Author/s:
Nguyen, Tram Thao Thanh

Title:
Total synthesis of effectors for modulating the human immune system

Date:
2020

Persistent Link:
https://hdl.handle.net/11343/252886

# Total synthesis of effectors for modulating the human immune system

**TRAM THAO THANH NGUYEN**

*Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy*

September 2020

School of Chemistry

Bio21 Institute

The University of Melbourne

## Abstract

Innate immunity is provided by a complex network of cells, soluble factors, and organs that respond immediately or within hours of the appearance of a stimulus in the body. Innate immune processes can involve recognition of metabolites, including those of self or microbial pathogens, by endogenously-expressed pattern recognition receptors. In this thesis, I explored the structure of different antigens that can be recognised by the innate immunity macrophage inducible C-type lectin receptor (Mincle), innate-like natural killer T (NKT) cell, and mucosal-associated invariant T (MAIT) cells.

Chapter 2 describes the synthesis of acyl variants of cholesteryl and ergosteryl α-mannoside (αCAMs and αEAMs, respectively), proposed structures for glycolipids from *Candida albicans*. These glycolipids were synthesised by mannosylation of cholesterol or ergosterol, followed by introduction of the acyl groups by esterification of the sugar primary alcohol. The synthetic glycolipids were assessed for their ability to agonize signaling by mouse and human Mincle. We showed that both αCAMs and αEAMs elicited strong signalling through both mouse and human Mincle.

Chapter 3 discloses the total synthesis of an α-galactosylceramide originally reported from *Bacteroides fragilis*: α-GalCer$_{Bf\text{-}716}$, as well as analogues bearing modified lipid side chains to allow exploration of structure-activity relationships for activation of CD1d-restricted iNKT cells. α-GalCer$_{Bf\text{-}716}$ was synthesised by galactosylation of a sphinganine acceptor, followed by an amide coupling with different acid side chains. All synthesised α-GalCer$_{Bf\text{-}716}$ analogues stimulated mouse and human iNKT cells.

In Chapter 4 we proposed a structure for an undefined antigen (5-F-7-RdX) arising from the reaction of 5-A-RU with a dicarbonyl compound. We synthesized a

simpler analogue, 7-RdX. 7-RdX was shown to act as a weak and selective agonist for MAIT TCR-Tyr94α cell lines, over Tyr95α cell lines. The electron density of 7-RdX from an X-ray structure in complex with the antigen presenting molecule MR1 is a close match to that for unknown antigen.

## Declaration

This is to certify that:

(i) The thesis comprises of only my original work towards the PhD except where indicated in the Preface

(ii) Due acknowledgement has been made in the text to all the material used

(iii) The thesis is less than 100,000 words in length, exclusive of tables, bibliographies and appendices.

Tram Nguyen

September 2020

## Preface

All work reported herein has been conducted by the candidate Tram Nguyen except where indicated. Mincle immunoassays were conducted under the supervision of Prof. Sho Yamasaki of Kyushu University, Japan. Invariant NKT cell immunoassays were conducted under the supervision of Prof. Dale Godfrey, and MAIT cell immunoassays was conducted under the supervision of Prof. James McCluskey, both at The University of Melbourne. X-ray crystallography of MAIT cell studies were conducted by Dr Wael Awad in the laboratory of Prof. Jamie Rossjohn at Monash University.

## Acknowledgements

I am using this opportunity to express my gratitude to Professor Spencer Williams who has given me an excellent opportunity to work with him in this project. I am thankful for his guidance, advice and encouragement during the course. I am sincerely grateful to him for sharing his knowledge and support on a number of issues related to my research.

I would like to thank my panel, Professor Mark Rizzacasa and Professor Jonathan White, for all their guidance and support throughout my PhD. Acknowledgements and thanks to the collaborators, particular Drs Garth Cameron for invariant NKT cell immunoassays, Takashi Shimizu for Mincle studies, Lars Kjer-Nelson, Alexandra Corbett and Michael Souter for MAIT cell immunoassays, and Wael Awad for obtaining X-ray crystal structures of the molecules.

I express my warm thanks to Drs. Phillip van der Peet, Ruwan Epa, Zal Halki, Janice Cheng and Sayali Shah for their expert advice and brilliance in the lab. I would like to thank past and present lab members, especially Thimali Kalanika Ariyarathne, Jinling Li and Amani Alhifthi for their contributions. Thank you for making my journey more memorable and enjoyable.

Thanks to the Bio21 Institute support staff for all the help, Dr Hamish Grant and Sunnia Rajput in the NMR facility, Dr Nicholas Williamson and Dr Ching-Seng Ang in the Mass Spectrometry facility. Thanks to the University of Melbourne and the Australian Research Council for supporting me financially, and for funding my visit to the International Carbohydrate Symposium, Portugal.

I want to give my forever thankfulness to my family and especially my partner Chris Pham for their unconditional love, faith and unwavering support. There is no way I could have passed these challenges without his unwavering support and understanding. Thanks for always inspiring me and I am glad to have you by my side for a new chapter of my life.

# Contents

## Abbreviations

| | |
|---|---|
| **αCAGs** | cholesteryl 6-*O*-acyl-α-glucosides |
| **αCAMs** | cholesteryl 6-*O*-acyl-α-mannosides |
| **αCG** | cholesteryl α-glucoside |
| **αCPGs** | cholesteryl 6-*O*-phosphatidyl-α-glucosides |
| **αEAMs** | ergosteryl 6-*O*-acyl-α-mannosides |
| **α-GalCer** | α-galactosylceramide |
| **α-GalCer$_{Bf}$** | *Bacteroides fragilis* α-galactosylceramide |
| **APC** | antigen presenting cell |
| **APM** | antigen presenting molecule |
| **5-A-RU** | 5-amino-6-D-ribitylaminouracil |
| **β2m** | β2-microglobulin |
| **Bcl10** | B-cell lymphoma/leukemia 10 |
| **BINAP** | 1,1'-binaphthalene-2,2'-diylbis(diphenylphosphine) |
| **CARD9** | caspase recruitment domain-containing protein 9 |
| **CD1d** | cluster of differentiation 1 class d |
| **CDR** | complementarity-determining region |
| **CID** | collision-induced dissociation fragmentation |
| **CLRs** | C-type lectin receptors |
| **CPE** | ceramide phosphorylethanolamine |
| **CRAC** | cholesterol recognition/interaction amino acid consensus motif |
| **CRD** | carbohydrate recognition domain |
| **CTLD** | C-type lectin-like domain |
| **DAMP** | damage associated molecular pattern |
| **DIAD** | diisopropyl azodicarboxylate |
| **DIBALH** | diisobutylaluminium hydride |
| **DIC** | N,N'-Diisopropylcarbodiimide |
| **DMAP** | 4-Dimethylaminopyridine |
| **EMCV** | encephalomyocarditis virus |
| **ESI** | electrospray ionisation |
| **FcRγ** | Fc receptor γ-chain |
| **6-FP** | 6-formylpterin |
| **5-F-7-RdX** | 5-formyl-7-D-ribityl-5-deazaxanthine |

| | |
|---|---|
| **G-CSF** | granulocyte colony-stimulating factor |
| **GMCM** | glucose monocorynomycolate |
| **GMM** | glucose monomycolate |
| **HPAEC** | high-pH anion-exchange chromatography |
| **HPLC-ELSD** | high performance liquid chromatography – evaporative light scattering detector |
| **iGb3** | isoglobohexosylceramide |
| **IFN-γ** | interferon-γ |
| **IL** | interleukins |
| **iNKT cell** | invariant natural killer T cell |
| **ITAM** | immunoreceptor tyrosine-based activation motif |
| **ITIM** | immunoreceptor tyrosine-based inhibition motif |
| **KC** | Keratinocyte chemoattractant |
| **KHMDS** | potassium bis(trimethylsilyl)amide |
| **LP** | lipopolysaccharide |
| **M-AcM-MAG** | α-mannosyl-1,3-(6-acyl-α-mannosyl)-1,1-monoacylglycerol |
| **MAIT** | mucosal-associated invariant T cells |
| **MALT1** | mucosa-associated lymphoid tissue lymphoma translocation protein 1 |
| **MAMP** | microbe-associated molecular pattern |
| **MCM** | monocorynomycolate |
| **MDA** | malondialdehyde |
| **Mincle** | macrophage inducible C-type lectin receptor |
| **MIP** | macrophage inflammatory protein |
| **MHC** | major histocompatibility complex |
| **mMDA** | methyl-MDA, 1,3-dicarbonyl 3-oxobutanal |
| **5-MOP-RU** | 5-(1-methyl-2-oxopropylidenamino)-6-D-ribitylaminouracil |
| **MPL** | monophosphoryl lipid A |
| **MR1** | monomorphic MHC class I-like antigen-presenting molecule |
| **MyD88** | Myeloid differentiation primary response 88 |
| **NaHMDS** | sodium bis(trimethylsilyl)amide |
| **NF-κB** | nuclear factor kappa-light-chain-enhancer of activated B cells |
| **NKT cell** | natural killer T cell |
| **NLR** | nucleotide oligomerisation receptor |

| | |
|---|---|
| **5-OE-RU** | 5-(2-oxoethylideneamino)-6-D-ribitylaminouracil |
| **5-OP-RU** | 5-(2-oxopropylideneaminio)-6-D-ribitylaminouracil |
| **PAMP** | pathogen-associated molecular pattern |
| **PPTS** | pyridinium *p*-toluenesulfonate |
| **PRR** | pattern recognition receptor |
| **7-RdX** | 7-D-ribityl-5-deazaxanthine |
| **RL** | 8-D-ribityllumazine |
| **RL-6,7-diMe** | 6,7-dimethyl-8-D-ribityllumazine |
| **RL-6-Me-7-OH** | 7-hydroxy-6-methyl-8-D-ribityllumazine |
| **RLR** | RIG-I-like receptor |
| **rRL-6-$CH_2$OH** | dihydrogen-reduced 6-hydroxymethyl-8-D-ribityllumazine |
| **rRL-6-Me-7-OH** | dihydrogen-reduced form of RL-6-Me-7-OH |
| **RL-7-Me** | 7-methyl-8-D-ribityllumazine |
| **SG** | steryl β-glucoside |
| **Syk** | spleen tyrosine kinase |
| **TBDMS-OTf** | *tert*-Butyldimethylsilyl trifluoromethanesulfonate |
| **TCR** | T cell receptor |
| **TDB** | trehalose dibenhenate |
| **TDCM** | trehalose dicorynomycolate |
| **TDM** | trehalose-6,6'-dimycolate |
| **TMP** | tetramethoxypropane |
| **$T_H1/T_H2$** | T helper 1/2 |
| **TLR** | Toll-like receptor |
| **TMB** | trehalose monobenhenate |
| **TMCM** | trehalose monocorynomycolate |
| **TMM** | trehalose monomycolate |
| **TNF** | tumor necrosis factor |
| **TPA** | 12-*O*-tetradecanoylphorbol-12-myristate-13-acetate |

# Chapter 1

# Introduction

## 1.1 Overview of the human immune system

The human immune system is a complex network of molecules, cells and organs that is essential for human health. It protects the body from infectious micro-organisms including viruses, bacteria and fungi, and helps clear damaged and dead cells. Important components of the immune system include white blood cells (B and T-lymphocytes), antibodies, the complement and lymphatic systems, the spleen, the thymus, and bone marrow. The human immune system can be classified into parts capable of adaptive and innate responses. Adaptive immunity is an antigen-specific defense mechanism that involves antigen processing, display and recognition. Adaptive immunity creates tailored immune cells and provides humoral immunity that is specific for only one type of antigen. These immune cells and highly specific soluble antibodies allow future infections to be recognized, thereby providing immunological memory. On the other hand, innate immunity is usually a non-specific defense mechanism. Innate immunity is a less specific, intrinsic response to pathogens. It occurs immediately or within hours of the appearance of a novel antigen, and can respond to chemical species that typically do not elicit antibodies such as lipids, nucleic acids and unusual metabolites.[1]

The paradigmatic innate immune processes involve recognition of molecules including those of self or microbial pathogens, by endogenously-expressed pattern recognition receptors (PRRs). PRRs are located either on the cell surface or inside the cell.[2] Innate receptors detect danger signals involved in altered self, known as damage associated molecular patterns molecules (DAMPs), or compounds derived from pathogens, termed pathogen-associated molecular patterns (PAMPs).[3] PRRs can also recognize molecular structures associated with related microbes regardless of pathogenicity, referred to as microbe-associated molecular patterns (MAMPs).[4-5] PRRs recognize nucleotides, lipids, sugars, peptides, lipopolysaccharides, and other pathogen

components.[6] When PRRs recognize DAMPs or PAMPs, intra- and intercellular signaling pathways are triggered, causing activation of specialized immune cells such as T cells, B cells or macrophages,[7-9] which then release cytokines including chemokines, interleukins and interferons that can kill or damage the infected host or pathogen cells.

PRRs can be classified into subgroups according to their ligand specificity, function, localization and evolutionary relationship. Based on their localization, PRRs have been divided into two subgroups: membrane-bound PRRs including Toll-like receptors (TLRs)[10-12] and C-type lectin receptors (CLRs)[13]; and cytoplasmic PRRs comprising NOD-like receptors (NLRs) and RIG-I-like receptors (RLRs).[14]

## 1.2 Case study: Toll-like receptor 4 recognizes liposaccharide (LPS)

Toll-like receptors (TLRs) are the best-understood players in the innate immune response. TLRs are transmembrane proteins that are present in most defense cells including macrophages, neutrophils, dendritic cells, mast cells and B cells.[15] TLRs contain leucine-rich repeats as part of their ectodomain (the domain that extends into the extracellular space), and which gives them their specific TLR functionality.[16] The Toll receptor was first described in the fruit-fly *Drosophila melanogaster*, and activates phagocytes and tissue dendritic cells to respond to pathogens.[17] So far, thirteen different TLRs termed TLR1-13 have been described with the first eleven expressed by humans.[18] In plants, insects and vertebrates including mammals, ligand binding to TLRs triggers the synthesis and secretion of cytokines and activation of host defense programs to fight infection.[19]

TLR4 was the first TLR found in humans, and is responsible for the recognition of lipopolysaccharides (LPS, also known as endotoxin), a cell wall component of Gram-negative bacteria.[20-21] LPS contains a core structure of a glucosamine-based phospholipid disaccharide called lipid A, a potent immune stimulator (Figure 1.1).[22]

When LPS is recognized by TLR4, signaling is initiated from the plasma membrane by the adaptor proteins MAL and MyD88 leading to activation of intracellular NF-κB and inflammatory cytokine production, up-regulating the production of pro-inflammatory cytokines.[23] This may cause serious septic shock,[24] leading to low blood pressure and abnormalities in cellular metabolism.[25]

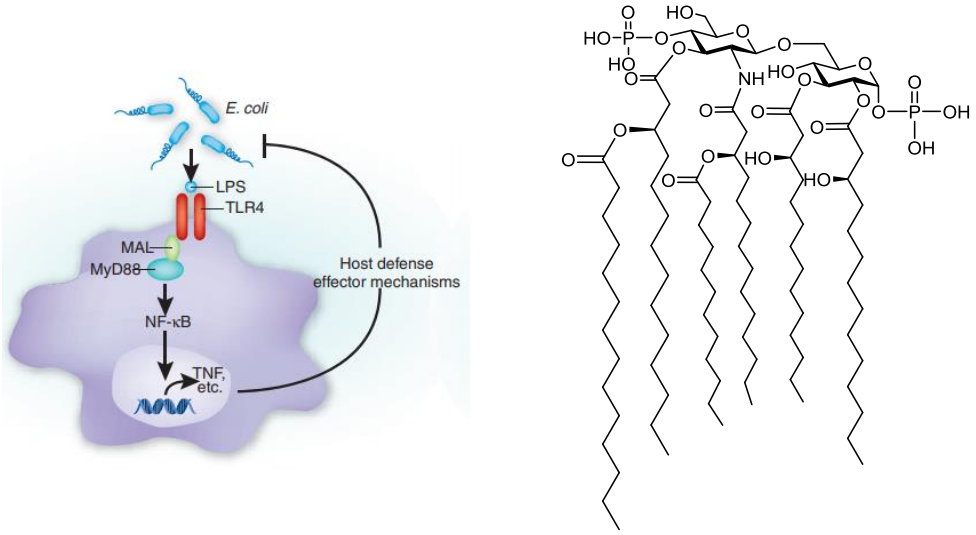



**Figure 1.1:** (Left) Recognition of LPS by TLR4 causes a series of immune response;[26] (Right) the structure of lipid A from *E. coli*.

Several TLR ligands are currently in clinical development or use as vaccine adjuvants.[21] Monophosphoryl lipid A (MPL), a derivative of lipid A from *Salmonella minnesota*, was discovered in the 1970s and is at least 100-fold less toxic and pyrogenic than lipid A (Figure 1.2).[27,28] Purified forms of MPL have been developed by GSK and are used as part of an adjuvant system in adjuvants AS03 and AS04.[29] Other examples of drugs that target TLRs include imiquimod and resiquimod. Imiquimod was discovered by scientists at 3M Pharmaceutical Division. Imiquimod is an immune response modifier that signals through TLR7, and is used to treat genital warts, superficial basal cell carcinoma, and actinic keratosis.[30] Resiquimod is an agonist of TLR7 and 8. It is used as a topical gel, in the treatment of skin lesions. Recently, resiquimod has been explored as an agent for cancer immunotherapy.[31]

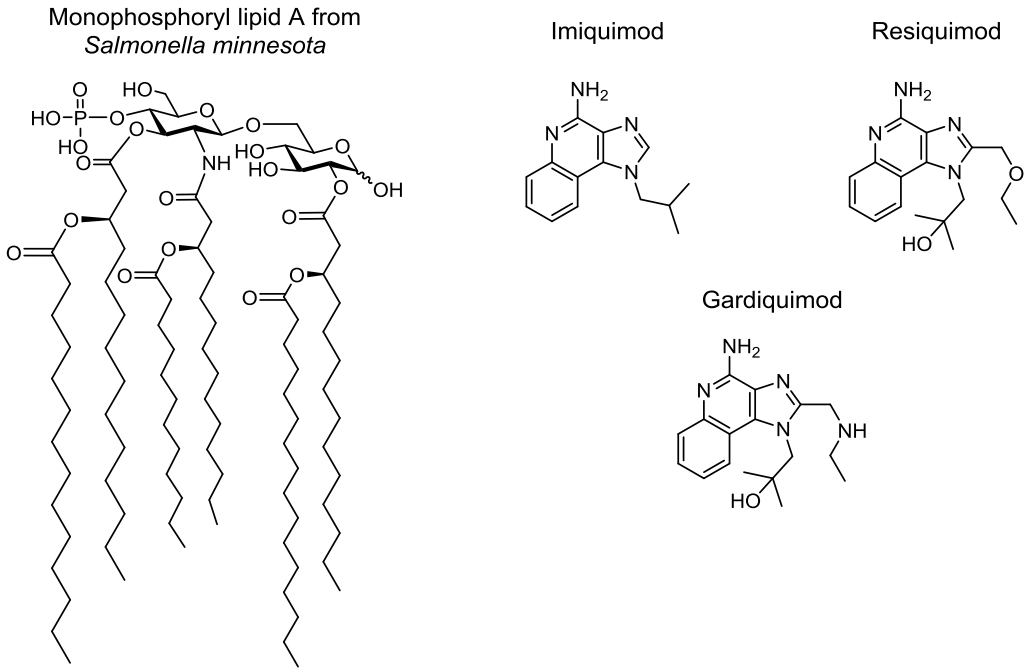



**Figure 1.2:** Structure of several TLR ligands in clinical use or development.

## 1.3 C-type lectin receptor (CLRs) recognizes lipid antigens

Another contributor to innate responses are carbohydrate-binding proteins termed lectin receptors. C-type lectin receptors (CLRs) are a large family of proteins that bind carbohydrates, in a calcium-dependent manner.[32-33] CLRs may display one or more carbohydrate recognition domains (CRDs),[34] or may display C-type lectin-like domains (CTLDs). CTLDs are structurally like CRDs but typically lack the ability to bind calcium and often bind to non-carbohydrate ligands. CLRs are classified into 3 groups: two groups of membrane-bound CLRs (Type I and II) and a group of soluble CLRs. Type I CLRs contain several CRDs or CRD-like domain, whereas Type II CLRs typically carry a single CRD domain.

CLRs can also be classified based on signaling potential, including activation Syk-coupled CLRs with immunoreceptor tyrosine-based activation motif (ITAM) domains, inhibitory CLRs with immunoreceptor tyrosine-based inhibition motif (ITIM) domains, or CLRs without ITAM or ITIM domains.[35]

### 1.3.1 Macrophage inducible C-type lectin, Mincle

Macrophage inducible C-type lectin (Mincle) is a novel type II transmembrane receptor that is expressed in macrophages, dendritic cells, and monocytes upon stimulation.[36] Mincle is a type II transmembrane protein, comprised of an extracellular C-type lectin domain (CRD), and a short cytoplasmic tail.[37-38] Mincle recognizes a variety of exogenous and endogenous stimuli, such as mycobacteria, certain fungi and a soluble protein factor released by necrotic cells.[39] Ligand binding to Mincle results in phosphorylation of the ITAM domain, which recruits Syk either directly or indirectly through Fc receptor γ-chain (FcRγ).[40] This results in the activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) through caspase recruitment domain-containing protein 9 (CARD9)-B-cell lymphoma/leukemia 10 (Bcl10)-mucosa-associated lymphoid tissue lymphoma translocation protein 1 (MALT1) signalosomes.[41] DNA transcription leads to the production of cytokines including tumor necrosis factor (TNF), interleukin-6 (IL-6), granulocyte colony-stimulating factor (G-CSF) and interleukin-11 (IL-10), and chemokines including macrophage inflammatory protein 2 (MIP-2 or CXCL2) and chemokine (C-X-C motif) ligand 1 (CXCL1), as well as stimulating expression of inducible nitric oxide synthase.[42-43] Cytokines produced upon agonism of Mincle signaling orient effector T helper ($T_H$) cell development into $T_H$ subtypes.[44] In mice, a protective $T_H1$ and $T_H17$ cell mediated immunity is induced,[45] while in humans, effects are more complicated (Figure 1.3).

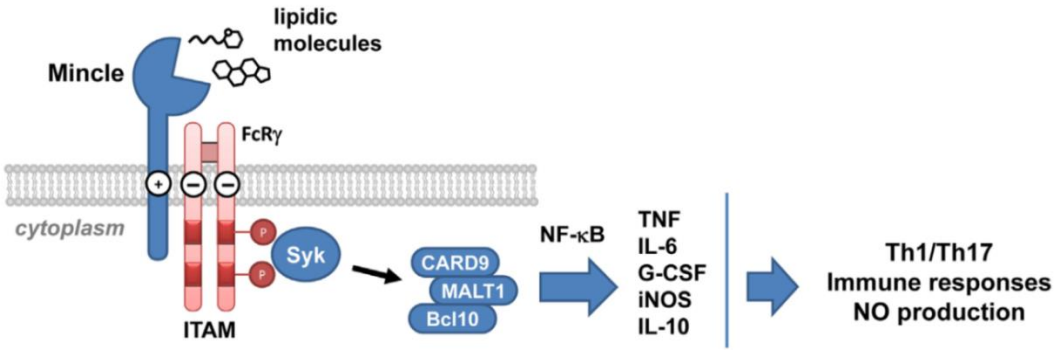



**Figure 1.3:** Signaling cascade when agonist ligands are recognized by Mincle.

### 1.3.2 3D structures of human and bovine Mincle CRDs reveal essential features required for ligand recognition

X-ray crystallography of human and bovine Mincle CRDs revealed similar 3D structural folds. Human Mincle CRD contains two $Ca^{2+}$ ions, while bovine Mincle CRD has an additional $Na^{+}$ ion (Figure 1.4).[46] In the bovine CRD structure, a $Na^{+}$ ion interacts with the main chain oxygens of Leu-172, Asp-177 and Val-175 residues, forming a bridge between Asp-193 and Glu-176, thus shifting the $Na^{+}$ ion closer to $Ca^{2+}$.[33]

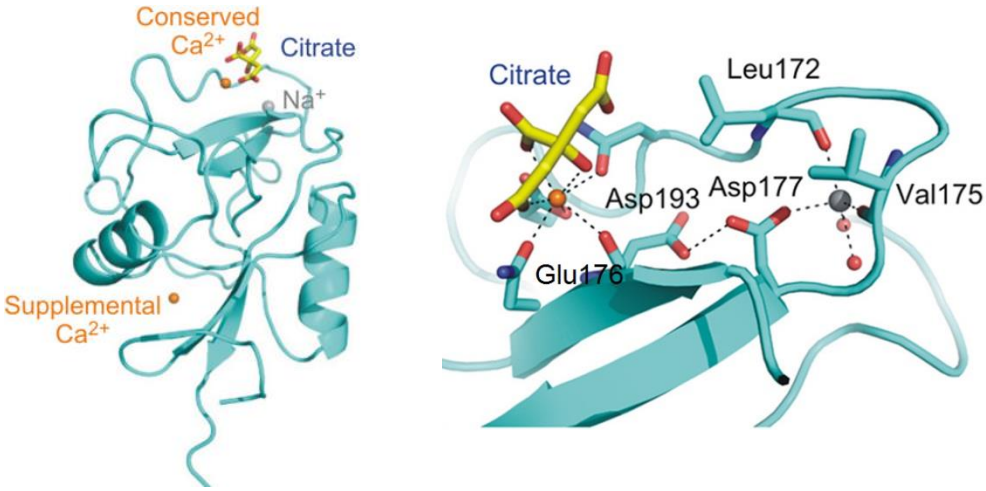



**Figure 1.4:** (Left) Overall structure of the Mincle CRD in which citrate is bound at the conserved $Ca^{2+}$ site; (Right) $Na^{+}$ interaction with the main chain of Leu-172 and Val-175, as well as with the side chain of Asp-177.[33]

The binding site of Mincle was identified through a complex of bovine Mincle with trehalose (Figure 1.5).[33] The CRD contains two carbohydrate-binding sites with one considerring as a "primary" site centered on a conserved $Ca^{2+}$. One glucose residue occupies this canonical primary binding site. The equatorial 3- and 4-OH groups coordinate to the $Ca^{2+}$ and form four hydrogen bonds with different amino acid side chains that also coordinate the $Ca^{2+}$, including Glu176, Asn192, Glu168 and Asn170. The beta face of the glucose is packed against the side chain of Leu172, which prevents

an alternative arrangement where the positions of 3-OH and 4-OH groups are switched by a rotation of 180°, because the presence of the second glucose residue would clash with the leucine side chain. This arrangement brings the second glucose residue into the secondary site.[41] The 2-OH is part of a cooperative hydrogen bonding network in which it accepts a hydrogen bond from Arg182 and donates to one Glu135, thereby bridging these two side chains. A mutant study suggested that the secondary sugar-binding site could function as part of the binding mechanism, since replacement of Glu135 and Arg182 by glutamine and lysine, respectively, resulted in a reduction of the enhanced binding of trehalose.[33]

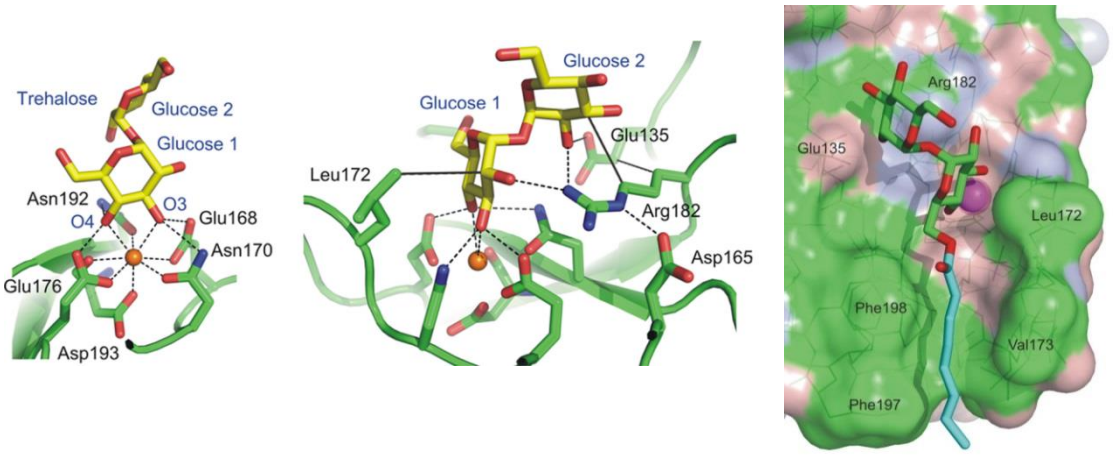



**Figure 1.5:** (Left) Primary binding site for glucose at the conserved $Ca^{2+}$ site; (Middle) trehalose interactions with both the primary and secondary binding sites; (Right) hydrophobic channel adjacent to the primary binding site.[33]

Examination of the surface near the trehalose binding site reveals a hydrophobic channel extending from the corresponding position of the 6-position of glucose.[47] Mincle binds the acyl groups at the 6-position using this hydrophobic region, leading to the conclusion that a single acyl chain at the 6 position is sufficient for Mincle binding.[46] The top of this channel is directly adjacent to the 6-OH group of the glucose residue in the primary binding site. Modelling of an octanoate ester attached to this 6-OH group suggests that the acyl chain sits within a hydrophobic groove formed by

residues Phe197 and Phe198 on one side and Leu172 and Val173 on the other side. This model suggests that at least six carbon atoms would be accommodated within the channel. Mutagenesis revealed that removal of either side of the hydrophobic channel resulted in loss of binding affinity.[41]

### 1.3.3 Mincle recognizes mycolic and corynomycolic acid-based glycolipids

Mycolic acids are long chain α-alkyl-β-hydroxy fatty acids with up to 90 carbons that contain unsaturation, cyclopropanation, ketones, esters and methoxy groups in the meso (main) chain, depending on the source organism.[48-51] Trehalose-6,6'-dimycolate (TDM) and trehalose monomycolate (TMM) are abundant glycolipids found in the cell wall of *Mycobacterium tuberculosis* and related species (Figure 1.6).[52] In 2009, Yamasaki and co-workers demonstrated that TDM could be recognized by Mincle, which then activated macrophages to produce inflammatory cytokines.[37]

Corynomycolic acids are related to mycolic acids, and are produced by a range of corynebacteria, including *Corynebacterium diphtheria, C. ulcerans*, and *C. glutamicum*.[41] Typically, corynomycolic acids are shorter than mycolic acids, and lack the more exotic functional groups such as cyclopropanes, ketones and esters. Lang *et al.* showed that different corynebacteria, including *C. diphtheria, C. glutamicum*, and *C. ulcerans* can bind a murine Mincle-Fc fusion protein.[53] While the precise glycolipids responsible for these effects were not unequivocally identified, it is likely that they comprise corynomycolate esters of trehalose. A mutant strain of *C. diphtheria* that lacks the ability to produce cell wall mycolates failed to bind murine Mincle. Williams and co-workers showed that the synthetic trehalose dicorynomycolate (TDCM) and trehalose monocorynomycolate (TMCM), each bearing C32-corynomycolic acid, stimulated signaling in reporter cells expressing human or murine Mincle.[54] In both

cases the monoesters caused less potent stimulation than the comparable diesters, a finding that has been replicated in other studies.[55-58]

Prior to the discovery that cord factor can signal through Mincle, significant effort was applied to the improvement of structurally simpler and safer immune adjuvants than whole heat-killed mycobacterial cells. Early work led to the development of trehalose dibenhenate (TDB) and trehalose monobenhenate (TMB),[59] which was subsequently shown to signal through Mincle.[58]

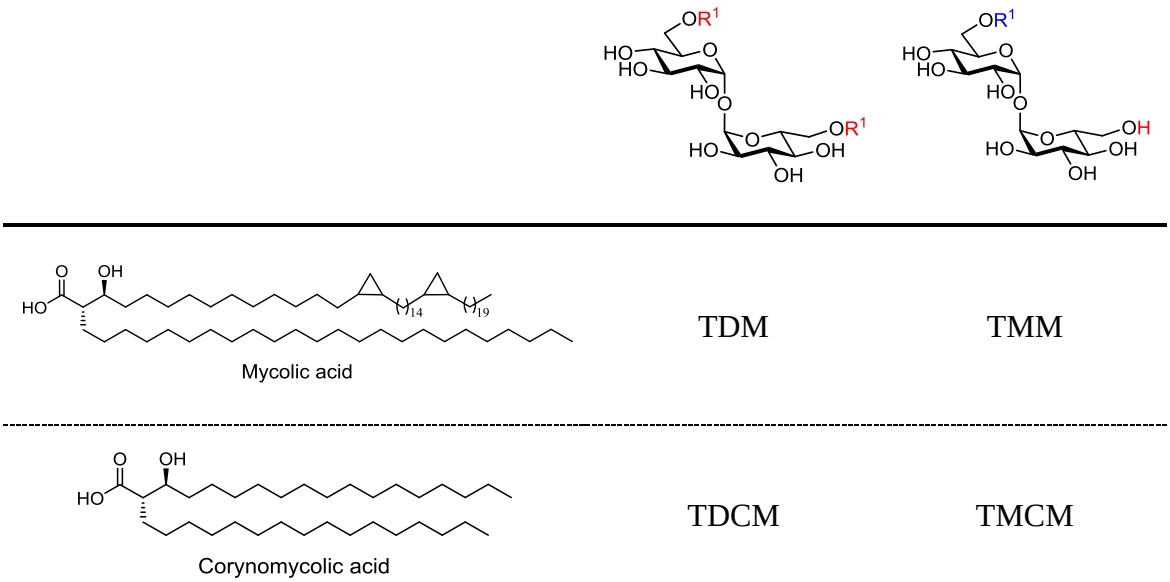



**Figure 1.6:** The prototypical Mincle signalling agonist TDM and analogues bearing mycolic acid and corynomycolic acid.

### 1.3.4 Mincle recognizes glucose- and mannose-based glycolipids

In addition to trehalose-based molecules, monoesters of glucose can signal through Mincle. Williams *et al*. showed that glucose monocorynomycolate (GMCM) could signal through Mincle,[60] and subsequently Decout *et al*. reported that glucose monomycolate (GMM) isolated from *M. tuberculosis* H37v strain was a powerful agonist of human and mouse Mincle with potency greater than TDM.[61] This suggested that the effective signaling is achieved with simply a glucose moiety, rather than requiring trehalose, and that the second carbohydrate is not essential for Mincle

signaling. Decout *et al*. prepared a range of analogues; interestingly, mannose monocorynomycolate (ManMCM) was almost as potent as GMCM at signaling through Mincle, showing that Mincle can recognize mannose-based glycolipids.[61]

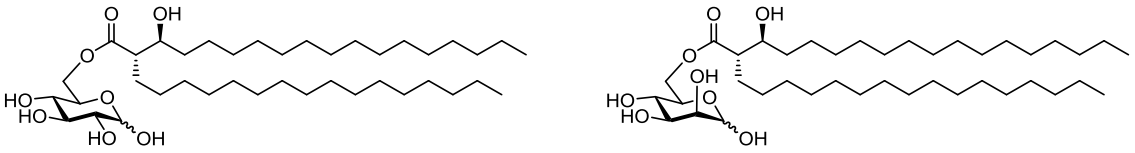



**Figure 1.7:** Structure of glucose monocorynomycolate (GMCM); mannose monocorymycolate (ManMCM).

### 1.3.5 Mincle recognizes cholesterol and cholesterol sulfate

Cholesterol is an essential membrane component for the maintenance of cell integrity and fluidity for a wide variety of cellular activities.[62-63] However, excessive intake of cholesterol leads to hypercholesterolemia and results in the formation of cholesterol crystals.[64] Cholesterol crystals have been defined as a DAMP.[65-67] Recognition by Mincle triggers inflammatory responses by inducing IL-1$\beta$ and IL-1$\alpha$.[66] Kiyotake and co-workers showed that human Mincle recognizes the endogenous ligand cholesterol in both crystalline and plate-coated forms, but not when detergent solubilized or in cell membranes. The multimeric nature of crystalline cholesterol or coated cholesterol may induce multimerization of Mincle in a mechanism similar to that proposed for multivalent TDM in the bacterial cell wall.[68]

Kostarnoy and colleagues demonstrated that cholesterol sulfate, another human sterol, is selectively recognised by human Mincle. Cholesterol sulfate is present in various tissues and body fluids, such as the epithelial cells of barrier tissues, respiratory tract and esophageal mucosa.[69-71] The interaction between cholesterol sulfate and Mincle was shown to induce a secretion of many proinflammatory mediators, such as

IL-1α, IL-1β, KC, MIP-1α and MIP-1β.[72] This interaction plays a role in the pathogenesis of allergic skin inflammation *in vitro* and *in vivo*.

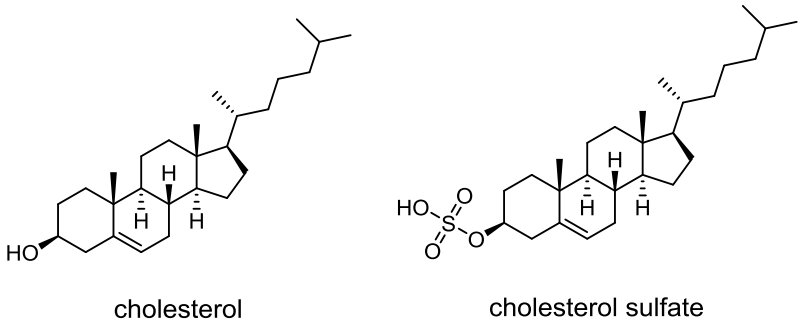



**Figure 1.8:** Mincle recognizes cholesterol and cholesterol sulfate.

### 1.3.6 Mincle recognizes fungal glycolipids

Yamasaki and co-workers were screened 50 species of pathogenic fungi for the ability to signal through Mincle.[38] Among them, only a range of *Malassezia* species, including *M. pachydermatis* and *M. dermatis*, was shown the activity. Fractionation of a lipid extract from *M. pachydermatis* identified two classes of Mincle agonists. It comprises the series of β-gentiobiosyl diglycerol with four lipoforms, and a complex mannosyloxystearyl mannitol (Figure 1.9).[73]

In 2015, Dr. Richarson from our group successfully synthesized a series of related β-gentiobiosides representing those isolated from *M. pachydermatis*, as well as closely related compounds that were reported from *Mycobacterium tuberculosis*.[74] All lipoforms were at best weak signaling agonists through mouse Mincle and did not signal through human Mincle; analogues including shorter lipid chains ($C_4$ and $C_8$) did not signal at all. Truncation of the gentiobiose to a single glucose unit resulted in potent signaling through mouse and human Mincle for the diglyceride bearing iso-$C_{17:0}$ fatty acids.

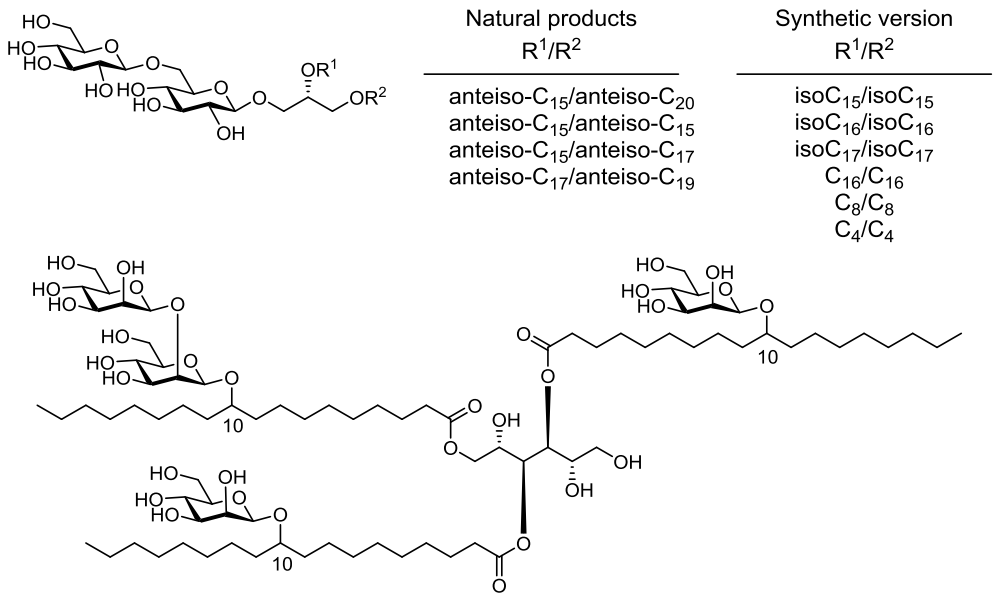



**Figure 1.9:** Structures of assorted mycobacterial glycolipids that signal through Mincle. (Top) β-gentiobiosyl diglycerol isolated from *Malassezia pachydermatis* and the structure activity relationship of synthetic analogues arising from *M. tuberculosis*; (Bottom) *β*-1,2-mannosyloxystearyl mannitol.

Our group reported the total synthesis of *β*-1,2-mannosyloxymannitol glycolipid from *Malassezia pachydermatis* and confirmed its ability to signal through murine Mincle.[75] In addition, it was identified as a potent agonist of human Mincle signaling and constitutes the first fungal metabolite identified that can signal through this human C-type lectin receptor hinting at a role for this glycolipid in antifungal immunity in humans.[75]

*Candida albicans* is a human fungal pathogen and is responsible for candidiasis disease. In 2008, Ashman and colleagues demonstrated that Mincle was a receptor for *C. albicans*, and plays a significant role in the mammalian immune response against this yeast.[76] Three strains of *C. albicans* isolated from patients suffering with candidiasis were shown to bind to Mincle and produced the inflammatory cytokine TNF-α in macrophages. However, a subsequent study by Yamasaki and co-workers demonstrated that three non-clinical strains from *C. albicans* did not activate Mincle-expressing

cells.[38] They speculated that a possible reason for the conflicting results with those of Ashman *et al.* might be the use of different strains by the two groups. Thus the molecular species responsible for Mincle signaling from *C. albicans* remains to be identified.

### 1.4 T cells and T cell receptors – the second arm of the immune system

Another important arm of the immune system involves T cells (T lymphocytes), a type of white blood cell that plays a critical role in cell-mediated immunity. T cells develop from a bone marrow derived lymphoid progenitor in the thymus. T cells are distinguished from other types of lymphocytes, such as B cells and natural killer cells, by the presence of T-cell receptors (TCRs) on their surface, which enables them to recognize antigens presented by antigen presenting molecules (APMs) and kill infected cells or assist other types of immune cells to fight infection.[77]

Each TCR is a heterodimer composed of two different protein chains, either an alpha (α) and a beta (β) chain, or a gamma (γ) and a delta (δ) chain.[78] Most T cells express an αβ heterodimer, while the γδ heterodimer is less common. The ability of T cells to recognize a range of antigens depends on the adaptability of T cell receptors, which are generated by recombination of variable (*V*), diversity (*D*), joining (*J*) and constant (*C*) gene segments, encoded in our genome. During αβ heterodimer development, the DNA encoding the α heterodimer is generated by fusing *V*, *J* and *C* gene segments by somatic recombination. The β chain is generated by fusing *V*, *D*, *J* and *C* gene segments. In the resulting TCRs, the *C* domain is proximal to the cell membrane, followed by a transmembrane region and a short cytoplasmic tail, while the *V* domain constitutes the antigen-binding site (Figure 1.10).

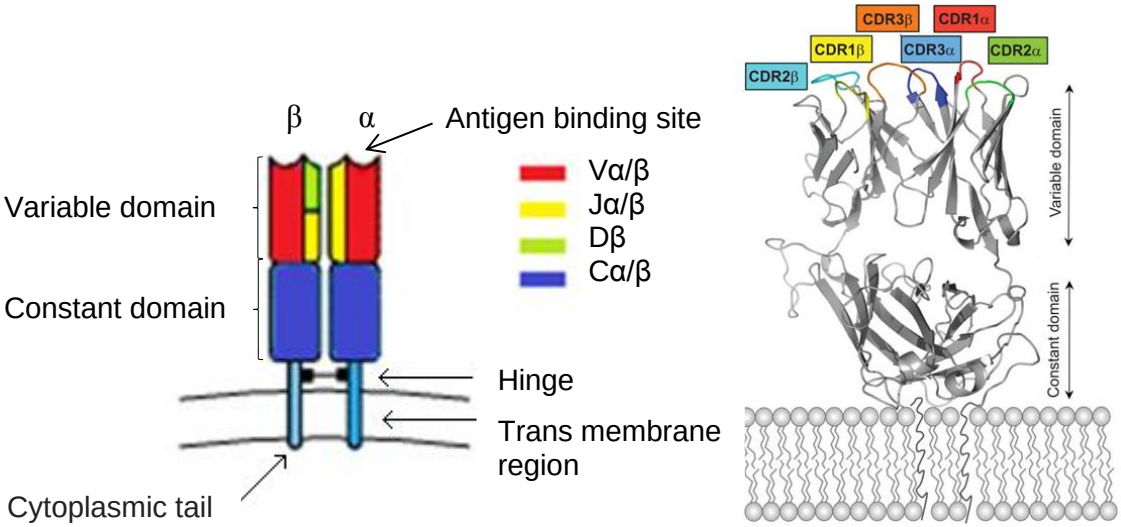



**Figure 1.10:** (Left) model structure of αβ T cell receptors; (Right) crystal structure of human αβ heterodimer TCR with diversity CDRs, and model for interaction with cell membrane.[79]

The variable domain of each chain has three hypervariable complementarity determining regions (CDRs): CDR1, CDR2 and CDR3. As the TCR is composed of two variable domains (α and β, or γ and δ), there are six CDRs of each TCR molecule that can collectively contact the antigen-APM complex. As the most variable parts of the TCRs, CDRs define the diversity of antigen specificity (Figure 1.10).[79]

Conventionally, TCRs allow recognition of antigenic peptides bound to major histocompatibility complex (MHC) molecules on the surface of antigen presenting cells.[80] When the TCR engages with a peptide/MHC complex, an immune synapse forms, the T cell is activated and cytokines are produced: a T helper 1 ($T_H1$)-type response, and immuno-suppressive, or a T helper 2 ($T_H2$)-type response.[81] $T_H1$ responses are pro-inflammatory and are characterized by the production of cytokines including interferon-γ (IFN-γ) and interleukin-2 (IL-2). $T_H2$ responses are immunosuppressive and release cytokines including IL-4 and IL-10.[82-83]

## 1.5 Invariant Natural Killer T cells – iNKT cells

Natural killer T (NKT) cells are a specialized class of T cells that share characteristics with natural killer cells, in particular expression of the NK1.1 cell surface marker.[84] NKT cells are considered to bridge adaptive and innate immune responses.[85-88] Stimulation of NKT cells leads to the releasing of signaling peptides, termed cytokines and chemokines that are recognized by other immune cells.

NKT cells are classified into two types. Type 1 (invariant) NKT (iNKT) cells, which have a limited repertoire of αβ chain diversity and recognize a glycolipid termed α-galactosylceramide (α-GalCer), and type 2 (diverse) NKT cells, which have a wide range of αβ chain diversity and recognize a wider range of lipidic antigens.[89-90] Unlike conventional T cells that recognize peptide antigens presented by MHC molecules, NKT cells recognize lipids and glycolipids when presented by the cluster of differentiation 1 class d (CD1d) APM. The CD1d-restricted iNKT cell lineage is the most extensively studied. iNKT cells are present in the thymus and in peripheral lymphoid organs, including the mucosal immune system. iNKT TCRs are comprised of an invariant α-chain encoded by the gene segment Vα14/Jα18 in mice and Vα24/Jα18 in humans, and is associated with a restricted set of β-chains, generally encoded by Vβ8 in mice and Vβ11 in human.[91]

NKT cell-mediated regulation of immune responses influence a wide range of disease states.[88] These include diseases such as viral infection, autoimmune diseases or cancer, where no exogenous lipids are presented. This suggests that endogenous lipid antigens play important roles in activating iNKT cells during immune responses.[92] Upon binding CD1d-glycolipid complexes, iNKT cells release variety of cytokines, which can promote $T_H1$ or $T_H2$ responses.

### 1.5.1 Discovery of iNKT cells antigens agelasphin-9b and α-GalCer (KRN7000)

The first iNKT cell antigen to be discovered was the sponge-derived sphingolipid agelasphin-9b. Agelasphin-9b is an α-galactosylceramide that was extracted from the sponge *Agelas mauritianus* isolated off the Okinawan coast of Japan by scientists at the Kirin Brewing company, and shown to exert powerful anticancer activity in mice (Figure 1.11).[93] This same group conducted an extensive structure-activity relationship study resulting in the development of α-GalCer (also known as KRN7000). Compared to agelasphin-9b, KRN7000 has a longer acyl chain (two extra carbons), no hydroxyl groups on the acyl chain, and no methyl branch in the sphingolipid tail. Both compounds are powerful agonists of iNKT cells *in vivo* and *in vitro*.[94-95] As the understanding of the immunoregulatory role of iNKT cells has grown, KRN7000 has risen to be the principal glycolipid used in the study of iNKT cells. Subsequent to the discovery of agelasphin-9b and the development of KRN7000 as an anticancer drug, they were shown to exert their anticancer activity by stimulation of CD1d-restricted iNKT cells, thereby mediating a variety of pro-inflammatory and immunoregulatory functions.[96]

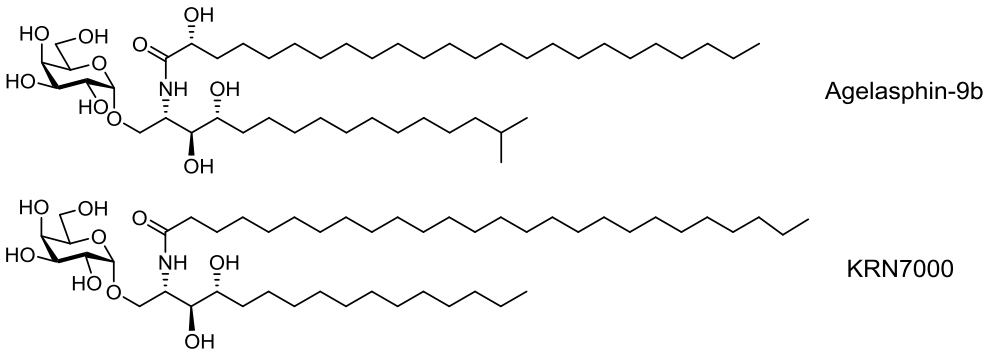



**Figure 1.11:** Structures of the natural product agelasphin-9b and its synthetic analogue KRN7000.

**Structure-activity studies of KRN7000**

A series of studies have determined structural requirements of an antigen for stimulating iNKT cells. These studies can be separated into two classes, and have focused on the investigation of the sugar head group, and the modification of the ceramide tails.

To investigate the effect of different sugars upon iNKT cell proliferation, Taniguchi and co-workers prepared a series of ceramides with different sugar head groups.[97] They found that KRN7000 was the most active compared, followed by α-glucosylceramide, while α-mannosylceramide and β-galactosylceramide were inactive. They concluded that the stereochemistry of hydroxyl groups and the glycosidic linkage, and especially the stereochemistry of the 2'-OH, are essential for iNKT cell recognition.

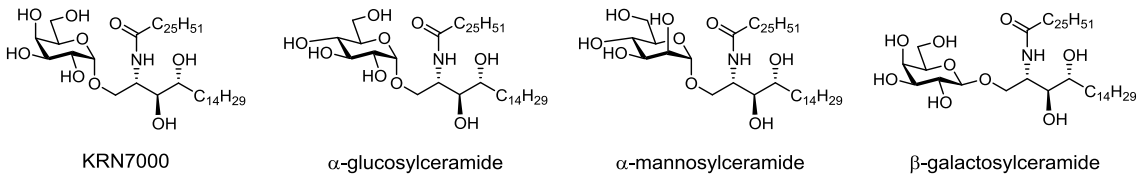



**Figure 1.12:** Modification of the sugar head group for iNKT cell recognition.

Modification of the ceramide tails revealed that the CD1d molecule can accommodate a variety of lipids and this variation can alter the polarization of iNKT cell activation.[98] Fukushima and co-workers demonstrated that the α-hydroxyl group in the acyl side chain did not significantly influence the anti-tumour activity, while the hydroxyl groups at C3 and C4 positions on the phytosphingosine tail were necessary for iNKT cell activation.[94] Kronenberg and co-workers studied modifications of the KRN7000 sphingolipid tail by deoxygenating the hydroxyl groups at different positions (Figure 1.13).[99] Loss of hydroxyl groups on the sphingolipid tail resulted in a loss of affinity; the C3-OH was essential for recognition by CD1d, while loss of the C4-OH resulted in only a small loss in activation or binding of the CD1d-lipid complex.

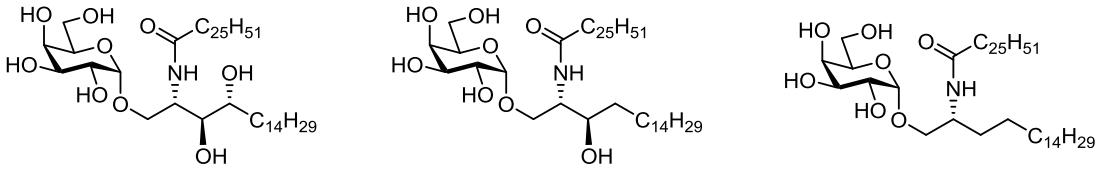



**Figure 1.13:** Modification of the hydroxyl groups in the ceramide tails.

### 1.5.2 The structural basis of interaction of iNKT TCR with the CD1d-KRN7000 complex

**Recognition of KRN7000 by CD1d**

Cerundolo and co-workers reported the 3D X-ray structure of KRN7000 bound to CD1d. KRN7000 binds with the head-group engaged in hydrogen bonds and with the two lipid tails occupying the A' and F' pockets (Figure 1.14).[100] The 2'-OH of the galactose ring is hydrogen-bonded to Asp151; the 3-OH on the sphingosine chain forms a hydrogen bond with Asp80; and the glycosidic oxygen forms a third hydrogen bond to Thr154.[101] These interactions anchor KRN7000 and position it within the lipid-binding groove. The lipid chains of KRN7000 reside deeply within the hydrophobic core of CD1d. KRN7000 seems to have the maximum alkyl chain lengths allowing compete occupation of the binding-groove channels of human CD1d; this tight fit may contribute to its high potency.

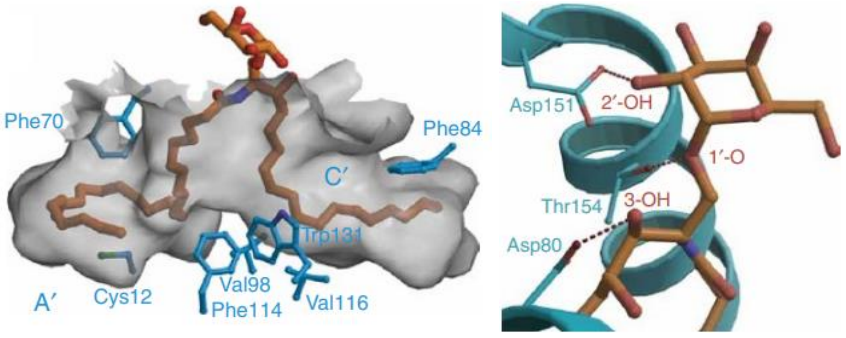



**Figure 1.14:** Binding of KRN7000 to CD1d. (Left) Side view of bound KRN7000 showing two lipid tails inside CD1d pocket; (Right) hydrogen bonds between CD1d and KRN7000.

**Interaction of iNKT TCR with the CD1d-KRN7000 complex**

Borg and co-workers reported the 3D structure of an iNKT TCR in complex with CD1d and the antigen KRN7000.[101] The TCR was positioned parallel over one end of the lipid-binding cleft of CD1d, directly above the F' pocket. Recognition of CD1d-KRN7000 by the TCR is mediated by only the CDR3α and CDR1α loops. Consistent with the structure of the CD1d-KRN7000 complex, only the galactose head group was exposed for recognition by the NKT TCR. The galactose head group of KRN7000 rested on the $V_\alpha$ side of the TCR, sited below the CDR1α loop whereas the CDR3α loop straddled the antigen-binding cleft, interacting with the α1 and α2 helix of CD1d and with KRN7000. KRN7000 made no contact with $V_\beta$ side of TCR (Figure 1.15).[102]

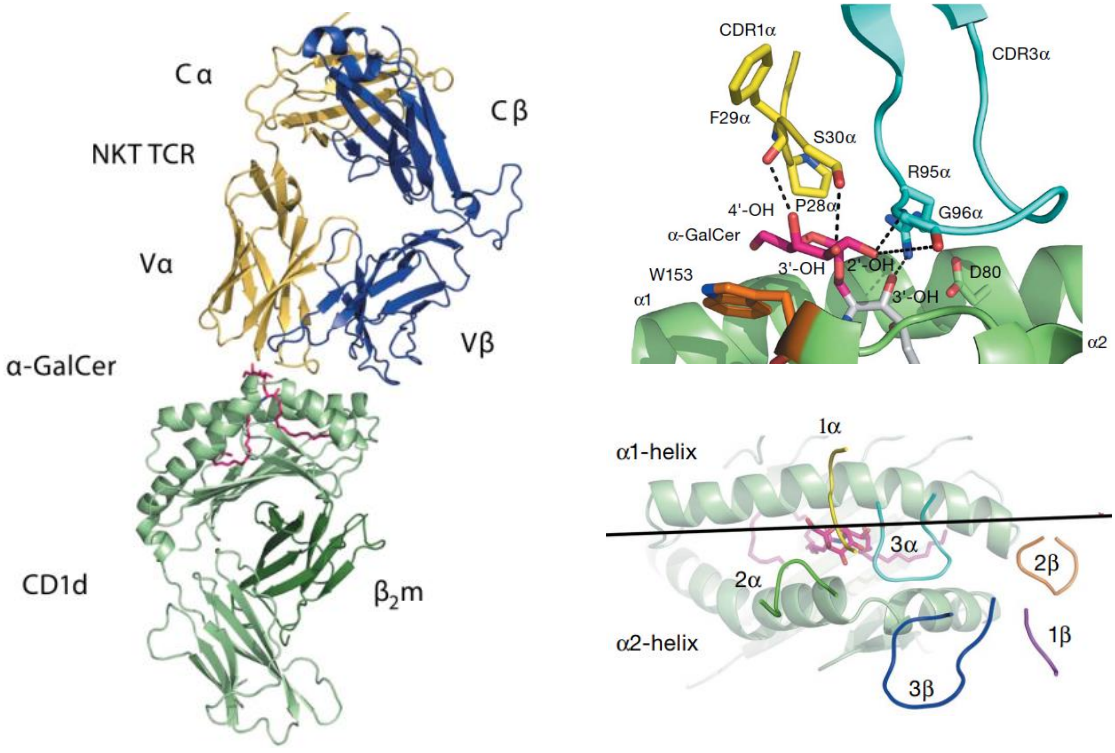



**Figure 1.15:** (Left) Interaction of iNKT-TCR with the complex of CD1d-KRN7000; (Top right) top view of TCR with complex of CD1d-KRN7000; (Bottom right) side view of essential hydrogen bonds for recognition of TCR with KRN7000.

In conclusion, structure-activity relationships and structural studies have revealed that iNKT TCRs recognize an α-galactosyl group linked to a relatively long ceramide

chain bearing a hydroxyl group at the C3 position. Hydroxyl groups at other positions as well as the iso-branching in each tail are not necessary for recognition, and may in fact be detrimental for optimum activity.

### 1.5.3 Endogenous and exogenous antigens for iNKT cells

As first glycolipid identified as a ligand for iNKT cells (agelasphin 9b) was isolated from a marine sponge, the assumption is that these glycolipids are not natural antigens for NKT cells. Consequently, there has been great interest in identifying antigens to which humans (and other mammals that produce NKT cells) would be exposed. Understanding the repertoire of antigens that can activate iNKT cells would support our knowledge of their role(s) and our ability to manipulate their function.

Natural iNKT cell antigens can be separated into two groups: antigens that are produced by the host (endogenous antigens), and antigens from foreign organisms (exogenous antigens). Strong evidence for the existence of an endogenous antigen arises from the observation that mice that lack the endogenous antigen isoglobohexosylceramide (iGb3) also lack iNKT cells.[103] Other examples of endogenous antigens include β-galactosylceramide, β-glucosylceramide[104-105] and the disialoganglioside GD3.[106]

A variety of microbial glycolipids that activate iNKT cells have been isolated from pathogenic organisms, including α-galacturonosylceramides from *Sphingomonas* bacteria,[107-109] α-galactosyl diacylglycerol from *Borrelia burgdorferi*,[110] and a galactosyl sphingolipid (α-Galcer$_{Bf}$) produced by *Bacteroides fragilis*.[111-112] This suggests that one role of CD1d is to survey cellular lipids for the presence of species arising from infectious agents and trigger iNKT-based innate immune responses.[113]

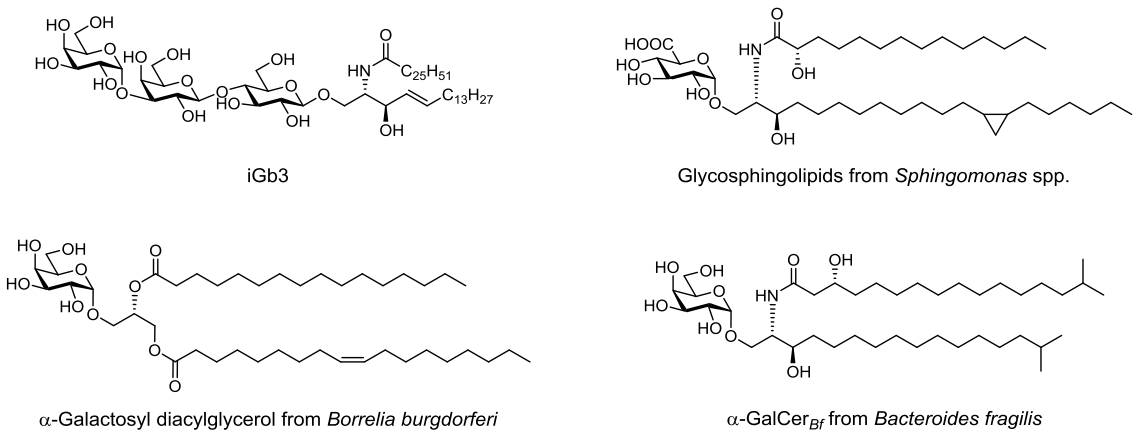



**Figure 1.16:** Some examples of endogenous and exogenous iNKT cell antigens.

The galactosyl sphingolipid α-Galcer$_{Bf}$ of *Bacteroides fragilis* is produced as a mixture of homologues of varying acyl chain length and branching patterns (iso and anteiso).[111-112] It is an interesting compound in that one study demonstrated it activated human iNKT cells in a CD1d-dependent manner,[111] whereas a second study showed it inhibited activation of mouse iNKT cells by KRN7000.[112] The apparent contradiction between these two studies has not been clarified, but may be a result of the use of different mixtures of impure compounds, or because of different concentrations are needed for activation versus inhibition. The activity of this compound could be clarified if chemically homogeneous material was available.

## 1.6 Introduction to MAIT cells

Mucosal-associated invariant T (MAIT) cells are an important subpopulation of innate T cells that recognize antigens when presented by the monomorphic MHC class I-like antigen-presenting molecule (MR1).[92, 114-115] MAIT cells are reminiscent of CD1d-restricted iNKT cells, although in humans MAIT cells are much more abundant.[116] They comprise up to 50% of liver T cells,[117-118] and up to 10% of the peripheral blood T-cell population of humans, and are detected in mesenteric lymph nodes and the gastrointestinal mucosa.[114]

Despite the abundance of MAIT cells, their role in health and disease is unclear.[119] They may have crucial functions in protective immunity, as recent data imply protective functions in autoimmunity and certain infections.[120-121] MAIT cells are conserved across species such as human, cattle and mice,[115, 121-124] and are activated by a wide range of microorganisms, including diverse strains of bacteria and yeast (but not viruses), suggesting that MAIT cells respond to a conserved antigen common to living microbes.[117, 125] Upon activation, MAIT cells release pro-inflammatory cytokines, such as IFN-γ, TNF-α and IL-17 in humans,[118, 126] and IFN-γ, IL-4, IL-5 and IL-10 in mice.[127] Furthermore, MAIT cells have been implicated in autoimmune diseases, such as multiple sclerosis, arthritis and inflammatory bowel disease.[126, 128] Curiously, MAIT cells are absent in the periphery of germ-free mice and their peripheral expansion depends on bacterial colonization, suggesting that they participate in the control of infection.[129-130]

### 1.6.1 MAIT-cell receptor (MAIT TCR) heterogeneity

An important parallel between MAIT and iNKT TCRs is that both constitute a subset of αβ T lymphocytes characterized by a semi-invariant T cell receptor α (TCRα) chain. The TCRα chain originates from the rearrangement of TCRα variable (*V*) and joining (*J*) gene segments. The MAIT TCR comprises Vα19Jα33, in combination with Vβ6 or Vβ8 (TRBV19 or TRBV13) in mice, or the homologous Vα7.2 (TRAV1-2) and Jα33 (TRAJ33), joined to Vβ2 or Vβ13 in humans.[131-133] The constrained gene usage of the MAIT TCR, which is distinct from that of the NKT-cell antigen receptor (Vα24Jα18-Vβ11 in humans), suggested that MAIT cells recognize a specific class of antigen.[117] Nevertheless, despite the limited TCR repertoire of MAIT cells, they respond to a surprisingly broad range of diverse bacteria and yeast strains, suggesting

the existence of conserved antigens common to these organisms that are presented to MAIT cells in an MR1-dependent manner.[121-122]

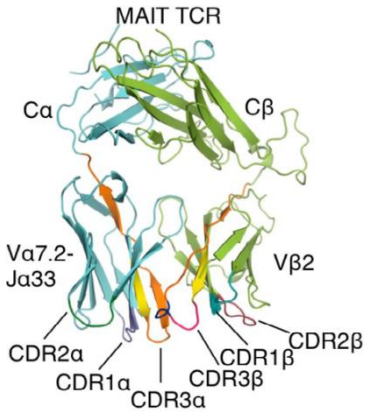



**Figure 1.17:** Structure of MAIT TCR.

To probe the molecular details of MR1-induced MAIT activation, Reantragoon and colleagues determined the 3D structure of a human MAIT TCR by X-ray crystallography, and used this structural information to generate a panel of MAIT TCR mutants.[116] 25 single-site alanine-scanning mutations were made, with 11 in the Vα chain and 14 in the Vβ chain. No difference was observed in the responses of β chain mutants to complexes of MR1 with antigen derived from *Salmonella typhimurium*, suggesting that no individual residue of the β chain was essential for the MAIT cell recognition, whereas a limited number of Vα residues of the MAIT TCR were important for the MR1-induced activation.[116]

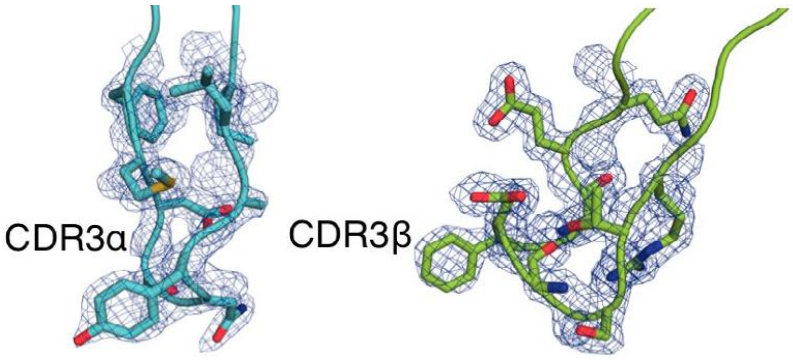



**Figure 1.18:** Electron density map shown as a blue mesh for the CDR3α and CDR3β of the MAIT TCR.[116]

In conclusion, mutant studies revealed highly conserved requirements for the MAIT recognition across human MAIT TCRs stimulated by distinct microbial sources, thereby suggesting that there is a common structural basis to the MAIT TCR-MR1 interaction and that MAIT TCRs recognize an antigen that is common to a wide range of microbes.

### 1.6.2 The structure of MR1 and its recognition of antigen

MR1 is monomorphic,[114, 134-135] and shares the same general construction as MHC-I but has evolved an antigen-binding groove that can recognise a series of non-peptide classes of ligands to specific T cell subsets.[89, 102, 136] MR1 structure consists of three domains (α1, α2 and α3), which is non-covalently bound to β2-microglobulin (β2m) and possesses an antigen-binding cleft with two distinct pockets (A' and F'). Their structures are distinct from the structure of peptide and lipid-based antigen-binding clefts in MHC-I, -II and CD1d.[136-138] The A'-pocket comprises of an 'aromatic cradle' composed of both charged and hydrophobic residues that form the antigen-binding pocket.[119] The high conservation of the MR1 antigen-binding site across species points to an evolutionarily conserved function for MR1 in immunity.[139-142] These observations, together with semi-invariant TCR α-chains and a limited array of TCR β-chain in the MAIT cell TCR, support the hypothesis that MAIT cells recognize a single antigen or a limited number of antigens when presented on MR1.[114, 133, 143] However, until less than 10 years ago, the identity of the MR1-restricted antigens represented an unanswered problem in MAIT-cell biology.

