University of Southampton

Faculty of medicine

Academic unit of Clinical and experimental sciences

Understanding the molecular interaction between unconventional T cell antigen receptors and their ligands; paving the way for future immunotherapy

by

Daniel Mark Burns

ORCID ID 0009-0000-5498-1308

Thesis for the degree of Doctor of Philosophy (PhD)

July 2024

University of Southampton

Abstract

Faculty of Medicine

Clinical and experimental sciences

Doctor of Philosophy

Understanding the molecular interaction between unconventional T cell antigen receptors and their ligands; paving the way for future immunotherapy

by

Daniel Mark Burns

The ability of the human immune system to coordinate an effective immune response to stress is a complex and tightly regulated process that brings together many different molecules and mechanisms from multiple interconnecting systems. Understanding these mechanisms and pathways is key in developing novel therapeutic targets to combat the wide and ever-growing array of human diseases. Unconventional T cells and their ligands have recently been established as targets for immunotherapy in diseases such as cancer, primarily due to their major histocompatibility (MHC)-unrestricted nature and strong anti-tumour response. They are often described to bridge the gap between innate and adaptive immunity, due to their ability to display characteristics and phenotypes of both these responses. Despite the growing interest in unconventional immune responses, limited structural data currently represents a bottleneck in understanding them fully, thus hindering the development of successful treatments in the field.

My work aims to investigate the hypothesis that CD1 antigen-presenting molecules display substantial conformational plasticity and are thus capable of presenting a wide variety of both stimulatory and inhibitory lipids to diverse TCRs. We utilise structural techniques to investigate the molecular mechanism underpinning CD1 plasticity and unconventional TCR recognition of two CD1 isoforms (CD1d and CD1c) presenting lipid ligands. We also develop an optimised pipeline to generate soluble, refolded gamma delta (γδ) TCRs that can be used to investigate CD1c recognition by the γδ TCR.

This work demonstrates that CD1d can alternate its conformation within the vicinity of the lipid binding site, in a lipid cargo-dependent fashion. We display, using X-ray crystallography and molecular dynamics (MD) simulations, that CD1d can ‘relax’ and ‘constrict’ its lipid binding groove providing potential mechanisms for its ability to accommodate a range of lipid sizes and properties, a feature

previously demonstrated in other CD1 isoforms. Furthermore, we use X-ray crystallography to produce structural data of a novel macaque CD1d structure, to explore the conservation of CD1d lipid-binding groove flexibility across species. We also demonstrate successful optimisation of a refolding and purification pipeline to produce soluble, stable CD1c-reactive γδ TCRs. Further to this, we utilise an in-house generated bead display system to confirm the CD1c-reactivity of the refolded γδ TCR. Our findings suggest CD1d shares the same conformational adaptability of other CD1 isoforms such as CD1b and CD1c, providing a potential molecular mechanism to explain CD1d's ability to bind ligands that exceed the standard groove size. We also show that we can refold stable, CD1c-reactive γδ TCRs that can be used to explore the molecular mechanisms underlying the recognition of CD1c lipid complexes by the γδ TCR. These results can be used to build on a greater understanding of how CD1 antigen-presenting molecules can modulate their conformation to perform different functions, and they pave the way for further studies to unravel CD1 recognition by TCRs.

Publications

Frontiers immunology (review paper)

Towards a better understanding of human iNKT cell subpopulations for improved clinical outcomes. Burns, Look et.al 2023.

Science advances

A peptide derived from the nuclear export protein XPO1 is recognised by the natural killer cell receptor KIR2DS2. Blunt et.al 2024

Vet Immunology immunopathology

Cattle killer immunoglobulin-like receptor expression on leukocyte subsets suggests functional divergence compared to humans. Hay et.al 2023

PNAS (under review)

CD1c-autoreactive T-cells exhibit potent responses to Mycobacterium tuberculosis. Farag et. al 2024 (Under review)

Table of Contents

List of Figures

List of Tables

Academic Thesis: Declaration of Authorship

I, Daniel Mark Burns, declare that this thesis and the work presented in it are my own,

and has been generated by me as the result of my original research:

Understanding the molecular interaction between unconventional T cell antigen receptors and their ligands; paving the way for future immunotherapy

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University.
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
3. Where I have consulted the published work of others, this is always clearly attributed.
4. Where I have quoted from the work of others, the source is always given. Except for such quotations, this thesis is entirely my own work.
5. I have acknowledged all main sources of help.
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself
7. None of this work has been published before submission.

Signed:

Date: 17/02/2025

Acknowledgments

I would like to first say a huge thank you to my primary and secondary supervisors, Dr Salah Mansour, and Dr Ivo Tews, for their immeasurable support and guidance throughout my project. Your passion for the subject and brilliant minds have been fundamental in providing me with a platform to achieve more than I could have ever hoped for when starting my PhD three years ago.

I would also like to thank my colleagues both within Southampton General Hospital and at the Highfield campus, Southampton University. During my PhD, I have made friendships for life and am extremely grateful for all the fun times we have been able to share. The interdisciplinary nature of my PhD project meant I was fortunate enough to meet a variety of scientists, all of whom have been incredibly helpful in providing support and knowledge throughout my project.

I would like to thank the many collaborators I was fortunate enough to work with throughout my project, including Professor Salim Khakoo, and Professor Paul Elkington, and our strong and fruitful collaboration with TCR Biotechnology company, Immunocore. Your guidance and willingness to share ideas to advance my project and strengthen our collaboration has been immensely helpful.

I would like to express my gratitude to Diamond Light Source and our international collaborators at the European Synchrotron Radiation Facility (ESRF), who provided me with the opportunity to visit your wonderful city of Grenoble, whilst collecting important data for my PhD project. I hope to return to Grenoble someday.

Finally, I would like to thank my family, my wife, and her family, for supporting me throughout my entire PhD journey. Without your guidance, belief, and support throughout, it would not have been possible for me to complete this degree.

Definitions and abbreviations

α1	Alpha 1
α2	Alpha 2
ADCC	Antigen dependant cellular cytotoxicity
αGC	Alpha Galactosylceramide
AGL	Altered Glycolipid ligand
AP2	Adaptor protein
APC	Antigen presenting cell
ARF-6	ADP ribosylation factor 6
β2m	Beta-2-microglobulin
BCR	B cell receptor
CD1	Cluster of differentiation 1
CDR	Complementarity determining region
CSM	Cell surface markers
DC	Dendritic cell
DMAPP	Dimethylallyl pyrophosphate
DTT	Dithiothreitol
ER	Endoplasmic reticulum
FASL	Fas ligand
FACS	Fluorescence-activated cell sorting
GEM	Granulocyte macrophage colony-stimulating factor
GMM	Glucose monomycolate
HIV	Human immunodeficiency virus
HMB-PP	(E)-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate (HMB-PP)
IGb3	Isoglobotriosylceramide
IFN-γ	Interferon gamma
iNKT cell	Invariant natural killer T cell
IL	Interleukin
IPP	Isopentenyl pyrophosphate
KDa	Kilodaltons
LPC	Lysophosphatidylcholine
LPS	Lipopolysaccharide
MAIT	Mucosal associated invariant T cells
Mg/ml	Milligram per millilitre
MHC	Major histocompatibility complex
MICA/B	MHC class 1 chain related protein A/B
MPM	Mannosyl-beta1-phosphomycoketides
MR1	MHC related protein 1
MyD88	Myeloid differentiation primary response 88
N-BP	Amino-bisphosphonate
NF-KB	Nuclear factor -KB
NK	Natural killer cells
NKG2D	Natural Killer Group 2D

PAMPs	Pathogen associated molecular patterns
PA	Phosphatidic acid
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PC	Phosphatidylcholine
PDB	Protein data bank
PD-L1	Programmed cell death ligand 1
PG	Phosphatidylglycerol
PI	Phosphatidylinositide
PM	Phosphomycoketides
PRRs	Pathogen recognition receptors
PS	Phosphatidylserine
PTM	Post translational modification
RAG	Recombination activating genes
SDS-PAGE	Sodium dodecyl-sulphate polyacrylamide gel electrophoresis
SM	Sphingomyelin
SPR	Surface plasmon resonance
Strep-PE	Streptavidin-Phycoerythrin
Strep-HRP	Streptavidin horseradish peroxidase
TCR	T cell receptor
TGF-β	Transforming growth factor beta
TGN	Trans Golgi network
TH1	T helper 1
TH2	T helper 2
TLRs	Toll-like receptors
TNF	Tumour necrosis factor
TRAIL	Tumour Necrosis Factor-Related Apoptosis Inducing Ligand
TMB	3,3',5,5'-Tetramethylbenzidine
VDJ	Variability, diversity, joining.
γδ	Gamma delta

Chapter 1 Introduction

1.1 The immune response

The immune system consists of a complex and interconnecting network of host defence mechanisms that can elicit varying responses to a vast array of foreign and pathogenic challenges that themselves regularly evolve. In addition to combatting invading foreign material, the body must also tightly regulate its internal pathways and responses to avoid unwanted interactions between the body's own machinery and immune surveillance mechanisms. The ability to discriminate between self and non-self is central to innate and adaptive immune function.

The immune response usually follows a 'loosely structured' sequence, whereby recognition of a pathogen leads to immune activation and, subsequently an adaptive response. Depending on the pathogen type and response, infection generally results in pathogen clearance, chronic persistence, or host death. The outcome of the host is generally dictated by its ability to detect and respond to foreign material. This is particularly relevant to Cluster of Differentiation 1 (CD1)-mediated antigen presentation, which plays a key role in recognising lipid-based antigens (1).

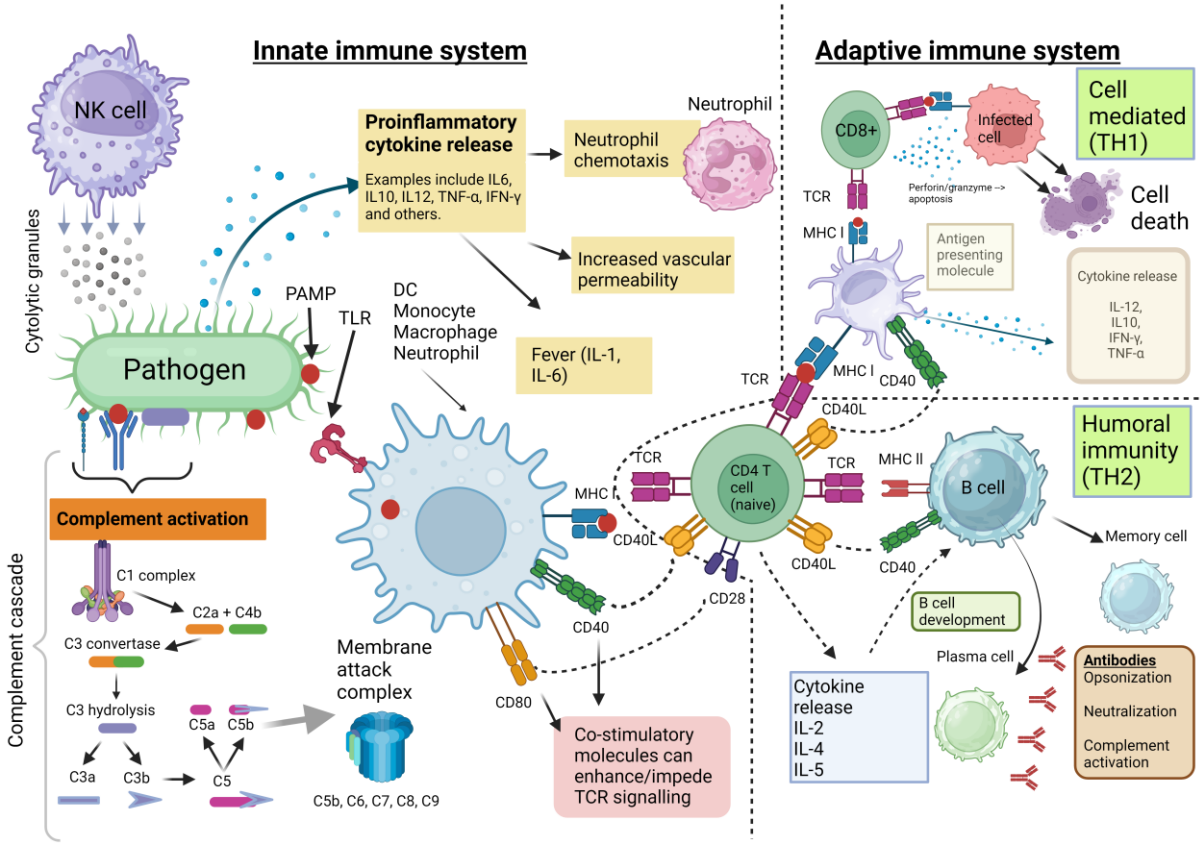


Figure 1.1 Overview of the human immune systems response to pathogenic challenge

Schematic diagram of the human immune response to infection. The response consists of a multi-faceted approach consisting of both innate and adaptive arms, both of which are interconnecting and tightly regulated to coordinate an efficient and effective immune response. Adaptive immunity can be further categorised into cell-mediated T helper 1 (TH1) and humoral T helper 2 (TH2). Central to the interconnecting immune pathways is a CD4 T cell, interacting with various cell types to elicit varying immune functions. Image generated using Biorender.com and adapted from 'https://geekymedics.com/immune-response/'.

The innate immune system provides a first-line defence against invading pathogens, utilising pathogen recognition receptors (PRRs) to detect common pathogenic features (2). This is a rapid response but lacks the memory features associated with adaptive immunity. Despite this, the innate immune system is fundamental in shaping the adaptive immune response by influencing antigen presentation. CD1 molecules are commonly involved at the interface between innate and adaptive immunity, and therefore, understanding the cross-talk between the two facets of the immune system is crucial in understanding the mechanism of antigen presentation to T cells (3).

1.1.1 The innate immune response

The innate immune system comprises both chemical and cellular mechanisms which work synergistically to eliminate invading pathogens (4). It consists of multiple components and requires extremely tight regulation to ensure rapid response times and cooperation between subcomponents of the larger system (5). The responses are typically very fast and are often triggered by direct physical contact with the invading material or other signals produced by environmental stress (4).

The cellular arm of the innate immune system involves phagocytes, such as neutrophils, monocytes, and macrophages which engulf pathogens to degrade them (6). An example of a common microbial feature that is recognised by these cells is lipopolysaccharides (LPS) (6, 7). Recognition of these features encourages an inflammatory response and subsequent recruitment of other cell types.

Neutrophils are a strong example of first responders and play important roles in signalling and foreign pathogenic clearance by releasing cytokines such as interleukin 1 (IL-1), interleukin 6 (IL-6), and tumour necrosis factor (TNF) (3, 8). The inability to efficiently produce the correct cytokines has been implicated in autoimmunity and inflammatory disease, likely due to a knock-on effect on local inflammation and cell recruitment (6). Nuclear factor-κB (NF-κB) is another molecule with important roles in local inflammation and cytokine release, brought about by PAMP-PRR interaction (9). Following the internalisation and ingestion of antigens, phagocytes then present the material to adaptive immune cells, demonstrating one way the synergy between the two systems is important.

Alongside neutrophils, natural killer (NK) cells utilise perforin and granzyme-mediated cytotoxicity to target infected cells (10, 11). NK cells are different in the sense that they utilise MHC class I sensing to detect changes in cellular expression levels and subsequent danger.

Chemical innate immunity is a synergistic mechanism that works in conjunction with the cellular system. It comprises a combination of the complement system and coordinated cytokine release to assist with both the uptake and ingestion of foreign material and the recruitment of other adaptive cell types. The complement system can be activated by three separate pathways: classical, mannose-binding lectin, and alternative. It is essentially made up of a network of plasma proteins that combine to opsonise pathogens.

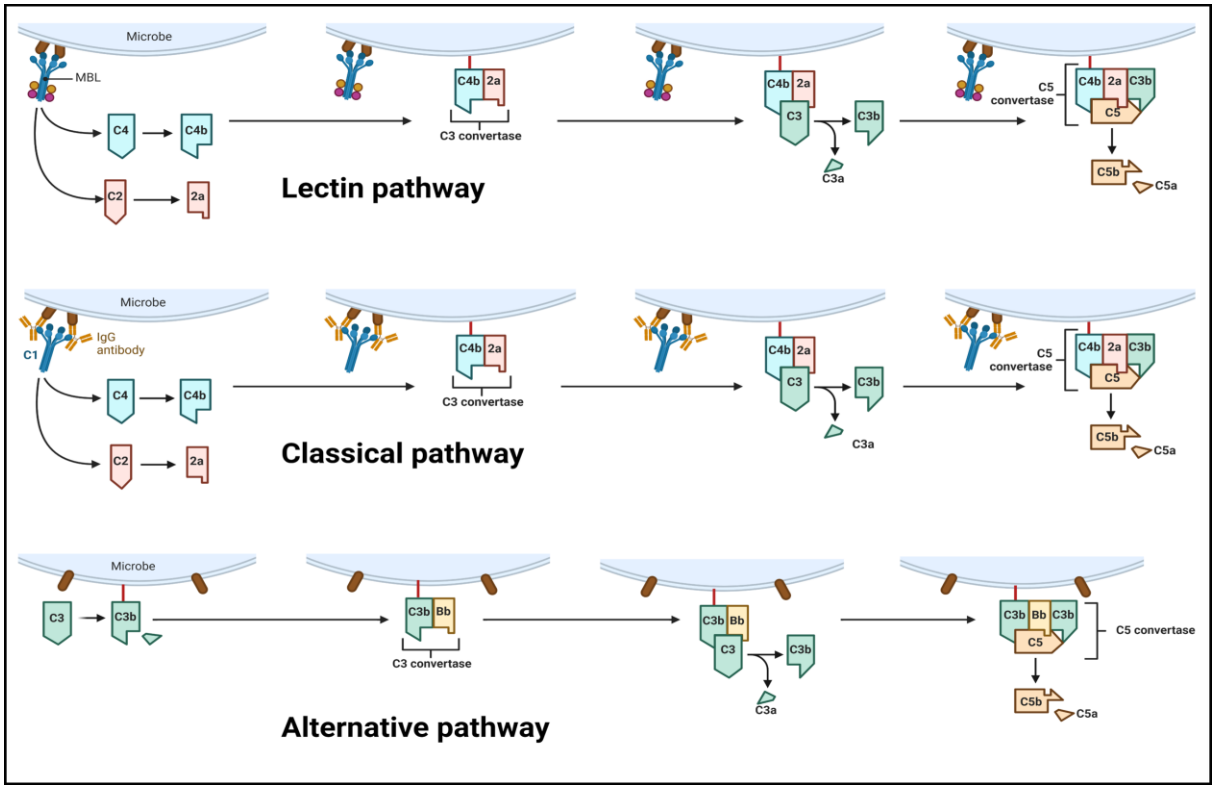


Figure 1.2 Schematic diagram of the three complement system pathways

Lectin, classical, and alternative complement system pathways utilise plasma proteins to opsonise pathogens and generate an inflammatory response. All three pathways use different initial mechanisms, but all lead to the production of C3 convertase and the generation of a membrane attack complex. Image generated using Biorender.com and adapted from *'https://geekymedics.com/immune-response/'*.

These pathways require antigen-antibody complexes, mannose carbohydrates, and direct surface interaction, respectively, for cascade initiation (12). Despite these pathways requiring different initiation molecules, the effector molecules are the same, with all three pathways generating C3 convertase and, subsequently a membrane attack complex (12). The complement system must be tightly regulated, as over- or under-activation has been associated with multiple infectious and non-infectious diseases (13). This system can also induce chemoattractant signals for phagocytes and damaged cells by penetration and, unsurprisingly, comes with rigorous regulation and checkpoints to prevent overstimulation and consequential host tissue or cell damage (14). Proinflammatory cytokines encompass the final arm of the chemical innate response, functioning not only to recruit adaptive cells to the site of inflammation but also to turn off the stimulus and prevent overstimulation. Like other areas of the immune response, cytokine and chemokine secretion must be tightly controlled to prevent over-activation and damage to self-tissue (15). The vast number of soluble mediators act both

individually and synergistically to activate and recruit various cell types. Examples of this include interleukin-12 (IL-12) activating NK and TH1 cells, interferon-gamma (IFN-γ) activating cell-mediated immunity, and interleukin-1 (IL-1) activating lymphocytes (15).

1.1.2 The adaptive immune response

The synergy between the two arms of the immune system is vital in coordinating a healthy immune response. This requires cell types that bridge the gap between innate and adaptive immunity, buying time for other immune cells to be activated. While the innate immune system provides immediate protection, the adaptive immune system generates pathogen-specific responses and establishes long-term immunological memory. This response is primarily driven by T and B lymphocytes, which undergo somatic recombination and subsequently produce a highly diverse antigen repertoire (16).

Although conventional antigen presentation primarily involves peptide recognition through major histocompatibility complex (MHC) molecules, CD1-mediated lipid antigen presentation plays a distinct role in shaping both humoral and cell-mediated immunity, making it a key focus of this thesis (17).

B cells are central to the humoral immune response and undergo a complex and well-regulated selection and maturation process, arising from pluripotent stem cells in the foetal liver and bone marrow (18). B cells contain a huge variety of B cell receptors (BCRs), which are essentially membrane-bound antibodies, achieved through variability, diversity, joining (VDJ) recombination (19). Upon antigen recognition within lymphoid tissues, B cells undergo clonal expansion and differentiate into antibody-secreting plasma cells, a process facilitated by CD4+ T helper (TH2) (20, 21). The high-affinity antibodies produced by this process contribute to pathogenic clearance using a combination of methods, namely opsonisation, neutralisation, and complement activation (22). Despite B cell activation and response predominantly being associated with protein or peptide antigens presented by MHC, there is growing evidence to suggest CD1-driven activation of B cells can directly contribute to an immune response, re-enforcing the importance of exploring non-peptide recognition (23).

Cell-mediated immunity, on the other hand, is driven by T lymphocytes. These rely on professional antigen presentation by antigen-presenting cells (APCs). For an effective response to be mounted, efficient and accurate translation of innate-like signals must occur, with DCs at the forefront of this process (24). Immature, tissue-resident DCs can utilise various methods to capture and digest antigens, such as micropinocytosis and receptor-mediated endocytosis (25). These DCs then migrate to lymphatic tissue, where they become activated, mature, and highly specialized antigen-presenting

cells (APCs) (26). DCs use cathepsins primarily to generate peptide antigens for loading onto MHC class II molecules, and to a lesser extent for MHC class I molecules via cross-presentation. These MHC-peptide complexes are then transported to the cell surface (25). The ability of these cells to efficiently activate naïve T cells requires the use of co-stimulatory molecules. An example of this is the B7 protein family, such as CD80 or CD86, which binds CD28 on T cells (27). These molecules allow activation of effector T cells and the subsequent initiation of a T cell response. A combination of antigen presentation, MHC presentation, and co-stimulation is required for the differentiation into TH1 and TH2 cells.

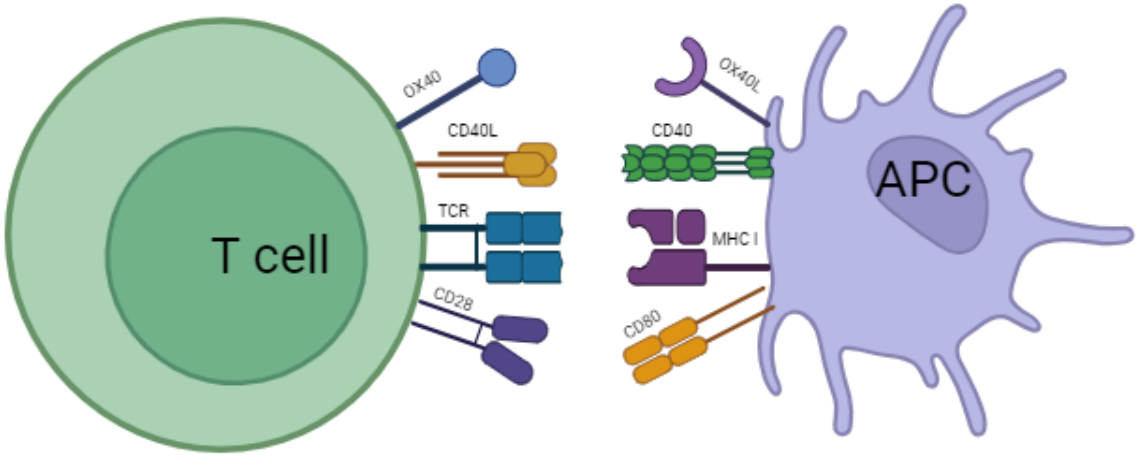


Figure 1.3 T-cell activation by antigen-presenting molecules.

T cells use a T cell receptor (TCR) and co-stimulatory molecules to interact with MHC class l and other co-receptors presented by APCs. CD28, OX40, and 4-1BB are co-stimulatory receptors that recognise co-stimulatory molecules and are required for full activation of the naïve T cell and APC. Other receptors expressed by APCs include CD80 and CD40. Image generated using Biorender.com and adapted from 'https://geekymedics.com/immune-response/'.

Activated CD8+ cytotoxic T cells eliminate infected cells through perforin-granzyme cytotoxicity, Fas-FasL apoptosis, and IFN-γ secretion (28-30).

Memory T cells remain in the blood and tissue long after infection clearance, allowing for a quicker and more efficient response upon re-infection. The differentiation pathways of memory T cells remain an area of active research, with significant implications for vaccine development and immunotherapeutic strategies (31). However, beyond the conventional peptide-MHC paradigm, unconventional T cell subsets such as natural killer T (NKT) cells and γδ T cells, interact with and

recognise lipid antigens presented by CD1 molecules. Given the structural and functional similarities between CD1d-mediated and MHC-restricted antigen presentation, further investigation into lipid-based immunity could provide key insights into how antigen presentation influences T-cell function and immune regulation. One question this thesis aims to address is how CD1d accommodates lipid ligands of varying properties and whether structural flexibility within the antigen-binding groove contributes to TCR recognition and immune responses.

1.2 Antigen presentation

1.2.1 CD1 vs MHC antigen presentation

The cluster of differentiation l (CD1) proteins consists of a family of glycosylated cell surface receptors that function to present lipid antigens to an array of T cells (32). Their discovery expanded with the paradigm of antigen presentation to T cells, demonstrating that lipids, as well as peptides, can be recognised (33). Unlike MHC, CD1 accommodates structurally diverse lipids within its binding groove, influencing T cell recognition and immune responses. There are five CD1 isoforms, which differ in terms of cellular expression patterns, lipid repertoire, and subsequent function (32). The isoforms also show different intracellular trafficking patterns, thus influencing the type of lipid antigens they encounter (34). In terms of cellular expression patterns, Group one CD1 expression is limited exclusively to professional APCs such as Langerhans cells, and thymocytes (35). CD1d however, is the most widely expressed CD1 isoform and has been found on cells of myeloid and lymphoid lineage (monocytes, macrophage, DC) and B cells and thymocytes, respectively (36).

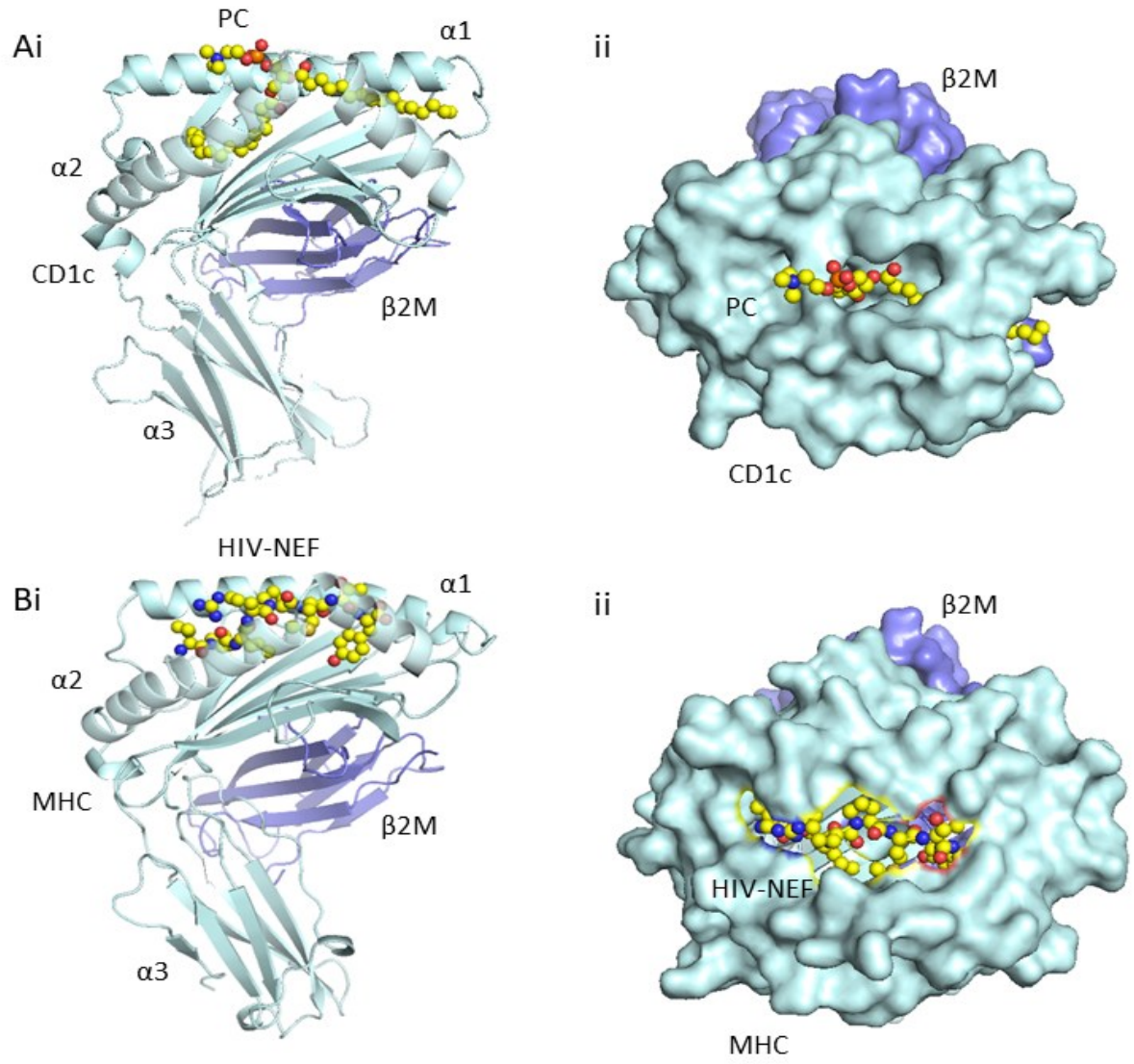


Figure 1. 4 Structural comparison of antigen presentation between CD1c and HLA-B*35:01 NEF75-82 (VPLRPMTY – 8mer).

CD1c (Ai) and MHC (Bi) share highly similar tertiary and quaternary structures with a 3-domain heavy chain non-covalently associated with beta-2-microglobulin β2M. Differences arise predominantly within the antigen presentation compartment of the molecules, with the CD1 bound lipid (PC) sat with its acyl chains buried deep within the antigen-binding cleft, and polar headgroup extending outwards above the CD1 heavy chain for TCR contact (Aii). In contrast, the peptide (VPLRPMTY) extends across the peptide binding groove of MHC, with a larger portion of the antigen available for TCR recognition compared with CD1-lipid (Bii). CD1/MHC heavy chain is shown in cyan, β2M in purple, and lipid in yellow (nitrogen and oxygen in blue and red respectively). Proteins are displayed in cartoon orientation using PyMOL. Protein data bank (PDB) entries 6C15 and 1A1N were downloaded for CD1c and MHC respectively.

The CD1 molecules are initially assembled in the endoplasmic reticulum (ER) where they bind stabilising lipids and traffic to the cell surface. They are then recycled to various intracellular compartments where the stabilising lipids are exchanged for more specific lipids for antigen presentation to T cells and B cells (37). While CD1 proteins share a similar three-domain heavy chain structure with MHC class I *(Figure 1.4*), the antigen-binding mechanisms employed by each are substantially different, with CD1 suited to lipid presentation. CD1 molecules utilise a tunnel network to accommodate lipid antigens, sequestering their acyl chains while exposing only the polar headgroup for TCR engagement. The extent to which CD1 can dynamically adjust its binding groove remains an open question and is a key focus of this thesis. The heavy chain consists of two antiparallel helical α sub-units (α1 and α2), which sit on top of a six-stranded β sheet (38, 39). These two subunits make up the extracellular antigen binding domain whilst also contacting the α3 transmembrane domain, which contains a cytoplasmic tail important for CD1 trafficking (39). One major difference between CD1 and MHC lies within the architecture of their respective binding grooves. Both antigen-presenting molecules contain pockets in which their respective antigens bind, however, they differ in terms of size, structure, and properties. MHC class I molecules present short peptides within a closed binding groove, differing from MHC class II which display longer more complex peptides in an open groove *(Figure 1.4).* Despite both using defined pockets to secure antigens, CD1 differs by the majority of the lipid being sequestered within its lipid binding tunnels, displaying predominantly the headgroup for interaction with the TCR (40).

Whilst MHC class l and ll utilise an F pocket and P1 pocket to dock the C and N terminal ends of their peptides respectively, the deeper and more complex interconnecting network of tunnels within CD1 allows for different isoforms to bind lipids of varying size and properties (40). The repertoire of lipid antigens available for presentation by each of the CD1 isoforms is dependent on the intracellular trafficking network and lipid binding groove architecture of each CD1 isoform. The intracellular trafficking dictates which cellular compartments each isoform localizes to, and therefore which lipids are available to be loaded.

1.2.2 Group 1 CD1

The CD1 proteins are categorised primarily by nucleotide and amino acid sequence homologies, with group one consisting of CD1a, CD1b, and CD1c. These proteins are expressed by mature hematopoietic cells, including B cells, APCs and thymocytes (41). Despite the molecular mechanisms of group one CD1-T cell interaction lagging slightly behind CD1d due to lack of expression in mice, some strong advances have been made.

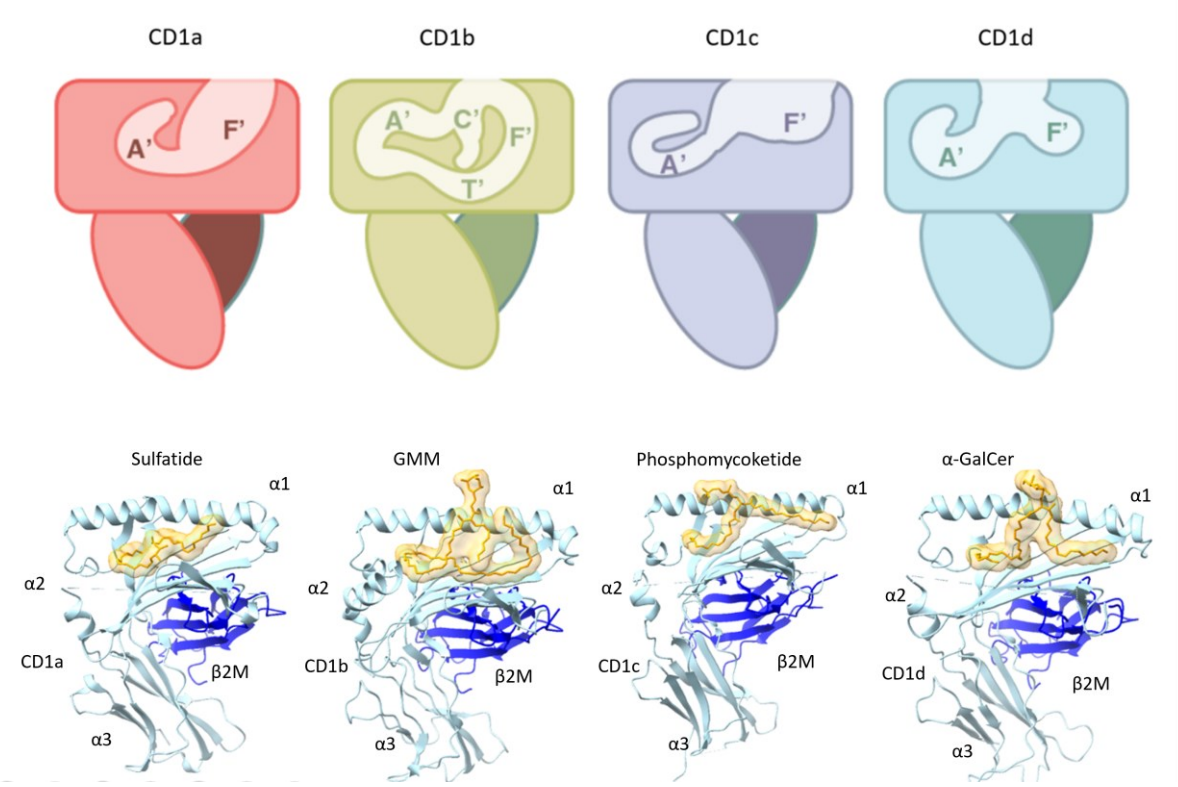


Figure 1. 5 Structural comparison of the four different antigen-presenting CD1 isotypes

Cartoon and ribbon representations of the four CD1 antigen-presenting molecules CD1a, CD1b, CD1c, and CD1d. Lipid cargo available for presentation by each molecule is dictated by binding groove architecture and the cell type each isoform is presented on, amongst other factors. Lipid binding grooves form a series of interconnecting tunnels, with CD1a, CD1c, and CD1d consisting of A' and F' grooves in different architectures, and CD1b also displaying an interconnecting T tunnel for accommodation of larger lipid moieties. Lipid antigen displayed in yellow, heavy chain in light blue, β2m in dark blue. Images were generated using PyMOL under PDB accession codes 1ONQ, 5L2J, 4ONO, and 1ZT4 from A to D respectively.

1.2.2.1 CD1a

The group one isoform, CD1a, is predominantly expressed on both myeloid DCs and a subpopulation of tissue-resident macrophages known as epidermal Langerhans cells, which form a web-like network within the skin epidermis (42). Due predominantly to its tissue expression pattern, CD1a has been implicated in a variety of autoimmune skin conditions, bacterial infections, and cancer (43, 44). Unlike other CD1 isoforms, CD1a lacks a tyrosine-based sorting motif, preventing it from trafficking to late endosomes and lysosomes (45, 46). Instead, it primarily resides at the cell surface, with limited recycling through early endosomal compartments, shaping its lipid antigen repertoire (47, 48).

As well as the recycling mechanism of CD1a, its structure (the lipid-specific binding groove) is thought to be another strong determinant of the lipid repertoire available for presentation. The first crystal structure of CD1a was solved in 2003, shown with a sulfatide self-antigen bound in the lipid-binding groove (49). As expected, common structural features of the CD1a fold were the same as other CD1 structures previously solved, with the α1,α2,α3 heavy chain domains non-covalently associated to β_2m, and the α1,α2 helices forming the deep lipid-binding groove for lipid loading (49).

The A' and F' channels, lined mainly with hydrophobic residues, branch out from the centre of the groove and form two deep channels which can each accommodate a lipid tail (49). The total volume of this binding groove is roughly 1,300Å, making it the smallest binding groove of all the CD1 isoforms (49). One defining characteristic of the CD1 A' channel is its conserved nature, appearing relatively similar in all isoforms compared to the F' channel. It also appears to be more selective in terms of the lipid it can bind, often being described as a molecular ruler by dictating the lipid chain length it can accommodate. The channel curves around a central pole in the structure is narrow and short in length, and terminates abruptly due to a series of specific residues preventing further access into the groove (49). The F' channel is much more varied in the type and size of lipids it can accommodate (49). The ability of the lipid to anchor within the two-channel groove has been demonstrated to have a profound effect on molecular mechanisms of T-cell recognition and antigen presentation by CD1a (50). This was demonstrated by lipopeptide and sulfatide comparison in which different pathways across the A' and F' junction (lipopeptide more direct compared to sulfatide) subsequently caused a lateral shift in the galactosyl headgroup, positioning it closer to the surface where the TCR is expected to contact (50).

More recently, CD1a has since been shown with various other lipids bound in the groove including lysophosphatidylcholine (LPC) and sphingomyelin (SM) (51). Structural data of CD1a bound in complex with an αβ TCR has also provided insight into the molecular mechanism by which T cells recognise lipid-bound CD1a (51). A more recent study has discovered a completely novel mechanism for γδ TCR binding, in which the γδ T cell not only required both γ and δ chains for binding (previous studies have shown a δ dominated mechanism of binding) but relied on a 'side end' docking mode (52).

1.2.2.2 CD1b

CD1b, another member of the group one CD1 proteins, is expressed predominantly on myeloid DCs, macrophages, and other peripheral immune cells (53). CD1b was the first CD1 isoform shown to present lipid antigens and is one of the better-understood CD1 proteins (33, 54). It has been demonstrated to bind three major classes of lipids, namely sphingolipids, phospholipids, and scaffold lipids, some of which require the protein to traffic through the endosomal network and some do not

(33, 55). CD1b appears to venture furthest into the endosomal compartments, with much evidence suggesting it reaches late endosomes and lysosomes to sample the lipid repertoire available within these areas (56). Surface plasmon resonance (SPR) binding studies have revealed that CD1b, like other CD1 isoforms, uses AP-2 and AP-3 mediated by the tyrosine-based motif found within the cytoplasmic tail to internalize from the cell surface and traffic to late endosomes/lysosomes (33). Multiple sources of evidence have provided the rationale that CD1b encountering lipid antigen at low pH is a requirement for successful and efficient antigen presentation to T cells. Firstly, the CD1b heavy chain has the strongest association with β_2m meaning it is unlikely to dissociate at low pH found within lysosomes. Secondly, deletion of the CD1b tail containing tyrosine motif, subsequently inhibiting efficient transfer to low pH endosomal compartments, decreases the ability of CD1b to present long-chained bacterial lipids to T cells including GMM and mycolates (33). CD1b has the largest and most complex binding groove of all the CD1 isoforms, with an interconnecting tunnel network (A', F', C', and T') that accommodates diverse antigens (*Figure 1.5).* This allows the binding of long lipid structures, including mycolates, making it uniquely capable of presenting microbial lipids to T cells (37).

The ever-growing library of CD1b crystal structures (including those loaded with phosphatidylcholine (PC), phosphatidylinositol (PI), and phosphatidylglycerol (PG) has allowed greater insight into the intricate mechanisms of both lipid antigen binding, and interaction with cognate TCRs (56). The majority of current structures include two-tailed lipids, in which the acyl chain and sphingosine chain are bound in the C' and A' pockets respectively (56). Similar to the A' pocket in CD1a acting as a molecular ruler, the exclusive C' portal in CD1b can represent an open and closed conformation which can limit the size of the lipid that the pocket can accommodate (typically around 16 amino acids) (50, 56). Due to the size of the CD1b lipid groove being larger than many of the lipids it can bind, scaffold lipids are used to maintain the stability of the molecule (49). These temporary lipids have also been suggested to populate the groove ready for exchange with longer foreign lipids such as mycolic acids (57).

Current co-crystal studies of CD1b with TCR include CD1b-PG-TCR, CD1b-PC-TCR, and CD1b-GMM-TCR (58-60). These three co-crystal structures allow insight into the molecular mechanism underpinning T-cell interaction with CD1b lipid complexes and provide potential implications as to how the TCR distinguishes between loaded lipids. All three TCRs exhibit commonalities in terms of CD1b recognition, with a 'hotspot' of residues across the α1 and α2 helix of CD1b predominating interactions with the three TCRs (60). The complementarity determining region (CDR)3β loop also plays a major role in interaction in all three cases, with salt bridge formation between specific CDR3β residues and the α1 helix of CD1b proving crucial for the stabilisation of the complexes (60). The main difference between the binding mechanisms becomes clear when the specificity of each TCR for the lipid antigen

is analysed. The most recent structure 'BC8B', shows less lipid specificity than the other two models, made apparent by an altered docking angle and TCR footprint across the binding cleft of CD1b (60). The TCRs that are highly specific for lipid antigens (namely GEM42 and PG90) elicit their functional specificity by essentially pinching either side of the lipid head groups and directing them towards the αβ chain interface (58, 60). The same mechanism cannot be described for the more recently discovered BC8B TCR, which contacts the lipid predominantly through the CDR1α and CDR3α loops (59).

1.2.2.3 CD1c

The last member of the group one CD1 molecules, CD1c, is arguably the least well-understood of this subcategory. As previously alluded to, cellular expression patterns as well as lipid groove architecture are two of the factors that can dictate lipid repertoire available for presentation. CD1c is expressed predominantly on marginal zone B cells and myeloid DCs, as well as macrophages and less frequently Langerhans cells (35). This protein's expression, along with CD1a, and CD1b, can also be induced on monocytes after treatment with Granulocyte-macrophage colony-stimulating factor (GMCSF) and IL-4 (61). CD1c, like other CD1 isoforms, utilises a YXXZ tyrosine motif in its cytoplasmic tail to associate with AP-2 which leads to subsequent clathrin-mediated internalization into 'sorting endosomes' (46, 62). CD1c recycles broadly through the endocytic pathway, surveying different endosomal compartments and therefore potentiating diversity in lipid presentation to T cells. It has been reported that tail motif deletion and increases in endosomal pH can abrogate CD1 antigen binding in CD1b, again providing evidence that efficient and correct endosomal recycling is crucial for lipid antigen presentation by CD1 (46, 56).

A combination of structural and functional data has given insight into the types of lipids – both self and non-self- that CD1c can bind. Despite an increase in understanding of the types and classes of lipids that can be presented by CD1c, an appreciation of the molecular mechanism by which T cells recognise CD1c and distinguish between lipid classes presented is still limited. This is in part due to a lack of structural data available, regarding CD1c in complex with TCR – particularly γδ TCR. Current structural data of CD1c includes complexes with phosphomycoketide (PM), PC, and an interesting 'spacer lipid' bound molecule (63-65). The first solved CD1c structure contained mannosyl-β1-phosphomycoketide (MPM) and was followed later by a similar lipid, PMPS (63, 66). These structures revealed a similar folding pattern to that of MHC class l and other CD1 isoforms, with the α3 domain heavy chain non-covalently associating with β2m, and the lipid-binding groove anchored in between the α1 and α2 helices (66). CD1c contains the second largest groove area at 1780 $Å^3$, perhaps unsurprisingly as its lipid sampling diversity is vast (66). Similar to other isoforms, CD1c contains an A'

and F' pocket, allowing access to two separate lipid tails, the longer of which is loaded clockwise into the A' portal, directly below the A' roof (66). The CD1c groove contains many of the conserved features of the other isoforms binding grooves, however, there are some noticeable differences, particularly some unique 'access portals' denoted D' and E' (66). Another major difference unique to CD1c is the open F groove, allowing greater solvent access and utilising a series of hydrophobic residues to set a preference for lipid properties within this groove (66). The CD1c-PM structure showed a similar mechanism of lipid loading into the groove of CD1c, with the A' channel containing the alkyl chain (which exits via the D' portal'), and the F' channel remaining in the open conformation (63). Interestingly, the headgroup is shifted substantially compared to the MPM structure, extending further and perhaps allowing the TCR to distinguish between these two relatively similar lipids (63).

One of the more interesting features of CD1c, highlighted by Mansour et al. 2016, is the ability of the molecule to adapt conformations depending on lipid occupancy within the antigen binding groove. The structure solved here, despite showing similarities in terms of A' channel binding, demonstrated markedly different F' channel conformation because of a parallel stacking of two 12-carbon spacer lipids (65). The radical difference in lipid occupancy of the F' groove translated to a tethering of the two α-helices and subsequently a closed F' roof conformation, not seen in previous CD1c structures, which shielded the F' groove from solvent *(Figure 1.6)* (65). These structural changes in the F' channel and F' roof further impacted the molecule by opening a new cavity, referred to as the 'G portal', and essentially restricting the E' portal seen in PM and MPM structures to the point it almost disappeared (65).

CD1c conformational changes on varied lipid occupancy provide a rationale as to a mechanism by which CD1c-restricted TCRs can differentiate between loaded lipids. In terms of complexes of CD1c with TCR, only one has been solved in which the molecular mechanism of interaction pointed towards CD1c autoreactivity being dictated by molecules with small/no lipid headgroups that weren't able to contact the TCR themselves, and hence the TCR bound the CD1c protein itself without making direct interactions with loaded lipid (64).

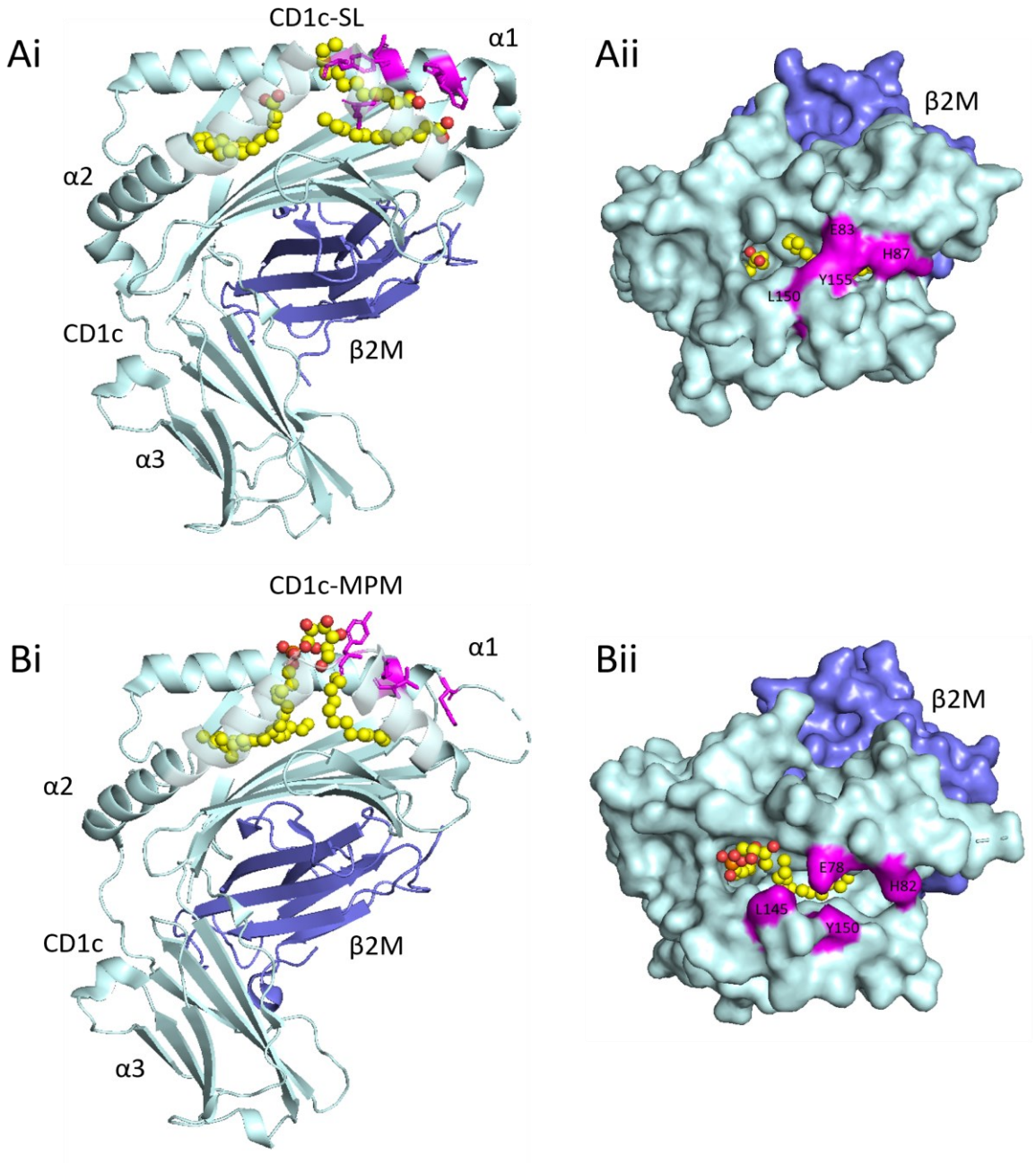


Figure 1. 6 Conformational adaptation of CD1c F' roof is dependent on lipid occupancy

CD1c can adjust its F' roof conformation depending on the occupancy of its lipid binding groove. A series of residues adjacent to the F' channel can adjust their conformation to open and close an F' roof above the F' channel. A i) The 'SL' spacer lipids (yellow) bound within the A' and F' channels of CD1c, with the four amino acids forming the F' roof (Y155, L150, H87, E83) coloured in magenta. CD1c-heavy chain is shown in light blue and β2M is shown in dark blue. ii) On-top view of the F' roof shielding the lipid constituents of the F' channel from solvent. B i) and ii) demonstrate the same things respectively for CD1c bound with MPM.

1.2.3 Group 2 CD1

1.2.3.1 CD1d antigen presentation

CD1d is the most extensively studied CD1 isoform due to wide cellular expression, structural conservation across species, and its unique presentation to iNKT cells (36). It is also widely expressed in B cells, underpinning its ability to present antigens to T cells (35, 36). The broad distribution contributes to its role in shaping the immune response via lipid antigen presentation to iNKT cells.

CD1d follows a well-characterised trafficking pattern, utilising the trans-Golgi network (TGN) for lipid exchange (67). Like group 1 isoforms, CD1d utilises AP-2 and AP-3 to internalise and traffic to late endosomes and lysosomes (68, 69). This process involves a tyrosine motif on the cytoplasmic tail which facilitates efficient antigen exchange. The importance of this has been demonstrated by a deletion leading to defective antigen presentation (70, 71).

1.2.3.2 Human CD1d

The crystal structure of human CD1d with and without αGC was solved in 2005, where the packing of the crystal lattice resulted in a 'lipid-bound' and 'lipid-unbound' form of CD1d (72). There is demonstrated overlap in recognition of both human and mouse CD1d by their corresponding invariant natural killer T (iNKT) cells and is therefore unsurprising that the sequence homology and structural architecture between the two molecules is relatively similar (65.4%) (72, 73). Consistent with that of the mouse CD1d structure, human CD1d contains a heavy chain comprising three α sub-units non-covalently linked to β_2M (72). A deep and hydrophobic lipid-binding groove was observed in between the α1 and α2 helices and formed two main channels, namely A' and F', similar to that of other CD1 isoforms (72). The overall groove size is the second smallest of the CD1 family and there are no accessory portals unlike in CD1c and CD1b (39, 74). It has a similar binding groove architecture to CD1b, however does not contain a T' tunnel due to occlusion by specific amino acid sidechains, therefore reducing the total lipid length able to bind (72). Interestingly, the acyl chain bound within the A' channel circulates the central pole in an opposite orientation to that of CD1b, a characteristic that is thought to be important in head group orientation and subsequent invariant natural killer T (iNKT) cell recognition (72). In the αGC lipid-bound structure of CD1d, the 26-carbon acyl chain and 18-carbon sphingosine chain *(Figure 1.7*) are shown to occupy the full length of the channels, implicating an ideal fit for αGC. The well-defined and rigid fit of the galactose headgroup atop of the

binding groove dictate the specific interactions between CD1d and iNKT cells, highlighting the CD1d glycolipid accommodation mechanism (38, 72). The lattice packing of this solved structure also allowed comparison of the lipid unbound form of CD1d, a feat which may not have otherwise been possible due to the inability of CD1d to refold in the absence of lipid (72). The main differences between the open and closed conformations (lipid unbound and lipid-bound respectively) are facilitated by a series of key residues which act to maintain a more open conformation of the CD1 molecule whilst preventing the complete collapse of the lipid-binding groove (72).

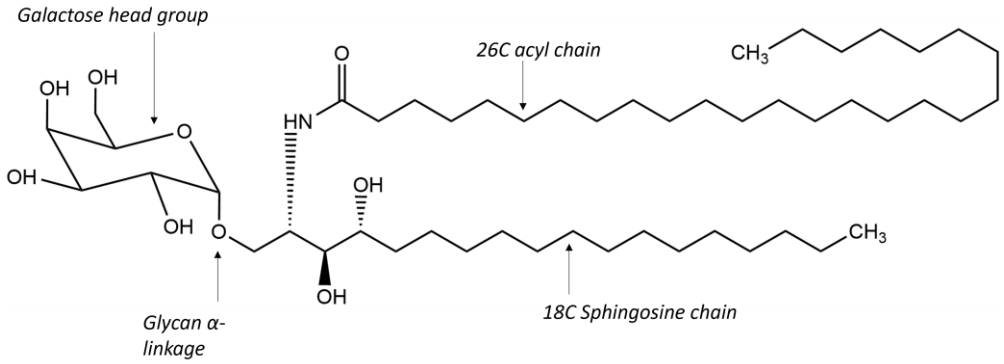


Figure 1. 7 Chemical structure of αGC

αGC is a synthetic glycolipid and a strong iNKT cell agonist, originally isolated from structure-activity relationships on marine sponges. It consists of a 26-carbon acyl chain and an 18-carbon sphingosine chain, connected via a glycan α-linkage to a galactose sugar headgroup. Image generated using ChemDraw 20.1.

1.2.3.3 CD1e - The odd one out

CD1e does not function as a classical antigen presenting molecule, instead playing a role in lipid processing and antigen loading within the endolysosomal network (39). It is expressed in both immature and mature DCs where it localises to the Golgi and lysosomes respectively; it is here it becomes a soluble protein (75, 76). Soluble CD1e facilitates lipid exchange in other CD1 isoforms, particularly CD1b. This role was first hypothesised due to the presence of soluble CD1e in CD1b+ DCs (77). The crystal structure of CD1e reveals a similar overall secondary, tertiary, and quaternary structure, despite vast differences in the lipid-binding portion of the protein (77). The first notable difference was discovered to reside within the α-helices, with a far greater gap between α1 and α2 observed in CD1e than in other CD1 isoforms *(Figure 1.8)* (77).

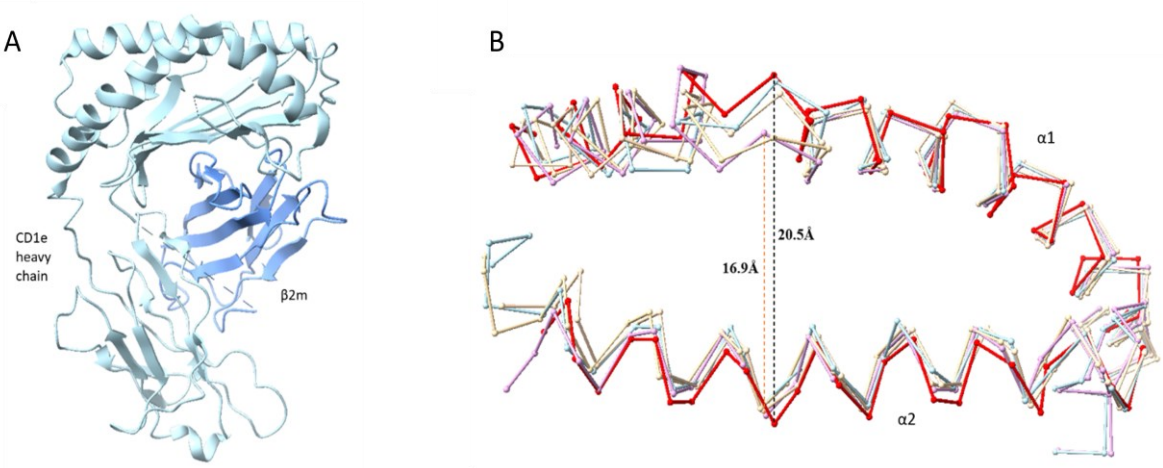


Figure 1. 8 CD1e displays a lipid-binding groove ideally suited for lipid transfer.

The quaternary structure of CD1e is structurally and visually identical to the other CD1 isoforms, with a 3-domain heavy chain non-covalently associated with β2M. Unlike other isoforms, however, the lipid binding groove is wider in CD1e than the other isoforms, with the widest part of the groove almost 21Å in diameter. A) Heavy chain shown in light blue and β2m in dark blue B) Comparison of structure and distances between the α1 and α2 helices of all five CD1 isoforms. CD1e is shown in red, CD1b in purple, CD1c in pale yellow, CD1a in blue, and CD1d in sand. There is a difference of 3.6Å between CD1e and CD1b in the wide part of the helix. Image generated using PyMOL under PDB accession codes 4X6F, 5L2J, 6C15, 1Z5L, and 3S6C for CD1a-CD1e respectively.

As expected, the highly conserved A' pocket showed only small differences in structure and was lined with predominantly hydrophobic residues. The F' pocket, however, showed far greater differences in structure and ligand binding capacity. A wider F' portal, pulled apart by the gap between the α-helices, construes CD1e's apparent ability to act as a chaperone and transfer lipids in and out of the CD1 grooves (77). Consistent with its chaperone characteristics, CD1e is found to facilitate more rapid lipid exchange in CD1b (77).

1.2.4 CD1 animal models

1.2.4.1 Murine CD1

The first crystal structure of CD1d was determined in the mouse in 1997, in which an MHC class I-like fold consisting of a 3-domain heavy chain associated with β2m was solved (38). This model was solved without a defined antigen present within the binding groove. It was suggested that the number of hydrogen bond donors and acceptors could be sufficient to engage a peptide ligand, despite the

positioning and distribution of these atoms suggesting a presentation mechanism distinct from what has been previously discovered with MHC (38, 78). The structure displayed a much deeper and narrower lipid-binding groove compared with that of MHC, predominantly due to the proximity of the two α-helices relative to one another (38). The deep lipid-binding groove appeared because of a shift in the α1 helix higher above the β sheet, resulting in a greater surface area for the groove to form (38). The groove itself consists of a series of interconnecting tunnels and pockets that function to capture and present lipid antigens to T cells, with the properties and size of the pocket determined by specific amino acid residues that line the groove (38). Specifically, an extensive hydrophobic network of interactions was found lining the binding groove, therefore dictating some of the properties required for lipid binding (38). Further mouse-CD1d structures have now been solved, including CD1d-GD3, mycobacterial lipoglycan, and CD1d-sulfatide (79-81). This also provides a rationale for the likely different binding mechanisms utilised by the lipid groove compared to that of the peptide binding groove seen in MHC I and II, which tend to use electrostatic interactions by basic residues at the bottom of the groove (38).

1.2.4.2 Macaque CD1 antigen presentation

Macaques provide a useful comparative model for investigating CD1 antigen presentation due to the strong similarities in immune systems (82). This allows for direct structural and functional comparison between species, which generates important insights into the conservation of lipid-binding properties and how CD1-restricted T cells distinguish between self and non-self. Early studies of CD1 expression in Rhesus Macaques found strong lipid presentation similarities to humans despite the divergence of lineages around 30 million years ago (83). Evidence of this conservation in structure and function was further backed up by a study published in immunogenetics in 2003, which also found the CD1d antigen presentation system is highly conserved between humans and rhesus Macaque (82). As well as group two CD1, it has also been demonstrated that Macaque group one CD1 isoforms contained the same crucial amino acid residues required for disulfide bond formation, N-linked glycosylation, and endosomal trafficking, as in the Human group one CD1 isoforms (84). On top of this, species cross-reactivity was first demonstrated between Macaque and humans through a macaque CD1b molecule presenting a mycobacterial glycolipid to human CD1b-restricted T-cells (84). The similarity in function between Human and Rhesus CD1b was also demonstrated through the conservation of specific amino acids thought to be fundamental for binding of specific TCRs (84).

1.3 T cell recognition of CD1

For T cell activation to be successful, antigens must be recognised precisely and specifically by TCRs. This recognition can be broadly categorised as MHC-restricted and MHC-non-restricted, the latter of which is described by CD1 molecules mediating lipid antigen presentation to both αβ and γδ T cells.

Multiple mechanisms have been described for interaction between antigen-presenting molecules, lipids or peptides, and TCR. These include the more conventional co-recognition model, as well as an unconventional left-right mismatch, buried ligand model, absence of interference, and more recently an atypical sideways recognition. A well-documented example of the co-recognition mechanism existing outside the MHC-peptide paradigm is the PG90-TCR interacting with and binding to CD1b loaded with PG (58). This interaction involves direct TCR-lipid contacts, where the CDR3β loop acts as a molecular tweezer, pinching and stabilising the lipid headgroup *(Figure 1.9 ii*) (58).

A similar tweezer mechanism of action was also seen in the glucose monomycolate (GMM)-TCR-CD1b structure in which the TCR CDR loops pinch together the GMM headgroup to stabilise the complex (60). There are also other co-recognition examples within the CD1 family, such as CD1d-bound phospholipids (66, 85, 86).

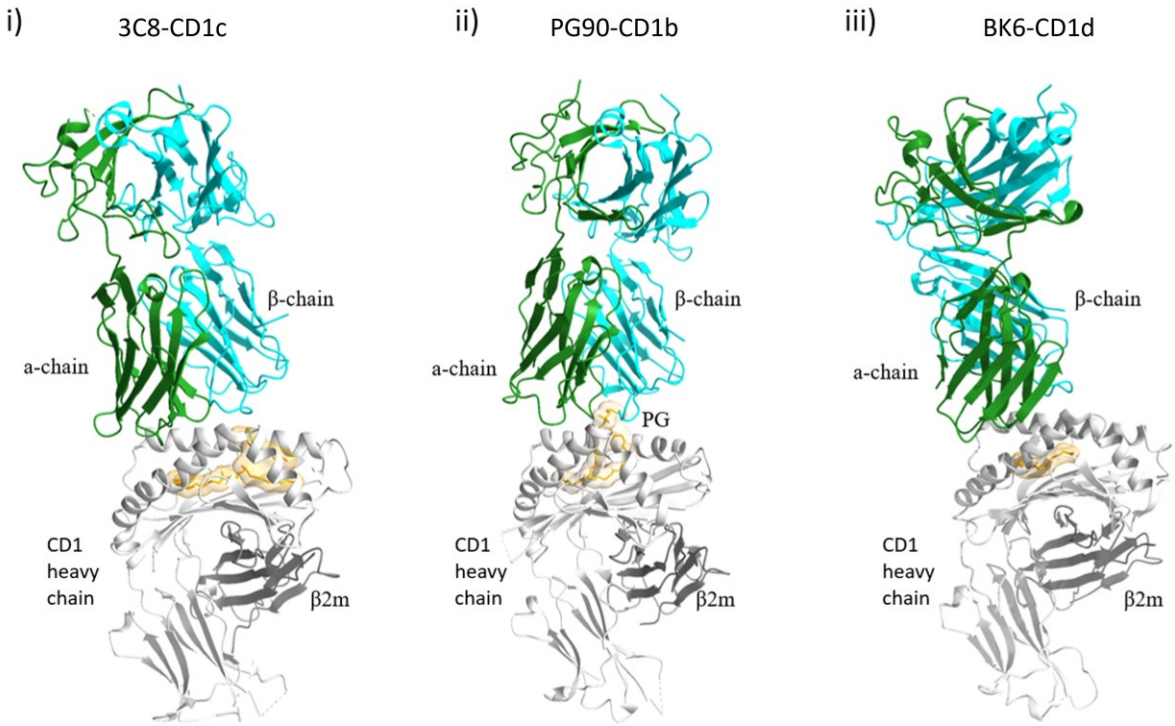


Figure 1. 9 Comparison of three αβ TCRs in complex with CD1

CD1 reactive αβ TCRs employ different mechanisms to recognise and interact with CD1- antigen complexes. Recognition mechanisms include the buried ligand model (absence of interference) (i), head-group discrimination (ii), and left-right mismatch (iii). Images i and iii differ from ii in that they don't interact with lipid antigens, and instead contact only the CD1 protein surface. TCR α and β chains are shown in dark green and cyan respectively. CD1 protein is shown in grey and lipid antigen in yellow. Images were generated using PyMOL under PDB four letter codes 6C09, 5WKI, and 4X6D from left to right respectively.

Absence of Interference

Interestingly, not all TCR CD1 interactions require direct lipid contact. Lipid agnostic binding has been documented, with two main mechanisms utilised by different TCRs. Left-right mismatch *(Figure 1.9, iii)* is demonstrated by the co-crystal structure of an αβ TCR in complex with CD1a, published in 2015 (51). This study included structural data of a BK6 TCR both on its own and in complex with CD1a-endo, and CD1a-LPC, to evaluate how changes in lipid antigen impacted TCR binding and mechanisms of CD1a antigen presentation to this autoreactive TCR.

In this CD1a-restricted TCR interaction, the BK6 TCR docks orthogonally over the A' roof of CD1a, without directly engaging the lipid antigen (Figure 1.9 iii). This interaction is highly sensitive to lipid-induced conformational changes in CD1a, as different lipid cargoes modulate the positioning of CD1a

residues critical for TCR binding (51). A well-conserved orthogonal docking mode was displayed between the two co-crystal structures, suggesting that autoreactivity and interaction mechanism by the BK6 TCR did not drastically change in the presence of alternative lipid antigens. Specifically, sulfatide loading within CD1a disrupts TCR contact residues Asn151 and His153, thereby reducing TCR affinity (45). This suggests that CD1-bound lipids can indirectly influence TCR recognition via structural remodeling of the CD1 platform.

Buried Ligand Model

A third mechanism (also lipid agnostic), the buried ligand model *(Figure 1.9, i)* is demonstrated by the co-crystal structure of CD1c-3C8 (αβ) TCR. The 3C8 TCR provided the first knowledge of an αβ TCR autoreactive to CD1c, creating interest as to whether the mechanism was like that of the previously solved CD1a-BK6 TCR complex. Structural data revealed a central binding footprint over CD1c with a TCR footprint that is predominantly A' roof mediated. Part of the TCR β chain bound to a portion further right of this platform over the F' portal, occluding the access portal from which lipids would normally protrude (64). The covering of this F' portal was predominantly carried out by the CDR3β loop, allowing it to make multiple wan der Waals and hydrogen bonds with CD1c (64).

The β chain of the TCR essentially acts as a molecular plug, preventing the escape of lipid head groups from the surface. This generates a question as to whether a lipid with a larger and therefore protruding headgroup would eradicate TCR binding or would adopt an induced fit mechanism whereby the TCR β chain and lipid head group adapt to accommodate each other. An example of this would be bacterial PM, a lipid known to be presented by CD1c and jut out of the F' portal to contact certain TCRs (66). This lipid, alongside others with large or charged headgroups, demonstrated functional blocking of TCR binding to CD1c using tetramer staining assays, providing evidence that it was indeed the lack of protruding headgroup allowing binding between the two molecules.

1.3.1 Unconventional T cells

Distinct from those T cells operating within the MHC-peptide paradigm, a subset of unconventional T-cells recognise non-peptide antigens through the presentation by monomorphic antigen-presenting molecules, including CD1. These T cells are described as bridging the gap between innate and adaptive immunity and are generally characterised by their donor-unrestricted nature and ability to be recruited rapidly to infection sites (87). Multiple reports suggest that these unconventional cells, unlike the normal 'conventional' T cells, circulate the body as populations with limited diversity and

are ready to initiate an effective and relatively specific response to a challenge, rather than the usual 7-10 day delay due to clonal expansion (87, 88).

Several subsets of unconventional T cells have been identified, including mucosal-associated invariant T (MAIT) cells, MHC-related protein 1 (MR1)-restricted T cells, CD1-restricted T cells, and γδ T cells, each with distinct antigen recognition properties. Among these subsets, invariant natural killer T (iNKT) cells and γδ T cells play key roles in lipid antigen recognition. These cells interact with CD1 molecules, which present self and foreign lipid antigens, and their ability to respond to lipid-based antigens suggest a unique mode of immune surveillance.

MR1 on the other hand is a different type of MHC-like antigen-presenting molecule and presents bacterial metabolites and self-antigens to MAIT cells and MR1 T cells respectively in the context of infection and cancer (89). It has been suggested that unconventional T cells may compete amongst each other for space and specific microenvironments within the body, an example of this being γδT17, NKT17, and MAIT17 cells, all of which utilise IL-17 for survival and expansion (88, 90). It has also been demonstrated that γδ T cell-deficient mice show an increased number of MAIT cells, again suggesting the two populations may be competing (91).

While MAIT and MR1-restricted T cells have been extensively studied in the context of bacterial metabolites, this thesis focuses on the structural and functional aspects of CD1-lipid-TCR interactions, particularly those involving CD1-restricted T cells. The following sections will explore the molecular basis of CD1 antigen presentation, with a specific emphasis on iNKT and γδ T cell recognition mechanisms.

1.3.2 iNKT cells

iNKT cells are an unconventional subset of T cells that can be subdivided into type l (classical) and type ll (non-classical) iNKT. They specifically recognise lipid-bound CD1d molecules. They are considered an 'innate-like' subset that can bridge the gap between innate and adaptive immunity, whilst also being described as attractive targets for cancer immunotherapy due to their proinflammatory and tumour-killing properties. They consist of a population of T lymphocytes that display properties of both T cells and NK cells, defined by specific cell surface markers and characteristic traits (92).

