Chemoenzymatic Synthesis of Glycoprotein Using the Peptide Ligation Chemistry

Yuya Asahina
Institute for Protein Research, Osaka University, 3–2 Yamadaoka, Suita, Osaka 565–0871, Japan
FAX: +81–6–6879–8603, E-mail: asahina@protein.osaka-u.ac.jp
(Received on January 8, 2015, accepted on February 5, 2015)

Key Words: glycoprotein, glycopeptide, thioester method, chemoenzymatic synthesis, Endo-M

Abstract
The chemoenzymatic method using glycosynthase is one of potent strategies for the synthesis of glycoproteins having a homogeneous glycan. In this method, the efficient preparation of the N-acetylglucosaminylated protein is of great importance. Thus, for producing this compound with a high efficiency, we developed the sequential thioester method and the fast preparation method of the N-acetylglucosaminyl peptide. These methods realized the efficient preparation of glycoproteins having GlcNAc. Finally, homogeneous glycoproteins, the Tim-3 Ig domain and emmprin were successfully synthesized by the transglycosylation using glycosynthase.

A. Introduction
A glycoprotein carbohydrate plays many significant roles, such as cell growth, adhesion and cancerization. However, due to the micro-heterogeneity, it is difficult to obtain glycoproteins in a pure glycoform from a natural source. Therefore, the relationship between glycan and its functions is clearly not known. Chemical synthesis of proteins has emerged as one of the potent strategies to obtain homogeneous glycoproteins. As one of these strategies, the chemoenzymatic method using endo-β-N-acetyl-glucosaminidase (Endo-M) (1) has been reported by Inazu and Yamamoto et al. (2).

In this first report, although the problem was the low efficiency of the transglycosylation by the competitive hydrolysis reaction, it was improved by the use of sugar oxazoline (3) and the glycosynthase (Endo-M mutant) (4), which realized to obtain the desired product in high yield. This method provides an efficient synthesis of various glycoproteins carrying the pure glycoform, if a corresponding N-acetylglycosaminylated protein can be prepared. Hence, this method can accommodate a variety of natural glycoproteins having a micro-heterogeneity (5). Recently, the efficient synthesis methods of glycosyl donors (6) and oxazoline (7) had been developed. In addition, the glycosynthase had been a commercially available enzyme. Thus, to realize the efficient synthesis of glycoproteins using a chemoenzymatic method, the preparation of the N-acetylglucosaminylated glycoprotein, which is the substrate for the transglycosylation by glycosynthase, is now important (Fig. 1). This key intermediate can be synthesized in two steps, which are the preparation step of the peptide having GlcNAc and the ligation step for leading peptide segments to a polypeptide. In the former step, two complicated deprotection treatments were needed at the end of the solid-phase peptide synthesis (SPPS). In the latter step, when multiple peptide condensations are required, the yield and synthetic efficiency became moderate by purification of the intermediate segment. To overcome these problems, we developed the efficient preparation of the N-acetylglucosaminylated protein for the chemoenzymatic synthesis.

B. Ag⁺-promoted and Ag⁺-free Thioester Method
SPPS is routinely used for a glycopeptide synthesis to assemble efficiently the peptide chain. However, in the preparation of the long polypeptide, it is difficult to isolate the desired product from byproducts generated during chain elongation. Therefore, a peptide ligation method is necessary for the glycoprotein synthesis.
more than 50 amino acid residues. One of peptide ligation methods, the thioester method (8a), is shown in Fig. 2. In this method, the N-terminal segment having an alkyl thioester is activated in the presence of Ag⁺ and the ligation proceeds by the direct aminolysis of the C-terminal segment having a free amine to form the entire polypeptide. Furthermore, the Ag⁺-free thioester method, that does not require the Ag⁺ activation by replacing the alkyl thioester with an aryl thioester, was newly reported (8b). Although these methods require the protections of amino and thiol groups on the peptide side chain, the condensation can be performed at any site. This is in contrast the native chemical ligation method (9) that is limited the condensation site. Thus, this method realized a flexible synthetic design which leads to various (glyco)protein syntheses.

