Capture and Molecular Control of CD1a-Autoreactive T Cells in Human Skin

Capture and molecular control of CD1a-autoreactive T cells in human skin

A dissertation presented

by

Rachel Nicole Cotton

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Immunology

Harvard University

Cambridge, Massachusetts

April 2020

Dissertation Advisor: D. Branch Moody, MD Rachel Nicole Cotton

Capture and molecular control of CD1a-autoreactive T cells in human skin

ABSTRACT

Functional studies of T cell response have revealed the existence of CD1a autoreactive T cells in human skin, and CD1a tetramers might directly enumerate T cells (TCRs) and dissect the molecular basis for activation. We employed CD1a tetramers without *a priori* selection of bound lipids, finding that untreated CD1a tetramers carrying unspecified cellular lipids stained a large skin T cell pool (0.1-13 %) in every subject tested (n=9). Tetramer-selected cells were CD4-biased and produced IL-22 and other cytokines in a CD1a-dependent manner. We performed mass spectrometry of eluted CD1a ligands, and found that tetramer staining occurred in the presence of ~100 distinct lipids. This pattern differs from CD1-lipid and MHC-peptide recognition models where defined antigen is an absolute requirement for binding. Tetramers treated with representative CD1a ligands - diacylglycerol, ceramide, hexosylceramide, phosphatidylcholine, or lysophosphatidylcholine – similarly stained skin T cell lines. Staining inversely correlated with lipid head-group size: sulfatide, sphingomyelin, and phosphatidylinositol inhibited CD1a binding to T cells. Further, mutational CD1a tetramer studies pointed to a candidate TCR binding surface on the outside of CD1a required for recognition. Thus, CD1a-autoreactive T cells are unexpectedly abundant in skin and mainly recognize the surface of CD1a, raising the question of what mechanisms regulate this T cell pool with autoimmune potential. By comparing the profiles of total cellular lipids against those captured by CD1a, we found that CD1a ligands were strongly biased toward sphingomyelins (SM), with >40-fold enrichment of a long chain, unsaturated (C42:2, combined lengths) SM over the C34:1

form dominant in cells. 42:2 SM strongly inhibited CD1a tetramer binding to skin derived CD1a autoreactive T cells from all donors tested, identifying 42:2 SM as an endogenous antagonist of CD1a recognition. Crystallography of CD1a-SM complexes revealed that longer lipid chains in 42:2 SM force the phosphocholine group through the roof of CD1a, disrupting the putative TCR contact surface. However, an SM just 6 carbons shorter was mostly contained within CD1a and stabilized CD1a surface residues. Our studies demonstrate molecular features of lipids preferentially captured by CD1a match those interfering with autoreactive CD1a-TCR binding, informing design of therapeutic inhibitors for CD1a-mediated skin inflammation.

ACKNOWLEDGMENTS

Above all, I would like to thank my thesis advisor, Branch Moody. You without a doubt care deeply about mentorship and training – thank you for the space to run with the CD1a work in skin and for facilitating the numerous interdisciplinary collaborations that have introduced me to new methods and answerable questions. I have especially enjoyed the process of writing several review articles together, because creating, discussing, and illustrating new models for T cell recognition of CD1 complexes that fly in the face of dogma is fun. Five years on, we may never ever quite understand each other's processes in writing or agree on the Oxford comma, but I value our similarly high standards for clear and aesthetically pleasing figure design. Thank you for pushing me to be a better scientific communicator, every week in Wednesday lab meeting, by reminding me to keep it simple with a "Texas 2-step" interpretation, to slow down, say what I really mean in fewer words (still working on that one), and that I need not reinvent the wheel (or experiment or figure design) if there is a collaborator to reach out to.

Thank you to all current and former members of the Moody Laboratory that I've had the pleasure of working and learning with, but especially Tan-Yun Cheng, Ildiko Van Rhijn, Annemieke de Jong, Josephine Fredereik "Fred" Reineveld, and Sara Suliman. Tan-Yun, thank you for your scientific contributions to the work herein, and more importantly thank you for your tireless support of all of the students and trainees. You go above and beyond to help others, to answer the experimental question as rigorously as possible in your own studies, and do so while managing the lab and multiple collaborations. We even moved the lab. I appreciate you.

Thank you to my dissertation advisory committee – Rachael Clark, Kai Wucherpfennig, and Michael Brenner – for bearing with me in the rather slow moving beginning of my PhD while I optimized tetramer assays, making plans for the single cell studies that ultimately shifted to CD1 ligand discovery studies. Thank you for the thoughtful and practical discussions in

prioritizing the several directions in which I could take the work. Thank you also to the members of my dissertation exam committee – Rachael Clark, Stephanie Dougan, Jim Moon, and Muzz Haniffa. To Rachael Clark in particular, thank you for all of the enthusiastic discussions about CD1a autoreactivity in skin disease over the years, potential TCR stereotypes, support for the work through the Human Skin Disease Resource Center, and the chances to share it with a broader dermatology audience. Jess Teague, thank you for all of your work in processing and organizing all of the human skin samples used in this study, and for invaluable technical advice.

The most enjoyable part of my PhD has been the experimental and interdisciplinary forward motion we've been able to do through The Wellcome Trust Collaborative Award group. There's no other group I'd rather wake up at 0430 on the 4th of July to talk science with. Special thanks to Graham Ogg and Jamie Rossjohn, who, with our bi-monthly Zoom calls the past 2 years, felt like additional mentors. Also a huge thank you to Marcin Wegrecki, not only for contributing CD1a sphingomyelin structural studies in Chapter 3, but also for the many discussions on interpreting certainty and novelty in these data and the many figure drafts over oddly timed emails between Boston and Monash. Its rather fitting then that in the wake of COVID-19 and the cancellation of our would-be Boston grant meeting, my dissertation defense seminar will also be a remote event via Zoom call.

To the global #CD1lyfe and 'CD1 Club' faculty, trainees, and students: you have been some of the most fun and supportive friends and colleagues, intellectually stimulating conversationalists, sometimes tough reviewers, and nocturnal experimentalists (looking at you Nelson LaMarche and Carlos Donado) a fellow scientist could hope for. To Mitch Kronenberg, thank you for giving me what was technically my first paying job in high school as a summer intern in your lab, and the unique opportunity to learn T cell and dendritic cell biology first

through the CD1 and NKT lens. Now going on 11 years later, I continue to be excited by the questions and therapeutic applications at the intersection of innate and adaptive immunity.

Thank you to my parents, who were unknowingly preparing me well for my immunology PhD. Mom, thank you for instilling a love of words and books, for making us kids do our summer math workbooks (SCIENCE IS STORYTELLING, but also MENTAL MATH), for letting me wander off and get lost more than a few times (JUST DO THE EXPERIMENT), and for recounting with great imagery at the dinner table your most cringe-worthy stories as an ER nurse (worse than cleaning the tissue culture waste vacuum traps). Thank you for never enforcing bed times, meal times, or being on-time, but always insisting on school uniform code, homework done well, writing thank you notes, and showing up. (DO THE LATE NIGHT EXPERIMENTS, but BE AT EARLY MORNING LAB MEETING). Dad, thank you for showing me that handstand contests are the solution to boredom and a quick way to make friends in any country or at any conference. Thank you for teaching me on-the-fly bicycle and car maintenance, fixing most anything with duct tape, bungee cords, or super glue, and how to find numerous ways in, around, through, or over when we inevitably misplaced keys (SCIENCE NEVER WORKS THE FIRST TIME, THINGS BREAK, TAKE CALCULATED RISKS, IMPROVISE). Thank you for teaching me through diving that when I crash or fail I need to get back on the board to do it 10 more times (STATISTICAL SIGNIFICANCE). Thanks for your sub-functional Factor XI allele, allowing us to brandish spectacular multicolored bruises for weeks when recounting admittedly minor sports injuries (TELL A GOOD STORY, but SHOW ME THE DATA, DATA DISPLAY MATTERS, and IF YOU DIDN'T DOCUMENT IT, IT DIDN'T HAPPEN).

Finally, I am so grateful to my circle of friends, mentors, and colleagues who have at times felt like family over my last decade of science training. Mary Ann McDowell and Roshanak Semnani, thank you for taking a chance on me early in my scientific career, all of your

mentorship and training in the laboratories at NIH and at Notre Dame, the encouragement to “keep pushing forward”, and the opportunities to drive projects and mentor other trainees, write papers, and give talks at national meetings. I would not have made it to graduate school without you. Thank you for everything. Kelsey Finn, my roommate throughout graduate school and my COVID-19 quarantine buddy, who will be defending her Immunology PhD a week before I do: I find it hard to imagine the last six years with anyone else. I appreciate you. Congratulations to you and thanks for being there. Peter Shankman, thank you for being my person during the difficult middle stretch of my PhD, for our many adventures together, and for being my early morning workout buddy, mostly remotely. Sarah Romano, Julie and Chas Miller, and Emily, Aurora, and Charlie Miller – evidence that one can find famly via Craigslist – thank you for being my home away from school since 2012. Aurora, now almost 6, is living evidence of just how long I have been in graduate school. I remember one early morning writing the 2018 Current Opinion in Immunology review from your kitchen, nervously watching then-3.5-year old Aurora try to pour cereal for herself, while on the phone with Branch discussing the original source of the BK6 T cell receptor, left-shifted binding, and Steve Porcelli’s 1995 predictions for the CD1 system. Kudos to all parents working from home with small children. Chas and Charlie, you know intimately well and are living reminders to me of how incredibly powerful the immune system can be, but also how tunable it is when we immunologists get it right. Julie, thank you for all of your questions that have helped me learn how to communicate immunology jargon in actionable and understandable ways.

DEDICATION

To Robert “Uncle Robby” Cotton,

and

my Grandmother, Yvonne Cotton

TABLE OF CONTENTS

CHAPTER 1: Introduction

I. The CD1 system

A brief history of discovery

In 1979 at Oxford, Andrew McMichael and Caesar Milstein reported studies of mice immunized with human thymocytes in order to produce monoclonal antibodies to differentiate classes of human T lymphocytes. Caesar Milstein was the discoverer of monoclonal antibodies, and among those they generated, the target of antibody NA1/34 became one of the first thymocyte differentiation antigens to be defined serologically. The NA1/34 antibody identified a family of proteins that were highly similar in size to Human Leukocyte Antigen (HLA) class I, but had a reciprocal expression pattern to class I in thymocytes in the sense that class I was found on thymic epithelium and mature T cells, and the class I-like NA1/34 target was found in cortical thymocytes[1]. In 1980 Reinherz and colleagues reported that antibody clone OKT6 bound a target with highly similar qualities to that bound by NA1/34: the OKT6 target was expressed on cortical thymocytes and some T cell malignancies, but absent from mature T cells in circulation or medullary thymocytes[2]. In 1984, the International Workshop on Human Leukocyte Differentiation Antigens[3], which used monoclonal antibodies as the organizing principle of cluster of differentiation (CD) antigens would give a name to the antigen(s) recognized by NA1/34, OKT6, D47, and M241: CD1. Now, the target of NA1/34 and OKT6 is known to be the heavy chain for CD1a, which binds non-covalently with the β2-microglobulin light chain.

In 1989, Porcelli, Brenner and colleagues first reported $CD4^{neg}CD8^{neg}$ $\alpha\beta$ T cell clone BK6, isolated from the blood of a patient with systemic lupus erythematosus, which was activated by and lysed target cells expressing CD1a[4]. In these studies, CD1 proteins themselves expressed on cells appeared to be sufficient for recognition, and did not require the addition of exogenous foreign antigen. Subsequently, proof of principle for CD1 as an antigen presenting molecule for lipids was reported in 1994 when Beckman, with Porcelli and

colleagues, discovered exogenous mycolic acid as a CD1b-restricted antigen specifically recognized by the DN1 T cell line[5]. Based on apparent sequence-predicted structural homologies[6] to the recently proven interactions of T cell receptors, peptide antigens, and HLA proteins, Porcelli [7] proposed that CD1 likely bound to self and foreign lipid antigens and displayed them to T cell receptors, but noted that other functions as a putative co-receptor could not yet be formally ruled out.

Between 2003 and 2005, Zajonc, Wilson and colleagues solved crystal structures of CD1a bound to a self or a bacterial lipid antigen[8, 9], proving that CD1a-lipid complexes exist and showing how a glycolipid and a lipopeptide could be exposed for TCR recognition. This core model of dual recognition of CD1 and lipid by a TCR has been the dominant model for antigen recognition in the CD1a system for two decades. In the experiments that follow, my data suggest a related but distinct molecular mechanism in which lipids are not always contributing epitopes that cause T cell activation. Chapter 2 provides evidence that lipids bound inside CD1 can be ignored and must act to get out of the way of the approaching TCR. In this model, it is CD1a itself rather than the carried lipid that is recognized. In Chapter 3, I describe the discovery of natural endogenous lipids that block responses by interfering with CD1a-TCR contact. First, I describe the broader cellular pathways for lipid antigen presentation by CD1a and how they differ from the related roles of human CD1b, CD1c and CD1d proteins.

Lipid capture and presentation by CD1

If CD1 proteins bind lipids and are recognized by T cell receptors, then what types of complexes are recognized and how frequent and prevalent are the T cells that recognize a particular specific antigen complex? In the discovery of antigens and antigen presenting molecule complexes that are targets of T cell recognition, it is useful to frame questions around two qualities with respect to the antigen presenting molecules themselves: location and structure. First, location deals with identifying what cellular or extracelluar compartments a particular antigen presenting molecule regularly visits and can survey. Second, differences in shapes and sizes of antigen binding grooves or clefts will favor binding of some molecules but exclude others, subject to factors like pH, protease or lipase activity, or binding partners to facilitate exchange.

MHC I and MHC II proteins traverse largely separate subcellular pathways to capture peptide antigens, and points of overlap are known as cross-presentation [10]. Likewise, cellular pathways of lipid antigen capture by the four types of human CD1 antigen presenting molecules show both overlap and divergence (**Fig 1.1**). CD1a, CD1b, CD1c and CD1d navigate the secretory pathway in a similar manner: nascent CD1 heavy chains fold in the endoplasmic reticulum (ER), bind β2 microglobulin (β2m), capture endogenous self lipids and egress to the cell surface [11, 12, 13]. Thereafter, the mechanisms of lipid antigen capture and trafficking diverge. While at the cell surface, CD1a [14] and CD1c [15, 16] readily capture exogenous lipids, but CD1b [17, 18] and CD1d [19, 20] do this to a lesser degree. All four protein types enter the endosomal network, but do so using distinct mechanisms. CD1a remains predominantly at the surface at steady state and undergoes an inefficient and shallow recycling pathway mediated by yet to be identified mechanisms of guidance [21, 22, 23]. CD1b, CD1c and CD1d traffic through endosomes using tyrosine residues in their cytoplasmic tails to bind to the μ-subunits of adaptor protein

complexes (AP). Human CD1d and CD1c bind AP-2, targeting them for delivery to early and late endosomal compartments [15, 16]. The particular tyrosine motif present in human CD1b (and mouse CD1d) mediates binding to AP-2 and AP-3, which provides strong redirection to late endosomes and lysosomes [17, 20, 24].

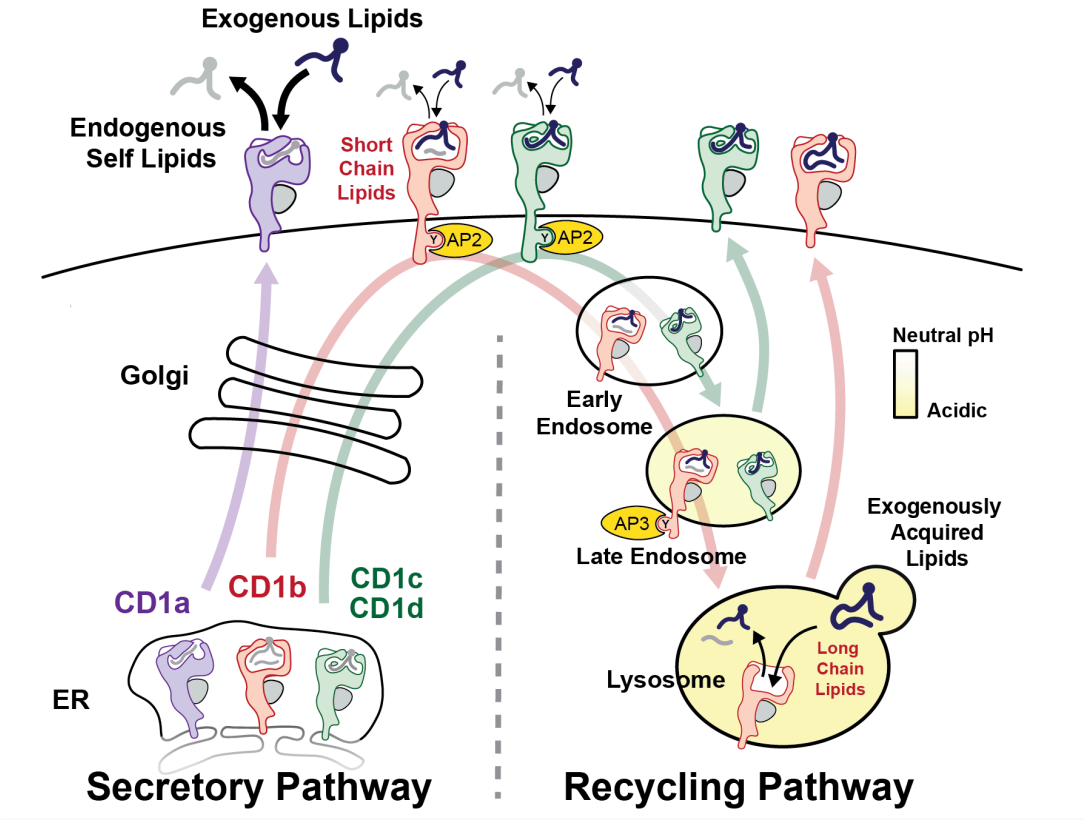

Figure 1.1. Pathways of lipid antigen capture by human CD1 proteins. All four types of human CD1 proteins bind endogenous self lipids and egress to the cell surface through the secretory pathway. Exchange of lipid cargo predominantly occurs at the surface for CD1a and has been observed for CD1c. CD1b, CD1c and CD1d proteins enter the endosomal network for specialized loading reactions prior to recycling to the surface for contact with TCRs. ER, endoplasmic reticulum; AP, adaptor protein complex. Figure modified from the original published in Current Opinion in Immunology (Moody DB & Cotton RN, 2017).

The extent to which each CD1 protein recycles through the endosomal network correlates with the degree to which an acidic pH is needed for lipid antigen capture. CD1c and CD1d capture antigens less efficiently when their endolyosomal targeting motifs are altered, and

CD1c and CD1d bind antigen more efficiently when at an acidic pH in loading assays [15, 16, 19]. CD1b has higher [17] or absolute requirements for an acidic pH when loading lipids, especially long chain lipids [18]. Acid protonates CD1b residues to release ionic bonds that form interdomain tethers that normally connect the α1 and α2 helices at neutral pH, and acid may otherwise reversably denature CD1 proteins to provide access to their clefts [25], allowing the formation of durable CD1-lipid complexes. Acid also permits the activity of acid-dependent proteases like cathepsins, which cleave prosaposin to activate this and other lipid transfer proteins [26, 27, 28]. Thus, the modern view is that the four CD1 proteins initially operate on the same track for capture of endogenous self lipids, but then each of the four human CD1 protein types fan out to distinct locations within the endosomal network for capture of exogenous lipids (**Fig 1.1**).

Efforts to map peptide epitopes are governed fundamentally by the limitations on antigen size and shape by the structure of MHC grooves. The closed ends on MHC I restrict binding to nonamer peptides [29, 30, 31, 32, 33], whereas the open-ended groove of MHC II binds longer peptides with ragged ends[34]. Further, key residues at named positions 1 through 9 of the peptide make specific charge-charge or hydrogen bonding interactions with the MHC I groove, so sequence specific motifs are known, and antigenic peptide binding to particular MHC alleles can be predicted algorithmically [35]. In the CD1 system, the size and architecture of the antigen-binding clefts present in the four types of human CD1 antigen presenting molecules have been described [9, 36, 37, 38]. They likewise impose size limits on and determine the ligands bound, which are lipids and small molecules with most having an overall mass between 200 and 1000 mu [39, 40]. Because CD1 genes are non-polymorphic, any motifs that specifically mediate binding to CD1a, CD1b, CD1c, or CD1d would presumably apply to all humans. Nevertheless, CD1 isoform specific motifs have been slow to emerge.

The clefts present in MHC proteins are known as ‘grooves’ because no surface covers the top of the cavity. The clefts of CD1 proteins however are substantially covered by structures known as A' roofs, which form cave-like cavities located within the globular α1-α2 superdomain of CD1 proteins. Owing to the roof structures, CD1 clefts have a defined molecular volume. CD1a, CD1c, and CD1d have two pockets, known as A’ and F’, which are analogous to the A and F pockets in MHC I [41]. Clefts present in these three proteins are somewhat similar in volume and generally capture lipids with an overall alkyl chain length of C36-42 [9, 36, 37, 38]. However, the shape and interconnectedness of A’ and F’ pockets differ. In CD1c and CD1d, the A' pocket is a toroid that encircles a pole (formed by residue F70) [37, 38]. Lipids can traverse this donut-shaped cavity in a clockwise or counterclockwise direction [42]. In CD1a, the A' pocket is somewhat truncated and the toroid is interrupted, such that the A' pocket is shaped like a bent tube (**Fig. 1.2**) [9]. Crystallography of CD1a-lligand complexes revealed an A’ pole (formed by F70 and V12)[43]. The size, flexibility, and chemistry of the particular ligand bound in CD1a determines whether or not that ligand will insert deeply into the the 'bent tube' and wrap partially around the A’ pole to sit deeply in the pocket. Alternatively, some ligands sit at the junction of the A' and F' pockets, and therefore, fail to insert deeply into the bent, A' pocket. This choke point has appeared to endow CD1a with the ability to bind small ligands that are not expected to fully occupy the cleft, such as urushiol and farnesol [43, 44]. CD1c adopts different conformations depending on ligands captured[45], and this occurs with CD1a to a degree as well [8, 9, 43, 46, 47].

T cell direct recognition of CD1 proteins

A hallmark of T cell mediated immunity is T cell receptor (TCR) co-recognition of peptide- MHC complexes. The term ‘co-recognition’ refers to TCRs that are highly specific for the peptide antigen and the MHC encoded antigen presenting molecule [48, 49]. TCRs are

restricted to a particular MHC allomorph. Unlike MHC-proteins that present peptide antigen in an open groove spanning the length of the molecule, CD1 proteins bind lipids and small molecules deep inside hydrophobic pockets. There is only a small opening, which is shifted off center and known as the F' portal, where a lipid head group moiety can emerge for TCR contact. Lipids protrude somewhat ectopically on the CD1 platform rather than the center. At the opposite end of the CD1 platform, a closure known as the A' roof over the CD1a hydrophobic pocket creates a large platform where TCRs might contact CD1 itself rather than protruding lipid epitopes (**Fig. 1.2**). If this mode of CD1-centric, lipid non-specific binding occurred broadly, than any T cell coming in contact with its cognate CD1 protein in circulation or tissue could hypothetically become activated and produce cytokine.

The 1989 report of the BK6 T cell clone recognizing CD1a proteins first raised the concept of CD1a autoreactivity for one T cell clone. Broader study of this concept on polyclonal T cell basis not resurface until >20 years later, when Annemieke de Jong and Claudia de Lalla published studies demonstrating that the recognition of CD1a proteins without addition of exogenous antigen was actually a frequent pattern of response among blood-derived T cells from healthy individuals without autoimmune disease[50, 51]. Indeed, both studies, using different methods of measure T cell response, found that polyclonal responses to CD1a in blood T cells were more frequent than responses to CD1b, CD1c or CD1d. Two general models of TCR recognition could potentially explain this finding: CD1 and lipid co-recognition or CD1-centric binding.

First, the TCR might substantially contact a common CD1-bound self lipid, as suggested in the earliest models and analogous to co-recognition in the peptide-MHC system. Binary structures of CD1-lipid complexes, in which the lipid head group protrudes substantially onto the surface of CD1a, CD1b, CD1c or CD1d, suggest that, in general, lipid head group moieties are

displayed for contact by approaching TCRs. Ternary structures have shown that TCRs contacted both lipid and CD1d [52], or lipid and CD1b [53], mechanistically ruling in CD1-restriction and lipid-specificity in a model designated as 'headgroup recognition' (**Fig. 1.2**).

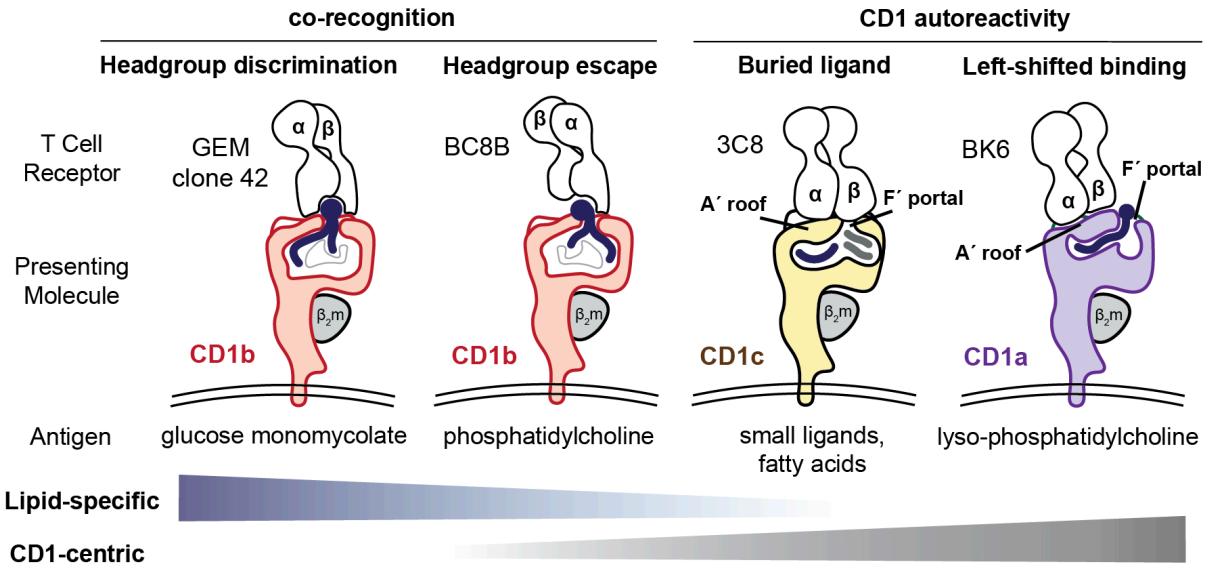

Figure 1.2. Four models for CD1 recognition by T cell receptors. In head group discrimination the TCR binds directly to the protruding carbohydrate present in glucose monomycolate bound to CD1b. In headgroup escape, the TCR binds CD1b and glycerophospholipids like phosphatidylcholine, but is cross-reactive to many phospholipids due to recognition of the common glycero-phosphate “neck” region, and an “escape” channel for the headgroup to avoid TCR contact or disruption[54]. In the left-shifted binding model, the CD1a-autoreactive BK6 TCR contacts the CD1a A’ roof but not the lysophosphatidylcholine ligand. Figure updated and modified from Current Opinion in Immunology, (Moody DB & Cotton RN, 2017), (Cotton RN, *et al,* 2018).

In a related model of semi-specific contact, one TCR was found to bind CD1b and the glycero-phosphate moiety at the “neck” region present in many types of membrane phospholipids. However, recognition was cross-reactive to different phospholipid head groups. Ternary structures revealed that the head group moiety sharply bends and escapes laterally so that the class defining moieties bound to phosphatidic acid avoid direct TCR contact [54, 55]. Thus, another model of co-recognition, known as head group escape, is that the TCR

recognizes core epitope common to many self lipids as well as some aspect fo the surface of CD1 proteins (**Fig. 1.2, red**).

A second general possibility for cargo-agnostic recognition of CD1 proteins is that the approaching TCR entirely fails to contact the lipid antigen, and binding is instead CD1-centric, or CD1-autoreactive. For CD1c, a ternary crystal was solved in which the lipids are buried deeply within CD1c, leading to the 'buried ligand' model (**Fig. 1.2**, **yellow**). Here the TCR binds residues on the outer surface of CD1c surrounding the F' portal, but no part of the carried lipid protrudes through the portal to make contact. For TCR recognition of CD1a proteins, the identification of small, hydrophobic, "headless" epidermal CD1a ligands first hinted that lipids might nest deeply within CD1 and not protrude to the surface[56]. Finally, the first CD1a-lipid-TCR ternary crystal structure, which was published at the time I began my studies[46], revealed an unexpected mode of CD1a-TCR contact: that a CD1a-autoreactive TCR, BK6, bound the closed CD1a platform (A' roof) but not a lysophosphatidylcholine moiety carried and presented by CD1a. The A' roof is typically shown on the left side of CD1 clefts, and the F' portal is shown in the right (**Fig. 1.2, purple**). Thus, I referred to this as a 'left shifted TCR' that fails to see an antigen located on the right side of CD1a. Here the lack of TCR contact occurs due to left-right mismatch rather burying of the ligand fully within the cleft of CD1a **(Fig 1.2)**.

Unlike the MHC-peptide 'co-recognition model' where the TCR simultaneously contacts the antigen and the antigen presenting molecule, direct recognition of CD1a or CD1c proteins is a new model that emerged from structural studies, as I was initiating the experiments in Chapter 2 and 3. Thus, a goal of the experimental work was to understand the extent to which lipiid antigen dependent or antigen independent modes of CD1a recognition were operative among polyclonal skin T cells.

II. Convergence of the CD1a system in human skin

In 1981, less than a year following the discovery of CD1 as a cortical thymocyte and T cell malignancy marker, Fithian and Edelson at Columbia reported concentrated expression of CD1 antigens on epidermal Langerhans cells via the OKT6 antibody [57]. Subsequent serologic studies revealed intermediate and low levels of expression on dermal dendritic cells and keratinocytes respectively[58, 59]. We focused on T cells from skin for several reasons. First, CD1a is highly expressed predominantly in skin[60]. Second, epidermal lipids and lipids from sebum have been shown to bind CD1a and favor the activation of CD1a autoreactive T cell clones[56]. Third, T cells expressing skin homing receptors are enriched for CD1a-autoreactive responses[50]. Based on these observations, we hypothesized that T cells residing in or migrating through skin would be enriched for binding CD1a tetramers.

