Available online at www.sciencedirect.com
ScienceDirect
Carbohydrate Research 343 (2008) 1509–1522
Carbohydrate RESEARCH

Perspective

Chemoenzymatic synthesis of glycopeptides and glycoproteins through endoglycosidase-catalyzed transglycosylation

Lai-Xi Wang*

Institute of Human Virology and Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, 725 West Lombard Street, Baltimore, MD 21201, United States

Received 29 February 2008; received in revised form 16 March 2008; accepted 18 March 2008
Available online 27 March 2008

Abstract—Homogeneous glycopeptides and glycoproteins are indispensable for detailed structural and functional studies of glycoproteins. It is also fundamentally important to correct glycosylation patterns for developing effective glycoprotein-based therapeutics. This review discusses a useful chemoenzymatic method that takes advantage of the endoglycosidase-catalyzed transglycosylation to attach an intact oligosaccharide to a polypeptide in a single step, without the need for any protecting groups. The exploration of sugar oxazolines (enzymatic reaction intermediates) as donor substrates has not only expanded substrate availability, but also has significantly enhanced the enzymatic transglycosylation efficiency. Moreover, the discovery of a novel mutant with glycosynthase-like activity has made it possible to synthesize homogeneous glycoproteins with full-size natural N-glycans. Recent advances in this highly convergent chemoenzymatic approach and its application for glycopeptide and glycoprotein synthesis are highlighted.

Keywords: Glycopeptides; Glycoproteins; Chemoenzymatic synthesis; Transglycosylation; Sugar oxazolines; Endoglycosidase

1. Introduction

Glycosylation is one of the most common posttranslational modifications of proteins in eukaryotes. Protein glycosylation plays an important role in protein folding and intracellular trafficking.[1,2] The covalently linked oligosaccharides of glycoproteins are also involved in many important cellular communication processes, including cell adhesion, host–pathogen interaction, development, and immune responses.[3–7] However, detailed structure–activity relationship studies and biomedical applications of glycoproteins are often hampered by their structural micro-heterogeneity. As protein glycosylation involves a series of posttranslational events that are not under direct genetic control, glycoproteins are usually produced as a mixture of glycoforms that have the same polypeptide backbone but differ in the pendant sugar chains, from which pure glycoforms are difficult to isolate. To meet the urgent need of homogeneous materials for basic research and for biomedical applications, many research laboratories worldwide have taken the challenge to develop sophisticated synthetic methods for constructing complex glycopeptides and glycoproteins carrying defined oligosaccharides.

A number of excellent reviews have been published on this topic in recent years.[8–20] In general, three major strategies have been applied for synthesizing homogeneous glycopeptides. The most common strategy is to incorporate pre-formed glycosyl amino acids as building blocks in conventional solid-phase or solution-phase peptide synthesis. This approach has been very successful for preparing glycopeptides carrying relatively small oligosaccharides. Recent work has also demonstrated that when combined with native chemical ligation, this approach is appropriate for constructing some large N-glycopeptides and even selected glycoproteins.[21–23]

* Tel.: +1 410 706 4982; fax: +1 410 706 5068; e-mail: LWang3@ihv.umaryland.edu

0008-6215/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2008.03.025

Nevertheless, a potential problem of this approach is that the bulky glycans attached in the building block may result in low-yield coupling in solid-phase or solution-phase peptide synthesis, and the O-glycosidic linkages in the oligosaccharide moiety are susceptible to acidic hydrolysis under strong acidic conditions (e.g., TFA or HF treatment) required for final global deprotection and the release of peptide from the resin.

Another strategy is the direct coupling between an oligosaccharide glycosylamine and a pre-assembled polypeptide containing a free or selectively activated aspartyl side chain.[24–27] The major advantage of this approach is its convergence. A number of large and complex glycopeptides have been constructed by this strategy.[28–31] However, a major concern of this approach is the efficiency of the key coupling step, which involves a large oligosaccharyl-amine and the protected polypeptide. In addition, global deprotection of the resulting side-chain protected glycopeptide with strong acids remains a long-standing problem that may result in partial hydrolysis of the attached oligosaccharide moiety.

To address these problems, an alternative strategy is to combine enzymatic elaboration of sugar chains with chemical polypeptide synthesis. The chemoenzymatic approach requires the preparation of only monosaccharide-tagged polypeptides, and the enzymatic sugar chain extension is performed in aqueous solutions with free polypeptides, without the need for protecting groups. Thus, this strategy avoids the problems associated with chemical glycopeptide synthesis, such as the 'incompatibility' of protecting group manipulations for glycosylation and final global deprotection.

Both glycosyltransferases and endoglycosidases have been explored to elaborate sugar chains for this chemoenzymatic strategy. Common glycosyltransferases can extend sugar chains by adding monosaccharides one at a time.[32] In contrast, the endoglycosidase-catalyzed transglycosylation reaction can attach a large intact oligosaccharide to a Glc*p*NAc polypeptide in a single step, thus providing a highly convergent approach.[33,34] A major disadvantage of this endoglycosidase-based chemoenzymatic approach is the relatively low transglycosylation efficiency and the issue of product hydrolysis, which is the focus of discussion of this review.

Another class of carbohydrate enzymes that also hold great potential for in vitro glycoprotein synthesis are the oligosaccharyl transferases (OST), which transfer an oligosaccharide precursor to the asparagine side chain of the nascent protein during translation in N-glycoprotein biosynthesis.[35–39] However, the practical application of OST for in vitro glycoprotein synthesis has not yet been fulfilled, mainly because of the complexity and instability of the multiple-subunit complex of the enzyme. In contrast, the single-subunit protein oligosaccharyl transferase PglB, which was recently found to be responsible for protein N-glycosylation in *Campylobacter jejuni* could be further developed for in vitro glycoprotein synthesis.[40–42]

In addition to chemical and chemoenzymatic synthesis, an alternative approach toward homogeneous or less heterogeneous glycoproteins is to perform engineering of the glycan biosynthetic pathway in the host expression system. Toward this end, a major advance has been made in engineering the yeast *Pichia pastoris* to produce glycoproteins with humanized glycosylation.[43,44] The biosynthetic engineering process involves the elimination of endogenous yeast glycosylation pathways, and incorporation with proper localization of a range of eukaryotic proteins essential for human glycan biosynthesis.[45]

It should be pointed out that each of the synthetic approaches so far explored has their own advantages and limitations. The construction of full-size homogeneous glycoproteins with defined oligosaccharides is still a challenging task. The present review intends to focus on recent advances in the endoglycosidase-catalyzed transglycosylation strategy for the synthesis of N-linked glycopeptides and glycoproteins. The scope, limitation, and applications of the chemoenzymatic synthetic strategy are discussed.

2. Endoglycosidase-catalyzed transglycosylation for glycopeptide synthesis

endo-β-*N*-Acetylglucosaminidases (ENGases) are an important class of endoglycosidases, which are able to release N-glycans from glycoproteins by hydrolyzing the β-(1→4)-glycosidic bond in the *N*,*N*′-diacetylchitobiose core. These enzymes are widely distributed in nature and have been found in microorganisms, plants, animals, and human cells.[46–48] Besides hydrolytic activity, several enzymes in this class have been found to possess transglycosylation activity, that is, the ability to transfer the released oligosaccharyl moiety to a suitable acceptor to form a new glycosidic linkage. These ENGases include Endo-M from *Mucor hiemalis*,[49] Endo-A from *Arthrobactor protophormiae*,[50,51] Endo-CE from *Caenorhabditis elegans*,[47] and Endo-BH from alkaliphilic *Bacillus halodurans* C-125.[52] The observation that a Glc*p*NAc-containing peptide could serve as an efficient acceptor for the ENGase-catalyzed transglycosylation to form a new glycopeptide has suggested a new avenue for a highly convergent assembly of glycopeptides.[49,53–56] Thus, a new glycopeptide can be constructed by a concise two-step approach: First, a Glc*p*NAc-containing polypeptide would be synthesized, usually through solid-phase peptide synthesis, and then an intact oligosaccharide would be transferred to the acceptor by an appropriate ENGase to give the target glycopeptide (Fig. 1).

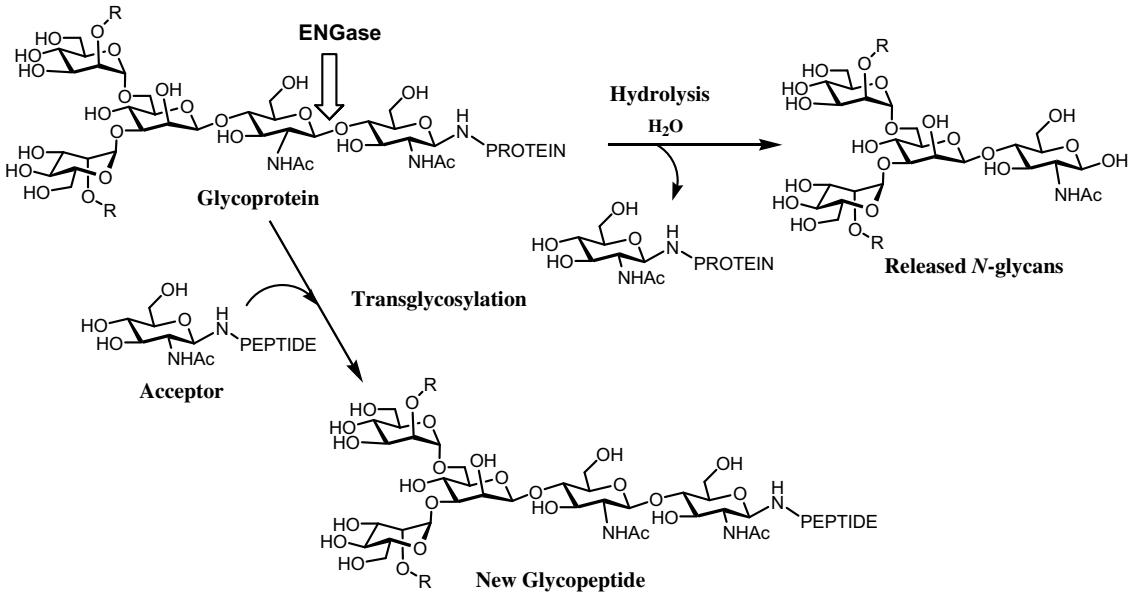


Figure 1. ENGase-catalyzed hydrolysis and transglycosylation.

