Tetrahedron 74 (2018) 2891–2903

Contents lists available at ScienceDirect

Tetrahedron

journal homepage: www.elsevier.com/locate/tet

Simplified beta-glycosylation of peptides

Yonglian Zhang, Spencer Knapp*

Department of Chemistry & Chemical Biology, Rutgers The State University of New Jersey, 610 Taylor Rd., Piscataway, NJ, USA

ARTICLE INFO

Article history:
Received 16 March 2018
Received in revised form
24 April 2018
Accepted 25 April 2018
Available online 27 April 2018

Keywords:
Lewis acid
Hydrogenolysis
Trichloroacetamide
Glycoprotein
Glycopeptide

ABSTRACT

A simple and effective activating system for *S*-phenyl thioglycosides, namely N-iodosuccinimide and catalytic copper(I) triflate, promotes beta-O-glycosylation at the serine and threonine hydroxyls of "mono-," di-, and tripeptides. The same activator combination promotes carboxamide beta-N-glycosylation of asparagine-containing mono-, di, and tri-peptides, as well as a nucleoside carboxamide and a sulfonamide. An important feature of the copper(I) triflate method is that undesired amide *O*-glycosylation is largely circumvented. For both sets of biologically important acceptor sites (HO— and —CONH$_2$), a *beta*-GlcNAc-equivalent donor is demonstrated to provide the linkages efficiently. Streamlined deprotection sequences have been developed based on global hydrogenolysis that lead smoothly to the parent glycopeptides. The core glycopeptide region for biological protein *N*-glycosylation, represented by N^4-(β-*N*-acetyl-D-2-glucosaminyl)-Asp-Gly-Thr-OH, has been prepared in solution, purified, and characterized as the fully deprotected (mono)glycosylated tripeptide.

1. Introduction

Enzyme mediated O- and N-glycosylation of proteins, particularly at the side chain residues of serine, threonine, and asparagine, is an important biological process that determines the properties, function, and fate of the resulting glycoproteins.[1–4] To take a recent example: cryo-EM and other evidence indicates that single N-glycosylated Asn154 of the Zika virus glycoprotein shell may function to enable attachment of the virus to host cells.[5–7] The analogous but simpler glycosylated peptides and amino acids, including the blood group antigens,[8,9] have their own distinctive biological roles.[10,11] The chemical preparation of N-glycopeptides[12–14] has developed along two parallel lines, with notable successes for both: (1) N-acylation of glycosylamines with activated aspartic acid derivatives,[15,16] and (2) direct O- and N-glycosylation of protected peptides with activated glycosyl donors.[17–21] In either case, elaboration of the glycosylated amino acid or small glycopeptide into a more complex glycopeptide can often be achieved by well precedented solution or resin-supported peptide coupling methods.[22–28] The direct glycosylation approach, while more versatile in principle, and more similar to the biological process itself, is limited by the difficulty of carrying out a chemical glycosylation on a complex acceptor with multiple Lewis basic sites. These can include the various amide and carbonyl-bearing protecting groups, which can undergo wasteful carbonyl-O-glycosylations[29–31] in competition with reaction at the desired acceptor site, namely serine/threonine —OH or asparagine —CONH$_2$. Thus, the chemical yields of these glycosylation reactions can be unacceptably low. The development of a simple and efficient process for O- and N- glycosylation of peptides that need not be optimized for each new specific case would benefit multistep preparations, especially those limited by availability of material, that cannot themselves be taken through extensive optimization. Part of this optimization should include mild and efficient deprotection protocols.

In addressing an analogous situation in nucleoside glycosylation, namely, how can glycosylation be carried out at an acceptor site that is not necessarily the most reactive one, we recently found a practical solution.[32] This features identifying an activator/Lewis acid combination that promotes reversibility of the initial glycoslation at a competing acceptor site, such as an amide carbonyl, and allows a slower but more permanent glycosylation at the desired site. In this paper we apply similar reasoning to develop procedures for peptide O- or N-glycosylation, and further sharpen the process by streamlining the deprotection steps.