Human skin is a particularly T cell-dense tissue site, harboring ~20 billion T cells in an average sized adult male, twice as many as in circulation[61, 62]. Among these are long-lived skin resident memory T cells ($CD69^+$) in the dermis ($CD103^{neg}$) and epidermis ($CD103^+$), as well as two populations of recirculating and migratory populations defined by CCR7 and L-selectin[63]. Recently, the human T cell population definied by $CD4^+CLA^+CD103^+$ and cutaneous residence was also found in circulation, with shared clonotypes between the $CD69^{neg}$ counterpart in blood and the $CD69^+$ T_{RM} population skin[64], and a common functional profile consistent with that of Th22 cells: IL-22 and IL-13 production[65, 66]. Notably, this population was 250-fold more frequent among skin T cells than in blood[64]. The antigen specificities of these major T cell pools remain largely undiscovered, but overalp with functions ascribed to T cells recognizing CD1a from several studies published to date.

At the outset of my work and in several recent studies in skin[67], data was emerging from the Ogg laboratory and others showing that CD1a-autoreactive T cells are sources of type 2

cytokines in allergic dermatitis and psoriasiform inflammation[44, 68, 69, 70, 71, 72] (**Fig 1.3**). In such activation assays of T cells *ex vivo*, the identity of any lipid antigen presented by CD1a is not typically known, although some studies implicate phospholipases and lyso-phopholipids in triggering response [68, 69]. Human CD1a-transgenic mice, as compared non-transgenic mice, show increased skin inflammation from certain small molecules or lipids including skin sensitizer imiquimod, and the poison ivy antigen, urushiol [44]. Recently lipidic and non-lipidic small molecule allergens present in skin creams, such as balsam of peru and farnesol, have been shown to activate CD1a-dependent T cell response [43]. Therefore, understanding how CD1a could be recognized and the extent to which antigens are involved has emerged as a major biological question in skin immunology.

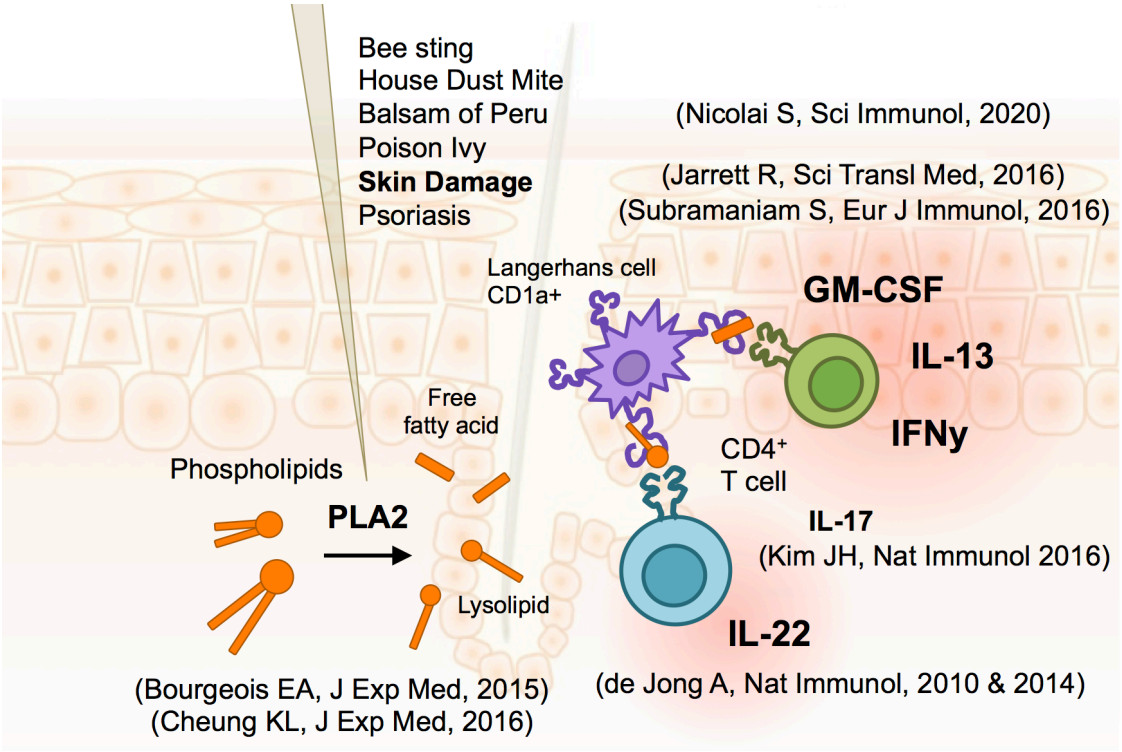

Figure 1.3. Summary of reported CD1a-autoreactive responses and fine-tuning by the endogenous or exogenous CD1a ligand milieu in skin.

III. Dissertation objectives

One could argue that finding definitive surface markers to track and subset immune cells of particular functionality is an immunologist's *raison d'être*, toward ultimately discovering causative drivers of response that might be targeted or manipulated therapeutically. To date, CD1a-autoreactive T cells in humans had been tracked and enumerated according to functional readouts like cytokine gene expression by RT-PCR or cytokine ELISPOT assays, and inhibition by the addition of CD1a-blocking antibodies. Reactivity of T cells to CD1a cannot currently be recognized by TCR sequence or any other cell surface marker on T cells. The paucity of validated and selective cell surface markers for CD1a-autoreactive T cells as a population limited study of CD1a autoreactive T cells to cumbersome functional methods that required *a priori* selection of activation read-outs. This approach is inherently prone to false negative detection, and cells are often destroyed in the process of measuring their responses. While CD1a-autoreactiivity is enriched among T cells expressing skin-homing receptors, like CCR4, CCR6, CCR10 and cutaneous lymphocyte antigen (CLA) [50], the antigen specificities of skin T cells remained largely unknown and are likely diverse. The fraction of T cells residing in or recirculating through skin that are CD1a-autoreactive was unknown. Recovering T cells from tissue also presented two key challenges. First, we needed to obtain sufficient numbers of pure T cells for study. Second, human skin T cells generally express intermediate levels of classical activation (CD69) and memory (CD45RO) markers at steady state, precluding the utility of certain activation-based read-outs for selection on T cell response.

Therefore in this thesis, we endeavored to capture CD1a-autoreactive cells from skin based on putative binding to T cell receptors themselves, using CD1a tetramer reagents. Human CD1a tetramers were validated in the Moody laboratory in collaboration with John Altman and first reported in 2013 in the setting of mycobacterial antigens and detecting T cells

in tuberculosis patients [73]. Based on incomplete information about the existence of any immunodominant self antigens that might bind to CD1a, initial questions focused on the choice of candidate autoantigens for loading onto CD1a. First, in **Chapter 2**, we optimize CD1a tetramer detection to enumerate and select skin T cells binding directly to the surface of CD1a, taking advantage of the theoretical direct CD1 detection models that predicted that the structure of CD1a itself might be more important than the loaded lipid antigen (**Fig. 1.2**). Accordingly, we attempted to stain T cells with CD1a tetramers without prior exchange of endogenous lipids for defined autoantigens. We then ask how individual lipid ligands in CD1a and the surface of CD1a itself influence TCR binding, again using tetramer stain intensity as readout. In **Chapter 3**, we ask if CD1a is selective in ligand binding from among the large cellular lipid pool to which it is exposed. We characterize the molecular basis for a highly enriched endogenous inhibitor of CD1a-autoreactive T cell binding, long chain sphingomyelins.

This page is intentionally left blank

CHAPTER 2: Detection and capture of CD1a-autoreactive T cells from skin

Parts of this chapter were submitted for publication as:

Rachel N. Cotton, Tan-Yun Cheng, Marcin Wegrecki, Jérôme Le Nours, Dennis P. Orgill Bohdan Pomahac, Simon G. Talbot, Richard A. Willis, John D. Altman, Annemieke de Jong, Jamie Rossjohn, Rachael A. Clark, D. Branch Moody. **Tetramers capture T cells from human skin that directly recognize CD1a with responses tuned by self lipids**

INTRODUCTION

CD1a is an MHC class I-like protein that binds lipids in its hydrophobic cleft and activates T cells via T cell receptors (TCR) [58, 74, 75, 76, 77]. CD1a-reactive T cell clones derived from blood [4] and evidence for CD1a expression in epidermal Langerhans cells [78] have been known for decades, providing indirect evidence for a role of CD1a in skin immune responses. More recently, CD1a-autoreactive T cells were detected at the polyclonal level among human blood T cells based on IL-22 transcript levels in response to CD1a expressing cell lines. Such responses were most frequent among T cells expressing skin homing receptors like cutaneous lymphocyte antigen (CLA), chemokine receptor 4 (CCR4), CCR6, and CCR10 [50, 51], suggesting that they preferentially home to skin. Further, functional studies of CD1a-dependent cytokine responses from skin-derived T cells suggest the presence of Th1, Th2 T cell types, with higher responses in patients with allergic dermatitis or psoriasis [44, 68, 69, 70, 71, 79]. Latest studies of skin T cells from human CD1a-transgenic mice show increased skin inflammation from contact sensitizers, including skin sensitizer imiquimod, and the poison ivy antigen, urushiol [44], as well as lipidic contact allergens present in skin creams, such as farnesol [43].

A fundamental difference between the MHC system for peptide presentation and the CD1 system for lipid presentation is that the former is the most polymorphic locus in the human

genome, whiile CD1 genes are nearly monomorphic [80]. The expression of nearly identical CD1 proteins by nearly all humans offers practical advantages for the basic study and therapeutic manipulation of such donor-unrestricted T cells. Experiments can be done with antigen presenting cells (APCs) that are not genetically matched, individuals are expected to bind and present the same kinds of antigens, and one kind of CD1 tetramer can be used for T cells from any donor [81]. Despite these advantages, current approaches to studying polyclonal skin T cells focus on cytokine-based assays and cannot broadly survey the CD1a reactive T cell repertoire. Activation based assays require *a priori* knowledge of both antigens and the cytokine to be measured. Although one study identified squalene and related skin oils as self antigens for several T cell clones [56], data do not currently address whether these or any other antigens are immunodominant such that responses to these molecules as tool compounds might broadly predict the human CD1a autoreactive repertoire. The cytokine profiles produced by CD1a autoreactive T cells are diverse [44, 68, 69, 70, 71, 79] and the optimal activation marker that could broadly detect CD1a autoreactive T cells is unknown. Further, the paucity of cell surface markers for CD1a-autoreactive T cells as a *bona fide* population is insufficient to enumerate and separate live CD1a-autoreactive T cells from the rest of the T cell panoply in skin. Last, activation assays used to detect CD1a-dependent responses necessarily destroy the cells under study, so that they are not available for selection and study at the individual cell level or for further *in vitro* based mechanistic studies.

To physically capture T cells expressing CD1a binding TCRs from their presumptive tissue resident site in human skin, we employed bivalent cell surface cytokine capture and CD1a tetramers. For MHC and CD1 [82], tetrameric antigen presenting molecules are typically treated with a single type of antigen, so that two or more arms of the tetramer carry the same ligand, allowing multimeric interactions with TCRs that increase the avidity of binding above the

threshold needed for detection in flow cytometry [83, 84, 85]. Whereas CD1a tetramers were recently validated using a foreign bacterial lipopeptide [73], they have not been used in studies of skin or other tissue-derived T cells. Further, the use of tetramers for detection of autoreactive T cells is hampered by the lack of any known immunodominant self antigen for CD1a. Therefore, testing a new molecular model that TCRs might recognize CD1a in preference to the lipid ligands carried by CD1a [46], we attempted to detect polyclonal skin T cells ex vivo with untreated CD1a tetramers carrying dozens of endogenous lipids (CD1a-endo). Surprisingly, this approach detected large numbers of T cells in every donor tested and the tetramer+ cells were found to have functional CD1a autoreactivity. Through lipidomic analysis of CD1a ligands and mutational analysis of the CD1a surface, our data indicate that the common mode of CD1a autoreactivity is strongly directed at CD1a itself with cross-reactivity for many types of self lipids. Most models of T cell response emphasize antigen recognition as a specific and essential on-switch for T cell activation. Here, the loaded lipids marginally affect CD1a-TCR binding, akin to fine-tuning rather than a binary switch. Further, direct capture of large numbers of CD1a autoreactive T cells from many donors ruled in IL-22 secretion, CD4 expression and other features of polyclonal T cells in the ex vivo state to support epidermal CD1a as a natural target of Th22 response in humans.

RESULTS

CD1a-autoreactive IL-22 responses in normal human skin

We sought to isolate polyclonal skin T cells in large numbers with high purity because our focus was detection of a putative CD1a autoreactive T cell population of unknown frequency. Conventional methods rely on tissue dissection, physical disaggregation, collagenases and DNAses, and chelating agents to release T cells from their natural tissue matrices, which necessarily damages cells and yields relatively low numbers, limiting downstream analyses [86, 87]. An alternative outgrowth method taking advantage of the migratory tendency of skin T cells enables the recovery of large numbers of pure skin T cells that retain surface marker expression and a diverse TCR repertoire [61, 62]. Skin explants were cultured on collagen-seeded three-dimensional growth matrices (3D method) for 21-28 days, allowing T cells to emigrate in response to chemoattractants from skin fibroblasts and supplemental IL-2 and IL-15. We directly compared T cell purity in a skin sample split between collagenase digestion and 3D culture, finding that the 3D method gave less subcellular debris and a 9-fold higher (8.6% versus 77% of total events) lymphocyte purity based on events entering live-dead stain exclusion, $CD3^+$, and lymphocyte size and granularity gates (**Fig. 2.1A**). Therefore, we employed the 3D method to surgical discarded skin from 25 unrelated donors for use in CD1a-dependent response assays, CD1 tetramer staining, or both (**Table 2.1**).

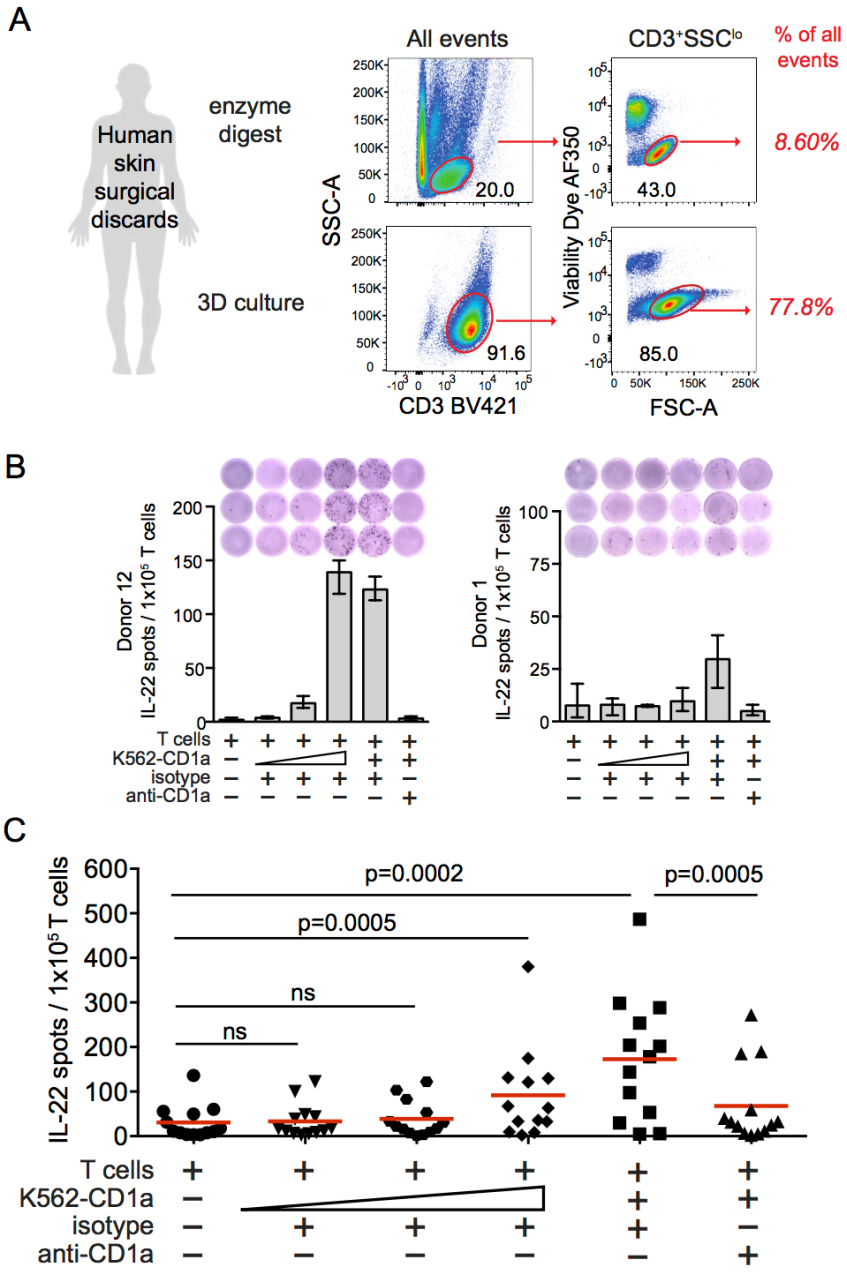

Figure 2.1. CD1a dependent IL-22 responses among human skin T cells. A) Human skin T cell collection recovered by a collagenase and DNAse digestion method at day 1 is shown in comparison to T cells collected at day 21 of 3D culture from the same donor. **B)** After co-culture with allogeneic *in vitro* derived LC-like cells + HLA-blocking antibodies (W632, L243), polyclonal skin T cells (n=13, 1x10^5/well) were tested in IL-22 ELISPOT assays for response to increasing numbers of CD1a-expressing K562 cells (400; 2,000; 10,000; or 50,000) and anti-CD1a (OKT6, 10ug/ml) or isotype control (P3, 10ug/ml) overnight. Two donors representing an intermediate and a low response pattern are shown. Error bars indicate mean with range of technical triplicates. **C**) Summary data from 13 donors. Each point represents the mean of technical triplicates from a single donor. P values were calculated using the Wilcoxon matched pairs signed rank test.

Table 2.1. Summary of human skin donor samples

Donor code	Skin site (if known)	Tested in:
A	unknown	IL-22 ELISPOT
B	unknown	IL-22 ELISPOT
C	unknown	IL-22 ELISPOT
1	face	IL-22 ELISPOT
2	breast	IL-22 ELISPOT
3	face	IL-22 ELISPOT
4	abdominal	IL-22 ELISPOT
7	face	IL-22 ELISPOT
9	abdominal	IL-22 ELISPOT
11	unknown	IL-22 ELISPOT
12	thigh	IL-22 ELISPOT
17	abdominal	IL-22 ELISPOT
18	abdominal	IL-22 ELISPOT
19	face	IL-22 capture
20	unknown	IL-22 capture
29	abdominal	IL-22 capture, CD1 tetramer
30	abdominal	IL-22 capture, CD1 tetramer
31	unknown	IL-22 capture, CD1 tetramer
32	unknown	IL-22 capture, CD1 tetramer
34	unknown	IL-22 capture
35	unknown	CD1 tetramer
36	abdominal	CD1 tetramer
37	abdominal	CD1 tetramer
38	abdominal	CD1 tetramer
39	abdominal	CD1 tetramer

In general, IL-22 producing T cells are enriched among circulating T cells that express skin homing receptors [65, 66], and Th22 cells are more frequent in human skin versus other tissue sites [88, 89]. While these $CD4^+$ Th22 cells might recognize MHC II, formally demonstrating MHC restriction at the polyclonal and interdonor level has been difficult, and one study implicated CD1a as a target, based on high rates of CD1a-dependent IL-22 mRNA expression by blood-derived T cells that expressed skin homing receptors [50]. However, the actual homing of CD1a-autoreactive IL-22-producing T cells from blood to skin and their contribution to the total skin T cell pool were unknown. Therefore, we tested polyclonal skin T cells from 13 healthy donors by IL-22 ELISPOT using CD1a-expressing K562 (K562-CD1a) antigen presenting cells (APCs) in the presence of CD1a-blocking antibody or an isotype-matched control. As shown in two

representative donors, IL-22 responses were higher in response to K562-CD1a cells as compared to baseline T cells alone, with stronger responses seen with increasing numbers of K562-CD1a cells. As a second criterion for restriction, IL-22 response was blocked by anti-CD1a (**Fig. 2.1B**).

The high absolute number of skin T cells obtained by 3D culture enabled triplicate measurements on a per donor basis (**Fig. 2.1C,** n=13), finding highly significant responses to K562-CD1a ($p=0.0002$) and blockade of responses with anti-CD1a ($p=0.0005$). Based on the two methods for measuring CD1a dependence, we calculated precursor frequency both as the difference in IL-22 spots appearing after addition of K562-CD1a cells versus T cells alone, and as the decrement in spots blocked by anti-CD1a versus isotype control. Both measurements gave similar estimates across the cohort of 13 patients, placing the frequency of IL-22 producing CD1a-autoreactive T cells at 0.10-0.14% of skin T cells. Thus, fulfilling a hypothesis based on study of blood T cells [56], IL-22 ELISPOT provides evidence that CD1a-autoreactive T cells enter skin and do so reproducibly when detected at the polyclonal level among unrelated donors. However, little is known about the cells themselves because they are destroyed during measurement. Therefore, we next employed a recently developed IL-22 cell surface capture reagent to detect and sort live CD1a-autoreactive cells.

Capture of polyclonal IL-22$^+$ CD1a-autoreactive T cells from skin

Using a bifunctional antibody that binds CD45 and IL-22 to preferentially coat cells that have secreted IL-22, we measured CD1a-dependent skin T cell responses from another 7 individuals. In all donors tested, we detected elevated frequencies of IL-22-secreting T cells in response to K562-CD1a cells versus T cells alone, or versus K562-CD1a cells pretreated with anti-CD1a (**Fig. 2.2A-B**). When calculated using the T cell only condition as background, the

median donor frequency was 0.35% of CD3 T cells. When calculated using the CD1a-blocked condition as background, the median frequency was 0.46% percent (**Fig. 2.2B**). Thus, the absolute frequencies were similar to those detected by IL-22 ELISPOT (**Fig. 2.1**). Further, while we did observe IL-22 producing cells that were CD4 negative, the majority of IL-22 producing cells detected in all donors were $CD4^+$ (**Fig. 2.2A-B**). Taken together with the 13 individuals tested by IL-22 ELISPOT, results from 20 total individuals rule in the presence of IL-22 producing CD1a-autreactive T cells in human skin.

Taking advantage of cytokine capture assays to recover cells of interest, we derived T cell lines as durable reagents to validate CD1a specificity and for molecular analysis of antigen response and TCR expression. From donor 34, we set a broad sort gate to include all IL-22-producing cells (**Fig. 2.2A, red**). We plated 60 sorted cells per well for expansion to oligoclonal lines. We considered the possibility that multiple rounds of activation on the K562 cell line, which expresses low but not necessarily absent levels of MHC class I [90], can lead to expansion and activation of alloreactive CD1a-nonspecific T cells. Since cells had already been selected based on response to CD1a expressed on K562, we used a second CD1 cellular expression system on C1R cells for secondary testing. Of 64 lines tested for response to C1R-CD1a and C1R-CD1b cells, we selected 10 lines that had detectable IL-22 responses to C1R-CD1a greater than those seen with C1R-CD1b (**Fig 2.2C**).

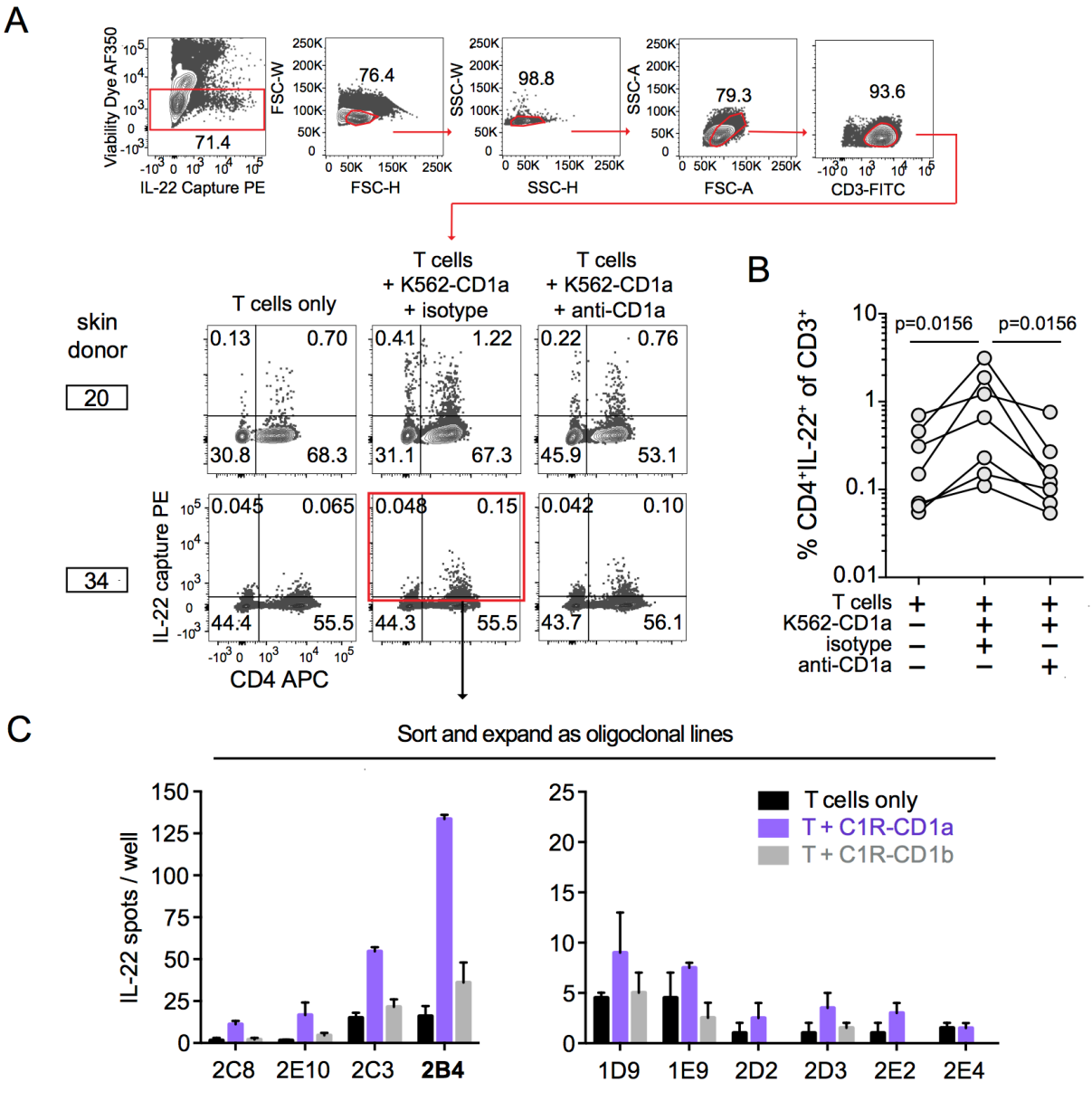

Figure 2.2. IL-22 surface capture selects CD1a autoreactive skin T cells. **A**) Polyclonal skin T cells cultured with K562 cells overnight were labeled with the IL-22 capture reagent, and pre-gated as shown for Donor 20. **B**) Frequencies of CD4+IL-22+ cells among CD3+ cells (top right quadrant gate as in A; n=7 donors). **C**) After oligoclonal expansion of Donor 34 sorted cells (red box indicates sort gate including CD4neg IL-22low cells), lines were screened by IL-22 ELISPOT on C1R cells. Data from 10 lines shown as mean with range of technical duplicates.

CD1 tetramer assays for autoreactive T cells from tissue

Assays based on T cell activation require *a priori* selection of cytokines that may not reflect the only or even the dominant effector function of responding T cells, so we next sought to optimize CD1a tetramers as a capture reagent that only requires CD1a-TCR binding. Human CD1a tetramers were recently developed, loaded with mycobacterial lipid to detect foreign antigen-reactive T cells [73]. However, human CD1 tetramers carrying endogenous lipids from the CD1 expression system (CD1a-endo) have not been applied to detection of T cells from tissue, and to do so required that we address several key issues relating to possible false positive or false negative detection. First, CD1b-autoreactive cytokine responses in blood are rare [50], and our pilot studies found infrequent staining of polyclonal skin T cells with CD1b-endo tetramers, allowing its use as a negative control. Second, tissue-derived T cells are often contaminated with tissue debris and dead cells, which are highly autofluorescent and non-specifically adhere to tetramer reagents, causing false positive detection [91]. Indeed, for T cells obtained by physical digestion of skin, we observed many suspected false positive events located outside of size parameters for lymphocytes (**Fig. 2.3**), which were less frequent in 3D cultures, which largely lacks subcellular debris. Further non-specific staining patterns could be reduced by removing autofluorescent events in the FITC channel and by dead cell exclusion reagents. Combined, both the 3D method and the gating strategy greatly reduce noise and false positive staining on skin T cells.

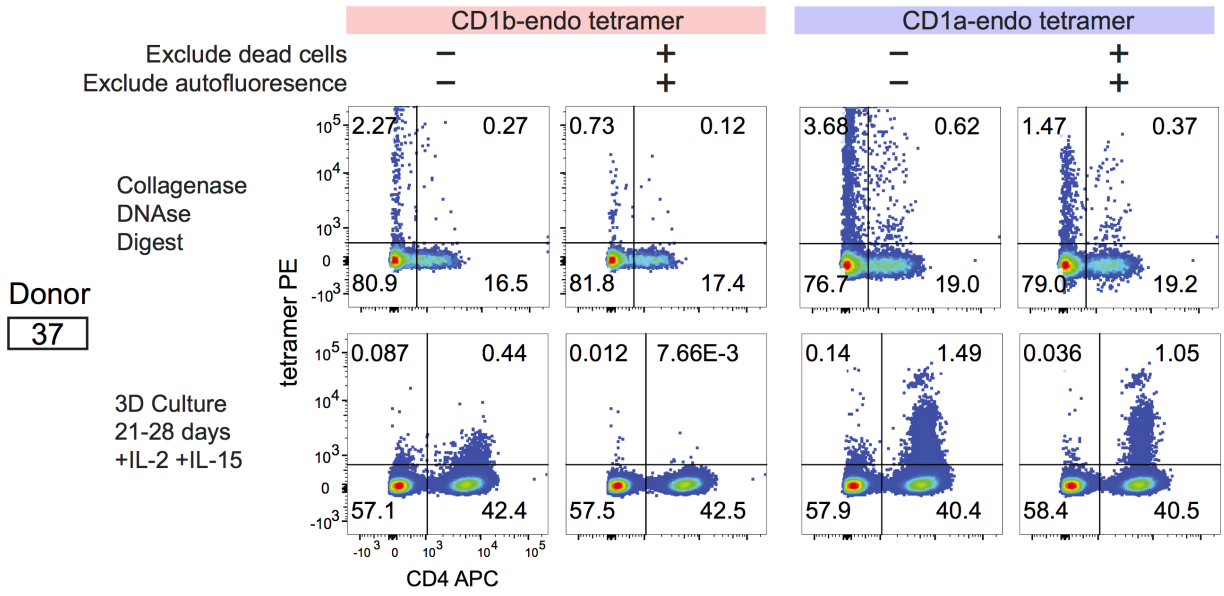

Figure 2.3. Exclusion of dead cells and autofluorescence improves signal to noise ratios of tetramer detection. Skin from donor 37 was allocated for enzymatic digestion or for 3D culture as in main figure 1. Cells were stained for flow cytometry with CD1b-endo or CD1a-endo tetramers, LIVE/DEAD amide-reactive dye, anti-CD3, and anti-CD4. Shown pre-gated on CD3+, lymphocyte size and granularity and singlets by FSC and SSC. (+) indicates additional gating to exclude cells based on autofluorescence in the FITC channel and uptake of amide-reactive dye.