The most frequently used ENGases in this chemoenzymatic approach are the bacterial enzyme Endo-A and the fungal enzyme Endo-M. These two *endo*-enzymes have distinct substrate specificity. Endo-A is specific for high-mannose type N-glycans, whereas Endo-M can act on N-glycans of the three major types (high-mannose type, hybrid type, and complex type), with preference to the complex type sugars. As summarized in our previous review,[34] a series of bioactive glycopeptides have been synthesized by this chemoenzymatic approach. Typical examples include the glycosylated calcitonin,[57] a *C*-linked analog of N-glycopeptides,[55,56] a glycosylated fragment of the nicotinic acetylcholine receptor (nAChR),[58] a glycosylated substance P,[59] large HIV-1 envelope glycoprotein fragments,[60,61] and homogeneous CD52 antigens carrying full-size high-mannose and complex type N-glycans.[62] These large, homogeneous glycopeptides are otherwise difficult to obtain by other means for structural and functional studies.

In comparison with the sequential extension of sugar chains by common glycosyltransferases that add monosaccharides one at a time, a unique advantage of the ENGase-catalyzed transglycosylation is the single-step attachment of an intact oligosaccharide to the polypeptide to form a new glycopeptide. However, the ENGase-based chemoenzymatic approach has encountered several problems that have hampered its broad application. First, ENGases are inherently glycosyl hydrolases. In comparison with their hydrolytic activity, the transglycosylation activity of ENGases is relatively low. Secondly, product hydrolysis will become a significant problem when the product is accumulated, as the resulting glycopeptide is a substrate for the enzyme. Although the incorporation of organic solvents in the reaction medium can enhance the transglycosylation yield to some extent,[56,61–63] the overall efficiency is generally low (5–20% in yield). Thirdly, the restriction of this chemoenzymatic approach to the use of only natural N-glycans or N-glycopeptides as the donor substrates for transglycosylation has hitherto limited its usefulness, as these natural substrates themselves are difficult to obtain. Recent exploration of synthetic sugar oxazolines (enzymatic reaction intermediates) as synthetic substrates for the enzymatic transglycosylation, together with the discovery of glycosynthase-like mutants of ENGases, has specifically addressed these problems and made the chemoenzymatic method a highly efficient one for synthesizing N-linked glycopeptides and glycoproteins.

3. Exploring transition-state analog substrates for endoglycosidase-catalyzed transglycosylation

To expand the substrate availability and to enhance the overall efficiency of enzymatic transglycosylation for glycopeptide synthesis, we and others have recently explored synthetic sugar oxazolines, the presumed enzymatic reaction intermediates, as donor substrates for the ENGase-catalyzed transglycosylation.[64–69] This development was based on the assumption that the ENGase-catalyzed reaction proceeds via a substrate-assisted mechanism through the participation of the 2-acetamido group to form a 1,2-oxazolinium ion intermediate, as demonstrated for some family-20 *N*-acetyl-β-hexosaminidases[70,71] and some family-18 chitinases[72–74] (Fig. 2).

There are precedents that disaccharide oxazoline derivatives could serve as transition-state analog

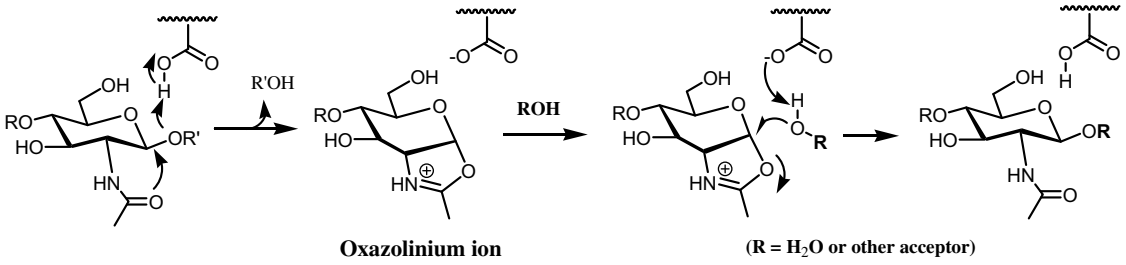
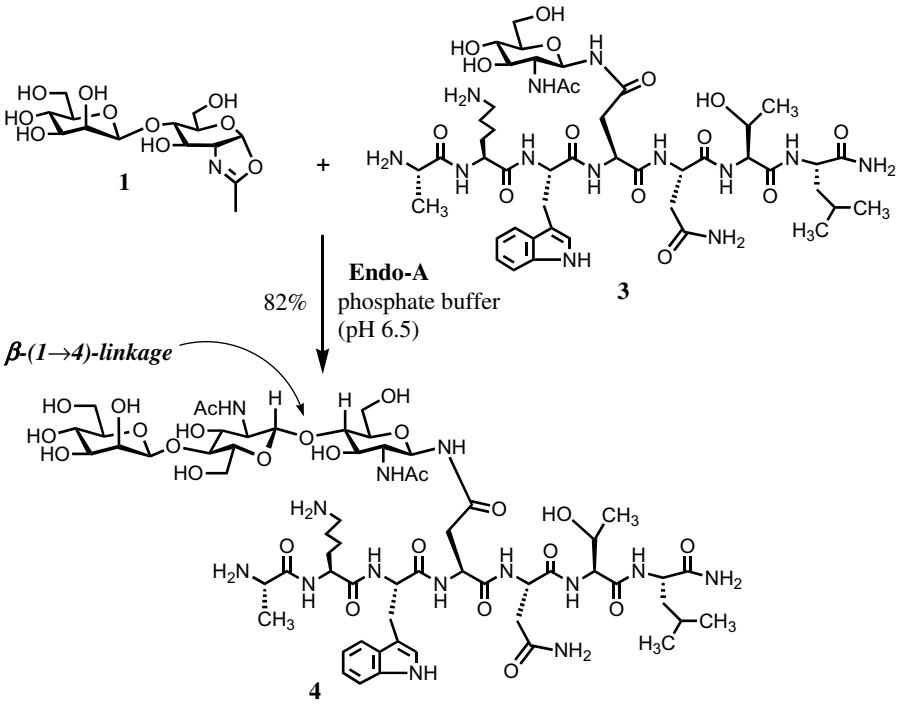


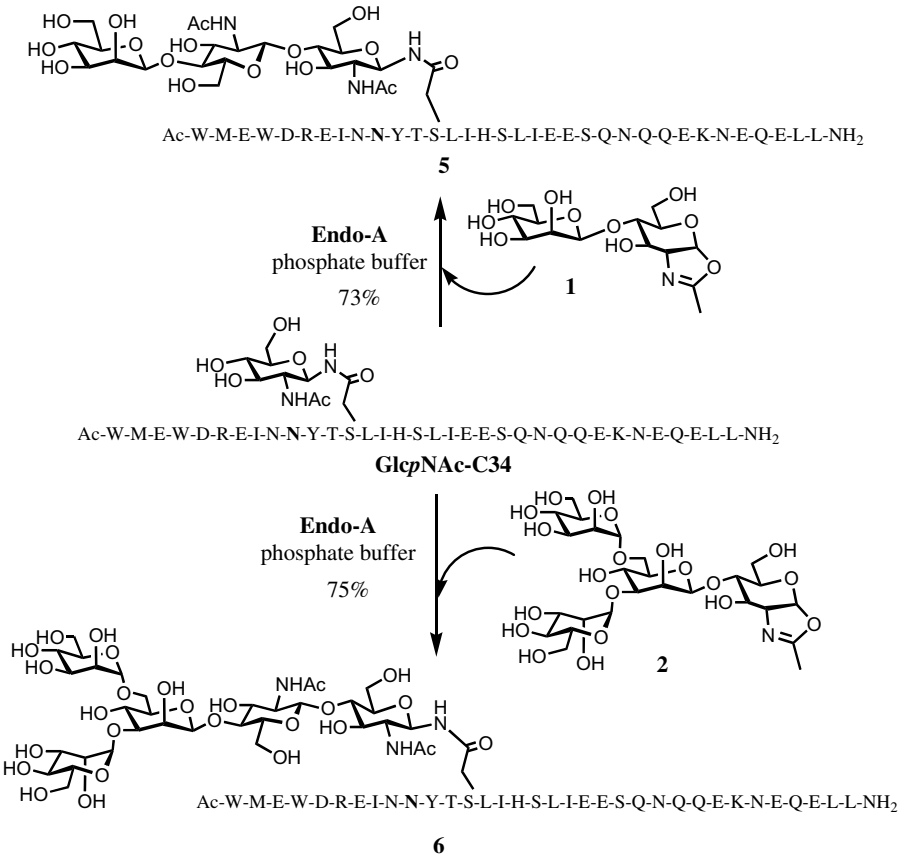
Figure 2. Substrate-assisted mechanism for ENGase-catalyzed reactions.

substrates for chitinase- and hyaluronidase-catalyzed polymerization for polysaccharide synthesis.[75–80] More direct evidence was provided by Fujita et al., who have demonstrated that a simple disaccharide oxazoline derived from β-Man*p*-(1→4)-Glc*p*NAc could serve as a substrate for detecting the transglycosylation activity of Endo-A and Endo-M through the formation of a PNP-tagged trisaccharide derivative.[81] These observations prompted us to examine the feasibility of synthetic sugar oxazolines as favorable substrates for chemoenzymatic synthesis of glycopeptides and glycoproteins.