Their TCR is semi-invariant, with the α chain consisting of a Vα24-Jα18 chain that is paired with a Vβ11 chain (93). In mice, the chain pairing is slightly different, with a Vα14-Jα18 chain pairing with one of three β chains (Vβ8, Vβ7, or Vβ2) (93). Although the structures are similar, the functional differences

between human and mouse iNKT are not so well defined, limiting translation of murine studies to clinical success in humans (94-96)

1.3.2.1 iNKT Cell-Mediated Immunity and Lipid Recognition

There are three main mechanisms by which iNKT cells mediate anti-tumour immunity, the first being recognition of CD1d presented antigens on tumour cells, the second being cytokine release and recruitment of other cell types, and the third being recognition of CD1d loaded lipid antigens on APCs (97). Structural studies reveal a relatively conserved iTCR docking mode, with the CDR3β loop functioning to fine-tune specificity (97-99). The ability of iNKT cells to distinguish between lipid variants has been explored using αGC analogues, some of which selectively skew responses towards TH1 or TH2 immunity, offering potential immunotherapeutic applications

1.3.2.2 Mouse iNKT

Murine iNKT cells have undergone extensive analysis to fully understand the relationship between TCR affinity and lipid specificity. Initial crystallographic studies in mice (Borg et.al 2007) aimed to dissect the molecular interaction between CD1d-αGalCer and iNKT cells with different Vβ chain usage. This study showed the interaction between a Vβ8.2 and Vβ7 TCR with CD1d-αGC and demonstrated a very similar docking footprint between the two, with both binding parallel and above the F' pocket to contact both the lipid antigen headgroup and CD1d itself (100). The main differences between the two docking modes were provided by the differential Vβ chain usage, subsequently conferring different contacts via the CDRβ loops and thus affecting the affinity of the TCR for CD1d-αGC. Differences in structure and sequence of the Vβ chains resulted in differential contact contributions in which the Vβ7 contacted CD1d more than Vβ8.2 (100). Mutagenesis data demonstrated that the Vβ7 chain was non-optimal for interaction with CD1d, therefore providing rationale as to the higher affinity binding seen by the Vβ8.2 TCR (127).

Subsequent studies (Wun et.al 2011) investigated how structural variations in lipid antigen effected the affinity of the iTCR for CD1d. The lipid antigen repertoire of CD1d is vast with examples including bacterial lipids such as glycosphingolipids (GSLs), glycosylceramides, PI mannosidase, and α-galactosyl-diacylglycerol, as well as mammalian lipids such as isoglobotriosylceramide (iGb3), and many derivatives of αGC (78, 80, 101-104). Many analogues of the strong agonist αGC have been generated, with many causing different iNKT cell responses.

A study published in 2011 by Malleavoy et al. demonstrated that recognition of CD1d-lipid complexes is influenced by specific CDR3β sequences (105). An in-depth analysis of specific CDR3β sequence elucidated a hydrophobic leucine/isoleucine occupying positions one and two of the CDR3β loop in those TCRs which were able to interact with CD1d, therefore mediating stronger autoreactivity towards CD1d. Further investigation of the binding mechanism between CD1d-self lipid and the iTCRs showed that despite a similar docking mechanism to that seen in previous CD1d αGC-iNKT co-crystal complexes (parallel, tilted and above the F' pocket), the two leucine residues at the top of the CDR3β loop played a crucial role in sequestering a surface exposed portion of the binding cleft, thus encouraging autoreactivity (105). This explains why preferential usage of leucine/isoleucine at positions one and two of the CDR3β was seen in the sequences conferring reactivity to CD1d-self lipid in the previous experiments.

Further structural studies indicated that small modifications to head groups and lipid tails can significantly influence TCR binding. Analogues such as OCH do not bind as strongly to the TCR due to chain modifications impacting the stability of the headgroup within the CD1d binding cleft (106). These alterations also impact cytokine response, despite the TCR having a similar binding footprint overall, due to lipid stability affecting headgroup orientation and subsequent TCR binding avidity and activation (107-110).

1.3.2.3 Human iNKT

Despite many clinical trial efforts and interesting discoveries within the mouse models of iNKT-cell-based studies including some extremely successful pre-clinical data, translation to humans has been far less promising. There are several reasons for this discrepancy, including different TME capabilities between humans and mice, different subsets of T cells, and variations in cytokine/chemokine profiles between the two species. The major difference between mouse and human iTCRs is that mice express three different versions of the Vβ chain, whereas humans express only one. The Vα24 chain is invariant, with the only variable section of the human iTCRs residing within the CDR3β loop (111). This suggests differences in TCR affinity that are independent of lipid occupancy within the CD1d binding groove, are likely to be modulated by sequence variations within the CDR3β loop. The two main factors considered when evaluating CD1-NKT biology, are the affinity of the lipid antigen for CD1d, and the affinity of the NKT TCR for the CD1d-antigen complex.

A study on human iNKT cells in 2007 (like that of the mouse study by Wun et al.) aimed to understand the molecular mechanism used by human iTCRs to distinguish between lipid antigens of varying chain

lengths and saturations (112). Recent crystal structures between CD1d and iTCRs demonstrated a conserved TCR binding footprint, consistent with the recognition of the lipid head group and antigen binding domains of CD1d, but not contacting the acyl or sphingosine chain directly (113). This suggests some degree of conformational plasticity between the lipid chain and polar headgroup, with re-orientation of the headgroup affecting the ability of the TCR to dock and stabilise. The study essentially found that shorter sphingosine chains (occupying the F' channel') decreased the affinity of the iTCRs for CD1d, as well as influencing important functional aspects of the T cell such as activation and synapse formation (113). Understanding whether the CDR3β loop plays a role in antigen specificity of CD1d interaction is important in understanding how the iTCRs can adapt to NKT cell response in different environments.

It has been demonstrated that human iNKT cells utilise the CDR3β loop to make important interactions with the CD1d protein, therefore suggesting its role in overall TCR affinity rather than lipid specificity (114). The CDR3β loop is known to be dispensable for CD1d-αGC binding, due to the optimal binding occupation of this lipid within the binding groove, however, the use of OCH tetramers generated distinct subset populations of iNKT cells capable of fluorescing to different intensities (114). Similar sub-populations were seen when staining these iNKT cell lines with CD1-self lipid (β-glucosylceramide) suggesting this difference in mean fluorescence intensity (MFI) was physiologically relevant (114). Comparisons of separate sub-groups i.e. OCH-High, and OCH-low, confirmed these differences, with OCH-high clones displaying greater cytokine release, proliferation, and cytotoxicity compared to the OCH-low clones (114).

A more recent study has shown that specific residues within the CDR3β loop are able to confer reactivity of the TCR to CD1d (115). Initially, a pan anti-Vα24MAb was used to stain Vα24 chains reconstituted with different Vβ11 chains expressing different CDR3β loops, two of which showed visibly lower staining despite similar CD3 expression levels. The difference in these two clones was their usage of Arginine at the third position within the CDR3β loop as opposed to serine or threonine, therefore indicating that an Arginine at position three facilitated structural rearrangement of the Vα24 chain and decreased staining by the mAb antibody (115). It has been demonstrated in previous studies that the CDR3α and CDR3β loops can form interactions with each other, often mediated by water molecules (116). This provides a rationale for the altered sequence of the CDR3β loop being able to directly alter the CDR3α loop and subsequently affect the docking of the TCR to the CD1d surface.

1.3.2.4 Macaque iNKT

Despite the extensive use of the Murine system to study human disease, differences in protein expression and TCR usage have provided limitations for using this system in the context of human disease. For this reason, other non-human primate systems have been studied, such as the Macaque. The Rhesus Macaque model is considered a very useful system for studying human disease and infection, due to its commonalities with the human immune system. Unlike mice, Rhesus Macaques have been shown to present group one CD1 molecules, however, less is known about their CD1d antigen presentation capability to iNKT cells.

An early study in 2003 used α-GC and IL-2 to demonstrate that Rhesus Macaques possessed Vα24TCR positive T cells, with very high sequence similarity to their human counterparts (117). Vα24+ T cells from Rhesus Macaques use the same CDR3 region as human iNKT cells but a different Vβ11 chain (118). This high sequence similarity is also consistent when comparing CD1d molecules between human and Rhesus, suggesting their antigen presentation mechanism is likely to follow a similar pattern (117). Likewise, group one CD1 molecules and CD1e have also been shown to be present in Macaques (119, 120). Transspecies activation has been demonstrated across the CD1 isoforms when activating human T cells, again suggesting high levels of similarity between the molecular recognition mechanism of human and Macaque T cells and CD1 (119). Macaque CD1b can present mycobacterial glycolipids to human T cells, such as GMM. A more recent study demonstrated that GMM presentation to T cells following BCG vaccination may be mediated predominantly via CD1c (121).

1.3.3 γδ T cells

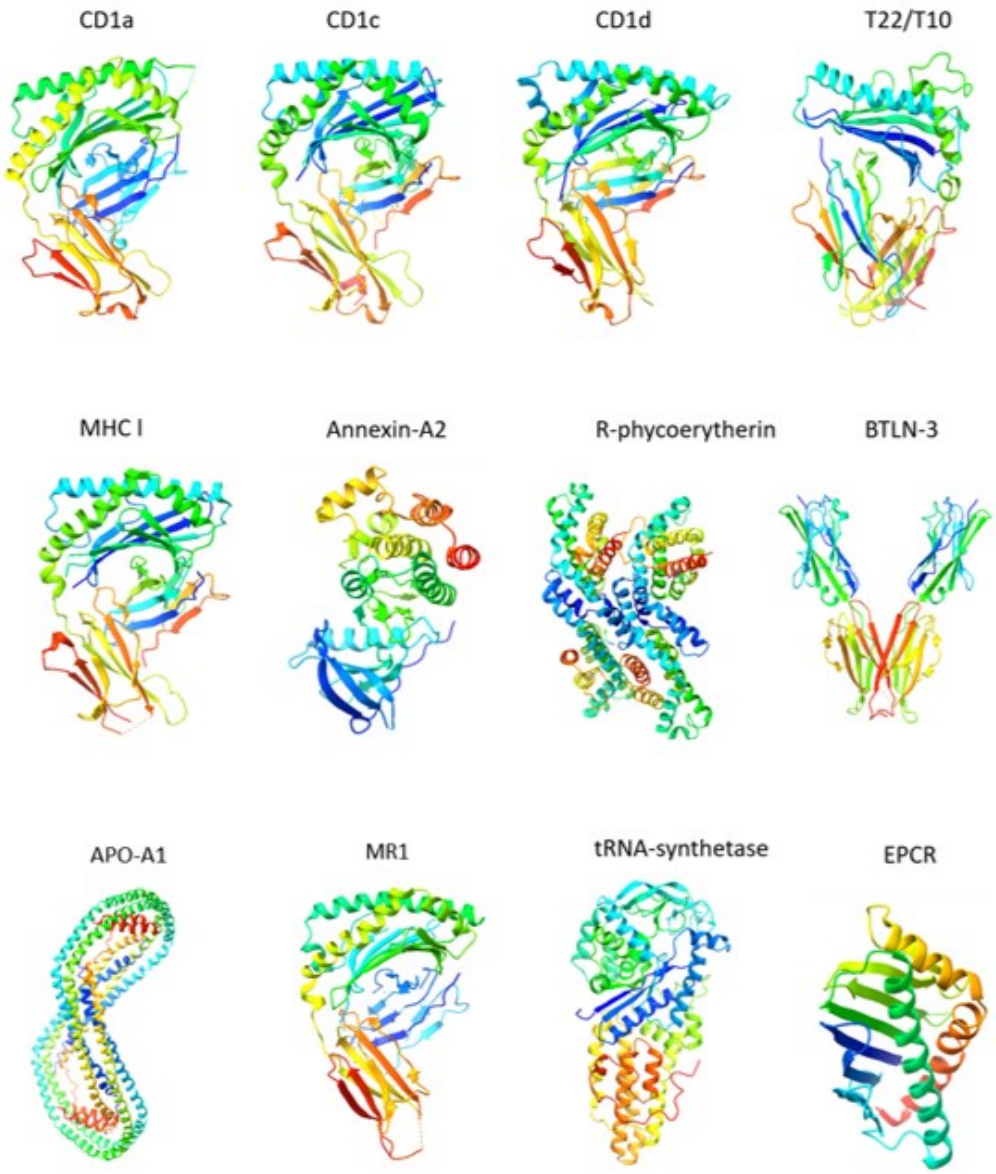


Figure 1. 10 Currently known γδ ligands show some variety but share common structural features.

Crystal structures of the currently known γδ ligands are displayed in cartoon rainbow orientation. Ligands display similar structural features and are generally dominated by α-helical folds sat above β sheets. All images were generated using PyMOL.

γδ T cells are a subset of T cells that reside in relatively small numbers in the peripheral blood and have been shown to demonstrate both innate and adaptive immune function. Due to them

demonstrating aspects of both the innate and adaptive immune system, γδ T cells are implicated to have roles in diverse disease states, including many different cancers as well as autoimmune diseases and infections (122). These T cells comprise roughly 2% of total T cells in humans and are enriched in many peripheral tissues such as lungs, skin, and intestines (123). They have been established as inflammation-promoting cells in the myeloid and lymphoid lineages, while also showing strong cytokine production ability and cytotoxic potential, including the ability to secrete large amounts of IFN-γ and IL-4 in response to pathogenic challenges (122, 124). They can directly lyse target cells, whilst also recruiting other types of immune cells (e.g., neutrophils) and acting as APCs. It is this combination that makes them an interesting target for cancer immunotherapy. As well as this, they can recognise tumour-associated ligands in an MHC-unrestricted manner, potentially unlocking a new class of tumour-associated ligand targets that αβ T cells are unable to recognise due to MHC restriction (125). Despite their apparent strong antitumour capability, success in clinical settings of cancer immunotherapy has been very limited. This is likely due to the polarization of antitumour γδ T cells into tumour cells within the tumour microenvironment (TME), making it difficult to control the specific repertoire of T cells present around the tumour (122).

The variety created by different γδ chain pairings is likely the reason they display diverse activity in many different cancers. There are four main categories of γδ T cells, denoted by their Vδ gene usage: δ1, δ2, δ3, and δ5. Each of these can pair with different Vγ chains to form an array of diverse TCRs (122, 126). In humans, Vγ2, Vγ3, Vγ4, Vγ5, Vγ8, Vγ9, and Vγ11 variations of the γ chain have been discovered, whilst in mice there are several more including Vγ1, Vγ5P, Vγ6, Vγ7, and Vγ10 (122, 127).

γδ T cells develop in the thymus, leaving as mature cells that do not require TCR signalling to display properties associated with differentiation and subsequent function (128). It is well documented that γδ T cells do not require MHC restriction to function. Evidence for this lies in the fact that their development within the thymus, unlike αβ T cells, is not affected by a lack of MHC class ll or β_2m (129, 130). In terms of tissue expression, γδ T cells generally demonstrate tissue-specific localisation of sub-populations, with limited diversity. Examples of this include skin epidermal γδ T cells predominantly expressing a Vγ5Vδ1 TCR, while the majority of γδ T cells found in the blood express a Vγ9Vδ2 TCR (131, 132). Ligands for γδ are relatively unknown, with a few under current study including MICA, MICB, phosphoantigens, and CD1 (133, 134). These cells have been found to exhibit several different phenotypes and functions, depending on multiple factors, including tissue localization and specific conditions or environmental stimuli from neighbouring cells. It is perhaps this functional plasticity and ability to represent any one of the TH1, TH2, Treg, and TH17 phenotypes that makes this subset of T cells so interesting yet challenging to work with (128). More recently IL-17 IL-17-producing γδ T cells have come under some scrutiny, with known roles in infection, autoimmunity, and antitumour

response (128). Certain characteristics of these cells differentiate their influence substantially from their αβ counterparts, namely the strength and size of their response and their ability to take effect without the need for antigen specificity (128). Examples of the potential negative effects of IL-17-producing γδ T cells have been demonstrated in mice, where the γδ T cells infiltrating the tumour were the prominent source of IL-17, increasing angiogenesis in the tumour microenvironment (135).

1.3.3.1 Vδ1

Humans contain two major subsets of γδ T cells, classified predominantly by their Vδ chain. The first of these is the Vδ1 subset, generally found in greater numbers within the thymus and peripheral tissue including the intestinal epithelium and epidermis (125, 136). Their presence in these types of tissue is thought to prime them for an initial immunosurveillance role against certain cancers, partly due to their known interaction with MHC class-1 related molecules which are upregulated on infection and in cancerous tissue (137, 138).

This subset of γδ T cells is no different in terms of its cytotoxic potential, demonstrating the ability to express key soluble mediators such as perforin and granzyme, as well as CD95L and tumour necrosis factor (CD27). Both are known to be important in influencing effector function by facilitating cellular death (136, 139). Specific cancers that this Vδ1 subset has shown anti-tumour capability against include lung, kidney, breast, and colon carcinoma, as well as melanoma. In these cancers, tumours expressing MICA and MICB were targeted by tumour-infiltrating Vδ1+ γδ T cells secreting IFN-γ and TNF-α and demonstrating cytotoxicity (140). Despite some knowledge of receptor expression, including natural killer group 2D (NKG2D) which is critical for some cancer cytotoxicity, and a vague understanding of potential tumour/infection responses, the mechanisms by which Vδ1 facilitate these responses are largely unknown. This is partly because their lack of presence in the peripheral blood makes them difficult to isolate and expand.

Despite this gap in knowledge, a study in 2015 demonstrated that tumour-infiltrating γδ T cells provide the most favourable anti-tumour immune population of any immune cells, in several different cancers (141). Briefly, the study analysed a large variety of human cancers in a large patient cohort (18000) and determined that tumour-infiltrating γδ T cells demonstrated the most favourable prognosis. This could be attributed to CD161 expression, a characteristic marker of innate immune function across a variety of T cell subsets (141, 142). Vδ1+ cells generally display limited diversity, a characteristic that is understood to represent their specificity to tumour antigens that are associated with cellular stress (125).

In terms of cell surface expression and receptor expression, Vδ1 T cells have a distinct profile compared to that of its Vδ2 counterpart, which is thought to dictate its mechanism in a tumour setting. This difference also allows preferential targeting of individual γδ subsets (136). Vδ1+ do not preferentially combine with a specific Vγ chain and yet have not displayed any reactivity to isopentenyl pyrophosphate (IPP) or amino-bisphosphonates (N-BP) (intermediates of the mevalonate pathway) (143). They also can circulate for a long time and resist activation-induced cell death, unlike their common Vδ2 counterpart (143). This is generally considered a positive trait; however, it does risk a longer exposure time for polarization to a tumour-promoting subtype. Despite their demonstrated tumour-killing capabilities, Vδ1 cells are activated in myeloid and lymphoid leukaemias. More recently, they have been shown to exhibit immunosuppressive characteristics in some disease settings (143).

As previously mentioned, Vδ1 can exhibit both anti-tumour and pro-tumour capabilities, with the latter emerging more recently. Tumour-infiltrating Vδ1 lymphocytes have been characterised; however, their mechanism of action and whether they promote or inhibit tumour growth is less well understood. Some studies suggest Vδ1 cells express high levels of the CXCR1 receptor, indicating a potential interaction with IL-8, a chemotactic factor associated with neutrophil recruitment and advanced disease states (144, 145). A different study found that stimulated Vδ1+ cells expressed high levels of CCR4 and CCR8, both of which recognise ligands expressed by lymphoid-derived malignancies (143). Further research is needed on chemokine expression and receptors in different tumour settings. Furthermore, a more detailed understanding is crucial for targeting specific sub-populations of T cells in the context of diseases such as cancer. Once localised to the TME, Vδ1 cells demonstrate both pro-tumour and anti-tumour responses depending on the context. The first studies demonstrating the anti-cancer potential for γδ T cells were in lymphoblastic leukaemia and acute myeloid leukaemia, where patients displayed significantly greater disease-free survival following αβ-depleted bone marrow grafts, leading to an increase in Vδ1+ T cells (146). Others include a Hodgkin's lymphoma study in which a raised Vδ1+ population in tandem with increased IL-4 led to better disease prognosis after one year. In a neuroblastoma study, Vδ1+ cells cultured with a human neuroblastoma cell line led to a decrease in tumour-promoting factors, an increase in TNF-α, and maintenance of IL-2 (147).

More recently, Vδ1+ T cells have been shown to have tumour-promoting properties in different tumour settings. The TME can polarize cells from one phenotype to another, sometimes resulting in a skew towards regulatory T cells. In some cases, Vδ1+ cells can induce tumour progression through means of IL-17a release, subsequently causing inflammation in a myeloid differentiation primary response 88 (MyD88)/ interleukin-23 (IL-23) dependant manner (148). Another mechanism by which Vδ1+ cells can indirectly promote tumour growth is through TGFβ secretion. This stimulates epithelial cells to undergo a transition in which they become mesenchymal cells, essentially creating an immune

evasion mechanism for the cancer (149). The ratio of specific subsets within the TME is critical in the facilitation of a successful immune response. It has been reported that an IL-4-mediated increase in Vδ1 cells relative to Vδ2, can increase IL-10 production by Vδ1 cells, in turn inhibiting Vδ2 further and modulating the cytokine balance. This then leads to an increase in IL-10, a decrease in IFN-y, and downregulation of NKG2D, an important co-marker in the immune response (122, 150). A common mechanism utilised by tumour cells to drive a tumour-enhancing TME is the inhibition of cells presenting a tumour-suppressing phenotype. An example of this is shown in breast infiltrating Vδ1 cells, which can prevent CD4+ and CD8+ T cell activation through inhibition of the antigen-presenting function of DCs (150). This mechanism is also demonstrated in pancreatic carcinoma, where high programmed cell death ligand 1 (PD-L1) expression leads to a reduction in the infiltration and effectiveness of αβ T cells. The removal or blockade of these γδ T cells was found to re-establish tumour immunogenicity (151).

1.3.3.2 Vγ9Vδ2

Another subset of γδ T cells with diverse antigen recognition is the Vγ9/Vδ2 population. These cells have been shown to respond to phosphoantigens and are well studied due to their distinct TCR-mediated activation mechanisms (152). They also expand in response to infection, sometimes reaching greater than 50% of peripheral T cells in a relatively short time post-infection (153). Vγ9Vδ2 T cells are well renowned for responding to metabolites produced by the mevalonate pathway, commonly referred to as phosphoantigens. Examples include (E)-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate (HMB-PP), IPP, and dimethylallyl pyrophosphate (DMAPP) (154). These by-products are generally found in gram positive and gram negative bacteria, as well as being over-synthesized by metabolically overactive tumour cells, potentially flagging them as targets for the localized Vγ9Vδ2 T cells (154). The killing mechanism and effector function of these cells is varied, with direct killing recruitment and priming of other cell types providing two mechanisms of action (155, 156).

Vγ9Vδ2 T cells are significant contributors to IFN-y and TNF-α production once stimulated. They facilitate the stimulation of cytotoxic activity and immature DCs (154). Once activated, these DCs can alter their chemokine production to that of 'γδ homing' chemokines, driving the process of antigen presentation to T cells in the lymphoid tissues to initiate a response (154). The Vγ9Vδ2 subset utilises the perforin granzyme pathway (sometimes dependant on the NKG2D receptor) to eradicate tumours in a multitude of tumour types, including renal cancer, breast cancer, and head and neck carcinomas (157-159). They are also able to induce tumour cell death through the tumour necrosis factor-related apoptosis-inducing ligand (TRAIL) and fas ligand (FasL), demonstrating a reduction in efficacy on the blockade of these pathways and an enhanced function when these ligands are upregulated (160). The

final mechanism by which Vγ9Vδ2 T cells have demonstrated killing is antigen-dependant cellular cytotoxicity (ADCC), in which ovarian carcinoma models have been used to show a tumour infiltrating Vγ9Vδ2 cells conjugated to IgG1 Fc can inhibit the growth of cancer (161).

Similarly, to that of its Vδ1 counterpart, Vγ9Vδ2 has demonstrated some pro-tumour function in recent years, initially in mice and more recently in a human setting. This has been shown to occur via both direct and indirect mechanisms (162). The ability of T cells to polarize and change phenotype in the presence of certain cytokine environments creates difficulty in targeting specific subsets for immune function and magnifies the importance of understanding both interaction mechanisms and ligand repertoires of these T cells and their surface receptors. Stimulation of Vγ9Vδ2 cells with IL-15 and TGF β in vitro triggered a polarization into FOXP3 γδ Treg cells, in turn causing a reduction in peripheral blood mononuclear cell (PBMC) proliferation (162, 163). This Vδ2 population also has negative consequences in rheumatoid arthritis, with IPP recognition leading to IL-17 secretion and disease aggravation (164). IL-10 and transforming growth factor beta (TGF-β) are two key players in the immunosuppressive function of γδ T cells with IL-10 predominantly working via inhibition of DC function to present antigen. More recently, IL-9 was presented as a potential pro-tumour cytokine, despite its known function in anti-tumour response. It was shown the Vγ9Vδ2 T cells can produce IL-9 in vitro, which can in turn promote lymphoma cell proliferation (165, 166). The immunomodulatory function of γδ T cells demonstrates vast functional plasticity, making it difficult to harness therapeutic potential from these cells, particularly in a difficult-to-predict TME.

1.3.4 CD1 restricted γδ T cells

Understanding the molecular interaction between γδ T cells and their ligands is pivotal in understanding their unique antigen recognition mechanisms. Structural data of γδ T cells in complex with CD1 remains limited, with further investigation required to fully elucidate the principles underpinning their ligand specificity and downstream signalling. Structural data of γδ T cells in complex with the antigen-presenting molecule CD1 is currently very limited, and far less understood than that of its αβ counterpart. Despite this gap in structural knowledge, progress has been made in other areas, with CD1 tetramers proving an important tool for investigating peripheral blood isolated γδ T cells that are restricted to the CD1 protein (167). As a result of this, the CD1 protein family is becoming an increasingly studied antigen-presenting molecule for γδ TCRs, with the majority of these TCRs belonging to the Vδ1 subset (167). This may suggest that the binding mechanism between this subset and CD1 may be conserved, however, more studies need to be conducted before this

relationship can be confirmed. Interestingly, the few known targets of γδ T cells are varied in terms of size and structure, suggesting these T cells can adapt to better interact with their target ligand.

The first co-crystal structures regarding a γδ T cell and CD1 molecule were achieved in 2013. Both structures consisted of a Vδ1+ TCR interacting with CD1d with a bound glycolipid (168, 169). Both TCRs were isolated from populations of autoreactive Vδ1+ cells and provided the first insight as to how γδ T cells can bind CD1d. The first of these TCRs, referred to as 9C2, bound orthogonally above the CD1 molecule at the far end of the A' pocket, with most contacts provided from the δ chain through the CDR1δ and CDR3δ loops (168). Key contacts across the CD1d-αGC antigen-binding cleft were made via the CDR3δ loop, primarily with residues spanning the α1 and α2 helix (168). The mechanism of binding is described as a 'lock and key' due to very limited conformational change in either the main chains or CDR loops upon contact with CD1d when compared to the unbound TCR (168). In terms of lipid reactivity, the CDR3γ chain mediated contact through the galactosyl headgroup, which is interesting as this loop was variable in sequence amongst CD1d-αGC restricted TCRs and could suggest it plays an important role in the affinity of different TCRs for the CD1d-lipid complex (168).

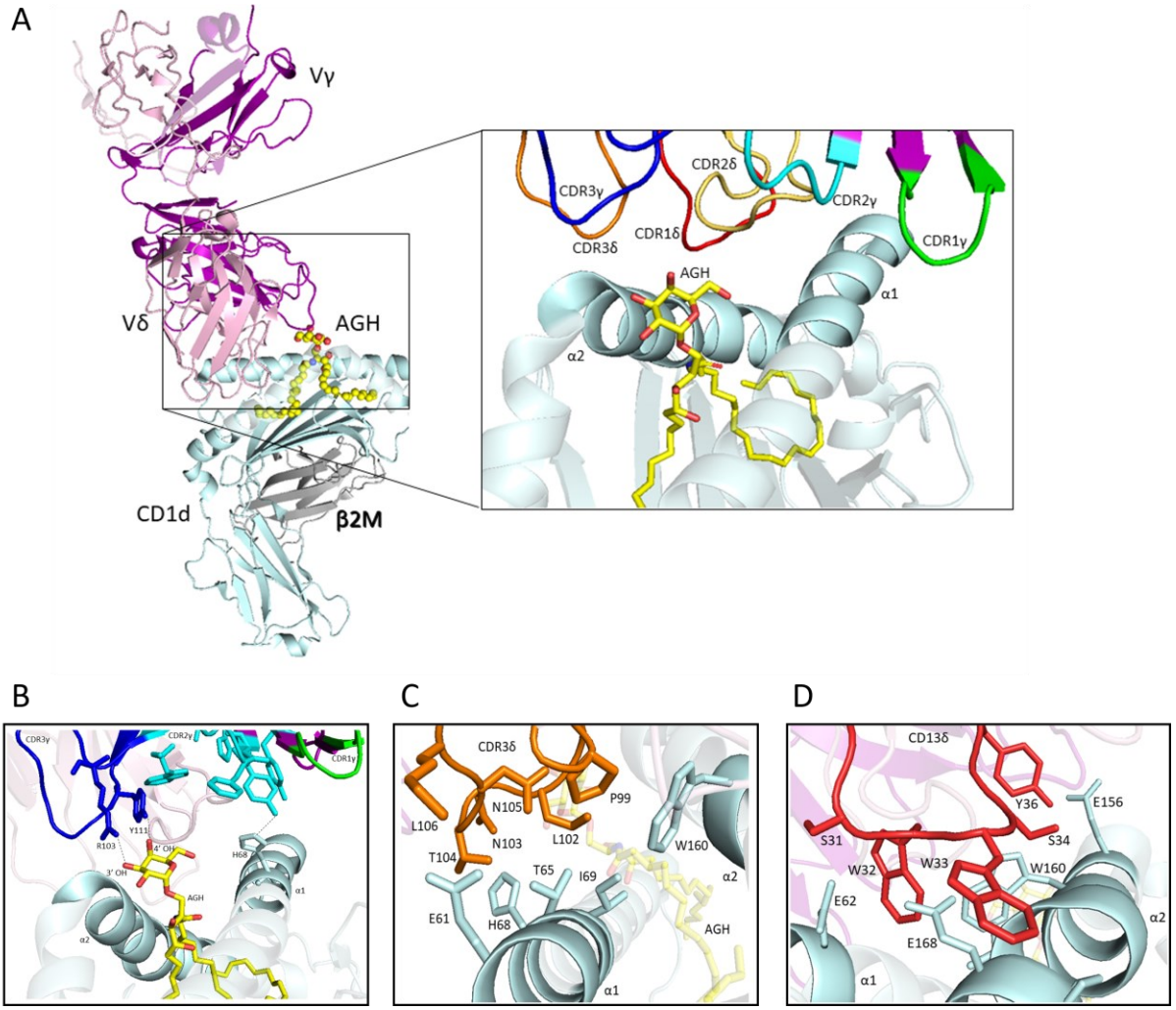


Figure 1. 11 9C2 γδ TCR in complex with CD1d-αGC

A co-crystal structure of the 9C2 γδ TCR in a complex with CD1d-sulfatide displays an unusual binding footprint compared to other known TCR-CD1 complexes. A) Orthogonal docking footprint demonstrates a CDRδ dominated interaction with the γ chain contacting lipid and shielding CDRδ loop. B) CDR3γ chain (dark blue) contacts the αGC head group (3' and 4'OH) but does not contact the CD1 protein itself. CDR2γ loop contacts H68 on the α1 helix and stabilises the Vδ chain. C) CDR3δ loop (orange) heavily contacts the CD1 α1 helix and contacts a tryptophan on the α2 helix. D) CDR1δ heavily contacts the α2 helix via a series of aromatic residues. CD1 is shown in cyan, β2M in grey, Vγ in purple and Vδ in pink. Lipid antigen is shown in yellow.

Another example of a CD1d- γδ TCR interaction demonstrated through structural data is the DP10.7 TCR, which recognises CD1d loaded with a Sulfatide molecule (169). This structure displays a distinct docking mechanism, with a sideways tilt displayed by the TCR, orienting the γ chain away from the CD1d binding surface so that only the δ chain makes contact (169). This example provides further support explaining the Vδ1 bias seen when isolating CD1d restricted γδ TCRs (170-172). Instead of the CDR3γ loop in the 9C2 structure, the DP10.7 solely utilises its variable CDR3δ loop to contact the sulfatide headgroup (169).

Until recently, TCR recognition of CD1 molecules has been confined solely to an 'on top' recognition and docking footprint, albeit with different orientations and lateral shifts across the top of the binding surface. A sideways 'atypical' recognition of CD1a by a γδ TCR has now been demonstrated through crystallographic studies (52). Previously, no γδ-CD1a co-crystal had been achieved due to a relatively poor understanding of antigen repertoire available to bind CD1a and difficulty in finding strong antigens i.e., αGC in CD1d. This study utilised two different γδ TCRs to investigate the interaction between CD1a-TCR at the molecular level, with initial SPR data demonstrating two different binding mechanisms for the TCRs (52).

The CO3 γδ TCR utilised an almost perpendicular binding footprint referred to as 'end to side' docking (52). This mechanism differs drastically from other known γδ binding footprints, with substantial contacts provided by β_2m on the CD1a molecule, as well as an insertion of the Vδ1 CDR loops into a cleft between the α1 and β_2m interface (52). It has been postulated that instead of direct contact with lipid antigen by the TCR, the affinity of the TCR for CD1a lipid can be modulated by indirect conformational changes caused by a rearrangement of the hydrophobic cavity depending on lipid occupation. This would explain the difference in TCR affinity between CD1a -DDM and CD1-endo (self-lipid derived from HEK cells) (52).

Interestingly, this 'sideways' docking mode, despite being the first documented example binding CD1a, has been demonstrated in similar MHC class I-like antigen-presenting molecule MR1 (173). Similarities between this and the recent CD1a-γδ structure can be seen, however, in the MR1 structure, the α3 domain provides contact points to the TCR instead of the β_2m (173). It is evident, despite the relatively poor understanding of γδ TCR-CD1 protein binding mechanism, that there is more than one mode of recognition utilised by the γδ TCR.

Although most CD1-γδ TCR structures support a 'top-to-bottom' docking mode, this new evidence indicates that alternative footprints may contribute to antigen specificity. Further structural studies are required to elucidate how different γδ TCRs distinguish between lipid antigens and whether additional docking modes exist.

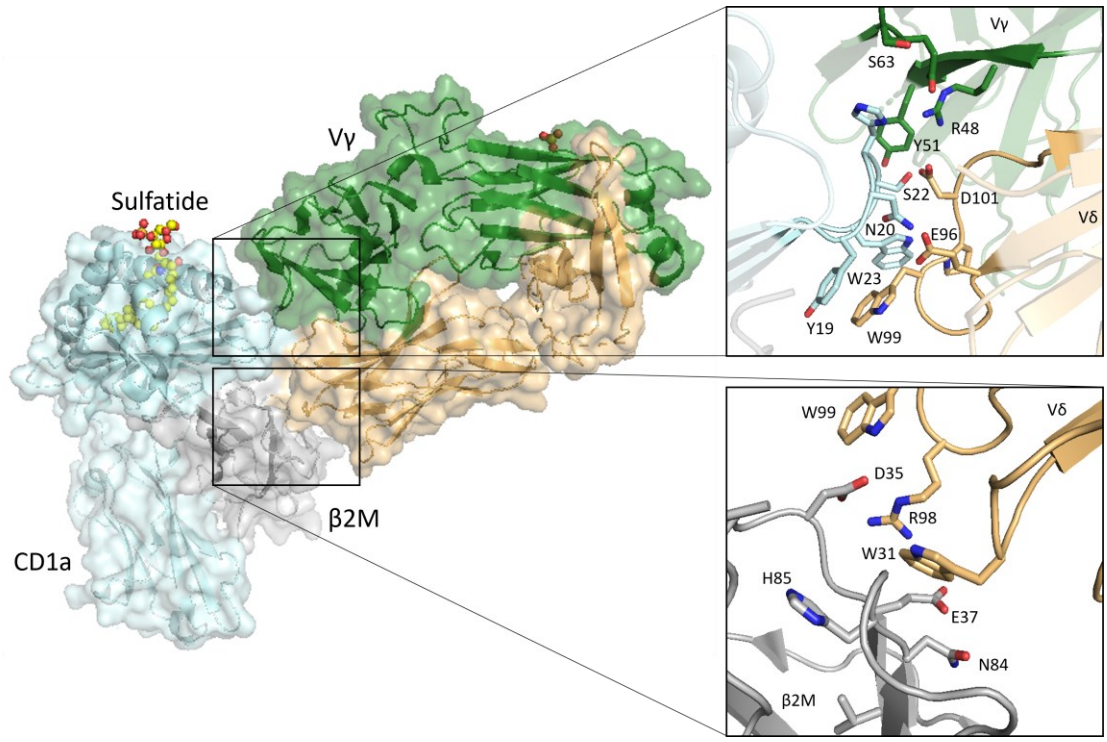


Figure 1. 12 The co-crystal complex of CD1a and the CO3 γδ TCR

CO3- γδ TCR binds in a sideways 'atypical' fashion to the side of CD1a, in a cleft between the CD1a main chain and β2m. CDR1δ loop inserts between the cleft, with a contribution of 40% and 60% overall between the γ and δ chains respectively. Zoomed-in images show the interaction site between CD1 and CO3, with β2M interacting solely with the Vδ chain and the CD1d heavy chain interacting with both the Vγ and Vδ chains. Complex shown in cartoon display with CD1d coloured in cyan, β2M in grey, Vγ in green and Vδ in orange. Sulfatide shown as spheres, coloured in yellow and coloured by heteroatom. Image generated using PyMOL with 4-letter code 7RYN.

This suggests that an internal conformational shift within CD1 molecules, influenced by groove occupation and lipid properties, may indirectly modulate γδ TCR binding affinity away from the antigen-binding groove. Unlike conventional MHC-restricted TCR interactions, where peptide antigen contacts are the primary determinant of specificity, this evidence supports an alternative mechanism of γδ TCR recognition, where lipid-induced structural changes in CD1 molecules could fine-tune TCR binding interactions.

While the CO3-CD1a structure provides the first crystallographic evidence of a sideways docking mode for a γδ TCR, it remains unclear whether this is a common feature among CD1-restricted γδ TCRs or specific to certain TCR-lipid interactions. Further structural studies are required to determine the extent of docking flexibility across CD1 isoforms, as well as the impact of lipid diversity on γδ TCR recognition mechanisms.

These findings highlight the importance of structural investigations in defining the molecular rules of CD1 antigen presentation. This thesis builds on these concepts by employing structural and biophysical approaches to investigate CD1c-γδ TCR and CD1d-iNKT TCR interactions, addressing key knowledge gaps in CD1-lipid-TCR binding mechanisms.

1.5 Summary

CD1 molecules are MHC class I-like antigen-presenting proteins that specialise in lipid antigen presentation to T cells, playing a critical role in immune surveillance. In contrast to conventional MHC molecules which primarily present peptide antigens, CD1 proteins bind a diverse array of lipid antigens and present them to unconventional T cells including iNKT and γδ T cells. These interactions facilitate immune responses to a diverse array of foreign and self-lipid antigens, however, the molecular mechanisms underpinning these reactions are poorly understood, particularly at the structural level.

Of the five CD1 isoforms, four present lipid antigens (CD1a, CD1b, CD1c, CD1d), with their specificity shaped by tissue distribution and the architecture of their lipid-binding grooves. Whilst there has been significant progress in using mouse models to study CD1d-TCR interactions, human co-crystal structural data remains limited. This lack of structural data leaves predominant gaps in our understanding of the mechanism by which unconventional T-cell recognition of lipid antigens governs immune responses.

Several key questions remain unanswered in the field, which would lead to dramatic increases in our understanding of how diverse lipid antigens shape TCR specificity and affinity for CD1 molecules.

- How do CD1 molecules accommodate lipids of significantly different structural and chemical properties and present them to T cells?
- What are the major structural features of iTCRs and CD1 that influence the affinity and specificity of TCRs for CD1?
- How do diverse lipids within the CD1 binding groove influence TCR engagement by adapting the CD1 groove shape and size?

To address these questions, this thesis aims to employ structural and functional approaches to investigate CD1c and CD1d interactions with γδ and iNKT TCRs, respectively. Through a combination of X-ray crystallography, biophysical, and functional analyses, this research aims to provide new insights into the molecular mechanisms of CD1-lipid-TCR recognition, advancing our understanding of lipid-based antigen presentation in humans.

1.6 Overarching Hypothesis

The molecular mechanism underpinning unconventional T cell recognition of antigen-presenting molecules is diverse and is influenced by several factors including the lipid cargo, CD1 binding groove flexibility, and TCR sequence.

1.7 Aims

The overall aims of this research are:

1. Understand CD1 lipid-binding groove flexibility and the influence different lipid cargo has on the TCR binding mechanism.

2. Investigate differences in iNKT TCR binding sequence, with emphasis on the CDR3β loop, and advance our understanding of the influence this has on CD1d recognition.

3. Develop and optimise a pipeline to investigate the lipid reactivity of CD1c-reactive γδ TCRs.

Chapter 2 Materials and Methods

2.1 Protein production

2.1.1 CD1 lipid monomer production – refolding method

Recombinant CD1 heavy chain and β2m protein were expressed as soluble inclusion bodies in *E. coli* strain Rosetta 2(DE3) pLysS (Novagen, #71403). Protein concentration was initially determined in 8M urea (SLS, #U5378), and inclusion bodies were denatured in 6M guanidine-HCl buffer (Sigma-Aldrich #SRE0066), aliquoted, and stored at -20°C. Before refolding, the inclusion body concentrations of individual aliquots were re-measured using a NanoDrop spectrophotometer in the presence of 8M urea. Proteins were denatured with 20 mM dithiothreitol (DTT) (Sigma-Aldrich, #D0632) at 37°C. Oxidative refolding of CD1 and β2m was carried out *in vitro* at 4°C under continuous stirring, either in the presence of lipid α-GC (Avanti, #867000) solubilized in vehicle (CD1d) or in vehicle alone (CD1c). The refolded protein was concentrated using a high-pressure nitrogen gas filtration system with a 400 mL stirred cell (Amicon, #UFSC40001) and Ultracel 10 kilodalton (kDa) ultrafiltration discs (Amicon # PLGC06210).

2.1.2 CD1c-endo production

CD1c-endo protein was expressed and purified using a mammalian expression system with the Expi293™ Expression System Kit (Thermo Fisher Scientific, #A14635). Expi293F cells were thawed and cultured in Nalgene™ Single-Use PETG Erlenmeyer Flasks (Thermo Fisher Scientific, #4115-0125) containing Expi293F medium. The cultures were maintained in a shaking incubator at 37°C with 8% CO_2. Following at least three passages, cells were transfected on **Day 0** at a viable cell density of 3 × 10^6 cells/mL using the ExpiFectamine™ 293 Transfection Kit (Thermo Fisher Scientific, #A14525). ExpiFectamine™ 293 Transfection Enhancer One and Enhancer Two were added on **Day 1**, along with Biotin (Thermo Fisher Scientific, #B20656). Protein was harvested on **Day 5** under continuous orbital shaking. After centrifugation, the supernatant was triple filtered using Millitop Steripore Vacuum Filters with pore sizes of 0.8 µm, 0.45 µm, and 0.22 µm (Merck, #S2HVU05RE; Fisher Scientific, #10064470). The filtered protein was stored at 4°C until purification.

2.1.3 αβ TCR production

NM4 (αβ) and iNKT (αβ) TCRs were generated as soluble inclusion bodies using the *Rosetta DE3*pLysS* *E. coli* expression system (lot number indicated above). Plasmid DNA from individual chains was

transformed, colony-selected, and grown in auto-induction media (Foremedium, #AIMSB0210). Inclusion bodies were washed in wash buffer and resuspension buffer before being stored at -20°C in 6Mguanidine buffer. Before refolding, chains were denatured in 20 mM DTT at 42°C for 40 minutes. The NM4 TCR used a ratio of 1:1 α to β inclusion bodies, respectively, while DN25 and 1369 iTCRs were mixed at a ratio of 4 α to 1 β, respectively. Denatured proteins were added to refolding buffer consisting of 4M urea (SLS, #U5378), 300 mM L-arginine (Thermo Fisher Scientific, #104990025), 2 mM EDTA (Sigma-Aldrich, #E5134), and 0.1 m Tris-HCl (Merck, #648-317), pH 8.1. Refolding was carried out overnight in pre-boiled 76 mm cellulose membrane dialysis tubing (Merck, #D9402) against water, followed by a second overnight dialysis against 20 mM Tris, pH 8.1. Samples were taken before refolding for reducing and non-reducing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis.

2.1.4 γδ TCR production

γδ TCRs were expressed using a methodology like that for αβ TCRs, but with the use of tryptone yeast peptone (TYP) media instead of auto-induction media (AIM). TYP media was supplemented with 20% glucose (Sigma-Aldrich, #G7021) prior to the addition of bacteria. The ratio of inclusion bodies for TCR refolding was dependent on the specific γδ TCR being expressed; however, inclusion bodies were generally incubated in 1:2, 1:1, and 2:1 ratios, with some TCRs also incubated in 5:1 and 8:1 ratios.

2.2 Protein purification

2.2.1 CD1 protein purification

Protein was purified using a two-step gel filtration protocol. Initial size exclusion purification was carried out using a preparatory-grade HI load S75, 26/600 Superdex® gel filtration column (GE Healthcare, #GE28-9893-34). Before loading, the protein was concentrated to <5 mL and loaded using a 5 mL superloop. The programme was run at 2 mL/min, and fractions were collected with an absorbance value of at least five mAU, ensuring a maximum peak width of 2 mL. Secondary purification was performed using an analytical S75, 10/300 GL column (GE Healthcare, #17517401). The protein was concentrated to <500 µL and injected using a 500 µL superloop. The programme was run at 0.15 mL/min. For CD1c, further purification of monomers was achieved using a MonoQ 4.6/100 PE column (GE Healthcare, #17517901), after which fractions were biotinylated and tetramerised.

2.2.2 TCR protein purification

Soluble, refolded TCRs were initially concentrated using a POROS HQ anion exchange column (Thermo Fisher Scientific, #82078). Dialysed protein was flowed through the column at a flow rate of 10 mL/min, resulting in binding to the column. Elution was achieved using a one-step, high-salt elution. Eluted protein was then subjected to multiple rounds of gel filtration chromatography, using a Hi-Prep S75, 26/600 size exclusion column, followed by an analytical S75, 10/300 GL size exclusion column (LOT numbers described above).

2.2.3 CD1c-SL purification

CD1c-SL protein was purified using a three-step purification process. Initial size exclusion purification was performed using a Hi-Prep S75, 26/600 gel filtration column, with fractions collected and analysed via SDS-PAGE. The protein of interest was then purified using a MonoQ 4.6/100 PE anion exchange column. The protein was then biotinylated and underwent a final analytical purification step using an S75, 10/300 GL gel filtration column.

2.2.4 CD1c-endo purification

Protein monomers were initially purified using an immobilised metal affinity chromatography (IMAC) column (Cytiva, #17371206). The protein was bound via a 6X-His tag to the IMAC column. Following the washing of weakly bound proteins with 20 mM imidazole (Merck, #5202), the target protein was eluted using a one-step process with 500 mM imidazole. The protein was then subjected to size exclusion chromatography using an analytical S75, 10/300 GL column (Cytiva).

2.3 Testing and validation of protein monomers

2.3.1 SDS-PAGE gel electrophoresis

SDS-PAGE analysis was carried out on purified protein fractions for protein molecular weight analysis. The gel was made from resolution and stacking gels (12%) consisting of Ultrapure H_2O, 30% acrylamide mix (Sigma-Aldrich, #A3699), 1.5 M Tris pH 8.8 (resolution), 1 m Tris pH 6.8 (stacking), 10% SDS (Fisher Scientific, #10593335), 10% APS (Sigma-Aldrich, #A3678), and TEMED (Fisher Scientific, #17919). 3 µL of loading dye (IM Tris pH 6.8, glycerol, SDS, bromophenol blue, 1 m DTT [reducing dye]) was incubated at 95°C with 10 µL protein for ten minutes before loading. The gel was run at 20 V for 20 minutes,

followed by 150 V for 80 minutes. The molecular weight ladder used was Precision Plus Protein Dual Colour (Bio-Rad, #1610374).

2.3.2 Biotin ELISA

A flat-bottomed 96-well plate (Fisher Scientific, #10334791) was used for biotin ELISA experiments. 150 µL of phosphate-buffered saline (PBS) (Thermo Fisher Scientific, #70011044) was added to each well, followed by 1 µg of irrelevant, unbiotinylated protein and 1 µg of biotinylated relevant protein (all conditions in triplicate). Samples were initially incubated at 37°C for one hour to allow protein adhesion to the plate. Wells were washed extensively with PBS before incubation for 30 minutes at 37°C with 100 µL blocking buffer containing PBS, 5% FBS (Thermo Fisher Scientific, #A5256701), and 0.1% sodium azide (Fisher Scientific, #15845188). The plate was washed again and then incubated for 30 minutes at room temperature with 50 µL of 1:2000 streptavidin HRP (Thermo Fisher Scientific, #SA10001). Following streptavidin incubation, wells were washed extensively to remove any unbound streptavidin. Finally, 50 µL of 3,3',5,5'-Tetramethylbenzidine (TMB) ELISA substrate (Fisher Scientific, #10544141) was added and left for 15 minutes in the dark. The reaction was stopped using 50 µL H_2SO_4 (reaction turns from blue to yellow). Final absorbance measurements were recorded using a plate reader at a 405 nm wavelength.

2.3.3 Tetramer generation

Tetramers were generated following overnight biotinylation, re-purification, and validation using biotin ELISA. Proteins were concentrated to approximately 1 milligram per millilitre (mg/mL) before tetramer generation. Individual protein monomers were added to a 1.5 mL Eppendorf tube, where they were conjugated to Streptavidin-R-phycoerythrin (Streptavidin-PE) (BioLegend, #405204). Streptavidin-PE was added in five equal additions, ten minutes apart, and kept on ice and in the dark during all additions. Following the additions, the tetramer was spun at 13,000 RPM for five minutes and stored at 4°C. CD1 and TCR tetramers were all generated using the same methodology.

2.4 Cell culture

2.4.1 EXPI 293F cell culture

Expi 293F cells were initially seeded to a density of 200,000 cells per mL in a 30 mL volume (125 mL flask) and allowed to recover for three to four days. Viability recovered to > 95% seven days post-thaw. Cells were cultured using PETG polycarbonate, sterile, non-baffled Erlenmeyer flasks in one of

either 125 mL, 250 mL, or 500 mL (Thermo Fisher Scientific, #10266432, #10045411, #10216821, respectively) on a shaking incubator at 125 RPM at 37°C with 8% CO_2. Cells were sub-cultured every three to four days and seeded to a final density of $0.2 \times 10^6 – 0.3 \times 10^6$ in fresh EXPI293F medium (Fisher Scientific, #A1435101). Transfection was not performed until viability was > 95% and viable cell density was between 3×10^6 and 5×10^6. Cells were passaged a minimum of three times before transfection. All cell manipulations were performed while gently shaking the flask in a clockwise direction.

2.4.2 PBMC isolation

Blood was removed from the donor directly into Li^+ heparin tubes (SLS, #VS367885) (carried out by a trained colleague). The blood was then diluted in PBS and layered over 15 mL of Ficoll-Paque PLUS (Merck, #GE17-1440-02) in a 50 mL tube. The blood was centrifuged at 2,000 RPM for 15 minutes with the brake and accelerator off. The Buffy coat was harvested, diluted 1:1 with PBS, and spun for seven minutes with the brake and accelerator on. The supernatant was discarded before resuspending the cells in 50 mL of PBS. The cells were spun at 1,500 RPM for five minutes before being resuspended in RPMI 1640 complete media and counted. The complete media consisted of 500ml RPMI 1640, 50ml Foetal Bovine Serum (FBS), and 5ml Penicillin-Streptomycin/Glutamine solution unless otherwise stated. Following PBMC isolation, the cells were either used fresh or cryopreserved at -80°C in FBS + 10% DMSO (Merck, #D8418) at a density of $1–2 \times 10^7$ cells per vial. *Study ethics number for use of human PBMCs =* ***14291***.

2.4.3 Freezing and thawing iNKT lines and clones

Cells were thawed by removing them from liquid nitrogen and swirling them in a 37°C water bath. The cells were then quickly added to 20 mL of pre-warmed RPMI 1640 complete media in a 5 mL Falcon tube. The cells were spun at 1,500 RPM for 10 minutes at 4°C before being resuspended in 1 mL of RPMI 1640 complete media in a 24-well plate. Following a two-hour incubation to revive the cells, they were counted and seeded at a density of 2×10^6 cells into individual wells of a 24-well plate, containing T cell media supplemented with 0.5 µg/mL of PHA (Sigma-Aldrich, #11249738001). One mL (2×10^6) of irradiated allogeneic PBMC feeders was added to each well, and the cells were left undisturbed. The cells were then monitored, and the media was replaced when it turned yellow. iNKT cells were frozen at 10^7 cells/mL. Before freezing, the cells were spun at 4°C for 10 minutes at 1,500 RPM. The cells were frozen in 1 mL aliquots in Mr Frosty and placed at -80°C for 24 hours. The cells were then removed from -80°C and placed in liquid nitrogen for long-term storage.

2.4.4 Generation and maintenance of iNKT cell lines

Blood samples were obtained from healthy local volunteers after informed consent. PBMCs were isolated from human peripheral venous blood by density gradient centrifugation (previously described). iNKT cell lines were generated from PBMCs by seeding at 2 × 10^6 cells/mL into 24-well plates in 1 mL of iNKT cell growth medium containing RPMI 1640 (Fisher Scientific #16497060), 10% foetal bovine serum (Thermo Fisher Scientific #A5256701), Penicillin/Streptomycin (100 µL/mL; Gibco #15140148), 1% non-essential amino acids (Thermo Fisher Scientific #11140050), 1% essential amino acids (Thermo Fisher Scientific #11130051), 1% L-glutamax (ThermoFisher Scientific #35050061), 55 µM 2-mercaptoethanol (Sigma-Aldrich #M6250), and 15 mM HEPES (Gibco #15630106). Before use, the solution was filtered using a Stericup Quick Release Durapore 0.22 µm PVDF filter (Merck #15780319). PBMCs were pulsed with KRN7000 (Avanti #867000) (100 ng/mL) on Day 0 and IL-2 (Sigma-Aldrich #I2655) (200 IU/mL) was added on Day 5. Cells were fed every two to three days with fresh iNKT cell growth medium containing IL-2 (200 IU/mL). Following the 14-day expansion, CD3+, CD1d-KRN7000 tetramer+, and Vβ11+ T cells were bulk sorted into tubes containing iNKT cell growth medium. Sorted lines were stimulated with 1 µg/mL PHA-L (Invitrogen #00-4977-93), IL-2 (200 IU/mL) in the presence of 2 × 10^6/mL autologous γ-irradiated (50 Gy) PBMCs and cultured in 2 mL in a 24-well plate. Following a two-week expansion, CD3+, CD1d-OCH tetramer+, and Vβ11+ cells were sorted to generate iNKT cell lines.

2.5 Investigating lipid reactivity.

2.5.1 Lyophilising and storage of lipid aliquots

Lipids used for reactivity experiments were ordered from Avanti Polar Lipids. The lipids were supplied in vials and required lyophilisation before use in reactivity experiments. Dried lipid was initially sonicated using a water bath sonicator in a 1:1 chloroform-methanol mixture. The volume of chloroform-methanol used depended on the total amount of lipid received, i.e., 100 mg of lipid was solubilised in 1 mL and divided into ten 100 µL aliquots. Individual aliquots were left to dry overnight before being stored at -20°C for long-term storage.

2.5.2 MACSI-bead assay

CD1c-endo protein was conjugated to Anti-Biotin MACSiBead™ Particles (Miltenyi #130-092-357) with 10% sodium azide (Fisher Scientific #15845188) and rotated overnight at 4°C. The beads were then

added to FACS tubes (100,000 per tube) before 10 µg of lipid was added, and the mixture was vortexed briefly. Before adding the lipid to the beads, it was sonicated at 37°C for 30 minutes in a 0.5% CHAPS (Sigma-Aldrich #220201) solution to ensure proper solubilisation. A variety of lipids were used, including SM (Avanti #860584), Liver Phosphatidylinositide (Liver PI) (Avanti #840042), and several PI analogues, including 08:0 (Avanti #85018), 16:0 (Avanti #850141), and 18:0 (Avanti #850143). Following lipid addition, the beads were incubated overnight at 37°C. The following morning, beads were washed several times in sterile PBS before being stained with NM4 (αβ) TCR tetramer for 45 minutes at 4°C. A separate FACS tube containing beads only and another containing beads + vehicle, but without added lipid, were used as controls. After 45 minutes of staining, the beads were washed extensively with sterile PBS before resuspending and analysing on the FACS Aria. Analysis was carried out using FlowJo VX (FlowJo LLC).

2.5.3 Plate bound assay.

Individual wells of a Maxisorp 96-well plate (ThermoFisher #442404) were coated with 5 µg of CD1c-endo protein in 50 µL of PBS, or 50 µL of PBS only, and incubated overnight at 4°C. Wells were washed extensively (three times with 200 µL of PBS) to remove unbound protein. Subsequently, 1 µg, 5 µg, and 10 µg of sonicated liver PI (Avanti #840042) in 0.5% CHAPS/citrate buffer (pH 6.5) were added in triplicate to the wells containing protein. The plate was incubated overnight at 37°C. Wells were washed three times with 200 µL PBS again to remove unbound lipids. Then, 100,000 Jurkat cells were added, followed by a further overnight incubation at 37°C. After incubation, the wells were mixed to resuspend the protein-lipid complex, transferred to FACS tubes, and stained with BV421-conjugated anti-human CD3 antibody (clone UCHT1) and APC-conjugated anti-human CD69 antibody (clone FN50) and streptavidin PE (all from BioLegend) for 30 minutes. Cells were then washed using MACS buffer and analysed using FACS Calibur.

2.6 Surface plasmon resonance (SPR)

Binding analysis was performed using a BIAcore® 3000 (GE Healthcare, Buckinghamshire, U.K.) equipped with a CM5 sensor chip. Briefly, CM5 chip coupling solutions containing 100 µL of 100 mM NHS and 100 µL of 400 mM EDC were used to activate the chip before streptavidin binding. Approximately 5000 response units (RU) of streptavidin (110 µL of 200 µg/mL in 10 mM acetate, pH 4.5) were covalently linked to the chip surface in all four flow cells. A 100 µL injection of 1 m

ethanolamine hydrochloride was used to deactivate any remaining reactive groups. Approximately 1000 RU of the CD1d-lipid complex was attached to the CM5 sensor chip at a slow flow rate of 10 µL/min to ensure uniform distribution on the chip surface. A blank flow cell was used as a negative control. The TCRs were purified and concentrated, and 12 serial dilutions (1/2) were prepared for each sample and injected over the sensor chips at 25°C. The TCRs were injected over the chip surface using kinetic injections at a flow rate of 30 µL/min. Results were analysed using BIAcore T200 Evaluation (GE Healthcare, Buckinghamshire, U.K.) and GraphPad Prism software. The equilibrium binding constant (K_D) values were calculated assuming a 1:1 interaction (A + B ↔ AB) by plotting specific equilibrium-binding responses against protein concentrations, followed by non-linear least squares fitting of the Langmuir binding equation: $AB = \frac{B \times ABmax}{K_D + B}$.

2.7 Structural studies

2.7.1 CD1d crystallisation (Human)

CD1d-Human protein was initially concentrated to 10 mg/ml and set up using the Oryx 8 crystallisation robot in screens Morpheus, TOPS, and JCSG+ (Molecular Dimensions). Initial crystallisation rounds yielded predominantly aggregated wells in all three screens with no evidence of crystals. The protein was subsequently re-purified and concentrated to >14Mg/ml. Large crystals grew in the MORPHEUS and TOPS screens (C12, G12, and H4 respectively). Three large crystals were harvested, flash-frozen in liquid nitrogen, and sent to ESRF for X-ray diffraction and data collection. The remaining, slightly smaller crystals were smashed up using seed beads and stored as seed stock at 1:10, 1:100, 1:1000, and 1:10000 dilutions for future crystallisation experiments.

2.7.2 CD1d crystallisation (Macaque)

CD1d Macaque was concentrated using Vivaspin 20 spin columns (Sigma-Aldrich #Z614602) to initial concentrations of 12 mg/ml. The protein was set up in initial screens Morpheus, TOPS, and JCSG+, in 96-well plates, using seed stock from the previous CD1d-Human purification. Ratios of seed stock to protein to mother liquor were 0.5:1:1, respectively. Crystals grew in the H4 wells of the TOPS screen, both with and without seed stock, but were larger without.

2.7.3 iTCR crystallisation

iTCRs were concentrated to ~10 mg/ml and set up using the Oryx-8 crystallisation robot in standard screens (Morpheus #MD1-91), (JCSG+ #MD1-37), and (TOPS #TOP96). Initial micro-crystals were used to generate seed stock, which was then used in subsequent rounds of crystallisation. The ratio of seed stock to protein to mother liquor was 0.5:1:1, respectively. Small crystals grew in multiple plates across the three screens, with two main morphologies appearing after roughly four days post-set-up. Crystals were harvested, flash-frozen in LN2, and sent to the Diamond Light Source for X-ray diffraction.

2.7.4 CD1c-endo crystallisation

CD1c-endo protein from the same batch used for tetramer staining (pre-biotinylation) was concentrated to 12 mg/ml and set up using the Oryx-8 crystallisation robot in 96-well plates. Initial crystallisation attempts yielded hundreds of small, needle-shaped crystals per well. These were too small to fish for structural studies. For our second attempt, CD1c-endo protein was treated with endoglycosidase F (Sigma-Aldrich #E9762) to remove N-linked glycan post-translational modifications before set-up. Seed stock from our initial crystallisation attempt was also used. Crystals grew slightly larger and fewer per well than the previous attempt. Subsequent rounds of crystallisation used more diluted seed stock and lower concentrations of protein; however, we were not able to increase protein size sufficiently for high-resolution structural data to be collected.

2.8 Structure determination

Human and Macaque crystals were then sent to the European Synchrotron Radiation Facility (ESRF) and Diamond Light Source, respectively. CD1d-Human crystals diffracted to 1.72 Å and CD1d-Macaque crystals diffracted to 1.8 Å. Both were then subjected to molecular replacement carried out using CCP4i2, with AlphaFold2.ipynb-Colaboratory predictions as the search models. Final models were then obtained by repeated refinement cycles in CCP4i2, until Rwork/Rfree values were 0.17/0.2 and 0.18/0.21, respectively. Data quality was then validated and analysed using the Worldwide Protein Data Bank (PDB) validation and deposition servers. Residues making contacts of interest were calculated using CCP4i2 and figures were generated using PyMOL.

2.9 Figure presentation

Lipid structures were prepared using ChemDraw software (CambridgeSoft). All figures were generated using either ChimeraX or PyMOL. Hydrogen bonding network comparisons were carried out using LigPlot+.

2.10 Flow cytometry

2.10.1 Staining and sorting iNKT cells

The following fluorescent reagents were used to analyse and/or sort iNKT cells: PE-conjugated human or macaque CD1d tetramers loaded with KRN7000, FITC-conjugated anti-human TCR Vα24 (clone C15; Beckman Coulter Ltd, High Wycombe, UK), FITC-conjugated anti-human TCR Vβ11 (clone C21; Beckman Coulter Ltd) and BV421-conjugated anti-human CD3 antibody (clone UCHT1). Cells were stained with mAbs in phosphate-buffered saline containing 1% FBS and EDTA 1 mM for 30 minutes at 4°C, acquired on BD FACS Aria and analysed on FlowJo V10.9.0 (FlowJo LLC, Oregon, USA) software.

2.10.2 Staining of NM4 Jurkat T cells

J.RT3-T3.5 Jurkat cells, which expressed the NM4 TCR (human CD1c-restricted), were counted and aliquoted into 2×10^5 cells per tube. As a control, parental line Jurkat cells not expressing the TCR were also aliquoted. Cells were washed in cold MACS buffer before adding 0.3 µg of CD1c tetramer and CD3-APC. Cells were then incubated on ice for 45 minutes. Following incubation, cells were washed in cold MACS buffer and taken to the flow cytometer for staining. The acquisition was carried out using a FACSAria, and analysis was carried out using FlowJo.

2.10.3 Staining Macaque PBMCs

Frozen Macaque PBMCs were received from PHE Porton Down in collaboration. Samples were then thawed, counted, and stained with PE-conjugated human or macaque CD1d tetramers loaded with KRN7000, and FITC-conjugated anti-human TCR Vα24 (clone C15; Beckman Coulter Ltd, High

Wycombe, UK). The acquisition was carried out using a FACSAria, and analysis was carried out using FlowJo.

2.11 TCR viral transduction of Jurkat T cells

The following is a summary of the 11-day protocol carried out to achieve TCR transduction:

Day 1 - HEK cells were initially plated in three different densities (1 x 10^6, 1.5 x 10^6, and 2 x 10^6) in 7 ml of DMEM complete media (500ml DMEM, 50ml FBS, 5ml Pen-Strep/L-glutamine) in a 100 mm tissue culture dish.

Day 2 - DNA media was then prepared as follows:

Concentration	**Volume**
Pol/Gag	1 µg/µl
pRSV.REV	1 µg/µl
pVSV-G	0.5 µg/µl
LV Plasmid DNA	1 µg/µl
DMEM (no FBS or PSG)	

DNA media was vortexed, mixed with 20 µl TurboFect reagent, and incubated for 10 minutes. The media was removed from the cells, followed by a wash with 4Ml PBS, and then DNA media supplemented with 3 ml DMEM was added. This was incubated for 3 hours, followed by another 4Ml PBS wash and replacement with 6 ml DMEM media.

Day 3 – Jurkat T cells to be transfected were plated at 0.5 x 10^6/ml and 1 x 10^6/ml in 3 ml complete DMEM media and incubated at 37°C overnight.

Day 4 – HEK cell media was removed, centrifuged, filtered, and added to the Jurkat cells at 0.5 x 10^6/ml and 1 x 10^6/ml. Plates were then spun at 2200 RPM for 90 minutes and incubated at 37°C.

Day 5 – The transduction procedure was repeated (see Day 4).

Day 11 – Jurkat cells were counted, sorted into 5 x 10^5/ml in tubes, and analysed via FACS for TCR expression using FITC anti-γδTCR (clone B1) (Biolegend #331209) or anti-Vδ1 antibodies (Miltenyi Biotec #130-135-325) – (Clone REA173).

Method acknowledgements

I would like to thank PHE (Public Health England) who kindly offered us access to Macaque PBMCs from their animal facility, without which we would have been unable to carry out some of the experiments and analysis seen in this thesis (chapter 3).

I would also like to express my gratitude towards Immunocore, who gave up their time and resources to positively impact my project through SPR studies and advice on the set-up of the mammalian expression system at Southampton General Hospital.

Finally, I would like to thank Diamond Light Source and the European Synchrotron Radiation Facility, (ESRF), without whom I would have been unable to obtain crystallographic structural data for my project.

Chapter 3 New insights into CD1d-αGC structure reveal a conserved flexibility in the F' roof.

3.1 Investigating isoform differences in CD1d F' roof conformation.

The ability of CD1 proteins to undergo conformational changes in response to lipid occupancy has been well documented through various structural and functional studies, including SPR, flow cytometry, and X-ray crystallography (104). This structural adaptability may facilitate the binding of diverse lipid antigens by modulating the lipid-binding groove.

This chapter aims to determine whether CD1d exhibits a dynamic conformational mechanism like that observed in other CD1 isoforms, where structural rearrangements within the F' roof alter the lipid-binding groove volume. Specifically, we focus on Phe84, a conserved aromatic residue within the F' roof, and assess whether its rotational flexibility alters lipid binding and TCR recognition.

We use CD1d as a model system to examine this, employing a combination of X-ray crystallography and functional assays. First, we conduct a comparative sequence and structural analysis of CD1a, CD1b, CD1c, and CD1d, identifying conserved features within the lipid-binding groove that may underpin flexibility. We then determine the high-resolution structure of CD1d-αGC and characterise its structural adaptability. To explore the functional relevance of this flexibility, we use tetramer staining, SPR binding studies, and molecular dynamics simulations to investigate its role in TCR recognition. Finally, we compare human, macaque, bovine, and mouse CD1d structures to assess whether Phe84 flexibility is a conserved feature across species.

By integrating structural and functional approaches, we aim to determine whether CD1d employs a dynamic mechanism for lipid antigen presentation, providing insight into how structural flexibility enables CD1 molecules to accommodate diverse lipid antigens.

3.1.1 Molecular comparison of CD1 antigen-presenting isoforms.

CD1a, CD1b, CD1c, and CD1d all present lipid antigens to TCRs at the cell surface. To investigate whether CD1d exhibits conformational flexibility within its lipid-binding groove, we carried out a comparative sequence and structural analysis of CD1 isoforms, focusing on residues in and around the antigen-binding site. Given that CD1c exhibits flexibility in the F' roof (Mansour et.al 2016), we assessed whether similar conformational adaptability is present in CD1d by analysing conserved aromatic residues within the lipid-binding region.

a)

α1 helix

A.A number	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88
CD1a	S	N	E	E	W	K	E	L	E	T	L	F	R	I	R	T	I	R	S	F	E	G	I	R	R	Y	A	H	E	L
CD1b	S	D	K	E	V	A	E	L	E	E	I	F	R	V	Y	I	F	G	F	A	R	E	V	Q	D	F	A	G	D	F
CD1c	S	N	E	E	L	S	D	L	E	L	L	F	R	F	Y	L	F	G	L	T	R	E	I	Q	D	H	A	S	Q	D
CD1d	S	D	Q	Q	W	E	T	L	Q	H	I	F	R	V	Y	R	S	S	F	T	R	D	V	K	E	F	A	K	M	L

α2 helix

A.A number	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183
CD1a	N	M	A	K	H	F	C	K	V	L	N	Q	-	N	Q	H	E	N	D	I	T	H	N	L	L	S	D	T	C	P	R	F	I	L	G	L	L	D	A	G	K	A	H	L	Q
CD1b	S	R	A	Q	K	F	C	A	L	I	I	-	Q	Y	Q	G	I	M	E	T	V	R	I	L	L	Y	E	T	C	P	R	Y	L	L	G	V	L	N	A	G	K	A	D	L	Q
CD1c	S	L	A	Q	S	V	C	H	L	L	N	H	Q	Y	E	G	V	T	E	T	V	Y	N	L	I	R	S	T	C	P	R	F	L	I	G	L	L	D	A	G	K	M	Y	V	H
CD1d	L	W	V	N	L	A	I	Q	V	L	N	Q	-	D	K	W	T	R	E	T	V	Q	W	L	L	N	G	T	C	P	Q	F	V	S	G	L	L	E	S	G	K	S	E	L	K

b)

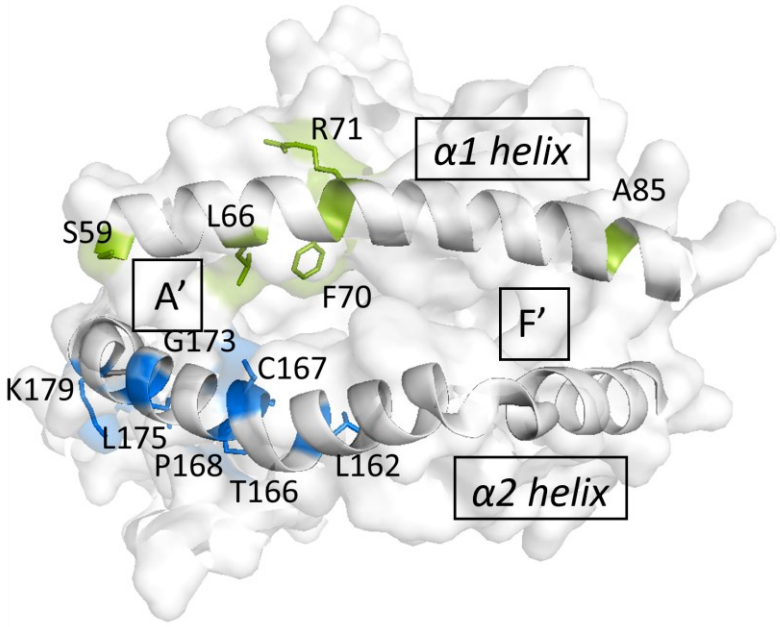


Figure 3. 1 Sequence analysis and structural representation of lipid presenting CD1 isoforms.

Sequence alignment of the four surface-expressed lipid-presenting CD1 isoforms, focusing on the α1 and α2 helices that comprise the lipid-presenting tunnel network. Amino acids conserved across all four species are highlighted in green and blue for the α1 and α2 helix respectively. b) Structural representation of the CD1 monomer, highlighting where the conserved amino acids sit within the α helices. The model used for this representation is CD1d (from PDB entry 1ZT4) with the image generated using PyMOL.

Sequence analysis revealed most of the conserved amino acids were situated above the A' tunnel, perhaps unsurprisingly as the A' tunnel is commonly described as the most conserved lipid binding pocket (174).