A known cause of false negative detection by tetramers is the gap between the relatively high affinities typically needed for detectable TCR binding to antigen complex tetramers, versus the lower affinity threshold needed for T cell activation by professional APCs. This affinity gap is broadly documented for MHC [92, 93] and CD1 tetramers [94, 95]. To mitigate these theoretical false negative results, unlabeled anti-CD3 monoclonal antibody (OKT3) was added after tetramer staining to cluster and retain surface TCR complexes, thereby increasing the stoichiometry of TCR interaction with tetramers. This approach produced striking improvements in sensitivity, revealing a previously obscured tetramer$^+$ population in the CD1a-autoreactive skin T cell line DermT [50] (**Fig. 2.4A**). Using a sorted CD1a-tetramer$^+$ population, we observed a steep dose response and a 10-fold higher CD1a tetramer staining intensity by this method with no substantial effect on background CD1b tetramer staining (**Fig. 2.4B**).

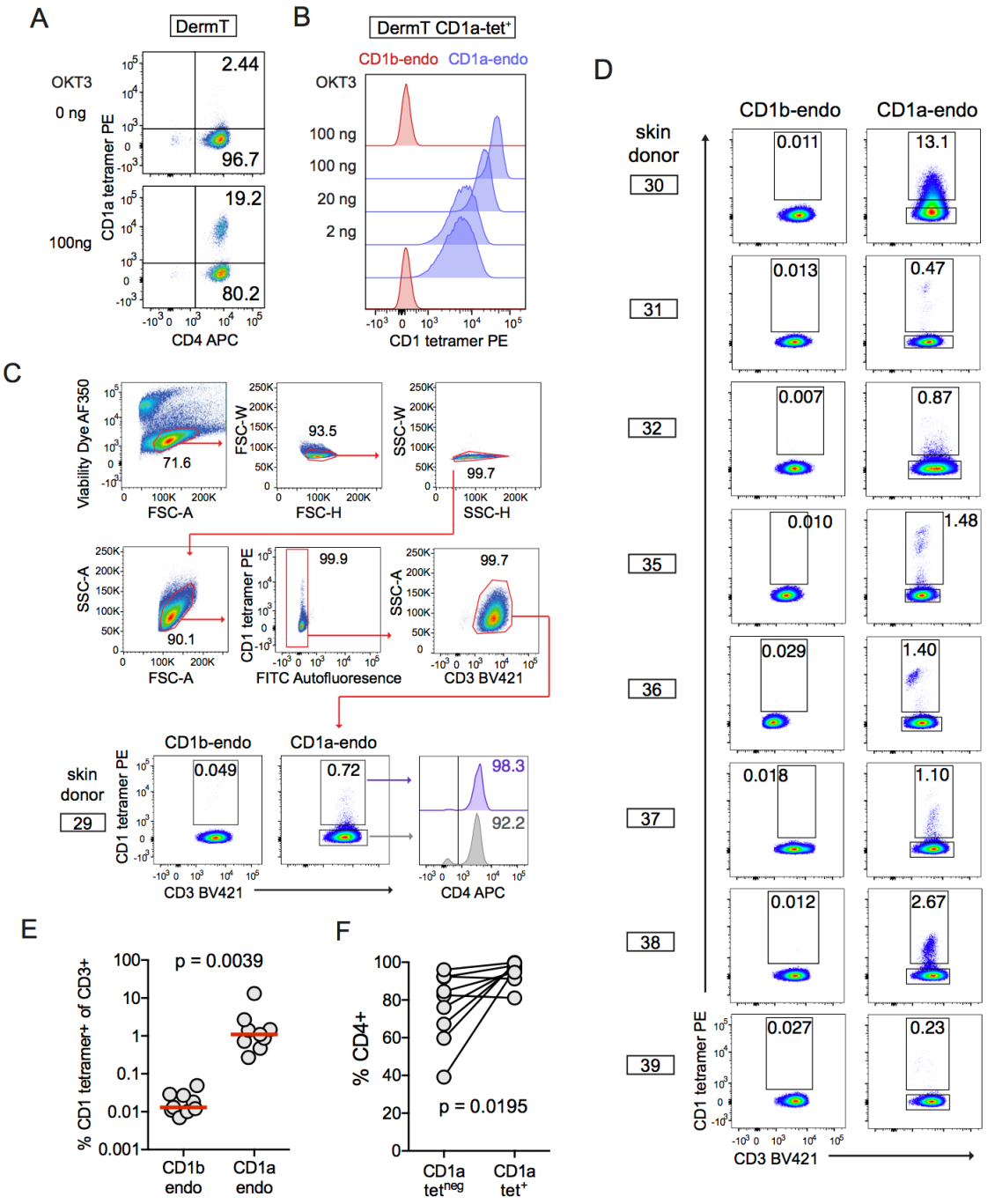

Figure 2.4. . CD1a-endo tetramer staining of polyclonal skin T cells. A) Unlabeled OKT3 was added during CD1a-endo tetramer staining of the oligoclonal CD1a-autoreactive T cell line DermT. Pregated: Live, Lymphocytes, Singlets, $CD3^{+}$. **B**) Dose response of CD1 tetramer staining intensity with added OKT3 on a sorted DermT CD1a tetramer+ subpopulation. **C-D**) Polyclonal skin T cells obtained from 3D culture were stained with CD1b-endo or CD1a-endo tetramers (n=9) and representative pregating is shown for donor 29. **E**) Summary data show the absolute tetramer detection rates, and **F**) rates of CD4 positivity among CD1a-endo $tetramer^{+}$ and tetramer negative gates.

CD1a-endo tetramers detect skin T cells

Both MHC [96] and CD1 [73, 83, 84, 85, 97] tetramers use proteins treated with one antigen to promote multivalent interactions with TCRs, so a basic feasibility question relates to the choice of self antigen to load. Although skin and plant oils activate T cell clones [43, 44, 56], no immunodominant lipid that broadly activates CD1-dependent polyclonal T cells is known. Individual T cell lines have been reported that can, however, recognize CD1a or CD1c without the addition of exogenous lipid [56, 98]. The only available ternary CD1a-lipid-TCR structure shows that the BK6 TCR contacts CD1a itself, rather than its carried lyso-phosphatidylcholine (LPC) ligand [46]. Although most TCRs require addition of a specific antigen to bind with high affinity to antigen presenting molecules, we reasoned that if CD1a-directed TCRs like BK6 are common *in vivo*, randomly loaded CD1a tetramers carrying endogenous lipids from the expression system (CD1a-endo) might detect such T cells at the polyclonal level (**Fig. 2.4C**).

Among polyclonal skin T cells from nine donors tested (**Table 2.1**), cells binding CD1a-endo tetramers were detected at rates above control CD1b-endo tetramer staining in all cases. The median rate of CD1a-endo tetramer staining as a percentage of T cells was high in absolute terms (1.1%), which was ~100-fold higher than rates of CD1b-endo tetramer staining (median = 0.013%, p=0.0039, Wilcoxon test) (**Fig. 2.4E**). Donor 30 was a notable outlier on the high end, where CD1a-endo tetramer$^+$ cells represented 13% of CD3$^+$ cells. The low rate of CD1b-endo tetramer staining in donor 30 and the low rate of CD1a-endo tetramer staining T cells from the individual (donor 29) that was tested in parallel in the same experiment both argue against non-specific binding or autofluorescence causing this high measurement (**Fig. 2.4E**). We examined CD4 expression because prior reports of functional analysis of CD1a-autoreactive T cells suggested a CD4 predominance [50, 51]. However, there is no mechanistic rationale to predict if or how CD1a-autoreactive T cells segregate based on CD4 or CD8 co-receptor expression. Indeed

the initial discovery of the BK6 CD1a-autoreactive T cell clone in 1989 was from among CD4 CD8 double negative cells [4]. Here, CD1a-endo tetramer staining of *ex vivo* skin T cells indicated that $CD3^+$ CD1a-endo $tetramer^+$ cells were enriched for CD4 positivity versus the $CD3^+$ CD1a-endo tetramer negative fraction (**Fig. 2.4E,** p=0.0195). In all donors tested, the CD1a-endo tetramer positive fraction was greater than 80% $CD4^+$ (**Fig. 2.4F**).

Tetramer$^+$ cells functionally recognize CD1a without added ligand

Overall, the pattern in which ~1% of total skin T cells stained with CD1a-endo tetramers represents a high rate for staining with any tetramer in polyclonal human T cells. CD1b tetramers ruled out many forms of non-specific staining of T cells by CD1a tetramers, but CD1c and CD1d can show specific binding to non-TCR surface ligands on T cells [99]. Therefore, to study CD1a-TCR binding, we first sorted the oligoclonal CD1a-autoreactive line DermT, which was 20.5 percent CD1a $tetramer^+$, to yield a clonal population (DermT2, 99.8 % $tetramer^+$). TCR sequencing yielded single productive TCRα and TCRβ chains, where TRBV2*01 in the ImMunoGeneTics (IMGT) international database nomenclature encodes a variable region recognized by a monoclonal antibody specific Vβ22 (**Fig. 2.5A**). Dual staining with tetramer and TCR Vβ22 identified only 1 diagonally oriented population in which the intensity of anti-TCR correlated with the intensity of tetramer staining, and the resulting clone was CD1a autoreactive by two different measures (**Fig. 2.5B**). Thus, one clonotypic TCR was responsible for a CD1a-mediated functional response and binding to CD1a tetramers.

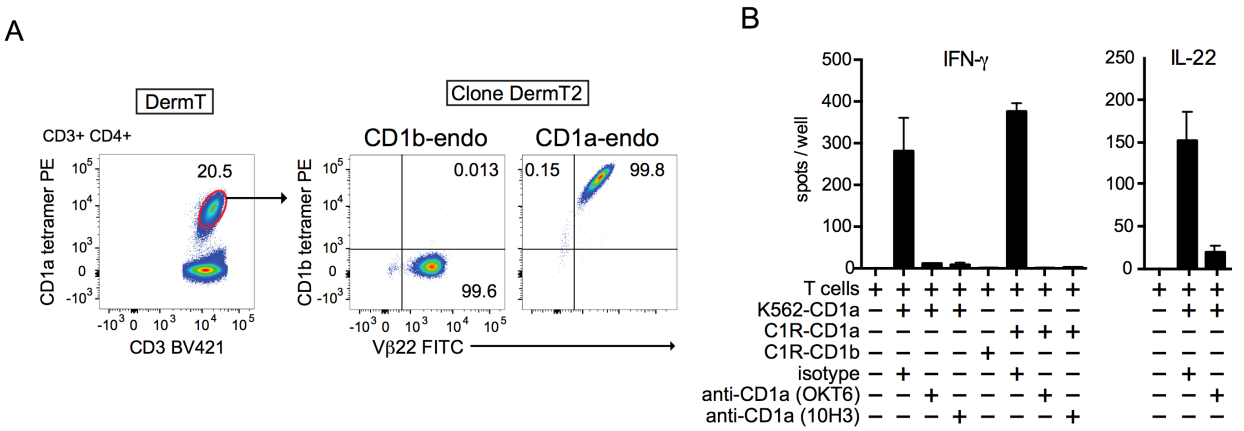

Figure 2.5. CD1a tetramers select for CD1a autoreactive T cell cytokine response. **A)** A CD1a tetramer$^+$ subpopulation from the CD1a-autoreactive DermT line was sorted (pre-gated live, lymphocyte and singlet FSC and SSC gates, CD3$^+$, Autofluoresence(FITC)^{neg}, CD4$^+$) and bulk expanded. The sorted line, DermT2, was stained with the TCR Vβ22 antibody and tetramers. (pregated: Live, lymphocytes, singlets, CD3$^+$, CD4$^+$). **B**) IL-22 and IFN-γ ELISPOTs of DermT2 cells.

To test more broadly whether CD1a tetramer$^+$ T cells expressed TCRs with functional recognition of CD1a, we sorted CD1a tetramer$^+$ cells from polyclonal skin T cells from 4 donors (**Fig. 2.6A**). Tested after expansion of sorted T cells, resulting T cell lines did not bind CD1b-endo tetramers, but were strongly enriched for CD1a-endo tetramer binding in all cases (**Fig. 2.6**). Next we plotted CD1a-endo tetramer staining versus anti-CD3 for patients in which TCRs were unknown (Donors 31, 32) or versus anti-TCR antibodies, where TCR sequencing (**Table 2.2**) or antibody screening identified the variable TCR regions, which were anti-Vβ22 TCR in Donor 30 and anti-Vβ23 TCR in Donor 36. This approach revealed distinct populations with patterns that suggested TCR binding of CD1a-endo tetramers (**Fig. 2.6**). First, populations separated by CD1a tetramer intensity are consistent with tetramer binding to clonally distributed ligands. Second, the diagonal staining pattern within each dually staining population indicates that tetramer staining correlated with anti-CD3 or particular anti-TCR Vβ staining intensity. Third, the limited, nearly clonal narrowing of TCR repertoire after sorting with CD1a-endo tetramers in

donor 30 and line 36 is suggests selection based on TCR binding. Last, TCRs typically determine the antigen specificity of T cells, and the sorted cells all showed IL-22 or interferon-γ release, which was completely or nearly completely blocked with anti-CD1a or use of cells not expressing CD1a in all cases (**Fig. 2.6B**). Therefore, measured on a multi-donor basis on polyclonal, oligoclonal, and clonal T cell levels, CD1a tetramers selected skin T cells with anti-CD3 and anti-TCR binding properties consistent with CD1a tetramer binding to the TCR complex, as well as functional reactivity towards CD1a. Thus, CD1a-autoreactive T cells are normally abundant in human skin and can be tracked and enumerated using tetramers without any knowledge of named lipid antigens. Further, on a molecular basis, these unexpected findings suggest that CD1a carrying highly diverse self ligands likely bind TCRs and that this is a common mechanism of CD1a reactivity, an hypothesis we subsequently investigated in detail.

Table 2.2. TCR sequences from DermT2 and line 36

Line	TRAV	CDR3α	TRAJ
DermT2	8-6	CAVSYGDYKLSFGAG	20
line 36	1-2	CAVRGPGGSYIPTFGRG	6
line 36	8-6	CAVSLSGNTPLVFGKG	29

Line	TRBV	CDR3β	TRBJ
DermT2	2	CASSEENQPQHFGDG	1-5
line 36	13	CASSLGWVAGVVETQYFGPG	2-5
line 36	27	CASPTAARWGNILYGYTFGSG	1-2

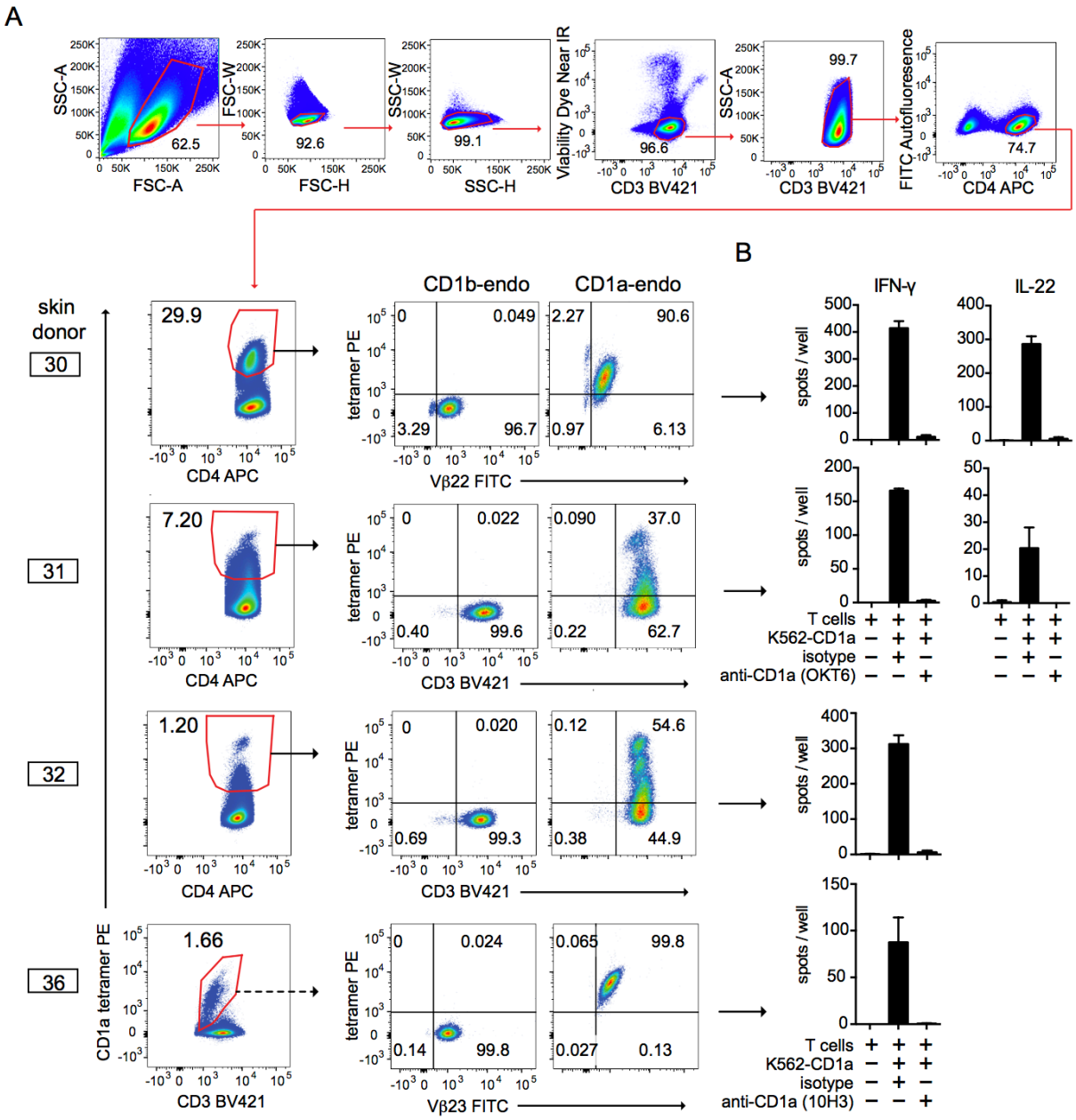

Figure 2.6. CD1a tetramers select for CD1a autoreactive T cell cytokine response. **A)** Polyclonal skin T cells were sorted based on binding to CD1a tetramers with representative pre-gating shown for Donor 30. Sorted cells were bulk expanded and subjected to tetramer and Vβ antibody staining by flow cytometry and **B**) CD1-dependent activation in IL-22 and IFN-γ ELISPOT assays. Dashed line arrow from Donor 36 indicates that a second round of sorting and expansion was done prior to shown tetramer staining versus Vβ23. Solid line arrows for all other lines reflect a single sort and expansion. Error bars indicate standard deviation from the mean of technical triplicates.

Molecular analysis of CD1a-endo tetramers

Next we sought to determine whether the spectrum of lipids bound to CD1a monomers during their production in human HEK293 cells were relatively homogenous, such that individual arms of a tetramer might carry the same or different lipids (**Fig. 2.7A**). Working with a biochemist in the laborotory, Tan-Yun Cheng, we measured the of lipids bound in CD1a-endo monomers from the NIH Tetramer Facility using a normal phase lipidomics method based on high performance liquid chromatography (HPLC)– time of flight (TOF) mass spectrometry (MS). We detected ~600 ion chromatograms in eluents of CD1a treated with chloroform and methanol, but this was possibly an overestimate of ligand heterogeneity. We censored ions corresponding to redundant detection of isotopes, recognizable lipid dimers with high mass and alternate adducts of any molecule M, such as ($[M+H]^+$, $[M+Na]^+$, $[M+NH_4]^+$). We further censored ions present in solvents, ions also eluting from an MHC II control protein or *m/z* values matching patterns for polyethylene glycol-based detergents. This process identified 98 high quality mass chromatograms with unique *m/z* and retention time values, as a conservative estimate the distinct lipids bound to CD1a (**Fig. 2.7B**). Whereas compound identification initially focused on the most abundant lipids, we specifically scanned for and identified lysophosphatidylcholine bound to CD1a. This targeted lipid search was conducted because the structure of CD1a-LPC complexes is known[46]. Normal phase chromatography results in separation of lipid classes with general structure, but not chain length and unsaturation of the alkyl chains. The range of separation in this system covers most known classes of cellular lipids, and retention time correlates directly with polarity [100]. Thus, considering the 98 unnamed compounds of known mass and retention time (**Fig. 2.6B, lower**) this result indicated that CD1a ligands broadly spanned the natural range of common lipids classes present in HEK293 cells and these singly charged ions differed in mass from ~200 to 1000 amu.

Assuming random tetramerization, these data predict on a statistical basis that differentially loaded monomers exist among the four arms of nearly all tetramers. Monomers of MHC and CD1 proteins rarely stain T cells [83], and TCRs only rarely cross-react with many highly diverse antigens. Therefore, CD1 or MHC tetramers are usually treated with one type of antigen prior to staining. In contrast, diversely liganded CD1a-endo tetramers used here (**Fig. 2.7B**) brightly stained DermT (**Fig. 2.4A-B**) and identified subpopulations among polyclonal skin T cells from all 9 donors tested (**Fig. 2.4C-D**). As hinted by the binding of one TCR to CD1a and the recent isolation of CD1c-restricted T cell lines [46], these results could be explained if CD1a autoreactive TCRs showed CD1a-centric binding and largely ignore the carried lipid.

To more directly investigate the molecular mechanism, we first identified a subset of the total CD1a-lipid interactome by matching the detected *m/z* values with masses of known lipids in databases (**Table 2.3**). This process, which was carried out by Tan-Yun Cheng and analyzed by me and other members of the group, generated tentative compound identifications, which were confirmed by measuring retention time in comparison with lipid standards (**Fig. 2.7B**). From 98 unknown lipids, we identified and annotated 30 molecular species across seven families: four diacylglycerols (DAG), one ceramide, one hexosylceramide, one sulfatide, sixteen phosphatidylcholines (PC), six sphingomyelins (SM), and one phosphatidylinositol (PI). Naming this subset of lipids confirmed that individual lipids do have differing polar head groups that span across known families of neutral lipids, glycolipids, phospholipids, sulfolipids, and sphingolipids. Also naming lipids in CD1a-complexes supported design of experiments to directly measure lipid cross-reactivity among CD1a-TCR binding events involving skin T cells.

Table 2.3. HPLC-MS identifications of detected ions and retention times from CD1a eluent

	detected *m/z*	RT (min)	Formula	predicted *m/z*	lipid chain
diacylglycerol (DAG)	638.5721	2.524	C39H76NO5 $[M+NH_4]+$	638.5718	C36:2
ceramide	648.6223	2.546	C42H82NO3 [M+H]+	648.6290	C42:2
hexosyl-ceramide	810.6756	3.964	C48H92NO8 [M+H]+	810.6818	C42:2
sulfatide	890.6376	19.955	C48H92NO11S [M+H]+	890.6386	C42:2
phosphatidyl-choline (PC)	760.5846	21.268	C42H83NO8P [M+H]+	760.5851	C34:1
sphingomyelin (SM)	813.6836	21.734	C47H94N2O6P [M+H]+	813.6844	C42:2
phosphatidyl-inositol (PI)	906.6064	23.611	C47H89NO13P $[M+NH_4]+$	906.6067	C38:3

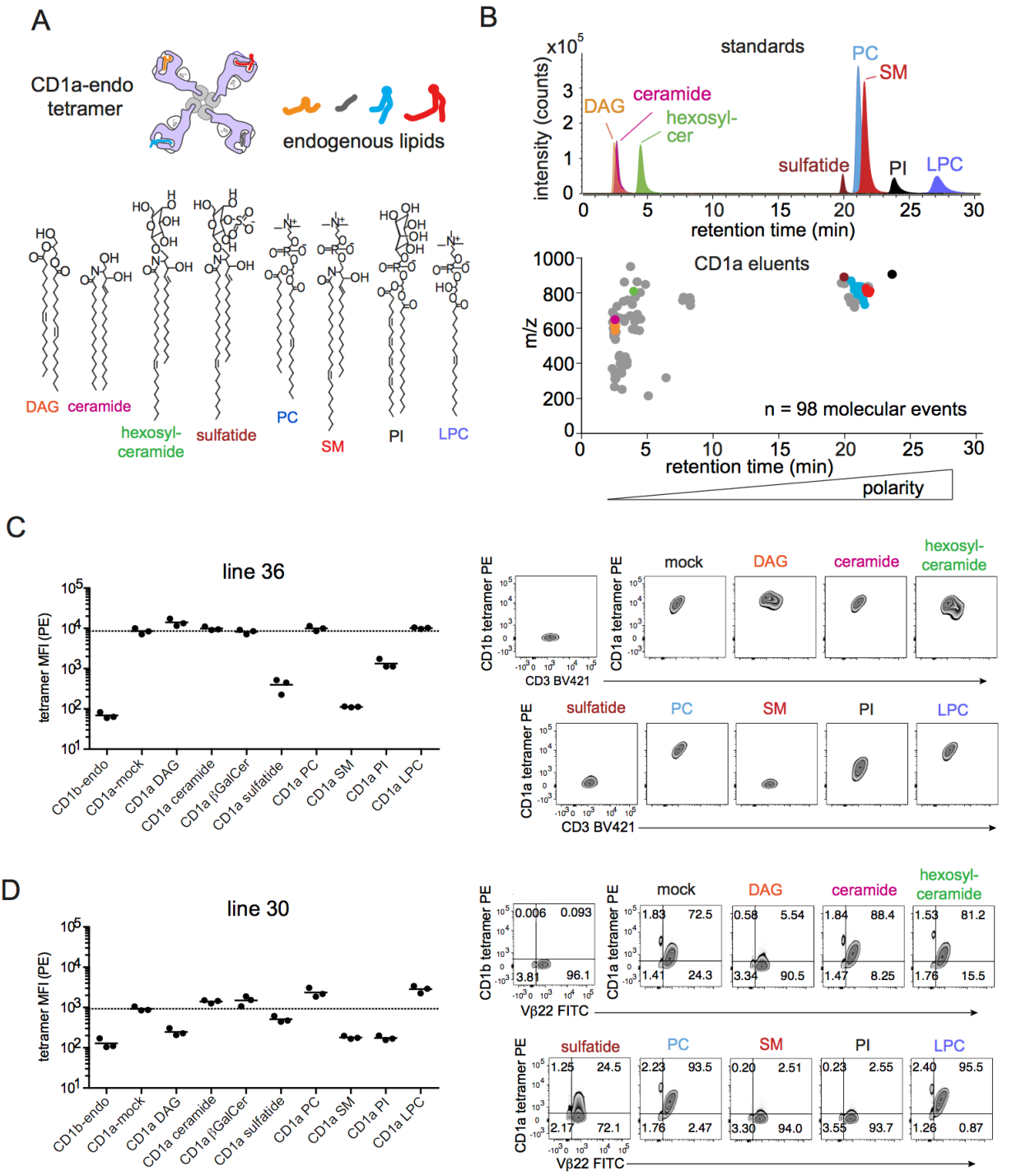

Figure 2.7. CD1a endogenous lipid tetramers bind autoreactive T cells despite lipid heterogeneity. **A**) Each arm of a CD1a endogenous lipid tetramer (CD1a-endo) may harbor a different lipid, representative lipids for 7 lipid families shown. **B**) Positive mode, normal phase HPLC-QToF-MS lipidomics of lipids eluted from CD1a monomer (bottom), and standards for each lipid class (top). Gray points indicate unique ion and retention time pairs of undetermined lipid identity. **C-D**) Characterization of CD1a endogenous lipid families (as shown in A) by tetramer staining of line 36 and line 30.

Broad cross-reactivity and three defined blocking lipids

Taking advantage of newly generated primary skin T cell lines (line 30, line 36), I treated CD1a tetramers with a representative neutral lipids (diacylglycerol, DAG; ceramide), a glycolipid (β-galactosyl ceramide, βGalcer), a sulfolipid (sulfatide), a sphingolipid (sphingomyelin, SM), phospholipids (phosphatidylcholine, PC; phosphatidylinositol, PI) and a lysophospholipid (LPC) (**Fig. 2.7C-D**). Relative changes in tetramer mean fluorescence intensity (MFI) or frequency of detection are interpreted as a correlate of ligand compatibility with CD1a-TCR binding. Staining of line 36 revealed a homogenous and high intensity staining population with equivalent staining by CD1a-endo versus tetramers treated with DAG, ceramide, βGalCer, PC and LPC. Partial reduction of staining was seen with PI and sulfatide, with nearly complete blockade of staining with SM. For line 30 sorted from donor 30, baseline staining was lower in intensity and reflects at least two distinct tetramer populations within this line. The pattern of lipid specificity was similar to that of line 36 with CD1a-endo showing unchanged or slightly augmented staining after treatment with ceramide, βGalCer, PC and LPC, and with partial blockade after treatment with DAG, sulfatide, SM, and PI. Overall, these patterns were similar except for DAG. Thus, these patterns of staining with tetramers treated with purified lipids corresponding to natural ligands suggests broad trends in ligand compatibility with CD1a-autoreactive T cell binding to CD1a. Not only is diversely loaded CD1a is sufficient to bind T cells, but also treatment with any one of several lipids with differing head groups do not alter the interaction. Finally, certain lipids, especially sphingomyelin, can block CD1a tetramer binding.