For the purpose, we have synthesized di- and tetrasaccharide oxazolines corresponding to the core of N-glycans, Man*p*-(1→4)-Glc*p*NAc-oxazoline (**1**) and Manp_3Glc*p*NAc-oxazoline (**2**), respectively.[64] Then, we tested Endo-A and Endo-M catalyzed transglycosylation using Glc*p*NAc-heptapeptide (**3**) derived from HIV-1 gp120 as the acceptor. It was found that the synthetic oxazoline (**1**) could serve as a good substrate for transglycosylation catalyzed by Endo-A to form the corresponding glycopeptide (**4**) in 82% yield, when a small excess of the donor substrate was used (Scheme 1). The enzymatic transglycosylation proceeded under very mild conditions (phosphate buffer, pH 6.5, room temperature). Detailed NMR studies of the product indicated that the transglycosylation was regio- and stereoselective, leading to the formation of the desired β-(1→4)-glycosidic linkage between the two Glc*p*NAc moieties in the newly formed N-glycan. In addition, more complex Glc*p*NAc-peptides were shown to be equally efficient as acceptors for the oxazoline-based transglycosylation. For example, two glycoforms (**5** and **6**) of HIV-1 gp41 glycopeptide (glyco-C34) carrying the N-linked core tri- and pentasaccharides were efficiently synthesized in good yields through a single-step enzymatic transglycosylation, using sugar oxazolines **1** and **2** as donor substrates, respectively (Scheme 2). Interestingly, while the di- and tetrasaccharide oxazoline



Scheme 1. ENGase-catalyzed synthesis of a gp120 fragment carrying the core trisaccharide.

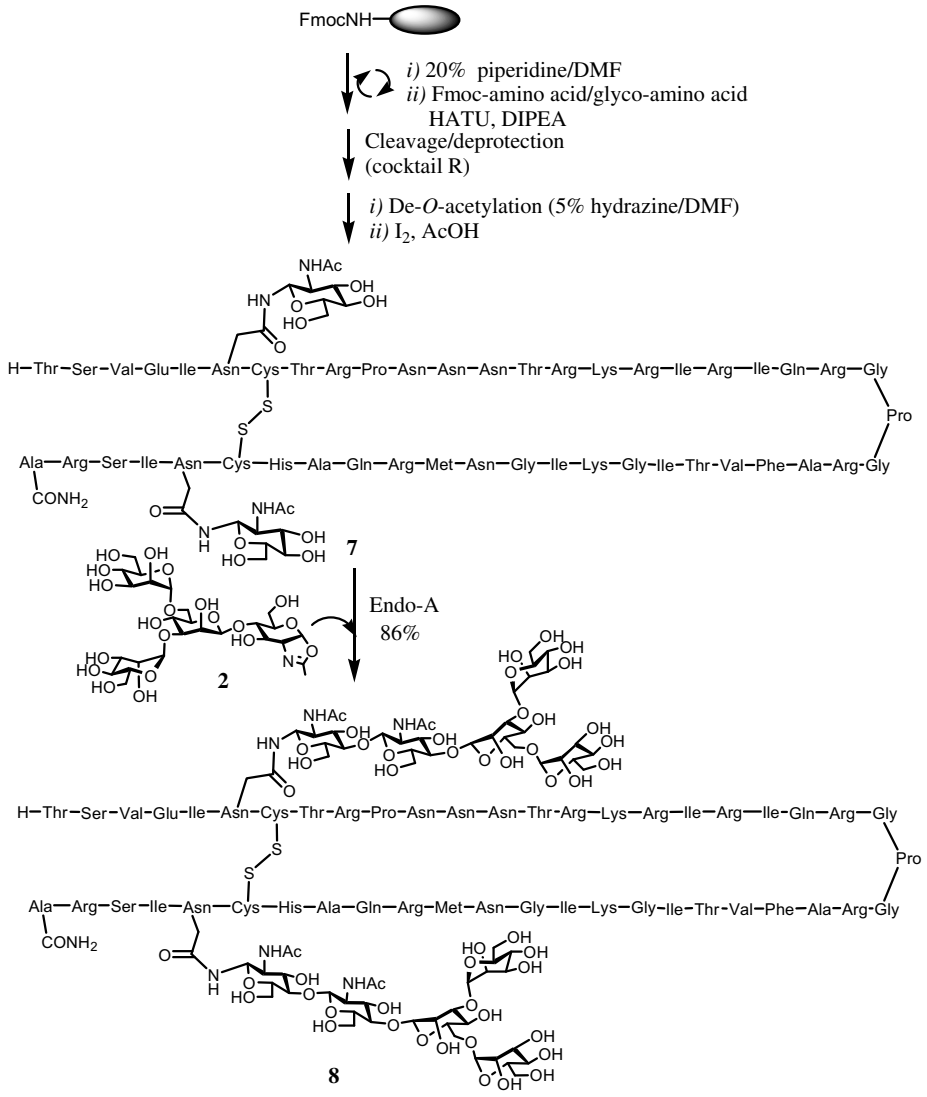


Scheme 2. ENGase-catalyzed transglycosylation with the di- and tetrasacharide oxazolines.

(**1** and **2**) were excellent substrates for Endo-A, the resulting glycopeptides (**5** and **6**) were completely resistant to hydrolysis catalyzed by Endo-A under reaction conditions.[64] The great difference in enzymatic reaction rates between the activated oxazoline substrates and the ground-state substrate (the product) favors the accumulation of the transglycosylation product.

This improved chemoenzymatic approach has been successfully applied to the synthesis of a large HIV-1 gp120 fragment, a 47-mer V3 domain glycopeptide carrying two core N-linked pentasaccharide.[65] The synthesis was achieved by a concise, two-step approach. First, the 47-mer polypeptide that contains two Glc*p*NAc moieties was prepared by an Fmoc-based solid-phase peptide synthesis, using Glc*p*NAc-Asn as building blocks to introduce the two Glc*p*NAc tags. Then, the Endo-A catalyzed transglycosylation was applied to attach two N-glycans simultaneously to the acceptor, providing the desired glycopeptide (**8**) in an excellent yield (Scheme 3). The synthetic V3 glycopeptides were used for probing the effects of glycosylation on the global conformations and the protease stability of the V3 domain.[65] The high-yield, simultaneous enzymatic double glycosylation with the sugar oxazolines on a large polypeptide attested to the power and potential of this chemoenzymatic approach for constructing large and complex glycopeptides that are difficult to obtain by other means. These experimental data have demonstrated that the use of sugar oxazolines as activated substrates for ENGase-catalyzed transglycosylation has not only expanded the substrate availability, but also led to substantial enhancement of the overall synthetic efficiency, allowing a high-yield assembly of large glycopeptides in a highly convergent manner, without the need for protecting groups.

To evaluate the donor substrate structural requirement in the enzymatic transglycosylation, we have synthesized an array of truncated and modified sugar oxazolines, and tested their activity toward Endo-A[66] (Scheme 4). It was found that the 6′-modified oxazolines (**9** and **10**) could serve as substrates for transglycosylation, allowing the synthesis of selectively modified glycopeptides (**15** and **16**). However, the change of the configurations at the 2′ and/or 4′-hydroxyl groups, as in the case of β-Glc*p*-(1→4)-Glc*p*NAc-oxazoline **11** and LacNAc-oxazoline **12**, or modification at the C-2′ position as in the case of oxazoline **13**, all resulted in a total loss of substrate activity to Endo A (Scheme 4). These results suggest that the β-Man*p*-(1→4)-Glc*p*NAc-oxazoline moiety is the minimum structure recognized by Endo-A for transglycosylation. However, selective modification on the mannose moiety such as