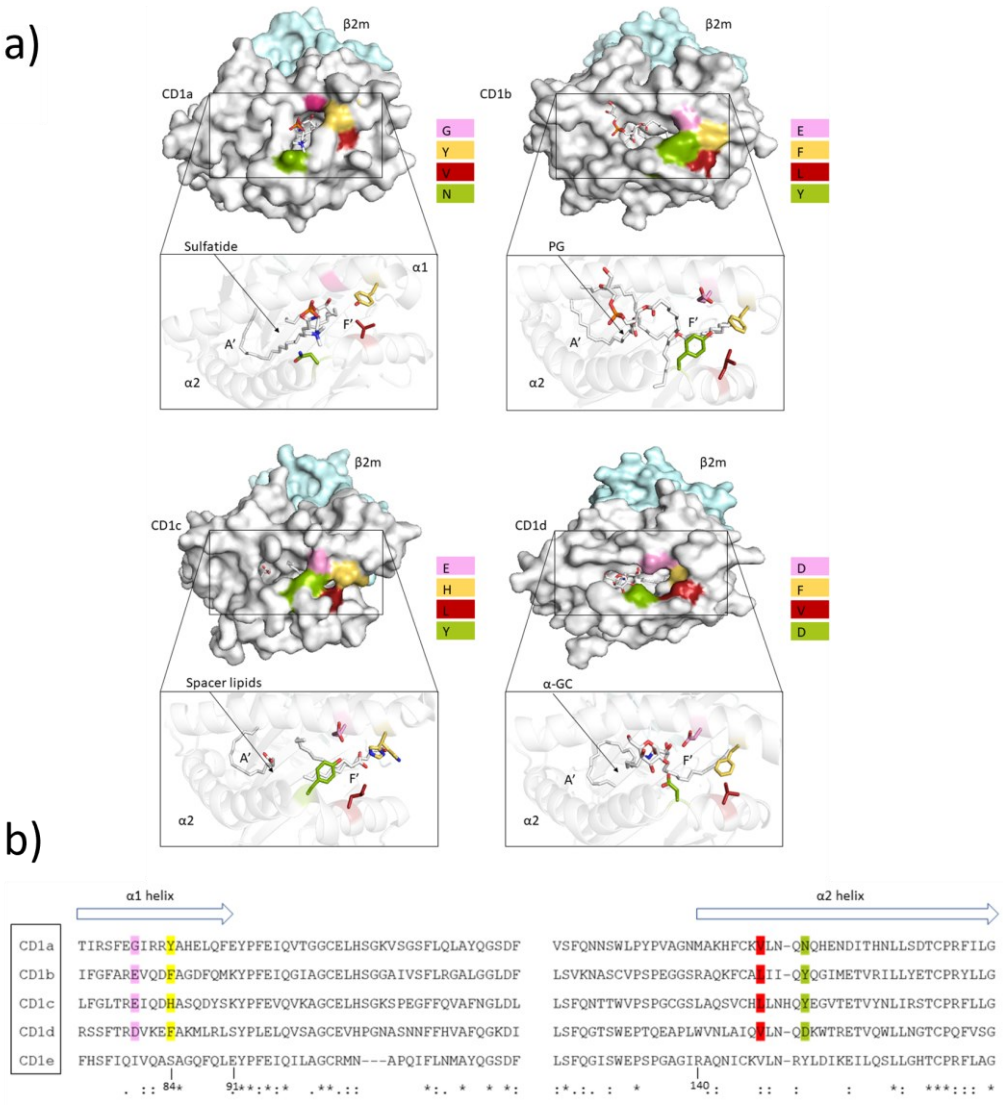


Figure 3.2 Structural comparison and sequence alignment of CD1a, CD1b, CD1c, and CD1d

a) Surface and cartoon representation of CD1a (i) CD1b (ii) CD1c (iii) and CD1d (iv) lipid-bound complexes, with emphasis on equivalent residues comprising the F' roof structure in CD1c-SL. The CD1 heavy chain is coloured in grey, β2m in light blue, and lipid in white and coloured by heteroatom (oxygen in red and nitrogen in blue). b) Sequence alignment of CD1 antigen-presenting isoforms with relevant residues highlighted in colour. Partial α-helices are represented by blue arrows and sequence similarities and differences are highlighted below the sequence. PDB codes 4X6F, 5WL1, 5C9J, and 1ZT4 were used for CD1a, CD1b, CD1c, and CD1d, respectively.

We hypothesised that amino acid side chain flexibility within the CD1 isoforms is likely to reside predominantly on the right-hand side of the lipid-binding pocket, surrounding the more variable F' portal. Closer analysis of the residues involved in opening and closing the CD1c-SL F' roof revealed that one of the residues in CD1c-SL (His87) was also aromatic at the equivalent position in the other antigen-presenting CD1 isoforms. This amino acid is positioned next to a conserved alanine residue at the C-terminal end of the α1 helix, as shown in *Figure 3.1.*

Following this, we examined the CD1c-SL electron density maps and identified a dual conformation of His87, allowing it to shift between open and closed states depending on lipid occupancy. Given this observed flexibility, we proceeded to investigate whether CD1d displays a similar structural feature using X-ray crystallography.

CD1d was selected for crystallographic investigation due to its stability, refolding capability, and crystallisation properties. Additionally, the TCR binding footprint is known to reside predominantly on the right-hand side of the molecule above the F' portal *(Figure 3.2)*, supporting the relevance of investigating flexibility in this region**.**

For continuity, the human CD1d- αGC structure will be referred to as CD1d(h) for the remainder of this thesis. This will represent the refolded human CD1d molecule with αGC as the lipid cargo, unless otherwise stated.

3.1.2 Expression and purification of human CD1d-αGC for structural studies

CD1 is a heterodimer consisting of a three-domain heavy chain (α1-α3) and β2M, with a lipid antigen bound to stabilise the complex and assist in TCR recognition. CD1d(h) was refolded in vitro using an oxidative refolding method with αGC as the loaded lipid. Initial inclusion body generation yielded suboptimal total yields, prompting an attempt to improve the process. To increase yield and reproducibility, inclusion body generation was tested using AI media, and growth time and total yield were compared to conventional inclusion body production methods.

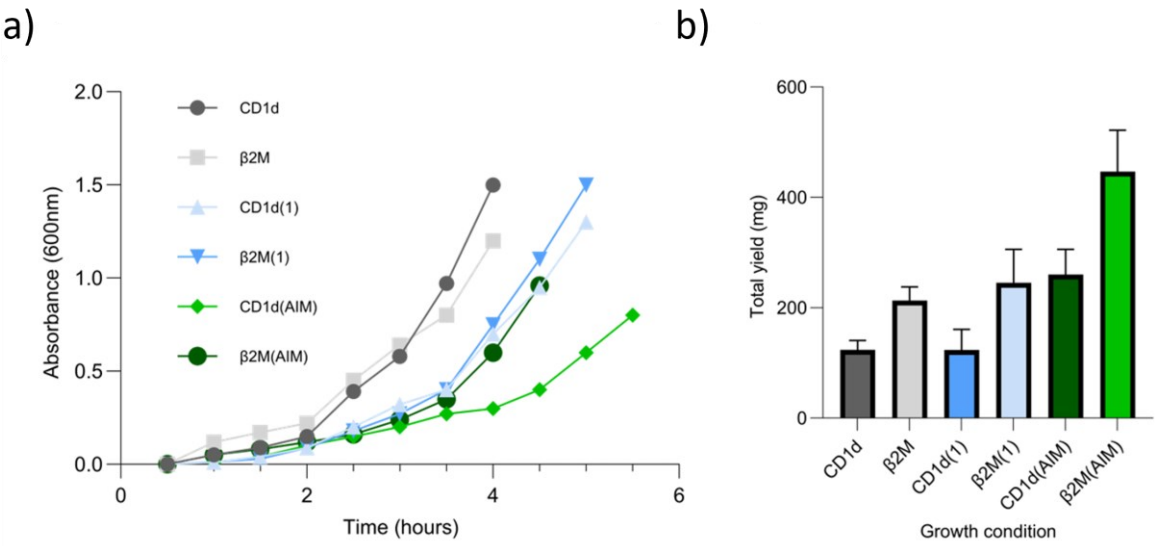


Figure 3. 3 Growth curves and total yield of CD1c and β2M inclusion bodies

CD1d and β2M inclusion bodies were grown under varied conditions to optimise total inclusion body yield. (a) Growth curves of CD1d and β2M cultures inoculated with 10ml overnight culture (grey), 1 ml overnight culture (blue), and using AI media (green), with measurements taken every 30 minutes until OD600 was reached. (b) Total inclusion body yield (mg) in triplicate for all growth conditions, showing AI media consistently producing higher total yield of inclusion bodies.

Following optimisation of inclusion body growth, we proceeded to investigate the purity of inclusion bodies using SDS-PAGE gel electrophoresis.

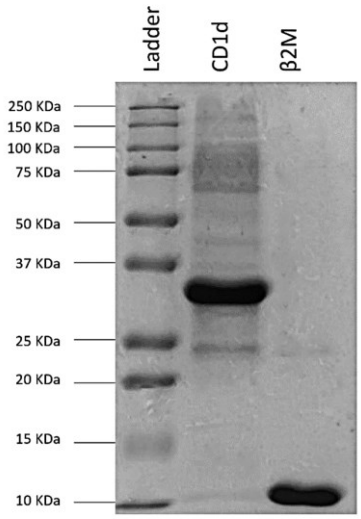


Figure 3. 4 Inclusion body purity of CD1d and β2M protein assessed by gel electrophoresis.

Representative SDS-PAGE gel showing CD1d and β2M inclusion bodies following bacterial expression with Rosetta (DE3) pLysS cells. Both sets of inclusion bodies were run with denaturing dye and clean bands are present at the correct molecular weight of roughly 33KDa and 11KDa for CD1d and β2M respectively. Both sets of inclusion bodies in this figure were produced using AI media.

Inclusion bodies were clean and migrated at the expected molecular weight relative to the Bio-Rad protein standard ladder. Bands generated from inclusion body production in low-salt broth were consistently smaller, indicating a lower yield, despite appearing equally as clean as those produced using AI media (data not shown).

Following successful inclusion body production, refolding of the CD1d heavy chain with β2M and αGC was attempted. The resulting protein complexes were successfully purified, with gel filtration yielding retention times consistent with previous CD1 refolds. Protein was concentrated using stirred cell concentrators and spin columns, followed by two-step gel filtration purification, which produced distinct peaks corresponding to CD1d protein fractions.

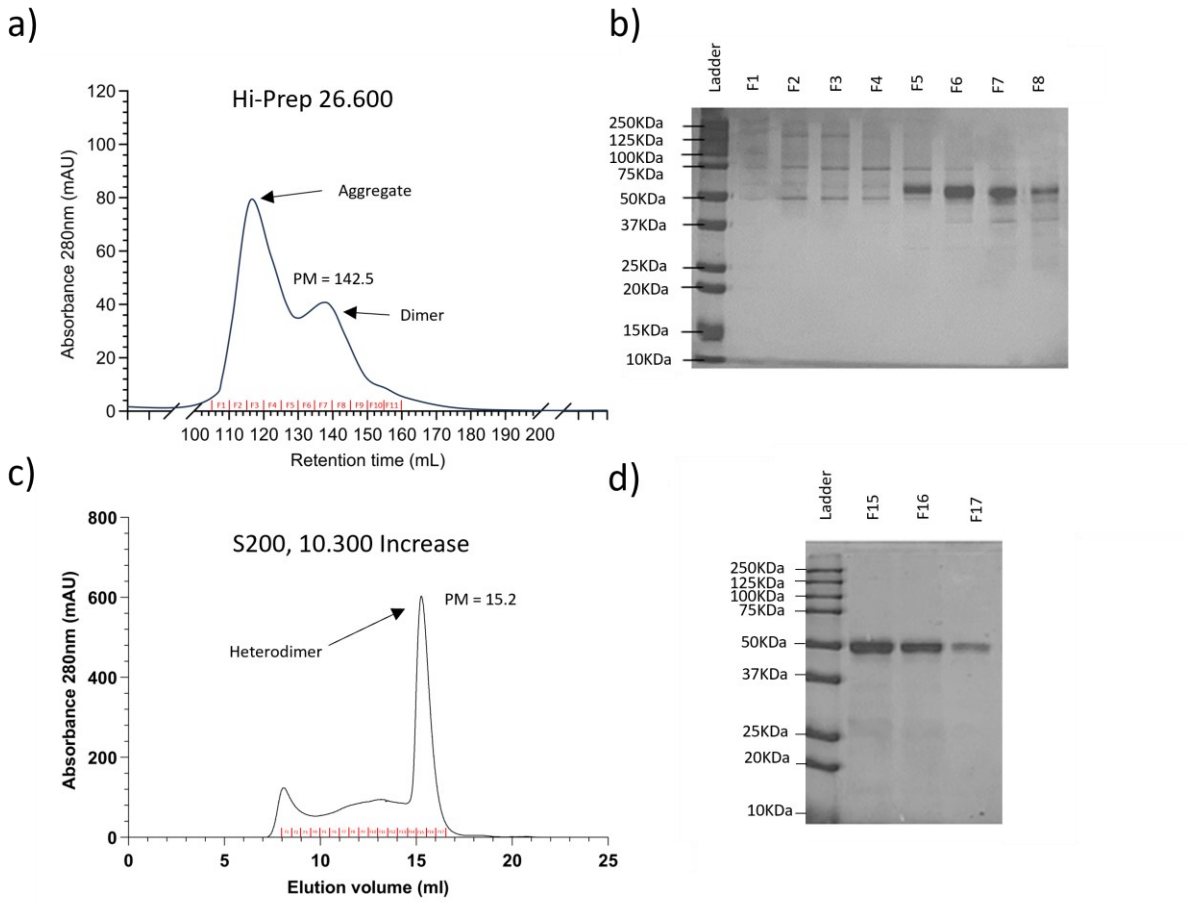


Figure 3. 5 Purification of CD1d(h)-αGC

a) Gel filtration/size exclusion chromatography trace of refolded human CD1d-αGC protein, post-concentration to <10ml. 5-10ml of protein was loaded using a 10mL super loop, and the program was run at 2.4mL/min. Fractions at intervals of 5ml were collected and run on an SDS-PAGE gel. b) SDS-PAGE gel of fractions F1-F8 from the Hi-Prep chromatography run, under reducing conditions (4X loading dye). c) Analytical gel filtration/size exclusion chromatography trace of fractions of interest collected previously (F5-F8). Combined fractions were concentrated to <500ml and run at 0.2mL/min on the S200 increase column. Fractions were loaded using a 0.5ml loop. The resulting peaks were collected in 0.5ml fractions and analysed via SDS-PAGE. d) SDS-PAGE gel of F15-F17 from the previous size-exclusion purification, run under non-reducing conditions with protein standards ladder from Bio-Rad in lane 1.

Before crystallisation, we tested the functionality of refolded CD1d(h) protein using tetramer staining assays. Biotinylated CD1d(h) was successfully re-purified, and biotin ELISA confirmed successful biotinylation following overnight incubation at 4°C.

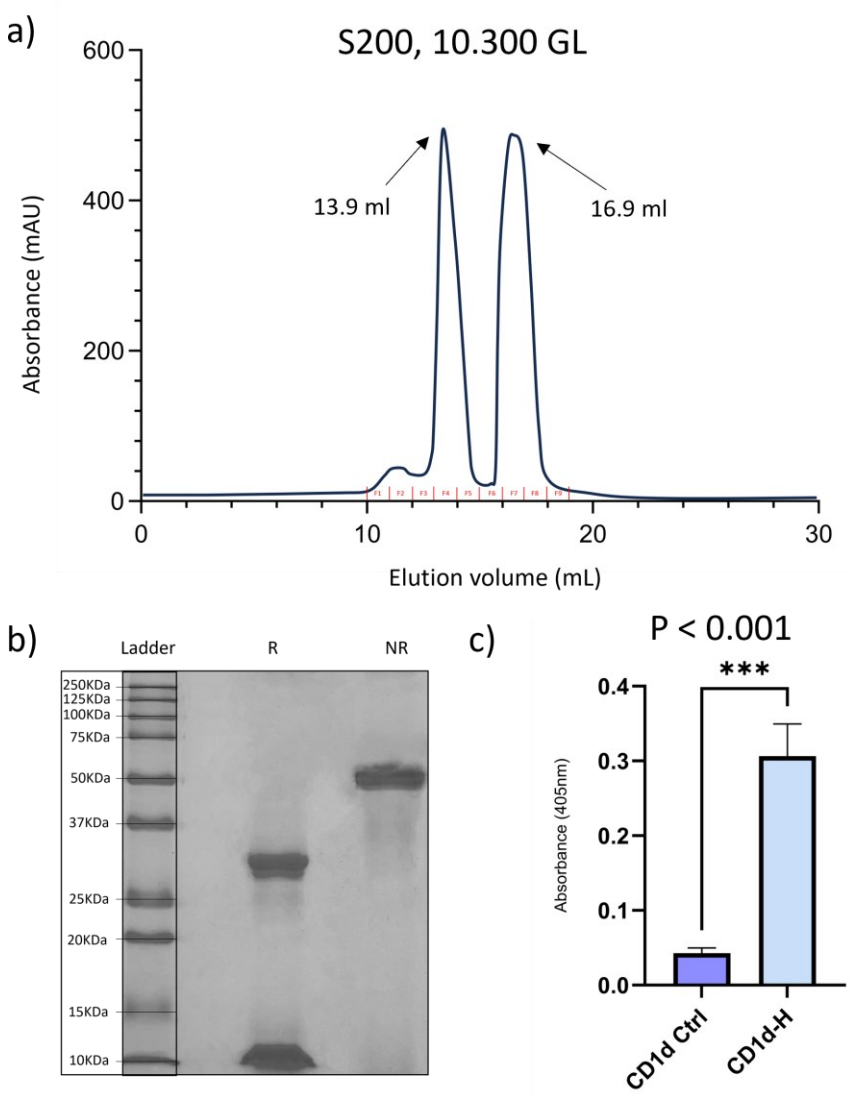


Figure 3. 6 Re-purification and analysis of biotinylated human CD1d-αGC

a) Gel filtration/size exclusion chromatography trace of biotinylated CD1d-αGC protein. Protein was loaded via a 0.5mL super loop and run at 0.2 mL/min, with 0.5ml fractions collected and run on SDS-PAGE. b) Reducing and non-reducing SDS-PAGE analysis of biotinylated CD1d-αGC from fraction four of the elution. c) Biotin ELISA from fraction four of the purified sample, with unbiotinylated CD1d used as a control (n=3 for all samples shown in the ELISA).

Following confirmation that our CD1d(h) protein was successfully biotinylated and remained stable throughout re-purification, we assessed its ability to bind to a CD1d-reactive TCR. Binding was successfully detected using multiple assay formats, confirming that the refolded CD1d(h) was functionally active.

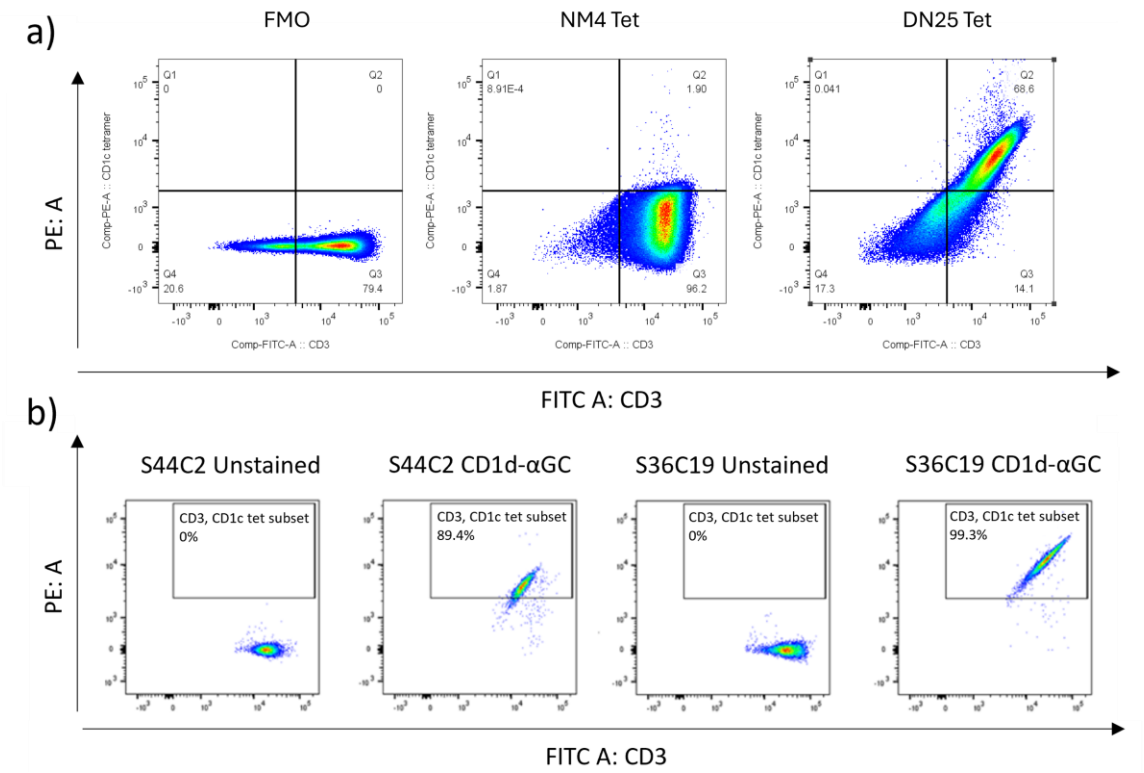


Figure 3. 7 Human CD1d-αGC binds to iTCRs.

a) Flow cytometry plots showing CD1d-αGC tetramer staining of a human polyclonal iNKT cell line expressing CD1d-αGC reactive iTCRs. Unstained cells and cells stained with an irrelevant TCR tetramer (NM4 tetramer) served as control. b) Flow cytometry plots showing CD1d-αGC tetramer staining of two iNKT-cell clones, S44C2 and S36C19. Both clones stain with CD1d-αGC tetramer, the latter to a greater extent relative to the unstained control.

We observed successful staining with our CD1d(h) tetramer, showing strong diagonal staining of polyclonal iNKT cells. The CD1d(h) tetramer also bound to iNKT clones, with stronger binding to the S36C19 line than the S44C2 line (*Figure 3.7*). Mean fluorescence intensity (MFI) values indicated a difference in binding between the two TCRs. This was expected, as the latter expresses a lower affinity iTCR and the former a higher affinity TCR. Despite GC demonstrating a weaker capability to discriminate between high and low affinity iTCRs in previous studies, there still appeared to be a difference in staining, suggesting different binding mechanisms by these two TCRs which subsequently alters the MFI value.

To further assess absolute binding affinity, we conducted SPR experiments using two in-house generated soluble iTCRs to evaluate CD1d(h) recognition.

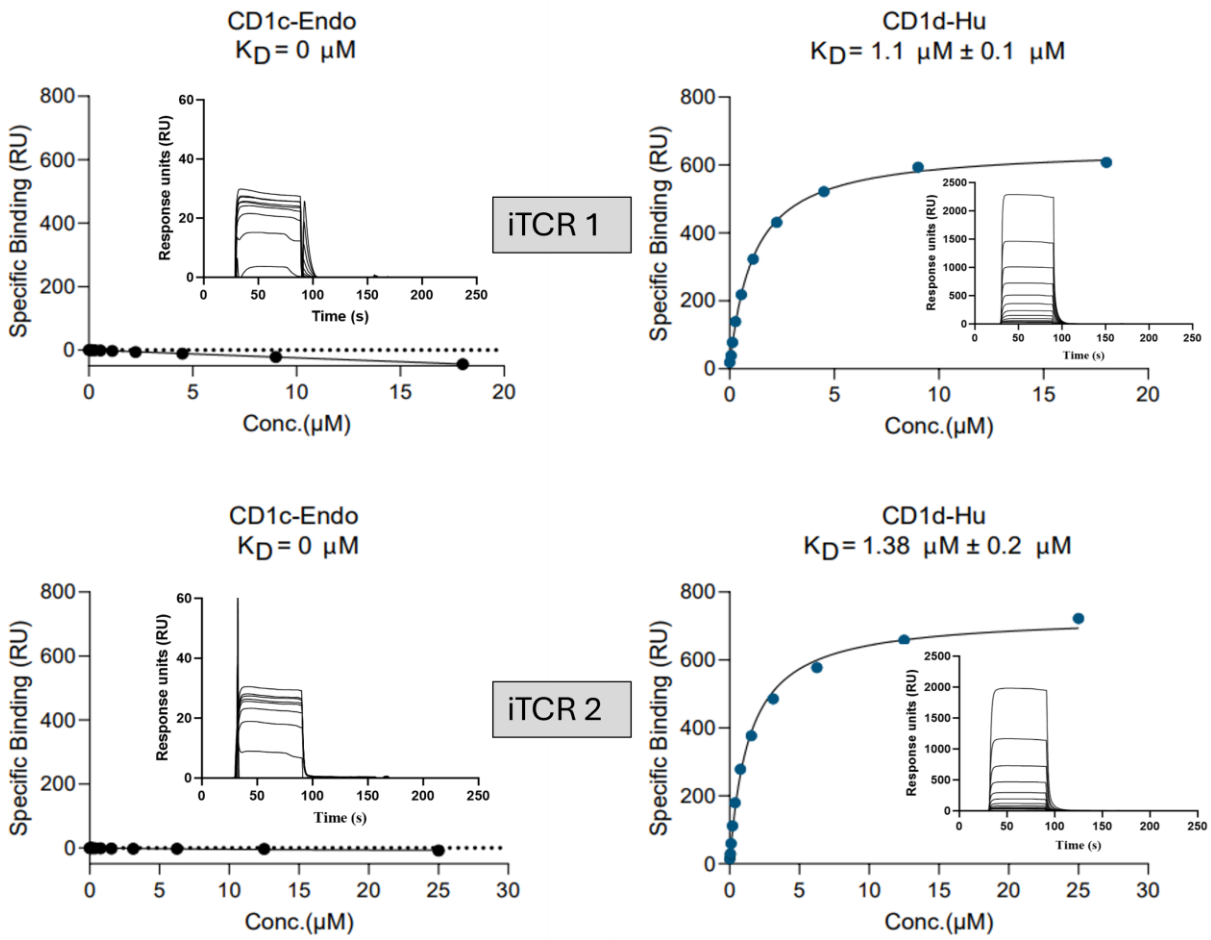


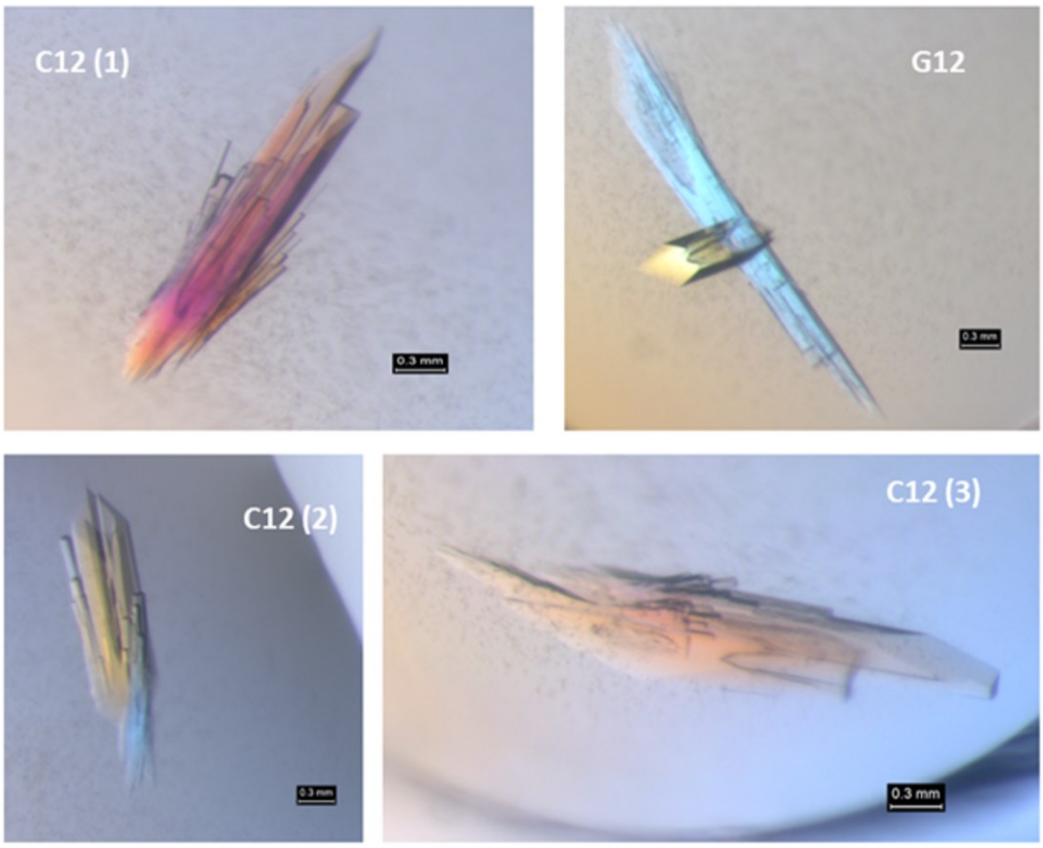
Figure 3. 8 Soluble iTCRs binding to CD1d-Hu and CD1c-Endo

SPR binding curves and calculated dissociation constants (KD_DD) for soluble iTCRs binding to CD1d-Hu and the CD1c-Endo control. Biotinylated CD1 monomers were immobilised on a CM5 chip, and iTCRs were flowed over the surface at 30 µL/min in 12 dilutions. Specific binding was plotted against concentration using GraphPad Prism, and equilibrium binding constants (KD_DD) were determined by non-linear regression fitting. Inset graphs display representative sensorgrams from Biacore experiments, showing real-time response units (RU) over time, which illustrate the association and dissociation kinetics for each interaction. RU, response units. *(SPR experiments were conducted in collaboration with Immunocore.)*

3.1.3 Investigating CD1d-αGC using X-ray crystallography.

With conformation that our CD1d(h) was fully functional, as demonstrated by SPR binding studies and tetramer staining experiments, we attempted crystallisation of the protein. Non-biotinylated CD1d(h) was thawed and concentrated using Viva-spin 20 columns before crystallisation. The protein was initially set up at a concentration of 10 mg/ml using standard crystallisation screens from Molecular Dimensions (JCSG+, TOPS, Morpheus).

Initial crystallisation attempts yielded mostly empty wells, with some showing aggregation, suggesting our initial concentration was too low. Based on these observations, the protein concentration was increased to 14Mg/ml for subsequent crystallisation attempts.



.

Figure 3. 9 Microscope images of CD1d(h) crystals

Microscope images showing CD1d(h) crystals grown in the Morpheus crystallisation screen from molecular dimensions. Crystals were grown using sitting drop vapour diffusion in 96-well plates after approximately 7-8 days of incubation at 21°c. Wells that grew crystals were C12 and G12 from the Morpheus screen, using a protein: mother liquor ratio of 1:1. The crystals were large and generally consisted of multiple crystals growing into each other, however appeared to separate on loop contact.

Crystallisation well	Conditions
Morpheus C12	12.5% w/v PEG 1000, 12.5% w/v PEG 3350, 12.5% v/v MPD 0.03M of each NPS 0.1 m bicine/Trizma base pH8.5
Morpheus G12	12.5% w/v PEG 1000, 12.5% w/v PEG 3350, 12.5% v/v MPD 0.02M of each carboxylic acid 0.1 m bicine/Trizma base pH8.5

Table 1 Crystallisation conditions of each well from which a CD1d-αGC crystal grew within the Morpheus crystallisation screen.

Increasing the protein concentration to 14Mg/ml resulted in large crystals forming in multiple wells of the Morpheus crystallisation screen *(Figure 3.9*). Crystals from wells C12 (1), C12 (2), and G12 were fished out, flash-frozen in liquid nitrogen, and sent to the European Synchrotron Radiation Facility (ESRF) for X-ray diffraction experiments.

Successful diffraction data was obtained from the crystal in well G12, and molecular replacement was completed using a model generated by AlphaFold2.pynb-Colaboratory as the search model and ran multiple refinement rounds using CCP4i2 software. The final refined structure was solved at 1.72 Å resolution, providing high-quality data for further analysis of lipid antigen presentation by CD1d. The structure was then assessed by a combination of ChimeraX and PyMOL protein modelling software, to investigate whether the high-resolution data revealed any new insights into lipid antigen presentation by the CD1d protein.

Table 2 Data collection and refinement statistics (CD1d-Hu)

Datset	**CD1d-Hu**
Data Collection	
Space Group	P 21 21 2
Cell Dimensions	
a (Å)	117.679
b (Å)	156.433
c (Å)	51.936
α°	90
β°	90
γ°	90
Resolution	49.29-1.78 (1.85-1.78)
Rmerge	0.58 (0.598)
I/σI	11.2 (1.79)
Wilson B	28.2
Completeness (%)	99.8 (98.62)
Refinement	
Resolution (Å)	1.78
No. of reflections	
all/free	96,529
Rwork/Rfree	0.186/0.212
Ramchandran plot	
Favoured	98.4
Allowed	100%
Total no. of atoms	13554
Protein	6336
Ligand/Ion	160
Water	736
RMSD	
Bond lengths (Å)	0.015
Bond angles (°)	1.415

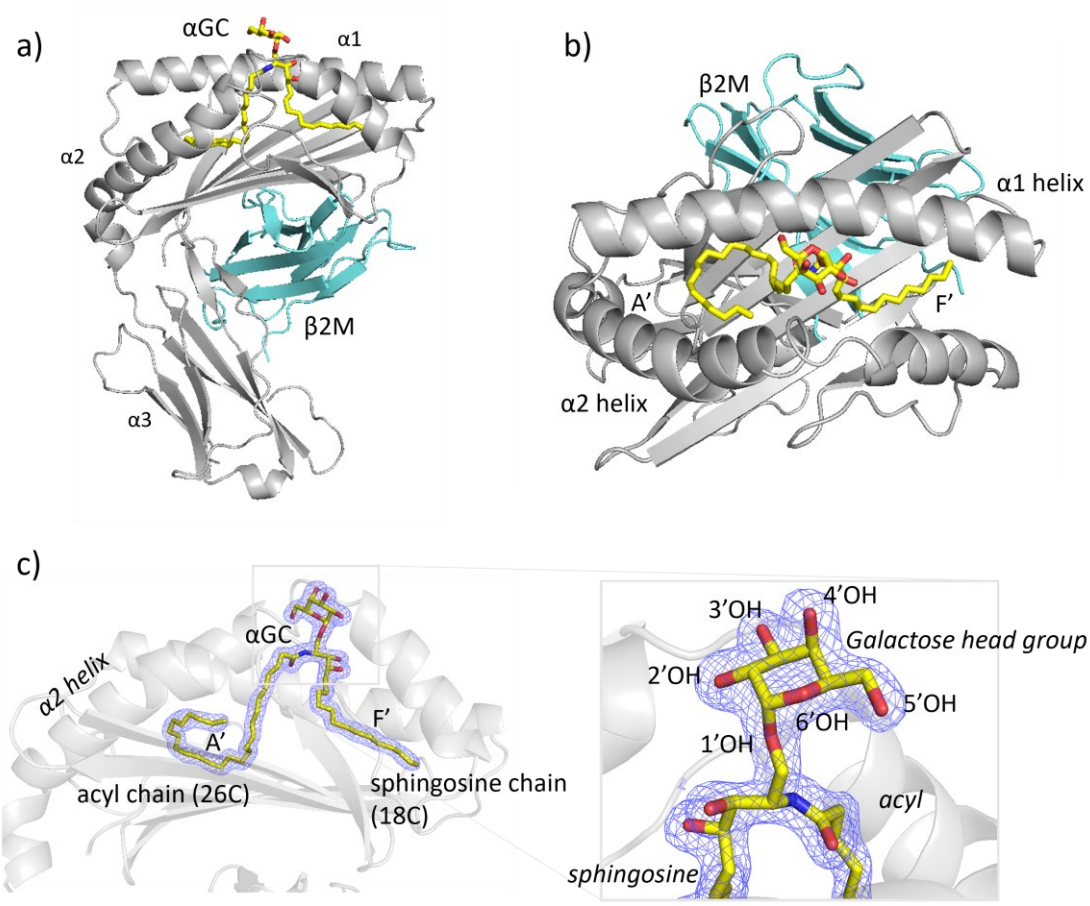


Figure 3. 10 High-resolution crystal structure of CD1d(h)

a) Front-facing view of the human CD1d-β2m complex bound with αGC in its lipid binding groove. The galactose head group is positioned above the CD1d molecule for contact with TCR, while the lipid chains are buried within the lipid biding grooves b) Birds-eye view of CD1d-αGC showing acyl and sphingosine chains in the A' and F' grooves, respectively. c) Strong αGC electron density calculated using an omit map, contoured to 3σ, is shown. CD1d heavy chain is illustrated in grey, β2m in cyan, and αGC in yellow, with nitrogen and oxygen coloured in blue and red, respectively. Images were generated using PyMOL.

The structure revealed the tertiary and quaternary organisation of CD1d(h), consisting of a three-domain heavy chain non-covalently associated with β2M. The lipid αGC was positioned within the CD1d binding groove in a typical fashion, with its acyl and sphingosine chains fully occupying the A' and F' channels, respectively, while the galactose head group sat atop the molecule, positioned for TCR contact.

Electron density was well resolved across the entire lipid moiety and surrounding amino acid side chains, enabling detailed molecular analysis of the differences between this CD1d(h) structure and the previously published structure (PDB accession code 1ZT4).

A comparison of the lipid moiety and binding groove between our structure and 1ZT4 revealed a high degree of similarity. The only noticeable difference was a slight reorientation at the start of the sphingosine chain, where the O2 oxygens of the sphingosine chain (*Figure 3.11b*) were positioned at a 90° angle to each other. This did not affect hydrogen bond formation between the lipid and surrounding residues.

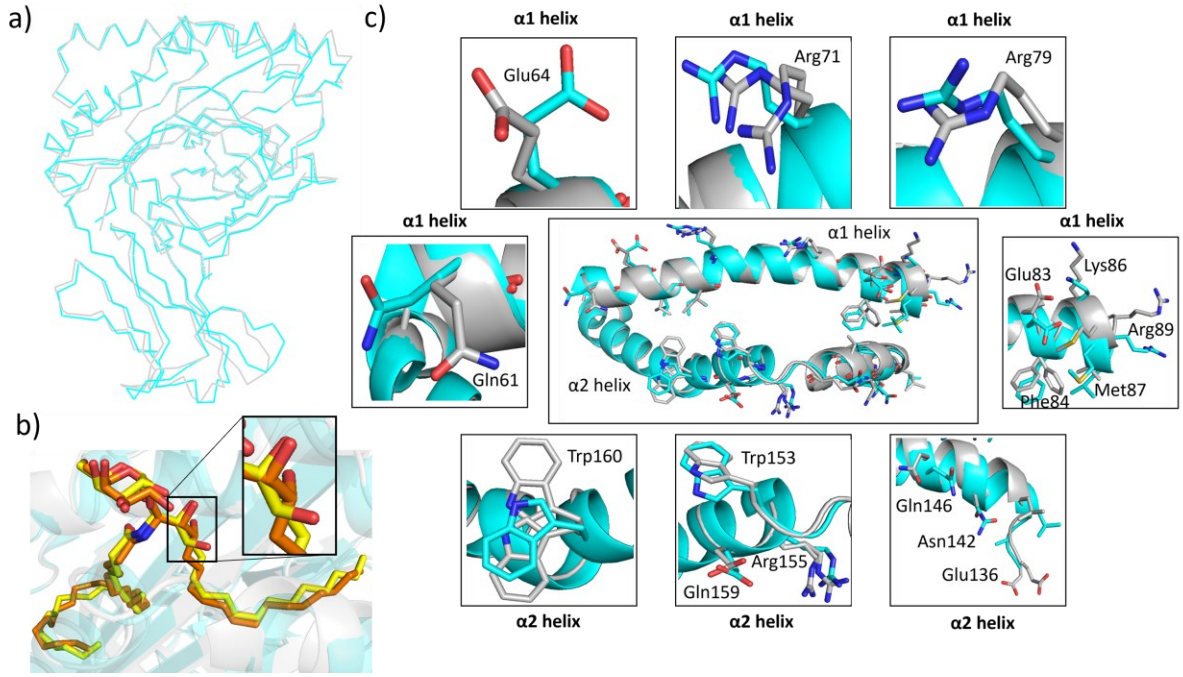


Figure 3. 11 Structural comparison of our CD1d(h) with 1ZT4

a) Overall comparison of our CD1d(h) structure (grey) with the published 1ZT4 lipid bound structure (cyan). Both molecules were overlayed using chimera and shown in ribbon view. b) Overlay of the αGC lipids in stick view, with 1ZT4 in orange and our model in yellow. c) Detailed molecular analysis of side chain differences within the α1 and α2 helices, with the three top, two middle, and three bottom panels showing zoomed-in areas of the α1 and α2 helices respectively. Amino acids showing substantial differences are labelled. Nitrogen atoms are shown in blue and oxygen in red. All images were generated using PyMOL.

A detailed molecular analysis of the helical segments in both structures revealed some side-chain differences, primarily in regions distant from the TCR binding sites or the lipid-binding groove. Most structural features were highly conserved, including Phe70, which forms a central pole around which the acyl chain curves in an anticlockwise direction, as previously described by Koch et al (72).

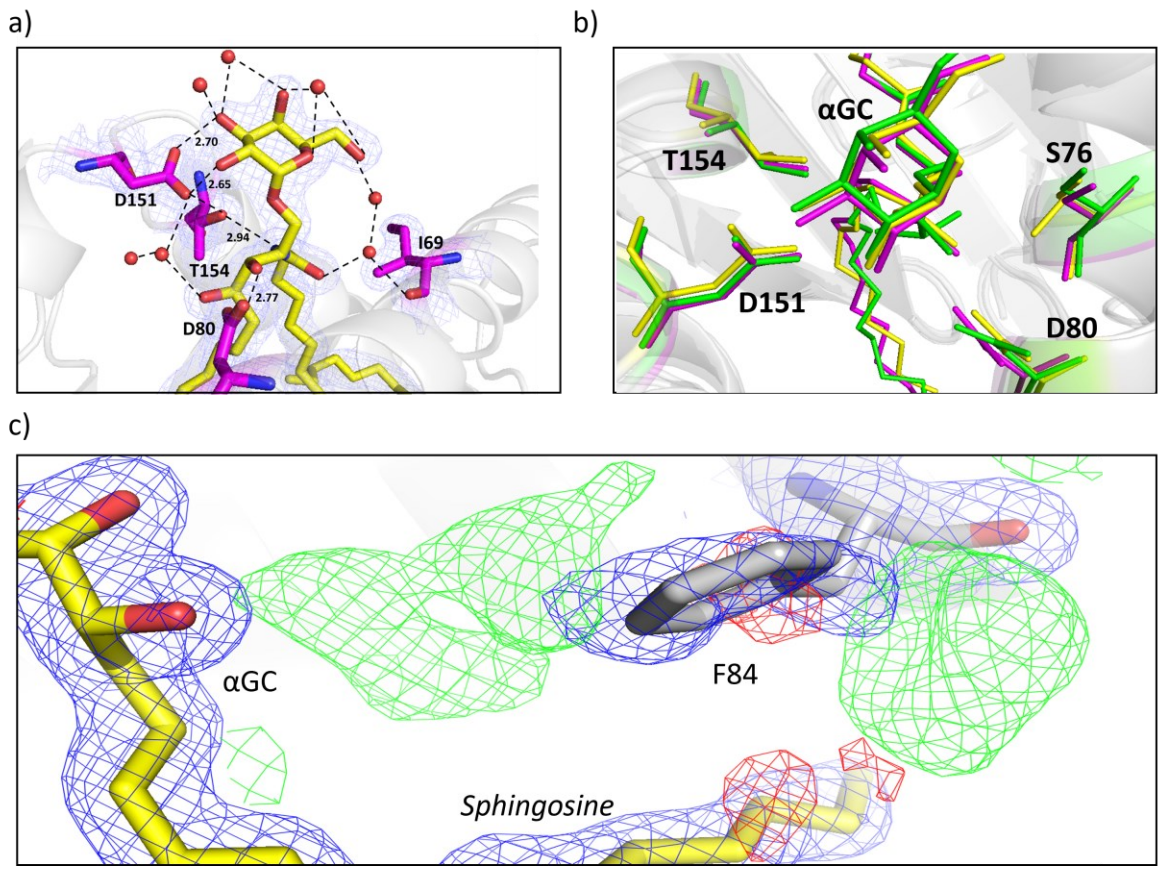


Figure 3. 12 New insights into CD1d(h) antigen presentation

Structural representation of the new insights from our high-resolution CD1d(h) structure a) Hydrogen bonding network surrounding the αGC lipid, with important amino acid residues labelled and coloured in magenta. Waters are shown as red spheres and hydrogen bonds as black dashed lines. b) Structural overlay and comparison of our new CD1d-αGC structure with 2PO6 (CD1d αGC-human iNKT complex) and 3HUJ (CD1d-αGC-mouse iNKT complex). c) Difference density maps illustrating discrepancies in density surrounding the Phe84 residue when in its commonly seen 'canonical' binding position. Green density indicates multiple conformations with another large patch of density in the cavity between lipid and Phe84. Blue density represents the 2Fo-Fc map and is contoured at 1σ, while red and green represent negative and positive Fo-Fc maps, respectively, contoured at 3σ. All images were generated using PyMOL.

A comparison of the hydrogen bonding networks between the two structures revealed strong similarities, with the only direct difference being an additional bond from D151 to the 3'OH group of the αGC headgroup (*Figure 3.12a*).

Ile69 was also observed to be coordinated into the hydrogen bonding network through two well-defined water molecules. Additionally, density variation around Ser76 suggested the presence of multiple conformations, as modelled in *Figure 3.12b*.

Further analysis of Phe84 in the 'canonical' conformation *(Figure 3.12c)* revealed differences in electron density on either side of the residue. We decided to investigate this in greater detail using multiple omit maps contoured to 3 (Fo-Fc).

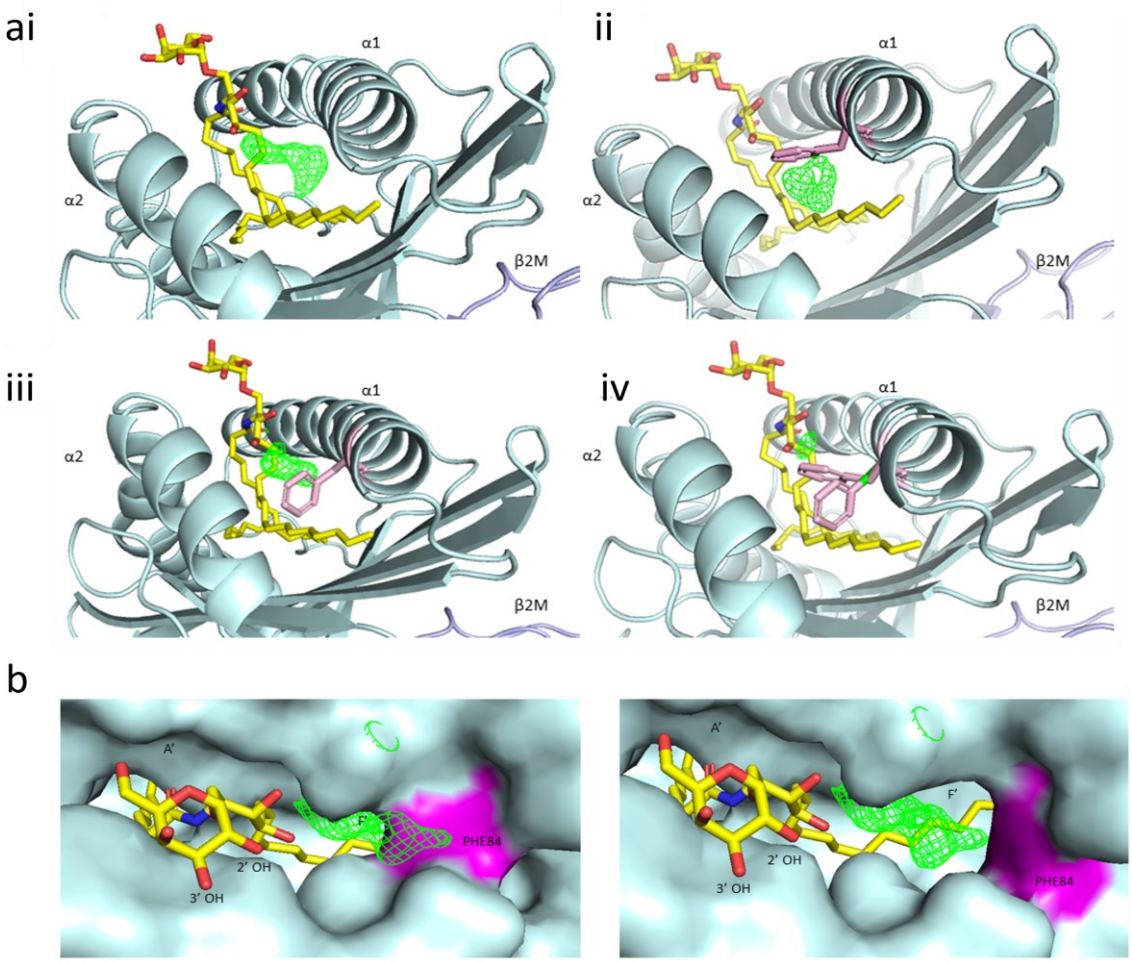


Figure 3. 13 Flexibility of Phe84 in CD1d(h)

PyMOL figures displaying the flexibility of Phe84 and the resulting electron density calculated using an omit map and contoured at 3σ. ai) Phe84 removed, and structure refined showing a difference in density in the absence of a sidechain. ii) and iii) show the surrounding difference in density from Phe84 in the 'canonical' and 'non-canonical' position respectively. iv) Phe84 is shown in the dual conformation with the difference in density disappearing. b) Surface representations showing Phe84 in the canonical and non-canonical positions, and the resulting cavity volume difference in each position. The difference in density from within the cavity space left by the outwards rotation of phenylalanine is shown in green. Phe84 is shown in magenta, lipid in yellow coloured by heteroatom, and CD1d heavy chain is shown in teal.

Electron density maps omitting Phe84 in either the canonical or non-canonical position, as well as both positions simultaneously (all contoured at 3σ), confirmed the bi-conformational flexibility of the Phe84 residue (*Figure 3.13a*). A distinct area of green density was consistently observed between the αGC lipid and Phe84, regardless of rotamer conformation (*Figure 3.13b)*.

Overlay analysis with previously published CD1c-SL and CD1c-MPM structures was then performed to compare the observed side-chain reorientation with that of a previously published CD1c-SL (65).

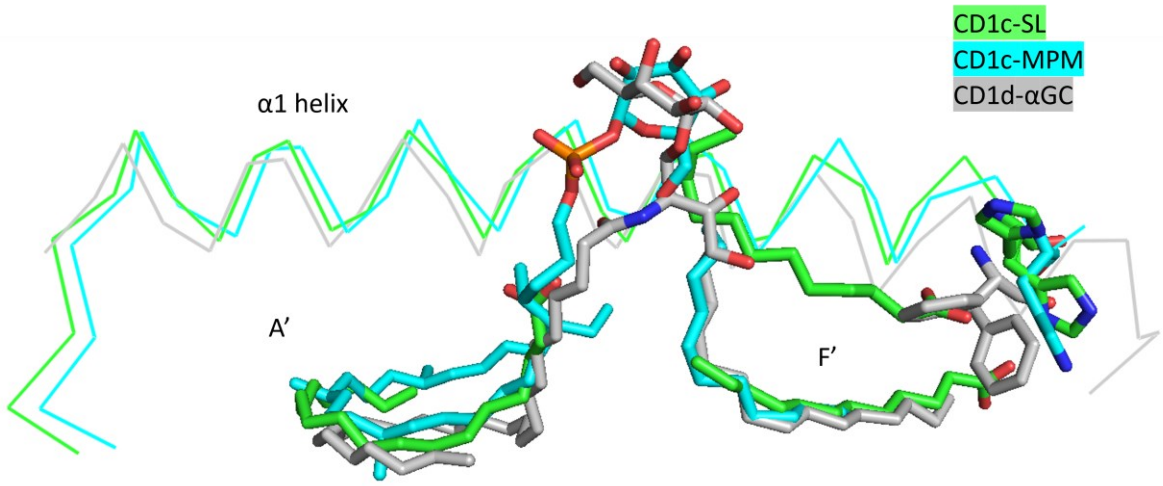


Figure 3. 14 Structural overlay of CD1d(h), CD1c-MPM, and CD1c-SL

Comparison of the α1 helices, lipid cargo, and aromatic residue in the CD1c-SL (green), CD1c-MPM (cyan) and human CD1d(h) (grey) structures. Lipids for each structure are shown in stick representation with hydrocarbon chains occupying the A' and F' grooves. Aromatic residues are shown in stick representation and α-helices are represented by ribbons. Lipids and residues are coloured by heteroatom, with oxygen in red and nitrogen in blue. CD1c-SL and CD1c-MPM structures were downloaded from the PDB under accession codes 5C9J and 30V6 respectively. The image was generated using PyMOL.

CD1c-SL and CD1c-MPM both show a significant upward shift in their α1 helix compared to CD1d. Similar conformational adaptability has been demonstrated in CD1b (175). We overlaid our CD1d-αGC structure with the CD1b structure depicted by Gadola et.al, to investigate how the two structures compared in the area surrounding this phenylalanine. Interestingly, CD1b shows a direct conservation at this residue *(Figure 3.1).*

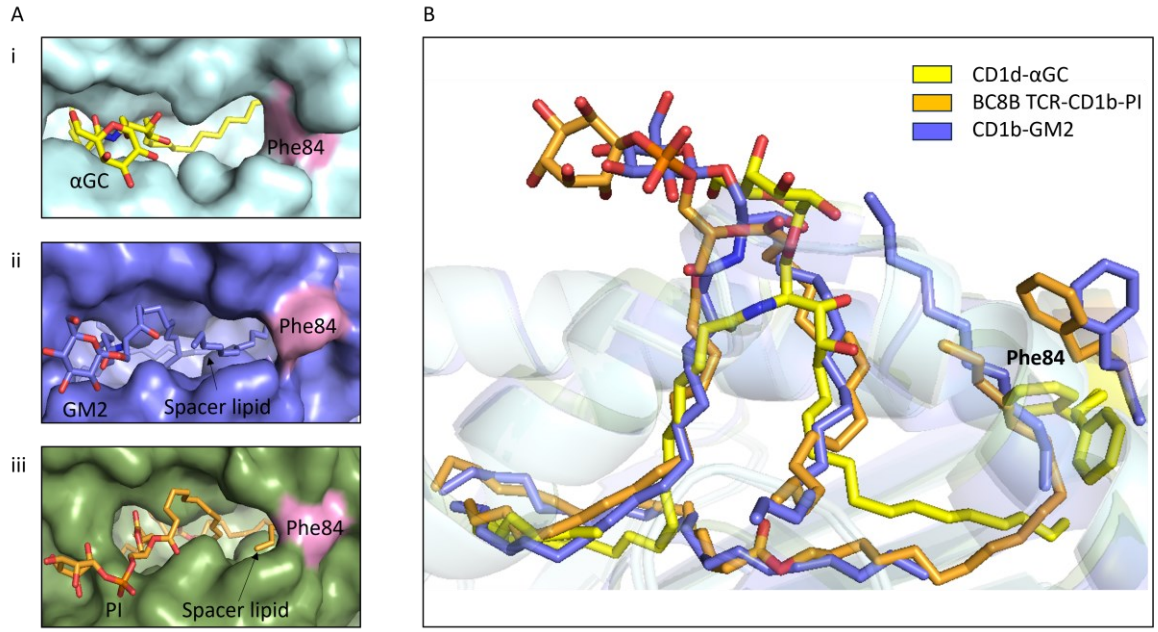


Figure 3. 15 Structural comparison of CD1d-αGC with CD1b(h)

A i) Surface representation of our human CD1d-αGC structure in the open conformation ii) CD1b in complex with GM2 ganglioside and spacer lipid iii) CD1b in complex with BC8B TCR, PI, and spacer lipid. All images are zoomed in on the lipid-binding site with Phe84 highlighted in pink and lipids highlighted in yellow, slate, and orange for αGC, GM2, and PI respectively. B) Stick representation of a structural overlay between our CD1d structure and the two previously highlighted CD1b structures, with emphasis on the Phe84 position which changes depending on lipid occupancy.

The overlay supports known differences in CD1b and CD1d binding groove architectures but demonstrates a similar mechanism by both molecules whereby the region surrounding Phe84 can adapt conformation to accommodate a combination of different lipids, buffer molecules, or detergents. This phenomenon is further supported by a CD1d(m) structure by Wun et.al, where a combination of binary and ternary complex data elucidates similar conformational adaptability in the Leu84 sidechain and surrounding area (106). This suggests that despite the smaller sidechain and potentially lesser impact on cavity volume, a similar underlying mechanism is employed by mouse CD1d to accommodate varying lipid moieties and spacer molecules.

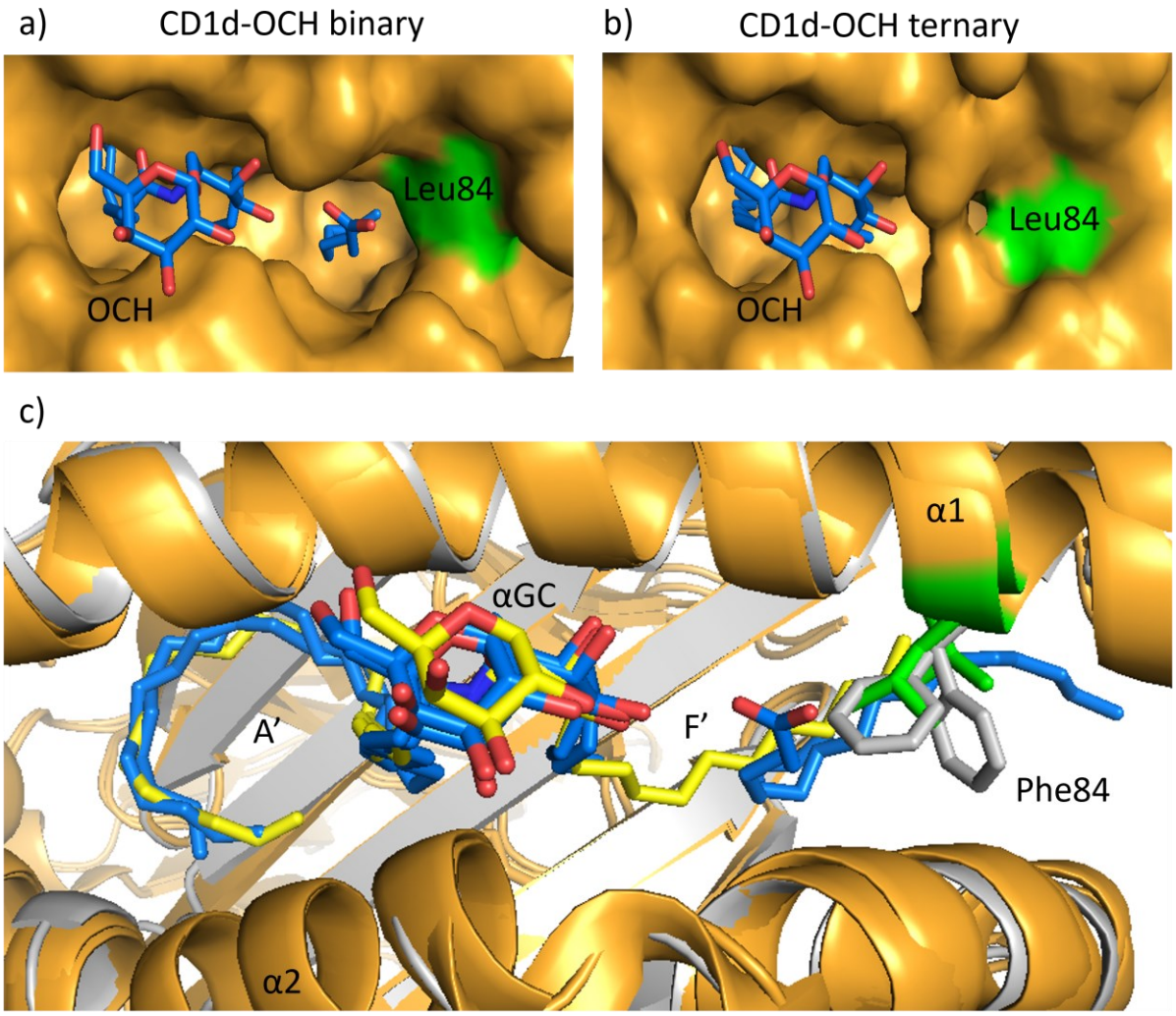


Figure 3. 16 Comparison of our human CD1d-αGC with mouse CD1d-OCH

Surface and cartoon overlays of our CD1d-αGC structure with the CD1d(m)-OCH binary and ternary structures. a) Surface representation of mouse CD1d-OCH with spacer lipid sat within the open cavity generated by Leu84. b) Surface representation of the CD1d-OCH ternary complex with iTCR, showing a closed cavity by an inward rotation of Leu84. c) Cartoon overlay of both CD1d-OCH structures with our human CD1d-αGC. OCH is coloured in blue, αGC in yellow, Phe84 in grey, and Leu84 in green.

Leu84 in the mouse CD1d-OCH structure displayed rotational flexibility, allowing for binding groove cavity volume alterations based on lipid occupancy. Overlay analysis with our human CD1d(h) structure showed a similar angle of rotation between Leu84 in mouse CD1d and the two Phe84 orientations observed in our electron density maps. The discovery of these similarities across both isoforms and in a different CD1d expressing species prompted further investigation into the stability of the Phe84 in an open conformation in our new CD1d(h) structure.

To further investigate the relative occurrence of each rotamer conformation, REST2 MD simulations were performed using each independent rotamer as a starting conformation. The two states were distinguished by their dihedral angles (N-Cα-Cβ-Cγ) of -84.4° (canonical) and -161.1° (non-canonical). Simulations showed that when initially modelled in the non-canonical conformation, Phe84 flipped immediately (within 20 ps of NVT equilibration). In contrast, the canonical form remained stable throughout the simulations.

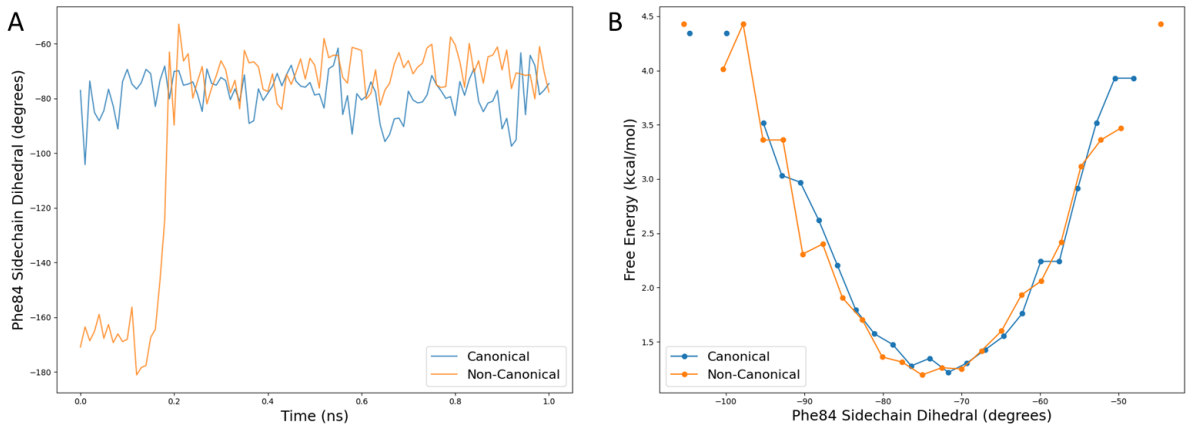


Figure 3.17 Molecular dynamics simulation graphs of Phe84 in CD1d(h)

Molecular dynamics simulations of CD1D-αGC. A) Phe84 Sidechain dihedral angle during NVT equilibration from both canonical and non-canonical Phe84 starting conformations. B) Potential of mean force plot for Phe84 sidechain dihedral across the subsequent REST2 simulation trajectory.

Electron density maps indicated an alternate conformation of Phe84, suggesting the presence of an undefined electron density in the lipid-binding groove. To further explore this, a hydrophobic, four-carbon spacer lipid (butane) was docked into the region, and five independent conventional MD simulations were performed. To investigate this further, we docked a hydrophobic, 4-carbon spacer lipid (butane) into the lipid-binding groove and ran 5 independent conventional MD simulations using SwissDock. The molecule was docked in the space indicated by x-ray crystallography map to contain unexplained density.

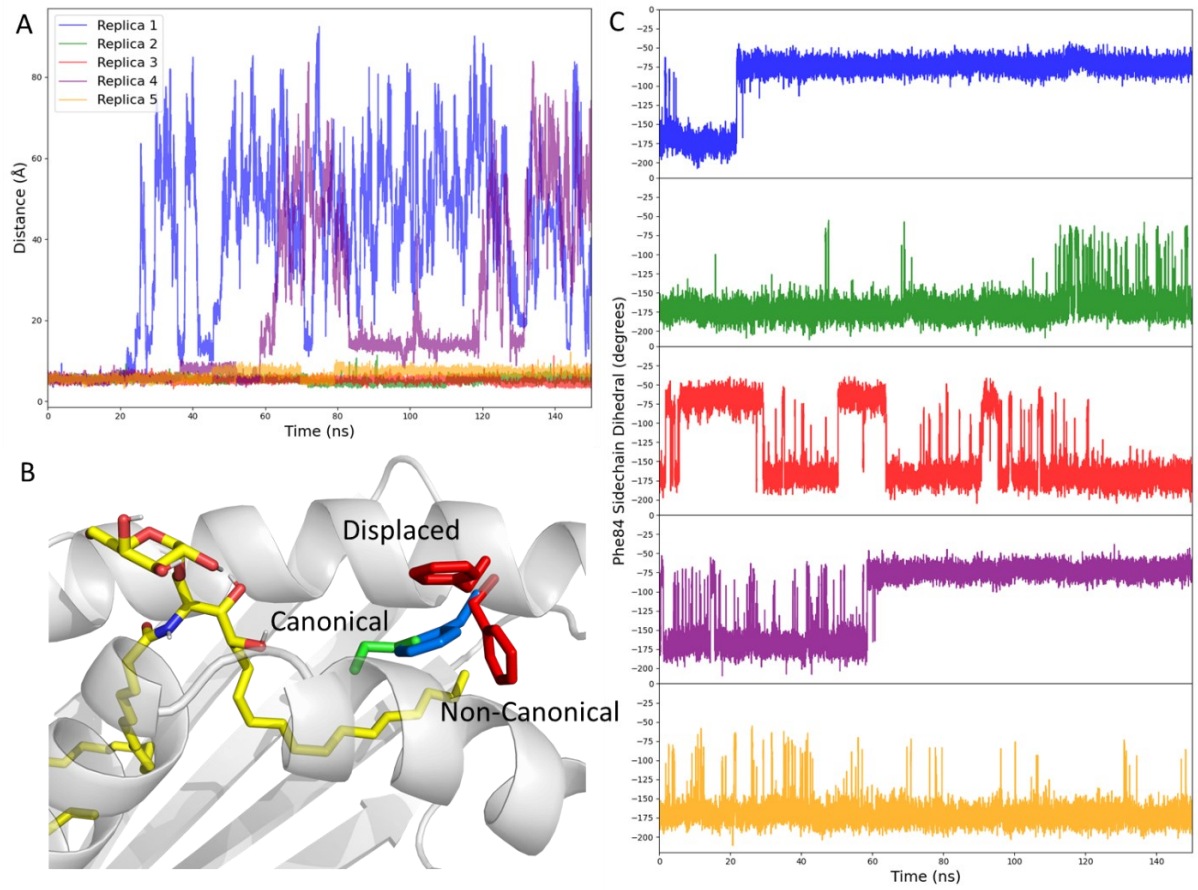


Figure 3. 18 Molecular dynamics simulations of CD1D-αGC-butane

Molecular dynamics simulations of CD1D-αGC-butane. A) Representative conformations of the Phe84 sidechain positions observed with butane bound, derived from replica three (Red), and butane unbound, derived from replica one (Blue). Representative conformations of αGC (Yellow) and butane (Green) are overlaid for reference. B) Distance between αGC (O4 atom) and butane centre-of-mass as a measure of butane dissociation from the complex for all five replicas. C) Phe84 N-Cα-Cβ-Cγ atom dihedral angle across the simulation trajectory for all five replicas.

MD simulation results indicated that Phe84 remained in the non-canonical conformation in three out of five separate runs. In replica one and replica four, where butane dissociated from CD1d, Phe84 flipped back to the canonical conformation (*Figure 3.18C*). This suggests the Phe84 residue was predominantly stable in the non-canonical conformation in the presence of the butane molecule.

In simulations where Phe84 remained in the non-canonical conformation, a displaced conformation was observed, in which the Phe84 side chain positioned itself above the butane molecule, further sealing the pocket *(Figure 3.18B*). *(All MD simulations were performed and analysed by Dr Stephen Turner – University of Southampton Chemistry department – collaborator). Figures 3.17 and 3.18 were also generated by Dr Stephen Turner with some input from myself.*

3.2 Investigating CD1 antigen presentation across species.

3.2.1 Sequence comparison of CD1d expressing species.

The apparent conservation of the aromatic residue at the same position across isoforms, as well as the similar functioning of the mouse CD1d Leu84, suggested wider implications of this flexible area of the binding groove. Often, residue or sequence conservation across species is an indicator of functional importance, hence we wanted to investigate this Phe84 residue across CD1d-expressing species, to see if it was also conserved. We conducted sequence alignment across ten different CD1d-expressing species and investigated the phylogeny of the species using phylogenetic tree analysis.

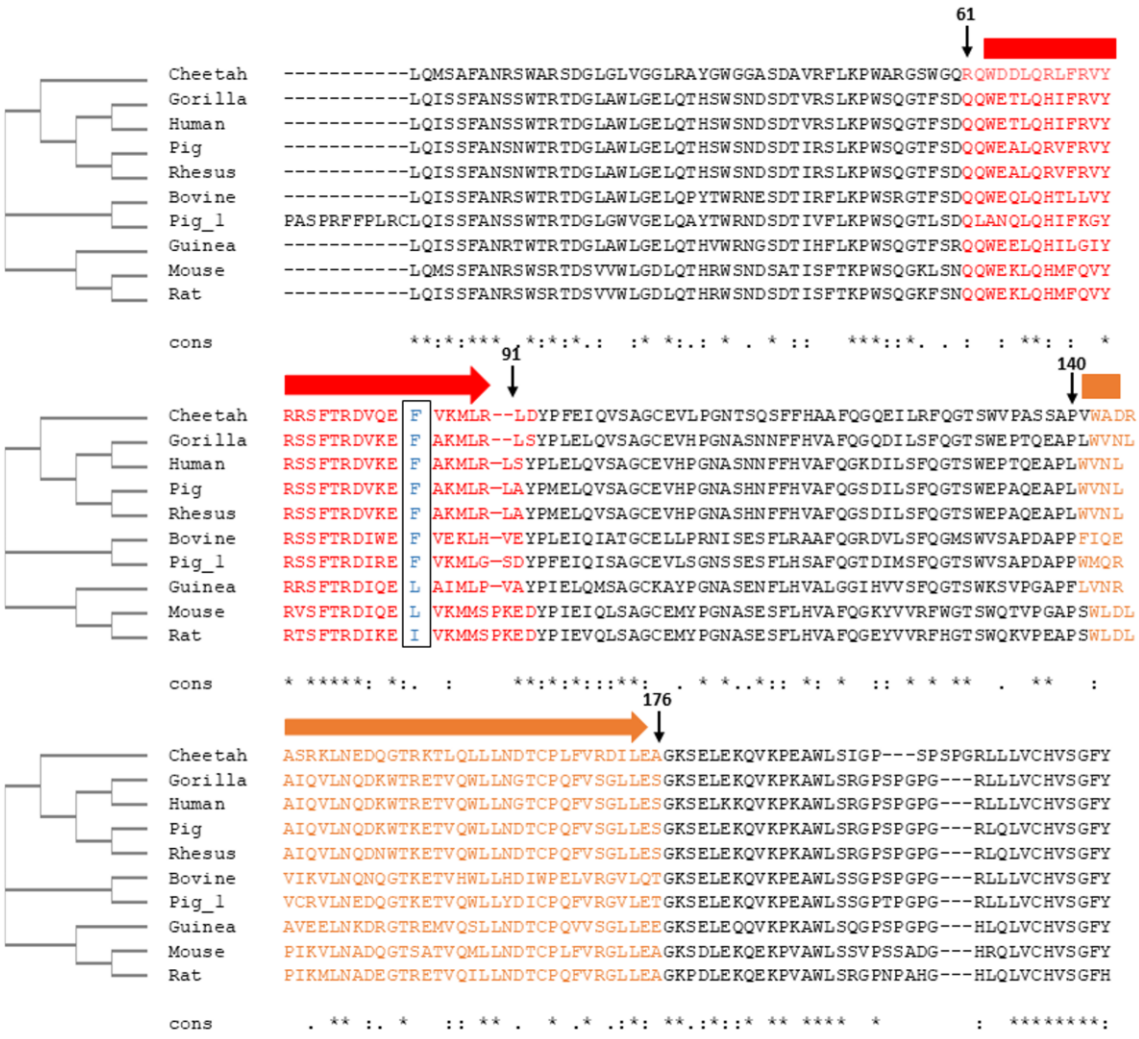


Figure 3. 19 Sequence alignment of CD1d expressing species.

Sequence alignment of ten different CD1d expressing species, with particular focus on the α1 helices, α2 helices, and the equivalent residue to Phe84 in our new CD1d-αGC structure. The phylogenetic tree illustrating species divergence is shown on the left-hand side. α1 helices are coloured in red, α2 helices are coloured in orange, and the amino acid at an equivalent position to Phe84 in our CD1d structure is highlighted in blue.

