THE UNIVERSITY OF CHICAGO

STRUCTURAL AND CELLULAR PROPERTIES OF ARCHAIC AND NON-CLASSICAL HLA MOLECULES

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

COMMITTEE ON GENETICS, GENOMICS, AND SYSTEMS BIOLOGY

BY

PHILIPP ROSS

CHICAGO, ILLINOIS

MARCH 2023

To my family and friends.

And all those who have supported me.

And reminded me that the meaning of life is life itself.

I do not strive for riches because I have always been.

Thank you.

Table of Contents

List of Abbreviations

BLCL - B-lymphoblastoid cell line

cAMP - Cyclic adenosine monophosphate

ESF - Endometrial stromal fibroblast

EVT - Extravillous trophoblasts

Fab - Fragment antigen-binding region

hCG - Human chorionic gonadotropin

HLA - Human leukocyte antigen

IFN-γ - Interferon gamma

IHC - Immunohistochemistry

IL1β - Interleukin-1 beta

KIR - Killer immunoglobulin-like receptor

LC-MS/MS - Liquid chromatography followed by tandem mass spectrometry

LIF - Leukemia inhibitory factor

LIR - Leukocyte immunoglobulin-like receptor

MHC - Major histocompatibility complex

MPA - Medroxyprogesterone acetate

NHP - Non-human primate

NK - Natural killer

PLC - Peptide-loading complex

TCR - T cell receptor

List of Definitions

HLA nomenclature - Using HLA-B*73:01 as an example, HLA class I allele nomenclature convention is to separate the HLA prefix by a dash (-) followed by the gene locus (B), then use an asterisk (*) to separate the HLA gene from the allele group number (73), followed by a colon (:) to separate the allele group number from the specific HLA protein number (01)

Non-human primate HLA nomenclature - *HLA-F* refers to the human *HLA-F* gene. Genes found to be orthologous to *HLA-F* in NHPs are named such that the first two letters of the genus and species are combined followed by a dash and the gene locus letter. So the *HLA-F* ortholog in *Pan troglodytes* (chimpanzee) is referred to as *Patr-F*.

C1, C2, Bw4 epitopes - Shorthand names given to non-linear amino acid epitopes found in particular HLA molecules that denote them either as ligands for KIR2DL1 (C2), KIR2DL2 and KIR2DL3 (C1), or KIR3DL1 (Bw4)

Archaic humans - Neanderthal and Denisovan species

Peptidome - the theoretically complete set of peptides presented by an HLA allele

W6/32 - Pan class I monoclonal antibody clone (Can recognize complexed HLA-A, -B, -C, -E, -F, and -G)

3D11 - An HLA-F specific monoclonal antibody clone (epitope unknown)

Fecundability - The probability of conception within a single menstrual cycle

Tetramer - Common staining reagent used to overcome low-affinity interactions between ligand:receptor complexes, which are common in immunology

List of Figures

List of Tables

List of Contributions

This thesis would never have been completed without help from the many collaborators and contributors listed below.

Shared reagents

The Parham Lab at Stanford University generously shared BLCLs of 2 chimpanzees, 2 gorillas, 2 orangutans, and 2 gibbons, and BLCLs overexpressing HLA-B*46:01 and HLA-B*73:01. The Gilad Lab at the University of Chicago generously shared 17 rhesus macaque BLCLs. THESC, St-T1b, HTR8, and JAR cell lines were generously shared by the Lynch Lab, then at the University of Chicago, now at SUNY Buffalo. The JEG3 cell line was generously shared by the Ober Lab at the University of Chicago. The Huh7.5 cell line was generously shared by the Randall Lab at the University of Chicago. The Swan71 cell line was generously shared by Caroline E. Dunk at the Lunenfeld Tanenbaum Research Institute who originally received it from Dr. Gil Mor at Wayne State University. Primary first-trimester and term-derived trophoblast cell lines were generously shared by Dr. Tamara Tilburgs, Dr. Sarika Kshirsagar, and Dr. Jack Strominger, formerly of Harvard University. The β2M KO Jurkat cell line with and without engineered KIR3DS1:CD3ζ was generously shared by the Altfeld Lab Leibniz Institute of Virology (LIV) in Hamburg, Germany. The Parham Lab at Stanford University also generously shared cDNA for soluble HLA-B*73:01, HLA-B*46:01, and KIR2DL3 bacterial expression constructs. Dr. Hisashi Arase of Osaka University generously shared plasmid DNA encoding KIR-Fc fusion constructs which were used to clone KIR2DL2 expression constructs.

Data and figures

Data and figures were generated by Philipp Ross (PR) unless noted otherwise below.

Chapter 2

For figure 1, PR generated the figure himself while design contributions were provided by Dr. Hugo Hilton (HH) of the Parham Lab. Data for figure 1 was generated by Dr. Curtis McMurtrey (CM) of the Hildebrand Lab. Data was originally analyzed by HH and reanalyzed by PR.

For figure 2, HH contributed to the final figure through design of panels B and F. Data used was a combination of data generated by CM, HH, and from another paper cited within. PR processed and analyzed the data and made figures for the rest of the panels.

Figures 3, 4, and 5 were generated by PR and data was generated and analyzed by PR. KTN helped express KIR2DL2 and KIR2DL2 for SPR binding experiments.

Chapter 3

Figure 2 uses data generated by Sophie Krahnke (SK), Dr. Kristof Nolan (KTN) of the Adams Lab, and Tomasz Ślężak (TS) of the Kossiakoff lab at the University of Chicago. Figures were originally made by TS and adapted by PR. PR cloned the expression construct, SK and KTN expressed and purified the proteins, and TS generated Fabs using phage display.

Figure 3 is based on preliminary structural data generated by Jane Lodwick, KTN and TS. Panel A was generated by KTN and Dr. James Fuller. Sequence alignment of HLA alleles was done by PR. Binding experiments and protein expression and purification were done by PR, KTN, and Max Chu, a former rotation student in the Adams Lab.

Figures 4, 5, 6, and 7 were made by PR and use data generated by PR. Biotinylated Fabs were expressed and purified by KTN.

Figure 8 was made by PR and uses data generated by PR. Biotinylated Fabs were expressed and purified by KTN. Protocol was originally shared by Dr. Angelique Hölzemer of the Altfeld Lab and adapted by PR.

Figure 9 was made by PR and uses data generated by PR. Biotinylated Fabs were expressed and purified by KTN.

Figure 10 was made by PR and uses data generated by PR and Brock Adams, a former rotation student in the Adams Lab.

Figure 14 was made by PR and uses data generated by PR. LC-MS/MS data was generated and processed by Dr. Samuel Weng of the Proteomics Core at the University of Chicago.

All supplementary figures were made by PR and use data generated by PR. Biotinylated Fabs were expressed and purified by KTN.

Chapter 4

Figure 2 was made by PR with contribution by Dr. Tamara Tilburgs and uses data generated by PR.

Acknowledgements

None of this would have been possible without the support, encouragement, and understanding of various people that have served as pillars of support during various stages of my life. Firstly, I would like to thank my advisor, Erin Adams, and the members of my thesis committee, Carole Ober, Bana Jabri, Allan Drummond, and Vinny Lynch. The guidance and feedback provided by each of you throughout this process was vital to my success and critically important to my development as a scientist. I was incredibly lucky to have such experienced and capable individuals supervise my work for all these years and I hope I will continue to utilize your wisdom in my future endeavors.

I would also like to thank my former advisor, Manuel Llinás, from whom I first learned the thrill of basic scientific research. He took a risk when he hired me, but his support never wavered as I cultivated my interest in science and learned what it took to transform experimental results into scientific insight. He continues to support me in my pursuits to this day.

A huge thank you to all previous and past members of the Adams lab at UChicago that I met while in graduate school. All of you were always willing to help and listen and your daily interactions were incredibly important to me. This includes May Gu, Marta Borowska, Chris Boughter, Ryan Duncombe, John Leonard, Caitlin Castro, Charlotte James, Augusta Broughton, Idongesit Ekpo, Sofi Maltseva, Amrita Ramesh, Sean Ryan, Mayuri Viswanathan, Sophie Kranke, Andrew Kelley, Sam Dubensky, Daniel Yong, Tala Azzam, Andrew Peev. I also want to give special thanks to three members not mentioned above: Charlie Dulberger, Kristof Nolan,

and Sobhan Roy. The three of you provided me with incredible mentorship that I was able to successfully apply to my thesis work while also showering me with heartfelt acts of friendship.

I am particularly thankful for Sue Levison who did so much more for me than her job title suggests. I witnessed how she treated myself and students that started before and after me with respect, dignity, and guardianship. Sue was incredibly patient and willing to guide me through the small, but necessary tasks that every student has to wrestle with during their graduate work. She also showed unwavering support during some of the most difficult times for me personally for which I am eternally grateful.

To the friends I have made while in Chicago–from former and current members of the Adams lab to those unaffiliated with the University of Chicago–I thank you for the many experiences we have been able to share together. One friend in particular, my former roommate and good friend, Alex Advani, did so much to ease my transition into graduate school and life in Chicago that I would struggled had it not been for his wisdom and understanding. I consider myself extremely fortunate to have Alex in my life. And to Bill Richter, Katie Mika, Max Novak, Jane Lodwick, Bruce Tammen, Mirna Marinic, Mel Yamsek, Hendrik Glauninger, Chris Watkins, Jordan Brown, Charlie Lang, Unjin Lee, and Linsin Smith–thank you and please know that I will value your friendship forever.

