Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 15 (2005) 895–898
Bioorganic & Medicinal Chemistry Letters

Chemoenzymatic synthesis of CD52 glycoproteins carrying native N-glycans

Hengguang Li, Suddham Singh, Ying Zeng, Haijing Song and Lai-Xi Wang*

Institute of Human Virology, University of Maryland Biotechnology Institute, University of Maryland, 725 West Lombard Street, Baltimore, MD 21201, USA

Received 1 November 2004; revised 20 December 2004; accepted 21 December 2004
Available online 19 January 2005

Abstract—A facile synthesis of homogeneous CD52 glycoproteins carrying native N-glycans was achieved using an endolycosidase-catalyzed oligosaccharide transfer as the key step. The synthesis consists of two steps: the solid phase synthesis of GlcNAc-CD52 and the transfer of a high-mannose type or complex type N-glycan from $Man_9GlcNAc_2$ Asn or a sialglycopeptide to the GlcNAc-CD52, under the catalysis of the endo-β-*N*-acetylglucosaminidases from *Arthrobacter* (Endo-A) and *Mucor hiemalis* (Endo-M), respectively.

Glycosylation is one of the most common posttranslational modifications of proteins. The oligosaccharide components of glycoproteins have been implicated to play important roles in modulating protein's structure and function, including stability, folding, immunogenicity, and various cellular activities.[1–4] However, a clear understanding of the structure–function relationships of glycoproteins is often hampered by the difficulties in obtaining homogeneous glycopeptides or glycoproteins. Natural glycoproteins usually appear as a mixture of glycoforms that have the identical polypeptide backbone but differ in the structure of the pendant oligosaccharide chains. CD52 is a GPI-anchored cell surface glycoprotein that is expressed on almost all human lymphocytes and sperm cells.[5–7] CD52 consists of only 12 amino acid residues but carries a large complex N-glycan at the Asn-3 residue. The protein core of the human lymphocyte CD52 and the sperm CD52 is identical, but the two antigens exhibit distinct functions in biological system. For example, lymphocyte CD52 is involved in human immune response, while sperm CD52 plays an essential role in sperm–egg interaction.[5] Structural analysis revealed that lymphocyte CD52 and sperm CD52 both carry highly heterogeneous N-glycans and are different in the composition of the N-glycans.[6,7] Together with the varied modifications on the GPI-anchor portion, the different glycosylation may contribute to the distinct biological functions of the two types of CD52. As an effort to provide homogeneous materials for further biological studies, we report here a facile chemoenzymatic synthesis of CD52 glycoproteins carrying a native, high-mannose type or complex type N-glycan (Fig. 1).

Synthesis of homogeneous glycopeptides bearing large native N-glycans remains a challenging task.[8,9] CD52 glycopeptides containing a tri- and penta-saccharide moiety was previously prepared by chemical synthesis.[10,11] Recently, a bi-antennary N-linked glycoform of CD52 carrying an undecasaccharide mimic was constructed using a convergent chemoselective ligation method.[12] In order to synthesize homogeneous glycoforms of CD52 carrying native N-glycans, we applied a chemoenzymatic approach using *endo*-β-*N*-acetylglucosaminidases as the key enzyme for oligosaccharide transfer. Some *endo*-β-*N*-acetylglucosaminidases possess significant transglycosylation activity, which has been explored for oligosaccharide and glycopeptide synthesis.[13,14] For example, Endo-A from *Arthrobacter* can transfer a high-mannose type N-glycan to a GlcNAc-containing peptide,[15,16] while Endo-M isolated from *Mucor hiemalis* prefers complex type N-glycan in transglycosylation.[17–19] Therefore, the distinct substrate specificity of the two endoglycosidases in transglycosylation

Keywords: Glycopeptide; Glycoprotein; CD52; Chemoenzymatic synthesis.
* Corresponding author. Tel.: +1 410 706 4982; fax: +1 410 706 5068; e-mail: wangx@umbi.umd.edu