Sequence analysis revealed the conservation of phenylalanine in seven out of ten of the investigated CD1d-expressing species, with mouse, rat, and guinea pig sequences showing differences. The phylogenetic tree (*Figure 3.19*) indicated that these species diverged earlier compared to those retaining phenylalanine.

To investigate how these sequence variations impact groove architecture, structural data from CD1d-expressing species available in the PDB was analysed. Currently, CD1d has only been crystallised in three species: *Bovine* (cow), *Mus musculus* (mouse), and *Homo sapiens* (human).

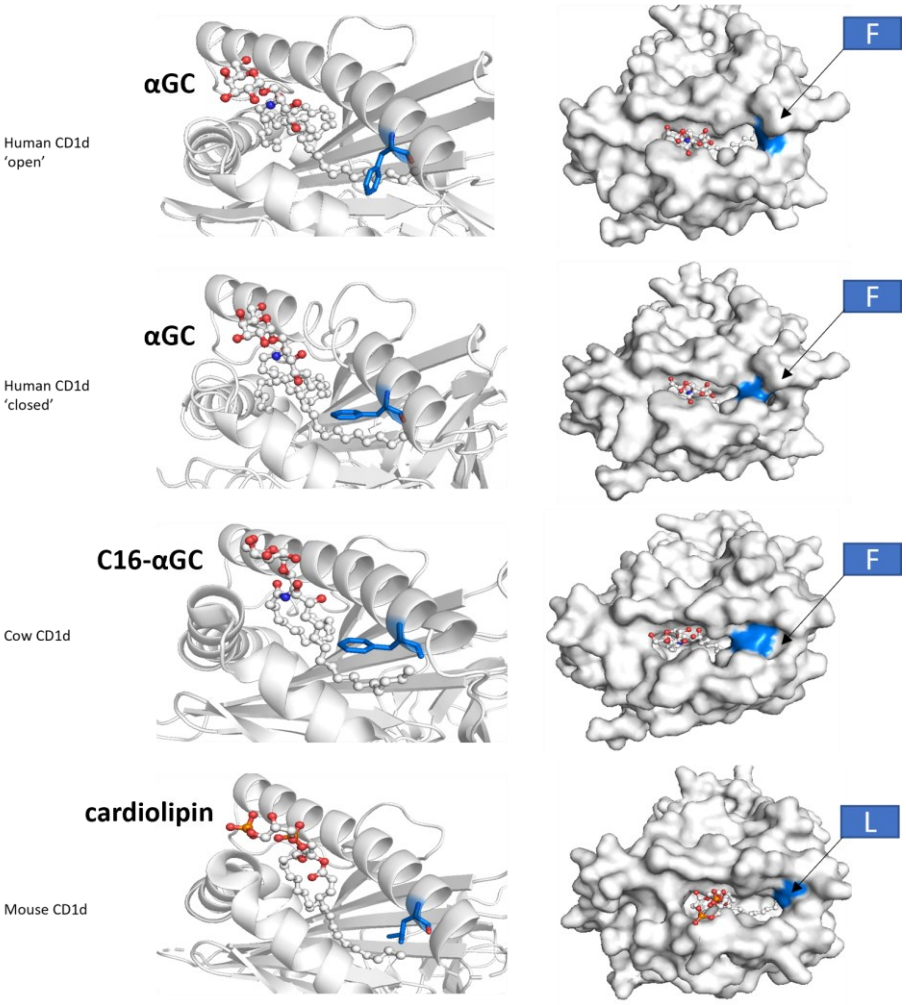


Figure 3. 20 Structural comparison of CD1d expressing species.

Cartoon and surface representation of human CD1d structures with Phe84 in the canonical and non-canonical position, Bovine CD1, and mouse CD1d. CD1d heavy chain is coloured in grey, lipids for each species are coloured by heteroatom, and phenylalanine/leucine residues are coloured in blue. All images were generated using PyMOL and PDB codes for Bovine and Mouse CD1d are 4F7E and 2AKR, respectively.

Surface representations of our human CD1d(h) structure with Phe84 in the canonical and non-canonical conformations closely resembled the bovine (phenylalanine) and mouse (leucine) structures, respectively. The smaller isoleucine created a larger binding cavity volume, exposing the lipid antigen to solvent compared to the bovine cavity, which was more enclosed by phenylalanine.

When investigating current structural data regarding CD1d-expressing species, we were surprised to find there was no non-human primate (NHP) structure of human CD1d presenting lipid antigens, either in complex with TCR or in its binary form. NHPs are considered excellent model organisms for studying

human disease due to their relatively similar sequence homology compared to other species, and similarities in immune function. Sequence analysis revealed the conservation of phenylalanine in the non-human primate (NHP) CD1d sequence. To further investigate structural similarities in CD1d antigen presentation, crystallisation of rhesus macaque CD1d was attempted. Due to extremely high sequence homology of macaque and human β2M sequences, coupled with the fact that we already had this sequence cloned into a mammalian expression vector, we decided to use human β2M and refold the inclusion bodies with macaque CD1d heavy chain.

For continuity, further mentions of our Macaque CD1d protein will be referred to as CD1d(mq). This represents our CD1d Macaque protein with lipid αGC bound within the lipid-binding groove.

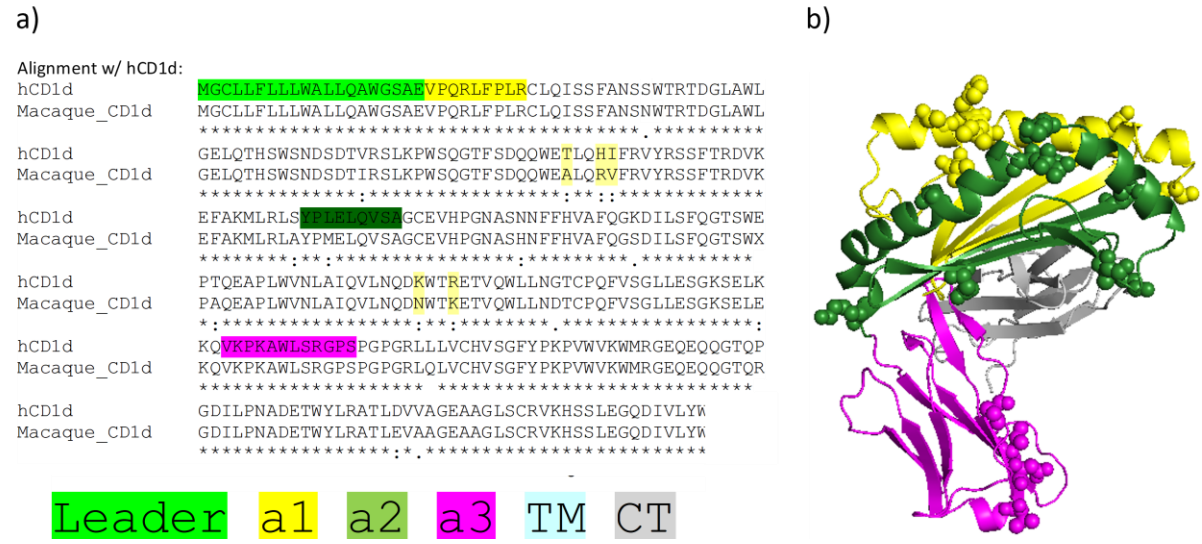


Figure 3. 21 Human and Macaque CD1d Protein Sequence Comparison

a) Sequence alignment of Human and Macaque CD1d heavy chains with conserved residues underlined with a '*', and differences underlined with ':' or '.'. b) Cartoon representation of human CD1d-αGC colour-coded for the individual heavy chain domains, with α1, α2, and α3 coloured in yellow, green, and magenta respectively. Amino acids differing between human and macaque CD1d are shown using sphere representation. The model used for the cartoon representation was our newly solved human CD1d-αGC structure, generated using PyMOL.

Sequence alignment between human and macaque CD1d heavy chains revealed that most differences were located on the left-hand side of the binding interface, away from the iTCR binding footprint and lipid-binding pocket. However, several amino acid differences were identified near the iNKT binding footprint and lipid-binding tunnels.

To further investigate these differences, soluble CD1d(mq) production was attempted, using previously determined retention times and SDS-PAGE analysis from CD1d(h) production as a reference.

3.2.2 Producing CD1d(mq) for structural and functional studies.

Due to previous work within our lab, we had access already to the Macaque CD1d sequence in a bacterial expression vector. We produced soluble inclusion bodies using the same expression media and protocols as for CD1d(h), with clean bands on our SDS-PAGE analysis at the correct molecular weight relative to the protein standard ladder (*Figure 3.22a*). CD1d(mq) is similar in length and sequence to CD1d(h), hence we expected similar data from expression and purification with regards to retention times and SDS-PAGE gels.

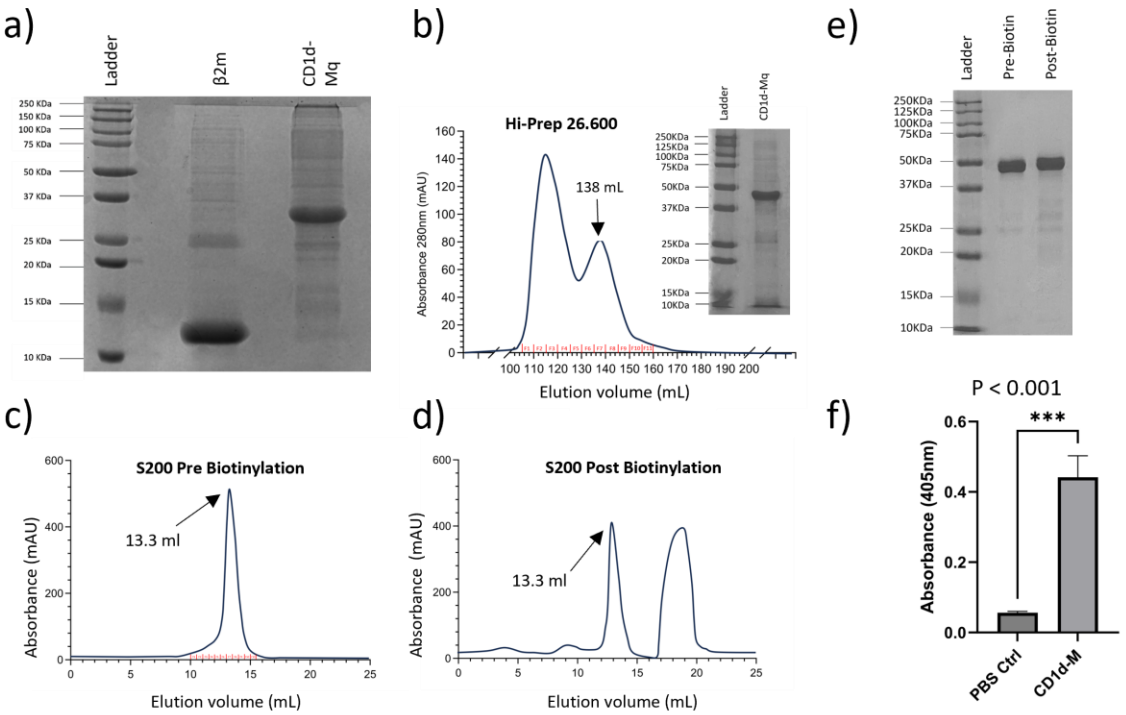


Figure 3. 22 Pipeline for Macaque CD1d-αGC protein production

a) SDS-PAGE gel showing inclusion bodies of Macaque CD1d and β2M proteins, revealing clean bands at the correct molecular weight relative to protein standards ladder (lane 1). b) Chromatogram showing refolded CD1d(mq) purification via size exclusion chromatography using a Hi-Prep 26.600 gel filtration column. SDS-PAGE analysis of the resulting fraction eight from the peak with 138ml PM shows clean protein eluted at ~50KDa corresponding to the Macaque dimer. c) and d) Chromatograms from S200 analytical size exclusion purification pre- and post-biotinylation, respectively. e) SDS-PAGE gel of fractions taken from the middle of the large peak at 13.3ml PM, pre- and post-biotinylation. f) Bar chart showing biotin ELISA results with un-biotinylated CD1d(mq) protein used as a control. PM = peak max

CD1d(mq) was refolded with αGC as the presenting lipid using the same method as CD1d(h), with concentration performed using the Amicon stirred cell and Viva-spin columns (Sartorius) before purification. The CD1d(mq) refold generated a greater yield than CD1d(h), producing ~3 mg of purified protein per 500 ml refold compared to ~2 mg for CD1d(h).

Initial purification using a 26.600 gel filtration column produced two main peaks, with the second peak corresponding to the expected retention time for refolded CD1d protein (*Figure 3.22b*). Reducing SDS-PAGE analysis confirmed the presence of clean bands at molecular weights corresponding to β2M and CD1d monomers, verifying that the peak contained the correctly refolded protein. Further purification

using an analytical S200 column yielded retention times nearly identical to human CD1d purification, with SDS-PAGE (*Figure 3.22e*) confirming the purity of the refolded protein (~50 kDa).

The purified protein was biotinylated for further functional testing, with biotin ELISA data (*Figure 3.22f*) confirming successful biotinylation relative to the unbiotinylated control.

Following successful purification and biotinylation, we assessed CD1d(mq) binding to iTCRs using tetramer staining and SPR.

3.2.3 Testing the functional validity of CD1d(mq).

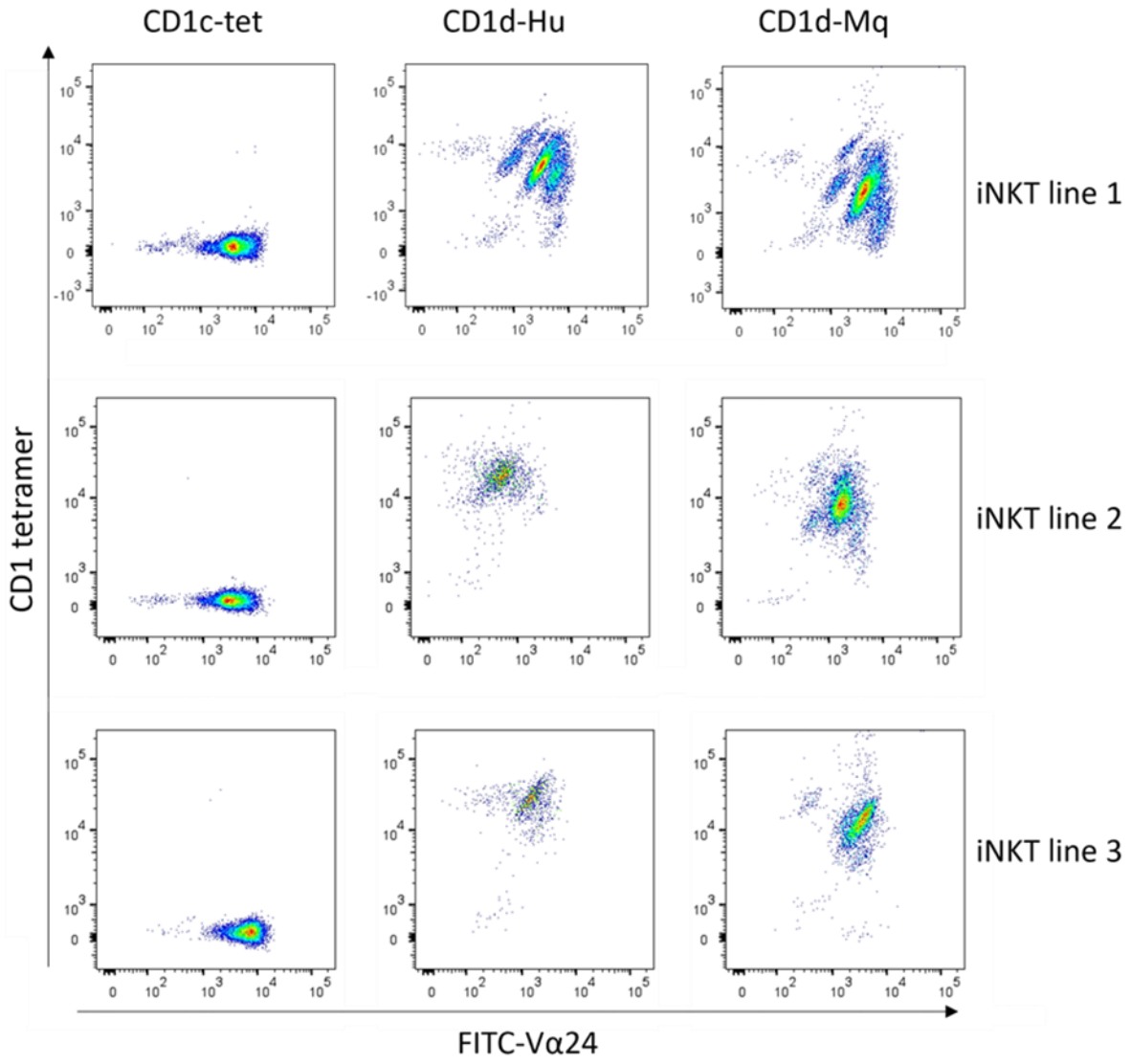


Figure 3. 23 Tetramer staining of human iNKT cell lines using different tetramers.

Representative flow cytometry dot plots showing Human and CD1d(mq) tetramer staining of three different CD1d restricted human iNKT cell lines. CD1c-endo was used as a negative control tetramer. All three iNKT lines stained well with both CD1d(h) and CD1d(mq) tetramers, however showed a slightly higher mean MFI value with the human tetramer. Both tetramers can distinguish multiple clonal populations likely representing differential affinities towards CD1d within the iNKT cell lines, with CD1d(mq) providing greater separation than its human counterpart.

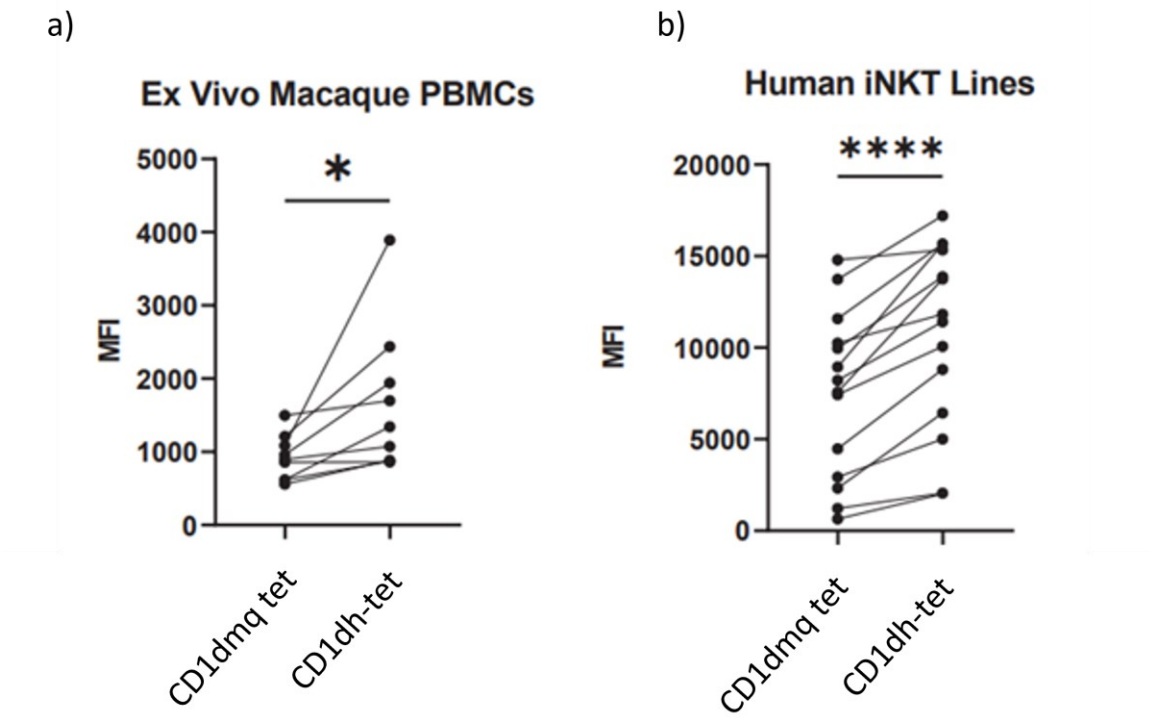


Figure 3. 24 CD1d(h) stains Human and Macaque T cells better than CD1d(mq)

a) Graph showing the mean MFI values of seven different Rhesus Macaque ex vivo PBMCs stained with CD1d(mq) and CD1d(h) tetramer. Human CD1d-αGC consistently stains with a higher MFI value. b) Tetramer staining of human iNKT lines with CD1d(h) and CD1d(mq) >10 different donors. All donors show stronger staining with the human CD1d tetramer compared to the macaque CD1d tetramer.

Figure 3.23 and *Figure 3.24* show that the CD1d(h) tetramer stained both human and macaque iNKT cells more strongly than the CD1d(mq) tetramer. To further assess binding interactions, SPR analysis was performed to determine the binding affinities of CD1d(mq) for the two human iTCRs used in the CD1d(h) SPR study (*Figure 3.8*).

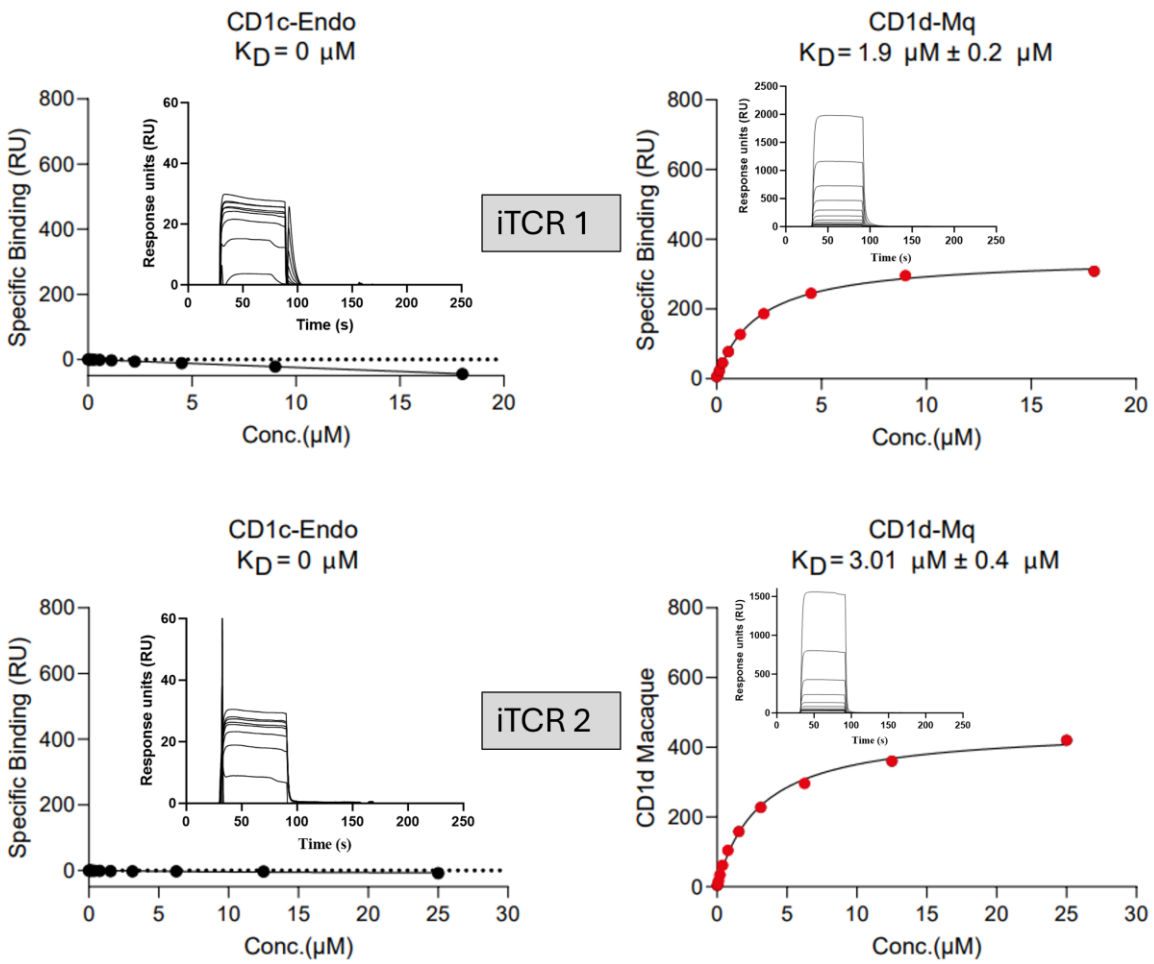


Figure 3. 25 CD1d-Mq binding to human iTCRs

SPR binding curves and calculated dissociation constants (KD_DD) for two different human iTCRs binding to macaque CD1d-αGC and the CD1c-Endo control. Biotinylated CD1 monomers were immobilised on a CM5 chip, and iTCRs were flowed over the surface at 30 µL/min in 12 dilutions. Specific binding was plotted against concentration using GraphPad Prism, and equilibrium binding constants (KD_DD) were determined by non-linear regression fitting. Inset graphs display representative sensorgrams from Biacore experiments, showing real-time response units (RU) over time, which illustrate the association and dissociation kinetics for each interaction. RU, response units. *(SPR experiments were conducted in collaboration with Immunocore.)*.

SPR results confirmed the binding of our human iTCRs to CD1d(mq) protein, confirming this protein was functional. The binding affinities also corroborated the slightly higher tetramer staining seen with the human vs macaque protein; however, this is speculative as the affinities were only slightly lower when using the CD1d(mq) protein. Following confirmation of successful binding, we proceeded to attempt the crystallisation of CD1d(mq), using the same initial screens as with our human protein. The macaque protein was concentrated to 14mg/ml, and seed stock from previous crystallisations of the

human protein was added at a ratio of 0.5:1:1 for seed stock, protein, and reservoir, respectively. Proteins were incubated at 21°c, and crystals grew after roughly one week.

3.2.4 Investigating the novel CD1d(mq) structure.

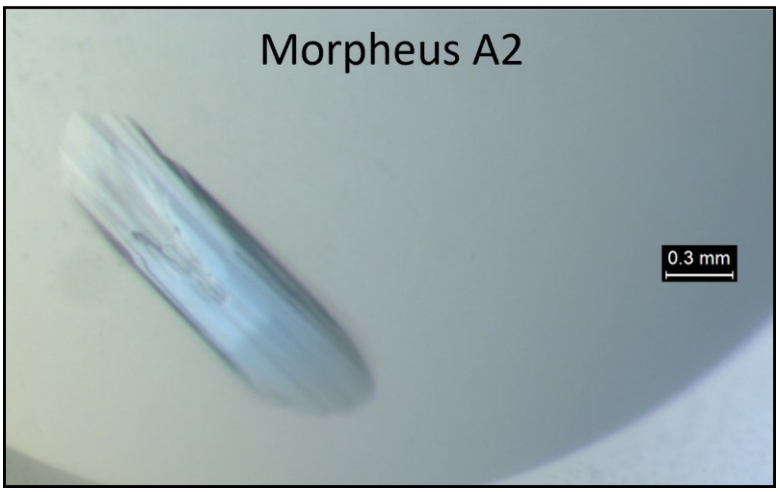


Figure 3. 26 CD1d(mq) crystal image

Microscope image of CD1d(mq) in A2 of the Morpheus crystallisation screen. The protein was crystallised using sitting drop vapour diffusion in 96-well plates and the crystal grew in ~7 days. Seed stock from previous CD1d-αGC crystals were used in combination with 14mg/ml CD1d-Mq protein and mother liquor in a ratio of 0.5:1:1 respectively. The scale bar is shown at 0.3mm for reference size (crystal measuring roughly 0.5mm in diameter and 1.5mm in length).

The resulting crystals were sent to Diamond light source for X-ray diffraction. The novel crystal structure of CD1d(mq) was determined using molecular replacement with our CD1d(h) model as the search model and refined using multiple rounds of refinement using CCP4i2 crystallographic software, with the use of data to 1.75Å.

Table 3 Data collection and refinement statistics (CD1d-Mq)

Datset	CD1d-Mq
Data Collection	
Space Group	P 31 2 1
Cell Dimensions	
a (Å)	87.08
b (Å)	87.08
c (Å)	232.79
α°	90
β°	90
γ°	120
Resolution	77.2-1.75 (1.82-1.75)
Rmerge	0.75 (0.762)
I/σI	10.9 (1.88)
Wilson B	26.9
Completeness (%)	98.2 (95.20)
Refinement	
Resolution (Å)	1.75
No. of reflections	
all/free	108023 / 5339
Rwork/Rfree	0.179 / 0.212
Ramchandran plot	
Favoured	98.9
Allowed	100%
Total no. of atoms	13536
Protein	6181
Ligand/Ion	177
Water	910
RMSD	
Bond lengths (Å)	0.0147
Bond angles (°)	1.785

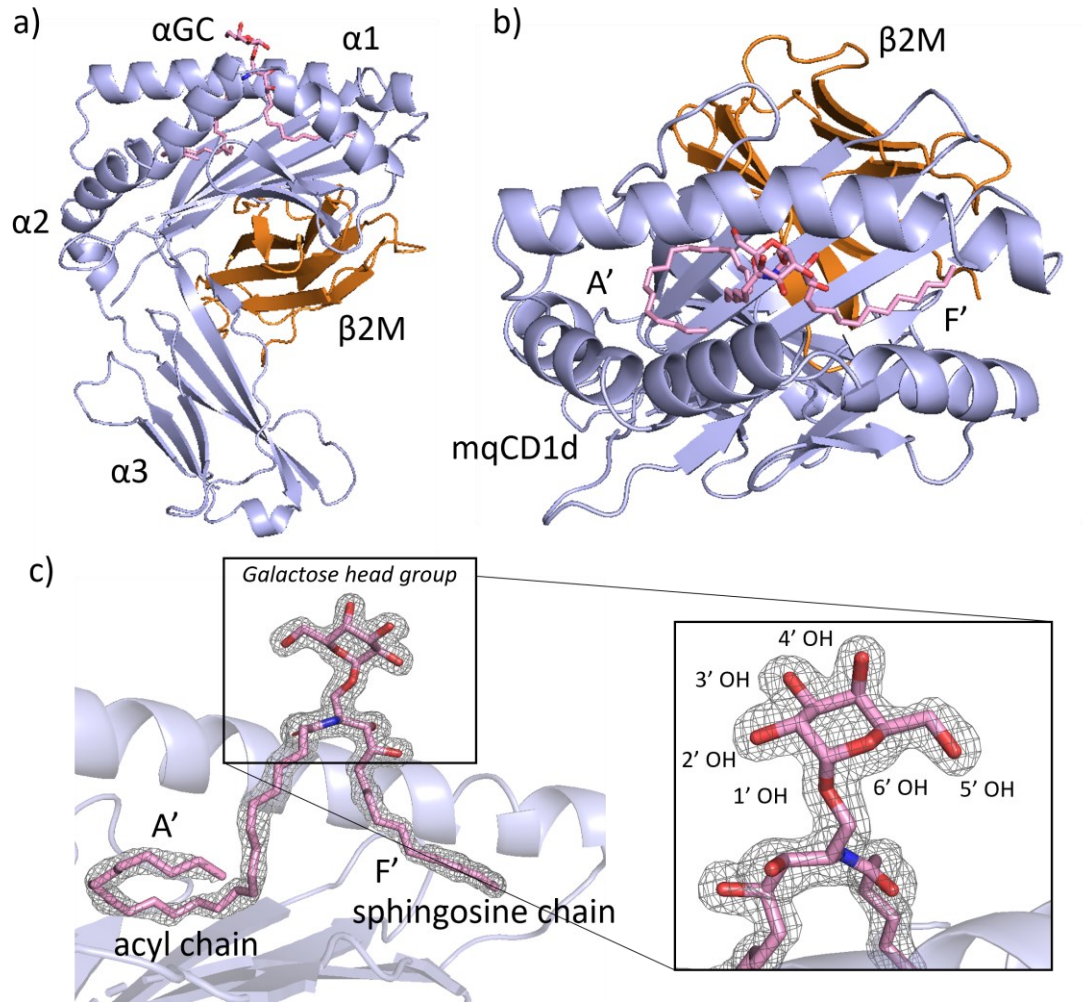


Figure 3. 27 Novel crystal structure of CD1d(mq)

Cartoon representation of CD1d(mq) protein crystallised to 1.8Å resolution, shown in two orientations. ai) Front overview of our novel CD1d(mq) crystal structure, with the lipid αGC protruding from the lipid-binding groove. b) Birds-eye view showing the lipid acyl and sphingosine chains sat in the A' and F' channels respectively. CD1d heavy chain is coloured in lilac, β2M in orange, and αGC in pink with nitrogen and oxygen in blue and red respectively. c) Electron density for lipid αGC shown over the lipid, calculated using an omit map and contoured to 3σ (density shown in grey). Image generated using PyMOL.

Our CD1d(mq) structure displays the common fold structure shared by all published CD1 models, with the 3-domain heavy chain non-covalently associated via the α3 domain to β2M. The two α-helices sit on top of a six-stranded β sheet, with the lipid head group protruding from the exit portal above the molecule for contact with TCR, and the 26C and 18C acyl and sphingosine chains sat in the A' and F' tunnels respectively. Electron density for the lipid was unambiguous, hence we carried out comparisons with other solved CD1d protein structures using Ligplus^{+}.

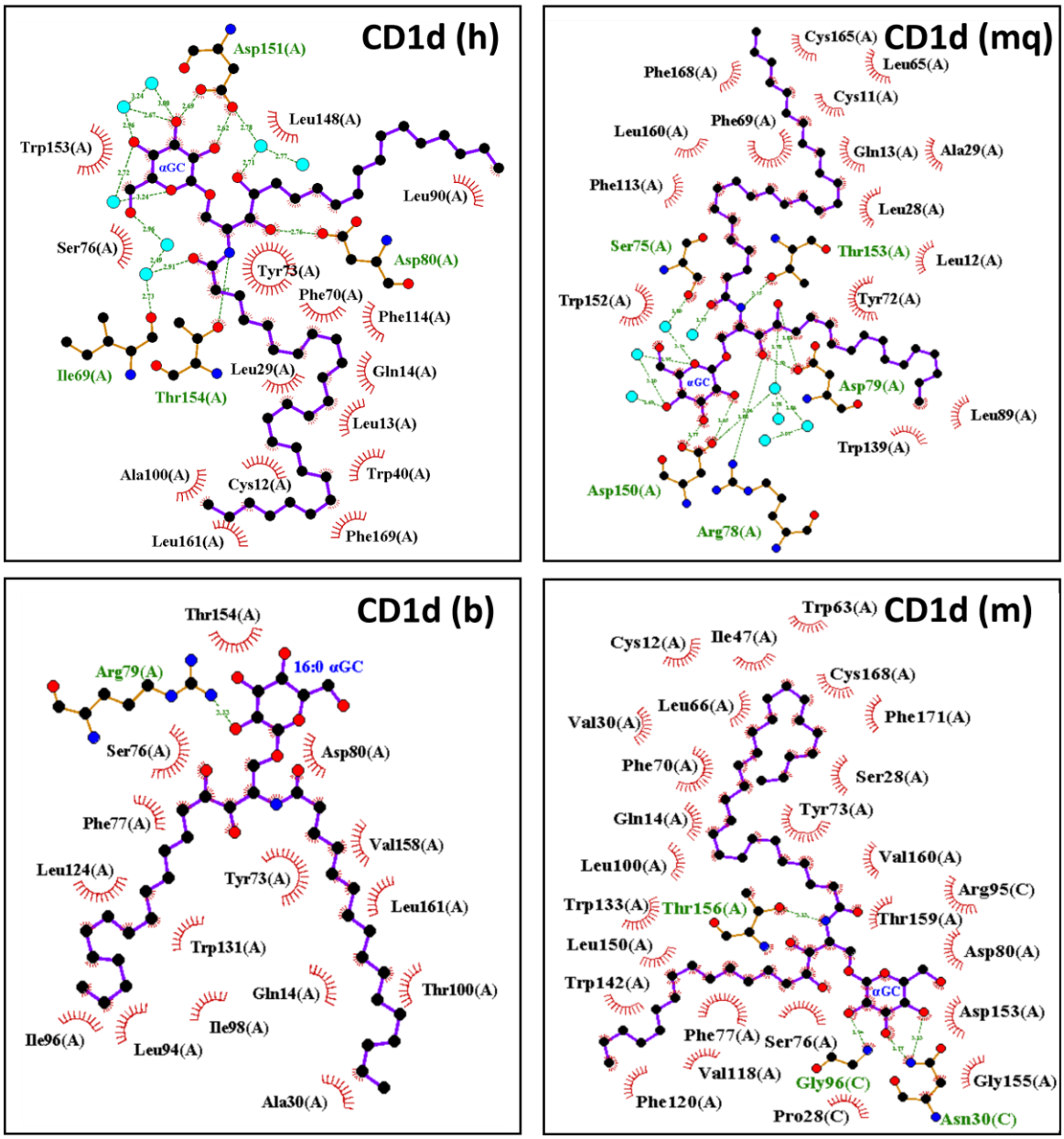


Figure 3. 28 Hydrogen network comparison of the four current CD1d structures

Ligplus+ plot images of CD1d human (h), CD1d macaque (mq), CD1d mouse (m), and CD1d Bovine (b) hydrogen bonding networks between the lipid and CD1d protein. Lipids and residues forming the hydrogen bonding network are shown in ball and stick representation. Sphingosine chain interactions are represented as red dashed lines and labelled in black. Hydrogen bonds are shown as green dashed lines and labelled to two decimal places. Lipid carbon atoms, oxygen, nitrogen, and water molecules are coloured in black, red, dark blue, and turquoise respectively.

The hydrogen bonding networks of CD1d(h), CD1d(mq), and CD1d(b) were compared. Despite conserved lipid interactions across three of the four proteins, differences in hydrogen bonding networks were observed.

In the NHP structure, Asp79 formed two hydrogen bonds with the primary and secondary oxygens of the sphingosine chain, whereas in the human structure, Asp80 formed only one bond with the secondary oxygen. Asp150 in the NHP model and Asp151 in the human model each formed two hydrogen bonds with the 5’ and 6’ OH groups of the galactose head group, while also interacting with the water network surrounding the headgroup.

Thr53 (CD1d(h)) and Thr54 (CD1d(mq)) both formed hydrogen bonds with the nitrogen at the start of the acyl chain. The NHP model displayed four amino acids making direct hydrogen bonds with the lipid (five interactions in total), whereas the human model involved three amino acids making direct hydrogen bonds (four interactions in total).

The total number of waters involved in the hydrogen bond network was seven in CD1d(h) and ten in CD1d(mq).

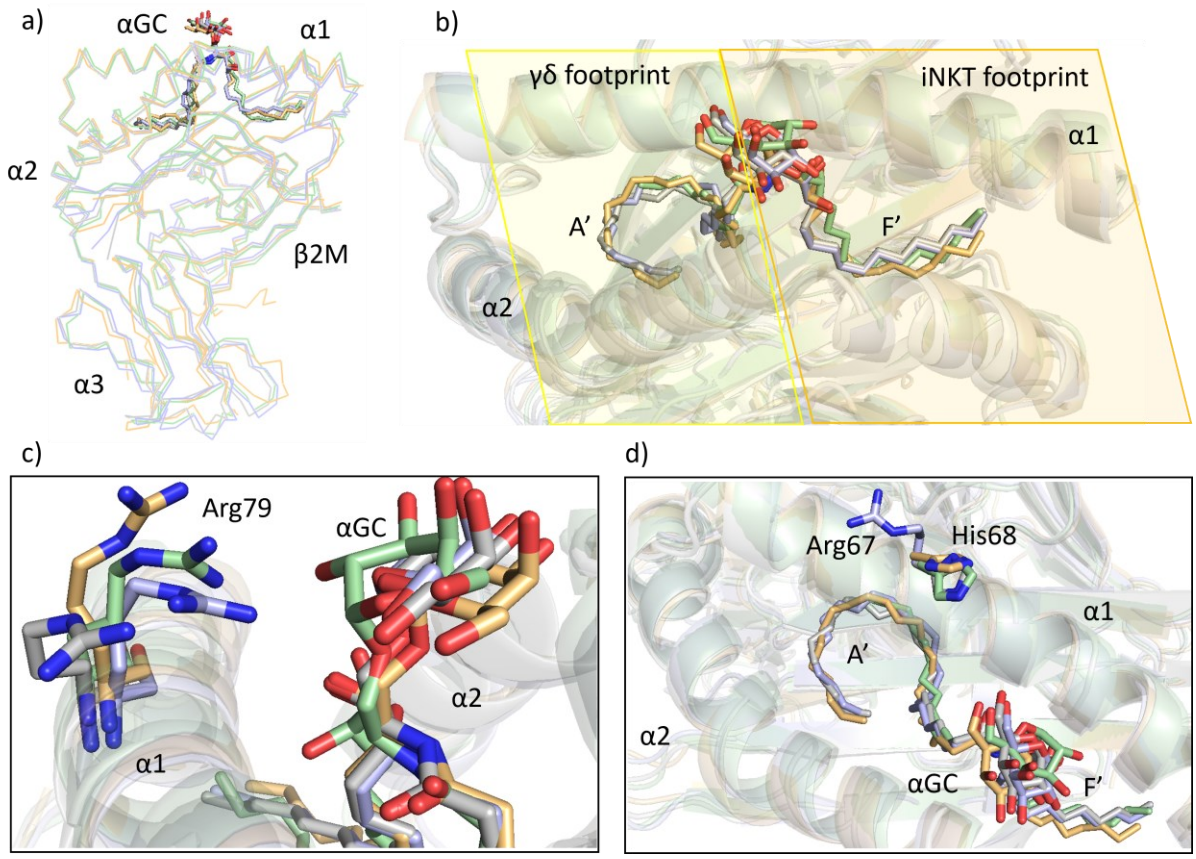


Figure 3. 29 Structural overlay of the four published CD1d structures derived from different species.

a) structural overlay in ribbon representation of CD1d(h), CD1d(mq), CD1d Bovine, and CD1d mouse. All structures contain 'wildtype' lipid αGC apart from the Bovine structure which has a 16:0 αGC. Lipid is shown in stick representation. b) Birds-eye view of the lipids bound within the CD1d groove, with sphingosine chains in the F' channel and acyl chain sin the A' channel. γδ and iNKT binding footprints are shown in pale yellow and pale orange respectively. c) Stick representation of the lipids and Arg79 residue, coloured by heteroatom with nitrogen in blue and oxygens in red. d) Birds-eye view of the lipids and the residue at position 68 (human, bovine, mouse) and 67 (macaque) showing differences in orientation. Species are coloured orange, purple, grey, and green for mice, macaques, humans, and bovines respectively.

Our structural comparison focuses on differences between the four different CD1d-expressing species that have currently been solved (including our novel Macaque structure). Initial comparisons reveal that lipid orientation in our new macaque model is highly like that of the CD1d(h) structure, but less so in mice and Bovine. We also compared the orientation of Arg79, a highly flexible residue demonstrated to play a role in both TCR binding and αGC stabilisation in different published CD1d models. The equivalent residue to Arg79 in our CD1d(mq) model (Arg78) most closely resembled that

of the Bovine, also playing a role in αGC stabilisation, unlike the human Arg79 which has an outward rotation of nearly 90°. We also looked at His68, another residue shown to be involved in TCR binding in multiple TCR-CD1d crystal structure complexes. The equivalent residue in macaque is Arg67, which is rotated 180° and points the opposite way to the His68 shown in the other three structures. Another area of focus was the Phe84 residue shown to be conserved amongst most of the CD1d-expressing species. We looked at the Fo-Fc and 2Fo-Fc maps to understand the densities around the residue and how they differed between our human and macaque structures.

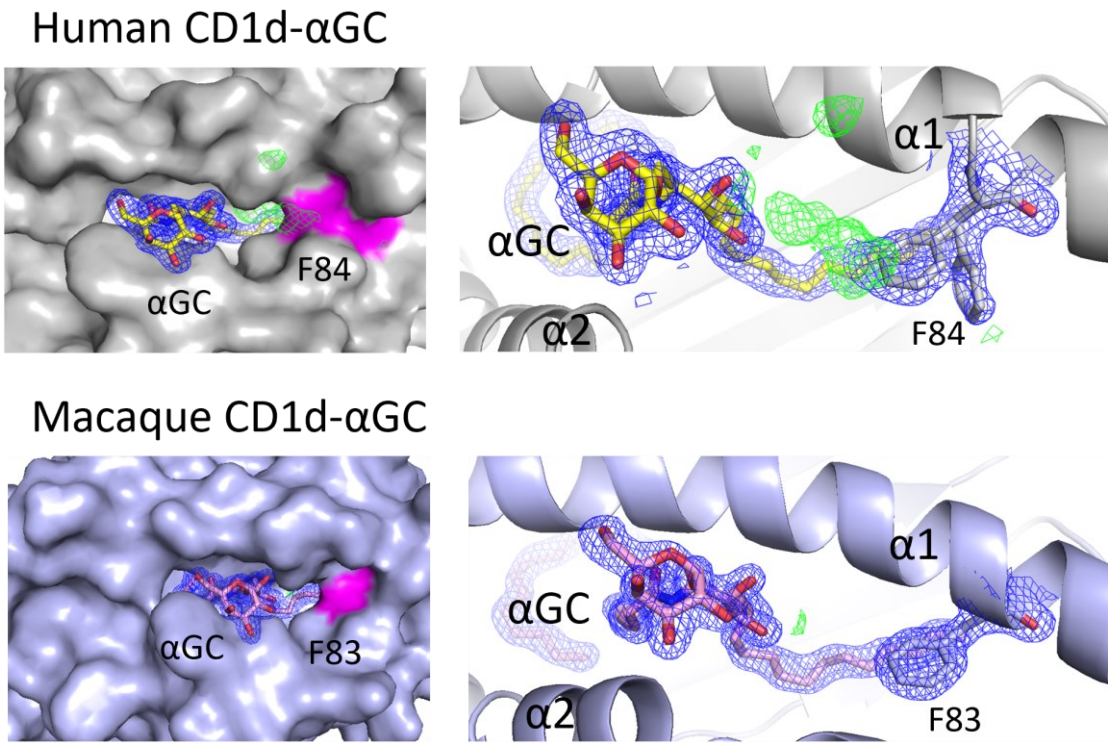


Figure 3. 30 CD1d(h) vs CD1d(mq) Phe84 and surrounding electron density

Surface and cartoon representation of the CD1d(h) and CD1d(mq) binding grooves with a focus on the Phe84 residue and surrounding electron density. The α-helices are shown in the 'cartoon' representation, with the αGC lipids shown in the stick representation and coloured by heteroatom (nitrogen in blue and oxygen in red). Human CD1d lipid is shown in yellow, and Macaque lipid is shown in pink. b) Surface (left) and cartoon (right) figures showing electron density immediately surrounding the Phe84/Phe83 and αGC lipid in both structures. 2Fo-Fc is shown in blue, and Fo-Fc is shown in green. Dual conformation of Phe84 in human CD1d is highlighted with Fo-Fc density shown surrounding both conformations. Images generated using PyMOL.

As demonstrated by *Figure 3.30*, the difference density seen in the human structure that indicates a multiple conformation, is not seen in the Macaque structure, with virtually no difference in density surrounding the residue. There is also no difference in density seen in the cavity between the Phe83 (equivalent to Phe84 in humans) and the lipid head group. Whilst this is not consistent with the CD1d(h) residue and surrounding pocket, we believe it suggests that the presence of buffer molecules to stabilise this residue in the open conformation is a necessity to visualise the flexibility.

Finally, we wanted to investigate the functional output of both proteins regarding activation and cytokine release. We investigated activation markers CD25 and CD69, as well as IL-2 release, as functional outputs following incubation of CD1d(h) and CD1d(mq) protein with iNKT cells.

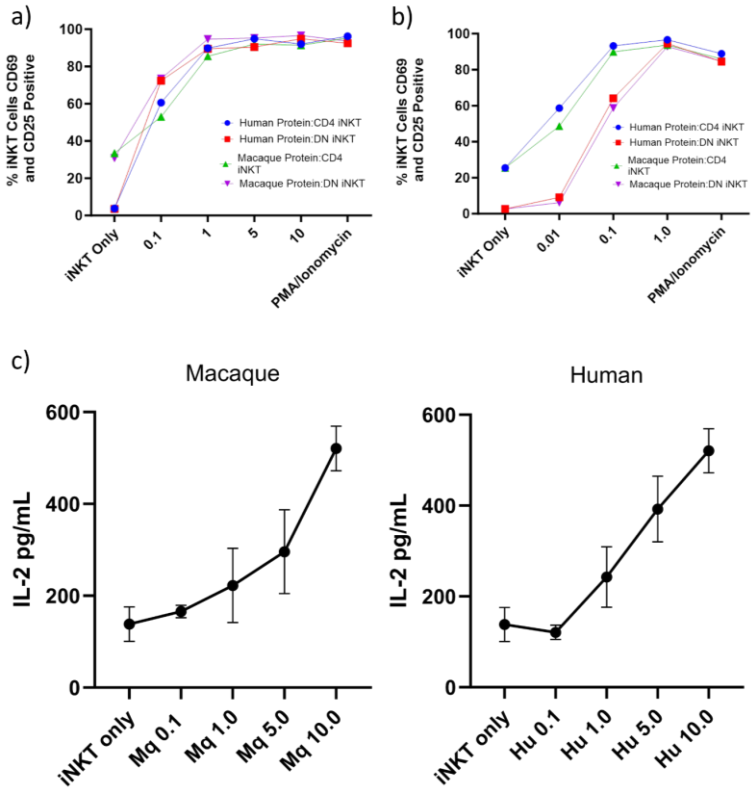


Figure 3. 31 Activation and cytokine release of iNKT cells following CD1d-αGC recognition.

a) Activation marker assay showing the proportion (%) of iNKT cells expressing CD69 and CD25 following incubation with CD1d(h) and CD1d(mq) protein. Double negative (DN) and CD4 iNKT cells were used. b) The same experiment was repeated using a range from 0.01-1µg of protein and showed the same trend. c) IL-2 ELISA showing increase in IL-2 in the supernatant post incubation with concentrations of CD1d(h) and CD1d(mq) ranging from 0.1µg to 10µg. Both proteins caused a dose-dependent increase in IL-2 concentration in the supernatant.

In a CD1d plate-bound assay, we measured the % of iNKT cells which were expressing CD69 and CD25 post-incubation with our two proteins of interest. We chose to measure the activation of both DN and CD4 iNKT cells, with initial concentrations of protein at 0.1, 1, 5, and 10µg. Despite indicating a dose-response, the initial increments were too high, with a plateau reached at 1µg of protein for both Human and Macaque CD1d. We therefore adjusted the protein dose increments to 0.01, 0.1, and 1µg. We saw a clear dose-response increase in CD69 and CD25 expression for both proteins. Interestingly, DN iNKTs appeared to require a greater concentration of protein to achieve the same % activation as the CD4 subset, suggesting that the CD4 subset was more sensitive to activation by the CD1d-αGC. We also carried out an IL-2 ELISA, which again demonstrated a dose-dependent increase in IL-2 secretion, which was comparatively similar for both proteins.

This data suggests that despite the subtle differences in hydrogen bonding network and apparent tetramer staining and SPR measurements between the two proteins, these are not enough to elicit a divergent functional response.

3.3 Discussion

3.3.1 Investigating CD1 isoform conservation and antigen presentation

Since the discovery of the CD1 gene family in 1986, the repertoire of lipid antigens has expanded significantly (176). Structural studies of the four antigen-presenting isoforms (CD1a–CD1d) have uncovered a distinct network of interconnecting tunnels, which partially accounts for isoform-specific lipid preferences (49, 66, 72, 175). While cell surface expression, recycling pathways, and scaffold molecules also contribute to lipid selection, the structural architecture of the binding groove remains a key determinant in defining lipid-binding specificity (17, 177, 178).

Despite these advances, less is known about the extent of conformational flexibility within a single CD1 isoform to accommodate lipids of varying chemical properties and structures. Co-crystal complex data has provided insights into how CD1 and TCRs structurally adapt to facilitate TCR binding however the intricate mechanisms underlying lipid selection and presentation remain unclear (52, 100, 179, 180).

Previous Investigations into CD1c structural plasticity have demonstrated conformational flexibility within the F' roof (65). The high variability in lipid structural and chemical properties presented by the same CD1 isoform suggests there is likely a significant degree of plasticity within the binding groove to accommodate for this variability (181, 182). It is also likely, due to the high degree of structural and sequence conservation amongst the antigen-presenting CD1 isoforms, that the fundamental mechanisms utilised by the CD1 isoforms to accommodate lipids of varying properties are similar (183). Sequence analysis identified an aromatic residue adjacent to a conserved alanine in CD1c, suggesting a shared mechanism among antigen-presenting CD1 isoforms that enables dynamic groove modulation in response to lipid occupancy *(Figure 3.1).* Given the high conservation of an aromatic residue at this position across multiple isoforms (*Figure 3.2*), it is likely that a similar structural adaptation mechanism exists within CD1d, enabling the accommodation of a diverse range of lipid antigens through conformational adaptation. The ability of this aromatic residue to dynamically constrict or expand the binding groove may be fundamental in shaping CD1d's antigen repertoire and influencing T cell responses.

This study set out to investigate whether CD1d shares the same structural adaptability seen in other CD1 isoforms, focusing on the conserved aromatic residue at the F' portal. By examining sequence and

structural conservation, we aimed to determine whether CD1d uses a similar dynamic mechanism to CD1c, potentially broadening its ability to present a diverse range of lipid antigens.

3.3.2 Generating Soluble CD1d-αGC for Structural and Functional Studies

To investigate the structural and functional properties of CD1d antigen presentation, we aimed to generate stable, soluble CD1d monomers refolded with the potent iNKT agonist lipid αGC. Due to its high stability and well-characterised TCR recognition profile, αGC is an ideal model antigen for studying CD1d-lipid-TCR interactions (72, 114, 184).

Optimising the bacterial expression system was crucial to obtaining sufficient protein yields for both crystallographic and functional assays. Several expression conditions were tested, ultimately leading to an approach that provided higher inclusion body yields, ensuring consistent refolding efficiency (*Figure 3.3*). Given that inclusion body purity correlates with successful refolding, these optimisations improved protein stability and lipid loading consistency (185, 186).

The observed binding affinities of our in-house TCRS for our refolded CD1d-αGC (1-2µM), align with previously reported values for human iNKT-CD1d interactions, reinforcing the stability of our refolded complex and supporting the physiological relevance of our structural data *(Figure 3.8*). This conservation of affinity across experiments also indicates that lipid presentation is relatively stable before TCR engagement and doesn't require much structural rearrangement upon TCR binding. The stability of αGC within the CD1d groove is well-documented, and its potency as an iNKT cell agonist has been attributed to its ability to bind CD1d with high stability, independent of TCR affinity (187). This suggests that lipid properties, rather than TCR sequence variation alone, may play a significant role in dictating antigen recognition dynamics.

By generating soluble, stable CD1d-αGC complexes, we established a robust platform for investigating the molecular mechanisms underlying CD1d-mediated antigen presentation. This provided the foundation for high-resolution structural analysis and subsequent functional studies to elucidate lipid-dependent conformational flexibility within the CD1d binding groove.

3.3.3 Crystallisation of CD1d-αGC to High Resolution

To further investigate the structural properties of CD1d-αGC, crystallisation trials were performed to obtain high-resolution structural data *(Figure 3.10, Table 2)*. Comparison with previously published CD1d-αGC structures confirmed that the overall fold of CD1d remained highly conserved, with αGC adopting a typical binding orientation within the antigen-binding groove (72, 188, 189). The structural conservation reinforces the fact that CD1 proteins are evolutionarily conserved and optimised for stable TCR interactions, and hence subtle adjustments rather than significant structural rearrangements dictate specificity for different lipids.

Despite the overall conservation of the CD1d fold, the increased resolution provided novel insights into how specific side-chain reorientations might fine-tune lipid anchoring and TCR recognition *(Figure 3.11c).* These structural refinements suggest a degree of plasticity in CD1d that may have functional implications for antigen presentation, particularly in shaping TCR specificity. The observed flexibility of Trp160 aligns with its established role in γδ TCR interactions and may contribute to a broader range of TCR recognition by CD1d. This suggests a potential mechanism by which CD1d adapts to diverse TCR footprints without requiring significant structural rearrangement (168). Similarly, the slight shift of Trp153 towards the binding groove may indicate an inherent flexibility that optimises antigen presentation while maintaining stable lipid-TCR interactions. Trp153 is also involved in iNKT binding, by packing against the galactosyl ring and stabilising the lipid head group within the CD1d binding cleft (190). The increased resolution in our structure and discovery of greater sidechain flexibility in CD1d suggest that certain residues are inherently dynamic, potentially influencing TCR binding affinities and recognition in a sequence-dependent manner. This flexibility is also likely to impact lipid complex binding within the groove, a phenomenon explored throughout this thesis.

Analysis of hydrogen bonding networks further highlighted potential stabilising interactions between αGC and CD1d. While key residues remained consistent with previously solved structures, additional water-mediated interactions were identified, which may contribute to lipid stability within the groove (*Figure 3.12a)*. This is particularly relevant in the context of TCR binding affinity modulation, where minor adjustments in hydrogen bonding networks could influence the duration of antigen presentation to T cells and subsequently change TCR recognition thresholds. One example of this is a study by Wun et.al which investigates fine TCR specificity for CD1d antigen complexes (191). The extra hydrogen bond by Asp151 to the 3'OH made clear by our high-resolution structure likely contributes to the higher tolerance of human iNKT binding to CD1d-αGC when the 4'OH of the galactose headgroup is modified, compared to the 3'OH, as the latter would impact head-group stability and subsequent recognition by the iTCRs (106, 190).

One of the most striking findings was evidence of conformational flexibility at Phe84 *(Figure 3.13),* a residue previously implicated in CD1b and CD1c lipid accommodation (65, 175). The presence of dual conformations within the electron density maps corresponding to different phenylalanine rotamers suggests that Phe84May play a role in modulating the size of the binding groove, therefore providing an adaptive mechanism that enables CD1d to present a more diverse range of lipid antigens, enhancing its role in modulating iNKT cell activation. Partial occupancy was observed, with neither rotamer fully occupying the observed density. Whilst this indicates not all monomers within the lattice displayed the Phe84 in the same orientation, it points to a mechanism by which the residue can adapt depending on the occupancy of molecules within the groove. This observation aligns with studies demonstrating bi-conformational flexibility in CD1c, reinforcing that CD1d may employ a similar adaptive mechanism (71). This conformational plasticity phenomenon has also been shown in fatty acid-binding proteins and human immunodeficiency virus (HIV) reverse transcriptase (192, 193). Overlay of our structure with other CD1 isoforms revealed an upward shift in the α1 helix of CD1c compared to CD1d, suggesting that flexibility extends beyond side-chain rotation to include helical displacement. A subtle shift in CD1d's α1 helix (relative to 1ZT4) hints at a potential adaptation to accommodate additional molecules, though this deviation is less pronounced than in CD1c, likely due to strong αGC-CD1d interactions. Computational studies in 2020 by Cuevas-Zuviria further support that CD1d helices remain stable with high-affinity ligands, whereas weaker lipid interactions, as seen in CD1c, may induce greater conformational changes (194). Collectively, the structural adaptability of CD1b, CD1c, and CD1d suggests a shared mechanism of groove modulation to accommodate lipids of varying sizes *(Figure 3.14, Figure 3.15)*.

Investigation into the functional relevance of this flexibility led to the discovery of prior studies indicating hydrophobic spacer molecules can occupy pockets within CD1 molecules, thus altering antigen presentation via induction of residue flexibility like that seen in our new structure (175). The positive difference density observed in *(Figure 3.13b)* suggests the presence of a stabilising molecule, like findings reported by Gadola et.al (2002). Another example of this is Wun et al. (2011), where multiple iNKT-CD1d ternary complexes with different lipids were solved (106). In one CD1d(m)-OCH structure (PDB: 3ARB), the shortened sphingosine chain left the F' tunnel partially empty, which was compensated for by an 8-carbon spacer lipid. This reoriented Leu84 away from the groove cavity, expanding the binding surface.

A related ternary complex of CD1d-OCH with an iTCR showed Leu84 adopting a more closed conformation, as TCR-induced lipid repositioning displaced the spacer lipid (190). Overlaying our CD1d(h)-αGC structure with these models revealed a similar conformational mechanism between Phe84 in CD1d(h) and Leu84 in CD1d(m). While CD1d(h)-αGC fully occupies the F' channel, preventing

additional hydrocarbon binding, Phe84's outward rotation could accommodate smaller stabilising molecules. This aligns with CD1b, CD1c, and CD1d(m), where structural flexibility expands groove volume to facilitate diverse lipid presentation.

We were unable to confidently assign a single molecule to fully satisfy the difference density, indicating partial occupancy within the crystal lattice. A recent mass spectrometry study (Huang et al., 2023) identified triacylglycerols (TAGs) as frequent CD1d ligands, despite their large hydrocarbon chains (183). While the mechanism of TAG presentation remains unclear, we propose that Phe84 rotation may expand the CD1d binding cavity, facilitating the accommodation of bulkier lipids. This is consistent with CD1c, where helix shifts facilitate lipid stacking. Interestingly, CD1d exhibited the highest chain length capture despite having a relatively small binding groove, again suggesting there is likely to be structural rearrangement upon ligand binding to accommodate the larger molecules.

To explore this, we performed molecular dynamics simulations on Phe84 flexibility in the absence of a stabilising molecule. As expected, Phe84 strongly favoured the canonical conformation, reinforcing the hypothesis that ligand or buffer molecules are required to stabilise its open state. Extensive docking experiments failed to fully satisfy the observed electron density, likely due to partial occupancy by a mix of buffer molecules. As a proof of principle, we modelled a short hydrocarbon chain, confirming that Phe84 remains stable in the open conformation when stabilised by a bound molecule. However, attempts to model a full TAG molecule were unsuccessful due to steric clashes with αGC, highlighting the need for further structural studies to determine how CD1d accommodates larger lipids. Future work involving mass spec analysis on crystallised protein samples or soaking experiments involving the utilisation of defined lipid candidates could help resolve the identity of the stabilising molecule and clarify its role in groove modulation.

These findings provided a structural basis for understanding how CD1d accommodates lipid diversity, supporting the hypothesis that antigen presentation by CD1d may involve dynamic conformational adjustments. Further comparative analysis was therefore conducted across CD1 isoforms and species, to determine whether this structural flexibility is a conserved feature of CD1d antigen presentation.

3.3.4 Structural and sequence comparisons across CD1d-expressing species

CD1d antigen presentation differs across species, but the extent to which structural conservation dictates functional similarities remains unclear. Sequence conservation is often indicative of functional importance for a given molecule (195, 196). Given this, we ran sequence analysis across multiple

CD1d-expressing species to investigate whether specific residues or areas of the protein were highly conserved, paying particular interest to the Phe84 residue.

The strong conservation of Phe84 amongst the higher species *(Figure 3.19)* suggests its aromatic ring structure plays a functionally significant role in modulating CD1d groove dynamics. The Leucine/Isoleucine in rodents contrast this, potentially reflecting an alternative role in antigen presentation. This alternative role is likely dictated by the lack of aromatic ring structure physically reducing the surface area of the residue and the subsequent effect on groove size modulation. This raises the question as to whether rodents are less reliant on a flexible groove and instead maintain a more open conformation, despite still demonstrating some flexibility depending on lipid occupancy.

Structural comparisons further highlight how these residue differences could translate to antigen presentation (*Figure 3.20*). The canonical conformation of Phe84 in human CD1d mirrors bovine CD1d, where the groove remains relatively constrained. However, our structural data demonstrates that Phe84 reorientation increases groove volume, resembling mouse CD1d, where a Leu84 substitution naturally increases groove volume. This may indicate that human and macaque CD1d regulate antigen presentation more dynamically, altering groove size in response to lipid occupancy.

These findings raise interesting questions regarding the impact of structural differences on TCR recognition across species. If higher species regulate antigen presentation dynamically, while rodents rely on a constitutively wider groove, this could suggest an evolutionary divergence in TCR recognition strategies. This is particularly relevant in the context of CD1d-mediated immune responses in disease models, where subtle variations in lipid accommodation may influence cross-species immunogenicity. Despite the CD1d groove in rodents appearing to have less capability to open and close to the same extent as in humans, there have been studies demonstrating conformational changes in rodent CD1d following binding of weak-affinity lipids (197). Rodents may also compensate for reduced groove flexibility by exhibiting different iTCR usage patterns, therefore shaping iTCR response differently to humans. Further work is needed to establish whether Phe84-dependent groove modulation directly affects lipid selectivity, TCR binding affinity, or subsequent immune activation. Undertaking molecular dynamics simulations of mouse and human CD1d loaded with a diverse range of lipid antigens would help to uncover the differences in groove flexibility and how either species is affected by differing lipid cargo.

Given the structural similarities between human and macaque CD1d, we aimed to determine whether subtle sequence variations influence antigen presentation. Investigating CD1d(mq) offered a valuable opportunity to assess the functional consequences of minor structural differences across species.

3.3.5 Novel CD1d(mq) Crystal Structure and Its Functional Implications

Non-human primates (NHPs) are widely regarded as valuable models for human immune research due to their high genetic and structural homology to humans. The close evolutionary relationship between the two meant determining the crystal structure of rhesus macaque CD1d (CD1d(mq)) in complex with αGC allowed us to investigate potential species-specific adaptations in antigen presentation.

Our structural analysis reiterates the strong conservation of CD1d architecture between humans and macaques *(Figure 3.27),* suggesting a shared functional role in lipid presentation. Despite this, subtle differences in hydrogen bonding networks and sidechain orientations indicate potential variations in lipid stability and TCR recognition, which could affect cross-species antigen presentation. These findings raise interesting questions about how minor structural differences influence immune responses in disease models, particularly in contexts where CD1d-restricted T cell activation plays a role in infection or cancer immunotherapy.

Notably, Arg78 in CD1d(mq) adopted a different conformation compared to its human counterpart (Arg79 in CD1d(h)), bringing it into closer proximity with the sphingosine chain of αGC. Given the flexible nature of this residue, this may influence lipid positioning within the groove, potentially altering ligand stability and recognition dynamics. Similar flexibility in Arg79 has been observed in bovine CD1d, where it plays a role in stabilising different lipid headgroup orientations, suggesting that this adaptability may be a conserved mechanism across species (188).

We conducted functional assays to investigate key differences in antigen presentation between the species *(Figure 3.23, Figure 3.24*). Despite strong sequence homology, the observed differences in CD1d(mq) tetramer staining and SPR binding affinities suggest that even minor alterations in the lipid groove environment can fine-tune iTCR interactions. This may indicate that macaque CD1d accommodates lipids slightly differently from its human counterpart, potentially influencing the repertoire of antigens it presents and its recognition by cross-species TCRs. The improved separation of different iTCR affinity subsets by CD1d(mq) tetramers mirrors the behaviour observed with weaker binding lipids like OCH, suggesting that the macaque CD1d groove may stabilise certain lipid conformations differently from its human counterpart (114). This raises the possibility that CD1d(mq) could be leveraged as a tool for isolating human iNKT subsets with distinct functional profiles, providing a valuable alternative to lipid variants with altered stability.

While minor, these differences align with previous findings demonstrating that lipid stability within the CD1d groove can modulate TCR engagement (225).

Taken together, these findings highlight the conserved nature of CD1d antigen presentation across species, while also revealing subtle adaptations that may fine-tune lipid recognition and TCR engagement *(Figure 3.30).* Understanding these differences provides a structural framework for cross-species comparisons of CD1d-mediated immunity, with potential implications for vaccine development and immunotherapy approaches targeting CD1d-restricted T cells.

3.3.6 Conclusion

This study reveals a conserved aromatic residue at position 84 in CD1d, first identified in CD1c-SL, and demonstrates its substantial flexibility within CD1d. The ability of Phe84 to conformationally adapt suggests a mechanism by which CD1d may modulate its lipid-binding cavity, allowing for structural adaptation in response to lipid occupancy.

The observed electron density within the expanded groove cavity in the non-canonical conformation of Phe84 suggests that stabilising molecules, potentially buffer components, may influence its conformational state. This suggests a previously unrecognised mechanism by which CD1d dynamically adjusts its binding groove volume, similar to other antigen-presenting CD1 molecules. Given the diversity of lipids presented by CD1d, including glycosphingolipids, phospholipids, and triacylglycerols, such structural plasticity may be essential for optimising antigen presentation and facilitating broad iNKT recognition.

Sequence conservation analysis across CD1d-expressing species reinforces the functional significance of this region. The aromatic residue at position 84 is conserved in higher species (humans, macaques, bovines), while rodents exhibit a leucine or isoleucine substitution, likely impacting binding groove adaptability. This suggests an evolutionary divergence in antigen presentation strategies, where higher species employ a dynamic groove adjustment, whereas rodents may rely on a constitutively wider groove.

The structural and functional comparison of human and macaque CD1d further supports the conserved role of antigen presentation within primates, despite subtle differences in lipid-TCR interactions. The lower affinity of human iTCRs for CD1d(mq), coupled with its ability to distinguish between different TCR affinities, highlights potential species-specific adaptations in antigen recognition. This reinforces the need for further investigation into the impact of lipid diversity and groove flexibility on antigen presentation across species.

Overall, this study provides new insights into the structural flexibility of CD1d, demonstrating that dynamic conformational changes within the binding groove may facilitate lipid accommodation and TCR recognition. These findings expand our understanding of CD1d-mediated immunity and highlight potential avenues for further research into antigen presentation mechanisms and immunotherapeutic applications.

Chapter 4 Investigating the interaction between high and low affinity iTCRs and the CD1d-lipid complex.

Invariant natural killer T (iNKT) cells are a subset of unconventional T cells that recognise lipid antigens presented by CD1d molecules (198). These cells play a pivotal role in bridging innate and adaptive immunity and have shown promise in cancer immunotherapy due to their cytokine release profiles and tumour surveillance capabilities (199, 200). However, despite promising preclinical findings, the translation of iNKT-based therapies to human clinical trials has been met with challenges, yielding inconsistent therapeutic outcomes in humans (201, 202). One contributing factor to this discrepancy is the variation in iNKT cell frequency, phenotype, and function between mice and humans, which may influence therapeutic efficacy (203).

A key feature of human iNKT cells is their semi-invariant TCR repertoire, which is largely conserved except for variability within the hypervariable complementarity-determining region 3 beta (CDR3β) loop (204). This loop is believed to play a critical role in modulating iTCR affinity for CD1d-lipid complexes, subsequently influencing the iNKT function. While the molecular basis of murine iNKT TCR interactions has been extensively characterised, the structural mechanisms underpinning human iTCR affinity for CD1d-lipid antigens remain poorly understood (205). Specifically, it is unclear to what extent CDR3β loop composition, length, and flexibility dictate TCR binding affinity, which may have implications for targeted therapeutics involving iNKT cells.

This chapter aims to address this knowledge gap by investigating the structural and functional differences between high- and low-affinity human iTCRs, focusing particularly on the role of the CDR3β loop in governing these interactions. To investigate these mechanisms, co-crystallisation studies were attempted to allow structural comparison between high and low-affinity iTCRs bound to CD1d-lipid complexes. However, due to experimental challenges preventing the successful generation of diffraction-quality co-crystal structures, AlphaFold modelling was employed as an alternative approach to analysing CDR3β loop structures across a range of iTCRs. Although computational modelling cannot directly replicate the accuracy of co-crystal diffraction data, they provide valuable insight into potential structural differences between different iTCRs, and how this translates to affinity. By combining structural predictions with functional binding data, this work aims to gain insight into how specific sequence motifs and structural features of the iTCR CDR3β loop contribute to the underlying binding mechanism and subsequent affinity seen for different iTCRS.

4.1 Sequence analysis of high- and low-affinity iTCRs

Previous studies within our lab investigating iTCRs with diverse affinities for human CD1d provided access to DNA plasmids encoding high- and low-affinity iTCRs. For this study, we selected one high-affinity and one low-affinity TCR, originally described by Matulis et al. (2010) as 1369 (high affinity) and DN25 (low affinity), respectively (114). These TCRs were used as templates for protein production, refolding, and purification to facilitate downstream structural studies.

Before refolding, we validated the TCR sequences within the plasmid vectors corresponding to the high-affinity (1369) and low-affinity (DN25) β chains, as well as the invariant Vα24 chain.

Figure 4. 1 iTCR sequence analysis

Annotated sequence alignment of the Vα24, DN25, and 1369 TCR chains. CDR1, CDR2, and CDR3 loops are highlighted in yellow, with individual residues colour-coded in red, green, and purple, respectively. The DN25 TCR contains an Avi-tag for biotinylation at the C-terminal end. The 1369 sequence is slightly longer than DN25, due to variability in the CDR3β loop and joining region.

Following sequence confirmation *(Figure 4.1)*, we proceeded to generate soluble iTCRs using chaotropic-based refolding, as described previously (114). Initial attempts to produce soluble inclusion bodies in *E. coli* for all three TCR chains using TYP media, with resulting inclusion bodies analysed via SDS-PAGE gel electrophoresis *(Figure 4.2).* We also attempted expression of inclusion bodies using

LSLB and AIM media, however corresponding SDS-PAGE analysis showed either low or dirty expression, for LSLB *(Figure 4.2*) and AIM (data not shown) respectively. We therefore decided to use TYP media in our expression system.