2. Results and discussion

2.1. Beta-glycosylation of serine hydroxyl

Thioglycoside donors have advantages for glycosylation that

* Corresponding author.
E-mail address: spencer.knapp@rutgers.edu (S. Knapp).

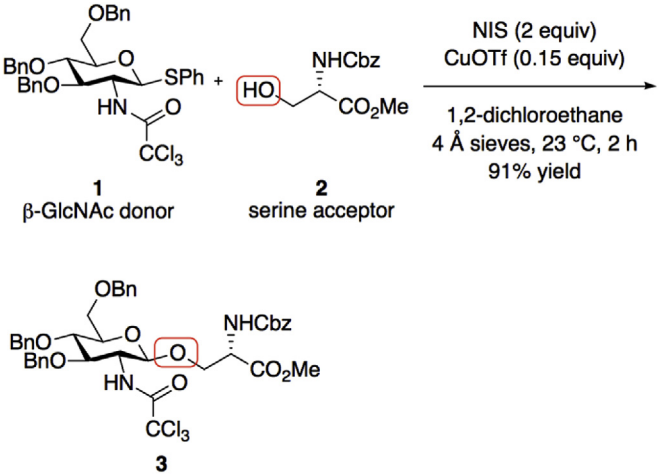
https://doi.org/10.1016/j.tet.2018.04.082
0040-4020/

include their simplicity and ease of preparation, their chemical stability, and the versatility of the various activation methods.[33–36] The beta-GlcNAc donor **1**, with phenylthio as the activatable leaving group, and trichloroacetamido as the participating group at C-2,[37] was evaluated as a glycosylation donor with commercial N-Cbz serine methyl ester (**2**) as the protected amino acid ("monopeptide") acceptor. A qualitative screen of acids (triflic acid, TMS-triflate) and metal triflates[38–40] in combination with N-iodosuccinimide (NIS) led to the selection of copper(I) triflate, commercially available as a 2:1 toluene complex, as an effective catalytic promoter (91% yield, Scheme 1). This yield compares favorably with the best literature methods for serine beta O-GlcNAc-ylation.[41–46] The best promoter for nucleoside O-glycosylation, namely, indium(III) triflate,[32] was not as effective for the peptide example. Although we have no direct evidence, it is possible that Cu(I) is better than the other promoters because it effectively isomerizes mis-glycosylated imidate-type byproducts.

By application of the same reaction conditions, with **1** and also with the analogous *galacto* donor[47] **4**, to the glycosylation of several other protected serine and threonine acceptors, the glycopeptides in Fig. 1 were prepared in the yields shown. H-1 NMR spectroscopic analysis (*J* at the anomeric H) indicates that products have exclusively the *beta* glycosyl stereochemistry. The original acceptor atoms are highlighted with red boxes.

Similarly, a dipeptide and a tripeptide acceptor were converted to glycopeptides **11**–**14** by using donors **1** and **4** (Fig. 2). Yields are somewhat reduced compared with the monopeptides, but are nevertheless acceptable for acceptors with this number of competing Lewis basic sites. The acceptor atoms are highlighted with red boxes.

The use of hydrogenolyzable protecting and precursor groups on both the donor and acceptor moieties enables a one step deprotection protocol, as shown in Scheme 2. No fewer than eight individual hydrogenolytic steps (as indicated by the red arrows) occur during the palladium hydroxide mediated deprotection of **6**, which leads to the hydrochloride of the GlcNAc serine derivative[48] **15** in 91% overall yield. Examination of the H-1 and C-13 NMR spectra indicates that hydrogenolysis is complete. There are, for example, no signals attributable to a mono-benzyl or a chloroacetamido byproduct. Likewise, no stereoisomers, such as might have been produced had partial epimerization occurred at the serine *alpha* carbon, are observed. The same efficient procedure provides glycopeptides **16**–**19** from the respective glycosylation products (Scheme 2). Because of the presence of aqueous HCl during



Scheme 1. Prototypical serine O-glycosylation.

