Synthesis of Complex Glycosphingolipids of the Globo Series
José M. Lassaletta* and Richard R. Schmidt§

Abstract
Versatile strategies were designed for the synthesis of biologically important glycosphingolipids of the globo series, as illustrated for the synthesis of the hexasaccharide of the human breast cancer antigen (Globo H, MBr1) and of the stage-specific embryonic antigen 4 (SSEA-4, Sialylgalactosylgloboside).

1 INTRODUCTION
Glycosphingolipids (GLS's) of the globo series are membrane-associated antigens possessing as structural feature a Galα1→4Gal (galabiose-like) linkage which is responsible for their specific roles in recognition and cell differentiation. They are recognized by proteins of pathogenic E. coli in human epithelial cells of the urinary tract[1]. They are also receptors for Shiga toxin[2] and verotoxin (Shiga-like toxin)[3] and constitute the basis of the P-blood group system[4]. Many of them have been described as tumor-associated antigens[5]. The great interest in this family of GLS's and the difficulties for their isolation in sufficient quantities from natural sources has stimulated interest in their chemical synthesis. Considering their particular significance, globo H (Human Breast Cancer Antigen, 1) and sialylgalactosylgloboside (SGG, Stage-Specific Embryonic Antigen-4, 2) were chosen as synthetic targets in this study (Figure 1).

2 GLOBO H (HUMAN BREAST CANCER) ANTIGEN
2.1 Biological Activity
Tumor-associated Globo H antigen was first isolated and chemically characterized by Bremer et al. from human breast cancer cell line MCF-7[6] and immunocharacterized by monoclonal antibodies MBr1[7], and, more recently, VK-9 [8]. It has been identified on several human cancers (including prostate and breast cancer). Synthetic vaccines based on globo H has been shown to produce immune response in mice[9] and induced a focused humoral response in prostate cancer patients[10].

2.2 Retrosynthetic Analysis
Our approach to globo H is based in the retrosynthetic analysis depicted in Scheme 1. According to this view, disconnections ➀-➃ suggested a synthesis based in five appropriately protected building blocks (3-7).

2.3 Building Blocks
The building block corresponding to the lactose unit should be carefully designed in order to minimize dramatic losses of material at the final stages of the synthesis. Compound 7 was chosen in this context taking into account the following ideal characteristics:
• The 4b-OH appears as the unique free hydroxy group, as this position has to be selectively glycosylated.
Easily removable benzyl protecting groups around the rather low-reactive position 4b will maintain its reactivity at higher levels than other commonly used electron withdrawing protecting groups.

The bulky stereocontrolling pivaloyl group at the 2a position grants the formation of β-glycosides, but should avoid the formation of undesired orthoesters at the azidosphingosine glycosylation step (see below).

Taking into account the importance of 7 for the synthesis of glycosphingolipids and the length of the synthesis previously described[11], we decided to investigate a shorter, more efficient route to this key compound. Starting from lactose, the desired compound was obtained in only 8 steps according to Scheme 2[12]. Interestingly, the key 11→12 transformation, based on 1,2-O-silyl rearrangement under basic conditions, has been found to be synthetically useful for the synthesis of related carbohydrates[13].

According to the retrosynthetic analysis designed, compounds 3 and 4 were connected to yield the disaccharide 15 (Scheme 3)[18]. This material was deacetylated (→16) and selectively benzoylated (→17) thanks to the higher reactivity of the 3b position.

"Inverse procedure"[19] fucosylation by donor 6 yielded the trisaccharide 18 in high yield, and this compound was then desilylated (→19) and activated as trichloroacetimidate (→20). Subsequent reaction with acceptor 5 was carried out taking advantage of the 'nitrile effect'[20] for the stereoselective formation of the β anomer 21. The tetrasaccharide was then transformed into donor 24 by successive benzoylation (→22), removal of the anomeric allyl group by iridium-promoted isomerization[21] followed by treatment with NBS[22] (→23) and activation as trichloroacetimidate. Finally, the hexasaccharide core of globo H was obtained by glycosylation of 7 with 24 by applying again the "inverse procedure"[19]. Hydrogenolytic [Pd(OH)2/C, H2] debenzylation and debenzylideneation of this material with concomitant azido group reduction afforded 25, which was fully acetylated to yield compound 26. This compound can be transformed into target globo H following the well-established azidosphingosine glycosylation procedure[23].
In summary, the hexasaccharide moiety of globo H can be efficiently synthesized in only eleven steps from readily available building blocks 3-7, thereby validating the whole strategy. In addition, compounds 21 and 25 possess an appropriate set of protecting groups, which should allow their use as intermediates for the synthesis of other interesting different glycosphingolipids.