4.2 Refolding high and low-affinity iTCRs for structural studies

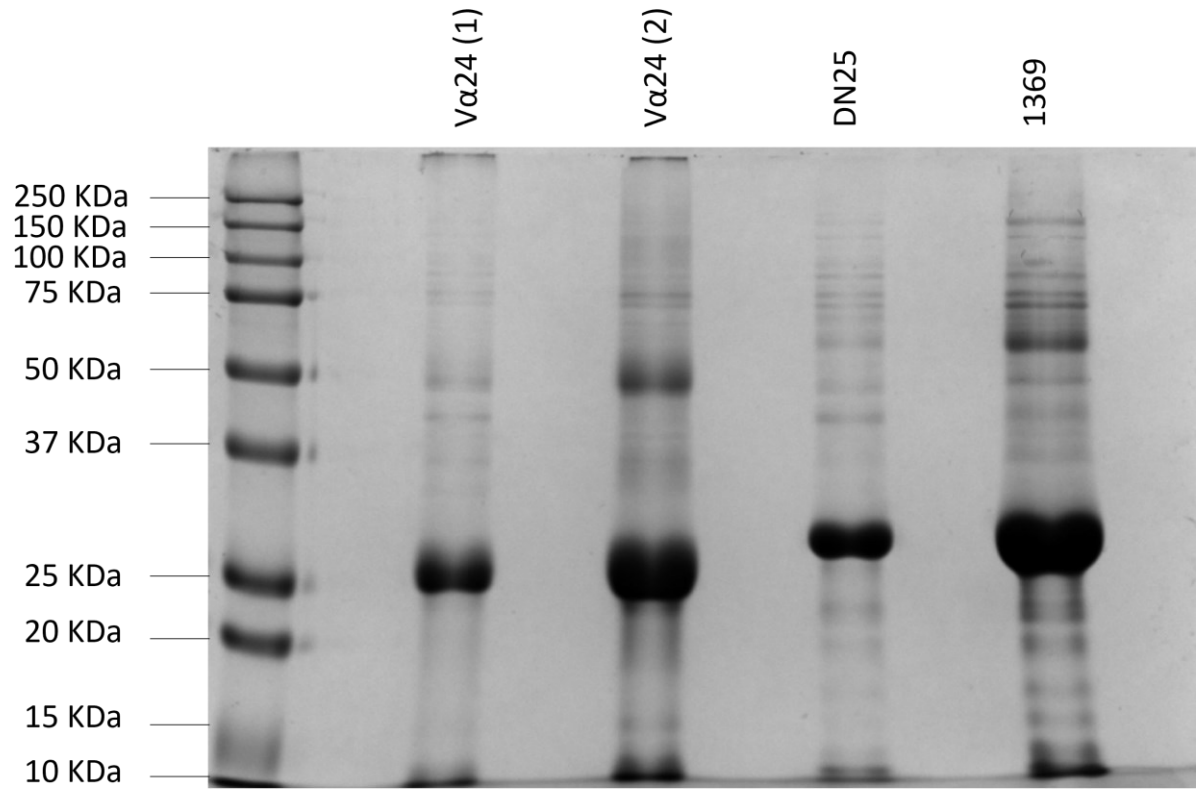


Figure 4. 2 SDS-PAGE Gel Analysis of Inclusion Bodies

SDS-PAGE analysis of inclusion bodies for the Vα24, DN25, and 1369 iNKT chains was performed to assess purity and yield following bacterial expression. Vα24 inclusion bodies were generated using LSLB (lane 1) and TYP (lane 2) media for comparison. All four inclusion bodies exhibited distinct bands, with the Vα24 α-chain appearing at the expected lower molecular weight (~25 kDa) compared to the β chains, DN25 and 1369 (~29 kDa). A standard protein ladder (Bio-Rad) was loaded in lane 1 for molecular weight comparison.

Large yields of inclusion bodies were generated, with SDS-PAGE analysis confirming distinct bands at the expected molecular weights (~24 kDa for the α chain and ~29 kDa for the β chain) *(Figure 4.2*).

However, initial refolding attempts of the Vα24 and Vβ11 chains (either DN25 or 1369) resulted in poor yields of refolded heterodimer.

To improve refolding efficiency and obtain higher yields of soluble iTCRs for structural studies, the refolding protocol was optimised. The dialysis buffer appeared cloudy following refolding, suggesting the presence of insoluble aggregates. To address this, we tested different α:β chain ratios, using 1:1, 2:1, 4:1, and 8:1 ratios. These ratios were selected based on previous studies in our lab, which suggested that β chains exhibit self-association, potentially leading to homodimer formation. Thus, an excess of α chain was introduced to promote heterodimer formation.

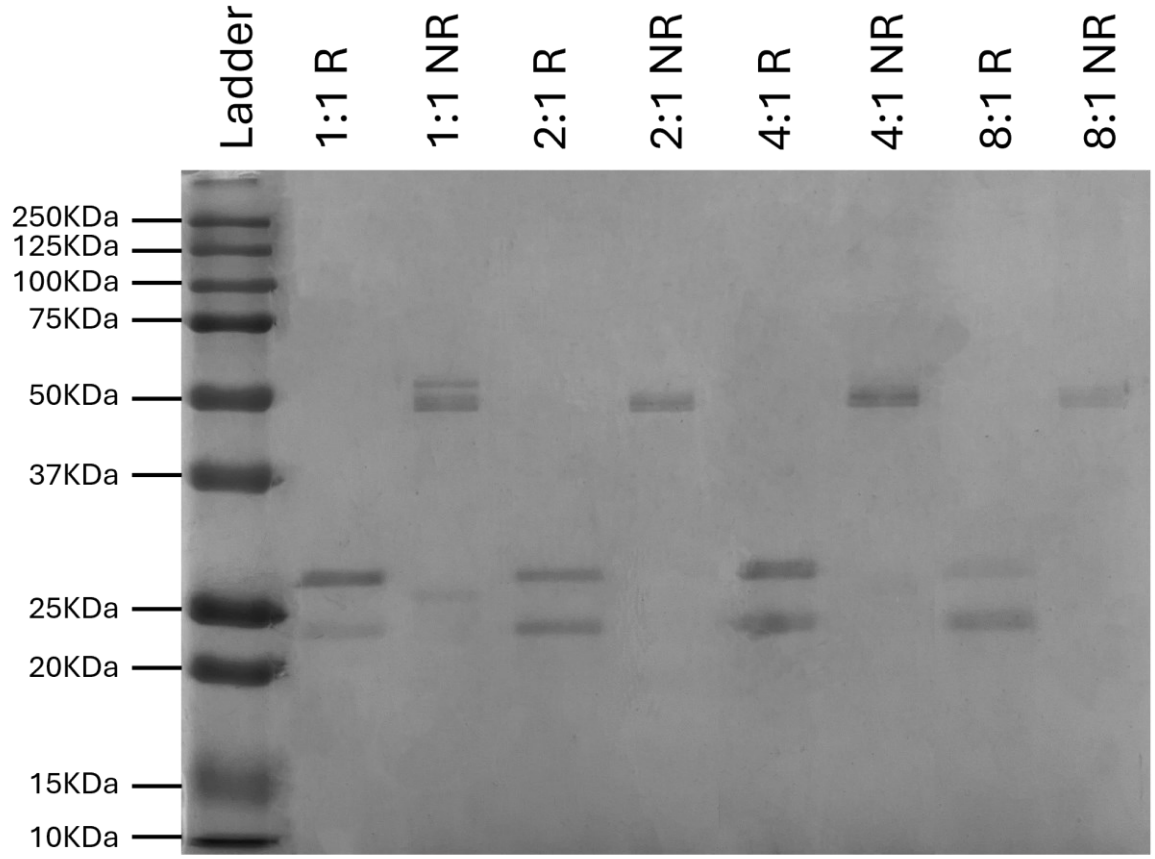


Figure 4. 3 SDS-PAGE Analysis of iTCR Refolding Using Different α:β Ratios (DN25)

SDS-PAGE analysis of refolded iTCRs using four different α:β chain ratios, post-dialysis in water and TRIS buffer. The gel suggests that a 2:1 or 4:1 ratio is optimal, as indicated by the presence of distinct bands at equal intensities in the reducing lane (representing individual chains) and a single band at ~50 kDa in the non-reducing lane, corresponding to the refolded heterodimer. A Bio-Rad molecular weight ladder was run in lane 1 for comparison.

The SDS-PAGE analysis revealed that the 4:1 α:β ratio yielded the highest amount of heterodimer. The 1:1 ratio appeared to produce an excess of β chain, indicated by two bands in the non-reducing lane and a darker band at a higher molecular weight in the reducing lane. The higher band in the non-reducing lane is likely a β homodimer, as the β chain has a higher molecular weight than the α chain (*Figure 4.3*).

At an 8:1 ratio, the formation of heterodimer was reduced, despite the elimination of the homodimer band. This condition also resulted in an excess α chain in the reducing lane. The 2:1 and 4:1 ratios appeared similar on the SDS-PAGE gel; however, the 4:1 ratio provided a greater overall yield of heterodimer (*Figure 4.3*). Based on these findings, we selected the 4:1 ratio for all subsequent experiments.

We then proceeded to refold both high- and low-affinity iTCRs using the 4α:1β ratio, at a scale sufficient to generate purified TCRs for structural and functional studies (500 mL).

The same dialysis protocol was used to initially refold both high- and low-affinity iTCRs. Conductivity measurements were taken after two days of dialysis, prior to POROS HQ concentration, to ensure maximum binding and elution of soluble protein from the ion exchange column. Flow-through fractions from the POROS concentration step were analysed by SDS-PAGE to confirm that heterodimer loss did not occur during the process (data not shown).

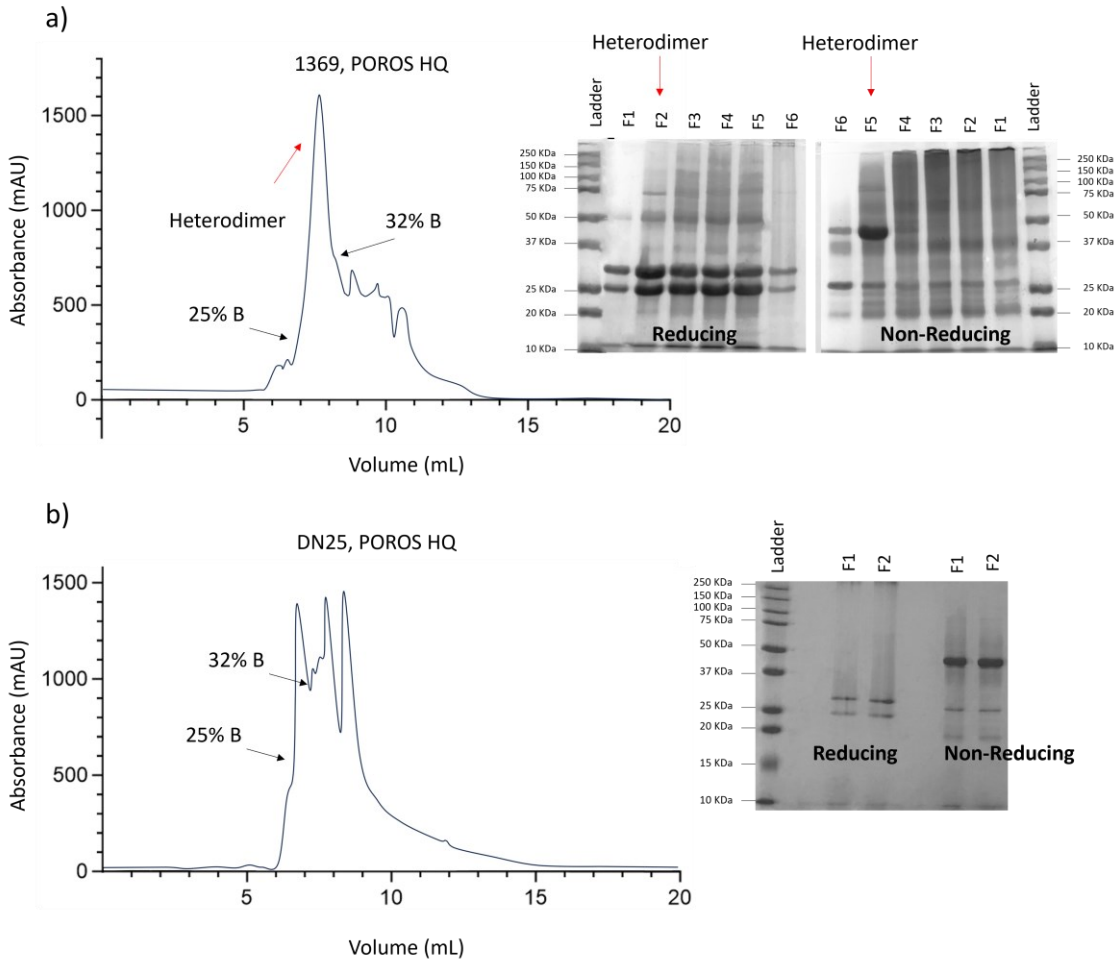


Figure 4. 4 POROS HQ concentration and SDS-PAGE gels of DN25 and 1369 iTCRs

iTCRs were concentrated using a POROS HQ column, and a salt gradient (Buffer B – 1 m NaCl) was used to elute protein from the anion exchange column. SDS-PAGE analysis showed that TCR elution occurred within the first sharp, large peak, starting at ~25% Buffer B and finishing at ~32% Buffer B. 1369 iTCR (a) refolded more efficiently than DN25 (b), which exhibited multiple peaks corresponding to aggregates and single-chain TCR fractions. A large peak corresponding to heterodimer was pooled for further purification.

Both iTCR heterodimers were eluted between 25% and 32% Buffer B, as confirmed by reducing and non-reducing SDS-PAGE. The non-reducing lane showed a single band at ~50 kDa, corresponding to the refolded heterodimer, while the reducing lane displayed two bands at equal intensities, representing the individual α and β chains (*Figure 4.4*).

The heterodimer-containing fractions were then concentrated to <5 mL for size exclusion purification using a HiPrep 26/60 Superdex 75 column. The purification was performed at a flow rate of 2.4ML/min, following the manufacturer's guidelines (Cytiva).

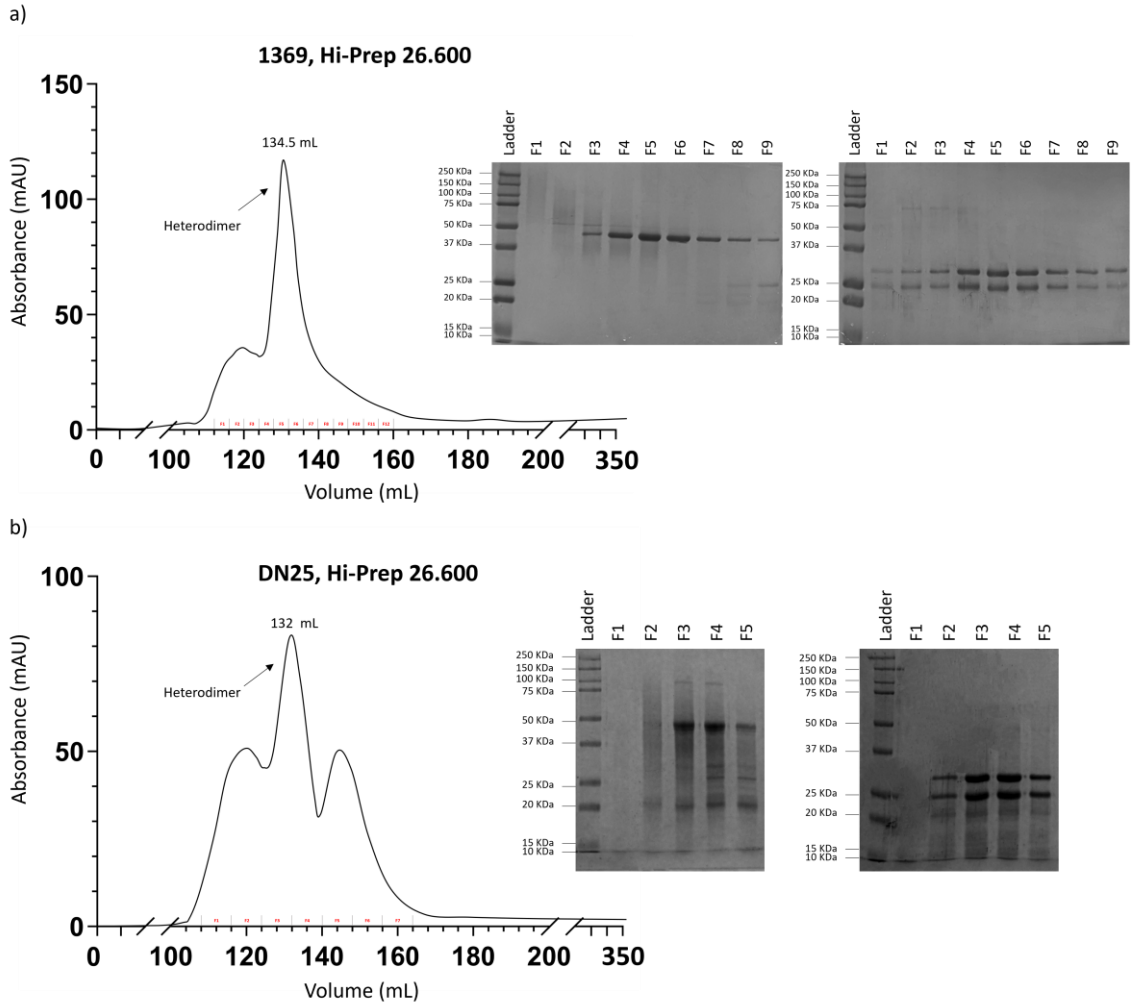


Figure 4. 5 Gel Filtration Purification of iTCRs

Pooled fractions from the POROS HQ column were purified using a HiPrep 26/60 Superdex 75 gel filtration column via size exclusion chromatography. Protein was loaded into the column using a 5 mL Superloop and run at 2.4ML/min, following guidance from the Cytiva user protocol. The resulting chromatograms were similar for both TCRs, with heterodimer elution corresponding to the major peak, which had a peak maximum at ~133 mL, as confirmed by SDS-PAGE analysis. The 1369 purification profile (a) appeared cleaner than DN25 (b), exhibiting less aggregated protein and minimal single-chain species in the non-reducing lanes. Fractions corresponding to heterodimer (F3–F9 for 1369 and F3–F5 for DN25) were pooled and concentrated for a final purification step.

As shown in *Figure 4.5*, chromatogram profiles from the size exclusion purification of DN25 and 1369 exhibited similar heterodimer retention times (134.5 mL for 1369 – *Figure 4.5a*, and 132 mL for DN25 – *Figure 4.5b*). The 1369 TCR displayed a cleaner elution profile, with a higher heterodimer-to-single-chain ratio compared to the low-affinity DN25 TCR. Additionally, 1369 exhibited lower levels of

aggregation, as indicated by the smaller initial peak preceding the heterodimer peak in both chromatograms. SDS-PAGE analysis corroborated these findings, with sharper, more defined bands present in both reducing and non-reducing gels for the 1369 TCR elution, in contrast to broader, less distinct bands in the DN25 gels. For X-ray crystallography, high-purity protein is essential for maximising the chances of successful diffraction (201, 202). To optimise protein purity, the DN25 TCR underwent further purification using an analytical Superdex 200 10/300 GL gel filtration column, via size exclusion chromatography. As shown in *Figure 4.1*, the DN25 TCR contains a C-terminal BirA motif, enabling biotinylation before analytical purification. The rationale for this additional step was to assess the binding capability of DN25 to a CD1d-expressing cell line using tetramer studies. This experiment was conducted to confirm successful TCR refolding and CD1d binding capacity before progressing to crystallographic studies.

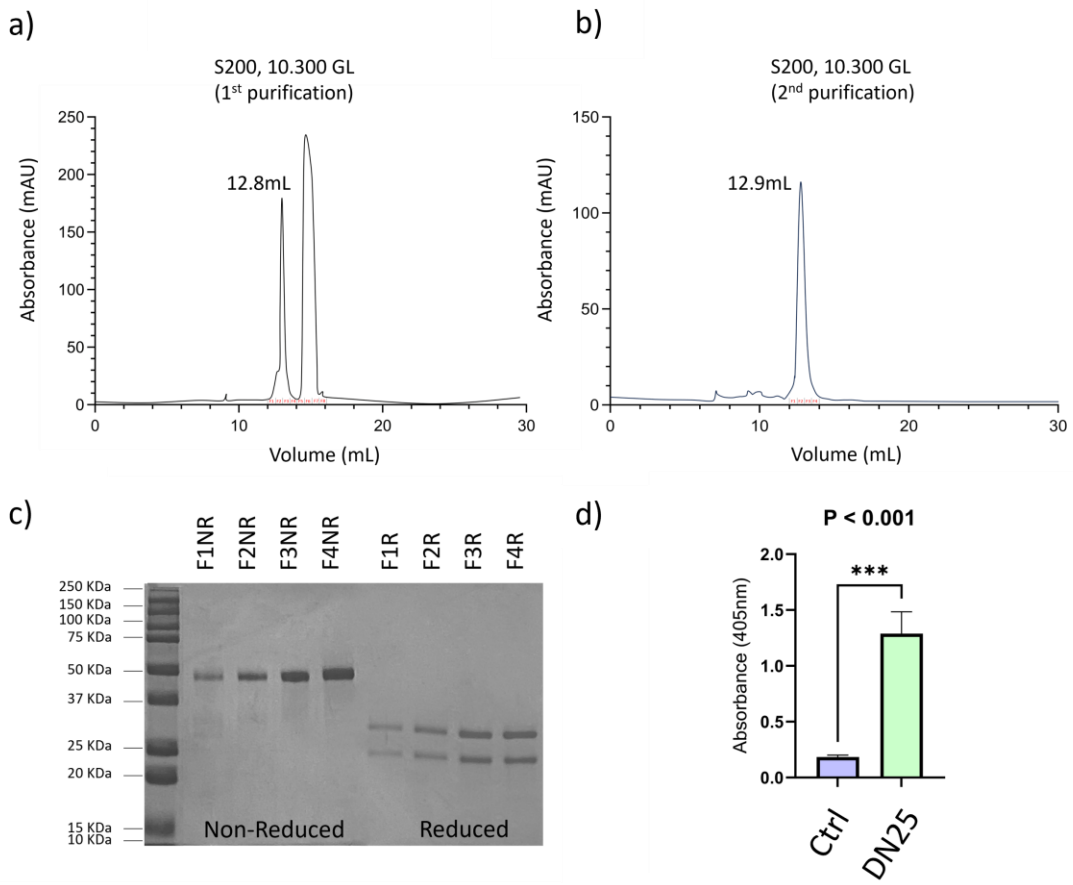


Figure 4. 6 Analysis pipeline for DN25 iTCR purification

The DN25 iTCR underwent two rounds of purification following overnight biotinylation to remove excess biotin. (a) Size exclusion chromatogram following Superdex 200 10/300 GL purification. (b) Fractions F1–F4 from the initial purification were pooled and re-purified using the same column to further eliminate excess biotin. (c) SDS-PAGE analysis of F1–F4 from the second purification was performed to assess sample purity. Reducing lanes showed distinct bands at ~25 kDa and ~30 kDa, corresponding to the Vα24 and Vβ11 chains, respectively. The non-reducing lane displayed a single band at ~50 kDa, corresponding to the heterodimer. (d) Biotin ELISA confirmed successful biotinylation of the purified TCR, as demonstrated by a significant signal increase compared to the unbiotinylated iTCR control. (n = 3 for all samples in the ELISA assay).

Consecutive purifications revealed almost identical retention times on the chromatogram, with the elimination of the second peak (likely corresponding to the biotin ligase enzyme) demonstrated in the second purification. The biotin ELISA illustrated a strong signal from the biotinylated protein, relative to the unbiotinylated DN25 control, suggesting the overnight biotinylation was successful. Consequently, we tested our DN25 TCR sample using CD1d-conjugated MACSI beads.

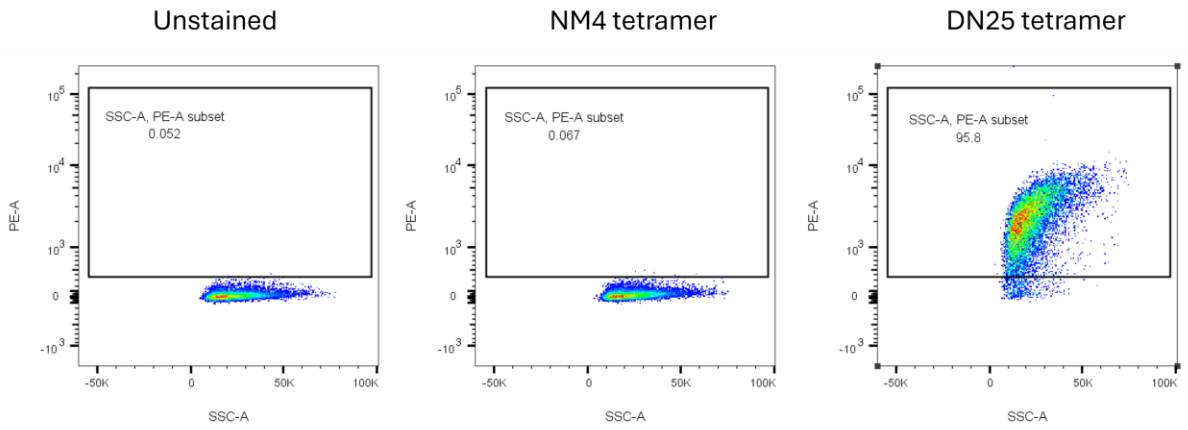


Figure 4. 7 Validation of DN25 TCR tetramer binding to CD1d-conjugated MACSi beads

FACS plots showing MACSi beads conjugated with biotinylated CD1d and stained with either the NM4 TCR tetramer (negative control), or the DN25 TCR tetramer. The left plot shows an unstained control. Representative flow cytometry plots show SSC-A vs PE-A signal for each condition.

The DN25 TCR tetramer successfully stained the CD1d conjugated MACSI beads, with the NM4 control showing negligible staining. This confirmed the DN25 TCR was binding to its target and was of high quality. Understanding the molecular mechanisms underpinning iTCR recognition of CD1d-lipid complexes is fundamental to the development of novel therapeutics that utilise these proteins. The role of the CDR3β loop in distinguishing between high- and low-affinity TCRs remains unclear. To address this, we attempted to crystallise these TCRs in both their binary form and in ternary complexes with CD1d. Before proceeding, we briefly assessed the stability of CD1d in complex with OCH, as a potential model system for investigating iNKT binding mechanisms using a weaker lipid agonist. Our rationale for this approach was that DN25 and 1369 iTCRs were originally isolated using CD1d-OCH, as the weaker lipid facilitated the identification of distinct iNKT clonal populations in tetramer studies (187). This suggests that structural findings relating to CDR3β loop contribution between iTCRs with diverse affinities are likely to be more pronounced when in complex with CD1d-OCH.

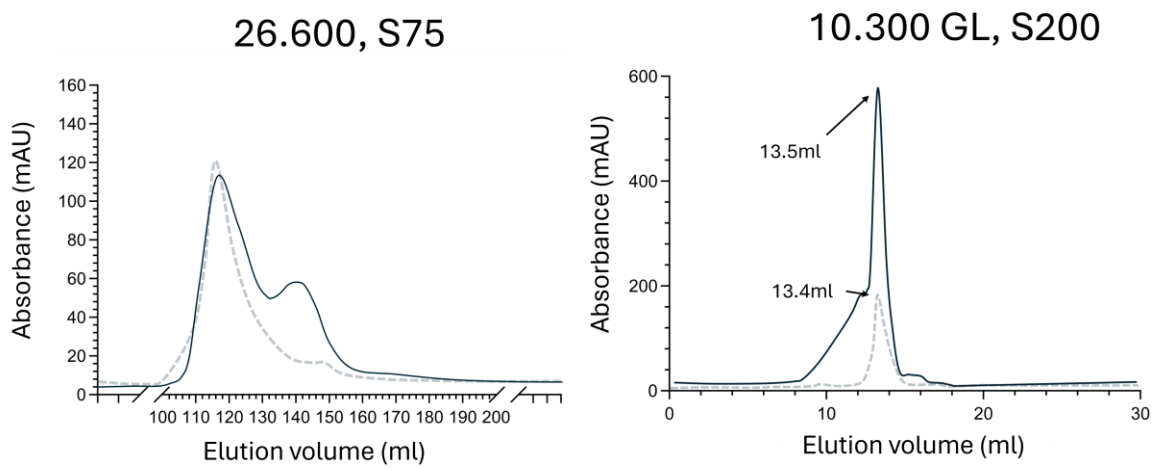


Figure 4. 8 CD1d-lipid complex purification

Chromatogram traces of Hi-prep (26.600, S75) and analytical (S200, 10.300 GL) size exclusion chromatography gel filtration runs of CD1d-OCH (dotted lines) and CD1d-αGC (solid lines). The instability of CD1d-OCH protein led to a lower total yield in both purification runs when compared to those of CD1d-αGC.

As shown in *Figure 4.8*, initial gel filtration using the preparative column revealed markedly different elution profiles, with a prominent shoulder in the αGC trace compared to OCH. The shoulder corresponding to the heterodimer was further purified using analytical SEC, which again displayed a large, homogeneous peak in the αGC trace, whereas the OCH trace remained more heterogeneous.

Due to the instability of the CD1d-OCH complex, it was deemed unsuitable for further structural studies. Consequently, CD1d-αGC was used for co-crystallisation attempts with both high- and low-affinity iTCRs.

Although the original iTCR clonal populations were identified using CD1d-OCH tetramers, SPR affinity measurements using CD1d-αGC as the ligand exhibited a similar affinity range to that observed in CD1d-OCH tetramer studies. Specifically, the 4C12 TCR (lowest affinity) and 1369 TCR (highest affinity) differed in KD values by approximately 15-fold. This suggested that structural data of the TCR-CD1d complex would still provide insight into the molecular basis of iNKT-TCR CDR3β recognition of CD1d, particularly in distinguishing between high- and low-affinity iTCRs.

Before attempting co-crystallisation, we first sought to crystallise DN25 and 1369 in their binary forms to investigate conformational changes in the TCR CDR loops upon complex formation and to determine whether these differed between high- and low-affinity iTCRs.

4.3 Investigating high- and low-affinity iTCR structure.

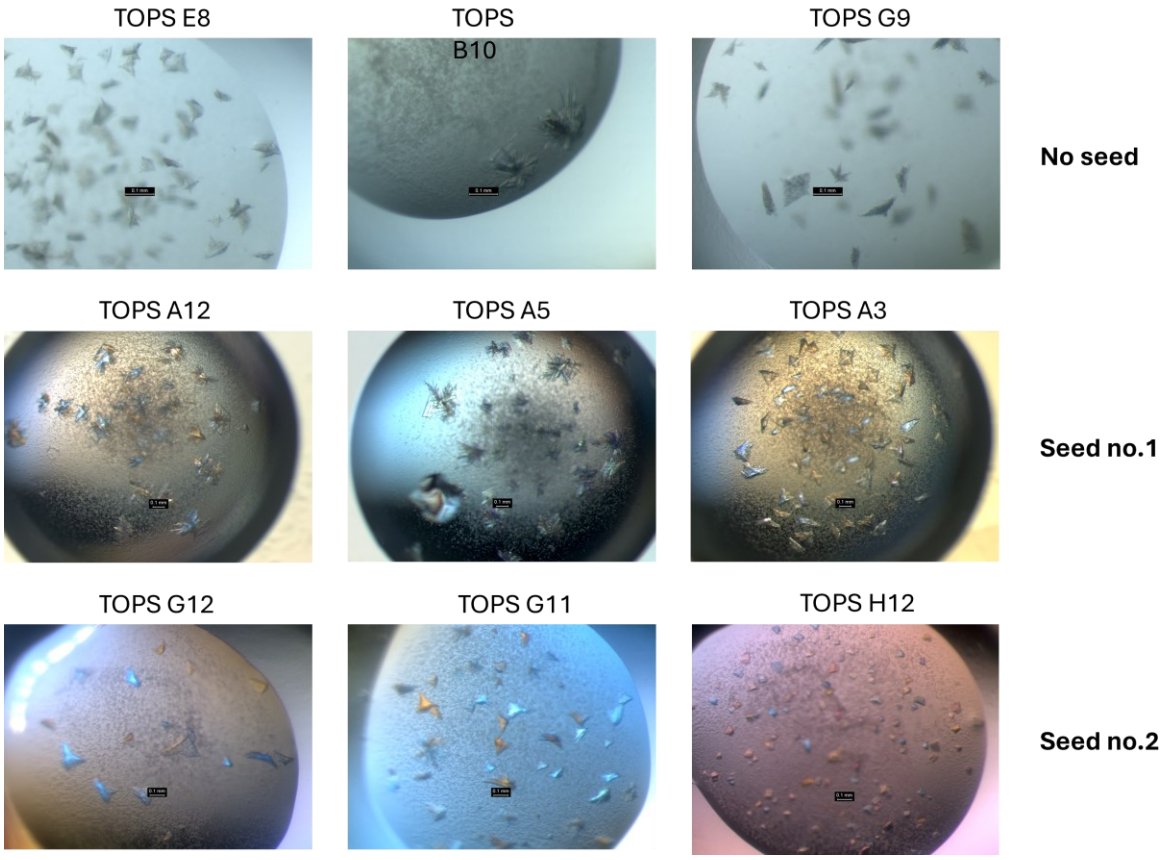


Figure 4. 9 Crystal images of DN25 TCR pre- and post-optimisation

Microscope images of DN25 TCR crystals under different conditions, with and without seed stock (from previous crystallisation attempts). A variety of morphologies can be seen in different wells of the TOPS crystallisation screen. Second and third rounds of crystallisation were carried out with varying concentrations of seed stock, to increase crystal size.

The DN25 iTCR was successfully crystallised, with initial screens (JCSG+, Morpheus, and TOPS) producing small, grainy crystals after ~4 days (*Figure 4.9*, top row). Their poor morphology suggested they were unlikely to diffract, prompting optimisation using seed stock generated from these initial crystals.

Secondary crystallisation attempts using seed stock (0.5:1:1, seed stock: mother liquor: protein) resulted in crystals with sharper edges and improved uniformity. Further optimisation (third round) produced high-quality crystals, which were harvested and frozen for X-ray diffraction.

Diffraction data were collected at the Diamond Light Source, producing a dataset with a resolution of 3.5 Å. Structural determination was performed using molecular replacement (PDB code 2CDE),

followed by multiple rounds of refinement. Despite the low-resolution presenting challenges, an iTCR structure was successfully determined, enabling comparisons with other iTCRs in the PDB.

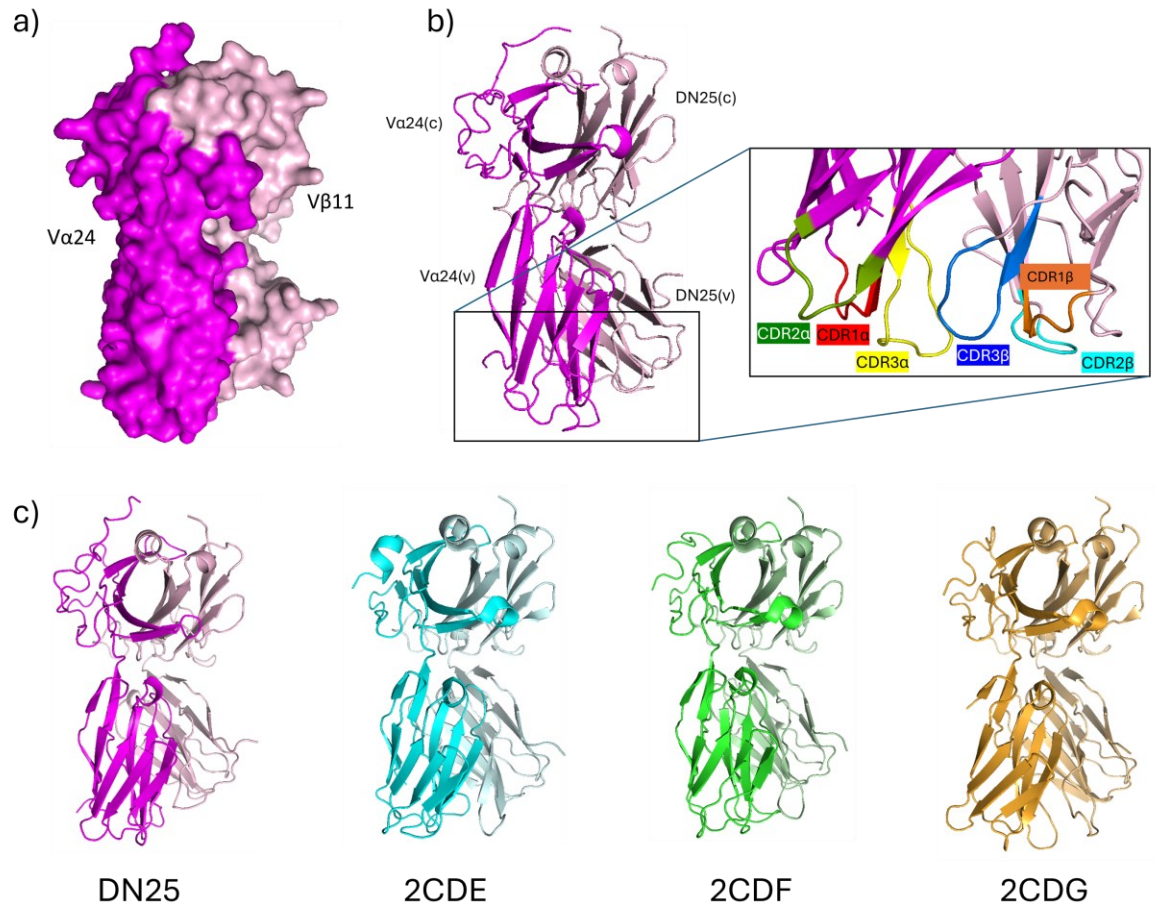


Figure 4. 10 Crystal structure of the DN25 TCR

a) Surface representation of the DN25 iTCR with the α chain coloured in magenta and the β chain coloured in pink. b) Cartoon representation of the DN25 TCR including a close-up of the CDR loops. CDR1, CDR2, and CDR3 of the α chain are coloured in red, green, and yellow respectively. The same loops from the β chain are coloured in orange, cyan and dark blue respectively. c) cartoon representation of DN25 overall structure compared to another iTCRs (2CDE) and two non Vα24 CD1d restricted TCRs.

The overall architecture of the DN25 TCR closely resembled previously published iTCR structures, with main chain conformations aligning well with known structures. Superimposition of the DN25 TCR onto PDB structures 2CDE, 2CDF, and 2CDG yielded RMSD values of 0.900, 1.572, and 1.586, respectively. This was expected, as 2CDF and 2CDG feature a non-Vα24 α chain, whereas 2CDE contains the canonical iNKT Vα24 chain, making it the closest structural match.

The highest crystallographic B-factors were observed within the α constant region, a phenomenon previously noted in other published iTCR structures (113, 206). This finding was further corroborated by RMSD alignment values between individual chains of DN25 and 2CDE, which were 0.811 and 0.594

for the α and β chains, respectively. Alignments of the DN25 α chain to 2CDF and 2CDG produced RMSD values of 1.669 and 1.969, while β chain alignments resulted in RMSD values of 0.772 and 1.083, respectively.

Due to the low resolution of our structure, individual residue-level analysis of the CDR3β loop was challenging. However, main chain comparisons revealed notable differences between DN25 and previously characterised iTCRs. Sequence analysis of the CDR3β loop between DN25 and 2CDE, 2CDF, and 2CDG revealed that DN25 had a significantly shorter CDR3β sequence (21 amino acids compared to 24, 25, and 24, respectively). Structural analysis further demonstrated that the DN25 CDR3β loop was substantially shorter in the binary structure, suggesting that it is unable to extend vertically to the same extent as other iTCRs. This potentially limits interactions between the CDR3β loop and the CD1d helices.

In parallel to the DN25 iTCR crystallisation pipeline, we also attempted to crystallise the high-affinity 1369 TCR for comparison. Despite extensive optimisation, including multiple crystallisation rounds, varying protein concentrations, and cross-seeding, no diffraction-quality crystals were obtained for the high-affinity TCR 1369, precluding structural determination via X-ray crystallography.

As an alternative approach to the crystal structure, AF2 modelling was utilised to generate crystal structures of the other TCRs discovered in the paper by Matulis et.al 2011(114).To validate the accuracy of AF2 models, we first compared an AF2-generated DN25 structure against our experimentally determined DN25 crystal structure.

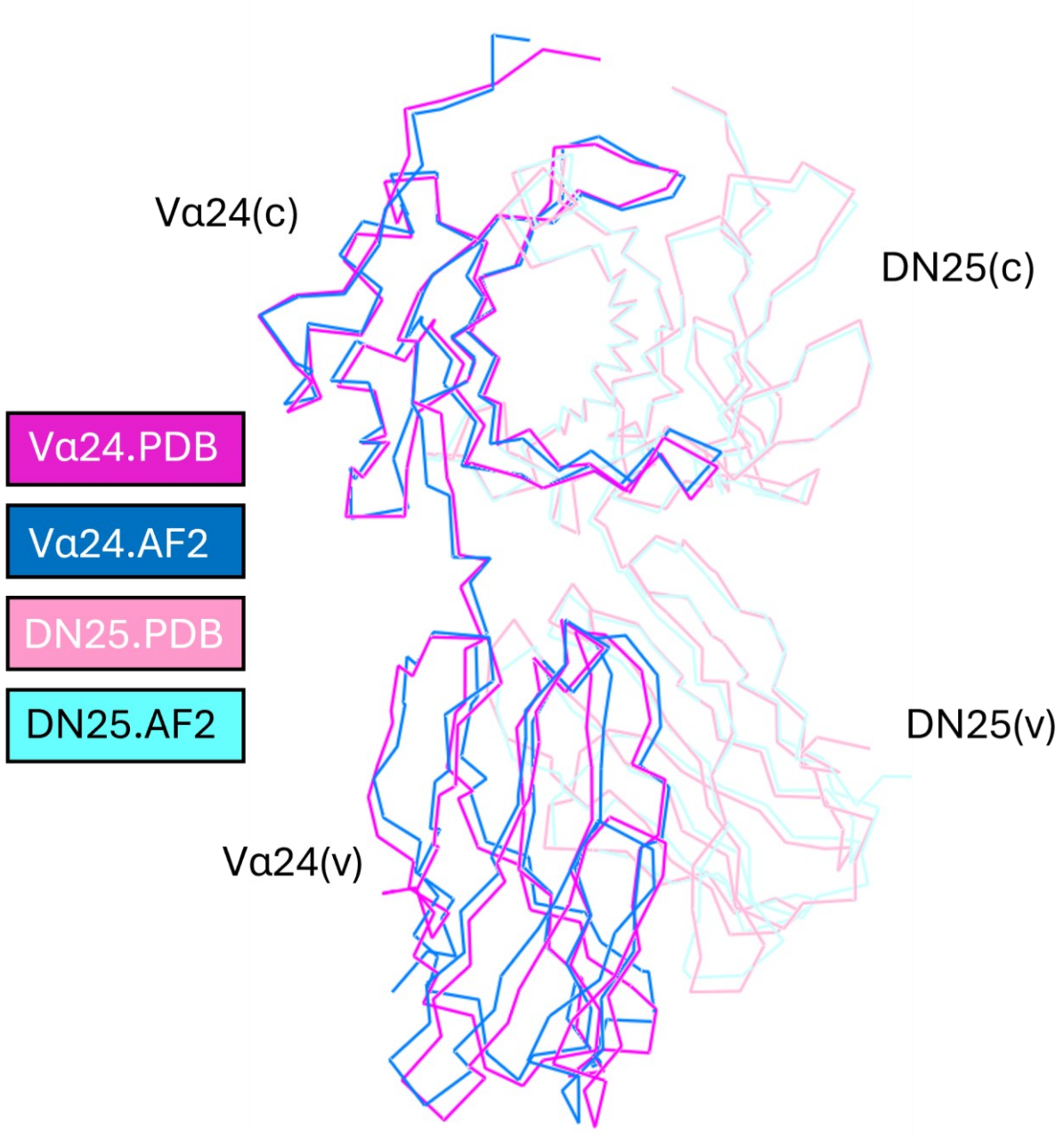


Figure 4. 11 DN25 crystal structure comparison with AF2 model

Cartoon representation and overlay between our DN25 crystal structure and the DN25 AF2 generated model. α and β chains from the crystal structure are coloured in magenta and pink respectively, whilst the same chains from the AF2 model are coloured in blue and cyan. Models overlay with surprising accuracy, with RMSD values generated in PyMOL of 0.941 from 3310 atoms.

Superposition of the DN25 crystal structure with the AF2-generated model revealed a highly similar overall structure, with an RMSD value of 0.941 across 3,150 atoms. The RMSD contribution from both chains was comparable, with α and β chain overlays yielding RMSD values of 0.9 and 0.8, respectively.

Closer analysis of the CDR3β loop conformation indicated a highly similar structural arrangement between the AF2 model and the experimentally solved structure, with most amino acid side chains oriented in similar directions. This provided confidence in using AF2-generated models to explore the structural differences between high- and low-affinity iTCRs.

To further investigate potential structural determinants of TCR affinity, AF2 modelling was used to generate structures for six additional iTCRs *(Figure 4.12*). These models revealed seven highly similar TCR structures, with CDR3β loop architecture being the primary distinguishing feature. To assess the likely impact of CDR3β sequence length and amino acid composition on CD1d binding affinity, we conducted a comparative analysis of both sequence and structural properties across all seven iTCRs.

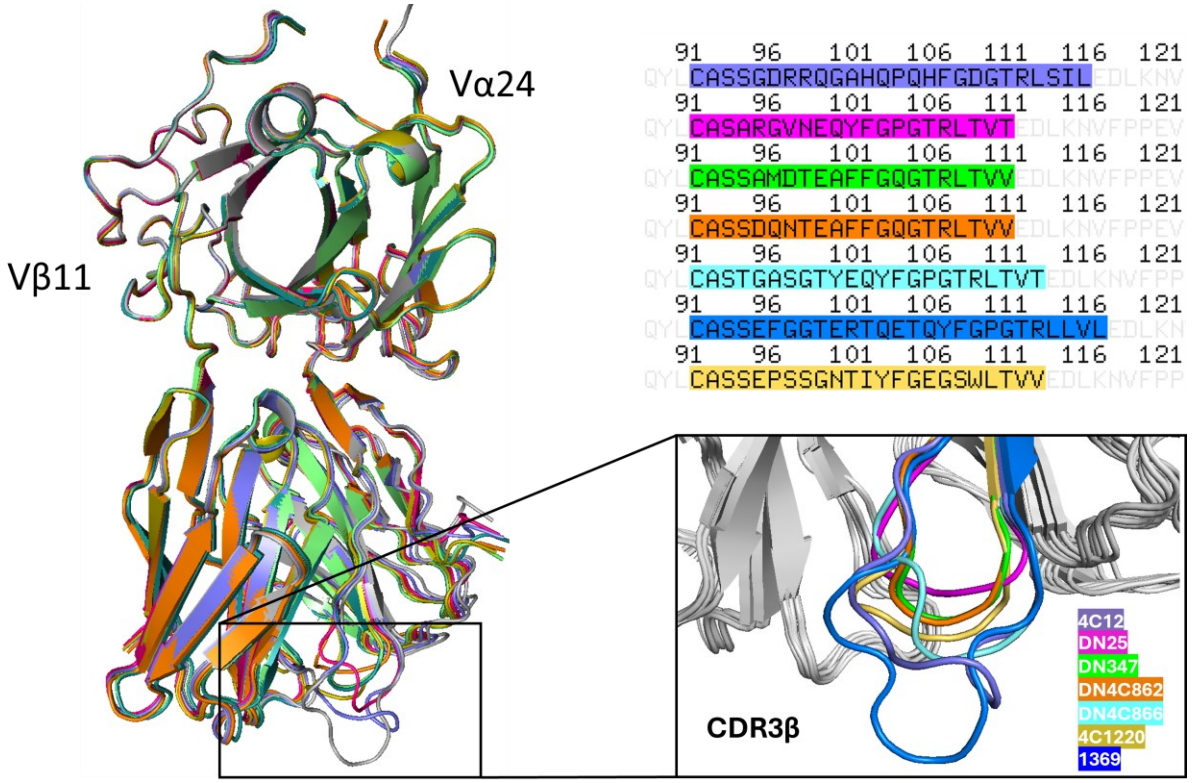


Figure 4. 12 CDR3β loop comparison of seven different iTCRs

Cartoon representation of seven different iTCRs structures generated using AF2 (colab). Individual TCRs are colour coded as denoted by the naming system in the bottom left corner of the zoomed in image. TCRs are named in order of affinity as seen in the publication by Matulis et.al, with the lowest affinity at the top (4C12) and the highest affinity at the bottom (1369). The zoomed-in image shows a substantial difference in CDR3β loop conformation existing primarily due to the length of the V,D and J regions of the CDR3β loop of each TCR. The image is generated using PyMOL.

Substantial differences in CDR3β loop conformations and sequences were observed across the seven iTCRs. The highest-affinity TCR (1369) exhibited the longest CDR3β loop, potentially allowing for greater interaction with CD1d-αGC. In general, CDR3β loop length correlated with TCR affinity for CD1d-αGC. However, 4C12 was a notable outlier, as it had the second-longest CDR3β loop but the lowest affinity for CD1d-αGC.

Given the variability in CDR3β loop conformations, co-crystallisation of CD1d-αGC with two TCRs at opposite ends of the affinity spectrum (DN25 – low affinity, 1369 – high affinity) was attempted. The hypothesis was that DN25 and 1369 exhibit differing CDR3β contributions at the TCR-CD1d binding interface, which in turn could affect the stability and affinity of the TCR for CD1d.

Co-crystallisation was attempted using CD1d-αGC protein (purified as described in Chapter 3) and DN25 protein (produced as described earlier in this chapter, *Figures 4.4 and 4.5*). Successful co-crystallisation requires high-purity, stable complexes and the correct monomer ratio within the crystallisation mixture.

The initial co-crystallisation method, based on Yu et al. (2011) (207), involved incubating the TCR and CD1d monomers together for one hour before crystallisation set-up. However, no crystal formation was observed, and most wells showed precipitation, suggesting that complex formation was insufficient to support ordered crystal lattice formation. A modified approach, based on Wegrecki & Rossjohn et al. (2022) was then tested (52). Here, the TCR and CD1d monomers were incubated at 4°C overnight, followed by a concentration of 8Mg/mL at a 1:1 ratio prior to crystallisation set-up. This method yielded large plate-like crystals after one week of incubation.

Although crystals were successfully obtained, they appeared fragile during harvesting for X-ray diffraction, and subsequent attempts to improve crystal quality were unsuccessful, despite trialling multiple optimisation methods from previously published literature. Optimisation strategies included varying the TCR: CD1d ratio, adjusting the TCR-CD1d complex concentration in the crystallisation set-up and incorporating seed stock from previous DN25 and CD1d-αGC crystals, both individually and together. Despite these efforts, high-quality diffraction-grade crystals were not obtained.

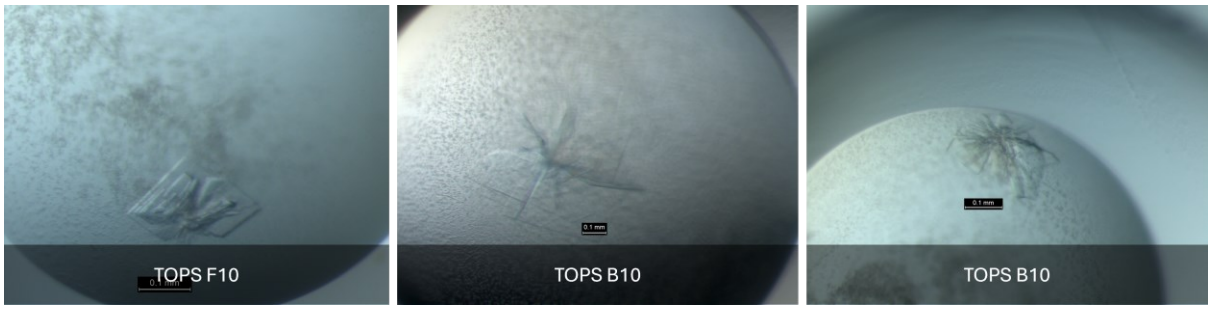


Figure 4. 13 Crystals of DN25-CD1d-αGC in the TOPS screen

Microscope images of DN25 TCR in complex with CD1d-αGC. Crystals formed in the TOPS screen under different conditions (B10 and F10) and had a thin, plate-like morphology. Crystals appeared ~8 days post-set-up and were tested under UV for protein conformation. Individual fragments were fished, and flash-frozen in liquid nitrogen for x-ray diffraction.

Unfortunately, the crystals did not diffract to a sufficiently high resolution for structural analysis. Given these limitations and time constraints, we adopted an alternative approach, docking AF2-generated iTCR models and our DN25 crystal structure onto a previously published iTCR-CD1d-αGC co-crystal structure using PyMOL overlay.

The rationale for this approach was to investigate the spatial proximity of the CDR3β loops in TCRs with differing affinities to the αGC lipid, providing insights into how CDR3β loop positioning may influence iNKT affinity.

4.4 Investigating differences in CD1d-αGC recognition between high and low-affinity iTCRs.

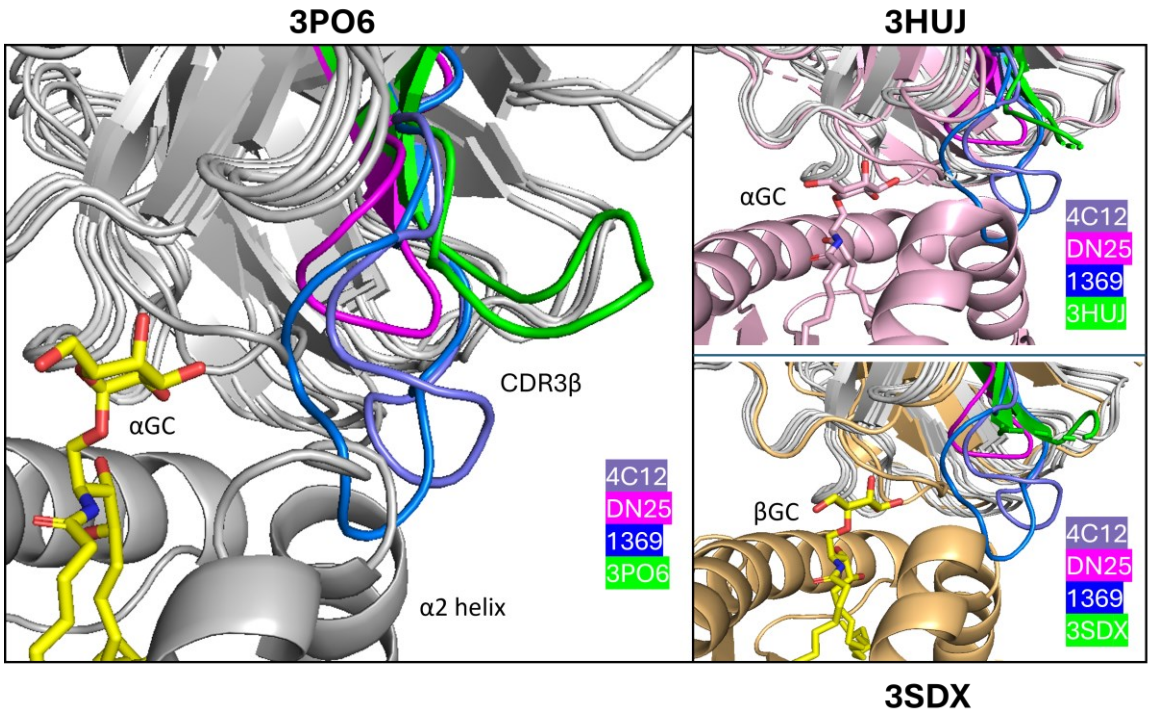


Figure 4. 14 Different affinity iTCRs predicted binding to CD1d-αGC.

PyMOL docking of the DN25 crystal structure and 4C12 and 1369 AF2 models, overlayed with 3PO6, 3HUJ, and 3SDX. CDR3β loop conformation is shown with 4C12 in purple, DN25 in magenta, and 1369 in dark blue. CDR3β loop of the respective co-crystal structure is shown in green for comparison. Substantial conformational change is required for the CDR3β loop to adopt suitable conformation upon interaction with the CD1d-lipid complex in all three cases.

Overlaying our DN25 crystal structure and AF2-generated models (1369 and 4C12) with three different human iTCR-CD1d-αGC co-crystal structures from the PDB revealed a consistent structural trend.

The CDR3β loop of 1369 exhibited a significant clash with the α2 helix, suggesting that substantial conformational changes would be required upon binding to CD1d-lipid complexes. In contrast, the DN25 TCR was positioned much further from the α2 helix, while the wild-type TCR from the original co-crystal structure displayed substantial conformational rearrangement relative to both the binary AF2 model and the DN25 structure.

Collectively, these results demonstrate key structural differences between high- and low-affinity iTCRs, with a particular focus on the role of the CDR3β loop in CD1d recognition. Sequence and structural analyses showed variation in CDR3β loop length and composition, which may contribute to differences in binding affinity and flexibility. Crystallisation trials led to the successful structure determination of DN25, but 1369 did not yield diffraction-quality crystals, leading to the use of AF2 modelling for comparative structural analysis. Computational overlays provided insights into CDR3β loop positioning relative to the CD1d-binding interface, identifying structural constraints that may influence iNKT recognition and function. These findings provide a basis for further discussion on how variations in iTCR structure impact antigen recognition and immune response mechanisms.

4.5 Discussion

4.5.1 Investigating iTCR Affinity for CD1d-Lipid Complexes

T-cell and TCR-based immunotherapies traditionally rely on conventional αβ TCRs interacting with peptide major histocompatibility complexes (pMHC) or antibody-based strategies targeting extracellular antigens (208, 209). While these approaches have shown clinical efficacy, they are limited by HLA polymorphism, which restricts their applicability across diverse patient populations. Similarly, antibody-based therapies target only a small subset of extracellular antigens, covering only a fraction of the proteome (210).

An alternative strategy involves targeting non-peptide antigens presented by CD1 molecules, which can bind and display lipid antigens (211, 212). CD1d, which presents lipid antigens to invariant natural killer T (iNKT) cells, provides an avenue for non-peptide antigen recognition (213, 214). However, despite its potential, significant gaps remain in understanding the molecular mechanisms governing iNKT cell recognition of CD1d-lipid complexes, particularly how variations in iTCR affinity influence antigen recognition and TCR specificity.

This study aimed to investigate the structural variability in human iTCRs, particularly the hypervariable CDR3β loop, and how this may impact TCR function. Whilst successful crystallisation of a low-affinity iTCR was achieved, co-crystallisation of iTCRs with CD1d-αGC and the high-affinity 1369 iTCR alone remained a challenge, preventing high-resolution structural comparisons. AlphaFold 2 (AF2) was employed to generate structural predictions of a range of different affinity TCRs, and this was combined with PyMOL docking to model how CDR3β loop length, flexibility, and sequence may impact TCR binding dynamics.

By integrating experimental and computation modelling data, this study highlights that loop sequence and flexibility, rather than solely length, appear to contribute to iTCR binding affinity for CD1d-αGC complexes. However, in the absence of high-resolution crystal structures, these findings remain predictive, and further structural studies are necessary to confirm the intricate contribution of the CDR3β loop to iTCR affinity and subsequent function.

4.5.2 Structural basis of human iTCR recognition of CD1d-αGC

Previous studies have demonstrated that despite iTCR interactions with CD1d being dominated by the vα24 chain, individual iTCRs display varying affinities for CD1d, indicating the hypervariable CDR3β

loops play a role in fine tuning the interaction (113, 189). We investigate this by attempting to crystallise both binary iTCR structures (DN25 and 1369) and ternary iTCR-CD1d-αGC complexes. Despite extensive efforts, only the DN25 low affinity iTCR crystallised, suggesting structural factors – potentially the increase in flexibility from the longer CDR3β loop of the 1369 – hindered crystallisation. This aligns with previous studies where flexible loops have hindered the formation of a stable crystal lattice (190). Whilst the absence of 1369 crystals prevented direct structural comparisons, the DN25 crystal structure enabled analysis of loop composition relative to previously solved structures.

Our DN25 structure overlayed with previous iNKT crystal structures from Gadola et.al revealed a strong docking conservation geometry, reinforcing the conserved binding orientation for these iTCRs (215). However, sequence and structural comparisons suggest that differences in CDR3β loop conformation contribute to affinity variation. While shorter loops, such as in DN25, may limit stabilising contacts, longer loops, such as in high affinity iTCRs, appear to enhance interactions with CD1d. Prior studies indicate that CDR3β loop flexibility allows repositioning to accommodate diverse lipid antigens, which may be a key factor in affinity modulation (216).

Sequence analysis further suggests that specific amino acid compositions, rather than loop length alone, influence affinity. The high-affinity 1369 iTCR contains glycine-rich motifs, which have been implicated in increased flexibility and adaptability upon antigen binding (217). In contrast, 4C12, despite having a similarly long CDR3β loop, exhibited reduced affinity, possibly due to steric hindrance introduced by arginine residues. These findings align with studies showing that CDR3β interactions with CD1d-lipid complexes are often stabilised through water-mediated networks rather than direct contacts (218).

Although the lack of high-resolution co-crystals limits definitive conclusions, these observations suggest that CDR3β-mediated affinity modulation is driven by a combination of sequence composition and flexibility rather than structural rearrangement alone. Future studies using alternative structural techniques, such as cryo-electron microscopy or mutagenesis approaches, are necessary to fully resolve the contribution of CDR3β variation to iTCR affinity.

4.5.3 Insights from Murine iNKT TCRs

While structural data on human iTCRs remains limited, murine studies provide valuable insights into how CDR3β sequence variability influences CD1d recognition. Murine iNKT TCRs exhibit a hierarchical preference for specific Vβ chain usage, correlating with differential affinity for CD1d-lipid complexes

(100, 191). Given these findings, we explored whether similar mechanisms apply to human iTCRs by comparing structural motifs across species.

Previous murine studies demonstrated that while the CDR2β loop plays a major role in defining affinity (219), specific CDR3β sequences can compensate for lower-affinity Vβ chain interactions, enhancing overall binding stability (105, 191). Conversely, some CDR3β sequences reduce affinity by spatially hindering otherwise optimal interactions. This suggests that the CDR3β loop acts as a fine-tuning mechanism for affinity modulation rather than directly dictating lipid selectivity (189, 220).

In human iTCRs, a similar pattern may explain the divergent affinities observed between 1369 and 4C12. Despite their comparable CDR3β loop lengths, differences in sequence composition likely contribute to their distinct binding affinities for CD1d-lipid complexes. The presence of specific motifs, such as glycine-rich sequences, may enhance flexibility and binding adaptability, whereas other residues may introduce steric hindrance, reducing affinity. These observations align with prior studies showing that affinity differences in iTCRs are more pronounced when the lipid antigen is less stable within CD1d, reinforcing the notion that CDR3β loops can stabilise weaker interactions rather than drive lipid selectivity (105, 114).

Further work has identified a hydrophobic motif within the CDR3β loop of murine iNKT TCRs that directly modulates affinity for CD1d-lipid complexes (114). While this provides a mechanistic explanation for affinity variation in murine models, its relevance to human iTCRs remains unclear, as no direct correlation between hydrophobicity and affinity has been established in human iNKT TCRs (114). Given these interspecies differences, future studies should focus on comparative structural and functional analyses to determine whether murine-derived mechanisms translate to human iNKT cell function.

4.5.4 Conclusion & Future Directions

This study provides new insights into the structural features governing iTCR affinity for CD1d-lipid complexes, with a focus on the role of the hypervariable CDR3β loop. Structural modelling and sequence analyses suggest that loop sequence composition and flexibility, rather than length alone, influence binding affinity. However, due to crystallisation challenges, these findings remain predictive.

Future work should focus on obtaining high-resolution co-crystal structures of human iTCRs in complex with CD1d presenting diverse lipid antigens. Cryo-electron microscopy and NMR spectroscopy may provide alternative structural insights, particularly for high-affinity TCRs that prove difficult to

crystallise. Additionally, functional binding studies, such as surface plasmon resonance and mutational analysis, could help validate the role of specific CDR3β motifs in modulating iTCR affinity.

A deeper understanding of CDR3β sequence determinants may contribute to improved design strategies for iNKT-based immunotherapies. Given that iTCR affinity influences iNKT cell activation and effector function, further investigations into the structural mechanisms of antigen recognition will be critical for refining therapeutic approaches targeting iNKT cells.

Chapter 5 Generating CD1c-endo and γδ TCRs for structural and functional studies.

5.1 Generation of mammalian expressed CD1c-endo for structural and functional studies.

The overarching aim of this chapter was to optimise the production, characterisation, and functional analysis of CD1c-endo and γδ TCR proteins to support structural and functional studies. This involved improving the expression and purification of CD1c-endo using a mammalian expression system to achieve higher yields suitable for crystallisation. Additionally, we sought to develop a refolding and characterisation pipeline for a panel of γδ TCRs hypothesised to interact with CD1c and to investigate γδ TCR recognition of CD1c-endo using functional assays.

CD1c, a group 1 CD1 molecule, has been implicated in various diseases and is recognised as a ligand for γδ T cells, a subset of unconventional T cells that bridge innate and adaptive immunity. However, the molecular basis of γδ TCR recognition of CD1c remains poorly understood, largely due to the technical challenges of producing enough of these proteins for structural studies. Previous work in our lab generated CD1c protein using oxidative refolding methods, but low yields limited its crystallographic applications. To overcome this, we employed a mammalian expression system, producing CD1c-endo loaded with endogenous lipids, ensuring a consistent and scalable platform for further structural and functional studies.

Soluble γδ TCRs present significant challenges due to their propensity for aggregation and dependence on proper refolding conditions. To establish a reliable refolding pipeline, we initially focused on NM4, an αβ TCR with known CD1c reactivity, before applying the process to a panel of γδ TCRs. This approach allowed us to optimise protein yield, stability, and functionality conditions. Additionally, we developed a bead-based assay to assess the lipid-specific binding of refolded γδ TCRs to CD1c-endo, providing a framework for functional characterisation.

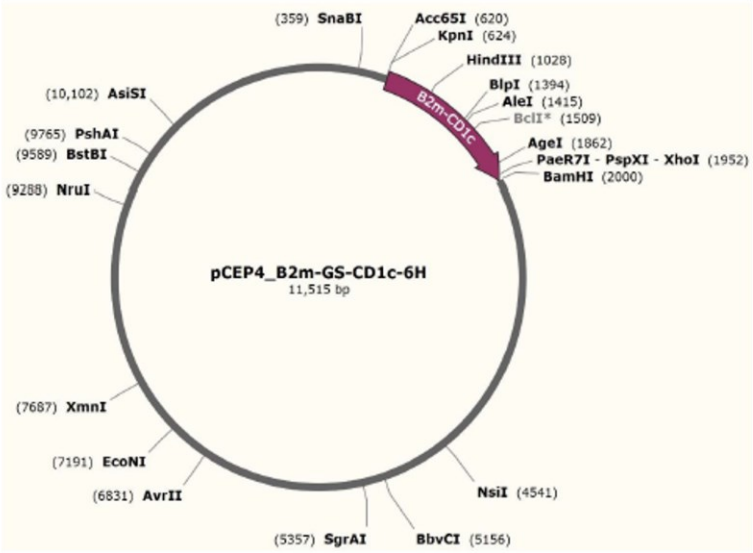
This chapter outlines the strategies used to overcome technical challenges in protein production and purification, evaluates the effectiveness of our optimisation efforts, and discusses their implications for future structural and functional studies of CD1c-reactive γδ TCRs.

5.1.1 CD1c-endo sequence and construct analysis

A common bottleneck in structural and functional studies investigating ligand-receptor complexes is the protein yield obtained from in vitro refolding protocols. These processes are often inefficient, as they struggle to replicate the in vivo refolding environment, which requires numerous chaperones and highly specific conditions. Even small deviations from these conditions can significantly impact final yield and protein quality.

Our lab previously produced CD1c protein via oxidative refolding, either in the presence of vehicle solution (NaCl and Tween-20) or with specific lipids in the refolding mixture. While this approach generated clean, well-folded protein suitable for functional studies, the low overall yield created a major barrier to pursuing structural studies. To overcome this, we pursued CD1c expression in a mammalian system, referred to hereafter as CD1c-endo, which is loaded with endogenous lipids and provides a consistent and scalable source of stable, soluble CD1c protein. Additionally, CD1c-endo has been widely used in the literature, including in recent structural studies (64).

a)



b)

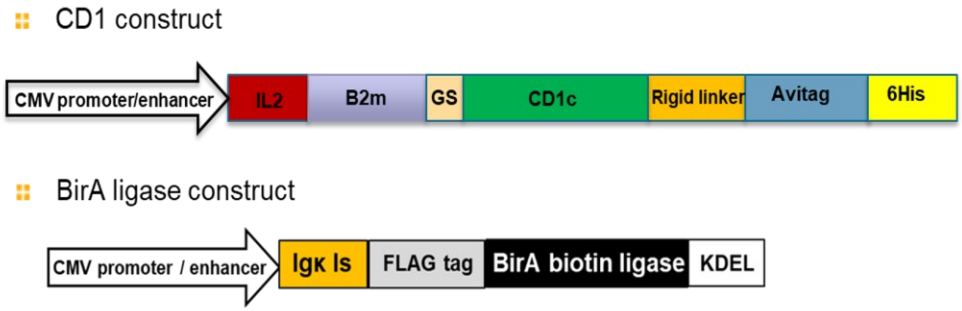


Figure 5. 1 Mammalian DNA plasmid and construct map of CD1c-endo and BirA ligase

Schematic diagrams of a) A simplified pCEP4 vector construct, showing the restriction sites surrounding the β2M-CD1c construct insert. b) Linear representation of the individual sub-compartments comprising the CD1 construct and the BirA ligase construct. Individual compartments are colour-coded with β2M and CD1c portions separated by a flexible linker. A His-tag is added to the end of the protein for purification and the Avi-tag is added for biotin conjugation.

Table 4 Amino acid sequences of the individual components making up the CD1c-endo plasmid construct.

Insert	Sequence
CD1c	KLNSEHVSFHVIQIFSFVNQSWARGQGSGWLDELQTHGWDSESGTIIFLHNWSKGNFSNEELSDLEL LFRFYLFGLTREIQDHASQDYSKYPFEVQVKAGCELHSGKSPEGFFQVAFNGLDLLSFQNTTWVPSPG CGSLAQSVCHLLNHQYEGVTETVYNLIRSTCPRFLLGLLDAGKMYVHRQVRPEAWLSSRPSLGSGQL LLVCHASGFYPKPVWVTWMRNEQEQLGTKHGDILPNADGTWYLQVILEVASEEPAGLSCRVRHSSL GGQDIILYWG
β2M	IQRTPKIQVYSRHPAENGKSNFLNCYVSGFHPSDIEVDLLKNGERIEKVEHSDLSFSKDWSFYLLYYTEF TPTEKDEYACRVNHVTLSQPKIVKWDRD
Linker	GGGGSGGGGSGGGGS
BirA	GLNDIFEAQKIEWHELEVLFQGP
HisTag	HHHHHH

5.1.2 Optimising CD1c-endo production via the mammalian expression system.

CD1c-endo was produced using the mammalian expression system (detailed in Methods). Briefly, EXPI-293F cells were cultured to a density of 3 × 10^6 cells/mL before being transfected using ExpiFectamine™ 293 transfection reagent and the CD1c-endo plasmid, both pre-diluted in Opti-MEM reduced serum media and added dropwise to the cells. Enhancers and biotin were added 20 hours post-transfection, and the protein was harvested five days later. CD1c-endo was then purified using a two-step process, beginning with His-trap affinity purification, followed by size exclusion chromatography. For protein intended for structural studies, the biotinylation step was omitted on day 1.