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.12.066

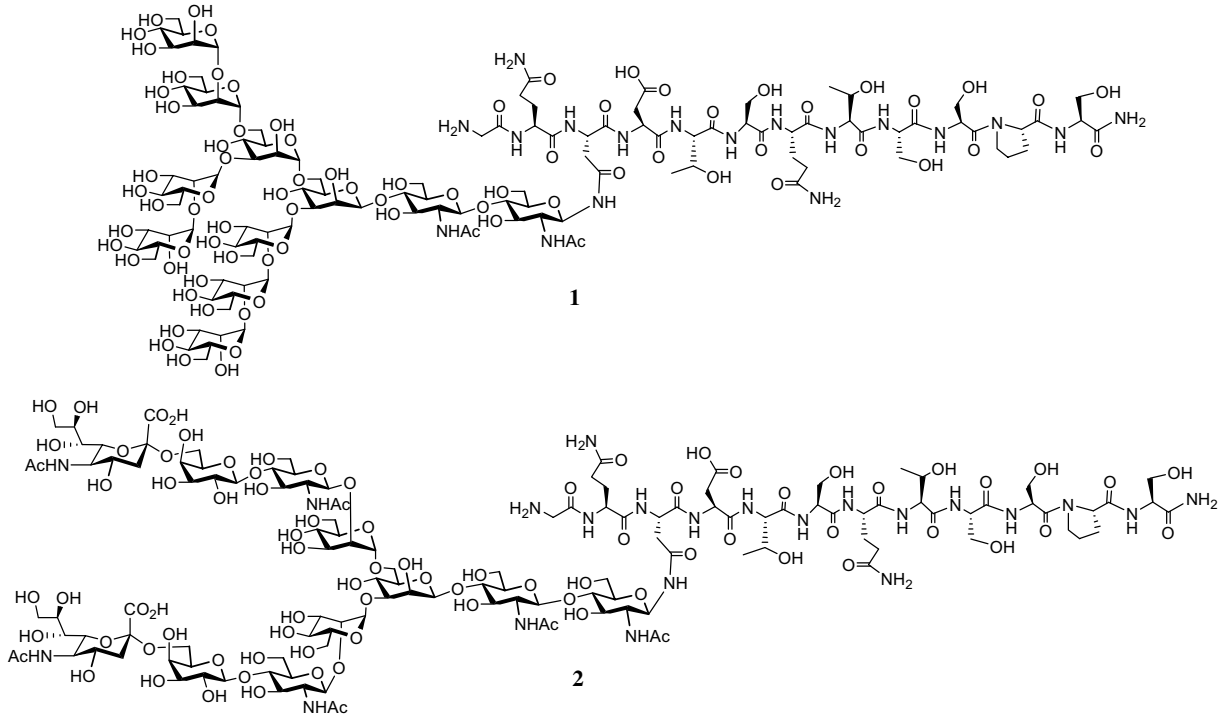
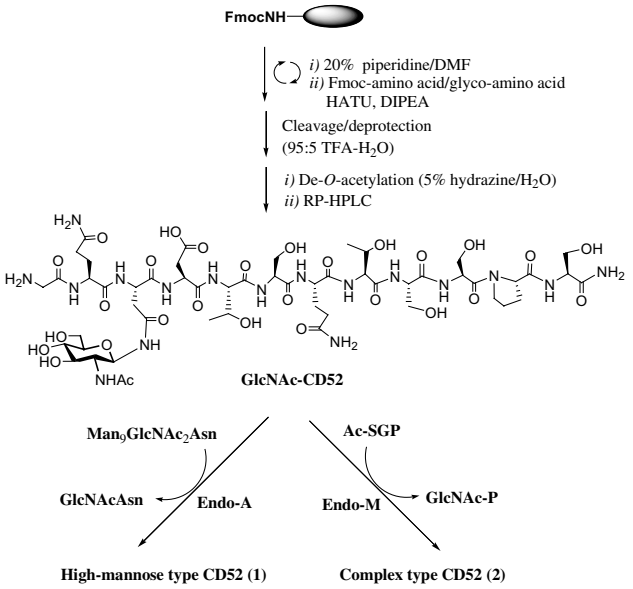


Figure 1. Structures of the synthetic high-mannose type (**1**) and complex type CD52 (**2**).

will allow the construction of different glycoforms of CD52. Our synthesis started with the preparation of the CD52 peptide containing a GlcNAc moiety at the Asn-3 N-glycosylation site. The GlcNAc-CD52 was synthesized on a PAL-PEG-PS resin using standard Fmoc-amino acid derivatives and Fmoc-(Ac_3GlcNAc)Asn-OH as the building blocks and HATU/DIPEA as the coupling reagents. The peptide was retrieved from the resin with simultaneous side-chain deprotection by treatment with 95% TFA. After de-*O*-acetylation with 5% aqueous hydrazine, the crude peptide was purified by RP-HPLC to give GlcNAc-CD52, the identity of which was confirmed by ESI-MS and ^{1}H NMR.[20]

Transglycosylation of a high-mannose type oligosaccharide chain to the GlcNAc-CD52 was achieved using $Man_9GlcNAc_2$ Asn[21] as the oligosaccharide donor under the catalysis of Endo-A (Scheme 1). The enzymatic reaction was performed in an acetate buffer containing 25% acetone and monitored by HPLC. The transglycosylation product Man_9-CD52 (**1**) was isolated by RP-HPLC and characterized by ESI-MS and ^{1}H NMR.[22] For the synthesis of the complex type CD52 antigen (**2**), we isolated a bi-antennary sialylglycopeptide (SGP), H-Lys-Val-Ala-Asn[(NeuAc-Gal-GlcNAc-Man)$_2$-Man-GlcNAc$_2$]-Lys-Thr-OH, from hen's egg yolk according to the literature.[23] However, initial attempt to use the SGP as oligosaccharide donor for the Endo-M catalyzed transglycosylation encountered a problem for both monitoring of the reaction and subsequent separation of the desired product, because the SGP and the desired transglycosylation product had similar retention times under various RP-HPLC conditions and could not be separated. To solve the problem, we performed selective *N*-acetylation (Ac_2O–aqueous



Scheme 1. Chemoenzymatic synthesis of the glycoforms of CD52.

$NaHCO_3$) at the three free amino groups in the sialylglycopeptide SGP to provide the *N*-acetylated derivative Ac-SGP, **Ac**-Lys(**Ac**)-Val-Ala-Asn[(NeuAc-Gal-GlcNAc-Man)$_2$-Man-GlcNAc$_2$]-Lys(**Ac**)-Thr-OH [ESI-MS: calcd. for $C_{118}H_{195}N_{15}O_{73}$, 2991.87; Found: 998.18 $(M+3H)^{3+}$, 1496.90 $(M+2H)^{2+}$], which demonstrated enhanced hydrophobicity. We found that the modified glycopeptide Ac-SGP was able to serve as a substrate of Endo-M for transglycosylation but was eluted more slowly than SGP, GlcNAc-CD52, and the product (**2**) under RP-HPLC conditions without interfering with the product separation. Thus, the