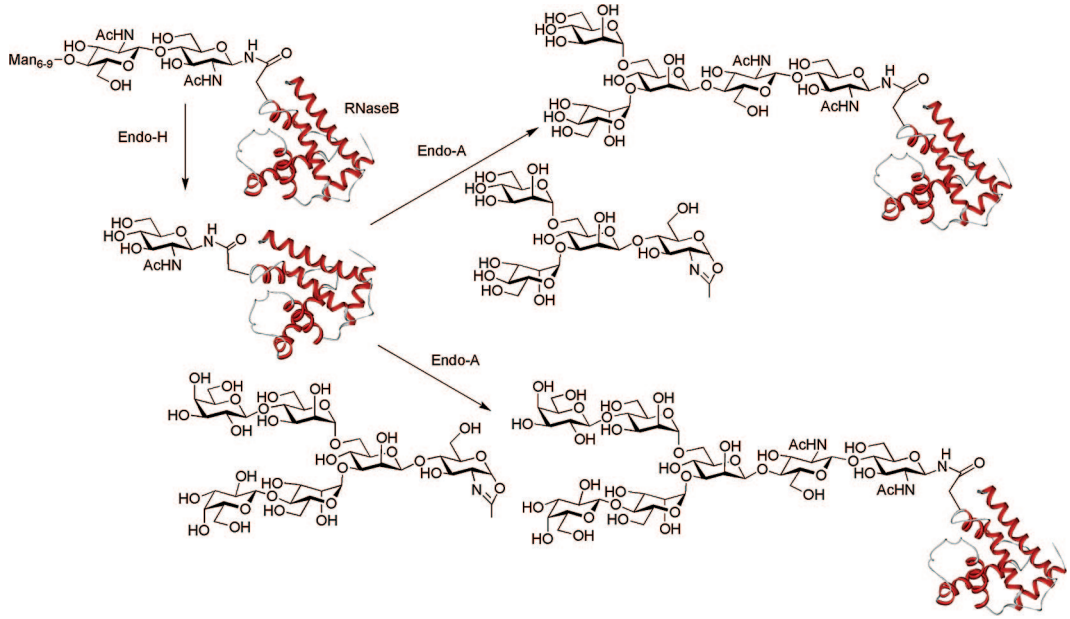
4.2. Glycoprotein Remodeling (GPR)

The combined trimming (typically enzymatic) of existing glycan structures followed by elaboration to alternative ones has been termed "glycoprotein remodeling" (GPR). Fundamental to this technique, due to the difficulty in making the natural glycan−protein link, at least one glycan must remain on the protein to act as a handle for elaboration. A useful "tag" of this kind can be obtained through the use of enzymes, such as endoglycosidase-H (Endo-H), to degrade heterogeneous glycoprotein glycans to a single *N*-linked GlcNAc residue. This can then be enzymatically extended by transglycosylation using either Endo-M, (from *Mucor hiemalis*) or Endo-A (from *Arthrobacter protophormiae*). This chemoenzymatic approach is capable of transferring large carbohydrate motifs onto glycosylated peptides and proteins in a single convergent step, unlike the more commonly used glycosyltransferase-based systems. The first example of such endoglycosidase-catalyzed transglycosylation was first demonstrated by Takegawa and co-workers in 1986 in the synthesis of a $Man_6GlcNAc_2$-modified glycoprotein.[304] Since then, Endo-A and Endo-M transglycosylation has allowed for addition of high-mannose structures (Endo-A) and complex *N*-glycans (Endo-M) onto a variety of peptide backbones.[305,237,235,164,238,240,306,307] Of particular significance is the synthesis of a 34-mer by Wang and co-workers,[241] representing a section of the HIV-1 envelope of gp41 and a recent example by Takegawa and co-workers which demonstrated the combined use of Endo-M and

Scheme 25. Use of Oxazoline Donors and Endo-A to Glycosylated a 34aa Peptide[311]



Scheme 26. Oxazoline Donors in Endoglycosidase-Catalyzed Glycoprotein Synthesis



Endo-A in the successful transglycosylation of eel calcitonin, a 32-mer peptide,[308] where *N*-transglycosylation was demonstrated at up to two sites with both high-mannose and complex glycans. In addition, the monoglycosylated peptides, following exposure to Endo-M, could be separated and subsequently exposed to Endo-A, the result being the first examples of differential multiple glycosylation of a glycopeptide using endoglycosidases. This procedure relied upon the efficiency (yields up to 38%) of Endo-M to transglycosylate complex glycans and the specificity of Endo-A toward high-mannose moieties, which prevented hydrolysis of the initially installed, complex *N*-glycan motif.

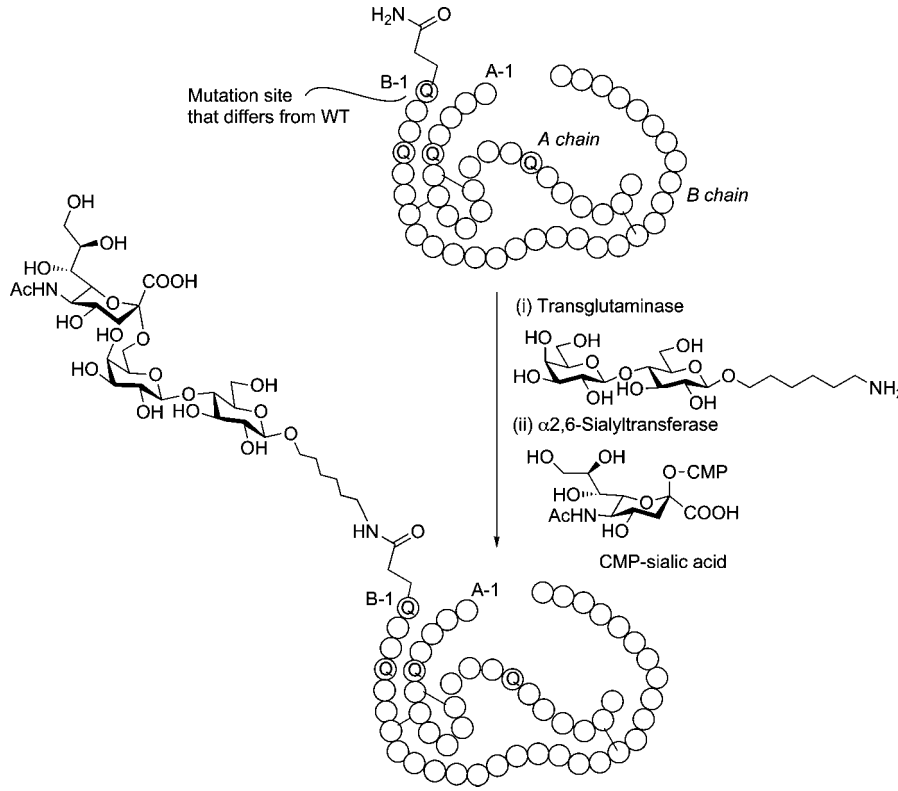
One potential limitation of endoglycosidase-mediated transglycosylation is that it suffers from low yields, typically 5−30%. Post-transglycosylation hydrolysis of glycopeptide product is a key factor in determining overall yield; efficiency can be improved through the use of part organic solvent systems, such as 35% aq. acetone.[309,310] However, more recently, Wang and co-workers largely overcame this potential limitation and synthesized 34-residue and 47-residue glycopeptides through an Endo-A-catalyzed glycosylation in yields of up to an impressive 86% (Scheme 25) using novel oligosaccharide donor substrates.[311,164]The key to this success has been the ingenious use of oxazolines, which act as good substrates for glycosylation; yet the resulting fully incorporated glycopeptide is a poor substrate for hydrolysis, thus minimizing the degradation observed with transglycosylation. A similar method has also been reported by Fairbanks and co-workers.[312]

This strategy has impressively been extended to glycosylation of the protein RNase following the action of Endo-H, allowing synthesis of homogeneous $Man_3GlcNAc_2$-RNase and $Gal_2Man_3GlcNAc_2$-RNase (Scheme 26).[313] Evaluation of the substrate specificity of Endo-A and Endo-M has shown that the minimum requirement for glycosylation is an oxazoline of disaccharide Manβ(1−4)GlcNAc. Substitution patterns are tolerated at both the 3′ and the 6′ portions of the central mannose moiety,[314] but glycosylation yields decrease with increased glycan size due to increased recognition of the product by Endo-A with respect to hydrolysis. The field of endoglycosidase-catalyzed transglycosylation has recently been reviewed.[315]

4.3. Enzymatic Formation of the Glycan−Protein/Peptide Link

An attractive approach to enzymatic glycoprotein synthesis is to directly exploit the enzymes responsible for formation

Scheme 27. TGase-Catalyzed Glycosylation of Insulin[48]



of the sugar−protein link. In *N*-linked glycoprotein biosynthesis[316] the enzyme responsible, oligosaccharyltransferase (OST), cotranslationally transfers a high-mannose core oligosaccharide from a fatty acid pyrophosphate carrier to the side-chain amide of an asparagine (Asn) residue in the consensus sequence Asn-X-Thr/Ser of the nascent glycoprotein, although other sequences are also rarely glycosylated (e.g., Asn-Ala-Cys).[317] Use of this enzyme in isolated form in in vitro glycoprotein synthesis has however met with only modest success. While transfer of carbohydrates to a 17-residue peptide containing an unusual Asn-Asn-Thr-Ser sequence was possible, direct glycan transfer to RNase-A failed.[318] In addition, transfer to sequences in which X = Pro is not possible and those in which X = Trp, Asp, Glu, and Leu are inefficient.[319]

More success has been achieved following the identification and isolation of an *N*-linked glycosylation system in the bacterium *Campylobacter jejuni*; Imperiali and co-workers obtained glycopeptides using the PglB from *C. jejuni*.[320] This enzyme was able to transfer bacterial disaccharides from prenylated donors (GalNAc-α1,3-bacillosamine-pyrophosphate undecaprenyl) onto heptapeptide KDFNVSK and an octapeptide (both are portions from known glycoproteins of *C. jejuni*). Furthermore, Aebi and co-workers demonstrated that PglB is capable of in vitro glycosylation of RNaseA containing a bacterial consensus site, again with unusual bacterial glycans.[321] Interestingly, it was found that fully folded RNaseA was inefficiently glycosylated, but partial or full denaturing of RNaseA greatly enhanced its activity as a substrate. This led to the conclusion that specific substrate conformations are required during the enzyme-catalyzed glycosylation and that glycosylation sites in bacterial glycoproteins will therefore need to be located in locally flexible sites. Investigations into the peptide specificity of prokaryotic PglB highlight the requirement of amino acid determinants beyond the eukaryotic sequon of N-X-S/T. The oligosaccharide component shows a partially relaxed specificity, although recently it has been demonstrated that a C-2 acetamido group in the reducing sugar terminus is a critical requirement for recognition and/or catalysis.[322]

Use of a microbial transglutaminase (TGase) to transamidate the side-chain γ-carboxamide group in the dipeptide Z-Gln-Gly with $-O(CH_2)_3S(CH_2)_2NH_2$-linked glycosides (shorter spacers were unsuccessful) to form an unnaturally *N*-linked glycopeptide has also been described.[323,324] Furthermore, in an attempt to find a universal enzymatic method for the synthesis of homogeneous glycoprotein a combination of site-directed mutagenesis and TGase-catalyzed enzymatic modification has been used to synthesize insulin bearing sialic acid residues (Scheme 27).[48] Nishimura and co-workers replaced the *N*-terminal phenylalanine with glutamine in the B chain of insulin and performed transamidation at this site. In contrast, TGase-catalyzed transamidation of naturally occurring glutamines in wild-type insulin was found to be highly inefficient. Interestingly, the F19 mutant protein successfully underwent TGase transamidation with lactose derivative onto not only mutation site Q1 but also naturally occurring Q4; it was reasoned that mutation at Q1 altered the environment of Q4. Purified Q1 glycoprotein was subsequently sialylated using CMP−sialic acid in the presence of an α2,6 sialyltransferase, giving ready access to Siaα2,6Lac-insulin. More recently, this work has been developed to include dendritic displays of sialic acid using a similar procedure.[296] The in vitro biological testing of these glycoproteins showed decreasing binding of insulin to its natural receptor as the proportion of sialic acid increased but an increased in vivo blood glucose lowering activity of the glycoproteins in STZ-mice. A similar chemoenzymatic approach has also allowed for preparation of glycopeptides as potential influenza inhibitors.[325]