To the friends who have shaped me into the person I am today, this is as much your accomplishment as it is mine. There are too many to list here, but I want to mention Stephen

Cholewka, Steven Nowicki, Anthony Frachioni, Edelio Bazan, Kaylin Mansley, Emily Underwood, Trish Luksich, David Curley, Nick Vernola, and Michael Fuller for all the cheerful years I have been lucky enough to call you my friends. You all deserve substantial recognition for what I have accomplished up until now.

To my family that is scattered around the globe, I hope you know I love you very much. Specifically, to my mother, Sabine, my sister, Katharina, her husband, Andrew, and my recently born niece, Sonia in Boston; my father, Uwe, his wife, Laura, my sister, Hanna, and brother, Anderson, in New York; to my grandmother, Margerete, in Moringen, Germany who will be turning 90 this year; to my aunt, Ute, and my uncle, Reinhard, in Frankfurt, Germany; to my cousin, Juliane, in Berlin, Germany; and to my late grandparents, Gertrud and Erhard Keppler, and late grandfather, Karl Wolfgang Roß; thank you for your sacrifice and kindness. I hope I have made you all proud.

Finally, I need to thank my partner, Nicole, for providing me with a foundation upon which I could build myself up during this uncertain time. A doctorate is a long and arduous journey that takes patience and substantial strength and you provided me with that and then some. Strength comes in many forms, but you were able to supply me with every configuration imaginable. I am lucky to have you in my life and I love you very much.

Abstract

Class I HLA molecules “present” short peptide antigens to NK and T cell receptors to convert intracellular information into an extracellular immune response. These molecules are encoded by thousands of different alleles, each one adept in presenting distinct peptides. Differences in peptide presentation are an important factor in determining which alleles are best at alerting lymphocytes to the presence of a viral infection. While *in silico* methods exist to predict which peptides a particular HLA molecule binds, none can accurately predict the likelihood of presentation on the cell surface, nor the three-dimensional structure of a particular peptide-HLA complex. Thus, *in vitro* molecular and cellular experiments are essential to understand the properties of uncharacterized HLA molecules. In this thesis, I experimentally studied the antigen-presenting capabilities of two class I HLA molecules: the rare and archaic classical molecule, HLA-B*73:01 and the unusual and highly conserved non-classical molecule, HLA-F. The investigation of HLA-B*73:01 consists of a high resolution description of the peptides it presents (referred to as its peptidome) and the structural mechanisms that dictate how it presents these peptides. For HLA-F, I pursued two separate but overlapping lines of inquiry: understanding the different protein forms of HLA-F and whether they can both be detected on the cell surface; and how HLA-F is regulated in endometrial stromal cells in the immune privileged context of human pregnancy. While the study of HLA-B*73:01 highlights how unique it is relative to other class I molecules, the reappraisal of cell surface expressed HLA-F actually suggests that it is more like other class I molecules than previously thought.

Chapter 1 - Introduction

1.1 The molecular diversity of class I HLA molecules and why it matters

1.1.1 How class I HLA molecules are presented

The major histocompatibility complex (MHC) is a vertebrate-specific genomic locus containing a set of closely linked genes that encode for proteins essential in the adaptive immune response (1). The encoded cell surface proteins, known as MHC molecules, can bind to endogenously or exogenously sourced molecules, referred to as antigens, and display or present them on the cell surface. Surveying leukocytes can then recognize antigen-bound MHCs using receptors which allow vertebrate immune systems to discern self from non-self. In humans, the six genes that encode for MHC molecules that present endogenously-sourced peptides are called class I human leukocyte antigen (HLA-I) molecules. HLA-I molecules canonically function as trimeric complexes made up of a heavy chain (the HLA polypeptide itself), a light chain (a separate protein called beta-2-microglobulin (β2M)), and an 8-15 amino acid peptide antigen (2,3). The peptide antigen is the most variable component of the complex and essential for communicating with CD8+ T cells, the cytotoxic representatives of adaptive lymphocytes (4). In order to ensure reactivity in the case of infection or malignancy, while also preventing unwanted autoreactivity, trafficking of HLA-I molecules to the cell surface is highly regulated.

Regulation of HLA-I molecules mostly takes place in the endoplasmic reticulum (ER) prior to trafficking through the Golgi before finally reaching the cell surface (**Figure 1.1A**). Peptide antigens are sourced from old, misfolded, or damaged proteins (5). Thus, peptide antigens can be

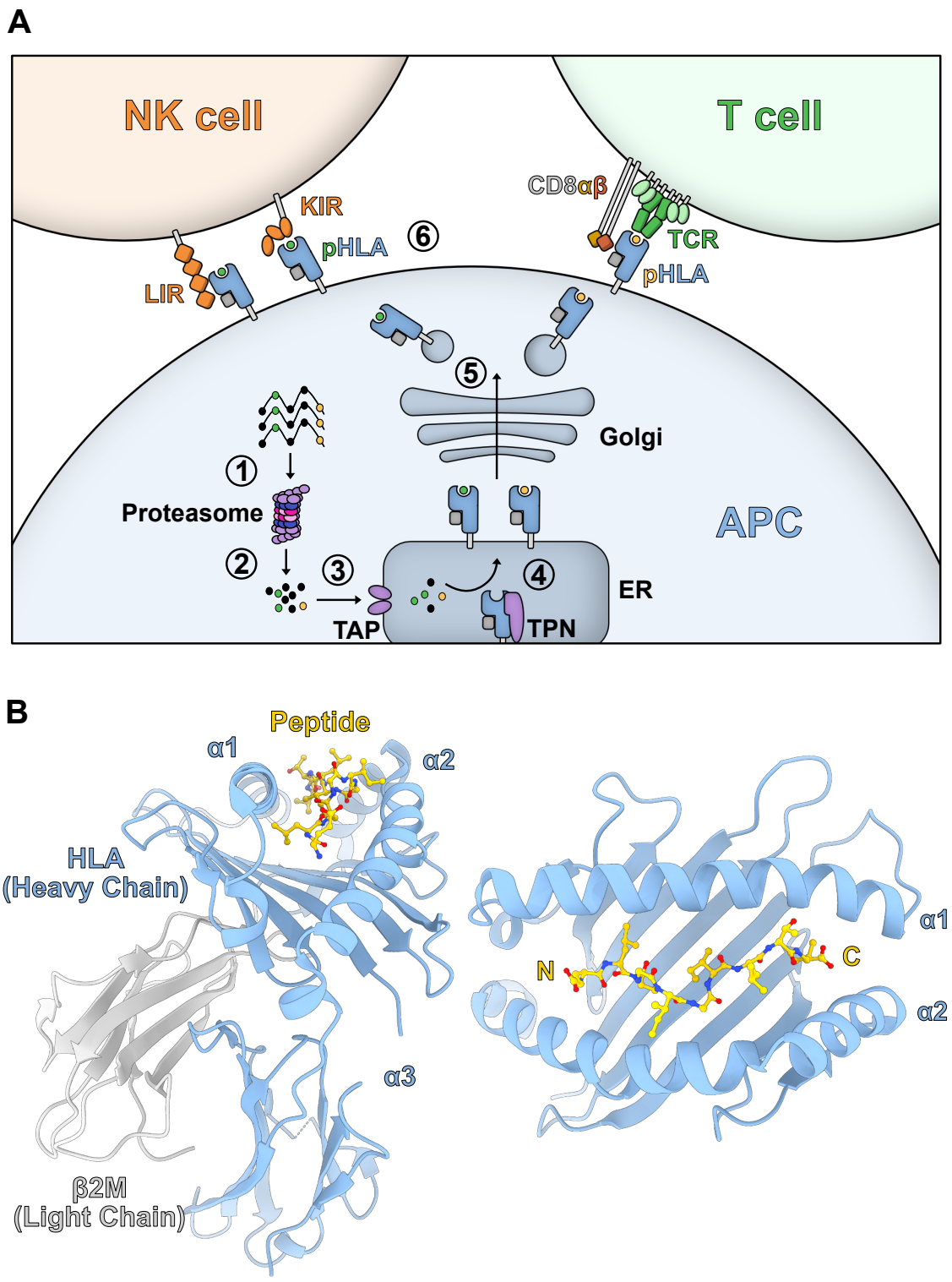

Figure 1.1 - The presentation pathway and structure of class I HLA molecules.