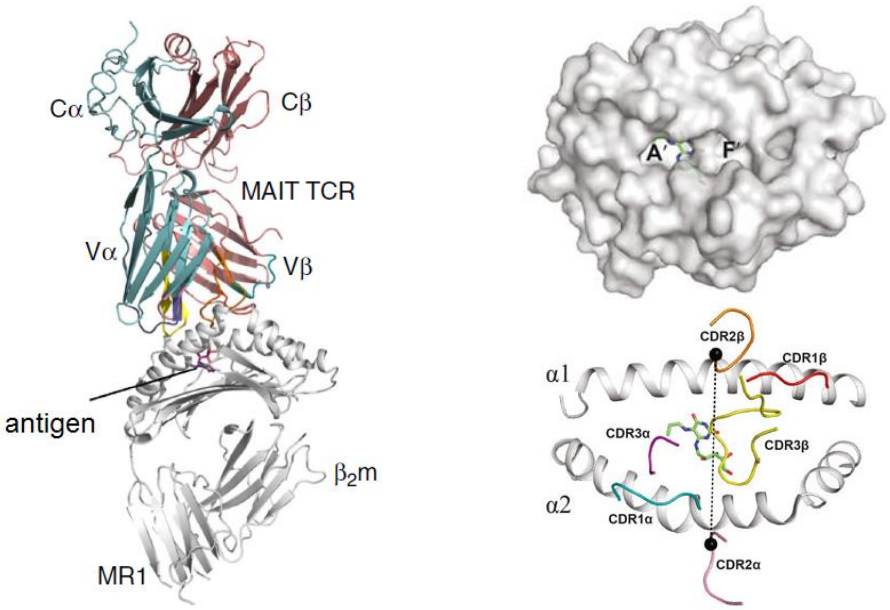



**Figure 1.19:** (Left) MAIT TCR-antigen-MR1 crystal structures of the ternary complex; (Top right) the molecular surface of MR1 with its ligand; (Bottom right) top view of the antigen-binding cleft of the MR1 molecules.[137]

### 1.6.3 Experimental models to identify the sources of MAIT cell antigens

To identify MR1 antigens, the gene encoding MAIT TCR α and β chains were introduced into either of human leukemia T cell lines SKW3[144-145] or the Jurkat T cell lines,[146] thus creating clonal cellular reagents for MAIT cell recognition.

To generate antigen-presenting cells, the MHC-I-deficient human lymphoblastoid cell line C1R was transduced with the gene for human MR1. T cell activation was determined by flow cytometry by monitoring CD69 up-regulation when co-incubated with *S. typhimurium* infected C1R cells expressing MR1. Both SKW3 and Jurkat T cell lines transduced with MAIT TCR genes were activated by C1R cells (expressing endogenous MR1) only when infected with *S. typhimurium*. Collectively, these cell lines were used to search for the MR1 presented antigen(s).

Jurkat cell lines expressing both human and mice MAIT TCRs responded in an MR1-dependent manner to CR1 antigen-presenting cells cultured with bacteria and yeast.[121] Gram-negative bacilli (*Pseudomonas aeruginosa* and *Klebsiella pneumoniae*),

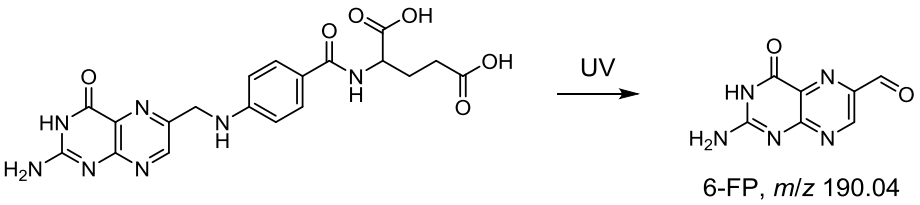
Gram-positive bacilli (*Lactobacillus acidophilus*, *S. typhimurium* and *Escherichia coli*), and Gram-positive cocci (*Staphylococcus aureus* and *Staphylococcus epidermidis*), as well as Mycobacteriaceae such as *Mycobacterium tuberculosis*, all induced a strong response in MAIT cells. Interestingly, MAIT $TCR^+$ cells were activated by phagocytes that were co-cultured with *Candida* and *Saccharomyces* spp. fungi.[147]

Several viruses, including encephalomyocarditis virus (EMCV), Sendai virus, Newcastle disease virus, herpes simplex virus and parainfluenza 3 viruses, all failed to activate MAIT $TCR^+$ cells. These results suggest that MAIT cells respond "preferentially" to bacteria and yeasts. [121] Yet, *Listeria monocytogenes*, *Enterococcus faecalis* and *Streptococcus* group A, which are closely related to S*taphylococci* spp., did not activate MAIT $TCR^+$ cell line.[147] This suggested that these viruses and bacterial strains lack the antigen that is presented by MR1. Shimamura and co-workers have proposed that MAIT cells are activated by the lipidic compound α-mannosylceramide presented by MR1,[148] however, Le Bourhis and co-workers could not independently reproduce this claim.[121]

### 1.6.4 Discovery of the MR1 ligand 6-FP, an antagonist of MAIT cells

MHC-I and MHC-I like molecules are only stable in the presence of ligand, and analogously MR1 could not be refolded efficiently in buffer alone. Kjer-Nielsen and co-workers discovered that small amounts of MR1 could fold in the presence of the cell culture medium RPMI-1640, but in the absence of a bacterial or fungal agonist. This culture medium was used for growing bacteria, and contains several vitamin supplements, some of which are uniquely synthesized in bacteria and plants, but not mammals. Ultimately, the ability of MR1 to fold in the presence of RPMI-1640 was traced to the presence of various folic acid sources (including RPMI-1640 medium, vitamin B complex tablets, or folic acid itself). MR1 that was folded in the presence of

folic acid or RPMI medium was analysed by mass spectroscopy, revealing identical data for a small molecule captured in the folding process. Two new molecular ions of *m*/*z* 190.03 and 147.03 were detected in negative mode, but surprisingly the molecular ion of folic acid, *m*/*z* of 440.13, was not detected. The captured ligand was identified as 6-formylpterin (6-FP, *m*/*z* 190.04), a UV-light degradation product of folic acid (Scheme 1.1) and collision-induced dissociation (CID) fragmentation of 6-FP revealed that the *m*/*z* 147.03 species is a daughter ion of the parent *m*/*z* of 190.03 species.




**Scheme 1.1:** Photodegradation of folic acid in ultraviolet light to 6-FP.

MR1 was refolded in the presence of 6-FP and the 3D structure of the MR1-6FP complex determined by X-ray crystallography. The crystal structure revealed that 6-FP was surrounded by a large number of aromatic residues and sequestered deep within the MR1 antigen-binding cleft.[147] The ligand is captured by a complex of interactions of polar (Lys43, Arg9 and Arg94), and non-polar residues (Tyr7, Tyr62, Trp69, Tyr152 and Trp156) located within the pocket.[149] 6-FP was covalently bound to MR1 through a Schiff base linkage between the formyl group of 6-FP and the side chain lysine 43 of MR1 (Figure 1.20).[119] 6-FP failed to refold MR1 modified with a Lys43 mutation, highlighting the importance of this Schiff base linkage.

Remarkably, even though 6-FP was presented by MR1, it failed to activate MAIT cells. In fact, when both 6-FP and *S. typhimurium* supernatant were treated together with MAIT cells, MAIT cell activation was inhibited, thus 6-FP was concluded to be an antagonist for MAIT cells.

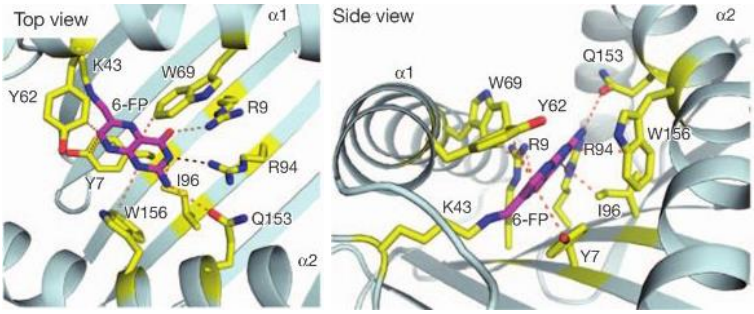
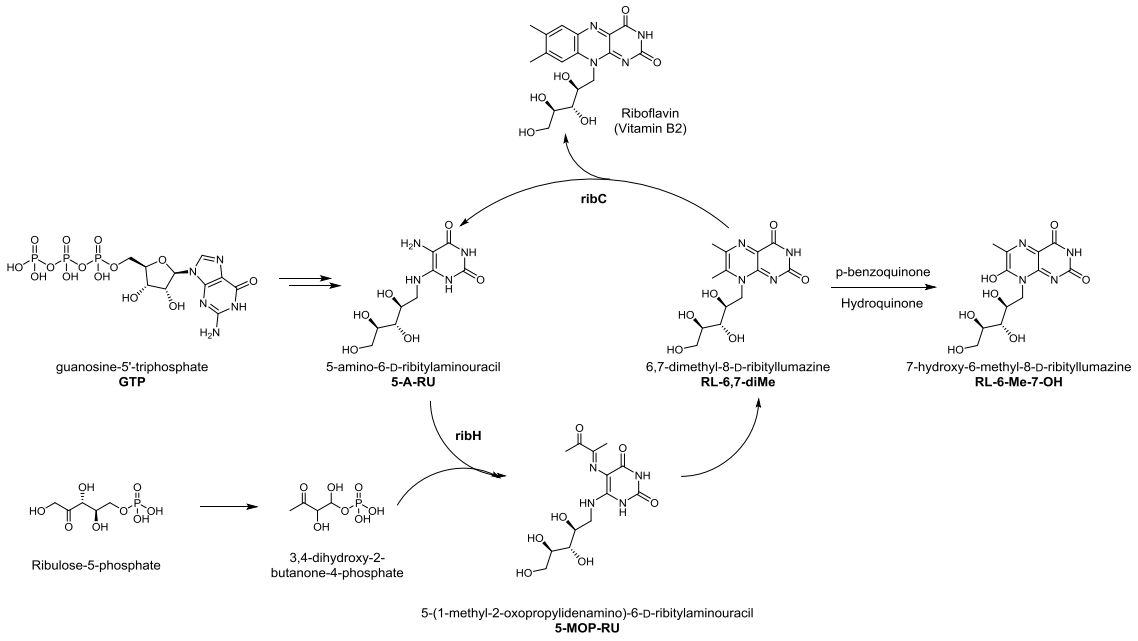



**Figure 1.20:** Top and side views of MR1-6-FP.[119] MR1 residues that contact 6-FP are in yellow. Hydrogen bond and van der Waals interactions are shown in black and red-dashed lines. MR1 is in cyan and 6-FP is in magenta.

### 1.6.5 Discovery of 5-OE-RU and 5-OP-RU: prototyple MAIT cell reactive antigens

The discovery that 6-FP can act as an antagonist of MAIT cells, suggested that capture and identification of an antigen might be more efficient if the culture medium had minimal levels of folic acid. M9 minimal medium, which lacks vitamin B9, was used for growth of *S. typhimurium*. High resolution mass spectrometry in negative mode of MR1 refolded in M9 media revealed a molecular ion of *m*/*z* 329.11, which was assigned as $C_{12}H_{18}N_4O_7$. A database search of this formula did not provide any candidates that matched the mass, but some ribityllumazine compounds from the riboflavin (vitamin B2) biosynthetic pathway were closely matched. It was noted[120-121] that only microbes that activated MAIT cells possessed a complete riboflavin biosynthetic pathway.[119] Thus it was proposed that a riboflavin metabolite may be an MAIT cell antigen.




**Scheme 1.2:** Riboflavin biosynthesis pathway.

Three new compounds could be associated to the experimental data $C_{12}H_{18}N_4O_7$, none of which had previously been described in the vitamin B2 biosynthesis pathway. They included the dihydrogen-reduced form of RL-6-Me-7-OH (rRL-6-Me-7-OH), a dihydrogen reduced form of 6-hydroxymethyl-8-D-ribityllumazine (rRL-6-$CH_2OH$) or 5-(2-oxopropylideneaminio)-6-D-ribitylaminouracil (5-OP-RU). The three candidates, RL-6-Me-7-OH, rRL-6-$CH_2OH$ and 5-OP-RU, were synthesized, and only the last could activate MAIT $TCR^+$ cells, with potency on par with the *S. typhimurium* supernatant.

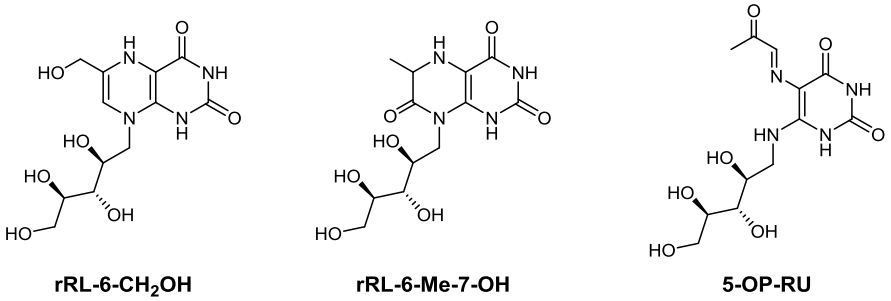
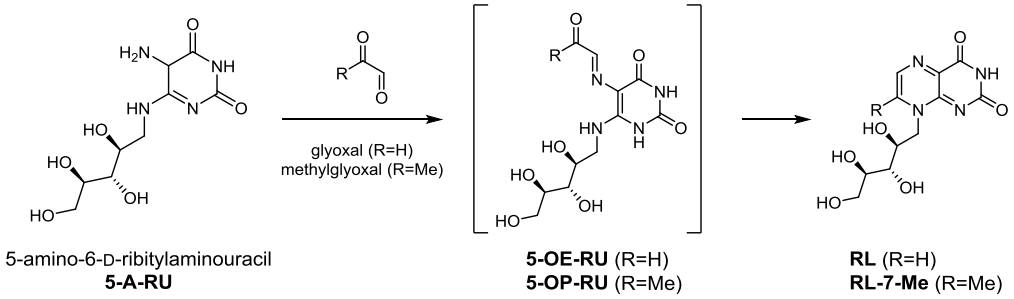



**Figure 1.21:** Structure of three isomers (*m*/*z* 329.11, negative ion mode) that share the molecular formula $C_{12}H_{18}N_4O_7$: rRL-6-$CH_2OH$, rRL-6-Me-7-OH and 5-OP-RU.

Bacher's team[150-151] predicted that *en route* to producing RL-6,7-diMe in the riboflavin pathway, 5-A-RU forms a Schiff base with 3,4-dihydroxy-2-butanone-4-phosphate or 2,3-butanedione to yield an unstable ring-opened pyrimidine intermediate 5-MOP-RU (Scheme 1.2).[152] This reaction occurs spontaneously, in the absence of any enzymatic catalysis. Methylglyoxal, lacking a methyl group compared to 2,3-butanedione, would also be expected to form a Schiff base with 5-A-RU to yield 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil (5-OP-RU, $C_{12}H_{18}N_4O_7$), matching the *m*/*z* 329.11 species, a compound theoretically generated *en route* to creating 7-methyl-8-D-ribityllumazine (RL-7-Me). Similarly, it was reasoned that glyoxal, which lacks two methyl groups compared to 2,3-butanedione, could form a Schiff base with 5-A-RU to yield 5-(2-oxoethylideneamino)-6-D-ribitylaminouracil (5-OE-RU, $C_{12}H_{18}N_4O_7$), matching the *m*/*z* of 315.09. By analogy with 5-OP-RU, 5-OE-RU was theoretically generated *en route* to creating 8-D-ribityllumazine (RL) (Scheme 1.3).[137]




**Scheme 1.3:** Formation of 5-OE-RU and 5-OP-RU from condensation of 1,2-dicarbonyl compound with 5-A-RU.

To form a complex of MR1 and 5-OP-RU or 5-OE-RU, a mixture of 5-A-RU and glyoxal or methylglyoxal was refolded with MR1. Mass spectrometry of the complexes revealed two abundant species with molecular mass *m/z* of 315.09 and 329.11. MR1 tetramers loaded with either 5-OP-RU or 5-OE-RU potently stained MAIT cell, showing that these two compounds are the natural MAIT cell antigens. Notably, both

glyoxal and methylglyoxal are formed from a number of metabolic pathways, including mammalian and bacterial glycolysis.[153]

High resolution 3D X-ray structures of the MAIT TCR complex with MR1 bound to 5-OP-RU or 5-OE-RU unambiguously showed the structures of the ligands.[136] A Schiff base was formed between 5-OP-RU or 5-OE-RU to Lys43 of MR1. The aliphatic moiety burrowed into the MR1 cleft and was stabilized by interactions with Tyr7 and Tyr62. Hydrogen bonds between the ribityl tail and two conserved MR1 residues (Arg9 and Arg94) fix the position of the ribityl tail for MAIT TCR recognition. The affinities of MR1-5-OP-RU complexes with MAIT TCRs, were comparable to those with conventional TCRs recognising MHC-peptide complexes, thus suggesting that the covalent bonded Schiff base is required for MAIT TCR binding and MAIT cell activation.

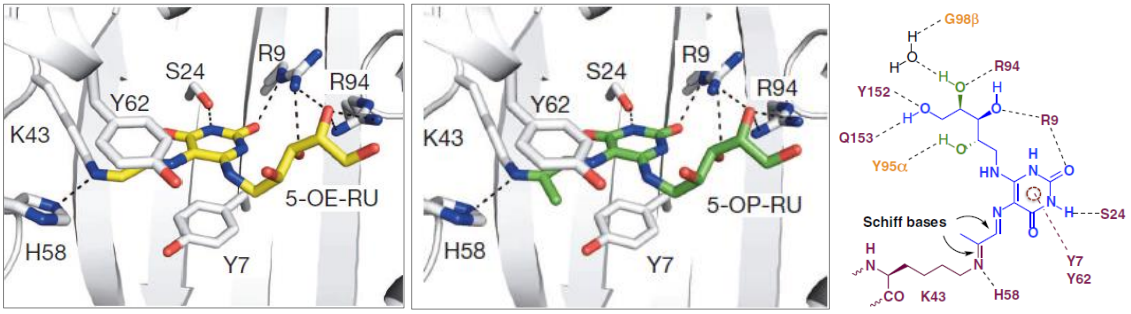



**Figure 1.22:** (Left) MR1 contacting 5-OE-RU; (Middle) MR1 contacting 5-OP-RU; (Right) interaction observed in a ternary crystal structure of a human MAIT TCR-MR1 complex.

The ribityl moieties of 5-OP-RU and 5-OE-RU directly contact the MAIT TCR α chain via hydrogen bonding with the mostly conserved Tyr95. It also formed a hydrogen bond via a bridging water molecule mediated interaction with the CDR3β loop of the TCR. As 6-FP lacks the ribityl moiety, this suggests that the ribityl tail on riboflavin-based precursors is required for MAIT cell activation.

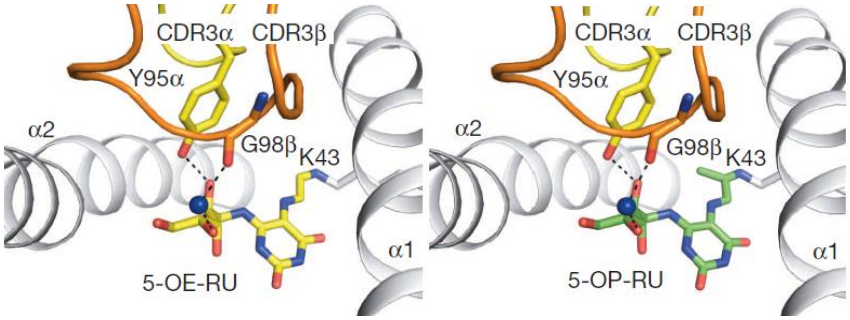



**Figure 1.23:** MAIT TCR contacting (Left) 5-OE-RU and (Right) 5-OP-RU.

Both 5-OP-RU and 5-OE-RU are very unstable in water and cyclize to ribityllumazines (RL) and 7-methyl-8-D-ribityllumazine (RL-7-Me) in minutes at 37 $^{o}$C.[136] However, instead of capturing the thermodynamically more stable ribityllumazine, MR1 sequesters the transiently formed and thermodynamically unstable pyrimidine formed through condensation of compounds arising from two distinct metabolic pathways, namely 5-A-RU from the bacterial biosynthesis of riboflavin and methylglyoxal or glyoxal from bacterial or mammalian glycolysis. While studies to date have identified 5-OP-RU and 5-OR-RU as bona fide MAIT cell antigens, it is unknown if there are other possible antigens that could be presented by MR1 and recognized by MAIT cells.

## 1.7 Outline of the thesis

The thesis is divided into three parts:

**Chapter 2: Discovery of Mincle agonists from *Candida albicans*: cholesteryl 6-*O*-acyl mannosides (αCAM) and ergosteryl 6-*O*-acyl mannosides (αEAM)**

*Candida albicans*, a commensal found in the gastrointestinal tract and mouth of 40–60% of healthy adults, is responsible for the majority of *Candida* bloodstream infections. A prior study showed that clinical strains of *C. albicans* could signal through the innate immune receptor Mincle,[76] but the molecular species have not been identified.[38] In this chapter, we described the synthesis of candidate acyl variants of cholesteryl 6-*O*-acyl mannosides (αCAM) and ergosteryl 6-*O*-acyl mannosides (αEAM)

from *Candida albicans.* In collaboration with Yamasaki and colleagues at Osaka University, we demonstrated that αCAMs and αEAMs can agonize Mincle signalling.

**Chapter 3: Total synthesis of the *Bacteroides fragilis* α-galactosyceramide (α-GalCer$_{Bf}$) and CD1d-restricted activation of iNKT cells**

*B. fragilis* is a member of the human gut microbiota and influences human health in many different ways. *B. fragilis* generates a mixture of α-galactosylceramides that are structurally related to KRN7000 and agelasphin-9b. In this chapter, we developed an approach to the total synthesis of α-Galcer$_{Bf\text{-}716}$. Additionally, we synthesized analogues with different acyl chains in the ceramide to explore structure-activity relationships. Assessment of the ability of α-Galcer$_{Bf\text{-}716}$ and its analogues to stimulate iNKT cell signaling was conducted in collaboration with Prof. Dale Godfrey at Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity.

**Chapter 4: Studies toward identification of a novel vitamin B2 derived T cell antigen**

MAIT cells are much more abundant than other types of T cells in humans; however, their roles and activation remain areas of active research. In this chapter we developed an approach for the preparation of a hypothetical structure (5-F-7-RdX) of a newly discovered antigen that is proposed to arise from the chemical reaction of the riboflavin intermediate 5-A-RU with the 1,3-dicarbonyl compound malondialdehyde (MDA). While 5-F-7-RdX was not synthesised we developed a synthetic approach allowing preparation of the des-formyl analogue 7-RdX. In collaboration with Prof. James McCluskey from the Peter Doherty Institute for Infection and Immunity, we investigated the ability of 7-RdX to stimulate MAIT cells signaling to determine whether it matches the immune-stimulating properties of the natural material, and in

collaboration with Prof. Jamie Rossjohn at Monash University, determined the 3D X-ray structure of this ligand bound to MR1.

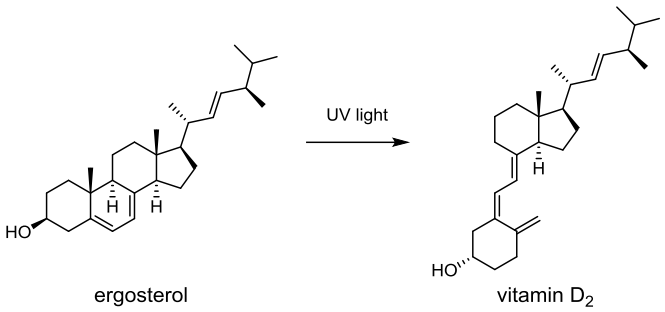
# Chapter 2

# Discovery of Mincle agonists from *Candida albicans*: cholesteryl 6-*O*-acyl mannosides (αCAM) and ergosteryl 6-*O*-acyl mannosides (αEAM)

## 2.1 Introduction

### 2.1.1 Ergosterol and ergosteryl glucoside induce immune responses

Ergosterol is an abundant sterol that resides in the cell membranes of fungi and protozoa, and which maintains membrane fluidity and permeability.[154] In humans, ergosterol is a precursor of vitamin $D_2$ (ergocalciferol); exposure of ergosterol to ultraviolet light causes a photochemical reaction.[155]




**Scheme 2.1:** Photochemical formation of vitamin $D_2$ upon exposure of ergosterol to UV light.

As ergosterol is present in the cell membranes of fungi but not in those of animals, it is an antifungal and anti-protozoal drug target. Polyene antimycotic agents such as amphotericin B and nystatin target ergosterol by physically binding to ergosterol and creating a polar pore in the cell membrane. Azole antifungal drugs such as fluconazole, ittaconazole and ketoconazole inhibit the synthesis of ergosterol by lanosterol 14α-demethylase.[156-157]

Ergosterol exhibits a range of immunological activities; however, the mechanism(s) by which it acts remain unclear. Yasukawa and co-workers showed that ergosterol inhibited tumour promotion by 12-*O*-tetradecanoylphorbol-12-myristate-13-acetate (TPA), a potent tumour promoter in ear oedema.[158] Ergosterol has been shown to suppress the LPS-induced production of TNF-α through inhibition of transcription factor NF-kB. It suppressed proinflammatory gene expression in macrophages and the growth of human colon adenocarcinoma cells.[159] Kou showed that down-regulation of

NF-κB cascade-associated proteins coincided with inhibition of TNF-α production in ergosterol-treated cells.[160]

Caroprese and colleagues showed that a purified extract from *Dunaliella tertiolecta*, characterized as a mixture of ergosterol and 7-dehydroporiferasterol (Figure 2.1), could reduce immune reactions resulting from inflammatory diseases in sheep.[161] It was suggested that it could be used as a feed supplement to modulate sheep immune responses.

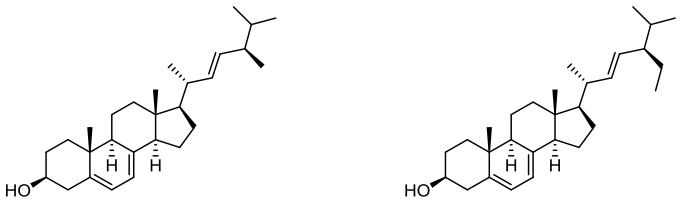



**Figure 2.1:** Structure of ergosterol and 7-dehydroporiferasterol.

Ergosteryl β-glucoside, a glycolipid produced by animals, plants, and various fungi,[162-163] modulates the immune system. *Cryptococcus* spp. are opportunistic fungal pathogens that cause cryptococcosis. Infection by *Cryptococcus neoformans* and *Crytococcus gattii* lead to more than 600,000 deaths per year,[164] especially among immunocompromised individuals. Wantanabe identified a protein capable of hydrolysing steryl β-glucosides (SGs) in *C. neoformans*.[163] Rella and colleagues showed that blocking this protein led to accumulation of steryl β-glucoside, and enhanced immunomodulatory properties.[165] Ergosteryl β-glucoside is linked to immunological protection of the host in an animal model of fungal disease, even though the underlying molecular mechanisms remain unknown.[165]

### 2.1.2 Several sterols are recognised by Mincle, inducing immunological effect to human health

Beside cholesterol[68] and cholesterol sulfate,[69-71] other sterols that can signal through human Mincle are the plant sterol sitosterol and desmonsterol (a cholesterol intermediate), but not sterol ergosterol or cholestanoic acid from yeast. These results suggest that a hydroxyl group at C3 and an alkyl chain at C17 appear to be minimally, but not exclusively required for recognition by human Mincle. A range of other steroids that do not signal through Mincle have been described: cortisone, progesterone, estradiol, testosterone, and aldosterone (Figure 2.2).

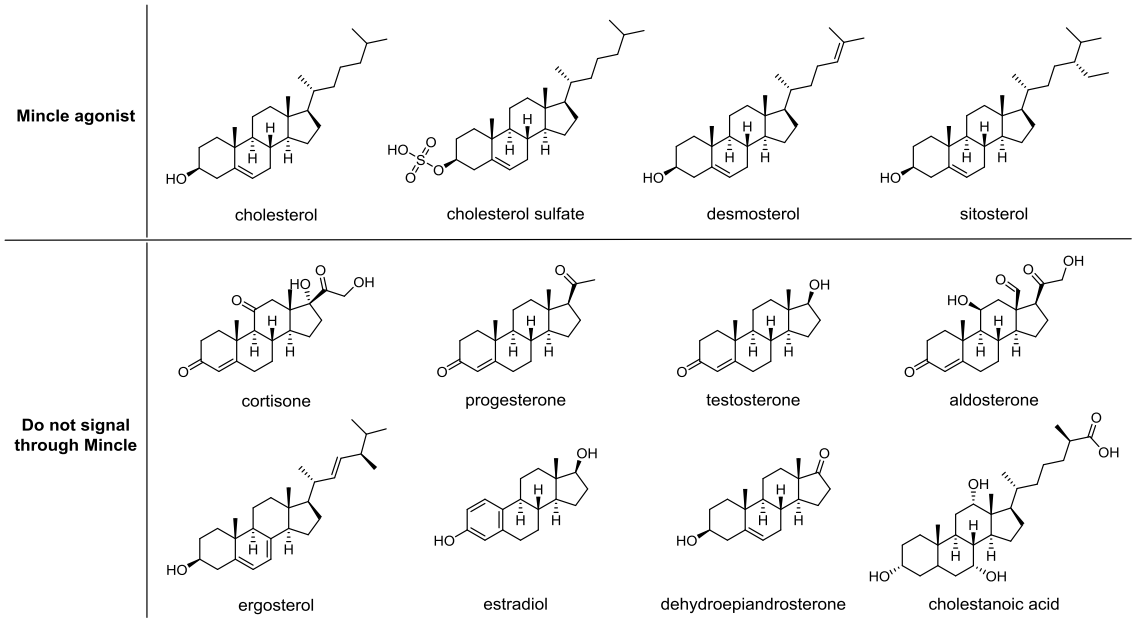



**Figure 2.2:** Structure of sterols that act as hMincle agonists, and those which do not.

Recognition of cholesterol is limited to human Mincle; rodent Mincle from rat and mouse cannot sense cholesterol.[68] The primary difference between human and mice Mincle is the presence of a hydrophobic cholesterol recognition/interaction amino acid consensus (CRAC) motif in the former.[166-168] An amino acid sequence comprising this motif is present in human Mincle ($L^{127}SYKKPKMR^{135}$) but not in rodent Mincle. According to a simulation based on the crystal structure of human Mincle, the CRAC-

like sequence appears to form a pocket that may interact with cholesterol. Mutagenesis of the CRAC consequence abolished cholesterol signalling through Mincle.[68]

Specifically, replacement of Arg[135] in this motif with Leu, which is known to abolish cholesterol binding,[166] gave a mutant human Mincle that expressed normally on the cell surface, but failed to recognise cholesterol while maintaining the ability to recognize TDM. This suggests that the binding site of cholesterol in human Mincle differs from the binding site for TDM.[169-170]

### 2.1.3 Cholesteryl 6-*O*-acyl-α-glucosides (αCAG) from *Helicobactor pylori* signal through Mincle

*Helicobacter* is Gram-negative, helically-shaped, microaerophilic bacterium. The family contains around 35 species, of which *Helicobactor pylori* (*H. pylori*) is the most prominent through its relevance to human health. It was discovered in 1982 to be the cause of chronic gastritis and gastric ulcers, a disease not previously believed to have a microbial cause.[171] This discovery led to the Nobel Prize in Physiology or Medicine (2005), awarded to Australian scientists Barry Marshall and Robin Warren.[172-173] *H. pylori* infects as much as 50% of the world's population yet only 20% show symptoms.[174]

Studies on the chloroform-methanol-extracted lipids of *H. pylori* led to the discovery of a range of cholesteryl glucosides that account for about 25% (w/w) of the total lipid.[175-177] These molecules include cholesteryl α-glucoside (αCG), and various modified glucosides such as cholesteryl 6-*O*-acyl-α-glucosides (αCAG) and cholesteryl 6-*O*-phosphatidyl-α-glucosides (αCPG).[175] The lipid chain composition in the αCAG and αCPG differs among the individual *Helicobacter* species, and includes saturated, and unsaturated and fatty acyl chains.

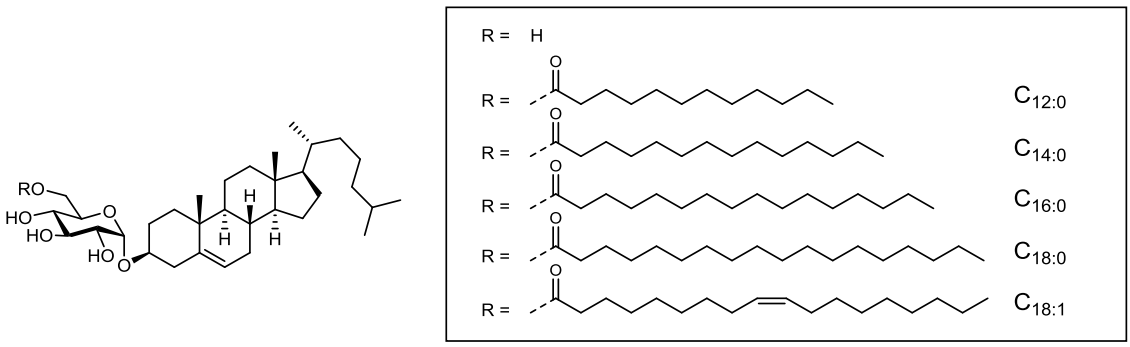
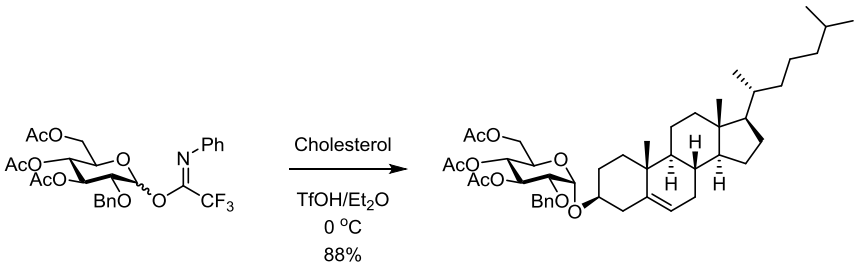



**Figure 2.3:** Structure of acyl variants of the αCAGs from the *Helicobacter* spp.

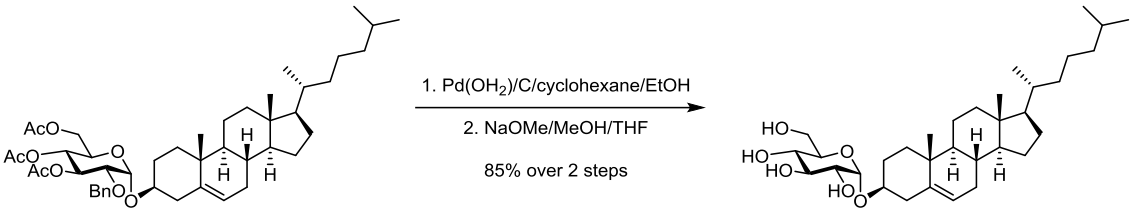
### 2.1.4 Synthesis of cholesteryl 6-*O*-acyl-α-glucoside (αCAGs) from *H. pylori*

The similarities between the known Mincle agonists and the cholesteryl glucosides produced by *Helicobacter* spp. are unmistakable. Dr Dylan Smith from our laboratory synthesized a range of αCAGs from *Helicobactor pylori*, including the $C_{12:0}$, $C_{14:0}$, $C_{16:0}$, $C_{18:0}$, $C_{18:1}$ variants to explore their potential for signalling through Mincle. Two synthetic challenges for the preparation of cholesteryl 6-*O*-acyl-α-glucosides (αCAG) include achieving α-anomer stereoselectivity in the glycosylation and the regioselective installation of the fatty acid ester on O-6. Dr. Dylan Smith in our group developed an improved method for the preparation of cholesteryl 6-*O*-acyl-α-glucosides (αCAG).[unpublished] A 3,4,6-tri-*O*-acetyl-2-*O*-benzylglucosyl *N*-phenyltrifluoroacetimidate donor provided good α-selectivity for cholesterol with a 14:1 α/β ratio. Recrystallization gave the pure α-anomer.



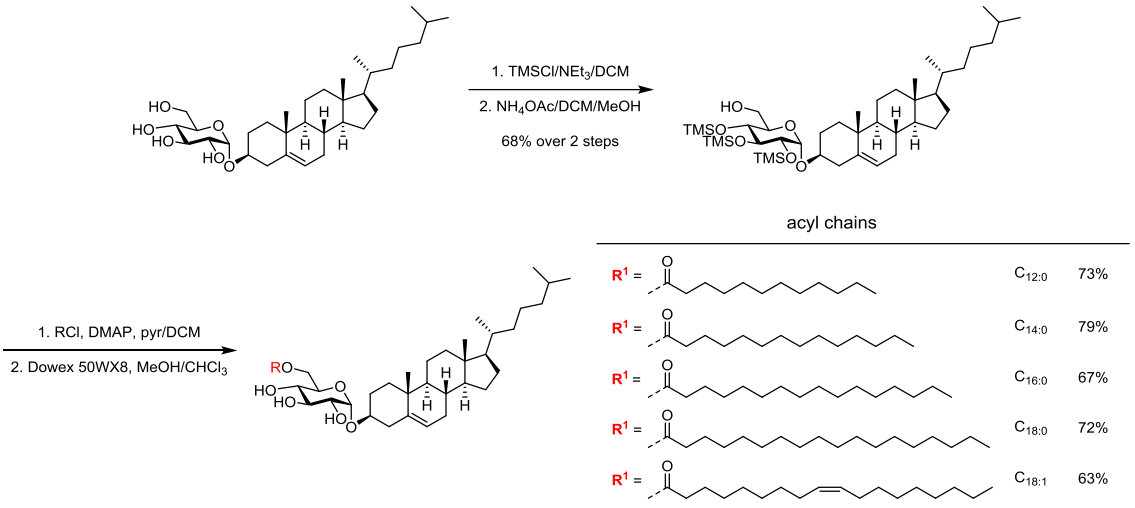

**Scheme 2.2:** Glucosylation of cholesterol with an optimised glucosyl donors.

The benzyl groups were removed by reduction with $Pd(OH)_2/C$ in cyclohexene and EtOH, followed by Zemplén transesterification of the acetates, affording 85% yield of the tetraol while maintaining the integrity of the cholesterol double bond.[178]




**Scheme 2.3:** Removal of the protecting groups.

A method developed by Gervay-Hague and colleagues was applied to gain access to a protcted derivative with a free 6-hydroxyl.[179-180] This involved installation of four TMS ethers and selective cleavage of the primary ether with $NH_4OAc$ (68% yield). Acylation with the appropriate acyl chloride in the presence of pyridine, followed by removal of the delicate TMS groups by treatment with Dowex 50WX8-200 strong acid resin in methanol, gave 63-79% yield of the $C_{12:0}$, $C_{14:0}$, $C_{16:0}$, $C_{18:0}$, $C_{18:1}$ αCAGs.




**Scheme 2.4:** Installation of the 6-*O* acyl chain.

The laboratory of Prof. Yamasaki showed that synthetic αCG is an agonist of human and, more weakly, mouse Mincle. On the other hand, all lipoforms of the αCAGs stimulated cells through both mouse and human Mincle with similar potency to TDM.

In conclusion, the αCGs and the αCAGs from *Helicobactor pylori* are the first naturally occurring Mincle agonists that contain both a carbohydrate and cholesterol moiety. In all cases signalling occurred through mouse and human Mincle, suggesting that cholesteryl glucosides are not being recognised in the same way as cholesterol, and do not require the CRAC motif.[68]

### 2.1.5 *Candida albicans*: a human gut pathogen, responsible for candidiasis disease

The ability of the αCG and αCAGs from *H. pylori* to agonize Mincle signalling brought our attention to another human pathogen, *Candida albicans*. *C. albicans* causes infections of the skin, genitals, throat, mouth and blood. More generally, the genus *Candida* spp. contains the most common yeasts to cause fungal infections in humans.[181] Over 200 species have been described within the *Candida* genus. *Candida* spp. are typically found on mucosal surfaces and the skin. Other significant pathogens of this genus include *Candida glabrata*, *Candida tropicalis*, *Candida parapsilosis* and *Candida krusie*.[182] Among these, *C. albicans* is the most significant species responsible for the majority of *Candida* bloodstream infection.[183]

*C. albicans* is found in the gastrointestinal tract and mouth of 40–60% of healthy adults without causing any overt symptoms of infection.[184] However, when the immune system is compromised, it can invade and cause opportunistic infections.[185] For immunocompromised individuals diagnosed with serious diseases such as HIV and cancer, it is an important nosocomial infection. Invasive candidiasis is responsible for 2,800 to 11,200 deaths yearly in US, with *C. albicans* alone accounting for approximately half of the mortality attributed to systemic forms of the disease.[186]

### 2.1.6 Immunological properties of the isolated lipids from *C. albicans*

In 2008, Ashman and colleagues demonstrated that Mincle was a receptor for *C. albicans*, and plays a significant role in the mammalian immune response against this yeast.[76] Three clinical strains of *C. albicans* isolated from patients suffering with candidiasis were shown to bind to Mincle and contributed to the production of the inflammatory cytokine TNF-α in macrophages. When Mincle was blocked by an affinity-purified polyclonal Ab targeting Mincle, the production of TNF-α was significantly reduced. Mice lacking Mincle under *C. albicans* infection challenge had a significantly greater fungal burden in the kidney than those in the control. From these results, they concluded that Mincle is a receptor for *C. albicans*, and rapidly induced inflammatory cytokines TNF-α in macrophages exposed to the yeast.

A contemporary study by Yamasaki and co-workers studied the ability of fungi to signal through Mincle.[38] Extracts of >50 species of pathogenic fungi were screened for agonism of Mincle signalling; only *Malassezia* fungi were active. Three different strains of *C. albicans* within the set did not activate Mincle-expressing cells. Yamasaki speculated that a possible reason for the conflicting results with the study of Ashman *et al.* might be the use of different strains by the two groups. An alternative explanation could be that differences in growth conditions exhibited a wide range of different morphological phenotypes, and therefore activated Mincle differently.

The ligands generated from *C. albicans* have not yet been identified; however, it is known that *C. albicans* contains cell wall mannan with terminal β-1,2-linkage[187] that share some similarity with known Malassezia ligands, and also contains the sugars *N*-acetyl-glucosamine (as part of chitin), *N*-acetylgalactosamine, as well as sialic acid and rhamnose.[188] Possibly, the CRD of Mincle may bind these carbohydrates, but this does not explain the differences between the two studies described above.

### 2.1.7 *C. albicans* possess a series of cholesteryl 6-*O*-acyl mannosides (αCAM) and ergosteryl 6-*O*-acyl mannosides (αEAM)

*C. albicans* exhibits a wide range of different morphological phenotypes that undergo high-frequency switching, most notably yeast-to-hyphae. Lipid production by the yeast and hyphal forms *C. albicans* at various temperatures can vary dramatically.[189-192] Radwan and co-workers co-cultured the yeast and mycelial forms of *C. albicans* at 37 ºC in SSV medium[191] for 96 h. The yeast and mycelial forms were collected separately and the lipids extracted and characterized.[193] The total mass of lipid-extract was 1.3% dry weight for yeast and 6.3% for the mycelial form. Both yeast and mycelial forms accumulated a range of steryl glycosides and esterified steryl glycosides, which made up 11-15% of the total lipids. The steryl glycosides were separated by preparative TLC and subjected to acid hydrolysis, generating a hexose and a sterol mixture in the ratio of 1:1. By combining GLC and mass spectroscopy, the sterol was identified mostly as cholesterol, with a small amount of ergosterol and zymosterol; whereas the hexose was determined as mannose by paper chromatographic analysis. Fatty acid analysis identified polyunsaturated lipids (14:1, 16:1, 18:1, 18:2, 18:3) and saturated lipids (12:0, 14:0, 16:0, 18:0). The major glycolipid components of the two morphological forms were almost identical with the *C. albicans* KCCC 14172, a clinical isolate from the oral cavity of a cancer patient.[193]

Steryl glycosides and esterified steryl glycosides are widely distributed in the lipids of higher plants and plant cell cultures,[194] but reports of their occurrence in lipids of micro-organisms are rare.[195] Therefore the identification of these lipids is unusual.

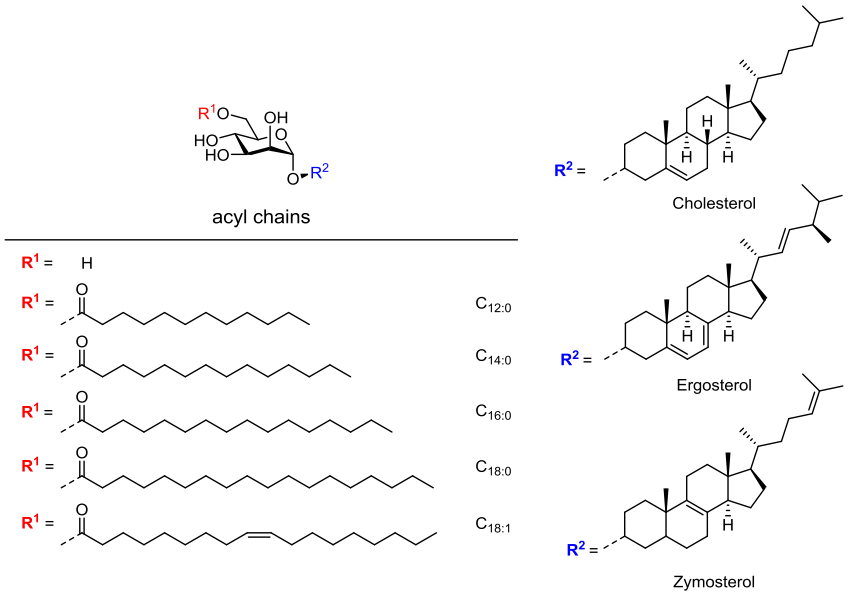



**Figure 2.4:** Proposed structure of esterified steryl glycosides isolated from *C. albicans*: cholesteryl 6-*O*-acyl-α-mannosides (αCAM), ergosteryl 6-*O*-acyl-α-mannosides (αEAM) and zymosteryl 6-*O*-acyl-α-mannosides.

### 2.1.8 Activation of iNKT cells stimulated with αCAMs

In 2011, Illarionov and co-workers demonstrated that cholesteryl 6-*O*-acyl glucosides (CAGs) extracted from the *Helicobacter pylori* could induce immune responses from iNKT cells in a CD1d-dependent manner.[196] It is not known if these CAGs can bind to CD1d and be recognized by NKT cell TCRs, or if other mechanisms of action apply.

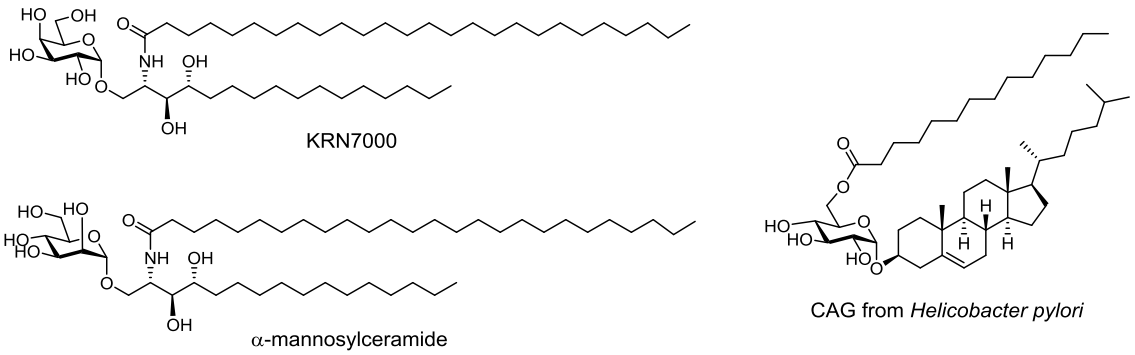



**Figure 2.5:** Structure of KRN7000, α-mannosylceramide, and CAG from *Helicobacter pylori*, which was previously studied the ability to stimulate iNKT cells activation.

As mentioned in the introduction, α-mannosylceramide is inactive toward iNKT cells, leading to the conclusion that an axial configuration of the 2'-hydroxyl group in the sugar is detrimental for their activity. In 2017, Illarionov and co-workers reported the first examples of mannose-containing glycolipid activators for iNKT cells. α-Mannosyl-1,3-(6-acyl-α-mannosyl) 1-monoacylglycerol (M-AcM-MAG) found in *Saccharopolyspora*[197] and cholesteryl 6-*O*-myristoyl α-mannoside (CAM-$C_{14:0}$)[193, 198] can induce the activation of iNKT cells through CD1d-dependent manner.[199]

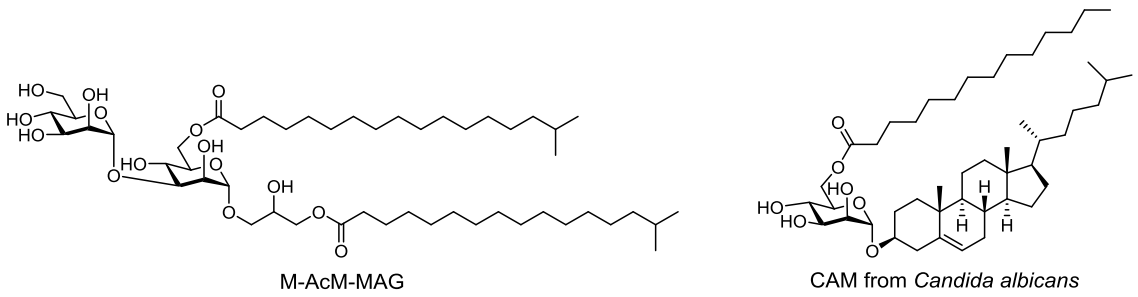



**Figure 2.6:** The first examples of α-mannosyl containing activators for iNKT cells.

### 2.1.9 Are *C. albicans* cholesteryl 6-*O*-acyl-α-mannosides (αCAM) and ergosterol 6-*O*-acyl-α-mannosides (αEAM) agonists for Mincle signalling?

The conflicting results between two studies by Ashman *et al*. and Yamasaki *et al*. raise questions as to the identity of the *C. albicans* metabolites that signal through Mincle. As *C. albicans* exhibits a range of different morphological phenotypes, the conflict between these two studies may have their origin in differences between the morphological forms studied by the two groups.

Our studies (with Yamasaki group) showed that cholesteryl α-glucosides (αCG) and cholesteryl 6-*O*-acyl-α-glucosides (αCAG) from *Helicobacter pylori* are agonists for Mincle signalling. Given the structural similarity between cholesteryl 6-*O*-acyl-α-glucosides (αCAG) from *H. pylori* and cholesteryl 6-*O*-acyl-α-mannosides (αCAM) from *C. albican*s, we hypothesized that αCAM could result in signalling through Mincle. This suggestion was bolstered by the fact that both mannose monomycolate a

complex mannose glycolipid from *Malassezia pachydermatis* are potent agonists for Mincle signaling.

### 2.1.10 Research aims

The goal of this chapter is to examine whether representative *C. albicans* mannosyl sterols can signal through Mincle. The specific aims are to:

**Aim 1**: Synthesize α-mannosides of cholesterol and ergosterol.

**Aim 2**: Synthesize acyl variants of the αCAM and αEAM from *C. albicans*: the $C_{12:0}$, $C_{14:0}$, $C_{16:0}$, $C_{18:0}$, $C_{18:1}$ lipoforms.

**Aim 3**: Investigate the ability of αCAMs and αEAMs to agonize Mincle signalling in collaboration with Yamasaki and colleagues at Osaka University.

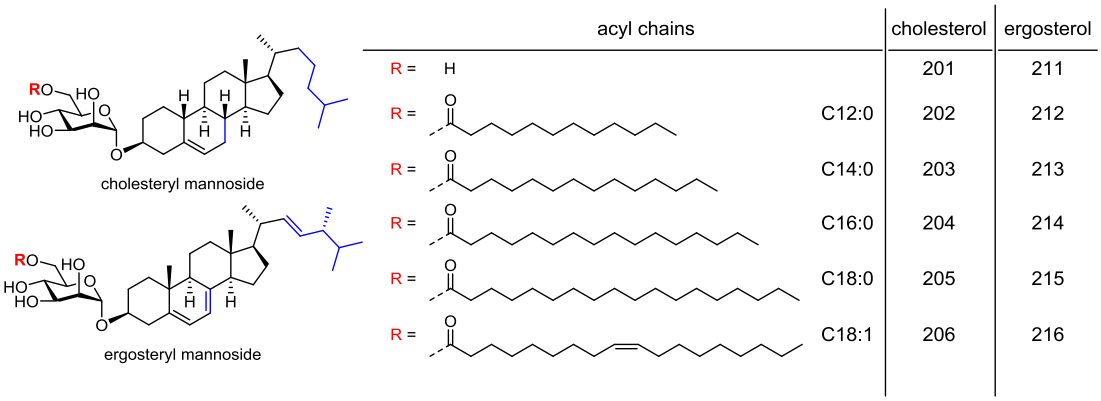


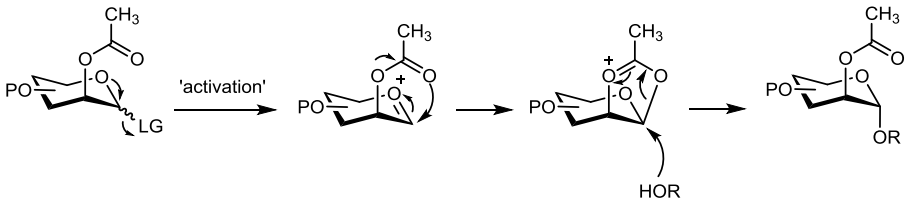

**Figure 2.7:** Proposed targets: cholesteryl α-mannosides (αCM), cholesteryl 6-*O*-acyl-α-mannosides (αCAMs), ergosteryl α-mannosides (αEM), and ergosterol 6-*O*-acyl-α-mannosides (αEAMs).

## 2.2 Results and discussion

### 2.2.1 Formation of *cis*-mannoside (α-mannosides)

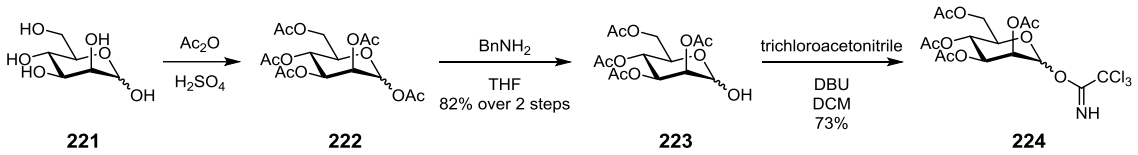
Following the elegant strategies developed by Dr Dylan Smith in our group, we proposed a synthetic route to prepare cholesteryl 6-*O*-acyl-α-mannosides (αCAM) and ergosterol 6-*O*-acyl-α-mannosides (αEAM). Unlike the preparation of a *cis*-glucoside, the formation of *cis*-mannosides (α-mannosides) is relatively straightforward utilizing

neighbouring group participation. Typically, an acetyl (Ac) or benzoyl (Bz) protecting group at the C2 position can promote α-selective glycosylation.[200] We proposed to explore both acetylated and benzoylated mannosyl trichloroacetimidates to evaluate which was optimal.




**Scheme 2.5:** Mechanism of the neighbouring participation by the C-2 acyl functionality.

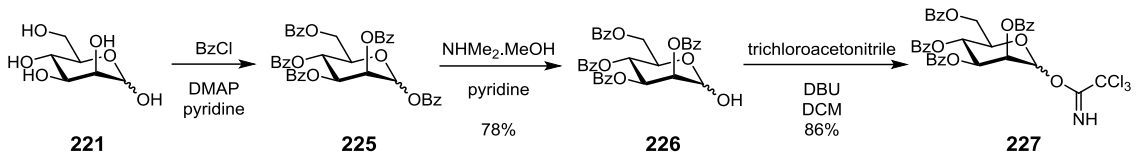
Preparation of the tetra-acetate mannosyl donor **224** started from the reaction of D-mannose **221** with acetic anhydride catalyzed by $H_2SO_4$, affording D-mannose pentaaceate **222**. Anomeric deacetylation utilizing benzylamine delivered the hemiacetal **223** in 82% yield. Treatment of **223** with trichloroacetonitrile and DBU in $CH_2Cl_2$ gave the tetraacetate trichloroacetimidate **224** in 73% yield (Scheme 2.6).




**221** **222** **223** **224**

**Scheme 2.6:** Preparation of the tetraacetate trichloroacetimidate **224**.

Preparation of the tetra-benzoate mannosyl donor started from the reaction of D-mannose **221** with benzoyl chloride in DMAP/pyr, affording the D-mannose pentabenzoate **225**. Anomeric debenzoylation utilizing dimethylamine/MeOH in pyridine afforded the hemiacetal **226** in 78% yield. Treatment with trichloroacetonitrile and DBU in $CH_2Cl_2$ gave the trichloroacetimidate donor **227** in 86% yield (Scheme 2.7).



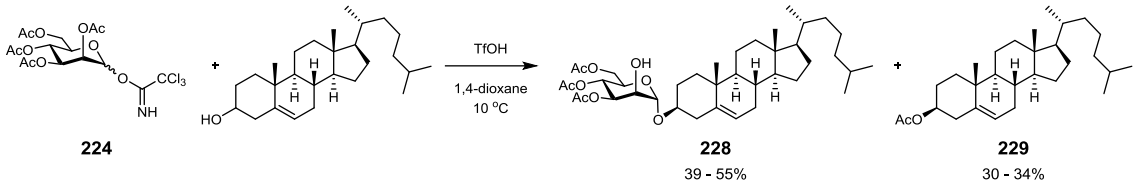

**221** **225** **226** **227**

**Scheme 2.7:** Preparation of the tetrabenzoate trichloroacetimidate.

### 2.2.2 Preparation of cholesteryl 6-*O*-acyl-α-mannosides (αCAM) with various fatty acids

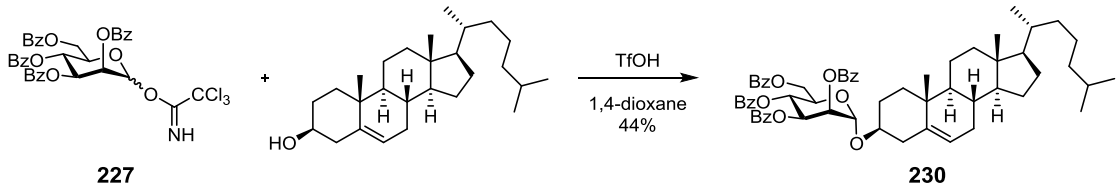
The solubility of cholesterol has long attracted attention, as this enigmatic molecule is highly crystalline and has poor solubility in most solvents at room temperature or below. Cholesterol contains a hydroxyl group and a double bond, thus having the potential for interaction with polar and aromatic solvents.[201] The solubility of cholesterol has been examined in many solvents, including various fats and oils,[202] n-alkanols from methanol to undecanol and their binary mixtures,[203] as well mixtures of benzene with hexane and 1,4-dioxane.[204] Effects of temperature on solubility of cholesterol have also been reported.[205] Unfortunately, none of the solvents identified are ideal for glycosylation reactions, which are usually performed at low temperature.

Our first glycosylation experiment using tetraacetyl donor **224** and TfOH at 0 ºC highlighted the poor solubility of cholesterol in $Et_2O$. We switched to dioxane and noted significantly improved solubility of cholesterol. Upon warming to 10 ºC the solubility of cholesterol improved sufficiently for the glycosylation reaction to proceed. Unfortunately, even after several attempts, no product was isolated. Instead, the deacetylated glycoside **228** was isolated in yields of 39-55% yield. We also isolated 30-34% of acetylated cholesterol **229**. The phenomenon of acetyl group transfer is often seen with acetylated donors, leading to modest yields of glycosides. It arises from decomposition of orthoesters formed by alcohol attack on intermediate dioxolenium ions.[206-207]



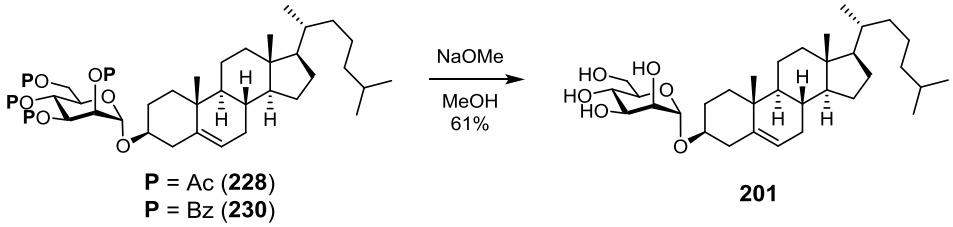

**Scheme 2.8:** Mannosylation with cholesterol afforded the 2-hydroxy glycoside **228** and acetylated cholesterol **229**.

The rate of acyl transfer can often be diminished by using more bulky benzoyl groups. Glycosylation of cholesterol using the tetrabenzoyl trichloroacetimidate **227** in dioxane gave 44% of the glycoside **230**, with no transfer of benzoyl groups observed. While this yield was lower than the best obtained with the tetraacetate trichloroacetimidate **224** (55%), the reaction was more reliable and easily purified, and was preferable for large scale reactions.




**Scheme 2.9:** Mannosylation of cholesterol with tetrabenzoate trichloroacetimidate **227** afforded **230** without formation of acyl transfer by-product.

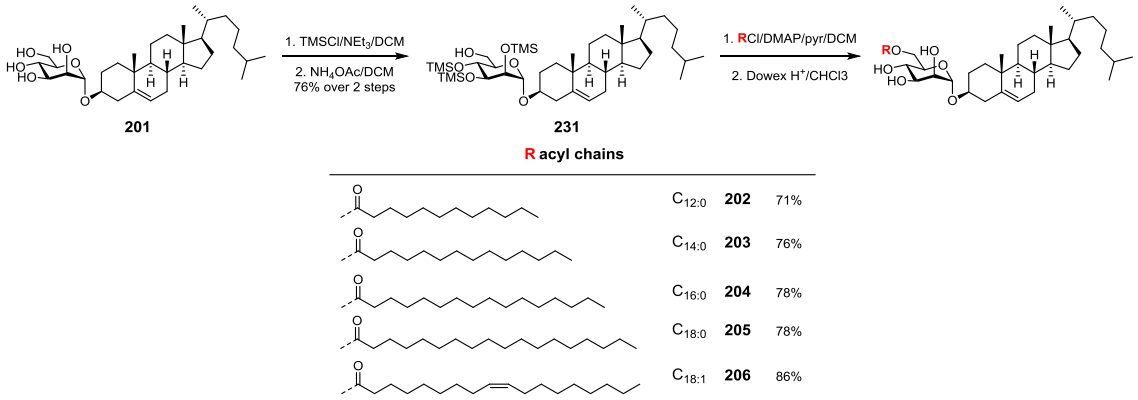
Treatment of **228** or **230** with NaOMe in MeOH, and recrystallization from EtOAc/MeOH afforded tetraol **201** in 61% yield.




**Scheme 2.10:** Deprotection of glycosides **228** and **230** afforded αCM **201**.

Completion of the synthesis of αCAMs required fatty acid acylations at the 6-position. Gervay-Hague and co-workers demonstrated the access to a free hydroxyl at O-6 of a glucoside by using a selective TMS-deprotection.[179-180] Treatment of αCM **201**

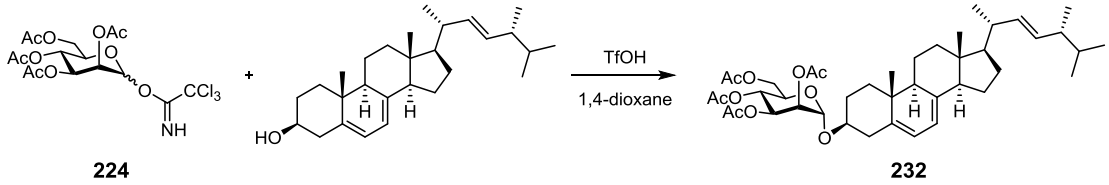
with TMSCl in $NEt_3$ afforded the tetra-TMS ether, which was immediately treated with $NH_4OAc$ in chloroform, effecting selective removal of the 6-*O*-TMS affording **231** in 76% yield. Acylation of **231** was achieved by reaction with the appropriate acyl chloride in the presence of pyridine. Due to the poor stability of TMS ethers, we immediately treated the protected compound with Dowex 50WX8-200 strong acid resin in methanol to cleave the TMS groups. By this means we synthesized the $C_{12:0}$, $C_{14:0}$, $C_{16:0}$, $C_{18:0}$, and $C_{18:1}$ α-cholesteryl mannosides in good yield (Scheme 2.11).




**Scheme 2.11:** Completion of the synthesis of αCAMs.

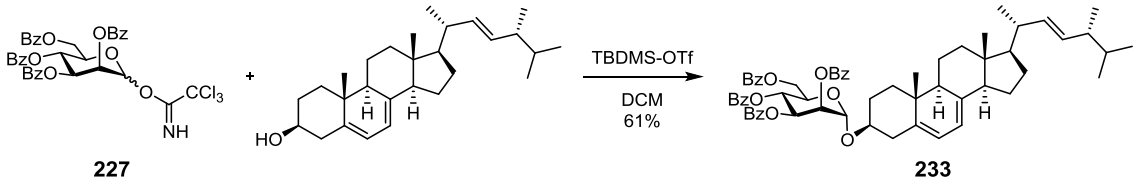
### 2.2.3 Preparation of ergosterol 6-*O*-acyl-α-mannosides (αEAMs) with various fatty acids

A similar approach was attempted for the synthesis of ergosterol 6-*O*-acyl-α-mannosides (αEAMs). Unfortunately, ergosterol exhibits even poorer solubility than cholesterol in dioxane. Nonetheless, in the hope that it would dissolve as the reaction proceeded, the mannosylation was attempted in this solvent under activation by TfOH. However, none of product **232** was isolated.



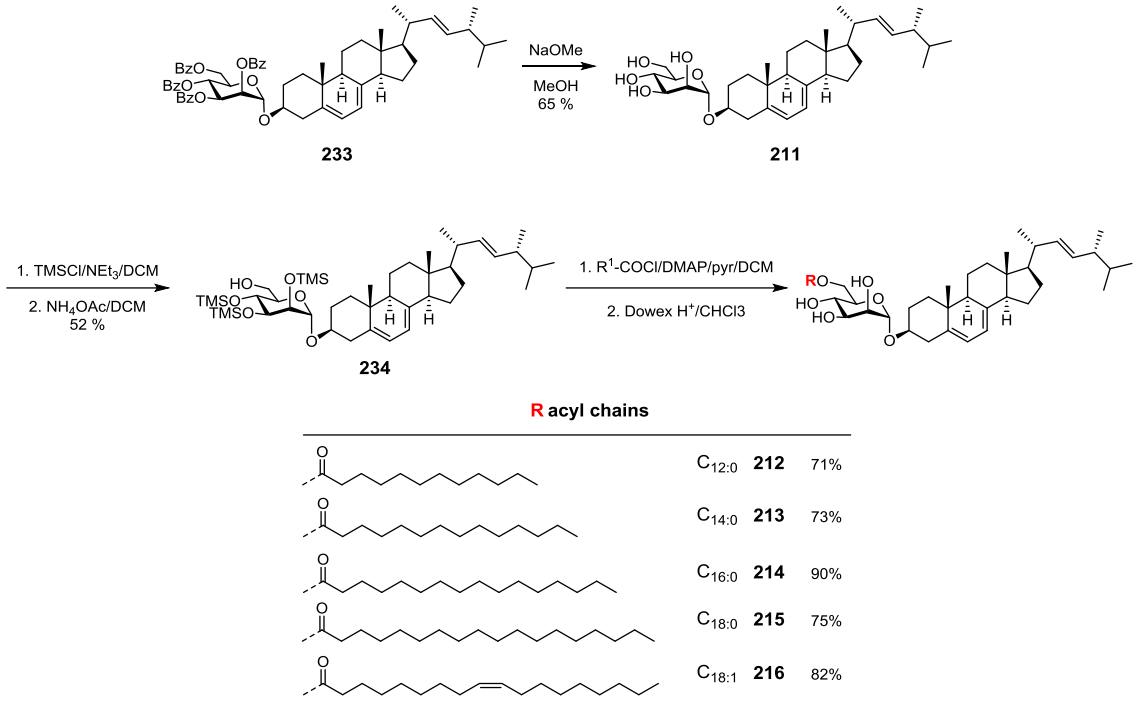

**Scheme 2.12:** Attempted mannosylation of ergosterol by tetraacetyl trichloroacetimidate **224**.

The solubility of ergosterol was examined in a range of solvents including $CH_2Cl_2$, THF, toluene, EtOAc, dioxane and $Et_2O$. Ergosterol completely dissolved in $CH_2Cl_2$ under the conditions of the glycosylation. A glycosylation using tetrabenzoate trichloroacetimidate **227** and ergosterol with TfOH in $CH_2Cl_2$ showed an encouraging result with 12% yield of the product **233.** No migration of the 2-*O*-benzoyl group was observed but despite the low yield, no unreacted ergosterol was recovered. We hypothesised that instability of ergosterol under acidic conditions could contribute to the low yield. The reaction was repeated by replacing TfOH with the more hindered Lewis acid TBDMS-OTf. After several trials and minor optimization, the highest yield obtained was 61% of **233** (Scheme 2.13).




**Scheme 2.13:** Mannosylation of ergosterol by tetrabenzyl donor **227** under promotion with TBDMS-OTf.

Benzoate groups were cleaved using NaOMe and MeOH and gave a 65% yield of the tetraol **211**. Applying the same method described for the αCAMs, tetra-*O*-TMS protected mannoside was obtained by treatment of tetraol **211** with TMSCl in $NEt_3$. This was immediately treated with $NH_4OAc$ in $CH_2Cl_2$ to effect selective removal of the 6-*O*-TMS, affording **234** in 52% yield (Scheme 2.14).