4.4. Glycopeptide Ligation

Use of proteases to catalyze ligation of peptide fragments under kinetic control is well documented.[326] The procedure is highly dependent upon the relative rates of hydrolysis vs aminolysis caused by the competition between water and the peptidic amine acceptor; hydrolysis can be reduced by increasing the percentage of an organic cosolvent. To this end, subtilisin peptidases have been used to catalyze the synthesis of glycopeptides[327,328] in spite of the fact that the natural specificity of these enzymes has hampered these peptide ligations. It was found through careful substrate tolerance studies that the S4, S3, S2′, S3′, and S4′ pockets of the subtilisin active site would accept glycosylated amino acid residues, while pockets proximal to the amide bond being formed/cleaved (S1, S1′) would not. Thus, ligation of Z-Gly-OBz with H-Gly-Ser(Ac_3GlcNAcβ)-NH_2 was successful, but no yield of product was obtained with H-Ser(Ac_3GlcNAcβ)-NH_2. A more substantial example involved a subtilisin-catalyzed peptide condensation to form a 15-residue glycopeptide from two smaller synthetic peptides.[329] This required the synthesis of a 12-residue peptide ester through SPPS using a PAM-modified Rink amide resin, giving a peptide ester. The 12-aa acyl donor peptide ester was then ligated to a 3-residue acyl acceptor glycopeptide using subtilisin in a buffered mixture of water and DMF. Use of this ligation method coupled with other enzyme-mediated strategies culminated in a truly elegant synthesis of a single unnatural glycoform of ribonuclease B (RNase-B) using a protease-catalyzed ligation of fragments of the protein backbone, including a fragment bearing a single GlcNAcβ-Asn followed by glycosyltransferase-catalyzed elaboration reactions of that glycan.[330] More recently, some of the limitations of the natural specificity of the subtilisins have been overcome through engineering of the enzyme ligation catalyst using a combination of site-directed mutagenesis and chemical modification with polar prosthetic groups. A triamino substituent in the S1 pocket of subtilisin led to the greatest substrate broadening, and the resulting biocatalyst was used in the synthesis of a variety of glycopeptides and analogues of compounds showing anti-HIV activity in yields of up to 90%.[265]

5. Molecular and Cell Biological Techniques

In vivo methods that amend the natural machinery of glycosylation present exciting opportunities in the preparation of homogeneous glycoproteins.[331−333] Although prokaryotes do not typically glycosylate proteins, use of eukaryotic systems can circumvent this problem. This approach is made difficult, however, by the large range of potential biosynthetic glycosylation products and thus the corresponding array of pathways that need to be controlled and/or adapted.[334] The consequence of each pathway will differ between cell type, species, and protein. Thus, expression of a particular glycoprotein in one organism will potentially produce different glycosylation patterns than those found by expression of the same protein in another. In this way glycosylation patterns may be channelled and regulated in a particular direction. Thus, expression in, e.g., plants[335] or mice[336] may allow production of patterns similar but subtly altered compared to those in mammalian systems. The research on the *N*-linked biosynthetic pathway in yeast and associated congenital diseases has been comprehensively reviewed.[34,337−340]

5.1. Biosynthesis Augmentation

Regulation of glycosyltransferases involved in the post-translational processing of protein-linked glycans, so-called "glycosylation engineering", can produce dramatic changes in glycosylation patterns. This can be achieved in cell lines or cultures where extra glycosyltransferase-expressing genes have been introduced to augment the presence of particular sugars in glycan structures. This idea has been tested in an attempt to avoid the immunogenicity associated with xenotransplants and therapeutic retroviruses, much of which is caused by the response of circulating human anti-Gal antibodies toward the αGal epitope (Galα1,3Galβ1,4GlcNAc).[341] Wirth and co-workers reduced the expression of αGal on retroviral coat proteins through expression of other glycosyltransferases that compete for the Galα1,4GlcNAc epitope.[342] In order to bias substrate competition, replacement of the localization domain of α2,3-sialyltransferase (ST3), a late-Golgi enzyme, was required with that of FucT1 or α2,6-sialyltransferase, enzymes that both reside in the earlier Golgi cisternae. As a result, the chimeric ST3 intercepted the Galα1,4GlcNAc epitope before it encountered the α1,3-galactosyltransferase (GGTA1) in the Golgi assembly line. Stable expression of this chimeric ST3 in murine cells resulted in significant reduction of αGal levels (34% to 8%), rendering the engineered cells as more suitable hosts for retrovirus production. Glycosylation of secreted glycoproteins has also has been regulated by glycosyltransferase engineering. The biological responses to IgG antibodies are strongly dependent upon the glycan linked to the Fc region of the protein. For example, the tetracycline-regulated expression of the GlcNAc-transferase (GnT-III) involved in "bisecting" (i.e., glycosylation of OH-4 of the β-Man residue in *N*-linked glycans) the core of *N*-linked glycans, thereby blocking elaboration of the branching mannose residues by other transferases (α1,6FucT, ManII, and GnT-II), resulted in an increase in bisected glycoforms from 25% to 50%. This, in turn, increased the antibody-dependent cellular cytotoxicity of the IgG.[343] Recently, Ferrara and co-workers offered further improvement via exchange of the localization domain of GnT-III with that of other Golgi resident enzymes.[344] Depending on the domain employed, expression of the chimera efficiently blocked further elaboration of *N*-linked glycans with the optimal chimera comprised of the localization domain of α-mannosidase II.

The re-engineering of the glycosylation pathway in the yeast *Pichia pastoris*,[338] to mirror the processing of human *N*-glycans, resulted in secretion of glycoproteins with complex *N*-glycosylation.[339,340,345−347] For example, Wildt and co-workers re-engineered strains of *P. pastoris* to mimic the early processing of *N*-glycans in humans to synthesize complex glycans bearing $GlcNAcMan_5GlcNAc_2$.[340] Furthermore, elimination of the "natural" glycosylation pathway and introduction and correct localization of mannosidase I and II, *N*-acetylglucosaminyl transferases I and II, and uridine 5′-diphosphate (UDP)—*N*-acetylglucosamine transporter produced essentially homogeneous glycoproteins bearing the complex human *N*-glycan $GlcNAc_2Man_3GlcNAc_2$.[339] More recently, galactosylation was also obtained in this system when a fusion protein containing the catalytic domain of GalT, UDP-galactose-4-epimerase and a yeast leader sequence, was anchored into the Golgi compartment of a host that was capable of producing biantennary structures that terminate in GlcNAc. This synthetic fusion protein was able

to mediate epimerization of UDP-glucose to UDP-galactose and catalyze transfer of the resulting donor to terminal GlcNAc residues, thereby producing a biantennary galactosyl-tipped glycan.[345] The final step in obtaining terminally sialylated glycoproteins has recently been reported by Gerngross and co-workers.[346] This procedure required knockout of four yeast-specific glycosylation genes and total introduction and localization of 14 heterologous genes. A total of five additional enzymes and proteins were added to the above re-engineered yeast strain: UDP-*N*-acetylglucosamine-2-epimerase/*N*-acetyl mannosamine kinase (GNE), CMP−sialic acid synthase, CMP−sialic acid transporter, *N*-acetylneuraminate-9-phosphate synthase (SPS), and a library of chimeric sialyltransferases (STs) bound to transmembrane anchoring peptide sequences. The best levels of sialylation with the ST library involved fusion of the catalytic domain of mouse α-2,6-sialyltransferase with the targeting peptide signal of mannosyltransferase 1 (MnT1). Yeast containing the above re-engineered expression vectors were cultivated to secrete recombinant EPO, which displayed glycans containing the motif of $Sia_2Gal_2GlcNAc_2Man_3GlcNAc_2$ in greater than 90%, with minor contributions from $SiaGal_2GlcNAc_2Man_3GlcNAc_2$ (7.9%). As expected, the recombinant EPO displayed a dose-dependent response consistent with a biologically active form of the protein. A similar strategy has also been employed in *P. pastoris* to provide a general platform for production of recombinant monoclonal antibodies with human *N*-glycosylation patterns with high homogeneity.[347] Further engineered glycosylation pathways have been comprehensively reviewed elsewhere.[338] An example of engineering *E. coli* to produce glycoproteins has been accomplished by combining protein *N*-glycosylation with lipopolysaccharide biosynthesis at the point were PglB, the oligosaccharyltransferase (OST) found in *C. jejuni*, transfers *O*-polysaccharides from their lipid carrier to asparagine side chains in acceptor proteins. The relaxed specificity of PglB OST (see section 4.3) enabled a diverse range of these glycans to be attached beyond the standard *C. jejuni* glycan.[348]

Glycosylation of larger biomolecular structures may also be accomplished by taking advantage of the often-relaxed specificities of biosynthetic pathways. For example, use of an unnatural *N*-levulinoylmannosamine (Man-Lev) as a precursor in preference to the natural precursor *N*-acetyl mannosamine (ManAc), required in sialic acid biosynthesis, can be achieved simply by feeding it to cells, thereby allowing introduction of a unique ketone tag into sialic acid residues found on the cellular surface.[349,350] This strategy allowed chemoselective introduction of further glycans through reaction with aminooxy and hydrazide-functionalized carbohydrates to form imines and hydrazones (see sections 2 and 3 and Scheme 17a).[351] This strategy is similar to one previously applied following chemical introduction of aldehydes to cell surfaces.[227] The powerful Man−Lev technique was also used to display biotin on cell surfaces.[352] This "cell surface engineering" utilized a biotinylated hydrazide to give cells that were far more readily transfected by adenovirus when pretreated with an antiadenovirus antibody that had been conjugated to avidin. Salvage pathways in CHO cells have also allowed incorporation of a 2-keto analogue of GalNAc into cell surface glycoproteins but not a 2-keto analogue of GlcNAc.[353] Neural cell surfaces have also been engineered by introducing an unnatural *N*-propanoylneuraminic acid precursor.[354] In a similar fashion, use of peracetylated *N*-azidoacetylmannosamine (Ac_4ManNAz) as a precursor allowed the cell surface display of azido-sialic acids.[355] It should be noted that use of acetyl protection is required in all examples to facilitate cell transport, once internalized intracellular deacetylation is thought to occur spontaneously or metabolically. Exposure of the resulting azide-modified cell surface to a biotinylated phosphine-ester resulted in an intramolecular amide formation via so-called Staudinger coupling and hence allowed biotinylation of the cell surface. The advantage of this method is that azides and phosphines are extraneous to nature and therefore unlikely to react in the absence of one another (although phosphine oxidation might compete); therefore, this method can also be used to execute *intra*cellular ligation, an accomplishment recently demonstrated by Bertozzi and co-workers within C57BLl/6 mice.[356] Only mice treated with both the Ac_4ManNAz- and FLAG-peptide-containing reagent phos-FLAG displayed any significant fluorescence when incubated with anti-FLAG modified by a fluorescein isothiocyanate, suggesting that Staudinger ligation occurred in vivo (Scheme 28).[356] A similar approach has also been reported by Harding and co-workers in which the GM3 antigen was altered on tumor cells.[357] Furthermore, a traceless variant of ligation which extrudes the phosphine oxide formed has also been published to allow for "clean" amide formation.[358]