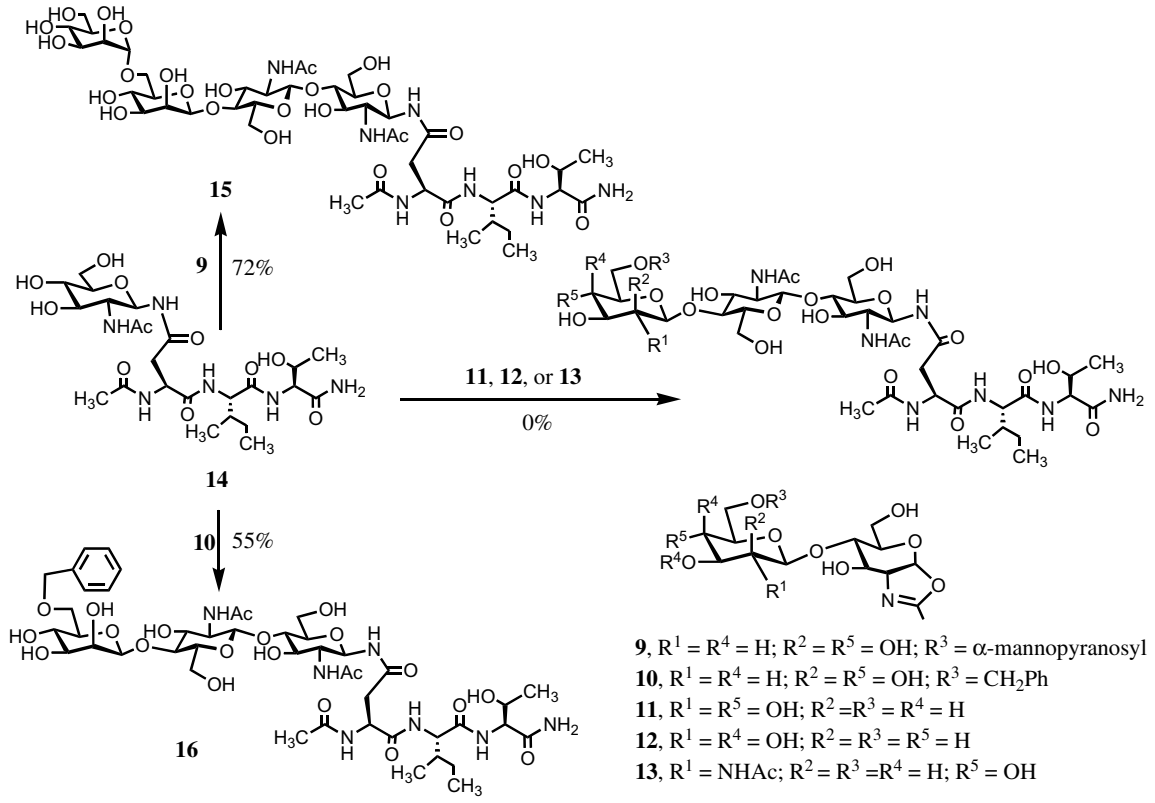
Scheme 3. Chemoenzymatic synthesis of a cyclic V3 glycopeptide carrying two N-glycans.

introducing a tag or additional sugar residue at the 6′-position could be tolerated without the loss of substrate activity toward Endo-A.

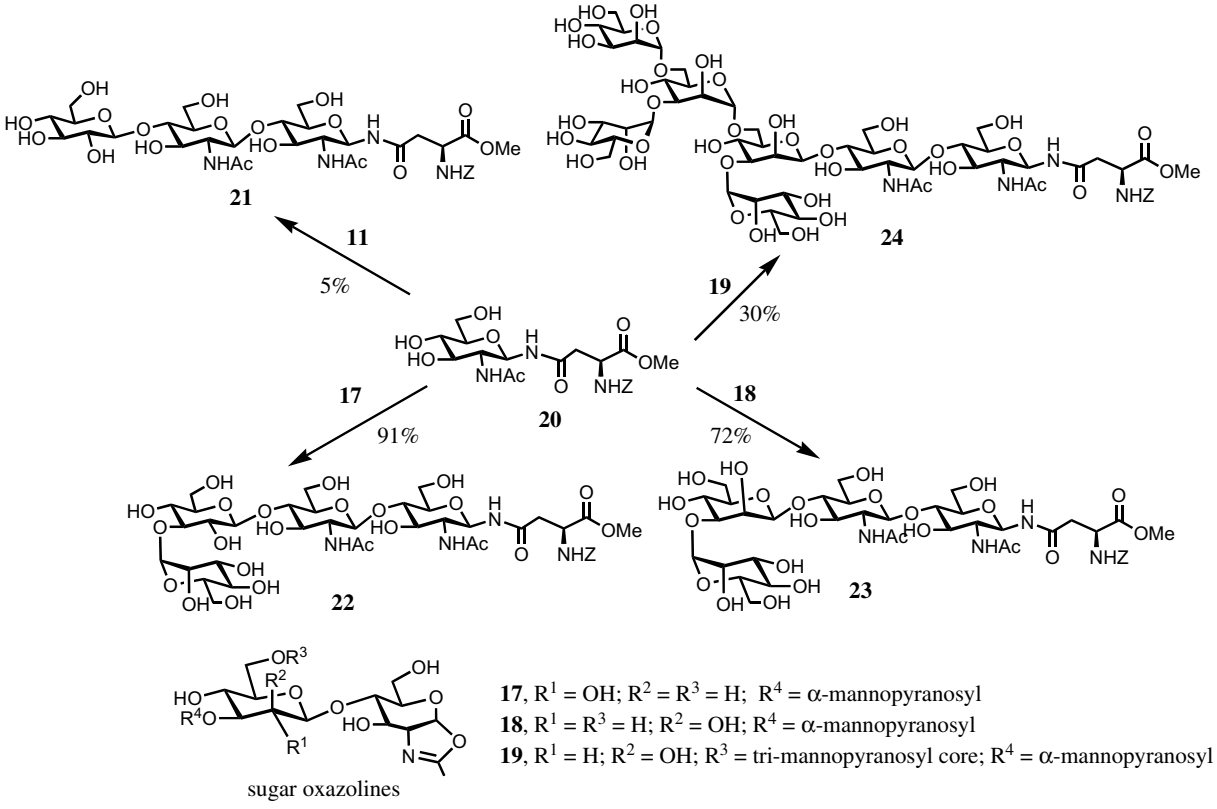
Independently, Fairbanks and co-workers have synthesized additional sugar oxazolines and tested their activity for transglycosylation with Endo-M[68,69] (Scheme 5). It was observed that the glucose-containing disaccharide oxazoline, β-Glc*p*-(1→4)-Glc*p*NAc-oxazoline (**11**), was a substrate of Endo-M for transglycosylation to form the product **21**, albeit at a low reaction rate with only a 5% yield. This activity was in contrast to Endo-A, which could not take compound **11** as a substrate for transglycosylation.[66]

Very interestingly, when disaccharide **11** was extended to a trisaccharide derivative with an additional α-(1→3)-linked mannosyl residue attached to the glucose moiety, the resulting trisaccharide oxazoline **17** became an excellent substrate for Endo-M, and the transglycosylation with acceptor **20** gave a 91% yield of the product **22**. Although the modified sugar oxazolines were active as donor substrates, the resulting glucose-containing products **21** and **22** were completely inactive to hydrolysis by Endo-M, due to the slight structural modifications.

However, when large sugar oxazolines corresponding to the natural structure of N-glycans were used as substrates, the chemoenzymatic synthesis became problematic because of the hydrolysis of the resulting glycopeptides by Endo-M. For example, when a Man-p_5Glc*p*NAc-oxazoline **19** was used as a donor substrate, the optimized yield for compound **24** was approximately 30% (Scheme 5). This low yield was attributed to fast competitive hydrolysis of the product rather than a slow enzymatic turnover of the oxazoline donor, as the resulting product turned to be an excellent substrate for Endo-M. Indeed, a prolonged incubation with Endo-M led to almost complete hydrolysis of the product **24**. Thus, product hydrolysis will become unavoidable when the glycans attached on the polypeptide are



Scheme 4. Endo-A catalyzed transglycosylation with modified sugar oxazolines.



Scheme 5. Endo-M catalyzed transglycosylation with modified sugar oxazolines.

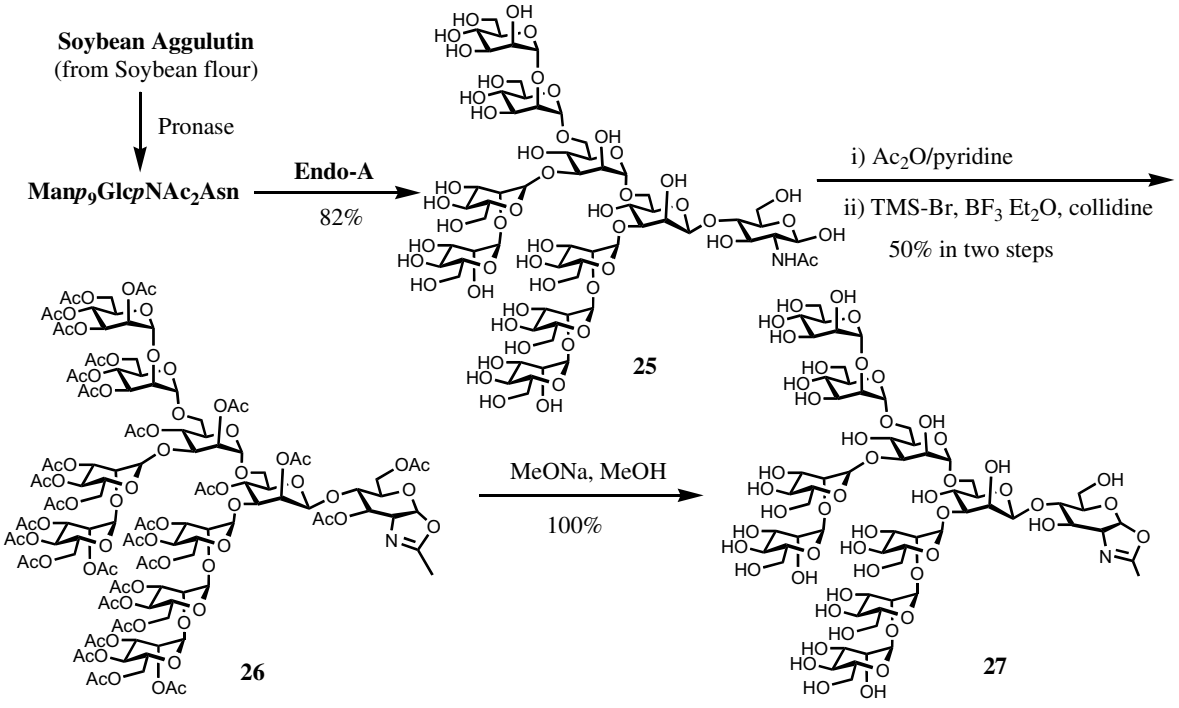
excellent substrates for the *endo*-enzymes. This weakness significantly limits the application of this chemoenzymatic method for synthesizing glycopeptides and glycoproteins carrying natural N-glycans.