Figure 1.1 Continued
A. The canonical class I presentation pathway and consequent interactions with innate and adaptive cell surface receptors. (1) Cytoplasmic polypeptides are targeted for degradation; (2) polypeptides targeted by the proteasome are fragmented into smaller polypeptides; (3) the TAP machinery selectively imports these peptide fragments; (4) chaperones such as TPN aid in the assembly of stable peptide-MHC complexes; (5) table peptide-MHC complexes escape the ER and are transported to the cell surface via the Golgi; (6) cell surface expressed peptide-HLA molecules can interact with cell surface receptors on NK and T cells. **B.** Class I peptide HLA complexes consist of three parts: a variable peptide (yellow), a heavy chain that encodes the HLA protein (blue), and an invariant light chain (grey), also known as beta-2-microglobulin (β2M).

derived from the cell's own proteins (self), from viruses that have hijacked the cellular machinery (non-self), or from proteins that are ectopically produced in transformed cells (altered-self) as is found in the context of cancer (6). As proteins are turned over, their polypeptide sequence is broken down in either lysosomes or by proteasomes (7). The peptide fragments generated by proteasome-targeted degradation can then be exploited by the transporter associated with antigen processing (TAP) complex (8). TAP selectively imports fragments into the ER where they can interact with freshly synthesized HLA:β2M complexes stabilized by multiple chaperones including, but not limited to, tapasin (TPN) (9–11). TAP, TPN, and other chaperones involved in this pathway are collectively referred to as the peptide loading complex (PLC) (12).

Some of the peptide fragments imported into the ER have a high enough affinity for the HLA:β2M complex that they outcompete chaperones such as TPN and generate stable pHLA complexes (13). These stable pHLA complexes then exit the ER and are transported to the cell surface through the Golgi. Once at the surface, pHLA complexes can interact with adaptive immune receptors, such as T cell receptors (TCRs), or innate immune receptors, such as killer

immunoglobulin-like receptors (KIRs) and leukocyte immunoglobulin-like receptors (LIRs). Importantly, in the absence of a properly functioning PLC and/or high affinity peptides, class I molecules are not efficiently expressed on the cell surface and cannot be used effectively to communicate with lymphocytes (5,14). However, previous studies have suggested that one HLA-I molecule, HLA-F (discussed in more detail in **Chapter 1.1.3**), may not require the PLC in order to reach the cell surface. In **Chapter 3** of this thesis, we present data that suggests that HLA-F may, in fact, be dependent on the PLC, in a manner similar to other HLA-I molecules. However, an important question that remains is: what determines a high affinity peptide for any particular HLA-I molecule? In order to address this question, decades of work has utilized crystal structures to reveal incredible plasticity within the region of the protein that binds to peptide antigens.

1.1.2 Structures of HLA-I molecules reveal the mechanisms behind peptide binding

As mentioned above, HLA-I molecules are displayed on the cell surface as trimeric complexes. Peptide antigens bind between the α1 and α2 segments of the heavy chain, which together make up the platform domain (**Figure 1.1B and Figure 1.2A**). The platform domain is stabilized by β2M, which mainly interacts with the α3 domain and the underside of the platform to stabilize the heavy chain enough for peptide loading (15). Conserved interactions between the peptide and HLA heavy chain are hydrogen bonds formed at the N-terminus of the peptide mediated by tyrosines at positions 7, 59, 159, and 171 and the C-termini mediated by tyrosine 84, lysine 146, and additional amino acids at positions 143 and 147 of the α2 helix (16,17). A closer inspection of the N- and C-termini also reveals why peptides bound by HLA-I molecules have such strict

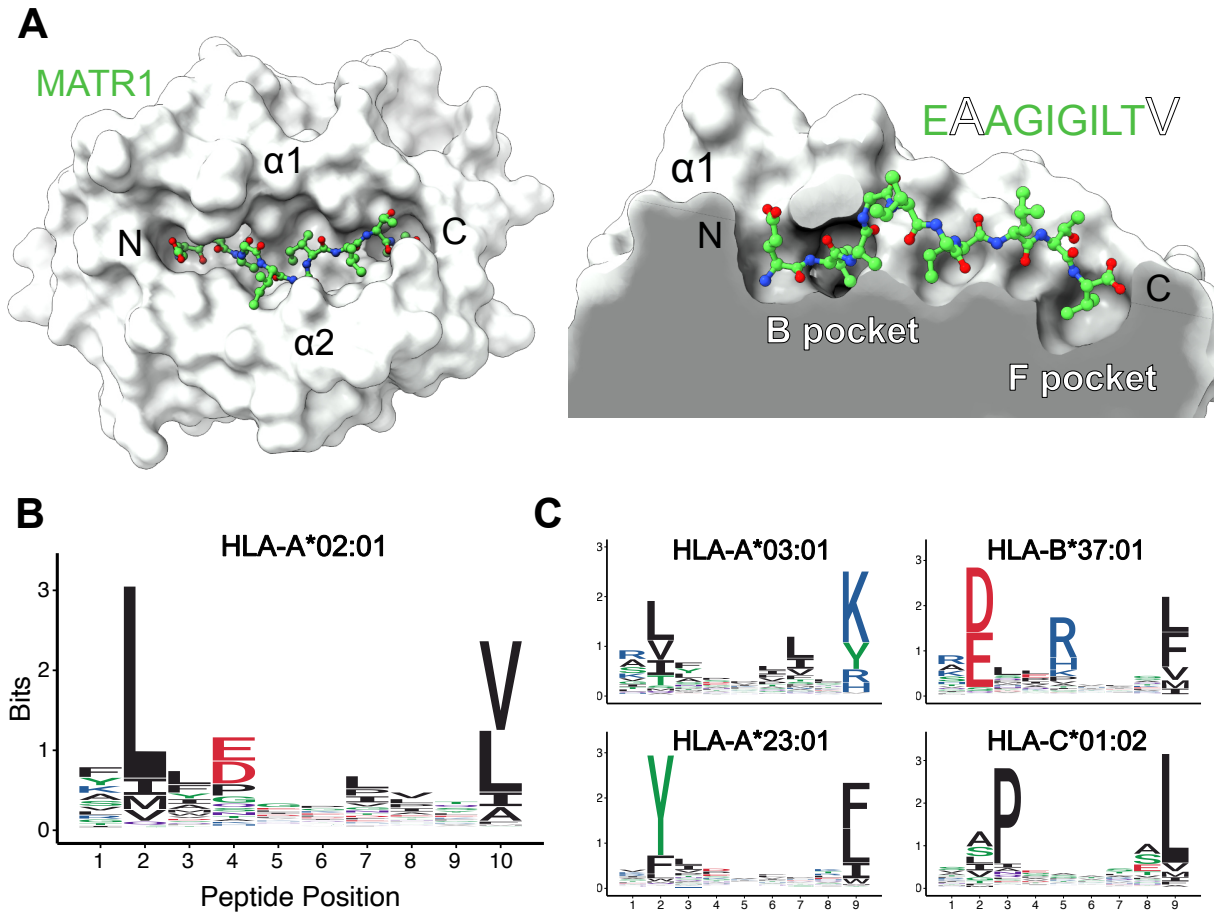

Figure 1.2 - Peptides are anchored into class I molecules at the N and C termini.
A. A top down view of HLA-A*02:01 represented as a surface presenting the MART1 peptide (left). On the right is a rotated view slicing the platform down the center of the binding groove in order to visualize the binding pockets indicated. The MART1 peptide sequence is indicated in white and green letters. White letters represent crucial anchor residues that are buried within the B and F pockets. **B.** A sequence logo representation of the peptide binding preferences of HLA-A*02:01 10mers. MART1 is also a 10mer. **C.** Sequence logo representations of the peptide binding preferences of four different alleles, indicated above.

length requirements. Peptides eluted from HLA-I molecules are typically 8-15 amino acids long, with the most common length being 9 amino acids. Residues at both ends of the platform domain, including, but not limited to residues 84 and 167, close off the peptide binding groove such that peptides are sterically hindered. This is in contrast to HLA-II molecules, which have open-ended peptide binding grooves to accommodate longer peptides (18).

In addition to conserved bonding interactions, structures of HLA molecules also revealed that peptides are anchored into variable pockets along the platform domain at canonical positions along the peptide (16,19–21). Anchors are defined as residues that are pointing downward into binding pockets accommodated by HLA-I platforms (**Figure 1.2B**). Most peptides are anchored into HLA-I binding grooves at P2–the second peptide residue from the N-terminus–and PΩ–the most C-terminal peptide residue (22). One HLA-I molecule, HLA-A*02:01, an allele of the HLA-A gene, prefers a leucine residue at P2 of peptide and a valine or leucine at PΩ of the peptide (**Figure 1.2B**). However, due to the malleable nature of HLA-I binding grooves, a peptide known as MART1, which encodes an alanine at P2 and a valine at PΩ, can bind as well.

Binding pockets are labeled from A through F, and the P2 anchor interacts with residues of the B pocket while the PΩ anchor interacts with residues of the F pocket. However, there are exceptions to this rule (**Figure 1.2C**). For example, another HLA-I molecule, HLA-B*37:01, an allele of the HLA-B gene, also has a preference for a basic residue at the P5 position while HLA-C*01:02, an allele of the HLA-C gene, is more strongly anchored at the P3 than the P2 position (22). The PΩ anchor is common to all HLA-I molecules studied to date, although the heterogeneity at this position varies by allele (22). Yet the alleles described so far represent only a small fraction of all HLA-I alleles, the vast majority of which are encoded by the HLA-A, HLA-B, and HLA-C loci, also known as the classical class-I molecules. The incredible degree of polymorphism in classical HLA-I genes can result in very different peptide binding properties, which can ultimately lead to predispositions to pathogen infections and autoimmune conditions.