3 STAGE-SPECIFIC EMBRYONIC ANTIGEN 4

3.1 Biological Activity
Sialylgalactosylgloboside (SGG, 1), a ganglioside belonging to the globo series, was first isolated in 1983 from chicken pectoral muscles[24]. In the same year, Solter et al.[25] identified SGG with a developmentally regulated antigen, the stage-specific embryonic antigen 4 (SSEA-4), and found SSEA-3 (previously found to be expressed as Gb5)[26] and SSEA-4 to be epitopes of this unique globo-series ganglioside isolated from human teratocarcinoma cells. More recent studies have shown that globo-series antigens SSEA-3 and SSEA-4 are expressed in different forms on human and murine teratocarcinoma cells[27] and are a hallmark of the former[28]. SSEA-4 has also been reported to bind human parvovirus B19 capsids[29].

3.2 Retrosynthetic Analysis and Building Blocks

The retrosynthetic view of the target compound 2 outlined in Scheme 4[30] suggests six convenient disconnections, designed in such a way that useful intermediates would be obtained within the course of the synthesis. Thus, the access to the target molecule is visualized from compound 28, having the oligosaccharidic skeleton of galactosylgloboside (Gb5, SSEA-3) and known building blocks 27[31] and 29[32] (disconnections ➀-➋). Likewise, compound 29 should be accessible starting from tetrasaccharide 30 (Gb4) and the readily available galactosyl donor 3[14] (disconnection ➃). A convergent retrosynthetic view for the former leads to compound 7 as the lactose building block and disaccharide donor 31 (disconnection ➄). Finally, the evident disconnection ➅ of the latter is proposed from acceptor 5 and azidogalactose donor 32.

3.3 Synthesis
The lactose building block 7[12] was selected for the reasons discussed above, and used in the two different approaches investigated for the synthesis of the globotetraose (Gb4) skeleton 30 (Scheme 5). For the first synthesis of the building block 31, the known 3-O-monochloroacetyl-protected azidogalactosyl donor 7a[33] was reacted with acceptor 5d. The latter was synthesized starting from allyl galactoside 5a[17] in
conditions (-40 °C, CH3CN) in which the 'nitrile effect' operates [20]; 31a was isolated in 88% yield and no α-product could be detected. The allyloxy group of 31a could not be isomerized under standard conditions (i.e., Wilkinson's catalyst in a protic solvent in the presence of a DBU-like base) because partial cleavage of the labile monochloroacetyl (MCA) group occurred (→31b). Use of the buffered Ogawa’s reagent [34] (PdCl2, AcOH/NaOAc), however, afforded directly the 1-0-deprotected disaccharide 31c in good yield, these conditions being mild enough to keep the acid-labile benzylidene protecting groups untouched. The following transformation of 31c into the corresponding trichloroacetimidate donor presented again problems due to the lability of the MCA group, which was partially cleaved under the basic conditions required. To overcome this problem, 31c was first quantitatively transformed into diol 31d (MeONa/MeOH), and subsequently transformed (Cl3CCN/DBU) into the bis-trichloroacetimidate 31e. Reaction of the latter with acceptor 7 furnished the desired tetrascarachide 30a in a yield of 62%, which can be considered as acceptable for this system. Ensuing HCl-mediated hydrolysis of the 3d-O-trichloroacetimidoyl group afforded acceptor 30b in moderate (51%) yield.

