Published in final edited form as:

Org Lett. 2018 August 17; 20(16): 5032–5035. doi:10.1021/acs.orglett.8b02182.

Simple and Efficient Preparation of O- and S-GlcNAcylated Amino Acids through InBr_3 -Catalyzed Synthesis of β -N-Acetylglycosides from Commercially Available Reagents

Cesar A. De Leon[†], Geoffrey Lang[†], Marcos I. Saavedra[†], and Matthew R. Pratt^{*,†,‡}

[†]Department of Chemistry, University of Southern California, Los Angeles, California 90089, United States

[‡]Department of Molecular and Computational Biology, University of Southern California, Los Angeles, California 90089, United States

Abstract

The facile synthesis of serine, threonine, and cysteine β -glycosides using commercially available peracetylated β -N-acetylglucosamine (β -Ac₄GlcNAc) and catalytic amounts of indium bromide (InBr_3) is described. This method involves only inexpensive reagents that require no further modification or special handling. The reagents are simply mixed, dissolved, and refluxed to afford the GlcNAcylated amino acids in great yields (70–80%). This operationally simple procedure should facilitate the study of O-GlcNAcylation without necessitating expertise in synthetic carbohydrate chemistry.

Graphical Abstract

O-GlcNAcylation is a dynamic post-translational modification (PTM) of intracellular proteins in which the single monosaccharide *N*-acetylglucosamine (GlcNAc) is β -linked to serine or threonine residues of substrate proteins.^{1,2} This highly abundant PTM is critical for

^{*}**Corresponding Author:** Address: University of Southern California, 840 Downey Way, LJS250, Los Angeles, CA 90089. Phone: (213) 740-3014. matthew.pratt@usc.edu.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.8b02182. Full experimental procedures and characterization data for all compounds, including NMR spectra (PDF)

Notes

The authors declare no competing financial interest.

several biological processes, and its misregulation has important implications in mammalian development, survival, and disease. Currently, the only approach for deciphering the biochemical and biophysical function of site-specific O-GlcNAcylation is through semi-synthetic preparation of the glycosylated target protein, which requires site-specifically modified glycopeptides.³ Given the heterogeneity that arises from in vitro enzymatic glycosylation⁴ and the racemization observed with post-translational mutagenesis,^{5,6} these glycopeptides must be prepared using solid-phase peptide synthesis (SPPS). Additionally, the key OGlcNAcylated amino acids required for SPPS must also be synthesized, rendering site-specific modification of proteins via chemical ligation methods out of reach for many nonspecialists.

The majority of published routes for O-GlcNAcylated amino acids require the prior synthetic preparation of complex glycosyl halides, trichloroacetimidates, acetates, or thioglycosides that may contain special protecting groups, such as trichloroethoxycarbonyl (Troc), at the C-2 amino group of glucosamine (Scheme 1).^{7–12} These protecting groups are introduced in order to enable neighboring group participation (NGP), which results in the formation of an intermediate oxazolinium that restricts the nucleophile's approach from the α -face and thus favors the formation of β -glycosides. Theoretically, the same stereochemical outcome can be achieved with the use of the native acetyl group at the C-2 amino group of glucosamine, but experimentally the glycosylation of this charged intermediate has proven to be quite challenging because of its poor reactivity and inclination to form a stable 2-methyloxazoline that results from proton abstraction (Scheme 2).¹³ As noted above, the acetyl group is therefore typically replaced by the more reactive carbamates that contain additional electronegative atoms. Even so, the major drawbacks to the use of these intricate donor sugars is the need for stoichiometric amounts of toxic promoters (e.g., 3 equiv of $\text{BF}_3 \cdot \text{OEt}_2$), extremely dry conditions (e.g., molecular sieves), and additional chemical transformations after the glycosylation event (e.g., *N*-Troc deprotection/*N*-acetylation). It is therefore imperative to develop an operationally simple synthetic route to O-GlcNAc-modified amino acids in order to expedite the generation of glycopeptides for both biological and functional studies.³

Recently, Polt and co-workers demonstrated the ability of indium bromide (InBr_3) to mediate the activation of glucose anomeric acetates and facilitate the formation of an acyloxonium ion that results from NGP of the C-2 acetate.^{14–16} This reactive intermediate was suggested to be responsible for the exclusive formation of β -glycosides. Thus, we hypothesized that InBr_3 may also be able to mediate the activation of 2-acetamido-2-deoxy-1,3,4,6-tetra-O-acetyl- β -D-glucopyranose (β -Ac₄GlcNAc) and favor β -glycosides through NGP of the C-2 amide (Scheme 2, step 1). As previously detailed, the only major concern was that the poor reactivity of the 2-methyloxazolinium and the stability of its neutral counterpart (2-methyloxazoline) that are expected to be generated in situ could hinder the glycosylation event. However, in support of our hypothesis, Braga and co-workers demonstrated the ability of indium(III) salts to cleave simple oxazolines, yielding β -selenoamides through coordination of the nitrogen heteroatom.¹⁷ Thus, it seemed plausible that InBr_3 could also promote the ring opening of the oxazoline derived from β -Ac₄GlcNAc (Scheme 2, step 2), which is generally considered an unreactive byproduct of glycosylation reactions.