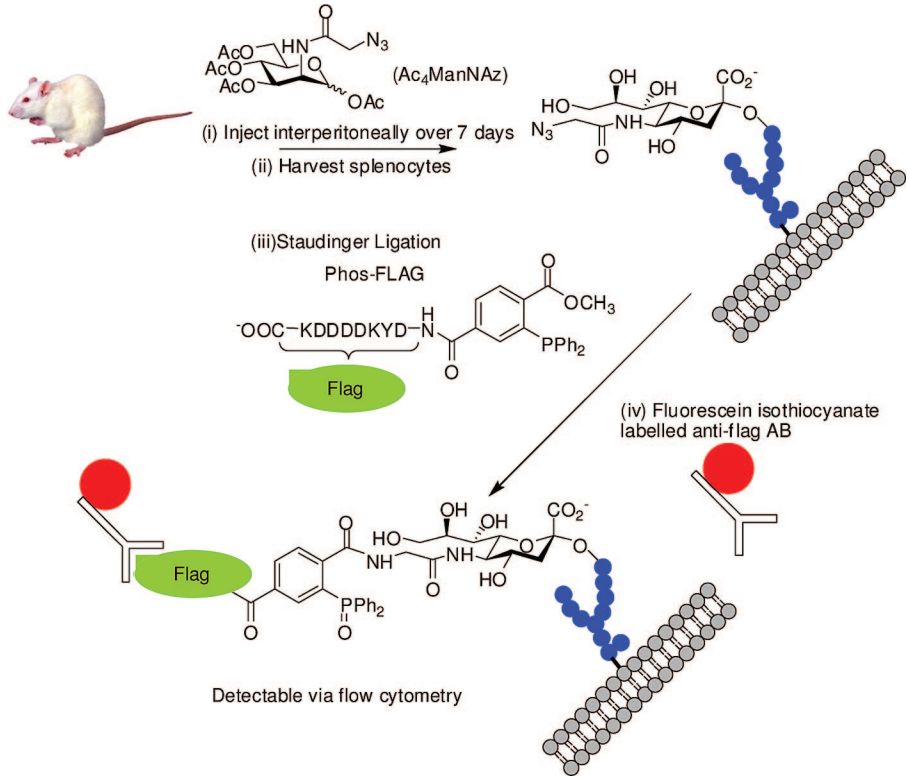
As an alternative reaction of azide, chemoselective 1,3-dipolar cycloaddition with alkynes[147] has also been successfully used in a chemical reporter strategy to evaluate glycoprotein fucosylation.[359] In this strategy, azide-containing fucose was displayed on the surface of Jurkat cells following uptake of modified peracetylated 6-azido-fucose through natural salvage pathways. However, it was noted that the 6-azido fucose analogue in concentrations > 100 μM resulted in cell death (2- and 4-azido fucosyl derivatives studied showed no cellular uptake). More recently, use of a ring-strain-promoted [3 + 2] coupling,[356,360] thereby removing the constraint of standard Cu(I) catalysis, has been proposed.[361] An essentially similar strategy has also been reported by Wong and co-workers.[362] This strategy has been applied to live cells by both Bertozzi[363] and Boons[364] and to map development in zebrafish by Bertozzi.[365] The relative merits and compatibilities of the Staudinger ligation, Cu(I)-catalyzed addition and strain-promoted [3 + 2] cycloaddition as azide-reactive strategies have recently been reviewed.[366]

Salvage pathways have also been exploited to identify *O*-GlcNAc-modified proteins from either living cells or proteins in vitro.[367] It was shown that the UDP-GlcNAc: polypeptidtyltransferase and *O*-GlcNAcase, the enzymes responsible for GlcNAcylation, were able to tolerate GlcNAc residues modified with an azide on the *N*-acyl side chain (GlcNAz). Treatment of Jurkat cells with Ac_4GlcNAz resulted in incorporation of the azido analogue into both nuclear and cytoplasmic proteins, which could be chemically highlighted via Staudinger ligation with FLAG phosphine. Detection of the FLAG phosphine-modified proteins was achieved by Western blot with mouse anti-FLAG-mAB conjugated to HRP. This method was further validated in the labeling of nuclear pore protein p62.

5.2. Biosynthesis Regulation

Regulation of the enzymes involved in the biosynthesis[316] of glycoproteins offers an alternative way of controlling glycan structure with the advantages that small molecule inhibition offers dynamic and tuneable control. For example,

Scheme 28. Cell Surface Modification via Salvage Pathways Allowed in Vivo Ligation[356]



Kohler et al. have employed the chemical inducer dimerization (CID) principle to modulate fucosylation through the noncovalent heterodimerization of protein subcomponents.[368] This technique relies upon the ability of the natural product rapamycin to heterodimerize the proteins FKBP and FRB,[369] which themselves can be fused to virtually any protein. The localization domain and catalytic domain of fucosyltransferase VII (FucT7) were fused to FKBP and FRB, respectively. Localization of the catalytic domain of FucT7 in the Golgi and therefore fucosylation of its cellular substrate was only observed in the presence of rapamycin. Thus, rapamycin was used to regulate production of sLe^x in living cells. In addition, CID has also been used to regulate α1,2-fucosyltransferase and therefore control competition between it and a 1,3-galactosyltransferase.[370] Use of CID has also been applied to carbohydrate sulfotransferases.[371]

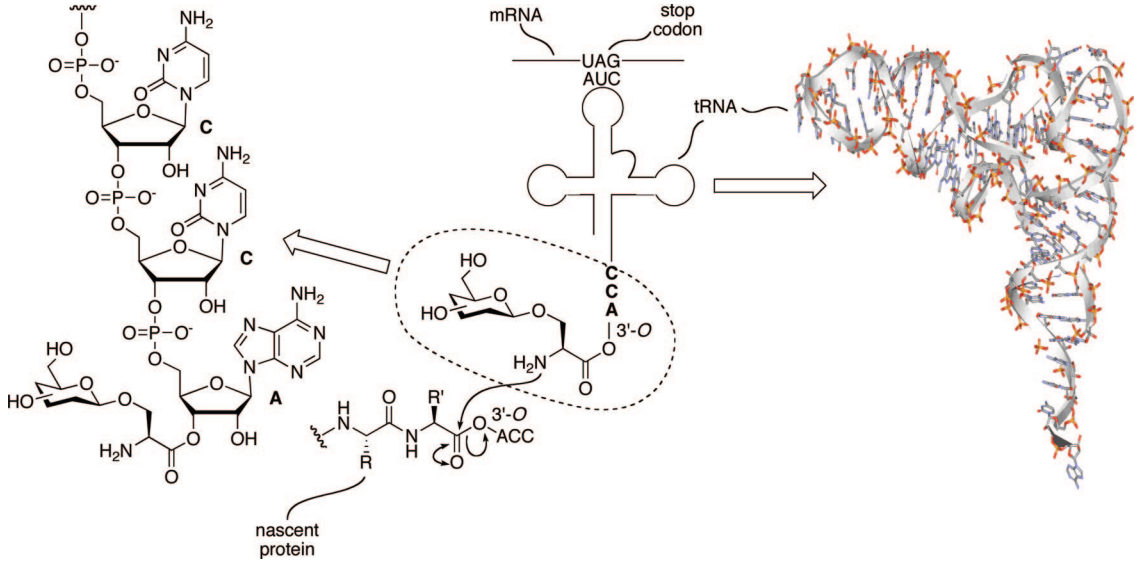
5.3. Codon Suppression Technology

Normal recombinant DNA-based expression is limited to the standard 20 amino acids; however, the ready use of nonsense codon suppression technology has allowed incorporation of novel building blocks through the natural mechanism of protein biosynthesis. Both Hecht[372−374] and Schmidt[375] proposed adoption of the in vitro use of misacylated tRNAs in nonsense codon suppression read-through techniques.[278] Unlike natural protein glycosylation, which occurs co- and post-translationally, this method can utilize, for example, ^{AUC}tRNA acylated with glycosylated amino acids (Scheme 29). A synthesis by Hecht and co-workers of tRNAs acylated with Gal,Glc,Man,GlcNAc-Ser has been reported, albeit in very low overall yields due to difficulties in achieving efficient acylation and deprotection steps.[376] These synthetic problems have recently been solved by Schmidt and co-workers in efficient syntheses of dinucleotides aminoacylated with protected glycosyl amino acids.[377] Such misacylated tRNAs have enabled incorporation of GlcβSer in place of Ser[286] of firefly luciferase[372,373] and allowed Schmidt and co-workers to exchange Lys with GlcNAc(Ac)$_3$-α-Ser in hARF-protein.[375] More recently, Hecht and co-workers extended the use of misacylated tRNAs to incorporate mono- and disaccharides via serine and tyrosine residues into firefly luciferase and DHFR.[378] The effects of glycosylation at position 284 in firefly luciferase on the wavelengths of light emitted are also reported, where red shifts were only observed when the carbohydrate was fully deprotected. Although they have low suppression efficiency (9−48%), these glycosylated amino acids, including the T_n-antigen, have been incorporated into human granulocyte-colony stimulating factor through noncoded read-through techniques. Impressively, this approach has been taken a step further by Schultz in the development of in vivo suppressor technology that works in an intact bacterial system. Here evolution of an orthogonal *Methanaococcus jannachii* Tyr-tRNA synthetase and Tyr-tRNA pair allowed for incorporation, in vivo, of the key glycosylamino acids, β-GlcNAc serine and α-GalNAc threonine, into myoglobin.[379,380] Although these procedures are still in their infancy and suppression levels are still low, they offer the scientific community a potentially universal tool for synthesis of homogeneous glycoproteins. Once a common method, it will be interesting to see if glycoproteins made via such 'pretranslational glycosylation' will vary from natural glycoproteins, which are co/post-translationally glycosylated, especially given the apparent role of glycosylation in correct protein folding. These read-through techniques have also allowed introduction of unnatural amino acids as tags for subsequent site-selective glycosylation (see section 3.3).[279]

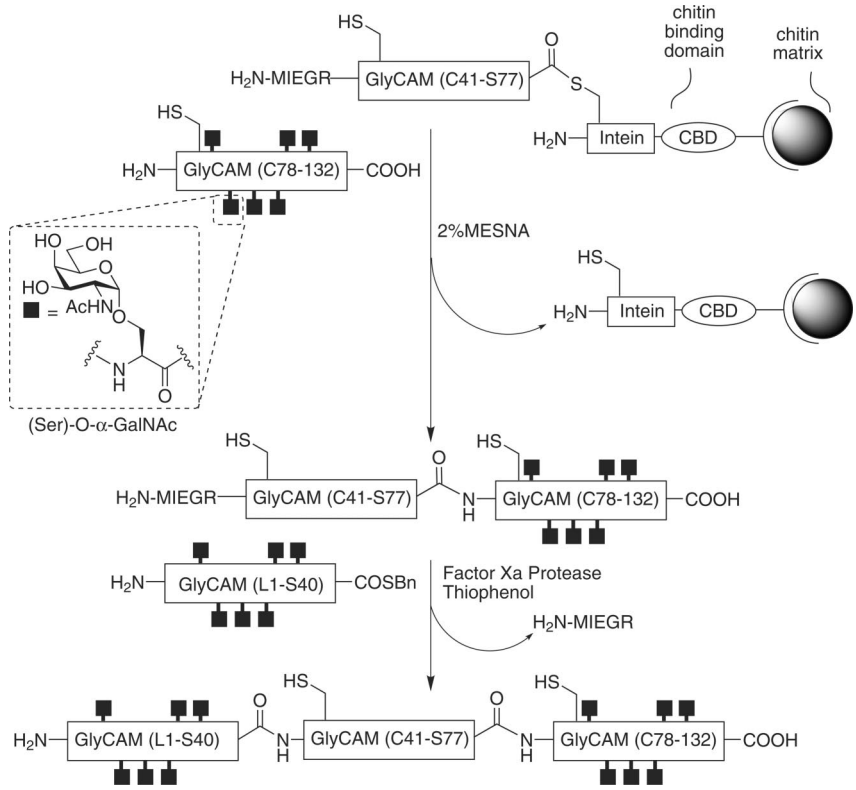
5.4. Expressed Protein Ligation (EPL)

Expressed protein ligation has been used to incorporate variously modified cysteines at the *C*- or *N*-terminus of bacterially expressed peptides.[183,381,382] This can allow access

Scheme 29. Codon Suppression Strategy for Protein Glycosylation Using Ribosomal Incorporation of Glycosylated Amino Acids



Scheme 30. Use of EPL in the Modular Synthesis of GlyCAM-1[387]



to the peptide thioester components needed for NCL (sections 2.2.4 and 3.4) through expression hosts. In this way, EPL can allow ready access, for example, to larger protein scaffolds in a more economic manner than standard SPPS methods. For example, mannan-binding protein (MBP) was expressed in *E. coli* as a fusion to the *N*-terminus of a widely used intein from the mutated *Saccharomyces cerevisiae* VMA gene; a chitin-binding domain was also incorporated at the *C*-terminus for ready purification. Once expressed, this portion self-spliced the binding domain, and the resulting peptidothioester was used in NCL (see section 2.2.4) with glycosylated peptides such as Cys-Asn(GlcNAcβ).[383] Imperali and co-workers have reported a semisynthetic route toward the immunity protein Im7. Initially, a glycopeptide bearing an *N*-terminal Cys, obtained from standard SPPS, was coupled to a thioester-containing fragment via NCL, obtained from recombinant methods using the IMPACT system, a technique based on a pH-dependent intein cleavage.[384]