**Scheme 2.14:** Final steps in the synthesis of αEM and αEAMs.

The completion of this work identified challenges in handling αEM **234**, which unlike the αCM analogue **231**, easily decomposed upon exposure to sunlight, and acid, and even upon storage at room temperature. To avoid these problems, reactions were conducted in the dark, and the workup and purification process was performed at night. Retention of the TMS groups required the use of basic silica preparing by pre-treating with 4% triethylamine. Solvents were evaporated under reduced pressure at r.t, instead of 40 $^{o}$C.

We observed decomposition of **234** after just a few hours of storage at different temperatures, varying from r.t to -80 $^{o}$C. Therefore, the material **234** need to be used immediately, by treatment with the appropriate acyl chloride in the presence of pyridine to generate the protected αEAMs. Due to the poor stability of TMS ethers, we immediately treated the TMS-protected αEAMs with Dowex 50WX8-200 strong acid resin in methanol to cleave the TMS groups. We were able to synthesise the $C_{12:0}$, $C_{14:0}$, $C_{16:0}$, $C_{18:0}$ and $C_{18:1}$ α-ergosteryl mannosides in good yield (Scheme 2.14).

The final purification required careful optimization owing to the propensity of the ester groups to migrate under acidic and basic conditions. Scrambling was suppressed when utilizing a solvent mixture 0.1% of AcOH in EtOAc for chromatographic purification, and the column was pre-treated with 1% AcOH in pet. spirit. The αEAMs **212-216** slowly decomposed over time and aside from the period during which they were couriered to Prof. Yamasaki's laboratory, were stored at -80 ˚C in the dark, prior to testing as Mincle agonists.

### 2.2.4 *C. albicans* steryl mannosides signal through Mincle

The ability of αCM, αCAM, αEM and αEAM to induce signalling through Mincle was studied in the laboratory of Sho Yamasaki (Division of Molecular Immunology, Kyushu University, Japan). Yamasaki and co-workers have engineered a T cell hybridoma with nuclear factor of activated T cells-green fluorescent protein (NFAT-GFP) reporter capability that expresses FcRγ and is activated by signalling through Mincle. NFAT are a group of transcription factors that are involved in immune responses. When an activating ligand binds and activates Mincle, the signal is transmitted through FCRγ and the hybridoma produces GFP. GFP production by each cell is essentially a binary outcome (that is, cells are either activated or are not), and flow cytometry is used to count the number of activated cells, allowing calculation of the percentage of activated cells as a measure of agonist potency. Reporter cell lines have been developed for human and mouse Mincle. R135L hMincle is a mutant form of human Mincle with replacement of Arg by Leu at position 135 in the amino acid sequence corresponding to the CRAC motif. This mutant can be used to assess whether an agonist binds in the cholesterol binding site.

We firstly investigated whether non-acylated αCM and αEM were agonists for human and mouse Mincle. Plates coated with αCM induced a strong response through human Mincle. αCM signalled more weakly through R135L hMincle, but still significantly greater than the control FcRγ. αCM signalled weakly through mouse Mincle. In all cases signalling peaked at 100 (pmol/well). On the other hand, αEM did not signal through mouse or human Mincle (or the R135L mutant) (Figure 2.8).

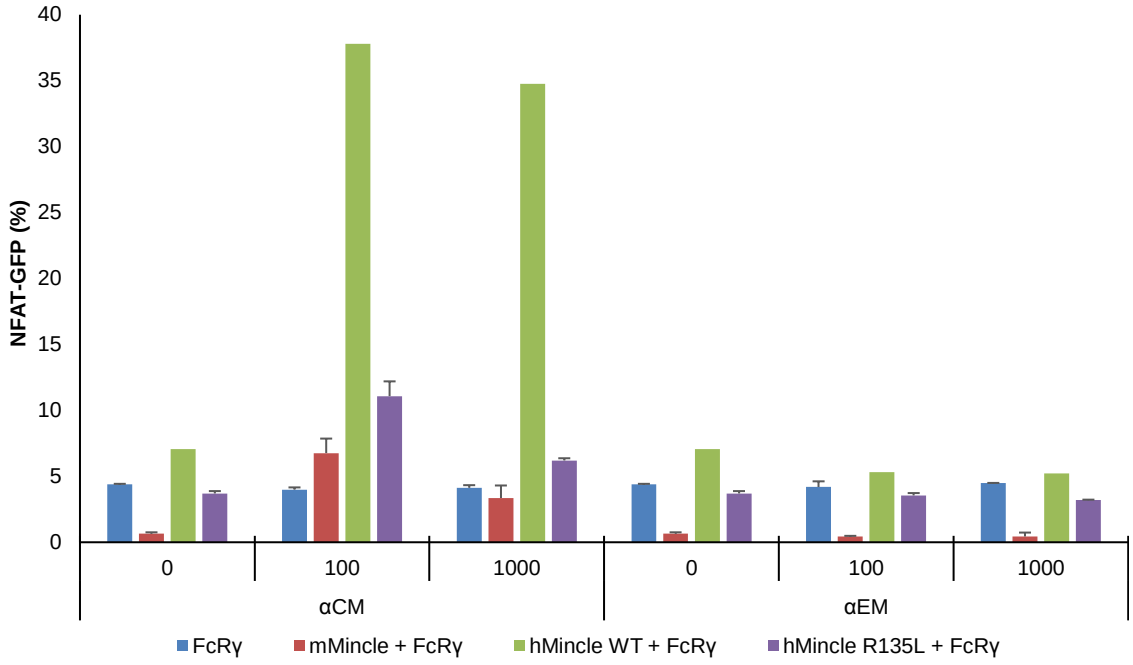



**Figure 2.8:** NFAT-GFP reporter cells expressing either human Mincle/FcRγ or mouse Mincle/FcRγ and mutant R135L hMincle/FcRγ, as well as those expressing FcRγ alone were tested for their reactivity to plate-bound non-acylated αCM and αEM. Assays were performed in duplicate; the mean values and standard derivation are shown (pmpl/well).

Next the ability of αCAMs to signal through mouse and human Mincle was investigated (Figure 2.9). All acyl lipoforms were agonists of signalling through mouse and human Mincle with similar potency. mMincle reporter cells were more weakly stimulated than hMincle reporter cells; however, this likely represents differences in the sensitivities of the reporter cell lines than intrinsic differences in recognition/signalling. In addition, all acyl lipoforms weakly signalled through mutant R135L hMincle. These

results demonstrate that the interaction site of human Mincle with αCAM is not the CRAC motif.

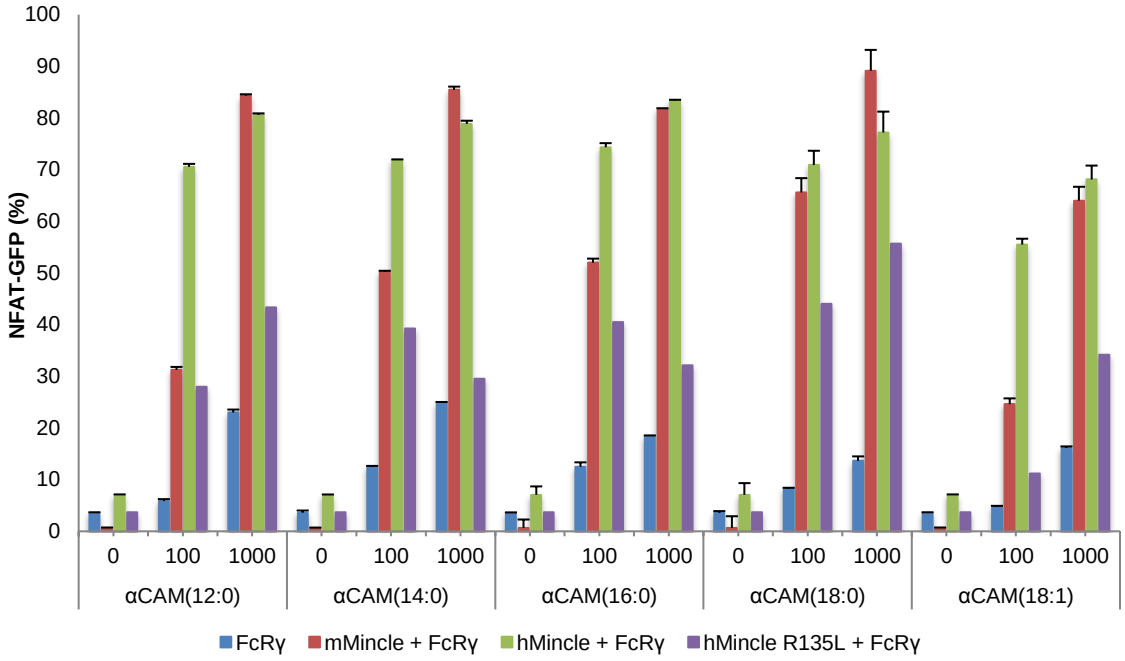



**Figure 2.9:** NFAT-GFP reporter cells expressing either human Mincle/FcRγ or mouse Mincle/FcRγ and mutant R135L hMincle/FcRγ, as well as those expressing FcRγ alone were tested for their reactivity to plate-bound αCAM($C_{12:0}$), αCAM($C_{14:0}$), αCAM($C_{16:0}$), αCAM($C_{18:0}$), αCAM($C_{18:1}$). Assays were performed in duplicate; the mean values and standard deviations are shown (pmpl/well).

All αEAM acyl variant lipoforms signalled through both mouse and human Mincle (Figure 2.10). The $C_{14:0}$-$C_{18:0}$ lipoforms produced robust responses, while the shorter chain $C_{12:0}$ and the unsaturated $C_{18:1}$ produced weaker signals. αEAM signalled more weakly through mutant R135L hMincle, compared to wildtype human Mincle. As αEAM can signal through human and mouse Mincle, while αEM cannot, the acyl chain is clearly required for signalling.

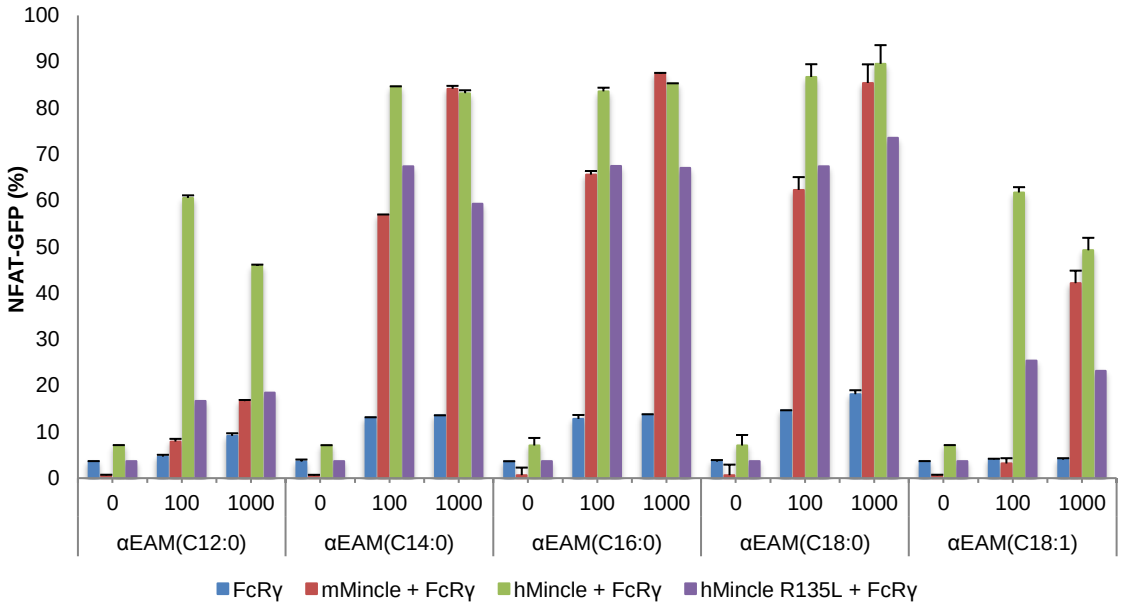



**Figure 2.10:** NFAT-GFP reporter cells expressing either human Mincle/FcRγ or mouse Mincle/FcRγ and mutant R135L hMincle/FcRγ, as well as those expressing FcRγ alone were tested for their reactivity to plate-bound αEAM($C_{12:0}$), αEAM($C_{14:0}$), αEAM($C_{16:0}$), αEAM($C_{18:0}$), αEAM($C_{18:1}$). Assays were performed in duplicate; the mean values and standard deviations are shown.

## 2.3 Conclusion

In this chapter we describe the synthesis of acyl variants of the cholesteryl and ergosteryl α-mannoside from *Candida albicans*. These molecules were synthesised by mannosylation of cholesterol and ergosterol using a mannosyl donor bearing benzoate protecting groups to promote α-selectivity. The highly crystalline nature of these materials allowed purification by crystallization. Deprotection afforded αCM and αEM, and subsequent elaboration of cholesteryl and ergosteryl α-mannoside to a series of 6-*O*-acyl derivatives afforded representative αCAMs and αEAMs from *C. albicans*.

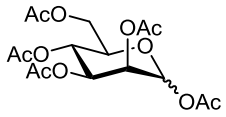
In collaboration with Yamasaki and colleagues at Osaka University, the analogues were assessed for their ability to agonize mouse and human Mincle signalling in NFAT-GFP reporter cell lines. αCM and αCAMs produced a very strong signalling through

both mouse and human Mincle. αEAMs signalled robustly through mouse and human Mincle; however, αEM did not. Thus, an acyl chain is required for signalling through Mincle for ergosteryl mannosides. All mannosyl sterols that could signal through human Mincle could also signal through a mutant human Mincle (R135L), showing that these compounds do not require a CRAC motif. That all acyl analogues signalled more strongly than the corresponding simple mannosides is consistent with structure activity relationships for glucosides that show a lipid chain at C1 and C6 provides superior signalling.[208] Binding of trehalose to Mincle involves the coordination of O3 and O4 with $Ca^{2+}$ in the CRD.[33] Presumably, these mannosides, which share the same configuration at O3 and O4, bind in the same way to this $Ca^{2+}$. Collectively, these data provide evidence for the ability of a previously unappreciated set of fungal metabolites to signal through Mincle, and provide the identity of the likely ligands responsible for the observation by Wells *et al*. that clinical strains of *C. albicans* can signal through Mincle.[76]

## 2.4 Experimental

The reactions were monitored using thin layer chromatography (t.l.c), performed with Merck Silica Gel 60 $F_{254}$. Detection was by charring with Hanessian' Stain [$Ce(SO_4)_2$ (5 g), $(NH_4)_6Mo_7O_{24}.4H_2O$) (25 g), 10% aq. $H_2SO_4$ (500 mL)] or potassium permanganate [$KMNO_4$ (3 g), $K_2CO_3$ (20 g) 5% aq. NaOH (5 mL) or by visualization in UV light. Flash chromatography was performed using Merck Silica Gel 60 according to Still method;[209] eluting solvents are quoted as volume/volume mixtures. Proton ($^1$H NMR) and proton-decoupled carbon-13 ($^{13}$C NMR) nuclear magnetic resonance spectra were recorded using Varian INOVA 400 MHz spectrometer (Melbourne, Australia). The residual solvent peaks were expressed in parts per millions (ppm) and used as references included deuterochloroform ($CDCl_3$, $^1$H NMR δ 7.26 ppm and $^{13}$C NMR δ

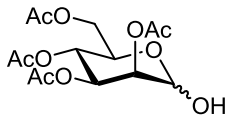
77.16 ppm); methanol-$d_4$ ($CD_3OD$, $^1H$ NMR δ 3.31 ppm and $^{13}C$ NMR δ 49.00 ppm); DMSO-$d_6$ ($^1H$ NMR δ 2.50 ppm and $^{13}C$ NMR δ 39.52 ppm). Splitting patterns are designated as s = singlet, d = doublet, dd = doublet of doublets, m = multiplet, q = quartet, t = triplet. High resolution electrospray ionization mass spectra (HRMS) were recorded using an Agilent ESI-TOF (chemistry) mass spectrometer (Agilent Technologies), The University of Melbourne, Australia. All samples were run using 0.1% formic acid. Optical rotations were obtained using a JASCO DIP-1000 Digital Polarimeter (Melbourne, Australia). $[\alpha]_D$ values are given in $10^{-1}$ $cm^2$ $g^{-1}$. Toluene and pyridine were distilled over $CaH_2$ and NaOH respectively. Acetonitrile, THF, diethyl ether, dichloromethane, methanol and DMF were dried using a Grubbs solvent purification apparatus as described by Pangborn.[210] Ethyl acetate, triethylamine were dried over activated 4 Å molecular sieves. Solvent was evaporated under reduced pressure using a rotary evaporator. Pet. spirits refers to petroleum ether, boiling range 40-60 $^oC$.




**1,2,3,4,6-Penta-*O*-acetyl-mannopyranose (222)**

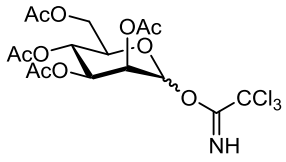
D-Mannose (5.0 g, 27.8 mmol) was added slowly into a stirred mixture of acetic anhydride (16.2 g, 158 mmol, 15 mL) containing drops of concentrated $H_2SO_4$ at 0 ºC and stirring was continued over 30 min. The reaction was concentrated and the residue was extracted with EtOAc, washed with aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). The combined organic layers were concentrated and purified by flash chromatography (EtOAc:pet. spirit, 40:60) to afford protected mannose **222** as the white solid (10.8 g, quantitative yield). $^1H$ NMR (400 Hz, $CDCl_3$) δ 6.08-6.09 (1 H, d, *J* = 1.7 Hz, H1α), 5.34-5.35 (2 H, m, H3, H4), 5.26 (1 H, m, H2), 4.26-4.30 (1 H, dd, *J* = 12.3, 4.9 Hz,

H6a), 4.08-4.12 (1 H, dd, *J* = 12.4, 2.3 Hz, H6b), 4.03-4.07 (1 H, m, H5), 2.18 (3 H, s), 2.17 (3 H, s), 2.09 (3 H, s), 2.05 (3 H, s), 2.00 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 170.8, 170.1, 169.9, 169.7, 168.2, 80.7, 70.8, 68.9, 68.5, 65.7, 62.3, 21.2, 21.0, 20.91, 20.86, 20.80, 20.78. HRMS (ESI) *m/z* calcd. for $[C_{16}H_{22}O_{11}+Na]^+$: 413.1064, obsd: 413.1054.




**2,3,4,6-Tetra-*O*-acetyl-mannopyranose (223)**

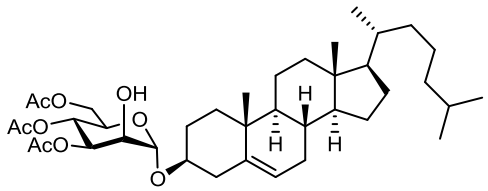
Benzylamine (0.42 g, 3.84 mmol) was added dropwise to the mixture of 1,2,3,4,6-penta-*O*-acetyl-mannopyranose (1.08 g, 2.56 mmol) in dried THF (25 mL) and stirred at rt overnight. The mixture was diluted with cold $H_2O$ and extracted with chloroform. Combined organic layers were washed with ice-cold HCl, aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). The organic layer was concentrated and purified flash chromatography (EtOAc:pet. spirit, 30:70) to afford hemiacetal **223** as the white solid (0.79 g, 82%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.40-5.43 (1 H, dd, *J* = 10.0, 3.3 Hz, H3), 5.24-5.32 (3 H, m, H1, H2, H4), 4.21-4.27 (2 H, m, H5, H6a), 4.10-4.15 (1 H, m, H6b), 3.50-3.51 (1 H, d, *J* = 3.6 Hz, OH), 2.16 (3 H, s), 2.10 (3 H, s), 2.04 (3 H, s), 2.00 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 171.0, 170.3, 170.2, 169.9, 92.3, 70.1, 68.9, 68.7, 66.3, 62.7, 21.1, 20.9, 20.84, 20.85. HRMS (ESI) *m/z* calcd. for $[C_{14}H_{20}O_{10}+Na]^+$: 371.0949, obsd: 371.0949.




**2,3,4,6-Tetra-*O*-acetyl-mannopyranosyl-trichloroacetimidate (224)**

Trichloroacetonitrile (17.47 g, 121 mmol, 12.13 mL) and DBU (2.76 g, 18.2 mmol, 2.7 mL) were added dropwise into the mixture of 2,3,4,6-tetra-*O*-acetyl-mannopyranose (4.21 g, 12.1 mmol) in dried $CH_2Cl_2$ (10 mLmmol$^{-1}$) at 0 ºC and continue stirred at rt

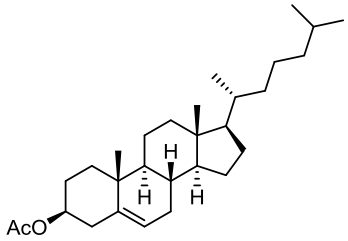
over 4 h. The solvent was evaporated and the residue was purified immediately by flash chromatography (EtOAc:pet. spirit, 50:50 + 1% $NEt_3$) to afford trichloroacetimidate **224** as the colourless oil (4.35 g, 73%). $^1$H NMR (400 Hz, $CDCl_3$) δ 8.79 (1 H, s, NH), 6.28 (1 H, d, *J* = 1.9 Hz, H1), 5.46-5.47 (1 H, m, H2), 5.39-5.41 (2 H, m, H3, H4), 4.26-4.27 (1 H, dd, *J* = 12.0, 4.7 Hz, H6a), 4.18-4.21 (1 H, m, H5), 4.15-4.17 (1 H, dd, *J* = 12.0, 2.3 Hz, H6b), 2.08 (3 H, s), 2.06 (3 H, s), 2.04 (3 H, s), 2.01 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 170.7, 169.94, 169.86, 169.8, 159.9, 95.7, 71.3, 68.9, 68.0, 65.5, 62.2, 20.9, 20.8, 20.7. HRMS (ESI) *m/z* calcd. for $[C_{16}H_{20}Cl_3NO_{10}+Na]^+$: 514.0045, obsd: 514.0054.




**Cholesteryl-3,4,6-ter-*O*-acetyl-α-D-mannopyranoside (228)**

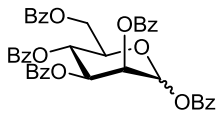
A solution of cholesterol (0.425 g, 1.10 mmol) and acetimidate **224** (0.493 g, 1.00 mmol) in dry dioxane (10 mL/mmol of the donor) was stirred with activated 3 Å molecular sieves at r.t for 30 min. The mixture was then cooled to 0 ºC before TfOH (0.030 g, 18 µL, 0.20 mmol) was added dropwise. The reaction mixture was stirred at 0 ºC for 30 min, at which stage TLC showed that the reaction was complete. The reaction was quenched with $NEt_3$ (0.202 g, 2.00 mmol, 0.279 mL) at 0 ºC. The resulting mixture was diluted with diethyl ether and the solid was removed by filtration through a pad of Celite and wash with diethyl ether. The combined filtrates were washed with aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). The solvent was removed under reduced pressure to give a crude residue which was then purified by flash chromatography (ether:pet. spirit, 5:95) to afford **228** the white solid (0.371 g, 55%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.35-5.36 (1 H, d, *J* = 5.1 Hz, double bond of cholesterol), 5.29-5.31 (2 H, m,

H3, H4), 5.02 (1 H, s, H1), 4.24-4.28 (1 H, dd, *J* = 12.2, 5.1 Hz, H6a), 4.08-4.13 (2 H, m, H5, H6b), 3.99-4.00 (1H, m, H2), 3.49-3.51 (1 H, m), 2.33-2.35 (2 H, d, *J* = 7.6 Hz), 2.08 (6 H, s, 2 x Ac), 2.04 (3 H, s, Ac), 1.82-2.04 (5 H, m), 0.83-1.60 (22 H, m), 1.00 (3 H, s), 0.90-0.92 (3 H, d, *J* = 6.5 Hz), 0.87 (3 H, d, *J* = 6.6 Hz), 0.85 (3 H, d, *J* = 6.6 Hz), 0.67 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 171.0, 170.1, 163.72, 140.8, 140.5, 126.2, 122.2, 121.8, 97.9, 78.1, 71.9, 70.1, 68.5, 66.6, 62.7, 56.9, 56.3, 50.2, 42.4, 40.0, 39.6, 39.4, 36.8, 35.8, 32.0, 28.1, 23.9, 22.9, 22.7, 21.2, 20.8, 19.5, 18.8, 11.9. HRMS (ESI) *m/z* calcd. for $[C_{39}H_{62}O_9+Na]^+$: 697.4286, obsd: 697.4282.




**Acetylated cholesterol** (**229)**

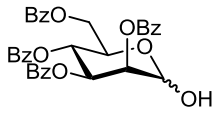
$^1$H NMR (400 Hz, $CDCl_3$) δ 5.36-5.37 (1 H, d, *J* = 4.6 Hz), 4.56-4.64 (1 H, m), 2.30-2.32 (2 H, m), 2.03 (3 H, s), 1.07-1.99 (26 H, m), 1.01 (3 H, s), 0.90-0.91 (3 H, d, *J* = 6.5 Hz), 0.86 (3 H, d, *J* = 6.6), 0.85 (3 H, d, *J* = 6.6 Hz), 0.67 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 170.5, 139.9, 122.6, 74.0, 56.7, 56.1, 50.0, 42.3, 39.7, 39.5, 38.1, 37.0, 36.6, 36.2, 35.8, 31.9, 31.8, 28.2, 28.1, 27.9, 24.3, 23.8, 22.8, 22.6, 21.4, 21.0, 19.3, 18.7, 11.9. HRMS (ESI) *m/z* calcd. for $[C_{29}H_{48}O_2+H]^+$: 429.3727, obsd: 429.3729.




**1,2,3,4,6-Penta-*O*-benzoyl-mannopyranose (225)**

DMAP (33.9 mg, 0.28 mmol) and benzoyl chloride (4.68 g, 33.3 mmol, 3.87 mL) were added respectively into the mixture of D-mannose (1.00 g, 5.55 mmol) in dried pyridine (11 mL) at 0 ºC and stirred at r.t overnight. The reaction was quenched with water and extracted with EtOAc, washed with HCl (1M), aq. $NaHCO_3$, aq. brine and dried

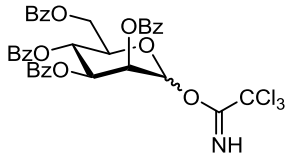
($MgSO_4$). The combined organic layers were concentrated and purified by flash chromatography (EtOAc:pet. spirit, 30:70) to afford protected mannose **225** as the white solid (3.89 g, quantitative yield). $^1$H NMR (400 Hz, $CDCl_3$) δ 8.19-8.21 (2 H, d, *J* = 7.5 Hz), 8.08-8.10 (4 H, d, *J* = 8.2 Hz), 7.95-7.97 (2 H, d, *J* = 7.7 Hz), 7.85-7.87 (2 H, d, *J* = 7.6 Hz), 7.29-7.67 (15 H, m), 6.62-6.63 (1 H, d, *J* = 1.6 Hz, H1), 6.26-6.31 (1H, t, *J* = 10.1 Hz, H4), 6.08 (1 H, dd, *J* = 10.2, 3.2 Hz, H3), 5.90-5.92 (1 H, m, H2), 4.68-4.71 (1 H, dd, *J* = 12.2, 2.3 Hz, H6a), 4.56-4.57 (1 H, m, H5), 4.48-4.52 (1 H, dd, *J* = 12.3, 3.7 Hz, H6b). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 165.9, 165.5, 165.3, 165.1, 165.0, 134.4, 133.5, 133.4, 133.21, 132.9, 130.0, 129.8, 129.63, 129.59, 128.6, 128.5, 128.3, 128.2, 128.2, 91.2, 21.0, 69.8, 69.3, 66.0, 62.2. HRMS (ESI) *m/z* calcd. for $[C_{41}H_{32}O_{11}+Na]^+$: 723.1837, obsd: 723.1834.




**2,3,4,6-Tetra-*O*-benzoyl-mannopyranose (226)**

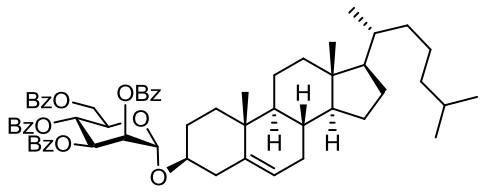
Dimethyamine (5.6 M in ethanol); (83.2 mmol, 14.9 mL) was added dropwise into the mixture of 1,2,3,4,6-penta-*O*-benzoyl-mannopyranose (3.89 g, 5.55 mmol) in dried pyridine (28 mL) and stirred over 1.5 h. The reaction was diluted with toluene and washed with aq. brine and dried ($MgSO_4$). The combined organic layers were concentrated and purified by flash chromatography (EtOAc:toluene, 10:90) to afford hemiacetal **226** as the white solid (2.58 g, 78%). $^1$H NMR (400 Hz, $CDCl_3$) δ 8.10-8.12 (2 H, d, *J* = 7.4 Hz), 8.01-8.03 (2 H, d, *J* = 7.5 Hz), 7.95-7.97 (2 H, d, *J* = 7.4 Hz, 7.83-7.85 (2 H, d, *J* = 7.5 Hz), 7.54-7.58 (2 H, t, *J* = 7.4 Hz), 7.47-7.51 (1 H, t, *J* = 7.3 Hz), 7.33-7.44 (8 H, m), 6.15-6.20 (1 H, t, *J* = 10.1 Hz, H4), 5.99-6.02 (1 H, dd, *J* = 10.2, 3.1 Hz, H3), 5.74-5.75 (1 H, m, H2), 5.33 (1 H, s, H1), 4.75-5.78 (1 H, dd, *J* = 12.2, 2.4 Hz, H6a), 4.65-4.69 (1 H, dt, *J* = 10.0, 2.9 Hz, H5), 4.41-4.45 (1 H, dd, *J* = 12.2, 3.5

Hz, H6b), 3.98 (1 H, s, OH). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 166.2, 165.4, 165.3, 133.2, 133.0, 132.9, 129.6, 129.55, 128.4, 128.3, 128.1, 92.1, 70.7, 69.6, 68.6, 66.7, 62.5. HRMS (ESI) *m/z* calcd. for $[C_{34}H_{28}O_{10}+Na]^+$: 619.1575, obsd: 619.1573.




**2,3,4,6-Tetra-*O*-benzoyl-mannopyranosyl-trichloroacetimidate (227)**

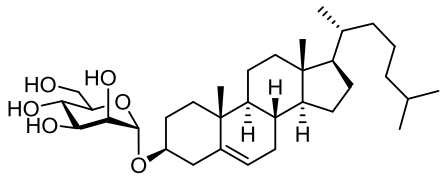
Trichloroacetonitrile (4.91 g, 34.0 mmol, 3.41 mL), and 1,8-diazabicyclo[5.4.0]undec-7-ene (0.259 g, 1.70 mmol, 0.254 mL) were added respectively into the mixture of 2,3,4,6-tetra-*O*-benzoyl-mannopyranose (2.03 g, 3.40 mmol) in dried $CH_2Cl_2$ (34.0 mL) at 0 ºC. The reaction was stirred at rt over 2 h then the solvent was evaporated and the residue was purified by flash chromatography immediately (EtOAc:pet. spirit, 40:60 + 1% $NEt_3$) to afford trichloroacetimidate **227** as a white solid (2.16 g, 86%). $^1$H NMR (400 Hz, $CDCl_3$) δ 8.87 (1 H, s, NH), 8.06-8.10 (4 H, m), 7.96-7.98 (2 H, d, *J* = 7.4 Hz), 7.83-7.85 (2 H, d, *J* = 7.4 Hz), 7.56-7.64 (2 H, m), 7.50-7.54 (1 H, t, *J* = 7.4 Hz), 7.35-7.47 (7 H, m), 7.26-7.30 (2 H, m), 6.57-6.58 (1 H, d, *J* = 1.5 Hz, H1), 6.21-6.26 (1 H, t, *J* = 10.1 Hz, H4), 5.94-5.99 (2 H, m, H2, H3), 4.71-4.75 (1 H, dd, *J* = 12.3, 2.3 Hz, H6a), 4.61-4.66 (1 H, m, H5), 4.49-4.53 (1 H, dd, *J* = 12.3, 4.1 Hz, H6b). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 165.9, 165.4, 165.3, 165.0, 159.8, 133.6, 133.5, 133.2, 133.0, 129.8, 129.7, 129.6, 128.8, 128.7, 128.6, 128.5, 128.4, 128.3, 128.2, 94.6, 71.4, 69.7, 68.7, 65.9, 62.3.




**Cholesteryl-2,3,4,6-tetra-*O*-benzoyl-α-D-mannopyranoside (230)**

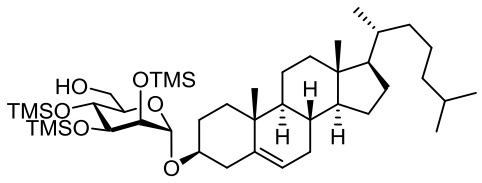
A solution of cholesterol (0.516 g, 1.34 mmol) and acetimidate **227** (0.90 g, 1.21 mmol) in dry dioxane (10 mL/mmol of the donor) was stirred with activated 3 Å molecular sieves at r.t for 30 min. The mixture was then cooled to 0 ºC before TfOH (0.036 g, 21 µL, 0.242 mmol) was added dropwise. The reaction mixture was stirred at 0 ºC for 30 min, at which stage TLC showed that the reaction was complete. The reaction was quenched with $NEt_3$ (0.245 g, 2.42 mmol, 0.337 mL) at 0 ºC. The resulting mixture was diluted with diethyl ether and the solid was removed by filtration through a pad of Celite and wash with diethyl ether. The combined filtrates were washed with aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). The solvent was removed under reduced pressure to give a crude residue which was then purified by flash chromatography (ether:pet. spirit, 5:95) to afford the white solid. This solid was recrystallised from ethanol to afford the pure α-anomer **230** as a white solid (0.514 g, 44%). $^1$H NMR (400 Hz, $CDCl_3$) δ 8.06-8.08 (4 H, t, *J* = 7.5 Hz), 7.97-7.99 (2 H, d, *J* = 8.3 Hz), 7.83-7.85 (2 H, d, *J* = 8.2 Hz), 7.49-7.62 (3 H, m), 7.35-7.45 (7 H, m), 7.25-7.29 (2 H, t, *J* = 7.7 Hz), 6.02-6.07 (1 H, t, *J* = 9.9 Hz, H4), 5.93-5.96 (1 H, dd, *J* = 10.1, 3.2, H3), 5.64-5.65 (1 H, m, H2), 5.27 (1 H, m, CH=), 5.25 (1 H, d, *J* = 1.5 Hz, H1), 4.65-4.68 (1 H, dd, *J* = 11.6, 1.5 Hz, H6a), 4.53-4.57 (1 H, m, H5), 4.47-4.51 (1 H, dd, *J* = 11.7, 5.4 Hz, H6b), 3.59-3.66 (1 H, m, cholesteryl-H3), 2.46-2.48 (2 H, m), 1.82-2.04 (5 H, m), 0.83-1.60 (21 H, m, cholesteryl-CH), 1.03 (3 H, s), 0.92-0.93 (3 H, d, *J* = 6.4 Hz), 0.87 (3 H, d, *J* = 6.6 Hz), 0.86 (3 H, d, *J* = 6.6 Hz), 0.69 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 166.0, 165.4, 165.3, 140.0, 133.3, 133.0, 132.8, 129.7, 129.6, 129.6, 129.3, 129.0,

128.8, 128.4, 128.3, 128.2, 128.1, 122.1, 95.8, 78.4, 71.0, 70.0, 68.8, 67.1, 63.1, 56.6, 56.0, 49.9, 42.2, 39.9, 39.6, 39.4, 36.8, 36.5, 36.0, 35.6, 31.8, 31.7, 28.1, 27.9, 27.7, 24.2, 23.7, 22.7, 22.4, 20.9, 19.3, 19.2, 18.6, 11.7. HRMS (ESI) *m/z* calcd. for $[C_{61}H_{72}O_{10}+H]^+$: 965.5198, obsd: 965.5199.



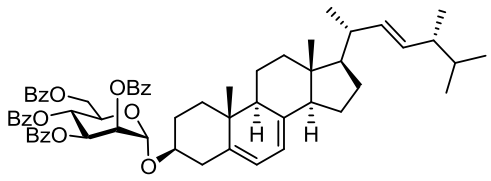

**Cholesteryl-α-D-mannopyranoside (201)**

NaOMe (25% w/w) was added dropwise into the mixture of benzoate protected cholesteryl α-D-mannopyranoside **230** (0.514 g, 0.532 mmol) in dried MeOH:THF (3 mL: 2 mL) to achieve pH = 10. The reaction was stirred at r.t for 1 h which was then quenched with Dowex resin (H+ form), filtered and concentrated to afford a white solid. This solid was recrystallised from EtOAc with drops of methanol to afford the cholesteryl mannoside **201** (0.178 g, 61%). $^1$H NMR (400 Hz, DMSO-$d_6$) δ 5.29-5.30 (1 H, d, *J* = 4.5 Hz, CH=), 4.75 (1 H, d, *J* = 1.5 Hz, H1), 4.67-4.68 (1 H, d, *J* = 4.8 Hz, OH-4), 4.63-4.65 (1 H, d, *J* = 4.4 Hz, OH-2), 4.51-4.52 (1 H, d, *J* = 5.9 Hz, OH-3), 4.39-4.42 (1 H, t, *J* = 5.8 Hz, OH-6), 3.61-3.66 (1 H, dd, *J* = 10.6, 6.1 Hz, H5), 3.54 (1 H, m, H2), 3.29-3.45 (5 H, m, H3,4,6a,6b,cholesteryl-H3), 2.31-2.35 (1 H, dd, *J* = 12.7, 4.3 Hz), 2.16-2.22 (1 H, t, *J* = 11.7 Hz), 1.76-1.97 (5 H, m), 0.99-1.55 (21 H, m, cholesteryl-CH), 0.95 (3 H, s), 0.89-0.90 (3 H, d, *J* = 6.4 Hz), 0.85-0.86 (3 H, d, *J* = 6.6 Hz), 0.83-0.84 (3 H, d, *J* = 6.6 Hz), 0.65 (3 H, s). $^{13}$C NMR (100 Hz, DMSO-$d_6$) δ 140.6, 121.1, 97.9, 75.2, 74.1, 70.9, 70.8, 67.1, 61.4, 56.2, 55.6, 49.5, 41.9, 36.6, 36.2, 35.7, 35.2, 31.4, 31.3, 27.8, 27.4, 23.9, 23.2, 22.7, 22.4, 20.6, 19.1, 18.6, 11.7. HRMS (ESI) *m/z* calcd. for $[C_{33}H_{56}O_6+H]^+$: 549.4150, obsd: 549.4150.




**Cholesteryl-2,3,4-tris-*O*-trimethylsilyl-α-D-mannopyranoside (231)**

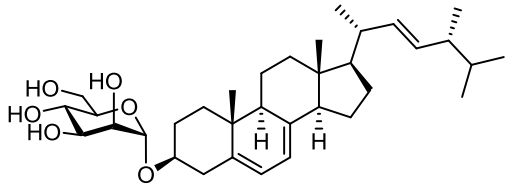
Trimethylsilyl chloride (0.176 g, 1.62 mmol, 0.206 mL) and $NEt_3$ (0.170 mg, 1.68 mmol, 0.235 mL) were added to a stirred mixture of cholesteryl-α-D-mannopyranoside **201** (0.178 g, 0.324 mmol) in $CH_2Cl_2$ (3 mL). The reaction was stirred overnight and then diluted with $CH_2Cl_2$ and aq. $NaHCO_3$. The combined organic layers were separated and washed with sat aq. $NaHCO_3$, water, aq. brine and dried ($Na_2SO_4$) and concentrated. The residue was dissolved in $CHCl_3$ (3 mL) and $NH_4OAc$ (55 mg, 0.71 mmol) was added. The reaction was stirred for next 48 h, then concentrated and purified by flash chromatography (pre-treated with 4% $NEt_3$/pet. spirits, EtOAc:pet spirits 5:95) to afford **231** as a clear glass (248 mg, 76%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.32-5.33 (1 H, dd, *J* = 3.4, 1.8 Hz, CH=), 4.73 (1 H, d, *J* = 1.9 Hz, H1), 3.84-3.88 (1 H, t, *J* = 9.1 Hz, H4), 3.76-3.78 (1 H, dd, *J* = 9.0, 2.5 Hz, H3), 3.73 -3.76 (1 H, dd, *J* = 11.4, 2.7 Hz, H6a), 3.69-3.70 (1 H, m, H2), 3.66-3.68 (1 H, m, H6b), 3.61-3.64 (1 H, m, H5), 2.28-2.31 (2 H, m), 1.94-2.01 (3 H, m), 1.83-1.85 (3 H, m), 0.86-1.60 (22 H, m, cholesteryl-CH), 1.00 (3 H, s), 0.90-0.91 (3 H, d, *J* = 6.2 Hz), 0.86-0.87 (3 H, d, *J* = 6.6 Hz), 0.85 (3 H, d, *J* = 6.6 Hz), 0.67 (3 H, s), 0.16 (9 H, s, 0.14 (9 H, s), 0.12 (9H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 140.8, 122.0, 99.1, 77.5, 74.4, 74.1, 72.6, 68.4, 62.4, 56.9, 56.3, 50.3, 42.5, 40.2, 39.9, 39.7, 37.2, 36.9, 36.3, 35.9, 32.1, 28.4, 28.2, 27.9, 24.4, 24.0, 23.0, 22.7, 21.2, 19.5, 18.9, 12.0, 0.83, 0.79, 0.52.




**Ergosteryl-2,3,4,6-tetra-*O*-benzoyl-α-D-mannopyranoside (233)**

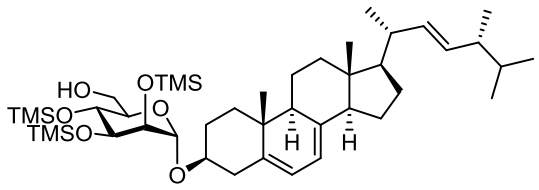
A solution of ergosterol (0.456 g, 1.15 mmol) and trichloroacetimidate **227** (0.741 g, 1.00 mmol) in dried $CH_2Cl_2$ (100 mL/mmol of the donor) was stirred with activated 3 Å molecular sieves at r.t for 30 min. The mixture was then cooled to 0 ºC before TBDMSOTf (0.053 g, 0.200 mmol, 46 µL) was added. The reaction mixture was stirred at 0 ºC for 30 min, at which stage TLC showed that the reaction was complete. The reaction was quenched with $NEt_3$ (0.202 g, 2.00 mmol, 0.278 mL) at 0 ºC. The resulting mixture was diluted with diethyl ether and the solids were removed by filtration through a pad of Celite and wash with diethyl ether. The combined filtrates were washed with aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). The solvent was removed under reduced pressure to give a crude residue which was then purified by flash chromatography (ether:pet. spirit, 5:95) to afford the white solid. This solid was recrystallised from ethanol to afford the pure α-anomer protected **233** as a white solid (0.537 g, 61%). $^1$H NMR (400 Hz, $CDCl_3$) δ 8.06-8.08 (4 H, d, *J* = 7.1 Hz), 7.97-7.99 (2 H, d, *J* = 8.3 Hz), 7.83-7.85 (2 H, d, *J* = 8.3 Hz), 7.29-2.62 (12 H, m), 6.03-6.07 (1 H, t, *J* = 9.8 Hz, H4), 5.93-5.96 (1 H, dd, *J* = 10.0, 2.6 Hz, H3), 5.66 (1 H, m, H2), 5.49-5.50 (1 H, dd, *J* = 5.3, 2.0 Hz, ring-B CH=), 5.36-5.37 (1 H, m, ring-B CH=), 5.38 (1 H, d, *J* = 1.7 Hz, H1), 5.19-5.35 (2 H, t, *J* = 7.1 Hz, C22/23 CH=CH), 4.64-4.66 (1 H, d, *J* = 10.9 Hz, H6a), 4.49-4.57 (2 H, m, H5,6b), 3.71-3.78 (1 H, ddd, *J* = 15.4, 11.0, 4.4 Hz, ergosteryl-H3), 1.20-2.10 (20 H, m, ergosteryl-CH), 1.04-1.05 (3 H, d, *J* = 6.6 Hz), 0.97 (3 H, s), 0.92-0.93 (3 H, d, *J* = 6.8 Hz), 0.83-0.85 (6 H, t, *J* = 7.1 Hz), 0.65 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 166.4, 165.7, 165.7, 141.6, 139.2, 135.7, 133.6, 133.3, 133.2,

133.1, 130.0, 130.0, 129.9, 128.7, 128.6, 128.5, 120.2, 116.5, 96.0, 71.3, 70.3, 69.2, 67.3, 63.4, 55.9, 54.8, 46.3, 43.0, 40.6, 39.2, 38.7, 38.2, 37.4, 33.3, 28.5, 28.3, 23.2, 21.3, 21.2, 20.1, 19.8, 17.8, 16.4, 12.2. HRMS ($ESI^+$) *m/z* calcd. for $[C_{62}H_{70}O_{10}+Na]^+$: 997.4861, obsd: 997.4856.




**Ergosteryl-α-D-mannopyranoside (211)**

NaOMe (25% w/w) was added dropwise into the mixture of benzoate protected ergosteryl α-D-mannopyranoside **233** (0.537 g, 0.550 mmol) in dried MeOH:THF (33 mL: 22 mL) to achieve pH = 10. The reaction was stirred at r.t for 1 h which was then quenched with Dowex resin (H+ form), filtered and concentrated to afford a white solid. This solid was recrystallised from EtOAc with drops of methanol to afford the crystallised **211** (0.200 g, 65%). $^1H$ NMR (400 Hz, DMSO-$d_6$) δ 5.50-5.51 (1 H, d, *J* = 4.2 Hz, ring-B CH=), 5.34 (1 H, m, ring-B CH=), 5.15-5.27 (2 H, qd, *J* = 15.4, 7.4 Hz, C22/23 CH=CH), 4.79 (1 H, s, H1), 4.67-4.69 (1 H, d, *J* = 4.9 Hz, OH-H4), 4.65-4.66 (1 H, d, *J* = 4.4 Hz, OH-H2), 4.52-4.53 (1 H, d, *J* = 5.9 Hz, OH-H3), 4.41-4.44 (1 H, t, *J* = 5.8 Hz, OH-H6), 3.62-3.67 (1 H, dd, *J* = 10.4, 6.0 Hz, H5), 3.55 (1 H, s, H2), 3.37-3.52 (5 H, m, H3,4,6a,6b,ergosteryl-H3), 2.21-2.33 (1 H, t, *J* = 12.9 Hz), 1.21-2.02 (22 H, m, ergosteryl-CH), 1.00-1.02 (3 H, d, *J* = 6.6 Hz), 0.87-0.90 (3 H, d, *J* = 6.8 Hz), 0.87 (3 H, s), 0.79-0.82 (3 H, t, *J* = 6.1 Hz), 0.59 (3 H, s). $^{13}C$ NMR (100 Hz, DMSO-$d_6$) δ 140.9, 140.2, 135.8, 131.9, 119.6, 116.5, 98.3, 74.6, 74.2, 71.3, 71.2, 67.6, 61.8, 55.5, 54.3, 46.0, 42.8, 42.5, 38.8, 38.5, 37.9, 37.1, 32.9, 28.4, 28.1, 23.0, 21.4, 21.0, 20.2, 19.9, 17.8, 16.4, 12.3. HRMS (ESI) *m/z* calcd. for $[C_{34}H_{54}O_6+H]^+$: 581.3813, obsd: 581.3810.



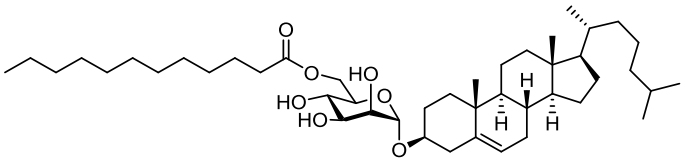

**Ergosteryl-2,3,4-*tert*-*O*-trimethylsilyl-α-D-mannopyranoside (234)**

Trimethylsilyl chloride (0.176 g, 1.62 mmol, 0.206 mL) and $NEt_3$ (0.170 mg, 1.68 mmol, 0.235 mL) were added to a stirred mixture of ergosteryl-α-D-mannopyranoside **201** (0.181 g, 0.324 mmol) in $CH_2Cl_2$ (6 mL). The reaction was stirred overnight and then diluted with $CH_2Cl_2$ and aq. $NaHCO_3$. The combined organic layers were separated and washed with sat aq. $NaHCO_3$, water, aq. brine and dried ($Na_2SO_4$) and concentrated. The residue was dissolved in $CHCl_3$ (6 mL) and $NH_4OAc$ (55 mg, 0.71 mmol) was added. The reaction was stirred for 48 h, then concentrated and purified by flash chromatography (pre-treated with 4% $NEt_3$/pet. spirits, EtOAc:pet spirits 5:95) to afford **234** as a clear glass (130 mg, 52%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.53-5.54 (1 H, d, *J* = 4.4 Hz, ring-B CH=), 5.37 (1 H, s, ring-B CH=), 5.13-5.25 (2 H, m, C22/23 CH=CH), 4.76 (1 H, s, H1), 3.52-3.89 (9 H, m), 2.46-2.51 (1 H, dd, *J* = 14.0, 3.1 Hz), 2.30-2.37 (1 H, m), 1.25-1.89 (17 H, m, ergosteryl-CH), 1.02-1.04 (3 H, d, *J* = 6.7 Hz), 0.94 (3 H, s), 0.91-0.92 (3 H, d, *J* = 6.7 Hz), 0.81-0.84 (6 H, t, *J* = 6.3 Hz), 0.62 (3 H, s), 0.16 (9 H, s), 0.15 (9 H, s), 0.13 (9 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 141.0, 139.6, 135.4, 131.8, 119.5, 116.2, 98.7, 75.1, 74.0, 73.9, 72.3, 68.0, 62.0, 58.6, 55.6, 54.5, 46.0, 45.5, 42.7, 40.3, 39.0, 38.4, 38.0, 37.1, 33.0, 28.1, 28.2, 27.9, 22.9, 21.0, 19.8, 19.5, 17.5, 16.2, 11.9, 10.4, 0.55, 0.52, 0.24.

**General procedure of acylation with cholesteryl and ergosteryl mannoside**

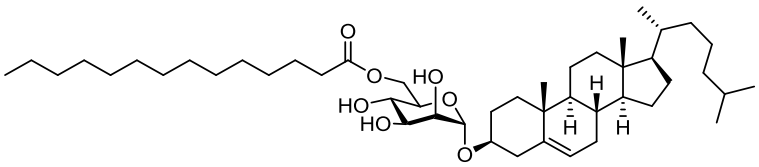
Acid chloride (2 equiv) was added a solution of DMAP (1 equiv) and TMS-protected cholesteryl mannoside **231** or TMS-protected ergosteryl mannoside **234** (1 equiv) in pyridine/ $CH_2Cl_2$ (3:97, 10 mL/mmol) at 0 ºC. The solution was stirred at r.t for 2 h then

MeOH (1 mL/mmol) was added and the mixture was stirred for 0.5 h. The solvent was evaporated and then the residue was redissolved in $CHCl_3$ (10 mL/mmol) and Dowex 50WX8-200 ($H^+$ form, approx. 100 mg) was added. After stirring for 1 h, the mixture was filtered, and the filtrate concentrated. Flash chromatography (EtOAc + 1% AcOH) afforded the product as a white solid.



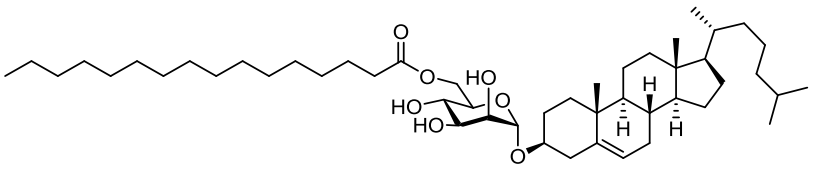
**Cholesteryl-6-*O*-dodecanoyl-α-D-mannopyranoside (202)**

The General Protocol for Acylation between **231** (30 mg, 0.039 mmol) and lauroyl chloride (17.1 mg, 0.078 mmol, 18 μL) after flash chromatography afforded **202** as white solid (20.2 mg, 71%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.33-5.35 (1 H, d, *J* = 4.3 Hz, CH=), 5.01 (1 H, s, H1), 4.55-4.59 (1 H, dd, *J* = 12.3, 3.9 Hz, H6a), 4.20-4.23 (1 H, d, *J* = 11.9 Hz, H6b), 3.91 (1 H, s, H2), 3.81-3.86 (2 H, dd, *J* = 19.4, 7.9 Hz, H3,5), 3.53-3.58 (1 H, t, *J* = 9.6 Hz, H4), 3.48-3.50 (1 H, m, cholesteryl-H3), 2.36-2.40 (2 H, t, *J* = 7.6 Hz), 2.29-2.31 (2 H, d, *J* = 7.8 Hz), 1.80 - 2.02 (6 H, m), 1.43-2.15 (12 H, m, cholesteryl-CH), 1.26 (18 H, m, lipid), 1.10 (8 H, m), 1.00 (3 H, s), 0.90-0.92 (3 H, d, *J* = 6.5 Hz), 0.87-0.89 (3 H, t, *J* = 6.7 Hz, ω-$CH_3$), 0.85-0.87 (3 H, d, 6.6 Hz), 0.86-0.87 (3 H, d, *J* = 6.6 Hz), 0.67 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 175.0, 140.4, 121.8, 97.6, 77.0, 71.2, 70.8, 70.3, 67.8, 63.2, 56.6, 56.0, 50.0, 42.2, 39.8, 39.6, 39.4, 36.8, 36.5, 36.0, 35.6, 34.10, 31.8, 31.8, 29.5, 29.4, 29.2, 29.1, 29.1, 28.1, 27.9, 27.5, 24.8, 24.2, 23.7, 22.7, 22.5, 22.4, 20.9, 19.2, 18.6, 14.0, 11.7.HRMS (ESI) *m/z* calcd. for $[C_{45}H_{78}O_7+Na]^+$: 753.5640, obsd: 753.5645.



**Cholesteryl-6-*O*-tetradecanoyl-α-D-mannopyranoside (203)**

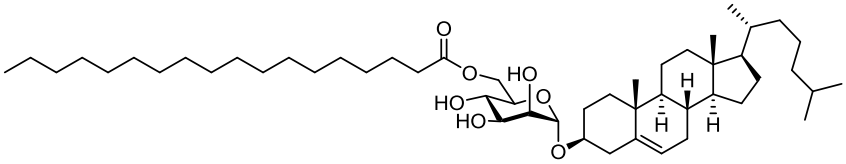
The General Protocol for Acylation between **231** (30 mg, 0.039 mmol) and myristoyl chloride (19.3 mg, 0.078 mmol, 21.2 µL) after flash chromatography afforded **203** as white solid (22.4 mg, 76%). $^{1}$H NMR (400 Hz, $CDCl_3$ δ 5.33-5.34 (1 H, d, *J* = 4.3 Hz, CH=), 4.99 (1 H, s, H1), 4.50-4.53 (1 H, dd, *J* = 11.8, 4.5 Hz, H6a), 4.25-4.27 (1 H, d, *J* = 11.7 Hz, H6b), 3.91 (1 H, s, H2), 3.82-3.85 (1 H, t, *J* = 8.4 Hz, H3), 3.47-3.49 (1 H, m, cholesteryl-H3), 3.33 (1 H, s, OH), 3.05 (1 H, s, OH), 2.70 (1 H, s, OH), 2.35-2.37 (2 H, t, *J* = 7.6 Hz), 2.30-2.31 (2 H, d, *J* = 7.0 Hz), 1.95-2.02 (2 H, m), 1.84-1.86 (2 H, m), 1.30-1.70 (18 H, m), 1.25 (22 H, m, lipid), 1.09-1.15 (6 H, m), 1.00 (3 H, s), 0.91-0.02 (3 H, d, *J* = 6.4 Hz), 0.87-0.89 (3 H, t, *J* = 6.4 Hz, ω-$CH_3$), 0.85-0.87 (6 H, *J* = 5.9 Hz), 0.67 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 174.8, 140.4, 121.8, 97.6, 77.0, 71.2, 70.9, 70.3, 67.8, 63.3, 56.6, 56.0, 50.0, 42.2, 39.8, 39.6, 39.4, 36.9, 36.5, 36.9, 36.5, 36.0, 35.7, 34.1, 31.8, 31.7, 29.6, 29.5, 29.4, 29.23, 29.19, 29.12, 28.1, 27.9, 27.5, 24.8, 24.1, 23.7, 22.7, 22.5, 22.4, 20.9, 19.2, 18.6, 14.0 11.7. HRMS (ESI) *m/z* calcd. for $[C_{47}H_{82}O_7+Na]^+$: 781.5953, obsd: 781.5950.



**Cholesteryl-6-*O*-hexadeconoyl-α-D-mannopyranoside (204)**

The General Protocol for Acylation between **231** (30 mg, 0.039 mmol) and palmitoyl chloride (21.4 mg, 0.078 mmol, 23.7 µL) after flash chromatography afforded **204** as white solid (23.9 mg, 78%)**.** $^{1}$H NMR (400 Hz, $CDCl_3$) δ 5.33-5.34 (1 H, d, *J* = 3.2 Hz, CH=), 5.00 (1 H, s, H1), 4.54-4.55 (1 H, dd, *J* = 11.8, 4.3 Hz, H6a), 4.21-4.24 (1 H,

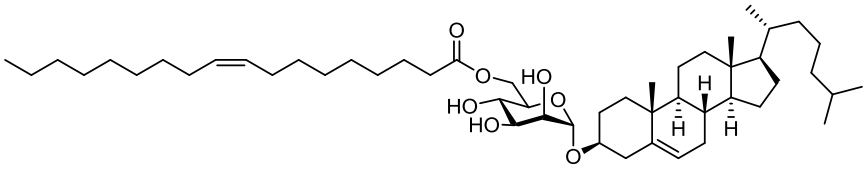
d, *J* = 11.8 Hz, H6b), 3.91 (1 H, s, H2), 3.81-4.84 (2 H, m, H3,5), 3.55 (1 H, t, *J* = 9.6 Hz, H4), 3.47-3.49 (1 H, m, cholesteryl-H3), 2.35-2.39 (2 H, t, *J* = 7.4 Hz), 2.29-2.31 (2 H, d, *J* = 7.8 Hz), 1.84-1.89 (6 H, m), 1.04-1.62 (12 H, m, cholesteryl-CH), 1.25 (26 H, s, lipid), 1.11 (8 H, m), 0.99 (3 H, s), 0.90-0.92 (3 H, d, *J* = 5.6 Hz), 0.85-0.87 (6 H, d, *J* = 6.6 Hz), 0.87-0.89 (3 H, t, *J* = 6.7 Hz, ω-$CH_3$), 0.67 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 175.2, 140.7, 122.1, 97.9, 77.2, 71.5, 71.2, 70.6, 68.1, 63.6, 56.9, 56.3, 50.3, 42.5, 40.1, 39.9, 39.7, 37.2, 36.9, 36.4, 36.0, 34.4, 32.1, 32.0, 29.9, 29.8, 29.7, 29.6, 29.4, 29.4, 28.4, 28.2, 27.8, 25.2, 24.4, 24.0, 23.0, 22.9, 22.7, 21.2, 19.5, 18.9, 14.3, 12.0. HRMS (ESI) *m/z* calcd. for $[C_{49}H_{86}O_7+Na]^+$: 809.6266, obsd: 809.6258.




**Cholesteryl-6-*O*-octadecanoyl-α-D-mannopyranoside (205)**

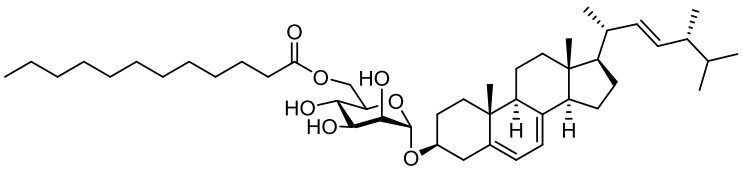
The General Protocol for Acylation between **231** (30 mg, 0.039 mmol) and stearoyl chloride (23.6 mg, 0.078 mmol, 26.3 μL) after flash chromatography afforded **204** as white solid (24.8 mg, 78%)**.** $^1$H NMR (400 Hz, $CDCl_3$) δ 5.33-5.34 (1 H, d, *J* = 4.1 Hz, CH=), 5.01 (1 H, s, H1), 4.54-4.58 (1 H, dd, *J* = 12.2, 4,4 Hz, H6a), 4.22-4.24 (1 H, d, *J* = 11.4 Hz, H6b), 3.91 (1 H, s, H2), 3.81-3.88 (2 H, m, H3,5), 3.53-3.58 (1 H, t, *J* = 9.6 Hz, H4), 3.47-3.51 (1 H, m, cholesteryl-H3), 2.35-2.39 (2 H, t, *J* = 7.6 Hz), 2.29-2.31 (2 H, d, *J* = 7.7 Hz), 1.95-2.02 (2 H, m), 1.80-1.89 (3 H, m), 1.05-1.64 (13 H, m, cholesteryl-CH), 1.25 (30 H, lipid), 1.10 (8 H, m), 1.00 (3 H, s), 0.90-0.92 (3 H, d, *J* = 6.4 Hz), 0.87-0.89 (3 H, t, *J* = 6.8 Hz, ω-$CH_3$), 0.87 (3 H, d, *J* = 1.2 Hz), 0.85 (3 H, d, *J* = 1.2 Hz), 0.67 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 175.1, 140.5, 122.0, 97.7, 71.3, 71.0, 70.5, 67.9, 63.3, 56.73, 56.2, 50.1, 42.3, 39.6, 39.7, 39.5, 37.0, 36.7, 36.2, 35.8, 34.3, 31.9, 31.8, 29.7, 29.6, 29.5, 29.4, 29.3, 28.7, 28.2, 28.0, 27.6, 25.0, 24.3,

23.8, 22.8, 22.7, 22.5, 21.0, 19.3, 18.7, 14.1, 11.8. HRMS (ESI) *m/z* calcd. for $[C_{51}H_{90}O_7+Na]^+$: 837.6579, obsd: 837.6580.



**Cholesteryl-6-*O*-oleyl-α-D-mannopyranoside (206)**

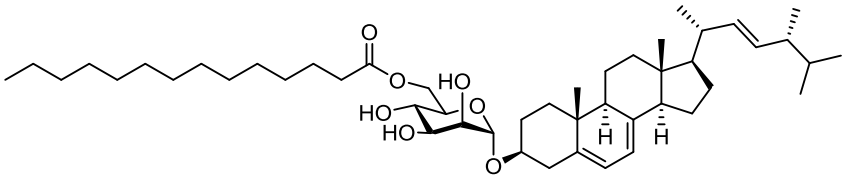
The General Protocol for Acylation between **231** (30 mg, 0.039 mmol) and oleoyl chloride (23.5 mg, 0.078 mmol, 25.8 µL) after flash chromatography afforded **204** as white solid (27.2 mg, 86%)**.** $^1$H NMR (400 Hz, $CDCl_3$) δ 5.32-5.35 (3 H, m), 4.99 (1 H, s, H1), 4.51-4.55 (1 H, dd, *J* = 11.8, 4.1 Hz, H6a), 4.23-4.25 (1 H, d, *J* = 11.6 Hz, H6b), 3.91 (1 H, s, H2), 3.81-3.87 (2 H, m, H3,5), 3.56-3.59 (1 H, m), 3.46-3.49 (1 H, m, cholesteryl-H3), 2.34-2.38 (2 H, t, *J* = 7.6 Hz), 2.29-2.31 (2 H, d, *J* = 7.5 Hz), 2.01-2.00 (6 H, m), 1.29 (14 H, m, lipid), 1.26 (12 H, m, lipid), 1.11-1.86 (12 H, m, cholesteryl-CH), 0.99 (3 H, s), 1.11 (6 H, m), 0.90-0.92 (3 H, d, *J* = 6.5 Hz), 0.87-0.89 (3 H, t, *J* = 6.9 Hz, ω-$CH_3$), 0.85-0.89 (6 H, d, *J* = 6.6 Hz), 0.67 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 174.9, 140.4, 129.9, 129.5, 121,8, 97.6, 77.0, 71.2, 70.8, 70.3, 67.8, 63.2, 56.6, 56.0, 50.0, 42.2, 39.8, 39.6, 39.4, 36.8, 36.5, 36.0, 35.6, 34.1, 31.8, 31.7, 29.6, 29.6, 29.4, 29.2, 29.0, 28.1, 27.9, 27.5, 27.1, 27.0, 24.8, 24.1, 23.7, 22.4, 20.9, 19.2, 18.6, 14.0, 11.7. HRMS (ESI) *m/z* calcd. for $[C_{51}H_{88}O_7+Na]^+$: 812.6530, obsd: 812.6538.



**Ergosteryl-6-*O*-dodecanoyl-α-D-mannopyranoside (211)**

The General Protocol for Acylation between **234** (23.5 mg, 0.030 mmol) and lauroyl chloride (13.1 mg, 0.060 mmol, 13.8 µL) after flash chromatography afforded **211** as

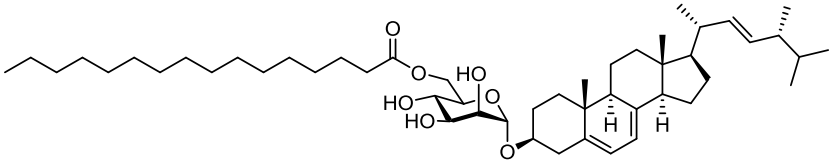
white solid (16.2 mg, 73%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 5.54-5.55 (1 H, d, *J* = 4.0 Hz, ring-B CH=), 5.36-5.37 (1 H, m, ring-B CH=), 5.13-5.25 (2 H, qd, *J* = 15.3, 7.3 Hz, C22/C23 CH=CH), 5.03 (1 H, s, H1), 4.52-4.56 (1 H, dd, *J* = 12.2, 4.6 Hz, H6a), 4.22-4.25 (1 H, d, *J* = 11.4 Hz, H6b), 3.92 (1 H, s, H2), 3.81-3.84 (2 H, m, H3,5), 3.57-3.59 (2 H, m, H4, ergosteryl-H3), 2.45-2.49 (1 H, dd, *J* = 14.5, 2.8 Hz), 2.34-2.38 (3 H, t, *J* = 7.7 Hz), 1.25 (18 H, m, lipid), 1.25-2.04 (18 H, m, ergsoteryl-CH), 1.02-1.04 (3 H, d, *J* = 6.6 Hz), 0.93 (3 H, s), 0.90-0.93 (3 H, d, *J* = 6.9 Hz), 0.86-0.89 (3 H, t, *J* = 6.8 Hz, ω-$CH_3$), 0.81-0.84 (6 H, t, *J* = 6.4 Hz), 0.62 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 175.2, 141.5, 139.6, 135.7, 132.1, 120.0, 116.4, 97.8, 75.7, 71.6, 71.2, 70.7, 68.0, 63.7, 55.9, 54.7, 46.3, 43.0, 42.9, 40.6, 39.2, 38.6, 38.2, 37.4, 34.5, 33.2, 32.1, 29.9, 29.8, 29.6, 29.5, 29.5, 28.5, 28.1, 25.2, 23.2, 22.9, 21.3, 21.2, 20.1, 19.8, 17.8, 16.3, 14.3, 12.2. HRMS (ESI) *m/z* calcd. for $[C_{46}H_{76}O_7+Na]^+$: 763.5483, obsd: 763.5480.




**Ergosteryl-6-*O*-tetradecanoyl-α-D-mannopyranoside (212)**

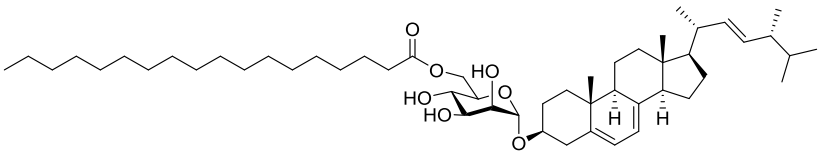
The General Protocol for Acylation between **234** (23.5 mg, 0.030 mmol) and myristoyl chloride (14.8 mg, 0.060 mmol, 16.3 μL) after flash chromatography afforded **212** as white solid (16.8 mg, 73%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 5.55-5.56 (1 H, d, *J* = 3.8 Hz, ring B CH=), 5.37-5.38 (1 H, m, ring B CH=), 5.15-5.25 (2 H, m, C22/23 CH=CH), 5.04 (1 H, s, H1), 4.54-4.56 (1 H, d, *J* = 11.6 Hz, H6a), 4.23-4.25 (1 H, d, *J* = 11.9 Hz, H6b), 3.92 (1 H, s, H2), 3.81-3.87 (2 H, m, H3,5), 3.54-3.63 (2 H, m, H4, cholesteryl-H3), 3.56 (2 H, m), 2.46-2.50 (1 H, dd, *J* = 14.8, 2.4 Hz), 2.36-2.39 (3 H, m), 1.34-2.05 (16 H, m, ergosteryl-CH), 1.25 (22 H, m, lipid), 1.03-1.04 (3 H, d, *J* = 6.6 Hz), 0.94 (3 H, s), 0.91-0.92 (3 H, d, *J* = 6.8 Hz), 0.87-0.88 (3 H, t, *J* = 6.9 Hz, ω-$CH_3$), 0.82-0.85 (6

H, t, *J* = 7.2 Hz), 0.63 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 175.2, 141.5, 139.6, 135.7, 132.1, 120.0, 116.4, 97.8, 75.7, 71.5, 71.1, 70.7, 68.1, 63.5, 58.3, 55.9, 54.7, 46.3, 43.0, 40.6, 39.2, 38.6, 38.2, 37.4, 34.4, 33.2, 32.1, 29.9, 29.8, 29.7, 29.5, 29.4, 29.4, 28.4, 28.1, 25.2, 23.2, 22.9, 21.3, 21.2, 20.2, 20.1, 19.8, 17.8, 16.4, 14.3, 12.2, 8.3. HRMS (ESI) *m/z* calcd. for $[C_{48}H_{80}O_7+Na]^+$: 791.5796, obsd: 791.5794.