Using Fmoc-Ser-OH as a model acceptor (Table 1), we attempted the glycosylation reaction at an analytical scale with stoichiometric amounts of InBr₃ and β -Ac₄GlcNAc in dichloroethane (DCE) (entry 1). After the reaction vessel was heated at 80 °C for 16 h, analysis of the crude mixture by reversed-phase high-performance liquid chromatography (RPHPLC) revealed a very complex mixture of signals (Figure S1). By the use of a previously synthesized standard, the product signal for β -Ac₃GlcNAc serine was identified from the crude trace and verified using mass spectroscopy. Despite the very low conversion of 29%, these results encouraged us to further optimize the reaction through manipulation of the amounts of sugar donor and Lewis acid, the concentration, and the solvent. As indicated in Table 1, we found that increasing the amount of β -Ac₄GlcNAc could improve the product conversion (entries 1–3) but did not prevent the formation of impurities (Figure S1). We postulated that stoichiometric InBr₃, together with the high temperature and acidity of the reaction mixture (pH ~2), facilitated caramelization of the sugar donor. This was supported by the presence of high-molecular-weight products and the dark-brown color of the reaction mixtures. Although the caramelization temperature is typically high, it is well-known that strongly acidic or basic conditions can accelerate the process.^{18,19} With that in mind, we hypothesized that lowering the amount of InBr₃ could resolve the formation of impurities. Excitingly, lowering the loading of InBr₃ with respect to the donor minimized the formation of impurities (Figure S2) and also improved the product conversion (entries 4–9). At lower loadings, we found there to be a good relationship between the amount of Lewis acid used and the product conversion.

However, the conversion did not notably improve with catalyst loadings above 20 mol %, indicating that catalytic amounts of InBr₃ are optimal to achieve maximal conversion and minimal side-product formation. The highest conversions were obtained with dichloroethane (DCE) compared with chloroform (CHCl₃) and other non-halogenated solvents such as toluene, acetonitrile (CH₃CN), tetrahydrofuran (THF), and 1,4-dioxane (entries 8–12). We also tested the reaction with the less-reactive α -Ac₄GlcNAc (entry 13). Interestingly, the two anomers display similar reactivities and result in nearly identical product conversions (Figure S4). However, given its commercial availability, we decided to use β -Ac₄GlcNAc for the remainder of the work described in this paper.

Having optimized the conditions to maximize the product conversion, we proceeded to run the reaction at a larger scale with a refluxing temperature of 84 °C for 16 h (Table 2). Following purification, Ac₃GlcNAc serine was isolated in 80% yield (entry 1) exclusively as the β -anomer as determined by NMR spectroscopy (Figure S9). Importantly, we also did not observe any racemization of the amino acid in the NMR spectra, although we did not characterize this possibility further. Extension of the same conditions to the secondary alcohol of Fmoc-Thr-OH yielded β -Ac₃GlcNAc threonine in 77% yield (entry 2). Notably, these yields were obtained without the need for dry solvents or special handling of the catalyst, suggesting a resilience to moisture. In order to investigate the effect of water on the catalyst, we purchased the monohydrate of Fmoc-Thr-OH and submitted it to the optimized conditions. Surprisingly, 1 equiv of water had only a very small effect, as β -Ac₃GlcNAc threonine was isolated in 71% yield (entry 3). Thus, this highlighted the ability of the reaction to tolerate the presence of water. Notably, these reactions can be easily scaled up to a gram with reproducible yields that range between 70 and 80%. Furthermore, the β -

glycosides generated according to Table 2 can be readily activated for SPPS as pentafluorophenyl (Pfp) esters in great yields (see the Supporting Information).