In a previous study, the condensations of three segments were carried out from the C to N-terminal direction to obtain the desired polypeptide. However, in this method, each condensation was performed in stepwise manner, since the intermediate segment must be isolated at the end of the ligation to remove an excess amount of piperidine used for the Fmoc deprotection. The purification step causes a significant reduction of the synthetic efficiency and yield due to the irreversible adsorption during reverse-phase chromatography. Thus, to overcome this problem, we developed the novel one-pot sequential thioester method, which excludes the isolation of the intermediate segment by using the difference in the alkyl and aryl thioester reactivities.

C. Synthesis of Tim-3 Ig Domain by the Sequential Thioester Method

Glycoprotein Tim-3 is located on the Th-1 cell surface. This extracellular Ig domain has an N-glycan that is recognized by galectin-9, which is suggested to induce the apoptosis signaling of the Th-1 cell (10). We attempted to synthesize this Ig domain by the sequential one-pot thioester method (Fig. 3). The entire sequence was divided into three segments, and each segment was prepared by the SPPS. Segments 1 and 2 were prepared as the aryl thioester and alkyl thioester, respectively. On the other hand, to segment 3 being highly insoluble, the O-acylisopeptide structure (11), which was developed by Sohma et al. and increases the peptide solubility by the reversible depsipeptide form, was introduced. After preparation of these segments, the condensations of segments 1 and 2 were first performed under Ag⁺-free thioester conditions. This reaction selectively activated the aryl thioester 1 to give the intermediate 4, while maintaining its alkyl thioester. In the next step, by the addition of segment 3 and AgCl, the second condensation was successively carried out. This Ag⁺-promoted condition proceeded the reaction through the activation of the alkyl thioester 4 and successfully produced the desired Ig domain polypeptide 5.

The isolated yield of 5 was 63%, which is a yield comparable to
by treating with ethylenediamine followed by acetylation. The obtained C-3 alcohol of was synthesized as shown Fig. 4 (right). To the labile groups and applied it to the synthesis of emmprin (22–104).}

In the next study, we aimed to investigate the efficient preparation of the N-acetylglycosaminyl peptide segment. In previous glycoprotein syntheses, we used the GlcNAc-Asn protected by a benzyl (Bn) group (6b, 13), which has a high stability and can be removed under acidic conditions without any damage to the peptide and the thioester. However, due to this stability, two de-protection treatments were required for the Bn-protected 8 at the end of SPPS to obtain the free glycopeptide as shown Fig. 4 (left); 1) TFA treatment for the deprotection of the peptide protecting groups, 2) Low-TfOH (14) treatment for removal of the Bn groups on the GlcNAc moiety. To avoid this complicated multiple deprotection, we newly designed the GlcNAc-Asn 9 protected by TFA-labile groups and applied it to the synthesis of emmprin (22–104). The derivative 9 was synthesized as shown Fig. 4 (right). To the C-3 alcohol of 10, the 4-methoxyphenylmethyl (MPM) group was introduced. The obtained 11 was converted to the acetoamide 12 by treating with ethylenediamine followed by acetylation. The condensation of the azide sugar 12 and amino acid 13 was carried out by the modified Staudinger reaction (15) in the presence of n-Bu3P and HOOBt in aqueous THF and efficiently produced the GlcNAc-Asn 14 in 83% yield. However, due to the poor solubility of 14, the deprotection of the phenacyl (Pac) group by the Zn/AcOH failed. This poor solubility might be caused by intermolecular aggregation, which was mediated by the hydrogen bonding at the C-1 and/or C-2 acetamide. Thus, to the nitrogen atom of C-2, 2,4,6-trimethoxybenzyl (TMB) group was introduced to block the hydrogen bonding. The TMB-substituted 16, which was derivatized from the azide 11, was also condensed with the amino acid 13. As a result, the solubility of the GlcNAc-Asn 17 drastically increased in CH2Cl2/AcOH compared to 14. Finally, the deprotection was carried out using Zn/AcOH and successfully gave the desired glycoamino acid 9 protected by TFA-labile groups.

