This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
A divergent approach to the synthesis of iGb3 sugar and lipid analogues via a lactosyl 2-azido-sphingosine intermediate

Janice M. H. Cheng, Emma M. Dangerfield, Mattie S. M. Timmer,* Bridget L. Stocker*

Isoglobotrihexosylceramide (iGb3, 1) is an immunomodulatory glycolipid that binds to CD1d and is presented to the T-cell receptor (TCR) of invariant natural killer T (iNKT) cells. To investigate how modifications to the lipid tail or terminal sugar residue of iGb3 influence iNKT cell activity, we developed an efficient and divergent synthetic route that provided access to both sugar and lipid iGb3 analogues which utilised a lactosyl 2-azido-sphingosine derivative as a common intermediate. In this way, iGb3 (1) and the unprecedented analogues 6´´´-deoxy-iGb3-sphingosine 2, 6´´´-deoxy-iGb3-sphinganine 3, C12 N-acyl iGb3 4 and C20:2 N-acyl iGb3 5 were prepared so that key structure-activity relationships can be explored.

1. Introduction

The glycolipid isoglobotrihexosylceramide (iGb3, 1, Figure 1) is a β-linked triglycosyl ceramide that activates a distinct population of T lymphocytes called invariant natural killer T (iNKT) cells when presented by CD1d on dendritic cells (DCs). Here, the lipid tails of the glycolipid lodge themselves into the deep hydrophobic pockets of the CD1d protein which leads to the presentation of the sugar headgroup for recognition by the T cell receptor (TCR) of the iNKT cell. β-Linked sphingolipids were initially studied for their potential to be the endogenous ligands responsible for the positive selection of iNKT cells, however, more recently, such glycolipids have been found to play a role in infection, as illustrated by β-glucosylceramide which accumulated in the lymphoid tissue of mice and humans following the stimulation of DCs with Toll-like receptor (TLR) agonists or live bacteria. Moreover, the manipulation of the immune response, by altering the structure of iGb3 and hence the activity of iNKT cells, has potential application in the treatment of a variety of diseases, such as cancer and autoimmune disorders. Historically, studies of iNKT cells focussed on the potent CD1d ligand α-galactosylceramide (α-GalCer), however, iNKT cell activation using α-GalCer can lead to anergy and accordingly, there has been growing interest in alternative CD1d-binding glycolipids.

Given their significant structural differences, it is remarkable that both α-GalCer and iGb3 can activate iNKT cells. To shed light on this phenomenon, Rossjohn and co-workers and Yu et al. concurrently solved the crystal structures...
of ternary TCR-iGb3-CD1d complexes and determined that the TCR of the iNKT cell pushes the protruding sugar headgroup of iGb3 flat against the α2-helix of CD1d, causing the internal glycosidic bond to resemble an α-configuration. Moreover, the 6′′′-hydroxyl of iGb3 forms a hydrogen bond with Thr159 of CD1d, and both the 4′′′- and 6′′′-hydroxyls partake in van der Waals interactions with Thr159 and Met162, which has been suggested to be crucial for activity.3 Changes to the glycolipid lipid chains, either at the N-acyl position or at the sphingoid backbone, also affect CD1d presentation and iNKT cell activation, resulting in the induction of altered cytokine profiles, a phenomenon mainly studied using α-GalCer analogues.8,19,20,21,22 To better understand the effects of structural modifications to iGb3 on CD1d binding and iNKT cell activation, we became interested in developing an efficient and divergent synthetic route that would allow for the synthesis of both sugar and lipid modified iGb3 analogues. In particular, we were interested in the synthesis of 6′′′-deoxy derivatives of iGb3 containing either the more typical sphingosine ceramide (2), or the saturated sphinganine backbone (3) (Figure 1). Biological evaluation of these compounds would allow for clarification about the importance of the iGb3 6′′′-hydroxyl for CD1d binding and iNKT cell activation and whether lessening the rigidity in the lipid backbone (i.e. 3) can compensate for the loss of the 6′′′-hydroxyl H-bond. Of the N-acyl derivatives, we were interested in the synthesis of the iGb3-C12 4, which will provide a platform to further explore how truncated lipids can affect the orientation of the sugar headgroup and hence activity, and C20:2 iGb3 5, as the C20:2 lipid has been found to have superior activity compared to the saturated C20 analogue in the case of α-GalCer and skews the immune system towards an anti-inflammatory (Th2) response.8,19,24 To date, both sphinganine iGb3 and α-iGb3 have been synthesised and found to activate iNKT cells more efficiently than the parent compound,25,26,27 and terminal sugar deoxy iGb3 derivatives containing a phytosphingosine have been synthesised28 and used to study iNKT cell TCR recognition of structurally diverse CD1d-restricted ligands.29 A selection of N-acyl chain homologues of iGb3 have also been prepared, such as C8,2,30 C16,31 C18,32 C2433 and C261, however, no systematic structure activity relationship studies have been performed on this series.