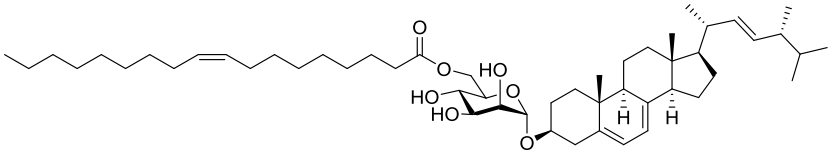

**Ergosteryl-6-*O*-hexadeconoyl-α-D-mannopyranoside (213)**

The General Protocol for Acylation between **234** (23.5 mg, 0.030 mmol) and palmitoyl chloride (16.5 mg, 0.060 mmol, 18.2 μL) after flash chromatography afforded **213** as white solid (21.5 mg, 90%)**.** $^1$H NMR (400 Hz, $CDCl_3$) δ 5.56 (1 H, d, *J* = 3.8 Hz, ring B CH=), 5.38-5.39 (1 H, m, ring B CH=), 5.21-5.26 (2 H, qd, *J* = 15.3, 7.5 Hz, C22/23 CH=CH), 5.05 (1 H, s, H1), 4.52-4.55 (1 H, dd, *J* = 12.0, 4.6 Hz, H6a), 4.26-4.28 (1 H, d, *J* = 11.2 Hz, H6b), 3.93 (1 H, s, H2), 3.83-3.85 (2 H, m, H3,5), 3.56-3.62 (4 H, m), 2.48-2.51 (1 H, dd, *J* = 14.9, 2.3 Hz), 2.36-2.38 (4 H, m), 1.42-2.06 (15 H, m, ergosteryl- CH), 1.26 (26 H, m, lipid), 1.04-1.05 (3 H, d, *J* = 6.6 Hz), 0.95 (3 H, s), 0.92-0.93 (3 H, d, *J* = 6.8 Hz), 0.88-0.90 (3 H, t, *J* = 6.9 Hz, ω-$CH_3$), 0.83-0.86 (6 H, t, *J* = 7.2 Hz), 0.64 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 175.1, 141.5, 139.6, 135.7, 132.7, 119.9, 116.4, 97.8, 75.7, 71.6, 71.1, 70.7, 68.1, 63.6, 58.6, 55.9, 54.7, 46.3, 43.0, 40.6, 39.2, 38.6, 38.2, 37.4, 34.4, 33.2, 32.1, 30.0, 29.8, 29.7, 29.5, 29.4, 29.3, 28.4, 28.1, 25.2, 23.2, 23.0, 21.2, 21.2, 20.1, 19.8, 17.8, 16.4, 14.3, 12.2, 8.4. HRMS (ESI) *m/z* calcd. for $[C_{50}H_{84}O_7+Na]^+$: 819.6109, obsd: 819.6113.



**Ergosteryl-6-*O*-octadecanoyl-α-D-mannopyranoside (214)**

The General Protocol for Acylation between **234** (23.5 mg, 0.030 mmol) and stearoyl chloride (18.2 mg, 0.060 mmol, 20.3 μL) after flash chromatography afforded **214** as white solid (18.5 mg, 75%)**.** $^{1}$H NMR (400 Hz, $CDCl_3$) δ 5.54-5.55 (1 H, d, *J* = 3.8 Hz, ring B CH=), 5.36-5.37 (1 H, m, ring B CH=), 5.14-5.25 (2 H, qd, *J* = 15.3, 7.5 Hz, C22/23 CH=CH), 5.02 (1 H, s, H1), 4.47-4.50 (1 H, dd, *J* = 11.9, 5.1 Hz, H6a), 4.27-4.30 (1 H, d, *J* = 11.4 Hz, H6b), 3.91 (1 H, m, H2), 3.84 (2 H, m, H3,5), 3.57-3.60 (2 H, m), 2.46-2.50 (1 H, dd, *J* = 14.7, 2.3 Hz), 2.34-2.37 (3 H, t, *J* = 7.7 Hz), 1.25-2.07 (18 H, m, ergosteryl-CH), 1.25 (30 H, m, lipid) 1.03-1.04 (3 H, d, *J* = 6.6 Hz), 0.93 (3 H, s), 0.91-0.92 (3 H, d, *J* = 7.0 Hz), 0.86-0.89 (3 H, t, *J* = 6.9 Hz, ω-$CH_3$), 0.82-0.85 (6 H, t, *J* = 7.1 Hz), 0.62 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 175.0, 141.4, 139.5, 135.7, 132.1, 120.0, 116.4, 97.9, 75.8, 71.6, 71.2, 70.7, 68.1, 63.7, 55.9, 54.7, 46.3, 43.0, 42.9, 40.6, 39.2, 38.6, 38.2, 37.4, 34.5, 33.2, 21.1, 29.88, 29.82, 29.79, 29.51, 29.49, 29.40, 28.4, 28.1, 23.1, 22.9, 21.3, 21.2, 20.1, 19.8, 17.8, 16.3, 14.3, 12.2. HRMS (ESI) *m/z* calcd. for $[C_{52}H_{88}O_7+Na]^+$: 847.6422, obsd: 847.6424.



**Ergosteryl-6-*O*-oleyl-α-D-mannopyranoside (215)**

The General Protocol for Acylation between **234** (23.5 mg, 0.030 mmol) and oleoyl chloride (18.1 mg, 0.060 mmol, 19.8 μL) after flash chromatography afforded **215** as white solid (20.5 mg, 82%)**.** $^{1}$H NMR (400 Hz, $CDCl_3$) δ 5.54-5.56 (1 H, dd, *J* = 5.2, 1.7 Hz, ring B CH=), 5.37-5.38 (1 H, m, ring B CH=), 5.33-5.35 (dt, 2 H, *J* = 5.6, 4.3

Hz), 5.15-5.25 (dq, 2 H, *J* = 15.3, 7.5 Hz, C22/23 CH=CH), 5.03 (1 H, s, H1), 4.53-4.56 (1 H, dd, *J* = 12.2, 4.5 Hz, H6a), 4.23-4.25 (1 H, d, *J* = 10.9 Hz, H6b), 3.92 (1 H, m, H2), 3.81-3.87 (2 H, m, H3,5), 3.55-3.62 (2 H, m), 2.46-2.50 (1 H, dd, *J* = 14.7, 2.4 Hz), 2.35-2.38 (3 H, m), 1.61-2.08 (18 H, m, ergosteryl-CH), 1.31 (14 H, m, lipid), 1.27 (12 H, m, lipid), 1.03-1.04 (3 H, d, *J* = 6.6 Hz), 0.93 (3 H, s), 0.91-0.92 (3 H, d, *J* = 6.8 Hz), 0.86-0.89 (3 H, t, *J* = 6.9 Hz), 0.82-0.85 (6 H, t, *J* = 7.2 Hz), 0.63 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 175.2, 141.5, 139.6, 135.7, 132.2, 130.2, 129.9, 120.0, 116.4, 97.8, 75.8, 71.6, 71.1, 70.7, 68.1, 63.6, 55.9, 54.7, 46.3, 43.0, 40.6, 39.2, 38.6, 38.2, 37.4, 34.4, 33.2, 32.1, 29.93, 29.88, 29.69, 29.49, 29.48, 29.35, 29.31, 28.4, 28.1, 27.39, 27.37, 25.1, 23.2, 22.8, 21.3, 21.2, 20.1, 19.8, 17.8, 16.4, 14.3, 12.2. HRMS (ESI) *m/z* calcd. for $[C_{52}H_{86}O_7+Na]^+$: 845.6266, obsd: 845.6270.

# Chapter 3

# Total synthesis of the *Bacteroides fragilis* α-galactosyceramide (α-GalCer$_{Bf}$) and CD1d-restricted activation of iNKT cells

## 3.1 Introduction

### 3.1.1 *Bacteroides fragilis,* a member of the human gut microbiota

"Humans contain more bacterial cells than human cells"! It is estimated that the cells in the human body are outnumbered 9:1 by bacterial cells that live within and upon us.[211] Consequently, resident bacteria play a major role in our body functions, including immunity, digestion and protection against disease.[212] The human colon contains the largest population of bacteria and the majority of these organisms are anaerobes, of which 25% are species of Bacteroidetes.[213] Bacteroidetes arose early in the evolutionary process and many of them have become indigenous to the host.[214] *Bacteroides* spp. are anaerobic, bile-resistant, non-spore forming and Gram-negative rods. *Bacteroides fragilis*, which comprises just 0.5% of the population in the human colonic flora, around 100-fold lower than other Bacteroidetes species, is the most frequent isolate from clinical specimens and is regarded as the most virulent Bacteroides species.

*B. fragilis* influences human health in many ways. *B. fragilis* is the only bacterium known to induce abscess formation as the sole infecting organism,[215] and causes skin and soft tissue infections. In addition, while rare, *B. fragilis* has been reported as a cause of meningitis and brain abscess.[216-218] *B. fragilis* is the most common anaerobe found in postsurgical wound infections in wounds relating to the gut flora.[219] At the sites of infection, it is speculated that *B. fragilis* may utilize host cell surface glycoproteins and glycolipids as a nutrient source.[220] However, the polysaccharide-utilizing ability of *B. fragilis* has not been extensively studied, in contrast to other bacteriodetes such as *B. thetaiotaomicron*.

### 3.1.2 The discovery of *B. fragilis* α-galactosylceramide

*B. fragilis* can influence the host immune system by activating the T-cell dependent immune responses.[221-226] However, for a long time the mechanism of this effect was unclear. To gain insight into the potential role of *B. fragilis* in mediating microbiota-host interactions, Brown and co-workers set out to define the complete set of sphingolipids produced by *B. fragilis*. Total lipids from wild-type were compared with those extracted from a sphingolipid-deficient mutant.[111] Upon purifying multi-milligram quantities of these compounds, high-resolution mass spectroscopy in the negative mode revealed the most abundant (pseudo) molecular ions of these molecules, including *m*/*z* of 677.5221, 554.5156 and 716.5698. The pseudo-molecular ion of *m*/*z* 677.5221 was consistent with the calculated *m*/*z* of 677.5234 $[M-H]^-$ for ceramide phosphorylethanolamine (CPE), and *m*/*z* 554.5156 was a match for the calculated *m*/*z* 554.5148 $[M-H]^-$ of a dihydroceramide base. This pseudo-molecular ion generated a daughter ion of *m*/*z* 328.2854, which was assigned as loss of a hydroxylated aliphatic chain, suggesting the sphingosine and fatty acid chain each have one methyl branch and one hydroxyl group. 1D and 2D NMR experiments on the purified compounds yielded resonances and couplings consistent with the assigned structures of CPE and ceramide$_{Bf}$, both known compounds.[111]

The pseudo-molecular ion of *m*/*z* 716.5698 was consistent with an empirical formula of $C_{40}H_{79}NO_9$, but did not correspond to a known compound. Detailed analysis of the 2D NMR and MS/MS spectra revealed that this compound and CPE harbor an identical dihydroceramide base, suggesting that the difference, $C_6H_{11}O_5$, corresponded to a different head group. Four lines of evidence suggest that this head group is an α-configured D-galactose. Firstly, MS/MS analysis revealed a fragment that was consistent with the elimination of a hexose group from a ceramide base. Secondly, by

using high-pH anion-exchange chromatography (HPAEC), since the hexose from the hydrolysate co-eluted with an authentic galactose standard at 12.6 mins.[112] Thirdly, the $^{1}$H NMR spectrum showed an anomeric proton with a chemical shift of δ 4.64 ppm, consistent with an α-linkage. Finally, chemically synthesized α-galactosylceramide, prepared by selective α-galactosylation of the isolated *B. fragilis* dihydroceramide, had a $^{1}$H NMR spectrum indistinguishable from that of the natural compound.

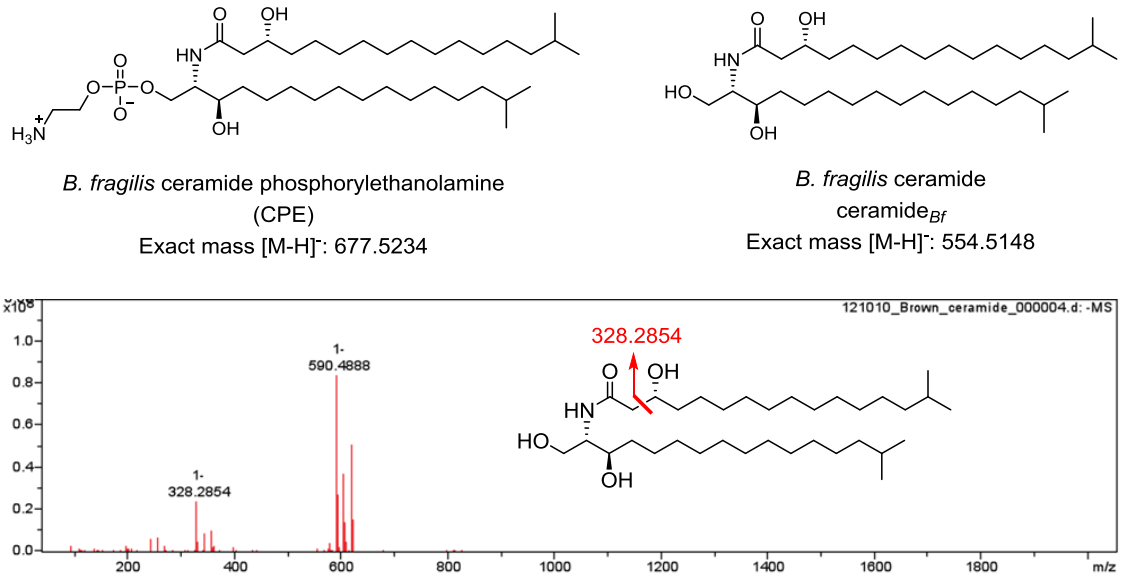
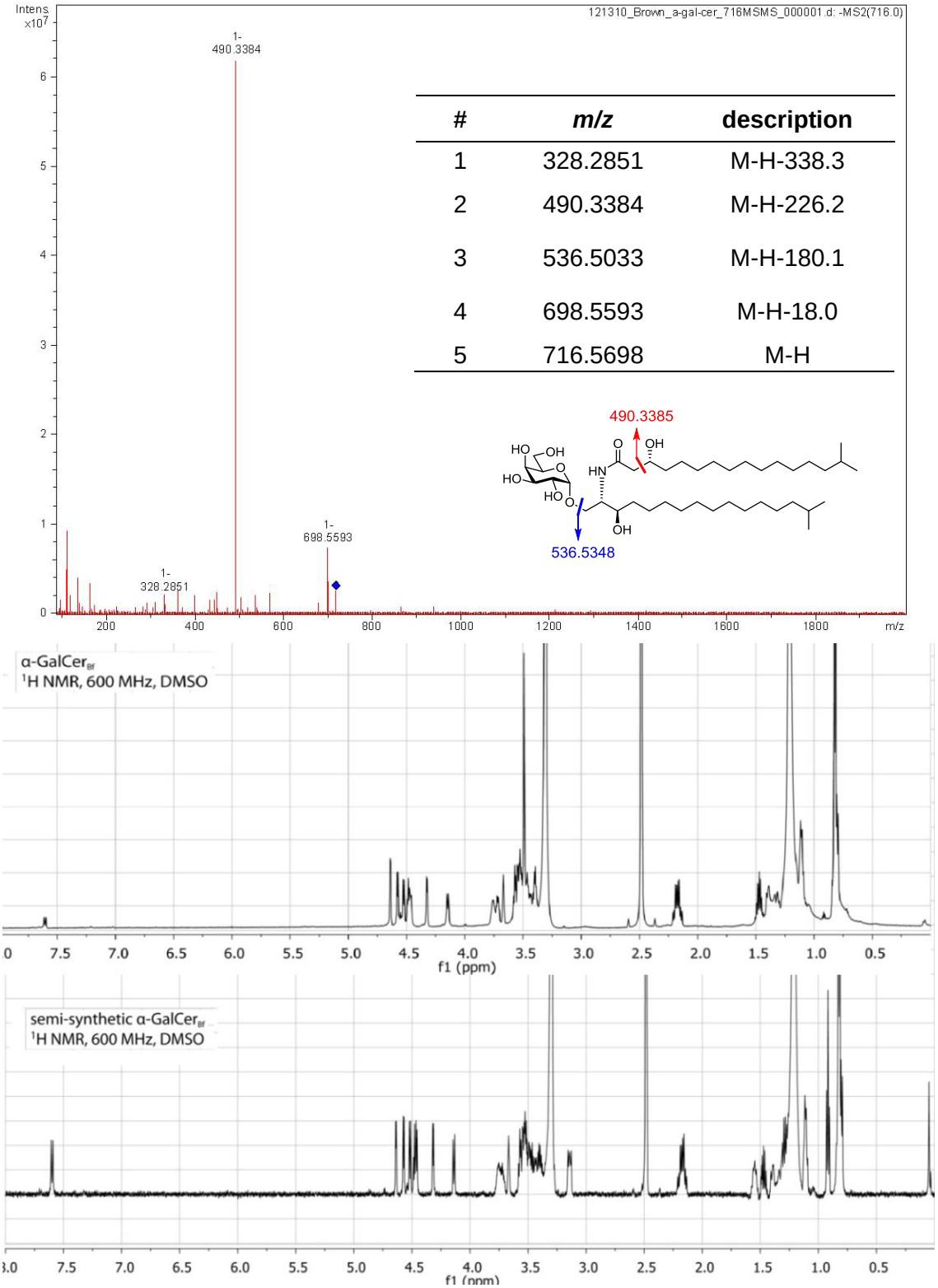


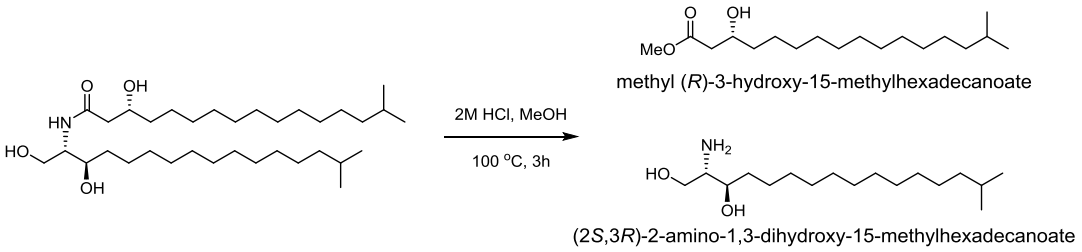

**Figure 3.1:** *B. fragilis* produces the phosphosphingolipid ceramide phosphoryl-ethanolamide (CPE, left) and the corresponding free ceramide (ceramide$_{Bf}$, right). High resolution mass spectra of crude ceramide extracted of *B. fragilis* and proposed fragment ion structures.




| # | *m/z* | description |
|---|---|---|
| 1 | 328.2851 | M-H-338.3 |
| 2 | 490.3384 | M-H-226.2 |
| 3 | 536.5033 | M-H-180.1 |
| 4 | 698.5593 | M-H-18.0 |
| 5 | 716.5698 | M-H |

**Figure 3.2:** (Top) MS fragmentation for pseudo-molecular ion of *m*/*z* 716.5698 $[M-H]^-$ demonstrating this compound harbours an identical dihydroceramide base with a different head group and its proposed fragment ion structures; (Bottom) high resolution $^1$H NMR spectrum of crude α-GalCer$_{Bf}$ extracted from *B. fragilis* overlaid with semi-synthetic α-GalCer$_{Bf}$ (600 MHz, DMSO).

One hydroxyl group needed to be located within each lipid chain and the absolute configuration established. Treatment of *B. fragilis* ceramide with 2 M HCl in MeOH at 100 ºC for 3 h afforded a hydroxylated fatty acid and the hydroxyl sphinganine side-chain. By comparison of the $^{1}$H NMR spectra and optical rotations of these two lipids with literature values, Brown and co-workers assigned the absolute configuration of these two hydroxyl groups as shown in Scheme 3.1.[227-228]




**Scheme 3.1:** Methanolysis of ceramide for determination of absolute configuration.

Collectively, these data suggest that the major *B. fragilis* glycolipid is composed of a ceramide base, with both the sphinganine and acyl chain each carrying one methyl branch and one hydroxyl group and bearing an α-D-galactosyl head group. Interestingly, beside the major *m*/*z* of 716.5683, other minor peaks were evident at *m*/*z* 730.5836 and 744.6011 that varied in mass by 14 Da. Further mass spectrometric studies demonstrated that the structural variability resided in the chain length of the sphinganine (Figure 3.3). Brown termed the mixture of glycolipids *B. fragilis* α-galactosylceramide (α-GalCer$_{Bf}$), and was used as a mixture in immunological experiments, as described subsequently.

A later paper by An *et al*. largely confirmed these structural assignments, but in addition to the three lipoforms described by Wieland Brown *et al*., three additional compounds were identified, namely the anteiso variants of the three acyl chain homologues.[112, 229]

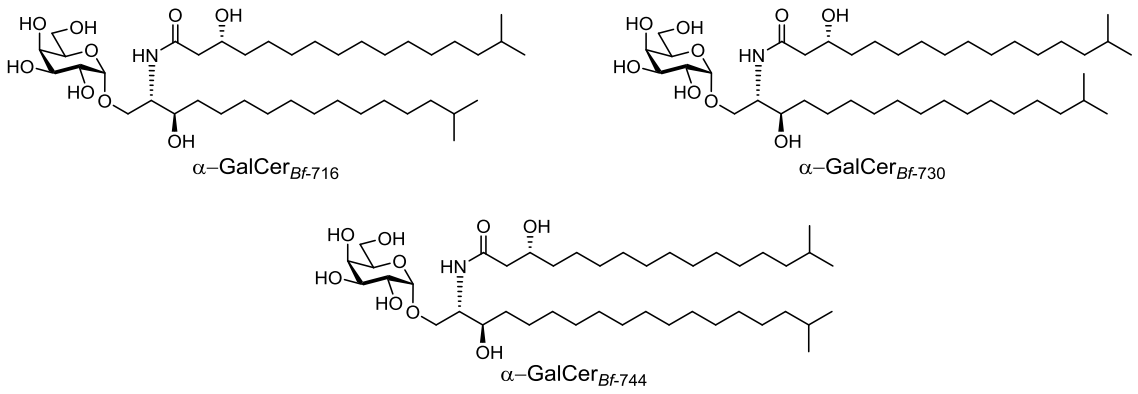



**Figure 3.3:** Structure of some components of the α-GalCer$_{Bf}$ mixture, named after their integer *m/z* value.

The α-GalCer$_{Bf}$ compounds are structurally related to KRN7000 and agelasphin-9b.[230] However, they differ through several important changes. The *N*-acyl chain is shorter within *B. fragilis* α-galactosyceramide, includes iso- or anteiso-branching, and bears a hydroxyl group on the β-carbon, rather than the α-carbon. Like agelasphin 9b, the *B. fragilis* α-galactosyceramide has an iso-branch in the sphinganine but lacks the C4-OH present in both agelasphin 9b and KRN7000.

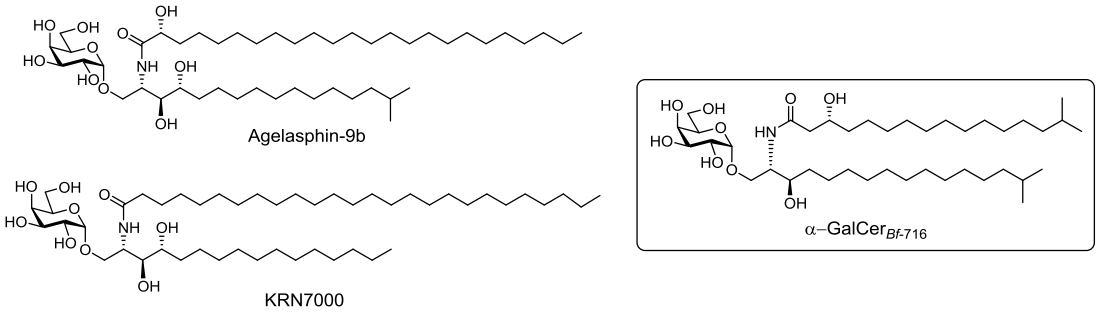



**Figure 3.4:** Chemical structures of agelasphin-9b, KRN7000 and *B. fragilis* α-galactosylceramide.

### 3.1.3 Immunological activity of *B. fragilis* α-galactosyceramide

KRN7000 is the prototypical agonist of iNKT cells and has become a critically important reagent for studying NKT cell biology in both *in vitro* and *in vivo*. The structural resemblance of α-GalCer$_{Bf}$ to KRN7000 and agelasphin-9b led to the obvious suggestion that α-GalCer$_{Bf}$ could be a CD1d-presented antigen for iNKT cells.

However, synthetic derivatives of KRN7000 that either have shorter chains or lack a C4 hydroxyl group have been shown to elicit reduced activation of iNKT cells and/or provide an altered cytokine response, an effect that might be due to a change in the conformation of the CD1d-lipid complex influencing the affinity of TCR binding.[231]

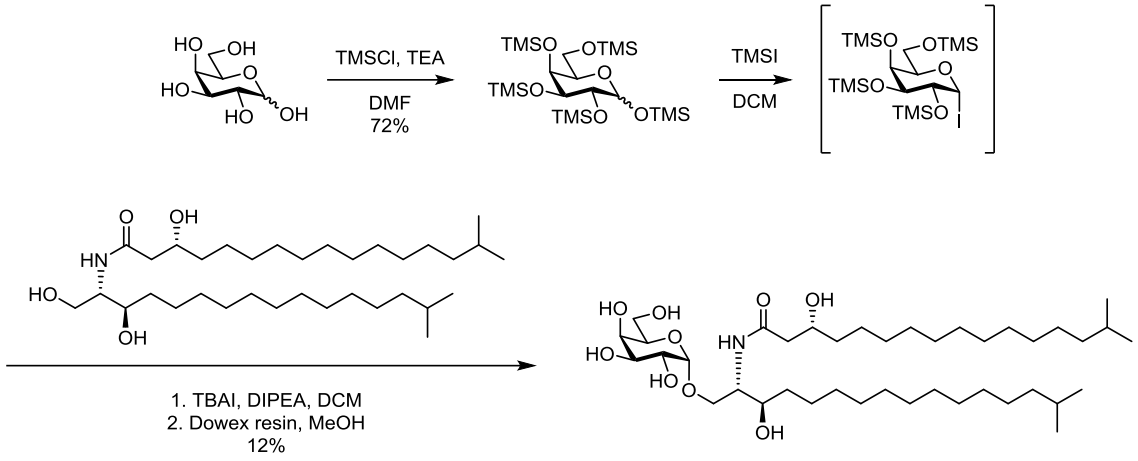
Two independent groups investigated the ability of α-GalCer$_{Bf}$ (as a mixture of lipoforms isolated from *B. fragilis*) to stimulate iNKT cells in a CD1d-dependent manner. Wieland Brown *et al.* showed that α-GalCer$_{Bf}$ can bind to CD1d and activate both mouse and human iNKT cells *in vitro* and *in vivo*.[111] However, subsequently, An and co-workers reported that α-GalCer$_{Bf}$ inhibits KRN7000 induced iNKT cell activation *in vivo*.[112] They showed that α-GalCer$_{Bf}$ can modulate host colonic iNKT cell homeostasis and protect the host from colitis challenge. It was suggested that α-GalCer$_{Bf}$ loads into CD1d but the resulting complex is unreactive to iNKT cells.

Variations in the precise composition of the α-GalCer$_{Bf}$ mixture may be a reason for the opposing effects of α-GalCer$_{Bf}$ on the immune system reported by these two studies. An *et al*. suggested that some of the lipid structures are stimulatory while others are inhibitory.[112] Thus the data may record the total net effect of these lipids or may be influenced by activities of individual components that differ between studies.

### 3.1.4 Semi-synthesis generation of α-GalCer$_{Bf}$

No total synthesis of α-GalCer$_{Bf}$ has been reported. Wieland Brown and colleagues reported a semi-synthetic route to prepare α-GalCer$_{Bf\text{-}716}$ by performing a halide-ion catalyzed glycosylation of TMS-protected galactosyl iodide and the ceramide isolated from *B. fragilis*.[111] The glycosyl iodide was prepared by reaction of D-galactose with TMSCl and $Et_3N$ affording penta-TMS galactose, followed by treatment with TMSI to introduce the iodine at the anomeric position. Promoted by

tetrabutylammonium iodide (TBAI) the glycosyl iodide and ceramide afforded the protected glycoside. Treatment with ion-exchange resin ($H^+$ form) gave α-GalCer$_{Bf}$.




**Scheme 3.2:** Semi-synthesis of α-GalCer$_{Bf}$.

### 3.1.5 Is *B. fragilis* α-galactosyceramide an agonist or antagonist of iNKT cells?

The structural similarities between KRN7000 and α-GalCer$_{Bf}$ produced by *B. fragilis* are unmistakable. Aside from α-GalCer$_{Bf}$ and the sponge-derived agelasphins, no other naturally occurring α-galactosylceramides have been discovered.[131] Substantial data have accumulated that α-GalCer$_{Bf}$ is a ligand for a subset of iNKT cells when presented by antigen-presenting molecule CD1d. Because of the conflicting results between two studies by Brown and co-workers[111] and An and co-workers,[112] we hypothesise that by providing access to a homogeneous sample of this glycolipid we may be able to clarify the contradictory reports on the immunological properties of α-GalCer$_{Bf}$. Additionally, by developing a means to synthesize α-GalCer$_{Bf}$, we also propose to synthesis analogues with different acyl chains in the ceramide to explore structure-activity relationships.

### 3.1.6 Research aims

*B. fragilis* influences human health in many ways, and the molecular mechanisms are under active investigation. *B. fragilis* generates a mixture of α-galactosylceramides

that modulate host signalling pathways, yet studies on the immunological properties of these molecules reveal conflicting results. Therefore, the aims of this chapter are to:

**Aim 1:** Develop an approach for the preparation of the acid side chain of the α-Galcer$_{Bf\text{-}716}$.

**Aim 2:** Develop an approach for the preparation of the sphinganine side chain of α-Galcer$_{Bf\text{-}716}$.

**Aim 3:** Prepare a galactosyl sphinganine by α-galactosylation of the sphinganine.

**Aim 4:** Couple various fatty acids to the galactosyl sphinganine to produce α-Galcer$_{Bf\text{-}716}$ **304** and analogues with different acid side chains. We propose to synthesise *Bf* $C_{16}$ **301**, *Bf* isoC$_{17}$ **302** *Bf* hydroxyl-$C_{16}$ **303** and *Bf* hydroxyl-isoC$_{17}$ **304** (Figure 3.5).

**Aim 5:** Investigate the ability of α-GalCer$_{Bf\text{-}716}$ and analogues to stimulate iNKT cell signalling in collaboration with Dr Garth Cameron and Prof. Dale Godfrey at Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity. This will help to elucidate the immunological interplay between human host and the *B. fragilis*.

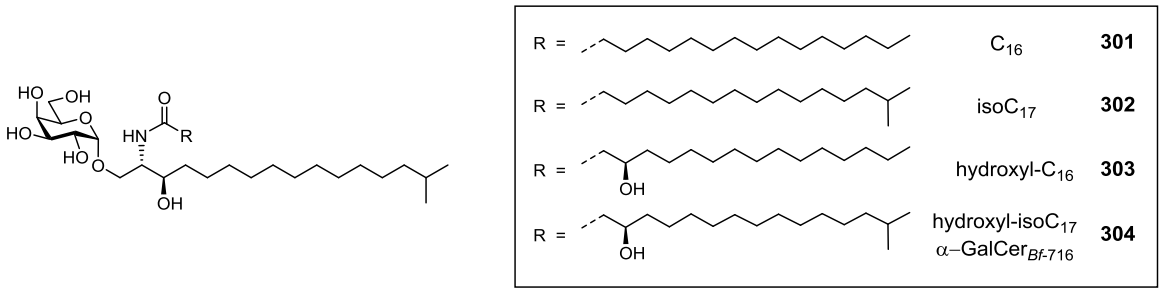


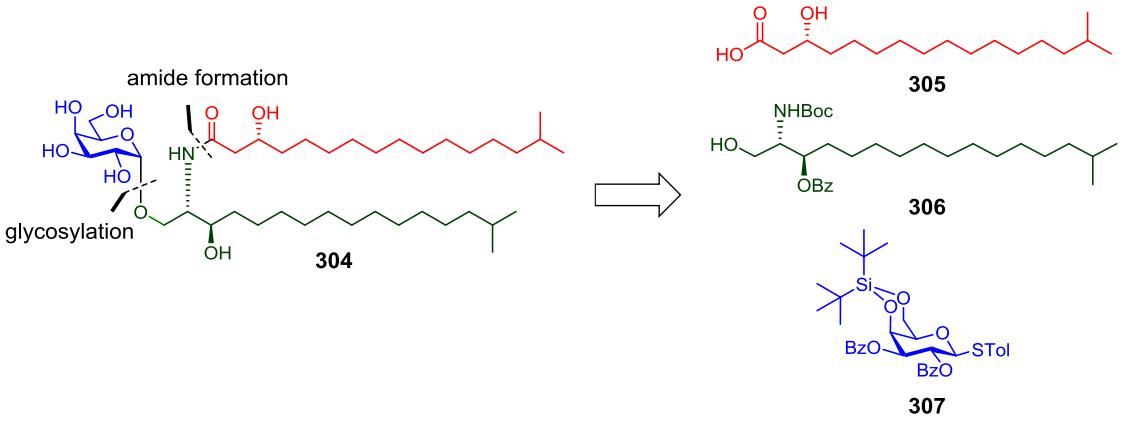

**Figure 3.5:** Proposed targets for total synthesis: α-GalCer$_{Bf\text{-}716}$ and acyl variants.

## 3.2 Results and discussion

### 3.2.1 Synthesis strategy

Our primary target is α-GalCer$_{Bf\text{-}716}$ **304**, but we also sought to prepare analogues **301**-**303** with variation in the fatty acid. Several elegant strategies have been developed for the synthesis of α-GalCer,[94, 232-233] with a key step being the glycosylation to form

the glycolipid skeleton. It can occur by the coupling of the functionalised sphingoid backbone and subsequent attachment of the acyl chain, or by direct coupling to the ceramide.[234-235] In our case, to provide an efficient route that allowed for the convergent synthesis of both α-GalCer$_{Bf\text{-}716}$ **304** and its analogues, we proposed that α-GalCer$_{Bf\text{-}716}$ **304** could be assembled from 3 fragments: (*R*)-3-hydroxy-15-methylhexadecanoic acid **305**, the sphinganine acceptor **306** and protected galactosyl donor **307**. Sphinganine acceptor **306** and galactosyl donor **307** could be coupled by a glycosylation reaction; subsequently the amide bond with the acid side chain **305** could be installed by an amide coupling to afford α-GalCer$_{Bf\text{-}716}$ **304** (Scheme 3.3).



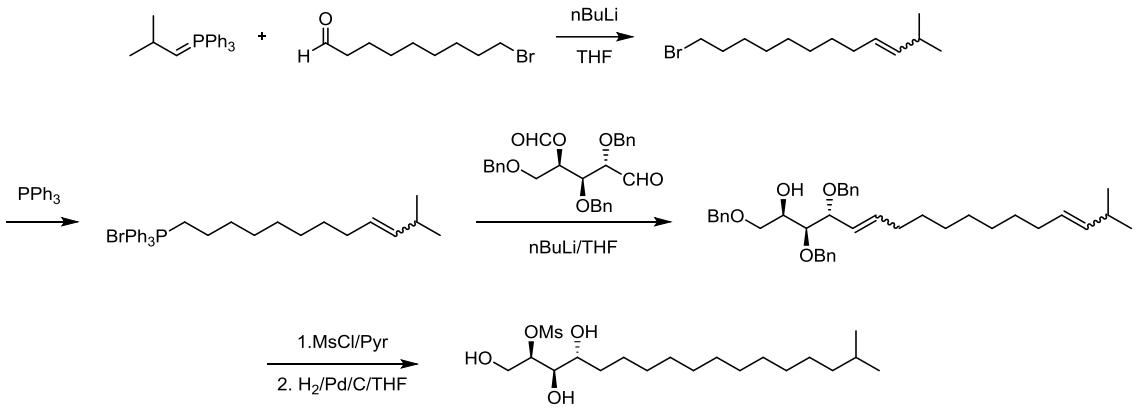

**Scheme 3.3**: Retrosynthetic disconnection of α-GalCer$_{Bf\text{-}716}$ **304**.

### 3.2.2 Preparation of acid side chains

**Preparation of (*R*)-3-hydroxy-15-methylhexadecanoic acid (305)**

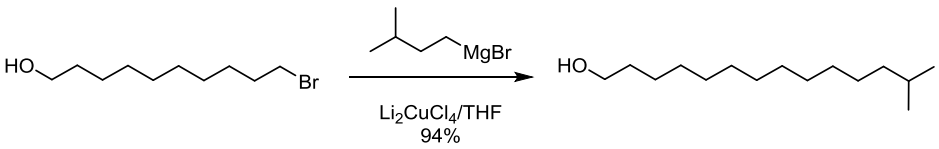
We first considered methods for the construction of the iso-branch. Common methods include the Wittig reaction and the copper(I)-catalyzed cross coupling of Grignard reagents. Morita and co-workers introduced an iso-branch into lipids by a Wittig reaction between isobutylidene triphenylphosphorane and 9-bromononanal (Scheme 3.4).[236] The Wittig reaction gave the adduct as a mixture of geometric isomers, which could not be separated. The mixture of isomers was carried forward until the very last step when reduction of the alkene resulted in convergence of the mixture of

diastereomers to a single product. However, this route has shortcomings in the complexity of analysis of mixtures of stereoisomers *en route*.




**Scheme 3.4:** Formation of the phytosphingosine of agelasphin-9b by Morita and co-workers.[236]

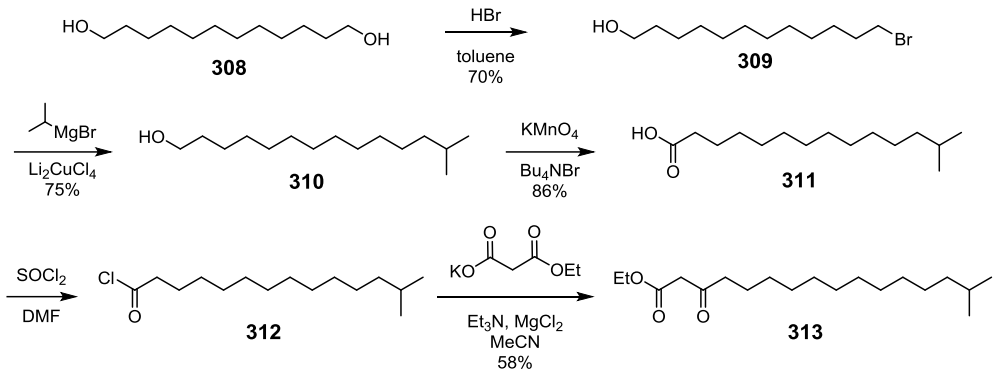
Mori and co-worker used copper(I)-catalyzed cross-coupling of 10-bromodecan-1-ol and isoamylmagnesium bromide to install an iso-branch into 13-methyltetradecan-1-ol **310**.[237] This cross-coupling provides saturated iso-branched alkanes directly in one step, avoiding the intermediate diasteromeric alkenes formed in a Wittig reaction, and the need to adjust oxidation state.




**Scheme 3.5:** The synthesis of 13-methyltetradecan-1-ol using a copper(I)-catalyzed cross coupling reaction.[237-238]

Our route to iso-branched hydroxy acid **305** was inspired by Mori's route. As iso-amylmagnesium bromide is expensive, we proposed to translocate the bond junction and instead couple 12-bromododecan-1-ol **309** and isopropylmagnesium bromide. Reaction of 1,12-dodecanediol **308** and aqueous HBr in toluene afforded mono-substituted bromide **309**,[239] which underwent copper(I)-catalysed cross coupling with isopropylmagnesium bromide to afford iso-$C_{15}$ alcohol **310** (Scheme 3.6).[237] Oxidation

of alcohol **310** with $KMnO_4$ and $Bu_4NBr$[240] in ether and water afforded 13-methyltetradecanoic acid **311**. Clay reported the synthesis the β-ketoesters by reaction of an acid chloride with ethyl potassium malonate in MeCN.[241] Applying Clay's protocol to carboxylic acid **311** afforded ethyl 15-methyl-3-oxohexadecanoate **313** in 58% yield over 2 steps.




**Scheme 3.6:** Preparation of ketoester **313**.

Noyori reported that halogen-containing BINAP-Ru(II) complexes are highly effective for enantioselective hydrogenation of ketones in homogeneous phase.[242,243] We selected a BINAP-Ru catalyst containing a *p*-cymene group as a candidate for reduction of β-ketoester **313**. *p*-Cymene Noyori catalysts are superior to Noyori's original benzene adducts as they are more stable and safer.[244] The catalytic cycle of the asymmetric reduction of β-ketoester with Ru-BINAP complexes in methanol is shown in Figure 3.6.[245]

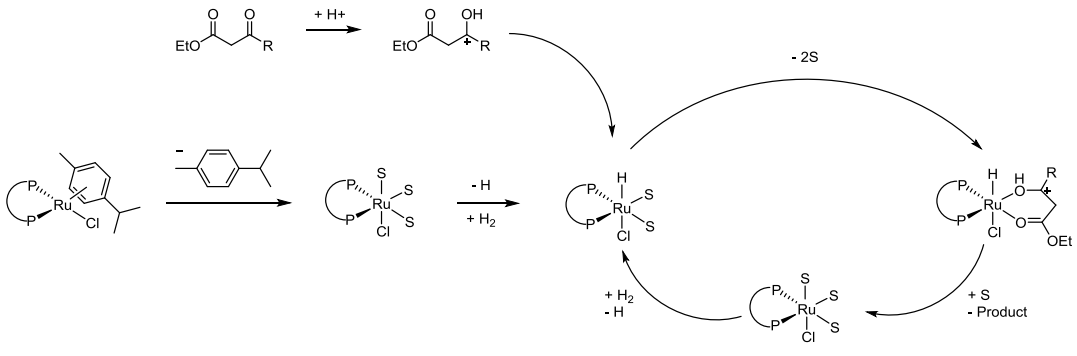
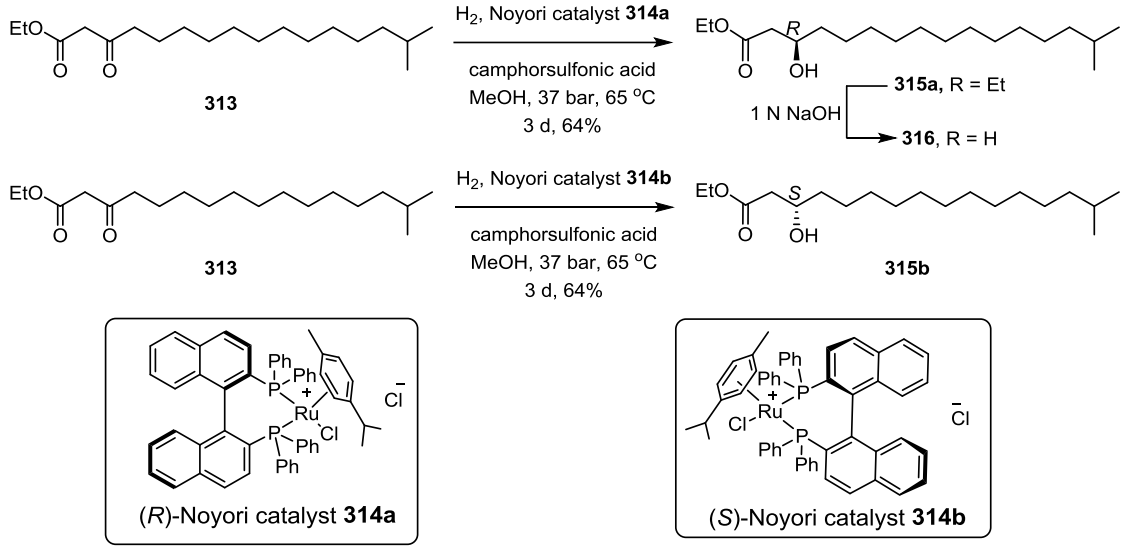



**Figure 3.6:** The catalytic cycle for Noyori enantioselective hydrogenation in the presence of acid. S = solvent

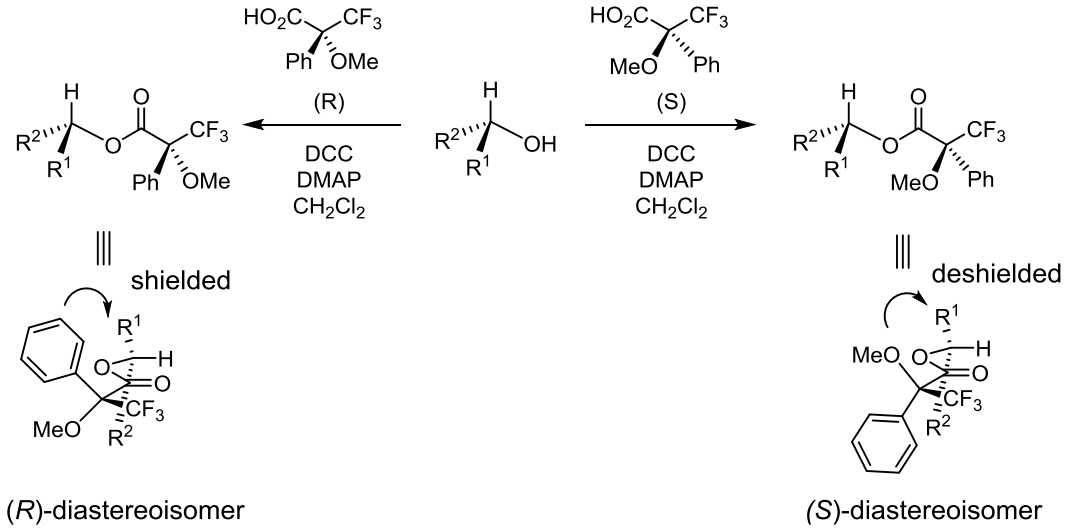
Reduction of β-keto esters typically require temperatures >80 ºC, hydrogen pressures >80 atm and specialized apparatus. However, in the presence of trace amounts of acid,[246] asymmetric hydrogenation can be achieved in a standard Parr apparatus at lower temperature and pressure. The increased reactivity in the presence of acid can be explained by the involvement of the acid in various stages of the catalytic cycle.[247] The first step in the catalytic cycle is ligand exchange, with loss of the *p*-cymene ligand from the catalyst precursor to give a coordinatively unsaturated species that is captured by solvent. One ligand position is then replaced by hydride, generated by heterolytic cleavage of hydrogen. The substrate then binds through its carbonyl groups at two coordination sites. This binding event ultimately determines the enantioselectivity of reduction, and the BINAP ligand dictates the orientation in which the substrate binds. In the presence of acid, the substrate is protonated prior to hydride insertion, which accelerates the hydride transfer. Finally, hydride transfer from the complex releases the chiral alcohol.




**Scheme 3.7:** Stereoselective reduction of β-ketoester **313** with enantiomeric Noyori catalysts.

Stereoselective reduction of β-ketoester **313** with either Noyori catalyst **314a** or **314b** afforded β-hydroxyesters tentatively assigned as ethyl (*R*)-3-hydroxy-15-methylhexadecanoate **315a** and (*S*)-3-hydroxy-15-methylhexadecanoate **315b**, respectively. Saponification of the (*R*)-hydroxy ester **315a** in using aq. NaOH/ MeOH gave the hydroxy acid **316** (Scheme 3.7).

The stereochemistry at the β position of **315a** and **315b**, predicted based on known catalyst stereoselectives, was confirmed using Mosher's ester analysis.[248] Mosher's ester analysis is a technique that uses $^{1}$H NMR spectroscopy to define the absolute stereochemistry of chiral alcohols.[249] Mosher's ester analysis can also be used to determine the ratio of two enantiomers: the enantiomeric excess.[250] The method involves reaction of a chiral alcohol with each enantiomer of a chiral Mosher's acid (Scheme 3.8). The resulting diastereomers have predictable differences in the chemical shifts of protons adjacent to the site of derivatization, which allow the stereochemistry of the chiral alcohol to be deduced.




**Scheme 3.8:** Model explaining Mosher's ester analysis for determination of stereochemistry of alcohols.

Mosher's ester analysis relies on the assumption that the major conformation of each diastereomeric ester formed from Mosher's acid is that where the system $CF_3$–C–C(=O)–O–C(H) lies in a plane. Considering the (*R*)-diastereomer, this assumes that $R^1$ and the phenyl group both lie above the plane, while $R^2$ and the $OCH_3$ group are both below the plane. Protons at the $R^1$ position will be shielded by the phenyl ring, causing an upfield chemical shift. For the (*S*)-diastereomer, $R^1$ and the phenyl group are on opposite sides, therefore the $R^1$ chemical shifts are deshielded by the $OCH_3$ group. By comparing the chemical shifts of the $R^1$ group for the two diastereomers, the stereochemistry can be assigned. Mosher's ester analysis can also be used to determine the enantiomeric excess of mixtures by comparison of the intensities of the signals of the diastereomers.

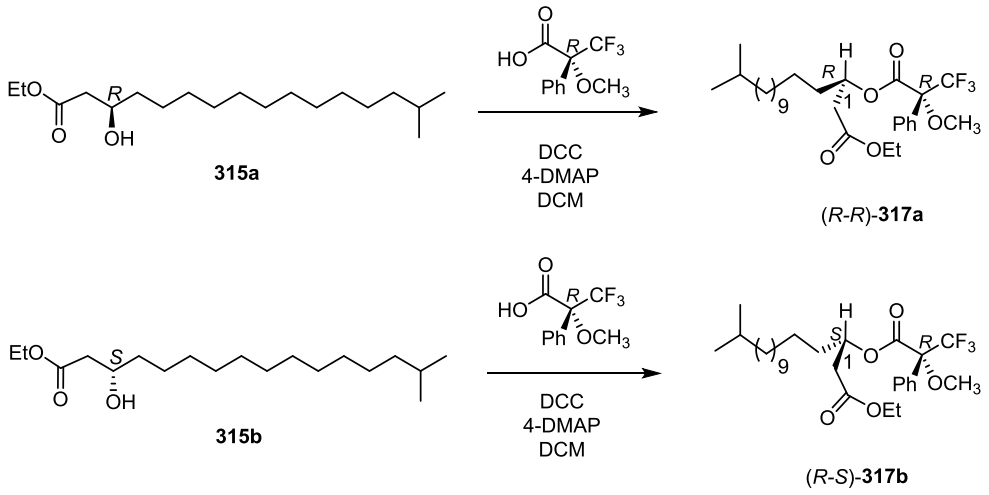



**Table 3.1:** Comparison of proton chemical shifts of two Mosher's ester diastereomers **317a** and **317b**.

| Compound | δ H1a (ppm) | δ H1b (ppm) |
|---|---|---|
| *R,R*-**317a** | 2.72 | 2.58 |
| *S,S*-**317b** | 2.67 | 2.54 |

We performed Mosher's ester analysis using two enantiomeric alcohols and a single Mosher's acid. Thus, the (*R*)-hydroxy ester **315a** and (*S*)-hydroxy ester **315b** were esterified under Steglich conditions[251] with the (*R*)-Mosher's acid. The resulting products are diastereomers *viz*., *R,R*-**317a** and *S,R*-**317b**. The α-protons of the ester appear at δ 2.72 and 2.58 ppm for *R,R*-**317a**; the equivalent protons appear at δ 2.67 and 2.54 ppm for *S,R*-**317b** (Table 3.1). The $R^1$ protons of *R,R*-**317a** are deshielded relative to the $R^1$ protons of *S,R*-**317b**, confirming the stereochemical assignment, and that the *R*-configured catalyst affords the *R*-hydroxyl ester and the *S*-configured catalyst afford the *S*-hydroxy ester, consistent with previously reported examples using the Noyori catalyst **314a** and **314b**.[244]

Mosher's ester analysis also allowed determination of the enantiomeric purity. Analysis of the NMR spectrum of **317a** reveals a major compound plus small amounts of a second compound with signals identical to the major compound of Mosher's ester of **315b**, namely **317b**. The reciprocal situation is true for the spectrum of **317b**. The optical purity of the products was determined to be 95% e.e by integration of the $^1$H NMR spectra of diastereoisomeric Mosher's esters **317a** and **317b**.

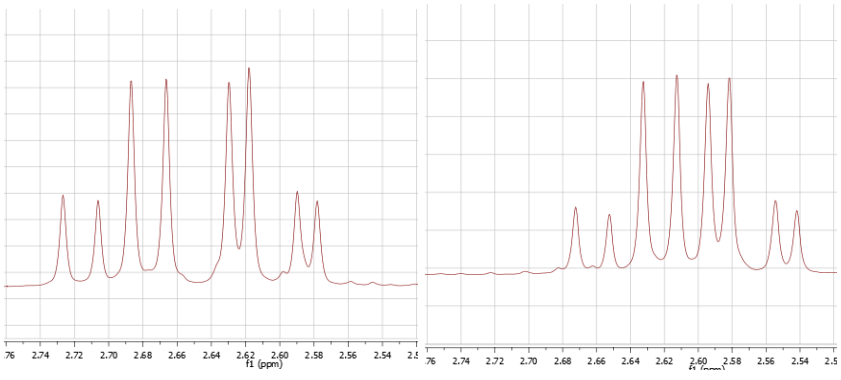
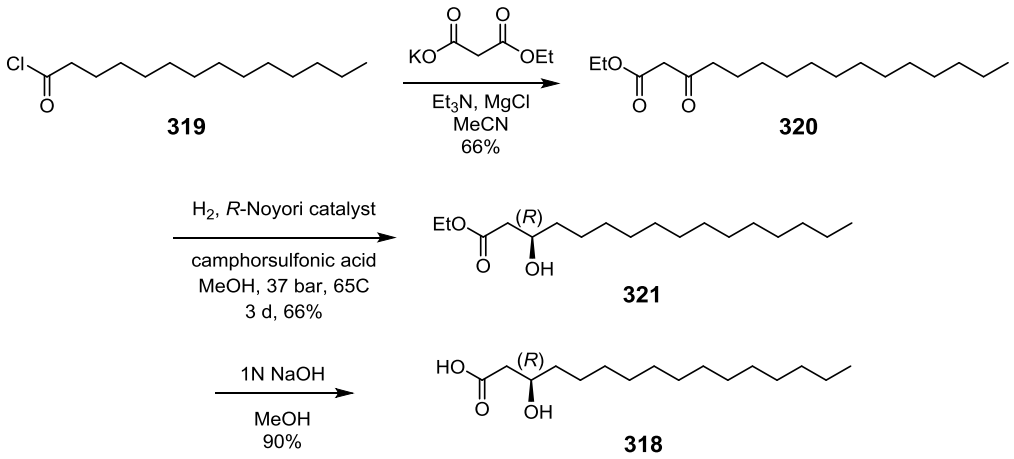



**Figure 3.7:** $^1$H NMR spectra of H1 proton from the Mosher's ester **317a (**left) and the Mosher's ester **317b** (right).

**Preparation of (*R*)-3-hydroxyhexadecanoic acid (318)**

To prepare the non-branched analogue **303**, the β-hydroxy acid **305** was prepared. Myristoyl chloride **319** was treated with ethyl potassium malonate and magnesium chloride in MeCN at 0 ºC to give ethyl 3-oxohexadecanoate **320** in 66% yield. Stereoselective reduction of β-ketoester **320** with (*R*)-Noyori catalyst **315a** gave ethyl (*R*)-3-hydroxyhexadecanoate **321**, assigned based on the previous results. Saponification gave the hydroxy acid **318** in 90% yield.

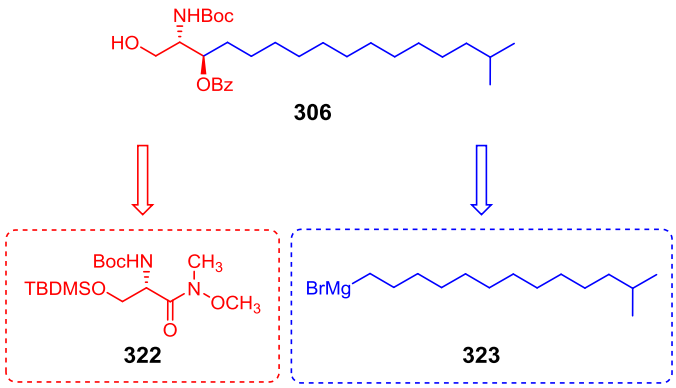



**Scheme 3.9:** Synthesis of (*R*)-3-hydroxy-hexadecanoic acid **318**.

### 3.2.3 Synthesis of sphinganine side chain

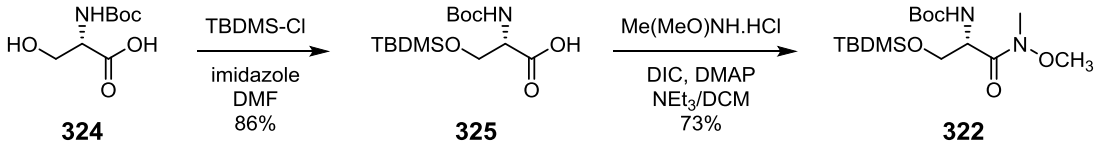
Several methods have been reported for the synthesis of sphinganines. Sphinganines have been synthesised from sugar precursors,[252-253] from the amino acid L-serine,[254-259] and from racemic or achiral precursors by a variety of asymmetric strategies.[260-263] We followed an approach developed by Howell and co-workers[227] which is a four-step procedure starting from a protected L-serine derivative.

We proposed that sphinganine **306** could be synthesized from two fragments: 12-methyltridecylmagnesium bromide **323** and Weinreb amide **322**. Cross coupling of the Weinreb amide **322** and Grignard reagent **323** could create the carbon skeleton, followed by reduction and then removal of TBDMS protecting group to give the primary-alcohol **306** (Scheme 3.10).



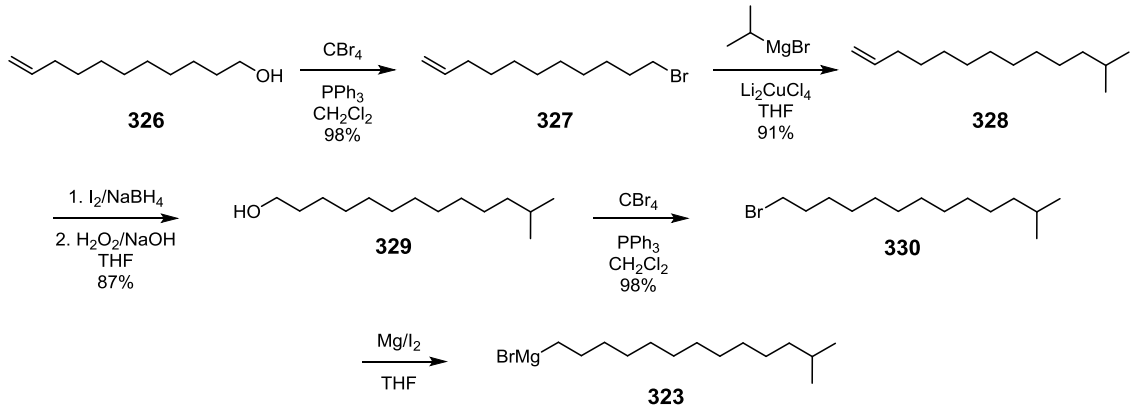

**Scheme 3.10:** Disconnection for the synthesis of sphinganine acceptor **306**.

We began with the formation of the Weinreb amide **322**. The Weinreb amide **322** was prepared by treatment of *N*-(*tert*-butoxycarbonyl)-L-serine **324** with TBDMS-Cl to protect the alcohol affording silyl ether **325**, which was reacted with Me(MeO)NH.HCl, promoted by DIC and DMAP in $CH_2Cl_2$, to afford the Weinreb amide **322** in 73% yield (Scheme 3.11).[264]



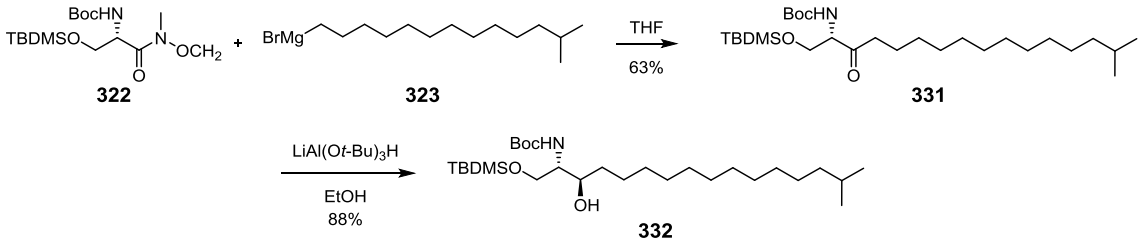

**Scheme 3.11:** Synthesis of Weinreb amide **324**.

Preparation of the Grignard reagent **323** started from 9-decen-1-ol **326**. Bromination using $CBr_4$ and $PPh_3$ in $CH_2Cl_2$ gave the bromide **327**, which underwent copper(I)-catalysed cross coupling with isopropylmagnesium bromide to afford iso-$C_{14}$ alkene **328**. Anti-Markovnikov hydration of the alkene **328** using $I_2$ and $NaBH_4$, followed by oxidative deborylation using hydrogen peroxide and sodium hydroxide, gave 12-methyl-1-tridecanol **329**.[265] Bromination of the primary alcohol using $CBr_4$ and $PPh_3$ afforded the iso-$C_{14}$ bromide **330**, which was heated under reflux with activated magnesium to generate 12-methyltridecylmagnesium bromide **323** (Scheme 3.12).




**Scheme 3.12:** Synthesis of the Grignard reagent **323**.

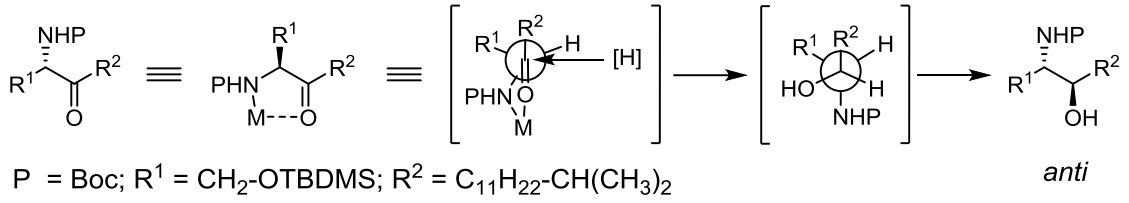
Treatment of the Weinreb amide **322** with the Grignard reagent **323** in THF at 0 ºC gave the ketone **331**. Reduction of the ketone functional group with $LiAl(O^tBu)_3H$,[266] afforded the alcohol **332** as a single stereoisomer (Scheme 3.13).




**Scheme 3.13:** Cross coupling reaction of the Weinreb amide and the Gringnard reagent, followed by the reduction to afford a single isomer alcohol.

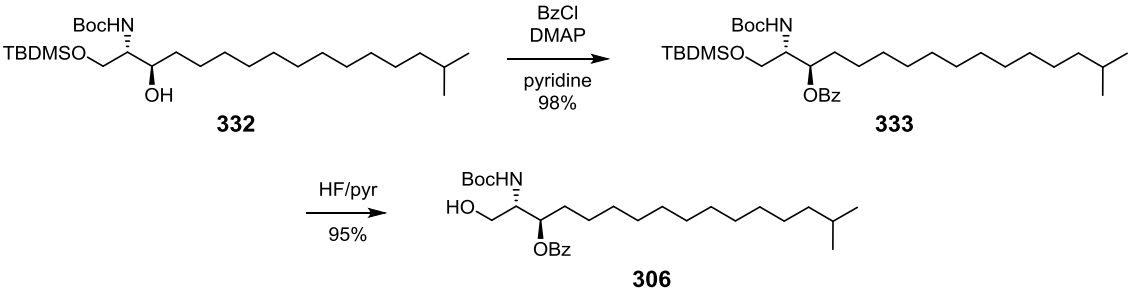
Chelation control, in which a Lewis acid or metal ion coordinates to the carbonyl oxygen and the amine nitrogen, enforces a *syn*-periplanar relationship between the amine and ketone groups, and leads to the *anti* diastereomer (Scheme 3.14). Hoffman and co-workers in 2009 found that reduction of related systems is efficient for reducing agents including $NaBH_4$, $LiBH_4$, $Zn(BH_4)_2$ and $LiAl(O^tBu)_3H$. Reduction with $LiAl(O^tBu)_3H$ in EtOH at -78 ºC gave the best *anti* stereoselectivity (>95:5) yet surprisingly the same reaction in THF gave a 1:1 mixture, indicating an important role for solvent in the diastereoselection. It was suggested that hydroxylic solvents promote the exchange and/or disproportionation of ligand attached to aluminium so that the

substrate can chelate to aluminium and thus allow delivery of hydride seteroselectively through chelation stereocontrol. In addition, temperature control is critical since at -50 ºC, an acid/base reaction between the aluminium hydride and ethanol begins to compete with the reduction.




**Scheme 3.14:** Rationalization of stereoselectivity through chelation control in the reduction of ketone with $LiAl(O^tBu)_3H$.

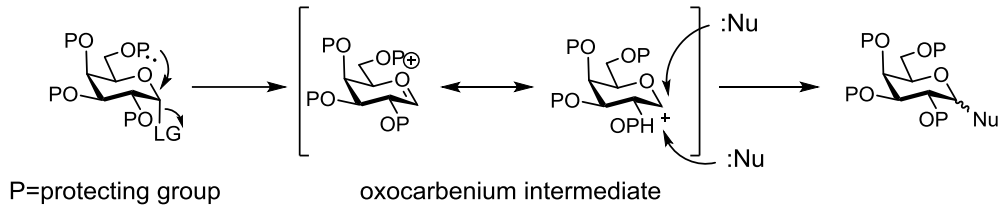
The alcohol **332** was protected with a benzoyl (Bz) group and then the TBDMS group was cleaved using HF/pyr to give the primary alcohol **306** (Scheme 3.15).




**Scheme 3.15:** Preparation of sphinganine alcohol **306**.

### 3.2.4 Synthesis of the α-selective glycosyl donor

Chemical glycosylation reactions involve the reaction of a glycosyl donor and a glycosyl acceptor.[267] Glycosyl donors are normally sugars with a good leaving group at the anomeric position. In a glycosylation reaction, the glycosyl donor is activated to form an oxocarbenium ion intermediate (Scheme 3.16). The glycosyl acceptor is an alcohol that reacts with the anomeric carbon of the glycosyl cation. Glycosylation reactions often result in a mixture of stereoisomers at the anomeric carbon.[268,269]

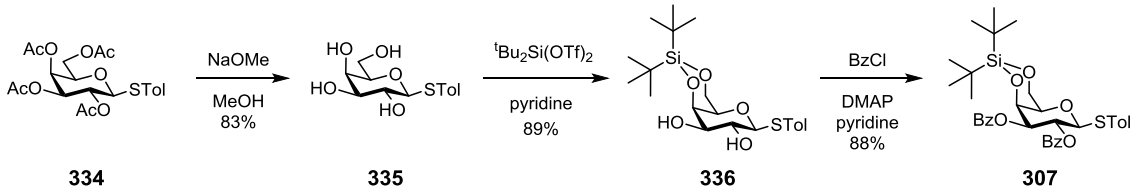



**Scheme 3.16:** Glycosylation via an oxocarbenium ion intermediate.

Stereoselective α-galactosylation of alcohol **306** requires careful selection of hydroxyl-protecting groups for the galactosyl donor. Importantly, the protecting groups should withstand the glycosylation condition, be easily removed subsequently, and promote α-selectively in the glycosylation. Suitable protecting groups on the donor and variation of the reaction conditions can greatly alter the reactivity and stereoselectivity of glycosylation reactions. Unfortunately, reactions that are highly α-selective are often not high yielding.[233] Typically, electron-donating groups such as benzyl ethers can improve the reactivity of the glycosyl donor. The application of 2,3,4,6-tetra-*O*-benzyl-galactosyl bromide in the presence of tetrabutylammonium bromide has been used successfully in the synthesis of α-galactosyl-ceramides.[270] However, these reactions required extended reaction times, are technically demanding, and often give low yields.[271]

Figueroa-Pérez and Schmidt showed that a 4,6-*O*-benzylidene group within a galactosyl trichloroacetimidate can afford α-galactosyl-ceramides in high yield and with good anomeric selectivity.[272] It was suggested that the *cis*-decalin ring with an equatorially-oriented phenyl group hinders nucleophilic attack from the β face. In addition, it can override neighbouring participation by the C-2 acyl functionality, which normally promotes β-galactosylation. Subsequently, Kiso and co-workers introduced the 4,6-*O*-di-*tert*-butylsilylene (DTBS) group,[273] which is a powerful α-directing group in galactosyl donors and, significantly, under circumstances in which the corresponding 4,6-*O*-benzylidene galactosyl donors give only β-products.