Role of the roof of CD1a

One testable hypothesis for how TCR binding could absolutely require CD1a, yet show cross-reactivity for lipids, was suggested by the crystal structure of the BK6 TCR bound to a CD1a-LPC complex. Here the LPC head group minimally protrudes to the CD1a surface through a portal on the right side of platform, whereas the TCR is shifted to the left. This TCR footprint missed direct contact with the LPC lipid, but extensively contacts the outer surface of CD1a, where tethering residues that link the α1 and α2 helices of the CD1a heavy chain create a flat, unliganded surface known as the A' roof (**Fig. 2.8A**). This left-right mismatch leading to CD1-centric recognition is known for only for one CD1a reactive TCR [46], and is clearly distinct from the many examples of dual TCR contact with CD1-lipid complexes involving CD1b, CD1c and CD1d [52, 53, 101], or MHC and peptide [48, 49].

To determine the role of the A' roof in CD1a-endo tetramer staining, the Rossjohn laboratory generated alanine mutations in CD1a, including triple mutants that cluster on the α1 (E62A, E65A, I72A) and α2 (R168A, T165A, I157A) helices, as well as a tetra-mutant at the extreme left margin of both helices (E62A, E65A, R168A, T165A) (**Fig. 2.8A**). I carried out staining analysis and worked with Tan-Yun Cheng to interpret lipid profiles released from CD1a proteins. For all three triple mutants, these mutations abrogated CD1a-endo tetramer staining of line DermT2 and Line 36. (**Fig. 2.8B-C**). This result is consistent with the TCR binding the A' roof. The CD1a pattern seen here is different from right-sided binding over the antigen portal as seen for the NKT TCR [52], or a newly identified mode of TCR binding on lateral face of MR1 proteins[102]. However, complete loss of binding in all three cases did not localize the interaction within differing areas of the roof. Also, simultaneous changes in 3 or 4 residues might have altered domain or tertiary structures somewhat distant from the sites of mutation.

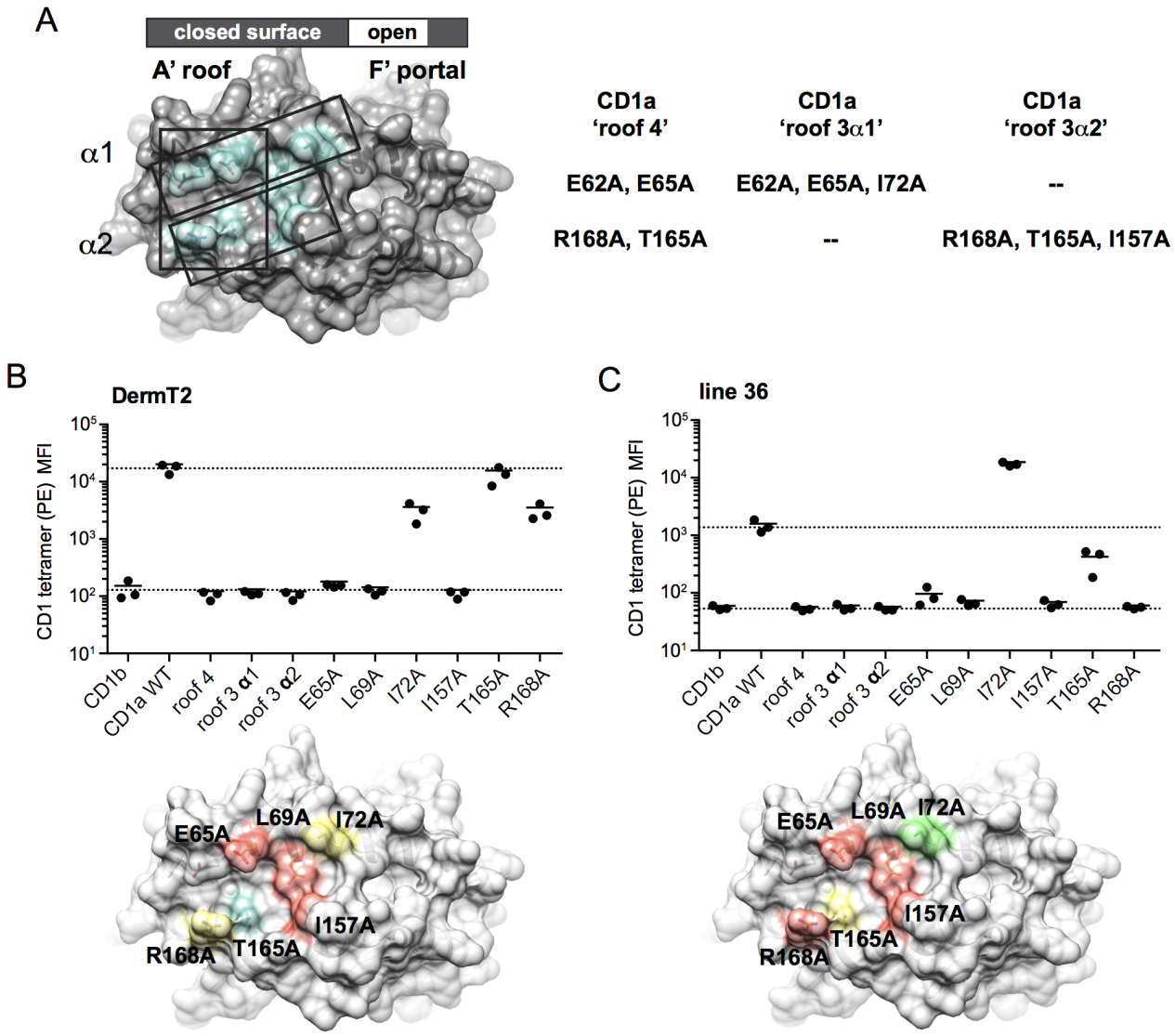

Figure 2.8. CD1a mutant tetramers point to direct CD1a-TCR contacts for recognition. **A**) Summary of CD1a A' roof mutants: roof 4, roof 3α1, roof 3α2 (aqua, overlaid on PDB:4X6E). **B-C**) CD1b, CD1a-WT and CD1a mutant tetramer staining on lines DermT2 (B) and line 36 (C), shown as tetramer mean fluorescence intensity (PE, MFI). Each point represents an individual experiment and solid lines indicate the mean of replicates. Dashed line is drawn from the mean of CD1a-WT staining. Relative contributions of each mutant to tetramer staining intensity are summarized as overlaid on PDB 4X6E (green = increased staining, red = low or absent staining, yellow = intermediate/reduced staining, aqua = similar to WT).

Therefore, the Rossjohn laboratory designed a second panel single site mutants dispersed throughout the CD1a roof (E65A, L69A, I72A, R168A, T165A, I157A), which I tested for tetramer binding. Because the main goal was to distinguish direct TCR contact with the CD1a roof versus the roles of bound lipids located inside CD1a or on the right side of the platform, Tan-Yun Cheng screened these single mutants for in HPLC-TOF-MS to determine the spectrum of ligands bound. All of the 6 single mutants bound to each the same spectrum of named lipids seen in CD1a-endo complexes, albeit with some differences in the relative signals for individual lipids (**Fig. 2.9**). For example, mutant I72A, which is located at the edge of the roof nearest the ligand portal, showed somewhat higher binding to DAG and TAG lipids (**Fig. 2.9**). Because line 36 had marginally preferential binding to DAG-treated tetramers (**Fig. 2.7C**) the small increase in DAG capture might have accounted for the slight increase in binding to line 36 by CD1a-I72A tetramers (**Fig. 2.8**). Otherwise, the profiles of named lipids were similar among the point mutants (**Fig. 2.9**), yet many single site mutants located near the center of the roof (E65A, L69A, I157A) showed a complete loss of tetramer binding with all lines tested. Two other mutants (T165A, R168A) located on the α2 helix and at the left-central portion of the A' roof, showed partial or complete binding loss, which varied according to the line tested (DermT2, line 36), suggesting that TCRs have somewhat different footprints.

Overall these data provided clear evidence for the central portion of the A' roof in mediating binding to T cells, most likely via effects of TCR binding to the A' roof. More generally, the patterns of staining with diversely loaded tetramers, patterns of staining with tetramers treated with alternate ligands and the pattern of protein mutation all suggest that CD1a autoreactive TCRs bind to the A' roof in a way that is tolerated by many types of lipids bound. Because these patterns of recognition were seen in all subjects tested, and in clones,

lines and polyclonal T cells, we conclude that CD1a-centric recognition by TCRs is a common mechanism of autoreactivity.

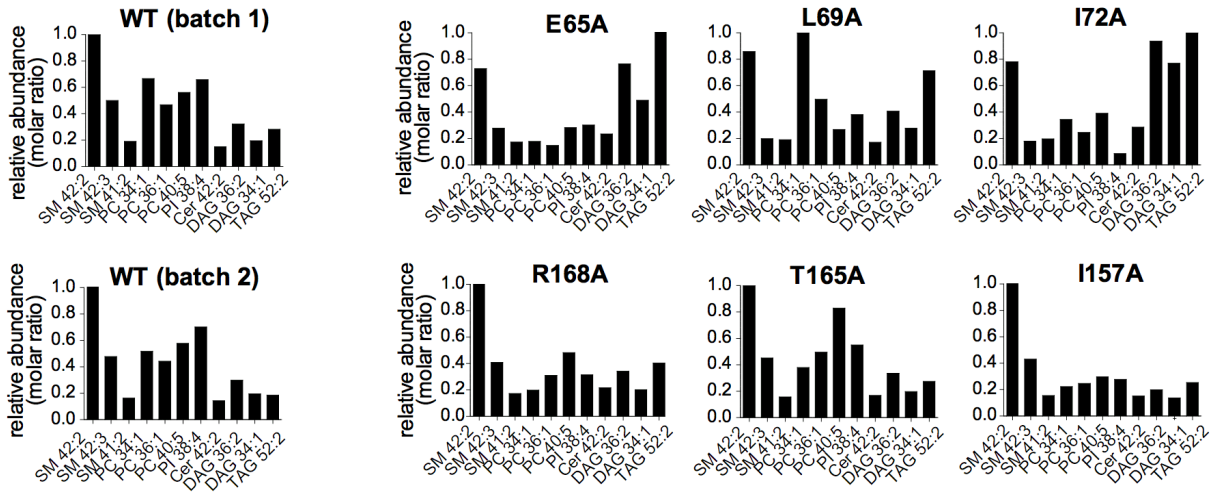

Figure 2.9. **Lipid elution from CD1a mutant tetramers.** Positive mode, reversed phase HPLC-QToF-MS analysis of lipids eluted from CD1a monomers. Data shown as relative abundance (molar ratio) relative to the most abundant of an 11-molecular species group comprised of 6 lipid families.

DISCUSSION

Human skin contains an estimated 20 billion resident T cells per person [61, 62], and while many recognize MHC-peptide, the full spectrum of antigenic targets of the large skin T cell repertoire remains to be elucidated experimentally. Prior studies suggested that skin T cells might recognize CD1a based on its high expression in the epidermis [74], existence of CD1a autoreactive cells in blood that express skin homing receptors [56] and blocking of CD1a to infer CD1a response [44, 68, 69, 70, 71]. In contrast, this study used CD1a tetramers to directly analyze tissue-based skin T cells, taking advantage of the interaction of CD1a and TCR to physically capture autoreactive cells for study over time and determine the molecular basis for recognition.

Surprisingly, CD1a-endo tetramers that were not treated with defined lipids brightly stained a considerable percentage, ~1% on average, of skin resident T cells. Further, CD1a-autoreactive or CD1a-endo tetramer positive T cell populations were present in all donors tested.

CD1a-lipid complexes used to stain CD1a-autoreactive TCRs carried nearly one hundred molecularly distinct ligands, dominated by common families of self lipids, including small hydrophobic lipids like diacylglycerols and ceramide as well as polar lipids with hydrophilic head groups like sphingomyelins and phosphatidylcholines. Lipid tetramer studies were initially viewed as infeasible based on lack of knowledge of any named self lipids that could dominate as antigenic targets. However, such studies are, in fact, readily feasible and reliably identify large T cell pools with functional specificity for CD1a. Thus, these results rule in CD1a-autoreactive T cells as a normal part of the human skin T cell repertoire and provide a straightforward approach to study this T cell population in any skin disease with untreated CD1a-endo tetramers.

Among human CD1 isoforms, CD1a has the highest measured rates of T cell autoreactivity using activation [50] or cloning assays [50, 51], and here we demonstrate high rates of CD1a-autoreactivity among human skin T cells using tetramers. Unlike CD1b, CD1c and CD1d, CD1a lacks any cytoplasmic adaptor protein binding domain that functions as a sorting motif for trafficking to late endosomes, where exogenous antigens are imported into cells and loaded onto CD1 proteins. Both observations support an emerging biological role wherein CD1a, more so than other CD1 isoforms, functions in autoreactivity or immunosurveillance of self lipids [75, 76, 103, 104]. Furthermore, CD1a-autoreactive T cell response measured ex vivo was similar when CD1a was expressed on the surface of human erythroleukemia K562 cells or B cell-like C1R cells, and CD1a tetramers that selected CD1a-autoreactive cells were produced in HEK293 lines. Thus, non-professional APCs support CD1a-autoreactive T cell responses, underscoring

that the CD1a-lipid complex itself in the absence of strong costimulation is necessary and sufficient for response in many conditions.

Based on the high absolute frequency (~1%) of CD1a-autoreactive T cells detected among skin T cells from randomly recruited donors, CD1a-autoreactive T cells might normally subject to some kind of negative regulation, serve some homeostatic immunoregulatory role, or both. In general, human Th22 cells are defined by expression of skin homing receptors including cutaneous lymphocyte antigen (CLA), CCR4, CCR6, and CCR10, as well as IL-22 production. Th22 cells can be polarized from naïve cells by culture with $CD1a^+$ Langerhans cells [65, 66, 88, 105]. In mice, IL-22 is produced with IL-17 by Th17 cells. However 'pure Th22 cells' are defined by production of IL-22 without IL-17. CD1a and pure Th22 cells are absent in mice but present in humans, providing correlative support for the possibility that CD1a could play a central role promoting Th22 function. Whereas a prior study showed enrichment of CD1a-autoreactive responses among blood-derived T cells expressing the Th22-definiting homing receptors [50], here we use CD1a-endo tetramer staining in polyclonal *ex vivo* skin T cells across individuals to directly rule skin residence, IL-22 production, and $CD4^+$ predominance occurring in human skin, further suggesting that CD1a is a natural target of Th22 response. Future tetramer-based sorting and single-cell sequencing studies will provide further definition of the phenotypic and functional profile of CD1a-autoreactive cells, without *a priori* selection of a cytokine or other markers in co-culture activation assays.

This new perspective regarding CD1a-centric response is plausible on a structural basis because CD1 proteins, unlike MHC I and II, lack an open groove on the left side of the platform. Instead, CD1a residues form tethers between the α1 and α2 helices, which form a roof over the A' pocket [8, 9]. This roof could block direct TCR contact to antigens located down under, simultaneously forming an outer antigen-free surface for TCRs to bind. Lipids are sequestered

inside CD1 proteins and emerge to the surface nearer to the right edge of the CD1 display platform [104, 106]. Currently, direct contact of lipid-free surfaces on CD1 but not lipid is known for two TCRs [46, 98], with many counterexamples that contact protruding lipid antigen [52, 53, 54, 107]. Therefore, a major unanswered question was the extent to which lipid-dependent and lipid-independent mechanisms are operative in humans. Here we show that the rate of CD1a-endo tetramer staining is 1% of skin T cells on (range 0.1 to 13%), which is extremely high for single tetramer defined epitope in polyclonal T cells.

Extending recent structural studies showing that adding sulfatide to CD1a can physically disrupt the A' roof, we found here that sulfatide and sphingomyelin are natural endogenous ligands bound to CD1a, but that each when additionally loaded, reduces CD1a tetramer binding to skin T cells. Thus, added lipids are not required for CD1 response in this model, but the presence of endogenous lipid blockers might dampen or tune this 'on until off' recognition system. We explicitly addressed this question in the next chapter. Finally, on a practical level, this CD1-centric mode of activation bypasses a central roadblock to implementing CD1a, and possibly tetramers of other CD1 proteins, in the study of human autoimmune disease. If no defined lipid is required for these frequently observed T cell responses, then the optimized CD1a-endo tetramer staining of tissue derived T cells now represents a straightforward approach to direct monitoring of T cells recognizing native CD1a in any tissue or disease state.

METHODS

Human Subjects

Discarded skin from cosmetic surgeries was obtained through the Human Skin Disease Resource Center at Harvard Medical School and Brigham and Women's Hospital under protocols approved by the Partners Institutional Review Board. Human peripheral blood mononuclear cells (PBMC) were obtained from leukoreduction collars provided by Brigham and Women's Hospital.

Recovery of skin T cells by three-dimensional culture or enzymatic digestion

Skin T cells were recovered after culture for 21-28 days on three-dimensional cell foam growth matrices (Cytomatrix, Australia) seeded with collagen I (Thermo Fisher, 354236). Skin T cell culture media [IMDM, 10-20% FCS, L-glutamine, penicillin-streptomycin, 2-mercaptoethanol] was supplemented with IL-2 (BWH or Peprotech) and recombinant human IL-15 (10ng/mL, Peprotech #200-15) as previously described [62]. Alternatively, fresh skin cells were isolated by enzymatic digestion using a mixture of 30 Kunitz units / ml DNase I from bovine pancreas (Sigma) and 0.2% Collagenase Type I in skin T cell culture media, incubated for up to 2 hrs at 37°C with shaking at 300-500 rpm.

T cell activation assays

Prior to IL-22 ELISPOT or IL-22 capture assays, polyclonal skin T cells were co-cultured 5:1 with irradiated allogeneic Langerhans-like cells and HLA-blocking antibodies for 7-10 days in complete T cell media with 2nM IL-2. *In vitro* Langerhans-like cells (LC) were derived from adherent PBMC over 6 days as previously described [108, 109], using recombinant human TGF-β1

(#100-21), GM-CSF (#300-3) and IL-4 (#200-4) from Peprotech. LC-like cells were irradiated to prevent proliferation of any remnant lymphocytes and treated with antibody clones W632 (10μg/ml) and L243 (10μg/ml) for 1 hr prior to co-culture with allogeneic skin T cells.

Cell lines and in-house antibodies. K562 [50] or C1R [4] cells transduced to express CD1a or CD1b were used as APCs. Antibodies to HLA-A,-B and -C (W6/32), HLA-DR (L243), CD1a (OKT6, 10H3), CD3 (OKT3), and a mouse IgG1 isotype control (P3) were made in-house and added at 10 μg/ml for 1 hour prior to adding T cells.

*ELISPOT***.** IL-22 and IFN-g ELISPOT assays were done according to manufacturer's instructions (Mabtech) using anti-IL-22 (MT12A3), biotin anti-IL-22 (MT7B27), anti-IFN-γ (1-D1K), biotin anti-IFN-γ (7-B6-1). Polyclonal T cells were tested in triplicate at $1x10^5$ T cells per well alone. T cell lines and clones were tested in ELISPOT assays directly without prior co-culture on LC-like cells and at lower cell numbers, typically 10,000 T cells : 20,000 K562 or C1R cells.

IL-22 Capture Assay. The IL-22 secretion assay (MACS Miltenyi) was used with skin T cells re-stimulated by overnight co-culture with K562-CD1a (2:1) in T cell media with 4% human AB serum. The next day, IL-22 secreting cells were labeled over a 45 minute secretion period at 37°C according to manufacturer's instructions with minor modifications, which included lower cell concentration ($1-5x10^5$ cells/ml) during the secretion period. Cells were also labeled with anti-CD3-BV421 and anti-CD4-APC during secondary labeling with anti-IL-22-biotin, and labeled just prior to flow cytometry analysis with a viability dye.

CD1-endo tetramer staining and flow cytometry

Biotinylated human CD1a and CD1b monomers (NIH Tetramer Core Facility and Monash University) were produced in HEK293-derived cell lines in 4 versions. Version 1 CD1a monomer from NIH was previously described [73, 97, 110], is produced in lentivirus-transduced HEK293T cells. Versions 2 and 3 of NIH CD1a monomers were expressed in HEK293T cells deficient in *MGAT* and *TAP2* genes and used a lentiviral vector for separate expression of β2m and CD1a with a double hexahistidine tag. Only version 3 contains a cysteine to serine swap meant to reduce secreted protein aggregation, and these were used interchangeably for polyclonal skin T cell staining and sorting. Version 4, produced at Monash University, was expressed in HEK293S GnTI- cells as previously described [46] and treated with endoglycosidase H (New England Biolabs).

CD1 monomers were adjusted to 0.2 μg/μl in 50mM Tris-buffered saline pH 8.0 (Sigma) with or without 0.5% CHAPS (Sigma) prior to tetramer assembly 1:5 w/w with streptavidin-PE (BD Biosciences) for a final concentration of 0.1 μg/μl with respect to CD1a monomer. Antibodies and reagents were Blue Live/Dead (Life Technologies), Near IR Live/Dead (Life Technologies), CD4-APC (clone RPA-T4, BD Biosciences), CD4-APC-Cy7 (clone OKT4), CD8α-BV711 (clone RPA-T8, Biolegend), CD3-BV421 (clone UCHT1, Biolegend), TCR Vb22-FITC (clone IMMU546, Beckman Coulter), and TCR Vβ23-FITC (clone REA497, Miltenyi). Cells were washed twice in PBS and pulsed with Live/Dead stain in PBS for 15 minutes RT, washed once with PBS 1% bovine serum albumin +/- 0.01% sodium azide (FACS Buffer), and resuspended in FACS buffer at a maximum of 1 million cells / 50μl. Per 50μl staining volume, 1ul tetramer PE was added to cells and incubated for 25-30 mins at room temperature in the dark. Without washing, 0.1 μg unlabeled anti-CD3 (OKT3) was added and incubated for an additional 10 minutes at room temperature. Labeled surface stains were then added for 15

minutes on ice. Cells were washed once in FACS buffer and resuspended in PBS for immediate acquisition on an LSR Fortessa or a custom-modified FACSAria II flow cytometer (BD Biosciences) and analyzed with FlowJo v10.5.3 (Tree Star).

Lipid sources

Phosphatidylcholine (PC 850475), C18:1 lyso-PC (845875), C24:1 sphingomyelin (SM, 860593), milk sphingomyelins (SMs, 860063), C24:1 β–galactosylceramide (860546), and phosphatidylinositol (PI, 840042) were purchased from Avanti Polar Lipids. Sulfatides (1049) were from Matreya. Diacylglycerol (D0138) was from Sigma.

Loading CD1a monomers with defined lipids for tetramers

Lipids stored in chloroform and methanol were transferred to new borosilicate glass tubes, dried under nitrogen gas and reconstituted to 400 μM in TBS pH 8.0, 0.5% CHAPS buffer by sonication in a 37°C water bath for ~1 hr. Lipid-buffer sonicates or a buffer-only control were transferred to 1.5ml eppendorf tubes on a 37°C heat block to which CD1a monomer was added to a final concentration of 0.2 μg/μl and incubated for 2hrs at 37°C, then overnight at room temperature. Loaded monomers were stored at 4°C and used for tetramer assembly and staining within 3 months.

Generation of mutant CD1a tetramers

Wild type and mutant CD1a proteins were recombinantly produced in human embryonic kidney cells (HEK293 GnTI$^-$). Both, β_2-microglobulin and the heavy chain of CD1a were expressed as a single construct from a modified pHLsec vector, in which the original signal peptide was substituted by the Igκ leader sequence and the P2A self-cleavage site was introduced between

both genes. Both chains carried leucine zippers Fos (β_2m) and Jun (CD1a). Additional BirA- and 6xHis- tags were fused to the carboxy-terminus of CD1a. The CD1a/β_2m heterodimer was purified by nickel affinity and size exclusion chromatography. Pure protein was biotinylated using BirA ligase followed by a gel filtration run.

Generation of T cell lines

The DermT parent line was previously derived by limiting dilution expansion from PBMC [50, 56]. Sorting from the DermT line to generate clone DermT2 and sorting from polyclonal skin T cells from 3D culture was done with LPC treated CD1a tetramers, which previous experiments showed to be similar to CD1a-endo. For two donors (29, 30) skin T cells were co-cultured with allogeneic LC-like cells as for activation assays prior to tetramer staining. Live, $CD3^+$ $CD4^+$ CD1a-LPC $tetramer^+$ cells were sorted and then expanded over 14-18 days by stimulation with $25x10^6$ irradiated allogeneic PBMC, $5x10^6$ irradiated Epstein-Barr virus-transformed B cells, 30 ng/ml anti-CD3 (OTK3), and with 2nM IL-2 added on day 2 of the culture.

T cell receptor PCR and sequencing

RNA was isolated from sorted T cell populations using the RNeasy kit (Qiagen) and cDNA synthesized using the Quantitect Reverse Transcription Kit (Qiagen). TCR V gene usage was determined by PCR using primer sets IPS000029 and IPS000030 as described at www.imgt.org/IMGTPrimerDB in combination with TCRα constant region reverse primer 5'GTGGTAGCAGCTTTCACCTCCTTGG 3 and TCRβ constant region reverse primer 5' GGTGGCAGACAGGACCCCTTGC 3'.

HPLC-MS analysis of eluents from CD1a monomers.

Version 2 CD1a protein and HLA-DRB1*03:01 protein (80μg) were extracted in triplicate in 15-ml glass tubes using chloroform, methanol and water [111]. The organic phase was recovered and dried under nitrogen gas. Eluent residue was redissolved in chloroform-methanol, normalized to 20 μM based on input protein and stored at -20°C. For the normal phase HPLC-MS We injected 20μl for Q-ToF HPLC-MS positive ion mode analysis using an Agilent 6530 Accurate-Mass Q-TOF and 1260 series HPLC system with a normal phase Inertsil Diol column (150 mm × 2 mm, GL Sciences), running at 0.15 ml/min as described [39, 100, 112].

For the reversed-phase HPLC-MS, an Agilent Poroshell 120 A, EC-C18, 3 x 50 mm, 1.9 *μ*m column coupled with an Agilent EC-C18, 3 x 5 mm, 2.7 *μ*m guard column were used based on the published methods[113] with slight modification as following: mobile phase (A) methanol /water (95/5; V/V) supplemented with 2 mM ammonium-formate and mobile phase (B) 1-propanol/cyclohexane/water (90/10/0.1; v/v/v) supplemented with 3 mM ammonium-formate. The solvent gradient changes in a 20-minute run: 0-4 min, 100% A; 4-10 min, from 100% A to 100% B; 10-15 min, 100%B; 15-16 min, from 100% B to 100% A; 16-20 min, 100% A. The columns were equilibrated by running 100% A for 5 min before the next run. The lipid quantification was determined by the individual ion chromatogram peak area compared to the external standard curves[100].

Statistics. Statistical analyses were done in GraphPad Prism 6, using the Wilcoxon matched pairs signed rank test unless otherwise stated.

CHAPTER 3. Sphingomyelins as natural blockers of CD1a autoreactive T cells

Parts of this chapter are currently in preparation for publication as:

Rachel N. Cotton*, Marcin Wegrecki*, Tan-Yun Cheng, Natacha Veerapen, Jérôme Le Nours, Dennis P. Orgill, Bohdan Pomahac, Simon G. Talbot, Richard Willis, John D. Altman, Annemieke de Jong, Rachael A. Clark, Gurdyal S. Besra, Jamie Rossjohn, D. Branch Moody. **Cellular CD1a selectively captures lipids that sterically hinder CD1a autorecognition by T cells in human skin**

* These authors contributed equally.

INTRODUCTION

CD1a binds many endogenous cellular self lipids to form CD1a-endo complexes that are recognized by T cells [39, 46, 56]. Whereas most studies of CD1 ligands emphasize their antigenic function in activating T cells, CD1 ligands could have diverse functions. For example, a lipid or class of lipids might serve a loose placeholder function to protect the groove during trafficking from the endoplasmic reticulum until it captures biologically relevant antigens in more distal cellular compartments, analogous to Class II associated invariant peptide (CLIP) in the MHC system [114, 115]. In addition, studies in the CD1b system had identified scaffolding lipids, which are space filling molecules that bind within the CD1 cleft and position other lipids for recognition [39, 116].

Given the high frequency of T cell autoreactivity to CD1a proteins found in human peripheral blood and skin, yet a lack of ubiquitous and commensurate T cell mediated inflammation, we considered possible mechanisms by which autoreactivity could be dampened to avoid autoimmunity. For example, CD1 ligands might serve a regulatory function by blockade

of recognition by autoreactive T cell receptors, or by out-competing other kinds of ligands that favor recognition and activation. We reasoned that lipids serving such scaffold or regulatory functions, if such ligands exist for CD1a, could be discovered in a new kind of cellular screen that emphasizes CD1a binding to cellular lipids, instead of using T cell activation assays as readouts. Our prior studies (**Chapter 2**) established HEK293T cells load at least 98 types of lipids onto cellular CD1a proteins as they fold and exit via the secretory pathway. However, this large number of lipids exceeded our capacity to identify and study them all by mass spectrometry. Therefore, using a comparative lipidomics platform, we asked if CD1a neutrally captured any type of lipid present in HEK293T cells or instead had intrinsic or cellular capture mechanisms to select out lipids based on their loading or binding to CD1a. In the latter case, we sought to identify the CD1a ligands and determine their biological functions.

RESULTS

Lipidomic analysis of cells versus CD1a

I worked with Tan-Yun Cheng to analyze comparative lipidomics of all ionizable lipids in HEK293T cells versus the spectrum of lipids captured by transmembrane truncated CD1a proteins expressed in this cell line. The comparative lipidomics platform takes advantage of low-resolution, normal phase high performance liquid chromatography (HPLC) to broadly separate lipids extractable into chloroform and methanol [117]. Positive mode quadrupole time of flight (QToF) mass spectrometry (MS) has been previously optimized to give sensitive detection of CD1 ligands over a broad dynamic range with low rates of false positivity as compared to eluents of MHC proteins [39, 56]. We focused on the brightest ions in the CD1 eluents, tentatively identifying lipids that matched the elution time of standards and the reported *m/z* values for major classes of neutral lipids, glycolipids, phospholipids and sphingolipids, including phosphatidylcholine (PC) and sphingomyelin (SM). PC was chosen as an example of an abundant anionic membrane lipid, and sphingomyelin (SM) is a less abundant cellular sphingolipid the forms microdomains.