### D. Development of GlcNAc-Asn Derivative Protected by TFA-labile Protecting Groups

In the next study, we applied GlcNAc-Asn 9 to the synthesis of the emmprin 1st Ig domain. This glycoprotein is related to cancer metastasis and possesses a complex-type N-glycan at Asn44 (16). The synthetic route of emmprin is shown Fig. 5. The overall sequence was divided into three segments, 18, 19 and 20, that were prepared by the Fmoc-SPPS. The condensation of these segments was attempted by the sequential thioester method to obtain the entire polypeptide in one-pot. After the deprotection and folding, the transglycosylation by glycosynthase was carried out onto the glycopeptide 21 to yield emmprin carrying asialo- and disialo-type N-glycan, 22 and 23. The glycopeptide 19 was derivatized as shown in Fig. 6 (left). The Rink amide resin having NAC (17), which was the N-to-S acyl shift device to prepare a peptide thioester after SPPS, was subjected to peptide elongation by the Fmoc method.

### E. Synthesis of Emmprin 1st Ig Domain Using a Novel GlcNAc-Asn 9 Derivative

In the next study, we applied GlcNAc-Asn 9 to the synthesis of the emmprin 1st Ig domain. This glycoprotein is related to cancer metastasis and possesses a complex-type N-glycan at Asn44 (16). The synthetic route of emmprin is shown Fig. 5. The overall sequence was divided into three segments, 18, 19 and 20, that were prepared by the Fmoc-SPPS. The condensation of these segments was attempted by the sequential thioester method to obtain the entire polypeptide in one-pot. After the deprotection and folding, the transglycosylation by glycosynthase was carried out onto the glycoprotein 21 to yield emmprin carrying asialo- and disialo-type N-glycan, 22 and 23. The glycopeptide 19 was derivatized as shown in Fig. 6 (left). The Rink amide resin having NAC (17), which was the N-to-S acyl shift device to prepare a peptide thioester after SPPS, was subjected to peptide elongation by the Fmoc method.
At the Asn^{44} glycosylation site, the GlcNAc-Asn 9 was manually introduced using the HOBt-DIC method. After the peptide chain assembly, the glycopeptidyl resin was treated with the usual TFA cocktail for 2 h at room temperature. As shown in Fig. 6 (right), the acid-labile groups on the GlcNAc moiety were removed with the peptide protecting groups by the one-step deprotection showing the usefulness of 9. By the thioester exchange with 3-mercaptopropionic acid, the crude glycopeptide was converted to the thioester 19 in a good yield. After preparation of the other segment, a threesegment condensation was performed by the sequential thioester method. After the aryl thioester 18 and alkyl thioester 19 were reacted under Ag^{+}-free conditions, the second condensation with
the C-terminal segment was then carried out under Ag⁺-promoted conditions to give the desired polypeptide. After the deprotection and the disulfide bond formation, the folded Ig domain 21 was obtained. Finally, the transglycosylation with oxazoline, which was derivatized by CDCMI (7b), successfully gave emmprin carrying the asialo- and disialo-type sugars, 22 and 23 (Fig. 7) (18). A structural analysis and the MMP activity of these glycoproteins will be investigated.

F. At the End

In this study, we developed an efficient and fast preparation method of a N-acetylglucosaminylated protein, which is the substrate of glycosynthase, for the chemoenzymatic method. The sequential thioester method enhanced the ligation efficiency by the one-pot condensation of three segments and produced the glycoproteins on a mg scale. The novel GlcNAc-Asn 9 protected by TFA-labile groups realized the fast preparation of the glycopeptide carrying GlcNAc by the one-step TFA treatment. These methods will be used for various glycoprotein syntheses to discover their function and nature. Finally, I am deeply grateful to Prof. Hironobu Hojo (Inst. Protein Res., Osaka Univ.), who was my supervisor and kindly raised my scientific philosophy. I would like to thank retired Prof. Yoshiaki Nakahara (Dept. Appl. Biochem., Tokai Univ.), who was my mentor of the glycoscience. I must also thank Chief Prof. Akemi Suzuki (Inst. Glycosci., Tokai Univ.) for his aids and kind advice.

References


Information of the Author

Yuya Asahina is assistant professor of Institute for Protein Research, Osaka University. He received a bachelor degree (2009) under the supervision of Professor Yoshiaki Nakahara, a master degree (2011) and a doctoral degree (2014) under the supervision of Professor Hironobu Hojo in Tokai University. His researches are synthesis of glycan, glycopeptide and glycoprotein including peptide ligation chemistry.