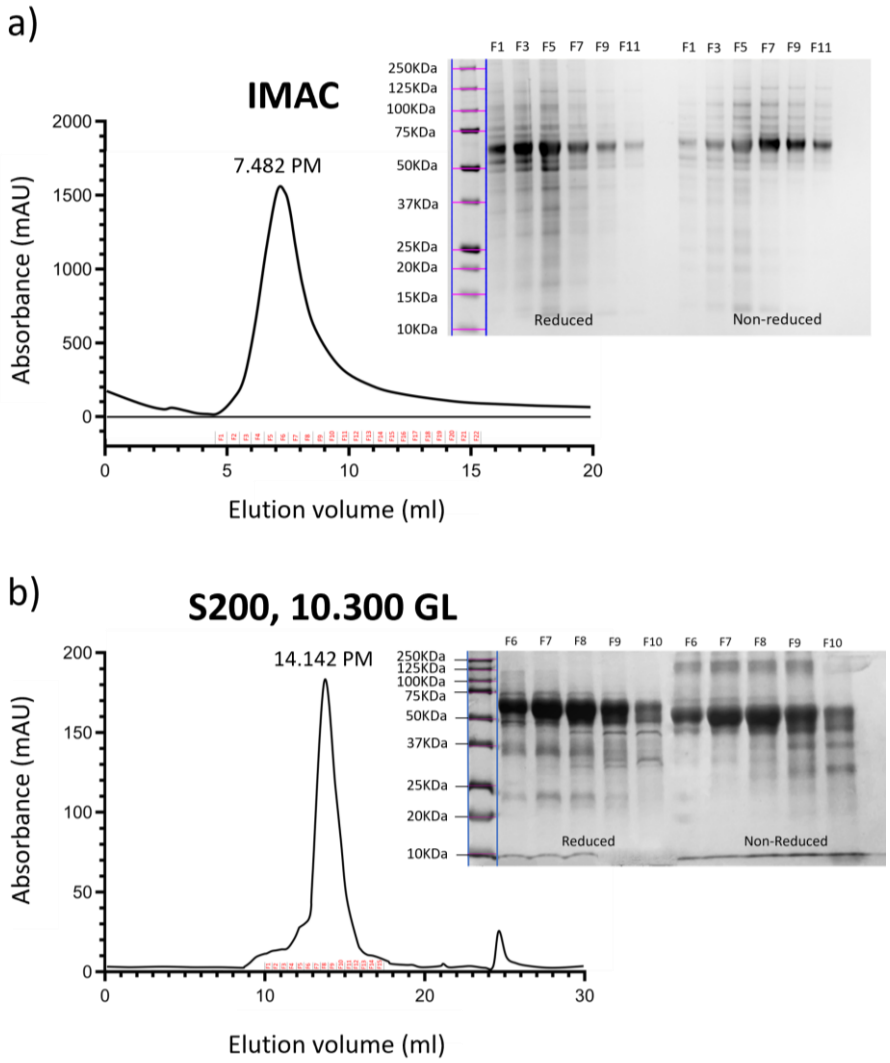


Figure 5. 2 Purification of mammalian expressed CD1c-endo.

Immobilized metal affinity chromatography (IMAC) and size exclusion chromatography (SEC) purifications and SDS-PAGE gels of CD1c-endo protein. a) Ni+ affinity HisTrap chromatogram and resulting reducing and non-reducing SDS-PAGE gel of CD1c-endo. Alternate fractions from the large peak were run on a reducing and non-reducing SDS-PAGE and displayed bands at ~50KDa when compared to the molecular weight ladder. b) Fractions corresponding to the large peak were pooled and run on an S200 analytical size exclusion column (10.300 increase) and fractions were analysed again via SDS-PAGE. Fractions corresponding to large peaks display strong bands at the correct molecular weight (~50KDa).

CD1c-endo bound to the Ni^{2+} affinity column via the His-tag located at the C-terminal end of the protein sequence and was eluted using a one-step imidazole elution. SDS-PAGE analysis of the eluted fractions showed strong bands at ~50 kDa, which were pooled and further purified using analytical size exclusion chromatography.

To confirm successful biotinylation of CD1c-endo, we incorporated biotin alongside enhancers at 18–22 hours post-transfection. As an alternative approach, post-purification biotinylation was performed using a Biotin Protein Ligase Kit (Tebu-Bio).

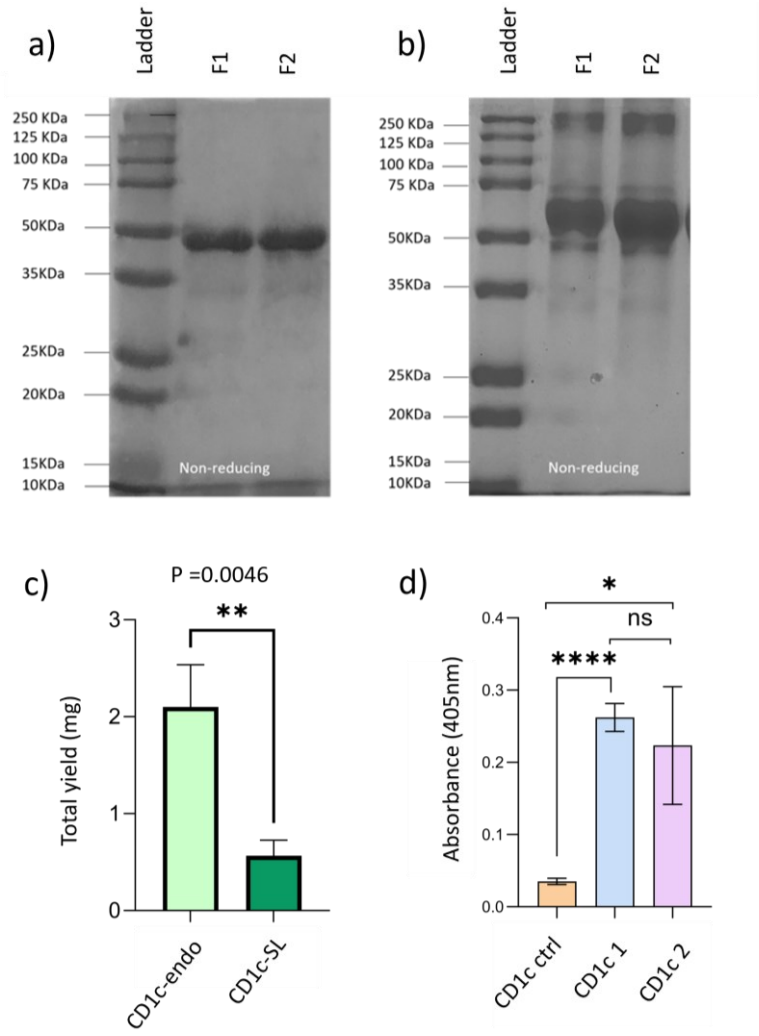


Figure 5. 3 Post-purification analysis of CD1c-endo and CD1c-SL

a) and b) show SDS-PAGE gels of purified CD1c-SL and CD1c-endo respectively, run using non-reducing loading dye, with Bio-Rad protein standard ladder in lane one on each gel. c) Bar chart showing total protein yield comparison between CD1c-endo and CD1c-SL (n=3 for each sample). d) Bar chart showing biotin ELISA results of CD1c-endo and CD1c-SL biotinylation relative to unbiotinylated control. Results are representative of three independent experiments, with P values of $P > 0.001$ and $p = 0.0161$ for ctrl vs CD1c one and ctrl vs CD1c two respectively.

CD1c-endo purification resulted in a threefold higher yield from a 100 mL transfection compared to a 500 mL CD1c-SL refold. Unlike CD1c-SL, which is refolded in the presence of a vehicle solution, CD1c-endo is naturally loaded with endogenous mammalian lipids. Previous structural data from Mansour et al. suggest that in CD1c-SL, the groove is occupied by 'spacer lipids', which stack in a parallel fashion within the F' pocket (65).

SDS-PAGE analysis of CD1c-endo revealed a slightly higher-than-expected molecular weight, likely due to post-translational modifications. Given the significantly higher yield and the retention of high-quality, pure protein bands in the CD1c-endo gel compared to CD1c-SL, we proceeded with CD1c-endo as our monomer of choice for investigating molecular interactions with γδ TCRs.

As part of our protein production and structural pipeline, we aimed to crystallise CD1c-endo to establish a seed stock for future co-crystallisation studies with γδ TCRs. The high yield of CD1c-endo provided sufficient protein to optimise crystallisation conditions, alongside efforts to generate γδ TCRs for co-crystallisation. To this end, an unbiotinylated batch of CD1c-endo was concentrated to 15 mg/mL and set up in standard crystallisation screens using an Oryx8 crystallisation robot.

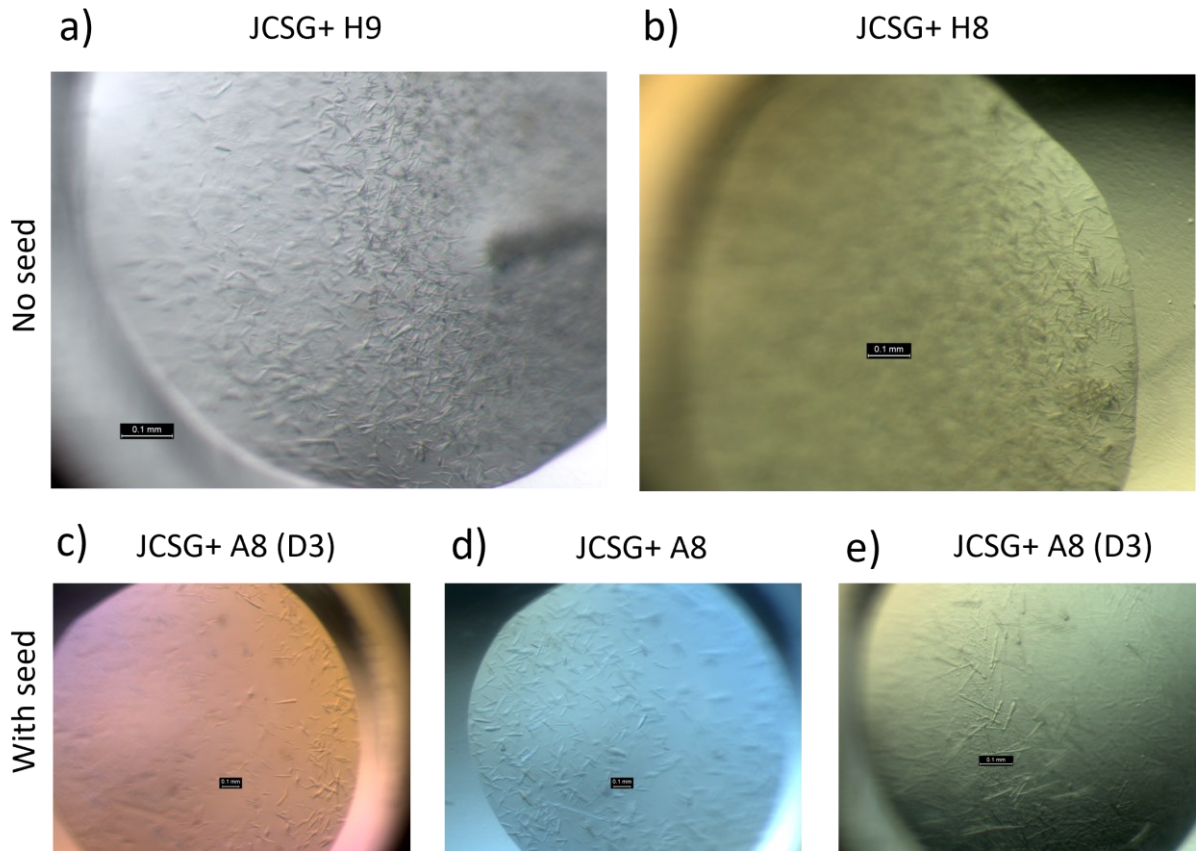


Figure 5. 4 CD1c-endo crystal image

Crystal images of CD1c-endo in 96 well plates set up by Oryx8 robot. CD1c-endo formed thin, needle-like crystals in the JCSG+ plate. Images a) and b) were taken from wells H8 and H9 respectively without seed stock. Crystals found in wells from images c) – e) were generated using seed stock from a) and b) at a ratio of 0.5:1:1 for seed:protein: mother liquor respectively. Crystals grew within three days and larger crystals were flash-frozen in liquid nitrogen for X-ray diffraction experiments.

Unfortunately, our crystallisation attempts did not produce diffraction-quality crystals, despite multiple optimisation efforts. As a result, we decided to pause standalone CD1c-endo crystallisation and instead wait until a γδ TCR production pipeline was established, allowing us to attempt co-crystallisation of γδ TCRs with CD1c-endo. A key objective of this study was to produce γδ TCRs that recognise CD1c, enabling us to crystallise the CD1c–γδ TCR complex and gain insight into the mechanism of interaction between γδ TCRs and CD1c. This mechanism remains unresolved, but recent structural studies of γδ TCRs in complex with CD1 and other antigen-presenting molecules suggest that the binding footprint is not conserved (52, 221).

5.2 Refolding and characterizing NM4 (αβ) TCR for functional studies

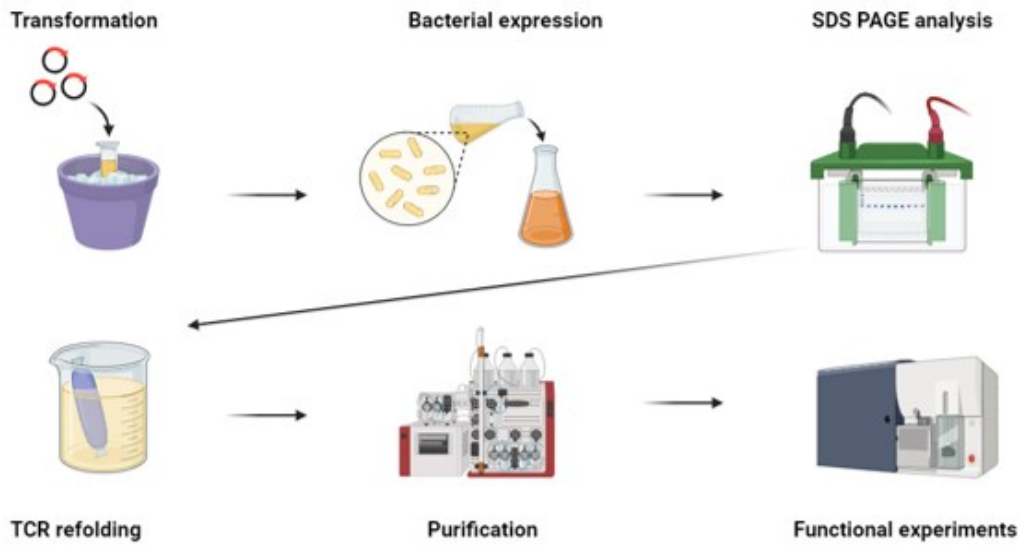


Figure 5. 5 Typical soluble TCR refolding pipeline

Schematic diagram of the pipeline set-up for refolding soluble TCRs. TCR single chain plasmid DNA are transformed into Rosetta *E. coli* and colonies are picked to be grown using TYP or AI media. Inclusion bodies generated through bacterial expression are refolded using a dialysis method, followed by multiple purification steps using anion exchange and size exclusion chromatography. Refolded TCRs can then be used for downstream functional experiments and structural studies.

We next wanted to generate CD1c-reactive γδ TCRs to use in structural and functional studies with CD1c-endo. Due to the recognised difficulty in generating soluble, refolded γδ TCRs, we initially set up and tested our pipeline using a CD1c-reactive αβ TCR, that we knew had been refolded before and recognised CD1c. Our rationale for this was to demonstrate the successful refolding of a soluble TCR that retained its function in binding to CD1c-endo, therefore providing reassurance that we could investigate the refolding and reactivity of a new panel of γδ TCRs and their interaction with CD1c-endo protein. This NM4 TCR, despite demonstrating the ability to bind CD1c, has never been crystallised in complex with it. Generation of this TCR would therefore allow for potential crystallographic studies in complex with the mammalian produced CD1c-endo, as well as provide a tool for optimising assays to test our panel of γδ TCRs further down the line.

The NM4 TCR used for the optimisation of the TCR refolding and purification pipeline, had previously been isolated from CD1c-SL tetramer studies. The NM4 TCR α and β chains were cloned into bacterial expression vectors for production via the bacterial expression system. We opted for bacterial expression of TCRs as our end goal was to pursue crystal structures of our refolded TCRs in complex with CD1c-endo. TCRs refolded in this way do not undergo the same post-translational modifications that mammalian TCRs do, therefore making them a more attractive target for co-crystal studies (222). Initially, we attempted to transform these TCR chains into Rosetta 2(*DE3*)*pLysS* cells (Merck) and express them as soluble inclusion bodies using TYP media.

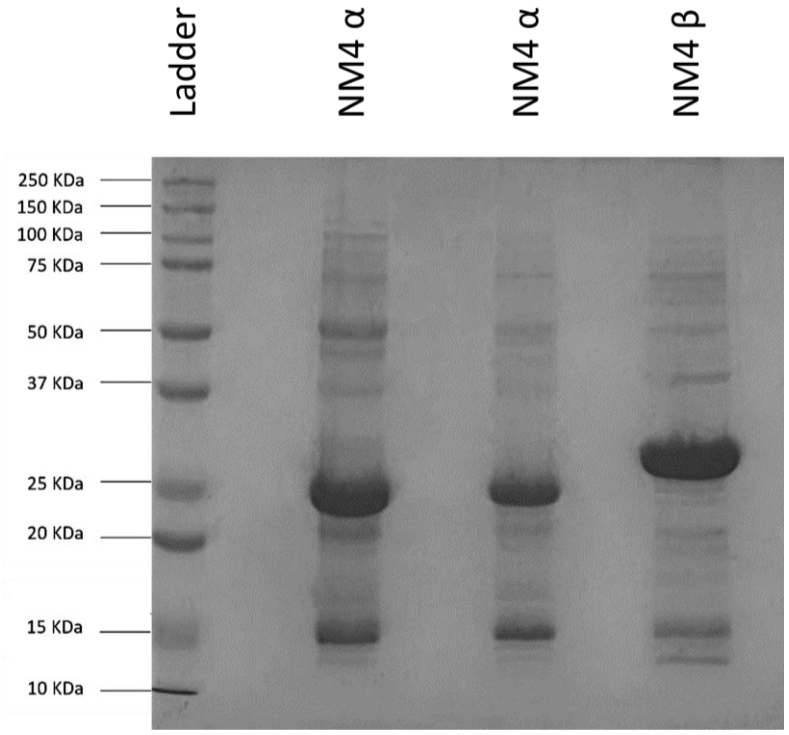


Figure 5. 6 SDS-PAGE gel showing inclusion body proteins for NM4 TCR chains

SDS-PAGE gel of NM4α and NM4β inclusion bodies. Inclusion bodies were generated separately using TYP (lanes two and four) or LSLB media (lane 3). 10µl of the inclusion body sample was incubated with 3µl of 4X reducing dye for ten minutes at 95°c. Inclusion body bands show correct molecular weight of approximately 24KDa and 30KDa for NM4α and NM4β respectively. The protein standards ladder from Bio-Rad was run in lane one for direct comparison of molecular weights.

Both TCR chains expressed efficiently, producing clear bands at the expected molecular weights (24 kDa for the α chain and 30 kDa for the β chain). The middle lane (NM4α) was generated using the same method as the first and last lanes, except LSLB media was used instead of TYP media. This

demonstrated that LSLB media is not optimal for bacterial TCR expression, as it resulted in reduced expression levels. TYP media, supplemented with 20% glucose and ampicillin, was used for bacterial culture inoculation. Although all three lanes contained low molecular weight contaminants, these were minor compared to the band of interest and were therefore not a concern.

Following successful expression, we proceeded with refolding the NM4 TCR using our standard refolding protocol. Briefly, the TCR chains were denatured in 6Mguanidine and 20 mM DTT, then refolded. Dialysis was carried out at 4°C overnight against cold water, followed by a second dialysis step the next night in Tris buffer (pH 8.1). Two 20 µL fractions were taken directly from the dialysis tubing and analysed by SDS-PAGE to assess refolding success. Properly refolded TCRs typically show a faint band at ~50 kDa in the non-reducing lane, with two bands in the reducing lane at approximately equal intensities, corresponding to the two single chains forming the heterodimer. SDS-PAGE analysis (data not shown) confirmed that the NM4 TCR met these criteria, allowing us to proceed with concentration and purification, as detailed in the Methods.

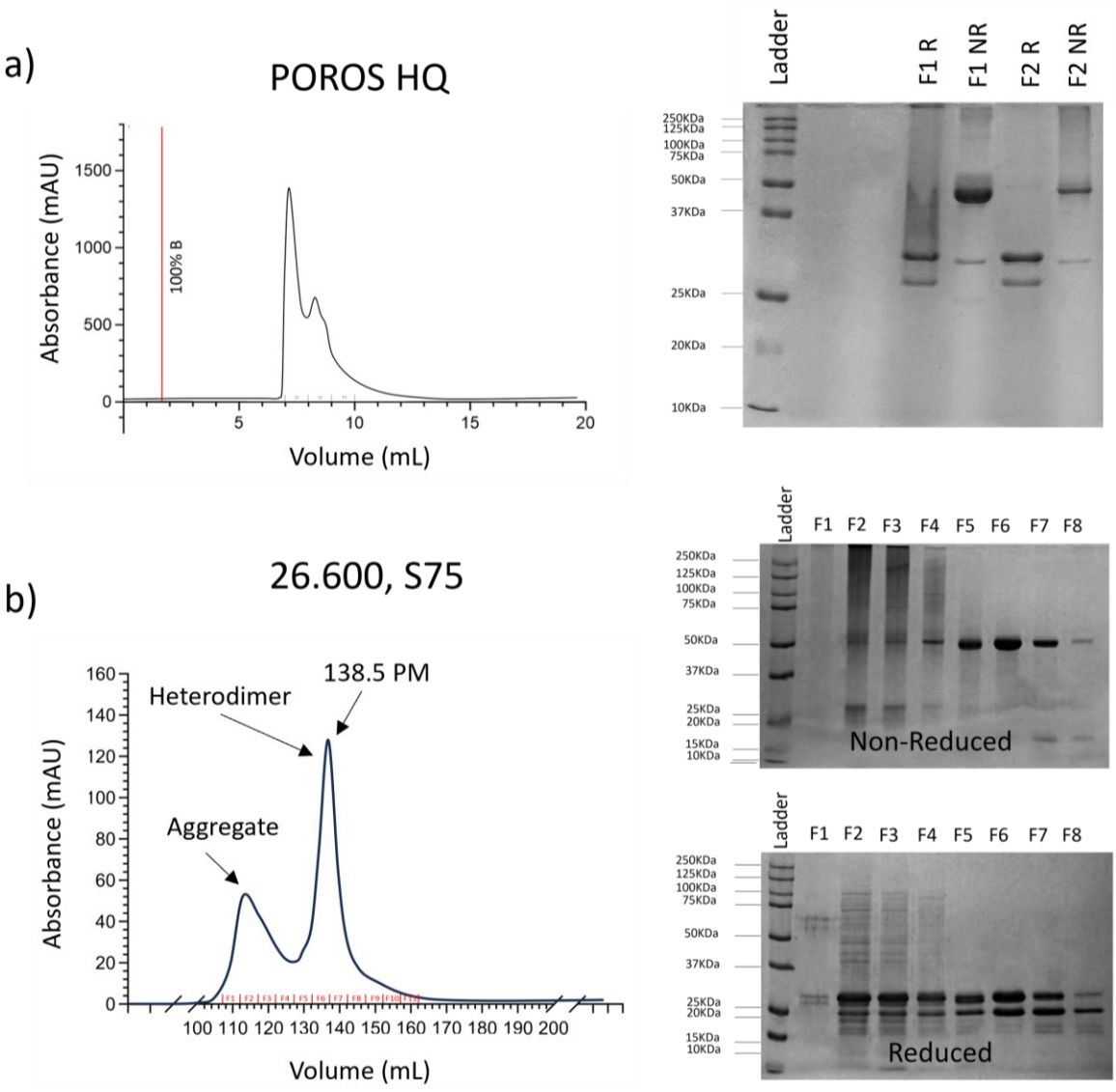


Figure 5. 7 Concentration and purification of αβ NM4 TCR

Ion exchange chromatography and size exclusion chromatography of NM4 TCR, with corresponding SDS-PAGE gels. a) Refolded NM4 TCR was concentrated using a POROS HQ 50 column using anion exchange chromatography. Dialysed TCR stuck to the column and was eluted using a one-step high salt elution (injection indicated by the straight red line). Fractions one and two were taken for SDS-PAGE and run under reducing and non-reducing conditions, with heterodimer in both non-reducing fractions (50KDa band) and two bands corresponding to a single chain in the reducing lanes. b) Peaks from the POROS elution were pooled and run on a 26.600 size exclusion column (*Cytiva*). Fractions across the peaks were analysed via SDS-PAGE with heterodimer eluting in the second peak with a 138.5ml peak max (denoted by the strong band at 50KDa in the non-reducing lanes F5-F7).

The NM4 TCR was concentrated and purified using ion exchange and gel filtration chromatography, yielding a properly refolded heterodimer with a strong 50 kDa band on SDS-PAGE. However, SDS-PAGE analysis of the POROS HQ elution revealed an excess of β chain in reducing conditions. The refolded TCR corresponded to the second peak on the chromatogram (peak max: 138.5 mL), with even ratios of single chains in the reducing lanes, confirming successful refolding.

To prepare the NM4 TCR for functional studies, we performed biotinylation, yielding 5 mg of biotinylated TCR and 5 mg of unbiotinylated TCR for structural and SPR studies. Excess biotin was removed, and biotinylation was confirmed via ELISA. The final protein was stored at -80°C in 50 µL aliquots (1 mg/mL) for downstream applications.

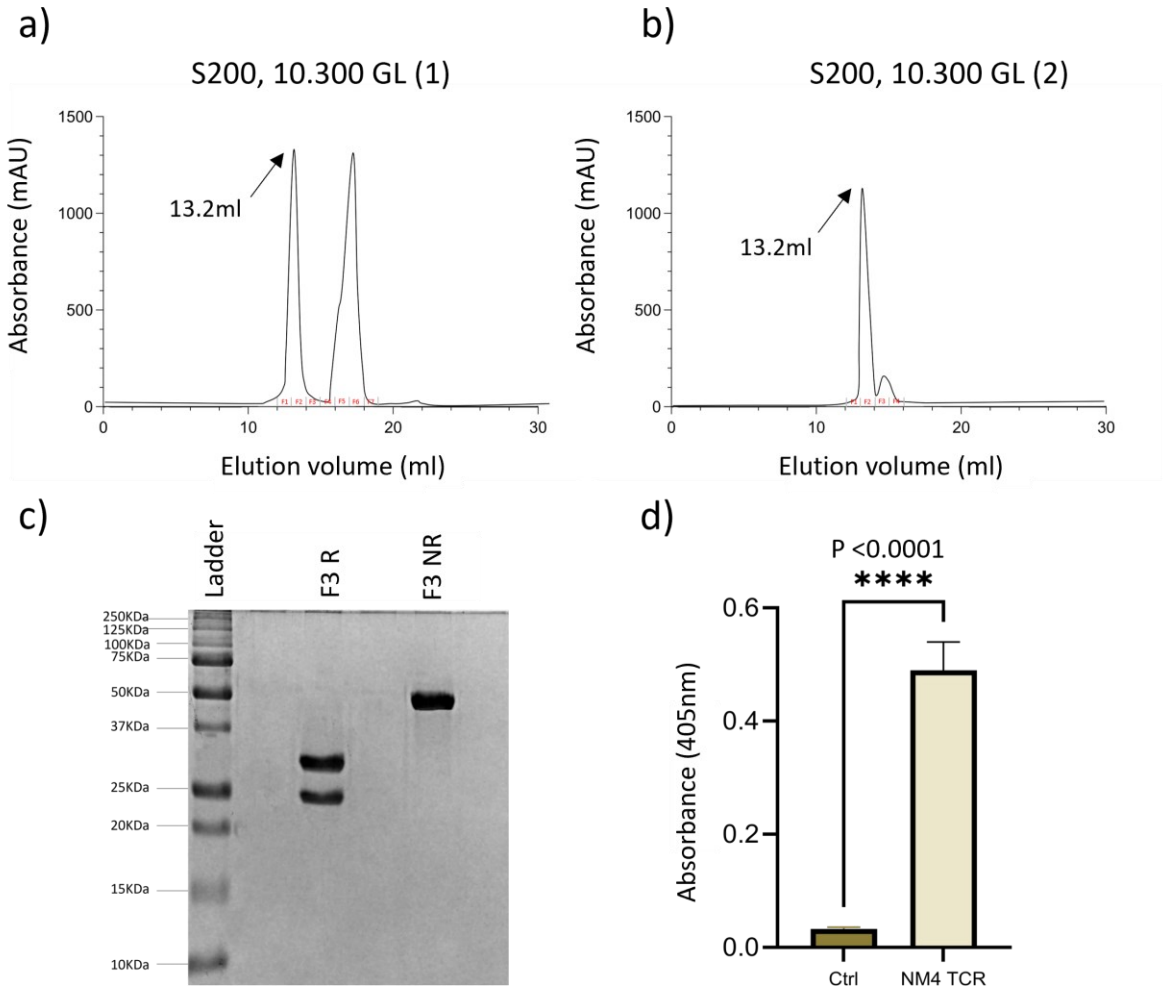


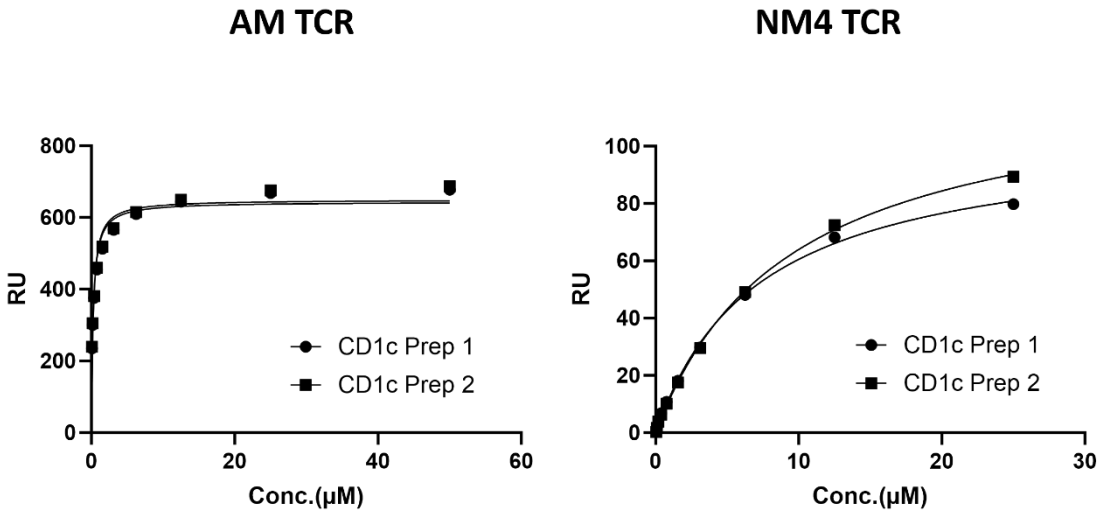
Figure 5. 8 Analysis of NM4 TCR post biotinylation

Purification, SDS-PAGE analysis, and Biotin ELISA of purified NM4 TCR post overnight biotinylation. a) NM4 TCR was purified using the S200, 10.300 GL size exclusion column (*Cytiva*). b) Protein was then re-purified using the same column after pooling fractions corresponding to the first peak on the initial chromatogram. A reducing and non-reducing fraction corresponding to the large sharp peak at 13.2ml peak max was run on SDS-PAGE. c) SDS-PAGE shows two bands at equal ratios corresponding to single chains in the reducing lane, and one band at 50KDa in the non-reducing lane corresponding to heterodimer. c) Biotin ELISA of NM4 TCR with non-biotinylated protein used as a positive control. Data is represented as the mean ± standard deviation for each sample (n=3). Background absorbance was removed from each sample by subtracting PBS-only control from each data point.

Confirmation that the NM4 TCR had been successfully refolded, biotinylated, and purified allowed us to proceed with testing its binding to CD1c-endo. To quantify this interaction, we performed SPR experiments using both our NM4 TCR and an affinity-matured NM4 TCR provided by collaborators at Immunocore. Two different batches of CD1c-endo were tested as ligands.

5.3 Optimising a pipeline to investigate TCR lipid reactivity.

5.3.1 SPR of NM4 TCR with CD1c



	CD1c Prep 1	CD1c Prep 2
KD (μM)	**0.253**	**0.249**
Bmax	643	650
R^2	0.9556	0.9533

	CD1c Prep 1	CD1c Prep 2
KD (μM)	**7.23**	**9.39**
Bmax	104	124
R^2	0.9986	0.9989

Figure 5. 9 NM4 TCR binds to CD1c-endo with modest affinity.

SPR binding curves of NM4 TCR and an affinity matured (AM) TCR, binding to two different batches of CD1c-endo protein. CD1c endo was coated to the chip and analyte (TCR) was injected over the chip using kinetic injections at a flow rate of 30μL/min. KD values were in the low μM range for the affinity-matured TCR as expected, and 7.23 and 9.39 for the NM4 TCR binding to Prep One and Prep Two respectively.

Strong binding of the affinity-matured NM4 TCR (AM TCR) to CD1c-endo was observed, with KD values in the expected range, confirming that the CD1c-endo protein was functional. The binding of our NM4 TCR was significantly lower, as anticipated, with KD values of 7.23 and 9.39 for the two different CD1c-endo batches. Despite binding to both CD1c-endo monomers coated on the SPR chip, the NM4 TCR tetramer did not stain CD1c-expressing THP1 cells effectively, with only a slight shift relative to the negative control (data not shown).

SDS-PAGE analysis of NM4 TCR purification *(Figure 5.8)* displayed clean bands and strong biotinylation, suggesting that protein quality was not an issue. However, low staining efficiency on THP1 cells indicated that alternative factors might be influencing binding. To explore this further, we implemented a bead-based binding assay as a more sensitive alternative to the live cell-based system.

Biotinylated CD1c-endo was loaded onto MACSI beads, as described in Methods (2.5.2), and incubated with NM4 TCR tetramer to assess its binding capacity relative to a non-binding irrelevant tetramer control. To optimise this approach, a titration series was performed to determine the optimal NM4 TCR tetramer concentration for FACS analysis.

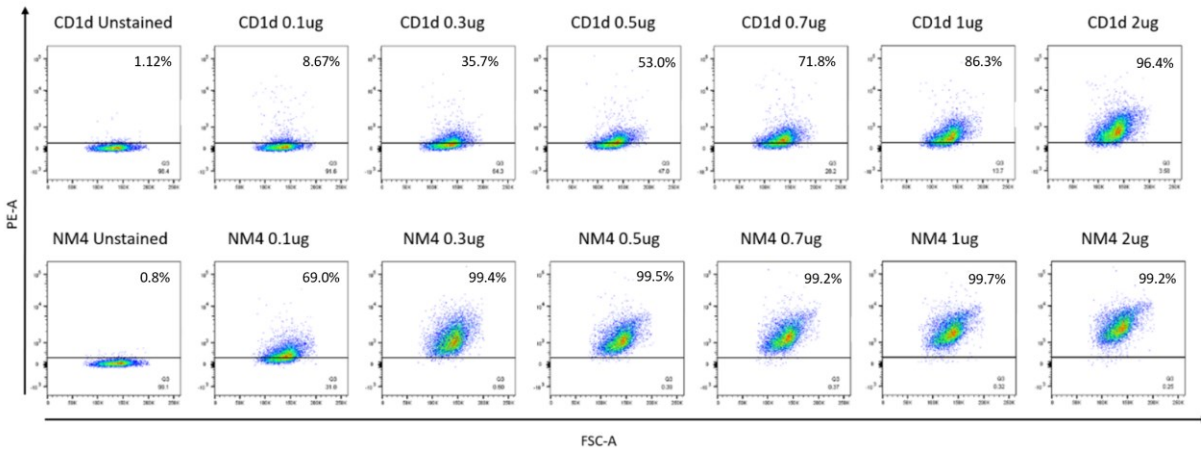


Figure 5. 10 NM4 TCR stains CD1c-endo loaded MACSI beads.

Tetramer titration to investigate the optimal concentration of TCR tetramer required to stain CD1c-endo MACSI beads. CD1c-endo tetramer was added to 100,000 MACSI beads in separate tubes at concentrations ranging from 0.1µg to 2µg. An irrelevant CD1d tetramer was used as a negative control. Data shows 0.3µg as the optimal tetramer concentration for bead staining.

Titration of NM4 TCR tetramer and the irrelevant CD1d tetramer control identified 0.3 µg of NM4 TCR tetramer as the optimal concentration, yielding the highest signal-to-background ratio between relevant NM4 TCR staining and irrelevant CD1d tetramer staining. With this optimised concentration, we proceeded to refine the assay by investigating methods to remove endogenous lipids from the CD1c-endo binding groove. Removing bound lipid would allow specific lipids to be introduced into the groove, enabling an assessment of lipid specificity for NM4 and the γδ TCR panel we aim to generate.

To achieve this, we tested a citrate-CHAPS detergent buffer for lipid removal, based on a combination of published literature (223, 224), and previous in-house data suggesting CHAPS may facilitate lipid

exchange within the CD1c binding groove. Three different CHAPS/citrate buffer concentrations were tested at pH 7.5, as previous data suggested this may be the optimal pH for lipid removal.

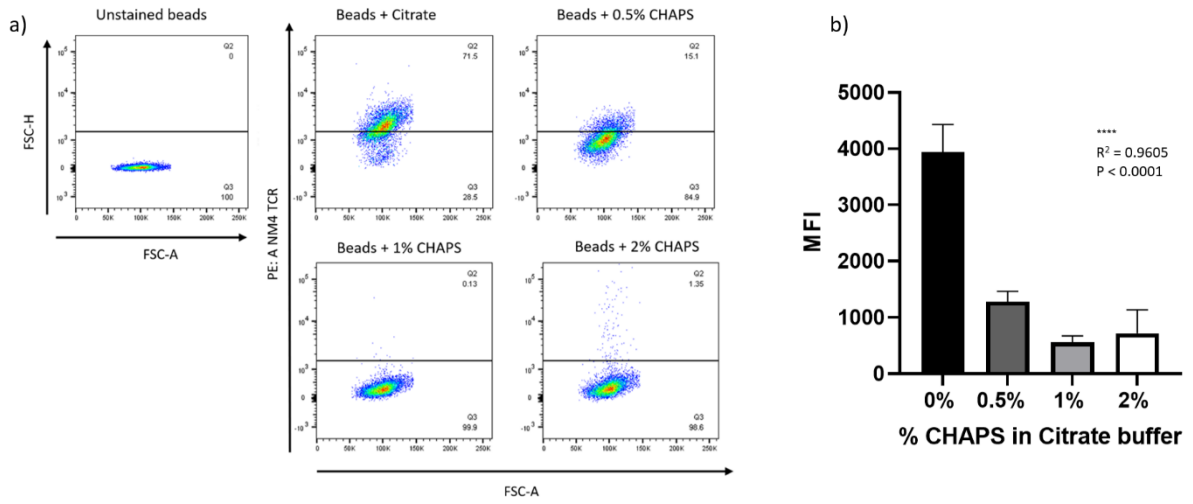


Figure 5. 11 Increasing CHAPS concentration reduces CD1c staining by NM4 TCR

CD1c-endo coated MACSI beads were stained with NM4 TCR tetramer following incubation with CHAPS-citrate buffer containing different CHAPS concentrations. a) FACS plots of NM4 treated MACSI beads incubated with citrate buffer only, and citrate + 0.5%, 1%, and 2% CHAPS. b) The bar graph shows triplicate MFI values calculated using the geometric mean for each concentration of CHAPS buffer with the best reduction of staining achieved using 1% CHAPS in citrate buffer.

Incubation of CD1c-endo coated beads in citrate buffer alone resulted in strong NM4 TCR tetramer staining, indicating that citrate alone was insufficient for removing endogenous lipids from the CD1c groove. Increasing the CHAPS detergent concentration to 0.5%, 1%, and 2% led to a progressive reduction in NM4 TCR staining, suggesting that CHAPS facilitated lipid removal from the CD1c groove.

To ensure that reduced staining was due to lipid removal rather than CHAPS interference, beads were extensively washed before staining. We further explored the optimal pH range for CHAPS/citrate buffer, testing pH 4.5 to 7.5, as the ideal pH for lipid exchange in vitro remained uncertain. Given that CD1c molecules survey the intracellular system to exchange endogenous lipids, we aimed to identify conditions that best replicate this process (34, 225).

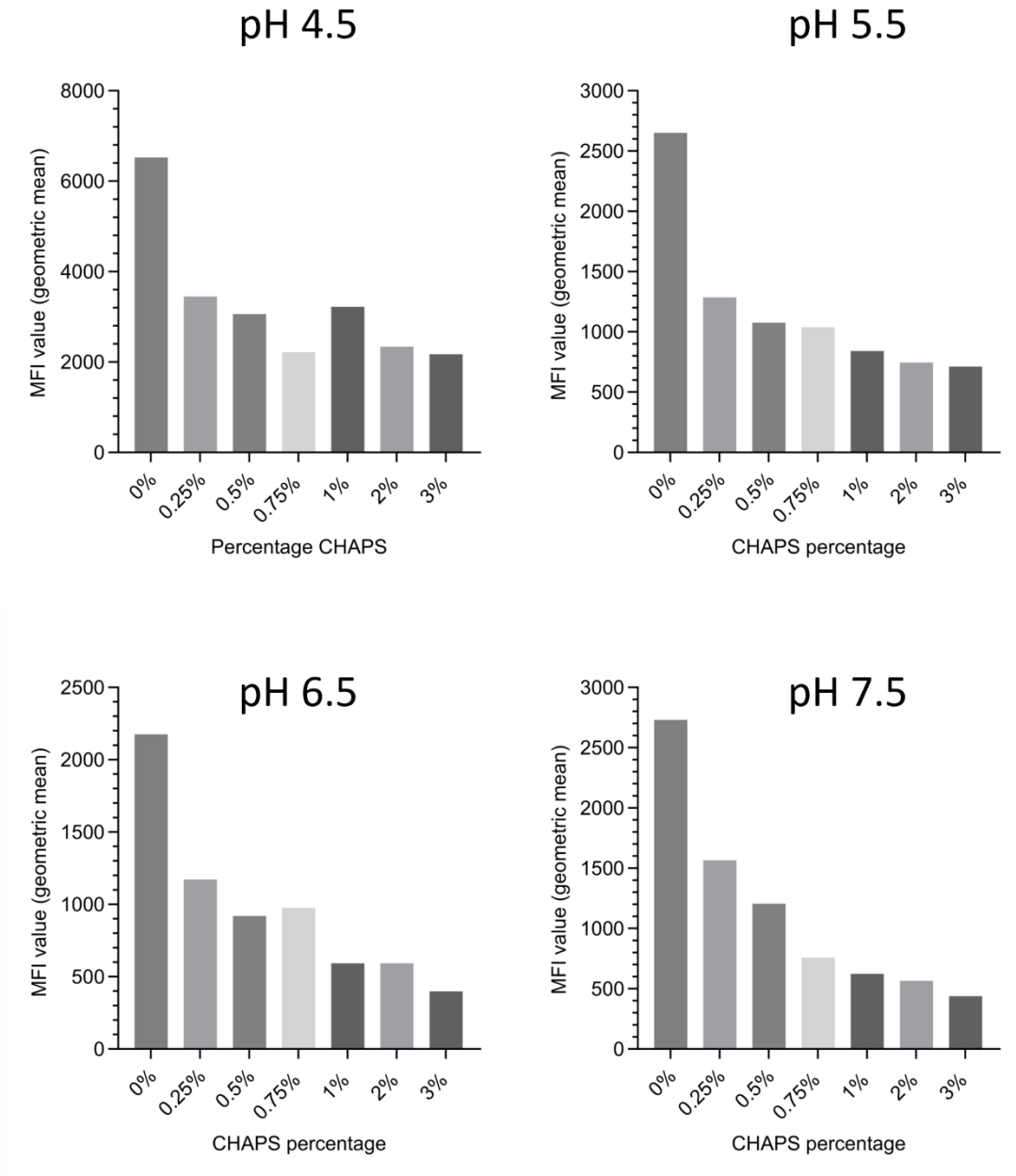


Figure 5. 12 Increasing CHAPS concentration reduces CD1c-endo staining by NM4 TCR

Graphs depicting the reduction in staining of CD1c-endo coated MACSI beads by NM4 TCR tetramer when incubated with increasing concentrations of CHAPS. The experiment was carried out at four x different pH values and data is representative of at least two repeats at each pH. 0% CHAPS consisted of solely citrate buffer at the pH referred to in each respective graph. All four pH values show a general downward trend as the concentration of CHAPS increases. Results representative of two independent experiments per pH.

All four pH values demonstrated a general decrease in binding and subsequent staining of the NM4 TCR to the CD1c-endo-coated MACSI beads. As well as reducing the staining by removing the bound endogenous lipid, we also wanted to incubate with an excess of specific lipid to replace the endogenous lipid, so we could then investigate lipid reactivity of the TCR. Initially, we attempted this using two lipids that we have previously demonstrated to be recognised when presented by CD1c (unpublished data from our lab). These lipids are both relatively soluble as they are not extreme hydrophobes and are readily available and are therefore ideal for optimisation experiments.

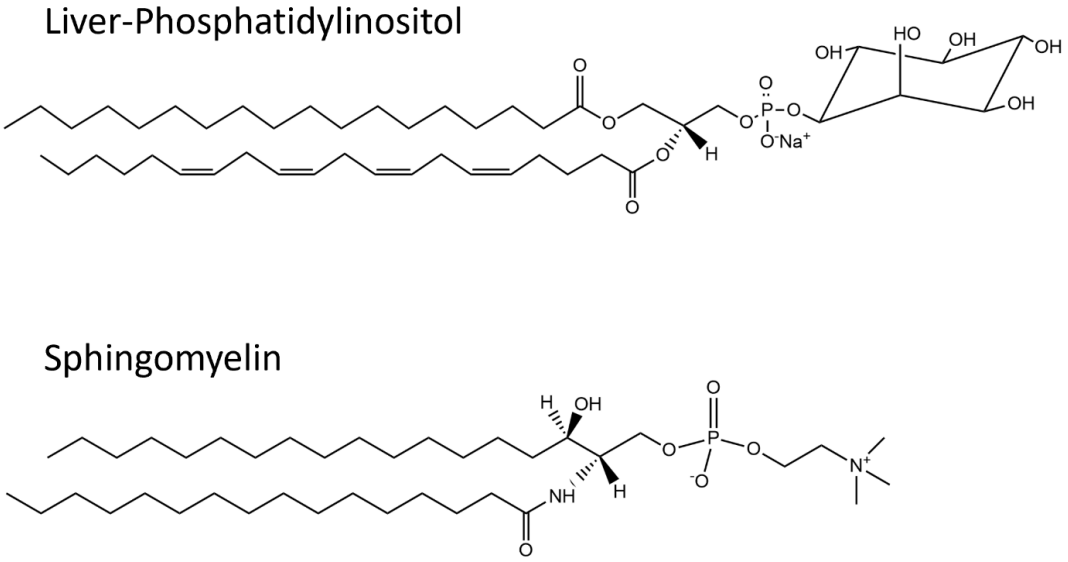


Figure 5. 13 Chemical structures of liver PI and SM

Liver phosphatidyl inositol and SM both contain two long hydrocarbon chains, with liver PI also containing a bulky alcohol head group. Chemical structures were generated using ChemDraw 20.1 and are representative of the structures ordered from Avanti polar lipids, under codes 840042 and 860061 for liver PI and SM respectively.

Based on a combination of previous experiments in our lab, and the data shown in *Figure 5.11*, we decided to investigate the lipid exchange between endogenous lipids and excess pulsed lipids, using 1% and 2% CHAPS/citrate as our exchange buffer. Before pulsing, lipids were sonicated at 37°c for one hour, followed by overnight incubation in either 1% CHAPS citrate buffer, or 2% CHAPS citrate buffer. We optimised this using liver PI due to its relatively high solubility, at a range of 0.1µg to 10µg.

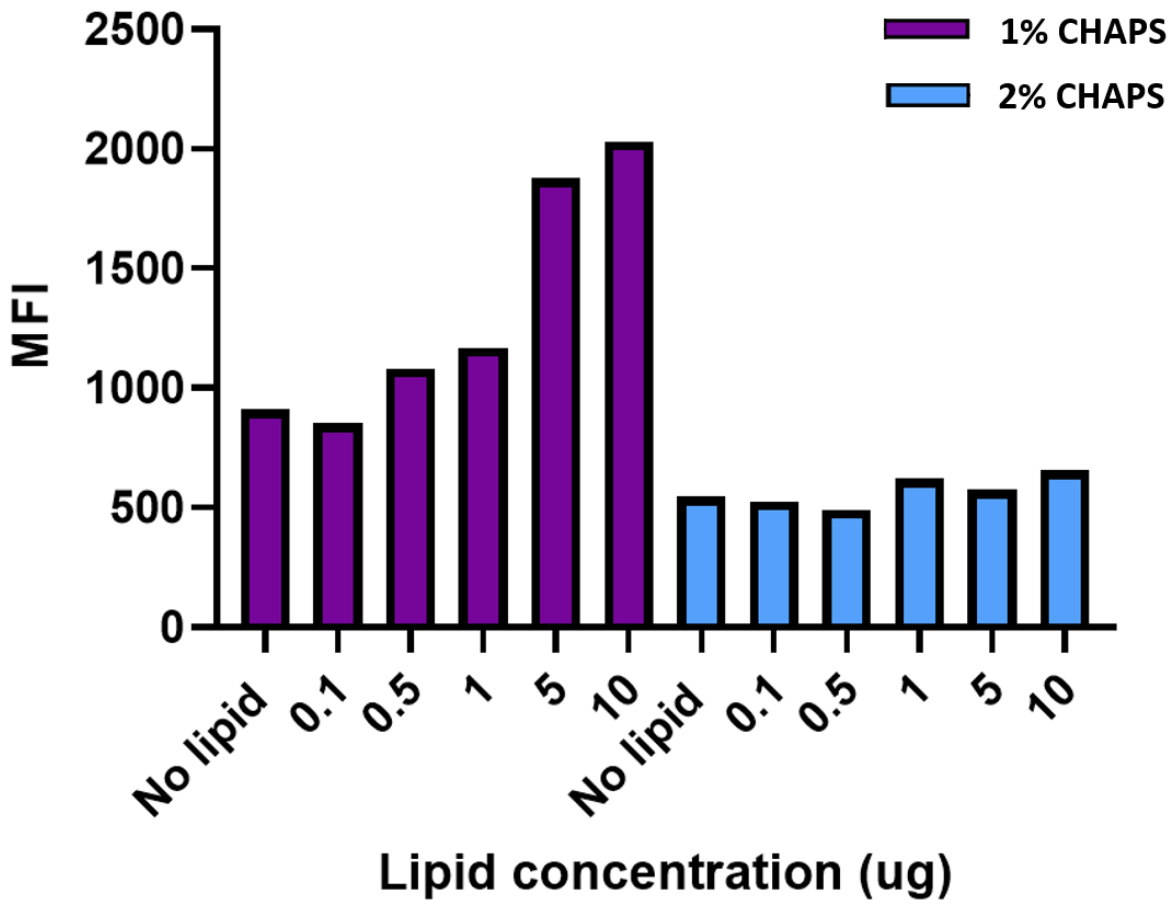


Figure 5. 14 Liver PI is successfully loaded at 1% CHAPS buffer concentration.

Bar graph demonstrating reconstitution of NM4 TCR staining when incubating CD1c-endo with different concentrations of solubilized PI lipid, at two different CHAPS concentrations. Liver PI was incubated with CD1c-endo coated MACSI beads at concentrations ranging from 0.1-10µg, at two different CHAPS concentrations (1% and 2%). Lipid loaded in 1% CHAPS buffer (purple) could successfully reconstitute NM4 TCR staining, whereas 2% CHAPS did not have this effect.

The data indicate that 1% CHAPS buffer was the most effective condition for reconstituting NM4 TCR binding to pulsed CD1c-endo. While additional concentrations were tested, higher and lower CHAPS levels were less successful (data not shown).

Given our broader aim of investigating CD1c reactivity by γδ TCRs, and the apparent success of our assay in removing endogenous lipids and replacing them with specific lipids, we decided to pause NM4 lipid reactivity testing until we had established our panel of γδ TCRs.

In parallel with optimising the bead-based assay, we also aimed to develop a plate-bound assay for assessing lipid reactivity once our γδ TCR panel was available. This alternative approach would provide

a different readout, enabling the measurement of Jurkat T cell activation upon incubation with CD1c-endo protein loaded with specific lipids.

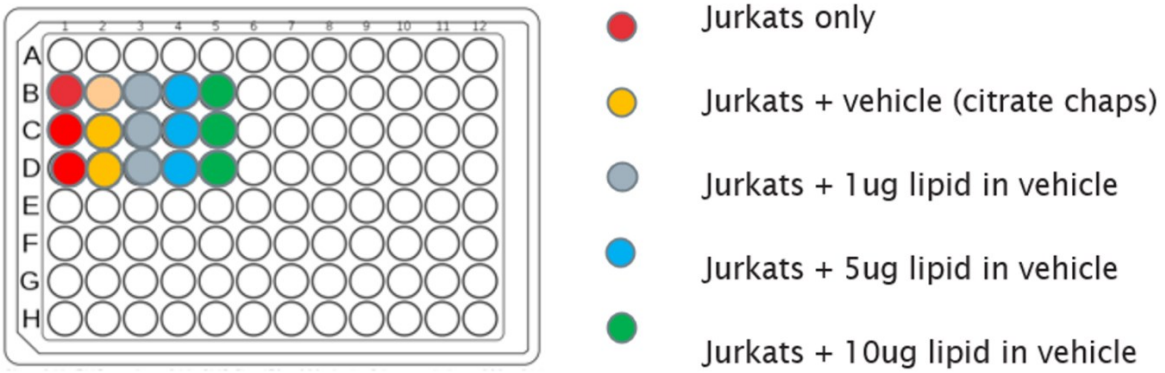


Figure 5. 15 Schematic diagram of plate-bound activation marker assay.

Representation of the 96 well plate set-up for CD1c-endo lipid pulsing. Wells were set up in triplicate with two control wells (Jurkat only and Jurkat + Vehicle). Three different concentrations of lipid were used in triplicate to attempt to obtain a dose response in terms of CD69 expression following incubation with lipid and protein.

This assay investigated the surface activation marker expression (CD69) on Jurkat T cells expressing a CD1c-reactive αβ TCR when incubated on CD1c-endo coated plates. The rationale was to run this assay in parallel with the bead assay to correlate the tetramer-based bead assay results with a functional readout, measured by CD69 expression.

For this, CD1c-endo protein was coated onto the plate and treated with either citrate buffer alone (control) or citrate-CHAPS buffer, followed by incubation with different sonicated lipids. The pulsed CD1c-endo proteins were then incubated with Jurkat cells, and CD69 expression was measured as an indicator of functional activation.

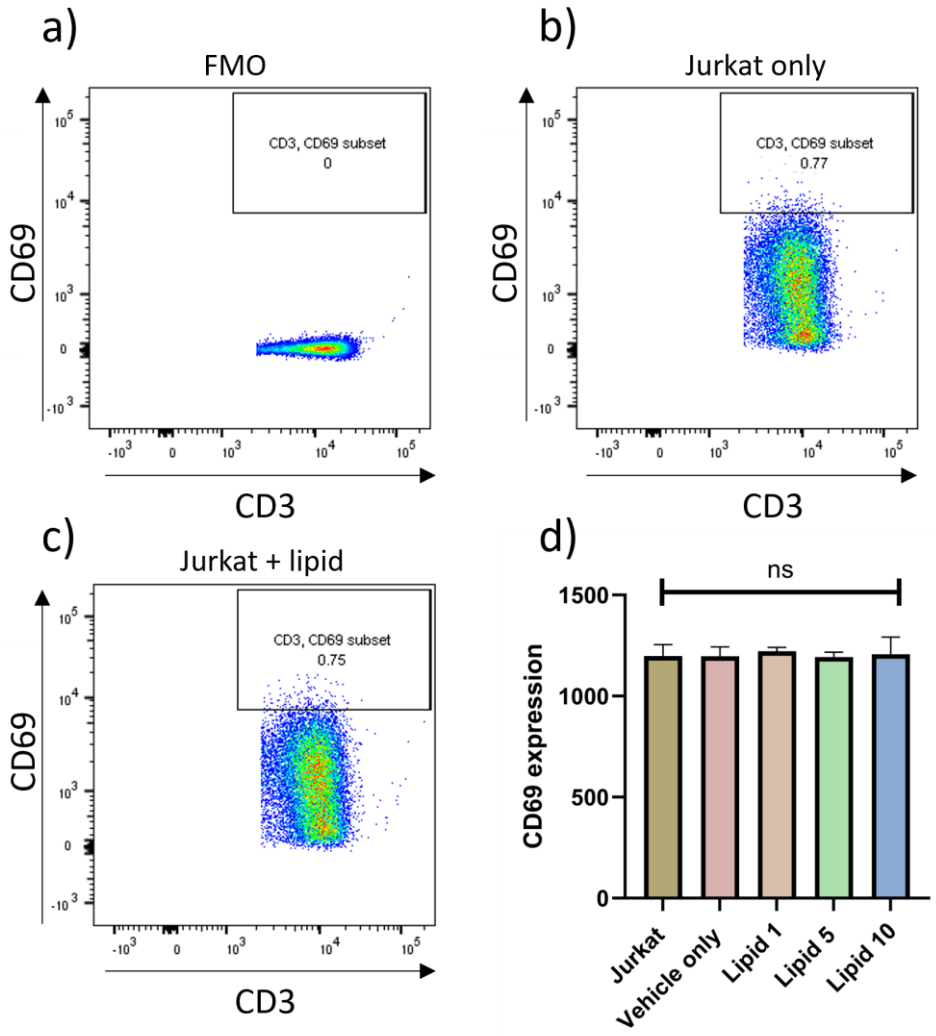


Figure 5. 16 NM4 Jurkats are activated in the absence of antigens.

a), b), and c) show representative flow cytometry dot plots of CD69 expression, both in the absence and presence of plate-bound CD1c-endo protein. d) shows a bar graph of CD69 expression of all measured conditions, including Jurkat only, vehicle only, and pulsed CD1c-endo wells at three different lipid concentrations. Each bar is representative of triplicate measurements and is non-significant following an ordinary one-way ANOVA, (P = 0.95).

Unexpectedly, strong upregulation of CD69 activation was observed in the absence of plate-bound CD1c-endo protein (Jurkat-only condition), suggesting that JRT3.5 Jurkat cells may express CD1c or another activating molecule on their surface *(Figure 5.17)*. This was evident in the control well, which showed high CD69 expression despite not being incubated with CD1c-endo protein, indicating that activation occurred independently of added antigen recognition.

As a result, we were unable to detect a further increase in CD69 activation following incubation of Jurkat cells with lipid-pulsed, CD1c-endo-coated wells at varying lipid concentrations. Given that CD1c expression has been reported in Jurkat T cells, this suggests that the Jurkat cells were activating themselves via CD1c-NM4 TCR interactions. To address this issue, we generated β2M knockout Jurkat T cells (JRT3.5-β2mKO) to eliminate CD1c expression and prevent self-activation.

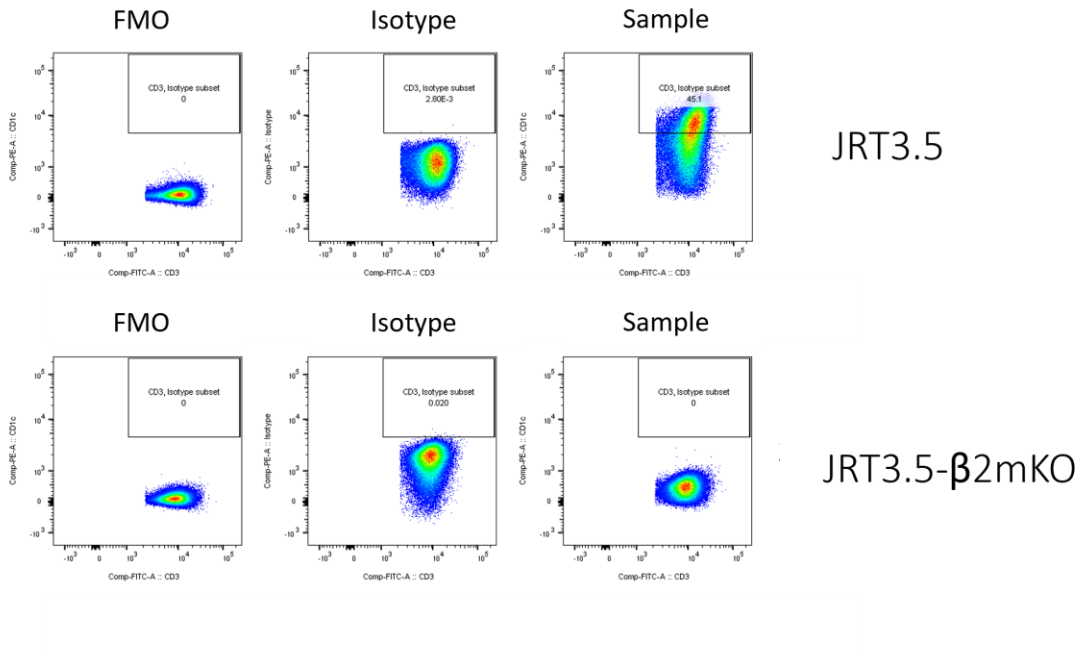


Figure 5. 17 CD1c-knockout Jurkats don't stain with anti-CD1c.

Dot plots of JRT3.5 Jurkats and JRT3.5-β2mKO Jurkat staining, comparing CD1c expression on the Jurkat cell surface. Data demonstrates that wild-type Jurkat T cells express CD1c while JRT3.5-β2mKO do not.

The lack of staining of our newly generated JRT3.5-β2mKO line with anti-CD1c mAB relative to the parental JRT3.5 line, suggests successful deletion of CD1c on this cell line. This will therefore provide a platform to investigate TCR autoreactivity of CD1c-restricted TCRs in those CD1c-KO Jurkats.

5.4 Refolding γδ TCRs for structural studies

5.4.1 Refolding and characterization of B4, hu20, and B10 γδ TCRs

Despite the knowledge that CD1c is a ligand for γδ TCRs, the molecular binding mechanisms have yet to be structurally determined. We aimed to investigate this interaction through structural studies using soluble TCRs. To produce these TCRs, we chose to utilise the bacterial expression system with oxidative refolding rather than the mammalian expression system (Expi-HEK or Expi-CHO). This decision was based on the fact that we already had an established in-house TCR refolding pipeline, and that the mammalian system introduces complications such as post-translational modifications (PTMs), including glycosylation.

One of our overarching goals was to generate structural data of CD1c-endo in complex with γδ TCRs, making production yield a key factor. TCRs produced via bacterial expression do not contain PTMs, making them immediately suitable for structural studies without requiring enzymatic deglycosylation, which could further reduce yield due to protein instability and additional purification steps.

Previous work within our lab led to the discovery and isolation of three potentially CD1c-reactive γδ TCRs. These were isolated from donor PBMCs using CD1c-SL tetramers, in a study conducted by a former PhD student. The TCR sequences were obtained through CD1c-SL tetramer-guided sorting (in collaboration with colleagues at King's College London) and were cloned into bacterial expression vectors via GenScript. These TCRs had already been cloned before the initiation of this PhD project, so we used them as a starting point for attempting the generation of stable γδ TCRs.

Table 5 Sequences of original γδ TCRs isolated using CD1c-SL tetramer guided sorting from donor PBMCs, in collaboration with colleagues at Kings College, London.

TCR	Vγ chain sequences	Vδ chain sequence
B4	HMSSNLEGRTKSVIRQTGSSAEITCDLAEGSTGYIHWYLHQEGKAPQRLLYYDS YTSSVVLESGISPGKYDTYGSTRKNLRMILRNLIENDSGVYYCATWDGPNYYKK LFGSGTTLVVTEDLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYPDHVELS WWVNGKEVHSGVCTDPQPLKEQPALNDSRYALSSRLRVSATFWQDPRNHFR CQVQFYGLSENDEWTQDRAKPVTQIVSAEAWGRADGSGGGLNDIFEAQKIE WH	HMAQKVTQAQSSVSMPVRKAVTLNCLYETSWWSYYIFWYKQLPSKEMIFLIRQGSD EQNAKSGRYSVNFKKAAKSVALTISALQLEDSAKYFCALGELDRAWGSGYTDKLIFGKG TRVTVEPIQNPDPAVYQLRDSKSSDKSVCLFTDFDSQTNVSQSKDSDVYITDKCVLDM RSMDFKSNSAVAWSNKSDFACANAFNNSIIPEDTFFPSPESS-
B6	HMSSNLEGRTKSVIRQTGSSAEITCDLAEGSTGYIHWYLHQEGKAPQRLLYYDS YTSSVVLESGISPGKYDTYGSTRKNLRMILRNLIENDSGVYYCATWDGYKKLFGS GTTLVVTEDLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYPDHVELSWW VNGKEVHSGVCTDPQPLKEQPALNDSRYALSSRLRVSATFWQDPRNHFRCQV QFYGLSENDEWTQDRAKPVTQIVSAEAWGRADGSGGGLNDIFEAQKIEWH	HMAQKVTQAQSSVSMPVRKAVTLNCLYETSWWSYYIFWYKQLPSKEMIFLIRQGSD EQNAKSGRYSVNFKKAAKSVALTISALQLEDSAKYFCALGPPLFYVLGYRKLIFGKGTRV TVEPIQNPDPAVYQLRDSKSSDKSVCLFTDFDSQTNVSQSKDSDVYITDKCVLDMRSM DFKSNSAVAWSNKSDFACANAFNNSIIPEDTFFPSPESS-
B10	HMSSNLEGRTKSVTRPTGSSAVITCDLPVENAVYTHWYLHQEGKAPQRLLYYD SYNSRVVLESGISREKYHTYASTGKSLKFILENLIERDSGVYYCATWDVTYPTGW FKIFAEGTKLIVTSPEDLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYPDHV ELSWWVNGKEVHSGVCTDPQPLKEQPALNDSRYALSSRLRVSATFWQDPRN HFRCQVQFYGLSENDEWTQDRAKPVTQIVSAEAWGRADGSGGGLNDIFEAQ KIEWH	HMAQKVTQAQSSVSMPVRKAVTLNCLYETSWWSYYIFWYKQLPSKEMIFLIRQGSD EQNAKSGRYSVNFKKAAKSVALTISALQLEDSAKYFCALGELGVSYMTGGYTDKLIFGK GTRVTVEPIQNPDPAVYQLRDSKSSDKSVCLFTDFDSQTNVSQSKDSDVYITDKCVLD MRSMDFKSNSAVAWSNKSDFACANAFNNSIIPEDTFFPSPESS-

We set up our γδ TCR refolds following the same template and experimental design previously used for the NM4 TCR refold. We initially focused on the B4 TCR (Vγ4Vδ1), starting with the inclusion body preparation in AI media to ensure high protein yield. The inclusion bodies were relatively clean and had molecular weights consistent with the protein standard ladder (25 kDa and 30 kDa for δ and γ chains, respectively).

We proceeded with refolding the inclusion bodies at an initial 1:1 ratio, following incubation at 42°C for 30 minutes in the presence of 20 mM DTT. The refolded protein was then dialysed against 20 mM Tris (pH 8.1), concentrated using a POROS ion exchange column, and eluted with 1 m NaCl buffer via single-step elution.

SDS-PAGE analysis of the eluted fractions showed a 50 kDa band in the non-reducing lane, corresponding to the dimer, while three separate reducing fractions displayed two distinct bands, corresponding to the single chains that had been denatured by DTT in the reducing dye.

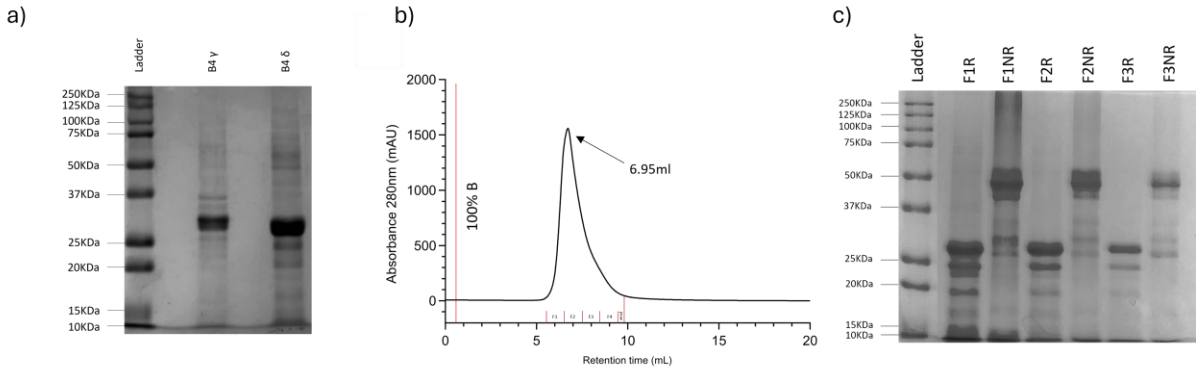


Figure 5. 18 B4 TCR Inclusion Bodies and Initial Purification

SDS-PAGE analysis of inclusion bodies and initial purification (and concentration) of the B4 γδ TCR via anion exchange chromatography. (a) SDS-PAGE gel showing B4 γ and δ inclusion body chains, with the molecular weight protein standard ladder on the left. (b) Chromatogram displaying the initial purification and refold concentration of the B4 TCR using a POROS HQ anion exchange column. The dialysed protein was eluted via a one-step high-salt elution protocol, with eluted fractions analysed by SDS-PAGE. (c) SDS-PAGE analysis of fractions collected under reducing (R) and non-reducing (NR) conditions. A band at ~50 kDa, corresponding to the correctly folded γδ TCR heterodimer, is visible in all non-reducing fractions (NR). Under reducing conditions (R), two distinct bands are observed, corresponding to the γ and δ single chains.

The entire POROS elution peak was concentrated using a VivaSpin 50 column with a 10 kDa molecular weight cut-off (MWCO). The concentrated sample was then loaded onto a Hi-Prep 26/60 size exclusion/gel filtration column at a volume of <5 mL and run at 2.4ML/min, as recommended by Cytiva. Fractions were collected and analysed by SDS-PAGE to assess purity.

SDS-PAGE analysis confirmed the presence of a stably refolded B4 heterodimer in fractions F3–F7, corresponding to a large, sharp peak with a peak max of 135.6. The heterodimer retention time was comparable to our control αβ NM4 TCR, and SDS-PAGE bands appeared relatively clean, though a higher molecular weight contaminant was observed in fractions F5 and F6.

Since we intended to use the γδ TCRs for both tetramer-based assays (MACSI bead assay) and structural studies, the TCR sample was split into fractions, with a portion used for biotinylation (shown in *Figure 5.19b*). Biotinylation was carried out overnight using a Tebu-Bio kit (see Methods). The subsequent purification resulted in a sharp, homogeneous peak at the same retention time as the control NM4 TCR, with SDS-PAGE analysis confirming clean protein bands at the expected molecular weights in both reducing and non-reducing lanes.

A second peak was collected and analysed by SDS-PAGE but did not show any protein bands, likely corresponding to the ligase enzyme or residual components from the biotinylation kit. Final fractions were stored at -80°C for later use.

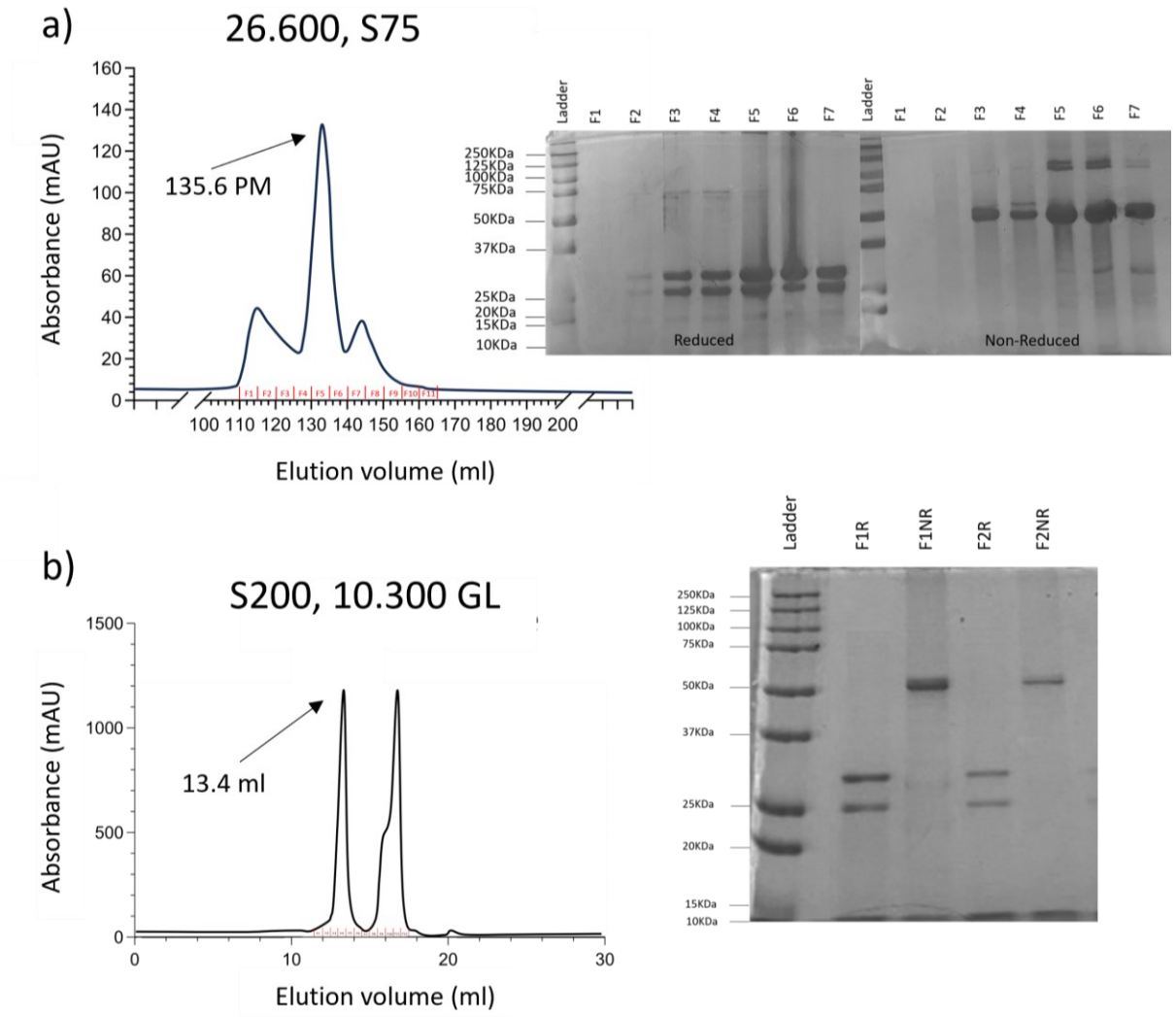


Figure 5. 19 Purification of B4 γδ TCR

a) size exclusion chromatography chromatogram, with purification performed on a Hi-load S75, 26.600 column. The resulting fractions were run on SDS-PAGE under reducing and non-reducing conditions. Analysis shows a strong 50KDa band in the non-reducing lanes corresponding to fractions F3-F7 and reducing lanes indicating a heterodimer due to an even ratio of single-chain bands. b) F3-F7 were combined, biotinylated, and re-purified using an analytical S200 10.300 GL column. Protein collected from the initial peak and run under reducing and non-reducing conditions via SDS-PAGE shows clean, purified protein with heterodimer band in the non-reducing lane at ~50KDa and monomers in the reducing lane at ~24KDa and ~30KDa for δ and γ respectively.