1.1.3 The classical HLA-I molecules are highly polymorphic

The six HLA-I encoding genes can be further broken down into classical and non-classical molecules. The classical molecules, which consist of HLA-A, HLA-B, and HLA-C, are the most polymorphic genes in the human genome and are expressed on the surface of nearly all nucleated cell types (23). Polymorphisms are concentrated in the platform domain and this gives different alleles the ability to present unique peptides and differentially associated with chaperones of the PLC (**Figure 1.3A**). Indeed, the correlation observed between the most polymorphic residues and platform residues that are known to interact with bound peptides relative to the rest of the molecule is striking (**Figure 1.3B**). In addition to differences in bound peptides, polymorphisms also contribute to variable HLA-I cell surface expression levels (39) and variable dependence on chaperones of the PLC (39–41). Remarkably, even a single substitution can lead to drastically divergent molecular properties as is the case for HLA-B*44:02 and HLA-B44:05, which differ at a single residue in their platform domains (24).

While the physical differences between pHLA complexes may be small, they can have important functional consequences. For example, several studies have shown that HLA-A*01:01 is able to present a handful of conserved SARS-CoV-2 peptides to CD8+ T cells, suggesting that it can effectively protect against severe COVID19 (25–27). In contrast, carriers of HLA-B*46:01 are reported to be vulnerable to the original SARS virus (28). Consistent with this observation, this allele is cannot efficiently present SARS or SARS-CoV-2 peptides, leading to the hypothesis that HLA-B*46:01 carriers are also at increased risk of severe COVID19 (29,30). Instead, HLA-B*46:01 is thought to have evolved for its ability to protect against leprosy due to evidence that

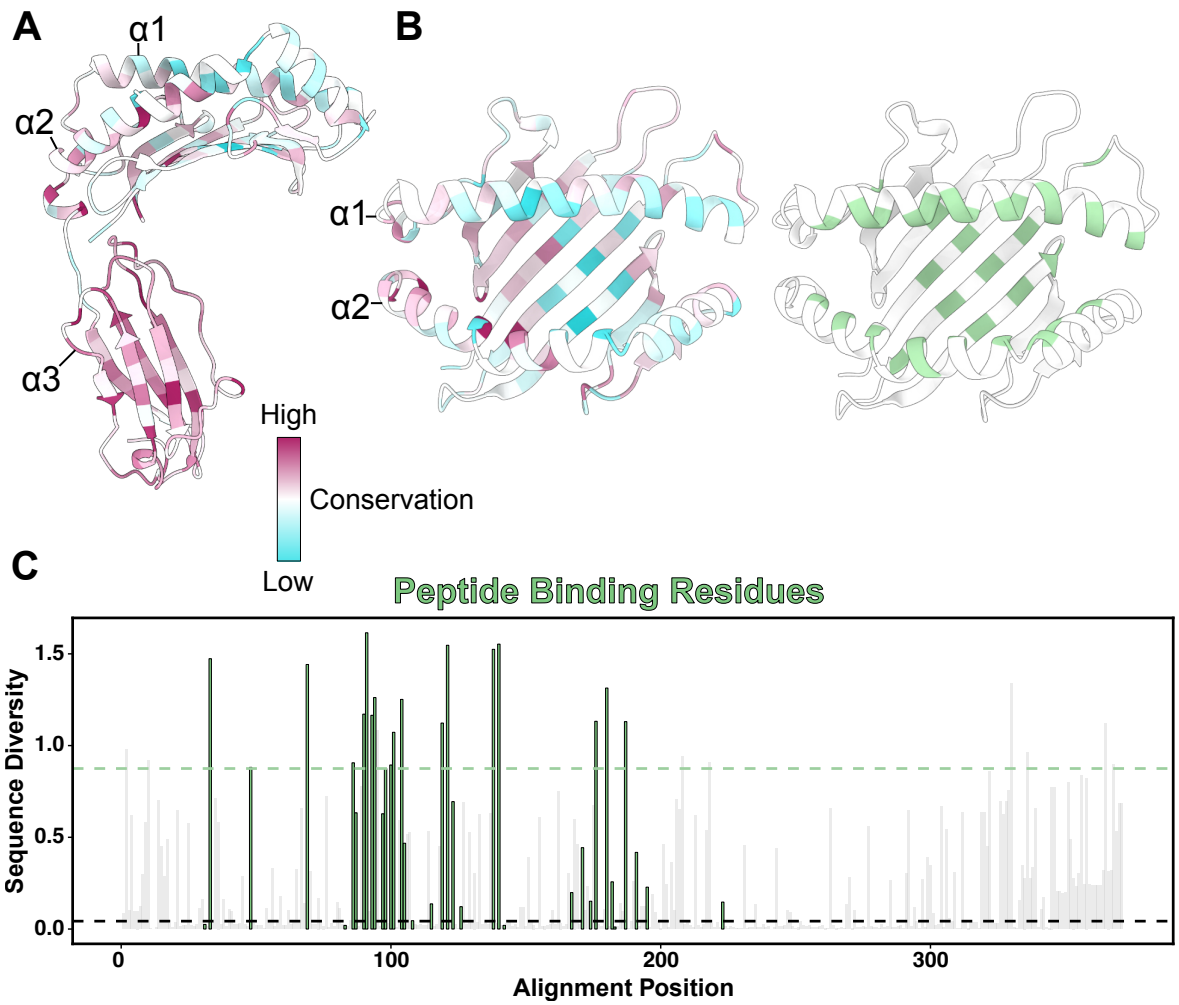

Figure 1.3 - Polymorphisms in HLA molecules are concentrated in the platform domains where peptides reside.
A. A heavy chain molecule colored by diversity in residue composition at each position of the mature protein. An alignment of over 7000 different protein alleles downloaded from IMGT of class I molecules was used to generate the conservation scores. The α3 domain is noticeably the most conserved domain. **B.** A top down view of the heavy chain shows much greater diversity (lower conservation) in the platform domain. Many of the positions with the lowest conservation are also positions associated with peptide binding (green). **C.** A bar plot quantifying the sequence diversity (Shannon) at each position. The black horizontal dashed line represents the mean diversity across the entire heavy chain. The green horizontal dashed line represents the mean diversity across residues shown in panel B.

it can present peptides derived from *Mycobacterium leprae* (31). Not surprisingly, specialization in protecting against certain pathogens may come with a negative trade-off in defending against

others, as appears to be the case for HLA-B*46:01.

In addition to differences in an individual's ability to mount an immune response to certain viral infections, carriers of certain HLA-I alleles can also be predisposed to autoimmune disorders (32). For example, while certain HLA-B*27 alleles may increase resistance to human immunodeficiency virus (HIV) progression and hepatitis C virus (HCV) infection (33), they also increase susceptibility to ankylosing spondylitis, an incurable form of arthritis that causes inflammation in the joints and ligaments of the spine (34). The association between Behçet's disease and HLA-B*51 is yet another example (35). BD is a chronic form of vasculitis characterized by repeated oral and genital ulcers and other multisystemic features. However, it is not yet known whether HLA-B*51 alleles also provide an advantage in fighting off pathogens similar to what is seen for HLA-B*27 alleles. It should be noted that most known associations between autoimmune disorders and HLA genes involve class II HLA molecules (32).

The exceptional polymorphism of classical HLA-I alleles is thought to be maintained due to selection pressure imposed mainly, although not exclusively, by microbial pathogens (36–40). Most current models suggest that HLA variants and the infectious agents they interact with become subject to Red Queen dynamics (41) where novel variants are consistently favored to keep up with the more rapid evolutionary rates of microbial pathogens (36,42). The result is a large number of HLA alleles maintained within populations to maximize the chances of carrying at least one allele that can efficiently present pathogen-derived peptides to NK and T cells and mount an effective immune response. However, there remains considerable debate over the

evolutionary processes that contribute to HLA polymorphism and to what extent they do contribute (36).

It has been suggested that several HLA alleles have been introduced into the modern human population through introgression with archaic humans (Denisovans and Neanderthals). Among these alleles was the exceptionally divergent HLA-B allele, HLA-B*73:01, which carries an amino acid epitope normally found in HLA-C alleles that allows it to be recognized by several innate NK cell receptors (43,44). Although Abi-Rached *et al.* proposed that HLA-B*73:01 was likely retained for adaptive reasons, genetic association studies that included HLA-B*73:01 carriers have failed to reach statistical significance potentially due to its low worldwide allele frequency (45,46). Interestingly, results from these studies suggest that HLA-B*73:01 may impede HIV disease progression and lower susceptibility to Behçet's disease. However, in the absence of established genetic associations, an investigation of the molecular properties of HLA-B*73:01 may help elucidate what makes HLA-B*73:01 distinct relative to other alleles. In **Chapter 2** of this thesis, we described a high resolution catalog of peptides presented by HLA-B*73:01 and determined the first crystal structure of HLA-B*73:01. We then used these data to better understand how HLA-B*73:01 presents the peptides it does and how this relates to its interesting evolutionary history.