4. Generation of glycosynthase-like mutants for an improved synthesis of glycopeptides carrying full-size natural N-glycans

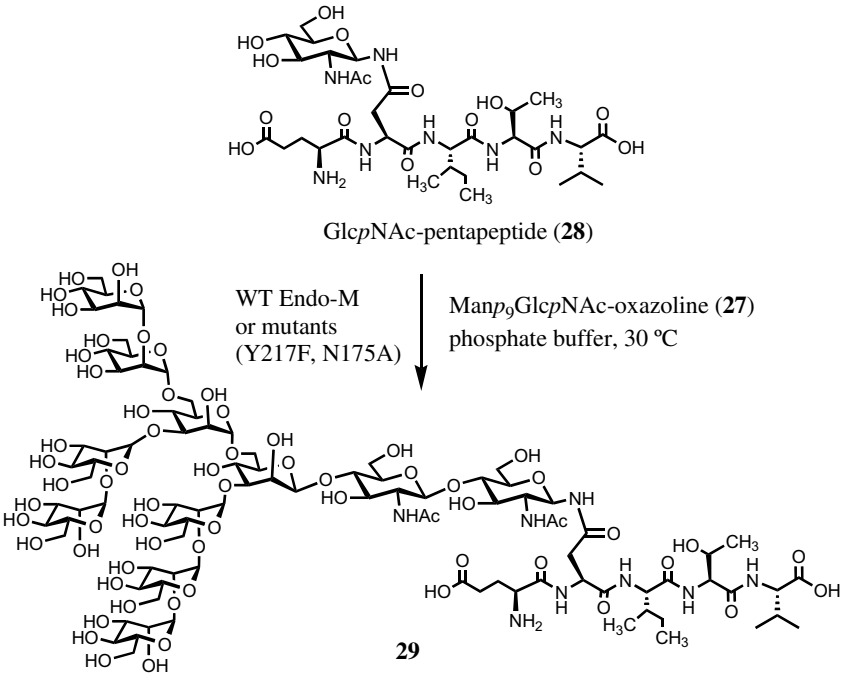
To address the low efficiency and product hydrolysis generally associated with glycosidase-catalyzed synthesis, a promising approach is to create novel glycosidase mutants termed glycosynthases, which lack hydrolytic activity because of the deletion of the nucleophilic residue, but can still take an activated glycosyl donor such as glycosyl fluoride with an opposite anomeric configuration for catalysis to form a new glycosidic bond.[82–84] As catalysis by ENGases proceed via a substrate-assisted mechanism in which the nucleophile is the 2-acetamido group in the substrate, the general method to create a glycosynthase by 'knocking out' the nucleophilic residue in the enzyme cannot apply to ENGases. Thus, we have teamed up with Professor Yamamoto's group to evaluate various Endo-M mutants aiming to diminish the hydrolytic activity but to keep or enhance the transglycosylation activity.[85]

A series of Endo-M mutants were created by site-directed mutagenesis on residues in, or around, the putative catalytic region of wild type Endo-M. Their transglycosylation activity was examined using synthetic sugar oxazolines as the activated substrate. For this purpose, a large sugar oxazoline corresponding to the natural high-mannose type N-glycan, Man*p*$_9$Glc*p*NAc-oxazoline (**27**), was synthesized. Briefly, the Asn-linked oligosaccharide Man*p*$_9$Glc*p*NAc$_2$Asn, which was prepared from soybean flour as previously described,[86] was treated with Endo-A to give the Man*p*$_9$Glc*p*NAc oligosaccharide **25**. This oligosaccharide was acetylated and the resulting peracetylated compound was treated with trimethylsilyl bromide and $BF_3{\cdot}Et_2O$ in the presence of collidine to give the peracetylated sugar oxazoline **26**. Finally, de-O-acetylation with sodium methoxide in methanol provided the free Man*p*$_9$Glc*p*NAc-oxazoline **27** (Scheme 6).

The screening of the Endo-M mutants was performed using Man*p*$_9$Glc*p*NAc-oxazoline **27** as the donor substrate and a small Glc*p*NAc-peptide **28** as the acceptor. The experiments led to the discovery of two interesting mutants, Y217F and N175A. The Y217F mutant was found to possess a much enhanced transglycosylation activity and yet relatively low hydrolytic activity, giving a maximum yield of 50% for the product **29**. In contrast, the wild type Endo-M gave only 8% yield at the optimal point, due to its relatively low transglycosylation activity and also its quick hydrolysis of the product (Scheme 7 and Fig. 3). The enhanced transglycosylation activity might be attributed to an increased affinity of the mutant enzyme for the acceptor substrate, as implicated by a smaller K_m value of the Y217F for a model



Scheme 6. Semi-synthesis of Man*p*$_9$Glc*p*NAc-oxazoline.



Scheme 7. Transglycosylation with Endo-M and its mutants using Manp_9Glc*p*NAc-oxazoline.

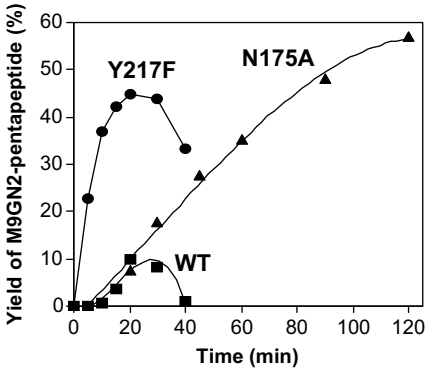


Figure 3. Time-course of the transglycosylation reactions using Manp_sGlc*p*NAc-oxazoline as the donor substrate. The reactions were carried out in a phosphate buffer (pH 6.6) at 30 °C, using the GlcN*p*Ac-pentapeptide as an acceptor (molar ratio of donor to acceptor, 2:1).

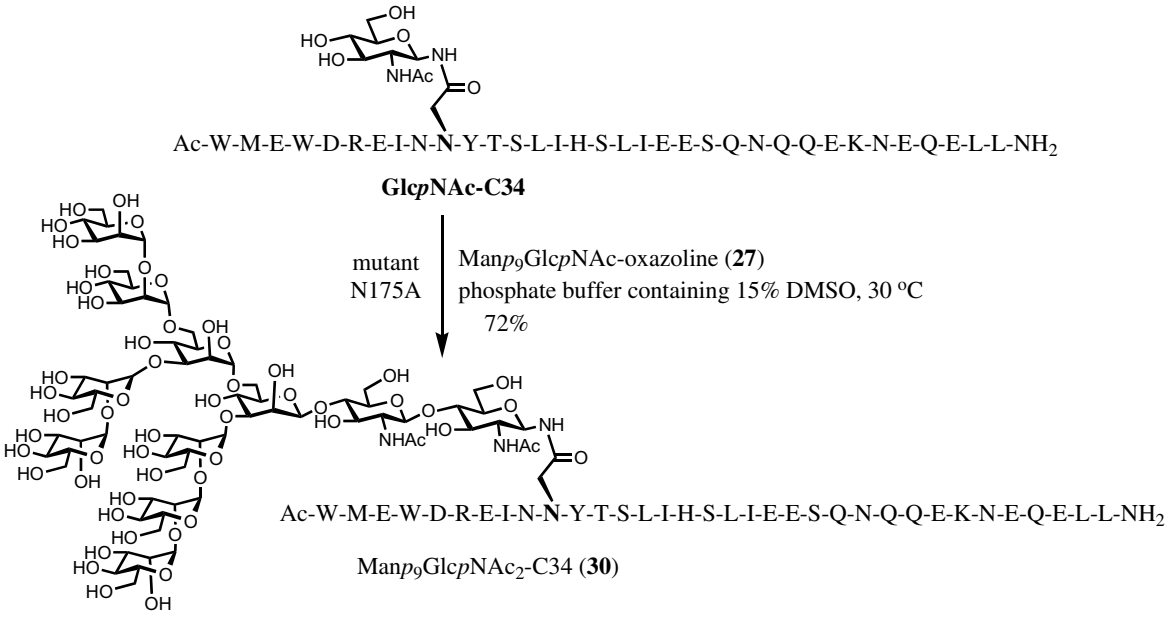
acceptor (Glc*p*NAc-MU) than that of the wild type Endo-M. However, the Y217F mutant can still hydrolyze the product, given a longer incubation time (Scheme 7 and Fig. 3).

On the other hand, the N175A mutant acted as a glycosynthase. That is, this mutant could take the highly activated sugar oxazoline **27** as the donor substrate for transglycosylation, but it lacked the activity to hydrolyze the glycopeptide product **29** that was formed, resulting in the accumulation of the product. The yield reached 60% within 2 h under the reaction condition, but the complete conversion of the starting material **28** to product **29** could be achieved when an excess of sugar oxazoline was added in several portions for a longer incubation time (unpublished data). The usefulness of this novel mutant for glycopeptide synthesis was exemplified by an efficient synthesis of an HIV-1 glycopeptide inhibitor Manp_9Glc*p*NAc$_2$-C34 (**30**) (Scheme 8).