To prepare iGb3 (1) and analogues 2-5, we envisioned using a lactosyl 2-azido-sphingosine intermediate 6 that could be orthogonally functionalised (Scheme 1). In this way, both the terminal sugar residue and lipid chain could be varied with a...
minimal number of synthetic steps. The synthesis of iGb3 (1) and the 6''''-deoxy-iGb3 derivatives 2 and 3 would involve the installation of the C26 N-acyl group after reduction of the azide on lactosyl 2-azido-sphingosine intermediate 6 to the amine, followed by selective orthoester opening to give the 4''''-O-acetyl derivative 24 and glycosylation with galactosyl donor 722 or D-fucosyl donor 8. The distinction between the two 6''''-deoxy-iGb3 derivatives would be achieved at the final deprotection step, whereby global deprotection using Birch conditions would afford sphingosine 2, while catalytic hydrogenation followed by ester hydrolysis would give sphenanine 3. For the synthesis of the N-acyl derivatives 4 and 5, however, lactosyl 2-azido-sphingosine intermediate 6 would first be transformed into its 4''''-O-acetyl derivative, followed by glycosylation of the 3''''-OH with donor 7 and installation of the C12 or C20:2 N-acyl chain after reduction of the azide to an amine. Lactosyl 2-azido-sphingosine 6 would be obtained by the coupling of lactosyl donor 9 and sphingosine acceptor 11, with donor 9 being prepared from D-lactose (10)35 and sphingosine acceptor 11 from D-arabinose (12).22

2. Results and Discussion

The synthesis of the iGb3 derivatives began with the preparation of key lactosyl 2-azido-sphingosine intermediate 6. To this end, a strategy similar to that developed by Xing et al.35 was adopted in order to prepare an orthogonally protected lactosyl donor (Scheme 2). D-Lactose (10) was thus peracetylated, the anomeric acetate converted into a thiophenyl glycoside, and the remaining acetate groups removed under Zemplén conditions to give thiolactoside 13 as a crystalline solid.36 Thiolactoside 13 was then subjected to 2,2-dimethoxypropane and a catalytic amount of p-toluensulfonic acid (pTsOH) to afford the crystalline 3''''-4''''-ketal in 56% yield,35 which was subsequently benzoylated to give the fully protected thioglycoside 14 in 94% yield, again as crystalline material. At this stage several glycosidation strategies were explored. The double bond in the sphingosine acceptor precluded the use of a thiophilic promoter, however, the anomeric thiophenol could be readily hydrolysed with NBS in aqueous acetone to give the lactol, which could be further functionalised. Conversion of the lactol into the corresponding lactosyl bromide using bromine and triethyl phosphite proceeded smoothly, however, the lactosyl bromide was not a suitable donor for the glycosylation of sphingosine acceptor 11 as coupling yields did not exceed 40%. Consequently, the lactol was equipped with a trichloroacetimidate leaving group to give lactosyl imidate donor 9 in 73% yield over the two steps. The coupling reaction between lactosyl imidate 9 and sphingosine acceptor 1122 was then performed and this reaction proceeded quantitatively when using 1.5 equivalents of donor to give β-lactosyl 2-azido sphingosine 6 as a single anomer, as confirmed by the 7.8 Hz H-NMR coupling constant between H-1'' and H-2''.

Scheme 2. Synthesis of key lactosyl 2-azido-sphingosine intermediate 6

Having successfully prepared the lactosyl 2-azido-sphingosine intermediate 6, the next synthetic targets were the donors 7 and 8 to install the terminal sugars. While D-galactosyl donor 7 could be readily prepared from D-galactose in 7 steps and 45% overall yield,22 a new strategy was required for the synthesis of the D-fucosyl donor 8. To this end, two strategies were envisioned – the first requiring 6-steps and starting from D-fucose (15), and the second commencing with the significantly less expensive D-galactose (18) though requiring an additional four steps to reach the same common intermediate 16 (Scheme 3). Accordingly, D-fucose (15) was peracetylated (→ 16) and converted into the ethyl thioglycoside via treatment...
of 16 with ethanethiol and BF$_3$•Et$_2$O to give thioether 17 in 72% yield over the two steps. To reach the same intermediate 16, D-galactose 18 was converted to the primary iodide 19 via selective isopropylidene protection and iodination, followed by reduction of the iodide with tributyltin hydride and protecting group manipulations. The acetate groups in 17 were then removed under Zemplén conditions and the non-participatory benzyl protecting groups installed to give benzylated donor 20 in 83% yield over two steps. Here, it should also be noted that thioethers (e.g. 20) are useful glycosyl donors which can be activated under mild conditions using CuBr$_2$ and Pr$_4$NBr, thus donor 20 was also used en route to the attempted synthesis of 6´´-deoxy-iGb3 derivatives 2 and 3 (vide infra).