1.1.4 The non-classical HLA-I genes are functionally divergent

The non-classical HLA-I genes, HLA-E, HLA-F and HLA-G, unlike the classical gHLA-I genes, are characterized by more restricted diversity within humans, with only a limited number of

alleles designated for each gene (3). The non-classical molecules are also more conserved evolutionarily, with orthologs identified in non-human primates. The least conserved member, HLA-G, has monomorphic orthologs in chimpanzee and gorilla lineages, but is polymorphic and functionally distinct in the orangutan and a pseudogene in the rhesus macaque (47). HLA-F has orthologs even in New World Monkeys, although they appear to be more polymorphic in the night monkey and have duplicated several times in the common marmoset (48). HLA-F orthologs in Old World Monkeys and hominids have low degrees of polymorphism, indicating that HLA-F and its orthologous loci are likely under purifying selection (48). HLA-E is thought to have the most well conserved orthologs within non-human primates and even has a functional analog in mice, suggesting that its particular properties are likely beneficial (47).

The non-classical molecules also have unique expression profiles and are more distinct in their peptide-presenting properties. HLA-E stands apart due to its ability to present HLA-I leader peptides and its reactivity with members of the NKG2 innate immune receptor family (49,50), while HLA-G is known for its exceptionally specific cell surface expression in placental cells known as extravillous trophoblasts (EVTs) (51) **(Figure 1.4**). HLA-F, on the other hand, remains the most puzzling of all HLA-I molecules. Despite reports that it too has a restricted expression profile (52,53), my own analysis of single-cell RNA sequencing (scRNA-seq) data from the Human Protein Atlas shows that its RNA expression profile is highly correlated with that of HLA-A, suggesting that at the RNA level, its expression is relatively ubiquitous across nucleated cells (**Figure 1.4**). However, this does not necessarily suggest the same for HLA-F protein

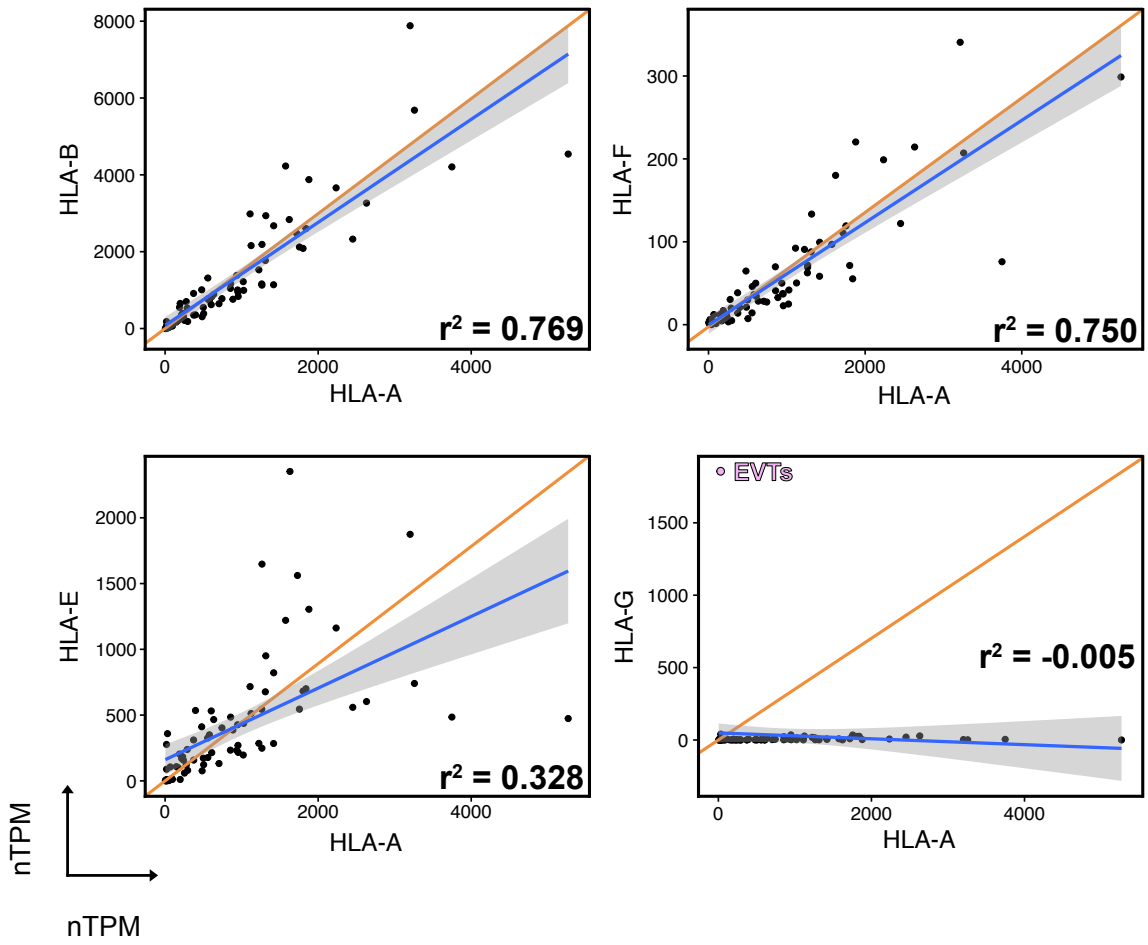

Figure 1.4 - The expression profiles of HLA-Ib molecules are distinct from HLA-Ia molecules.
Data from the Human Protein Atlas (https://www.proteinatlas.org/) was downloaded and used to generate scatter plots. Along the x-axis are the RNA expression values for HLA-A in different cell types as assayed by single-cell RNA-seq. The y-axis varies by quadrant: HLA-B is plotted on the top left, HLA-F on the top right, HLA-E on the bottom left, and HLA-G on the bottom right. Orange lines indicate a one-to-one relationship between x and y variables. Blue lines indicate linear regressions with standard error in grey. R^2 values represent the fits for estimated linear regressions.

expression within or on the cell. Indeed, HLA-F protein has historically been the most difficult HLA-I molecule to detect.

This has been in part due to the use of different antibodies to detect HLA-F protein. Early studies used non-specific reagents such as the the pan-HLA-I antibody, W6/32 (54), and the anti-β2M antibody, BBM.1 (55), to detect HLA-F, assuming that it behaves similarly to other HLA-I molecules by complexing with β2M and presenting peptide. Antibodies designed to be specific to HLA-F such as Fpep1.1 (56), FG1 (57), EPR6803 (Abcam ab126624; (58)), and 3D11 and 4A11 (52) have also been utilized. Not surprisingly, the use of different antibodies has at times led to conflicting conclusions regarding HLA-F and its properties. One example of this is in regard to HLA-F's association with the TAP complex of the PLC. Where one study using Fpep1.1 concluded that HLA-F interacts with TAP (56), another study using 3D11 and 4A11 concluded that HLA-F does not (53). Nonetheless, as the amount of data on HLA-F has grown, a unifying model to describe HLA-F relative to other class I molecules has emerged: unlike other classical and non-classical molecules, HLA-F can be expressed on the cell surface without peptide and perhaps even without β2M (59). Furthermore, HLA-F is expressed only on specific cell types, such as lymphocytes (60), under conditions that induce its expression, such as viral infection (58,61–63). And while these properties are unique among human class I molecules, they are reminiscent of the human MHC-related molecules, MICA and MICB (64,65), and the murine non-classical MHC molecules, T22 (66) and MILL2 (67).

This empty form of HLA-F is commonly referred to as its "open conformer" state (ocHLA-F) and more recent work has continued to expand upon its notable properties. For instance, multiple studies have now demonstrated that ocHLA-F can interact with lineage II KIR molecules KIR3DS1, KIR3DL1, and KIR3DL2 (58,68–70). Even more remarkable is that HLA-F protein