The acyl acceptor components for NCL can also be accessed via expression: Wong and co-workers employed tobacco etch virus NIa protease (TEV protease), a highly selective cysteine protease, which recognizes and cleaves six residues from the *N*-terminal cysteine.[385] This protease system was successfully used in the ligation of a truncated section of human interlukin-2, a T-cell growth factor that is used in the treatment of renal cell carcinoma, and a peptide segment of gp41 (C37−H6). Fusion constructs that contained

TEV cleavable linkers were constructed for both a truncated form of interleukin-2 and the gp41 segment and readily expressed in *E. coli*. TEV protease cleavage of both fusion proteins resulted in an *N*-terminal cysteine peptides that were subsequently coupled to synthetic thioester glycopeptides via NCL. *N*-Terminal cysteine-containing peptides have also been cleanly produced through cyanogen bromide-mediated cleavage at methionine sites.[386] Production of *N*-terminal cysteines is also possible using factor Xa, which also exploits a protease-catalyzed cleavage via recognition of a Ile-Glu-Gly-Arg (IEGR) cleavage site.[387]

Macmillan and co-workers have used EPL to construct three well-defined model GlyCAM-1 glycoproteins in the first reported modular syntheses of biologically relevant glycoproteins.[388] The mucin-like GlyCAM-1 glycoprotein serves as a ligand during leukocyte homing and is comprised of two mucin domains separated by a central unglycosylated core domain. Semisynthetic variants of GlyCAM-1 were obtained displaying the glycosylated *N*-terminal domain only (1), the glycosylated *C*-terminal domain only (2), and glycosylation in both domains (3). The *N*-terminal glycosylated-only variant (1) was obtained via NCL between a glycosylated thioester peptide and GlyCAM-1 41−132, bearing an *N*-terminal cysteine expressed as an intein-chitin binding domain fusion protein. The latter was purified on chitin beads and was subsequently cleaved from its *C*-terminal intein-chitin binding domain using the factor Xa protease. The *C*-terminal variant (2) required an inverse procedure to (1) with a bacterially derived thioester (GlyCAM-1 1−77)[388] and a synthetic *N*-terminal cysteine glycopeptide (78−132), bearing an *N*-terminal cysteine constructed from both SPPS and NCL. The final *N*- and *C*-glycosylated variant (3) was constructed from the same glycopeptide 78−132 and ligated to a bacterially expressed central core unit (C41−S77) using 2% MESNA. The resulting intermediate was subsequently ligated in a similar fashion to that of variant (1) to produce variant (3) with an impressive presentation of 13 *N*-acetylglucosamines in predetermined positions (Scheme 30).[387]

6. References

(1) Davis, B. G. *Chem. Rev.* **2002**, *102*, 579.
(2) Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357.
(3) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683.
(4) Rudd, P. M.; Elliot, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A. *Science* **2001**, *291*, 2370.
(5) Talbot, P.; Shur, B. D.; Myles, D. G. *Biol. Reprod.* **2003**, *68*, 1.
(6) Varki, A. *Glycobiology* **1993**, *3*, 97.
(7) Helenius, A. *Mol. Biol. Cell* **1994**, *5*, 253.
(8) Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364.
(9) Trombetta, E. S.; Helenius, A. *Curr. Opin. Struct Biol.* **1998**, *8*, 587.
(10) Hebert, D. N.; Garman, S. C.; Molinari, M. *Trends Cell. Biol.* **2005**, *15*, 364.
(11) Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364.
(12) Park, H.; Suzuki, T.; Lennarz, W. J. *Proc. Natl. Acad. Sci., U.S.A.* **2001**, *98*, 11163.
(13) Imperiali, B.; O'Connor, S. E. *Curr. Opin. Chem. Biol.* **1999**, *3*, 643.
(14) Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Arulanandam, A. R.; Smolyar, A.; Reinherz, E. L.; Wagner, G. *Science* **1995**, *269*, 1273.
(15) Opdenakker, G.; Rudd, P. M.; Ponting, C. P.; Dwek, R. A. *FASEB J.* **1993**, *7*, 1330.
(16) Arnold, U.; Ulbrich-Hofmann, R. *Biochemistry* **1997**, *36*, 2166.
(17) Arnold, U.; Schierhorn, A.; Ulbrich-Hofmann, R. *Eur. J. Biochem.* **1999**, *259*, 470.
(18) Rudd, P. M.; Joao, H. C.; Coghill, E.; Fiten, P.; Saunders, M. R.; Opdenakker, G.; Dwek, R. A. *Biochemistry* **1994**, *33*, 17.
(19) Hansen, T. N.; Carpenter, J. F. *Biophys. J.* **1993**, *64*, 1843.
(20) Matsumoto, S.; Matsusita, M.; Morita, T.; Kamachi, H.; Tsukiyama, S.; Furukawa, Y.; Koshida, S.; Tachibana, Y.; Nishimura, S.-I.; Todo, S. *Cryobiology* **2006**, *52*, 90.
(21) Tachibana, Y.; Fletcher, G. L.; Fujitani, N.; Tsuda, S.; Monde, K.; Nishimura, S.-I. *Angew. Chem., Int. Ed.* **2004**, *43*, 856.
(22) Davis, B. G. *Science* **2004**, *303*, 480.
(23) van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G. *Nature* **2007**, *446*, 1105.
(24) Schreiber, S. L. *Nat. Chem. Biol.* **2007**, *3*, 352.
(25) Laine, R. A. *Glycobiology* **1994**, *4*, 759.
(26) Grogan, M. J.; Pratt, M. R.; Marcaurelle, L. A.; Bertozzi, C. R. *Annu. Rev. Biochem.* **2002**, *71*, 593.
(27) Brooks, S. A.; Dwek, M. V.; Schumacher, U. *Functional and Molecular Glycobiology*; Garland Science: 2002.
(28) Gabius, H.-J.; Siebert, H.-C.; Andre, S.; Jimenez-Barbero, J.; Rudiger, H. *ChemBioChem* **2004**, *5*, 740.
(29) Wong, S. Y. C. *Curr. Opin. Struct. Biol.* **1995**, *5*, 599.
(30) Yarema, K. J.; Bertozzi, C. R. *Curr. Opin. Chem. Biol.* **1998**, *2*, 49.
(31) Davis, B. G. *J. Chem. Soc., Perkin Trans. 1* **1999**, 3215.
(32) Lloyd, R. C.; Davis, B. G.; Jones, J. B. *Bioorg. Med. Chem.* **2000**, *8*, 1537.
(33) Doores, K. J.; Gamblin, D. P.; Davis, B. G. *Chem. Eur. J.* **2006**, *12*, 656.
(34) Aebi, M.; Hennet, T. *Trends Cell. Biol.* **2001**, *11*, 136.
(35) Ashwell, G.; Harford, J. *Annu. Rev. Biochem.* **1982**, *51*, 531.
(36) Wadhwa, M. S.; Rice, K. G. *J. Drug Targeting* **1995**, *3*, 111.
(37) Hassan, N.; Wu, G. Y. *Nat. Med.* **1995**, *1*, 210.
(38) Wu, G. Y.; Wu, C. H. *Adv. Drug Deliv. Rev.* **1993**, *12*, 159.
(39) Meijer, D. K. F.; Molema, G.; Moolenaar, F.; deZeeuw, D.; Swart, P. J. *J. Controlled Release* **1996**, *39*, 163.
(40) Franssen, E. J. F.; Jansen, R. W.; Vaalburg, M.; Meijer, D. K. F. *Biochem. Pharmacol.* **1993**, *45*, 1215.
(41) Robinson, M. A.; Charlton, S. T.; Garnier, P.; Wang, X.-T.; Davis, S. S.; Perkins, A. C.; Frier, M.; Duncan, R.; Savage, T. J.; Wyatt, D. A.; Watson, S. A.; Davis, B. G. *Proc. Natl. Acad. Sci., U.S.A.* **2004**, *101*, 14527.
(42) Lee, Y. C.; Stowell, C. P.; Krantz, M. J. *Biochemistry* **1976**, *15*, 3956.
(43) Beck, M. *Hum. Genet.* **2007**, *121*, 1.
(44) Zhu, Y.; Li, X.; Kyazike, J.; Zhou, Q.; Thurberg, B. L.; Raben, N.; Mattaliano, R. J.; Cheng, S. H. *J. Biol. Chem.* **2004**, *279*, 50336.
(45) Verez-Bencomo, V.; Fernandez-Santana, V.; Hardy, E.; Toledo, M. E.; Rodriguez, M. C.; Heynngnezz, L.; Rodriguez, A.; Baly, A.; Herrera, L.; Izquierdo, M.; Villar, A.; Valdes, Y.; Cosme, K.; Deler, M. L.; Montane, M.; Garcia, E.; Ramos, A.; Aguilar, A.; Medina, E.; Torano, G.; Sosa, I.; Hernandez, I.; Martinez, R.; Muzachio, A.; Carmenates, A.; Costa, L.; Cardoso, F.; Campa, C.; Diaz, M.; Roy, R. *Science* **2004**, *305*, 522.
(46) Schofield, L.; Hewitt, M. C.; Evans, K.; Siomos, M.-A.; Seeberger, P. H. *Nature* **2002**, *418*, 785.
(47) Hewitt, M. C.; Snyder, D. A.; Seeberger, P. H. *J. Am. Chem. Soc.* **2002**, *124*, 13434.
(48) Sato, M.; Sadamoto, R.; Niikura, K.; Monde, K.; Kondo, H.; Nishimura, S.-I. *Angew. Chem., Int. Ed.* **2004**, *43*, 1516.
(49) Markussen, J.; Diers, I.; Engesgaard, A.; Hansen, M. T.; Hougaard, P.; Langkjaer, L.; Norris, K.; Ribel, U.; Soerensen, A. R.; et al. *Protein Eng.* **1987**, *1*, 215.
(50) Markussen, J.; Diers, I.; Hougaard, P.; Langkjaer, L.; Norris, K.; Snel, L.; Soerensen, A. R.; Soerensen, E.; Voigt, H. O. *Protein Eng.* **1988**, *2*, 157.
(51) Egrie, J. C.; Browne, J. K. *Br. J. Cancer* **2001**, *84*, 3.
(52) Egrie, J. C.; Dwyer, E.; Browne, J. K.; Hitz, A.; Lykos, M. A. *Exp. Hematol.* **2003**, *31*, 290.
(53) Davis, B. G. *Chem. Commun.* **2001**, 351.
(54) Johansson, S. M. C.; Arnberg, N.; Elofsson, M.; Wadell, G.; Kihlberg, J. *ChemBioChem* **2005**, *6*, 358.
(55) Watt, G. M.; Lund, J.; Levens, M.; Kolli, V. S. K.; Jefferis, R.; Boons, G.-J. *Chem. Biol.* **2003**, *10*, 807.
(56) Rademacher, T. W.; Parekh, R. B.; Dwek, R. A. *Annu. Rev. Biochem.* **1988**, *57*, 785.
(57) Bill, R. M.; Flitsch, S. L. *Chem. Biol.* **1996**, *3*, 145.
(58) Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong, C. H. *Chem. Rev.* **1998**, *98*, 833.
(59) Buskas, T.; Ingale, S.; Boons, G.-J. *Glycobiology* **2006**, *16*, 113R.
(60) Evans, P. G.; Gemmell, N.; Osborn, H. M. I. *Carbohydrates*; Academic Press: New York, 2003.
(61) Hojo, H.; Nakahara, Y. *Biopolymers* **2007**, *88*, 308.
(62) Nakahara, Y. *Trends Glycosci. Glycotechnol.* **2003**, *15*, 257.
(63) Specker, D.; Wittmann, V. *Top. Curr. Chem.* **2007**, *267*, 65.
(64) Demchenko, A. V. *Lett. Org. Chem.* **2005**, *2*, 580.
(65) Fairbanks, A. J. *Synlett* **2003**, 1945.
(66) Jahn, M.; Withers, S. G. *Biocatal. Biotrans.* **2003**, *21*, 159.