Given that we had previously demonstrated the enzymatic stability of S-GlcNAc (β -Ac₃GlcNAc cysteine) and its potential utility for in vivo studies, we decided to further extend the reaction to Fmoc-Cys-OH.²⁰ Unfortunately, attempts with the commercial hydrate afforded β -Ac₃GlcNAc cysteine in 50% yield (entry 4). We attributed the lower yield to the oxidation of cysteine in the presence of stoichiometric amounts of water. To overcome the relatively poor yield obtained from the hydrate, we sought to protect the thiol with the acid-labile triphenylmethyl (trityl) group. We reasoned that the acidic conditions generated from the activation of the anomeric acetate could facilitate the in situ deprotection, thereby unmasking the thiol for nucleophilic attack (Scheme 3). To prevent the addition of the trityl cation back onto to the thiol, triisopropylsilane (TIPS) could be added to scavenge the cation. However, given the ability of tertiary silanes to reduce carboxylic acids in the presence Lewis acids,²¹ we chose to protect the C-terminal acid with a Pfp ester. Since it is a common practice to activate the β -glycosides prior to SPPS, the use of the Pfp ester would remove the need for additional chemical transformations following the glycosylation. As postulated, treatment of Fmoc-Cys(Trt)-OPfp with the optimized reaction conditions and excess TIPS led to the formation of β -Ac₃GlcNAc cysteine Pfp ester in 73% yield (Scheme 3).

In light of the exclusive formation of β -glycosides using various acceptors, we reasoned that NGP of the C-2 amide to generate the 2-methyloxazoline must be occurring as postulated in Scheme 2 (step 1). In support of this, we were able to detect formation of the 2-methyloxazoline during the course of the reactions by mass spectroscopy (Figure S31), but its isolation from the crude mixtures proved to be very difficult. To test whether InBr₃ could also catalyze the ring opening of the 2-methyloxazoline as hypothesized in Scheme 2 (step 2), we synthesized the 2-methyloxazoline according to the published literature procedure²² and submitted it to the optimized conditions using Fmoc-Ser-OH as the amino acid. HPLC analysis of the crude reaction mixture revealed a chromatogram almost identical to that for the reaction with β -Ac₄GlcNAc (Figure S5), and NMR analysis confirmed the exclusive formation of the β -anomer. Importantly, no product formation was detected in the absence of InBr₃ (Figure S27). These data provide evidence for the 2-methyloxazoline as a key intermediate and suggest that InBr₃ is a dual catalyst in these reactions with roles in anomeric activation and ring opening of the 2-methyloxazoline. In order to evaluate other 2-acetamido donors, we synthesized α -Ac₄GalNAc and submitted it to the conditions developed in Table 2 using Fmoc-Ser-OH. These conditions resulted in 69% product conversion and exclusive formation of the β -glycoside as determined by NMR analysis (Figures S6, S21, and S30).

In summary, we have reported the facile synthesis of serine, threonine, and cysteine β -glycosides using commercially available β -Ac₄GlcNAc and catalytic amounts of InBr₃. The present route yields exclusively the β -glycosides in yields comparable to those for published methods without the need for several chemical transformations. Unlike other Lewis acids used in glycosylation reactions, indium can tolerate moderate levels of water, as exemplified by the glycosylation of amino acid hydrates. Lastly, we have provided evidence that supports

neighboring group participation of the C-2 amide and demonstrates the dual catalyst feature of InBr₃ in the activation and glycosylation event. Overall, this synthetic route represents an excellent alternative to current procedures and does not require extensive experience in carbohydrate chemistry. We believe that these convenient features will make this synthetic route the standard choice in the preparation of β -Ac₃GlcNAc amino acids.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

C.A.D.L. is a National Science Foundation Graduate Research Fellow (DGE-0937362). This research was supported by the National Institutes of Health (R01GM114537 to M.R.P.). The authors thank Narek Darabedian (USC) for assistance with figure preparation and compound characterization. C.A.D.L. thanks his family for their unwavering support and Dr. Ricardo Pacheco (USC) for insightful discussions.