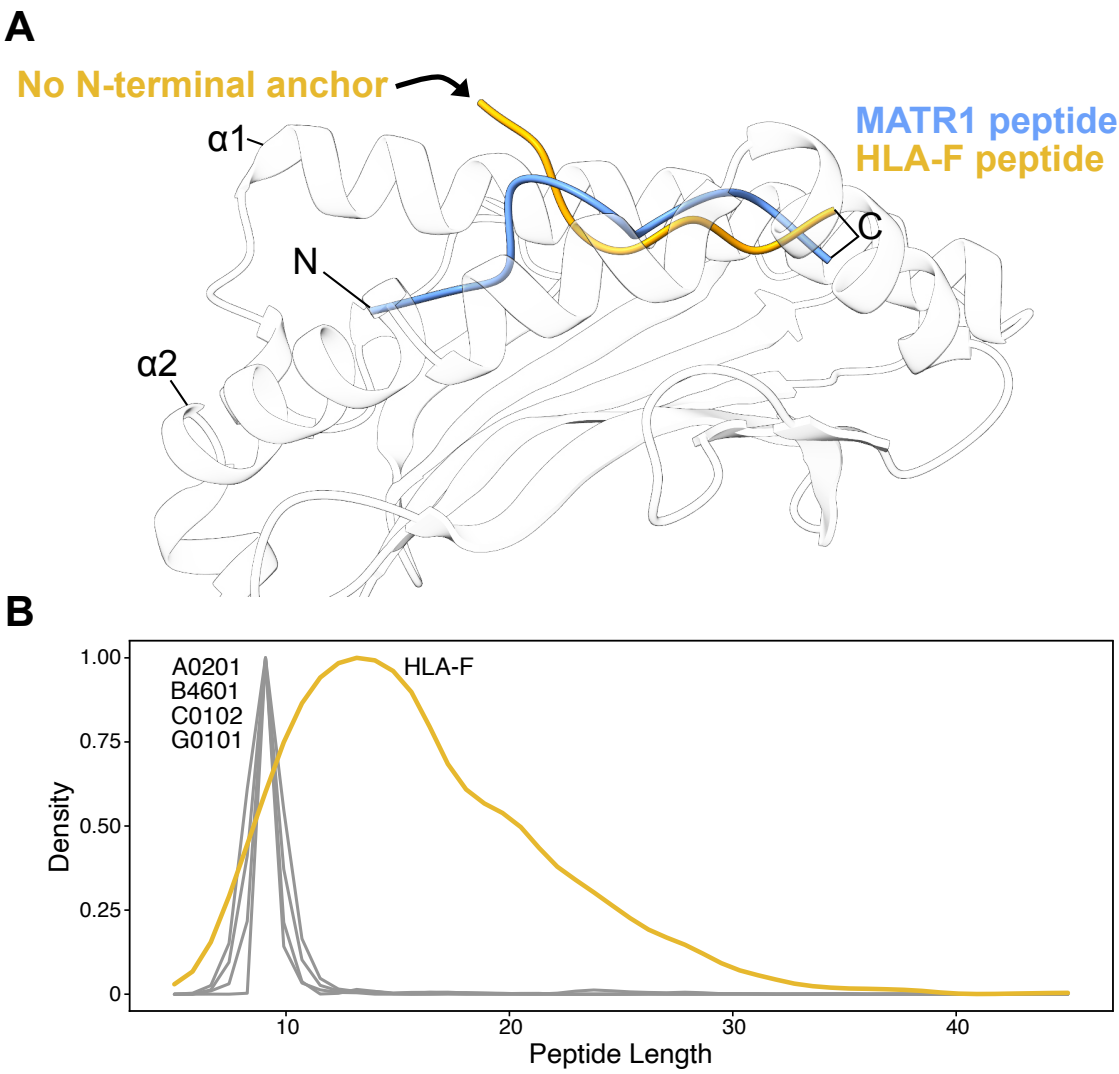

Figure 1.5 - HLA-F presents peptides that are not N-terminally anchored and that are unusually long.
A. A side view of the MHC binding groove showing ribbon representations of the MART1 peptide bound by HLA-A*02:01 and the peptide predicted to bind HLA-F. **B.** Peptide length distributions of peptides eluted from the corresponding HLA molecules and sequenced by LC-MS/MS. HLA-F data is derived from Dulberger *et al.* (68) using peptides eluted from recombinant protein produced in HEK293T cells.

has been shown to be induced in many different kinds of viral infections *in vitro*, including in human immunodeficiency virus (58), hepatitis C virus (62), BK polyomavirus (61), and adenovirus (63) infection. Moreover, its expression on various cell lines immortalized with

Epstein–Barr virus (EBV) signifies that EBV may also induce its expression (53). Together, these data indicate that one role of HLA-F may be to act as a virally-induced stress signaling molecule that can interact with an activating receptor such as KIR3DS1 (more details on KIRs can be found in **Chapter 1.2.2**) in order to inhibit viral replication. However, *in vivo* validation of this model is still lacking.

One important question this work raises is: what is structurally different about HLA-F that allows it to be expressed at the cell surface without peptide? In 2017, the Adams Lab published its first paper on HLA-F trying to address this question (68). To their surprise, Dulberger *et al.* showed that a recombinant form of HLA-F fused to β2M was able to present peptides when expressed in insect or mammalian cells, but not refolded from bacterial inclusion bodies. A crystal structure of HLA-F revealed what appeared to be peptide-like electron density within the peptide-binding groove. Furthermore, the peptides presented by HLA-F were unique in length and appearance relative to other class I molecules (**Figure 1.5**). In **Chapter 3** of this thesis, we review the findings of Dulberger *et al.* in greater detail and discuss the implications of HLA-F existing in multiple forms. We also examine the relevant forms of HLA-F found on the cell surface using newly generated reagents specific to different forms of HLA-F, generated by the Adams Lab in collaboration with the Kossiakoff Lab.

1.2 HLA-I molecules are ligands for several innate immune receptors

1.2.1 Leukocyte immunoglobulin-like receptors are peptide-agnostic

In addition to T cell receptors, HLA-I molecules are also ligands for several families of innate immune receptors. The LIR family of receptors consists of both activating and inhibitory receptors, but I will focus exclusively on LIR1 and LIR2, two structurally well-described inhibitory receptors. Although LIR1 and LIR2 both bind HLA-I molecules, their preferences slightly differ. LIR1 reacts more strongly with complexed peptide-loaded HLA molecules, whereas LIR2 is less biased and can bind to uncomplexed HLA heavy chains as well (71). Additionally, while LIR1 can be found on both myeloid cells and lymphocytes, LIR2 expression is more preferential to myeloid cells (71). Both molecules are polymorphic in humans and in the case of LIR1, polymorphisms have been shown to impact HLA recognition (72–74).

Structurally, LIR1 and LIR2 recognize pHLA in a peptide-agnostic manner, with an interaction supported entirely by β2M and the less polymorphic α3 domain (**Figure 1.6A**). LIR1's physiological importance is well understood as it has been implicated in a variety of contexts including malaria, viral infection, cancer, and pregnancy (75–80). In pregnancy, the function of LIR1 has only recently been characterized. Its expression on decidual NK (dNK) cells allows it to interact with HLA-G dimers expressed on EVTs leading to the secretion of essential fetal growth factors (79). Additionally, it was shown that dNK cell populations in multigravid women become "trained" to express significantly higher levels of LIR1, presumably to better prepare for future EVTs expressing HLA-G (78). These data suggest that LIR1's specification as an

inhibitory receptor is likely oversimplified. Finally, while it was previously shown that LIR1 also recognizes HLA-F (81), the Adams Lab showed that LIR1 actually binds HLA-F with even greater affinity than HLA-G *in vitro* (68). Whether an HLA-F:LIR1 signaling axis plays an important role in any physiological contexts remains to be determined. However, in **Chapter 4**

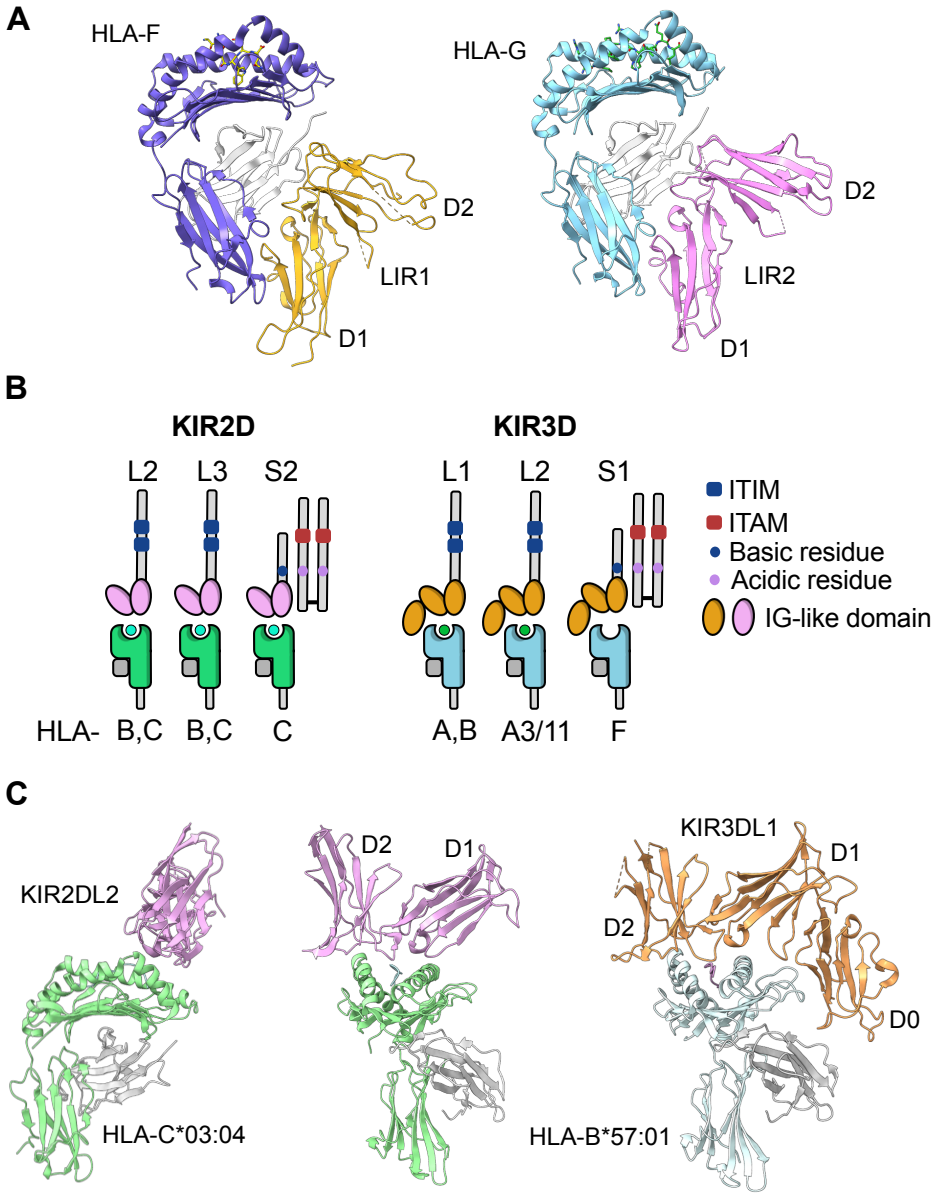

Figure 1.6 - Structural depictions of LIR and KIR engagement with class I HLA molecules.

