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).