Following the successful refolding of our B4 TCR, we continued to refold the B10 TCR, using the same refolding procedure.

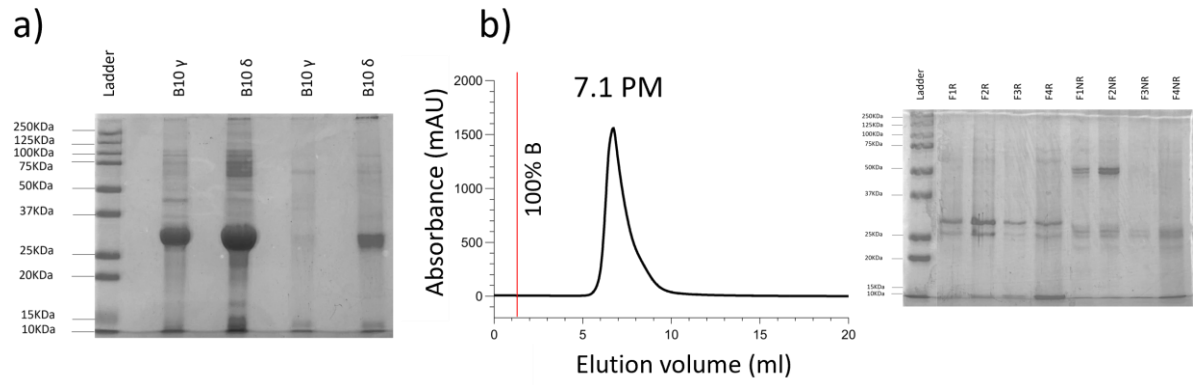


Figure 5. 20 Initial purification of B10 γδ TCR

B10 TCR inclusion bodies and initial purification of refolded TCR. a) Inclusion bodies generated using AI media (Lanes 2+3) compared to LSLB media. AI media produces stronger, cleaner bands compared to LSLB, with all bands appearing at the rough correct molecular weight relative to the protein ladder. b) TCR was initially purified and concentrated via the POROS HQ column, using a one-step high salt concentration elution. Fractions 1-4 were run separately on a reducing and non-reducing SDS-PAGE gel. Data suggests that F1 + F2 contain TCR heterodimer.

The B10 TCR inclusion bodies expressed efficiently in E. coli when using AI media (lanes 2 and 3) but failed to express in TYP media (lanes 4 and 5). The reason for this discrepancy is unclear but may be due to protein induction timing, as AI media induces expression at the peak of the growth phase, whereas TYP media requires manual IPTG induction at OD 0.4–0.6.

The B10 refold was initially performed using a 1:1 inclusion body ratio, but poor yield, indicated by weak bands in the non-reducing SDS-PAGE lanes (data not shown), prompted an optimised refold strategy. This included additional clean-up steps, high-speed centrifugation, and filtration of inclusion bodies before refolding. These adjustments resulted in a higher refold yield, as demonstrated by the gel in *Figure 5.20b*, where strong bands are visible in the F1NR and F2NR lanes.

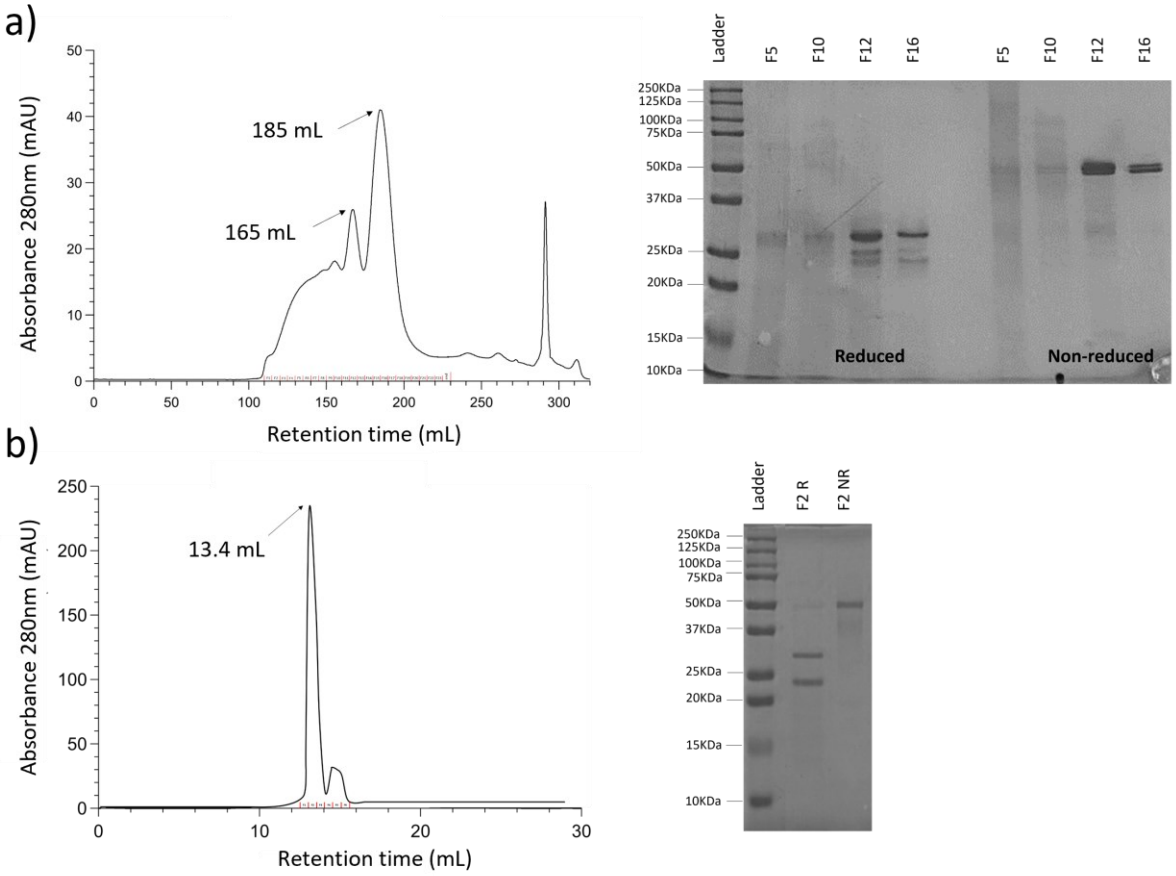


Figure 5. 21 Two-step purification of B10 γδ TCR

Chromatograms and SDS-PAGE gels of B10 γδ TCR purification. a) B10 was initially purified using a Hi-prep 26.600, S200 gel filtration column, with fractions collected and run on SDS-PAGE. F5, F10, F12, and F16 were run on the SDS-PAGE gel, with the peak at 165ml corresponding to most of the heterodimer. b) Fractions F10-F14 were pooled, concentrated, and run on the analytical S200, 10.300 GL gel filtration column. The resulting chromatogram yielded a sharp peak corresponding to heterodimer with peak max at 13.4ml, followed by a small single chain peak. SDS-PAGE analysis of the large peak shows clean, purified TCR at the correct molecular weight.

Fractions were concentrated to <5 mL and run on the Hi-Prep S26/600, S200 column at 2.4ML/min. A different column was used compared to the B4 TCR refold, as the S75 column was unavailable at the time. This accounted for the later retention time and peak max of 165 mL, compared to 138ML for the B4 TCR. SDS-PAGE analysis was performed to compare individual peaks and their corresponding molecular weights.

Based on previous protein standard runs on the S200 column, the expected retention time of a 55 kDa protein (TCR heterodimer) was around 160–170 mL. However, an unexpected large peak was

observed following this, corresponding to a single chain, suggesting that this TCR may be less stable than the previously refolded NM4 and B4 TCRs. Further purification on the S200 analytical column resulted in a single, homogeneous peak at an identical retention time to the B4 refold (13.4ML). SDS-PAGE analysis confirmed pure, clean bands at the correct molecular weights.

Once both TCRs demonstrated successful refolding via SDS-PAGE, we proceeded to assess their ability to bind CD1c-endo. This was a crucial step prior to crystallography studies, ensuring that the TCRs bound to CD1c before attempting co-crystallisation experiments. To investigate this, we used our previously optimised bead-based tetramer assay alongside SPR to measure the binding affinity of the TCRs.

Prior to investigating whether the TCRs stained the CD1c-endo coated MACSI beads, we were required to biotinylate both TCRs. These biotinylated TCRs could then be tetramerised for the bead-based binding assay.

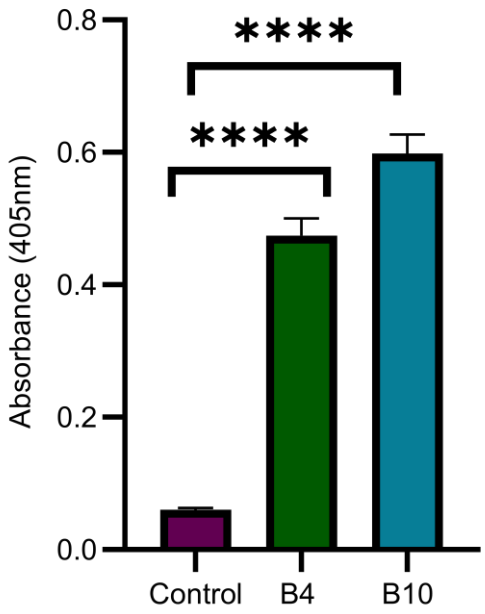


Figure 5. 22 B4 and B10 TCR biotin ELISA

Bar graph showing the Biotin ELISA results following overnight biotinylation of B4 and B10 γδ TCRs. Unbiotinylated B4 protein was used as a control. Data is represented as the mean ± standard deviation for each sample (n=3). Background absorbance was removed from each sample by subtracting PBS-only control from each data point. P values from independent T-tests were $P < 0.001$ for both 'ctrl vs B4' and 'ctrl vs B10'.

Biotin ELISA showed that both proteins were successfully biotinylated and revealed significant absorbance relative to the non-biotinylated control. Unfortunately, neither of the γδ TCRs stained the CD1c-coated MACSI beads, whereas the positive control TCR NM4 did. We also wanted to investigate binding using SPR, in case the lack of staining with tetramer assays was not sensitive to reveal weak affinity binding.

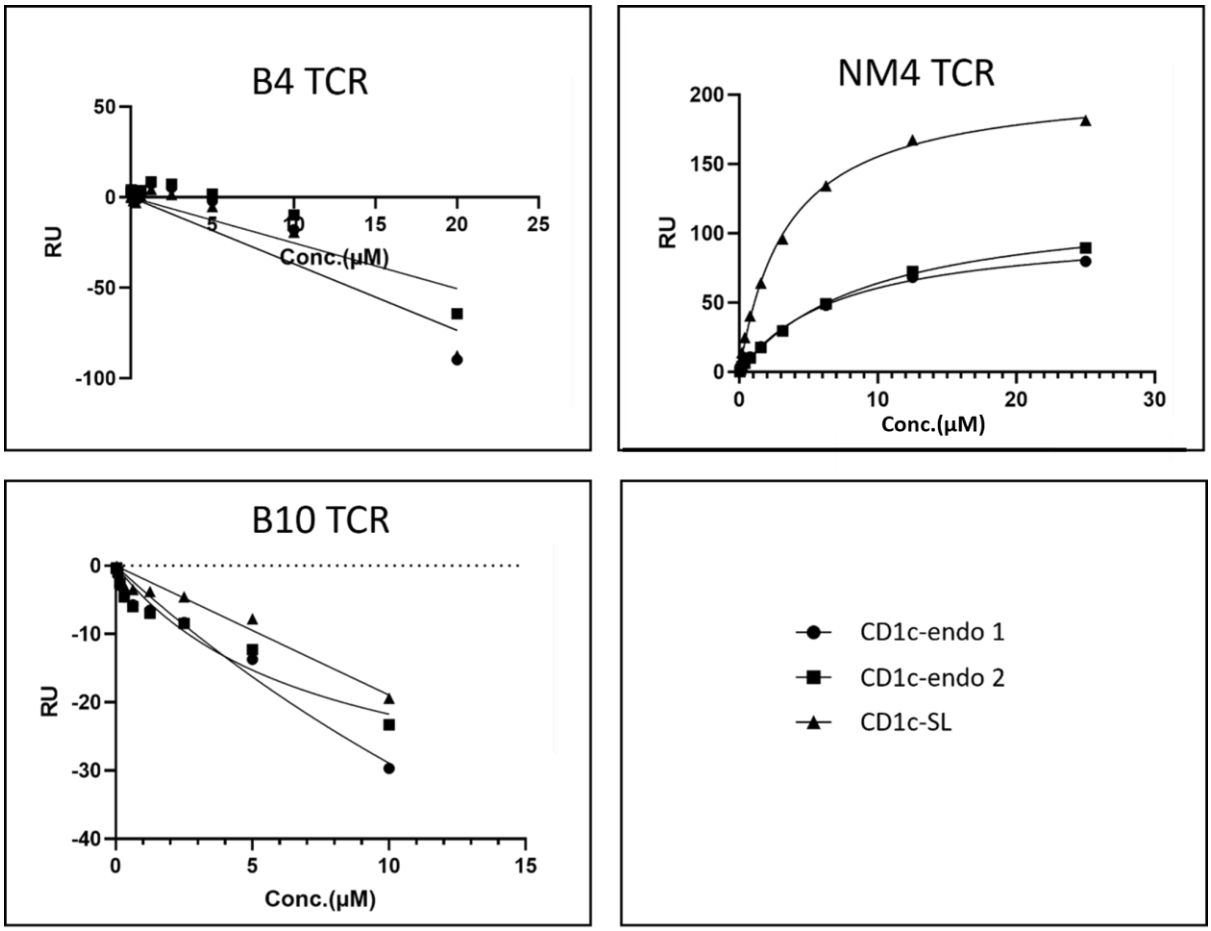


Figure 5. 23 SPR measurements of B4 and B10 γδ TCRs

Two different batches of CD1c-endo and one batch of CD1c-SL (spacer lipids) were loaded on the chip at roughly 1000 RU. Five different analytes were used, a TCR affinity matured for high CD1c-reactivity, and two known CD1c reactive αβ TCRs, B4, and B10. Data showed that **γδ** TCRs did not bind CD1c. TCR flow rate over the chip was 30μl/min.

Unfortunately, neither of the refolded γδ TCRs bound to CD1c-endo or to the CD1c-SL protein coated on the SPR chip. Although these TCRs were originally isolated using CD1c-SL protein, our modified CD1c protein production method aimed at high-yield expression for structural studies may have

altered their ability to recognise the presented lipid. This initially led us to hypothesise that the TCRs might be lipid-specific, failing to recognise the endogenous lipid cargo of CD1c-endo. However, this theory was disproven by their failure to bind CD1c-SL, which typically presents spacer lipids.

SDS-PAGE analysis confirmed that the TCR proteins were pure and stable, suggesting that both TCRs did not recognise CD1c. Following these results, we proceeded to refold the third Vγ4Vδ1 TCR (hu20), which had been isolated and cloned in collaboration with colleagues at King's College London. This TCR had previously demonstrated CD1c reactivity via tetramer staining of Jurkat cells (226).

Initial inclusion body production proved challenging, with strong expression of the γ chain but weak expression of the δ chain. We optimised inoculation amounts, glucose supplementation concentrations, and culture times in a small-scale culture format to address this. A re-transformation was also performed, and subsequent expressions yielded significantly better results, suggesting that the original transformation may have been faulty.

Following successful inclusion body generation, we carried out a 1:1 refold. The resulting elution profile *(Figure 5.24c)* and SDS-PAGE analysis *(Figure 5.24d)* displayed a peak at a similar retention time to the B4 and B10 refolds. However, a large aggregate peak and a wide single-chain peak were observed on either side of the dimer. Additionally, the initial SDS-PAGE gel appeared distorted, making it difficult to evaluate the refold constituents and indicating that further purification was necessary.

Purification on the S200 analytical column resulted in a relatively homogeneous protein, though its retention time was approximately 1 mL earlier than expected, based on the previous heterodimer refolds. SDS-PAGE analysis revealed the presence of a homodimer rather than a heterodimer, as only a single band was visible in the reducing lanes.

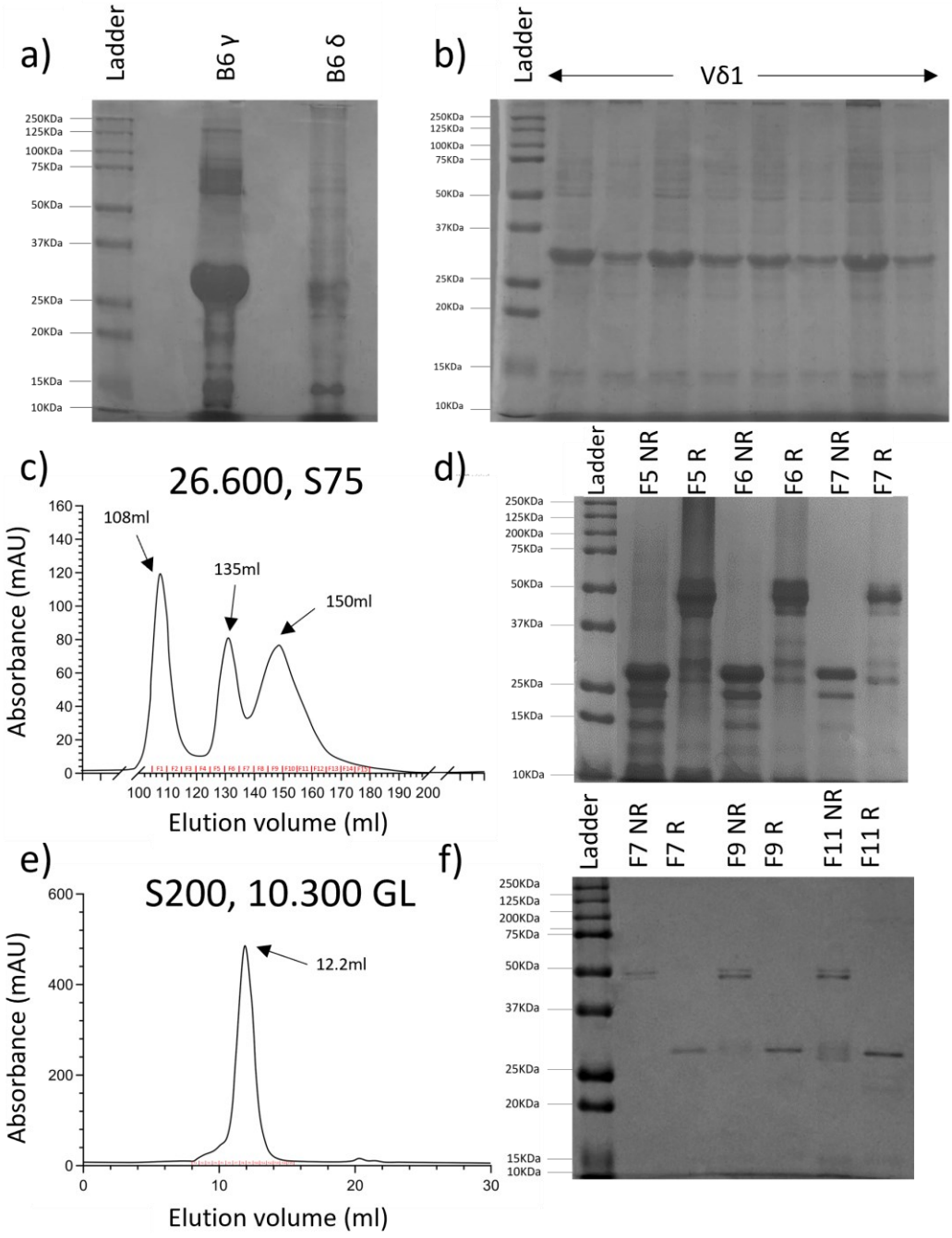


Figure 5. 24 Inclusion body gel and purification of hu20 TCR

a) SDS-PAGE of hu20 γ and δ chain inclusion bodies. b) SDS-PAGE gel of δ chain small-scale re-expression. c) Size exclusion chromatogram of hu20 TCR refold using a Hi-Prep S75, 26.600 gel filtration column, and d) Reducing and non-reducing SDS-PAGE of fractions five, six, and seven from the 26.600 purification e) Size exclusion purification of the pooled fractions F5-F7 from the previous purification using an analytical S200, 10/300 GL column. f) Reducing and non-reducing SDS-PAGE from fractions F7, F9, and F11 from the sharp peak seen in the prior S200 purification. A molecular weight ladder from Bio-Rad was used for molecular weight comparison in all SDS-PAGE gels.

5.4.2 Refolding a hybrid γδ TCR.

Following the discovery of homodimer formation during purification of the hu20 TCR, we attempted to refold a hybrid TCR, combining a γ chain capable of pairing with its corresponding δ chain (Vγ4 from B4 TCR) with a δ chain known to recognise CD1c-endo (Vδ1 from hu20). The rationale for this approach was based on evidence that the δ chain, particularly the CDR3δ loop, plays a key role in γδ TCR recognition of CD1c (227). By pairing hu20 Vδ1 with the B4 Vγ4 chain, which had not previously shown a tendency for homodimer formation, we aimed to generate a TCR that was both stable as a heterodimer and CD1c-reactive.

To test this, we used previously generated inclusion bodies for B4γ and hu20δ and refolded them at different ratios (1:3, 1:1, and 3:1), following the same refolding method described earlier.

Initial concentration via POROS ion exchange chromatography produced several fractions, which were analysed by SDS-PAGE. Given that TCR dimers typically elute between 22% and 27% salt gradient, we focused on the first two fractions from the POROS elution. SDS-PAGE analysis initially appeared promising, showing a 50 kDa band in the non-reducing lanes and two chains in the reducing lanes at approximately equal ratios (*Figure 5.24*).

Further purification using Hi-Prep 26/600 size exclusion chromatography also appeared successful, with a peak at a retention time of 132.4ML, comparable to previous heterodimeric TCRs. SDS-PAGE analysis showed a single band in the non-reducing lanes, while the reducing lanes displayed two bands, corresponding to the individual γ and δ chains. The δ chain (lower band) was fainter than the γ chain, with the most even ratio observed in fraction three, which corresponded to the large peak at 132.4ML.

Unfortunately, further purification using the S200 analytical column (*Figure 5.24e*) demonstrated that the hybrid TCR was unstable, as multiple fractions eluted with no evidence of a stable, refolded heterodimer in SDS-PAGE analysis.

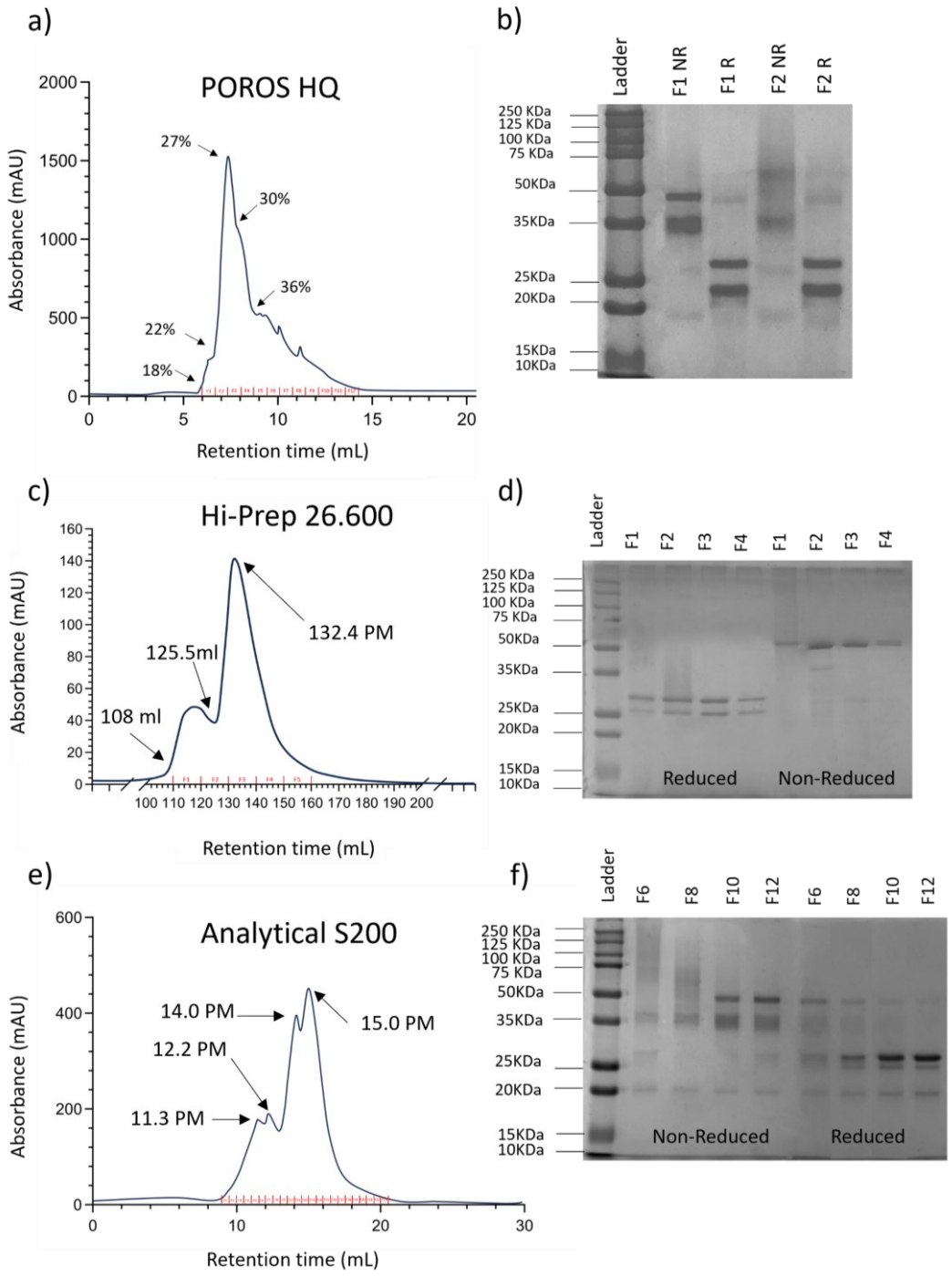


Figure 5. 25 Purification of the hybrid γδ TCR

a) A chromatogram from an anion exchange chromatography purification using a POROS HQ column. b) Reducing and non-reducing SDS-PAGE analysis of the POROS HQ purification of our hybrid TCR refold. c) Size exclusion purification trace and d) SDS-PAGE analysis of the relevant hybrid TCR fractions from the previous analysis. e) Analytical size exclusion chromatogram performed using an S200, 10.300 GL column. Relevant fractions were run on reducing and non-reducing SDS-PAGE (f). A molecular weight standards ladder was run in lane one in all gels for comparison (Bio-Rad).

5.4.3 Sequence and structural analysis of the γδ TCR dimer interface

Following the disappointment of the hu20 TCR failing to refold correctly, we sought to elucidate the underlying mechanism dictating the differences in chain pairing preference among our panel of γδ TCRs. To investigate this, we performed sequence analysis, comparing the γδ TCRs we had attempted to refold alongside several additional TCRs for which we had access to frozen samples.

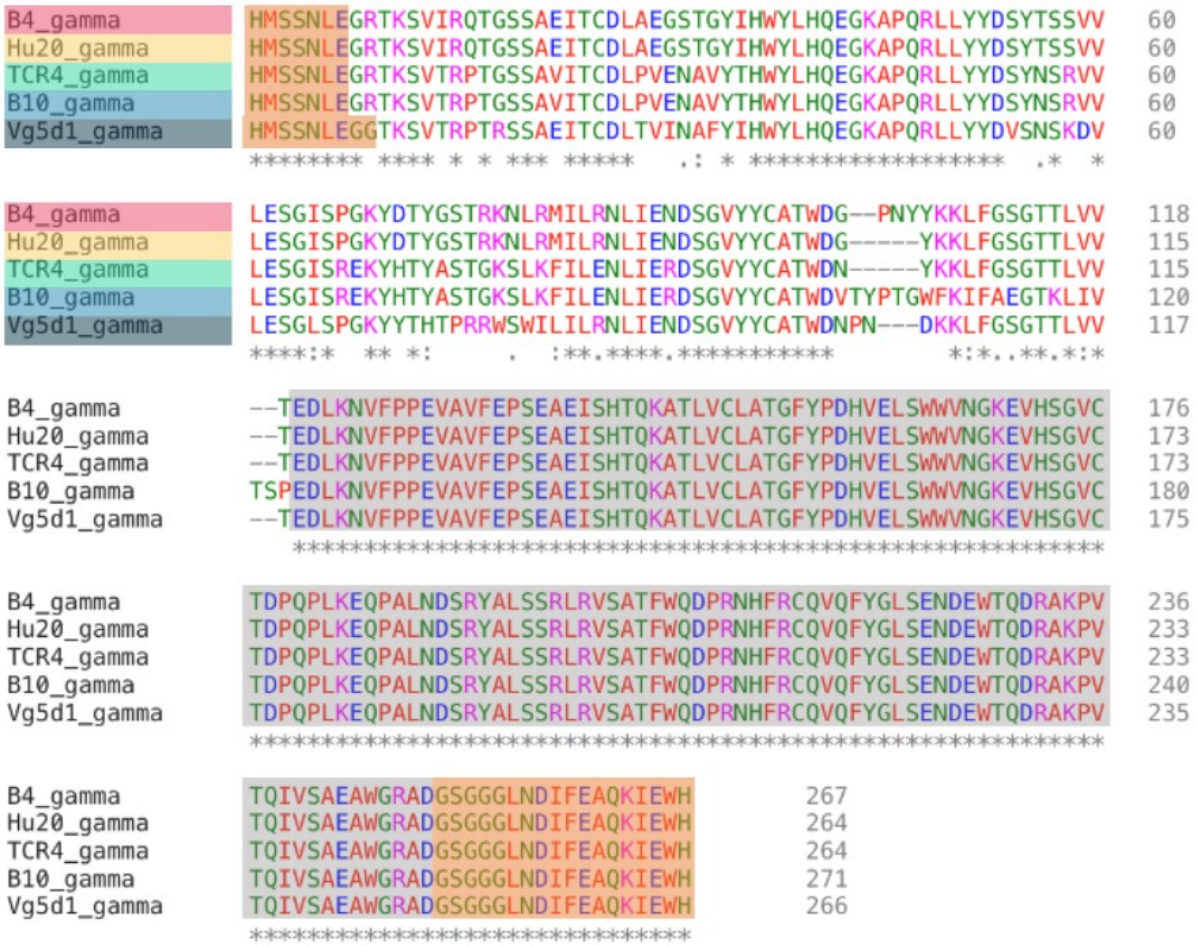


Figure 5. 26 γ chain sequence analysis of five γδ TCRs

Comparative sequence analysis of the TCRs we investigated, with the constant region highlighted in grey and the variable region unhighlighted. Sequence conservation is represented as an asterisk whilst differences are represented as a semi-colon, full stop, or blank space.

We conducted a sequence alignment and analysis focusing on the γ chain from five TCRs in our panel. The γ chain was chosen due to its apparent propensity for homodimer formation in some TCRs but

not others. We expected to identify clear differences in one of the metrics listed below, which could explain the homodimer formation observed in TCRs such as hu20. However, neither sequence alignment nor structural analysis revealed any obvious differences in residues forming the dimer interface.

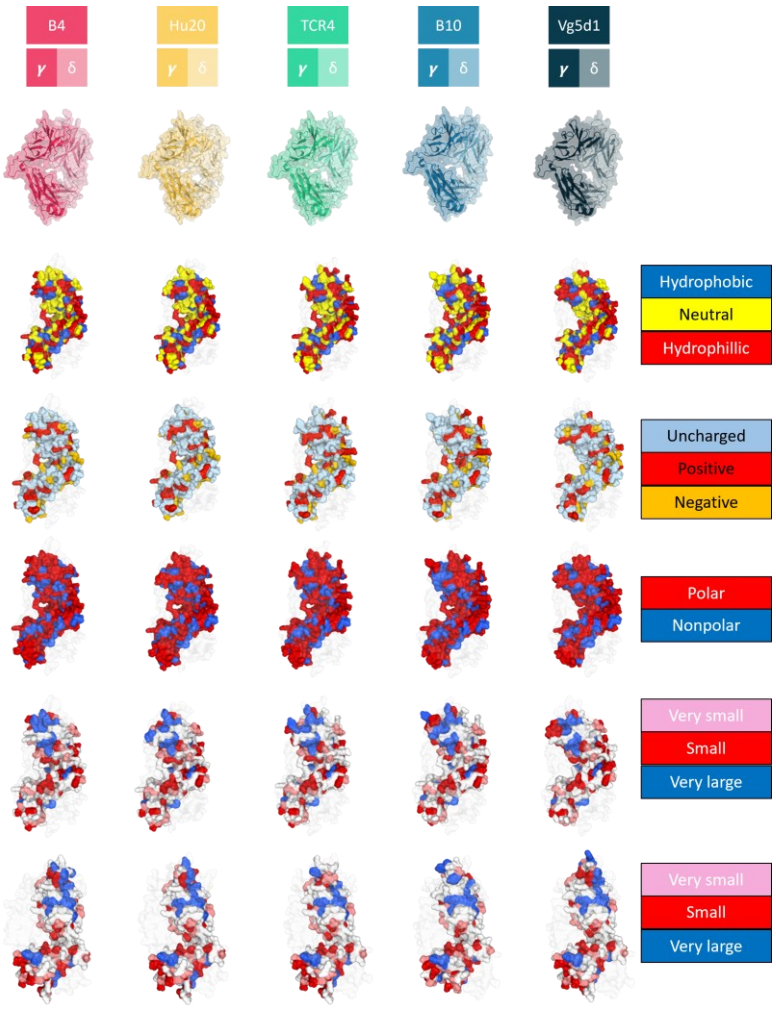


Figure 5. 27 γδ TCR interface comparison

Structural representation of five of the different γδ TCRs, analysed using several different metrics. The TCRs are colour-coded and amino acid composition is compared in terms of hydrophobicity, charge, polarity, and size. Both the γ and δ chains are analysed for residue sizes at the interface.

5.4.4 A protein engineering approach to encourage TCR heterodimer formation.

As no obvious differences were identified at the interface between the two chains of our TCRs of interest, we decided to take a different approach. Previous studies have demonstrated that protein engineering can be used to promote correct chain pairing in cases where mispaired interactions occur, both in TCRs and antibody therapies (228-230).

Based on this, we attempted to introduce mutations into the γ and δ chains, applying a 'knob-in-hole' engineering strategy to encourage heterodimer formation. This approach was modelled on a previously described methodology, where an αβ TCR (PDB entry 1TCR) was engineered to incorporate a 'knob-in-hole' mutation.

Despite poor sequence homology between 1TCR and hu20 TCR, their structural homology was significant, with the hu20 γ chain aligning to 1TCR with an RMSD of 0.832 and the hu20 δ chain aligning to the 1TCR α chain with an RMSD of 1.237. This structural similarity allowed us to model the same mutation into our γδ TCR sequence.

The final mutations introduced were Ser180δ → Arg180δ and Arg195γ → Gly195γ.

Table 6 Mutated γ and δ chain sequences with 'knob in hole' mutations are highlighted in yellow.

Mutated γ chain sequence	GRTKSVIRQTGSSAEITCDLAEGSTGYIHWYLHQEGKAPQRLLYYD SYTSSVVLESGISPGKYDTYGSTRKNLRMILRNLIENDSGVYYCAT WDGYKKLFGSGTTLVVTEDLKNVFPPEVAVFEPSEAEISHTQKATL VCLATGFYPDHVELSWWVNGKEVHSGVCTDPQPLKEQPALNDS RYALSS**G**LRVSATFWQDPRNHFRCQVQFYGLSENDEWTQDRAK PVTQIVSAEAWGRAD
Mutated δ chain sequence	QKVTQAQSSVSMPVRKAVTLNCLYETSWWSYYIFWYKQLPSKEMI FLIRQGSDEQNAKSGRYSVNFKKAAKSVALTISALQLEDSAKYFCAL GPPLFYVLGYRKLIFGKGTRVTVEPIQNPDPAVYQLRDSKSSDKSVC LFTDFDSQTNVSQSKDSDVYITDKCVLDMRSMDFKSN**R**AVAWSN KSDFACANAFNNSIIPEDTFFPS

Due to time constraints, the hu20 mutated γ and δ sequences were mutated via the Biotechnology company, Gene script. The sequences were cloned into bacterial expression vectors which we then utilised to produce soluble inclusion bodies. Successful inclusion body production preceded refolding, and purification via the same method as documented with previous TCRs.

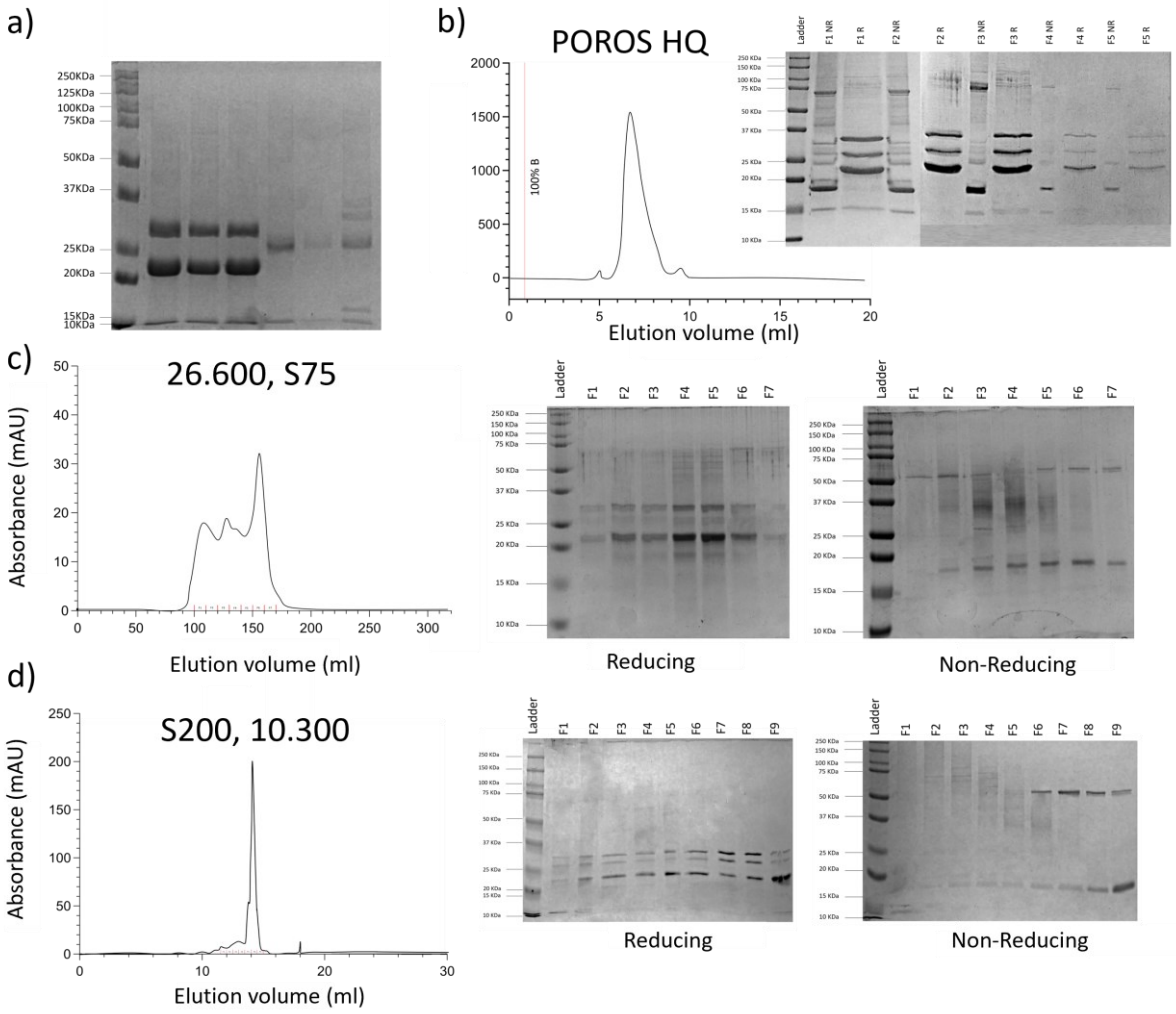


Figure 5. 28 Refolding engineered hu20 γδ TCR

a) SDS-PAGE gel of the mutated **hu20** inclusion bodies from three different preparations. The γ chain shows an unexpected extra band below the correct molecular weight. b) Anion exchange chromatography trace using a POROS HQ column with corresponding SDS-PAGE analysis. c) Size exclusion chromatography trace and corresponding reducing and non-reducing SDS-PAGE analysis of the relevant fractions from the Poros HQ elution, run using a 26.600 S75 preparatory column. d) Analytical size exclusion chromatography trace and resulting reducing and non-reducing SDS-PAGE analysis. Purification was performed using an S200, 10.300 GL column.

Despite again looking promising, the mutant hu20 refold was unstable, with biotinylation and re-purification following the purification seen in *Figure 5.28d* resulting in little to no stable, purified heterodimer. Due to time constraints making it difficult to spend a long time engineering a specific TCR, alongside the clear demonstration of γδ TCRs refolding using our method (B4 and B10), we decided to attempt refolding of some other CD1c-reactive γδ TCRs. These were isolated in a different way to our previous panel of TCRs (B4, hu20, and B10), using CD1c-endo tetramer staining of a γδ T cell line generated in-house.

5.4.5 Refolding of γδ TCRs isolated using CD1c-endo.

Following isolation using CD1c-endo tetramer and single-cell sorting, TCRs were then cloned using gene script and transduced onto Jurkat T cells to investigate CD1c-reactivity prior to bacterial expression and refolding. Cells were gated on single, live, CD3 positive CD1c-endo positive cells.

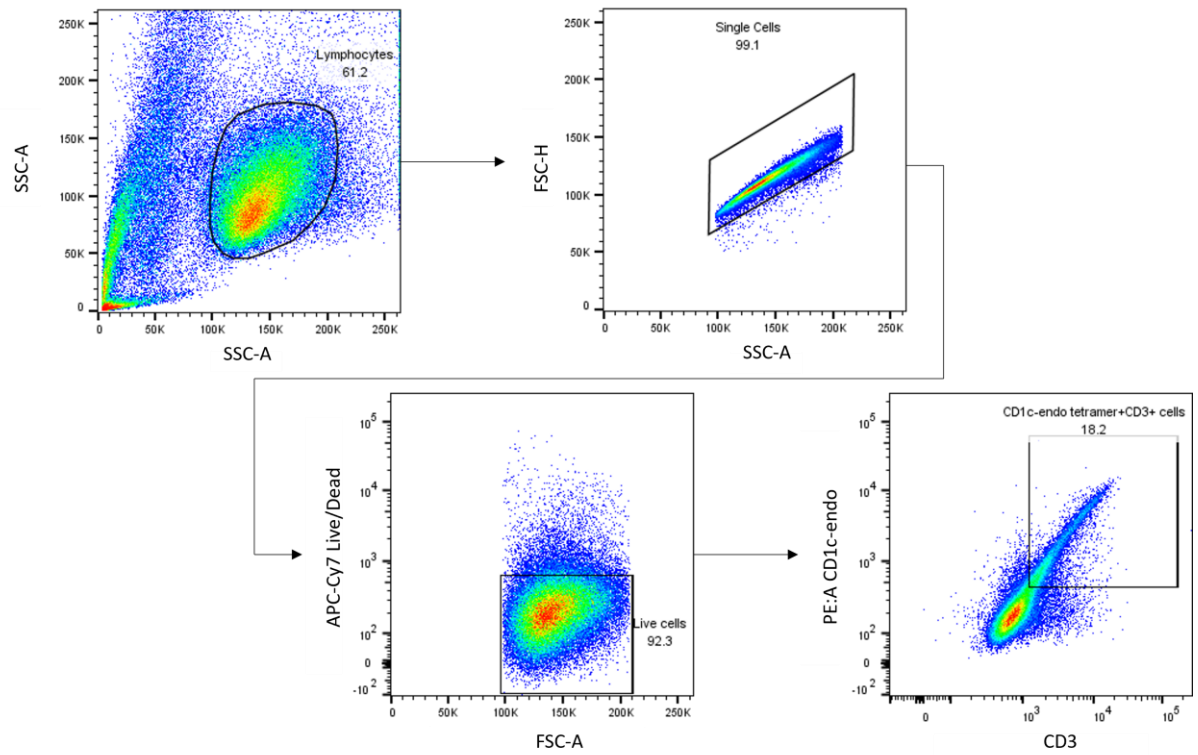


Figure 5. 29 Gating strategy for testing CD1c-reactivity of γδ TCR transduced Jurkats

Representative dot plots showing the gating strategy for investigating the CD1c-reactivity of the γδ TCR transduced Jurkat T cells. Lymphocytes were initially gated on, followed by single cells and live/dead, and finally CD1c-tetramer positive cells. The representative example consists of our positive control TCR, showing strong diagonal staining (tetramer vs CD3) using CD1c-endo tetramer, thus confirming CD1c-reactivity.

The transduction and tetramer staining process was repeated for all γδ TCRs that were originally isolated using our CD1c-endo tetramer staining of the in-house generated γδ T cell line. *Figure 5.30* below shows representative examples of successful transductions. Our data suggests that two out of three cloned TCRs were CD1c reactive as Jurkats stained brightly with CD1c-endo tetramer.

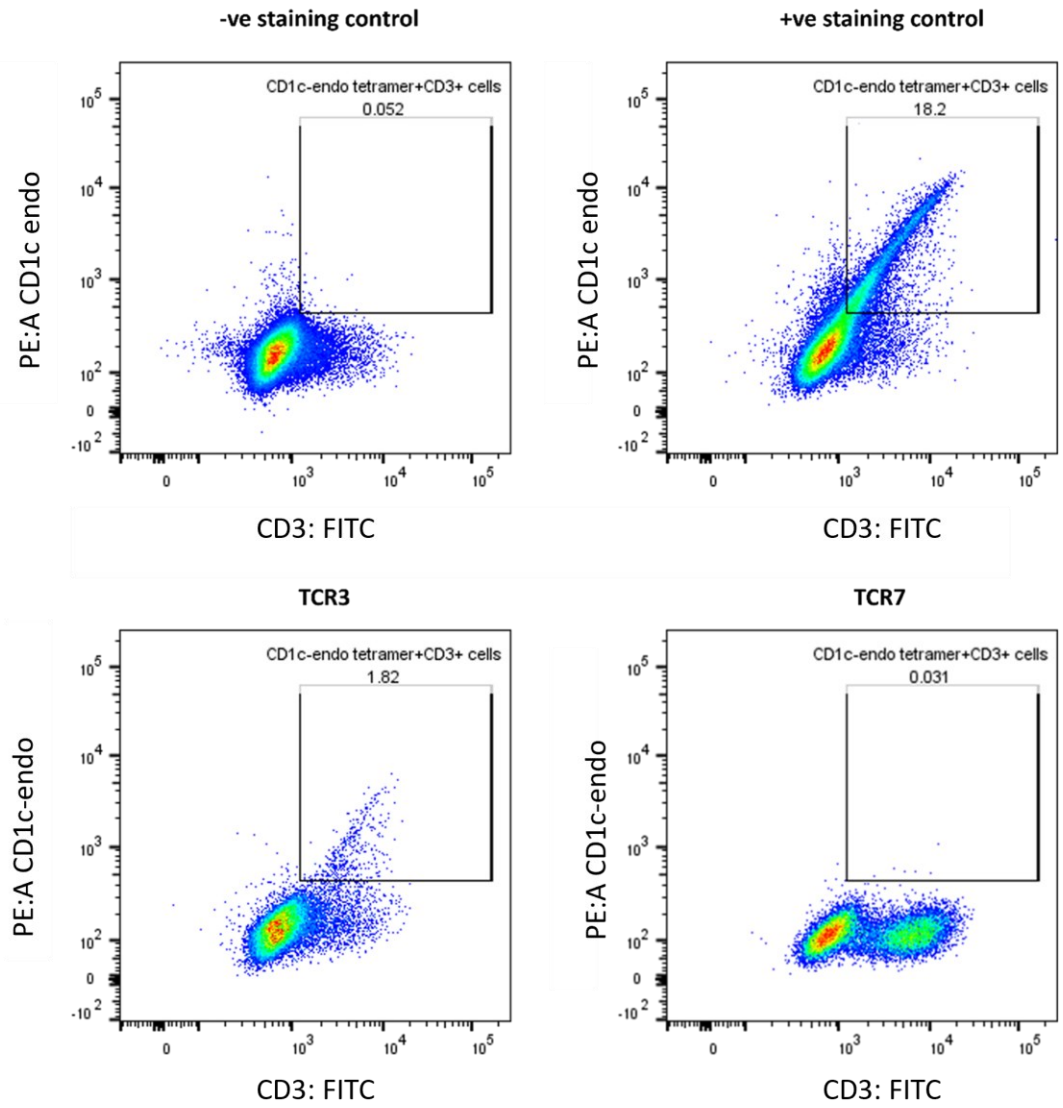


Figure 5. 30 Representative CD1c-endo tetramer staining of γδ TCR transduced Jurkat lines.

Flow cytometry dot plots of negative staining control, positive staining control, TCR3, and TCR7 transduced Jurkat T cell lines. TCR3 showed poor transduction efficiency but exhibited clear staining with CD1c-endo tetramer, confirming CD1c-reactivity. TCR7 demonstrated higher transduction efficiency than TCR3 but failed to stain with CD1c-endo tetramer, suggesting that TCR7 is not CD1c-reactive.

Transduction onto Jurkats and staining with CD1c-endo tetramer confirmed two new CD1c-reactive γδ TCR targets for refolding (***see table 5.4***)

Table 7 Sequence of the gamma and delta chains of TCR3 and TCR4, both isolated using CD1c-endo tetramer from an in-house generated γδ T cell line.

TCR	Vγ chain sequence	Vδ chain sequence
TCR3	MAWALAVLLAFLSPASQKSSNLEGRTKSVIRQTGSSAEITCDLAE GSTGYIHWYLHQEGKAPQRLLYYDSYTSSVVLESGISPGKYDTYG STRKNLRMILRNLIENDSGVYYCATWDVKRGWIKTFAKGTRLIV TSPDKQLDADVSPKPTIFLPSIAETKLQKAGTYLCLLEKFFPDVIKI HWQEKKSNTILGSQEGNTMKTNDTYMKFSWLTVPEKSLDKEH RCIVRHENNKNGVDQEIIFPPIKTDVITMDPKDNCSKDANDTLLL QLTNTSAYYMYLLLLLKSVVYFAIITCCLLRRTAFCCNGEKS	MLFSSLLCVFVAFSYSGSSVAQKVTQAQSSVSMPVRKAVTLNCLYET SWWSYYIFWYKQLPSKEMIFLIRQGSDEQNAKSGRYSVNFKKAAKSV ALTISALQLEDSAKYFCALGEGSSLKYWGIPADKLIFGKGTRVTVEPRS QPHTKPSVFVMKNGTNVACLVKEFYPKDIRINLVSSNKITEFDPAIVIS PSGKYNAVKLGKYEDSNSVTCSVQHDNKTVHSTDFEVKTDSTDHVK PKETENTKQPSKSCHKPKAIVHTEKVNMMSLTVLGLRMLFAKTVAV NFLLTAKLFFL
TCR4	MVLALALLLAFLPPASQKSSNLEGRTKSVTRPTGSSAVITCDLPV ENAVYTHWYLHQEGKAPQRLLYYDSYNSRVVLESGISREKYHTY ASTGKSLKFILENLIERDSGVYYCATWDNYKKLFGSGTTLVVTDK QLDADVSPKPTIFLPSIAETKLQKAGTYLCLLEKFFPDIIKIHWQEK KSNTILGSQEGNTMKTNDTYMKFSWLTVPEESLDKEHRCIVRH ENNKNGIDQEIIFPPIKTDVTTVDPKDSYSKDANDVTTVDPKYNY SKDANDVITMDPKDNWSKDANDTLLLQLTNTSAYYMYLLLLLK SVVYFAIITCCLLRRTAFCCNGEKS	MLFSSLLCVFVAFSYSGSSVAQKVTQAQSSVSMPVRKAVTLNCLYET SWWSYYIFWYKQLPSKEMIFLIRQGSDEQNAKSGRYSVNFKKAAKSV ALTISALQLEDSAKYFCALGTSFLRGIRLASADKLIFGKGTRVTVEPRSQ PHTKPSVFVMKNGTNVACLVKEFYPKDIRINLVSSKKITEFDPAIVISPS GKYNAVKLGKYEDSNSVTCSVQHDNKTVHSTDFEVKTDSTDHVKPK ETENTKQPSKSCHKPKAIVHTEKVNMMSLTVLGLRMLFAKTVAVNFL LTAKLFFL

We initially focused on TCR3, following the same refolding and purification protocol as previously described. Due to the dual expression of bands in the γ chain, we refolded at a 2γ:1δ ratio, assuming the contaminant band was present at similar levels to the correct molecular weight band. While this approach inevitably increased the concentration of the contaminant band in the refold, we hypothesised that we would still detect some indication of refolded protein, which could later be optimised by adjusting chain ratios and redox conditions.

The refolded protein was first concentrated using an AMICON stir cell concentrator, followed by purification on the Hi-Prep 16/600 size exclusion column. The resulting chromatogram and SDS-PAGE analysis revealed a complex mixture of components in the eluent.

Despite evidence of a dimer in the non-reducing gel, the reducing gel was more difficult to interpret. Heterodimer bands were observed in lanes three and four, while a contaminant band was present in lanes five and six, along with additional γ chain appearing in the later lanes (seven to nine). An excess of γ chain was also detected in fractions four and five, suggesting the presence of a combination of heterodimer and homodimer in these fractions.

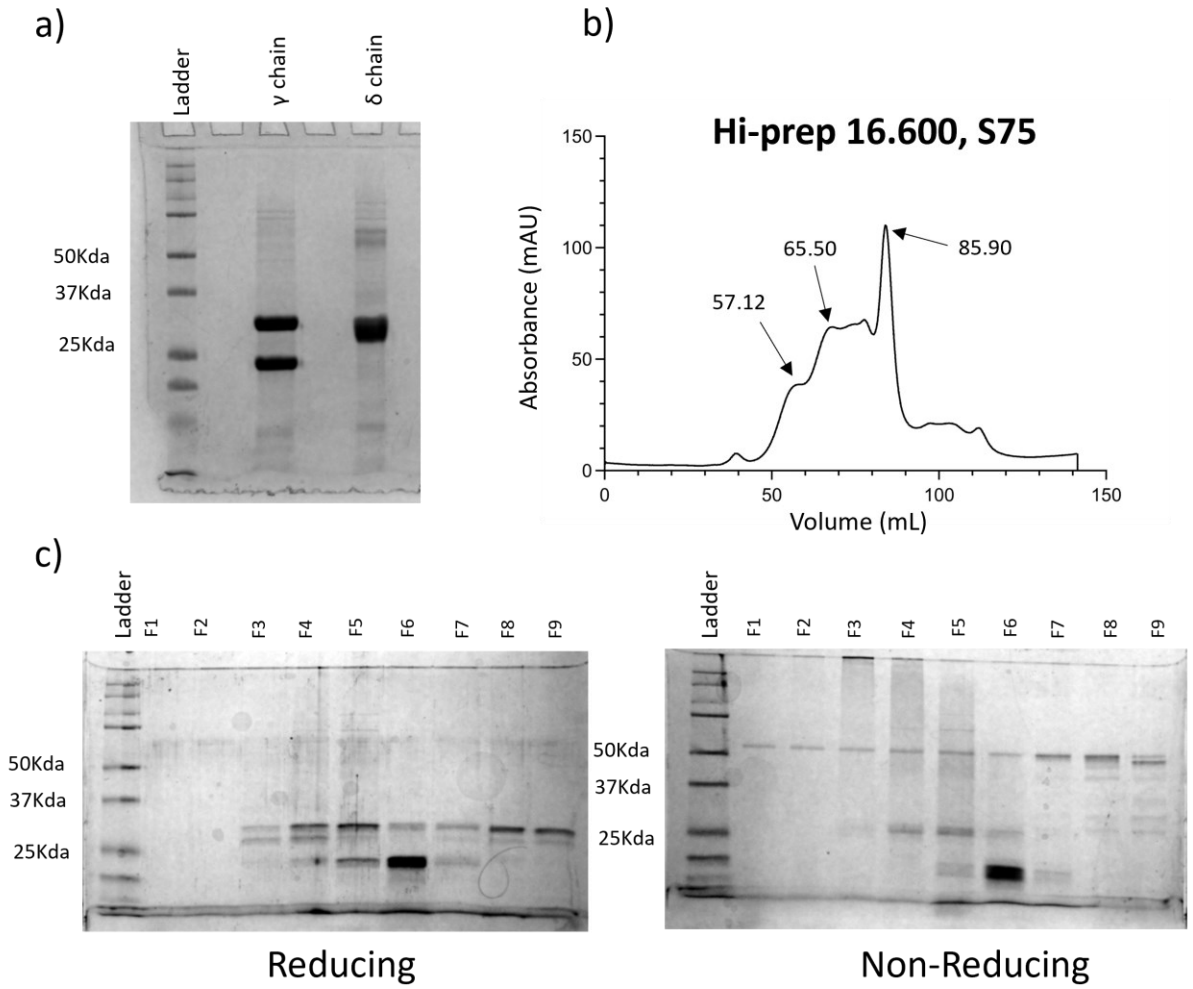


Figure 5. 31 Purification and SDS-PAGE analysis of TCR3

a) SDS-PAGE gel showing inclusion body proteins of the TCR3 γ and δ chains. The γ chain appears to be expressed at two main molecular weights, one at the expected 30KDa and one at a truncated molecular weight of less than 25KDa. b) Size exclusion chromatogram of refolded protein run on a Hi-Prep 16.600, S75 column. c) Reducing and non-reducing SDS-PAGE analysis of the resulting chromatogram peaks, with a molecular weight standard ladder in lane one for comparison.

Due to the inability to efficiently separate the components from the refolded mixture and obtain our pure heterodimer fraction, we decided to adapt our refolding protocol to introduce a His-tag to the δ chain. We hypothesized that we could remove the majority of unwanted samples i.e. single γ chain, γ homodimer, and the contaminant band seen in the γ inclusion body gel, prior to the size exclusion step by using an IMAC column. This should therefore eliminate the majority of contaminants prior to the size exclusion and improve the efficiency of the process.

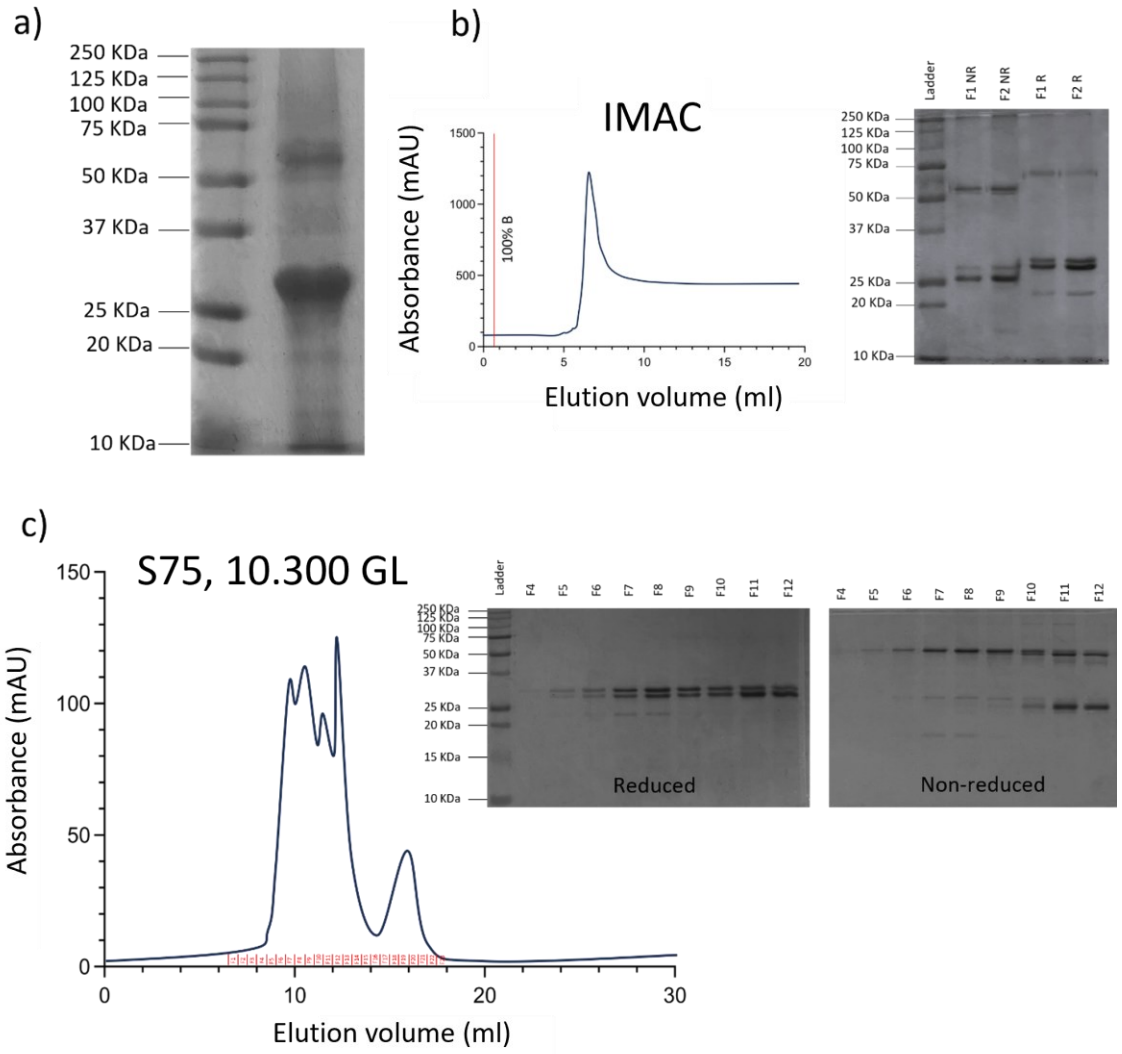


Figure 5. 32 Purification and SDS-PAGE analysis of 6X-His-tag TCR3

a) Inclusion body SDS-PAGE analysis of the new sequence TCR3 δ chain with a His-tag attached, showing a clear band at the expected molecular weight relative to the molecular weight ladder. b) IMAC purification of the refolded TCR3, with a 50KDa band indicating dimer formation in the non-reducing lanes and two bands in the reducing lanes. c) Analytical size exclusion purification of TCR3 pooled fractions from the IMAC purification. The resulting SDS-PAGE analysis shows clean bands in the non-reducing lanes 4-9, with two bands in the reducing lanes at even ratios, indicating successful heterodimer formation.

Our TCR3 6x-His-tag purification looked promising, with stable heterodimer appearing to remain throughout the IMAC elution and size exclusion chromatography run. We decided to pool different fractions along the peak, make tetramer with each fraction pool, and stain the CD1c-endo coated MACSI beads to analyse which fractions corresponded to the heterodimer. The multi-peak chromatogram from *Figure 5.32c* made it difficult to correctly analyse which peak corresponded to which lanes on the gel, hence we decided the best approach was to make multiple tetramers and investigate the fractions using our bead-based system that has previously been described.

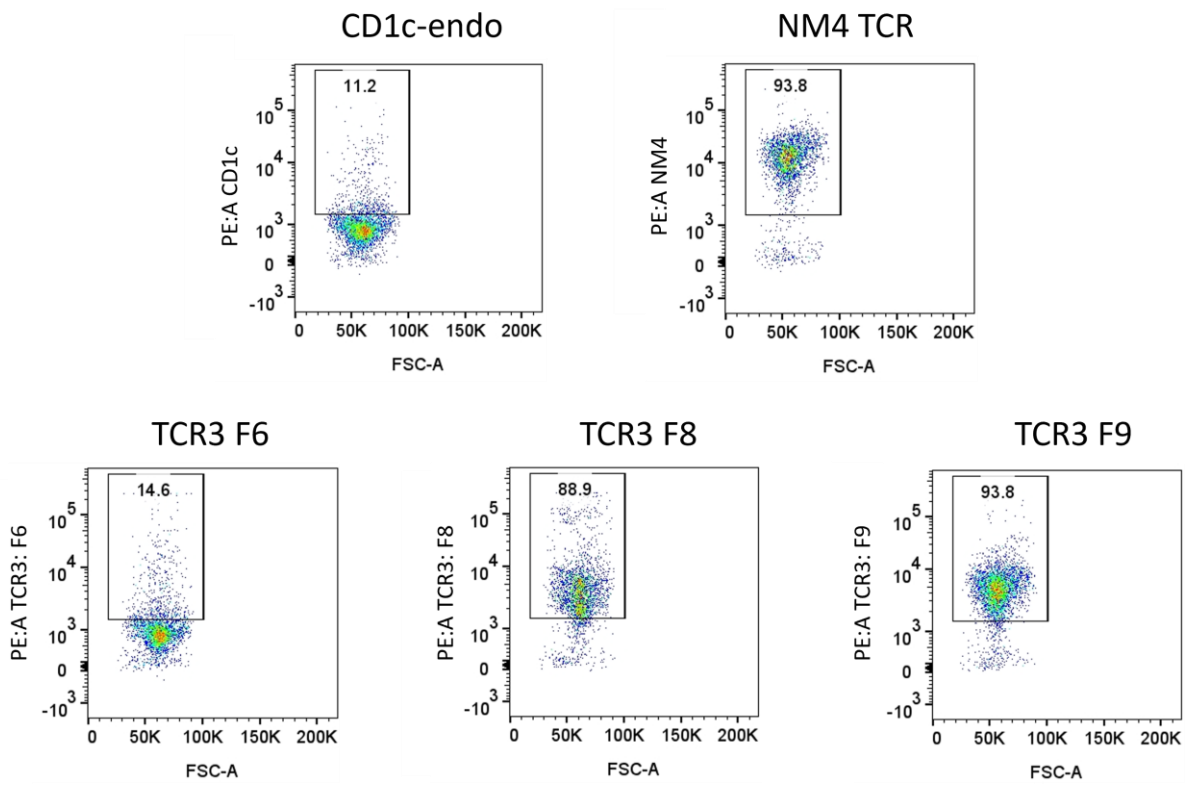


Figure 5. 33 TCR3 stains CD1c-endo coated MACSI beads

Flow cytometry dot plots showing tetramer staining of three different TCR3 fractions from the previous TCR3 His-tag purification. The CD1c-endo tetramer was used as a negative control, and NM4 TCR served as a positive control. The PE-positive gate was set above the CD1c-endo tetramer negative control to account for background staining.

We generated tetramers from individual fractions collected during S200 analytical purification, using F6, F8, and F9 to stain CD1c-coated MACSI beads. NM4 TCR was used as a positive control, with fraction six failing to stain, while fractions eight and nine showed reasonably good staining, despite not being 100% purified.

Using the IMAC column as an initial purification step appeared to simplify the refolding mixture, primarily by eliminating the lower contaminant band observed in inclusion body gels, as well as removing excess single γ chain and any potential γ homodimer that may have formed during refolding.

Due to time constraints, we were unable to conduct SPR experiments with this newly refolded TCR, but these will now be pursued for testing and crystallisation by our new PhD student. Additionally, we have identified another TCR of interest (TCR4), which exhibits clear staining with CD1c-endo tetramer when transduced into Jurkat T cells. We aim to refold this TCR using the same methodology, alongside the previously refolded TCR, to set up co-crystallisation experiments with CD1c-endo and further investigate γδ TCR recognition mechanisms.

Although γδ TCR refolding has proven challenging, we successfully refolded and purified a CD1c-reactive γδ TCR, which can now be utilised for structural and functional studies. Furthermore, we developed several assays, including the plate-bound assay and bead-based assay, which can be used to assess the reactivity of TCR3 and any additional γδ TCRs that we successfully refold and purify. Previous unsuccessful TCR refolding attempts may also benefit from our new His-tag purification system, providing an opportunity to expand the number of CD1c-reactive γδ TCRs available for our structural pipeline.

5.5 Discussion

5.5.1 Optimising CD1c-endo and γδ TCR protein production

This study aimed to optimise the production and purification of CD1c-endo and γδ TCR proteins to facilitate structural and functional investigations. Given the technical challenges associated with producing these proteins in sufficient purity and yield, we refined expression and purification workflows to generate high-quality protein for downstream applications. CD1c-endo was successfully expressed in a mammalian system, providing a scalable and reproducible method for producing functional protein, while γδ TCRs were expressed in bacterial systems followed by in vitro refolding.

Optimising the CD1c-endo production pipeline increased protein yield and purity while streamlining workflows. However, glycosylation and lipid heterogeneity complicated crystallisation and binding studies, with lipid variability likely influencing TCR recognition. Despite attempts to mitigate glycosylation effects through endoglycosidase treatment, structural heterogeneity remained an obstacle, highlighting the need for alternative glycosylation control strategies.

γδ TCR refolding was similarly challenging, with several candidates displaying poor stability or failing to bind CD1c-endo. The TCRs, originally isolated using CD1c-SL tetramers, did not consistently demonstrate specificity for CD1c-endo, suggesting alternative binding mechanisms. Despite these challenges, we successfully refolded a CD1c-reactive γδ TCR, providing a foundation for future structural studies. Further improvements in refolding conditions and engineering strategies may enhance stability and overall yield.

These findings highlight both the progress made and the remaining obstacles in studying CD1c-reactive γδ TCRs, particularly the impact of glycosylation, lipid heterogeneity, and TCR stability. The following sections explore these challenges in greater detail.

5.5.2 Challenges and Optimisation of CD1c-endo production

CD1c-endo production presented several technical challenges, requiring optimisation to generate sufficient yield and quality for structural and functional studies. The complex glycosylation patterns and variable lipid repertoire of CD1c contribute to batch-to-batch variability, influencing both protein behaviour and TCR binding dynamics. Previous studies have successfully expressed CD1c in bacterial and mammalian systems, yet achieving consistency, purity, and functional integrity remains a

significant hurdle. This section discusses glycosylation heterogeneity and lipid variability, and the optimisations undertaken to mitigate these issues.

A mammalian expression system was selected due to its ability to retain native post-translational modifications (PTMs), such as glycosylation, which are critical for proper protein folding and function. Purification was streamlined through His-tag affinity and size-exclusion chromatography, yielding a more consistent final product compared to previous multi-step protocols. The incorporation of biotin during transfection further improved efficiency and total protein yield, avoiding losses due to precipitation and aggregation (*Figure 5.3*).

The reproducibility of mammalian-expressed CD1c-endo facilitated its use in SPR, bead-based assays, and lipid-exchange experiments. However, glycosylation heterogeneity posed a major challenge for crystallisation. Similar issues have been reported for lipid-presenting proteins, where glycan heterogeneity interferes with crystal lattice formation and diffraction (222).

Glycosylation challenges in CD1c-endo production

CD1c-endo contains four N-linked glycosylation sites at asparagine residues 38, 70, 75, and 146 (Uniprot), which significantly impact protein behaviour. Glycosylation has been shown to influence CD1 molecule function and expression, with previous studies highlighting a trade-off between maintaining protein integrity and achieving successful crystallisation (231, 232). The chemical diversity and heterogeneity of glycoproteins at these N-linked glycosylation sites likely hindered the formation of an ordered crystal lattice.