Figure 1.6 Continued
A. LIR1 and LIR2 bind class I HLA molecules agnostic to the peptide. Significant surface area of the binding interaction between LIR1 and LIR2 and HLA molecules is buried between the LIR and β2M (PDB: 5KNM and 2DYP). **B.** Cartoon representations of KIRs relevant for this thesis. KIRs are broken up into molecules with 2 (KIR2D) and 3 (KIR3D) IG-like domains. KIR2D and KIR3D molecules can be further broken down into inhibitory (this with long tails and ITIM motifs) and activating (those with short tails and ITAM-containing adaptor molecules) receptors. Shown are HLA genes with alleles that are capable of binding the different KIR molecules. Only HLA-F unloaded with peptide has been shown to interact with KIR3DS1. **C.** KIR molecules bind towards the C-terminal end of HLA molecules above the F pocket. All KIR structures solved to date bind in a similar manner and have significantly overlapping binding footprints. The D0 domain of KIR3D molecules add additional interactions with loops near the α1 helices (PDB: 1EFX and 3VH8).

of this thesis, we present a model proposing that HLA-F interacts with LIR1 in the context of pregnancy establishment and maintenance.

1.2.2 Killer immunoglobulin-like receptors are peptide-selective

Killer immunoglobulin-like receptors (KIRs) are germline-encoded innate immune receptors expressed mainly on NK cells, but also on certain T cell subsets (82–85). The polymorphic nature of classical HLA-I genes is rivaled only by the KIR locus (23,86). Individuals often carry not only unique haplotypes, but unique alleles within those haplotypes. Since HLA-I molecules serve as ligands for KIRs, ligand:receptor pairs likely co-evolve with one another, due to a combination of selective pressures put on humans by infectious diseases and successful reproduction (86). Similar to LIRs, KIRs also come in activating and inhibitory forms (**Figure 1.6B**). However, unlike LIRs, KIRs are not peptide-agnostic as KIRs bind to HLA-I platform domains, in a manner similar to TCRs, but biased towards the C-terminal side of the peptide ligand (87–89). Interestingly, activating and inhibitory forms have been shown to be more and less peptide-selective, respectively (90–92).

From a structural point of view, all KIR:HLA complex structures solved thus far demonstrate binding directly over the F pocket of HLA-I (**Figure 1.6C**) (87,88,93–96). Depending on the P7 and P8 residues of the bound peptide, KIR affinity can vary by several orders of magnitude and affinity directly correlates with downstream signaling (92). KIR recognition of HLA-I molecules is determined by the presence of specific KIR epitopes, which are not typically constrained to specific HLA-I genes. For example, the Bw4 epitope recognized by KIR3DL1 alleles is found in HLA-A and HLA-B alleles (97). Furthermore, the C1 epitope recognized by KIR2DL2 and KIR2DL3 can be found in both HLA-B and HLA-C alleles (97). However, to be clear, not all HLA-A and HLA-B alleles confer reactivity to KIR3DL1, nor do all HLA-B or HLA-C alleles confer reactivity to KIR2DL2 or KIR2DL3. Rather, a mix of KIR-reactive and non-reactive HLA-I alleles segregate within the human population and as a result, not everyone encodes HLA-I molecules that can interact with their own endogenously encoded KIRs (97). Physiologically, this can lead to differences in NK cell education, which can have functional consequences such as NK cells with an enhanced or reduced cytotoxic reaction to the loss of inhibitory MHC signaling (82).

Out of all HLA-I molecules, HLA-C is the only gene of which all alleles carry a KIR binding epitope (97). While the C1 epitope confers reactivity to KIR2DL2, KIR2DL3, and KIR2DS2, the C2 epitope confers reactivity to KIR2DL1 and KIR2DS1 (90,98). A small number of substitutions between KIR2DL2/L3 and KIR2DL1 and between HLA-C alleles containing C1 and C2 epitopes divide their reactivities (90,98). As mentioned before, the C1 epitope can also be

found in some HLA-B alleles. However, while greater than 10% of HLA-B alleles in chimpanzee, bonobo, gorilla, and orangutan encode C1 epitopes, only two allele groups (<2% of all HLA-B alleles) have retained the C1 epitope: HLA-B*46 and HLA-B*73 (99). Importantly, reactivity between HLA-B*46:01 and HLA-B*73:01 and KIR2DL2/L3-Fc fusion proteins has been confirmed empirically (88,98,100–102).

While, the reason for the difference in the number of C1-containing (C1+) HLA-B alleles between humans and non-human primates is unknown, the consequences of carrying HLA-B*46:01 were recently described in individuals with HIV progression (103). The main finding in this study was that HLA-B*46:01 is associated with rapid HIV disease progression. Furthermore, this study found that NK cells in HIV-infected HLA-B*46:01 carriers were hypoactivated relative to NK cells in non-HLA-B*46:01 carriers, suggesting that the NK cells within these individuals were likely altered due to the interactions between HLA-B*46:01 and its cognate KIRs. These data suggest that in humans there are fitness costs associated with carrying a C1+ HLA-B allele. While the interaction between HLA-B*73:01 and KIR2DL2/L3 has not been comprehensively characterized, in **Chapter 2** of this thesis, we present a crystal structure of HLA-B*73:01 with KIR2DL2 and show that its unique mode of interaction may contribute to a distinct relationship between ligand and receptor among C1+ alleles.

1.3 Pregnancy establishment and maintenance relies heavily on HLA-I molecules

1.3.1 A unique combination of HLA-I molecules mediate the relationship between mother and fetus

Although mammalian immune systems have evolved to detect and remove what they consider foreign, maternal immune systems have also had to adapt to a semi-allogeneic fetus. Despite Peter Medawar's appreciation of this connection 65 years ago, the genes and molecular mechanisms that govern fetomaternal tolerance remain poorly understood (104). The uterus is considered an 'immune privileged' organ thanks to its main function in establishing fetal immune tolerance leading up to and during pregnancy (105). However, rather than excluding or suppressing its local immune environment in order to achieve this goal, the uterus harbors multiple populations of specialized maternal immune cells, many of which are known to extensively remodel maternal tissue and interact directly with fetal cells (106–111). Ultimately, every healthy pregnancy relies on the appropriate balance of maternal immune activation and inhibition in order to achieve fetal tolerance and for both mother and fetus to thrive.

The maternal-fetal interface defines the boundary between mother and fetus: the tissue including and surrounding where the placenta directly implants into the uppermost uterine epithelial and stromal layers, referred to as the endometrium (112). Placental villi are made up of cells that anchor into the endometrium during implantation and invade the endometrium throughout pregnancy similar to how cancer cells permeate nearby healthy tissue (105). These cells, known as villous and extravillous trophoblasts (VTs and EVTs, respectively), have been the subjects of

intense scientific investigation due to their unique immunosuppressive properties (113–116) and crucial roles in establishing fetomaternal tolerance (105). These abilities are partially due to their unique HLA expression pattern: while most nucleated cells in the human body express HLA-A, -B and -C, EVTs preferentially express HLA-C, HLA-E and HLA-G (117). Consequently, they express several ligands for various LIRs and KIRs expressed on maternal leukocytes including decidual NK (dNK) cells.

Unlike in peripheral blood, NK cells make up the largest proportion of uterine leukocytes (118). Following decidualization and differentiation of uterine NK cells into dNK cells, they represent up to 70% of all leukocytes during the first trimester of pregnancy (119). Furthermore, it is now well understood that dNK cells play vital roles in fetal development through the secretion of growth-promoting factors and regulation of EVT invasion (79,107). Thus, it has been hypothesized that the interactions between EVTs and dNK cells are essential for establishing a balanced maternal immune response, although these hypotheses remain difficult to test in humans.