For each class, one lead compound was formally identified using collision-induced dissociation (CID) MS. For example, CID-MS of the ion at *m/z* 760.59 matched the expected mass of phosphatidylcholine (PC), whose two fatty acyl units combined for 34 methylene groups and 1 unsaturation (34:1 PC). The ion of *m/z* 813.68 matched the expected mass of sphingomyelin (SM), where the combined fatty acyl and sphingosine base 42 methylene units and two unsaturations (42:2 SM) (**Fig. 3.1A**). Both PC and SM contain phosphocholine head groups, and both ions showed fragments that lost a mass interval of m/z 184.07, the diagnostic fragment for phosphocholine (**Fig. 3.1A**). Next we counted the number and assigned the

molecular variants within each lipid class. We could infer the total number of methylene units (X) and unsaturations (Y) of the two combined alkyl chains present in every member of each class (X:Y). Considering ions released from CD1a that nearly co-elute with 34:1 PC (20.4 min-21.5 min), we deduced 15 additional PC variants ranging from C34 to C42 with up to 5 unsaturations (**Fig. 3.1B, blue**). Using the same approach we identified 5 additional SMs eluting from CD1a that nearly co-eluted with 42:2 SM (**Fig. 3.1B, red**). The names, formulas and experimental mass errors for the 21 CD1a ligands are detailed in **Table 3.1**.

Table 3.1. Phosphatidylcholines and sphingomyelins identified by CID-MS.

phosphatidylcholine (PC)

detected m/z	RT (min)	Formula	predicted m/z	lipid chain
870.6946	20.4807	C50H97NO8P [M+H]+	870.6947	C42:2
842.6625	20.7123	C48H93NO8P [M+H]+	842.6634	C40:2
816.6458	20.7345	C46H91NO8P [M+H]+	816.6477	C38:1
788.6158	20.9596	C44H87NO8P [M+H]+	788.6164	C36:1
814.6308	20.9754	C46H89NO8P [M+H]+	814.6321	C38:2
812.6154	21.0952	C46H87NO8P [M+H]+	812.6164	C38:3
838.6302	21.1123	C48H89NO8P [M+H]+	838.6321	C40:4
786.6005	21.1556	C44H85NO8P [M+H]+	786.6008	C36:2
760.5846	21.2677	C42H83NO8P [M+H]+	760.5851	C34:1
810.5992	21.2677	C46H85NO8P [M+H]+	810.6008	C38:4
836.6164	21.2725	C48H87NO8P [M+H]+	836.6164	C40:5
808.5845	21.3798	C46H83NO8P [M+H]+	808.5851	C38:5
784.5850	21.3834	C44H83NO8P [M+H]+	784.5851	C36:3
758.5699	21.4230	C42H81NO8P [M+H]+	758.5695	C34:2
782.5675	21.4277	C44H81NO8P [M+H]+	782.5695	C36:4
732.5520	21.5386	C40H79NO8P [M+H]+	732.5538	C32:1

sphingomyelin (SM)

detected m/z	RT (min)	Formula	predicted m/z	lipid chain
827.7029	21.6214	C48H96N2O6P [M+H]+	827.7001	C43:2
815.6946	21.6448	C47H96N2O6P [M+H]+	815.7001	C42:1
801.6868	21.7086	C46H94N2O6P [M+H]+	801.6844	C41:1
813.6836	21.7335	C47H94N2O6P [M+H]+	813.6844	C42:2
799.6680	21.8713	C46H92N2O6P [M+H]+	799.6688	C41:2
811.6670	21.9329	C47H92N2O6P [M+H]+	811.6688	C42:3

Quantitation of cellular lipids with mass spectrometry

Next Tan-Yun Cheng deconvoluted the ion chromatograms corresponding to the *m/z* of each variant and integrated the area under the curve to rank compounds corresponding to relative abundance (**Fig. 3.1C, insets**). She tested the mass spectrometry response factor to injections of known lipids over a broad does range with intensity tested up to 3 x 10^7 count-seconds (**Fig. 3.1D**). Signal was linearly related to the mass injected for PC and MS. Further, the response factors for PC and SM in positive electrospray ionization mass spectrometry (ESI-MS) were equivalent at both low and high dose ranges, as expected based on their similar head group structures (**Fig. 3.1D**). Therefore, for ion chromatograms, the relative areas under the curve for each compound provide a good approximation of the relative absolute concentration of each lipid present. The most abundant PC species (34:1 PC) was ~ 8-fold more intense than the most abundant SM (34:1 SM) (**Fig. 3.1C, upper**). This ratio is in general agreement with prior reports [118], where PC is a major structural component of anionic bilayers and SM is broadly present in cellualr membranes but a much less abundant lipid with specialized roles [118, 119].

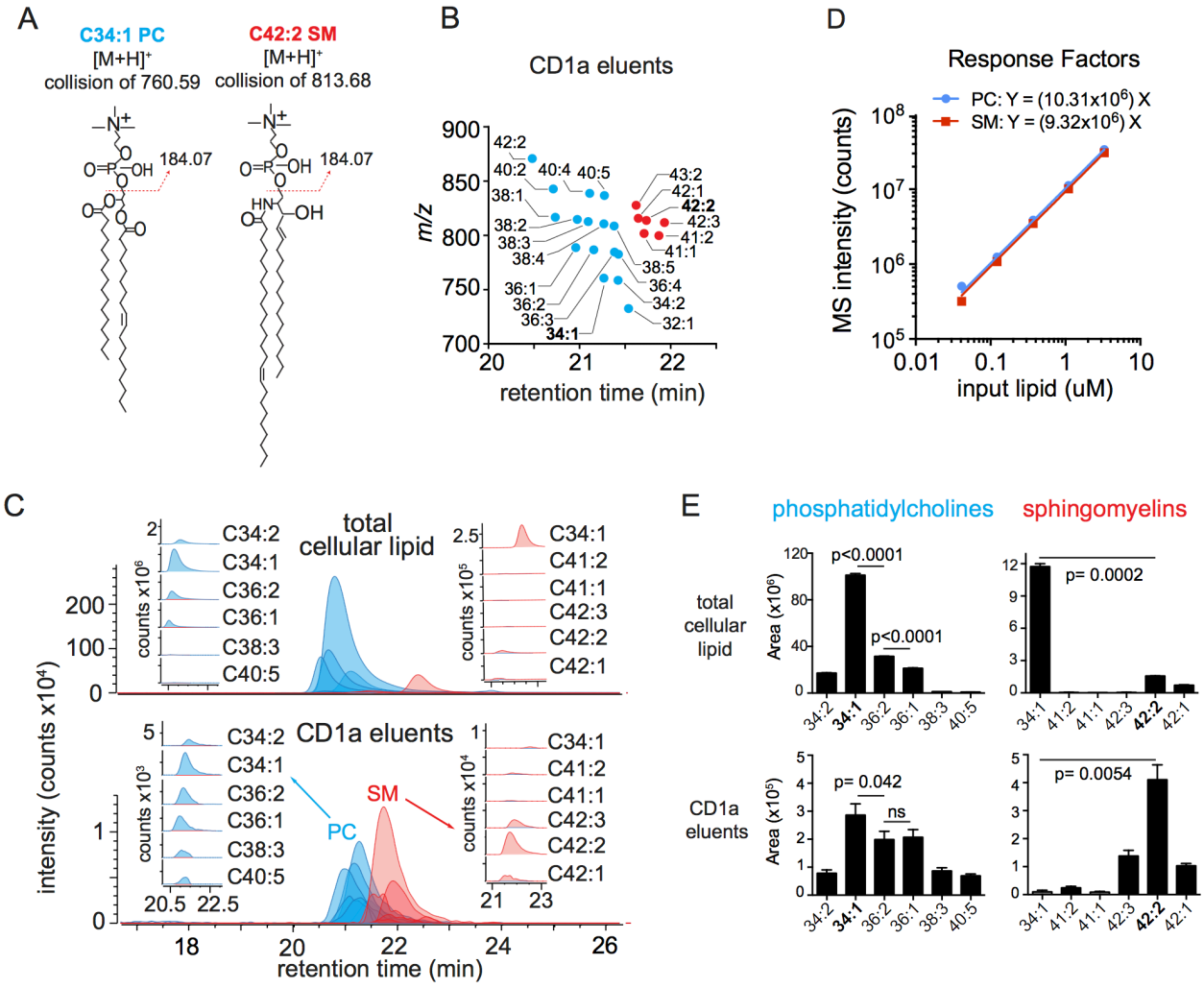

Figure 3.1. Targeted HPLC-MS analysis of CD1 monomer eluents and cellular lipids. **A)** Collision induced fragmentation for CD1a ligands matching the expected m/z of C34:1 PC and C42:2 SM. **B)** Chain length and saturation variants identified within the same lipid class have similar retention times and *m/z* that vary by an integer number of methylene units (chain length variants) or H2 (unsaturations). We identified 16 molecular variants of PC and 6 SMs. **C)** Mass chromatograms of the 6 most abundant PC and SM family members eluted from NIH CD1a monomer (top) and in the total lipid extract from matched CD1a producing HEK293T MGAT- cells (bottom). **D**) Response factors for SM and PC were highly similar based on a synthetic standard curve. **E**) Phosphatidylcholines, sphingomyelins, quantified as integrated area under the curve (count-seconds) for each lipid chain variant detected in triplicate at a diagnostic *m/z* value and retention time window. Error bars indicate standard deviation from the mean. P values were calculated using Welch's corrected t test.

CD1a-bound lipids versus cellular lipids

Comparison of triplicate lipidomic profiles from HEK293T cells to lipid eluents of CD1a produced in the same cell type generated two striking findings. First, whereas PCs dominated in cells, SM species showed higher signals in CD1a eluents, representing an inversion of the PC:SM ratio. This apparently strong effect was quantitatively validated by determining chromatogram areas for the 6 most abundant molecular species in each class (**Fig. 3.1C insets, Fig. 3.1E**). Second, the SMs bound to CD1a proteins skewed toward significantly longer and more unsaturated alkyl chains as compared to the cellular pool of SMs (**Fig. 3.1E**). Among cellular SMs, a short chain 34:1 SM species was predominant, with its area under the ion chromatogram measuring 7.5-fold more than 42:2 SM, which was scarcely detectable ($p=0.0002$) (**Fig. 3.1C, inset, blue**). In CD1a eluents, 34:1 SM was so low in intensity that it was not initially picked up by the automated peak-picking software. By manually searching for this specific mass and integrating the peak area, we found that 34:1 SM peak area was 44-fold less than the peak area for 42:2 SM ($p=0.0054$). Longer chain lipids might remain preferentially bound to CD1 proteins during the protein purification process, and other aspects of the experimental system might have influenced capture. However, the preference for longer chain doubly unsaturated lipids was not seen among PC species, as 34:1 PC was the most abundant species in both cells and CD1a eluent (**Fig. 3.1D**).

We next sought to test for selective lipid capture in an independent set of eluents, analyzed in triplicate, from secreted CD1a, CD1b, and CD1c proteins that had also been produced HEK293T cell lines (**Fig 3.2A**). This second CD1a protein batch confirmed both of the prior findings, as SMs were greater than PCs, and 42:2 SM was the most frequent SM species, followed by 42:1 and 42:3 SM variants. In CD1b and CD1c, some skewing toward longer SMs was seen, but it was much weaker than for CD1a, and the short chain 34:1 SM was the

predominant SM released from CD1c. Finally, in three separate CD1a preparations with a different protein design and produced in HEK293S cells, both the SM predominance and the skewing to longer 42:2 SM in CD1a eluents were confirmed (**Fig. 3.2B**). These experiments were performed with independently engineered CD1 proteins with differing linking mechanisms, capture tags, buffers, and purification methods. These specificity controls suggested that SM predominance over PC, and the trend toward longer SM chain lengths for CD1a capture, were not a function of the CD1 protein production, elution, and MS detection systems, but instead a reproducible effect, found only for the particular combination of CD1a and SM. We interpret the over-representation of lipids in the CD1a monomer eluent as preferential loading or retention in CD1a proteins traversing the secretory pathway.

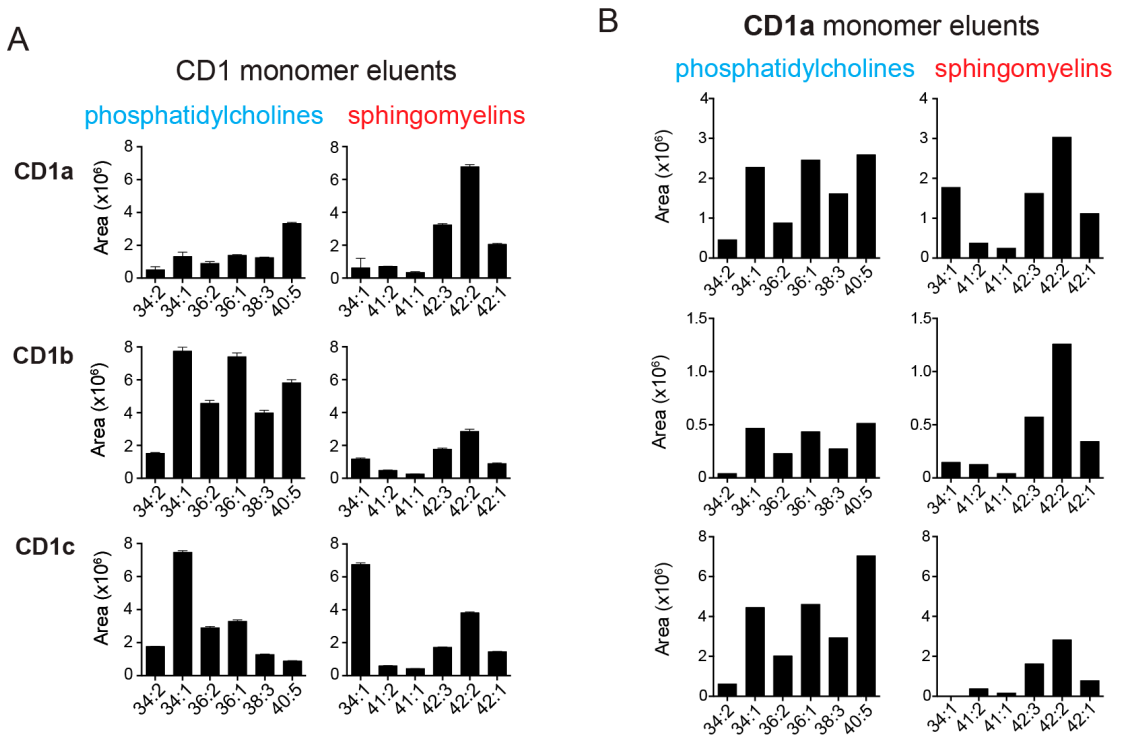

Figure 3.2. Targeted HPLC-MS analysis of CD1a, CD1b, and CD1c eluents.
A) PCs and SMs, eluted from CD1a, CD1b, and CD1c protein monomers produced separately, quantified as integrated area under the curve (count-seconds) for each lipid chain variant detected in triplicate at a diagnostic *m/z* value and retention time window. Error bars indicate standard deviation from the mean. **B**) PCs and SMs, eluted from three separate CD1a preparations produced at Monash University, independently from A, quantified as in A.

Molecular analysis of natural and synthetic SMs

To determine the molecular basis of SM capture by CD1a, we first assembled a panel of synthetic sphingomyelins that differed systematically only in alkyl chain length and unsaturations. ESI-MS provides certain advantages in sensitivity and breadth of detection for the micro-analysis of lipids present in trace amounts in protein eluents (**Fig. 3.1**), but this method could not establish all aspects of the structures of SM species eluted from CD1a proteins (**Fig. 3.3A**). For example, the total number of methylene units and unsaturations can be reliably inferred, but the position and stereochemistry of unsaturations in both the sphingosine base and alkyl chains not formally known. Sphingosine, also known as D-erythro-sphingosine or (E)-sphing-4-enine in reference to the *trans* unsaturation at the 4th carbon (abbreviated d18:1) is the major sphingoid base in mammals, and so the sphingosine base unsaturation is inferred [120]. However the location of the fatty acyl unsaturation is unknown and might be varied (**Fig. 3.3A**). Also, for any single ion predicting chain total chain length of C42, the actual composition likely contains isobaric lipids, where incremental increases in the sphingosine base are matched to shorter fatty acyl units and vice versa. In contrast, the actual length as well as position and stereochemistry of synthetic molecules is known (**Fig. 3.3A, right**). Accordingly, we obtained SMs with the d18:1 $\Delta 4$ *trans* unsaturated sphingosine base in without (18:1 SM) an amide linkage or with amide linked fatty acids up to C24:0 (known as 42:1 SM, which corresponds to length and unsaturation in both chains), or C24:1 (42:2 SM) *cis* unsaturation at the indicated position (**Fig. 3.3B**).

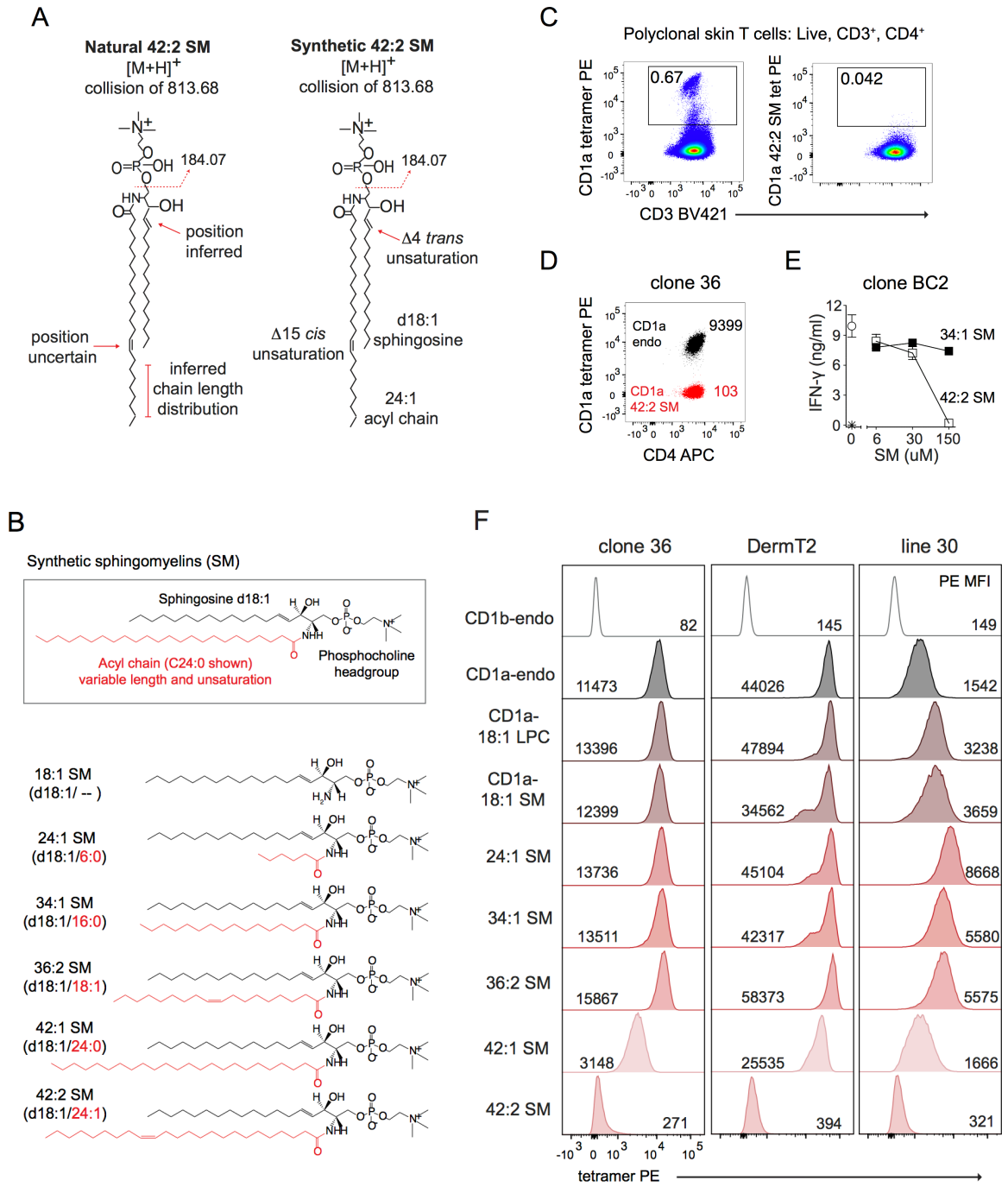

Figure 3.3. Lipid chain length and unsaturation determine CD1a tetramer binding. **A)** Natural versus synthetic 42:2 SM. **B**) Synthetic SM acyl chain variants. **C)** Polyclonal skin T cells stained with CD1a tetramer or tetramer treated with 42:2 SM. **D)** CD1a autoreactive skin T cell clone 36 staining with CD1a-endo or CD1a-42:2 SM. Insets indicate tetramer mean fluorescence intensity (MFI). **D)** IFN-γ ELISA of clone BC2 T cells (star indicates T cells only) and plate-bound CD1a treated with synthetic 34:1 SM (black) or 42:2 SM (white square) or untreated (white circle). **F)** CD1a-SM tetramer staining of skin T cell lines DermT2, Line 30, and Line 36. Pre-gated: live, $CD3^+$, $CD4^+$ $Autofluoresence(FITC)^{neg}$. Histograms normalized to mode.

Sphingomyelin influence on CD1a-TCR binding

Next we considered the immunological consequences of selective capture of long chain unsaturated sphingomyelins. Sphingolipids, including the NKT cell agonist α-galactosyl-ceramide[121] are antigens in the CD1 system, but sphingomyelin itself has not been broadly or prominently identified as a T cell activating lipid. To search for 42:2 SM- specific T cells, we tetramerized C42:2 SM-treated CD1a monomers and stained polyclonal T cells (**Fig. 3.3C**). The frequency of $CD3^+CD4^+$ T cells staining a control CD1a tetramer that carried endogenous lipids (CD1a-endo) from the CD1 expression system, was high in absolute terms (0.67%). This finding was consistent with our prior observations where 1% percent of T cells stain with CD1a-endo complexes comprised of many lipids, which was found to result from direct CD1a-TCR interactions (**Chapter 2**). Notably, staining with 42:2 SM treated tetramers reduced the rate of staining by 14-fold (**Fig. 3.3C**). Thus, 42:2 SM itself was not an immunodominant antigen in the patient tested and instead strongly blocked staining that could otherwise be seen with CD1a-endo tetramers. This result, along with hints from in vitro T cell assays from prior studies[56], suggested that 42:2 SM might inhibit CD1a-TCR contact.

To assess this hypothesis in more defined systems with single TCRs, we used the CD1a autoreactive skin T cell line, clone 36, which is likely clonal because it stains > 99% by the TCR Vβ-specific antibody that recognizes Vβ23. Clone 36 binds brightly to CD1a-endo tetramer, and treating CD1a monomers with synthetic 42:2 SM reduced staining intensity (MFI) by 90-fold (**Fig. 3.3E**). Next, when we stripped and treated plate-bound CD1a proteins with either 34:1 SM or 42:2 SM and co-cultured with the CD1a-autoreactive T cell clone BC2 (**Fig 3.3D**), we found that increasing concentrations of 42:2 SM, but not 34:1 SM, reduced the magnitude of IFN-y response. This hinted that long chain length, which also determined the selective capture of SM by CD1a, might be the molecular feature that controls inhibition of CD1a and TCR contact.

Molecular fine mapping of CD1a-TCR blockade

To formally determine the molecular basis of alkyl chain length and unsaturation roles in CD1a-TCR interaction, I tested the panel of chemically defined synthetic sphingomyelins comprised of a 18:1 sphingosine base with the Δ4 *trans* unsaturation and various fatty acids (**Fig. 3.3B**). I treated CD1a monomers with these 6 candidate blockers, tetramerized and used to stain 3 human T cell lines with partially characterized or known TCRs: clone 36, DermT2 (TRBV2*01 TRBJ1-5*01, TRAV8-6 TRAJ20*01) and Line 30 (>90% Vβ22+) (**Fig. 3.3F**). I chose these three T cell lines for their low (Line 30, MFI=1,542), intermediate (Line 36, MFI=11,473) or high (DermT2, MFI=44.026) staining with CD1a-endo tetramers, creating a high range of baseline values from which each line might show differing sensitivity to alternatively loaded tetramers (**Fig 3.3F**). These three lines have non-identical TCRs, which could have different modes of docking on CD1a. I used two negative controls: CD1b-endo tetramers, which are not expected to bind to CD1a autoreactive lines, and CD1a-lysophosphatidylcholine (LPC), which is a small phosphocholine-containing ligand that does not generally block CD1a-TCR [46, 122] or CD1c-TCR interactions [98]. CD1a tetramers formed with shorter chain lipids, including lyso-phosphatidylcholine (LPC), 18:1 SM, 24:1 SM, 34:1 SM, and 36:2 SM stained lines at a similar or slightly higher mean fluorescence intensity than the heterogeneously loaded CD1a-endo tetramers (**Fig. 3.3F**). For line 30, LPC and the SMs up to C36 in length all augmented staining compared to CD1a-endo. In contrast, both C42 SM ligands blocked CD1a staining for all lines tested. CD1a-42:1 SM tetramers showed decreased levels of staining in all three lines at 2-5-fold lower MFI than CD1a-endo tetramers. Strikingly, the doubly unsaturated variant 42:2 SM blocked tetramer binding with 50-100-fold reduced MFI, so that staining was reduced to the background levels seen with CD1b tetramers, suggesting blockade of CD1a-TCR interactions.

These clear patterns supported several conclusions. First, smaller lipids could augment staining above that seen in CD1a-endo in some cases, consistent with the idea that shorter lipids and small molecules do not protrude to cover the A' roof [43, 46] (**Chapter 2**). In contrast, long chain SMs were inhibitory for CD1a-TCR contact. Second, C42 but not C36 SMs inhibited staining, indicating a length difference of just C6 conferred the blocking effect. Finally, blockade of CD1a tetramer binding was greatest when the long-chain SM had an additional unsaturation in the fatty acyl chain. We interpret this loss of staining to suggest that these similar SMs may in fact generate CD1a-SM complexes that are structurally distinct. A favored molecular hypothesis is that the longer chain length and unsaturation could in some way position SMs so that they more stably bind CD1a, or more substantially disrupt a TCR-CD1a binding interface.

Poor capture and retention in CD1a of the SM form dominant in cells

To determine if and how different sphingomyelins affect CD1a-lipid complex architecture, Marcin Wegrecki in the Rossjohn group used SM that I identified to treat CD1a with 34:1 SM, 36:2 SM, 42:1 SM, 42:2 SM, with the goal of achieving uniformly loaded CD1a-SM complexes for crystallization. Based on our initial finding that 34:1 SM was the dominant SM in HEK293 cells and therefore physiologically relevant, the Rossjohn group first attempted to solve a CD1a-34:1 SM binary structure. However, isoelectric focusing (IEF) performed on 34:1 SM-treated CD1a showed multiple, poorly focused bands suggesting poor loading or retention of 34:1 SM (**Fig. 3.4A**). In agreement with this interpretation, the resulting crystal structure was poorly resolved (2.7Å), and with no defined electron density within the cleft. These results from *in vitro* exchange experiments were consistent with prior cellular studies (**Fig. 3.1**), where this shorter chain 34:1 SM shows the highest MS intensity among PC species in cells, yet is the least

abundant form eluted from CD1a monomer produced in those cell lines (**Fig. 3.1C-D**). We interpret this to mean 34:1 SM is poorly captured or retained in CD1a at steady state.

Overview of CD1a-sphingomyelin structures

For 36:2 SM, 42:1 SM, and 42:2 SM, we observed that CD1a bands in in isoelectric focusing gels were more homogenous relative to CD1a-endo, suggesting high occupancy of CD1a with these sphingomyelins (**Fig. 3.4A**). The three binary complexes all crystallized leading to high resolution structures: CD1a-42:1 SM complex at 2.4 Å resolution, CD1a-42:2 SM complex at 2.0 Å resolution, and CD1a-36:2 SM complex to 2.1 Å resolution. Unlike CD1a-34:1 SM complexes, the intra-cleft densities were well defined, leading to the unambiguous assignment of the lipid ligand positions in all three cases (**Fig 3.4B**). The CD1a cleft has two pockets, A' and F', which connect to the outer surface of CD1a via the F' portal.

The internal volume of CD1a has been measured at various values ranging from 1280 to 1650Å^3, corresponding to a chain length capacity of ~C34-C39 [8, 9]. Therefore the combined fatty acyl and sphingosine chains in these three SM molecules of C36-42 were expected to just fill or perhaps overfill the CD1a cleft and protrude through the F' portal. Indeed, for C42:1 SM and C42:2 SM complexes respectively, 16% and 17% of the ligand protrudes and is solvent exposed on the outer surface of CD1a. In contrast, the 36:2 SM is sequestered within CD1a so that it is only 4% is solvent exposed (**Fig. 3.4B**).

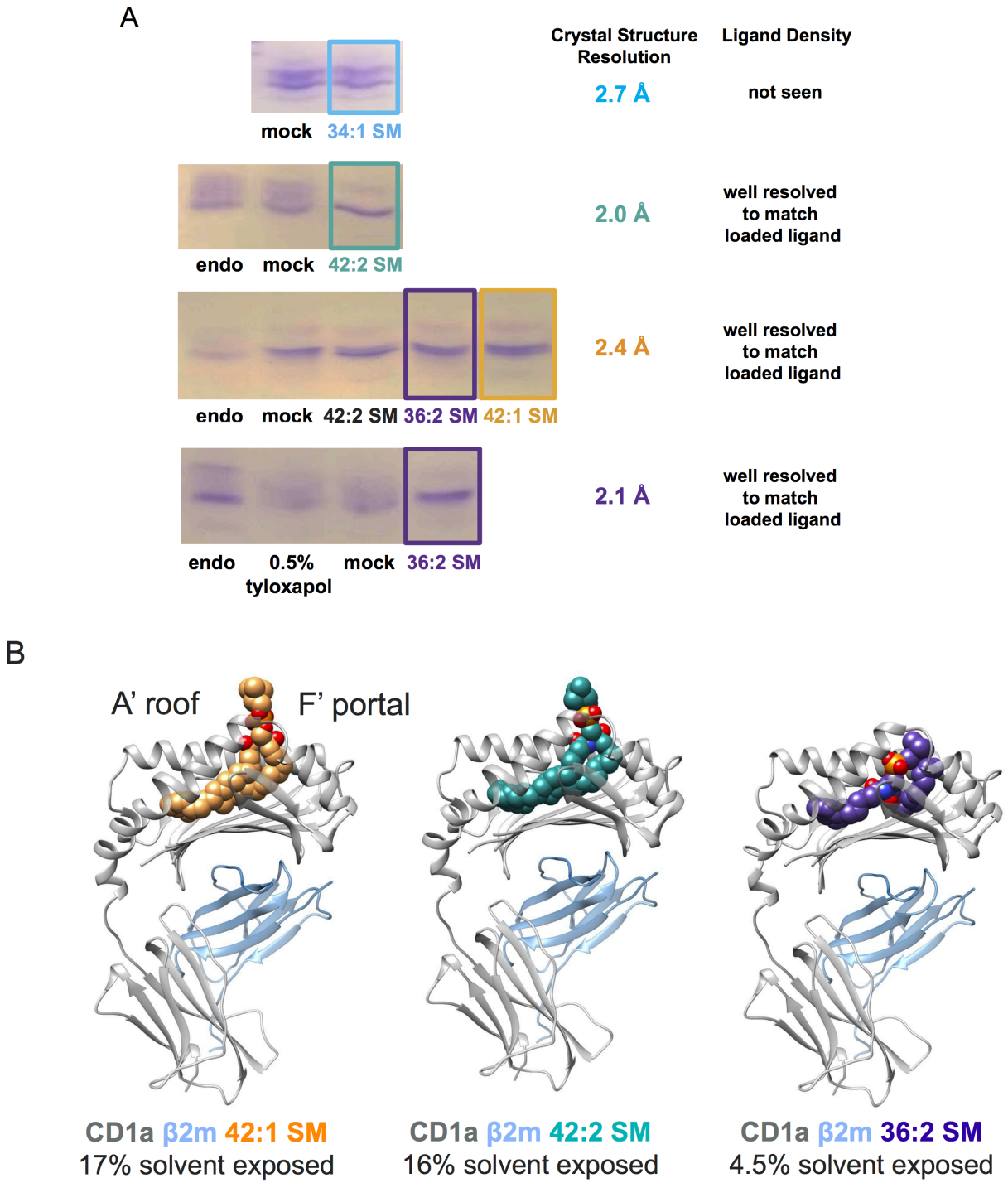

Figure 3.4. Loading and structure elucidation of CD1a-SM complexes. A) CD1a monomers treated and loaded for crystallography studies were assessed qualitatively by isoelectric focusing gel for uniform ligand loading. Prior to C36:2 SM loading, dzdgCD1a-endo was washed with 0.5% tyloxapol. C36:2 SM was then loaded using CHAPS as above. This method yeilded a CD1a-36:2 SM structure of 2.1 Å resolution. **B**) CD1a-42:1 SM (orange), CD1a-42:2 SM (teal), and CD1a- 36:2 SM (purple).