To mitigate this, enzymatic deglycosylation using endoglycosidase F was attempted, a method previously used successfully in CD1 structural studies (REF). However, this approach did not improve crystal quality, likely due to incomplete deglycosylation. Alternative strategies include the use of glycosylation inhibitors such as kifunensine or swainsonine, which restrict glycan processing to produce a more uniform glycan profile (222). Another potential approach involves alternative expression systems, such as bacterial or insect cell expression, though these often result in low yields or a loss of lipid-loading capability (REF). Despite these efforts, diffraction-quality crystals were not obtained, indicating that further modifications to the CD1c construct or expression system may be necessary.

Lipid Heterogeneity and Its Impact on TCR Binding

Unlike CD1c-SL, which presents a well-characterised set of spacer lipids, CD1c-endo displays a heterogeneous array of endogenous lipids (233). This presents a significant challenge due to variation

in lipid cargo significantly affecting TCR binding and affinity (64, 65). Autoreactivity of some TCRs has been demonstrated in previous studies, with the TCR CDR loops and main chain primarily contacting the CD1 protein itself, regardless of the lipid cargo (64). In contrast, some TCRs are highly specific and recognise distinct lipid structures protruding from the binding groove (65, 234). This variability presents challenges when predicting TCR-CD1 binding patterns. This is particularly relevant to our NM4 TCR, which was originally isolated using CD1c-SL and demonstrated a lower binding affinity for CD1c-endo compared to CD1c-SL (~9 µM vs. ~6 µM) (71). One explanation for this could be the differences in lipid content between CD1c-SL and CD1c-endo, as bulky headgroup lipids such as PC, PG, and sphingomyelin are more prevalent in CD1c-endo and may interfere with TCR recognition (235).

To address this, we explored lipid-exchange strategies to allow the controlled replacement of endogenous lipids with defined ligands. The most effective method involved CHAPS detergent in citrate buffer, which successfully removed bound lipids without destabilising CD1c-endo. Further refinement of this method could facilitate the systematic characterisation of lipid-dependent TCR recognition, particularly for γδ TCRs that show specificity for self or foreign lipids presented by CD1c.

5.5.3 γδ TCR Refolding and Optimisation for Structural and Functional Studies

γδ TCRs exhibit diverse antigen recognition mechanisms (236), differing significantly from the canonical peptide-MHC interactions of αβ TCRs. Instead of conforming to the MHC peptide co-recognition paradigm, they engage a broader range of antigens, including protein, lipids and metabolites (236). Often, these molecules are also unconventionally recognised by the TCR (52). Aside from the unusual variability in antigen recognition, characterisation and analysis of these γδ TCRs are further complicated by an inherent instability, chain pairing variability, and misfolding/aggregation propensity. A common view is that the chain pairing variability exhibited by these TCRs contributes heavily to the differences in ligand specificity when compared to αβ TCRs. Structural studies in the past have been further complicated by homodimer formation (particularly amongst the γ chain), which likely contributes to the lack of γδ TCR/CD1 co-crystal data currently in the PDB.

We optimised various parameters to overcome these issues, including chain ratios, refolding buffer composition, and purification strategy. His-tag purification demonstrated the cleanest and most efficient approach, enabling correctly folded γδ TCR generation and isolation. This method also allowed for the reduction in single chain and homodimer, thus efficiently cleaning up the final product. Despite promising results, we were unable to refold some of our target TCRs, perhaps suggesting optimisation should be personalised for the TCR and carried out on a case-by-case basis.

One key finding was the inability of some of our γδ TCR panel to bind CD1c-endo, despite being originally isolated using a CD1c (SL) tetramer. This may reflect the difference in lipid composition between CD1c-endo and CD1c-SL, perhaps suggesting that the characteristics of the mammalian lipids that encouraged TCR recognition are not compatible for other lipid-expressing CD1c molecules. Additionally, CD36 has been reported to interact with CD1c, raising the possibility that some TCRs initially identified as CD1c-reactive may have been cross-reacting with CD36 or another associated ligand during tetramer staining. This underlines the importance of implementing validation checkpoints during the protein production pipeline such as transducing candidate TCRs onto a Jurkat cell line and confirming CD1c reactivity via tetramer staining, prior to initiating the full TCR cloning and production pipeline.

To further enhance γδ TCR refolding, protein engineering strategies designed to stabilise heterodimer formation could be explored. The "knob-in-hole" approach, which has been successfully applied to stabilise antibody Fc heterodimers, could be adapted for γδ TCRs to minimise homodimer formation and improve overall yield (230). Another potential strategy involves grafting γδ variable domains onto αβ constant regions, as demonstrated in recent studies using mammalian expression systems, which may provide an alternative method for producing soluble γδ TCRs with increased stability (237). While bacterial expression remains the preferred method for generating soluble TCRs, the potential benefits of mammalian expression, particularly when post-translational modifications influence stability, warrant further investigation.

Despite the difficulties associated with γδ TCR refolding, this study successfully produced a soluble, CD1c-reactive γδ TCR, demonstrating the feasibility of generating γδ TCRs for structural and functional studies. Future work should focus on refining refolding conditions, incorporating additional validation steps to confirm antigen specificity, and applying engineering strategies to enhance stability and yield. Addressing these challenges will be crucial in achieving co-crystal structures of γδ TCRs bound to CD1c, providing key insights into the molecular basis of γδ TCR recognition of lipid-presenting molecules.

5.5.4 Validation of γδ TCR Binding to CD1c and Functional Implications

Establishing whether γδ TCRs recognise CD1c is crucial for understanding their molecular interactions with lipid-presenting antigen-presenting molecules. Previous studies have shown that γδ TCRs exhibit highly diverse binding mechanisms, with some recognising CD1c in a lipid-dependent manner while others interact autoreactively with the protein itself (65, 238). Given the heterogeneity of lipids presented by CD1c, defining the precise requirements for γδ TCR binding remains a major challenge. This study sought to validate γδ TCR reactivity against CD1c using tetramer-based bead assays and

surface plasmon resonance (SPR), allowing for an assessment of binding affinity and specificity. However, inconsistencies in staining intensities and binding profiles indicated that additional factors, including lipid composition and glycosylation, may influence TCR recognition.

One significant observation was that certain γδ TCRs failed to bind CD1c-endo despite being originally isolated using a CD1c-SL tetramer. This suggests that these TCRs may require specific lipid ligands for recognition or that their selection was influenced by an alternative binding mechanism. CD1c-SL is known to present well-defined spacer lipids, whereas CD1c-endo displays a diverse spectrum of endogenous lipids, which may significantly alter the conformation of the antigen-binding groove (71, 305). The presence of bulky headgroup lipids such as PC, PG, and sphingomyelin in CD1c-endo may interfere with TCR docking (306), providing a plausible explanation for the reduced or absent binding of some γδ TCRs.

Another potential factor affecting γδ TCR recognition is CD36 cross-reactivity. CD36 has been identified as a ligand for CD1b, CD1c, and CD1d, and previous reports suggest that TCRs initially isolated using CD1c tetramers may have been inadvertently binding to CD36 rather than CD1c itself (320). This highlights the need for rigorous validation when identifying CD1c-reactive TCRs, particularly when working with unconventional T cells that may engage multiple ligands. Future studies should consider blocking CD36 interactions during TCR selection to eliminate the possibility of non-specific binding events influencing γδ TCR discovery.

Tetramer-based assays proved useful for assessing γδ TCR interactions with CD1c, yet the observed variability in staining intensities across different CD1c-endo batches underscores the impact of lipid heterogeneity on binding outcomes. To mitigate this issue, a lipid-exchange system was implemented to allow the controlled replacement of endogenous lipids with defined ligands, providing a more standardised approach to lipid presentation. While this system facilitated a degree of consistency, some γδ TCRs still failed to bind both CD1c-endo and CD1c-SL, raising the possibility that their original selection was influenced by factors unrelated to CD1c specificity.

SPR experiments provided further insight into γδ TCR binding properties, revealing that interactions with CD1c-endo were generally weaker compared to those observed with CD1c-SL. This is consistent with reports showing that certain CD1-restricted TCRs exhibit higher affinity for specific lipid-loaded CD1 isoforms, reinforcing the importance of lipid cargo in determining TCR specificity. Post-translational modifications such as glycosylation may also contribute to binding variability, as glycan moieties on CD1c-endo could mask key TCR contact sites, reducing binding affinity compared to the more uniformly glycosylated CD1c-SL. Future studies should investigate these variables using

glycosylation-deficient CD1c constructs to assess the extent to which post-translational modifications impact γδ TCR recognition.

To further validate γδ TCR binding to CD1c, additional functional assays are required. One approach would be to transduce candidate TCRs into a Jurkat reporter cell line, enabling the measurement of TCR activation upon CD1c engagement. This would provide a functional readout of whether CD1c binding leads to downstream signalling, distinguishing biologically relevant interactions from weak or non-specific binding events. Additionally, the use of mutational analysis of both γδ TCRs and CD1c could help pinpoint the structural determinants of their interactions, providing deeper insights into the rules governing γδ TCR-CD1c specificity.

Understanding the factors driving γδ TCR binding to CD1c has implications beyond fundamental immunology, particularly in cancer immunotherapy. CD1c expression is upregulated in certain haematological malignancies, including leukaemia and lymphoma, suggesting a potential role for CD1c-targeted γδ TCR therapies (238, 239). By refining TCR selection, lipid-loading protocols, and validation pipelines, future studies will be better positioned to harness γδ TCRs for therapeutic applications, providing a framework for developing novel TCR-based immunotherapies.

5.5.5 Conclusion and Future Directions

This study optimised CD1c-endo and γδ TCR production to enable structural and functional investigations of γδ TCR-CD1c interactions. Refinements in mammalian expression significantly improved CD1c-endo yields, though glycosylation heterogeneity and lipid variability remained barriers to crystallisation. Similarly, γδ TCR refolding proved challenging, with many candidates demonstrating poor stability or failing to bind CD1c-endo. However, iterative optimisation of chain ratios, refolding buffers, and purification workflows led to the successful production of a soluble, CD1c-reactive γδ TCR, providing a platform for future studies.

Despite these advances, challenges remain. γδ TCR heterodimer stabilisation is a key hurdle, as chain pairing variability and homodimer formation hinder yields and structural studies. Protein engineering strategies, such as the "knob-in-hole" (KiH) approach used in antibody engineering, could be adapted to promote correct pairing and solubility. Similarly, grafting γδ variable regions onto αβ constant domains could improve structural stability while maintaining CD1c recognition.

Further refinement of functional validation pipelines is also critical. While tetramer staining and SPR provided insight into binding specificity, they do not confirm TCR activation. Jurkat reporter assays could be used to assess functional responses, distinguishing biologically relevant interactions from

non-specific binding. Additionally, optimising lipid-exchange protocols could enable more systematic investigation of lipid-dependent γδ TCR binding.

A broader challenge in γδ TCR research is the lack of high-resolution co-crystal structures. While this study contributed to the pipeline for soluble γδ TCR generation, achieving diffraction-quality crystals remains difficult. Alternative strategies, including glycosylation engineering, co-crystallisation with stabilising ligands, or Fab-assisted lattice formation, may improve structural resolution. Given these challenges, cryo-electron microscopy (cryo-EM) could provide a viable alternative for studying γδ TCR-CD1c interactions, particularly where crystallisation is unsuccessful.

Finally, CD1c specificity of γδ TCRs warrants further investigation, as some TCRs isolated using CD1c-SL tetramers failed to bind CD1c-endo. Differences in lipid presentation likely contributed, but alternative interactions—such as binding to CD36—may have influenced selection. Future studies should integrate Jurkat transduction assays to confirm CD1c specificity before proceeding with soluble TCR expression.

In summary, this study established a framework for generating soluble γδ TCRs and investigating CD1c recognition, highlighting glycosylation, lipid heterogeneity, and sequence variability as key factors. The successful production of a CD1c-reactive γδ TCR provides a foundation for future structural and functional studies. Moving forward, efforts should focus on stabilising γδ TCR heterodimers, refining antigen validation methods, and leveraging cryo-EM to obtain high-resolution insights into γδ TCR recognition of CD1c-lipid complexes. Addressing these challenges will be crucial for advancing unconventional T cell biology and its potential applications in disease pathogenesis and immunotherapy.

Chapter 6 Future work and concluding remarks.

6.1 Future work

6.1.1 Investigating CD1d F' roof flexibility.

Our CD1d(h) structure revealed a flexible Phe84 residue, which, in combination with the flexibility of the flanking α helices, can likely modulate the binding groove cavity area for the accommodation of the vast spectrum of lipid antigens known to be presented by CD1d. Despite conclusive electron density demonstrating these multiple conformations, as well as difference density filling the gap between the loaded αGC and the Phe84 in the canonical position, we were unable to decipher exactly what ligand occupied the groove to stabilise this Phe84 in the open conformation. This is likely due to the complex array of purification and crystallisation buffer molecules that the protein was exposed to during the experimental pipeline, leading to overlapping electron densities of various molecules. Using molecular dynamics, we were able to demonstrate that the Phe84 residue is stable in the non-canonical conformation when a small 4-carbon butane molecule is sat in the space normally occupied by Phe84 in the 'canonical state'. This suggests the molecule can adapt to a range of lipid occupancies, increasing and decreasing the cavity volume to best suit the lipid cargo.

In terms of the biological relevance of CD1d plasticity, we postulate the molecule can expand and constrict its groove to a greater extent than is currently described to accommodate the growing spectrum of lipids presented by CD1d. This includes molecules such as TAG, which would otherwise be too big to fit in the groove. Modelling a molecule such as this may be challenging; however, it would likely indicate a potential mechanism for the accommodation of this lipid in the groove. Further to this, despite being potentially challenging, it would be great to gain structural data of lipids within the groove, which causes the stabilisation of the open conformation of this residue. This would increase our understanding of the exact mechanism used to accommodate these larger lipids and demonstrate the other areas in the CD1 roof that can adapt to accommodate these larger molecules. As well as structural data, it would also be interesting to gain functional insight as to the relevance of this residue and how it can alter the interaction with TCR. Using a combination of tetramer-based assays, SPR, and mutagenesis studies, it may be possible to elucidate exactly which residues surrounding this flexible region of CD1d are most important for CD1d lipid complexes and their recognition by TCRs. One way to do this could be to use alanine substitutions of Phe84 and surrounding areas and investigate the effects of binding of these TCRs on mutated CD1d proteins. Adopting this approach in combination with pulsing experiments using lipids that could potentially alter CD1d cavity structure, would allow us to build a comprehensive understanding of the mechanisms employed by CD1d to accommodate a variety of different lipids.

The challenges faced when attempting to gain a structural insight into the mechanism by which CD1d accommodates larger, more complex lipids are substantial however refolding CD1d in the presence of these lipids is likely not the optimal method due to the relative insolubility of these lipids in solution. One potential method to investigate this would be to use a CD1d endo-based system by which the CD1d molecules could be crystallised in complex with their 'natural' cellular lipids. This could be used in combination with pulsing optimisation experiments to explore the potential binding capacity of CD1d. We hypothesise that CD1d structure is altered substantially following the binding of larger lipid molecules and may develop exit portals not currently described in the literature. However, this phenomenon is hard to prove without the use of structural data.

6.1.2 Investigating the interaction between high and low-affinity iTCRs and CD1d-lipid complexes.

The difference in binding affinity between human iTCRs can likely be attributed to the differential sequences and properties of their hypervariable CDR3β loops. Unfortunately, we were unable to generate co-crystal complexes of CD1d-αGC with a high and a low-affinity TCR and were unable to see the exact contribution of the CDR3 loop to both TCRs. Future work should involve further attempts to co-crystallise these structures to understand the binding mechanism differences. This would likely shed light on whether the CDR3β loop is directly influencing the TCR affinity by physically binding the protein/lipid itself or indirectly via modulating the structure and subsequent affinity of other CDR loops. As well as understanding how the CDR3β sequence partially dictates affinity across different TCRs, we would also like to understand the impact of the CDR3β loop of the same TCR on different CD1d-lipid complexes. There is evidence in the literature suggesting the hierarchy of affinity of different TCRs towards CD1d-lipid complexes remains the same irrespective of the stability of the lipid-bound within CD1d, suggesting the CDR3β is not interacting with the lipid itself. It is unclear, however, whether this phenomenon remains the same for TCRs of high, medium, and low affinity or whether the affinity hierarchy remains the same amongst TCRs of different affinities. To fully understand this, a greater amount of structural data from TCRs of ranging affinities would be required.

Although crystallography would elucidate the physical binding mechanism used by specific TCRs, this may not be applicable across all iTCRs. It is possible that within the pool of high-affinity iTCRs, there are different mechanisms by which the high-affinity is established, be it through the CDR3β loop stabilising other chains or interacting directly with CD1. One way to investigate this further would be to utilise bioinformatic methodologies such as 10X genomics to analyse single-cell RNA sequencing data to investigate differences in CDR3β sequence across large pools of high, medium, and low-affinity

TCRs, something we are currently investigating in-house-. Another future study would be to investigate *in-vitro* and *in-vivo* responses to αGC treatment when comparing high and low-affinity iTCRs.

6.1.3 Generating CD1c-endo and γδ TCRs for structural and functional studies.

Our investigation into our panel of γδ TCRs saw several challenges, including inclusion body expression, refolding capability, and purification issues. We established, through a methodical approach, an optimal process for the purification of γδ TCR heterodimers from complex mixtures of aggregate, contaminants, single chain, homodimer, and heterodimer. We demonstrated successful refolding and purification of a CD1c-reactive γδ TCR, which can now further be characterised in terms of affinity, tested for lipid specificity using our bead-based assay, and set up in complex with CD1c for structural studies. The next steps for the continuation of this project and pipeline are refolding of other CD1c-reactive γδTCRs using our His-tag approach, followed by characterization of TCR affinities for CD1c-endo. These TCRs can then be set up in complex with CD1c-endo for structural studies to further understand the molecular mechanism underpinning γδ TCR reactivity with CD1c. These TCRs can also be investigated for lipid reactivity using our bead assay to understand the types of lipids they can bind to and which lipids abrogate their binding.

Despite soluble, refolded γδ TCRs having their advantages, such as lack of PTM, e.g., glycosylation, the yield and stability remain a problem. There is growing evidence in the literature that γδ TCRs can be successfully produced in mammalian cells, allowing for a potential fast screening approach for unconventional TCRs against a given target. Using this method, CD1c-reactive γδ TCR sequences could be 'plugged in' to a mammalian expression construct and expressed in plate format to screen large amounts of TCRs for CD1c-reactivity. This would be a more efficient way of investigating the interaction between CD1c and γδ TCRs, however, questions remain around stability and TCR produced in this way would likely require treatment with endonucleases prior to crystallisation.

6.2 Conclusions

The capability of the CD1 isoforms to present a wide spectrum of lipid antigens has been described using a combination of structural and functional studies. However, the mechanisms by which individual isoforms can conformationally adapt are less well understood. A greater understanding of the intricate details surrounding CD1 antigen presentation is fundamental in development of TCR based therapeutics utilising these molecules to initiate an immune response. Using x-ray

crystallography, we uncover a new insight into CD1d-αGC antigen presentation, with potential relevance to the presentation of other lipid molecules. Furthermore, we analyse this phenomenon in relation to other CD1 isoforms and species and discover a potentially conserved mechanism across these molecules that incorporates a flexible area of the antigen binding groove to intricately alter the cavity volume and thus generate the potential for a large variety of lipid antigens to bind. Our results suggest that the flexibility in the CD1 grooves, particularly surrounding the Phe84 residue discussed in this thesis, is conserved across isoforms and species and is a fundamental property of these molecules and their function. We also present a new NHP CD1d structure, which demonstrates subtle structural differences to the other CD1d presenting species but predominantly shows conservation of CD1d antigen presentation and functional recognition by TCRs. An upsurge in TCR-CD1 co-crystal data will likely increase our knowledge of the exact contribution of this PHe84 and surrounding residues both to lipid accommodation and TCR recognition and the interplay between these two factors.

As well as the investigation into CD1d structure and antigen presentation, we also investigated differences in the binding affinity of iTCRs for CD1d, with a particular interest in the CDR3β loops. A combination of previous structural and functional data, both in mice and humans, attributed an important but largely indirect role of the CDR3β loop in stabilising other CDR loops involved in direct CD1d binding and subsequently modulating the affinity of the TCR for CD1d. We attempted co-crystallisation of CD1d-αGC (chosen primarily due to stability) with high and low-affinity iTCRs. Unfortunately, we were unable to crystallise the complex. However, we were able to crystallise a low-affinity iTCR along with six other AF2-produced models of varying affinity. Analysis with previously published structures revealed some interesting patterns, which may partially explain the mechanism by which the 1369 TCR was able to elicit a higher affinity for CD1d-αGC. More work needs to be done to fully characterise the interaction between CD1d and TCRs of varying affinity, with an increase in the amount of complex data generation likely proving critical.

Finally, we investigated the interaction between CD1c and another subset of unconventional T cells (γδ). There is evidence of cross-reactivity between CD1 isoforms and unconventional TCRs, with an example being CD1d binding both γδ TCRs and αβ TCRs such as iNKT. This suggests similarities in binding mechanisms across the CD1 isoforms and points towards a complex web of interactions between CD1 lipid-presenting molecules and unconventional T cells to elicit an important part of the immune response. Currently, CD1c is arguably the least understood isoform amongst the CD1 family, with no current structural data of CD1c in complex with a γδ TCR. To further explore this interaction, we attempted to refold CD1c-reactive γδ TCRs using a bacterial expression system to investigate the molecular mechanism underpinning CD1c reactivity. Through use of a highly logical approach, we carefully optimised the γδ refolding and purification procedure by systematically investigating several

factors, including IB purity, chain ratios, and engineering approaches. Through this approach, we developed a method by which we could purify a CD1c-reactive γδ TCR, which we could tetramerise and demonstrate the staining of beads coated in CD1c-endo protein. This process can now be scaled up and replicated for other TCRs we believe to be reactive to CD1c. Co-crystallisation can then be attempted to further investigate the molecular interaction mechanism underpinning CD1c reactivity by unconventional γδ TCRs.

References

1. Dutronc Y, Porcelli SA. The CD1 family and T cell recognition of lipid antigens. Tissue Antigens. 2002;60(5):337-53.
2. Jang JH, Shin HW, Lee JM, Lee HW, Kim EC, Park SH. An Overview of Pathogen Recognition Receptors for Innate Immunity in Dental Pulp. Mediators Inflamm. 2015;2015:794143.
3. Chaplin DD. Overview of the immune response. J Allergy Clin Immunol. 2010;125(2 Suppl 2):S3-23.
4. Hato T, Dagher PC. How the Innate Immune System Senses Trouble and Causes Trouble. Clin J Am Soc Nephrol. 2015;10(8):1459-69.
5. Cui J, Chen Y, Wang HY, Wang RF. Mechanisms and pathways of innate immune activation and regulation in health and cancer. Hum Vaccin Immunother. 2014;10(11):3270-85.
6. Marshall JS, Warrington R, Watson W, Kim HL. An introduction to immunology and immunopathology. Allergy, Asthma & Clinical Immunology. 2018;14(2):49.
7. Kumar H, Kawai T, Akira S. Pathogen Recognition by the Innate Immune System. International Reviews of Immunology. 2011;30(1):16-34.
8. Mortaz E, Alipoor SD, Adcock IM, Mumby S, Koenderman L. Update on Neutrophil Function in Severe Inflammation. Front Immunol. 2018;9:2171.
9. Dev A, Iyer S, Razani B, Cheng G. NF-κB and innate immunity. Curr Top Microbiol Immunol. 2011;349:115-43.
10. Vivier E, Tomasello E, Baratin M, Walzer T, Ugolini S. Functions of natural killer cells. Nature Immunology. 2008;9(5):503-10.
11. Prager I, Watzl C. Mechanisms of natural killer cell-mediated cellular cytotoxicity. J Leukoc Biol. 2019;105(6):1319-29.
12. Rus H, Cudrici C, Niculescu F. The role of the complement system in innate immunity. Immunol Res. 2005;33(2):103-12.
13. Merle NS, Church SE, Fremeaux-Bacchi V, Roumenina LT. Complement System Part I - Molecular Mechanisms of Activation and Regulation. Front Immunol. 2015;6:262.
14. Haapasalo K, Meri S. Regulation of the Complement System by Pentraxins. Front Immunol. 2019;10:1750.
15. Lacy P, Stow JL. Cytokine release from innate immune cells: association with diverse membrane trafficking pathways. Blood. 2011;118(1):9-18.
16. Chi H, Pepper M, Thomas PG. Principles and therapeutic applications of adaptive immunity. Cell. 2024;187(9):2052-78.
17. Van Kaer L, Wu L, Joyce S. Mechanisms and Consequences of Antigen Presentation by CD1. Trends Immunol. 2016;37(11):738-54.
18. Burrows PD, Cooper MD. B cell development and differentiation. Curr Opin Immunol. 1997;9(2):239-44.
19. Elhanati Y, Sethna Z, Marcou Q, Callan CG, Jr., Mora T, Walczak AM. Inferring processes underlying B-cell repertoire diversity. Philos Trans R Soc Lond B Biol Sci. 2015;370(1676).
20. Sebina I, Pepper M. Humoral immune responses to infection: common mechanisms and unique strategies to combat pathogen immune evasion tactics. Curr Opin Immunol. 2018;51:46-54.
21. Qi H. T follicular helper cells in space-time. Nature Reviews Immunology. 2016;16(10):612-25.
22. Forthal DN. Functions of Antibodies. Microbiol Spectr. 2014;2(4):Aid-0019-2014.
23. Rastogi I, Jeon D, Moseman JE, Muralidhar A, Potluri HK, McNeel DG. Role of B cells as antigen presenting cells. Front Immunol. 2022;13:954936.
24. Dempsey PW, Vaidya SA, Cheng G. The art of war: Innate and adaptive immune responses. Cell Mol Life Sci. 2003;60(12):2604-21.

25. Palucka K, Banchereau J. Dendritic cells: a link between innate and adaptive immunity. J Clin Immunol. 1999;19(1):12-25.
26. Janeway C. Immunobiology 5 : the immune system in health and disease. 5th ed ed. New York: Garland Pub. New York; 2001.
27. Harris NL, Ronchese F. The role of B7 costimulation in T-cell immunity. Immunol Cell Biol. 1999;77(4):304-11.
28. Groscurth P, Filgueira L. Killing Mechanisms of Cytotoxic T Lymphocytes. News Physiol Sci. 1998;13:17-21.
29. Peter ME, Krammer PH. The CD95(APO-1/Fas) DISC and beyond. Cell Death Differ. 2003;10(1):26-35.
30. Kang S, Brown HM, Hwang S. Direct Antiviral Mechanisms of Interferon-Gamma. Immune Netw. 2018;18(5):e33.
31. Clark RA. Resident memory T cells in human health and disease. Sci Transl Med. 2015;7(269):269rv1.
32. Zajonc DM, Wilson IA. Architecture of CD1 proteins. Curr Top Microbiol Immunol. 2007;314:27-50.
33. Moody DB, Porcelli SA. Intracellular pathways of CD1 antigen presentation. Nat Rev Immunol. 2003;3(1):11-22.
34. Moody DB, Cotton RN. Four pathways of CD1 antigen presentation to T cells. Curr Opin Immunol. 2017;46:127-33.
35. Dougan SK, Kaser A, Blumberg RS. CD1 expression on antigen-presenting cells. Curr Top Microbiol Immunol. 2007;314:113-41.
36. Chaudhry MS, Karadimitris A. Role and regulation of CD1d in normal and pathological B cells. J Immunol. 2014;193(10):4761-8.
37. Maitra R. CD1b in Review: High TCR Specificity Limits Auto-Reactivity. Biomedical Journal of Scientific & Technical Research. 2019.
38. Zeng Z, Castaño AR, Segelke BW, Stura EA, Peterson PA, Wilson IA. Crystal structure of mouse CD1: An MHC-like fold with a large hydrophobic binding groove. Science. 1997;277(5324):339-45.
39. Ly D, Moody DB. The CD1 size problem: lipid antigens, ligands, and scaffolds. Cell Mol Life Sci. 2014;71(16):3069-79.
40. Wieczorek M, Abualrous ET, Sticht J, Álvaro-Benito M, Stolzenberg S, Noé F, et al. Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. Front Immunol. 2017;8:292.
41. Dellabona P, Consonni M, de Lalla C, Casorati G. Group 1 CD1-restricted T cells and the pathophysiological implications of self-lipid antigen recognition. Tissue Antigens. 2015;86(6):393-405.
42. West HC, Bennett CL. Redefining the Role of Langerhans Cells As Immune Regulators within the Skin. Front Immunol. 2017;8:1941.
43. Van Rhijn I, Ly D, Moody DB. CD1a, CD1b, and CD1c in immunity against mycobacteria. Adv Exp Med Biol. 2013;783:181-97.
44. Coventry B, Heinzel S. CD1a in human cancers: a new role for an old molecule. Trends Immunol. 2004;25(5):242-8.
45. Barral DC, Cavallari M, McCormick PJ, Garg S, Magee AI, Bonifacino JS, et al. CD1a and MHC class I follow a similar endocytic recycling pathway. Traffic. 2008;9(9):1446-57.
46. Sugita M, Grant EP, van Donselaar E, Hsu VW, Rogers RA, Peters PJ, et al. Separate pathways for antigen presentation by CD1 molecules. Immunity. 1999;11(6):743-52.
47. Salamero J, Bausinger H, Mommaas AM, Lipsker D, Proamer F, Cazenave JP, et al. CD1a molecules traffic through the early recycling endosomal pathway in human Langerhans cells. J Invest Dermatol. 2001;116(3):401-8.

48. Cernadas M, Cavallari M, Watts G, Mori L, De Libero G, Brenner MB. Early recycling compartment trafficking of CD1a is essential for its intersection and presentation of lipid antigens. J Immunol. 2010;184(3):1235-41.
49. Zajonc DM, Elsliger MA, Teyton L, Wilson IA. Crystal structure of CD1a in complex with a sulfatide self antigen at a resolution of 2.15 A. Nat Immunol. 2003;4(8):808-15.
50. Zajonc DM, Crispin MD, Bowden TA, Young DC, Cheng TY, Hu J, et al. Molecular mechanism of lipopeptide presentation by CD1a. Immunity. 2005;22(2):209-19.
51. Birkinshaw RW, Pellicci DG, Cheng TY, Keller AN, Sandoval-Romero M, Gras S, et al. αβ T cell antigen receptor recognition of CD1a presenting self lipid ligands. Nat Immunol. 2015;16(3):258-66.
52. Wegrecki M, Ocampo TA, Gunasinghe SD, von Borstel A, Tin SY, Reijneveld JF, et al. Atypical sideways recognition of CD1a by autoreactive γδ T cell receptors. Nat Commun. 2022;13(1):3872.
53. Lopez K, Iwany SK, Suliman S, Reijneveld JF, Ocampo TA, Jimenez J, et al. CD1b Tetramers Broadly Detect T Cells That Correlate With Mycobacterial Exposure but Not Tuberculosis Disease State. Front Immunol. 2020;11:199.
54. Porcelli S, Morita CT, Brenner MB. CD1b restricts the response of human CD4-8- T lymphocytes to a microbial antigen. Nature. 1992;360(6404):593-7.
55. Shahine A. The intricacies of self-lipid antigen presentation by CD1b. Mol Immunol. 2018;104:27-36.
56. Jackman RM, Stenger S, Lee A, Moody DB, Rogers RA, Niazi KR, et al. The tyrosine-containing cytoplasmic tail of CD1b is essential for its efficient presentation of bacterial lipid antigens. Immunity. 1998;8(3):341-51.
57. Huang S, Cheng TY, Young DC, Layre E, Madigan CA, Shires J, et al. Discovery of deoxyceramides and diacylglycerols as CD1b scaffold lipids among diverse groove-blocking lipids of the human CD1 system. Proc Natl Acad Sci U S A. 2011;108(48):19335-40.
58. Shahine A, Van Rhijn I, Cheng TY, Iwany S, Gras S, Moody DB, et al. A molecular basis of human T cell receptor autoreactivity toward self-phospholipids. Sci Immunol. 2017;2(16).
59. Shahine A, Reinink P, Reijneveld JF, Gras S, Holzheimer M, Cheng TY, et al. A T-cell receptor escape channel allows broad T-cell response to CD1b and membrane phospholipids. Nat Commun. 2019;10(1):56.
60. Gras S, Van Rhijn I, Shahine A, Cheng TY, Bhati M, Tan LL, et al. T cell receptor recognition of CD1b presenting a mycobacterial glycolipid. Nat Commun. 2016;7:13257.
61. Van Rhijn I, Moody DB. CD1 and mycobacterial lipids activate human T cells. Immunol Rev. 2015;264(1):138-53.
62. Briken V, Jackman RM, Dasgupta S, Hoening S, Porcelli SA. Intracellular trafficking pathway of newly synthesized CD1b molecules. Embo j. 2002;21(4):825-34.
63. Roy S, Ly D, Li NS, Altman JD, Piccirilli JA, Moody DB, et al. Molecular basis of mycobacterial lipid antigen presentation by CD1c and its recognition by αβ T cells. Proc Natl Acad Sci U S A. 2014;111(43):E4648-57.
64. Wun KS, Reijneveld JF, Cheng TY, Ladell K, Uldrich AP, Le Nours J, et al. T cell autoreactivity directed toward CD1c itself rather than toward carried self lipids. Nat Immunol. 2018;19(4):397-406.
65. Mansour S, Tocheva AS, Cave-Ayland C, Machelett MM, Sander B, Lissin NM, et al. Cholesteryl esters stabilize human CD1c conformations for recognition by self-reactive T cells. Proc Natl Acad Sci U S A. 2016;113(9):E1266-75.
66. Scharf L, Li NS, Hawk AJ, Garzón D, Zhang T, Fox LM, et al. The 2.5 Å structure of CD1c in complex with a mycobacterial lipid reveals an open groove ideally suited for diverse antigen presentation. Immunity. 2010;33(6):853-62.
67. Shin JH, Park SH. The effect of intracellular trafficking of CD1d on the formation of TCR repertoire of NKT cells. BMB Rep. 2014;47(5):241-8.
68. Cernadas M, Sugita M, van der Wel N, Cao X, Gumperz JE, Maltsev S, et al. Lysosomal localization of murine CD1d mediated by AP-3 is necessary for NK T cell development. J Immunol. 2003;171(8):4149-55.

69. Lawton AP, Prigozy TI, Brossay L, Pei B, Khurana A, Martin D, et al. The mouse CD1d cytoplasmic tail mediates CD1d trafficking and antigen presentation by adaptor protein 3-dependent and -independent mechanisms. J Immunol. 2005;174(6):3179-86.
70. Jayawardena-Wolf J, Benlagha K, Chiu YH, Mehr R, Bendelac A. CD1d endosomal trafficking is independently regulated by an intrinsic CD1d-encoded tyrosine motif and by the invariant chain. Immunity. 2001;15(6):897-908.
71. Prigozy TI, Naidenko O, Qasba P, Elewaut D, Brossay L, Khurana A, et al. Glycolipid antigen processing for presentation by CD1d molecules. Science. 2001;291(5504):664-7.
72. Koch M, Stronge VS, Shepherd D, Gadola SD, Mathew B, Ritter G, et al. The crystal structure of human CD1d with and without alpha-galactosylceramide. Nat Immunol. 2005;6(8):819-26.
73. Brossay L, Chioda M, Burdin N, Koezuka Y, Casorati G, Dellabona P, et al. CD1d-mediated recognition of an alpha-galactosylceramide by natural killer T cells is highly conserved through mammalian evolution. J Exp Med. 1998;188(8):1521-8.
74. Zajonc DM. The CD1 family: serving lipid antigens to T cells since the Mesozoic era. Immunogenetics. 2016;68(8):561-76.
75. Angénieux C, Fraisier V, Maître B, Racine V, van der Wel N, Fricker D, et al. The cellular pathway of CD1e in immature and maturing dendritic cells. Traffic. 2005;6(4):286-302.
76. de la Salle H, Mariotti S, Angenieux C, Gilleron M, Garcia-Alles LF, Malm D, et al. Assistance of microbial glycolipid antigen processing by CD1e. Science. 2005;310(5752):1321-4.
77. Garcia-Alles LF, Giacometti G, Versluis C, Maveyraud L, de Paepe D, Guiard J, et al. Crystal structure of human CD1e reveals a groove suited for lipid-exchange processes. Proceedings of the National Academy of Sciences. 2011;108(32):13230-5.
78. Brutkiewicz RR. CD1d ligands: the good, the bad, and the ugly. J Immunol. 2006;177(2):769-75.
79. Zajonc DM, Maricic I, Wu D, Halder R, Roy K, Wong CH, et al. Structural basis for CD1d presentation of a sulfatide derived from myelin and its implications for autoimmunity. J Exp Med. 2005;202(11):1517-26.
80. Rossjohn J, Pellicci DG, Patel O, Gapin L, Godfrey DI. Recognition of CD1d-restricted antigens by natural killer T cells. Nat Rev Immunol. 2012;12(12):845-57.
81. Zajonc DM, Ainge GD, Painter GF, Severn WB, Wilson IA. Structural characterization of mycobacterial phosphatidylinositol mannoside binding to mouse CD1d. J Immunol. 2006;177(7):4577-83.
82. Kashiwase K, Kikuchi A, Ando Y, Nicol A, Porcelli SA, Tokunaga K, et al. The CD1d natural killer T-cell antigen presentation pathway is highly conserved between humans and rhesus macaques. Immunogenetics. 2003;54(11):776-81.
83. Angénieux C, Salamero J, Fricker D, Wurtz JM, Maître B, Cazenave JP, et al. Common characteristics of the human and rhesus macaque CD1e molecules: conservation of biochemical and biological properties during primate evolution. Immunogenetics. 2003;54(12):842-9.
84. Morita D, Katoh K, Harada T, Nakagawa Y, Matsunaga I, Miura T, et al. Trans-species activation of human T cells by rhesus macaque CD1b molecules. Biochemical and Biophysical Research Communications. 2008;377(3):889-93.
85. Mori L, Lepore M, De Libero G. The Immunology of CD1- and MR1-Restricted T Cells. Annu Rev Immunol. 2016;34:479-510.
86. Cotton RN, Shahine A, Rossjohn J, Moody DB. Lipids hide or step aside for CD1-autoreactive T cell receptors. Curr Opin Immunol. 2018;52:93-9.
87. Godfrey DI, Uldrich AP, McCluskey J, Rossjohn J, Moody DB. The burgeoning family of unconventional T cells. Nature immunology. 2015;16(11):1114-23.
88. Pellicci DG, Koay HF, Berzins SP. Thymic development of unconventional T cells: how NKT cells, MAIT cells and γδ T cells emerge. Nat Rev Immunol. 2020;20(12):756-70.
89. Vacchini A, Chancellor A, Spagnuolo J, Mori L, De Libero G. MR1-Restricted T Cells Are Unprecedented Cancer Fighters. Front Immunol. 2020;11:751.

90. Corpuz TM, Stolp J, Kim HO, Pinget GV, Gray DH, Cho JH, et al. Differential Responsiveness of Innate-like IL-17- and IFN-γ-Producing γδ T Cells to Homeostatic Cytokines. J Immunol. 2016;196(2):645-54.
91. Constantinides MG, Link VM, Tamoutounour S, Wong AC, Perez-Chaparro PJ, Han SJ, et al. MAIT cells are imprinted by the microbiota in early life and promote tissue repair. Science. 2019;366(6464).
92. Hung J-T, Huang J-R, Yu AL. Tailored design of NKT-stimulatory glycolipids for polarization of immune responses. Journal of Biomedical Science. 2017;24(1):22.
93. Shissler SC, Webb TJ. The ins and outs of type I iNKT cell development. Mol Immunol. 2019;105:116-30.
94. Ishikawa A, Motohashi S, Ishikawa E, Fuchida H, Higashino K, Otsuji M, et al. A Phase I Study of α-Galactosylceramide (KRN7000)–Pulsed Dendritic Cells in Patients with Advanced and Recurrent Non–Small Cell Lung Cancer. Clinical Cancer Research. 2005;11(5):1910-7.
95. Toyoda T, Kamata T, Tanaka K, Ihara F, Takami M, Suzuki H, et al. Phase II study of α-galactosylceramide-pulsed antigen-presenting cells in patients with advanced or recurrent non-small cell lung cancer. J Immunother Cancer. 2020;8(1).
96. Schneiders FL, Scheper RJ, von Blomberg BME, Woltman AM, Janssen HLA, van den Eertwegh AJM, et al. Clinical experience with α-galactosylceramide (KRN7000) in patients with advanced cancer and chronic hepatitis B/C infection. Clinical Immunology. 2011;140(2):130-41.
97. King LA, Lameris R, de Gruijl TD, van der Vliet HJ. CD1d-Invariant Natural Killer T Cell-Based Cancer Immunotherapy: α-Galactosylceramide and Beyond. Front Immunol. 2018;9:1519.
98. Brigl M, Bry L, Kent SC, Gumperz JE, Brenner MB. Mechanism of CD1d-restricted natural killer T cell activation during microbial infection. Nat Immunol. 2003;4(12):1230-7.
99. Fernandez CS, Cameron G, Godfrey DI, Kent SJ. Ex-vivo α-galactosylceramide activation of NKT cells in humans and macaques. J Immunol Methods. 2012;382(1-2):150-9.
100. Pellicci DG, Patel O, Kjer-Nielsen L, Pang SS, Sullivan LC, Kyparissoudis K, et al. Differential recognition of CD1d-alpha-galactosyl ceramide by the V beta 8.2 and V beta 7 semi-invariant NKT T cell receptors. Immunity. 2009;31(1):47-59.
101. Girardi E, Zajonc DM. Molecular basis of lipid antigen presentation by CD1d and recognition by natural killer T cells. Immunol Rev. 2012;250(1):167-79.
102. Erausquin E, Morán-Garrido M, Sáiz J, Barbas C, Dichiara-Rodríguez G, Urdiciain A, et al. Identification of a broad lipid repertoire associated to the endothelial cell protein C receptor (EPCR). Sci Rep. 2022;12(1):15127.
103. Brennan PJ, Cheng T-Y, Pellicci DG, Watts GFM, Veerapen N, Young DC, et al. Structural determination of lipid antigens captured at the CD1d–T-cell receptor interface. Proceedings of the National Academy of Sciences. 2017;114(31):8348-53.
104. Chancellor A, Gadola SD, Mansour S. The versatility of the CD1 lipid antigen presentation pathway. Immunology. 2018;154(2):196-203.
105. Mallevaey T, Clarke AJ, Scott-Browne JP, Young MH, Roisman LC, Pellicci DG, et al. A molecular basis for NKT cell recognition of CD1d-self-antigen. Immunity. 2011;34(3):315-26.
106. Wun KS, Cameron G, Patel O, Pang SS, Pellicci DG, Sullivan LC, et al. A molecular basis for the exquisite CD1d-restricted antigen specificity and functional responses of natural killer T cells. Immunity. 2011;34(3):327-39.
107. Davitt CJH, Longet S, Albutti A, Aversa V, Nordqvist S, Hackett B, et al. Alpha-galactosylceramide enhances mucosal immunity to oral whole-cell cholera vaccines. Mucosal Immunology. 2019;12(4):1055-64.
108. Wang Y, Bhave MS, Yagita H, Cardell SL. Natural Killer T-Cell Agonist α-Galactosylceramide and PD-1 Blockade Synergize to Reduce Tumor Development in a Preclinical Model of Colon Cancer. Frontiers in Immunology. 2020;11.

109. Uchida T, Nakashima H, Yamagata A, Ito S, Ishikiriyama T, Nakashima M, et al. Repeated administration of alpha-galactosylceramide ameliorates experimental lupus nephritis in mice. Scientific Reports. 2018;8(1):8225.
110. Guillaume J, Wang J, Janssens J, Remesh SG, Risseeuw MDP, Decruy T, et al. Galactosylsphingamides: new α-GalCer analogues to probe the F'-pocket of CD1d. Scientific Reports. 2017;7(1):4276.
111. Chamoto K, Guo T, Imataki O, Tanaka M, Nakatsugawa M, Ochi T, et al. CDR3β sequence motifs regulate autoreactivity of human invariant NKT cell receptors. J Autoimmun. 2016;68:39-51.
112. McCarthy C, Shepherd D, Fleire S, Stronge VS, Koch M, Illarionov PA, et al. The length of lipids bound to human CD1d molecules modulates the affinity of NKT cell TCR and the threshold of NKT cell activation. J Exp Med. 2007;204(5):1131-44.
113. Gadola SD, Koch M, Marles-Wright J, Lissin NM, Shepherd D, Matulis G, et al. Structure and binding kinetics of three different human CD1d-alpha-galactosylceramide-specific T cell receptors. J Exp Med. 2006;203(3):699-710.
114. Matulis G, Sanderson JP, Lissin NM, Asparuhova MB, Bommineni GR, Schümperli D, et al. Innate-like control of human iNKT cell autoreactivity via the hypervariable CDR3beta loop. PLoS Biol. 2010;8(6):e1000402.
115. Chamoto K, Guo T, Scally SW, Kagoya Y, Anczurowski M, Wang CH, et al. Key Residues at Third CDR3β Position Impact Structure and Antigen Recognition of Human Invariant NK TCRs. J Immunol. 2017;198(3):1056-65.
116. López-Sagaseta J, Sibener LV, Kung JE, Gumperz J, Adams EJ. Lysophospholipid presentation by CD1d and recognition by a human Natural Killer T-cell receptor. Embo j. 2012;31(8):2047-59.
117. Kashiwase K, Kikuchi A, Ando Y, Nicol A, Porcelli SA, Tokunaga K, et al. The CD1d natural killer T-cell antigen presentation pathway is highly conserved between humans and rhesus macaques. Immunogenetics. 2003;54(11):776-81.
118. Gansuvd B, Hubbard WJ, Hutchings A, Thomas FT, Goodwin J, Wilson SB, et al. Phenotypic and Functional Characterization of Long-Term Cultured Rhesus Macaque Spleen-Derived NKT Cells 1. The Journal of Immunology. 2003;171(6):2904-11.
119. Morita D, Katoh K, Harada T, Nakagawa Y, Matsunaga I, Miura T, et al. Trans-species activation of human T cells by rhesus macaque CD1b molecules. Biochem Biophys Res Commun. 2008;377(3):889-93.
120. Angénieux C, Salamero J, Fricker D, Wurtz J-M, Maître B, Cazenave J-P, et al. Common characteristics of the human and rhesus macaque CD1e molecules: conservation of biochemical and biological properties during primate evolution. Immunogenetics. 2003;54(12):842-9.
121. Morita D, Hattori Y, Nakamura T, Igarashi T, Harashima H, Sugita M. Major T cell response to a mycolyl glycolipid is mediated by CD1c molecules in rhesus macaques. Infect Immun. 2013;81(1):311-6.
122. Zhao Y, Niu C, Cui J. Gamma-delta (γδ) T cells: friend or foe in cancer development? J Transl Med. 2018;16(1):3.
123. Wu HJ, Wu E. The role of gut microbiota in immune homeostasis and autoimmunity. Gut Microbes. 2012;3(1):4-14.
124. Sebestyen Z, Prinz I, Déchanet-Merville J, Silva-Santos B, Kuball J. Translating gammadelta (γδ) T cells and their receptors into cancer cell therapies. Nat Rev Drug Discov. 2020;19(3):169-84.
125. Donia M, Ellebaek E, Andersen MH, Straten PT, Svane IM. Analysis of Vδ1 T cells in clinical grade melanoma-infiltrating lymphocytes. Oncoimmunology. 2012;1(8):1297-304.
126. Kabelitz D, Marischen L, Oberg HH, Holtmeier W, Wesch D. Epithelial defence by gamma delta T cells. Int Arch Allergy Immunol. 2005;137(1):73-81.
127. Adams EJ, Gu S, Luoma AM. Human gamma delta T cells: Evolution and ligand recognition. Cell Immunol. 2015;296(1):31-40.
128. Paul S, Shilpi, Lal G. Role of gamma-delta (γδ) T cells in autoimmunity. J Leukoc Biol. 2015;97(2):259-71.

129. Correa I, Bix M, Liao NS, Zijlstra M, Jaenisch R, Raulet D. Most gamma delta T cells develop normally in beta 2-microglobulin-deficient mice. Proc Natl Acad Sci U S A. 1992;89(2):653-7.
130. Bigby M, Markowitz JS, Bleicher PA, Grusby MJ, Simha S, Siebrecht M, et al. Most gamma delta T cells develop normally in the absence of MHC class II molecules. J Immunol. 1993;151(9):4465-75.
131. Macleod AS, Havran WL. Functions of skin-resident γδ T cells. Cell Mol Life Sci. 2011;68(14):2399-408.
132. Hoeres T, Smetak M, Pretscher D, Wilhelm M. Improving the Efficiency of Vγ9Vδ2 T-Cell Immunotherapy in Cancer. Front Immunol. 2018;9:800.
133. Kabelitz D, Glatzel A, Wesch D. Antigen recognition by human gammadelta T lymphocytes. Int Arch Allergy Immunol. 2000;122(1):1-7.
134. Wu J, Groh V, Spies T. T cell antigen receptor engagement and specificity in the recognition of stress-inducible MHC class I-related chains by human epithelial gamma delta T cells. J Immunol. 2002;169(3):1236-40.
135. Silva-Santos B. Promoting angiogenesis within the tumor microenvironment: the secret life of murine lymphoid IL-17-producing gammadelta T cells. Eur J Immunol. 2010;40(7):1873-6.
136. Lawand M, Déchanet-Merville J, Dieu-Nosjean MC. Key Features of Gamma-Delta T-Cell Subsets in Human Diseases and Their Immunotherapeutic Implications. Front Immunol. 2017;8:761.
137. Groh V, Rhinehart R, Secrist H, Bauer S, Grabstein KH, Spies T. Broad tumor-associated expression and recognition by tumor-derived gamma delta T cells of MICA and MICB. Proc Natl Acad Sci U S A. 1999;96(12):6879-84.
138. Groh V, Steinle A, Bauer S, Spies T. Recognition of stress-induced MHC molecules by intestinal epithelial gammadelta T cells. Science. 1998;279(5357):1737-40.
139. Spada FM, Grant EP, Peters PJ, Sugita M, Melián A, Leslie DS, et al. Self-recognition of CD1 by gamma/delta T cells: implications for innate immunity. J Exp Med. 2000;191(6):937-48.
140. Paul S, Lal G. Regulatory and effector functions of gamma-delta (γδ) T cells and their therapeutic potential in adoptive cellular therapy for cancer. Int J Cancer. 2016;139(5):976-85.
141. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med. 2015;21(8):938-45.
142. Fergusson JR, Smith KE, Fleming VM, Rajoriya N, Newell EW, Simmons R, et al. CD161 defines a transcriptional and functional phenotype across distinct human T cell lineages. Cell Rep. 2014;9(3):1075-88.
143. Siegers GM, Lamb LS, Jr. Cytotoxic and regulatory properties of circulating Vδ1+ γδ T cells: a new player on the cell therapy field? Mol Ther. 2014;22(8):1416-22.
144. Bernhard S, Hug S, Stratmann AEP, Erber M, Vidoni L, Knapp CL, et al. Interleukin 8 Elicits Rapid Physiological Changes in Neutrophils That Are Altered by Inflammatory Conditions. J Innate Immun. 2021;13(4):225-41.
145. Zarogoulidis P, Katsikogianni F, Tsiouda T, Sakkas A, Katsikogiannis N, Zarogoulidis K. Interleukin-8 and interleukin-17 for cancer. Cancer Invest. 2014;32(5):197-205.
146. Godder KT, Henslee-Downey PJ, Mehta J, Park BS, Chiang KY, Abhyankar S, et al. Long term disease-free survival in acute leukemia patients recovering with increased gammadelta T cells after partially mismatched related donor bone marrow transplantation. Bone Marrow Transplant. 2007;39(12):751-7.
147. Catellani S, Poggi A, Bruzzone A, Dadati P, Ravetti JL, Gobbi M, et al. Expansion of Vdelta1 T lymphocytes producing IL-4 in low-grade non-Hodgkin lymphomas expressing UL-16-binding proteins. Blood. 2007;109(5):2078-85.
148. Kimura Y, Nagai N, Tsunekawa N, Sato-Matsushita M, Yoshimoto T, Cua DJ, et al. IL-17A-producing CD30(+) Vδ1 T cells drive inflammation-induced cancer progression. Cancer Sci. 2016;107(9):1206-14.
149. Yang J, Weinberg RA. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Dev Cell. 2008;14(6):818-29.

150. Mao Y, Yin S, Zhang J, Hu Y, Huang B, Cui L, et al. A new effect of IL-4 on human γδ T cells: promoting regulatory Vδ1 T cells via IL-10 production and inhibiting function of Vδ2 T cells. Cell Mol Immunol. 2016;13(2):217-28.
151. Daley D, Zambirinis CP, Seifert L, Akkad N, Mohan N, Werba G, et al. γδ T Cells Support Pancreatic Oncogenesis by Restraining αβ T Cell Activation. Cell. 2016;166(6):1485-99.e15.
152. Walwyn-Brown K, Pugh J, Cocker ATH, Beyzaie N, Singer BB, Olive D, et al. Phosphoantigen-Stimulated γδ T Cells Suppress Natural Killer-Cell Responses to Missing-Self. Cancer Immunol Res. 2022;10(5):558-70.
153. Morita CT, Jin C, Sarikonda G, Wang H. Nonpeptide antigens, presentation mechanisms, and immunological memory of human Vgamma2Vdelta2 T cells: discriminating friend from foe through the recognition of prenyl pyrophosphate antigens. Immunol Rev. 2007;215:59-76.
154. Tyler CJ, Doherty DG, Moser B, Eberl M. Human Vγ9/Vδ2 T cells: Innate adaptors of the immune system. Cell Immunol. 2015;296(1):10-21.
155. Moser B, Eberl M. gammadelta T cells: novel initiators of adaptive immunity. Immunol Rev. 2007;215:89-102.
156. Bonneville M, O'Brien RL, Born WK. γδ T cell effector functions: a blend of innate programming and acquired plasticity. Nature Reviews Immunology. 2010;10(7):467-78.
157. Viey E, Fromont G, Escudier B, Morel Y, Da Rocha S, Chouaib S, et al. Phosphostim-activated gamma delta T cells kill autologous metastatic renal cell carcinoma. J Immunol. 2005;174(3):1338-47.
158. Dhar S, Chiplunkar SV. Lysis of aminobisphosphonate-sensitized MCF-7 breast tumor cells by Vγ9Vδ2 T cells. Cancer Immun. 2010;10:10.
159. Alexander AA, Maniar A, Cummings JS, Hebbeler AM, Schulze DH, Gastman BR, et al. Isopentenyl pyrophosphate-activated CD56+ {gamma}{delta} T lymphocytes display potent antitumor activity toward human squamous cell carcinoma. Clin Cancer Res. 2008;14(13):4232-40.
160. Li Z, Xu Q, Peng H, Cheng R, Sun Z, Ye Z. IFN-γ enhances HOS and U2OS cell lines susceptibility to γδ T cell-mediated killing through the Fas/Fas ligand pathway. Int Immunopharmacol. 2011;11(4):496-503.
161. Zheng J, Guo Y, Ji X, Cui L, He W. A novel antibody-like TCRγδ-Ig fusion protein exhibits antitumor activity against human ovarian carcinoma. Cancer Lett. 2013;341(2):150-8.
162. Caccamo N, Todaro M, Sireci G, Meraviglia S, Stassi G, Dieli F. Mechanisms underlying lineage commitment and plasticity of human γδ T cells. Cell Mol Immunol. 2013;10(1):30-4.
163. Casetti R, Agrati C, Wallace M, Sacchi A, Martini F, Martino A, et al. Cutting edge: TGF-beta1 and IL-15 Induce FOXP3+ gammadelta regulatory T cells in the presence of antigen stimulation. J Immunol. 2009;183(6):3574-7.
164. Hu C, Qian L, Miao Y, Huang Q, Miao P, Wang P, et al. Antigen-presenting effects of effector memory Vγ9Vδ2 T cells in rheumatoid arthritis. Cell Mol Immunol. 2012;9(3):245-54.
165. Qiu L, Lai R, Lin Q, Lau E, Thomazy DM, Calame D, et al. Autocrine release of interleukin-9 promotes Jak3-dependent survival of ALK+ anaplastic large-cell lymphoma cells. Blood. 2006;108(7):2407-15.
166. Hoelzinger DB, Dominguez AL, Cohen PA, Gendler SJ. Inhibition of adaptive immunity by IL9 can be disrupted to achieve rapid T-cell sensitization and rejection of progressive tumor challenges. Cancer Res. 2014;74(23):6845-55.
167. Reijneveld JF, Ocampo TA, Shahine A, Gully BS, Vantourout P, Hayday AC, et al. Human γδ T cells recognize CD1b by two distinct mechanisms. Proc Natl Acad Sci U S A. 2020;117(37):22944-52.
168. Uldrich AP, Le Nours J, Pellicci DG, Gherardin NA, McPherson KG, Lim RT, et al. CD1d-lipid antigen recognition by the γδ TCR. Nat Immunol. 2013;14(11):1137-45.
169. Luoma AM, Castro CD, Mayassi T, Bembinster LA, Bai L, Picard D, et al. Crystal structure of Vδ1 T cell receptor in complex with CD1d-sulfatide shows MHC-like recognition of a self-lipid by human γδ T cells. Immunity. 2013;39(6):1032-42.
170. Van Rhijn I, Le Nours J. CD1 and MR1 recognition by human γδ T cells. Mol Immunol. 2021;133:95-100.

171. Bai L, Picard D, Anderson B, Chaudhary V, Luoma A, Jabri B, et al. The majority of CD1d-sulfatide-specific T cells in human blood use a semiinvariant Vδ1 TCR. Eur J Immunol. 2012;42(9):2505-10.
172. Russano AM, Bassotti G, Agea E, Bistoni O, Mazzocchi A, Morelli A, et al. CD1-restricted recognition of exogenous and self-lipid antigens by duodenal gammadelta+ T lymphocytes. J Immunol. 2007;178(6):3620-6.
173. Le Nours J, Gherardin NA, Ramarathinam SH, Awad W, Wiede F, Gully BS, et al. A class of γδ T cell receptors recognize the underside of the antigen-presenting molecule MR1. Science. 2019;366(6472):1522-7.
174. Garcia-Alles LF, Giacometti G, Versluis C, Maveyraud L, de Paepe D, Guiard J, et al. Crystal structure of human CD1e reveals a groove suited for lipid-exchange processes. Proc Natl Acad Sci U S A. 2011;108(32):13230-5.
175. Gadola SD, Zaccai NR, Harlos K, Shepherd D, Castro-Palomino JC, Ritter G, et al. Structure of human CD1b with bound ligands at 2.3 A, a maze for alkyl chains. Nat Immunol. 2002;3(8):721-6.
176. Calabi F, Milstein C. A novel family of human major histocompatibility complex-related genes not mapping to chromosome 6. Nature. 1986;323(6088):540-3.
177. Barral DC, Brenner MB. CD1 antigen presentation: how it works. Nature Reviews Immunology. 2007;7(12):929-41.
178. Cohen NR, Garg S, Brenner MB. Chapter 1 Antigen Presentation by CD1: Lipids, T Cells, and NKT Cells in Microbial Immunity. Advances in Immunology. 102: Academic Press; 2009. p. 1-94.
179. Almeida CF, Sundararaj S, Le Nours J, Praveena T, Cao B, Burugupalli S, et al. Distinct CD1d docking strategies exhibited by diverse Type II NKT cell receptors. Nat Commun. 2019;10(1):5242.
180. Reijneveld JF, Ocampo TA, Shahine A, Gully BS, Vantourout P, Hayday AC, et al. Human γδ T cells recognize CD1b by two distinct mechanisms. Proceedings of the National Academy of Sciences. 2020;117(37):22944-52.
181. Paterson NM, Al-Zubieri H, Barber MF. Diversification of CD1 Molecules Shapes Lipid Antigen Selectivity. Molecular Biology and Evolution. 2021;38(6):2273-84.
182. Garzón D, Anselmi C, Bond PJ, Faraldo-Gómez JD. Dynamics of the Antigen-binding Grooves in CD1 Proteins: REVERSIBLE HYDROPHOBIC COLLAPSE IN THE LIPID-FREE STATE*. Journal of Biological Chemistry. 2013;288(27):19528-36.
183. Huang S, Shahine A, Cheng TY, Chen YL, Ng SW, Balaji GR, et al. CD1 lipidomes reveal lipid-binding motifs and size-based antigen-display mechanisms. Cell. 2023;186(21):4583-96.e13.
184. Schiefner A, Fujio M, Wu D, Wong C-H, Wilson IA. Structural Evaluation of Potent NKT Cell Agonists: Implications for Design of Novel Stimulatory Ligands. Journal of Molecular Biology. 2009;394(1):71-82.
185. Batas B, Schiraldi C, Chaudhuri JB. Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography. Journal of Biotechnology. 1999;68(2):149-58.
186. Chaudhuri JB, Batas B, Guise AD. Improving protein refolding yields by minimizing aggregation. Ann N Y Acad Sci. 1996;782:495-505.
187. Matulis G, Sanderson JP, Lissin NM, Asparuhova MB, Bommineni GR, Schümperli D, et al. Innate-Like Control of Human iNKT Cell Autoreactivity via the Hypervariable CDR3β Loop. PLOS Biology. 2010;8(6):e1000402.
188. Wang J, Guillaume J, Pauwels N, Van Calenbergh S, Van Rhijn I, Zajonc DM. Crystal structures of bovine CD1d reveal altered αGalCer presentation and a restricted A' pocket unable to bind long-chain glycolipids. PLoS One. 2012;7(10):e47989.
189. Borg NA, Wun KS, Kjer-Nielsen L, Wilce MC, Pellicci DG, Koh R, et al. CD1d-lipid-antigen recognition by the semi-invariant NKT T-cell receptor. Nature. 2007;448(7149):44-9.
190. Wun KS, Ross F, Patel O, Besra GS, Porcelli SA, Richardson SK, et al. Human and mouse type I natural killer T cell antigen receptors exhibit different fine specificities for CD1d-antigen complex. J Biol Chem. 2012;287(46):39139-48.

191. Wun KS, Cameron G, Patel O, Pang SS, Pellicci DG, Sullivan LC, et al. A molecular basis for the exquisite CD1d-restricted antigen specificity and functional responses of natural killer T cells. Immunity. 2011;34(3):327-39.
192. Hopkins AL, Ren J, Esnouf RM, Willcox BE, Jones EY, Ross C, et al. Complexes of HIV-1 reverse transcriptase with inhibitors of the HEPT series reveal conformational changes relevant to the design of potent non-nucleoside inhibitors. J Med Chem. 1996;39(8):1589-600.
193. Sacchettini JC, Scapin G, Gopaul D, Gordon JI. Refinement of the structure of Escherichia coli-derived rat intestinal fatty acid binding protein with bound oleate to 1.75-A resolution. Correlation with the structures of the apoprotein and the protein with bound palmitate. J Biol Chem. 1992;267(33):23534-45.
194. Cuevas-Zuviría B, Mínguez-Toral M, Díaz-Perales A, Garrido-Arandia M, Pacios LF. Dynamic plasticity of the lipid antigen-binding site of CD1d is crucially favoured by acidic pH and helper proteins. Sci Rep. 2020;10(1):5714.
195. Ponting CP. Biological function in the twilight zone of sequence conservation. BMC Biology. 2017;15(1):71.
196. Capra JA, Singh M. Predicting functionally important residues from sequence conservation. Bioinformatics. 2007;23(15):1875-82.
197. Patel O, Pellicci DG, Gras S, Sandoval-Romero ML, Uldrich AP, Mallevaey T, et al. Recognition of CD1d-sulfatide mediated by a type II natural killer T cell antigen receptor. Nat Immunol. 2012;13(9):857-63.
198. Zhang Y, Springfield R, Chen S, Li X, Feng X, Moshirian R, et al. α-GalCer and iNKT Cell-Based Cancer Immunotherapy: Realizing the Therapeutic Potentials. Front Immunol. 2019;10:1126.
199. Saka H, Kitagawa C, Ichinose Y, Takenoyama M, Ibata H, Kato T, et al. A randomized phase II study to assess the effect of adjuvant immunotherapy using α-GalCer-pulsed dendritic cells in the patients with completely resected stage II–IIIA non-small cell lung cancer: study protocol for a randomized controlled trial. Trials. 2017;18(1):429.
200. Janssens J, Bitra A, Wang J, Decruy T, Venken K, van der Eycken J, et al. 4"-O-Alkylated α-Galactosylceramide Analogues as iNKT-Cell Antigens: Synthetic, Biological, and Structural Studies. ChemMedChem. 2019;14(1):147-68.
201. McPherson A, Gavira JA. Introduction to protein crystallization. Acta Crystallogr F Struct Biol Commun. 2014;70(Pt 1):2-20.
202. Kim Y, Bigelow L, Borovilos M, Dementieva I, Duggan E, Eschenfeldt W, et al. Chapter 3. High-throughput protein purification for x-ray crystallography and NMR. Adv Protein Chem Struct Biol. 2008;75:85-105.
203. Look A, Burns D, Tews I, Roghanian A, Mansour S. Towards a better understanding of human iNKT cell subpopulations for improved clinical outcomes. Front Immunol. 2023;14:1176724.
204. Mallevaey T, Scott-Browne JP, Matsuda JL, Young MH, Pellicci DG, Patel O, et al. T cell receptor CDR2 beta and CDR3 beta loops collaborate functionally to shape the iNKT cell repertoire. Immunity. 2009;31(1):60-71.
205. Krovi SH, Loh L, Spengler A, Brunetti T, Gapin L. Current insights in mouse iNKT and MAIT cell development using single cell transcriptomics data. Seminars in Immunology. 2022;60:101658.
206. Garcia KC, Teyton L, Wilson IA. Structural basis of T cell recognition. Annu Rev Immunol. 1999;17:369-97.
207. Yu ED, Girardi E, Wang J, Zajonc DM. Cutting edge: structural basis for the recognition of β-linked glycolipid antigens by invariant NKT cells. J Immunol. 2011;187(5):2079-83.
208. Tsimberidou A-M, Van Morris K, Vo HH, Eck S, Lin Y-F, Rivas JM, et al. T-cell receptor-based therapy: an innovative therapeutic approach for solid tumors. Journal of Hematology & Oncology. 2021;14(1):102.
209. Chandran SS, Klebanoff CA. T cell receptor-based cancer immunotherapy: Emerging efficacy and pathways of resistance. Immunol Rev. 2019;290(1):127-47.

210. Duan Z, Ho M. T-Cell Receptor Mimic Antibodies for Cancer Immunotherapy. Mol Cancer Ther. 2021;20(9):1533-41.
211. Schiefner A, Wilson IA. Presentation of lipid antigens by CD1 glycoproteins. Curr Pharm Des. 2009;15(28):3311-7.
212. Huang S, Shahine A, Cheng T-Y, Chen Y-L, Ng SW, Balaji GR, et al. CD1 lipidomes reveal lipid-binding motifs and size-based antigen-display mechanisms. Cell. 2023;186(21):4583-96.e13.
213. Rakhshandehroo M, Gijzel SM, Siersbæk R, Broekema MF, de Haar C, Schipper HS, et al. CD1d-mediated presentation of endogenous lipid antigens by adipocytes requires microsomal triglyceride transfer protein. J Biol Chem. 2014;289(32):22128-39.
214. Matsuda JL, Mallevaey T, Scott-Browne J, Gapin L. CD1d-restricted iNKT cells, the 'Swiss-Army knife' of the immune system. Curr Opin Immunol. 2008;20(3):358-68.
215. Mallevaey T, Selvanantham T. Strategy of lipid recognition by invariant natural killer T cells: 'one for all and all for one'. Immunology. 2012;136(3):273-82.
216. Reiser J-B, Grégoire C, Darnault C, Mosser T, Guimezanes A, Schmitt-Verhulst A-M, et al. A T Cell Receptor CDR3β Loop Undergoes Conformational Changes of Unprecedented Magnitude Upon Binding to a Peptide/MHC Class I Complex. Immunity. 2002;16(3):345-54.
217. Yu K, Shi J, Lu D, Yang Q. Comparative analysis of CDR3 regions in paired human αβ CD8 T cells. FEBS Open Bio. 2019;9(8):1450-9.
218. Chamoto K, Guo T, Scally SW, Kagoya Y, Ancruzowski M, Wang C-H, et al. Key Residues at Third CDR3β Position Impact Structure and Antigen Recognition of Human Invariant NK TCRs. The Journal of Immunology. 2017;198(3):1056-65.
219. Mallevaey T, Scott-Browne JP, Matsuda JL, Young MH, Pellicci DG, Patel O, et al. T Cell Receptor CDR2β and CDR3β Loops Collaborate Functionally to Shape the iNKT Cell Repertoire. Immunity. 2009;31(1):60-71.
220. Pellicci DG, Patel O, Kjer-Nielsen L, Pang SS, Sullivan LC, Kyparissoudis K, et al. Differential Recognition of CD1d-α-Galactosyl Ceramide by the Vβ8.2 and Vβ7 Semi-invariant NKT T Cell Receptors. Immunity. 2009;31(1):47-59.
221. Rice MT, von Borstel A, Chevour P, Awad W, Howson LJ, Littler DR, et al. Recognition of the antigen-presenting molecule MR1 by a Vδ3(+) γδ T cell receptor. Proc Natl Acad Sci U S A. 2021;118(49).
222. Chang VT, Crispin M, Aricescu AR, Harvey DJ, Nettleship JE, Fennelly JA, et al. Glycoprotein structural genomics: solving the glycosylation problem. Structure. 2007;15(3):267-73.
223. Cotton RN, Wegrecki M, Cheng TY, Chen YL, Veerapen N, Le Nours J, et al. CD1a selectively captures endogenous cellular lipids that broadly block T cell response. J Exp Med. 2021;218(7).
224. Chancellor A, Tocheva AS, Cave-Ayland C, Tezera L, White A, Al Dulayymi JaR, et al. CD1b-restricted GEM T cell responses are modulated by Mycobacterium tuberculosis mycolic acid meromycolate chains. Proceedings of the National Academy of Sciences. 2017;114(51):E10956-E64.
225. Sugita M, van der Wel N, Rogers RA, Peters PJ, Brenner MB. CD1c molecules broadly survey the endocytic system. Proceedings of the National Academy of Sciences. 2000;97(15):8445-50.
226. Melandri D, Zlatareva I, Chaleil RAG, Dart RJ, Chancellor A, Nussbaumer O, et al. The γδTCR combines innate immunity with adaptive immunity by utilizing spatially distinct regions for agonist selection and antigen responsiveness. Nature immunology. 2018;19(12):1352-65.
227. Roy S, Ly D, Castro CD, Li NS, Hawk AJ, Altman JD, et al. Molecular Analysis of Lipid-Reactive Vδ1 γδ T Cells Identified by CD1c Tetramers. J Immunol. 2016;196(4):1933-42.
228. Gunasekaran K, Pentony M, Shen M, Garrett L, Forte C, Woodward A, et al. Enhancing Antibody Fc Heterodimer Formation through Electrostatic Steering Effects: APPLICATIONS TO BISPECIFIC MOLECULES AND MONOVALENT IgG. Journal of Biological Chemistry. 2010;285(25):19637-46.
229. Liu P, Gao X, Lundin V, Shi C, Adem Y, Lin K, et al. Probing the Impact of the Knob-into-Hole Mutations on the Structure and Function of a Therapeutic Antibody. Analytical Chemistry. 2020;92(1):1582-8.

230. Xu Y, Lee J, Tran C, Heibeck TH, Wang WD, Yang J, et al. Production of bispecific antibodies in "knobs-into-holes" using a cell-free expression system. MAbs. 2015;7(1):231-42.
231. Mazumder R, Morampudi KS, Motwani M, Vasudevan S, Goldman R. Proteome-wide analysis of single-nucleotide variations in the N-glycosylation sequon of human genes. PLoS One. 2012;7(5):e36212.
232. Sriram V, Willard CA, Liu J, Brutkiewicz RR. Importance of N-linked glycosylation in the functional expression of murine CD1d1. Immunology. 2008;123(2):272-81.
233. Haig NA, Guan Z, Li D, McMichael A, Raetz CR, Xu XN. Identification of self-lipids presented by CD1c and CD1d proteins. J Biol Chem. 2011;286(43):37692-701.
234. Cao TP, Shahine A, Cox LR, Besra GS, Moody DB, Rossjohn J. A structural perspective of how T cell receptors recognize the CD1 family of lipid antigen-presenting molecules. J Biol Chem. 2024;300(8):107511.
235. Szoke-Kovacs R, Khakoo S, Gogolak P, Salio M. Insights into the CD1 lipidome. Front Immunol. 2024;15:1462209.
236. Born WK, Kemal Aydintug M, O'Brien RL. Diversity of γδ T-cell antigens. Cell Mol Immunol. 2013;10(1):13-20.
237. Serra P, Garabatos N, Singha S, Fandos C, Garnica J, Solé P, et al. Increased yields and biological potency of knob-into-hole-based soluble MHC class II molecules. Nature Communications. 2019;10(1):4917.
238. Guo T, Koo MY, Kagoya Y, Anczurowski M, Wang CH, Saso K, et al. A Subset of Human Autoreactive CD1c-Restricted T Cells Preferentially Expresses TRBV4-1(+) TCRs. J Immunol. 2018;200(2):500-11.
239. Xiang X, Gao LM, Zhang Y, Zhu Q, Zhao S, Liu W, et al. Identifying CD1c as a potential biomarker by the comprehensive exploration of tumor mutational burden and immune infiltration in diffuse large B cell lymphoma. PeerJ. 2023;11:e16618.