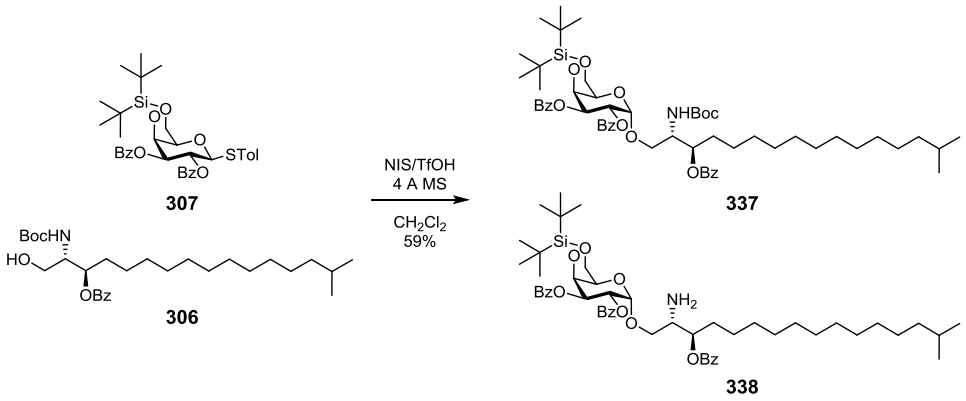
Preparation of thioglycoside **307** started from 4-methylphenyl 1-thio-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside **334**.[273] Deacetylation with NaOMe in MeOH afforded the tetraol **335** in 83% yield. Treatment of tetraol **335** with $^{t}Bu_2Si(OTf)_2$ in pyridine gave the 4,6-*O*-DTBS protected diol **336**. Benzoylation of **336** gave thioglycoside **307** as colourless syrup in 88% yield (Scheme 3.17).




**Scheme 3.17:** Preparation of the thioglycoside **307**.

### 3.2.5 Glycosylation of sphinganine

The coupling of the thioglycoside **307** and sphinganine acceptor **306** was achieved using NIS/TfOH. Because of the possiblility of cleavage of the acid sensitive Boc group on the alcohol **306**, we explored different conditions for the glycosylation (Scheme 3.18, Table 3.2).




**Scheme 3.18:** Glycosylation of the sphinganine acceptor **306** by the galactosyl donor **307**.

**Table 3.2**: Summary of optimisation of glycosylation.

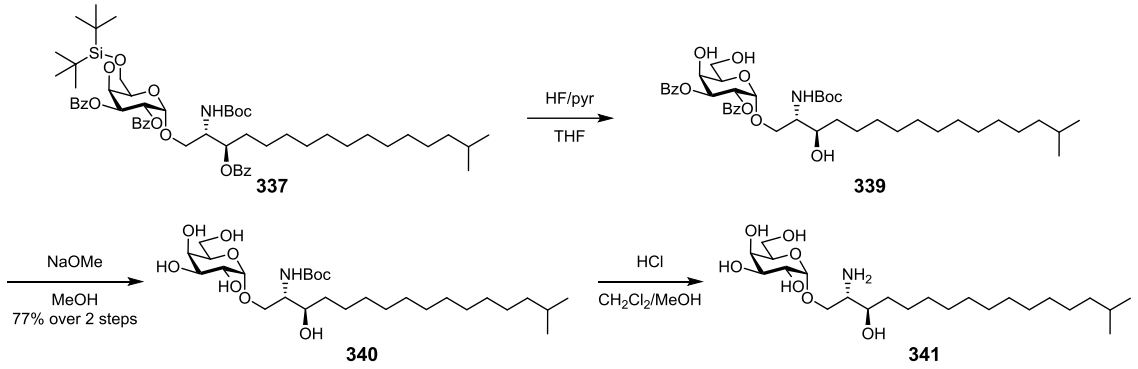
| Entry | Equiv. of TfOH | Temperature (ºC) | Time (h) | Yield (%) |
|---|---|---|---|---|
| 1 | 0.1 | 0 → rt | 3 | 44 |
| 2 | 0.2 | 0 → rt | 3 | 49 |
| 3 | 0.5 | 0 → rt | 3 | 40 |
| 4 | 0.3 | 0 | 2 | 59 |
| 5 | 0.1 | 0 | 2 | 37 |

Using 0.1 eq of TfOH and a generous amount of activated molecular sieves as desiccant, the reaction was initiated at 0 ºC then slowly warmed to rt over 3 h, giving a 44% yield of the glycoside **337**, solely as the α-anomer (entry 1). Several by-products were isolated including the hydrolysed donor and material arising from the migration of the benzoate-protecting group to the primary alcohol of the acceptor.

Molecular sieves, which were used in the glycosylation reaction as a desiccant, can neutralize TfOH, therefore reducing the amount of acid available to activate the donor in the reaction. Increasing the amount of acid used could overcome this limitation. In the second attempt, 0.2 eq of TfOH was used, improving the yield of **337** from 44% to 49% (entry 2). However, increasing the amount of acid used risks cleaving the Boc group. When 0.5 eq of TfOH was used, only 40% yield of **337** was isolated, along with 31% of the amine **338** (entry 3).

To reduce the rate of Boc-group cleavage, 0.3 eq of TfOH was used and the reaction was kept at 0 ºC. Monitoring the reaction every 30 min showed that the Boc group starts to be lost after 3 h. Therefore, the reaction was repeated and quenched at 2 h, even though some of the starting material had not been consumed. The best condition obtained for this glycosylation gave 59% yield of **337** with 4% of the amine **338** (entry 4). One final attempt was made without molecular sieves using just 0.1 eq of TfOH to minimize the cleavage of Boc group (entry 5). However, **337** was obtained in only 37% yield.

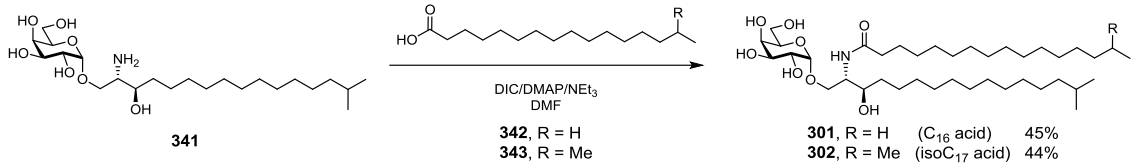
Stepwise deprotection of compound **337** required careful ordering of the steps. We elected to remove the Boc group last to avoid base-induced migration of the benzoyl group from the ester to the amine. Treatment of **337** with HF/pyr followed by treatment with NaOMe and MeOH generated the tetraol **340**. The Boc group was cleaved from **340** with 3 M HCl and gave amine **341** (Scheme 3.19).




**Scheme 3.19:** Stepwise deprotection to avoid migration of other protecting groups to the free amine.

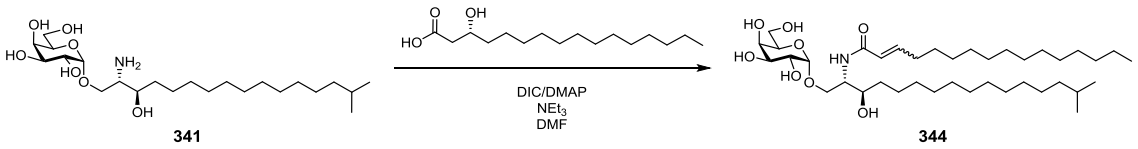
### 3.2.6 Acylation of galactosyl sphinganine

Under Steglich conditions using DMAP and $Et_3N$,[274] coupling of amine **341** and palmitic acid **342** or 15-methylhexadecanoic acid **343** gave the galactosyl ceramides **301** and **302** in 45% and 44% yield, respectively.



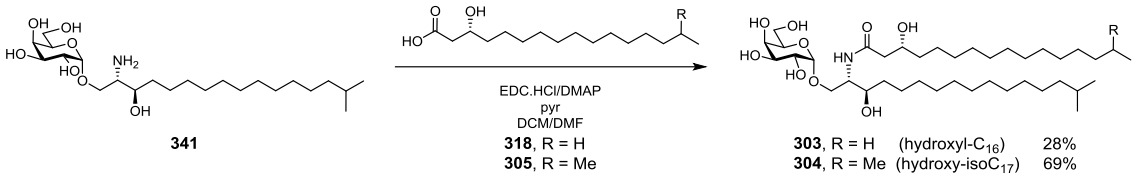

**Scheme 3.20:** Preparation of α-GalCer$_{Bf}$ analogue **301** and **302**.

Applying the same protocol using the (*R*)-3-hydroxy-hexadecanoic acid **318**, resulted in elimination to give **344**. Amide coupling in the Steglich reaction benefits from a base to promote the reaction; however, we presume that the strong triethylamine caused elimination of water from the β-hydroxyl acid or ester.




**Scheme 3.21:** Attempted preparation of α-GalCer$_{Bf}$ analogue **303**.

To avoid elimination, the reaction was repeated using a weaker base, pyridine. In addition, DIC was replaced by EDC.HCl to generate a buffer for the reaction to reduce basicity, and to improve solubility, $CH_2Cl_2$ was added. Gratifyingly, a 28% yield of **303** was obtained, with no elimination observed. Variation of the ratio of EDC.HCl, pyridine and DMAP was undertaken to maintain the pH of the reaction around 7. Ultimately, with 1 eq of DMAP and 2 eq of EDC.HCl and pyridine, 69% yield of **304** was obtained.




**Scheme 3.22:** Steglich reaction under milder conditions generated the α-GalCer$_{Bf\text{-}716}$ **304** and the analogues **303** without elimination.

### 3.2.7 Comparison between synthetic α-GalCer$_{Bf\text{-}716}$ and α-GalCer$_{Bf\text{-}716}$ isolated from *B. fragilis*

Prof. Jon Clardy (Harvard University) supplied original NMR spectra for natural α-GalCer$_{Bf\text{-}716}$ isolated from *B. fragilis*. Comparison of our synthetic α-GalCer$_{Bf\text{-}716}$ with the NMR data of natural α-GalCer$_{Bf\text{-}716}$ from *B. fragilis* confirmed that the synthetic α-GalCer$_{Bf\text{-}716}$ has the same structure as the natural material. The structure of the synthetic α-GalCer$_{Bf\text{-}716}$ was supported by key peaks in the $^1$H NMR spectrum including the amide NH at 7.61 ppm (1 H, d, $J$ = 9.0 Hz), the anomeric proton at 4.64 ppm (1 H, d, $J$ = 3.2 Hz), the α proton of the acid side chain at 2.15 ppm (2 H, qd, $J$ = 13.8, 6.5 Hz), and the two methyl-branched at the end of the tails at 0.83 ppm (6 H, d, $J$ = 6.6 Hz).

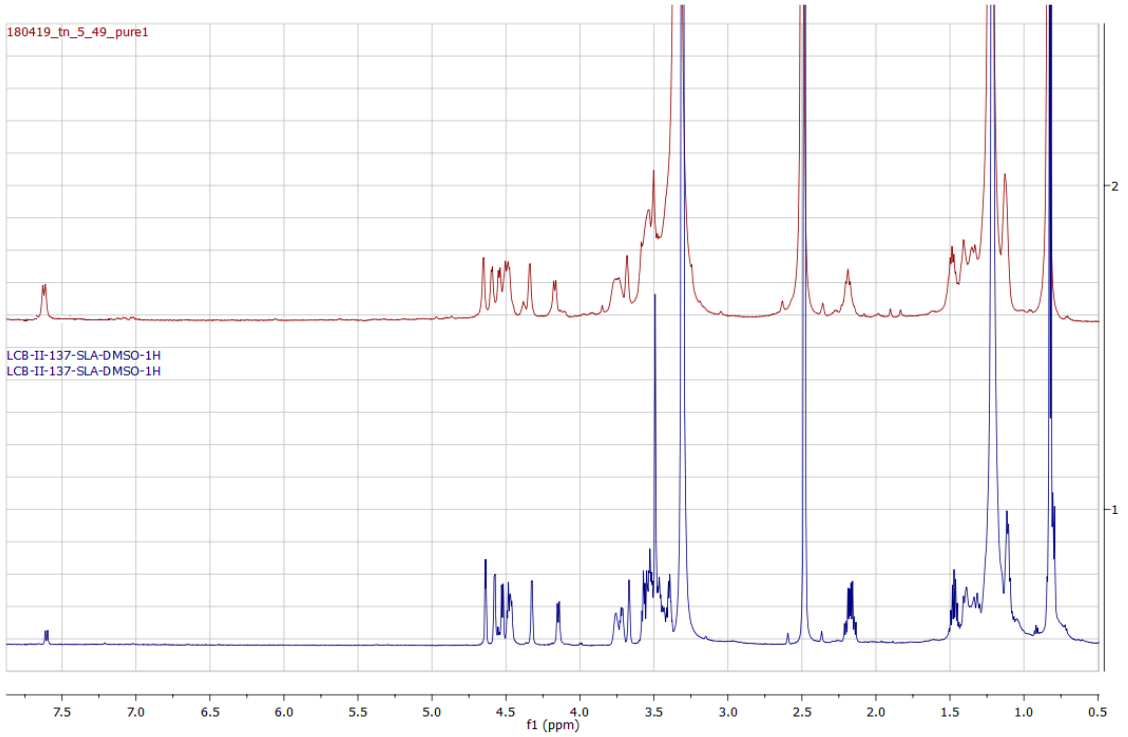



**Figure 3.8:** High resolution $^{1}$H NMR spectra of (top) synthetic **α-GalCer$_{Bf}$** and (bottom) purified α-GalCer$_{Bf\text{-}716}$ extracted from *B. fragilis*, courtesy of Prof. John Clardy, Harvard University. Both spectra are at 600 MHz in DMSO.

### 3.2.8 Immunological study of α-GalCer$_{Bf\text{-}716}$ and its analogues

In collaboration with the Prof. Dale Godfrey's laboratory at the Peter Doherty Institute for Infection and Immunity, University of Melbourne, α-GalCer$_{Bf\text{-}716}$ and its analogues were assessed for their ability to stimulate iNKT cells in both human and mouse CD1d-restricted manner.

Firstly, we investigated whether synthetic α-GalCer$_{Bf\text{-}716}$ was immunogenic. C57BL/6 mice were injected intravenously with KRN7000, α-GalCer$_{Bf\text{-}716}$ or an equivalent volume of vehicle solution. After 3 days, the lymph nodes were harvested and used to prepare a single cell suspension following the protocol developed by Matsuda and co-workers.[275] Fluorescently labelled CD1d tetramers loaded with KRN7000 were used to stain the single cell suspension, and stained cells were detected by flow cytometry. As KRN7000 loaded CD1d tetramers are specific for iNKT cells,

this approach can be used to determine if the injected glycolipids cause expansion of iNKT cells.

Mice stimulated with either KRN7000 or α-GalCer$_{Bf\text{-}716}$ resulted in expansion of iNKT cells to levels higher than for mice injected with vehicle alone (Figure 3.9). Mice treated with 100 ng/ml KRN7000 responded strongly with 8.6% of total cells staining brightly; whereas mice treated with 10 µg/ml α-GalCer$_{Bf\text{-}716}$ had 3.2% stained. In both cases, the population stained was much higher than mice treated with vehicle (0.25%). This indicates that α-GalCer$_{Bf\text{-}716}$ acts as an agonist for iNKT cells *in vivo*, driving their expansion.

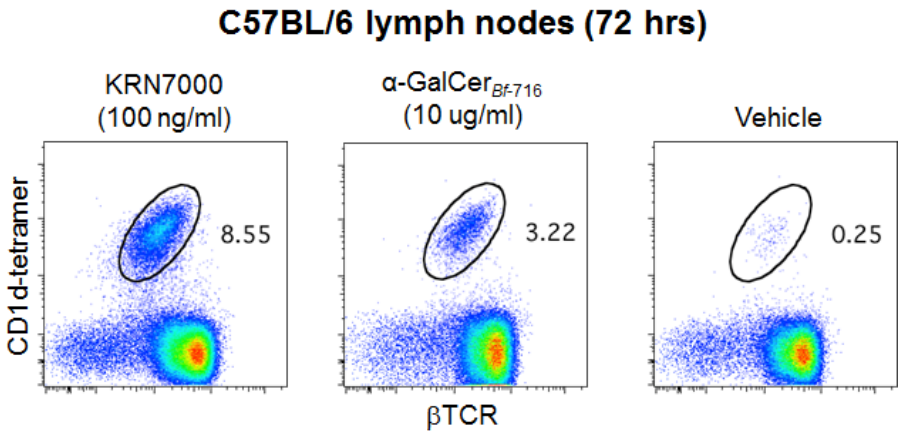



**Figure 3.9:** Representative plots of TCR-transduced cell lines stained with CD1d tetramers loaded with either KRN7000 (100 ng/mL), α-GalCer$_{Bf\text{-}716}$ (10 µg/mL) or vehicle. Plots present the intensity of CD1d tetramer staining (on *y*-axis) relative to a non-agonist control CD1d tetramer (on *x*-axis). Cells have been gated on an equivalent level of TCR expression during analysis; number indicates the geometric mean fluorescence intensity (GMFI) of each CD1d tetramer.

Next, we explored the ability α-GalCer$_{Bf\text{-}716}$ and KRN7000 to be presented by CD1d molecules to type 1 NKT cell-derived TCRs. BW58 cell lines (a TCR-deficient mouse thymoma cell line) were retrovirally transduced with TCR-β chains. CD1d tetramers were produced using a baculovirus expression system, as described by

Matsuda *et al.*[275] Glycolipids were incubated with purified biotinylated CD1d at different lipid:CD1d protein molar ratio (6:1 for KRN7000 and 30:1 for α-GalCer$_{Bf\text{-}716}$) at 37 ˚C overnight. The TCR-transduced NKT cell lines were stained with anti-CD3, and then by glycolipid-loaded CD1d tetramers. The DNA of cells were stained with fluorescently labelled 7-aminoactinomycin D. Fluorescence was assessed by flow cytometer. The results indicate α-GalCer$_{Bf\text{-}716}$ can stain cells in a CD1d dependent manner.

TCR-transduced BW58 cell lines were then tested with CD1d tetramers loaded with the α-GalCer$_{Bf\text{-}716}$ analogues including hydroxyl-C$_{16}$, iso-C$_{17}$ and C$_{16}$. Lipid-loaded CD1d tetramer staining intensity was measured on cells gated for a comparable level of TCR expression by each cell line. CD1d tetramers loaded with iso-C$_{17}$ and C$_{16}$ analogues stained cells at a magnitude comparable with KRN7000. On the other hand, the hydroxyl-C$_{16}$ and the α-GalCer$_{Bf\text{-}716}$ were stained weaker. These results may be explained by differences in ability of glycolipids to load into CD1d, or by differences in affinity of TCR binding to the glycolipid-loaded CD1d molecules, or a combination of both.

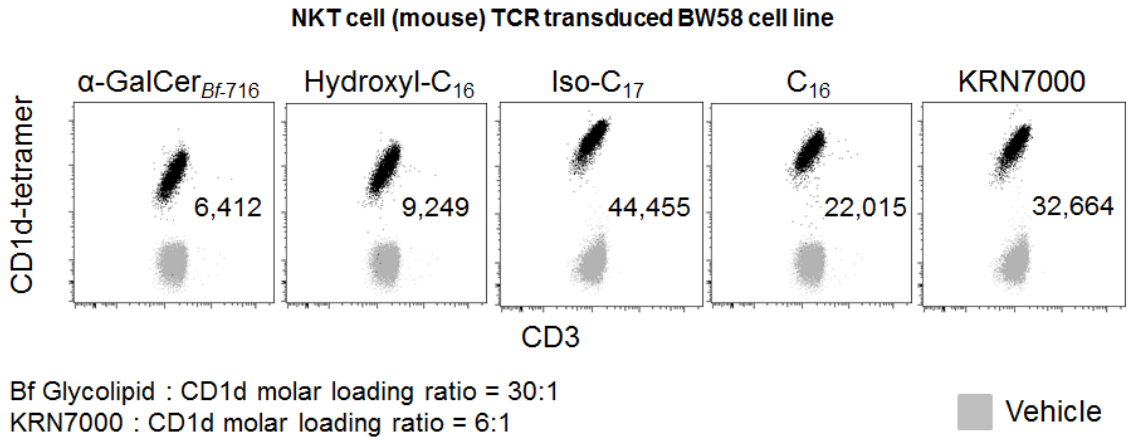



**Figure 3.10:** Representative plots of TCR-transduced cell lines stained with CD1d tetramers loaded with KRN7000, α-GalCer$_{Bf\text{-}716}$ and with its analogues: hydroxyl-C$_{16}$, iso-C$_{17}$ and C$_{16}$. Grey = vehicle control.

Finally, we explored whether naturally circulating human iNKT cells displayed reactivity to CD1d loaded with α-GalCer$_{Bf\text{-}716}$ and its analogues. Peripheral blood cells (PBMCs) from a healthy human donor were isolated by density gradient centrifugation and stained with antibodies specific for CD3, CD14 and CD19, and 7-aminoacetinomycin D (7-AAD) as a test of viability.[276] Glycolipids were prepared in 0.5% tyloxapol and loaded into CD1d. In vitro-expanded PBMC donor cells were enriched with CD1d-KRN7000 and were stained with equal concentrations of CD1d-tetramers loaded with various glycolipids, and analyzed using a flow cytometer.

Human CD1d 'unloaded' tetramer, which contains endogenous lipid antigen, stained just 0.01% of lymphocytes, indicating a very low level of CD1d autoreactivity. Using CD1d tetramers loaded with glycolipids, we detected populations of iNKT cells varying from 0.16% to 0.23% in NKT cell-enriched samples. This indicates that this healthy human donor has naturally occurring iNKT cells that are reactive to the range of α-GalCer$_{Bf\text{-}716}$ analogues.

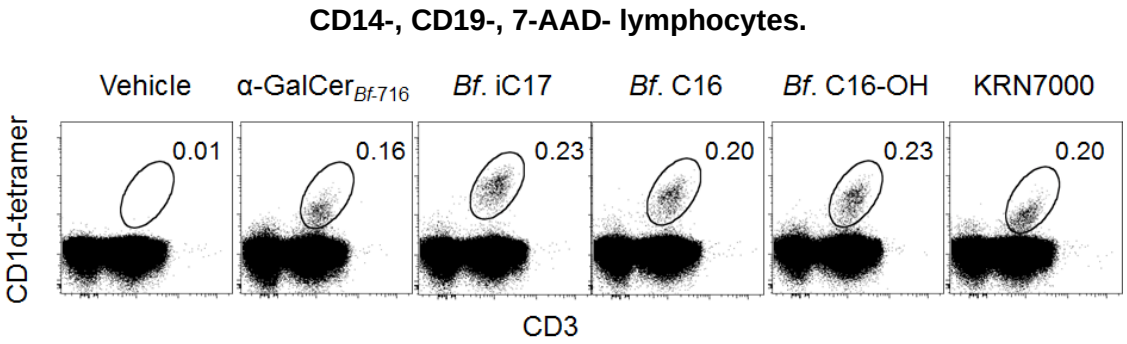


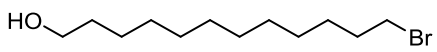

**Figure 3.11:** Flow cytometry analysis showing human CD1d tetramer, loaded with KRN7000, α-GalCer$_{Bf\text{-}716}$ and its analogues or 'endo', staining on CD3$^+$ cells from *in vitro* – expanded CD1d-KRN7000 tetramer-enriched PBMCs from human donor.

## 3.3 Conclusion

In this chapter we report the synthesis of α-GalCer$_{Bf\text{-}716}$ from *B. fragilis*. α-GalCer$_{Bf\text{-}716}$ and analogues were successfully synthesised by glycosylation reaction of a sphinganine acceptor and a galactose donor, followed by an amide coupling with different acid side chains. The $^{1}$H NMR data was an excellent match for the natural material isolated by the laboratory of Prof. Jon Clardy (Harvard University). By way of the penultimate intermediate, a galactosyl sphinganine, we synthesized a series of galactosyl ceramide analogues bearing modified lipid side-chains to allow exploration of structure-activity relationships for binding and activation of CD1d-restricted T cells.

By providing access to a homogeneous sample of this glycolipid, we gained a better understanding of the immunological properties of α-GalCer$_{Bf}$. Synthetic α-GalCer$_{Bf\text{-}716}$ caused expansion of iNKT cells *in vivo*. CD1d-tetramers loaded with straight-chain α-GalCer$_{Bf}$ $C_{16}$ analogues stained iNKT cells brightest, while the branched chain congeners, α-GalCer$_{Bf}$ iso$C_{17}$, $C_{16}$-OH and α-GalCer$_{Bf\text{-}716}$, stained more weakly. Finally, in vitro-expanded human PBMC donor cells displayed reactivity to CD1d loaded with α-GalCer$_{Bf\text{-}716}$ and its analogues

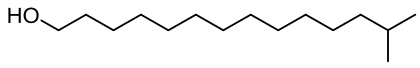
## 3.4 Experimental




**12-Bromododecan-1-ol (309)**

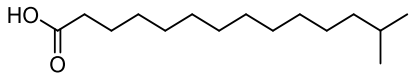
Aqueous HBr (48%) (3.58 mL, 31.6 mmol) was added to a solution of 1,12-dodecanediol **308** (4.00 g, 19.8 mmol) in toluene (100 mL). The mixture was heated under reflux for 72 h. The solution was allowed to cool to r.t and diluted with $Et_2O$, washed with aq. NaOH (2 M), followed by aq. HCl (10%), water and aq. brine, dried ($MgSO_4$) and concentrated. Purification by flash chromatography (EtOAc:pet. spirits, 20:80) afforded **309** as a white solid (3.62 g, 69%). $^{1}$H NMR (400 MHz, $CDCl_3$) δ 3.63-

3.65 (2 H, t, *J* = 6.6 Hz), 3.39-3.42 (2 H, t, *J* = 6.9 Hz), 1.83-1.87 (2 H, m), 1.54-1.58 (2 H, m), 1.39-1.44 (2 H, m), 1.24-1.36 (14 H, m). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 63.2, 34.2, 32.98, 32.95, 29.7, 29.67, 29.65, 29.56, 28.9, 28.3, 25.9.




**13-Methyltetradecan-1-ol (310)**

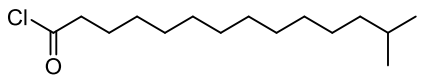
Isopropylmagnesium bromide (2 M in THF; 9.42 mmol, 4.70 mL) and dilithium tetrachlorocuprate (0.1 M in THF; 0.0220 mmol, 0.220 mL) was added into the stirring mixture of **309** (0.500 g, 1.89 mmol) in dry $Et_2O$ (5 mL) at -78 ºC. The mixture was warmed to r.t and stirred overnight. The mixture was then quenched with sat. $NH_4Cl$ and extracted with EtOAc. The combined organic extracted were washed with water, aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). The organic layer was concentrated under reduced pressure to give colourless oil. Purification by flash chromatography (EtOAc:pet. spirits, 20:80) afforded **310** as a colourless oil (0.324 g, 75%). $^1$H NMR (400 MHz, $CDCl_3$) δ 3.63-3.65 (2 H, t, *J* = 6.6 Hz), 1.50-1.58 (3 H, m), 1.22-1.32 (18 H, m), 1.15 (2 H, m), 0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 63.2, 39.1, 32.9, 30.0, 29.7, 29.74, 29.73, 29.71, 29.66, 29.64, 29.5, 28.0, 27.5, 25.8, 22.7.




**13-Methyltetradecanoic acid (311)**

A mixture of **310** (0.324 g, 1.42 mmol), tetrabutylammonium bromide (0.229 g, 0.710 mmol), $KMnO_4$ (1.11 g, 7.02 mmol), AcOH (4.00 ml), water (14.0 mL) and $CH_2Cl_2$ (14.0 mL) was heated under reflux for 16 h. The dark purple mixture was cooled to r.t then acidified by the addition of aq. HCl (3 M). Small portions of $Na_2SO_3$ were added carefully to discharge the colour. The organic phase was collected and the remaining aqueous phase was extracted with $CH_2Cl_2$ and dried ($MgSO_4$), filtered and evaporated

under reduced pressure. Purification by flash chromatography (EtOAc:pet. spirits, 20:80) afforded **311** as a white solid (0.288 g, 85%). $^{1}$H NMR (400 MHz, $CDCl_3$) δ 2.33-2.35 (2 H, t, *J* = 7.5 Hz), 1.64 (2 H, m), 1.50 (1 H, m), 1.24-1.26 (16 H, m), 1.15 (2 H, m), 0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 179.5, 39.2, 34.1, 30.1, 29.85, 29.80, 29.75, 29.69, 29.4, 29.2, 28.1, 27.6, 24.8, 22.8.

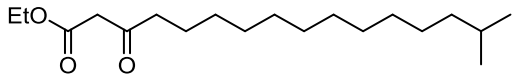



**13-Methyltetradecanoyl chloride (312)**

A mixture of **311** (0.288 g, 1.19 mmol), 5 drops of DMF and thionyl chloride (3.00 ml) was heated under reflux for 2 h. Thionyl chloride was removed by evaporation to afford **312** as a colourless oil that was used immediately without further purification. $^{1}$H NMR (400 MHz, $CDCl_3$) δ 2.86-2.89 (2 H, t, *J* = 7.1 Hz), 1.64 (2 H, m), 1.50 (1 H, m), 1.24-1.26 (16 H, m), 1.15 (2 H, m), 0.86 (6 H, d, *J* = 6.6 Hz).

**General procedure of making β-keto esters (313) and (320)**

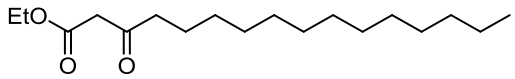
$Et_3N$ (1.10 equiv) and $MgCl_2$ (2.5 equiv) was added to a solution of potassium ethyl malonate (2.1 eq) in dry MeCN (10 mL $mmol^{-1}$) stirred under nitrogen. Stiring was continued at r.t for 2 h and then the solution was cooled to 0 ºC and acid chloride (1 eq) in driy MeCN (10 mL $mmol^{-1}$) was added dropwise. The mixture was warmed to r.t and stirred for 16 h. It then was cooled to 0 ºC, diluted with aq. HCl (1 M) and extracted with $Et_2O$ (3 times). The combined organic layers were washed with aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). The solvent was evaporated and purified by flash chromatography (EtOAc:pet. spirits, 20:80) to afford β-ketoester as a white solid.




**Ethyl-15-methyl-3-oxohexadecanoate (313)**

The General Protocol for making β-keto ester with **312** (0.310 g,1.19 mmol) after flash chromatography afforded **313** as a white solid (0.215 g, 58%). $^{1}$H NMR (400 MHz,

$CDCl_3$) δ 4.17-4.21 (2 H, q, *J* = 7.1 Hz), 3.42 (2 H, s), 2.51-2.54 (2 H, t, *J* = 7.4 Hz), 1.64 (2 H, m), 1.50 (1 H, m), 1.24-1.26 (16 H, m), 1.15 (2 H, m), 0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 202.9, 170.4, 61.2, 49.3, 43.2, 38.2, 30.1, 29.87, 29.82, 29.81, 29.72, 29.67, 28.1, 27.6, 25.6, 22.8.

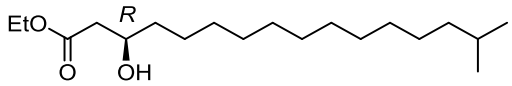



**Ethyl-3-oxohexadecanoate (320)**

The General Protocol for making β-keto ester with **319** (0.492 g, 2.00 mmol, 0.543 mL) after flash chromatography afforded **320** as a white solid (0.257 g, 43%).$^1$H NMR (400 MHz, $CDCl_3$) δ 4.16-4.20 (2 H, t, *J* = 7.1 Hz), 3.41 (2 H, s), 2.50-2.53 (2 H, t, *J* = 7.4 Hz), 1.56-1.59 (2 H, m), 1.24-1.27 (23 H, m), 0.85-0.88 (3 H, t, *J* = 7.0 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 202.8, 170.4, 61.5, 49.5, 43.3, 32.1, 29.82, 29.79, 29.75, 29.60, 29.51, 29.18, 23.63, 22.85, 14.27. HRMS (ESI) *m/z* calcd. for $[C_{18}H_{34}O_3+H]^+$: 299.2581, obsd: 299.2579.

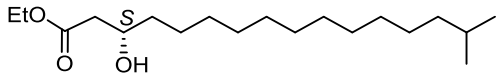
**General procedure for stereoselective Noyori ketone reduction forming 315a, 315b and 321**

Chiral Noyori' catalyst: chloro[(*R*)-2,2'-bis(diphenylphosphino)-1-1'-binaphthyl](*p*-cymene) ruthenium(II) chloride) **314a** or chloro[(*S*)-2,2'-bis(diphenylphosphino)-1-1'-binaphthyl](*p*-cymene) ruthenium(II) chloride) **314b** (0.01 eq) and camphorsulfonic acid (1 eq) were added to a solution of β-keto ester **313** or **320** (1 eq) in dry methanol (10 mL $mmol^{-1}$). The solution was placed under $H_2$ at 39 atm, and stirred at 60 ºC for 3 d. The solvent was evaporated and the crude material was purified by flash chromatography (EtOAc:pet. spirits, 10:90) afford the alcohol.



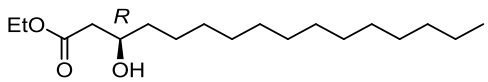
**Ethyl-(*R*)-3-hydroxy-15-methylhexadecanoate (315a)**

The General Protocol for stereoselective reduction with **313** (0.100 g, 0.320 mmol) and **314a** (0.0030 g, 0.0032 mmol) after flash chromatography afforded **315a** as a white solid (0.0644 g, 64%). $[\propto]_{\mathrm{D}}^{25^{\circ}C}$ = - 11.0. $^{1}$H NMR (400 MHz, $CDCl_3$) δ 4.00 (1 H, m), 3.72 (2 H, s), 2.50-2.54 (1 H, dd, *J* = 16.4, 3.1 Hz), 2.39-2.44 (1 H, dd, *J* = 16.4, 9.0 Hz), 1.50 (2 H, m), 1.43 (2 H, m), 1.24-1.26 (16 H, m), 1.15 (2 H, m), 0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 173.7, 68.2, 51.9, 41.2, 39.2, 36.7, 30.1, 29.9, 29.82, 29.81, 29.72, 29.67, 28.1, 27.6, 25.6, 22.8.



**Ethyl-(*S*)-3-hydroxy-15-methylhexadecanoate (315b)**

The General Protocol for stereoselective reduction with **313** (0.078 g, 0.248 mmol) and **314b** (0.0023 g, 0.0025 mmol) after flash chromatography afforded **315b** as a white solid (0.0500 g, 64%). $[\propto]_{\mathrm{D}}^{25^{\circ}C}$ = + 11.3. $^{1}$H NMR (400 MHz, $CDCl_3$) δ 4.00 (1 H, m), 3.71 (2 H, s), 2.49-2.53 (1 H, dd, *J* = 16.4, 3.1 Hz), 2.37-2.43 (1 H, dd, *J* = 16.4, 9.0 Hz), 1.50 (2 H, m), 1.43 (2 H, m), 1.24-1.26 (16 H, m), 1.15 (2 H, m), 0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 173.7, 68.3, 51.9, 41.3, 39.3, 36.7, 30.3, 29.9, 29.82, 29.81, 29.72, 29.67, 28.1, 27.6, 25.6, 22.8.



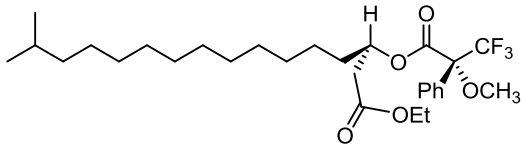
**Ethyl-(*R*)-3-hydroxy-hexadecanoate (321)**

The General Protocol for stereoselective reduction with **320** (0.106 g, 0.355 mmol) and **314a** (0.0033 g, 0.0036 mmol) after flash chromatography afforded **321** as a white solid (0.074 g, 66%). $^{1}$H NMR (400 MHz, $CDCl_3$) δ 3.99-4.02 (1 H, m), 3.71 (2 H, s), 2.85 (1 H, s), 2.50-2.53 (1 H, dd, *J* = 16.4, 3.0 Hz), 2.38-2.44 (1 H, dd, *J* = 16.4, 9.1 Hz), 1.50-

1.55 (1 H, m), 1.42-1.44 (2 H, m), 1.25-1.30 (24 H, m), 0.86-0.89 (3 H, t, *J* = 6.9 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 173.7, 68.2, 51.9, 41.2, 36.7, 32.1, 29.81, 29.73, 29.67, 29.51, 25.6, 22.9, 14.3. HRMS (ESI) *m/z* calcd. for $[C_{19}H_{38}O_3+H]^+$: 301.2737, obsd: 301.2736.

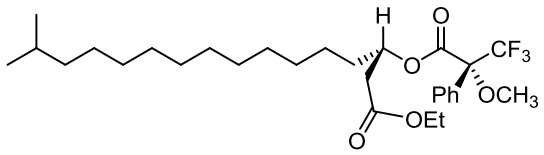
**General procedure for making Mosher's ester compound 317a and 317b**

A mixture of chiral hydroxyl ester **315a** or **315b** (1 eq), chiral α-methoxy-α-trifluoromethylphenylacetic acid **314a** or **314b** (1.1 eq), DCC (1.1 eq) and DMAP (1.1 eq) in dry $CH_2Cl_2$ (10 mL mmol$^{-1}$) was stirred at r.t for 17 h. Aq. HCl (1 M) was added dropwise to the stirring mixture, which was extracted with $CH_2Cl_2$ (3 times), washed with water, aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). Evaporation of the solvent under reduced pressure and purification by flash chromatography (EtOAc:pet. spirits, 10:90) gave a white solid product.




**Ethyl-(*R*)-15-methyl-3-α-methoxy-α-trifluoromethylphenylacetoyl)oxy)hexadecanoate (317a)**

The General Protocol for making Mosher's ester with **315a** (0.0500 g, 0.159 mmol) and the (*R*)-Mosher's ester **314a** (40.9 mg, 0.175 mmol) after flash chromatography afforded (*R*)-Mosher's ester **317a** as a white solid (0.0548 g, 65%). $^1$H NMR (400 MHz, $CDCl_3$) δ 7.51 (2 H, m), 7.37-7.39 (3 H, m), 5.47 (1 H, m), 3.66 (2 H, s), 2.66-2.72 (1 H, dd, *J* = 15.9, 8.2 Hz), 2.57-2.62 (1 H, dd, *J* = 15.9, 4.7 Hz), 1.50 (2 H, m), 1.43 (2 H, m), 1.24-1.26 (16 H, m), 1.15 (2 H, m), 0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 170.7, 166.1, 132.5, 129.7, 128.5, 127.5, 73.5, 55.6, 52.0, 39.2, 38.7, 33.8, 30.1, 29.9, 29.80, 29.75, 29.6, 29.5, 29.3, 28.1, 27.6, 24.8, 22.8.

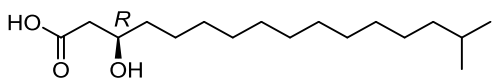



**Ethyl-(*S*)-15-methyl-3-α-methoxy-α-trifluoromethylphenylacetoyl)oxy)hexadecanoate (317b)**

The General Protocol for making Mosher's ester with **315b** (0.0500 g, 0.159 mmol) and the (*S*)-Mosher's ester **314b** (40.9 mg, 0.175 mmol) after flash chromatography afforded (*S*)-Mosher's ester **317b** as a white solid (0.0548 g, 65%). $^{1}$H NMR (400 MHz, $CDCl_3$) δ 7.51 (2 H, m), 7.38-7.42 (3 H, m), 5.47 (1 H, m), 3.58 (2 H, s), 2.61-2.65 (1 H, dd, *J* = 15.9, 8.0 Hz), 2.55-2.58 (1 H, dd, *J* = 15.9, 5.0 Hz), 1.50 (2 H, m), 1.43 (2 H, m), 1.24-1.26 (16 H, m), 1.15 (2 H, m), 0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 170.5, 166.0, 129.8, 129.7, 128.9, 128.5, 127.6, 127.0, 73.6, 55.9, 39.2, 38.6, 35.1, 34.1, 33.9, 30.1, 29.8, 29.79, 29.75, 29.62, 29.55, 29.4, 28.1, 27.6, 22.8.

**General protocol of saponification of the hydroxy ester 315a, 321**

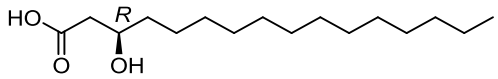
Aq. NaOH (1 M, 1 eq) was added dropwise into the mixture of hydroxy ester **315a** or **321** (1 eq) in MeOH (10 mL mmol$^{-1}$) at 0 ºC and stirred for 4 h. When the reaction finished, aq. HCl (1 M) was added dropwise to quench excess NaOH. The mixture was extracted with $Et_2O$ (3 times), washed with aq. brine and dried ($MgSO_4$). Evaporation of the solvent under reduced pressure and purification by flash chromatography (EtOAc:pet. spirits, 20:80) gave a white solid product.




**(*R*)-3-Hydroxy-15-methylhexadecanoic acid (316)**

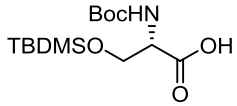
The General Protocol for saponification with **315a** (0.011 mg, 0.0353 mmol) after flash chromatography afforded **316** as a white solid (9.4 mg, 93%). $^{1}$H NMR (400 MHz, $CDCl_3$) δ 4.03 (1 H, m), 2.54-2.58 (1 H, dd, *J* = 16.6, 3.0 Hz), 2.47-2.50 (1 H, dd, *J* =

16.6, 9.0 Hz), 1.43-1.54 (3 H, m), 1.26-1.28 (20 H, m), 1.14-1.15 (3 H, m) 0.87-0.89 (3 H, t, *J* = 6.9 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 177.3, 68.0, 41.1, 39.1, 36.5, 30.0, 29.72, 29.68, 29.66, 29.59, 29.57, 29.49, 28.0, 27.4, 25.5, 22.7. HRMS (ESI) *m/z* calcd. for $[C_{17}H_{34}O_3+H]^+$: 287.2581, obsd: 287.2579.




**(*R*)-3-Hydroxyhexadecanoic acid (318)**

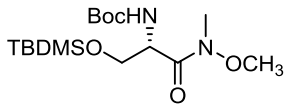
The General Protocol for saponification with **321** (7.36 mg, 0.0245 mmol) after flash chromatography afforded **318** as a white solid (0.00601 g, 90%). $^1$H NMR (400 MHz, $CDCl_3$) δ 4.02-4.05 (1 H, m), 2.56-2.60 (1 H, dd, *J* = 16.6, 3.0 Hz), 2.45-2.50 (1 H, dd, *J* = 16.6, 9.0 Hz), 1.53-1.58 (1 H, m), 1.43-1.49 (2 H, m), 1.26-1.28 (20 H, m), 0.87-0.89 (3 H, t, *J* = 6.9 Hz). $^{13}$C NMR (100 MHz, $CDCl_3$) δ 173.4, 67.9, 51.6, 41.0, 36.4, 31.8, 29.56, 29.48, 29.42, 29.3, 25.4, 22.6, 14.02. HRMS (ESI) *m/z* calcd. for $[C_{16}H_{32}O_3+H]^+$: 273.2424, obsd: 273.2423.




**(2*S*)-2-*tert*-butoxycarbonylamino-3-(*tert*-butyldimethylsilanyloxy)propionic acid (325)**

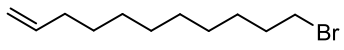
Imidazole (1.02 g, 15.0 mmol) and TBDMS-Cl (2.26 g, 15.0 mmol) were added respectively into the stirring mixture of (*S*)-N-(*tert*-butoxycarbonyl) serine **324** (1.03 g, 5.00 mmol) in dry DMF (25 mL) at 0 ºC and stirred overnight. The mixture was extracted with EtOAc, washed with water, aq. brine and dried ($MgSO_4$). Combined organic solvents were evaporated under reduce pressure and purified by flash chromatography (EtOAc) to afford **325** as the colourless oil (1.39 g, 87%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.32-5.34 (1 H, d, *J* = 8.0 Hz), 4.36 (1 H, bs), 4.09-4.11 (1 H, d, *J* = 7.5 Hz), 3.80-3.84 (1 H, dd, *J* = 10.0, 4.4 Hz), 1.46 (9 H, s), 0.88 (s, 9 H), 0.07 (3 H, s),

0.06 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 174.6, 155.8, 80.5, 63.4, 55.2, 28.5, 26.0, 18.4, -5.40. HRMS (ESI) *m/z* calcd. for $[C_{14}H_{29}NO_4Si+H]^+$: 320.1889, obsd: 320.1887.




**Methyl-(2*S*)-*N*-methyl-2-(*tert*-butoxycarnonylamino)-3-(*tert*-butyldimethylsilanyloxy) propanohydroxamate (322)**

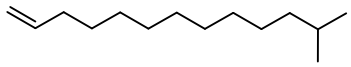
Me(MeO)NH.HCl (0.64 g, 6.53 mmol), DIC (0.82 g, 6.53 mmol, 1.01 mL), DMAP (0.87 mmol, 0.11 g) and $NEt_3$ (0.66 mg, 6.53 mmol, 0.91 mL) was added respectively into the mixture of **325** (1.39 g, 4.35 mmol) in dry $CH_2Cl_2$ (15 mL) at r.t and stirred overnight. The mixture was then filtered, the organic layer was extracted with EtOAc, washed with water, aq. brine, and dried ($MgSO_4$). The organic layer was concentrated and purified by flash chromatography (EtOAc) to afford **322** as the colourless oil (1.15 g, 73%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.33-5.35 (1 H, d, *J* = 8.6 Hz), 4.73 (1 H, s), 3.82-3.83 (1 H, dd, *J* = 9.9, 7.4 Hz), 3.77-3.78 (1 H, dd, *J* = 10.0, 5.2 Hz), 3.73 (3 H, s), 3.19 (3 H, s), 1.41 (9 H, s), 0.84 (9 H, s), 0.01 (3 H, s), 0.00 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 170.8, 155.5, 79.6, 63.6, 61.6, 52.6, 32.3, 28.5, 25.9, 18.4, - 5.40. HRMS (ESI) *m/z* calcd. for $[C_{16}H_{35}N_2O5Si+H]^+$: 363.2310, obsd: 363.2311.




**10-bromo-1-tridecene (327)**

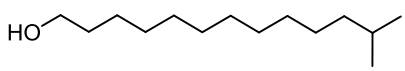
Triphenylphosphine (7.87 g, 30.0 mmol) and tetrabromo carbon (7.96 g, 24 mmol) were added portion into the stirring mixture of 9-decen-1-ol **323** (3.41 g, 4.01 mL, 20.0 mmol) in $CH_2Cl_2$ (100 mL) at 0 ºC and continuing stirred for 2 h. The solvent was concentrated and triphenylphosphine oxide was precipitated with pet. spirits. The precipitate was filtered and washed with pet. spirits. The combined organic solvents were concentrated and purified by flash chromatography (EtOAc:pet. spirits, 10:90) to

afford **327** as the colourless oil (4.57 g, 98%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 5.76-5.86 (1 H, ddt, *J* = 16.9, 10.2, 6.7 Hz), 4.92-5.02 (2 H, dd, *J* = 17.1, 1.5 Hz), 3.99-3.42 (3 H, t, *J* = 6.9 Hz), 2.01-2.06 (2 H, q, *J* = 6.9 Hz), 1.81-1.88 (3 H, m), 1.36-1.44 (4 H, m), 1.29 (8 H, m). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 139.3, 114.3, 34.2, 33.9, 33.0, 29.5, 29.2, 29.1, 28.9, 28.3.



**11-Methyl-1-tridecene (328)**

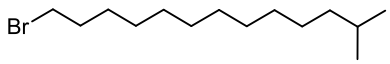
A mixture of **327** (4.57 g, 19.6 mmol) in THF (100 mL) was cooled to -78 ºC before isopropyl magnesium bromide (2.0 M in THF; 58.8 mmol, 29.4 mL) and $Li_2CuCl_4$ (0.1 M in THF; 0.23 mmol, 2.31 mL) were respectively added dropwise into the reaction. After the addition, the reaction was warmed up to r.t and stirred overnight. Sat. $NH_4Cl$ was added to quench the excess of the reagents. The organic layer was extracted with pet. spirits, washed with aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). Combined organic layers were evaporated under reduced pressure and the crude material was purified by flash chromatography (pet. spirits) to afford **328** as a colourless oil (3.50 g, 91%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 5.77-5.87 (1 H, ddt, *J* = 16.9, 10.2, 6.7 Hz), 4.92-5.02 (2 H, dd, *J* = 17.1, 1.4 Hz), 2.02-2.07 (2 H, dd, *J* = 14.2, 6.9 Hz), 1.49-1.54 (1 H, m), 1.38 (1 H, m), 1.26- 1.27 (12 H, m), 1.14-1.16 (2 H, m), 0.86-0.87 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 139.4, 114.2, 39.2, 34.0, 30.1, 29.9, 29.8, 29.7, 29.3, 29.1, 28.1, 27.6, 22.8.




**12-Methyl-1-tridecanol (329)**

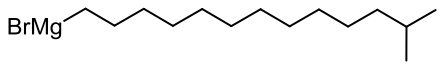
Iodine (5.42 g, 21.4 mmol) in dry THF (40 mL) was added dropwise into the stirring mixture of $NaBH_4$ (1.68 g, 44.5 mmol) in dry THF (20 mL) at 0 ºC for 2.5 h. The mixture was heated under reflux for 1 h until the brown colour disappeared. **328** (3.50 g,

17.8 mmol) in dry THF (20 mL) was added dropwise into the colourless mixture at 0 ºC. The reaction was warmed to r.t and stirred overnight. After that, the mixture was cooled down to 0 ºC again before the mixture of 30% aq. $H_2O_2$/ 3N aq. NaOH (1:1) (60 mL) was added. Each addition released large volume of gas resulting in exothermic reaction. After cooling, the mixture was extracted with EtOAc (3 times), washed with water, aq. $NaHCO_3$, aq. brine, dried ($MgSO_4$) and concentrated under reduced pressure to afford a light-yellow liquid. Purification by flash chromatography (EtOAc:pet. spirits, 20:80) to afford **329** as the light-yellow liquid (3.40 g, 89%). $^1$H NMR (400 MHz, $CDCl_3$) δ 3.61-3.65 (2 H, t, *J* = 6.6 Hz), 1.47-1.57 (3 H, m), 1.25 (16 H, m), 1.15 (2 H, m), 0.85-0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 63.3, 39.2, 33.0, 30.1, 29.9, 29.8, 29.77, 29.76, 29.6, 28.1, 27.6, 29.5, 22.8.



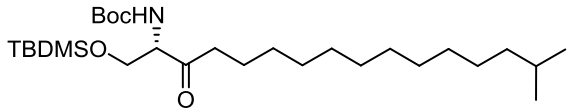

**1-Bromo-12-methyltridecane (330)**

Triphenylphosphine (5.74 g, 21.9 mmol) and tetrabromomethane (5.80 g, 17.5 mmol) were added portion into the mixture of **329** (3.40 g, 15.9 mmol) in dry $CH_2Cl_2$ (100 mL) respectively at 0 ºC. The reaction was continuing stirred at 0 ºC for 2 h. The solvent was concentrated and triphenylphosphine oxide was precipitated with pet. spirits. The mixture was filtered, washed with pet. spirits and the combined organic layers were concentrated. The residue was purified by flash chromatography (EtOAc:pet. spirits, 10:90) to afford **330** as the colourless oil (4.32 g, 98%). $^1$H NMR (400 MHz, $CDCl_3$) δ 3.39-3.42 (2 H, t, *J* = 6.9 Hz), 1.82-1.89 (2 H, m), 1.49-1.53 (1 H, m), 1.42 (2 H, m), 1.26 (15 H, m), 1.16 (2 H, m), 0.85-0.87 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 39.2, 34.2, 33.0, 30.1, 29.8, 29.8, 29.7, 29.6, 28.9, 28.4, 28.1, 27.6, 22.8.



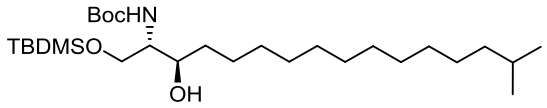

**12-Methyltridecane magnesium bromide (323)**

1-Bromo-12-methyltridecane **330** (3.97 g, 14.3 mmol) in dry THF was added slowly to the mixture of activated Mg (3.47 g, 0.143 mol) and $I_2$ (quantitate amount) to avoid an exothermic reaction. The mixture was stirred at r.t for 1 h before heated under reflux for next 2 h. When the reaction finished, the mixture was used immediately to the next reaction.



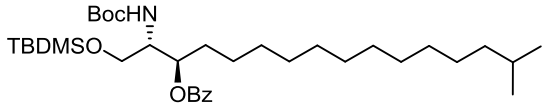

***tert*-Butyl-(*S*)-(1-(((*tert*-butyldimethosilyl)oxy)-15-methyl-3-oxohexadecan-2-yl) carbamate (331)**

Weinreb amide **322** (1.04 g, 2.86 mmol) in dry THF (28 mL) was added slowly to the mixture of **323** freshly prepared at -15 ºC. Reaction was slowly warmed up to rt and stirred overnight. Aq. HCl (1 M) was added dropwise to quench excess Grignard reagent. The mixture was diluted with EtOAc, washed with aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). The combined organic layers were concentrated and purified by flash chromatography (EtOAc:pet. spirits, 5:95) to afford **331** as the colourless oil (0.901 g, 63%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.48-5.49 (1 H, d, *J* = 7.3 Hz), 4.06-4.27 (1 H, bs), 4.03-4.05 (1 H, dd, *J* = 10.3, 2.6 Hz), 3.80-3.83 (1 H, dd, *J* = 10.3, 3.8 Hz), 2.54-2.59 (1 H, ddt, *J* = 46.2, 17.3, 7.4 Hz), 2.44-2.49 (1 H, ddt, *J* = 46.2, 17.3, 7.4 Hz), 1.50 -1.57 (3 H, m), 1.45 (9 H, s), 1.25 (18 H, m), 1.14-1.15 (2 H, m), 0.85-0.87 (6 H, d, *J* = 6.1 Hz), 0.85 (9 H, s), 0.02 (6 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 208.1, 155.5, 79.9, 63.6, 61.4, 40.3, 39.2, 30.1, 29.86, 29.81, 29.76, 29.61, 29.56, 29.4, 28.5, 28.1, 27.6, 25.9, 23.5, 22.8, 18.3, -5.4. HRMS (ESI) *m/z* calcd. for $[C_{28}H_{57}NO_4Si+H]^+$: 500.4130, obsd: 500.4138.




***tert*-Butyl-(2*S*,3*R*)-1-(((*tert*-butyldimethosilyl)oxy)-3-hydroxy-15-methylhexadecan-2-yl) carbamate (332)**

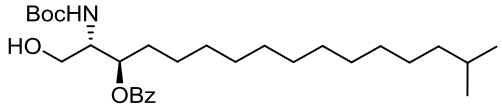
Lithium tri-*tert*-butoxyalmuniohydride (1.01 g, 3.96 mmol) was added slowly into the mixture of **331** (0.901 g, 1.80 mmol) in dry EtOH (18 mL) at -78 ºC. Reaction was warmed up to r.t and stirred for 4 h. When the reaction finished, the mixture was filtered, diluted with EtOAc, washed with aq. brine and dried ($MgSO_4$). The combined organic layers were concentrated and purified by flash chromatography (EtOAc:pet. spirits, 10:90) to afford secondary alcohol **332** as the colourless oil (0.796 g, 88%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.31-5.33 (1 H, d, *J* = 7.8 Hz), 3.94-3.97 (1 H, dd, *J* = 10.6, 2.7 Hz), 3.83 (1 H, d, *J* = 10.1 Hz). 3.62-3.64 (1 H, dd, *J* = 8.4, 4.2 Hz), 3.50 (1 H, bs), 3.02-3.04 (1 H, d, *J* = 8.8 Hz), 1.50-1.53 (4 H, m), 1.45 (9 H, s), 1.25 (19 H, m), 1.15 (2 H, m), 0.90 (9 H, s), 0.85-0.87 (6 H, d, *J* = 6.6 Hz), 0.08 (6 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 155.8, 79.5, 74.5, 63.7, 53.9, 39.2, 35.1, 30.1, 29.9, 29.8, 29.81, 29.75, 29.72, 28.6, 28.1, 27.6, 26.1, 26.0, 22.8, 18.3, -5.4, -5.0. HRMS (ESI) *m/z* calcd. for $[C_{28}H_{59}NO_4Si+H]^+$: 502.4286, obsd: 502.4293.




**(2*S*,3*R*)-2-((*tert*-Butoxycarbonyl)amino)-1-((*tert*-butyldimethylsilanyl)oxy-15-methylhexadecan-3-yl) benzoate (333)**

DMAP (97.1 mg, 0.80 mmol) and benzoyl chloride (0.891 g, 6.34 mmol, 0.736 mL) were added respectively in to the mixture of **332** (0.796 g, 1.59 mmol) in dried pyridine (2 mL) at r.t and stirred over 3 h. The reaction was extracted with EtOAC, washed with aq. $NaHCO_3$, aq. brine and dry ($MgSO_4$). The combined organic layers was concentrated and purified by flash chromatography (EtOAc:pet. spirits, 10:90) to afford

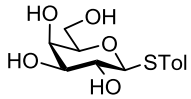
silyl ether **333** as the colourless oil (0.944 g, 98%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 8.03-8.04 (2 H, d, *J* = 7.8 Hz), 7.54-7.57 (1 H, t, *J* = 7.4 Hz), 7.42-7.45 (2 H, d, *J* = 7.7 Hz), 5.20 (1 H, s), 4.91-4.93 (2 H, d, *J* = 9.1 Hz), 3.97 (1 H, s), 3.73-3.75 (1 H, dd, *J* = 10.2, 3.5 Hz), 3.66-3.69 (1 H, dd, *J* = 10.2, 4.3 Hz), 1.47-1.53 (1 H, m), 1.44 (9 H, s), 1.22-1.25 (21 H, m), 1.13-1.16 (2 H, m), 0.87 (9 H, s), 0.85-0.86 (6 H, d, *J* = 6.6 Hz), 0.00 (3 H, s), - 0.02 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 166.0, 155.7, 133.0, 130.6, 129.8. 128.5, 79.6, 74.4, 62.1, 53.9, 39.2, 31.3, 30.1, 29.9, 29.80, 29.77, 29.73, 29.6, 29.5, 28.5, 28.1, 27.6, 26.0, 25.4, 22.8, 18.4, 14.1, -5.4, -5.5. HRMS (ESI) *m/z* calcd. for $[C_{35}H_{63}NO_5Si+H]^+$: 606.4548, obsd: 606.4557.




**(2*S*,3*R*)-2-((*tert*-Butoxycarbonyl)amino)-1-hydroxy-15-methylhexadecan-3-yl) benzoate (306)**

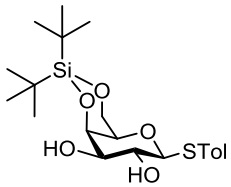
(2*S*,3*R*)-2-((*tert*-butoxycarbonyl)amino)-1-((*tert*-butyldimethylsilanyl)oxy-15-methylhexadecan-3-yl) benzoate **333** (0.944 g, 1.56 mmol) in dry THF (16 mL) was added dropwise into the buffer solution of HF/pyr (70:30 v/v; 82.7 mmol, 2.15 mL), in dried pyridine (9.27 mL) at 0 ºC and stirred overnight. The reaction was quenched by slowly addition of aq. $NaHCO_3$, extracted with EtOAC, washed with aq. brine and dried ($MgSO_4$). The combined organic layers were concentrated and purified by flash chromatography (EtOAc:pet. spirits, 10:90) to afford sphinganine alcohol **306** as the colourless oil (0.729 g, 95%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 8.04-8.06 (2 H, d, *J* = 7.8 Hz), 7.58-7.62 (1 H, t, *J* = 7.4 Hz), 7.45-7.49 (2 H, t, *J* = 7.7 Hz), 5.21-5.23 (1 H, d, *J* = 9.1 Hz), 5.04-5.08 (1 H, dd, *J* = 13.5, 7.1 Hz), 3.82 (1 H, m), 3.64 (2 H, bs), 2.84 (1 H, bs), 1.77-1.81 (2 H, m), 1.50 (1 H, m), 1.45 (9 H, s), 1.22-1.29 (18 H, bs), 1.13-1.14 (2 H, m), 0.85-0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 167.7, 156.0, 133.7,

130.0, 129.6, 128.7, 80.0, 74.4, 61.9, 54.6, 39.2, 31.5, 30.1, 29.9, 29.79, 29.77, 29.68, 29.55, 29.5, 28.5, 28.1, 27.6, 25.7, 22.8, 14.0. HRMS (ESI) *m/z* calcd. for $[C_{29}H_{49}NO_5+H]^+$: 492.3684, obsd: 492.3689.




**4-Methylphenyl 1-thio-β-D-galactopyranoside (335)**

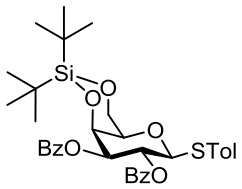
NaOMe (25% wt) was added dropwise into the mixture of 4-methylphenyl 2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-galactopyranoside **334**[277] (7.00 g, 15.4 mmol) in dry MeOH (77 mL) to adjust pH = 10. The reaction was stirred overnight before Dowex $H^+$ was added to neutralize the reaction. The mixture was filtered and washed with MeOH. The combined solvent was evaporated under pressure and the residue was recrystallised with EtOAc and drops of MeOH to afford **335** as a white solid (3.66 g, 83%). $^1$H NMR (400 Hz, $CD_3OD$) δ 7.44-7.46 (2 H, d, *J* = 8.0 Hz), 7.10-7.12 (2 H, d, *J* = 8.0 Hz), 4.49-4.51 (1 H, d, *J* = 9.6 Hz, H1), 3.88-3.89 (1 H, d, *J* = 2.4 Hz, H4), 3.68-3.77 (2 H, qd, *J* = 11.4, 6.1 Hz, H6a, H6b), 3.51-3.60 (2 H, m, H2, H5), 3.46-3.49 (1 H, dd, *J* = 9.2, 3.3 Hz, H3), 2.31 (3 H, s). $^{13}$C NMR (100 Hz, $CD_3OD$) δ 138.4, 132.9, 132.1, 130.5, 90.7. 80.6, 76.4, 71.0, 70.4, 62.6, 21.1.




**4-Methylphenyl-4,6-*O*-di-*tert*-butylsilylene-1-thio-β-D-galactopyranoside (336)**

Di-*tert*-butylsilyl-bis(trifluromethanesulfonate) (1.69 g, 3.84 mmol, 1.25 mL) was added dropwise into the mixture of compound **335** (1.0 g, 3.49 mmol) in dry pyridine (35 mL) at 0 ºC for 5 mins and continue stirred for next 10 mins at same temperature before MeOH was added dropwise to quench excess $^t$Bu$_2$Si(OTf)$_2$. The solvent was

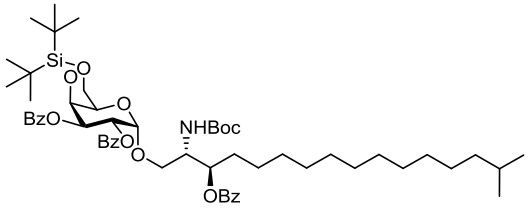
evaporated, and the residue was dissolved in EtOAc, washed with water, aq. brine and dried ($MgSO_4$). The combined organic solvents were concentrated and purified by flash chromatography (EtOAc) to afford diol **336** as the colourless oil (1.33 g, 89%). $^1$H NMR (400 Hz, $CDCl_3$) δ 7.43-7.45 (2 H, d, *J* = 8.0 Hz), 7.08-7.10 (2 H, d, *J* = 7.9 Hz), 4.46-4.48 (1 H, d, *J* = 9.8 Hz, H1), 4.42-4.43 (1 H, d, *J* = 3.2 Hz, H4), 4.25 (2 H, s, H6a, H6b), 3.69-3.74 (1 H, t, *J* = 9.3 Hz, H2), 3.50-3.52 (1 H, dd, *J* = 8.5, 2.5 Hz, H3), 3.42 (1 H, s, H5), 2.34 (3 H, s), 1.05 (9 H, s), 1.02 (9 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 138.4, 133.5, 129.9, 119.2, 89.6, 75.4, 75.2, 72.7, 70.8, 67.2, 27.7, 27.6, 23.5, 21.3, 20.8. HRMS (ESI) *m/z* calcd. for $[C_{14}H_{27}O_5Si+H]^+$: 427.1969, obsd: 427.1971.




**4-Methylphenyl 2,3-*O*-di-benzoyl-4,6-*O*-di-*tert*-butylsilylene-1-thio-β-D-galactopyranoside (307)**

DMAP (0.19 g, 1.55 mmol) and benzoyl chloride (2.18 g, 15.5 mmol, 1.80 mL) were added slowly into the mixture of compound **336** (1.33 g, 3.10 mmol) in dry pyridine (15 mL) and stirred at r.t over 3 h. The reaction was extracted with EtOAc, washed with water, aq. $NaHCO_3$, aq. brine and dried ($MgSO_4$). The combined organic solvent was concentrated and purified by flash chromatography (EtOAc) to afford dibenzoate **307** as the colourless oil (1.73 g, 88%). $^1$H NMR (400 Hz, $CDCl_3$) δ 7.99-8.00 (2 H, d, *J* = 7.2 Hz), 7.49-7.54 (2 H, m), 7.37-7.41 (6 H, m), 7.06-7.08 (2 H, d, *J* = 7.8 Hz), 5.88-5.93 (1 H, t, *J* = 10.0 Hz), 5.17-5.20 (1 H, dd, *J* = 9.8, 1.9 Hz), 4.86-4.89 (2 H, m), 4.30-4.32 (2 H, d, *J* = 4.7 Hz), 3.62 (1 H, s), 2.32 (3 H, s), 1.15-1.16 (d, 9 H, *J* = 1.0 Hz), 0.95-0.96 (d, 9 H, *J* = 0.7 Hz). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 166.3, 165.6, 138.3, 133.36, 133.34, 133.27, 130.1, 129.6, 129.94, 129.90, 129.6, 128.53, 128.48, 88.0, 75.7, 75.1, 70.6,

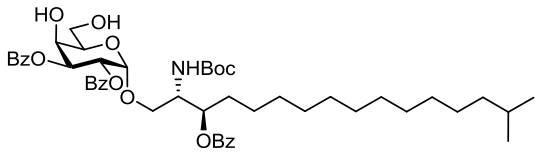
68.4, 67.3, 27.65, 27.60, 23.4, 21.3, 20.9. HRMS (ESI) *m/z* calcd. for $[C_{28}H_{35}O_7Si+H]^+$: 635.2504, obsd: 635.2496.




**(2*S*,3*R*)-3-(Benzoyloxy)-2-((*tert*-butoxycarbonyl)amino)-15-methylhexadecanyl-2,3-di-*O*-benzoyl-4,6-*O*-di-*tert*-butylsilylene-1-thio-β-D-galactopyranoside (337)**

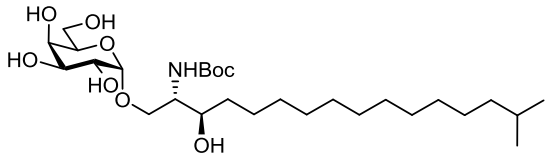
Thioglycoside **307** (0.586 g, 1.14 mmol) and the sphinganine alcohol **306** (0.469 g, 0.954 mmol) were co-evaporated twice with dry toluene (2.00 mL) and then dissolved in dry $CH_2Cl_2$. Activated molecular sieves (3 Å) was added and the mixture was stirred at r.t for 30 mins. N-Iodosuccinimide (300.5 mg, 1.34 mmol) was added at r.t. After 5 mins, the mixture was cooled down to 0 ºC before TfOH (42.9 mg, 0.286 mmol, 25 µL) was added dropwise; the mixture slowly turned dark pink and red. The reaction was stirred at 0 ºC for 2 h, before diluted with $CH_2Cl_2$, and quenched with aq. $Na_2S_2O_3$, aq. $NaHCO_3$, washed with water, aq. brine and dried ($MgSO_4$). The organic layers were combined and concentrated. The residue was purified by flash chromatography (EtOAc:Toluene, 5:95) to afford **337** as the colourless oil (0.564 g, 59%). $^1$H NMR (400 Hz, $CDCl_3$) δ 7.96-8.02 (6 H, m), 7.48-7.55 (3 H, m), 7.33-7.42 (6 H, m), 5.44-5.80 (1 H, dd, *J* = 3.7, 10.6 Hz), 5.52-5.54 (1 H, dd, *J* = 3.1, 10.6 Hz), 5.25 (1 H, d, *J* = 3.2 Hz), 5.17 (1 H, m), 4.97-4.99 (1 H, d, *J* = 8.9 Hz), 4.88-4.89 (1 H, d, *J* = 2.9 Hz), 4.15-4.21 (2 H, m), 4.07-4.12 (1 H, m), 3.90 (1 H, s), 3.80-3.83 (1 H, dd, *J* = 4.6, 10.5 Hz), 3.60-3.63 (1 H, dd, *J* = 10.7, 4.9 Hz), 1.49-1.51 (1 H, m), 1.43 (9 H, s), 1.18-1.21 (19 H, m), 1.13- 1.14 (3 H, m), 1.11 (9 H, s), 0.95 (9 H, s), 0.85-0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 166.3, 166.2, 166.0, 155.5, 133.26, 133.22, 130.1, 129.97, 129.89, 129.85, 129.80, 129.1, 128.8, 128.57, 128.51, 128.47, 97.8, 80.0, 79.4, 74.8,

71.3, 71.1, 68.5, 67.8, 67.4, 67.0, 39.2, 30.1, 29.86, 29.82, 29.78, 29.71, 29.63, 29.59, 28.5, 28.1, 27.64, 27.57, 27.4, 25.5, 23.4, 22.8, 20.9, 14.3. HRMS (ESI) *m/z* calcd. for $[C_{57}H_{83}NO_{12}Si+H]^+$: 1002.5757, obsd: 1002.5765.