REFERENCES

- (1). Bond MR; Hanover JA J. Cell Biol. 2015, 208, 869. [PubMed: 25825515]
- (2). Yang X; Qian K Nat. Rev. Mol. Cell Biol. 2017, 18, 452. [PubMed: 28488703]
- (3). Chuh KN; Batt AR; Pratt MR Cell Chem. Biol 2016, 23,86. [PubMed: 26933738]
- (4). Yuzwa SA; Yadav AK; Skorobogatko Y; Clark T; Vosseller K; Vocadlo DJ Amino Acids 2011, 40, 857. [PubMed: 20706749]
- (5). Wright TH; Bower BJ; Chalker JM; Bernardes GJL; Wiewiora R; Ng WL; Raj R; Faulkner S; Vallee MRJ; Phanumartwiwath A; Coleman OD; Thezenas ML; Khan M; Galan SRG; Lercher L; Schombs MW; Gerstberger S; Palm-Espling ME; Baldwin AJ; Kessler BM; Claridge TDW; Mohammed S; Davis BG Science 2016, 354, aag1465. [PubMed: 27708059]
- (6). Yang A; Ha S; Ahn J; Kim R; Kim S; Lee Y; Kim J; Soll D; Lee HY; Park HS Science 2016, 354, 623. [PubMed: 27708052]
- (7). Enugala R; Carvalho LCR; Dias Pires MJ; Marques MMB Chem. - Asian J 2012, 7, 2482. [PubMed: 22753270]
- (8). Meinjohanns E; Meldal M; Bock K Tetrahedron Lett. 1995, 36 (50), 9205.
- (9). Dullenkopf W; CastroPalomino J; Manzoni L; Schmidt R Carbohydr. Res 1996, 296, 135. [PubMed: 9008845]
- (10). Saha UK; Schmidt RR J. Chem. Soc., Perkin Trans. 1 1997, 12, 1855.
- (11). Mitchell SA; Pratt MR; Hruby VJ; Polt R J. Org. Chem 2001, 66, 2327. [PubMed: 11281773]
- (12). Marotta NP; Cherwien CA; Abeywardana T; Pratt MR ChemBioChem 2012, 13 (18), 2665. [PubMed: 23143740]
- (13). Arsequell G; Krippner L; Dwek RA; Wong SYC J. Chem. Soc., Chem. Commun 1994, 20 (20), 2383.
- (14). Lefever MR; Szabò LZ; Anglin B; Ferracane M; Hogan J; Cooney L; Polt R Carbohydr. Res 2012, 351, 121. [PubMed: 22342206]
- (15).) Coss C; Carrocci T; Maier RM; Pemberton JE; Polt R Helv. Chim. Acta 2012, 95, 2652.
- (16). Szabò LZ; Hanrahan DJ; Jones EM; Martin E; Pemberton JE; Polt R Carbohydr. Res 2016, 422, 1. [PubMed: 26795078]
- (17). Braga AL; Vargas F; Galetto FZ; Paixão MW; Schwab RS; Taube PS Eur. J. Org. Chem 2007, 2007, 5327.
- (18). Hrynets Y; Ndagijimana M; Betti M J. Agric. Food Chem 2015, 63, 6249. [PubMed: 26114422]
- (19). Hong PK; Betti M Food Chem. 2016, 212 (C), 234. [PubMed: 27374528]
- (20). De Leon CA; Levine PM; Craven TW; Pratt MR Biochemistry 2017, 56, 3507. [PubMed: 28627871]

- (21). Bezier D; Park S; Brookhart M *Org. Lett* 2013, 15, 496. [PubMed: 23317512]
(22). Nakabayashi S; Warren CD; Jeanloz RW *Carbohydr. Res* 1986, 150, c7. [PubMed: 3756952]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Scheme 1.
Previous Methods for β -Ac₃GlcNAc Amino Acids

Scheme 2.
Postulating InBr₃-Mediated Glycosylation

Scheme 3.
In Situ Deprotection and Glycosylation

Table 1.Screening Conditions for Glycosylation of Fmoc-Ser-OH^a

entry	equiv of donor	InBr ₃ (mol %)	solvent	conv. (%) ^b
1	1	100	DCE	29
2	2	100	DCE	30
3	3	100	DCE	63
4	3	5	DCE	49
5	3	10	DCE	64
6	3	20	DCE	75
7	3	50	DCE	77
8	3	20	THF ^c	2
9	3	20	Toluene	64
10	3	20	CH ₃ CN	61
11	3	20	CHCl ₃ ^c	22
12	3	20	dioxane	2
13	3 ^d	20	DCE	77

^a Unless otherwise noted, reactions were run at 80 °C and at 200 mM concentration of the acceptor.^b Determined by HPLC.^c The reaction was run at 60 °C.^d $\alpha\text{-Ac}_4\text{GlcNAc}$ was used.

Table 2.

Evaluation of the Substrate Scope under the Optimized Conditions

entry	Fmoc-amino acid	product	yield (%) ^a	$\alpha:\beta$ ^b
1	Fmoc-Ser-OH	β -Ac ₃ GlcNAc serine	80	β
2	Fmoc-Thr-OH	β -Ac ₃ GlcNAc threonine	77	β
3	Fmoc-Thr-OH ^c	β -Ac ₃ GlcNAc threonine	71	β
4	Fmoc-Cys-OH ^c	β -Ac ₃ GlcNAc cysteine	50	β

^a Isolated yields.^b Determined by ¹H NMR analysis.^c Monohydrate.