The substantial protrusion of the phosphocholine head groups of 42:1 SM and 42:2 SM to the outer surface of CD1a, where the TCR is expected to bind, provides a straightforward explanation for their observed ability to block CD1a tetramer binding to T cells. The simplest structural explanation might be that the C6 length increment between C36 SM and C42 SM creates protrusion on a size basis of otherwise similarly seated molecules, pushing the phosphocholine groups through the F' portal towards the surface. However, while there is some size effect contributing to protrusion, comparative analysis of the C42 and C36 CD1a-SM structures show marked differences in lipid seating within CD1a, revealing specific mechanisms of lipid and CD1a remodeling after binding different SMs.

An ionic platform stabilizing CD1a

For 42:1 SM and 42:2 SM ligands, the lipid chains lie in nearly overlapping positions, with the fatty acyl and sphingosine chains lying in a parallel orientation within the A' pocket. The C24 fatty acyl chain inserts more deeply than the sphingosine base and wraps around the A' pole, a structure formed by CD1a residues F70 and V12 (**Fig. 3.5**). In contrast, sphingosine and fatty acyl chains in 36:2 SM orient antiparallel (**Fig. 3.5B**). The C18:1 acyl chain is seated in the A’ pocket and fully wraps around the A’ pole. The d18:1 sphingosine threads in the opposite direction into the F’ pocket such that the sphingosine chain effectively shares the F’ pocket with the phosphocholine head group (**Fig. 3.5B**). Interestingly, in the C42 versus C36 SM structures, we observe a flipped orientation of the W14 residue at the internal floor of the hydrophobic cleft, as if to differentially allocate cleft volume between A’ and F’ pockets for an induced fit based on SM chain length.

As compared to the 42:2 SM, the shorter acyl chain in 36:2 SM and the deep seating of the chain into the A' pocket plunges the phosphate group ~7Å deeper into CD1a. This lower

position buries the phosphate of SM within CD1a, where it makes ionic interactions with R73 and R76, as well as a hydrogen bond between the oxygen on the acyl chain and R73 (**Fig. 3.5B**). These three charge-charge interactions create a stable network that acts like a platform to lock 36:1 SM in more deeply within CD1a at a defined position within the neck of the A' pocket. Also, the interaction creates a bend in the phosphocholine head group so that it runs parallel to and below the presumed CD1a-TCR contact surface (**Fig. 3.5B**). The longer C24 acyl chains in 42:1 SM and 42:2 SM push these ligands higher in the cleft so R73 does not interact with SM, and the oxygen on the SM sphingosine chain forms a hydrogen bond with R76 (**Fig. 3.5**).

These interactions, force the phosphocholine group into a vertical position running perpendicular to the presumed CD1a-TCR contact surface, and unlike 36:1 SM, swings upwards towards the expected TCR binding site on CD1a (**Fig 3.5**). The phosphocholine unit protrudes through the F' portal, disrupting the R73-E154-R76 salt bridge and the A' roof (**Fig. 3.5**). In both 42:1 SM and 42:2 SM-bound structures, the bulky phosphate group is positioned close to the side chain of E154 and forces R76 to orient vertically. This molecular arrangement was previously proposed to negatively impact the recognition of CD1a by the BK6 TCR, and illustrates the situation in which an inhibitory ligand renders a surface epitope of CD1a unavailable to TCR binding [46]. In contrast, when CD1a harbors the shorter 36:2 SM ligand, the A' roof integrity is maintained (**Fig. 3.4B. 3.5B,D**). The lipid does not substantially impact the surface of CD1a, leaving a broad lipid-free stretch of the CD1a platform on which autoreactive TCRs may dock.

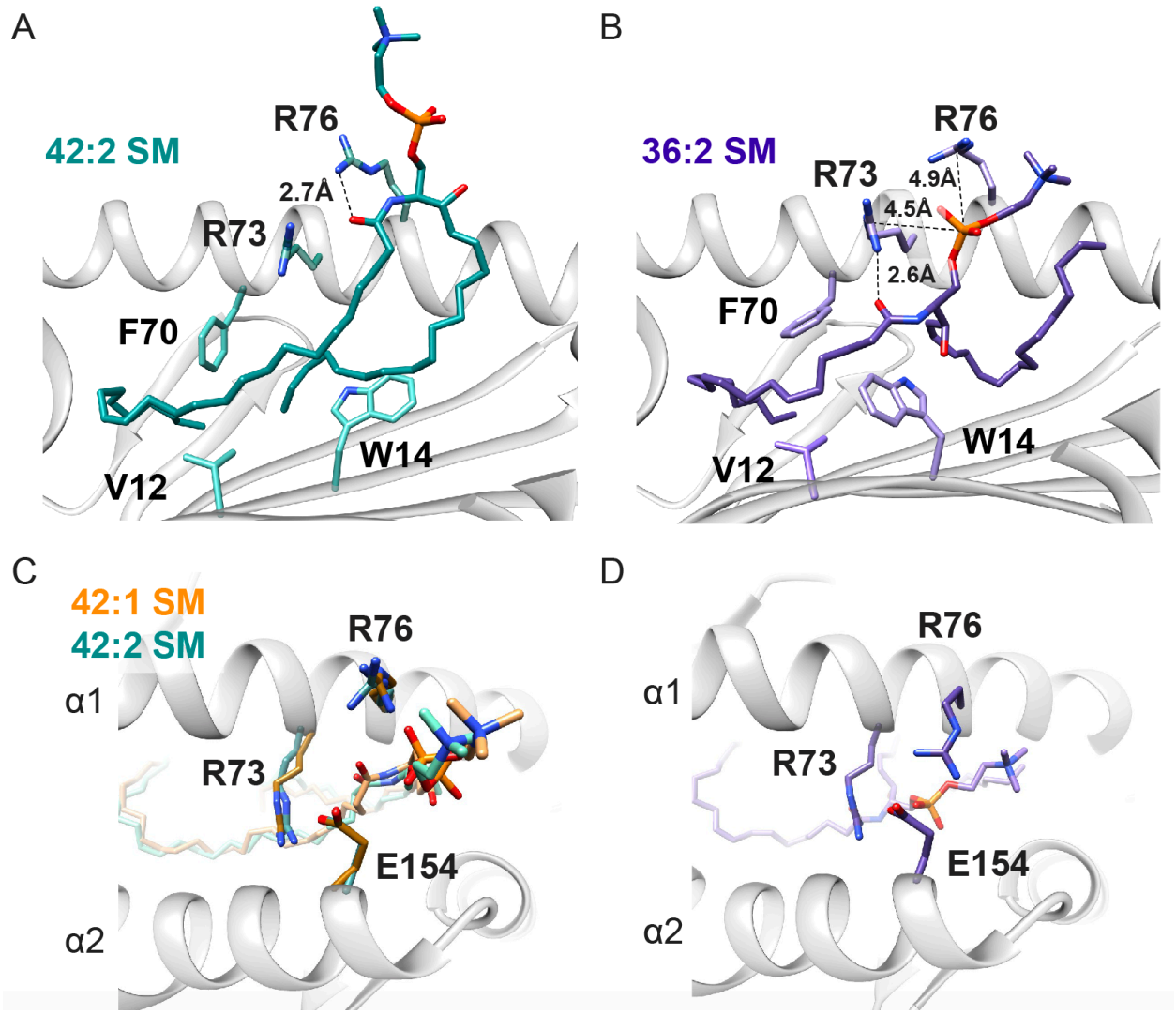

Figure 3.5. Remodeling of CD1a cleft and surface with short and long chain sphingomyelin. **A)** Side view of CD1a-42:2 SM with acyl and sphingosine chains in parallel in the A' pocket. F70 and V12 residues indicate A' pole. **B**) In contrast to 42:2 SM, 36:2 SM is seated 7 angstroms lower within the CD1a cleft and its two chains are oriented antiparallel, with the acyl chain in the A' pocket and the sphingosine chain in the F' pocket**.** Angstrom distances shown indicate an ionic platform stabilizing R76 and R73. **C)** Superimposed, top-down views of the CD1a roof for CD1a-42:1 SM and CD1a-42:2 SM complexes, indicating positioning of CD1a F' portal adjacent residues R73, R76, E154, in which R76 orients vertically**. D)** CD1a-36:2 SM complex oriented as in C, in which the R76 maintains a lateral, "closed" orientation**.**

Analysis of skin lipids

Although CD1a is expressed in the thymus and on dermal dendritic cells throughout the body, the brightest CD1a staining is seen on epidermal Langerhans cells in human skin [59]. Further, recent studies from our group and others have identified large numbers of CD1a autoreactive T cells in healthy (**Chapter 2**) and diseased human skin [68, 69, 70, 71, 123]. After identification of antagonist sphingomyelins in the HEK293 cell line (**Fig. 3.1**), next we examined whether this antagonist also exists in other cell types or tissues with a physiological role in CD1a autoreactivity, focusing on the ratio of C42 SM blockers relative to 36:1 and 34:1 SM permissive lipids that allow CD1a-TCR contact. I obtained and we analyzed model antigen presenting cells (APCs) that have been used to discover and quantify CD1a autoreactivity (reference HEK293 cells, C1R cells, K562 cells, THP-1 cells), as well as skin T cells, *in vitro* derived Langerhans-like cells (LCs), epidermal lipids from three donors, and total lipid extracts from full-thickness biopsies from breast, scalp, or abdominal skin (**Fig. 3.6**).

Sphingomyelin profiles from three genetically modified HEK293 sub-lines cultured separately were similar to one another (**Fig. 3.6**) and to prior results (**Fig 3.1**): 34:1 SM was the dominant sphingomyelin species with lower but readily detectable C42 SM forms present in all cases (**Fig. 3.6A**). Indeed among all samples tested, 34:1 SM, which showed weak or absent binding to CD1a (**Figs. 3.1, 3.2, 3.4A**), was the dominant SM species in cells and tissues. 42:2 SM and other long chain unsaturated SMs were detected in every case, but their relative abundance among SM profiles differed among cells and tissues. Notably, in full thickness skin samples and *in vitro* derived Langerhans cells, C42 SMs, including 42:1 SM, 42:2 SM and 42:3 SM were enriched as compared to most other non-skin sources (**Fig. 3.6B**). For example, in full-thickness skin samples from differing sites, C42 SMs made up an increased fraction of the SM pool, with similar levels of 42:2 and 42:1 SM, each with areas >25% of the level of 34:1 SM.

However, among all three epidermal lipid samples, C42 SMs were comparatively low as a fraction of SMs. This is consistent with the initial discovery that epidermal lipids and lipids and oils from sebum are enriched for species that favor CD1a-autoreactive T cell response [56]. Among non-professional antigen presenting cells (**Fig. 3.6C**), we found that C1R and THP-1 cells had highly similar sphingomyelin length profiles compared to the reference HEK293 lines, but that the two K562 cell lines had relatively increased fractions of C42 SMs. Overall, the long chain length variants of SM, which were potent blockers for CD1a-TCR interactions were found in physiologically relevant cells and tissues with some evidence for overexpression of blockers in full thickness skin and LC samples.

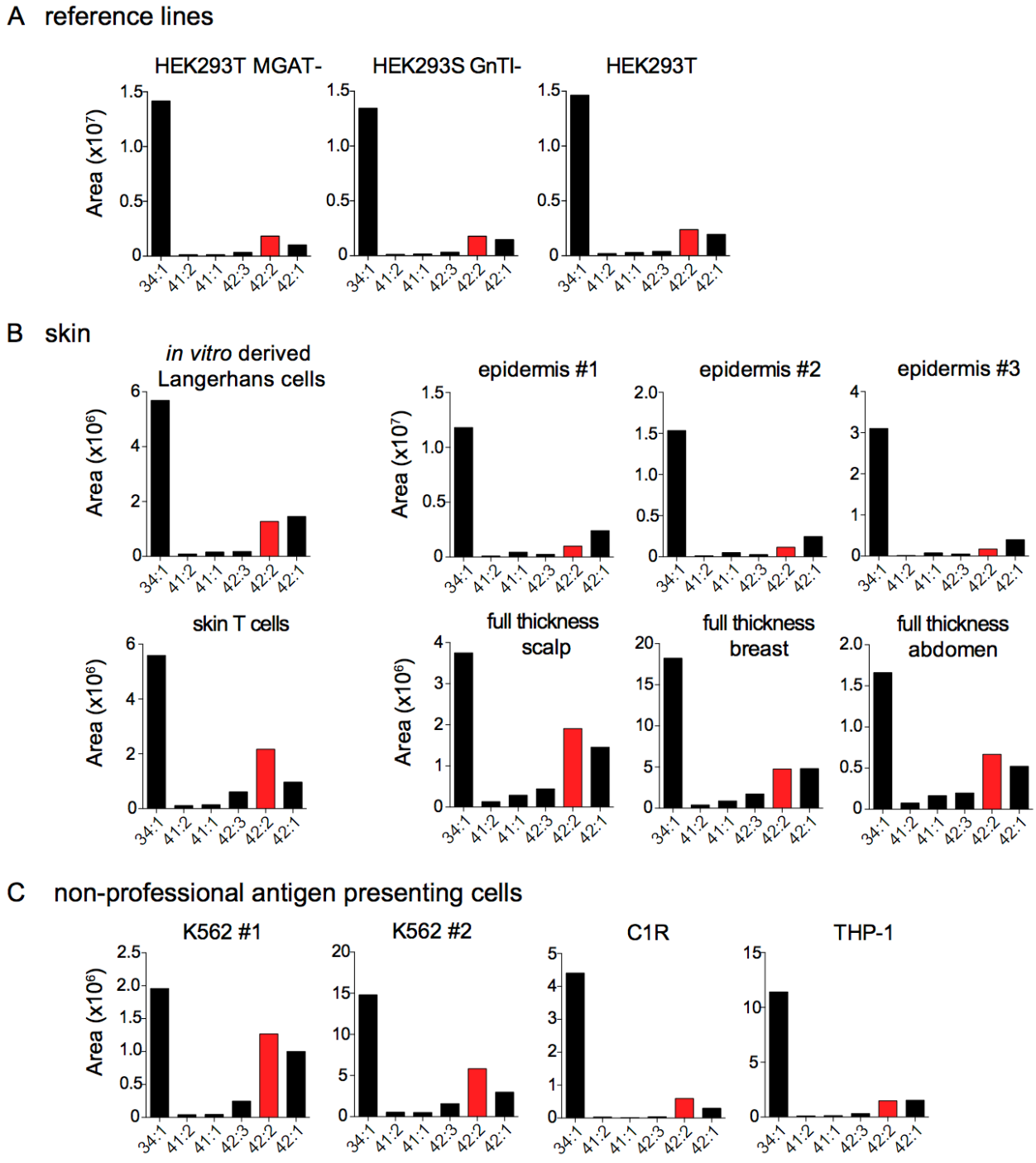

Figure 3.6. **Combined lipid length distribution of sphingomyelins in cells and skin.** Integrated areas under the curve (count-seconds) for 6 molecular species of sphingomyelins identified by m/z and retention time in the total lipid extracts from **A**) three HEK293 reference cell lines, **B**) skin cells, cell lines, and tissues, and **C**) non-professional antigen presenting cell lines.

Polyclonal skin-derived CD1a autoreactive T cells

Having seen blockade of tetramer staining in three T cell lines (**Fig. 3.3**), we next asked if long chain SM has broad effects on polyclonal CD1a autoreactive T cells measured *ex vivo*. This experiment was possible based on the implementation of a three-dimensional culture system, whereby viable T cells passively crawl into collagen matrices in the presence of T cell growth factors over 21-28 days (**Chapter 2**) [61, 62]. The inclusion of IL-2 and IL-15 represents a cytokine- but not antigen-driven expansion and allows us to estimate CD1a autoreactive T cells in human skin in the *ex vivo* state. We focused on $CD4^+$ T cells because previous studies have suggested higher rates of CD1a autoreactivity in this population [56]. Among polyclonal skin T cells from 8 unrelated donors, the 36:2 SM treated CD1a tetramers stained a greater frequency of $CD3^+CD4^+CD8\alpha^{neg}$ cells than did CD1a-endo tetramers in 6 of 7 patients (p=0.0313). Based on prior measurements of more than 90 lipids in CD1a-endo complexes (**Chapter 2**) and our detection of both short and long chain SMs in CD1a-endo (**Fig. 3.3**), this finding was consistent with displacement of a mixture of agonist and antagonist lipids with a short chain SM that allows CD1a-TCR contact.

Furthermore, similar to results seen in individual lines (**Fig. 3.3E**) 42:1 SM treated CD1a tetramers provided partial blockade of CD1a staining of T cells, staining a ~10-fold lower percentage of cells than CD1a-36:2 SM tetramers (p=0.0156) (**Fig. 3.7**) More strikingly, CD1a-42:2 SM tetramer staining was ~80-fold lower in frequency than CD1a-36:2 staining (p=0.0156), such that frequencies of CD1-42:2 SM staining were not significantly different from CD1b-endo tetramer staining (p=0.0781). This result implicates 42:2 SM as a potentially global inhibitor of CD1a interactions with diverse CD1a autoreactive TCRs.

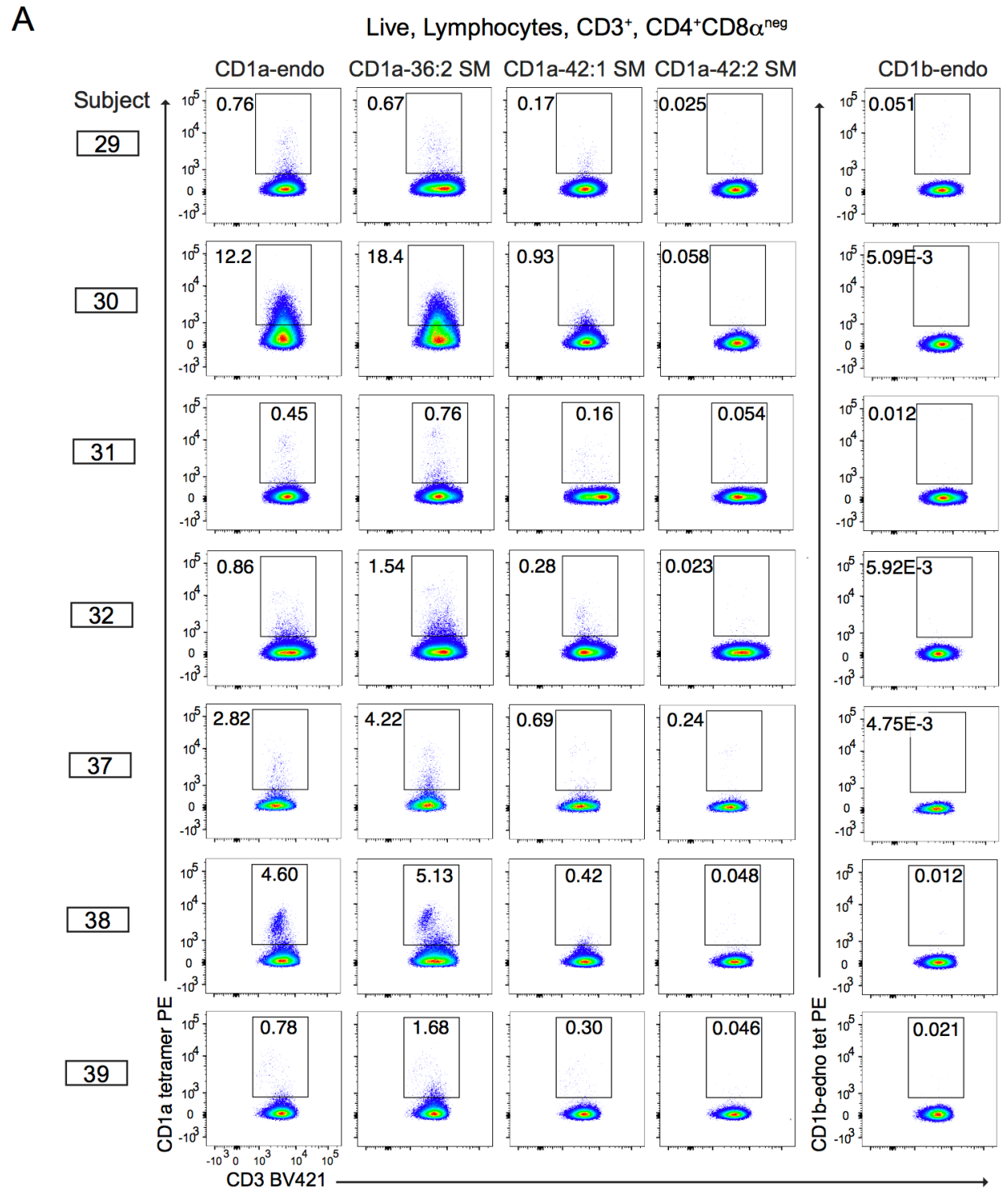

Figure 3.7. **CD1a-SM tetramer staining of polyclonal skin T cells. A)** Polyclonal T cells from 7 donors were stained with CD1a or CD1b tetramers carrying the indicated lipids, pre-gated Live, $CD3^{+}CD4^{+}CD8\alpha^{neg}$

B

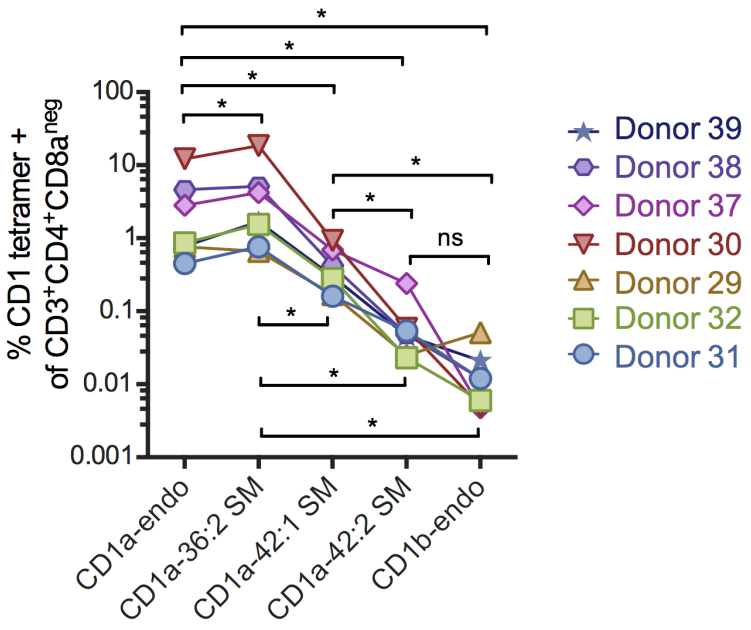

Figure 3.7 (continued). **CD1a-SM tetramer staining of polyclonal skin T cells.**
B) Summary data of polyclonal T cells from 7 donors stained with CD1 tetramers carrying the indicated lipids, pre-gated as in A. * = $p<0.05$, Wilcoxon matched-pairs signed rank test.

DISCUSSION

Overall, these results provide evidence that cellular CD1a proteins favor selective capture of shingolipids over membrane phospholipids, including specialized capture of an unsaturated long chain 42:2 sphingomyelin. This natural ligand is produced in multiple cell types and tissues, including human skin. This naturally occurring CD1a ligand is a potent blocker of CD1a interactions with various autoreactive cells and TCRs, measured at the clonal and polyclonal level and among genetically unrelated donors. The mechanism of blockade involves disruption of the CD1a surface by interactions with the the sphingmyelin head group, which is positioned within the CD1a cleft by a unique mechanism whereby the two anchoring lipids lie in a parallel orientation within the A' pocket.

Our ongoing work considers how not only lipid length, but also the number and position of unsaturations impact blockade of CD1a-TCR binding. We found that an acyl chain Δ15 *cis* unsaturation present in 42:2 SM but absent from 42:1 SM conferred an approximately 10-fold increase in blockade, observed as CD1a-SM tetramer staining intensity on T cell lines. Both 42:1 and 42:2 SM have nearly identical seating within CD1a, where the two chains fully occupy the cleft and establish an extensive network of hydrophobic interactions with apolar residues of CD1a. The unsaturation is likely sequestered deeply within the cleft, so its strong effect in augmenting tetramer blockade does not likely involve impeding TCR contact with CD1a. If the inferred location of the *cis* unsaturation that is known in synthetic lipids matches the location for natural 42:2 SM, it would localize at the distal end of the A' pocket, where the alkyl chain bends sharply to curve and encircle the A’ pole formed by F70 and V12. Thus, the *cis* unsaturation might provide a covalent constraint to stabilize the bent lipid conformation in the bent A' pocket and could thereby increase loading or retention of 42:2 SM in CD1a.

Our data add to existing data on lipid antigen-specific, CD1-restricted recognition by T cells provide proof of principle TCRs can indirectly sense sequestered lipid chain factors like length, unsaturation, or other modifications when bound in CD1 proteins. In the iNKT-CD1d system, shorter acyl chain length α-galactosyl ceramide from *Bacteroides fragilis* (<C19 versus C24:1 in KRN7000) was found to preferentially bind CD1d, but cannot be recognized by iNKT cells. One hypothesis is that the shorter chain length α-galactosyl ceramide sits deeper in the CD1d groove, precluding sufficient TCR recognition of the hexose headgroup for activation[124]. Indeed, a similar mechanism was previously shown in iNKT cells, in which glycopshinoglipid chain length and unsaturation promote optimal CD1d occupancy, TCR-CD1d synapse formation, and iNKT TCR signaling[125]. In the CD1a system, the CD1a-restricted, antigen-specific recognition of the mycobacterial lipopeptide didehydroxymycobactin (DDM) is most potent with DDM containing a C20:1 acyl chain. C20:0, C18:1, and C18:0 DDM were substantially less potent in activation assays, [126]. In all three of these examples, longer acyl chains and one or more *cis* unsaturations promoted optimal antigen-specific recognition of a moiety protruding from the CD1 cleft.

A recent study in the iNKT-CD1d system reported that sphingomyelin, a non-antigenic lipid for iNKT cells, serves a regulatory role in mice and humans by competing with antigenic α-galactosyl ceramide for CD1d binding. However, this effect on CD1d responses was shown to be largely independent of SM lipid chain length, while dependent on overall SM levels regulated by acid sphingomyelinase, which cleaves the headgroup from lipid tails [127]. Here, we demonstrate that long chain 42:2 SM is enriched for occupancy in CD1a proteins and serves the role of an endogenous blocker of CD1a-TCR contacts because the longer lipid chain forces the lipid head group up and out of the CD1a pocket, actively disrupting the putative contact face

shared by polyclonal skin T cells. Taken together, these results point to the feasibility of exploring molecular design for optimal binders and disruptors of CD1a-autoreactive T cells.

METHODS

Human Subjects

Discarded skin from cosmetic surgeries was used for T cell assays and analysis of skin lipids and was obtained through the Human Skin Disease Resource Center at Harvard Medical School and Brigham and Women's Hospital under approved protocols by Partners Institutional Review Board. Human peripheral blood mononuclear cells (PBMC) were obtained from leukoreduction collars provided by Brigham and Women's Hospital.

Recovery of skin T cells by three-dimensional culture

Skin T cells were recovered after culture for 21-28 days on three-dimensional cell foam growth matrices (Cytomatrix, Australia) seeded with collagen I (Thermo Fisher, 354236). Skin T cell culture media [IMDM, 10-20% FCS, L-glutamine, penicillin-streptomycin, 2-mercaptoethanol] was supplemented with IL-2 (BWH or Peprotech) and recombinant human IL-15 (10ng/mL, Peptrotech #200-15) as described [62].

Loading CD1a monomers with defined lipids for tetramers

Lipids stored in chloroform and methanol were transferred to new borosilicate glass tubes, dried under nitrogen gas and reconstituted to 400μM in TBS pH 8.0, 0.5% CHAPS buffer by sonication in a 37°C water bath for ~ 1 hr. Lipid-buffer sonicates or a buffer-only control were transferred to 1.5ml eppendorf tubes on a 37°C heat block to which CD1a monomer was added to a final concentration of 0.2 μg/μl and incubated for 2hrs at 37°C, then overnight at room temperature. Loaded monomers were stored at 4°C and used for tetramer assembly

Lipid sources

The following synthetic lipids were obtained from Avanti: C34:1 phosphatidylcholine (PC, 850475), C18:1 Lyso-PC (845875), C18:1 Lyso-SM (860600), C24:1 SM (860582), C34:1 SM (860584), C36:2 SM (860587), C42:1 SM (860592), and C42:2 SM (860593).