The N175A mutant represents the first glycosynthase in the class of endo-β-*N*-acetylglucosaminidases that proceed via a substrate-assisted mechanism. This glycosynthase holds great potential for the synthesis of large glycopeptides and glycoproteins that carry natural, full-size N-glycans. Nevertheless, the catalytic mechanism of the glycosynthase is still to be characterized, and it is still puzzling how a single mutant at the 'inert' asparagine residue N175 to an 'inert' alanine residue could lead to a dramatic change in the hydrolytic and transglycosylation activity of the endoenzyme. In addition, this mutant enzyme seems to have a much lower turnover of substrate than the wild type Endo-M or its mutant Y217F.[85] Further structural and mutational studies are required to enhance the efficiency of the mutant.

5. Endoglycosidase-catalyzed transglycosylation for glycoprotein synthesis and glycosylation engineering

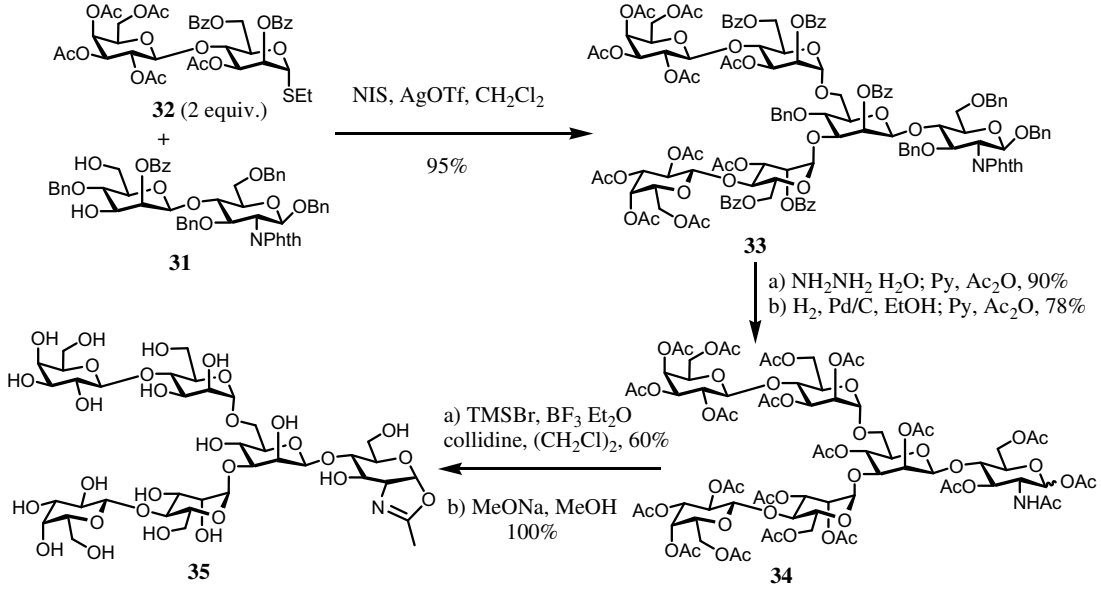
Previous attempts to perform glycosylation remodeling of glycoproteins using natural N-glycans as the donor substrates and wild type Endo-A or Endo-M as the enzyme met with little success. The yield of the glycoprotein product is very low (usually less than 5%) even when



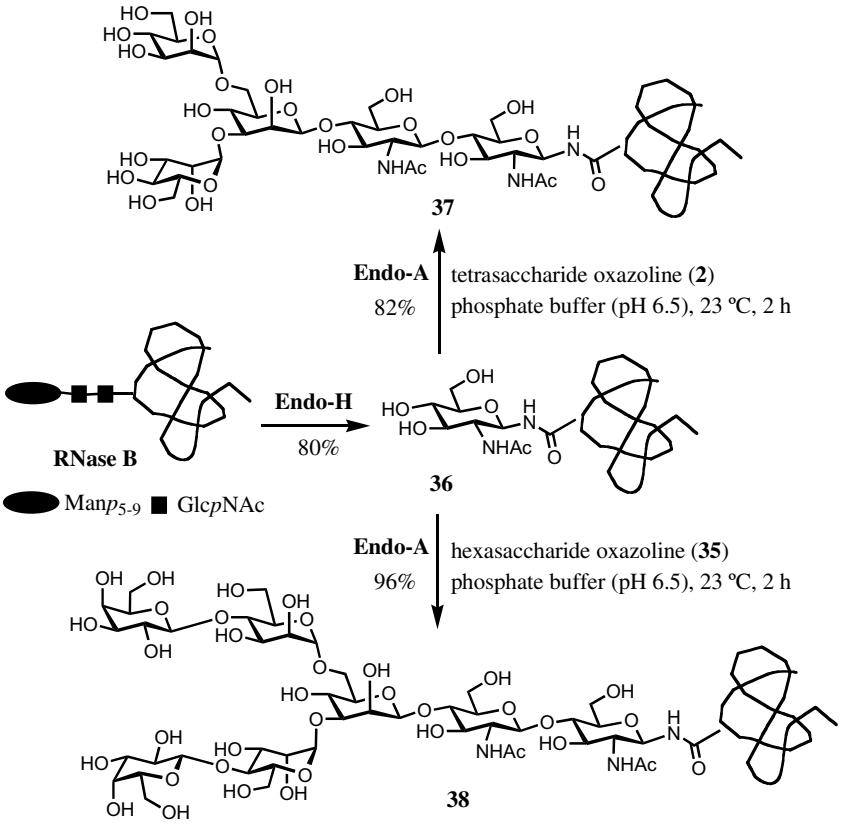
Scheme 8. Enzymatic synthesis of a HIV-1 gp41 glycopeptide using mutant N175A.

a large excess of donor substrates was used.[53,87] This situation prompted us to extend the sugar oxazoline-based chemoenzymatic method to glycoprotein synthesis, as exemplified by the successful glycosylation remodeling of bovine ribonuclease B.[67] For the purpose, a novel hexasaccharide oxazoline (**35**) (Galp_2Manp_3Glc*p*NAc-oxazoline) was synthesized, which has two galactose residues β-(1→4)-linked to the terminal mannose residues in the core Manp_3Glc*p*NAc (Scheme 9). This hexasaccharide derivative can be regarded as a mimic of a bi-antennary complex type N-glycan lacking the internal Glc*p*NAc moieties.

Natural bovine ribonuclease B is a mixture of glycoforms of the high-mannose type N-glycans ranging from Manp_5 to Manp_9. For glycosylation remodeling, the commercially available glycoprotein was treated with Endo-H to remove the heterogeneous N-glycans but leave only the innermost Glc*p*NAc at the Asn-34 site. The resulting Glc*p*NAc-containing ribonuclease, Glc*p*NAc-RNase (**36**), was then used as acceptor substrate for ENGase-catalyzed transglycosylation with the synthetic sugar oxazolines (**2** and **35**). It was found that the transglycosylation with sugar oxazoline **2** proceeded very fast in the presence of Endo-A under mild



Scheme 9. Synthesis of an unnatural hexasaccharide oxazoline.



Scheme 10. Glycosylation remodeling of bovine ribonuclease B.

conditions (phosphate buffer, pH 6.5, 23 °C), giving the homogeneous glycoprotein **37** carrying the core pentasaccharide in 82% isolated yield. The larger sugar oxazoline **35** was found to be an even better substrate for transglycosylation catalyzed by Endo-A, giving the corresponding homogeneous glycoprotein **38** in almost a quantitative yield (Scheme 10).

Again, the glycoprotein, once formed, was completely resistant to hydrolysis by Endo-A. This could be explained by the fact that Endo-A hydrolyzes only high-mannose type natural N-glycans, but glycoprotein **38** carries an unnatural N-glycan mimicking the complex type sugar chain. The enzymatic transglycosylation could take place under very mild, neutral conditions without the need to denature the protein acceptor. These results have clearly demonstrated that the sugar oxazoline-based chemoenzymatic approach is equally applicable for synthesizing homogeneous glycoproteins with defined oligosaccharides, when the acceptor substrate is a Glc*p*NAc-protein instead of a Glc*p*NAc-peptide. It is expected that homogeneous glycoproteins carrying natural N-glycans could also be efficiently synthesized by this approach, when the novel glycosynthase (mutant N175A), which lacks the hydrolytic activity on natural N-glycans, is used for the transglycosylation. The application of the sugar oxazoline-based chemoenzymatic method for total synthesis and glycosylation engineering of an array of biologically interesting glycoproteins is currently under way in our laboratory.

6. Preparation of Glc*p*NAc-functionalized polypeptides and proteins

The discovery of the glycosynthase-like ENGase mutant, in combination with the use of synthetic oligosaccharide oxazolines as substrates, has now made it possible to synthesize full-size homogeneous glycoproteins in a highly convergent manner. Nevertheless, a successful application of this approach for constructing a range of homogeneous glycoproteins will also rely on an easy access to Glc*p*NAc-functionalized proteins. Several approaches are available for the preparation of Glc*p*NAc-functionalized proteins, which are summarized in Figure 4.