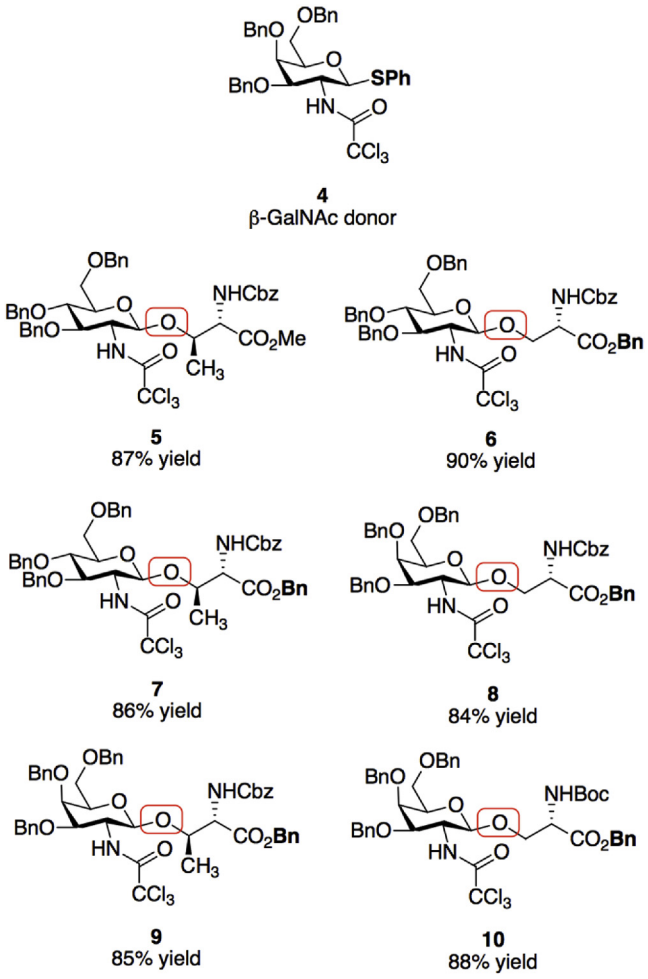


Fig. 1. O-Glycosylation of serine hydroxyls.

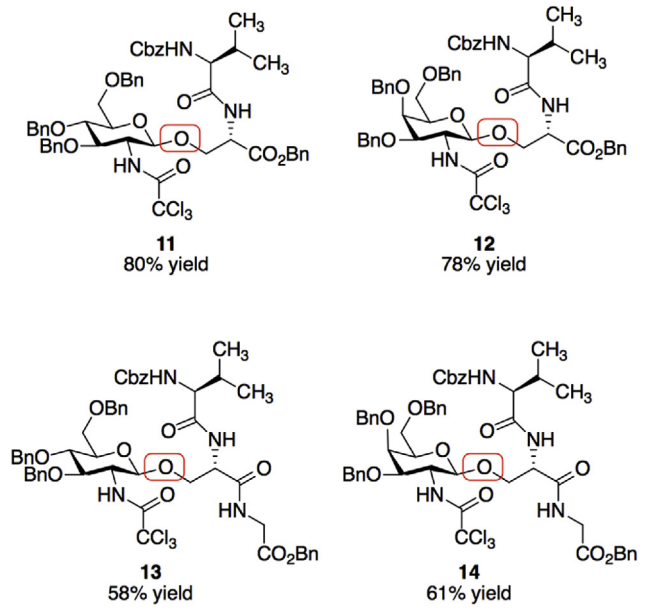


Fig. 2. Dipeptide and tripeptide serine O-glycosylations.