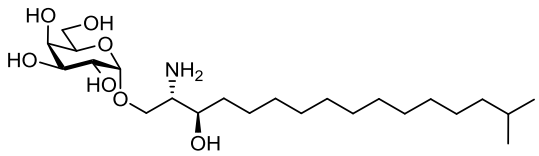

**(2*S*,3*R*)-3-(Benzoyloxy)-2-((*tert*-butoxycarbonyl)amino)-15-methylhexadecanyl-2,3-di-*O*-benzoyl-1-thio-β-D-galactopyranoside (339)**

HF/pyr (70:30 v/v; 70.4 mmol, 1.83 mL) was added dropwise into the stirring mixture of **337** (0.564 g, 0.563 mmol) in dry THF (5.6 mL) at 0 ºC and stirred for 4 h. The reaction was quenched by slowly addition of aq. $NaHCO_3$, diluted with EtOAC, washed with water, aq. brine and dried ($MgSO_4$). The organic layer was concentrated and purified by flash chromatography (EtOAc:$CH_2Cl_2$, 30:70) to afford **339** as the colourless oil (0.412 g, 85%). $^1$H NMR (400 Hz, $CDCl_3$) δ 7.94-8.00 (6 H, m), 7.47-7.57 (3 H, m), 7,31-7.44 (6 H, m), 5.69-5.73 (1 H, dd, *J* = 10.7, 3.3 Hz), 5.63-5.67 (1 H, dd, *J*-10.7, 2.6 Hz), 5.30-5.31 (1 H, d, *J* = 3.0 Hz), 5.17-5.20 (1 H, m), 5.04-5.06 (1 H, d, *J* = 9.2 Hz), 4.44 (1 H, s), 4.14 (1 H, s), 4.07 (1 H, m), 3.93-3.96 (1 H, m), 3.82-3.87 (2 H, m), 3.69-3.73 (1 H, m), 2.76 (1 H, s ), 2.47 (1 H, m), 1.70-1.73 (2 H, m), 1.49-1.54 (1 H, m), 1.44 (9 H, s), 1.19-1.21 (20 H, m), 0.85-0.86 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 166.1, 165.98, 165.95, 155.7, 133.6, 133.3, 133.3, 130.0, 129.96, 129.93, 129.8, 129.5, 128.63, 128.59, 128.5, 98.3, 74.9, 71.3, 69.9, 69.6, 68.7, 63.2, 61.7, 53.0, 39.2, 31.6, 30.1, 29.86, 29.82, 29.78, 29.71, 29.6, 29.5, 28.1, 27.6, 25.5, 22.8. HRMS (ESI) *m/z* calcd. for $[C_{49}H_{67}NO_{12}+H]^+$: 862.4736, obsd: 862.4716.




**(2*S*,3*R*)-2-((*tert*-butoxycarbonyl)amino)-3-hydroxy-15-methylhexadecanyl-1-thio-β-D-galactopyranoside (340)**

NaOMe (25% wt in MeOH) was added dropwise into the stirring mixture of **339** (0.412g, 0.478 mmol) dissolved in MeOH to adjust the pH = 10. The mixture was stirred at r.t overnight before Dowex ($H^+$) was added to neutralise the mixture. The mixture was filtered, washed with MeOH and the solvent was evaporated and the crude was purified by flash chromatography (MeOH:$CHCl_3$, 20:80) to afford **35** as the colourless oil (0.202 g, 89%). $^1$H NMR (400 Hz, $CDCl_3$) δ 5.24 (1 H, s), 4.96 (1 H, s), 4.10 (1 H, s), 3.96 (2 H, m), 3.84 (3 H, m), 3.79 (1 H, m), 3.69 (3 H, m), 3.05 (1 H, s), 2.70 (1 H, s), 2.47 (1 H, s), 2.34 (1 H, s), 1.49 (1 H, hex), 1.45 (9 H, s), 1.25 (20 H, m), 1.15 (2 H, m), 0.86-0.87 (6 H, d, *J* = 6.6 Hz). $^{13}$C NMR (100 Hz, $D_2O$) δ 158.2, 101.2, 80.20, 72.4, 71.9, 71.5, 71.71, 71.0, 79.4, 68.8, 62.7, 56.6, 40.2, 34.9, 31.0, 30.8, 30.7, 29.1, 28.8, 28.5, 26.7, 23.0. HRMS (ESI) *m/z* calcd. for $[C_{28}H_{55}NO_9+H]^+$: 550.3950, obsd: 550.3956.



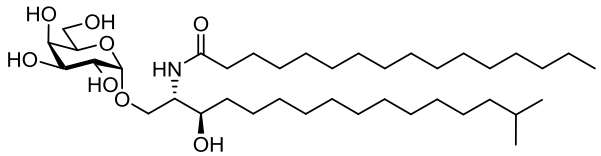

**(2*S*,3*R*)-2-Amino-3-hydroxy-15-methylhexadecanyl-1-thio-β-D-galactopyranoside (341)**

HCl (3 M, 1.22 mL) was added dropwise into the mixture of starting material **339** (0.202 g, 0.367 mmol) dissolved in $CH_2Cl_2$/MeOH (1:1, 7.34 mL) at 0 ºC. The mixture was warmed up to r.t and stirred until all the starting material was consumed. Solvent was evaporated and co-evaporated with toluene (0.6 mL) 3 times to remove all the

water. The crude material was pushed through the next reaction without any purification. HRMS (ESI) *m/z* calcd. for $[C_{23}H_{47}NO_7+H]^+$: 450.3425, obsd: 450.3425.

**General protocol of the coupling with palmitic acid and isoC17-acid**

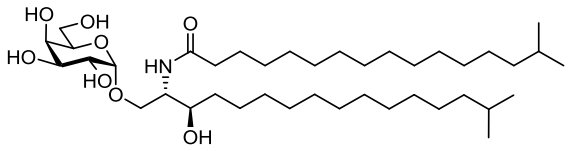
DMAP (2 eq) and DIC (2 eq) were added respectively into the mixture of acid lipid (1.1 eq) in dried $CH_2Cl_2$ (10 mL $mmol^{-1}$) at 0 ºC and stirred at r.t for 15 mins. At the same time, $Et_3N$ (1.2 eq) was added to the mixture of amine (1 eq) in dry DMF, and then transferred into the mixture of acid lipid at 0 ºC. The resulted mixture was warmed up to r.t and stirred overnight. The solvent was evaporated and the crude material was purified immediately by flash chromatography (MeOH:EtOAc, 5:95) to afford yellow solid which was purified once more time with RC ($H_2O$:MeOH, 20:80) to afford product as the white solid.




**(2*S*,3*R*)-3-Hydroxy-2-palmitoylamino-15-methylhexadecanyl-1-thio-β-D-galactopyranoside** (**301)**

The General Protocol for coupling between the amine **341** (0.020 g, 0.045 mmol) and palmitic acid **342** (0.0125 g, 0.049 mmol). After purification, the product was isolated as white solid (0.014 g, 45 %). $^1$H NMR (400 Hz, DMSO-$d_6$) δ 7.56-7.58 (1 H, d, *J* = 9.2 Hz, NH), 4.65-4.66 (1 H, d, *J* = 2.3 Hz, a), 4.54-4.55 (1 H, d, *J* = 6.4 Hz, 3-OH), 4.50-4.51 (2 H, d, *J* = 4.7 Hz, d-OH, f-OH), 4.35-4.36 (1 H, d, *J* = 3.6 Hz, c-OH), 4.15-4.16 (1 H, d, *J* = 7.8 Hz, b-OH), 3.72 (1 H, m, H2), 3.68 (1 H, m, H3), 3.45-3.58 (8 H, m, b/d/e/f'/H1a/H1b/3), 2.03-2.10 (1 H, dq, *J* = 13.7, 6.7 Hz, H2'), 1.46-1.52 (3 H, m, H15, H15'), 1.41 (2 H, m, H4), 1.23 (41 H, m, lipid chain), 1.13 (4 H, m, H14, H14'), 0.83-0.86 (3 H, m, H16'), 0.83-0.84 (6 H, d, *J* = 6.6 Hz, H16). $^{13}$C NMR (100 MHz, DMSO-$d_6$) δ 170.8, 99.9, 71.2, 69.8, 69.5, 69.3, 68.9, 68.7, 60.6, 53.4, 44.1, 38.4, 36.6,

33.6, 31.3, 29.3, 29.2, 29.1, 29.0, 28.9, 28.7, 28.6, 28.5, 28.3, 27.4, 26.9, 25.0, 22.5, 22.1, 14.0. HRMS (ESI) *m/z* calcd. for $[C_{39}H_{77}NO_8+H]^+$: 688.5722, obsd: 688.5728.



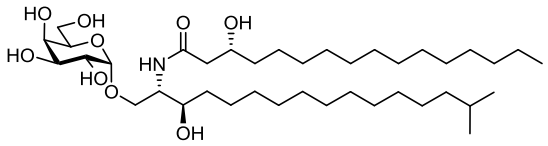

**(2*S*,3*R*)-3-Hydroxy-2-(15-methylhexadecanoyl)amino-15-methylhexadecanyl-1-thio-β-D-galactopyranoside** (**302)**

The General Protocol for coupling between the amine **341** (0.020 g, 0.045 mmol) and 15-methylhexadecanoic acid **343** (0.0134 g, 0.0495 mmol). After purification, the product **302** was isolated as white solid (0.014 g, 44%). $^1$H NMR (400 Hz, DMSO-$d_6$) δ 7.56-7.58 (1 H, d, *J* = 8.9 Hz, NH), 4.65-4.66 (1 H, d, *J* = 3.2 Hz, a), 4.54-4.55 (1 H, d, *J* = 6.5 Hz, OH-f), 4.50-4.51 (2 H, d, *J* = 5.1 Hz, OH-d, OH-f), 4.35-4.36 (1 H, d, *J* = 4.1 Hz, OH-c), 4.15-4.17 (1 H, d, *J* = 7.1 Hz, OH-b), 3.71 (1 H, m, H2), 3.68 (1 H, m, c), 3.45-3.59 (8 H, m, b/d/e/f/f'/H1a/H1b/H3), 2.03-2.09 (1 H, td, *J* = 13.6, 7.1 Hz, H2'), 1.46-1.51 (2 H, m, H15, H15'), 1.41 (1 H, m, H4a), 1.23 (41 H, m, lipid chain), 1.12-1.13 (4 H, m, H14, H14'), 0.83-0.84 (12 H, d, *J* = 6.6 Hz, H16, H16'). $^{13}$C NMR (100 MHz, DMSO-$d_6$) δ 170.0, 100.0, 71.3, 69.9, 69.5, 69.1, 68.9, 68.7, 60.9, 53.4, 44.3, 38.5, 36.6, 33.6, 31.3, 29.4, 29.2, 29.1, 29.0, 28.9, 28.8, 28.6, 28.5, 28.4, 27.4, 26.9, 24.9, 22.5, 22.0, 14.1. HRMS (ESI) *m/z* calcd. for $[C_{40}H_{79}NO_8+H]^+$: 702.5878, obsd: 702.5888.

**General protocol of the coupling with hydroxy acid**

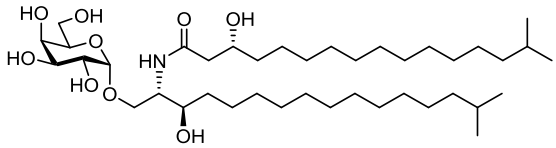
DMAP (1 eq) and EDC.HCl (2 eq) were added respectively into the mixture of acid lipid (2 eq) in dry $CH_2Cl_2$ (5mL mmol$^{-1}$) at 0 ºC and stirred at r.t for 15 mins. At the same time, pyridine (1 eq) was added to the mixture of amine (1 eq) in dry DMF: $CH_2Cl_2$ (1:1, 10 mL mmol$^{-1}$), and then transferred into the mixture of acid lipid at 0 ºC. The resulted mixture was warmed up to r.t and stirred for 6 h. The solvent was

evaporated and the crude material was purified immediately by flash chromatography (MeOH:EtOAc, 5:95) to afford a yellow solid that was purified once more time with reverse column ($H_2O$:MeOH, 20:80) to afford product as the white solid.



**(2*S*,3*R*)-3-Hydroxy-2-((*R*)-3-hydroxyhexadecanoyl)amino-15-methylhexadecanyl-1-thio-β-D-galactopyranoside** (**303**)

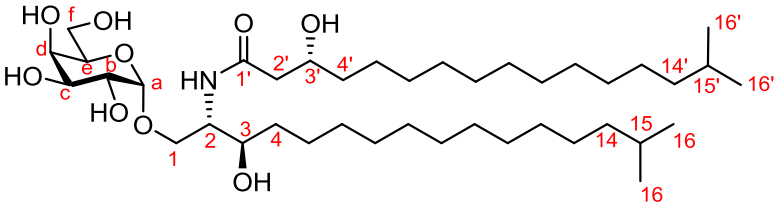
The General Protocol for coupling between the amine **341** (0.020 g, 0.045 mmol) and (*R*)-3-hydroxyhexadecanoic acid **318** (0.0245 g, 0.090 mmol). After purification, the product **303** was isolated as white solid (0.0088 g, 28%). $^1$H NMR (400 Hz, DMSO-$d_6$) δ 7.60-7.62 (1 H, d, *J* = 9.1 Hz, NH), 4.64-4.65 (1 H, d, *J* = 3.3 Hz, a), 4.58-4.60 (1 H, d, *J* = 4.8 Hz, 3'-OH), 4.52-4.53 (1 H, d, *J* = 6.4 Hz, 3-OH), 4.47-4.48 (1 H, d, *J* = 5.1 Hz, d-OH ), 4.49-4.50 (1 H, d, *J* = 5.5 Hz, f-OH), 4.33-4.34 (1 H, d, *J* = 4.0 Hz, c-OH), 4.15-4.16 (1 H, d, *J* = 7.5 Hz, b-OH), 3.41-3.76 (9 H, m, b,c,d,e,f,1,2,3,3'), 2.17-2.24 (1 H, qd, *J* = 13.8, 6.5 Hz, H2'), 1.46-1.52 (3 H, m, H4,H4'), 1.41 (2 H, m, H15,H15'), 1.12-1.19 (36 H, m, lipid chain), 0.83-0.86 (3 H, m, H16'), 0.83-0.85 (3 H, d, *J* = 6.6 Hz, H16). $^{13}$C NMR (100 MHz, DMSO-$d_6$) δ 170.3, 99.5, 71.2, 69.8, 69.7, 69.2, 68.9, 68.7, 67.5, 60.6, 53.0, 44.1, 38.4, 36.6, 33.6, 31.2, 29.3, 29.3, 29.2, 29.1, 29.0, 28.9, 28.7, 28.6, 28.50, 28.3, 27.4, 26.8, 25.1, 22.5, 22.1, 14.0. HRMS (ESI) *m/z* calcd. for $[C_{39}H_{77}NO_9+H]^+$: 704.5671, obsd: 704.5672.



**(2*S*,3*R*)-3-Hydroxy-2-((*R*)-3-hydroxy-15-methylhexadecanoyl)amino-15-methylhexadecanyl-1-thio-β-D-galactopyranoside (*α*-GalCer$_{Bf\text{-}716}$; 304)**

The General Protocol for coupling between the amine **341** (0.020 g, 0.045 mmol) and (*R*)-3-hydroxy-15-methylhexadecanoic acid **316** (0.0258 g, 0.090 mmol). After purification, the product **304** was isolated as white solid (0.022 g, 68%). $^{1}$H NMR (400 Hz, DMSO-d$_6$) δ 7.61-7.63 (1 H, d, *J* = 9.0 Hz, NH), 4.65-4.66 (1 H, d, *J* = 2.5 Hz, a), 4.59-4.60 (1 H, d, *J* = 4.7 Hz, 3'-OH), 4.53-4.54 (1 H, d, *J* = 6.3 Hz, 3-OH), 4.48-4.49 (1 H, d, *J* = 4.9 Hz, d-OH ), 4.50-4.51 (1 H, d, *J* = 5.6 Hz, f-OH), 4.33-4.34 (1 H, d, *J* = 3.9 Hz, c-OH), 4.15-4.17 (1 H, d, *J* = 7.4 Hz, b-OH), 3.41-3.76 (9 H, m, b,c,d,e,f,1,2,3,3'), 2.15-2.24 (1 H, qd, *J* = 13.8, 6.5 Hz, H2'), 1.46-1.52 (3 H, m, H4,H4'), 1.41 (2 H, m, H15,H15'), 1.12-1.19 (36 H, m, lipid chain), 0.83-0.85 (6 H, d, *J* = 6.6 Hz, H16,H16'). $^{13}$C NMR (100 MHz, DMSO-d$_6$) δ 170.5, 99.9, 71.3, 69.8, 69.7, 69.2, 68.9, 68.7, 67.9, 67.5, 60.8, 53.3 44.4, 38.5, 36.6, 34.0, 31.2, 29.3, 29.3, 29.2, 29.1, 29.0, 28.9, 28.7, 28.6, 28.50, 28.3, 27.4, 26.8, 25.1, 22.5, 22.1, 14.0. HRMS (ESI) *m/z* calcd. for $[C_{40}H_{80}NO_9+H]^+$: 718.5828, obsd: 718.5827.

**Table 3.3:** Assignment and comparison between NMR data of purely synthesis **α-GalCer$_{Bf}$** and purified α-GalCer$_{Bf}$ extracted from *B. fragilis* (600 MHz, DMSO).




| Hydrogen | $^1$H NMR data: δ ppm (multiplicity, *J*, integration) | |
|---|---|---|
| | **Natural (Lit)**[111] | **Synthetic** |
| NH | 7. 60 (d, *J* = 9.1 Hz, 1 H) | 7.61-7.63 (d, *J* = 9.0 Hz, 1 H) |
| a | 4.64 (d, *J* = 3.3 Hz, 1 H) | 4.65-4.66 (d, *J* = 3.2 Hz, 1 H) |
| 3'-OH | 4.58 (d, *J* = 4.9 Hz, 1 H) | 4.59-4.60 (d, *J* = 4.7 Hz, 1 H) |
| 3'-OH | 4.52 (d, *J* = 6.4 Hz, 1 H) | 4.53-4.54 (d, *J* = 6.3 Hz, 1 H) |
| d-OH | 4.47 (d, *J* = 5.4 Hz, 1 H) | 4.48-4.57 (d, *J* = 5.6 Hz, 1 H) |
| f-OH | 4.48 (d, *J* = 5.4 Hz, 1 H) | 4.48-4.57 (d, *J* = 4.9 Hz, 1 H) |
| c-OH | 4.33 (d, *J* = 4.2 Hz, 1 H) | 4.33-4.34 (d, *J* = 3.9 Hz, 1 H) |
| b-OH | 4.15 (d, *J* = 7.6 Hz, 1 H) | 4.15-4.17 (d, *J* = 7.4 Hz, 1 H) |
| b, c, d, e, f<br>1, 2, 3<br>3' | 3.79-3.37 (m, 9 H) | 3.41-3.76 (m, 9 H) |
| 2' | 2.17 (ddd, J = 26.3, 13.8, 6.6 Hz, 2 H) | 2.15-2.24 (qd, J = 13.8, 6.5 Hz, 2 H) |
| 4, 4' | 1.38-1.43 | 1.46-1.52 (m, 4 H) |
| 15, 15' | 1.42 (m, 2 H) | 1.41 (m, 2 H) |
| lipid | 1.17-1.24 (m, 36 Hz) | 1.12-1.19 (m, 36 H), |
| 16, 16' | 0.82 (d, *J* = 6.6 Hz, 12 H) | 0.83-0.85 (d, *J* = 6.6 Hz, 12 H) |

# Chapter 4

# Studies toward identification of a novel vitamin B derived T cell antigen

## 4.1 Introduction

### 4.1.1 MR1 presents other ligands distinct from vitamin B2 based derivatives

Le Bourhis *et al.* and Gold *et al.* demonstrated that a wide range of variety of microbes (bacteria and yeasts) but not viruses induce MR1-restricted MAIT cell responses.[120-121] Kjer-Nielsen and co-workers identified three classes of vitamin B-derived compounds that are presented by MR1 and activate MAIT cells to varying degrees, including the pterin 6-FP and the pyrimidines 5-OE-RU and 5-OP-RU.[119] However, several bacteria are known to activate MAIT cells even though their riboflavin synthesis pathways have not been documented in the strains tested. Several examples includes *Bifidobacterium animalis*, *Lactobacillus acidophilus*, *Lactobacillus casei*,[121] *Streptococcus pyogenes*[278] and *Nocardia asteroids*.[278] Similarly, riboflavin synthesis does not occur in humans, therefore riboflavin metabolites presented in the context of MR1 have been suggested to be a bacteria-induced or -derived endogenous antigen,[121-122] that acts as a pathogen-associated molecular pattern (PAMP). This raises the question whether there could be other antigens, which do not arise from riboflavin pathway, but which can activate MAIT cells.

Certain lines of evidence suggest the existence of as yet undescribed MR1 ligands, and structural analysis suggests that plasticity in the MR1-binding groove can accommodate a range of different ligands.[116, 136-138, 279-280] As the pterin, pyrimidine and lumazine rings occur commonly in nature, it is feasible that other microbial or host molecules with related chemotypic properties could bind to MR1 and function as antigens for MR1-restricted T cells.

Keller *et al*. performed a combined *in silico* docking, structural investigation and functional assay to explore whether MR1 could present ligands distinct from riboflavin and folate-based metabolites.[281] Various mono- and bicyclic molecules were identified

that could bind MR1 including 2-hydroxy-5-methoxybenzaldehyde (HMB), salicylates (3-formyl-salicylic acid, 3-F-SA), 2-hydroxy-1-naphthaldehyde (2-OH-1-NA), a photodegradation product of the chemotherapeutic aminopterin (2,4-diamino-6-formylpteridine, 2,4-DA-6-FP) and the anti-inflammatory drug diclofenac (DCF) and one of its metabolites, 5-hydroxy-diclofenac (5-OH-DCF). However, while these compounds bound MR1, most did not stimulate MAIT cells. However, 5-OH-DCF exhibited stimulatory properties, and the 3D structure of the MR1-TCR complex showed a hydrogen bond of 5-OH-DCF with CDR3β loop of the TCR. This suggests that a direct contact between the MAIT TCR and the compound bound within the MR1 cleft is required to activate MAIT cells.

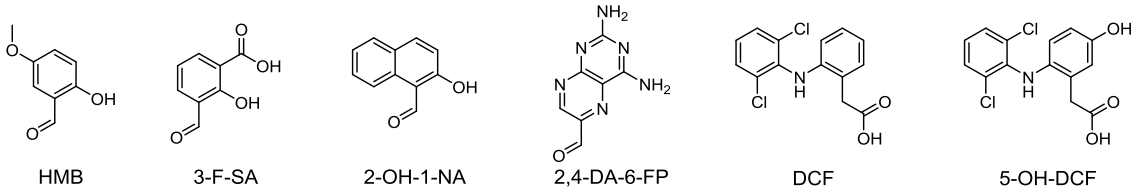



**Figure 4.1:** A series of mono- and bicyclic molecules that bind MR1.

In conclusion, the MR1 binding cleft can present a broad range of chemical scaffolds, beyond those that arise from folic acid and the riboflavin pathway. A central question in the field is what other naturally occurring ligands can bind MR1 and modulate MAIT cell functions?

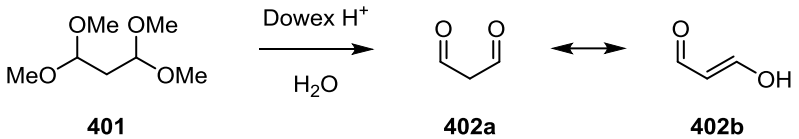
### 4.1.2 Discovery of new antigens by treatment of 5-A-RU with malondialdehyde

This chapter describes work in progress towards the discovery of a new MR1-restricted MAIT cell antigen, in collaboration with the laboratories of Prof. James McCluskey (The University of Melbourne) and Prof. Jamie Rossjohn (Monash University). To understand the origin of MAIT cell activating ligands from microbial riboflavin synthesis,[150] a chemical-genetic approach was used. Bacterial strains deficient in genes of the riboflavin pathway were examined for their ability to generate the

antigen. Ultimately, multiple strands of evidence demonstrated that MAIT cell antigen(s) are formed from 5-A-RU.[136] The currently known MAIT cell antigens, 5-OE-RU and 5-OP-RU, arise from reaction of 5-A-RU with the 1,2-dicarbonyls glyoxal and methylglyoxal, which form double Schiff bases linking MR1 and 5-A-RU.[136] Possibly, other antigens might arise from related reactions.

Malondialdehyde (MDA) is a 1,3-dicarbonyl molecule that arises from peroxidation of lipids and oxidative stress.[282] Like glyoxal, MDA can form covalent adducts with proteins (lysine, arginine) and DNA (deoxyadenosine and deoxyguanosine).[283-284] Possibly, MDA could form a Schiff base with 5-A-RU to yield an adduct that might react with MR1 and be presented to MAIT cells.

MDA is an unstable molecule, but can be purchased as the enolate, tetrabutylammonium malondialdehyde enolate. Alternatively, we developed an approach to synthesize MDA by treatment of tetramethoxypropane (TEP) with ion-exchange resin ($H^+$ form) in water. MDA undergoes facile keto-enol tautomerism, and has been reported to be the *trans* enol in water,[285] ethanol[286] or diethyl ether,[287] and the *cis* enol in hexane,[287] carbon tetrachloride[285] and dichloromethane.[285, 287] We performed NMR analysis of a solution of MDA prepared in the same way, but in $D_2O$, showed it to be the *trans* enol. MDA prepared in $D_2O$ decomposed in 1 h at room temperature. Thus, freshly prepared MDA **402** should be used immediately, by incubation with 5-A-RU, or alternatively can be stored at -80 $^oC$ freezer for several weeks.




**Scheme 4.1:** Preparation of 1,3-dicarbonyl compound, malondialdehyde (MDA).

Jurkat cells transduced with genes encoding semi-invariant MAIT TCRs comprising with either Tyr94α (A7 and A9) or Tyr95α (B1 and E3) cell lines, were

tested for activation by C1R antigen presenting cells expressing MR1 and mixtures of 5-A-RU and methylglyoxal or MDA. In this system, when activating ligands bind MR1 and are presented on the C1R cells, the TCR expressing cells upregulate CD69, which are recognized by anti-CD69 antibodies, and the number of activated cells can be counted by flow-cytometry.

Kjer-Neilsen and co-workers examined the activation of MAIT cells when treated with MR1 and 5-OP-RU and upon co-incubation of 5-A-RU and MDA.[(unpublished)] Both 5-OP-RU and the new antigen activated MAIT cells, but with different potency. Plates coated with MR1 and loaded with 5-OP-RU induced a strong response at high concentration for all MAIT cell lines, namely those expressing Tyr94 and Tyr95 in their α-chains. On the other hand, the new antigen generated from 5-A-RU and MDA selectively induced strong responses to only the A7 and A9 cell lines that express the Tyr94α MAIT TCRs, and a very weak signal through the B1 and E3 Tyr95α MAIT TCRs. The unknown antigen thus represents the first MAIT cell antigen with selective activation of the Tyr94α *versus* Tyr95α MAIT TCRs.

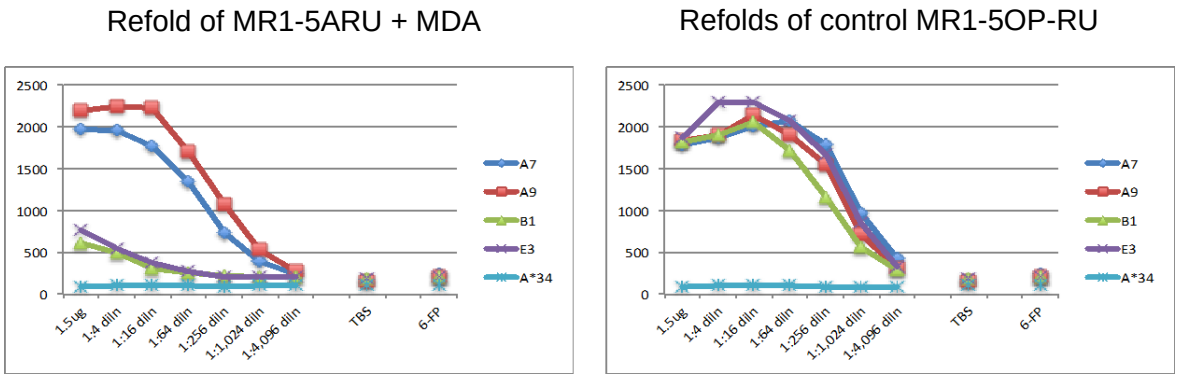



**Figure 4.2:** Cell activation of refolded MR1 with 5-A-RU and MDA (left), and the control refolded MR1 with 5OP-RU (right). A7 and A9 are Tyr 94α MAIT TCRs, B1 and E3 are Tyr95α MAIT TCRs.

In an attempt to define the structure of the unknown ligand, a low-resolution 3D structure was obtained by X-ray crystallography of MR1 folded in the presence of 5-A-

RU and MDA. The electron density at the site of antigen-binding pocket showed a new ligand bound but at a resolution that was insufficient to unambigously define the chemical structure of the antigen. However, it possessed similarities to 5-OP-RU and 5-OE-RU. A planar, aromatic moiety bound deep in the MR1 cleft, and appeared to form a Schiff base with Lys43 of MR1. This aromatic moiety bound in the cleft by interaction with Tyr7 and Tyr62. A ribityl group was unambigously defined that bound in essentially identical positions to that seen for 5-OP-RU and 5-OE-RU within their respective complexes and which formed a hydrogen bond to Tyr95α of the MAIT TCR.[137]

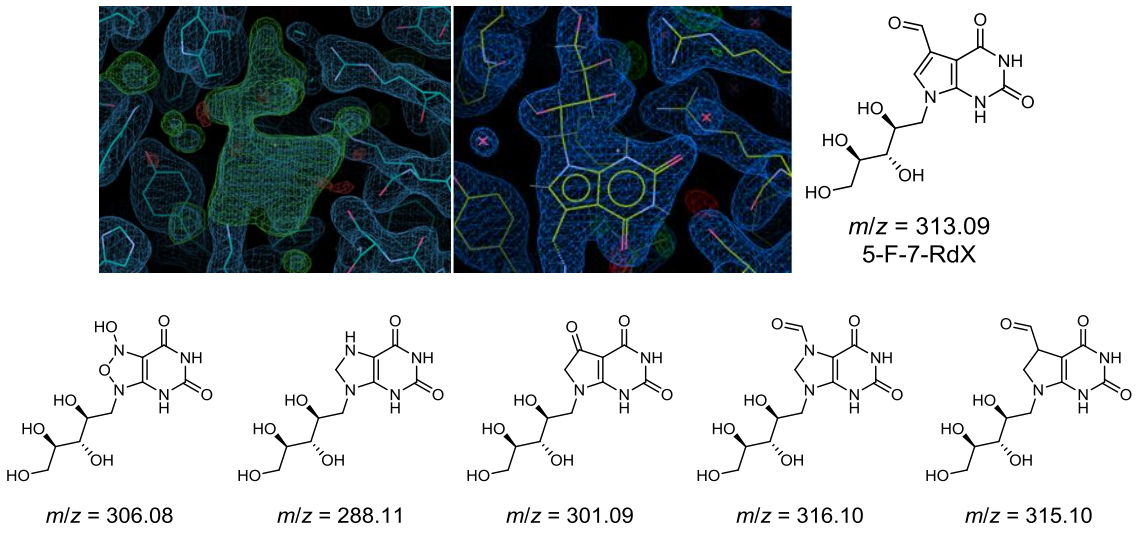


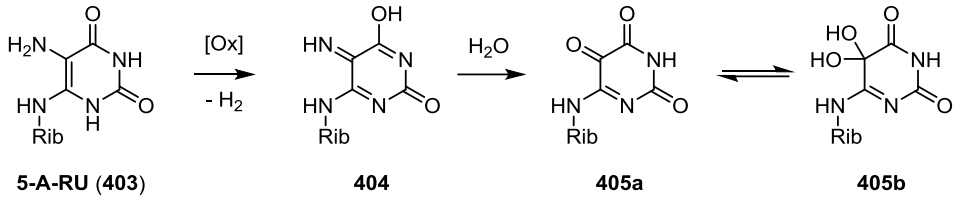

**Figure 4.3:** (Left) Electron density at the site of antigen-binding pocket; (Right) modelling of 5-F-7-RdX into the electron density; (Bottom) candidate structures of the unknown ligand.

Considering the electron density at the site of antigen-binding pocket, we proposed six candidate structures for the novel MR1 antigens. Mass spectrometric analysis of MR1 refolded MR1 in the presence of 5-A-RU and MDA revealed an ion with *m/z* of 314.10 in positive mode as the most abundant molecular mass, which is consistent with just one of these compounds: 5-formyl-7-D-ribityl-5-deazaxanthine (5-

F-7-RdX) (Figure 4.3). Therefore, we proposed that the unknown antigen for MAIT cells is 5-F-7-RdX.

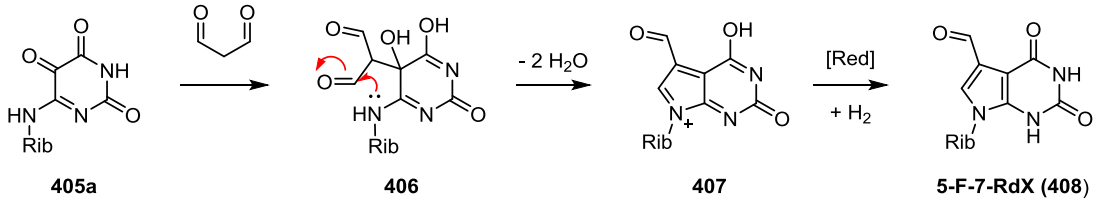
### 4.1.3 Possible mechanism of reaction between 5-A-RU and MDA

5-A-RU is notably unstable, and it decomposes to a 'multitude of products', within seconds to minutes. Lange and co-workers showed that 5-A-RU is oxidized with loss of dihydrogen to generate azaquinone **404** at r.t over two min. Hydrolysis gave intermediate **405a**, which exists as its hydrate **405b**.[288]




**Scheme 4.2:** Oxidative decomposition of 5-A-RU in dilute aqueous solution.[288]

Applying this proposed sequence of events to our system, we predicted that 5-A-RU would be quickly oxidised to azaquinone **404**, followed by hydrolysis to generate intermediate **405a**. At the same time, malondialdehyde could deprotonate to generate an enolate, which could attack to the 1,2-dicarbonyl system of the azaquinone **405** to generate intermediate **406** (possibly, direct nucleophilic attack by MDA enol upon **405** may also occur). Cyclization, followed by dehydration, generates cation **407**. Reduction of **407** restores aromaticity and would deliver 5-F-7-RdX **408**.




**Scheme 4.3:** Proposed formation of 5-F-7-RdX by treatment of 5-A-RU with MDA.

### 4.1.4 Research aims

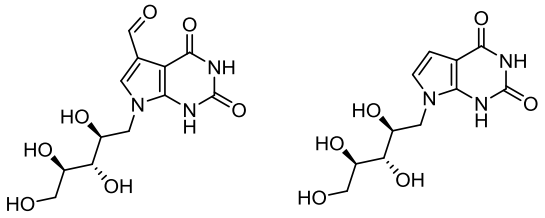
MAIT cells are much more abundant than other types of T cells in humans; however, their roles and antigens leading to their activation are still poorly defined. A

new, structurally undefined selective MAIT cell antigen is formed from the reaction of 5-A-RU with MDA but is available in such small amounts that its structure cannot be conclusively determined. Therefore, the aims for this chapter are to:

**Aim 1:** Develop an approach for the preparation of the des-formyl analogue of 5-F-7-RdX, 7-RdX.

**Aim 2:** Synthesize the proposed structure, 5-F-7-RdX.

**Aim 3:** Structurally characterize the binding of 7-deazaxanthine derivatives to MR1 and investigate their ability to stimulate MAIT cell signaling to determine whether it matches the immune-stimulating properties of the natural material. This aim will be done in collaboration with Prof. James McCluskey at Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity, and Prof. Jamie Rossjohn, Monash University.




5-F-7-RdX (**408**) 7-RdX (**409**)

**Figure 4.4:** Candidate antigen molecule 5-F-7-RdX and the analogue 7-RdX.

## 4.2 Discussion

### 4.2.1 Synthesis strategy

Sella and co-workers developed and optimized an elegant strategy for the synthesis of *N*-alkylated deazaxanthine analogues, with the key step an alkylation to form the C-N bond skeleton.[289-297] Their preferred approached utilized reaction of a nucleobase-anion[291, 298] with a sugar halide in acetonitrile. A mixture of powdered KOH and tris[2-(2-methoxyethoxy)ethyl]amine was used to generate the nucleobase anion.[299] Following this lead, we proposed that 5-F-7-RdX **408** and 7-RdX **409** could be

assembled from 3 fragments: a formyl group, a protected ribose, and a protected 2,4-dihydroxy-pyrrolopyrimidine (Figure 4.5). The sugar substituent could be introduced by an alkylation, and the formyl group could be introduced by Vilsmeier-Haack reaction utilizing an acid chloride such as oxalyl chloride or phosphorus oxychloride and DMF. However, there is still a question of whether the formyl or the *N*-9 substituent should be introduced first, and what protecting groups should be employed, as they must be easily removed in the final step without affecting any other functionality.

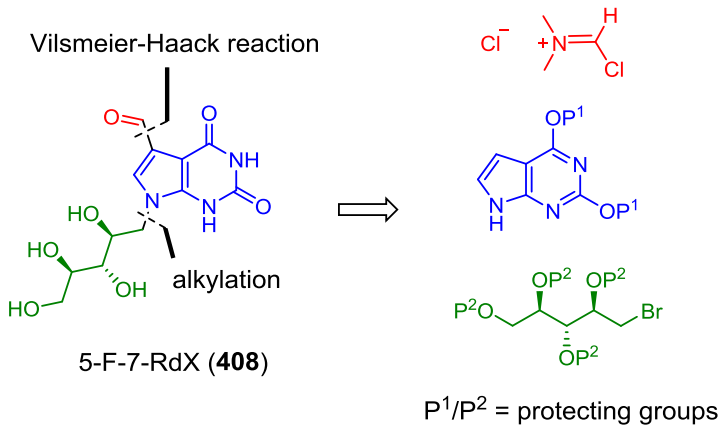
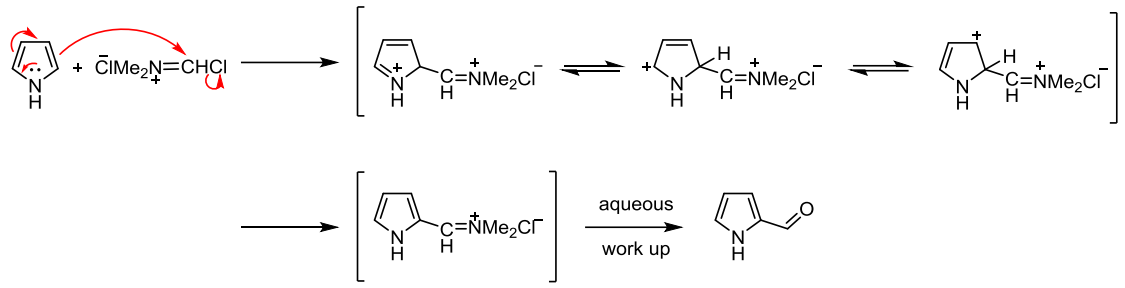



**Figure 4.5:** Restrosynthetic disconnection of 5-formyl-7-D-ribityl-5-deazaxanthine (5-F-7-RdX) **408**.

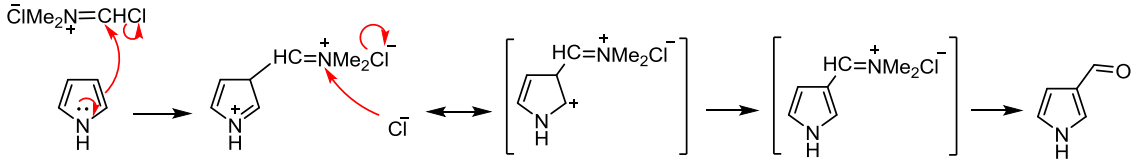
### 4.2.2 Installation and stability of the formyl group

Pyrroles undergo predominant or exclusive kinetic electrophilic substitution at the α (C2) position[300] with most electrophilic reagents.[301-302] Acylpyrroles can be prepared by the direct acylation of the unsubstituted pyrrole with an acid chloride in the presence or absence of a Lewis acid catalyst,[300] or using the Vilsmeier-Haack reagent,[300] nitrilium salts[303] or active esters.[304-305] When the electrophile attacks the α position, it generates a cationic intermediate with three resonance forms that assist in accommodating the positive charge, and therefore substitution at the α (C2) position is preferable (Scheme 4.4).



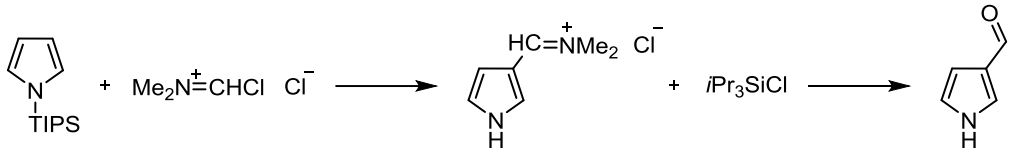

**Scheme 4.4:** Electrophilic substitution reaction at α (C2) position using Vilsmeier-Haack reagent.

β (C3) acyl pyrroles are less readily accessible. Direct acylation affords a cationic intermediate with just 2 reasonable resonance structures (Scheme 4.5). Approaches to access 3-acylpyrroles include the use of a removable deactivating 2-substituent (usually acyl) to direct the entry of an electrophile to the β position, developed by Anderson;[306] the acid-mediated isomerization of the easily prepared α isomers, developed by Muchowski;[307] and the direct substitution of *N*-(phenylsulfonyl)pyrrole with certain electrophiles, described firstly by Anderson[308] and later by Rokach.[309] The first method is intensively used; however, the removal of the 2-substituent is frequently difficult. The second method utilizing acid-mediated isomerization is remarkably general and can be conducted under mild conditions.[310] However, its main drawback is that it delivers an equilibrium mixture of α and β isomers, which limits yields and which may be difficult to separate. The third method is not very general: some electrophilic reagents do not react at all while others give unfavorable α:β product ratios.



**Scheme 4.5:** Electrophilic substitution reaction at β (C3) position using Vilsmeier-Haack reagent.

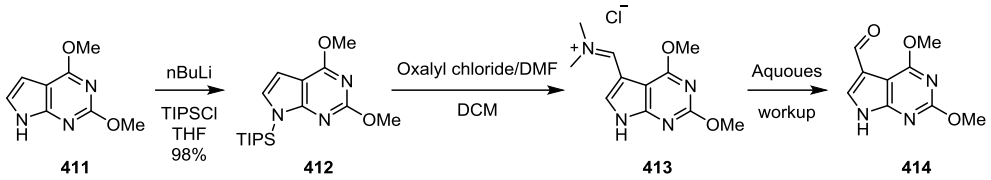
Another approach is to direct electrophilic attack to the β position through the use of a bulky group on nitrogen, such as a *tert*-butyl[311] or a trityl moiety (Scheme 4.6).[312] Even though selectivity for C-3 was observed, the process was of limited value because removal of the nitrogen substituent was impossible or was incompatible with the survival of many functional groups. A more promising alternative is to use the triisopropylsilyl (TIPS) group, which can be easily removed using tetraalkylammonium fluoride.[313-315] A model of the lowest energy conformation of TIPS-pyrrole with van der Waals surface shows that the hydrogen atoms of the isopropyl groups virtually encapsulate the pyrrole α-position and that access of an electrophile is therefore strongly impeded.[316] Bray showed that Vilsmeier-Haack formylation of *N*-(triisopropylsilyl)-pyrrole afforded >96% of the β-substituted iminium salt.[316-317] The high regioselectivity of this reaction and the absence of the silyl moiety in the final product indicated that the rate of formylation is substantially greater than the hydrogen chloride induced desilylation of TIPS group.




**Scheme 4.6:** Vilsmeier-Haack reaction of the *N*-(triisopropylsilyl)-pyrrole provides β-substitution.

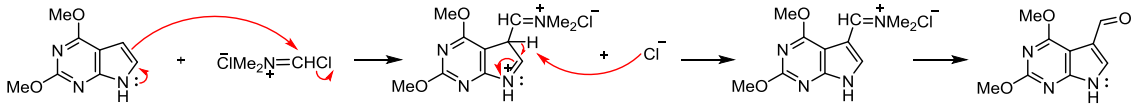
We commenced our work by applying Bray's protocol to the dimethyl pyrrolopyrimidine ring **411**. The TIPS-pyrrolopyrimidine **411** was easily prepared, in high yield, from the lithium salt of pyrrolopyrimidine **411** and TIPS-Cl.[318] Reaction of TIPS-pyrrolopyrimidine **412** with a slight excess of *N*,*N*-dimethylchloroformiminium chloride in $CH_2Cl_2$ at reflux temperature gave the *β*-substituted iminium salt, **413**, which was isolated and its structure confirmed by NMR and mass spectroscopy.

Unfortunately, upon aqueous workup utilizing strong base NaOH and water, **414** could not be isolated.




**Scheme 4.7:** Vilsmeier-Haack reaction of *N*-(triisopropylsilyl)-pyrrolopyrimidine.

Close examination of the structure of 5-F-7-RdX **408** reveals it to be much more interesting than just a 'normal' pyrrole. Compound **408** is a fused pyrrole with substituted pyrimidine ring system on C4 and C5, so that during acylation, only 2 positions at C2 and C3 are available for electrophile attack. Vilsmeier-Haack reaction of **408** should follow the mechanism below and only always give the β (C3) substituted product (Scheme 4.8). This analysis simplified our proposed synthetic approach to 5-F-7-RdX, as the order of formylation should not be important as the β substituted product should be the major product. Therefore, our primary concern was that the formyl group should survive any required reaction conditions.



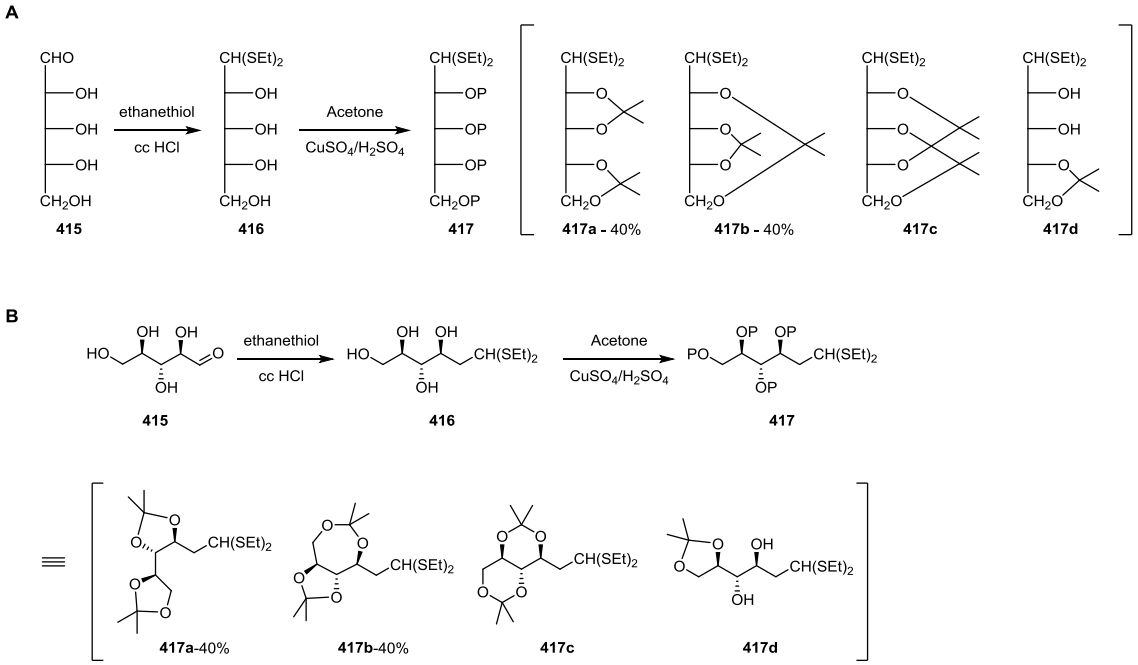

**Scheme 4.8:** Mechanism of Vilsmeier-Haack reaction on the pyrrolopyrimidine ring.

In 5-F-7-RdX, the formyl group is a vinylogous amide with respect to the pyrrole nitrogen. Vinylogous amides are generally much less reactive toward nucleophiles than simple aldehydes as a result of the contribution of the zwitterion canonical form to the resonance hybrid of 3-acylpyrroles.[319] Nonetheless, the decomposition of the iminium salt **413** noted above is worrisome, and therefore to avoid unwanted reactions, we decided to install the formyl group last.

**Scheme 4.9:** Resonance hybrid of 3-acyl pyrrolopyrimidine.

### 4.2.3 Preparation of D-ribityl chain

We first considered methods for the preparation of D-ribityl chain **420**. Several methods have been reported that use commercially available D-ribose and L-ribose. Aslani-Shotorbani synthesized the D-ribityl bromide **420** starting from D-ribose **415** (Scheme 4.10).[320]

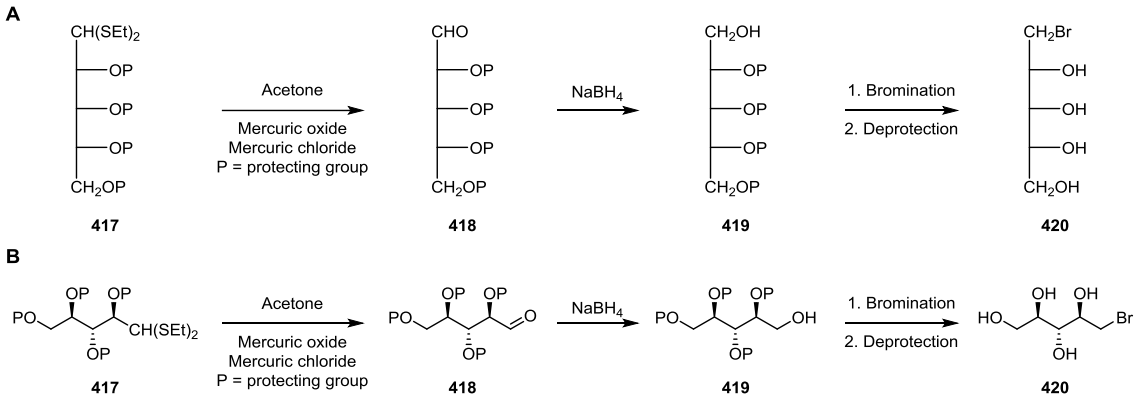



**Scheme 4.10:** Isopropylidenation of **416** gave a mixture of acetals of dimethyl dithioacetal **417a-d**. Transformation and products illustrated using A) Fischer projections and B) standard line drawing.

Firstly, treatment of **415** with ethanethiol in acidic conditions gave D-ribose diethyl dithioacetal **416**.[321] Isopropylidenation[322-325] of **416** has been variously report to give 2,3,4,5-di-*O*-isopropylidene acetal **417a**,[200, 322, 325-326] 2,5:3,4-di-*O*-isopropylidene acetal **417b**,[327] and 2,4:3,5-di-*O*-isopropylidene acetal **417c**.[328] Aslani-Shotorbani

reported that isopropylidenation of **416** with acetone in the presence of sulfuric acid and anhydrous cupric sulfate, gave two major products, **417a** and **417b**, in a 1:1 ratio.[320] An interesting note that if the isopropylidenation performed without sulfuric acids, the 4,5-*O*-isopropylidene monoacetal **417d** is the major product.

Thioacetal hydrolysis of **417** in aqueous acetone gave the aldehyde **418**, which underwent reduction with $NaBH_4$ to give the alcohol **419**. Bromination utilizing tetrabromomethane and triphenylphosphine in $CH_2Cl_2$ and deprotection gave the D-ribityl bromide **420**. Overall, this 5-step route is complex and the conflicting reports on how to conduct the isopropylidenation and the nature of the products discouraged us from pursuing this approach.

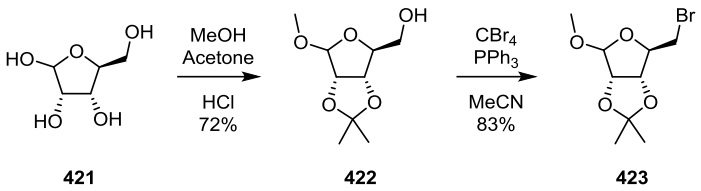



**Scheme 4.11:** Final steps of preparation of D-ribityl bromide **420** using A) Fischer projections and B) standard line drawing.

However, we noticed that D- and L-ribose are enantiomers. They differ only in the groups at C1 and C5, with C2, C3 and C4 being unchanged. Therefore, the conversion of D-ribose into L-ribose requires only a head-to-tail switch by interconversion of the two end groups: reduction of the aldehyde at C1 and oxidation of hydroxyl group at C5 to an aldehyde.

Thus, we proposed to prepare D-ribityl bromide using L-ribose as starting material and proceeding through a head-to-tail swap. Our synthetic route commenced by

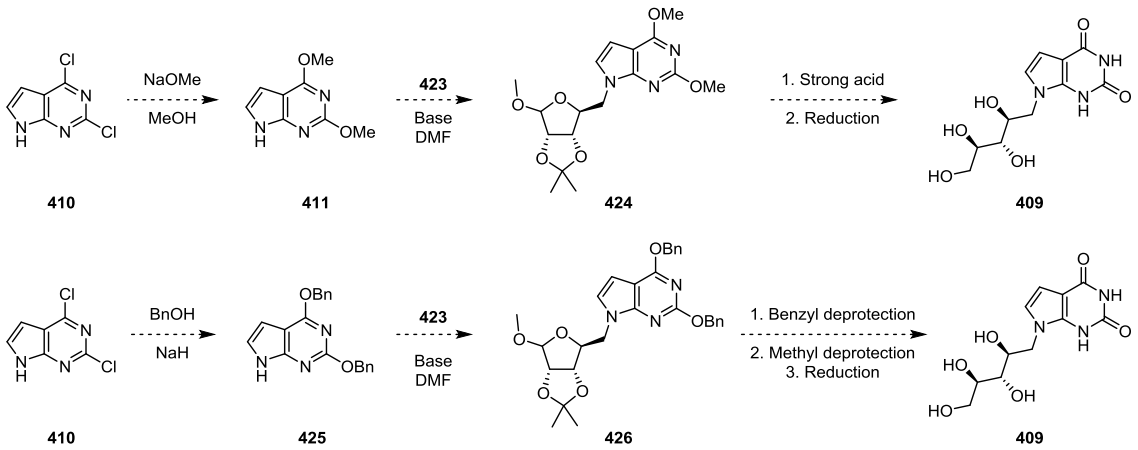
treatment of L-ribose with a 1:1 mixture of methanol and acetone in the presence of concentrated HCl, which gave methyl 2,3-*O*-isopropylidene-*β*-L-ribofuranoside **422**.[300, 329] Bromination of the primary alcohol with $CBr_4$ and $PPh_3$ afforded the bromide **423** in good yield (Scheme 4.12).[330] This compound is poised for reaction with a suitable pyrrolopyrimidine nucleophile.




**Scheme 4.12:** Synthesis of fully protected bromo-L-ribofuranoside **423**.

### 4.2.4 Proposal for preparation of 7-D-ribityl-deazaxanthine (409)

After a survey of possible starting materials, we elected to utilize commercially available 2,4-dichloropyrrolopyrimidine for synthesis of the deazaxanthine nucleus of 5-F-7-RdX **408** and 7-RdX **409**. We proposed the synthesis of 7-D-ribityl-deazaxanthine **409** in a four-step or five-step sequence utilizing two different protecting groups: methyl (Me) and benzyl (Bn), to systematically evaluate the reactivity profile of the protecting groups. In the case of methyl groups on the deazaxanthine nucleus, our proposed route would require four steps. Nucleophilic substitution of 2,4-dichloro-pyrrolopyrimidine **410** with sodium methoxide in methanol would give 2,4-dimethyl-pyrrolopyrimidine **411**, followed by alkylation with bromo L-ribofuranoside **423** would give the adduct **424**. Deprotection could in theory be carried out using strong acid to cleave all methyl groups, and finally reduction with sodium borohydride would give the 7-D-ribityl-deazaxanthine **409**.

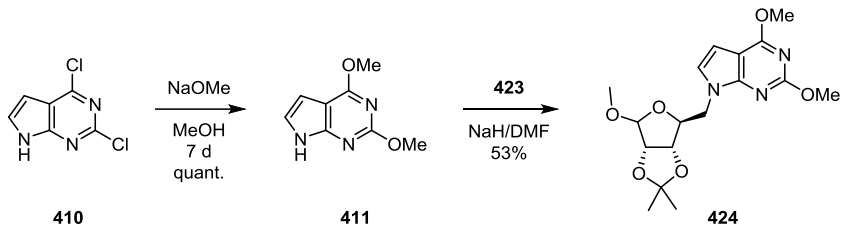



**Scheme 4.13:** Two strategies for synthesizing 7-D-ribityl-deazaxanthine **409** utilizing methyl (top) and benzyl (bottom) protecting groups.

Alternatively, an approach involving benzyl groups could start by nucleophilic aromatic substitution of **410** with sodium hydride and benzyl alcohol to give 2,4-dibenzyl-pyrrolopyrimidine **425**.[331] Alkylation with 5-bromo-L-ribofuranoside **423** would afford adduct **426**. This synthetic route requires stepwise deprotection, with the benzyl group removed first by hydrogenolysis, followed by cleavage of the methyl glycoside and isopropylidene group by treatment with strong acid. Finally, reduction using sodium borohydride would give 7-D-ribityl-deazaxanthine **409** (Scheme 4.13).

### 4.2.5 Preparation of 7-D-ribityl-deazaxanthine using methyl protecting groups

Commercially available 2,4-dichloro-pyrrolopyrimidine **410** was heated under reflux with sodium methoxide in methanol for 16 h to give a monomethoxy-pyrrolopyrimidine as the major product. More extended heating for 7 days eventually afforded the dimethoxy pyrrolopyrimidine **411** in quantitative yield.[332] Alkylation at nitrogen was investigated under a range of conditions (Table 4.1). We explored a range of bases, drawing our inspiration from the literature.[333-334]




410 411 424

**Table 4.1:** Optimization for *N*-alkylation of the 2,4-dimethoxy-pyrrolopyrimidine **411**.

| Entry | Base/Reagent | Solvent | Temp (º C) | Time (h) | Yield (%) |
|---|---|---|---|---|---|
| 1 | KOH/18-crown-6 | MeCN | rt | 24 | < 5 |
| 2 | $K_2CO_3$ | DMF | rt | 24 | < 5 |
| 3 | NaOH | DMF | rt | 72 | 21 |
| 4 | NaOH | DMF | 45 | 72 | 60 (10) |
| 5 | NaH | DMF | 45 | 72 | 61 (53) |
| 6 | NaOH | Dioxane | 45 | 72 | 0 |
| 7 | NaH | THF | 45 | 72 | 10 |

The first attempt applied the method developed by Seela, Westermann and Bindig.[298] The alkylation used protected 5-bromo-L-ribofuranoside **423** and employed MeCN as solvent and powdered KOH/18-crown-6 to generate the nucleobase anion, but less than 5% of the product **424** was isolated (entry 1). Gunderson developed a method for *N*-alkylation using $K_2CO_3$ as a base.[335] However this approach yielded at most traces of **424** (entry 2). When a strong base, NaOH, was used the yield improved to 21% after 3 days (entry 3). Heating the reaction at 45 ºC significantly improved the yield of the reaction (entry 4). However when this protocol was applied on a larger scale, the yield dropped to 10%. NaH was effective on both small (61%) and large (53%) scales (entry 5). Replacing the solvent with dioxane or THF reduced the yield of the reaction (entry 6, 7). In conclusion, the highest yield for this alkylation was achieved using NaH in DMF heated at 45 ºC for 3 d.

As part of this work we investigated running this reaction at >45 ºC but noted the formation of several by-products: enol ether **427**, *N*-methylpyrrolopyrimidine **428**,

bromide **429**, and an undefined dimer, detected by mass spectrometry (Figure 4.6). Enol ether **427** is proposed to arise from base mediated elimination of the 5-bromo-L-ribofuranoside. *N*-Methylpyrrolopyrimidine **428** is proposed to form by nucleophilic attack of the nucleobase anion on another 2,4-dimethylpyrrolopyrimidine molecule. The third by-product, bromopyrrolopyrimidine **429**, with *m*/*z* of 414.0659 in positive ion mode, is believed to arise from the substitution reaction of the alkylated product **424** by bromide liberated in the alkylation reaction (the regioisomer is also a possible contender for its structure). The final by-product was formed in trace amounts and was 186 *m*/*z* units greater than the third by-product **429**. We could not conclusively propose a structure for this product or a rationale for its formation, and simply propose structure **430** to highlight the structural elements that could assemble to generate a product of the appropriate mass. The trace amounts of this compound did not allow useful NMR data to be acquired to better define its structure.

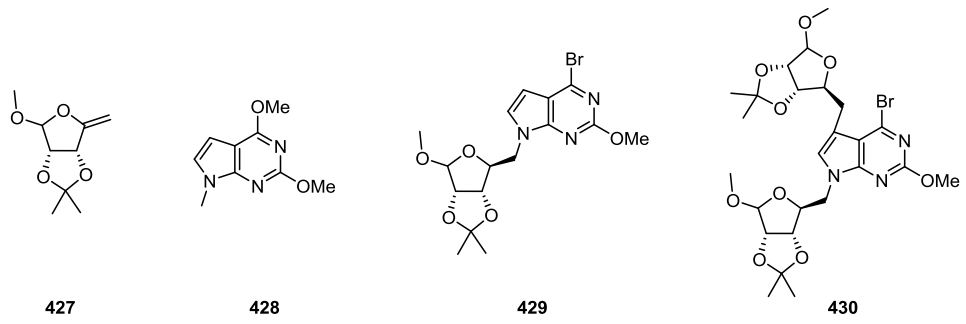


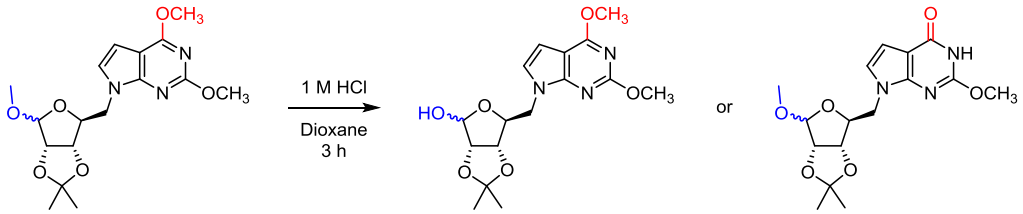

**Figure 4.6:** Several by-products identified in the reaction of **411** and **423**.

### 4.2.6 Removal of protecting groups utilizing strong acid

We next planned to remove all the protecting groups, followed by reduction of the hemiacetal, to afford 7-RdX **409**. Our initial attempts focused on the use of strong acid to cleave the methyl glycoside and methyl imidates. Reactions were monitored by mass spectroscopy as the products did not stain effectively with TLC staining solutions.

The first attempt utilized 1M HCl in dioxane at r.t. After 3 h, two abundant molecular ions could be detected by high resolution mass spectroscopy in the positive

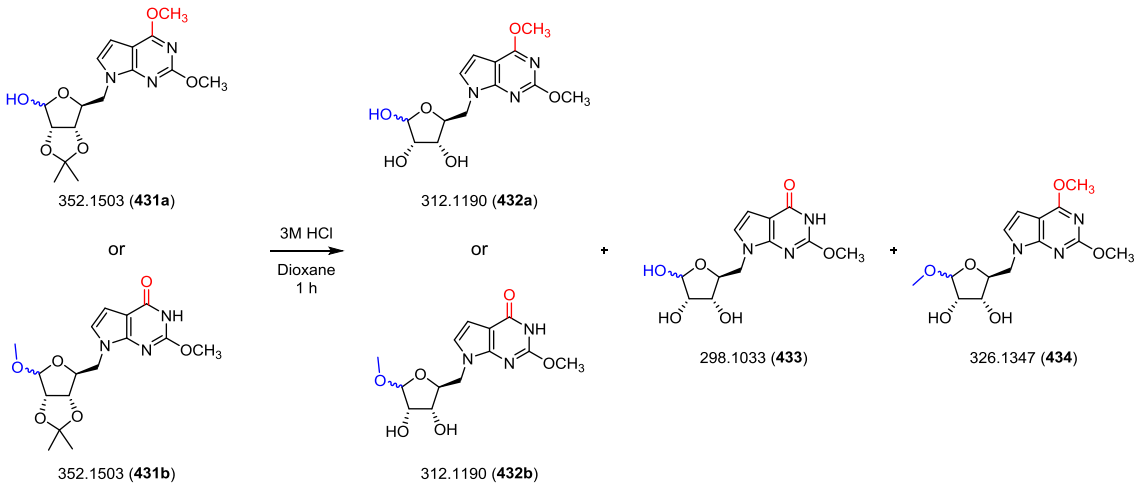
mode: *m*/*z* of 366.1660 and 352.1503. The species contributing to the ion at *m*/*z* of 366.1660 is the starting material. The ion of *m*/z of 352.1503, which was 14 *m*/*z* units smaller than the starting material, is consistent with the molecular mass of loss of one methyl from the starting material; however, the site of loss could not be determined by mass spectrometry alone. Two possible positions for this methyl groups are either at the hemiacetal position to generate **431a**, or at the C4-position on the pyrimidine ring to create **431b**. This position is known to be more labile than C2 (Scheme 4.14).[336]




366.1660 (**424**) 352.1503 (**431a**) 352.1503 (**431b**)

**Scheme 4.14:** Two proposed structures for the molecular ions *m*/z of 352.1503.

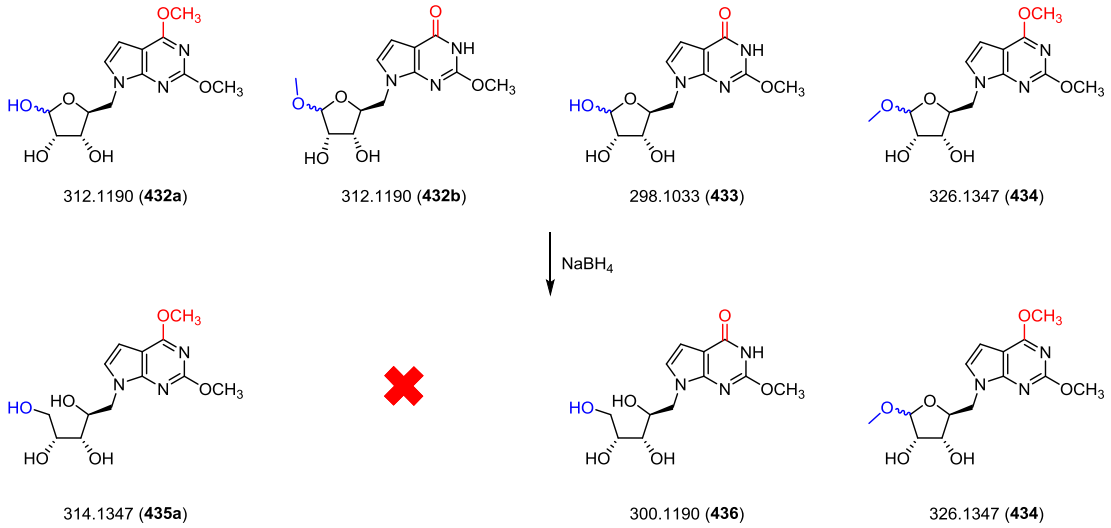
3M HCl was added to drive the reaction to completion. One hour after addition, mass spectrometry revealed an abundant molecular ion of *m*/*z* 312.1190, which is consistent with loss of one methyl group, and the isopropylidene group. Therefore, we predict *m*/*z* 312.1190 would have the structure of **432a** or **432b**. Two other minor ions included *m*/*z* 298.1033, which has lost two methyl groups and an isopropylidene group, possibly with structure **433**, and *m*/*z* of 316.1344, which lost only the isopropylidene group, and which must have structure **434** (Scheme 4.15). The mixture of **432a**/**b**, **433** and **434** could not be separated; therefore, none of the structures was confirmed. While these promising results showed partial progress towards our objective, additional of more concentrated HCl or a longer reaction time did not result in further progression of the reaction; rather decomposition occurred.




**Scheme 4.15:** Proposed structures of three abundant molecular ions after 3M HCl was added.

Even though we could not purify the material to identify structure of products, we next attempted reduction of the mixture to establish whether reduction of the ribose to a ribitol was feasible. Treatment of the mixture with $NaBH_4$ gave a new mixture that was analyzed by mass spectrometry. After the reduction, the ion *m*/*z* 326.1347 was unchanged, and the signals of *m*/*z* 312.1190 and 298.1033 disappeared and were replaced by new ions of approximately 2 mass units higher, namely *m*/*z* 314.1347 and *m*/*z* 300.1190. This result was very promising as this suggested that signal of *m*/*z* 312.1190 corresponds to the **432a** structure (and not **432b**), and that the signal of *m*/*z* 352.1503 corresponds to structure **431a** and not **431b** (Scheme 4.16).