(67) Kanemitsu, T.; Kanie, O. *Combi. Chem. High Throughput Screening* **2002**, *5*, 339.
(68) Karst, N. A.; Linhardt, R. J. *Curr. Med. Chem.* **2003**, *10*, 1993.
(69) Perugino, G.; Cobucci-Ponzano, B.; Rossi, M.; Moracci, M. *Adv. Synth. Catal.* **2005**, *347*, 941.
(70) Seeberger, P. H. *Chem. Commun.* **2003**, 1115.
(71) Seibel, J.; Joerdening, H.-J.; Buchholz, K. *Biocatal. Biotrans.* **2006**, *24*, 311.
(72) Davis, B. G. *Chem. Rev.* **2000**, 2137.
(73) Muir, T. W.; Dawson, P. E.; Kent, S. B. H. *Methods Enzymol.* **1997**, *289*, 266.
(74) Wilken, J.; Kent, S. B. H. *Curr. Opin. Biotechnol.* **1998**, *9*, 412.
(75) Kochendoerfer, G. G.; Kent, S. B. H. *Curr. Opin. Chem. Biol.* **1999**, *3*, 665.
(76) Robertson, N.; Ramage, R. *J. Chem. Soc., Perkin Trans. 1* **1999**, 1015.
(77) Borgia, J. A.; Fields, G. B. *TIBTECH* **2000**, *18*, 243.
(78) Brik, A.; Ficht, S.; Wong, C.-H. *Curr. Opin. Chem. Biol.* **2006**, *10*, 638.
(79) Grogan, M. J.; Pratt, M. R.; Marcaurelle, L. A.; Bertozzi, C. R. *Annu. Rev. Biochem.* **2002**, *71*, 593.
(80) Guo, Z.; Shao, N. *Med. Res. Rev.* **2005**, *25*, 655.
(81) Khmelnitsky, Y. L. *J. Mol. Catal. B* **2004**, *31*, 73.
(82) Liu, L.; Bennett, C. S.; Wong, C.-H. *Chem. Commun.* **2006**, 21.
(83) Marcaurelle, L. A.; Bertozzi, C. R. *Glycobiology* **2002**, *12*, 69R.
(84) Pratt, M. R.; Bertozzi, C. R. *Chem. Soc. Rev.* **2005**, *34*, 58.
(85) Bennett, C. S.; Wong, C.-H. *Chem. Soc. Rev.* **2007**, *36*, 1227.
(86) Arsequell, G.; Valencia, G. *Tetrahedron: Asymmetry* **1997**, *8*, 2839.
(87) Arsequell, G.; Valencia, G. *Tetrahedron: Asymmetry* **1999**, *10*, 3045.
(88) Kunz, H. *Angew. Chem., Int. Ed Engl.* **1987**, *26*, 294.
(89) Paulsen, H. *Angew. Chem., Int. Ed Engl.* **1990**, *102*, 851.
(90) Meldal, M. *Curr. Opin. Struct. Biol.* **1994**, *4*, 710.
(91) Meldal, M.; Bock, K. *Glycoconjugate J.* **1994**, *11*, 59.
(92) Garg, H. G.; Bruch, K. v. d.; Kunz, H. *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 277.
(93) Meldal, M.; St. Hilaire, P. M. *Curr. Opin. Chem. Biol.* **1997**, *1*, 552.
(94) Taylor, C. M. *Tetrahedron* **1998**, *54*, 11317.
(95) Osborn, H. M. I.; Khan, T. H. *Tetrahedron* **1999**, *55*, 1807.
(96) Herzner, H.; Reipen, T.; Schultz, M.; Kunz, H. *Chem. Rev.* **2000**, *100*, 4495.
(97) Haneda, K.; Inazu, T.; Mizuno, M.; Iguchi, R.; Tanabe, H.; Fujimori, K.; Yamamoto, K.; Kumagai, H.; Tsumori, K.; Munekata, E. *Biochim. Biophys. Acta* **2001**, *1526*, 242.
(98) Pachamuthu, K.; Schmidt, R. R. *Chem. Rev.* **2006**, *106*, 160.
(99) Vliegenthart, J. F. G.; Casset, F. *Curr. Opin. Struct. Biol* **1998**, *8*, 565.
(100) Shibata, S.; Takeda, T.; Natori, Y. *J. Biol. Chem.* **1988**, *263*, 12483.
(101) Wieland, F.; Heizer, R.; Schaefer, W. *Proc. Natl. Acad. Sci., U.S.A.* **1983**, *80*, 5470.
(102) Paul, G.; Lottspeich, F.; Wieland, F. *J. Biol. Chem.* **1986**, *261*, 1020.
(103) Messner, P.; Sleytr, U. B. *FEBS Lett.* **1988**, *228*, 317.
(104) Strahl-Bolsinger, S. S.; Gentzsch, M. M.; Tanner, W. W. *Biochim. Biophys. Acta* **1999**, *1426*, 297.
(105) DeBeer, T.; Vliegenthart, J. F. G.; Löffler, A.; Hofsteenge, J. *Biochemistry* **1995**, *34*, 11785.
(106) Deras, I. L.; Takegawa, K.; Kondo, A.; Kato, I.; Lee, Y. C. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1763.
(107) Inazu, T.; Kobayashi, K. *Synlett* **1993**, 869.
(108) Mizuno, M.; Muramoto, I.; Kobayashi, K.; Yaginuma, H.; Inazu, T. *Synthesis* **1999**, 162.
(109) He, Y.; Hinklin, R. J.; Chang, J.; Kiessling, L. L. *Org. Lett.* **2004**, *6*, 4479.
(110) Doores, K. J.; Mimura, Y.; Dwek, R. A.; Rudd, P. M.; Elliott, T.; Davis, B. G. *Chem. Commun.* **2006**, 1401.
(111) Vizvardi, K.; Kreytz, C.; Davis, A. S.; Lee, V. P.; Philmus, B. J.; Simo, O.; Michael, K. *Chem. Lett.* **2003**, *32*, 348.
(112) Kaneshiro, C. M.; Michael, K. *Angew. Chem., Int. Ed.* **2006**, *45*, 1077.
(113) Tsuda, T.; Nishimura, S.-I. *Chem. Commun.* **1996**, 2779.
(114) Shimawaki, K.; Fujisawa, Y.; Fumihiro, S.; Fujitani, N.; Masaki, K.; Hiroko, H.; Hiroshio, H.; Shin-Ichiro, N. *Angew. Chem., Int. Ed.* **2007**, *46*, 3047.
(115) Mukhopadhyay, B.; Maurer, S. V.; Rudolph, N.; van Well, R. M.; Russell, D. A.; Field, R. A. *J. Org. Chem.* **2005**, *70*, 9059.
(116) Cato, D.; Buskas, T.; Boons, G.-J. *J. Carbohydr. Chem.* **2005**, 503.
(117) Svarovsky, S. A.; Barchi, J. J. *Carbohydr. Res.* **2003**, *338*, 1925.
(118) Manabe, S.; Marui, Y.; Ito, Y. *Chem. Eur. J.* **2003**, *9*, 1435.
(119) Baldwin, J. E.; Spivey, A. C.; Schofield, C. J. *Tetrahedron: Asymmetry* **1990**, *1*, 881.
(120) Cohen, S. B.; Halcomb, R. L. *Org. Lett.* **2001**, *3*, 405.
(121) Zhu, Y.; vanderDonk, W. A. *Org. Lett.* **2001**, *3*, 1189.
(122) Galonic, D. P.; van der Donk, W. A.; Gin, D. Y. *Chem. Eur. J.* **2003**, *9*, 5997.
(123) Seebeck, F. P.; Szostak, J. W. *J. Am. Chem. Soc.* **2006**, *128*, 7150.
(124) Wang, J.; Schiller, S. M.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2007**, *46*, 6849.
(125) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 5052.
(126) Galonic, D. P.; Ide, N. D.; van der Donk, W. A.; Gin, D. Y. *J. Am. Chem. Soc.* **2005**, *127*, 7359.
(127) Galonic, D. P.; van der Donk, W. A.; Gin, D. Y. *J. Am. Chem. Soc.* **2004**, *126*, 12712.
(128) Jobron, L.; Hummel, G. *Org. Lett.* **2000**, *2*, 2265.
(129) Ohnishi, Y.; Ichikawa, M.; Ichikawa, Y. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1289.
(130) Zhu, X.; Schmidt, R. R. *Chem. Eur. J.* **2004**, *10*, 875.
(131) Zhu, X.; Schmidt, R. R. *Tetrahedron Lett.* **2003**, *44*, 6063.
(132) Thayer, D. A.; Yu, H. N.; Galan, M. C.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2005**, *44*, 4596.
(133) Zhu, X.; Pachamuthu, K.; Schmidt, R. R. *J. Org. Chem.* **2003**, *68*, 5641.
(134) Zhu, X.; Haag, T.; Schmidt, R. R. *Org. Biomol.Chem.* **2004**, *2*, 31.
(135) Whalen, L. J.; Halcomb, R. L. *Org. Lett.* **2004**, *6*, 3221.
(136) Zhu, X.; Pachamuthu, K.; Schmidt, R. R. *Org. Lett.* **2004**, *6*, 1083.
(137) Crich, D.; Zou, Y.; Brebion, F. *J. Org. Chem.* **2006**, *71*, 9172.
(138) Dondoni, A.; Massi, A.; Aldhoun, M. *J. Org. Chem.* **2007**, *72*, 7677.
(139) Ichikawa, Y.; Nishiyama, T.; Isobe, M. *Synlett* **2000**, 1253.
(140) Ichikawa, Y.; Ohara, F.; Kotsuki, H.; Nakano, K. *Org. Lett.* **2006**, *8*, 5009.
(141) Ichikawa, Y.; Matsukawa, Y.; Isobe, M. *J. Am. Chem. Soc.* **2006**, *128*, 3934.
(142) Lemieux, G. A.; Bertozzi, C. R. *TIBTECH* **1998**, *16*, 506.
(143) Zhao, Y.; Kent, S. B. H.; Chait, B. T. *Proc. Natl. Acad. Sci., U.S.A.* **1997**, *94*, 1629.
(144) Durieux, P.; Fernandez-Carneado, J.; Tuchscherer, G. *Tetrahedron Lett.* **2001**, *42*, 2297.
(145) Peluso, S.; Imperiali, B. *Tetrahedron Lett.* **2001**, *42*, 2085.
(146) Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057.
(147) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2596.
(148) Groothuys, S.; Kuijpers, B. H. M.; Quaedflieg, P. J. L. M.; Roelen, H. C. P. F.; Wiertz, R. W.; Blaauw, R. H.; van Delft, F. L.; Rutjes, F. P. J. T. *Synthesis* **2006**, 3146.
(149) Kuijpers, B. H. M.; Groothuys, S.; Keereweer, A. R.; Quaedflieg, P. J. L. M.; Blaauw, R. H.; van Delft, F. L.; Rutjes, F. P. J. T. *Org. Lett.* **2004**, *6*, 3123.
(150) Macmillan, D.; Blanc, J. *Org. Biomol. Chem.* **2006**, *4*, 2847.
(151) Wan, Q.; Chen, J.; Chen, G.; Danishefsky, S. J. *J. Org. Chem.* **2006**, *71*, 8244.
(152) Lin, H.; Walsh, C. T. *J. Am. Chem. Soc.* **2004**, *126*, 13998.
(153) Fu, X.; Albermann, C.; Zhang, C.; Thorson, J. S. *Org. Lett.* **2005**, *7*, 1513.
(154) Fu, X.; Albermann, C.; Jiang, J.; Liao, J.; Zhang, C.; Thorson, J. S. *Nat. Biotechnol.* **2003**, *21*, 1467.
(155) Wang, J.; Li, H.; Zou, G.; Wang, L.-X. *Org. Biomol.Chem.* **2007**, *5*, 1529.
(156) Pietrzik, N.; Schips, C.; Ziegler, T. *Synthesis* **2008**, 519.
(157) McGarvey, G. J.; Benedum, T. E.; Schmidtmann, F. W. *Org. Lett.* **2002**, *4*, 3591.
(158) Dondoni, A.; Marra, A. *Chem. Rev.* **2000**, *100*, 4395.
(159) Leeuwenburgh, M. A.; van der Marel, G. A.; Overkleeft, H. S. *Curr. Opin. Chem. Biol.* **2003**, *7*, 757.
(160) Dominique, R.; Liu, B.; Das, S. K.; Roy, R. *Synthesis* **2000**, 862.
(161) Biswas, K.; Coltart, D. M.; Danishefsky, S. J. *Tetrahedron Lett.* **2002**, *43*, 6107.
(162) Nolen, E. G.; Kurish, A. J.; Wong, K. A.; Orlando, M. D. *Tetrahedron Lett.* **2003**, *44*, 2449.
(163) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642.
(164) Li, H.; Li, B.; Song, H.; Breydo, L.; Baskakov, I. V.; Wang, L.-X. *J. Org. Chem.* **2005**, *70*, 9990.
(165) Yamamoto, N.; Takayanagi, A.; Yoshino, A.; Sakakibara, T.; Kajihara, Y. *Chem. Eur. J.* **2007**, *13*, 613.
(166) Shao, N.; Xue, J.; Guo, Z. *J. Org. Chem.* **2003**, *68*, 9003.
(167) Baumann, K.; Kowalczyk, D.; Kunz, H. *Angew. Chem., Int. Ed.* **2008**, *47*, 3445.
(168) Liu, L.; Hong, Z.-Y.; Wong, C.-H. *ChemBioChem* **2006**, *7*, 429.
(169) Ansfield, S. T.; Lansbury, P. T. *J. Org. Chem.* **1990**, *55*, 5560.
(170) Cohen-Anisfeld, S. T.; Lansbury, P. T., Jr *J. Am. Chem. Soc.* **1993**, *115*, 10531.
(171) Miller, J. S.; Dudkin, V. Y.; Lyon, G. J.; Muir, T. W.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2003**, *42*, 431.
(172) Dudkin, V. Y.; Miller, J. S.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 736.