One approach is to apply chemical protein synthesis techniques to ligate Glc*p*NAc-peptide and peptide fragments to form a full-length Glc*p*NAc-protein. The synthesis of Asn-linked Glc*p*NAc-peptides is not particularly more difficult than common polypeptides, as the N-linked Glc*p*NAc residue in the Asn-Glc*p*NAc building block can be viewed as a special side chain, which is actually stable toward the relatively strong acid or alkaline treatment required for global polypeptide

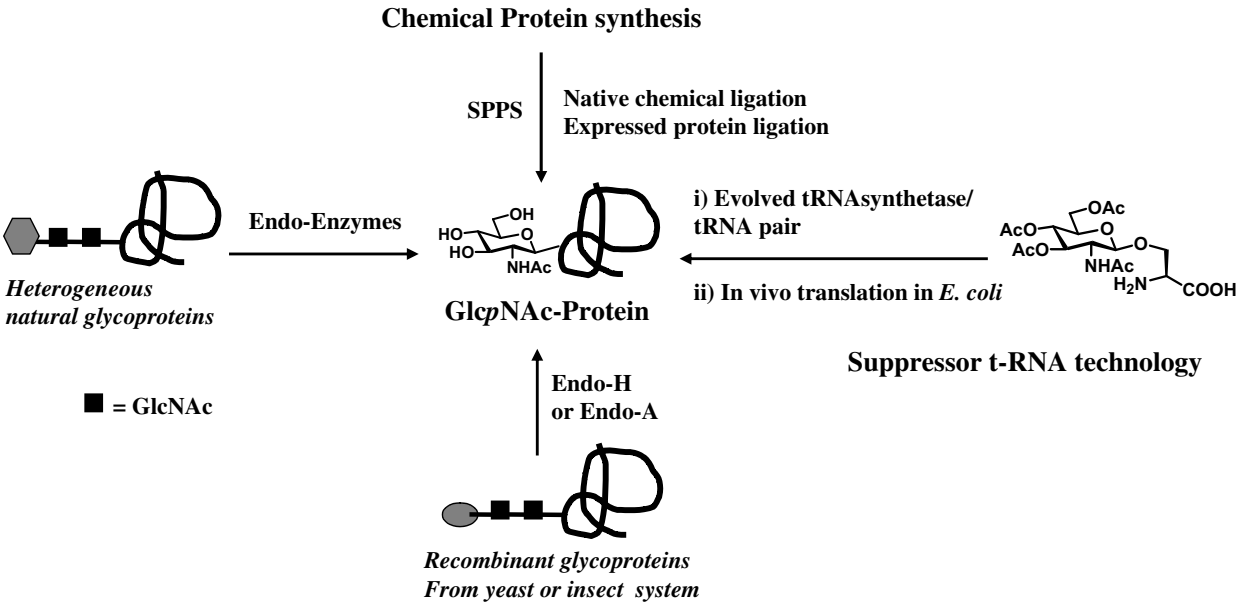


Figure 4. Preparation of Glc*p*NAc-functionalized proteins.

deprotection. When native chemical ligation or expressed protein ligation is used for the conjugation of Glc*p*NAc-peptide fragments, full-length proteins containing one or more Glc*p*NAc residues at predetermined sites could be fulfilled.[88,89]

Another approach is to obtain Glc*p*NAc-proteins from naturally existing glycoproteins through endoglycosidase treatment. In general, natural glycoproteins exist as a mixture of glycoforms, but heterogeneous glycans can be selectively removed by specific endo-*N*-acetylglucosaminidases to leave only the innermost Glc*p*NAc residue attached at the glycosylation sites. Several endoglycosidases that act on different N-glycan substrates are available. Thus, through appropriate choice of the *endo*-enzymes, it is possible to selectively remove high-mannose type N-glycans (e.g., using Endo-F_1, Endo-H) or complex type N-glycans (e.g., using Endo-F_2 and Endo-M).

Alternatively, well-established yeast expression systems can be used to overproduce glycoproteins carrying the yeast glycoforms, then the heterogeneous yeast-type N-glycans (with or without hypermannosylation) can be removed by Endo-A or Endo-H treatment to give homogeneous, Glc*p*NAc-functionalized proteins. Finally, the recently reported in vivo suppressor tRNA technology also holds promise for introducing monosaccharide moieties such as Glc*p*NAc residues into proteins in a novel bacterial expression system.[90,91] Taken together, recent advances in chemical protein synthesis and the exploration of novel protein expression system have provided the flexibility for preparing various Glc*p*NAc-functionalized proteins, thus making the sugar oxazoline-based chemoenzymatic approach feasible for constructing full-size homogeneous glycoproteins for different applications.

7. Conclusion

Homogeneous glycopeptides and glycoproteins are indispensable for detailed structural and functional studies of glycoproteins. Putting in place the correct glycosylation pattern is also fundamentally important for developing glycoprotein-based therapeutics. The chemoenzymatic method described in this review, which is based on endoglycosidase-catalyzed transglycosylation, provides a highly efficient approach for constructing N-linked complex glycopeptides and glycoproteins when synthetic sugar oxazolines are used as the donor substrates.

Both natural and modified N-glycans can be introduced via an appropriate choice of either wild type or mutant ENGases. A particularly appealing feature of the chemoenzymatic approach is the highly convergent ligation between a pre-formed oligosaccharide moiety and a polypeptide moiety in a regio- and stereo-specific manner, without the need for any protecting groups. Thus, this chemoenzymatic approach allows totally independent synthetic manipulations of the oligosaccharide and polypeptide/protein portions, and may provide an ultimate solution to the long-standing problem of 'incompatibility' of protecting group manipulations in glycopeptide synthesis.

Future studies should be directed to mechanistic studies of the ENGase-catalyzed transglycosylation to create more efficient mutants as glycosynthases. In addition,

diverse natural and selectively modified N-glycan oxazolines should be synthesized and tested, aiming to expand the scope of the chemoenzymatic method. It is expected that this novel chemoenzymatic approach will find wide application in constructing glycoproteins with defined oligosaccharides for structural and biological studies and for biomedical applications as well.

Acknowledgment

The chemoenzymatic glycoprotein synthesis project in our laboratory was supported by the National Institutes of Health (NIH Grant R01 GM080374).

References

1. Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364–2369.
2. Petrescu, A. J.; Wormald, M. R.; Dwek, R. A. *Curr. Opin. Struct. Biol.* **2006**, *16*, 600–607.
3. Varki, A. *Glycobiology* **1993**, *3*, 97–130.
4. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.
5. Dwek, R. A.; Butters, T. D.; Platt, F. M.; Zitzmann, N. *Nat. Rev. Drug Discov.* **2002**, *1*, 65–75.
6. Haltiwanger, R. S.; Lowe, J. B. *Annu. Rev. Biochem.* **2004**, *73*, 491–537.
7. Dube, D. H.; Bertozzi, C. R. *Nat. Rev. Drug Discov.* **2005**, *4*, 477–488.
8. Koeller, K. M.; Wong, C. H. *Nat. Biotechnol.* **2000**, *18*, 835–841.
9. Seitz, O. *ChemBioChem* **2000**, *1*, 214–246.
10. Herzner, H.; Reipen, T.; Schultz, M.; Kunz, H. *Chem. Rev.* **2000**, *100*, 4495–4538.
11. Hang, H. C.; Bertozzi, C. R. *Acc. Chem. Res.* **2001**, *34*, 727–736.
12. Davis, B. G. *Chem. Rev.* **2002**, *102*, 579–601.
13. Grogan, M. J.; Pratt, M. R.; Marcaurelle, L. A.; Bertozzi, C. R. *Annu. Rev. Biochem.* **2002**, *71*, 593–634.
14. Wong, C. H. *J. Org. Chem.* **2005**, *70*, 4219–4225.
15. Guo, Z.; Shao, N. *Med. Res. Rev.* **2005**, *25*, 655–678.
16. Pratt, M. R.; Bertozzi, C. R. *Chem. Soc. Rev.* **2005**, *34*, 58–68.
17. Liu, L.; Bennett, C. S.; Wong, C. H. *Chem. Commun. (Camb.)* **2006**, 21–33.
18. Buskas, T.; Ingale, S.; Boons, G. J. *Glycobiology* **2006**, *16*, 113R–136R.
19. Brik, A.; Ficht, S.; Wong, C. H. *Curr. Opin. Chem. Biol.* **2006**, *10*, 638–644.
20. Bennett, C. S.; Wong, C. H. *Chem. Soc. Rev.* **2007**, *36*, 1227–1238.
21. Kajihara, Y.; Yamamoto, N.; Miyazaki, T.; Sato, H. *Curr. Med. Chem.* **2005**, *12*, 527–550.
22. Hojo, H.; Matsumoto, Y.; Nakahara, Y.; Ito, E.; Suzuki, Y.; Suzuki, M.; Suzuki, A. *J. Am. Chem. Soc.* **2005**, *127*, 1372–1375.
23. Yamamoto, N.; Tanabe, Y.; Okamoto, R.; Dawson, P. E.; Kajihara, Y. *J. Am. Chem. Soc.* **2008**, *130*, 501–510.
24. Cohen-Anisfeld, S. T.; Lansbury, P. T., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 10531–10537.
25. Roberge, J. Y.; Beebe, X.; Danishefsky, S. J. *Science* **1995**, *269*, 202–204.
26. Miller, J. S.; Dudkin, V. Y.; Lyon, G. J.; Muir, T. W.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2003**, *42*, 431–434.
27. Kaneshiro, C. M.; Michael, K. *Angew. Chem., Int. Ed.* **2006**, *45*, 1077–1081.
28. Mandal, M.; Dudkin, V. Y.; Geng, X.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2004**, *43*, 2557–2561.
29. Geng, X.; Dudkin, V. Y.; Mandal, M.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2004**, *43*, 2562–2565.
30. Dudkin, V. Y.; Miller, J. S.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 736–738.
31. Warren, J. D.; Miller, J. S.; Keding, S. J.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 6576–6578.
32. Witte, K.; Sears, P.; Martin, R.; Wong, C. H. *J. Am. Chem. Soc.* **1997**, *119*, 2114–2118.
33. Yamamoto, K. *J. Biosci. Bioeng.* **2001**, *92*, 493–501.
34. Wang, L. X.; Singh, S.; Ni, J. In *Synthesis of Carbohydrates through Biotechnology*; Wang, P. G., Ichikawa, Y., Eds.; ACS Symposium Series 873; American Chemical Society: Washington, DC, 2004; pp 73–92.
35. Imperiali, B.; Hendrickson, T. L. *Bioorg. Med. Chem.* **1995**, *3*, 1565–1578.
36. Tai, V. W.; Imperiali, B. *J. Org. Chem.* **2001**, *66*, 6217–6228.
37. Fang, X.; Gibbs, B. S.; Coward, J. K. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2701–2706.
38. Gibbs, B. S.; Coward, J. K. *Bioorg. Med. Chem.* **1999**, *7*, 441–447.
39. Dempski, R. E., Jr.; Imperiali, B. *Curr. Opin. Chem. Biol.* **2002**, *6*, 844–850.
40. Glover, K. J.; Weerapana, E.; Numao, S.; Imperiali, B. *Chem. Biol.* **2005**, *12*, 1311–1315.
41. Weerapana, E.; Imperiali, B. *Glycobiology* **2006**, *16*, 91R–101R.
42. Kowarik, M.; Numao, S.; Feldman, M. F.; Schulz, B. L.; Callewaert, N.; Kiermaier, E.; Catrein, I.; Aebi, M. *Science* **2006**, *314*, 1148–1150.
43. 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*, 1244–1246.
44. 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*, 1441–1443.
45. Wildt, S.; Gerngross, T. U. *Nat. Rev. Microbiol.* **2005**, *3*, 119–128.
46. Tarentino, A. L.; Plummer, T. H., jr. *Methods Enzymol.* **1994**, *230*, 44–57.
47. Kato, T.; Fujita, K.; Takeuchi, M.; Kobayashi, K.; Natsuka, S.; Ikura, K.; Kumagai, H.; Yamamoto, K. *Glycobiology* **2002**, *12*, 581–587.
48. Suzuki, T.; Yano, K.; Sugimoto, S.; Kitajima, K.; Lennarz, W. J.; Inoue, S.; Inoue, Y.; Emori, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9691–9696.
49. Yamamoto, K.; Kadowaki, S.; Watanabe, J.; Kumagai, H. *Biochem. Biophys. Res. Commun.* **1994**, *203*, 244–252.
50. Takegawa, K.; Yamaguchi, S.; Kondo, A.; Iwamoto, H.; Nakoshi, M.; Kato, I.; Iwahara, S. *Biochem. Int.* **1991**, *24*, 849–855.
51. Takegawa, K.; Yamaguchi, S.; Kondo, A.; Kato, I.; Iwahara, S. *Biochem. Int.* **1991**, *25*, 829–835.
52. Fujita, K.; Takami, H.; Yamamoto, K.; Takegawa, K. *Biosci., Biotechnol., Biochem.* **2004**, *68*, 1059–1966.
53. Takegawa, K.; Tabuchi, M.; Yamaguchi, S.; Kondo, A.; Kato, I.; Iwahara, S. *J. Biol. Chem.* **1995**, *270*, 3094–3099.