To make the imidate donor, however, the anomeric thioethyl group of 20 was removed via the agency of NBS and resultant lactol 21 was reacted with trichloroacetonitrile and DBU to afford fucose donor 8, with $^1$H and $^{13}$C NMR spectral data matching that of the enantiomer and the optical rotation being equal in magnitude but with opposite sign.

With the terminal building block prepared, we next focussed on the synthesis of iGb3 (1) to explore the validity of our route and to also provide iGb3 as a control for subsequent biological testing. Accordingly, the azide in lactosyl 2-azido-sphingosine 6 was reduced with PPh$_3$ to produce the corresponding amine, which was used without further purification in the EDCI/DMAP-mediated condensation with hexacosanoic (C26) acid (Scheme 4). Under these conditions the fully protected LacCer 22 was produced in 72% yield over two steps. Here, it should be noted that the C26 acyl lipid is notorious for being poorly soluble and unreactive and therefore in the syntheses of glycosphingolipids, the original C26 acyl chain, as found in α-GalCer and iGb3, is often substituted by shorter variants. This presumably has the advantage of increasing reaction yields, however, it can lead to differences in the solubility of the glycolipids and variations in in vivo distribution and pharmacology.

From LacCer 22, the acetonide was then removed via hydrolysis under acidic conditions to afford the corresponding diol and, using a versatile approach to acetylate the axial hydroxyl in a cis-diol system first reported by Lemieux et al., the 4´´-OH was regioselectively protected using trimethyl orthoacetate and camphorsulfonic acid to first install a methyl orthoacetate protecting group across the 3´´-OH and 4´´-OH followed by ring opening during an acidic work-up to afford 4´´-OAc LacCer 23. The position of the acetate ester was confirmed via an HMBC between the carbonyl carbon of the acetate and the proton at the 4´´-position. In addition, the small coupling constant ($J_{3´´,4´´} = 3.5$ Hz) characteristic of H-4´´ observed at a downfield shift (δ 5.22 ppm) relative to the doublet of doublets ($J_{2´´,3´´} = 9.8$ Hz, $J_{3´´,4´´} = 3.5$ Hz) for H-3´´ (δ 3.82 ppm) also corroborated our assignment. A TMSOTf

\[ \text{Scheme 3. D-Fucose building block synthesis} \]
mediated glycosylation reaction between 4”’-OAc LacCer 23 and \( \alpha \)-galactosyl imidate donor 7 was then undertaken to give the fully protected \( \alpha \)-GSL 24, which was deprotected under Birch conditions to give \( \alpha \)-GSL (1).

With \( \alpha \)-GSL (1) in hand, attempts were then made to prepare the 6’’’-deoxy-\( \alpha \)-GSL derivatives 2 and 3 from 4’’’-OAc LacCer 23 and the \( \alpha \)-fucosyl thiethyl donor 20 using the CuBr₂ and \( \text{Pr}_3\text{NBr} \) promoter system. Unfortunately, despite several attempts at this reaction, none of the desired product was observed and unreacted lipid acceptor and hydrolysed donor were isolated upon work-up (Scheme 4). Accordingly, the glycosylation reaction was repeated using the \( \alpha \)-anomer of imidate donor 8, so as to follow Schmidt’s original coupling conditions. Gratifyingly, this reaction proceeded smoothly to give only the \( \alpha \)-anomer (\( J_{1,2} = 3.5 \) Hz) of the fully protected 6’’’-deoxy-\( \alpha \)-GSL 25. An HMBC between C-3’’’ and H-1’’’ confirmed the attachment of the terminal fucose moiety to the 3’’’-position of LacCer. The 69% yield for this reaction was encouraging as glycosylation reactions incorporating \( \beta \)- or \( \Lambda \)-fucose donors are often poor yielding due to the increased acid lability of the \( \alpha \)-fucosyl linkage. To provide the 6’’’-deoxy-\( \alpha \)-GSL derivatives 2 and 3, the fully protected trisaccharide 25 was either treated under Birch conditions or hydrogenated with Pearlman’s catalyst followed by the hydrolysis of the esters to afford 6’’’-deoxy-\( \alpha \)-GSL-sphingosine 2 or the 6’’’-deoxy-\( \alpha \)-GSL-sphinganine 3, respectively.