(173) Likhosherstov, L. M.; Novikova, O. S.; Derevitskaya, V. A.; Kochetkov, N. K. *Carbohydr. Res.* **1986**, *146*, C1.
(174) Bejugam, M.; Flitsch, S. L. *Org. Lett.* **2004**, *6*, 4001.
(175) Tachibana, Y.; Matsubara, N.; Nakajima, F.; Tsuda, T.; Tsuda, S.; Monde, K.; Nishimura, S.-I. *Tetrahedron* **2002**, *58*, 10213.
(176) Tachibana, Y.; Monde, K.; Nishimura, S.-I. *Macromolecules* **2004**, *37*, 6771.
(177) Bang, D.; Pentelute, B. L.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2006**, *45*, 3985.
(178) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *226*, 776.
(179) Shin, Y.; Winans, K. A.; Backes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1999**, *121*, 11684.
(180) Macmillan, D. *Angew. Chem., Int. Ed.* **2006**, *45*, 7668.
(181) Yeo, D. S. Y.; Srinivasan, R.; Chen, G. Y. J.; Yao, S. Q. *Chem. Eur. J.* **2004**, *10*, 4664.
(182) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923.
(183) Muir, T. W. T. W. *Annu. Rev. Biochem.* **2003**, *72*, 249.
(184) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91.
(185) Macmillian, D. *Angew. Chem., Int. Ed.* **2006**, *45*, 7668.
(186) Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H.; Schafer, W. *Justus Liebigs Ann. Chem.* **1953**, *583*, 129.
(187) Mandal, M.; Dudkin, V. Y.; Geng, X.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2004**, *43*, 2557.
(188) Geng, X.; Dudkin, V. Y.; Mandal, M.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2004**, *43*, 2562.
(189) Hojo, H.; Matsumoto, Y.; Nakahara, Y.; Ito, E.; Suzuki, Y.; Suzuki, M.; Suzuki, A.; Nakahara, Y. *J. Am. Chem. Soc.* **2005**, *127*, 13720.
(190) Yamamoto, N.; Tanabe, Y.; Okamoto, R.; Dawson, P. E.; Kajihara, Y. *J. Am. Chem. Soc.* **2008**, *130*, 501.
(191) Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526.
(192) Pentelute, B. L.; Kent, S. B. H. *Org. Lett.* **2007**, *9*, 687.
(193) Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248.
(194) Crich, D. D.; Banerjee, A. A. *J. Am. Chem. Soc.* **2007**, *129*, 10064.
(195) Offer, J.; Boddy, C. N. C.; Dawson, J. *J. Am. Chem. Soc.* **2002**, *124*, 4642.
(196) Wu, B.; Chen, J.; Warren, D. J.; Chen, G.; Hua, Z.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2006**, *45*, 4116.
(197) Warren, J. D.; Miller, J. S.; Keding, S. J.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 6576.
(198) Mezzato, S.; Schaffrath, M.; Unverzagt, C. *Angew. Chem., Int. Ed.* **2005**, *44*, 1650.
(199) Bernatowicz, M. S.; Daniels, S. B.; Koster, H. *Tetrahedron Lett.* **1989**, *30*, 4645.
(200) Brik, A.; Yang, Y.-Y.; Fricht, S.; Chi-Huey, W. *J. Am. Chem. Soc.* **2006**, *128*, 5626.
(201) Yang, Y.-Y.; Ficht, S.; Brik, A.; Wong, C.-H. *J. Am. Chem. Soc.* **2007**, *129*, 7690.
(202) Brik, A.; Ficht, S.; Yang, Y.-Y.; Bennett, C. S.; Wong, C.-H. *J. Am. Chem. Soc.* **2006**, *128*, 15026.
(203) Ficht, S. S.; Payne, R. J. R. J.; Brik, A. A.; Wong, C. H. C.-H. *Angew. Chem., Int. Ed.* **2007**, *46*, 5975.
(204) Ingale, S.; Buskas, T.; Boons, G. J. *Org. Lett.* **2006**, *8*, 5785.
(205) Ingale, S.; Wolfert, M. A.; Gaekwad, J.; Buskas, T.; Boons, G.-J. *Nat. Chem. Biol.* **2007**, *3*, 663.
(206) Chen, G.; Wan, Q.; Tan, Z.; Kan, C.; Hua, Z.; Ranganathan, K.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 7383.
(207) Payne, R. J.; Ficht, S.; Greenberg, W. A.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2008**, *47*, 4411.
(208) Brik, A.; Ficht, S.; Yang, Y.-Y.; Bennett, C. S.; Wong, C.-H. *J. Am. Chem. Soc.* **2006**, *128*, 15026.
(209) Pearce, O. M. T.; Fisher, K. D.; Humphries, J.; Seymour, L. W.; Smith, A.; Davis, B. G. *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 1057.
(210) Bernstein, M. A.; Hall, L. D. *Carbohydr. Res.* **1980**, *78*, C1.
(211) Pan, Y.; Chefalo, P.; Nagy, N.; Harding, C.; Guo, Z. *J. Med. Chem.* **2005**, *48*, 875.
(212) Chefalo, P.; Pan, Y.; Nagy, N.; Harding, C.; Guo, Z. *Glycoconjugate J.* **2004**, *20*, 407.
(213) Ragupathi, G.; Koganty, R. R.; Qiu, D. X.; Lloyd, K. O.; Livingston, P. O. *Glycoconjugate J.* **1998**, *15*, 217.
(214) Tietze, L. F.; Schroter, C.; Gabius, S.; Brinck, U.; Goerlach-Graw, A.; Gabius, H. J. *Bioconjugate. Chem.* **1991**, *2*, 148.
(215) Kamath, V. P.; Diedrich, P.; Hindsgaul, O. *Glycoconjugate J.* **1996**, *13*, 315.
(216) Saksena, R.; Zhang, J.; Kovac, P. *Tetrahedron: Asymmetry* **2005**, *16*, 187.
(217) Saksena, R.; Ma, X.; Wade, T. K.; Kovac, P.; Wade, W. F. *FEMS Immunol. Med. Microbiol.* **2006**, *47*, 116.
(218) Meeks, M. D.; Saksena, R.; Ma, X.; Wade, T. K.; Taylor, R. K.; Kovac, P.; Wade, W. F. *Infect. Immun.* **2004**, *72*, 4090.
(219) Bundle, D. R.; Rich, J. R.; Jacques, S.; Yu, H. N.; Nitz, M.; Ling, C.-C. *Angew. Chem., Int. Ed.* **2005**, *44*, 7725.
(220) Wu, X. X.; Ling, C. C. C.-C.; Bundle, D. R. D. R. *Org. Lett.* **2004**, *6*, 4407.
(221) Pozsgay, V.; Vieira, N. E.; Yergey, A. *Org. Lett.* **2002**, *4*, 3191.
(222) Berkin, A.; Coxon, B.; Pozsgay, V. *Chem. Eur. J.* **2002**, *8*, 4424.
(223) Kubler-Kielb, J.; Pozsgay, V. *J. Org. Chem.* **2005**, *70*, 6987.
(224) Gupta, S. S.; Kuzelka, J.; Singh, P.; Lewis, W. G.; Manchester, M.; Finn, M. G. *Bioconjugate Chem.* **2005**, *16*, 1572.
(225) Ghosh, S. S.; Kao, P. M.; McCue, A. W.; Chappelle, H. L. *Bioconjugate Chem.* **1990**, *1*, 71.
(226) Rodriguez, E. C.; Winans, K. A.; King, D. S.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1997**, *119*, 9905.
(227) Orr, G. A.; Rando, R. R. *Nature* **1978**, *272*, 722.
(228) Shimaoka, H.; Kuramoto, H.; Furukawa, J.-i.; Miura, Y.; Kurogochi, M.; Kita, Y.; Hinou, H.; Shinohara, Y.; Nishimura, S.-I. *Chem. Eur. J.* **2007**, *13*, 1664.
(229) Grigalevicius, S.; Chierici, S.; Renaudet, O.; Lo-Man, R.; Deriaud, E.; Leclerc, C.; Dumy, P. *Bioconjugate Chem.* **2005**, *16*, 1149.
(230) Cervigni, S. E.; Dumy, P.; Mutter, M. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1230.
(231) Marcaurelle, L. A.; Rodriguez, E. C.; Bertozzi, C. R. *Tetrahedron Lett.* **1998**, *39*, 8417.
(232) Davis, N. J.; Flitsch, S. L. *Tetrahedron Lett.* **1991**, *32*, 6793.
(233) Marcaurelle, L. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2001**, *123*, 1587.
(234) Pratt, M. R.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2003**, *125*, 6149.
(235) Wang, L.-X.; Tang, M.; Suzuki, T.; Kitajima, K.; Inoue, Y.; Inoue, S.; Fan, J.-Q.; Lee, Y. C. *J. Am. Chem. Soc.* **1997**, *119*, 11137.
(236) Wang, Z.-W.; Zhang, X.; Visser, M.; Live, D.; Zatorski, A.; Iserloh, U.; Lloyd, K. O.; Danishefsky, S. J. *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 1728.
(237) Wang, L.-X.; Fan, J.-Q.; Lee, Y. C. *Tetrahedron Lett.* **1996**, *37*, 1975.
(238) Li, H.; Singh, S.; Zeng, Y.; Song, H.; Wang, L.-X. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 895.
(239) Ni, J.; Powell, R.; Baskakov, I. V.; DeVico, A.; Lewis, G. K.; Wang, L.-X. *Bioorg. Med. Chem.* **2004**, *12*, 3141.
(240) Singh, S.; Ni, J.; Wang, L.-X. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 327.
(241) Wang, L.-X.; Song, H.; Liu, S.; Lu, H.; Jiang, S.; Ni, J.; Li, H. *ChemBioChem* **2005**, *6*, 1068.
(242) Macmillan, D.; Daines, A. M.; Bayrhuber, M.; Flitsch, S. L. *Org. Lett.* **2002**, *4*, 1467.
(243) Wong, S. Y.; Guile, G. R.; Dwek, R. A.; Arsequell, G. *Biochem. J.* **1994**, *300* (Pt 3), 843.
(244) Yamamoto, N.; Sakakibara, T.; Kajihara, Y. *Tetrahedron Lett.* **2004**, *45*, 3287.
(245) Ito, Y.; Hagihara, S.; Matsuo, I.; Totani, K. *Curr. Opin. Struct Biol.* **2005**, *15*, 481.
(246) Suzuki, T.; Hara, I.; Nakano, M.; Zhao, G.; Lennarz, W. J.; Schindelin, H.; Taniguchi, N.; Totani, K.; Matsuo, I.; Ito, Y. *J. Biol. Chem.* **2006**, *281*, 22152.
(247) Macindoe, W. M.; vanOijen, A. H.; Boons, G.-J. *Chem. Commun.* **1998**, 847.
(248) Watt, G. M.; Boons, G.-J. *Carbohydr. Res.* **2004**, *339*, 181.
(249) Shin, I.; Jung, H.-J.; Lee, M.-R. *Tetrahedron Lett.* **2001**, *42*, 1325.
(250) Ni, J.; Singh, S.; Wang, L.-X. *Bioconjugate Chem.* **2003**, *14*, 232.
(251) Ni, J.; Song, H.; Wang, Y.; Stamatos, N. M.; Wang, L.-X. *Bioconjugate Chem.* **2006**, *17*, 493.
(252) Calarese, D. A.; Scanlan, C. N.; Zwick, M. B.; Deechongkit, S.; Mimura, Y.; Kunert, R.; Zhu, P.; Wormald, M. R.; Stanfield, R. L.; Roux, K. H.; Kelly, J. W.; Rudd, P. M.; Dwek, R. A.; Katinger, H.; Burton, D. R.; Wilson, I. A. *Science* **2003**, *300*, 2065.
(253) Totani, K.; Matsuo, I.; Ihara, Y.; Ito, Y. *Bioorg. Med. Chem.* **2006**, *14*, 5220.
(254) Totani, K.; Ito, Y. *Trends Glycosci. Glycotechnol.* **2005**, *17*, 121.
(255) Totani, K.; Matsuo, I.; Ito, Y. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2285.
(256) Lin, P.-C.; Ueng, S.-H.; Tseng, M.-C.; Ko, J.-L.; Huang, K.-T.; Yu, S.-C.; Adak, A. K.; Chen, Y.-J.; Lin, C.-c. *Angew. Chem., Int. Ed.* **2006**, *45*, 4286.
(257) Davis, B. G.; Maughan, M. A. T.; Green, M. P.; Ullman, A. *Tetrahedron: Asymmetry* **2000**, *11*, 245.
(258) Davis, B. G.; Lloyd, R. C.; Jones, J. B. *J. Org. Chem.* **1998**, *63*, 9614.
(259) Davis, B. G.; Lloyd, R. C.; Jones, J. B. *J. Org. Chem.* **1998**, *63*, 9614.
(260) Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. *Org. Biomol. Chem.* **2003**, *1*, 3642.
(261) Prante, O.; Einsiedel, J.; Haubner, R.; Gmeiner, P.; Wester, H.-J.; Kuwert, T.; Maschauer, S. *Bioconjugate Chem.* **2007**, *18*, 254.
(262) Davis, B. G.; Maughan, M. A. T.; Green, M. P.; Ullman, A.; Jones, J. B. *Tetrahedron: Asymmetry* **2000**, *11*, 245.
(263) Matsumoto, K.; Davis, B. G.; Jones, J. B. *Chem. Commun.* **2001**, 903.