These results provided solid support for our general strategy, even though we could not continue with a mixture of materials. Additional attempts using 3M HCl after 3 h gave a mixture of 4 compounds **431a**, **432a**, **433** and **434**, which was no different from the trial with 1 M HCl. Longer reaction times or utilizing stronger acid such as HBr in AcOH and HBr in $H_2O$ resulted in extensive decomposition.




**Scheme 4.16:** Mass spectrometry of the products formed by borohydride reduction of the partially hydrolyzed mixture.

### 4.2.7 Step-wise deprotection approach

Based on the failure to achieve a global deprotection under acidic conditions, we next considered a step-wise deprotection approach. We aimed to remove the methyl protecting groups on the pyrimidine ring first, and then remove the methyl protecting groups on the sugar. Various protocols have been described for the conversion of a methoxy group to a carbonyl group in nucleosides, including acidic cleavage, HBr in AcOH,[337] TMSCl/NaI in MeCN,[338] and NaOH reflux.[338] As our efforts to use acidic cleavage failed, we decided to investigate basic conditions to cleave the first methoxy group at the C4 position, followed by treatment with TMSCl/NaI to cleave the methoxy group at C2.

The dimethoxy compound **423** was converted into the corresponding monomethoxy derivative **437** by treatment with strong base (Table 4.2). The first attempt using powdered NaOH in $H_2O$ at reflux after 3 d gave only trace amounts of **437** (entry 1). Next, dioxane was explored to improve the solubility of **423**, however, the yield did not improve (entry 2). When the reaction was performed in pyridine with powdered KOH, the yield leaped to 87% of product **437** (entry 3).

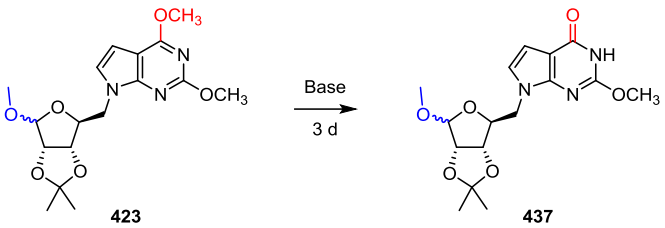



**Table 4.2:** Optimization for the cleavage of methoxy group under basic condition.

| Entry | Base | Solvent | Temp ($^{O}$C) | Yield |
|---|---|---|---|---|
| 1 | NaOH | $H_2O$ | 100 | <5 |
| 2 | NaOH | dioxane | 100 | <5 |
| 3 | KOH | pyridine | 110 | 87 |

The conversion of compound **437** to the 7-L-ribopyranosyl-deazanxanthine **438** was studied using TMSCl and NaI (as a source of TMSI). Using 2 equiv of TMSCl and NaI (as a source of TMSI) at r.t after 4 h gave 60% of **438** and 10% of the deisopropylidenated by-product **439**. When more concentrated conditions were used, 60% of **439** was isolated as the major product with 15% of **438** unchanged (entry 3). As these results showed that TMSCl and NaI could be used to remove both the methyl and isopropylidene groups, we predicted that a higher concentration or longer reaction time could remove all the remaining protecting groups. However, when the reaction was run for 24 h, all material decomposed (entry 3).[338] The final attempt employed double the amount of all the reagents while running the reaction for a very short time; unfortunately this also caused extensive decomposition and no product could be isolated. Ultimately, entry 1, yielding compound **438**, was preferred as the reaction is cleaner and easier to purify.[299]

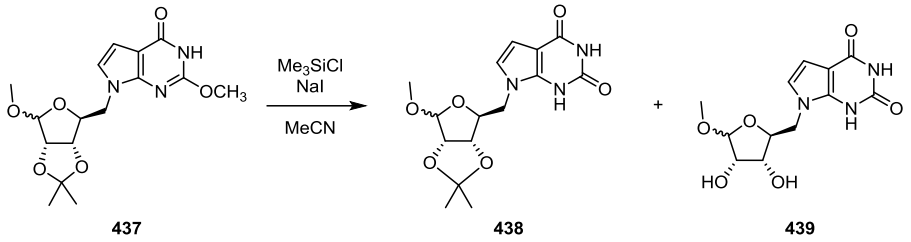
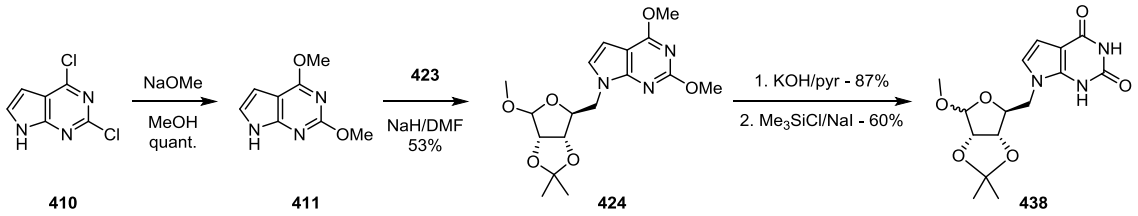



**Table 4.3:** Removal of the 2nd methoxy protecting group on the pyrollopyrimidine ring.

| Entry | Equiv of 437 | Equiv of TMSCl | Equiv of NaI | Time (h) | Yield of 438 (%) | Yield of 439 (%) |
|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 2 | 4 | 60 | 10 |
| 2 | 1 | 6 | 5 | 4 | 15 | 60 |
| 3 | 1 | 6 | 5 | 24 | 0 | 0 |
| 4 | 1 | 11 | 10 | 1 | 0 | 0 |

To summarize the progress thus far, we had at this stage developed a four-step approach using methyl protecting groups that allowed the preparation of ribose-protected deazaxanthine **438.** This synthetic route requires a total of 20 days for preparation and purification with overall yield of 28%.

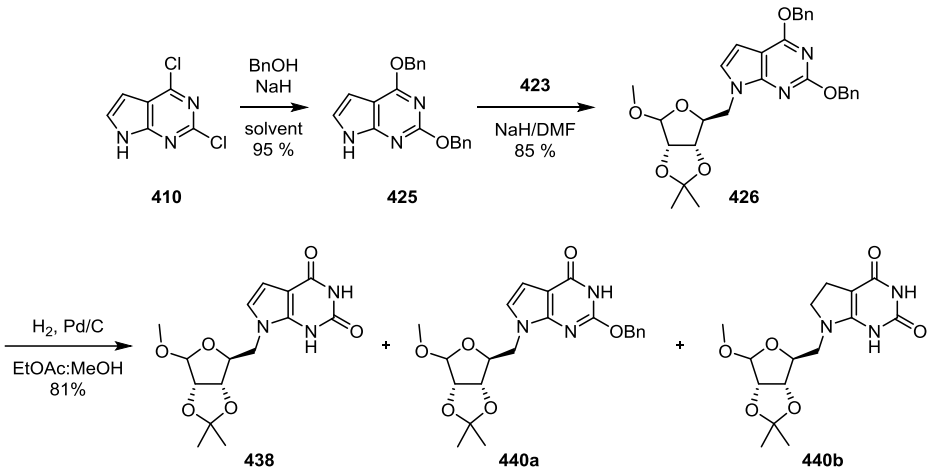



**Scheme 4.17:** Overall synthetic route for preparation of ribose-protected **438**.

### 4.2.8 Preparation of 7-L-ribopyranosyl-deazanxanthine using benzyl protecting groups

Although reasonable results had eventually been obtained using methyl protection on the pyrrolopyrimidine, we sought to investigate whether benzyl protecting groups may provide an advantage. Substitution of 2,4-dichloro-pyrrolopyrimidine **410** with sodium hydride and benzyl alcohol gave the 2,4-dibenzyl-pyrrolopyrimidine **425** in good yield after 16 h.[339] This reaction occurred much faster than the 7-day reflux

required for methoxide substitution. The previously described alkylation conditions proved effective for alkylation of 2,4-dibenzyl-pyrrolopyrimidine **425** by the protected bromo-L-ribofuranoside **423**, giving **426** in 89% yield after 3 d. Hydrogenolysis of the benzyl ethers using hydrogen at high pressure and Pd/C as catalyst for 1 h gave 81% yield of ribose-protected deazaxanthine **438**, along with a small amount of mono-benzyl protected product **440a** and the hydrogenated product **440b**.




**Scheme 4.18:** Preparation of ribose-protected deazaxanthine **438**.

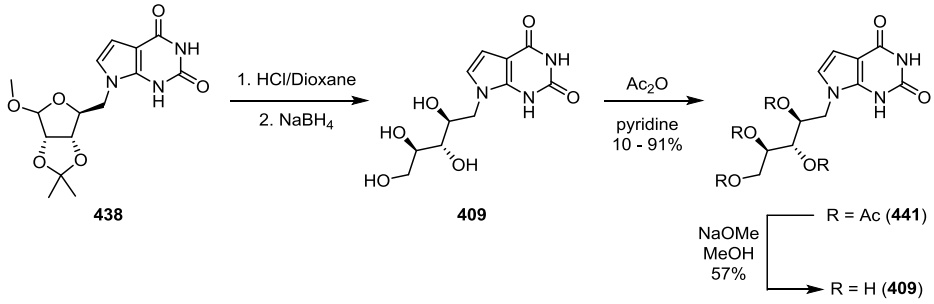
In summary, this new approach affords ribose-protected deazanxanthine **438** in three steps and requires just 7 d for preparation and purification, with an overall yield of 65%. This is clearly advantageous compared to the four-step method using methyl group protection of the pyrrolopyrimidine, which gave 28% yield of **438** and required 20 d.

### 4.2.9 Preparation of 7-L-ribit-5-yl deazaxanthine (7-RdX)

With ribose-protected deazaxanthine **438** in hand, we then tried to remove the protecting groups on the sugar. Treatment of **438** with 3 M HCl at 60 ºC for 1 h resulted in the unprotected L-ribose deazaxanthine, followed by reduction using $NaBH_4$ afforded 7-L-ribit-5-yl-deazaxanthine (7-RdX) **409**. Unfortunately, due to its high polarity and poor solubility, pure material could not be isolated at this stage. To overcome these

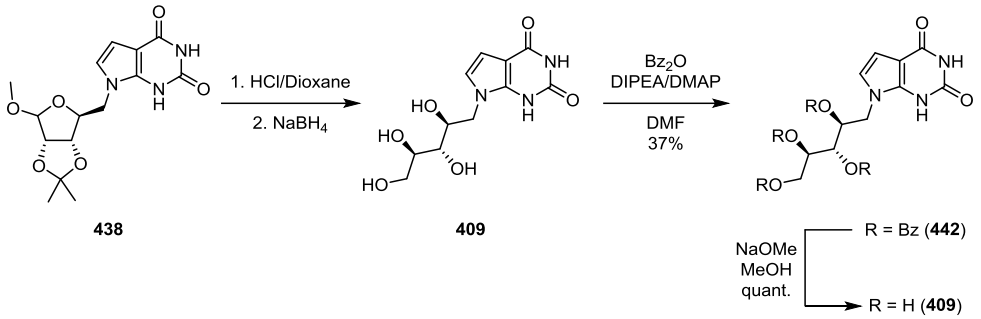
problems, we proposed the temporary introduction of a protecting group, to introduce onto the D-ribityl chain, and to this end explored acetyl and benzoyl protecting groups.

To install the acetyl protecting groups, immediately after reduction, crude 7-RdX **409** was treated with acetic anhydride in pyridine to afford the tetra-*O*-acetyl-D-ribityl-deazaxanthine **441** (Scheme 4.19). A small-scale trial gave **441** in 91% yield. However, on a larger scale, the yield dropped to 10%. Purified tetraacetyl-D-ribityl deazaxanthine **441** was treated with mixture of NaOMe in MeOH, and then cation exchange resin ($H^+$ form) to afford 7-RdX **409** in 57% yield without any purification.




**Scheme 4.19:** Preparation of 7-L-ribityl deazaxanthine (7-RdX) using acetyl protecting groups.

To investigate the use of benzoyl groups, crude 7-RdX **409** was treated with benzoyl chloride, DMAP and pyridine, but no product could be isolated. Better results were obtained with benzoic anhydride, DIPEA and DMAP in DMF, which gave 35% of the tetrabenzoate **442** over 3 steps, with consistent yields on a larger scale. Purified **442** was treated with NaOMe in methanol, followed by treatment with acid resin and extraction of methyl benzoate with methanol, to give 7-RdX **409** without further purification. Even though the yield for benzoylation was low, the benzoyl approach was preferable. The overall yield was more reproducibile and purification was more readily achieved.




**Scheme 4.20:** Preparation of 7-L-ribityl deazaxanthine (7-RdX) using benzoyl protecting groups.

### 4.2.10 Immunological study of 7-RdX

The ability of 7-RdX to induce MAIT cells signaling through MR1 was studied in the laboratory of Prof. James McCluskey (Department of Microbiology and Immunology, Peter Doherty Institute for Infection and Immunity). In his laboratory, Lars Kjer-Nelson and co-workers have engineered a CD69 gMFI cell line that is activated by signaling through MAIT TCR recognition of ligands presented by MR1 expressed by the antigen presenting cell C1R. Four lines of MAIT TCR expressing cells have been tested, including A7, A9 cell line (which possess the Tyr94α in the TCR α-chain) and B1, E3 (which possess Tyr 95α) cell line. Four different ligands were tested, including 6-FP (an antagonist of MAIT cells), 5-OP-RU (a powerful agonist of MAIT cell), 7-RdX **409** and the MR1-RF-DX (MR1 presenting cell with a mixture of 5-A-RU and MDA).

The antagonist 6-FP did not cause signaling through MAIT cells. 5-OP-RU induced a very strong signal at low concentrations. When this protocol was applied to the mixture of MR1 with MDA and 5-A-RU, the unknown ligand selectively activated MAIT TCR-Tyr94α expressing cells, but not those expressing MAIT TCR-Tyr95α, and it induced a similar level of activation as 5-OP-RU but at a 4-fold higher concentration. Synthetic 7-RdX induced a very weak signal, requiring $10^4$ higher concentrations and

was a selective agonist for Tyr94α MAIT cells like the unknown ligand. It has previously been demonstrated that a formyl group (for ligation through a Schiff base to MR1) is an essential structural element for potent signaling through MAIT cells, thus the absence of formyl group in 7-RdX could explain the weak response of 7-RdX.

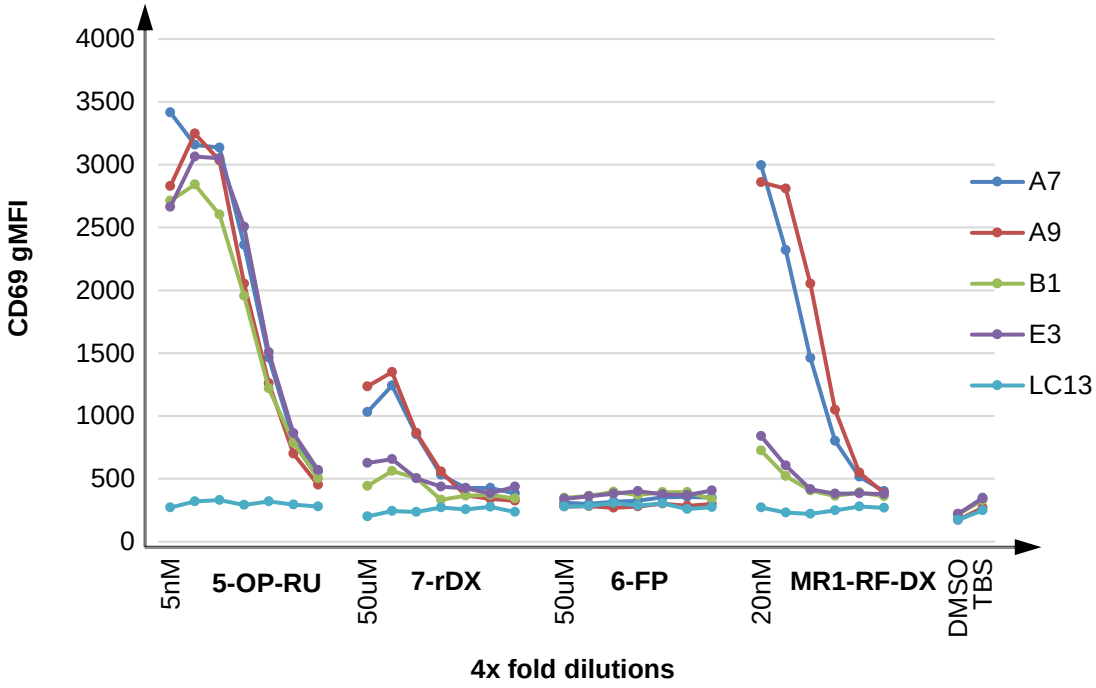



**Figure 4.7:** 7-RdX **409** preferentially activated Tyr94α-MAIT cells at high concentration, with selectivity similar to a mixture of MDA and 5-A-RU.

### 4.2.11 Crystal structure of synthetic 7-RdX

The 3D structure of synthetic material 7-RdX in a complex with antigen presenting molecule MR1 was determined by X-ray crystallography in Prof Jamie Rossjohn's laboratory (Figure 4.8). The electron density at the site of the antigen-binding pocket is almost identical with the electron density of the unknown ligand. The only difference between them is the absence of (a putative) Schiff base with Lys 43 of MR1, as the formyl group is absent in 7-RdX. However, this crystal structure supports our hypothesis that 5-F-7-RdX is the true structure of the antigen generated by reaction of 5-A-RU and MDA.

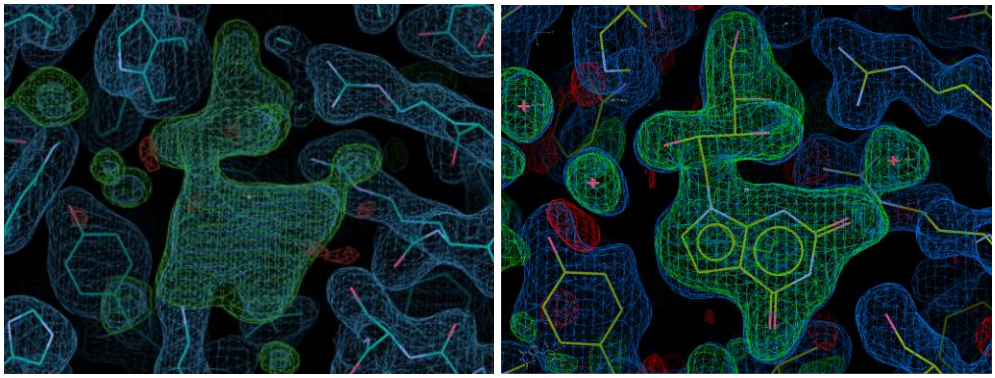


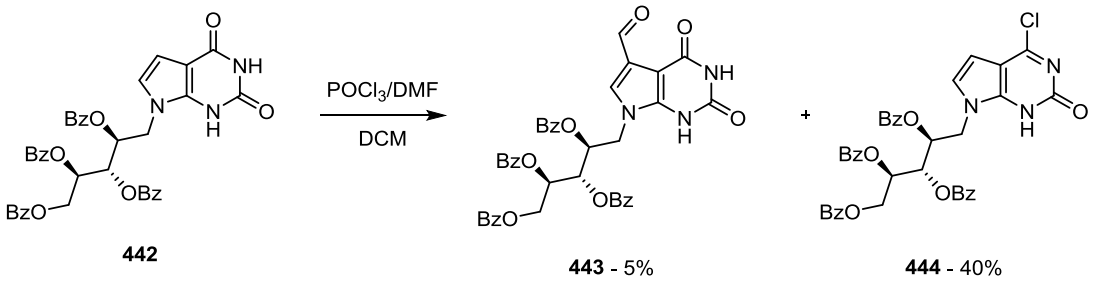
**Figure 4.8:** (Left) Electron density at the site of antigen-binding pocket of MR1-unknown antigen-MAIT TCRs; (Right) crystal structure of synthetic 7-RdX bound to MR1.

### 4.2.12 Attempted preparation of 5-formyl-7-L-ribityl deazaxanthine (408)

With 7-RdX successfully synthesized, we next attempted the Vilsmeier-Haack reaction to introduce the formyl group. The Vilsmeier-Haack reaction uses DMF, and an acid chloride such as oxalyl chloride or phosphorus oxychloride to convert an electron rich aromatic ring to an aryl aldehyde. Reaction of DMF with the acid chloride forms an iminium salt, known as the Vilsmeier reagent. This is an electrophile that reacts with aromatic rings in an electrophilic aromatic substitution to form an iminium intermediate. Aqueous work-up can hydrolyze the iminium group, leading to the aryl aldehyde.[340] Inert, non-polar solvents like $CH_2Cl_2$ are used for the Vilsmeier-Haack reaction, which often causes the iminium salt to precipitate, allowing its isolation by filtration. Direct hydrolysis of the isolated iminium salt can deliver the aryl aldehyde without requiring purification.

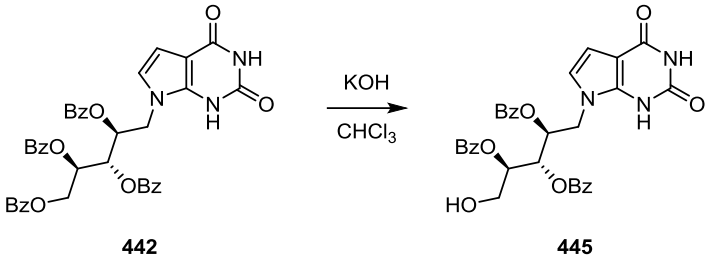
The first attempt using tetrabenzoate **442** with oxalyl chloride and DMF did not yield any iminium salt. When phosphorus oxychloride was used, only a trace amount of iminium salt was detected by mass spectroscopy.[341] The major product isolated was the chloride **444**, which arose through reaction of an aminocarbonyl function in the pyrimidinone ring. In retrospect this was unsurprising as Vilsmeier reagents have been

applied for the preparation of chloropyrimidines from the corresponding pyrimidinones.[342]




**Scheme 4.21:** Preparation of the second analogue 5-F-7-RdX.

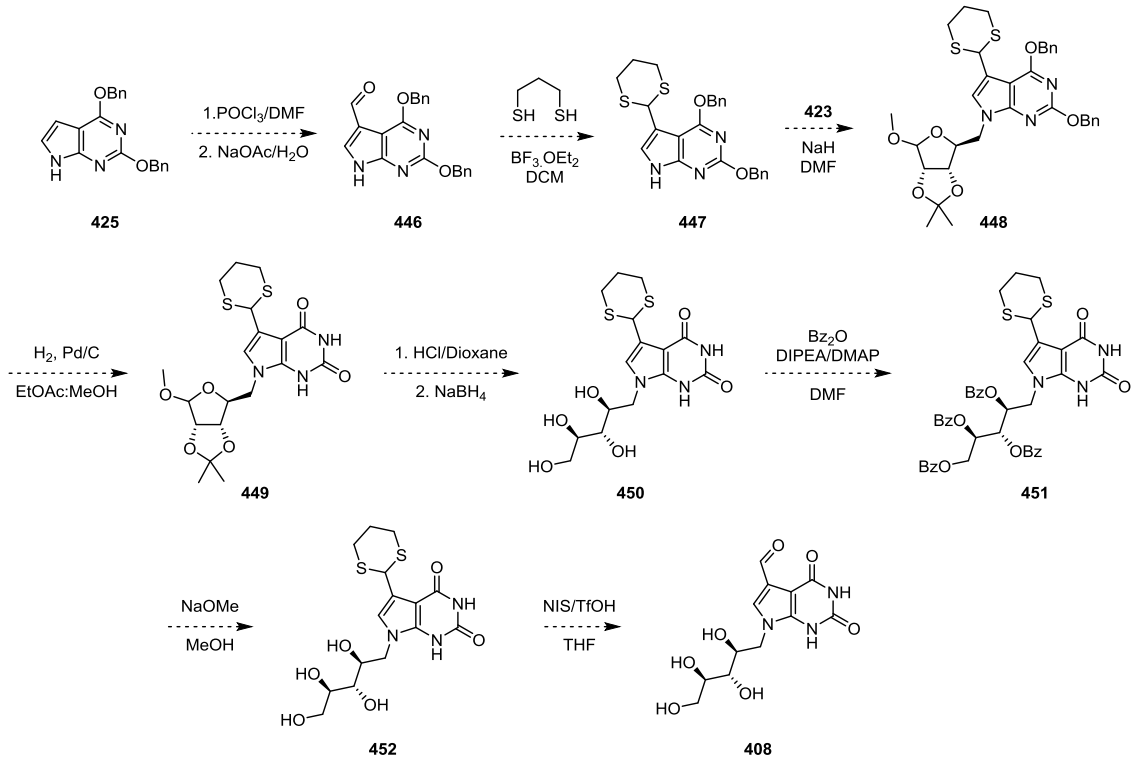
Before abandoning this approach, we also attempted the Reimer-Tiemann reaction, which uses a strong base such as KOH in chloroform.[343] Unfortunately, instead of formylation of the pyrrolopyrimidine ring of **442**, the strong base cleaved the benzoate protecting group, to give **445**.




**Scheme 4.22:** Major product isolated from the Reimer-Tiemann reaction.

### 4.2.13 Redesign of the approach to prepare 5-F-7-RdX (408)

The failure to install a formyl group after assembly of the ribityl-deazaxanthine structure prompted us to reconsider our approach to 5-F-7-RdX, and perform the formylation prior to cleavage of the alkoxy groups on the pyrimidine ring. However, because reduction of the ribose to the ribityl group was incompatible with the presence of a formyl group, we needed to reconsider the use of protecting groups.

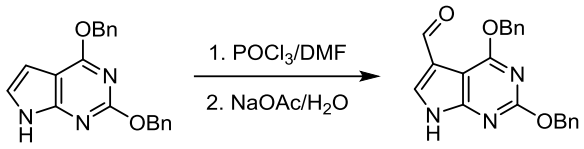



**Scheme 4.23:** Proposed synthesis of 5-formyl-7-D-ribityl-5-deazaxanthine (5-F-7-RdX).

We hoped that Vilsmeier-Haack reaction on **425** might proceed more smoothly. As noted previously, the formyl group of **446** is a vinylous amide, and might withstand the reduction of the ribose to ribityl group and the deprotection of benzyl groups by hydrogenolysis. If not, we proposed that the formyl group could be protected with 1,3-propandithiol to generate **447** (alternatively, ethylene glycol could be used to generate a dioxolane protecting group). Next, applying the protocol developed for alkylation between **447** and L-ribose bromide **423** will afford the adduct **448**. Removal of benzyl groups by hydrogenolysis will give deazaxanthine **449**, followed by removal of the sugar protecting groups (e.g. with 3M HCl), and reduction (with $NaBH_4$) would afford tetraol **450**. Protection as a tetrabenzoate would provide **451**, which is easier to purify compared with tetraol **450**. We hoped that the benzoylation (Bz) using benzoic anhydride would give protected-5-F-7-RdX **452**. Lastly, the 1,3-propandithiol group

would be removed by treatment with NIS and TfOH to give 5-F-7-RdX **408** (Scheme 4.23).

We initially sought to optimize the Vilsmeier-Haack formylation on 2,4-dibenzyl-pyrrolopyrimidine. Treatment of **425** with Vilsmeier-Haack reagent, generated by reaction of phosphorus oxychloride and DMF, resulted in the formation of an iminium salt as detected by mass spectroscopy. Different conditions were used to firstly optimize of formation of iminium salt, at the same time minimize the decomposition of the hydrolyzed product during work-up.




**425** **446**

**Table 4.4:** Optimization for Vilsmeier-Haack reaction of **411**.

| Entry | Equivalents of $POCl_3$ | Solvent | Temp ($^oC$) | Time (h) | Workup | Yield |
|---|---|---|---|---|---|---|
| 1 | 1.2 | $CH_2Cl_2$ | 60 | 24 | $NaOH/H_2O$ | <5 |
| 2 | 1.2 | $CH_2Cl_2$ | 60 | 24 | $KOH/H_2O$ | 9 |
| 3 | 1.5 | $CH_2Cl_2$ | 60 | 4 | $NaOAc/H_2O$ | 26 |
| 4 | 2 | $CH_2Cl_2$ | 60 | 4 | $NaOAc/H_2O$ | 8 |
| 5 | 3 | $CH_2Cl_2$ | 60 | 4 | $NaOAc/H_2O$ | 30 |
| 6 | 1.5 | $(C_2H_2)_2Cl_2$ | 80 | 4 | $NaOAc/H_2O$ | 10 |
| 7 | 2 | $(C_2H_2)_2Cl_2$ | 80 | 4 | $MeOH/H_2O$ | 0 |

Using 1.2 equivalents of phosphorus oxychloride in $CH_2Cl_2$ at reflux for 24 h gave reasonable amount of the iminium salt, detected by mass spectroscopy and TLC. However, less than 5% of **446** was isolated after workup process utilizing NaOH (entry 1). The yield improved slightly to 9% when KOH was used for the work-up (entry 2). Increasing the amount of phosphorus oxychloride to 1.5 equivalents and using sodium acetate for the workup process gave an improved yield of 26% (entry 3). However,

increasing the amount of phosphorus oxychloride to 2 equivalents reduced the yield of **452** to 8% (entry 4). When 3 equivalents of phosphorus oxychloride was used, the yield of **452** increased again to 30%. The yield of **452** did not improve when the solvent was changed to dichloroethane and conducted at higher temperature (entries 6, 7). In all attempts, the mass recovered was poor, with decomposition of material during work-up. Overall, while a Vilsmeier-Haack formylation is feasible on a 2,4-dibenzylpyrrolopyrimidine, the low yields did not bode well for successful completion of the proposed approach.

## 4.3 Conclusion

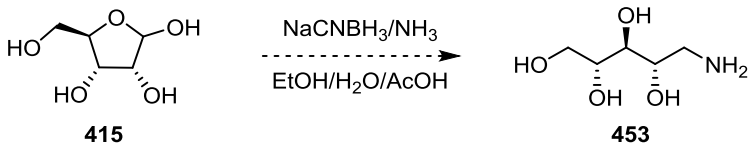
In this chapter, we proposed a structure, 5-F-7-RdX, for an unknown antigen arising from reaction of MDA and 5-A-RU, and report the synthesis of a close analogue, 7-RdX. This antigen selectively activates MAIT Tyr94α-TCRs over Tyr-95α-TCRs. Capture of the unknown antigen by MR1 and X-ray crystallography allowed identification of its electron density but at insufficient resolution to unambiguously define its structure. 7-RdX was synthesized by alkylation of methyl or benzyl protected 2,4-dihydroxy pyrrolopyrimidine by methyl 5-bromo-L-ribofuranoside. Stepwise deprotection followed by reduction afforded crude 7-RdX. To purify this molecule, benzoyl groups were temporarily introduced into the sugar tails, enabling purification and isolation of pure 7-RdX. 7-RdX was shown to act as a weak and selective agonist for MAIT TCR-Tyr94α cell lines, over Tyr95α cell lines. The potency was much lower than that achieved by the antigen formed in the mixture of MDA and 5-A-RU, which is proposed to be as a result of the lack of a formyl group, which would engage in Schiff base formation within the MR1 cleft. An X-ray structure of the complex of 7-RdX and MR1 gave a structure strinkingly similar to that of the unknown antigen, giving confidence that it is in fact 5-F-7-RdX.

As yet, the synthesis of 5-F-7-RdX has not been achieved. Direct formylation of protected 7-RdX using either the Vilsmeier-Haack or Riemer-Tiemann reactions were unsuccessful.

### 4.3.1 Future work

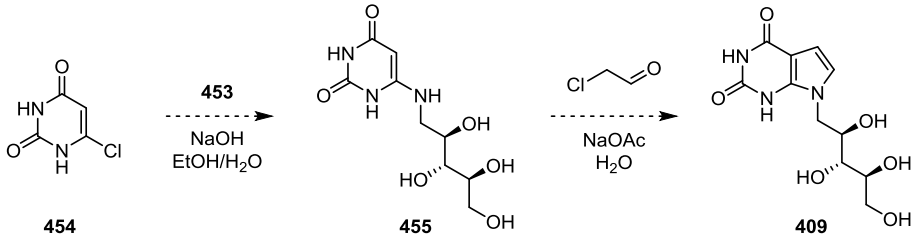
**New synthetic strategy for preparation of 5-F-7-RdX**

We proposed a new route to 5-F-7RdX **408** inspired by the procedure developed by Plaut,[344] which was used to prepare the series of riboflavin metabolites RL-6,7-diMe, RL-6-Me-7-OH and the reduced form rRL-6-$CH_2OH$ mentioned in the introduction.[119] The new approach involves formation of 5-A-RU and annulation of the pyrrole ring. Initially, direct reductive amination of D-ribose **415** will afford 1-amino-1-deoxy-D-ribitol **453**.[345]




**Scheme 4.24:** Protecting-group-free reductive amination reaction of D-ribose.

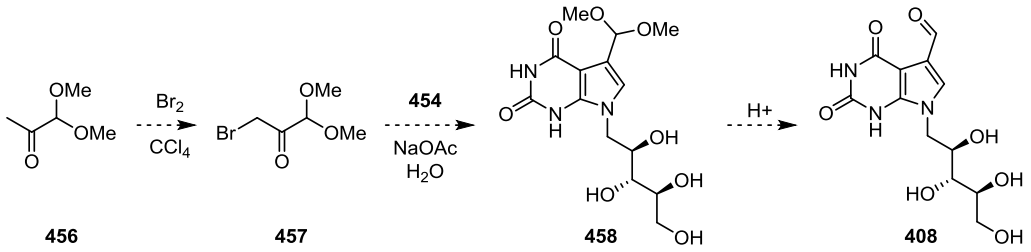
Nucleophilic aromatic substitution of 6-chlorouracil **454** with ribitylamine **453** under basic conditions will give 6-ribityluracil **455**.[346] We propose that condensation of **455** and chloroacetaldehyde will afford 7-RdX **409** (Scheme 4.25). This three-step method should be more efficient than our previously described approach.




**Scheme 4.25:** New preparation method of 7-RdX **409**.

If chloroacetaldehyde annulation of **455** is successful, we predict that similar chemistry can occur between 6-ribityluracil **454** and 3-bromo-1,1-dimethoxypropane-2-

one **457** to afford the dimethyl acetal of 5-F-7-RdX **458**. 3-Bromo-1,1-dimethoxypropane-2-one **457** can be prepared by bromination of methylglyoxal 1,1-dimethyl acetal **456** with bromine in carbon tetrachloride.[347] After cyclization, the dimethyl acetal could be cleaved by treatment with acid to yield 5-F-7-RdX **408**.

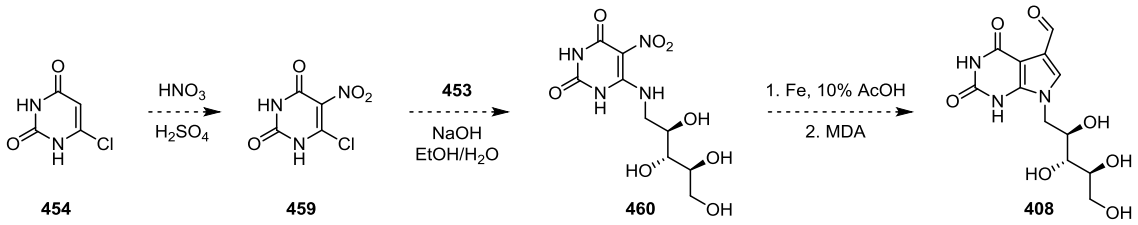



456 457 458 408

**Scheme 4.26:** Preparation of 5-F-7-RdX **408** via 6-ribityluracil **458**.

The proposed new synthetic route using 6-chlorouracil **454** provides a new approach for preparation of 5-F-7-RdX **408** and 7-RdX **409**. It avoids the need for protecting groups on both the pyrimidine ring and the sugar. This new route gives access to 7-RdX **409** in three steps and to 5-F-7-RdX **408** in five steps.

It has not escaped our notice that successful synthesis of 5-F-7-RdX **408** would not necessarily demonstrate that the unknown product of MDA and 5-A-RU possesses this structure. With an authentic sample of **408** in hand, we would seek to demonstrate whether this compound arises from the reaction of MDA and 5-A-RU. To obtain useful quantities of 5-A-RU, it will also be prepared. Nitration of commercially available 6-chlorouracil **454** with fuming $HNO_3$ in the presence of conc. $H_2SO_4$ will afford 6-chloro-5-nitrouracil **459**.[348] Nucleophilic aromatic substitution of **459** with D-ribitylamine **453** will give 6-ribitylaminouracil **460**.[349] Reduction of the nitro group using Fe powder in 10% AcOH (aq.) will give 5-A-RU. Our synthetically prepared 5-F-7-RdX will be used to conduct LC/MS and NMR analysis of the mixture of 5-A-RU and MDA to establish whether it is present, and confirm our hypothesis about the structure of 5-F-7-RdX **408** (Scheme 4.27).




**Scheme 4.27:** Preparation of 5-F-7-RdX **408** via 5-A-RU.

**Discovery of new novel MR1 ligands associated with different 1,3-dicarbonyls?**

Exploring different types of 1,3-dicarbonyl molecules could be a new and fruitful direction for this project. As mentioned in the Introduction, 5-A-RU reacts with glyoxal and methylglyoxal to generate 5-OE-RU and 5-OP-RU, respectively, which are agonists for MAIT cells. Therefore we predict that if the reaction of MDA **402** and 5-A-RU provides 5-F-7-RdX, which is a MAIT cell antigen, the related 1,3-dicarbonyl 3-oxobutanal **462** (methyl-MDA, mMDA) may produce 5-acetyl-7-RdX **463**, which might also prove to be a MAIT cell antigen. Towards this end we prepared mMDA **462** by treatment of 4,4-dimethylbutan-2-one **461** with resin ($H^+$ form) in deionized water. **462** decays at r.t within hours, therefore material needs to be freshly prepared, and can be stored at -80 $^oC$ freezer for subsequent use.

In collaboration with Prof. James McCluskey, we aim to investigate whether refolding of MR1 by loading a mixture of 5-A-RU and mMDA generates an antigen that can activate MAIT cells. A crystal structure of MR1 folded in the presence of 5-A-RU and mMDA could be obtained to help identify the structure of the new ligand, and mass spectrometry could be enlisted. If successful, this might expand further our knowledge of MAIT cell antigens. Based on the work outlined in this chapter, we predict that treatment of 5-A-RU with mMDA **462** will result in the formation of 5-acetyl-7-deazaxanthine (5-A-7-RdX, **463**). In **463** the formyl group of 5-F-7-RdX **408** is replaced by an acetyl group, the ketone of which should be able to form a Schiff-base with the lysine residue in the cleft of MR1.

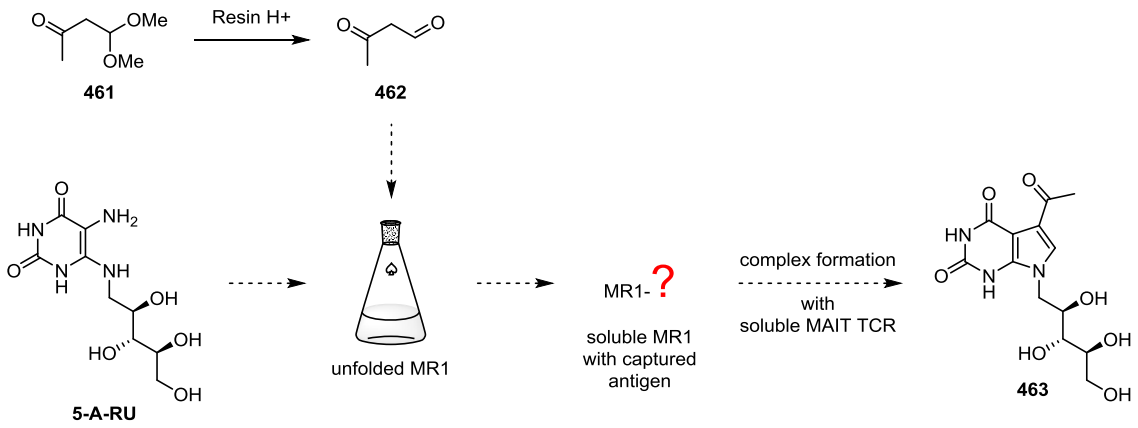
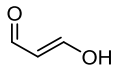


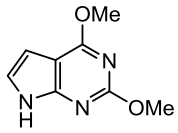

**Figure 4.9:** Develop and identify new ligands of MAIT cells.

## 4.4 Experimental



**Malondialdehyde (402)**

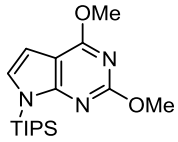
Dowex 50WX8-200 resin ($H^+$ form, approx 100 mg) was added to a mixture of tetramethoxypropane (TEP) (0.231 mg, 1.05 mmol, 0.262 mL) in deionized water (13.1 mL). The mixture was stirred at rt for 1 h, filtered to remove the resin, and additional deionized water (13.1 mL) used to wash the resin. $^1$H NMR (400 Hz, $D_2O$) δ 8.49 (2 H, m), 5.63-5.68 (1 H, t, *J* = 10.3 Hz). $^{13}$C NMR (100 Hz, $D_2O$) δ 168.4, 111.2, 88.9.



**2,4-Dimethoxy-7*H*-pyrrolo[2,3-d]pyrimidine (411)**

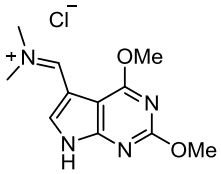
NaOMe (1 M in MeOH); (20 mmol, 28 mL) was added dropwise to the solution of 2,4-dichloro-7H-pyrrolo[2,3-d]pyrimidine **410** (2.63 g, 14 mmol) in MeOH (14 mL) and heated under reflux for 7 d. The mixture was concentrated, and water was added to the residue. The precipitate was filtered, and the crude product was purified by crystallization from MeOH to afford **411** as a white solid (2.50 g, quant yield). $^1$H NMR (400 Hz, DMSO-$d_6$) δ 11.74 (1 H, s), 7.09 (1 H, d, *J* = 2.8 Hz), 6.35 (1 H, s, *J* = 3.0 Hz),

3.97 (3 H, s), 3.87 (3 H, s). $^{13}$C NMR (100 Hz, DMSO-$d_6$) δ 163.5, 161.1, 153.8, 121.9, 99.6, 98.1, 54.1, 53.3. HRMS (ESI) *m/z* calcd. for $[C_8H_9N_3O_2+H]^+$: 180.0768, obsd: 180.0769.




**2,4-Dimethoxy-7-(triisopropylsilyl)-7*H*-pyrrolo[2,3-d]pyrimidine (412)**

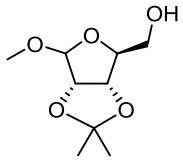
n-BuLi (2.5 M in ether; 0.307 mmol, 0.122 mL) was added dropwise into the mixture of **411** (50 mg, 0.279 mmol) in dried THF (5.6 mL) at -78 °C and stirred for 10 mins. Subsequently, triisopropylsilyl chloride (59.2 mg, 0.307 mmol, 65.7 µL) was added dropwise and the reaction was warmed up to r.t over 30 mins. Solvent was evaporated; water was added to the residue and extracted with diethyl ether. The organic layer was then dried over anhydrous $MgSO_4$ and concentrated in vacuo to afford product as colorless oil (93.1 mg, 98%). $^1$H NMR (400 Hz, $CDCl_3$) δ 6.91 (1 H, d, *J* = 3.5 Hz), 6.52 (1 H, d, *J* = 3.5 Hz), 4.07 (3 H, s), 3.94 (3 H, s), 1.79-1.85 (3 H, dt, *J* = 15.0, 7.5 Hz), 1.12 (18 H, d, *J* = 7.5 Hz). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 164.4, 161.3, 159.4, 126.7, 102.7, 101.2, 54.6, 53.8, 18.2, 12.3.




**Iminium salt of 2,4-dimethoxy-7*H*-pyrrolo[2,3-d]pyrimidine (413)**

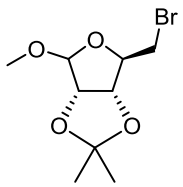
A mixture of DMF (9.3 mg, 0.128 mmol, 10 µL) and DCM (2.6 mL) was added dropwise into the stirring solution of oxalyl chloride (15 mg, 0.116 mmol, 10 µL) in DCM (2.3 mL) and stirred at 0 °C over 20 mins. **412** (37.7 mg, 0.111 mmol) in DCM (1.1 mL) was added dropwise into the Vilsmeier-Haack reagent at 0 °C and then heated under reflux for 30 mins. The reaction was then cooled down to 0 °C and the precipitate

was filtered, washed with diethyl ether and dry. Iminium salt was isolated as white solid (30 mg, quant yield). $^{1}$H NMR (400 Hz, DMSO-$d_6$) δ 8.84 (1 H, s), 8.53 (1 H, s), 5.60 (1 H, s), 4.10 (3 H, s), 3.95 (3 H, s), 3.77 (3 H, s), 3.60 (3 H, s).



**Methyl 2,3-*O*-isopropylidene-L-ribofuranoside (422)**

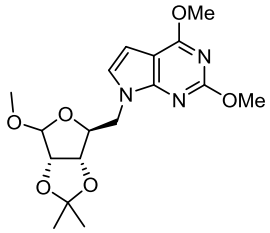
Concentrated HCl (0.7 mL) was added dropwise to the mixture of L-ribose (6.50 g, 43.3 mmol) in methanol:acetone mixture (1:1, 50 mL) at r.t and heated under reflux for 4 h. The solution was cooled down to room temperature and aq. $NaHCO_3$ was added dropwise to neutralize the reaction. The solvent was concentrated and the residue was dissolved in $CHCl_3$, washed with water, dried over anhydrous $MgSO_4$ and concentrated in vacuo. The crude material was purified by flash chromatography (EtOAc:pet. Spirits, 20:80) to afford a yellowish oil **422** as the product (6.39 g, 72%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 4.97 (1 H, s, H1), 4.83-4.84 (1 H, d, *J* = 5.9 Hz, H5a), 4.58-4.59 (1 H, d, *J* = 5.9 Hz, H5b), 4.42-4.43 (1 H, t, *J* = 2.6 Hz, H2), 3.67-3.71 (1 H, dt, *J* = 12.5, 2.4 Hz, H3), 3.59-3.64 (1 H, td, *J* = 12.5, 10.7, 3.2 Hz, H), 3.21-3.24 (1 H, dd, *J* = 10.6, 1.9 Hz, 1H), 3.43 (3 H, s), 1.48 (3 H, s), 1.31 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 112.3, 110.1, 88.5, 85.5, 86.0, 81.6, 64.1, 55.7, 26.5, 24.8. HRMS (ESI) *m/z* calcd. for $[C_9H_{16}O_5+H]^+$: 205.1071, obsd: 205.1072.



**Methyl 5-bromo-5-deoxy-2,3-O-isopropylidene-L-ribofuranoside (423)**

Tetrabromomethane (7.24 g, 21.84 mmol) and triphenylphosphine (5.73 g, 21.84 mmol) were added portion to the solution of **422** (2.23 g, 10.92 mmol) in dried $CH_2Cl_2$ (101

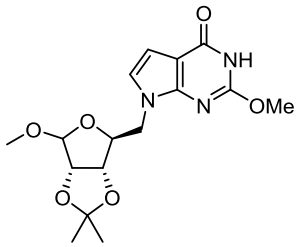
mL) at 0 ºC and stirred under nitrogen overnight. The mixture was then concentrated and the residue was purified by flash chromatography (EtOAc:pet. spirits, 10:90) to give a colorless oil **423** as the product (1.84 g, 63%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 5.01 (1 H, s, H1), 4.78-4.76 (1 H, d, *J* = 5.9 Hz, H5a), 4.61- 4.62 (1 H, d, *J* = 5.9 Hz, H5b), 4.37-4.41 (1 H, dd, *J* = 10.0, 5.8 Hz, H2), 3.41-3.45 (1 H, ddd, *J* = 10.1, 5.8, 0.8 Hz, H3), 3.30-3.32 (1 H, dd, *J* = 10.1, 0.9 Hz, H4), 3.35 (3 H, s), 1.48 (3 H, s), 1.33 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 112.8, 109.7, 86.8, 85.3, 82.7, 55.2, 32.6, 26.5, 25.1. HRMS (ESI) *m/z* calcd. for $[C_9H_{15}BrO4+Na]^+$: 289.0046, obsd: 289.0043.




**7-(Methyl 5-deoxy-2,3-*O*-isopropylidene-L-ribofuranos-5-yle)-2,4-dimethoxy-7*H*-pyrrolo[2,3-d]pyrimidine (424)**

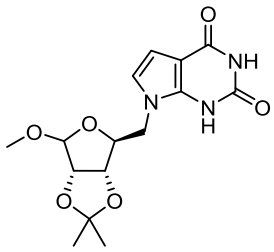
Sodium hydride (160 mg, 6.68 mmol) was added portion to the solution of 2,4-dimethoxy-7*H*-pyrrolo[2,3-d]pyrimidine **411** (599 mg, 3.34 mmol) and 5-bromo-methyl-2,3-*O*-isopropylidene-L-ribose **423** (982 mg, 3.68 mmol) in dried DMF (70 mL) at r.t and the resulting mixture was heated to 45 ºC and stirred vigorously for 3 d. The solvent was concentrated and the residue was purified by flash chromatography (EtOAc:pet. spirits, 20:80) to give the white solid **424** (646.8 mg, 53%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 6.88-6.89 (1 H, d, *J* = 3.4 Hz), 6.44-6.45 (1 H, d, *J* = 3.4 Hz), 5.01 (1 H, s, H1), 4.77-4.79 (1 H, d, *J* = 5.9 Hz, H5a), 4.68-4.70 (1 H, d, *J* = 5.9 Hz, H5b), 4.60-4.64 (1 H, t, *J* = 7.4 Hz, H2), 4.37-4.43 (1 H, dd, *J* = 14.0, 8.2 Hz, H3), 4.11-4.12 (1 H, dd, *J* = 14.0, 6.7 Hz, H4), 4.08 (3 H, s), 4.01 (3 H, s), 3.99 (3 H, s), 1.43 (3 H, s), 1.27 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 164.5, 161.8, 153.4, 124.1, 112.7, 110.1, 100.8, 99.2,

85.5, 85.4, 88.3, 55.5, 54.7, 53.9, 47.8, 26.5, 25.1. HRMS (ESI) *m/z* calcd. for $[C_{17}H_{23}N_3O_6+H]^+$: 366.1660, obsd: 366.1660.



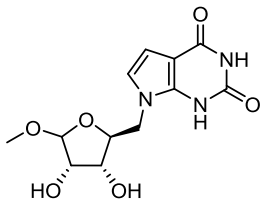

**7-(Methyl 5-deoxy-2,3-*O*-isopropylidene-L-ribofuranos-5-yle)-3,7-dihydro-2-methoxy-4*H*-pyrrolo[2,3-d]pyrimidine (437)**

Potassium hydroxide (4.96 g, 88.5 mmol) was added to the solution of **424** (646.8 mg, 1.77 mmol) in dried pyridine (177 mL) before the resulting mixture was heated under reflux for 3 d. The solvent was evaporated and the residue was purified by flash chromatography (EtOAc) to obtain a white solid **437** as product (540.7 mg, 87%). $^1$H NMR (400 Hz, $CDCl_3$) δ 10.45 (1 H, s, NH), 6.74-6.75 (1 H, d, *J* = 3.4 Hz), 6.63-6.64 (1 H, d, *J* = 3.4 Hz), 5.02 (1 H, s, H1), 4.75-4.76 (1 H, d, *J* = 5.9 Hz, H5a), 4.68-4.69 (1 H, d, *J* = 5.9 Hz, H5b), 4.59-4.63 (1 H, t, *J* = 7.5 Hz, H4), 4.26-4.31 (1 H, dd, *J* = 14.0, 8.1 Hz, H3), 4.04-4.10 (1 H, dd, *J* = 14.0, 6.7 Hz, H2), 4.02 (3 H, s), 3.40 (3 H, s), 1.46 (3 H, s), 1.29 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 160.6, 154.6, 148.5, 122.1, 112.8, 110.2, 103.1, 85.5, 85.4, 82.2, 55.6, 55.1, 48.2, 26.6, 25.1. HRMS (ESI) *m/z* calcd. for $[C_{16}H_{21}N_3O_6+H]^+$: 352.1503, obsd: 352.1503.



**7-(Methyl 5-deoxy-2,3-*O*-isopropylidene-L-ribofuranosid-5-yle)-1*H*-pyrrolo[2,3-d]pyrimidine-2,4(3*H*,7*H*)-dione (438)**

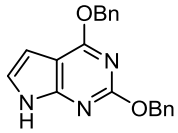
NaI (255 mg, 1.70 mmol) was added to the solution of **437** (298.8 mg, 0.851 mmol) in dried MeCN (85.1 mL) before distilled TMSCl (277 mg, 2.55 mmol, 324 µL) was added dropwise. The mixture was stirred at r.t for 4 h before MeOH was added. The solvent was concentrated and the residue was purified by flash chromatography (MeOH:EtOAc, 5:95) to obtain a white solid as the product (172 mg, 60%). $^{1}$H NMR (400 Hz, DMSO-$d_6$) δ 11.70 (1 H, s, NH), 10.59 (1 H, s, NH), 6.71-6.72 (1 H, d, *J* = 3.3 Hz), 6.28-6.29 (1 H, d, *J* = 3.3 Hz), 4.95 (1 H, s, H1), 4.64-4.69 (dd, 2 H, *J* = 14.6, 5.9 Hz, H2, H3), 4.30-4.34 (1 H, t, *J* = 7.1 Hz, H4), 4.10-4.15 (1 H, dd, *J* = 14.3, 6.2 Hz, H5a), 3.93-3.98 (1 H, dd, *J* = 14.5, 8.2 Hz, H5b), 3.29 (3 H, s), 1.37 (3 H, s), 1.26 (3 H, s). $^{13}$C NMR (100 Hz, DMSO-$d_6$) δ 159.7, 151.3, 120.5, 111.7, 109.5, 102.8, 98.9, 84.5, 81.4, 54.9, 48.2, 26.3, 24.7. HRMS (ESI) *m/z* calcd. for $[C_{15}H_{19}N_3O_6+H]^+$: 338.1347, obsd: 338.1345.



**7-(Methyl 5-deoxy-L-ribofuranosid-5-yle)-1,7-dihydro-2*H*-pyrrolo[2,3-d]pyrimidine-2,4(3H,7H)-dione (439)**

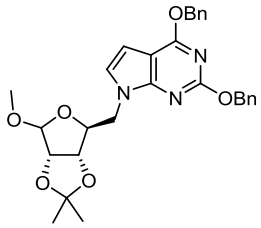
NaI (439 mg, 2.93 mmol) was added to the mixture of **437** (205.8 mg, 0.586 mmol) dissolved in dried MeCN (85.1 mL) before distilled TMSCl (382 mg, 3.516 mmol, 446

μL) was added dropwise into the solution. The mixture was stirred at rt for 4 h before MeOH was added. Solvent was evaporated and the residue was purified by flash chromatography (MeOH:EtOAc, 5:95) to obtain a white solid as the product (104.5 mg, 60%). $^{1}$H NMR (400 Hz, $CD_3OD$) δ 6.70 (1 H, d, *J* = 3.4 Hz, H from pyrimidine), 6.42-6.43 (1 H, d, *J* = 3.3 Hz), 4.71 (1 H, s, H1), 4.33-4.36 (1 H, dd, *J* = 14.8, 1.5 Hz, H5a), 4.12-4.15 (1 H, m, H4), 4.04-4.08 (1 H, dd, *J* = 14.9, 5.8 Hz, H5b), 3.96-3.98 (1 H, m, H3), 3.78-3.79 (1 H, d, *J* = 4.5 Hz, H2), 3.26 (3 H, s, $OCH_3$). $^{13}$C NMR (100 Hz, $CD_3OD$) δ 159.7, 151.3, 122.9, 110.3, 104.1, 104.0, 99.1. 84.5, 82.0, 75.9, 72.9, 56.0. HRMS (ESI) *m/z* calcd. for $[C_{12}H_{15}N_3O_6+H]^+$: 298.1034, obsd: 298.1032.



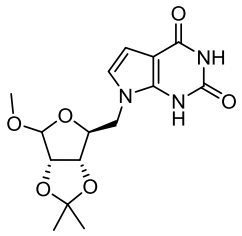

**2,4-Dibenzyl-7*H*-pyrrolo[2,3-d]pyrimidine (425)**

Benzyl alcohol (8.65 g, 80 mmol, 8.32 mL) was added dropwise into a suspension solution of NaH (60% wt; 1.60 g, 40 mmol) in anhydrous toluene (100 mL). When the hydrogen evolution stopped, 2,4-dichloro-7*H*-pyrrolo[2,3-d]pyrimidine (1.88 g, 10 mmol) was added to the stirring solution and then heated under reflux for 16 h. The solution was cooled down to room temperature, neutralized with AcOH and the solvent was evaporated. The crude material was purified by flash chromatography (EtOAc:pet. spirits, 20:80) followed by crystallization from MeOH to give a white solid **425** (3.15 g, 95%). $^{1}$H NMR (400 Hz, DMSO-$d_6$) δ 9.66 (1 H, s, NH), 7.47-7.50 (4 H, m), 7.31-7.40 (6 H, m), 6.92-6.94 (1 H, dd, *J* = 3.1, 2.3 Hz), 6.48-6.50 (1 H, dd, *J* = 3.1, 2.2 Hz), 5.57 (2 H, s, $CH_2$), 5.47 (2 H, s, $CH_2$). $^{13}$C NMR (100 Hz, DMSO-$d_6$) δ 163.9, 161.1, 154.5, 137.3, 136.8, 128.9, 128.2, 128.0, 123.3, 120.8, 100.9, 99.7, 69.1, 68.2. HRMS (ESI) *m/z* calcd. for $[C_{20}H_{17}N_3O_2+H]^+$: 332.1394, obsd: 332.1394.



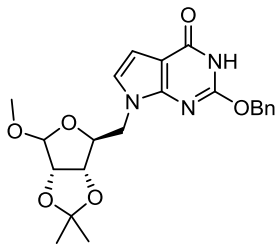
**7-(Methyl 5-deoxy-2,3-*O*-isopropylidene-L-ribofuranosid-5-yle)-2,4-dibenzyl-7*H*-pyrrolo[2,3-d]pyrimidine (426)**

Sodium hydride (232 mg, 5.81 mmol) was added portionwise to the solution of **425** (1.284 g, 3.87 mmol) in anhydrous DMF (39 mL). When the hydrogen evolution stopped, 5-bromo-methyl-2,3-*O*-isopropylidene-L-ribose (1.552 g, 5.81 mmol) in dried DMF (58 mL) was added into the stirring mixture and heated up to 60 ºC for 16 h. The solvent was evaporated and the residue was purified by flash chromatography (EtOAc:pet. spirits, 10:90) to afford white solid product **426** (1.70 g, 85%). $^{1}$H NMR (400 Hz, $CDCl_3$) δ 7.51-7.53 (2 H, m, Ph), 7.46-7.48 (2 H, m), 7.28-7.41 (6 H, m), 6.89-6.90 (1 H, d, *J* = 3.5 Hz), 6.47-6.48 (1 H, d, *J* = 3.5 Hz), 5.56 (2 H, s, $CH_2$), 5.47-5.48 (2 H, d, *J* = 6.6 Hz, $CH_2$), 5.03 (1 H, s, H1), 4.71-4.72 (1 H, d, *J* = 6.0 Hz, H2), 4.67-4.68 (1 H, d, *J* = 5.9 Hz, H3), 4.59-4.62 (1 H, m, H4), 4.40-4.44 (1 H, dd, *J* = 14.0, 8.4 Hz, H5a), 4.07-4.11 (1 H, dd, *J* = 14.0, 6.6 Hz, H5b), 3.41 (3 H, s), 1.44 (3 H, s), 1.26 (3 H, s). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 163.9, 161.0, 153.5, 137.4, 136.8, 128.6, 128.5, 128.3, 128.2, 128.1, 127.9, 124.2, 112.7, 110.1, 101.1, 99.4, 85.5, 85.4, 82.2, 69.1, 68.2, 55.5, 47.7, 26.5, 25.1. HRMS (ESI) *m/z* calcd. for $[C_{29}H_{31}N_3O_6+H]^+$: 518.2286, obsd: 518.2286.



**7-(Methyl 5-deoxy-2,3-*O*-isopropylidene-L-ribofuranosid-5-yle)-1,7-dihydro-2*H*-pyrrolo[2,3-*d*]pyrimidine-2,4(3*H*)-dione (438)**

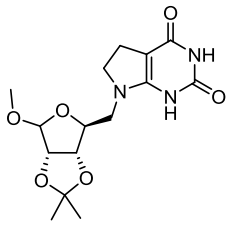
A mixture of **426** (1.70 g, 3.29 mmol) and Pd/C (85 mg, 5% wt) in the mixture of EtOAc:MeOH (1:1, 33 mL) was stirred under hydrogen at 20 bar for 1 h. The solvent was evaporated and the residue was purified by flash chromatography (MeOH: DCM, 10:90) to afford the white solid product **438** (899 mg, 81%). $^{1}$H NMR (400 Hz, DMSO-$d_6$) δ 11.70 (1 H, s, NH), 10.56 (1 H, s, NH), 6.71 (1 H, d, *J* = 3.3 Hz), 6.28-6.29 (1 H, dd, *J* = 3.3 Hz), 4.95 (1 H, s, H1), 4.68-4.70 (1 H, d, *J* = 5.8 Hz, H2/3), 4.64-4.66 (1 H, d, *J* = 5.9 Hz, H2/3), 4.30-4.34 (1 H, t, *J* = 7.0 Hz, H4), 4.10-4.15 (1 H, dd, *J* = 14.4, 6.1 Hz, H5a), 3.92-3.98 (1 H, dd, *J* = 14.4, 8.2 Hz, H5b), 3.29 (3 H, s), 1.37 (3 H, s), 1.25 (3 H, s). $^{13}$C NMR (100 Hz, DMSO-$d_6$) δ 159.7, 151.3, 138.9, 120.5, 111.7, 109.5, 102.8, 98.9, 84.5, 81.4, 54.9, 48.2, 26.3, 24.7. HRMS (ESI) *m/z* calcd. for $[C_{15}H_{19}N_3O_6+H]^+$: 338.1347, obsd: 338.1345.



**2-Benzyloxy-7-(Methyl 5-deoxy-2,3-*O*-isopropylidene-L-ribofuranosid-5-yle)-3,7-dihydro-4*H*-pyrrolo[2,3-*d*]pyrimidine-4-one (440a)**

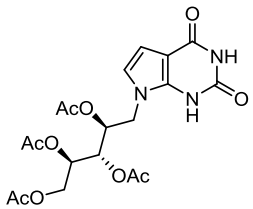
$^{1}$H NMR (400 Hz, DMSO-$d_6$) δ 11.85 (1 H, s, NH), 7.46-7.48 (2 H, d, *J* = 7.2 Hz), 7.31-7.40 (3 H, m), 7.01-7.02 (1 H, d, *J* = 3.4 Hz), 6.35 (1 H, d, *J* = 3.4 Hz), 5.35-5.43 (2 H, m), 4.97 (1 H, s, H1), 4.70-4.71 (1 H, d, *J* = 6.0 Hz, H2/3), 4.63-4.65 (1 H, d, *J* = 5.9

Hz, H2/3), 4.42-4.46 (1 H, t, *J* = 7.5 Hz, H4), 4.14-4.20 (1 H, dd, *J* = 13.8, 8.5 Hz, H5a), 4.02-4.07 (1 H, dd, *J* = 13.8, 6.7 Hz, H5b), 3.28 (3 H, s), 1.33 (3 H, s), 1.18 (3 H, s). $^{13}$C NMR (100 Hz, DMSO-$d_6$) δ 158.9, 154.2, 147.1, 135.9, 128.4, 128.3, 122.4, 111.7, 109.0, 102.8, 101.8, 84.7, 84.60, 81.3, 68.5, 54.7, 47.1, 26.2, 24.7. HRMS (ESI) *m/z* calcd. for $[C_{22}H_{25}N_3O_6+H]^+$: 428.1816, obsd: 428.1817.



**7-(Methyl 5-deoxy-2,3-O-isopropylidene-L-ribofuranosid-5-yle)-1,5,6,7-tetrahydro-2*H*-pyrrolo[2,3-*d*]pyrimidine-2,4(3H)-dione (440b)**

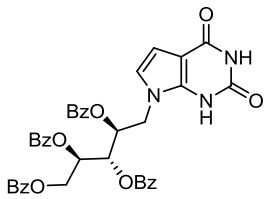
$^1$H NMR (400 Hz, DMSO-$d_6$) δ 11.14 (1 H, s, NH), 9.97 (1 H, s, NH), 4.94 (1 H, s, H1), 4.20-4.23 (1 H, t, *J* = 6.9 Hz, H2/3), 3.69-3.76 (1 H, dd, *J* = 17.3, 9.8 Hz), 3.48-3.54 (3 H, m), 3.27 (3 H, s,$OCH_3$), 3.09-3.17 (2 H, m), 2.55-2.61 (2 H, dd, *J* = 15.2, 7.9 Hz), 1.38 (3 H, s, $CH_3$), 1.26 (3 H, s, $CH_3$). $^{13}$C NMR (100 Hz, DMSO-$d_6$) δ 160.2, 152.2, 111.6, 109.4, 84.5, 84.4, 83.6, 82.0, 54.8, 51.4, 49.0, 26.3, 24.7, 22.4. HRMS (ESI) *m/z* calcd. for $[C_{15}H_{21}N_3O_6+H]^+$: 338.1347, obsd: 338.1345.



**7-(2,3,4,5-Tetra-*O*-acetyl)-D-ribityl-7-deazaxanthine (441)**

HCl (3 M; 9.0 mmol, 3.0 mL) was added dropwise in to the solution of **438** (101.2 mg, 0.3 mmol) in dioxane (15 mL) and was heated up to 60 ºC for 1 h. After that, the mixture was cooled down to 10 ºC before $NaBH_4$ (0.397 g, 10.5 mmol) was added portion until the pH of the reaction turned basic. The mixture was warmed up to room

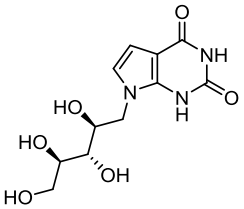
temperature and continues stirred vigorously for 1 h before HCl (1 M) was added dropwise to neutralize the reaction. The solvent was evaporated and the residue was co-evaporated twice with pyridine. The residue was dissolved again in pyridine (15 mL) followed by the addition of $Ac_2O$ (1.531 g, 15 mmol, 1.42 mL) and stirred at room temperature for 4 h. The solvent was evaporated and the residue was washed with mixture of DCM:methanol (1:1) and filtered through Celite. The solvent was evaporated and the residue was purified by flash chromatography (EtOAc:MeOH; 20:80) to afford white solid as a product. $^1$H NMR (400 Hz, DMSO-$d_6$) δ 11.71 (1 H, s, NH), 10.59 (1 H, s, NH), 6.59-6.60 (1 H, d, *J* = 3.3 Hz), 6.22 (1 H, d, *J* = 3.3 Hz), 5.16-5.22 (2 H, m), 4.29-4.32 (2 H, m), 4.13-4.23 (2 H, m), 2.06 (3 H, s), 2.03 (3 H, s), 1.98 (3 H, s), 1.88 (3 H, s). $^{13}$C NMR (100 Hz, DMSO-$d_6$) δ 170.1, 169.8, 169.3, 169.0, 159.6, 151.2, 120.6, 103.2, 98.8, 70.0, 69.6, 69.4, 61.3, 45.1, 20.7, 20.5, 20.4. HRMS (ESI) *m/z* calcd. for $[C_{19}H_{23}N_3O_{10}+H]^+$: 454.1456, obsd: 454.1453.




**7-(2,3,4,5-Tetra-*O*-benzoyl)-D-ribityl-7-deazaxanthine (442)**

Hydrochloric acid (3M, 79.8 mmol, 26.6 mL) was added dropwise into the mixture of SM (899 mg, 2.66 mmol) in dioxane (133 mL) before heated at 60 ºC for 1 h. After that, reaction was cooled down to 10 ºC before $NaBH_4$ (3.53 g, 93.3 mmol) was added portionwise. pH changed from acidic to basic environment. Reaction was warmed up to r.t and the reaction stirred for a further 1 h. When the reaction had reached completion, 1 M HCl was added dropwise to neutralize the condition of the reaction, then solvent was evaporated to dryness. The residue was dissolved in dried DMF (133 mL), before 4-DMAP (32.5 mg, 0.266 mmol), DIPEA (1.72 g, 13.3 mmol, 2.32 mL) and $Bz_2O$ (2.71

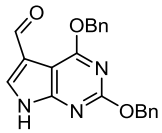
g, 12.0 mmol, 2.26 mL) were added respectively and the reaction was stirred at room temperature for 4 h. After that, water (1.60 mmol, 29 µL) was added to quench the excess of the reagent. The solvent was evaporated and the residue was diluted with water and extracted into EtOAc. The organic layer was washed with brine, dried ($MgSO_4$), filtered and concentrated in vacuo. The residue was purified by flash chromatography (EtOAc:pet. spirits, 80:20) to afford white solid as a product (692 mg, 37%). $^1$H NMR (400 Hz, $CDCl_3$) δ 12.11 (1 H, s, NH), 8.54 (1 H, s, NH), 8.06-8.08 (2 H, d, *J* = 7.5 Hz), 7.99- 8.01 (2 H, d, *J* = 7.5 Hz), 7.94-7.96 (2 H, d, *J* = 7.3 Hz), 7.93-7.95 (2 H, d, *J* = 7.3 Hz), 7.33-7.60 (15 H, m), 6.41-6.44 (2 H, dd, *J* = 7.2, 3.4 Hz), 6.11-6.14 (1 H, m, H3), 6.00-6.03 (1 H, m, H4), 5.85-5.88 (1 H, dd, *J* = 8.2, 4.3 Hz, H2), 4.99-5.03 (1 H, dd, *J* = 12.3, 2.7 Hz, H5a), 4.56-4.61 (1 H, dd, *J* = 12.3, 6.5 Hz, H5b), 4.50-4.53 (1 H, dd, *J* = 15.2, 2.8 Hz, H1a), 4.39-4.45 (1 H, dd, *J* = 15.2, 9.2 Hz, H1b). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 166.4, 166.3, 165.3, 165.0, 159.0, 153.0, 138.7, 134.2, 134.1, 133.9, 133.5, 130.2, 130.1, 129.9, 129.4, 128.9, 128.8, 128.7, 128.6, 120.8, 105.1, 100.5, 71.5, 71.3, 70.8, 63.0, 46.4. HRMS (ESI) *m/z* calcd. for $[C_{39}H_{31}N_3O_6+H]^+$: 702.2082, obsd: 702.2081.