CD1a recombinant expression and purification for crystallization studies

CD1a was expressed in HEK293S GnTI- (lacking N-acetylglucosaminyltransferase I activity) cells and purified as previously described [46]. Following endoglycosidase H (New England BioLabs) and thrombin (Sigma) treatment, the purified CD1a was loaded with sphingomyelin C42:2, C42:1 and C36:2 (Avanti) dissolved in Tris-buffered saline (TBS)/ 0.5% CHAPS (3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate hydrate, Sigma). CD1a was incubated overnight at pH 5.5 with a 40-fold molar excess of each lipid. The detergent and unbound sphingomyelin were removed by ion exchange chromatography (MonoQ 10/100 GL, glass, GE Healthcare). Because the initially obtained crystals of CD1a-36:2 SM did not diffract, probably due to sample heterogeneity, we modified the loading procedure. First CD1a was incubated overnight at low pH with 0.5% tyloxapol (Sigma) to remove endogenous lipids. The detergent was removed by gel filtration chromatography (Superdex 200, GE Healthcare). In a second step, the tyloxapol-treated CD1a was loaded overnight with 40-fold excess of SM 36:2 dissolved in TBS/0.5% CHAPS and the mixture was separated by ion exchange chromatography (MonoQ 10/100 GL, GE Healthcare).

Isoelectric Focusing (IEF)

Lipid loading was assessed by incubating 10μg of CD1a with a 30X molar excess of the corresponding lipid solubilized in Tris Buffer saline and 0.5% CHAPS for 16h at room temperature. Next, differently charged species of CD1a were separated by isoelectric focusing (IEF). Briefly, 1μg of each protein was loaded on a PhastGel IEF 5-8 (GE Life Sciences) that was then run on the PhastSystem electrophoresis unit (Pharmacia) with default settings. The gel was then stained with Coomassie Brilliant Blue and copper sulphate for 10 minutes.

HPLC-MS analysis

Wild-type or mutant CD1a protein (80 μg) was extracted in triplicate in 15-ml glass tubes using chloroform, methanol and water [111]. The organic phase was recovered and dried under nitrogen gas. Eluent residue was redissolved in chloroform-methanol, normalized to 20 μM based on input protein and stored at –20°C. We injected 20μl for Q-ToF HPLC-MS positive ion mode analysis using an Agilent 6520 Accurate-Mass Q-TOF and 1200 series HPLC system with a normal phase Inertsil Diol column (150 mm × 2 mm, GL Sciences), running at 0.15 ml/min as described [39]. For lipid extraction from cells, CD1a-transduced HEK293T cells (with *MGAT1* and *TAP2* genes knocked out) were obtained from the NIH Tetramer Core Facility and matched to the CD1a monomer batches used for tetramer staining and sorting. Cells were grown up in DMEM 20% fetal calf serum according to tetramer core facility culture conditions. Culture medium was removed and cells were washed twice with PBS. Cell pellets were extracted successively with chloroform-methanol (at a ratio of 1:2 for 2h and 2:1 for 1h), successive extractions were combined and dried under nitrogen gas, and lipids were

weighted and stored at 1mg/mL in chloroform-methanol at –20 °C. For cellular lipidomic analysis, HEK293T *MGAT- TAP2-* cells were extracted 3 times for triplicate HPLC-MS runs.

Crystallization, structure determination and refinement of CD1a-SM complexes

Seeds obtained from previous binary CD1a-antigen crystals [44] were used to grow crystals of the CD1a-36:2 SM, CD1a-42:1 SM, and CD1a-42:2 SM binary complexes in 20-25% polyethylene glycol 1500 / 10% MMT buffer pH 5-6. The crystals were flash-frozen and data were collected at the MX2 beamline (Australian Synchrotron) to 2.1 Å, 2.4 Å, and 2.0 Å resolutions respectively. All the data were processed with the program *XDS* [128] and were scaled with *SCALA* from the *CCP4* programs suite [129]. Upon successful phasing by molecular replacement using the program *PHASER* [130] and the CD1a-farnesol structure as the search model, the electron densities of the ligands were clear in respective unbiased electron density maps. Initial rigid body refinement was performed using *phenix.refine* [131], and iterative model improvement was performed using *COOT* [132] and *phenix.refine*. The quality of the structure was confirmed at the Research Collaboratory for Structural Bioinformatics Protein Data Bank Data Validation and Deposition Services website. All presentations of molecular graphics were created with UCSF-Chimera [133]

This page is intentionally left blank

CHAPTER 4: Conclusions and Future Directions

I. Overview

Current models of MHC-peptide [48, 49] and CD1-lipid response [52, 53] emphasize TCR co-recognition of antigens and antigen presenting molecules. For MHC proteins, the peptide antigen broadly spans the lateral dimension of the display platform, so peptide-free regions are small, and more than 100 solved ternary structures show that TCR footprints directly contact peptide antigen in all or nearly all cases [134, 135]. Similarly, the CD1 literature is dominated by TCRs that bind both CD1 and lipid [52, 53], and most responses studied to date require an added antigen like α-galactosyl ceramide or bacterial lipids [136]. However, all CD1 proteins have a roof-like structure, which we propose can function as a landing pad for TCRs that do not necessarily contact the carried lipid. In this case, it is CD1a itself that functions as a cognate antigen for the TCR.

Typically, in the MHC system, T cells scan among cell surface antigen complexes until encounter with a rare cognate antigen causes activation through high affinity binding of specific antigen complexes to the T cell receptor (TCR). In contrast, several lines of evidence point to a different mode of activation of CD1a autoreactive T cells, whereby the usual interaction of TCR with most cellular CD1a proteins leads to activation as the default response. For example, recently published studies [69, 70, 71] and data in Chapters 2 and 3 both indicate that non-professional antigen-presenting cells that lack known costimulation pathways, including K562 cells and C1R cells, can activate polyclonal CD1a autoreactive T cells with broad specificity when transduced to express CD1a. Indeed, even recombinant CD1a proteins bound to a plate can activate some CD1a autoreactive T cell clones.[56] Thus, opposite to the 'off until on' nature of TCR-peptide-MHC system, our studies raise the general possibility that encounter of CD1a autoreactive T cells with CD1a frequently leads to activation.

Specifically, our data now firmly support both the existence of a CD1a-centric mechanism of recognition in which lipids play only a minor role (**Chapter 2**) or instead block TCR contact with CD1a (**Chapter 3**). The capture of large T cell populations among *ex vivo* skin T cells by CD1a-endo tetramers, along with demonstration that the cells functionally recognize CD1a in the absence of added antigen, rules in the existence of a large skin resident population defined by responses to CD1a itself. Mutational studies directly implicated the center of the A' roof as the target of responses. Further, TCR binding can occur to tetramers treated with diverse neutral lipids, phospholipids, sphingolipids or sulfolipids, strongly supporting two key aspects of this new mechanism. While individual lipid families are differentially compatible with tetramer binding, the pool of endogenous lipids normally captured by CD1a proteins in human cellular expression systems is sufficient for TCR binding without ligand exchange, ruling out certain models in which some kind of rare self antigen is necessary for activation. Rather, the default response of autoreactive TCR contact is activation based on CD1a-centric binding. This 'on until off' type mechanism is opposite to the usual mechanisms of rare and specific activating antigens seen for combinations of peptide-MHC and previously studied CD1-lipid combinations.

Combined, our studies in Chapters 2 and 3 provide evidence for the the existence of a large T cell population, present in nearly all individuals tested, that responds directly to a common protein that is highly expressed on Langerhans cells and other dendritic cells. Also, the CD1a-endo tetramer has now been optimized as a feasible approach to quantitative analysis and selection of CD1a autoreactive T cells themselves or as a screening tool for CD1a ligand effect on binding.

Our data raise basic questions of what mechanisms normally restrain potentially pathologic autoreactivity in the absence of disease, and how such autoreactive T cells might survive thymic negative selection in the first place. Here we have identified an inhibitor of CD1a-

TCR interactions that is natural, broadly expressed in cells, selectively CD1a-associated, and appears to globally block CD1a interactions with skin T cells at the polyclonal level. Looking forward, certain unanswered questions regarding the action of inhibitory sphingomyelins will likely be adressed. In particular, we want to know if CD1a-42:2 SM complexes are sufficiently frequent in thymus or skin to downregulate CD1a autoreactive T cell responses *in vivo*. Our data show a greater than 40-fold increase in the concentration of 42:2 SM among lipids eluted from CD1a protein monomers compared to its concentration in the total lipid pool. Also, the inhibitor is present in key cell types and tissues in which CD1a functions, and there is some evidence for increased basal abundance in full-thickness skin. Yet, we do not understand if such complexes are sufficiently frequent in thymus or skin APCs to downregulate activation *in vivo*. Thus, investigation of 42:2 SM levels in skin or CD1a-expressing thymocytes is a key goal.

I. Roles of the CD1a system in health and disease

Co-recognition models emphasize antigen specificity and a regulated 'off until on' mode of T cell response in which TCRs scan many cellular antigen complexes before contacting the rare cognate antigen. The extreme lipid polyspecificity in CD1-centric models predict that T cells do not require a rare, defined autoantigen. Instead T cells might be directly activated by any antigen presenting cell expressing the correct CD1 protein, where activation may be the default outcome of CD1-TCR interactions. Thus, many CD1 autoreactive T cells are 'on until off,' an observation that raises the questions of if and how such responses might be regulated in the periphery.

CD1-TCR contact as a rare and regulated event

MHC I is expressed on nearly all cells, and MHC II is expressed mainly on specialized APCs and rarely on some epithelia. Even more so than MHC II, CD1 proteins are comparatively rare and show regulated expression in the periphery. The constitutive expression of CD1a, CD1b, or CD1c, is mostly limited to individual APC types[74]. CD1a is expressed mainly on epidermal Langerhans cells[57, 58, 59]. CD1c is found on marginal zone B cells and the major population of classical dendritic cells[137], and CD1b, while rarely expressed on unactivated cells in the periphery, may be found on macrophages[74]. Generally however, CD1a, CD1b, and CD1c expression is induced in inflammatory states [74]. As a natural brake on this process, serum lipids have been shown to inhibit expression of CD1a, CD1b, and CD1c on monocytes [138]. Once expressed, alternate decoy receptors may still block CD1-TCR contacts, as Ig-like transcript 4 has been shown to bind CD1c and CD1d, and it acts to dampen CD1-dependent T cell activation [139]. These observations suggest that circulating T cells would come into contact with CD1 proteins much less often than MHC I and II proteins. The inducible expression of CD1 by Toll-like receptor ligands and cytokines could plausibly be a mechanism to restrict autoreactivity to inflammatory states[140, 141]. Building on clear evidence that human αβ T cells nearly colocalize with $CD1a^+$ Langerhans cells in the skin[142, 143, 144], studies described below now document increased CD1a expression in lesional skin, rule in a specific contribution of CD1a to inflammation, and point to cell populations that are candidates for CD1a-autreactive recognition in tissue and circulation, in health and disease. My studies demonstrating large numbers of CD1a autoreactive T cells in the skin and the constitutive expression of CD1a in skin show the limits of the idea that T cells rarely contact CD1a. They invite consideration of active negative regulation of response, or responses that are not classically proinflammatory.

IL-22 and Th22 cells

Skin-homing memory Th22 cells from blood ($CD45RO^+$ $CD4^+$ $CCR4^+$ $CCR6^+$ $CCR10^+$ $CXCR3^{neg}$) are enriched for CD1a-autoreactivity[50]. The Th22 or Tc22 phenotype may be a default activation pathway for CD1a-autoreactive skin T cells, though whether this programming occurs in skin, in skin-draining lymph nodes, or as a consequence of antigen-reactivity or local cytokine environment remains unknown. IL-22, a member of the IL-10 cytokine family, has been reported to have both pro- and anti-inflammatory properties based on the tissue site or disease state, with roles in homeostasis and diease. At homeostasis, CD1a-autoreactive T cells of the Th22 phenotype produce factors that signal to non-immune cells and have roles in epithelial cell activation. IL-22 stimulates wound healing responses upon barrier disruption, and stimulates the production of antimicrobial peptides by keratinocytes[145]. Low level production of IL-22 could be involved in communication with non-immune cells in promoting epithelial activation, tissue repair, and protection via signaling through the heterodimeric IL-22 receptor, comprised of IL-22R1 and IL-10R2 subunits [65, 66, 88, 146].

In our studies, we test IFN-γ and IL-22 production in T cell lines and clones in response to CD1a, finding that lines often produced both cytokines. This may occur physiologically, but IFN-γ release may also be a consequence of repeated rounds of T cell expansion in *ex vivo* systems, where in our experience lines will often produce IFN-y. IL-22 expression is thought to be more stable than other cytokine responses over multiple expansions and to represent the outcome of *in vivo* programming that persists *ex vivo*[88]. Our observation that polyclonal skin T cells *ex vivo* produce IL-22 in a CD1a-dependent manner rules in IL-22 as a physiologically relevant response across individuals without skin disease.

There are multiple levels of regulation for IL-22 that might further serve as a check on pathologic production or signaling, including special organization of the skin itself. One level of

regulation is driven in part by TGF-β. In the skin, a gradient of TGF-β expression increasing from low levels in the dermis to high concentrations in epidermis[147] may counterregulate the expression of IL-22 by $CD4^+$ T cells[148] and localization of $CD1a^+$ Langerhans cells and skin dendiritic cells. TGF-β signaling induces expression of the transcriptional regulator c-maf in $CD4^+$ T cells, which inhibits IL-22 transcription[148]. Thus, TGF-β is expressed at the same site as CD1a would be encountered, so it now becomes a candidate inhibitor of baseline IL-22 responses. In skin, keratinocytes and dermal fibroblasts are the major IL-22 responsive cells via expression of IL-22 receptor, and receptor expression can be modulated by TNF-α and IFN-γ, increasing activity in inflammatory microenvironments, as well as the proliferative capacity of keratinocytes. Thus heightened levels of IL-22, lower levels of the soluble decoy IL-22 binding protein (IL-22BP), higher expression levels of IL-22 receptor, or all of the above, may contribute to skin disease[149]. That CD1a-lipid complexes are a common target of Th22 cells points to another node in which the activity of these cells might be regulated.

Allergic inflammation and PLA2

Work by Graham Ogg and others indicate that CD1a-autoreactive T cells, via overproduction of IL-22, IL-4, GM-CSF, and IL-13 also participate in allergic and inflammatory skin disease in humans [44, 68, 69, 70, 71]. Recent studies on human CD1a-dependent T cell responses and CD1a-dependent antigen discovery have highlighted this system in insect venom allergy[68, 71] , house dust mite sensitivity[70], and psoriasis[44, 69]. A common mechanism for CD1a-dependent response in these pathologies is an increased activity of phospholipases like PLA2, which generates an abundance of lyso-phospholipids and free fatty acids, which are both known antigens for CD1a-reactive T cells. Recently, in an *in vivo* skin challenge model of atopic

dermatitis, type 2 innate lymphoid cells (ILC2) were shown to upregulate CD1a and produced the phospholipase PLA2G4A[150]. CD1a was upregulated on ILC2 cells in response to TSLP, at levels physiologic to atopic dermatitis lesional skin[150]. Atopic dermatitis skin and psoriatic lesional skin have altered ceramide lipid compositions[151, 152], not only underscoring lipid barrier defects in these diseases, but also suggesting alterations in epidermal lipid generation or processing generally. If and to what degree these factors influence the lipids bound in CD1a and thereby CD1a-autoreactive T cell responses is an active area of research.

It is important to note however, that studies ruling in Th22 and Th2 cytokines were carried out in a hypothesis driven manner; that is by T cell activation assays requiring *a priori* selection of readouts, in ELISPOT, ELISA, intracellular cytokine staining, or cytokine gene expression. Other phenotypes or surface markers may emerge upon testing in high throughput single-cell RNA sequencing of CD1a tetramer sorted cells. Recently, numerous studies combinding scRNA-seq and TCR sequencing approaches have been carried out in healthy skin and lesional versus non-lesional human skin, and while the CD1a dependency in these studies remains unknown, they point to relevant T cell phenotypes consistent with those attributed to CD1a-autoreactive cells, and provide a larger picture of these potentially CD1a-associated phenotypes. For example, in blood and skin from healthy individuals, Klicznik *et al* described that the $CD4^+CLA^+CD103^+$ T cell population in blood may have come from the otherwise long-resident $CD4^+CD69^+CD103^+$ T_{RM} T cell populations in the dermis and epidermis because T cell clones in these populations overlapped by ~30%[64]. Both populations produced IL-22, IL-13, and other factors associated with antimicrobial defense and barrier repair[64]. Because these cells phenotypically match key aspects of CD1a autoreactive T cells, a new hypothesis is that the cells are actually the same, and CD1a is their natural target.

III. Selection, clonality and repertoire of CD1a autoreactive T cells

Due to the nearly monomorphic nature of CD1a expression in humans, it is tempting to speculate that CD1a autoreactive T cells may include TCRs that are shared among individuals. A striking feature of donor unrestricted T cells is the expression of invariant TCRs across individuals[81, 153, 154]. Some well-known known examples are NKT cells that recognize CD1d and mucosa-associated invariant T (MAIT) cells that recognize MR1. Using CD1b tetramers, two more invariant T cell types recognizing a mycobacterial lipid in CD1b were identified from genetically unrelated tuberculosis patients: LDN5-like T cells and Germline Encoded Mycolyl-lipid reactive (GEM) T cells. GEM T cells are characterized by a nearly invariant TCR-α chain (TRAV1-2, TRAJ9). LDN5-like T cells, though recognizing the same CD1b-lipid complex, are characterized by a TRBV4-1 β chain [53, 155]. Thus, there can be several interdonor conserved TCR stereotypes for a single antigenic complex.

Furthermore in the MHC system, it is a known phenomenon and even expected that a single TCR may recognize multiple pMHC complexes and be cross-reactive to multiple self or foreign peptide antigens[156]. In the thymus, CD1a is expressed on TCR-rearranged double positive committed thymocyte progenitors, but CD1a is downregulated on mature T cells in the periphery[157]. An alternative hypothesis is that CD1a autoreactive TCRs may dually recognize CD1a and MHC at different affinities, such that CD1a expression on thymocytes serves a minimal tonic signaling role supporting positive selection[158] . The binding topology on CD1a antigen complex versus an MHC class II peptide antigen complex would not predict such cross-reactivity. Perhaps more likely is that if the CD1a autoreactive TCR repertoire is diverse, then we may see repertoire groups based on HLA background. The development and thymic selection of CD1a-autoreactive T cells remains largely a black box. To this end, hCD1a-transgenic mouse studies but particularly human fetal tissue, fetal thymic organ cultures, and

post-natal thymus studies will be informative in elucidating the role of CD1a in T cell selection. For example, while CD1a-autoreactive γδT cells in humans have not yet been reported, a recent study found that the human fetal thymus generates γδ T cells with invariant human CMV-reactive TCRs and programmed effector functions[159]. This was found to be due to the suppression of the enzyme responsible for N-nucleotide additions (TdT) by Lin28b[159]. I and others have previously proposed the possible connection between Vδ1-containing TCRs and CD1a recognition[106, 153], based on published termary crystal structures in which TCR binding to CD1d-lipid complexes is mediated exclusively or predominantly by Vδ1-CD1 contacts [160, 161, 162], and based on reports of Vδ1 bias among γδ TCRs targeting CD1c-lipid[163, 164] and CD1d-lipid complexes[165].

IV. CD1a tetramer sorting for single-cell phenotyping and TCR discovery

A broad TCR sequencing and T cell phenotyping campaign for CD1a-autoreactive T cells is now possible with CD1a-endo tetramer reagents, but this has not been done in skin. For CD1a-autoreactive T cells, our understanding of the TCR repertoire is limited to a small number of sequenced T cell clones, and no patterns of restricted TCR V and J gene usage emerge from this short list above what could occur randomly. We see TRAV8-6 and TRBV2 genes usage in more than one CD1a-autoreactive line or clone from different individuals, but this is among only a tiny sampling of the vast diversity of possible TCR sequence estimates present in any individual, ~10^8 per imporved sequencing techniques[166]. My studies, which have enabled tracking and capture of CD1a autoreactive T cells now provide a feasible approach to transcriptioinally profiling the skin-resident T cell repertoire for TCRs and effector functions. Prior methods of CD1a autroreactive T cell capture necessarily involved longer term T cell

culture, so methods reported here offer new opportunities to rule in CD1a-autoreactive T cells as a *bona fide* human T cell type with physiologic roles in skin health and disease. Because CD1a is highly expressed in the epidermis and is non-polymorphic across the human population, CD1a-autoreactive T cells may be present in and respond to the same antigenic targets in any individual.

The 10X Genomics Chromium system partitions single cells with a uniquely barcoded bead for reagent delivery in oil. Critically, each bead delivers a cell-specific barcode and a 5' template switch oligo with a unique molecular identifier (UMI) barcode to control for sequencing error and enable transcript quantification[167]. After the RT reaction, TCR transcripts are enriched in 2 rounds of amplification with a universal primer to the 5' adaptor and successive nested primers to TCRα and TCRβ gene constant regions. There are several advantages to this system. First, there is no need for V- or J-specific primers that can introduce bias in a multiplex PCR. Second, fragmentation and sequencing are optimized for assembly of the full-length VDJ sequence (5' UTR to constant region) from short paired-end reads. This outcome leaves the option to also assay other transcripts to correlate TCR sequences with T cell phenotype for a single cell. There are now several published studies that have employed this system for combined paired TCR sequencing and single-cell transcriptomics[168, 169]. Recently, oligo-tagged antibodies that would permit detection and quantification of surface marker expression at the single cell level also became available with 5' chemistry, known as TotalSeq C. Thus in the Chromium system, transcript levels, surface protein expression levels, and paired TCR sequences can be assayed simultaneously.

The ideal minimum input cell number for the 10X Genomics Chromium system is ~1700 sored cells per well of the partitioning chip, highly concentrated at 700-1,200 cells/μl. I initially expected CD1a tetramer binding skin T cells to be rare, so considered methods to increase the

starting T cell pool and increase detection rate of low affinity TCRs. First, While the 3D culture method of skin T cell recovery yields a high number and high purity of T cells for tetramer selection, potential concerns over this approach relate to introduction of phenotypic bias by cell handling. These confounders are of less concern for TCR gene sequence discovery than for other transcripts. For example, including IL-15 and low-dose IL-2 during in the 3D culture results in 5-10-fold higher numbers of recovered T cells than without cytokines, but this may introduce bias to the T cell transcriptome, especially activation or proliferation related transcripts. This issue can be avoided simply by eliminating IL-2 and IL-15 from this step at the expense of total cells recovered[62]. Second, optimal tetramer staining requires the addition of unlabeled anti-CD3 during the staining step. In combination with the tetramer, this aggregates the TCR complex at the surface, increasing avidity for tetramer binding, increasing rates of detection, and increasing the fluorescence intensity range from which to sort CD1a tetramer binding events. Anti-CD3 may upregulate some activation-induced transcripts in the time course of staining and cell sorting. A final potential limitation inherent to both 10X Genomics 5' and CEL-Seq2 single-cell RNAseq approaches is that for lymphocytes, reads map to only ~1000 genes per cell. Even with 10X Genomics' 5' amplification step of TCR α and β genes, in a fraction of cells only the TCR β chain sequence is determined, because the TCR α locus is expressed at a lower level [168].

Since both psoriasis and atopic dermatitis are thought to be mediated by the antigen-driven expansion and activity of T cells in lesional skin, one might expect to identify T cell clones that are highly expanded in lesional skin versus non-lesional skin by comparing ranked lists of the most abundant TCR sequences. Studies profiling the TCR repertoire in psoriatic and atopic dermatitis lesional skin using Adaptive Biotechnologies system have revealed that the T cell infiltrate is highly polyclonal [170] and overlaps significantly with the TCR repertoire in matched nonlesional skin[171]. This finding is not inconsistent with the antigen-driven expansion of T cell

clones, but rather suggests that the number of those cells may be low in the context of a dense T cell infiltrate [170]. Graham Ogg has shown that CD1a-autoreactive T cells are present in lesional skin, and my methods allow direct sorting of CD1a autoreactive T cells from diseased skin. This approach could be used in the future to determine if TCR stereotypes may be present in disease or shared across donors.

IV. Therapeutic application

Certain CD1 ligands block the formation of stable TCR-CD1 complexes, pointing to a natural mechanism to limit T cell activation [46, 56, 107]. CD1a tetramer assays with polyclonal skin T cells and T cell clones in this thesis also revealed that the carried lipids may serve a blocking function. Blocking lipids generally have head groups that are larger than those on antigenic ligands, such as fatty acids, monoacylglycerols, squalene. These outcomes suggest that stearic hindrance of the TCR is the blocking mechanism. However, one structure of CD1a-sphingomyelin demonstrated that the ligand can interact with a triad of residues in the A' roof, which alters the TCR binding surface on CD1a [46]. Similarly, cholesterol esters and other ligands alter the overall shape of CD1 proteins in ways that affect TCR binding[45, 125]. We hypothesize that a sudden abundance of activating 'headless' CD1 ligands in the local milieu could increase productive CD1-autoreactive TCR contacts. For example, recent studies have demonstrated suggest that increased phospholipase activity in viral infection[172] and in autoimmune and allergic skin disease generates an excess of small ligands that activate CD1-autoreactive T cells [68, 69, 70]. The evidence in **Chapter 3** for selective binding of the inhibitory 42:2 SM makes it tempting to speculate that this lipid or a similar optimized CD1a binder might be able to outcompete CD1a ligand contact allergens like uroshiol [44] , farnesol, and balsam of peru[43] in the context of a

delayed type hypersensitivity reaction, or outcompete CD1a-activating products of PLA2 (**Fig 4.1**).

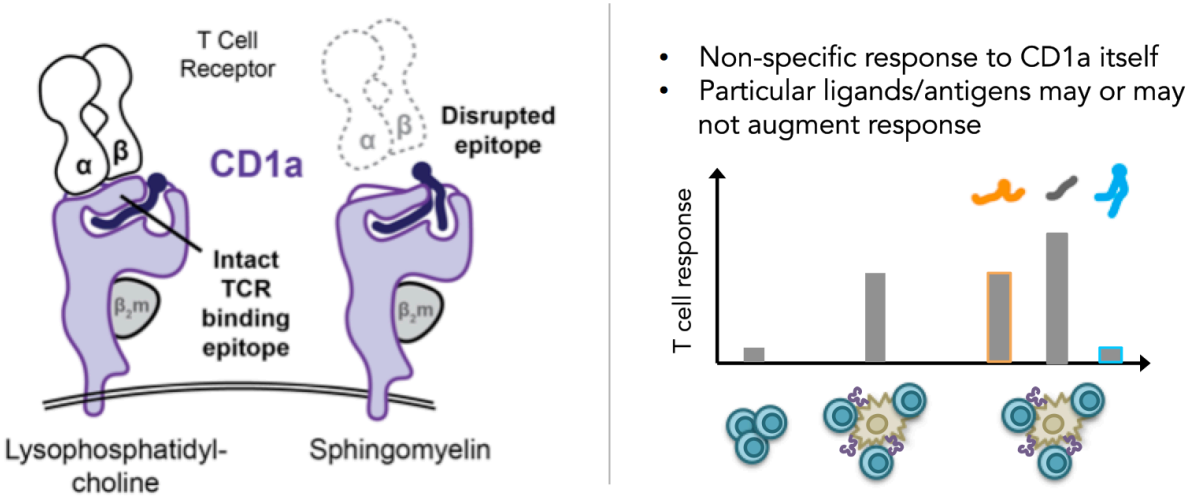

Fig 4.1. Models for manipulating CD1a autoreactivity. Left: Cartoon based on the BK6 TCR -LPC-CD1a ternary crystal structure PDB: 4X6E, in which the TCR binds to the closed roof of CD1a, left-shifted, while failing to contact the lipid itself. Right: Simulated data for a model of autoreactivity in which cytokine production by CD1a-autoreactive T cells (blue) co-cultured with a $CD1a^{+}$ antigen presenting cell line (beige) can be tuned up or tuned down by the addition of particular lipids. Increase in magnitude of activation would signify that a particular lipiid (orange, gray, blue), sufficiently exchanged with cell endogenous lipiids for CD1a occupancy, favors a CD1a conformation that is bound more strongly by TCRs than the average CD1a-TCR avidity achieved with the cellular lipid mileu.

The consequences of CD1 autoreactive T cells being turned on or off hinges on the nature of their responses. While CD1-autoreactive cells have thus far been shown to produce pro-allergic or pro-inflammatory cytokines, the predominance in healthy individuals of IL-22 production by CD1a-autoreactive T cells [50] and the killing of $CD1c^+$ leukemic cells by CD1c-autoreactive T cells [173] point to homeostatic roles in wound healing and cellular damage surveillance. Alternatively, there is evidence that CD1 autoreactivity can serve an instructive or regulatory role via cross talk with CD1-expressing DCs or B cells [174, 175, 176]. Low TCR specificity for the carried ligand, is consistent with immunosurveillance, puts more emphasis on CD1 expression and the overall balance of activating versus blocking ligands displayed on the surface of the antigen presenting cell.

We now want to know whether synthetically optimized molecular species of sphingomyelins or sulfatides could be designed based on altered chain length and unsaturation to show optimal CD1a binding and block T cell responses to CD1a expressing cells. Design of inhibitory peptides in the MHC system has been challenging, but the monomorphic nature of CD1 proteins predicts that optimized ligands for CD1a would act similarly among all persons, providing a candidate route to therapeutic applications of this basic work, including skin topicals formulated with CD1a-binding sphingomyelins or otherwise optimized ligands.

REFERENCES

1. McMichael, A.J. *et al.* A human thymocyte antigen defined by a hybrid myeloma monoclonal antibody. *Eur J Immunol* **9**, 205-210 (1979).

2. Reinherz, E.L., Kung, P.C., Goldstein, G., Levey, R.H. & Schlossman, S.F. Discrete stages of human intrathymic differentiation: analysis of normal thymocytes and leukemic lymphoblasts of T-cell lineage. *Proc Natl Acad Sci U S A* **77**, 1588-1592 (1980).

3. Bernard, A. & Boumsell, L. The clusters of differentiation (CD) defined by the First International Workshop on Human Leucocyte Differentiation Antigens. *Hum Immunol* **11**, 1-10 (1984).

4. Porcelli, S. *et al.* Recognition of cluster of differentiation 1 antigens by human CD4-CD8-cytolytic T lymphocytes. *Nature* **341**, 447-450 (1989).

5. Beckman, E.M. *et al.* Recognition of a lipid antigen by CD1-restricted alpha beta+ T cells. *Nature* **372**, 691-694 (1994).

6. Blumberg, R.S., Gerdes, D., Chott, A., Porcelli, S.A. & Balk, S.P. Structure and function of the CD1 family of MHC-like cell surface proteins. *Immunol Rev* **147**, 5-29 (1995).