(264) Matsumoto, K.; Davis, B. G.; Jones, J. B. *Chem. Eur. J.* **2002**, *8*, 4129.
(265) Doores, K. J.; Davis, B. G. *Chem. Commun.* **2005**, 168.
(266) Davis, B. G.; Sala, R. F.; Hodgson, D. R. W.; Ullman, A.; Khumtaveeporn, K.; Estell, D. A.; Sanford, K.; Bott, R. R.; Jones, J. B. *ChemBioChem* **2003**, *4*, 533.
(267) Davis, B. G. *Chem. Commun.* **2001**, 351.
(268) Rendle, P. M.; Seger, A.; Rodrigues, J.; Oldham, N. J.; Bott, R. R.; Jones, J. B.; Cowan, M. M.; Davis, B. G. *J. Am. Chem. Soc.* **2004**, *126*, 4750.
(269) Gamblin, D. P.; Garnier, P.; van Kasteren, S.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 828.
(270) Bernardes, G. J. L.; Gamblin, D. P.; Davis, B. G. *Angew. Chem., Int. Ed. Engl.* **2006**, *45*, 4007.
(271) Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. *Angew. Chem., Int. Ed.* **2008**, *47*, 2244.
(272) Macmillan, D.; Bill, R. M.; Sage, K. A.; Fern, D.; Flitsch, S. L. *Chem. Biol.* **2001**, *8*, 133.
(273) Swanwick, R. S.; Daines, A. M.; Tey, L.-H.; Flitsch, S. L.; Allemann, R. K. *ChemBioChem* **2005**, *6*, 1338.
(274) Swanwick, R. S.; Daines, A. M.; Flitsch, S. L.; Allemann, R. K. *Org. Biomol. Chem.* **2005**, *3*, 572.
(275) Mullegger, J.; Chen, H. M.; Warren, R. A. J.; Withers, S. G. *Angew. Chem., Int. Ed. Engl.* **2006**, *45*, 2585.
(276) Macindoe, W. M.; van Oijen, A. H.; Boons, G.-J. *Chem. Commun.* **1998**, 847.
(277) Cornish, V. W.; Hahn, K. M.; Schultz, P. G. *J. Am. Chem. Soc.* **1996**, *118*, 8150.
(278) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182.
(279) Liu, H.; Wang, L.; Brock, A.; Wong, C.-H.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 1702.
(280) Van Kasteren, S. I.; Kramer, H. B.; Gamblin, D. P.; Davis, B. G. *Nat. Protocols* **2007**, *2*, 3185.
(281) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci., U.S.A.* **2002**, *99*, 19.
(282) Van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 1282.
(283) Merkel, L.; Beckmann, H. S. G.; Wittmann, V.; Budisa, N. *ChemBioChem* **2008**, *9*, 1220.
(284) Kansas, G. S. *Blood* **1996**, *88*, 3259.
(285) Spevak, W.; Foxall, C.; Charych, D. H.; Dasgupta, F.; Nagy, J. O. *J. Med. Chem.* **1996**, *39*, 1018.
(286) Campbell, S. J.; Carlotti, F.; Hall, P. A.; Clark, A. J.; Wolf, C. R. *J. Cell. Sci.* **1996**, *109*, 2619.
(287) Sibson, N. R.; Blamire, A. M.; Bernades-Silva, M.; Laurent, S.; Boutry, S.; Muller, R. N.; Styles, P.; Anthony, D. C. *Magn. Reson. Med.* **2004**, *51*, 248.
(288) Ferrari, C. C.; Depino, A. M.; Prada, F.; Muraro, N.; Campbell, S.; Podhajcer, O.; Perry, V. H.; Anthony, D. C.; Pitossi, F. J. *Am. J. Pathol.* **2004**, *165*, 1827.
(289) Grau, G. E.; Mackenzie, C. D.; Carr, R. A.; Redard, M.; Pizzolato, G.; Allasia, C.; Cataldo, C.; Taylor, T. E.; Molyneux, M. E. *J. Infect. Dis* **2003**, *187*, 461.
(290) Winans, K. A.; King, D. S.; Rao, V. R.; Bertozzi, C. R. *Biochemistry* **1999**, *38*, 11700.
(291) Paulson, J. C.; Hill, R. L.; Tanabe, T.; Ashwell, G. *J. Biol. Chem.* **1977**, *252*, 8624.
(292) Totani, K.; Ihara, Y.; Matsuo, I.; Koshino, H.; Ito, Y. *Angew. Chem., Int. Ed.* **2005**, *44*, 7950.
(293) Naruchi, K.; Hamamoto, T.; Kurogochi, M.; Hinou, H.; Shimizu, H.; Matsushita, T.; Fujitani, N.; Kondo, H.; Nishimura, S.-I. *J. Org. Chem.* **2006**, *71*, 9609.
(294) Matsushita, T.; Hinou, H.; Fumoto, M.; Kurogochi, M.; Fujitani, N.; Shimizu, H.; Nishimura, S.-I. *J. Org. Chem.* **2006**, *71*, 3051.
(295) Hanashima, S.; Manabe, S.; Ito, Y. *Angew. Chem., Int. Ed.* **2005**, *44*, 4218.
(296) Sato, M.; Furuike, T.; Sadamoto, R.; Fujitani, N.; Nakahara, T.; Niikura, K.; Monde, K.; Kondo, H.; Nishimura, S. *J. Am. Chem. Soc.* **2004**, *126*, 14013.
(297) Murata, T.; Usui, T. *Biosci., Biotechnol., Biochem.* **2006**, *70*, 1049.
(298) Perugino, G.; Trincone, A.; Rossi, M.; Moracci, M. *Trends Biotechnol.* **2004**, *22*, 31.
(299) Koeller, K. M.; Smith, M. E. B.; Huang, R.-F.; Wong, C.-H. *J. Am. Chem. Soc.* **2000**, *122*, 4241.
(300) Leppänen, A.; Mehta, P.; Ouyang, Y.-B.; Ju, T.; Helin, J.; Moore, K. L.; vanDie, I.; Canfield, W. M.; McEver, R. P.; Cummings, R. D. *J. Biol. Chem.* **1999**, *274*, 24838.
(301) Schuster, M.; Wang, P.; Paulson, J. C.; Wong, C.-H. *J. Am. Chem. Soc.* **1994**, *116*, 1135.
(302) Fumoto, M.; Hinou, H.; Ohta, T.; Ito, T.; Yamada, K.; Takimoto, A.; Kondo, H.; Shimizu, H.; Inazu, T.; Nakahara, Y.; Nishimura, S. *J. Am. Chem. Soc.* **2005**, *127*, 11804.
(303) Fumoto, M.; Hinou, H.; Matsushita, T.; Kurogochi, M.; Ohta, T.; Ito, T.; Yamada, K.; Takimoto, A.; Kondo, H.; Inazu, T.; Nishimura, S.-I. *Angew. Chem., Int. Ed.* **2005**, *44*, 2534.
(304) Trimble, R. B.; Atkinson, P. H.; Tarentino, A. L.; Plummer, T. H.; Maley, F.; Tomer, K. B. *J. Biol. Chem.* **1986**, *261*, 12000.
(305) Deras, I. L.; Takegawa, K.; Kondo, A.; Koto, I.; Lee, Y. C. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1763.
(306) Yamamoto, K.; Kadowaki, S.; Watanabe, J.; Kumagai, H. *Biochem. Biophys. Res. Commun.* **1994**, *203*, 244.
(307) Haneda, K.; Inazu, T.; Yamamoto, K.; Kumagai, H.; Nakahara, Y.; Kobata, A. *Carbohydr. Res.* **1996**, *292*, 61.
(308) Haneda, K.; Takeuchi, M.; Tagashira, M.; Inazu, T.; Toma, K.; Isogai, Y.; Hori, M.; Kobayashi, K.; Takeuchi, M.; Takegawa, K.; Yamamoto, K. *Carbohydr. Res.* **2006**, *341*, 181.
(309) Fan, J. Q.; Takegawa, K.; Iwahara, S.; Kondo, A.; Kato, I.; Abeygunawardana, C.; Lee, Y. C. *J. Biol. Chem.* **1995**, *270*, 17723.
(310) Akaike, E.; Tsutsumida, M.; Osumi, K.; Fujita, M.; Yamanoi, T.; Yamamoto, K.; Fujita, K. *Carbohydr. Res.* **2004**, *339*, 719.
(311) Li, B.; Zeng, Y.; Hauser, S.; Song, H.; Wang, L.-X. *J. Am. Chem. Soc.* **2005**, *127*, 9692.
(312) Rising, T. W. D. F.; Claridge, T. D. W.; Moir, J. W. B.; Fairbanks, A. J. *ChemBioChem* **2006**, *7*, 1177.
(313) Li, B.; Song, H.; Hauser, S.; Wang, L.-X. *Org. Lett.* **2006**, *8*, 3081.
(314) Zeng, Y.; Wang, J.; Li, B.; Hauser, S.; Li, H.; Wang, L.-X. *Chem. Eur. J.* **2006**, *12*, 3355.
(315) Wang, L.-X. *Carbohydr. Res.* **2008**, *343*, 1509.
(316) Kornfield, R.; Kornfield, S. *Annu. Rev. Biochem.* **1985**, *54*, 631.
(317) Stenflo, J.; Fernlund, P. *J. Biol. Chem.* **1982**, *257*, 12180.
(318) Liu, Y.-L.; Hoops, G.; Coward, J. K. *Bioorg. Med. Chem.* **1994**, *2*, 1133.
(319) Shakin-Eshleman, S. H.; Spitalnik, S. L.; Kasturi, L. *J. Biol. Chem.* **1996**, *271*, 6363.
(320) Glover, K. J.; Weerapana, E.; Numao, S.; Imperiali, B. *Chem. Biol.* **2005**, *12*, 1311.
(321) Kowarik, M.; Numao, S.; Feldman, M. F.; Schulz, B. L.; Callewaert, N.; Kiermaier, E.; Catrein, I.; Aebi, M. *Science* **2006**, *314*, 1148.
(322) Wacker, M.; Feldman, M. F.; Callewaert, N.; Kowarik, M.; Clarke, B. R.; Pohl, N. L.; Hernandez, M.; Vines, E. D.; Valvano, M. A.; Whitfield, C.; Aebi, M. *Proc. Natl. Acad. Sci., U.S.A.* **2006**, *103*, 7088.
(323) Ramos, D.; Rollin, P.; Klaffke, W. *Angew. Chem., Int. Ed. Engl.* **2000**, *39*, 396.
(324) Ramos, D.; Rollin, P.; Klaffke, W. *J. Org. Chem.* **2001**, *66*, 2948.
(325) Ohta, T.; Miura, N.; Funitani, N.; Nakajima, F.; Niikura, K.; Sadamoto, R.; Guo, C.-T.; Suzuki, T.; Suzuki, Y.; Monde, K.; Nishimura, S.-I. *Angew. Chem., Int. Ed.* **2003**, *42*, 5186.
(326) Bordusa, F. *Chem. Rev.* **2002**, *102*, 4817.
(327) Wong, C. H.; Schuster, M.; Wang, P.; Sears, P. *J. Am. Chem. Soc.* **1993**, *115*, 5893.
(328) Witte, K.; Seitz, O.; Wong, C. H. *J. Am. Chem. Soc.* **1998**, *120*, 1979.
(329) Tolbert, T. J.; Wong, C.-H. *Methods Mol. Biol.* **2004**, *283*, 267.
(330) Witte, K.; Sears, P.; Martin, R.; Wong, C.-H. *J. Am. Chem. Soc.* **1997**, *119*, 2114.
(331) Cumming, D. A. *Glycobiology* **1991**, *1*, 115.
(332) Stanley, P. *Glycobiology* **1992**, *2*, 99.
(333) Grabenhorst, E.; Schlenke, P.; Pohl, S.; Nimtz, M.; Conradt, H. S. *Glycoconjugate J.* **1999**, *16*, 81.
(334) Sears, P.; Wong, C.-H. *Cell. Mol. Life. Sci.* **1998**, *54*, 223.
(335) Moffat, A. S. *Science* **1995**, *268*, 658.
(336) Hennet, T.; Ellies, L. G. *Biochim. Biophys. Acta* **1999**, *1473*, 123.
(337) Jenkins, N.; Parekh, R. B.; James, D. C. *Nat. Biotechnol.* **1996**, *14*, 975.
(338) Wildt, S.; Gerngross, T. U. *Nat. Rev. Microbiol.* **2005**, *3*, 119.
(339) Hamilton, S. R.; Bobrowicz, P.; Bobrowicz, B.; Davidson, R. C.; Li, H.; Mitchell, T.; Nett, J. H.; Rausch, S.; Stadheim, T. A.; Wischnewski, H.; Wildt, S.; Gerngross, T. U. *Science* **2003**, *301*, 1244.
(340) Choi, B.-K.; Bobrowicz, P.; Davidson, R. C.; Hamilton, S. R.; Kung, D. H.; Li, H.; Miele, R. G.; Nett, J. H.; Wildt, S.; Gerngross, T. U. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5022.
(341) Kobata, A.; Amano, J. *Immunol. Cell Biol.* **2005**, *83*, 429.
(342) Hansen, W.; Grabenhorst, E.; Nimtz, M.; Mueller, K.; Conradt, H. S.; Wirth, M. *Metab. Eng.* **2005**, *7*, 221.
(343) Umaña, P.; Jean-Mairet, J.; Moudry, R.; Amstutz, H.; Bailey, J. E. *Nat. Biotechnol.* **1999**, *17*, 176.
(344) Ferrara, C.; Brunker, P.; Suter, T.; Moser, S.; Puntener, U.; Umana, P. *Biotechnol. Bioeng.* **2006**, *93*, 851.