**7-D-Ribityl-deazaxanthine (409)**

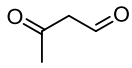
Sodium methoxide (1 M in MeOH; 0.375 mmol, 0.375 mL) was added dropwise to the solution of **442** (26.3 mg, 0.0375 mmol) in dried MeOH (5.6 mL) and stirred at room temperature for 1 h. The reaction was then quenched with Dowex resin (H+ form), filtered and concentrated. The residue was washed with methanol, filtered, and concentrated to afford product as a brown oil (10 mg, 93%). $^1$H NMR (400 Hz, DMSO-

$d_6$) δ 11.11 (1 H, s, NH), 10.51 (1 H, s, NH), 6.64-6.65 (1 H, d, *J* = 3.3 Hz), 6.24-6.25 (1 H, d, *J* = 3.3 Hz), 4.97-4.99 (1 H, d, *J* = 5.8 Hz, C2(OH)), 4.85-4.86 (1 H, d, *J* = 5.4 Hz, C4(OH)), 4.82 (1 H, br s, C3(OH)), 4.46 (1 H, s, C5(OH)), 4.07-4.11 (1 H, dd, *J* = 14.5, 2.4 Hz, C1a), 3.95-4.01 (1 H, dd, *J* = 14.5, 8.3 Hz, C1b), 3.80-3.81 (1 H, m, C2), 3.57-3.60 (4 H, m, C3, C4, C5a, C5b). $^{13}$C NMR (100 Hz, DMSO-$d_6$) δ 159.7, 151.0, 139.4, 120.9, 102.4, 98.5, 73.1, 72.5, 71.2, 63.1, 48.0. HRMS (ESI) *m/z* calcd. for $[C_{11}H_{15}N_3O_6+H]^+$: 286.1034, obsd: 286.1034.




**2,4-(Dibenzyloxy)-7-formyl-7*H*-pyrrolo[2,3-d]pyrimidine (452)**

Phosphorus oxychloride (230 mg, 1.5 mmol, 140 µL) was added dropwise into the mixture of DMF (109.6 mg, 1.5 mmol, 116 µL) and $CH_2Cl_2$ (15 mL) at 0 $^o$C and stirred over 20 mins. After that, **425** (165.7 mg, 0.5 mmol) in $CH_2Cl_2$ (5 mL) was added dropwise into the reaction at 0 $^o$C before it was heated under reflux for 4 h. The reaction was cooled down to 0 $^o$C before NaOAc.$3H_2O$ (245 mg, 1.8 mmol) and $H_2O$ (0.5 mL) were added and stirred for 30 mins. The reaction was extracted with $CH_2Cl_2$, The organic layer was washed with $H_2O$, $NaHCO_3$, dried $MgSO_4$, and filtered and concentrated in vacuo. The residue was purified by flash chromatography (EtOAc:pet. spirits, 10:90) to afford white solid as a product (53.9 mg, 30%). $^1$H NMR (400 Hz, $CDCl_3$) δ 7.24-7.48 (10 H, m, Ph), 6.92-6.94 (1 H, dd, *J* = 3.1, 2.3 Hz), 6.49-6.50 (1 H, dd, *J* = 3.7 Hz), 5.51 (2 H, s, $CH_2$), 5.38-5.49 (2 H, dd, *J* = 33.3, 12.3 Hz, $CH_2$). $^{13}$C NMR (100 Hz, $CDCl_3$) δ 163.87, 161.20, 153.86, 137.19, 136.67, 128.63, 128.57, 128.41, 128.19, 127.93, 120.77, 101.11, 100.83, 69.28, 68.24. HRMS (ESI) *m/z* calcd. for $[C_{21}H_{17}N_3O_3+H]^+$: 360.1343, obsd: 360.1337.



**3-Oxobutanal (mMDA; 462)**

Dowex 50WX8-200 resin ($H^+$ form, approx 100 mg) was added to a mixture of 3-Oxobutanal **462** (20 mg, 0.15 mmol, 20 µL) in deionized water (1 mL). The mixture was stirred at rt for 1 h, filtered to remove the resin, and additional deionized water (3 mL) used to wash the resin. $^1$H NMR (400 Hz, $D_2O$) δ 10.2 (1 H, s), 5.26 (2 H, m), 2.21 (3 H, s).

## 5. Bibliography


1. Pancer, Z.; Cooper, M. D., *Annu. Rev. Immunol.* **2006,** *24*, 497-518.
2. Mogensen, T. H., *Clin. Microbiol. Rev.* **2009,** *22*, 240-273.
3. Takeuchi, O.; Akira, S., *Cell* **2010,** *140*, 805-820.
4. Boller, T.; Felix, G., *Annu. Rev. Plant Biol.* **2009,** *60*, 379-406.
5. Mackey, D.; McFall, A. J., *Mol. Microbiol.* **2006,** *61*, 1365-1371.
6. Geijtenbeek, T. B. H.; Gringhuis, S. I., *Nat. Rev. Immunol.* **2009,** *9*, 465.
7. André, S.; Tough, D. F.; Lacroix-Desmazes, S.; Kaveri, S. V.; Bayry, J., *Am. J. Pathol.* **2009,** *174*, 1575-87.
8. Shevach, E. M., *Immunity* **2006,** *25*, 195-201.
9. Tang, Q.; Bluestone, J. A., *Nat. Immunol.* **2008,** *9*, 239-244.
10. Barton, G. M.; Kagan, J. C., *Nat. Rev. Immunol.* **2009,** *9*, 535+.
11. Akira, S.; Takeda, K., *Nat. Rev. Immunol.* **2004,** *4*, 499+.
12. Choe, J.; Kelker, M. S.; Wilson, I. A., *Science* **2005,** *309*, 581.
13. Sancho, D.; Reis e Sousa, C., *Annu. Rev. Immunol.* **2012,** *30*, 491-529.
14. Ahn, J.; Barber, G. N., *Exp. Mol. Med.* **2019,** *51*, 1-10.
15. Muzio, M.; Mantovani, A., *Microbes Infect.* **2000,** *2*, 251-255.
16. Botos, I.; Segal, David M.; Davies, David R., *Structure* **2011,** *19*, 447-459.
17. Lennarz, W. J.; Lane, M. D., *Encyclopedia of Biological Chemistry*. Elsevier Science: 2013.
18. Takeda, K.; Kaisho, T.; Akira, S., *Annu. Rev. Immunol.* **2003,** *21*, 335-376.
19. Ozinsky, A.; Underhill, D. M.; Fontenot, J. D.; Hajjar, A. M.; Smith, K. D.; Wilson, C. B.; Schroeder, L.; Aderem, A., *Proc. Nat. Acad. Sci. U. S. A.* **2000,** *97*, 13766.
20. Chow, J. C.; Young, D. W.; Golenbock, D. T.; Christ, W. J.; Gusovsky, F., *J. Biol. Chem.* **1999,** *274*, 10689-10692.
21. Toussi, D. N.; Massari, P., *Vaccines* **2014,** *2*, 323-353.
22. Moran, A. P.; Prendergast, M. M.; Appelmelk, B. J., *FEMS Immunol. Med. Microbiol.* **1996,** *16*, 105-115.
23. Tanimura, N.; Saitoh, S.; Matsumoto, F.; Akashi-Takamura, S.; Miyake, K., *Biochem. Biophys. Res. Commun.* **2008,** *368*, 94-99.
24. Kilár, A.; Dörnyei, Á.; Kocsis, B., *Mass Spectrom. Rev.* **2013,** *32*, 90-117.
25. Singer, M.; Deutschman, C. S.; Seymour, C.; et al., *JAMA* **2016,** *315*, 801-810.
26. O'Neill, L. A. J., *Nat. Med.* **2008,** *14*, 370-372.

27. Cluff, C. W., Monophosphoryl Lipid A (MPL) as an Adjuvant for Anti-Cancer Vaccines: Clinical Results. In *Lipid a in Cancer Therapy*, Jeannin, J. F., Ed. 2009; Vol. 667, pp 111-123.
28. Baldridge, J. R.; Crane, R. T., *Methods* **1999,** *19*, 103-107.
29. Garçon, N.; Di Pasquale, A., *Hum. Vaccines Immunother.* **2017,** *13*, 19-33.
30. Vasilakos, J. P.; Tomai, M. A., *Expert Rev. Vaccines* **2013,** *12*, 809-819.
31. Rodell, C. B.; Arlauckas, S. P.; Cuccarese, M. F.; Garris, C. S.; Li, R.; Ahmed, M. S.; Kohler, R. H.; Pittet, M. J.; Weissleder, R., *Nat. Biomed. Eng.* **2018,** *2*, 578-588.
32. Furukawa, A.; Kamishikiryo, J.; Mori, D.; Toyonaga, K.; Okabe, Y.; Toji, A.; Kanda, R.; Miyake, Y.; Ose, T.; Yamasaki, S.; Maenaka, K., *Proc. Nat. Acad. Sci. U. S. A.* **2013,** *110*, 17438.
33. Feinberg, H.; Jégouzo, S. A. F.; Rowntree, T. J. W.; Guan, Y.; Brash, M. A.; Taylor, M. E.; Weis, W. I.; Drickamer, K., *J. Biol. Chem.* **2013,** *288*, 28457-28465.
34. Zelensky, A. N.; Gready, J. E., *FEBS J.* **2005,** *272*, 6179-6217.
35. Hoving, J. C.; Wilson, G. J.; Brown, G. D., *Cell. Microbiol.* **2014,** *16*, 185-194.
36. Matsumoto, M.; Tanaka, T.; Kaisho, T.; Sanjo, H.; Copeland, N. G.; Gilbert, D. J.; Jenkins, N. A.; Akira, S., *J. Immunol.* **1999,** *163*, 5039.
37. Ishikawa, E.; Ishikawa, T.; Morita, Y. S.; Toyonaga, K.; Yamada, H.; Takeuchi, O.; Kinoshita, T.; Akira, S.; Yoshikai, Y.; Yamasaki, S., *J. Exp. Med.* **2009,** *206*, 2879-2888.
38. Yamasaki, S.; Matsumoto, M.; Takeuchi, O.; Matsuzawa, T.; Ishikawa, E.; Sakuma, M.; Tateno, H.; Uno, J.; Hirabayashi, J.; Mikami, Y.; Takeda, K.; Akira, S.; Saito, T., *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 1897-1902.
39. Yamasaki, S.; Matsumoto, M.; Takeuchi, O.; Matsuzawa, T.; Ishikawa, E.; Sakuma, M.; Tateno, H.; Uno, J.; Hirabayashi, J.; Mikami, Y.; Takeda, K.; Akira, S.; Saito, T., *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 1897-1902.
40. Zajonc, D. M.; Elsliger, M. A.; Teyton, L.; Wilson, I. A., *Nat. Immunol.* **2003,** *4*, 808-15.
41. Williams, S. J., *Front. Immunol.* **2017,** *8*.
42. Yamasaki, S.; Ishikawa, E.; Sakuma, M.; Hara, H.; Ogata, K.; Saito, T., *Nat. Immunol.* **2008,** *9*, 1179-1188.

43. Werninghaus, K.; Babiak, A.; Groß, O.; Hölscher, C.; Dietrich, H.; Agger, E. M.; Mages, J.; Mocsai, A.; Schoenen, H.; Finger, K.; Nimmerjahn, F.; Brown, G. D.; Kirschning, C.; Heit, A.; Andersen, P.; Wagner, H.; Ruland, J.; Lang, R., *J. Exp. Med.* **2009,** *206*, 89-97.
44. Geijtenbeek, T. B. H.; Gringhuis, S. I., *Nat. Rev. Immunol.* **2016,** *16*, 433-448.
45. Geijtenbeek, T. B. H.; Gringhuis, S. I., *Nat. Rev. Immunol.* **2009,** *9*, 465+.
46. Furukawa, A.; Kamishikiryo, J.; Mori, D.; Toyonaga, K.; Okabe, Y.; Toji, A.; Kanda, R.; Miyake, Y.; Ose, T.; Yamasaki, S.; Maenaka, K., *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110*, 17438-17443.
47. Feinberg, H.; Rambaruth, N. D. S.; Jégouzo, S. A. F.; Jacobsen, K. M.; Djurhuus, R.; Poulsen, T. B.; Weis, W. I.; Taylor, M. E.; Drickamer, K., *J. Biol. Chem.* **2016,** *291*, 21222-21233.
48. Watanabe, M.; Aoyagi, Y.; Ridell, M.; Minnikin, D. E., *Microbiology* **2001,** *147*, 1825-1837.
49. Watanabe, M.; Aoyagi, Y.; Mitome, H.; Fujita, T.; Naoki, H.; Ridell, M.; Minnikin, D. E., *Microbiology* **2002,** *148*, 1881-1902.
50. Yang, Y.; Shi, F.; Tao, G.; Wang, X., *J. Microbiol.* **2012,** *50*, 235-240.
51. Collins; Collins, M., *J. Gen. Microbiol.* **1982,** *128*, 129-149.
52. Ishikawa, E.; Ishikawa, T.; Morita, Y. S.; Toyonaga, K.; Yamada, H.; Takeuchi, O.; Kinoshita, T.; Akira, S.; Yoshikai, Y.; Yamasaki, S., *J. Exp. Med.* **2009,** *206*, 2879-2888.
53. Schick, J.; Etschel, P.; Bailo, R.; Ott, L.; Bhatt, A.; Lepenies, B.; Kirschning, C.; Burkovski, A.; Lang, R., *Infect. Immun.* **2017,** *85*, e00075-17.
54. van der Peet, P. L.; Gunawan, C.; Torigoe, S.; Yamasaki, S.; Williams, S. J., *Chem. Comm.* **2015,** *51*, 5100-5103.
55. Tima, H. G.; Al Dulayymi, J. a. R.; Denis, O.; Lehebel, P.; Baols, K. S.; Mohammed, M. O.; L'Homme, L.; Sahb, M. M.; Potemberg, G.; Legrand, S.; Lang, R.; Beyaert, R.; Piette, J.; Baird, M. S.; Huygen, K.; Romano, M., *J. Innate Immun.* **2017,** *9*, 162-180.
56. Khan, A. A.; Chee, S. H.; McLaughlin, R. J.; Harper, J. L.; Kamena, F.; Timmer, M. S. M.; Stocker, B. L., *ChemBioChem* **2011,** *12*, 2572-2576.
57. Stocker, B. L.; Khan, A. A.; Chee, S. H.; Kamena, F.; Timmer, M. S. M., *ChemBioChem* **2014,** *15*, 382-388.

58. Huber, A.; Kallerup, R. S.; Korsholm, K. S.; Franzyk, H.; Lepenies, B.; Christensen, D.; Foged, C.; Lang, R., *Innate Immun.* **2016,** *22*, 405-418.
59. Pimm, M. V.; Baldwin, R. W.; Polonsky, J.; Lederer, E., *Int. J. Cancer* **1979,** *24*, 780-785.
60. van der Peet, P. L.; Nagata, M.; Shah, S.; White, J. M.; Yamasaki, S.; Williams, S. J., *Org. Biomol. Chem.* **2016,** *14*, 9267-9277.
61. Decout, A.; Silva-Gomes, S.; Drocourt, D.; Barbe, S.; André, I.; Cueto, F. J.; Lioux, T.; Sancho, D.; Pérouzel, E.; Vercellone, A.; Prandi, J.; Gilleron, M.; Tiraby, G.; Nigou, J., *Proc. Nat. Acad. Sci. U. S. A.* **2017,** *114*, 2675.
62. Simons, K.; Vaz, W. L. C., *Annu. Rev. Biophys. Biomol. Struct.* **2004,** *33*, 269-295.
63. Tabas, I., *J. Clin. Invest.* **2002,** *110*, 905-911.
64. Ory, D. S., *Have the Orphans Found a Home?* **2004,** *95*, 660-670.
65. Lusis, A. J., *Nature* **2000,** *407*, 233.
66. Duewell, P.; Kono, H.; Rayner, K. J.; Sirois, C. M.; Vladimer, G.; Bauernfeind, F. G.; Abela, G. S.; Franchi, L.; Nuñez, G.; Schnurr, M.; Espevik, T.; Lien, E.; Fitzgerald, K. A.; Rock, K. L.; Moore, K. J.; Wright, S. D.; Hornung, V.; Latz, E., *Nature* **2010,** *464*, 1357.
67. Lim, R. S.; Suhalim, J. L.; Miyazaki-Anzai, S.; Miyazaki, M.; Levi, M.; Potma, E. O.; Tromberg, B. J., *J. Lipid Res.* **2011,** *52*, 2177-2186.
68. Kiyotake, R.; Oh-hora, M.; Ishikawa, E.; Miyamoto, T.; Ishibashi, T.; Yamasaki, S., *J. Biol. Chem.* **2015,** *290*, 25322-25332.
69. Rearick, J. I.; Hesterberg, T. W.; Jetten, A. M., *J. Cell. Physiol.* **1987,** *133*, 573-578.
70. Rearick, J. I.; Stoner, G. D.; George, M. A.; Jetten, A. M., *Cancer Res.* **1988,** *48*, 5289.
71. Strott, C. A.; Higashi, Y., *J. Lipid Res.* **2003,** *44*, 1268-78.
72. Kostarnoy, A. V.; Gancheva, P. G.; Lepenies, B.; Tukhvatulin, A. I.; Dzharullaeva, A. S.; Polyakov, N. B.; Grumov, D. A.; Egorova, D. A.; Kulibin, A. Y.; Bobrov, M. A.; Malolina, E. A.; Zykin, P. A.; Soloviev, A. I.; Riabenko, E.; Maltseva, D. V.; Sakharov, D. A.; Tonevitsky, A. G.; Verkhovskaya, L. V.; Logunov, D. Y.; Naroditsky, B. S.; Gintsburg, A. L., *Proc. Natl. Acad. Sci. U. S. A.* **2017,** *114*, E2758-E2765.

73. Ishikawa, T.; Itoh, F.; Yoshida, S.; Saijo, S.; Matsuzawa, T.; Gonoi, T.; Saito, T.; Okawa, Y.; Shibata, N.; Miyamoto, T.; Yamasaki, S., *Cell Host Microbe* **2013,** *13*, 477-488.
74. Richardson, M. B.; Torigoe, S.; Yamasaki, S.; Williams, S. J., *Chem. Commun.* **2015,** *51*, 15027-15030.
75. van der Peet, P. L.; Gunawan, C.; Watanabe, M.; Yamasaki, S.; Williams, S. J., *J. Org. Chem.* **2019,** *84*, 6788-6797.
76. Wells, C. A.; Salvage-Jones, J. A.; Li, X.; Hitchens, K.; Butcher, S.; Murray, R. Z.; Beckhouse, A. G.; Lo, Y.-L.-S.; Manzanero, S.; Cobbold, C.; Schroder, K.; Ma, B.; Orr, S.; Stewart, L.; Lebus, D.; Sobieszczuk, P.; Hume, D. A.; Stow, J.; Blanchard, H.; Ashman, R. B., *J. Immunol.* **2008,** *180*, 7404-7413.
77. Glusman, G.; Rowen, L.; Lee, I.; Boysen, C.; Roach, J. C.; Smit, A. F. A.; Wang, K.; Koop, B. F.; Hood, L., *Immunity* **2001,** *15*, 337-349.
78. Felix, N. J.; Allen, P. M., *Nat. Rev. Immunol.* **2007,** *7*, 942-953.
79. Danska, J. S.; Livingstone, A. M.; Paragas, V.; Ishihara, T.; Fathman, C. G., *J. Exp. Med.* **1990,** *172*, 27-33.
80. Sewell, A. K., *Nat. Rev. Immunol.* **2012,** *12*, 669-677.
81. Godfrey, D. I.; Kronenberg, M., *J. Clin. Invest.* **2004,** *114*, 1379-1388.
82. Godo, M.; Sessler, T.; Hamar, P., *Curr. Med. Chem.* **2008,** *15*, 1778-1787.
83. Pilones, K. A.; Aryankalayil, J.; Demaria, S., *Clin. Dev. Immunol.* **2012,** *2012*, 11.
84. Makino, Y.; Kanno, R.; Ito, T.; Higashino, K.; Taniguchi, M., *Int. Immunol.* **1995,** *7*, 1157-1161.
85. Kronenberg, M.; Rudensky, A., *Nature* **2005,** *435*, 598-604.
86. Kronenberg, M., Toward an understanding of NKT cell biology: Progress and paradoxes. In *Annu. Rev. Immunol.*, Annual Reviews: Palo Alto, 2005; Vol. 23, pp 877-900.
87. Luc Van, K., *Nat. Rev. Immunol.* **2005,** *5*, 31-42.
88. Yu, K. O. A.; Porcelli, S. A., *Immunol. Lett.* **2005,** *100*, 42-55.
89. Godfrey, D. I.; Uldrich, A. P.; McCluskey, J.; Rossjohn, J.; Moody, D. B., *Nat. Immunol.* **2015,** *16*, 1114.
90. Kumar, V.; Delovitch, T. L., *Immunology* **2014,** *142*, 321-336.
91. Kronenberg, M.; Gapin, L., *Nat. Rev. Immunol.* **2002,** *2*, 557.
92. Gapin, L., *Nat. Rev. Immunol.* **2010,** *10*, 272-277.

93. Natori, T.; Morita, M.; Akimoto, K.; Koezuka, Y., *Tetrahedron* **1994,** *50*, 2771-2784.
94. Morita, M.; Motoki, K.; Akimoto, K.; Natori, T.; Sakai, T.; Sawa, E.; Yamaji, K.; Koezuka, Y.; Kobayashi, E.; Fukushima, H., *J. Med. Chem.* **1995,** *38*, 2176-2187.
95. Morita, M.; Natori, T.; Akimoto, K.; Osawa, T.; Fukushima, H.; Koezuka, Y., *Bioorganic Med. Chem. Lett.* **1995,** *5*, 699-704.
96. Yu, K. O. A.; Im, J. S.; Illarionov, P. A.; Ndonye, R. M.; Howell, A. R.; Besra, G. S.; Porcelli, S. A., *J. Immunol. Methods* **2007,** *323*, 11-23.
97. Kawano, T.; Cui, J.; Koezuka, Y.; Toura, I.; Kaneko, Y.; Motoki, K.; Ueno, H.; Nakagawa, R.; Sato, H.; Kondo, E.; Koseki, H.; Taniguchi, M., *Science* **1997,** *278*, 1626.
98. Banchet-Cadeddu, A.; Henon, E.; Dauchez, M.; Renault, J.-H.; Monneaux, F.; Haudrechy, A., *Org. Biomol. Chem.* **2011,** *9*, 3080-3104.
99. Sidobre, S.; Hammond, K. J. L.; Bénazet-Sidobre, L.; Maltsev, S. D.; Richardson, S. K.; Ndonye, R. M.; Howell, A. R.; Sakai, T.; Besra, G. S.; Porcelli, S. A.; Kronenberg, M., *Proc. Nat. Acad. Sci. U. S. A.* **2004,** *101*, 12254.
100. Koch, M.; Stronge, V. S.; Shepherd, D.; Gadola, S. D.; Mathew, B.; Ritter, G.; Fersht, A. R.; Besra, G. S.; Schmidt, R. R.; Jones, E. Y.; Cerundolo, V., *Nat. Immunol.* **2005,** *6*, 819-826.
101. Borg, N. A.; Wun, K. S.; Kjer-Nielsen, L.; Wilce, M. C. J.; Pellicci, D. G.; Koh, R.; Besra, G. S.; Bharadwaj, M.; Godfrey, D. I.; McCluskey, J.; Rossjohn, J., *Nature* **2007,** *448*, 44-49.
102. Van Rhijn, I.; Godfrey, D. I.; Rossjohn, J.; Moody, D. B., *Nat. Rev. Immunol.* **2015,** *15*, 643-654.
103. Zhou, D.; Mattner, J.; Cantu, C.; Schrantz, N.; Yin, N.; Gao, Y.; Sagiv, Y.; Hudspeth, K.; Wu, Y.-P.; Yamashita, T.; Teneberg, S.; Wang, D.; Proia, R.; Levery, S.; Savage, P.; Teyton, L.; Bendelac, A., *Science* **2004,** *306*, 1786-1789.
104. Parekh, V.; Singh, A.; Wilson, M.; Olivares Villagómez, D.; Bezbradica, J.; Inazawa, H.; Ehara, H.; Sakai, T.; Serizawa, I.; Wu, L.; Wang, C.-R.; Joyce, S.; Van Kaer, L., *J. Immunol.* **2004,** *173*, 3693-3706.
105. Ortaldo, J.; Young, H.; Winkler Pickett, R.; Bere, E.; Murphy, W.; Wiltrout, R. H., *J. Immunol.* **2004,** *172*, 943-953.
106. Wu , D. Y.; Segal , N. H.; Sidobre , S.; Kronenberg , M.; Chapman , P. B., *J. Exp. Med.* **2003,** *198*, 173-181.

107. Kinjo, Y.; Wu, D.; Kim, G.; Xing, G.-W.; Poles, M.; Ho, D.; Tsuji, M.; Kawahara, K.; Wong, C.-H.; Kronenberg, M., *Nature* **2005,** *434*, 520-525.
108. Mattner, J.; DeBord, K.; Ismail, N.; Goff, R.; Cantu, C.; Zhou, D.; Saint Mezard, P.; Wang, V.; Gao, Y.; Yin, N.; Hoebe, K.; Schneewind, O.; Walker, D.; Beutler, B.; Teyton, L.; Savage, P.; Bendelac, A., *Nature* **2005,** *434*, 525-529.
109. Sriram, V.; Du, W.; Gervay Hague, J.; Brutkiewicz, R., *Eur. J. Immunol.* **2005,** *35*, 1692-1701.
110. Kinjo, Y.; Tupin, E.; Wu, D.; Fujio, M.; Garcia-Navarro, R.; Mohammed Rafii-El-Idrissi, B.; Zajonc, D. M.; Ben-Menachem, G.; Ainge, G. D.; Painter, G. F.; Khurana, A.; Kasper, H.; Behar, S. M.; Beutler, B.; Wilson, I. A., *Nat. Immunol.* **2006,** *7*, 978-86.
111. Brown, L. C. W.; Penaranda, C.; Kashyap, P. C.; Williams, B. B.; Clardy, J.; Kronenberg, M.; Sonnenburg, J. L.; Comstock, L. E.; Bluestone, J. A.; Fischbach, M. A., *PLOS Biol.* **2013,** *11*, 9.
112. An, D. D.; Oh, S. F.; Olszak, T.; Neves, J. F.; Avci, F. Y.; Erturk-Hasdemir, D.; Lu, X.; Zeissig, S.; Blumberg, R. S.; Kasper, D. L., *Cell* **2014,** *156*, 123-133.
113. Savage, P. B.; Teyton, L.; Bendelac, A., *Chem. Soc. Rev.* **2006,** *35*, 771-779.
114. Treiner, E.; Duban, L.; Bahram, S.; Radosavljevic, M.; Wanner, V.; Tilloy, F.; Affaticati, P.; Gilfillan, S.; Lantz, O., *Nature* **2003,** *422*, 164-169.
115. Le Bourhis, L.; Guerri, L.; Dusseaux, M.; Martin, E.; Soudais, C.; Lantz, O., *Trends Immunol.* **2011,** *32*, 212-218.
116. Reantragoon, R.; Kjer-Nielsen, L.; Patel, O.; Chen, Z.; Illing, P. T.; Bhati, M.; Kostenko, L.; Bharadwaj, M.; Meehan, B.; Hansen, T. H.; Godfrey, D. I.; Rossjohn, J.; McCluskey, J., *J. Exp. Med.* **2012,** *209*, 761-774.
117. Godfrey, D.; Rossjohn, J.; McCluskey, J., *Nat. Immunol.* **2010,** *11*, 693-695.
118. Dusseaux, M.; Martin, E.; Serriari, N.; Péguillet, I.; Premel, V.; Louis, D.; Milder, M.; Le Bourhis, L.; Soudais, C.; Treiner, E.; Lantz, O., *Blood* **2011,** *117*, 1250-1259.
119. Kjer-Nielsen, L.; Patel, O.; Corbett, A. J.; Le Nours, J.; Meehan, B.; Liu, L.; Bhati, M.; Chen, Z.; Kostenko, L.; Reantragoon, R.; Williamson, N. A.; Purcell, A. W.; Dudek, N. L.; McConville, M. J.; apos; Hair, R. A. J.; Khairallah, G. N.; Godfrey, D. I.; Fairlie, D. P.; Rossjohn, J.; McCluskey, J., *Nature* **2012,** *491*, 717+.

120. Gold, M.; Cerri, S.; Smyk Pearson, S.; Cansler, M.; Vogt, T.; Delepine, J.; Winata, E.; Swarbrick, G.; Chua, W.-J.; Yu, Y. Y. L.; Lantz, O.; Cook, M.; Null, M.; Jacoby, D.; Harriff, M.; Lewinsohn, D.; Hansen, T.; Marrack, P., *PLOS Biol.* **2010,** *8*, e1000407.
121. Le Bourhis, L.; Martin, E.; Péguillet, I.; Guihot, A.; Froux, N.; Coré, M.; Lévy, E.; Dusseaux, M.; Meyssonnier, V.; Premel, V.; Ngo, C.; Riteau, B.; Duban, L.; Robert, D.; Huang, S.; Rottman, M.; Soudais, C.; Lantz, O., *Nat. Immunol.* **2010,** *11*, 701-708.
122. Gold, M. C.; Cerri, S.; Smyk-Pearson, S.; Cansler, M. E.; Vogt, T. M.; Delepine, J.; Winata, E.; Swarbrick, G. M.; Chua, W.-J.; Yu, Y. Y. L.; Lantz, O.; Cook, M. S.; Null, M. D.; Jacoby, D. B.; Harriff, M. J.; Lewinsohn, D. A.; Hansen, T. H.; Lewinsohn, D. M., *PLOS Biol.* **2010,** *8*, e1000407.
123. Miyazaki, Y.; Miyake, S.; Chiba, A.; Lantz, O.; Yamamura, T., *Int. Immunol.* **2011,** *23*, 529-535.
124. Chiba, A.; Tajima, R.; Tomi, C.; Miyazaki, Y.; Yamamura, T.; Miyake, S., *Arthritis Rheum.* **2012,** *64*, 153-161.
125. Laurent, G.; Gapin, L.; Gapin, L., *PLOS Biol.* **2009,** *7*, e1000070.
126. Hinks, T. S. C., *Immunology* **2016,** *148*, 1-12.
127. Kawachi, I.; Maldonado, J.; Strader, C.; Gilfillan, S., *J. Immunol.* **2006,** *176*, 1618.
128. Bianchini, E.; De Biasi, S.; Simone, A. M.; Ferraro, D.; Sola, P.; Cossarizza, A.; Pinti, M., *Immunol. Lett.* **2017,** *183*, 1-7.
129. Treiner, E.; Duban, L.; Moura, I. C.; Hansen, T.; Gilfillan, S.; Lantz, O., *Microb. Infect.* **2005,** *7*, 552-559.
130. De Libero, G.; Singhal, A.; Lepore, M.; Mori, L., *Cold Spring Harb. Perspect. Med.* **2014,** *4*.
131. Bendelac, A.; Savage, P.; Teyton, L., *Ann. Rev. Immunol.* **2007,** *25*, 297-336.
132. Godfrey, D. I.; Pellicci, D. G.; Patel, O.; Kjer-Nielsen, L.; McCluskey, J.; Rossjohn, J., *Semin. Immunol.* **2010,** *22*, 61-67.
133. Tilloy, F.; Treiner, E.; Park, S.-H.; Garcia, C.; Lemonnier, F.; de la Salle, H.; Bendelac, A.; Bonneville, M.; Lantz, O., *J. Exp. Med.* **1999,** *189*, 1907-1921.
134. Riegert, P., *J. Immunol.* **1998,** *161*, 4066-4077.
135. Parra, C.; Parra Cuadrado, J. F.; Navarro, P.; Mirones, I.; Setien, F.; Oteo, M.; Martinez-Naves, E., *Tissue Antigens* **2000,** *56*, 170-172.

136. Corbett, A. J.; Eckle, S. B. G.; Birkinshaw, R. W.; Liu, L.; Patel, O.; Mahony, J.; Chen, Z.; Reantragoon, R.; Meehan, B.; Cao, H.; Williamson, N. A.; Strugnell, R. A.; Van Sinderen, D.; Mak, J. Y. W.; Fairlie, D. P.; Kjer-Nielsen, L.; Rossjohn, J.; McCluskey, J., *Nature* **2014,** *509*, 361.
137. Patel, O.; Kjer-Nielsen, L.; Le Nours, J.; Eckle, S. B. G.; Birkinshaw, R.; Beddoe, T.; Corbett, A. J.; Liu, L.; Miles, J. J.; Meehan, B.; Reantragoon, R.; Sandoval-Romero, M. L.; Sullivan, L. C.; Brooks, A. G.; Chen, Z.; Fairlie, D. P.; McCluskey, J.; Rossjohn, J., *Nat. Commun.* **2013,** *4*, 2142.
138. Eckle, S. B. G.; Birkinshaw, R. W.; Kostenko, L.; Corbett, A. J.; McWilliam, H. E. G.; Reantragoon, R.; Chen, Z.; Gherardin, N. A.; Beddoe, T.; Liu, L.; Patel, O.; Meehan, B.; Fairlie, D. P.; Villadangos, J. A.; Godfrey, D. I.; Kjer-Nielsen, L.; McCluskey, J.; Rossjohn, J., *J. Exp. Med.* **2014,** *211*, 1585-1600.
139. Hansen, T.; Huang, S.; Arnold, P.; Fremont, D. H., *Nat. Immunol.* **2007,** *8*, 563-568.
140. Huang, S.; Gilfillan, S.; Cella, M.; Miley, M.; Lantz, O.; Lybarger, L.; Fremont, D.; Hansen, T. H., *J. Biol. Chem.* **2005,** *280*, 21183-21193.
141. Huang, S.; Martin, E.; Kim, S.; Yu, L.; Soudais, C.; Fremont, D. H.; Lantz, O.; Hansen, T. H., *Proc. Nat. Acad. Sci. U. S. A.* **2009,** *106*, 8290-8295.
142. Nick, G.; Goldfinch, N.; Goldfinch, N.; Reinink, P.; Connelley, T.; Koets, A.; Morrison, I.; Van Rhijn, I., *Vet. Res.* **2010,** *41*, 62.
143. Porcelli, S.; Yockey, C. E.; Brenner, M. B.; Balk, S. P., *J. Exp. Med.* **1993,** *178*, 1-16.
144. Hundhausen, T.; Laus, R.; Müller-Ruchholtz, W., *J. Immunol. Methods* **1992,** *153*, 21-29.
145. Gras, S.; Chen, Z.; Miles, J. J.; Liu, Y. C.; Bell, M. J.; Sullivan, L. C.; Kjer-Nielsen, L.; Brennan, R. M.; Burrows, J. M.; Neller, M. A.; Khanna, R.; Purcell, A. W.; Brooks, A. G.; McCluskey, J.; Rossjohn, J.; Burrows, S. R., *J. Exp. Med.* **2010,** *207*, 1555-1567.
146. Gillis, S.; Watson, J., *J. Exp. Med.* **1980,** *152*, 1709-1719.
147. Kjer-Nielsen, L.; Corbett, A. J.; Chen, Z.; Liu, L.; Mak, J. Y.; Godfrey, D. I.; Rossjohn, J.; Fairlie, D. P.; McCluskey, J.; Eckle, S. B., *Immunol. Cell Biol.* **2018,** *96*, 573-587.
148. Shimamura, M.; Huang, Y.-Y.; Okamoto, N.; Suzuki, N.; Yasuoka, J.; Morita, K.; Nishiyama, A.; Amano, Y.; Mishina, T., *Eur. J. Immunol.* **2007,** *37*, 1836-1844.

149. Awad, W.; Le Nours, J.; Kjer-Nielsen, L.; McCluskey, J.; Rossjohn, J., *Immunol. Cell Biol.* **2018,** *96*, 588-597.
150. Bacher, A.; Bacher; Eberhardt, S.; Fischer, M.; Kis, K.; Richter, G. A., *Annu. Rev. Nutr.* **2000,** *20*, 153-167.
151. Kis, K.; Kugelbrey, K.; Bacher, A., *J. Org. Chem.* **2001,** *66*, 2555-2559.
152. Cushman, M.; Yang, D.; Gerhardt, S.; Huber, R.; Fischer, M.; Kis, K.; Bacher, A., *J. Org. Chem.* **2002,** *67*, 5807-5816.
153. Yu, W.; Wang, Y.; Wang, Y.; Ho, C.-T., *Chem. Soc. Rev.* **2012,** *41*, 4140-4149.
154. Douglas, L. M.; Konopka, J. B., *Annu. Rev. Microbiol.* **2014,** *68*, 377-393.
155. Chick, H.; Dalyell, E.; Hume, M.; Smith, H. H.; Mackay, H. M., *The Lancet* **1922,** *200*, 7-11.
156. Katzung, B. G.; Trevor, A. J., *Basic & clinical pharmacology. [electronic resource]*. 13th ed. ed.; McGraw-Hill Education LLC.: 2015.
157. Dowd, F. J.; Dowd, F. J.; Johnson, B. S.; Mariotti, A. J., *Pharmacology and therapeutics for dentistry*. 7th edition. ed.; Elsevier: 2017.
158. Yasukawa, K.; Aoki, T.; Takido, M.; Ikekawa, T.; Saito, H.; Matsuzawa, T., *Phytother. Res.* **1994,** *8*, 10-13.
159. Kobori, M.; Yoshida, M.; Ohnishi-Kameyama, M.; Shinmoto, H., *Brit. J. Pharmacol.* **2007,** *150*, 209-19.
160. Kuo, C.-F.; Hsieh, C.-H.; Lin, W.-Y., *Food Chem.* **2011,** *126*, 207-212.
161. Caroprese, M.; Albenzio, M.; Ciliberti, M. G.; Francavilla, M.; Sevi, A., *Vet. Immunol. Immunopathol.* **2012,** *150*, 27-35.
162. Warnecke, D.; Erdmann, R.; Fahl, A.; Hube, B.; Müller, F.; Zank, T.; Zähringer, U.; Heinz, E., *J. Biol. Chem.* **1999,** *274*, 13048-13059.
163. Watanabe, T.; Ito, T.; Goda, H.; Ishibashi, Y.; Miyamoto, T.; Ikeda, K.; Taguchi, R.; Okino, N.; Ito, M., *J. Biol. Chem.* **2015,** *290*, 1005-1019.
164. Park, B. J.; Wannemuehler, K. A.; Marston, B. J.; Govender, N.; Pappas, P. G.; Chiller, T. M., *AIDS* **2009,** *23*.
165. Rella, A.; Mor, V.; Farnoud, A. M.; Singh, A.; Shamseddine, A. A.; Ivanova, E.; Carpino, N.; Montagna, M. T.; Luberto, C.; Del Poeta, M., *Front. microbiol.* **2015,** *6*, 836-836.
166. Li, H., *Endocrinology* **1998,** *139*, 4991-4997.
167. Epand, R., *Prog. Lipid. Res.* **2006,** *45*, 279-294.
168. Gimpl, G., *Subcel. Biochem.* **2010,** *51*, 1-45.

169. Furukawa, A.; Kamishikiryo, J.; Mori, D.; Toyonaga, K.; Okabe, Y.; Toji, A.; Kanda, R.; Miyake, Y.; Ose, T.; Yamasaki, S.; Maenaka, K., *Proc. Nat. Acad. Sci. U. S. A.* **2013,** *110*, 17438-17443.
170. Jégouzo, S. A. F.; Harding, E. C.; Acton, O.; Rex, M. J.; Fadden, A. J.; Taylor, M. E.; Drickamer, K., *Glycobiology* **2014,** *24*, 1291-1300.
171. Watts, G., *Br. Med. J.* **2005,** *331*, 795-795.
172. Robin Warren, J.; Marshall, B., *The Lancet* **1983,** *321*, 1273-1275.
173. Marshall, B.; Warren, J. R., *The Lancet* **1984,** *323*, 1311-1315.
174. Blaser, M. J., *Indigenous microbes and the ecology of human diseases* **2006,** *7*, 956-960.
175. Hirai, Y.; Haque, M.; Yoshida, T.; Yokota, K.; Yasuda, T.; Oguma, K., *J. Bacteriol.* **1995,** *177*, 5327-33.
176. Haque, M.; Hirai, Y.; Yokota, K.; Oguma, K., *J. Bacteriol.* **1995,** *177*, 5334.
177. Haque, M.; Hirai, Y.; Yokota, K.; Mori, N.; Jahan, I.; Ito, H.; Hotta, H.; Yano, I.; Kanemasa, Y.; Oguma, K., *J. Bacteriol.* **1996,** *178*, 2065.
178. Presser, A.; Kunert, O.; Pötschger, I., *Monatshefte für Chemie* **2006,** *137*, 365-374.
179. Nguyen, H. Q.; Davis, R. A.; Gervay-Hague, J., *Angew. Chem. Int. Ed.* **2014,** *53*, 13400-13403.
180. Davis, R. A.; Fettinger, J. C.; Gervay-Hague, J., *J. Org. Chem.* **2014,** *79*, 8447-8452.
181. Manolakaki, D.; Velmahos, G.; Kourkoumpetis, T.; Chang, Y.; Alam, H. B.; De Moya, M. M.; Mylonakis, E., *Virulence* **2010,** *1*, 367-375.
182. Jespersen, L.; Nielsen, D.; Honholt, S.; Jakobsen, M. L., *FEMS Yeast Res.* **2005,** *5*, 441-453.
183. Gow, N. A. R.; Yadav, B., *Microbiology* **2017,** *163*, 1145-1147.
184. Kourkoumpetis, T. K.; Velmahos, G. C.; Ziakas, P. D.; Tampakakis, E.; Manolakaki, D.; Coleman, J. J.; Mylonakis, E., *Mycopathologia* **2011,** *171*, 85-91.
185. Erdogan, A.; Rao, S. S. C., *Curr. Gastroenterol. Rep.* **2015,** *17*, 16.
186. Pfaller, M. A.; Diekema, D. J., *Clin. Microbiol. Rev.* **2007,** *20*, 133.
187. Mille, C.; Janbon, G.; Delplace, F.; Ibata-Ombetta, S.; Gaillardin, C.; Strecker, G.; Jouault, T.; Trinel, P.-A.; Poulain, D., *J. Biol. Chem.* **2004,** *279*, 47952-47960.
188. López Ribot, J. L.; Martínez, J. P.; Monteagudo, C.; Alloush, H. M.; Mattioli, N. V.; Chaffin, W. L., *Rev. iberoam. Micol.* **1999,** *16*, 23-26.

189. Ballmann, G. E.; Chaffin, W. L., *Mycopathologia* **1979,** *67*, 39-43.
190. Bianchi, D. E., *Antonie van Leeuwenhoek* **1967,** *33*, 324-332.
191. S, M.; Marriott, M. S.; Marriott, M. S., *J. Gen. Microbiol.* **1975,** *86*, 115-132.
192. Yano, K.; Yamada, T.; Banno, Y.; Sekiya, T.; Nozawa, Y., *Jpn. J. Med. Mycol.* **1982,** *23*, 159-165.
193. Ghannoum, M.; Ghannoum, *J. Gen. Microbiol.* **1986,** *132*, 2367-2375.
194. Radwan; Radwan, S. S.; Spener, F.; Mangold, H. K.; Staba, E. J., *Chem. Phys. Lipids.* **1975,** *14*, 72-80.
195. Thompson, E. D.; Knights, B. A.; Parks, L. W., *Biochim. Biophys. Acta.* **1973,** *304*, 132-141.
196. Chang, Y.-J.; Kim, H. Y.; Albacker, L. A.; Lee, H. H.; Baumgarth, N.; Akira, S.; Savage, P. B.; Endo, S.; Yamamura, T.; Maaskant, J.; Kitano, N.; Singh, A.; Bhatt, A.; Besra, G. S.; den Elzen, P. v.; Appelmelk, B.; Franck, R. W.; Chen, G.; DeKruyff, R. H.; Shimamura, M.; Illarionov, P.; Umetsu, D. T., *J. Clin. Invest.* **2011,** *121*, 57+.
197. Gamian, A.; Mordarska, H.; Ekiel, I.; Ulrich, J.; Szponar, B.; Defaye, J., *Carbohydr. Res.* **1996,** *296*, 55-67.
198. Grille, S.; Zaslawski, A.; Thiele, S.; Plat, J.; Warnecke, D., *Prog. Lipid. Res.* **2010,** *49*, 262-288.
199. Shimamura, M.; Yamamura, M.; Nabeshima, T.; Kitano, N.; van den Elzen, P.; Yesilkaya, H.; Andrew, P.; Illarionov, P., *Sci. Rep.* **2017,** *7*, 9703.
200. Capon, B., *Chem. Rev.* **1969,** *69*, 407-498.
201. Domańska, U.; Klofutar, C.; Paljk, Š., *Fluid Phase Equilibr.* **1994,** *97*, 191-200.
202. Kritchevsky, D., *Proc. Soc. Biol. Med.* **1964,** *116*, 104-107.
203. Flynn, G. L.; Shah, Y.; Prakongpan, S.; Kwan, K. H.; Higuchi, W. I.; Hofmann, A. F., *J. Pharm. Sci.* **1979,** *68*, 1090-1097.
204. Weichherz; Weichherz, J., *Biochemische Zeitschrift* **1932,** *249*, 312-322.
205. Bar, L. K.; Garti, N.; Sarig, S.; Bar, R., *J. Chem. Eng. Data.* **1984,** *29*, 440-443.
206. McGill, N. W.; Williams, S. J., *J. Org. Chem.* **2009,** *74*, 9388-9398.
207. Hettikankanamalage, A. A.; Lassfolk, R.; Ekholm, F. S.; Leino, R.; Crich, D., *Chem. Rev.* **2020,** *120*, 7104-7151.
208. Smith, D. G. M.; Hosono, Y.; Nagata, M.; Yamasaki, S.; Williams, S. J., *Chem. Comm.* **2020,** *56*, 4292-4295.
209. Still, W. C.; Kahn, M.; Mitra, A., *J. Org. Chem.* **1978,** *43*, 2923-2925.

210. Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J., *Organometallics* **1996,** *15*, 1518-1520.
211. Xu, J.; Gordon, J. I., *Proc. Nat. Acad. Sci. U. S. A.* **2003,** *100*, 10452-10459.
212. Reid, G., *Clin. Infect. Dis.* **2004,** *39*, 827-830.
213. Salyers, A. A., *Annu. Rev. Microbiol.* **1984,** *38*, 293-313.
214. Russo, T. A.; Thompson, J. S.; Godoy, V. G.; Malamy, M. H., *J. Bacteriol.* **1990,** *172*, 2594-2600.
215. Tzianabos, A. O.; Kasper, D. L.; Onderdonk, A. B., *Clin. Infect. Dis.* **1995,** *20*, S132-S140.
216. Aucher, P.; Saunier, J. P.; Grollier, G.; Sebald, M.; Fauchère, J. L., *Eur. J. Clin. Microbiol. Infect. Dis.* **1996,** *15*, 820-823.
217. Feuillet, L.; Carvajal, J.; Sudre, I.; Pelletier, J.; Thomassin, J. M.; Drancourt, M.; Cherif, A. A., *J. Clin. Microbiol.* **2005,** *43*, 1467-1469.
218. Montejo Baranda, M.; Gonzalez de Zarate Apiñamiz, P.; Barron Fernandez, J.; Aguirre Errasti, C., *Can. Med. Assoc. J.* **1984,** *131*, 184-186.
219. Brook, I., *Pediatr. Rehabil.* **2002,** *5*, 171-176.
220. Hooper, L. V.; Midtvedt, T.; Gordon, J. I., *Annu. Rev. Nutr.* **2002,** *22*, 283-307.
221. Troy, E. B.; Kasper, D. L., *Front. Biosci.* **2010,** *15*, 25-34.
222. Dasgupta, S.; Erturk-Hasdemir, D.; Ochoa-Reparaz, J.; Reinecker, H.-C.; Kasper, D. L., *Cell host microbe.* **2014,** *15*, 413-423.
223. Wang, Y.; Telesford, K. M.; Ochoa-Repáraz, J.; Haque-Begum, S.; Christy, M.; Kasper, E. J.; Wang, L.; Wu, Y.; Robson, S. C.; Kasper, D. L.; Kasper, L. H., *Nat. Commun.* **2014,** *5*, 4432-4432.
224. Shen, Y.; Giardino Torchia, M. L.; Lawson, G. W.; Karp, C. L.; Ashwell, J. D.; Mazmanian, S. K., *Cell host microbe.* **2012,** *12*, 509-520.
225. Johnson, J. L.; Jones, M. B.; Cobb, B. A., *J. Biol. Chem.* **2015,** *290*, 5007-5014.
226. Deng, Z.; Mu, J.; Tseng, M.; Wattenberg, B.; Zhuang, X.; Egilmez, N. K.; Wang, Q.; Zhang, L.; Norris, J.; Guo, H.; Yan, J.; Haribabu, B.; Miller, D.; Zhang, H.-G., *Nat. Commun.* **2015,** *6*, 6956-6956.
227. So, R. C.; Ndonye, R.; Izmirian, D. P.; Richardson, S. K.; Guerrera, R. L.; Howell, A. R., *J. Org. Chem.* **2004,** *69*, 3233-3235.
228. Chênevert, R.; Simard, M.; Bergeron, J.; Dasser, M., *Tetrahedron: Asymmetry* **2004,** *15*, 1889-1892.
229. Miyagawa, E.; Azuma, R.; Suto, T.; Yano, I., *J. Biochem.* **1979,** *86*, 311-320.

230. Kuwahara, T.; Yamashita, A.; Hirakawa, H.; Nakayama, H.; Toh, H.; Okada, N.; Kuhara, S.; Hattori, M.; Hayashi, T.; Ohnishi, Y., *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101*, 14919-14924.
231. Savage, P. B.; Teyton, L.; Bendelac, A., *Chem. Soc. Rev.* **2006,** *35*, 771-779.
232. Du, W.; Gervay-Hague, J., *Org. Lett.* **2005,** *7*, 2063-2065.
233. Ndonye, R. M.; Izmirian, D. P.; Dunn, M. F.; Yu, K. O. A.; Porcelli, S. A.; Khurana, A.; Kronenberg, M.; Richardson, S. K.; Howell, A. R., *J. Org. Chem.* **2005,** *70*, 10260-10270.
234. Banchet-Cadeddu, A.; Hénon, E.; Dauchez, M.; Renault, J.-H.; Monneaux, F.; Haudrechy, A., *Org. Biomol. Chem.* **2011,** *9*, 3080-3104.
235. Morales-Serna, J. A.; Boutureira, O.; Díaz, Y.; Matheu, M. I.; Castillón, S., *Carbohydr. Res.* **2007,** *342*, 1595-1612.
236. Akimoto, K.; Natori, T.; Morita, M., *Tetrahedron Lett.* **1993,** *34*, 5593-5596.
237. Takikawa, H.; Nozawa, D.; Kayo, A.; Muto, S.-e.; Mori, K., *J. Chem. Soc. Perkin Trans. 1* **1999**, 2467-2477.
238. Takikawa, H.; Muto, S.; Nozawa, D.; Kayo, A.; Mori, K., *Tetrahedron Lett.* **1998,** *39*, 6931-6934.
239. Chong, J. M.; Heuft, M. A.; Rabbat, P., *J. Org. Chem.* **2000,** *65*, 5837-5838.
240. Herriott, A. W.; Picker, D., *Tetrahedron Lett.* **1974**, 1511-1514.
241. Clay, R. J.; Collom, T. A.; Karrick, G. L.; Wemple, J., *Synthesis-Stuttgart* **1993**, 290-292.
242. Ikariya, T.; Ishii, Y.; Kawano, H.; Arai, T.; Saburi, M.; Yoshikawa, S.; Akutagawa, S., *J. Chem. Soc. Chem. Commun.* **1985**, 922-924.
243. Noyori, R.; Ohkuma, T.; Kitamura, M.; Takaya, H.; Sayo, N.; Kumobayashi, H.; Akutagawa, S., *J. Am. Chem. Soc.* **1987,** *109*, 5856-5858.
244. Mashima, K.; Kusano, K.-h.; Ohta, T.; Noyori, R.; Takaya, H., *J. Chem. Soc. Chem. Commun.* **1989**, 1208-1210.
245. Girard, C.; Genêt, J.-P.; Bulliard, M., *Eur. J. Org. Chem.* **1999,** *1999*, 2937-2942.
246. King, S. A.; Thompson, A. S.; King, A. O.; Verhoeven, T. R., *J. Org. Chem.* **1992,** *57*, 6689-6691.
247. Wolfson, A.; Vankelecom, I. F. J.; Geresh, S.; Jacobs, P. A., *J. Mol. Catal. A-Chem* **2004,** *217*, 21-26.
248. Hoye, T. R.; Jeffrey, C. S.; Shao, F., *Nat. Protoc.* **2007,** *2*, 2451-2458.
249. Hoye, T. R.; Jeffrey, C. S.; Shao, F., *Nat. Protoc.* **2007,** *2*, 2451-2458.

250. Dale, J. A.; Dull, D. L.; Mosher, H. S., *J. Org. Chem.* **1969,** *34*, 2543-2549.
251. Neises, B.; Steglich, W., *Angew. Chem. Int. Ed. Engl.* **1978,** *17*, 522-524.
252. Wild, R.; Schmidt, R. R., *Liebigs Ann.* **1995,** *1995*, 755-764.
253. Reist, E. J.; Christie, P. H., *J. Org. Chem.* **1970,** *35*, 3521-3524.
254. Cook, G. R.; Pararajasingham, K., *Tetrahedron Lett.* **2002,** *43*, 9027-9029.
255. Azuma, H.; Tamagaki, S.; Ogino, K., *J. Org. Chem.* **2000,** *65*, 3538-3541.
256. De Jonghe, S.; Van Overmeire, I.; Poulton, S.; Hendrix, C.; Busson, R.; Van Calenbergh, S.; De Keukeleire, D.; Spiegel, S.; Herdewijn, P., *Bioorg. Med. Chem. Lett.* **1999,** *9*, 3175-3180.
257. Hoffman, R. V.; Tao, J., *J. Org. Chem.* **1998,** *63*, 3979-3985.
258. Newman, H., *J. Org. Chem.* **1974,** *39*, 100-103.
259. Thum, O.; Hertweck, C.; Simon, H.; Boland, W., *Synthesis* **1999,** *1999*, 2145-2150.
260. Roush, W. R.; Adam, M. A., *J. Org. Chem.* **1985,** *50*, 3752-3757.
261. Fernandes, Rodney A.; Kumar, P., *Eur. J. Org. Chem.* **2000,** *2000*, 3447-3449.
262. Kobayashi, S.; Furuta, T., *Tetrahedron* **1998,** *54*, 10275-10294.
263. Masui, M.; Shioiri, T., *Tetrahedron Lett.* **1998,** *39*, 5199-5200.
264. Collier, P. N.; Campbell, A. D.; Patel, I.; Raynham, T. M.; Taylor, R. J. K., *J. Org. Chem.* **2002,** *67*, 1802-1815.
265. Prasad, A. S. B.; Kanth, J. V. B.; Periasamy, M., *Tetrahedron* **1992,** *48*, 4623-4628.
266. Hoffman, R. V.; Maslouh, N.; Cervantes-Lee, F., *J. Org. Chem.* **2002,** *67*, 1045-1056.
267. Crich, D.; Lim, L. B. L., Glycosylation with Sulfoxides and Sulfinates as Donors or Promoters. In *Organic Reactions*, John Wiley & Sons, Inc.: 2004.
268. Juaristi, E.; Cuevas, G., *Tetrahedron* **1992,** *48*, 5019-5087.
269. Xia, C. F.; Yao, Q. J.; Schumann, J.; Rossy, E.; Chen, W. L.; Zhu, L. Z.; Zhang, W. P.; De Libero, G.; Wang, P. G., *Bioorg. Med. Chem. Lett.* **2006,** *16*, 2195-2199.
270. Lee, A.; Farrand, K. J.; Dickgreber, N.; Hayman, C. M.; Jürs, S.; Hermans, I. F.; Painter, G. F., *Carbohydr. Res.* **2006,** *341*, 2785-2798.
271. Murata, K.; Toba, T.; Nakanishi, K.; Takahashi, B.; Yamamura, T.; Miyake, S.; Annoura, H., *J. Org. Chem.* **2005,** *70*, 2398-2401.
272. Figueroa-Pérez, S.; Schmidt, R. R., *Carbohydr. Res.* **2000,** *328*, 95-102.

273. Imamura, A.; Ando, H.; Korogi, S.; Tanabe, G.; Muraoka, O.; Ishida, H.; Kiso, M., *Tetrahedron Lett.* **2003,** *44*, 6725-6728.
274. Neises, B.; Steglich, W., *Angewandte Chemie International Edition in English* **1978,** *17*, 522-524.
275. Matsuda, J. L.; Naidenko, O. V.; Gapin, L.; Nakayama, T.; Taniguchi, M.; Wang, C. R.; Koezuka, Y.; Kronenberg, M., *J. Exp. Med.* **2000,** *192*, 741-54.
276. Uldrich, A. P.; Le Nours, J.; Pellicci, D. G.; Gherardin, N. A.; McPherson, K. G.; Lim, R. T.; Patel, O.; Beddoe, T.; Gras, S.; Rossjohn, J.; Godfrey, D. I., *Nat. Immunol.* **2013,** *14*, 1137-1145.
277. Wang; Wang, Z.; Zhou, L.; El Boubbou, K.; Ye, X.-s.; Huang, X., *J. Org. Chem.* **2007,** *72*, 6409-6420.
278. Meermeier, E. W.; Laugel, B. F.; Sewell, A. K.; Corbett, A. J.; Rossjohn, J.; McCluskey, J.; Harriff, M. J.; Franks, T.; Gold, M. C.; Lewinsohn, D. M., *Nat. Commun.* **2016,** *7*, 12506.
279. López-Sagaseta, J.; Dulberger, C. L.; McFedries, A.; Cushman, M.; Saghatelian, A.; Adams, E. J., *J. Immunol.* **2013,** *191*, 5268.
280. Rossjohn, J.; Gras, S.; Miles, J. J.; Turner, S. J.; Godfrey, D. I.; McCluskey, J., *Annu. Rev. Immunol.* **2015,** *33*, 169-200.
281. Keller, A. N.; Eckle, S. B. G.; Xu, W.; Liu, L.; Hughes, V. A.; Mak, J. Y. W.; Meehan, B. S.; Pediongco, T.; Birkinshaw, R. W.; Chen, Z.; Wang, H.; D'Souza, C.; Kjer-nielsen, L.; Gherardin, N. A.; Godfrey, D. I.; Kostenko, L.; Corbett, A. J.; Purcell, A. W.; Fairlie, D. P.; McCluskey, J.; Rossjohn, J., *Nat. Immunol.* **2017,** *18*, 402-411.
282. Ayala, A.; Muñoz, M. F.; Argüelles, S., *Oxid. Med. Cell. Longev.* **2014,** *2014*, 360438-360438.
283. Zhou, X.; Taghizadeh, K.; Dedon, P. C., *J. Biol. Chem.* **2005,** *280*, 25377-25382.
284. Dedon, P. C.; Plastaras, J. P.; Rouzer, C. A.; Marnett, L. J., *Proc. Nat. Acad. Sci. U. S. A.* **1998,** *95*, 11113-11116.
285. George, W. O.; Mansell, V. G., *J. Chem. Soc. B* **1968**, 132-134.
286. Brown, R. S.; Tse, A.; Nakashima, T.; Haddon, R. C., *J. Am. Chem. Soc.* **1979,** *101*, 3157-3162.
287. Noi; Noi, R., *Zh. Org. Chem.* **1975,** *11*, 1778.

288. Lange, J.; Anderson, R. J.; Marshall, A. J.; Chan, S. T. S.; Bilbrough, T. S.; Gasser, O.; Gonzalez-Lopez, C.; Salio, M.; Cerundolo, V.; Hermans, I. F.; Painter, G. F., *ACS Chem. Biol.* **2020,** *15*, 437-445.
289. Seela, F.; Kehne, A., *Liebigs Ann. Chem.* **1983,** *1983*, 876-884.
290. Seela, F.; Thomas, H., *Helv. Chim. Acta.* **1994,** *77*, 897-903.
291. Winkeler, H. D.; Seela, F., *J. Org. Chem.* **1983,** *48*, 3119-3122.
292. Seela, F.; Menkhoff, S.; Behrendt, S., *J. Chem. Soc., Perkin Trans. 2* **1986**, 525-530.
293. Seela, F.; Steker, H.; Driller, H.; Bindig, U., *Liebigs Ann. Chem.* **1987,** *1987*, 15-19.
294. Seela, F.; Zulauf, M., *Synthesis* **1996,** *1996*, 726-730.
295. Ramzaeva, N.; Mittelbach, C.; Seela, F., *Helv. Chim. Acta.* **1999,** *82*, 12-18.
296. Seela, F.; Wei, C., *Helv. Chim. Acta.* **1999,** *82*, 726-745.
297. Frank, S.; Xiaohua, P.; Simone, B., *Curr. Org. Chem.* **2007,** *11*, 427-462.
298. Seela, F.; Westermann, B.; Bindig, U., *J. Chem. Soc. Perkin Trans. 1* **1988**, 697-702.
299. Seela, F.; Xu, K., *Helv. Chim. Acta.* **2008,** *91*, 1083-1105.
300. Lerner, L. M., *Carbohydr. Res.* **1977,** *53*, 177-185.
301. Frick, U.; Simchen, G., *Synthesis* **1984,** *1984*, 929-930.
302. Majchrzak, M. W.; Simchen, G., *Tetrahedron* **1986,** *42*, 1299-1304.
303. Eyley, S. C.; Giles, R. G.; Heaney, H., *Tetrahedron Lett.* **1985,** *26*, 4649-4652.
304. Nicolaou, K. C.; Claremon, D. A.; Papahatjis, D. P., *Tetrahedron Lett.* **1981,** *22*, 4647-4650.
305. Kozikowski, A. P.; Ames, A., *J. Am. Chem. Soc.* **1980,** *102*, 860-862.
306. Loader, C. E.; Anderson, H. J., *Can. J. Chem.* **1981,** *59*, 2673-2683.
307. DeSales, J.; Greenhouse, R.; Muchowski, J. M., *J. Org. Chem.* **1982,** *47*, 3668-3672.
308. Anderson, H. J.; Loader, C. E.; Xu, R. X.; Lê, N.; Gogan, N. J.; McDonald, R.; Edwards, L. G., *Can. J. Chem.* **1985,** *63*, 896-902.
309. Kakushima, M.; Hamel, P.; Frenette, R.; Rokach, J., *J. Org. Chem.* **1983,** *48*, 3214-3219.
310. Carmona, O.; Greenhouse, R.; Landeros, R.; Muchowski, J. M., *J. Org. Chem.* **1980,** *45*, 5336-5339.
311. Candy, C. F.; Jones, R. A.; Wright, P. H., *J. Chem. Soc. C* **1970**, 2563-2567.

312. Chadwick, D. J.; Hodgson, S. T., *J. Chem. Soc. Perkin Trans. 1* **1983**, 93-102.
313. Corey, E. J.; Cho, H.; Rücker, C.; Hua, D. H., *Tetrahedron Lett.* **1981,** *22*, 3455-3458.
314. Corey, E. J.; Snider, B. B., *J. Am. Chem. Soc.* **1972,** *94*, 2549-2550.
315. Corey, E. J.; Venkateswarlu, A., *J. Am. Chem. Soc.* **1972,** *94*, 6190-6191.
316. Bray, B. L.; Mathies, P. H.; Naef, R.; Solas, D. R.; Tidwell, T. T.; Artis, D. R.; Muchowski, J. M., *J. Org. Chem.* **1990,** *55*, 6317-6328.
317. Bray, B. L.; Muchowski, J. M., *J. Org. Chem.* **1988,** *53*, 6115-6118.
318. Katritzky, A. R.; Suzuki, K.; Singh, S. K.; He, H.-Y., *J. Org. Chem.* **2003,** *68*, 5720-5723.
319. Jones, R. A., **1992**.
320. Aslani-Shotorbani, G.; Buchanan, J. G.; Edgar, A. R.; Shahidi, P. K., *Carbohydr. Res.* **1985,** *136*, 37-52.
321. Williams, D. T.; Jones, J. K. N., *Can. J. Chem.* **1966,** *44*, 412-415.
322. Ogura, H.; Ogiwara, M.; Itoh, T.; Takahashi, H., *Chem. Pharm. Bull.* **1973,** *21*, 2051-2056.
323. Bukhari, M. A.; Foster, A. B.; Lehmann, J.; Webber, J. M.; Westwood, J. H., *J. Chem. Soc.* **1963**, 2291-2295.
324. Lance, D. G.; Szarek, W. A.; Jones, J. K. N., *Can. J. Chem.* **1969,** *47*, 2889-2891.
325. Blumberg, K.; Fuccello, A.; van Es, T., *Carbohydr. Res.* **1979,** *70*, 217-232.
326. Ducruix, A.; Pascard-Billy, C.; Eitelman, S. J.; Horton, D., *J. Org. Chem.* **1976,** *41*, 2652-2653.
327. Buck, K. W.; Foster, A. B.; Rees, B. H.; Webber, J. M.; Hardy, F. E., *Carbohydr. Res.* **1966,** *2*, 115-121.
328. Clode, D. M., *Chem. Rev.* **1979,** *79*, 491-513.
329. Gyepes, A.; Schäffer, R.; Bajor, G.; Woller, Á.; Fodor, P., *Polyhedron* **2008,** *27*, 2655-2661.
330. Ivanova, N. A.; Valiullina, Z. R.; Shitikova, O. V.; Miftakhov, M. S., *Russ. J. Org. Chem.* **2007,** *43*, 742-746.
331. Ruíz-Pérez, K. M.; Quiroz-García, B.; Hernández-Rodríguez, M., *Eur. J. Org. Chem.* **2018,** *2018*, 5763-5772.
332. Sakamoto, T.; Satoh, C.; Kondo, Y.; Yamanaka, H., *Chem. Pharm. Bull.* **1993,** *41*, 81-86.
333. Langli, G.; Gundersen, L.-L.; Rise, F., *Tetrahedron* **1996,** *52*, 5625-5638.

334. Brændvang, M.; Gundersen, L.-L., *BioOrg. Med. Chem.* **2007,** *15*, 7144-7165.
335. Bartoccini, F.; Piersanti, G.; Mor, M.; Tarzia, G.; Minetti, P.; Cabri, W., *Org. Biomol. Chem.* **2012,** *10*, 8860-8867.
336. Brown, D. J., *Chemistry of Heterocyclic Compounds: The Pyrimidines*. **1994**.
337. Seela, F.; Driller, H.; Liman, U., *Liebigs Ann. Chem.* **1985,** *1985*, 312-320.
338. Seela, F.; Shaikh, K., *Helv. Chim. Acta* **2004,** *87*, 1325-1332.
339. Ruíz-Pérez, K. M.; Quiroz-García, B.; Hernández-Rodríguez, M., *Eur. J. Org. Chem.* **2018,** *2018*, 5763-5772.
340. Vilsmeier, A.; Haack, A., *Ber. Dtsch. Chem. Ges.* **1927,** *60*, 119-122.
341. Ilyin, P. V.; Pankova, A. S.; Kuznetsov, M. A., *Synthesis* **2012,** *44*, 1353-1358.
342. Krivopalov, V. P.; Nikolaenkova, E. B., *Russ. Chem. Bull.* **1993,** *42*, 1059-1062.
343. Reimer, K.; Tiemann, F., *Ber. Dtsch. Chem. Ges.* **1876,** *9*, 1268-1278.
344. Plaut, G. W. E.; Smith, C. M.; Alworth, W. L., *Annu. Rev. Biochem.* **1974,** *43*, 899-922.
345. Dangerfield, E. M.; Plunkett, C. H.; Win-Mason, A. L.; Stocker, B. L.; Timmer, M. S. M., *J. Org. Chem.* **2010,** *75*, 5470-5477.
346. Dickens, M. P.; Roxburgh, P.; Hock, A.; Mezna, M.; Kellam, B.; Vousden, K. H.; Fischer, P. M., *Bioorg. Med. Chem.* **2013,** *21*, 6868-6877.
347. Wang, J.; Hu, X.; Wang, D.; Xie, C.; Lu, W.; Song, J.; Wang, R.; Gao, C.; Liu, M., *Org. Biomol. Chem.* **2018,** *16*, 4464-4470.
348. Robson, P.; Stacey, M.; Stephens, R.; Tatlow, J. C., *J. Chem. Soc.* **1960**, 4754-4760.
349. Schaefer, K.; Albers, J.; Sindhuwinata, N.; Peters, T.; Meyer, B., *ChemBioChem* **2012,** *13*, 443-450.