Next, we turned our attention towards the synthesis of the \( N \)-acyl \( \alpha \)-GSL analogues 4 and 5 (Scheme 5). Here, the
isopropylidene group in lactosyl 2-azido-sphingosine 6 was removed, and the corresponding orthoester selectively opened in excellent yield to give the 3’-OH lactoside 26. TMSOTf-mediated glycosylation of 26 with d-galactosyl imidate donor 7 yielded triglycosyl 2-azido-sphingosine 27 in 56% yield and as the α-anomer only. The sphingosine backbone was then functionalised with fatty acids bearing the C12 or the C20:2 lipid, first via reduction of the azide to the amine using triphenylphosphine and water, followed by an EDCI/DMAP-mediated coupling of the crude amine intermediate with either dodecanoic (C12) or 11Z,14Z-eicosadienoic (C20:2) acid to give the fully protected iGb3-C12 28 or iGb3-C20:2 29, respectively. Here, the difference in the esterification yield for the different acids (75% yield for C12 vs 28% yield for C20:2) was attributed to the different solubilities and reactivities of the two derivatives, with the coupling of the longer lipid to the trisaccharide being more problematic. In addition, it has also been reported that 1,4-dienes (skipped-dienes) are particularly prone to autoxidation to form conjugated diene hydroperoxides and indeed, we observed autoxidation of 11Z,14Z-eicosadienoic acid by HRMS with molecular ion peaks corresponding to the 1,3-conjugated diene alcohol (m/z calcd. for C_{28}H_{35}O_{3}: 323.2592, obsd.: 323.2587) and the 1,3-conjugated diene ketone (m/z calcd. for C_{29}H_{36}O_{3}: 321.2435, obsd.: 321.2435) if the acid was exposed to oxygen for too long. To minimise this, the compounds were kept under an inert, oxygen-free atmosphere as often as was practically feasible and fortunately, none of the by-products could be observed by HRMS or NMR spectroscopy in the purified fractions of the fully protected iGb3-C20:2 29, or in the final product. Global deprotection of the fully protected iGb3-homologues 28 and 29 containing the sphingosine backbone was then performed using a Birch reaction to give the target compounds, iGb3-C12 4 and iGb3-C20:2 5. At this point, we were somewhat disappointed with the modest yield of the Birch reaction, although we were aware that the yields of Birch reactions of complex carbohydrates vary in the literature. Thus, we chose to investigate further and upon careful HRMS analyses of aliquots taken from the crude reaction mixtures, observed that glycolipid fragmentation had occurred during this final deprotection with the hydrolysed by-products LacCer-C12, GlcCer-C12, and Cer-C12, and the analogous C20:2 homologues, being observed. As glycosidic linkages are generally considered to be stable under basic conditions, this result was somewhat unexpected and while others have noticed similar glycosidic cleavages, this is an area that needs to be more thoroughly explored. The target iGb3 derivatives, however, could be readily purified by silica gel and reverse phase column chromatography to give N-acyl iGb3 homologues 4 and 5.

Taken as a whole, the strategy presented herein differs from all other iGb3 and analogue syntheses insomuch as lactosyl 2-azido-sphingosine glycolipid 6, itself readily accessible from D-lactose, is used as the key intermediate. In other reported routes, a trisaccharide donor is coupled directly to either the complete ceramide or in two steps to the sphingosine (or similar) backbone followed by N-acylation. While these reported strategies may be more convergent for the synthesis of one specific target, the route designed in our studies is robust and versatile and allows for the synthesis of multiple targets of either the terminal sugar or the acyl chain series. The yield and number of steps for the syntheses of all iGb3 homologues reported to date are very similar and where one gain is made, another may be lost – as illustrated by the coupling of the sphingosine backbone which requires an additional N-acylation step in our strategy, but which occurs in better yield than the coupling of the full lipid. Thus far, iGb3 plus four unprecedented analogues have been prepared to probe key aspects of CD1d-glycolipid-iNKT cell interactions, however, it is envisioned this strategy will prove valuable in the synthesis of further derivatives in the near future.

Scheme 5. Synthesis of iGb3-C12 4 and iGb3-C20:2 5.
3. Conclusions

In summary, iGb3 (1), 6’’’-deoxy derivative 2, 6’’’-deoxy-iGb3-sphinganine 3 and the two the N-acyl homologues C12 iGb3 4 and C20:2 iGb3 5 were prepared via a single synthetic strategy involving a common lactosyl 2-azido-sphingosine intermediate. This intermediate was prepared by coupling a sphingosine acceptor with a lactosyl imidate donor, which proceeded in excellent yield and β-selectivity. Functionalisation of this key lactosyl 2-azido-sphingosine intermediate at either the 3’’-position or the sphingoid nitrogen provided facile entry into a range of iGb3 analogues, which will allow for the future assessment of structure activity relationships in terms of iNKT cell activation. These findings will be reported in due course.

Acknowledgements

The authors would like to thank Cancer Society of New Zealand for financial support. B. L. S. also thanks the Health Research Council of New Zealand (Hercus Fellowship, 13/049).

References