7. Porcelli, S.A. The CD1 family: a third lineage of antigen-presenting molecules. *Adv Immunol* **59**, 1-98 (1995).

8. Zajonc, D.M. *et al.* Molecular mechanism of lipopeptide presentation by CD1a. *Immunity* **22**, 209-219 (2005).

9. Zajonc, D.M., Elsliger, M.A., Teyton, L. & Wilson, I.A. Crystal structure of CD1a in complex with a sulfatide self antigen at a resolution of 2.15 A. *Nature immunology* **4**, 808-815 (2003).

10. Ackerman, A.L. & Cresswell, P. Cellular mechanisms governing cross-presentation of exogenous antigens. *Nat Immunol* **5**, 678-684 (2004).

11. McMichael, A.J. *et al.* A human thymocyte antigen defined by a hybrid myeloma monoclonal antibody. *Eur.J.Immunol.* **9**, 205-210 (1979).

12. Park, J.J. *et al.* Lipid-protein interactions: biosynthetic assembly of CD1 with lipids in the endoplasmic reticulum is evolutionarily conserved. *Proc Natl Acad Sci U S A* **101**, 1022-1026 (2004).

13. Briken, V., Jackman, R.M., Dasgupta, S., Hoening, S. & Porcelli, S.A. Intracellular trafficking pathway of newly synthesized CD1b molecules. *EMBO Journal* **21**, 825-834 (2002).

14. Manolova, V. *et al.* Functional CD1a is stabilized by exogenous lipids. *Eur.J.Immunol.* (2006).

15. Briken, V., Jackman, R.M., Watts, G.F., Rogers, R.A. & Porcelli, S.A. Human CD1b and CD1c isoforms survey different intracellular compartments for the presentation of microbial lipid antigens. *J.Exp.Med.* **192**, 281-288 (2000).

16. Sugita, M., van Der, W., Rogers, R.A., Peters, P.J. & Brenner, M.B. CD1c molecules broadly survey the endocytic system. *Proc.Natl.Acad.Sci.* **97**, 8445-8450 (2000).

17. Jackman, R.M. *et al.* The tyrosine-containing cytoplasmic tail of CD1b is essential for its efficient presentation of bacterial lipid antigens. *Immunity* **8**, 341-351 (1998).

18. Moody, D.B. *et al.* Lipid length controls antigen entry into endosomal and nonendosomal pathways for CD1b presentation. *Nature Immunol.* **3**, 435-442 (2002).

19. Spada, F.M., Koezuka, Y. & Porcelli, S.A. CD1d-restricted recognition of synthetic glycolipid antigens by human natural killer T cells. *J.Exp.Med.* **188**, 1529-1534 (1998).

20. Sugita, M. *et al.* Failure of trafficking and antigen presentation by CD1 in AP-3-deficient cells. *Immunity.* **16**, 697-706 (2002).

21. Sugita, M. *et al.* Separate pathways for antigen presentation by CD1 molecules. *Immunity.* **11**, 743-752 (1999).

22. Hunger, R.E. *et al.* Langerhans cells utilize CD1a and langerin to efficiently present nonpeptide antigens to T cells. *J.Clin.Invest* **113**, 701-708 (2004).

23. Barral, D.C. *et al.* CD1a and MHC class I follow a similar endocytic recycling pathway. *Traffic* **9**, 1446-1457 (2008).

24. Lawton, A.P. *et al.* The mouse CD1d cytoplasmic tail mediates CD1d trafficking and antigen presentation by adaptor protein 3-dependent and -independent mechanisms. *J Immunol* **174**, 3179-3186 (2005).

25. Relloso, M. *et al.* pH-dependent interdomain tethers of CD1b regulate its antigen capture. *Immunity* **28**, 774-786 (2008).

26. Kang, S.J. & Cresswell, P. Saposins facilitate CD1d-restricted presentation of an exogenous lipid antigen to T cells. *Nat.Immunol.* **5**, 175-181 (2004).

27. Winau, F. *et al.* Saposin C is required for lipid presentation by human CD1b. *Nat.Immunol.* **5**, 169-174 (2004).

28. Zhou, D. *et al.* Editing of CD1d-bound lipid antigens by endosomal lipid transfer proteins. *Science* **303**, 523-527 (2004).

29. Hunt, D.F. *et al.* Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **255**, 1261-1263 (1992).

30. Rammensee, H.G., Falk, K. & Rotzschke, O. Peptides naturally presented by MHC class I molecules. *Annu Rev Immunol* **11**, 213-244 (1993).

31. Rotzschke, O. *et al.* Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells. *Nature* **348**, 252-254 (1990).

32. Bjorkman, P.J. *et al.* Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329**, 506-512 (1987).

33. Bjorkman, P.J. *et al.* The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* **329**, 512-518 (1987).

34. Brown, J.H. *et al.* Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**, 33-39 (1993).

35. Rammensee, H., Bachmann, J., Emmerich, N.P., Bachor, O.A. & Stevanovic, S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **50**, 213-219 (1999).

36. Gadola, S.D. *et al.* Structure of human CD1b with bound ligands at 2.3 A, a maze for alkyl chains. *Nature Immunol.* **3**, 721-726 (2002).

37. Koch, M. *et al.* The crystal structure of human CD1d with and without alpha-galactosylceramide. *Nat.Immunol.* **6**, 819-826 (2005).

38. Scharf, L. *et al.* The 2.5 a structure of CD1c in complex with a mycobacterial lipid reveals an open groove ideally suited for diverse antigen presentation. *Immunity* **33**, 853-862 (2010).

39. Huang, S. *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* **108**, 19335-19340 (2011).

40. Cox, D. *et al.* Determination of cellular lipids bound to human CD1d molecules. *PloS one* **4**, e5325 (2009).

41. Saper, M.A., Bjorkman, P.J. & Wiley, D.C. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. *J Mol Biol* **219**, 277-319 (1991).

42. Wang, J. *et al.* Lipid binding orientation within CD1d affects recognition of Borrelia burgorferi antigens by NKT cells. *Proc Natl Acad Sci U S A* **107**, 1535-1540 (2010).

43. Nicolai, S. *et al.* Human T cell response to CD1a and contact dermatitis allergens in botanical extracts and commercial skin care products. *Sci Immunol* **5** (2020).

44. Kim, J.H. *et al.* CD1a on Langerhans cells controls inflammatory skin disease. *Nature immunology* **17**, 1159-1166 (2016).

45. Mansour, S. *et al.* Cholesteryl esters stabilize human CD1c conformations for recognition by self-reactive T cells. *Proc Natl Acad Sci U S A* **113**, E1266-1275 (2016).

46. Birkinshaw, R.W. *et al.* alphabeta T cell antigen receptor recognition of CD1a presenting self lipid ligands. *Nature immunology* **16**, 258-266 (2015).

47. Le Nours, J., Shahine, A. & Gras, S. Molecular features of lipid-based antigen presentation by group 1 CD1 molecules. *Semin Cell Dev Biol* (2017).

48. Garboczi, D.N. *et al.* Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* **384**, 134-141 (1996).

49. Garcia, K.C. *et al.* An alphabeta T cell receptor structure at 2.5 A and its orientation in the TCR-MHC complex. *Science* **274**, 209-219 (1996).

50. de Jong, A. *et al.* CD1a-autoreactive T cells are a normal component of the human alphabeta T cell repertoire. *Nature immunology* **11**, 1102-1109 (2010).

51. de Lalla, C. *et al.* High-frequency and adaptive-like dynamics of human CD1 self-reactive T cells. *Eur J Immunol* **41**, 602-610 (2011).

52. Borg, N.A. *et al.* CD1d-lipid-antigen recognition by the semi-invariant NKT T-cell receptor. *Nature* **448**, 44-49 (2007).

53. Gras, S. *et al.* T cell receptor recognition of CD1b presenting a mycobacterial glycolipid. *Nat Commun* **7**, 13257 (2016).

54. Shahine, A. *et al.* A T-cell receptor escape channel allows broad T-cell response to CD1b and membrane phospholipids. *Nat Commun* **10**, 56 (2019).

55. Van Rhijn, I. *et al.* Human autoreactive T cells recognize CD1b and phospholipids. *Proc Natl Acad Sci U S A* **113**, 380-385 (2016).

56. de Jong, A. *et al.* CD1a-autoreactive T cells recognize natural skin oils that function as headless antigens. *Nature immunology* **15**, 177-185 (2014).

57. Fithian, E. *et al.* Reactivity of Langerhans cells with hybridoma antibody. *Proc Natl Acad Sci U S A* **78**, 2541-2544 (1981).

58. Furue, M. *et al.* Epitope mapping of CD1a, CD1b, and CD1c antigens in human skin: differential localization on Langerhans cells, keratinocytes, and basement membrane zone. *J Invest Dermatol* **99**, 23S-26S (1992).

59. Furue, M. *et al.* Epitopes for CD1a, CD1b, and CD1c antigens are differentially mapped on Langerhans cells, dermal dendritic cells, keratinocytes, and basement membrane zone in human skin. *J Am Acad Dermatol* **27**, 419-426 (1992).

60. Mizumoto, N. & Takashima, A. CD1a and langerin: acting as more than Langerhans cell markers. *J Clin Invest* **113**, 658-660 (2004).

61. Clark, R.A. *et al.* The vast majority of CLA+ T cells are resident in normal skin. *J Immunol* **176**, 4431-4439 (2006).

62. Clark, R.A. *et al.* A novel method for the isolation of skin resident T cells from normal and diseased human skin. *J Invest Dermatol* **126**, 1059-1070 (2006).

63. Watanabe, R. *et al.* Human skin is protected by four functionally and phenotypically discrete populations of resident and recirculating memory T cells. *Sci Transl Med* **7**, 279ra239 (2015).

64. Klicznik, M.M. *et al.* Human CD4(+)CD103(+) cutaneous resident memory T cells are found in the circulation of healthy individuals. *Sci Immunol* **4** (2019).

65. Duhen, T., Geiger, R., Jarrossay, D., Lanzavecchia, A. & Sallusto, F. Production of interleukin 22 but not interleukin 17 by a subset of human skin-homing memory T cells. *Nature immunology* **10**, 857-863 (2009).

66. Trifari, S., Kaplan, C.D., Tran, E.H., Crellin, N.K. & Spits, H. Identification of a human helper T cell population that has abundant production of interleukin 22 and is distinct from T(H)-17, T(H)1 and T(H)2 cells. *Nature immunology* **10**, 864-871 (2009).

67. Kronenberg, M. & Havran, W.L. Immunology: oiling the wheels of autoimmunity. *Nature* **506**, 42-43 (2014).

68. Bourgeois, E.A. *et al.* Bee venom processes human skin lipids for presentation by CD1a. *J Exp Med* **212**, 149-163 (2015).

69. Cheung, K.L. *et al.* Psoriatic T cells recognize neolipid antigens generated by mast cell phospholipase delivered by exosomes and presented by CD1a. *J Exp Med* **213**, 2399-2412 (2016).

70. Jarrett, R. *et al.* Filaggrin inhibits generation of CD1a neolipid antigens by house dust mite-derived phospholipase. *Sci Transl Med* **8**, 325ra318 (2016).

71. Subramaniam, S. *et al.* Elevated and cross-responsive CD1a-reactive T cells in bee and wasp venom allergic individuals. *Eur J Immunol* **46**, 242-252 (2016).

72. Chen, Y.L. *et al.* Re-evaluation of human BDCA-2+ DC during acute sterile skin inflammation. *J Exp Med* **217** (2020).

73. Kasmar, A.G. *et al.* Cutting Edge: CD1a tetramers and dextramers identify human lipopeptide-specific T cells ex vivo. *J Immunol* **191**, 4499-4503 (2013).

74. Dougan, S.K., Kaser, A. & Blumberg, R.S. CD1 expression on antigen-presenting cells. *Current topics in microbiology and immunology* **314**, 113-141 (2007).

75. Hunger, R.E. *et al.* Langerhans cells utilize CD1a and langerin to efficiently present nonpeptide antigens to T cells. *J Clin Invest* **113**, 701-708 (2004).

76. Pena-Cruz, V., Ito, S., Dascher, C.C., Brenner, M.B. & Sugita, M. Epidermal Langerhans cells efficiently mediate CD1a-dependent presentation of microbial lipid antigens to T cells. *J Invest Dermatol* **121**, 517-521 (2003).

77. Rosat, J.P. *et al.* CD1-restricted microbial lipid antigen-specific recognition found in the CD8+ alpha beta T cell pool. *J Immunol* **162**, 366-371 (1999).

78. van de Rijn, M. *et al.* Human cutaneous dendritic cells express two glycoproteins T6 and M241 which are biochemically identical to those found on cortical thymocytes. *Hum.Immunol.* **9**, 201-210 (1984).

79. Bagchi, S. *et al.* CD1b-autoreactive T cells contribute to hyperlipidemia-induced skin inflammation in mice. *J Clin Invest* **127**, 2339-2352 (2017).

80. Joosten, S.A. *et al.* Harnessing donor unrestricted T-cells for new vaccines against tuberculosis. *Vaccine* **37**, 3022-3030 (2019).

81. Van Rhijn, I. & Moody, D.B. Donor Unrestricted T Cells: A Shared Human T Cell Response. *J Immunol* **195**, 1927-1932 (2015).

82. Karadimitris, A. *et al.* Human CD1d-glycolipid tetramers generated by in vitro oxidative refolding chromatography. *Proc.Natl.Acad.Sci.U.S.A* **98**, 3294-3298 (2001).

83. Altman, J.D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94-96 (1996).

84. Matsuda, J.L. *et al.* Tracking the response of natural killer T cells to a glycolipid antigen using CD1d tetramers. *J.Exp.Med.* **192**, 741-754 (2000).

85. Benlagha, K., Weiss, A., Beavis, A., Teyton, L. & Bendelac, A. In vivo identification of glycolipid antigen-specific T cells using fluorescent CD1d tetramers. *J.Exp.Med.* **191**, 1895-1903 (2000).

86. Campbell, J.J. *et al.* CCR7 expression and memory T cell diversity in humans. *J Immunol* **166**, 877-884 (2001).

87. Schaerli, P. *et al.* A skin-selective homing mechanism for human immune surveillance T cells. *J Exp Med* **199**, 1265-1275 (2004).

88. Eyerich, S. *et al.* Th22 cells represent a distinct human T cell subset involved in epidermal immunity and remodeling. *J Clin Invest* **119**, 3573-3585 (2009).

89. Wong, M.T. *et al.* A High-Dimensional Atlas of Human T Cell Diversity Reveals Tissue-Specific Trafficking and Cytokine Signatures. *Immunity* **45**, 442-456 (2016).

90. Klein, E. *et al.* Properties of the K562 cell line, derived from a patient with chronic myeloid leukemia. *Int J Cancer* **18**, 421-431 (1976).

91. Moon, J.J. *et al.* Tracking epitope-specific T cells. *Nat Protoc* **4**, 565-581 (2009).

92. Dolton, G. *et al.* Optimized Peptide-MHC Multimer Protocols for Detection and Isolation of Autoimmune T-Cells. *Front Immunol* **9**, 1378 (2018).

93. Rius, C. *et al.* Peptide-MHC Class I Tetramers Can Fail To Detect Relevant Functional T Cell Clonotypes and Underestimate Antigen-Reactive T Cell Populations. *J Immunol* **200**, 2263-2279 (2018).

94. Van Rhijn, I. *et al.* CD1d-restricted T cell activation by nonlipidic small molecules. *Proc Natl Acad Sci U S A* **101**, 13578-13583 (2004).

95. Sanderson, J.P. *et al.* CD1d protein structure determines species-selective antigenicity of isoglobotrihexosylceramide (iGb3) to invariant NKT cells. *Eur J Immunol* **43**, 815-825 (2013).

96. Altman, J.D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94-96 (1996).

97. Kasmar, A.G. *et al.* CD1b tetramers bind alphabeta T cell receptors to identify a mycobacterial glycolipid-reactive T cell repertoire in humans. *J Exp Med* **208**, 1741-1747 (2011).

98. Wun, K.S. *et al.* T cell autoreactivity directed toward CD1c itself rather than toward carried self lipids. *Nature immunology* **19**, 397-406 (2018).

99. Li, D. *et al.* Ig-like transcript 4 inhibits lipid antigen presentation through direct CD1d interaction. *J Immunol* **182**, 1033-1040 (2009).

100. Layre, E. *et al.* A comparative lipidomics platform for chemotaxonomic analysis of Mycobacterium tuberculosis. *Chem Biol* **18**, 1537-1549 (2011).

101. Roy, S. *et al.* Molecular basis of mycobacterial lipid antigen presentation by CD1c and its recognition by alphabeta T cells. *Proc Natl Acad Sci U S A* **111**, E4648-4657 (2014).

102. Le Nours, J. *et al.* A class of gammadelta T cell receptors recognize the underside of the antigen-presenting molecule MR1. *Science* **366**, 1522-1527 (2019).

103. Cernadas, M. *et al.* Early recycling compartment trafficking of CD1a is essential for its intersection and presentation of lipid antigens. *J Immunol* **184**, 1235-1241 (2010).

104. Moody, D.B. & Cotton, R.N. Four pathways of CD1 antigen presentation to T cells. *Curr Opin Immunol* **46**, 127-133 (2017).

105. Fujita, H. *et al.* Human Langerhans cells induce distinct IL-22-producing CD4+ T cells lacking IL-17 production. *Proc Natl Acad Sci U S A* **106**, 21795-21800 (2009).

106. Cotton, R.N., Shahine, A., Rossjohn, J. & Moody, D.B. Lipids hide or step aside for CD1-autoreactive T cell receptors. *Curr Opin Immunol* **52**, 93-99 (2018).

107. Shahine, A. *et al.* A molecular basis of human T cell receptor autoreactivity toward self-phospholipids. *Sci Immunol* **2** (2017).

108. Geissmann, F. *et al.* Transforming growth factor beta1, in the presence of granulocyte/macrophage colony-stimulating factor and interleukin 4, induces differentiation of human peripheral blood monocytes into dendritic Langerhans cells. *J Exp Med* **187**, 961-966 (1998).

109. Cotton, R.N. *et al.* Brugia malayi infective larvae fail to activate Langerhans cells and dermal dendritic cells in human skin. *Parasite Immunol* **37**, 79-91 (2015).

110. Ly, D. *et al.* CD1c tetramers detect ex vivo T cell responses to processed phosphomycoketide antigens. *J Exp Med* **210**, 729-741 (2013).

111. Bligh, E.G. & Dyer, W.J. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* **37**, 911-917 (1959).

112. Madigan, C.A. *et al.* Lipidomic discovery of deoxysiderophores reveals a revised mycobactin biosynthesis pathway in Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **109**, 1257-1262 (2012).

113. Carette, X. *et al.* Multisystem Analysis of Mycobacterium tuberculosis Reveals Kinase-Dependent Remodeling of the Pathogen-Environment Interface. *mBio* **9** (2018).

114. Cresswell, P. Chemistry and functional role of the invariant chain. *Curr Opin Immunol* **4**, 87-92 (1992).

115. Roche, P.A. & Cresswell, P. Invariant chain association with HLA-DR molecules inhibits immunogenic peptide binding. *Nature* **345**, 615-618 (1990).

116. Garcia-Alles, L.F. *et al.* Structural reorganization of the antigen-binding groove of human CD1b for presentation of mycobacterial sulfoglycolipids. *Proc Natl Acad Sci U S A* **108**, 17755-17760 (2011).

117. Layre, E. *et al.* A comparative lipidomics platform for chemotaxonomic analysis of Mycobacterium tuberculosis. *Chemistry & biology* **18**, 1537-1549 (2011).

118. van Meer, G. Cellular lipidomics. *EMBO Journal* **24**, 3159-3165 (2005).

119. Fahy, E. *et al.* A comprehensive classification system for lipids. *J.Lipid Res.* **46**, 839-862 (2005).

120. Merrill, A.H., Jr. Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. *Chem Rev* **111**, 6387-6422 (2011).

121. Kawano, T. *et al.* CD1d-restricted and TCR-mediated activation of valpha14 NKT cells by glycosylceramides. *Science* **278**, 1626-1629 (1997).

122. de Jong, A. Activation of human T cells by CD1 and self-lipids. *Immunol Rev* **267**, 16-29 (2015).

123. Jarrett, R. & Ogg, G. Lipid-specific T cells and the skin. *Br J Dermatol* **175 Suppl 2**, 19-25 (2016).

124. An, D. *et al.* Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. *Cell* **156**, 123-133 (2014).

125. McCarthy, C. *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* **204**, 1131-1144 (2007).

126. Moody, D.B. *et al.* T cell activation by lipopeptide antigens. *Science* **303**, 527-531 (2004).

127. Melum, E. *et al.* Control of CD1d-restricted antigen presentation and inflammation by sphingomyelin. *Nature immunology* **20**, 1644-1655 (2019).

128. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132 (2010).

129. Winn, M.D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242 (2011).

130. McCoy, A.J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

131. Afonine, P.V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352-367 (2012).

132. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).

133. Pettersen, E.F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).

134. Wucherpfennig, K.W. & Strominger, J.L. Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. *Cell* **80**, 695-705 (1995).

135. Rossjohn, J. *et al.* T Cell Antigen Receptor Recognition of Antigen-Presenting Molecules. *Annu Rev Immunol* (2014).

136. Salio, M., Silk, J.D., Jones, E.Y. & Cerundolo, V. Biology of CD1- and MR1-restricted T cells. *Annu Rev Immunol* **32**, 323-366 (2014).

137. Roura-Mir, C. *et al.* CD1a and CD1c activate intrathyroidal T cells during Graves' disease and Hashimoto's thyroiditis. *J Immunol* **174**, 3773-3780 (2005).

138. Leslie, D.S. *et al.* Serum lipids regulate dendritic cell CD1 expression and function. *Immunology* **125**, 289-301 (2008).

139. Li, D. *et al.* A novel role of CD1c in regulating CD1d-mediated NKT cell recognition by competitive binding to Ig-like transcript 4. *Int Immunol* **24**, 729-737 (2012).

140. Moody, D.B. TLR gateways to CD1 function. *Nature immunology* **7**, 811-817 (2006).

141. Roura-Mir, C. *et al.* Mycobacterium tuberculosis regulates CD1 antigen presentation pathways through TLR-2. *J Immunol* **175**, 1758-1766 (2005).

142. Pigozzi, B., Bordignon, M., Belloni Fortina, A., Michelotto, G. & Alaibac, M. Expression of the CD1a molecule in B- and T-lymphoproliferative skin conditions. *Oncol Rep* **15**, 347-351 (2006).

143. Seneschal, J., Clark, R.A., Gehad, A., Baecher-Allan, C.M. & Kupper, T.S. Human epidermal Langerhans cells maintain immune homeostasis in skin by activating skin resident regulatory T cells. *Immunity* **36**, 873-884 (2012).

144. Taylor, R.S., Baadsgaard, O., Hammerberg, C. & Cooper, K.D. Hyperstimulatory CD1a+CD1b+CD36+ Langerhans cells are responsible for increased autologous T lymphocyte reactivity to lesional epidermal cells of patients with atopic dermatitis. *J Immunol* **147**, 3794-3802 (1991).

145. Wolk, K. *et al.* IL-22 regulates the expression of genes responsible for antimicrobial defense, cellular differentiation, and mobility in keratinocytes: a potential role in psoriasis. *Eur J Immunol* **36**, 1309-1323 (2006).

146. Ouyang, W. & O'Garra, A. IL-10 Family Cytokines IL-10 and IL-22: from Basic Science to Clinical Translation. *Immunity* **50**, 871-891 (2019).

147. Morikawa, M., Derynck, R. & Miyazono, K. TGF-beta and the TGF-beta Family: Context-Dependent Roles in Cell and Tissue Physiology. *Cold Spring Harb Perspect Biol* **8** (2016).

148. Rutz, S. *et al.* Transcription factor c-Maf mediates the TGF-beta-dependent suppression of IL-22 production in T(H)17 cells. *Nature immunology* **12**, 1238-1245 (2011).

149. Voglis, S. *et al.* Regulation of IL-22BP in psoriasis. *Sci Rep* **8**, 5085 (2018).

150. Hardman, C.S. *et al.* CD1a presentation of endogenous antigens by group 2 innate lymphoid cells. *Sci Immunol* **2** (2017).

151. Di Nardo, A., Wertz, P., Giannetti, A. & Seidenari, S. Ceramide and cholesterol composition of the skin of patients with atopic dermatitis. *Acta Derm Venereol* **78**, 27-30 (1998).

152. Motta, S. *et al.* Ceramide composition of the psoriatic scale. *Biochim Biophys Acta* **1182**, 147-151 (1993).

153. Godfrey, D.I., Uldrich, A.P., McCluskey, J., Rossjohn, J. & Moody, D.B. The burgeoning family of unconventional T cells. *Nature immunology* **16**, 1114-1123 (2015).

154. Van Rhijn, I., Godfrey, D.I., Rossjohn, J. & Moody, D.B. Lipid and small-molecule display by CD1 and MR1. *Nat Rev Immunol* **15**, 643-654 (2015).

155. Van Rhijn, I. *et al.* A conserved human T cell population targets mycobacterial antigens presented by CD1b. *Nature immunology* **14**, 706-713 (2013).

156. Nelson, R.W. *et al.* T cell receptor cross-reactivity between similar foreign and self peptides influences naive cell population size and autoimmunity. *Immunity* **42**, 95-107 (2015).

157. Sanchez, M.J., Muench, M.O., Roncarolo, M.G., Lanier, L.L. & Phillips, J.H. Identification of a common T/natural killer cell progenitor in human fetal thymus. *J Exp Med* **180**, 569-576 (1994).

158. Colonna, M. Skin function for human CD1a-reactive T cells. *Nature immunology* **11**, 1079-1080 (2010).

159. Tieppo, P. *et al.* The human fetal thymus generates invariant effector gammadelta T cells. *J Exp Med* **217** (2020).

160. Luoma, A.M. *et al.* Crystal structure of Vdelta1 T cell receptor in complex with CD1d-sulfatide shows MHC-like recognition of a self-lipid by human gammadelta T cells. *Immunity* **39**, 1032-1042 (2013).

161. Pellicci, D.G. *et al.* The molecular bases of delta/alphabeta T cell-mediated antigen recognition. *J Exp Med* **211**, 2599-2615 (2014).

162. Uldrich, A.P. *et al.* CD1d-lipid antigen recognition by the gammadelta TCR. *Nature immunology* **14**, 1137-1145 (2013).

163. Spada, F.M. *et al.* Self-recognition of CD1 by gamma/delta T cells: implications for innate immunity. *J Exp Med* **191**, 937-948 (2000).

164. Roy, S. *et al.* Molecular Analysis of Lipid-Reactive Vdelta1 gammadelta T Cells Identified by CD1c Tetramers. *J Immunol* **196**, 1933-1942 (2016).

165. Bai, L. *et al.* The majority of CD1d-sulfatide-specific T cells in human blood use a semiinvariant Vdelta1 TCR. *Eur J Immunol* **42**, 2505-2510 (2012).

166. Qi, Q. *et al.* Diversity and clonal selection in the human T-cell repertoire. *Proc Natl Acad Sci U S A* **111**, 13139-13144 (2014).

167. Zheng, G.X. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat Commun* **8**, 14049 (2017).

168. Patil, V.S. *et al.* Precursors of human CD4(+) cytotoxic T lymphocytes identified by single-cell transcriptome analysis. *Sci Immunol* **3** (2018).

169. Zemmour, D. *et al.* Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nature immunology* **19**, 291-301 (2018).

170. Kirsch, I.R. *et al.* TCR sequencing facilitates diagnosis and identifies mature T cells as the cell of origin in CTCL. *Sci Transl Med* **7**, 308ra158 (2015).

171. Harden, J.L., Hamm, D., Gulati, N., Lowes, M.A. & Krueger, J.G. Deep Sequencing of the T-cell Receptor Repertoire Demonstrates Polyclonal T-cell Infiltrates in Psoriasis. *F1000Res* **4**, 460 (2015).

172. Zeissig, S. *et al.* Hepatitis B virus-induced lipid alterations contribute to natural killer T cell-dependent protective immunity. *Nature medicine* **18**, 1060-1068 (2012).

173. Lepore, M. *et al.* Targeting leukemia by CD1c-restricted T cells specific for a novel lipid antigen. *Oncoimmunology* **4**, e970463 (2015).

174. Fox, L.M. *et al.* Expression of CD1c enhances human invariant NKT cell activation by alpha-GalCer. *Cancer Immun* **13**, 9 (2013).

175. Sieling, P.A. *et al.* Human double-negative T cells in systemic lupus erythematosus provide help for IgG and are restricted by CD1c. *J Immunol* **165**, 5338-5344 (2000).

176. Vincent, M.S. *et al.* CD1-dependent dendritic cell instruction. *Nat Immunol* **3**, 1163-1168 (2002).

PUBLICATIONS

Primary Author

Cotton RN*, Wegrecki M*, Cheng TY, Le Nours J, Orgill DP, Pomahac B, Talbot SG, Willis RA, Altman JD, de Jong A, Clark RA, Rossjohn J*, Moody DB*. *Negative regulation of human skin T cells by cellular lipid antagonists in CD1a.* (In preparation for 2nd submission)

Cotton RN, Cheng TY, Wegrecki M, Le Nours J, Orgill DP, Pomahac B, Talbot SG, Willis RA, Altman JD, de Jong A, Rossjohn J, Clark RA, Moody DB. *Human skin* αβ *T cells recognize CD1a independently of presented lipid antigens.* (In preparation)

Cotton RN, Shahine A, Rossjohn J, Moody DB. *Lipids hide or step aside for CD1-autoreactive T cell receptors.* Current Opinion in Immunology. **2018.** PMID: 29738961

Moody DB & **Cotton RN**. *Four Pathways of CD1 Antigen Presentation to T cells.* Current Opinion in Immunology. **2017.** PMID: 28756303

Co-Author

Nicolai S, Wegrecki M, Cheng TY, Bourgeois EA, **Cotton RN**, Mayfield JA, Monnot GC, Le Nours J,Van Rhijn I, Rossjohn J, Moody DB, de Jong A. *Human T cell response to CD1a and contact dermatitis allergens in botanical extracts and commercial skin care products.* Science Immunology. **2020**. DOI: 10.1126/sciimmunol.aax5430

Buter J, *et al. Mycobacterium tuberculosis releases an antacid that remodels phagosomes.* Nat Chem Biol. **2019.** PMID: 31427817

Cheng JM, *et al. Total Synthesis of Mycobacterium tuberculosis Dideoxymycobactin-838 and Stereoisomers: Diverse CD1a-Restricted T Cells Display a Common Hierarchy of Lipopeptide Recognition.* Chemistry. **2017.** PMID: 27925318