(345) Bobrowicz, P.; Davidson, R. C.; Li, H.; Potgieter, T. I.; Nett, J. H.; Hamilton, S. R.; Stadheim, T. A.; Miele, R. G.; Bobrowicz, B.; Mitchell, T.; Rausch, S.; Renfer, E.; Wildt, S. *Glycobiology* **2004**, *14*, 757.
(346) Hamilton, S. R.; Davidson, R. C.; Sethuraman, N.; Nett, J. H.; Jiang, Y.; Rios, S.; Bobrowicz, P.; Stadheim, T. A.; Li, H.; Choi, B.-K.; Hopkins, D.; Wischnewski, H.; Roser, J.; Mitchell, T.; Strawbridge, R. R.; Hoopes, J.; Wildt, S.; Gerngross, T. U. *Science* **2006**, *313*, 1441.
(347) Li, H.; Sethuraman, N.; Stadheim, T. A.; Zha, D.; Prinz, B.; Ballew, N.; Bobrowicz, P.; Choi, B.-K.; Cook, W. J.; Cukan, M.; Houston-Cummings, N. R.; Davidson, R.; Gong, B.; Hamilton, S. R.; Hoopes, J. P.; Jiang, Y.; Kim, N.; Mansfield, R.; Nett, J. H.; Rios, S.; Strawbridge, R.; Wildt, S.; Gerngross, T. U. *Nat. Biotechnol.* **2006**, *24*, 210.
(348) Feldman, M. F.; Wacker, M.; Hernandez, M.; Hitchen, P. G.; Marolda, C. L.; Kowarik, M.; Morris, H. R.; Dell, A.; Valvano, M. A.; Aebi, M. *Proc. Natl. Acad. Sci., U.S.A.* **2005**, *102*, 3016.
(349) Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125.
(350) Luchansky, S. J.; Goon, S.; Bertozzi, C. R. *ChemBioChem* **2004**, *5*, 371.
(351) Yarema, K. J.; Mahal, L. K.; Bruehl, R. E.; Rodriguez, E. C.; Bertozzi, C. R. *J. Biol. Chem.* **1998**, *273*, 31168.
(352) Lee, J. H.; Baker, T. J.; Mahal, L. K.; Zabner, J.; Bertozzi, C. R.; Wiemer, D. F.; Welsh, M. J. *J. Biol. Chem.* **1999**, *274*, 21878.
(353) Hang, H. C.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2001**, *123*, 1242.
(354) Schmidt, C.; Stehling, P.; Schnitzer, J.; Reutter, W.; Horstkorte, R. *J. Biol. Chem.* **1998**, *273*, 19146.
(355) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007.
(356) Prescher, J. A.; Dube, D. H.; Bertozzi, C. R. *Nature* **2004**, *430*, 873.
(357) Chefalo, P.; Pan, Y.; Nagy, N.; Guo, Z.; Harding, C. V. *Biochemistry* **2006**, *45*, 3733.
(358) Saxon, E.; Armstrong, J. I.; Bertozzi, C. R. *Org. Lett.* **2000**, *2*, 2141.
(359) Rabuka, D.; Hubbard, S. C.; Laughlin, S. T.; Argade, S. P.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2006**, *128*, 12078.
(360) Dube, D. H.; Prescher, J. A.; Quang, C. N.; Bertozzi, C. R. *Proc. Natl. Acad. Sci., U.S.A.* **2006**, *103*, 4819.
(361) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046.
(362) Hsu, T.-L.; Hanson, S. R.; Kishikawa, K.; Wang, S.-K.; Sawa, M.; Wong, C.-H. *Proc. Natl. Acad. Sci., U.S.A.* **2007**, *104*, 2614.
(363) Baskin, J. M.; Prescher, J. A.; Laughlin, S., T.; Agard, N.; Chang, P. V.; Meiller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci., U.S.A.* **2007**, *104*, 16793.
(364) Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G.-J. *Angew. Chem., Int. Ed.* **2008**, *47*, 2253.
(365) Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R. *Science* **2008**, *320*, 664.
(366) Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem. Biol.* **2006**, *1*, 644.
(367) Vocadlo, D. J.; Hang, H. C.; Kim, E.-J.; Hanover, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci., U.S.A.* **2003**, *100*, 9116.
(368) Kohler, J. J.; Bertozzi, C. R. *Chem. Biol.* **2003**, *10*, 1303.
(369) Spencer, D. M.; Wandless, T. J.; Schreiber, S. L.; Crabtree, G. R. *Science* **1993**, *262*, 1019.
(370) Kohler, J. J.; Czlapinski, J. L.; Laughlin, S. T.; Schelle, M. W.; de Graffenried, C. L.; Bertozzi, C. R. *ChemBioChem* **2004**, *5*, 1455.
(371) de Graffenried, C. L.; Laughlin, S. T.; Kohler, J. J.; Bertozzi, C. R. *Proc. Natl. Acad. Sci., U.S.A.* **2004**, *101*, 16715.
(372) Mamaev, S. V.; Laikther, A. L.; Arslan, T.; Hecht, S. M. *J. Am. Chem. Soc.* **1996**, *118*, 7243.
(373) Arslan, T.; Mamaev, S. V.; Mamaeva, N. V.; Hecht, S. M. *J. Am. Chem. Soc.* **1997**, *119*, 10877.
(374) Dedkova Larisa, M.; Fahmi Nour, E.; Golovine Serguei, Y.; Hecht Sidney, M. *J. Am. Chem. Soc.* **2003**, *125*, 6616.
(375) Schmidt, R. R.; Castro-Palomino, J. C.; Retz, O. *Pure Appl. Chem.* **1999**, *71*, 729.
(376) Fahmi, N. E.; Golovine, S.; Wang, B.; Hecht, S. M. *Carbohydr. Res.* **2001**, *330*, 149.
(377) Roehrig, C. H.; Retz, O. A.; Hareng, L.; Hartung, T.; Schmidt, R. R. *ChemBioChem* **2005**, *6*, 1805.
(378) Fahmi, N. E.; Dedkova, L.; Wang, B.; Golovine, S.; Hecht, S. M. *J. Am. Chem. Soc.* **2007**, *129*, 3586.
(379) Xu, R.; Hanson, S. R.; Zhang, Z.; Yang, Y.-Y.; Schultz, P. G.; Wong, C.-H. *J. Am. Chem. Soc.* **2004**, *126*, 15654.
(380) Zhang, Z.; Gildersleeve, J.; Yang, Y.-Y.; Xu, R.; Loo, J. A.; Uryu, S.; Wong, C.-H.; Schultz, P. G. *Science* **2004**, *303*, 371.
(381) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci., U.S.A.* **1998**, *95*, 6705.
(382) David, R.; Richter, M. P. O.; Beck-Sickinger, A. G. *Eur. J. Biochem.* **2004**, *271*, 663.
(383) Tolbert, T. J.; Wong, C.-H. *J. Am. Chem. Soc.* **2000**, *122*, 5421.
(384) Hackenberger, C. P. R.; Friel, C. T.; Radford, S. E.; Imperiali, B. *J. Am. Chem. Soc.* **2005**, *127*, 12882.
(385) Tolbert, T. J.; Franke, D.; Wong, C.-H. *Bioorg. Med. Chem.* **2005**, *13*, 909.
(386) Macmillan, D.; Arham, L. *J. Am. Chem. Soc.* **2004**, *126*, 9530.
(387) Macmillan, D.; Bertozzi, C. R. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 1355.
(388) Macmillan, D.; Bertozzi, C. R. *Tetrahedron* **2000**, *56*, 9515.

CR078291I