54. Haneda, K.; Inazu, T.; Yamamoto, K.; Kumagai, H.; Nakahara, Y.; Kobata, A. *Carbohydr. Res.* **1996**, *292*, 61–70.
55. Wang, L. X.; Fan, J. Q.; Lee, Y. C. *Tetrahedron Lett.* **1996**, *37*, 1975–1978.
56. 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*, 11137–11146.
57. Mizuno, M.; Haneda, K.; Iguchi, R.; Muramoto, I.; Kawakami, T.; Aimoto, S.; Yamamoto, K.; Inazu, T. *J. Am. Chem. Soc.* **1999**, *121*, 284–290.
58. O'Connor, S. E.; Pohlmann, J.; Imperiali, B.; Saskiawan, I.; Yamamoto, K. *J. Am. Chem. Soc.* **2001**, *123*, 6187–6188.
59. 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*, 242–248.
60. Singh, S.; Ni, J.; Wang, L. X. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 327–330.
61. Wang, L. X.; Song, H.; Liu, S.; Lu, H.; Jiang, S.; Ni, J.; Li, H. *ChemBioChem* **2005**, *6*, 1068–1074.
62. Li, H.; Singh, S.; Zeng, Y.; Song, H.; Wang, L. X. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 895–898.
63. Akaike, E.; Tsutsumida, M.; Osumi, K.; Fujita, M.; Yamanoi, T.; Yamamoto, K.; Fujita, K. *Carbohydr. Res.* **2004**, *339*, 719–722.
64. Li, B.; Zeng, Y.; Hauser, S.; Song, H.; Wang, L. X. *J. Am. Chem. Soc.* **2005**, *127*, 9692–9693.
65. Li, H.; Li, B.; Song, H.; Breydo, L.; Baskakov, I. V.; Wang, L. X. *J. Org. Chem.* **2005**, *70*, 9990–9996.
66. Zeng, Y.; Wang, J.; Li, B.; Hauser, S.; Li, H.; Wang, L. X. *Chem. Eur. J.* **2006**, *12*, 3355–3364.
67. Li, B.; Song, H.; Hauser, S.; Wang, L. X. *Org. Lett.* **2006**, *8*, 3081–3084.
68. Rising, T. W.; Claridge, T. D.; Davies, N.; Gamblin, D. P.; Moir, J. W.; Fairbanks, A. J. *Carbohydr. Res.* **2006**, *341*, 1574–1596.
69. Rising, T. W.; Claridge, T. D.; Moir, J. W.; Fairbanks, A. J. *ChemBioChem* **2006**, *7*, 1177–1180.
70. Mark, B. L.; Vocadlo, D. J.; Knapp, S.; Triggs-Raine, B. L.; Withers, S. G.; James, M. N. *J. Biol. Chem.* **2001**, *276*, 10330–10337.
71. Williams, S. J.; Mark, B. L.; Vocadlo, D. J.; James, M. N.; Withers, S. G. *J. Biol. Chem.* **2002**, *277*, 40055–40065.
72. Brameld, K. A.; Shrader, W. D.; Imperiali, B.; Goddard, W. A., 3rd *J. Mol. Biol.* **1998**, *280*, 913–923.
73. Terwisscha van Scheltinga, A. C.; Armand, S.; Kalk, K. H.; Isogai, A.; Henrissat, B.; Dijkstra, B. W. *Biochemistry* **1995**, *34*, 15619–15623.
74. Tews, I.; Terwisscha van Scheltinga, A. C.; Perrakis, A.; Wilson, K. S.; Dijkstra, B. W. *J. Am. Chem. Soc.* **1997**, *119*, 7954–7959.
75. Kobayashi, S.; Kiyosada, T.; Shoda, S. *J. Am. Chem. Soc.* **1996**, *118*, 13113–13114.
76. Kobayashi, S.; Morii, H.; Itoh, R.; Kimura, S.; Ohmae, M. *J. Am. Chem. Soc.* **2001**, *123*, 11825–11826.
77. Kobayashi, S.; Fujikawa, S.; Ohmae, M. *J. Am. Chem. Soc.* **2003**, *125*, 14357–14369.
78. Ochiai, H.; Ohmae, M.; Kobayashi, S. *Carbohydr. Res.* **2004**, *339*, 2769–2788.
79. Kobayashi, S.; Ohmae, M.; Ochiai, H.; Fujikawa, S. *Chem. Eur. J.* **2006**, *12*, 5962–5971.
80. Ochiai, H.; Fujikawa, S. I.; Ohmae, M.; Kobayashi, S. *Biomacromolecules* **2007**.
81. Fujita, M.; Shoda, S.; Haneda, K.; Inazu, T.; Takegawa, K.; Yamamoto, K. *Biochim. Biophys. Acta* **2001**, *1528*, 9–14.
82. Perugino, G.; Trincone, A.; Rossi, M.; Moracci, M. *Trends Biotechnol.* **2004**, *22*, 31–37.
83. Hancock, S. M.; Vaughan, M. D.; Withers, S. G. *Curr. Opin. Chem. Biol.* **2006**, *10*, 509–519.
84. Faijes, M.; Planas, A. *Carbohydr. Res.* **2007**, *342*, 1581–1594.
85. Umekawa, M.; Huang, W.; Li, B.; Fujita, K.; Ashida, H.; Wang, L. X.; Yamamoto, K. *J. Biol. Chem.* **2008**, *283*, 4469–4479.
86. Wang, L. X.; Ni, J.; Singh, S.; Li, H. *Chem. Biol.* **2004**, *11*, 127–134.
87. Fujita, K.; Yamamoto, K. *Biochim. Biophys. Acta* **2006**, *1760*, 1631–1635.
88. Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
89. Schwarzer, D.; Cole, P. A. *Curr. Opin. Chem. Biol.* **2005**, *9*, 561–569.
90. Wang, L.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2004**, *44*, 34–66.
91. Zhang, Z.; Gildersleeve, J.; Yang, Y. Y.; Xu, R.; Loo, J. A.; Uryu, S.; Wong, C. H.; Schultz, P. G. *Science* **2004**, *303*, 371–373.