The general importance of the tetrascarachide 28, carrying the core structure of globoside Gb4 and, thanks to the set of protecting groups incorporated, being the key for all higher globosides, merits an improved synthesis. Therefore, a second approach was designed to overcome the limitations referred to above. In this case, the known l-butylidimethylsilyl 2-azido-4,6-O-benzylidene-β-D-galactopyranoside 32b [15] was converted into donor 32c by treatment with i) TBAF and ii) Cl3CCN/DBU, without isolation of the intermediate hemiacetal. TMSOTf-catalyzed glycosylation of diol 5a [17] with donor 32c proceeded regio- and stereoselectively, again taking advantage of the 'nitrile effect' [20] and the much higher reactivity of 3-OH than 2-OH in the latter. Thus, the desired disaccharide 31f was obtained in high yield as a single isomer, and then effectively transformed into trichloroacetimidate donor 31i by benzylolation (Ag2O/BnBr, →31g), allyl cleavage (PdCl2/AcOH/NaOAc→31h), and imidate formation under established conditions. Reaction of this compound with acceptor 7 [12], again applying the "inverse procedure" [19], afforded tetrascarachide 30c (63%, αβ = 4:1). Selective removal of the 3d-O-acetyl group to obtain 30b was then accomplished quantitatively by means of methanolic ammonia, keeping the 2a-O-pivaloyl group untouched. This material proved to be identical with product 30b obtained from 30a. The second approach, however, proved to be more efficient, since the synthesis required fewer steps (6 vs. 8), and the overall yield (27% vs. 8%) was higher.

For the transformation of fragment 30b into the target 2 (Scheme 6), β-galactosylation was first achieved by using trichloroacetimidate donor 3 [14] under established conditions, yielding the globopentaose derivative 28a in almost quantitative yield. This material was converted into acceptor 28b again using NH4/MeOH for selective de-O-acetylation. It is known [31] that sterically demanding sialyl donors can be attached in good yields to the 3 position of 2,3,4- or 2,3,4-O-protected galactose residues. Therefore, compound 28b reacted with phosphite donor 27, affording an easily separable mixture 33aαβ in good yield (62%, αβ = 4:1). This compound was then converted into trichloroacetimidate 33d in three convenient and high-yielding steps: i)
hydrogenolytic [Pd(OH)$_2$/C, H$_2$] cleavage of benzyl and benzylidene protecting groups with simultaneous reduction of the 2d-azido group followed by peracetylation of the crude product ($\rightarrow 33b$).

![Scheme 6](image)

Scheme 6

### 4 CONCLUSIONS

In summary, a versatile strategy has been described for the synthesis of complex globo-series glycolipids, based on powerful synthetic tools such as the trichloroacetimidate glycosylation procedure, the recently developed phosphite method for sialylation[30], the azidosphingosine-based protocol for glycosphingolipid synthesis[23], and the stereocontrolling properties of a neighbor pivaloyl protecting group. An advantage of the second strategy (developed for target 2) is the construction of the difficult galabiose-type $\alpha(1\rightarrow4)$ glycosidic linkage in the first stages of the synthesis. It should be also stressed that both 30b and 28b, bearing appropriate protecting groups, are not only precursors for the synthesis of Gb$_4$ and Gb$_5$ globo-series glycolipids, but also for further elongation to several higher globosides other than the target SGG. Compound 30b, for instance, could be alternatively glycosylated with an $\alpha$-galactosamine donor for the synthesis of the Forssman antigen. Likewise, the considerable difference in reactivity between the 3d-OH and 2d-OH of 28b can be used for the synthesis of Globo-H and Globo-A structures. Additionally, the lysoglycosphingolipid obtained after reduction of 34a is another interesting intermediate, since it may have different biological properties to that of the corresponding ceramide derivative[35].
5 REFERENCES AND NOTES


DEFINITIONS

Aglycon. The non-carbohydrate part of a glycoconjugate.

Anomer. Stereoisomer differing in the configuration at C-1 in carbohydrates.

Glycosyl donor: A mono- or oligosaccharide in which the anomeric position bears a functionality (leaving group) able to suffer, upon the action of a promoter or catalyst, the attack of nucleophiles.

Glycosyl acceptor: A mono- or oligosaccharide possessing at least one free OH group, which acts as nucleophile in the glycosylation reaction.

Inverse glycosylation procedure: Consists of slow addition of sensitive glycosyl donors to a mixture of the glycosyl acceptor and the catalyst. Minimum interaction of the former with the catalyst is allowed.

LysolysoglycoSPHINGOlipid: A glycosphingolipid which contains sphingosine instead of ceramide as the aglycon.

Nitrile effect: A kinetic effect which allows the obtention of β-glycosides when nitriles are used as solvents in glycosylation reactions.

Orthoesters: Common by-products in the glycosylation reaction. Formally results from the ketalation of the carbonyl of an ester protecting group at C-2 by the anomeric OH and the reacting alcohol.

TMSOTf: Trimethylsilyl trifluoromethanesulfonate, common catalyst for glycosylations via trichloroacetimidates.

Trichloroacetimidates: One of the most popular kind of glycosyl donors. Allow the use of catalysts instead of...