1.3.2 HLA-F's role in pregnancy warrants further investigation

HLA-F has also been shown to be expressed on EVTs, although the findings are considered controversial. Evidence that HLA-F may have a role to play at the maternal-fetal interface was first put forth in 2003 in a study using the 3D11 antibody (52). The authors used immunohistochemistry (IHC) to show that cells during the first trimester of pregnancy that expressed HLA-F also expressed HLA-G, strongly suggesting that these were true EVTs. Several

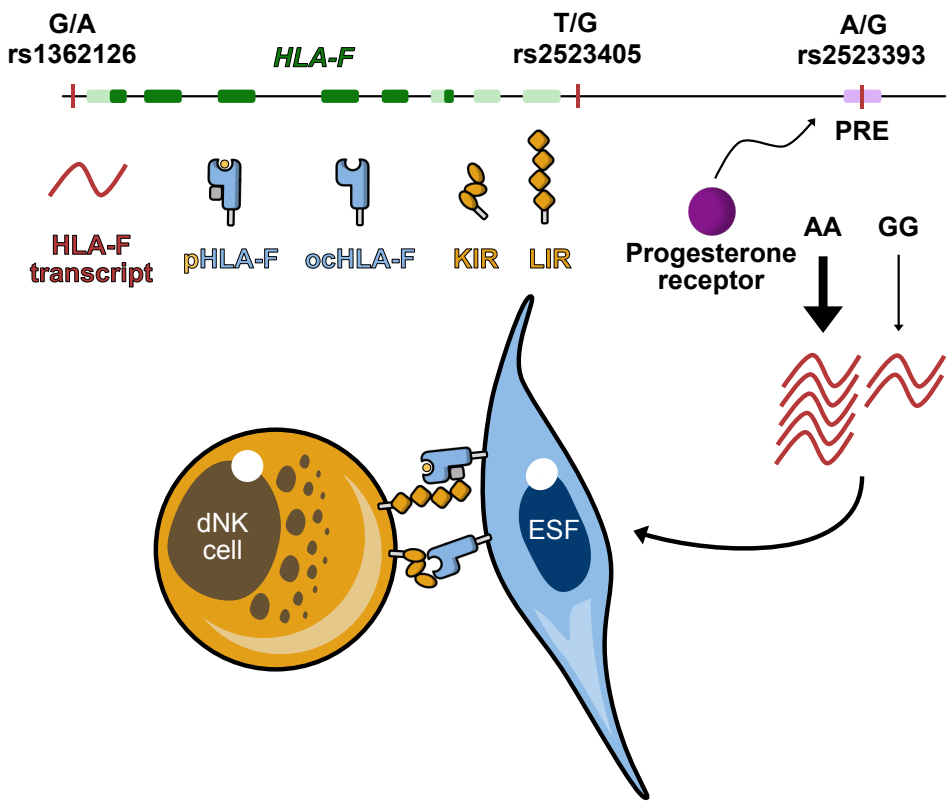

Figure 1.7 - HLA-F transcript expression levels are inversely associated with fecundability.
A model depicting putative mechanisms underlying HLA-F's association with fecund ability adapted from (130). SNPs found in the promoter, terminator, and enhancer region of HLA-F are associated with more or less HLA-F transcript. In particular, rs2523393 is located in a progesterone responsive enhancer (PRE), which is influenced by *in vitro* decidualization. HLA-F transcript then leads to the expression of HLA-F on the cell surface of ESFs where it could directly interact with receptors expressed on uterine NK cells, specifically LIR1 and KIR3DS1, KIR3DL1, or KIR3DL2.

studies have since claimed to replicate these findings using IHC and additional HLA-F specific antibodies (120–122). However, other groups have come to different conclusions. One study determined that HLA-F protein is detectable in EVTs, but not at the cell surface (123) while two other studies failed to detect HLA-F expression on EVTs by flow cytometry (64,124). For now, it remains unclear why different groups have come to different conclusions and thus whether HLA-

F is truly expressed on EVTs. Despite this, HLA-F protein expression has been detected in endometrial tissue.

Surprisingly little is currently known about HLA expression in the endometrium and their role(s) in priming uterine leukocytes prior to and following implantation. Interestingly, work in mice has shown that maternally expressed MHCs do play important roles in regulating fetal growth and decidual vascular remodeling (125,126). Although similar studies have not yet been performed in humans, *HLA-F* transcript expression in mid-secretory endometrium was found to be associated with time to pregnancy (127).

Additionally, studies have shown that HLA-F protein expression can be detected in the endometrium by IHC (128,129). One study was also able to detect HLA-F cell surface expression on endometrial stromal fibroblasts (ESFs) by flow cytometry (129). Together, these data suggest that *HLA-F* expression in endometrial tissue may play an important role in regulating pregnancy establishment and/or maintenance (**Figure 1.7**; adapted from (130)). In **Chapter 4** of this thesis, we elaborate on data implicating HLA-F as an important regulator of human pregnancy and present the results of an investigation into HLA-F expression at the maternal-fetal interface. We clarified where HLA-F protein is being expressed, how HLA-F protein expression is being regulated, and tested whether decidual leukocytes can recognize HLA-F protein.

Chapter 2 - The unique peptidome and structure of the archaic HLA-B*73:01

2.1 Archaic HLA alleles and their killer immunoglobulin-like receptors

When modern humans migrated out of Africa, they hybridized with archaic humans, including Neanderthals and Denisovans, that were already resident in Europe and Asia (131,132). As a result, 1.5-6% of modern non-African human genomes derive from archaic humans (133). In general, purifying selection has acted to remove archaic human DNA from modern humans (134). However, a limited number of introgressed (i.e. gained through interbreeding) archaic human genes have been preferentially retained (43,135,136). Amongst these positively selected loci are several dedicated to immune function, including multiple alleles of the highly polymorphic classical HLA-I genes, HLA-A, -B and -C (43).

Admixture with archaic humans is hypothesized to have restored HLA diversity in modern humans following the population bottleneck that occurred during the “out-of-Africa " migration. Consequently, admixture may have been a route to acquire advantageous HLA variants already adapted to local pathogens (43,137). One archaic HLA allele of particular interest is HLA-B*73:01. This exceptional allele, the only member of a deeply divergent lineage (MHC-BII), is distinct from the MHC-BI lineage to which all other human HLA-B alleles belong and is more closely related to subsets of chimpanzee and gorilla MHC-B than other human HLA-B alleles (43,44). Further distinguishing HLA-B*73:01, is that it is one of only two of the greater than 3000 HLA-B alleles that encodes the C1 NK cell binding epitope formed by residues V76 and N80 in the α1 helix (98,138). This epitope confers reactivity with KIR2DL3 and KIR2DL2,

inhibitory receptors that regulate the function of human NK cells predominantly through engagement with HLA-C alleles. HLA-B*46:01, the other C1+ HLA-B allele, acquired the C1 epitope more recently through a mini-gene conversion with HLA-C*01:02 (31,139).

Different combinations of the C1 epitope and KIR2D receptors are also known to be associated with various disease states (140). One of the earliest examples of this was the discovery that individuals that are homozygous for the C1 epitope and KIR2DL3 are better protected from chronic HCV than individuals without C1 epitope-containing HLAs or C1 carriers that instead encoded KIR2DL2 (141). Although this association is thought to be mediated by NK cells, KIR2DL2 has also been shown to amplify the immune response to various viral infections in a T cell-dependent manner, suggesting that KIR2DL2 also has a role to play in adaptive immunity (142). In another example, an in depth analysis of HLA-B*46:01 carriers in Southeast Asia showed that HIV-infected carriers of HLA-B*46:01 more rapidly progressed to AIDs than individuals without HLA-B*46:01 (103). A single-cell transcriptome analysis revealed this may be caused by a less activated NK cell subpopulation found exclusively in HLA-B*46:01 carriers prior to infection. Considering that each individual encodes either KIR2DL2, KIR2DL3, or both, individuals that also carry HLA-B*73:01 may have unique NK cell phenotypes, as is evident for carriers of HLA-B*46:01. Interestingly, the mechanisms that underlie these disease associations remain unknown, although it likely involves a combination of HLA-bound peptide ligands and their affects on KIR recognition.

While the peptidome of HLA-B*46:01 has been characterized (31), an unbiased investigation of HLA-B*73:01 peptides has not been performed. Additionally, due to the low frequency of HLA-B*73:01 in modern day human populations (43), genetic associations between HLA-B*73:01 and disease, similar to those described for other HLA alleles, would be extremely underpowered and consequently unreliable. Therefore, to better understand why HLA-B*73:01 was retained in modern humans, we began by investigating its molecular properties, such as its peptide-binding profile and three-dimensional atomic structure. We show here that HLA-B*73:01 presents unique peptides and has above average cell surface expression relative to other class I alleles. Using the first determined structure of HLA-B*73:01, we go on to describe the structural mechanisms that contribute to its unique peptide-binding profile. Finally, we also determined the first crystal structure of an HLA-B allele in complex with a KIR2D molecule to provide insight into the strategies used by HLA-B*73:01, and potentially also HLA-B*46:01, to engage with KIR2DL2. Our findings indicate that HLA-B*73:01 encodes a protein with several distinctive features that may underlie its adaptive introgression into modern human genomes.

2.2 HLA-B*73:01 presents rare peptides with an unusual length distribution

Although previous studies have characterized a small number of peptides bound by HLA-B*73:01 (143,144), a comprehensive atlas of HLA-B*73:01 peptides presented on cells (i.e. the HLA-B*73:01 peptidome) is currently lacking. Thus, to investigate its peptidome, we purified heterogeneously loaded HLA-B*73:01, eluted the bound peptides, and determined their sequences by LC-MS/MS. We then compared the HLA-B*73:01 peptidome with those of other well characterized alleles—HLA-B*15:01, HLA-C*01:02, and HLA-B*46:01—which were

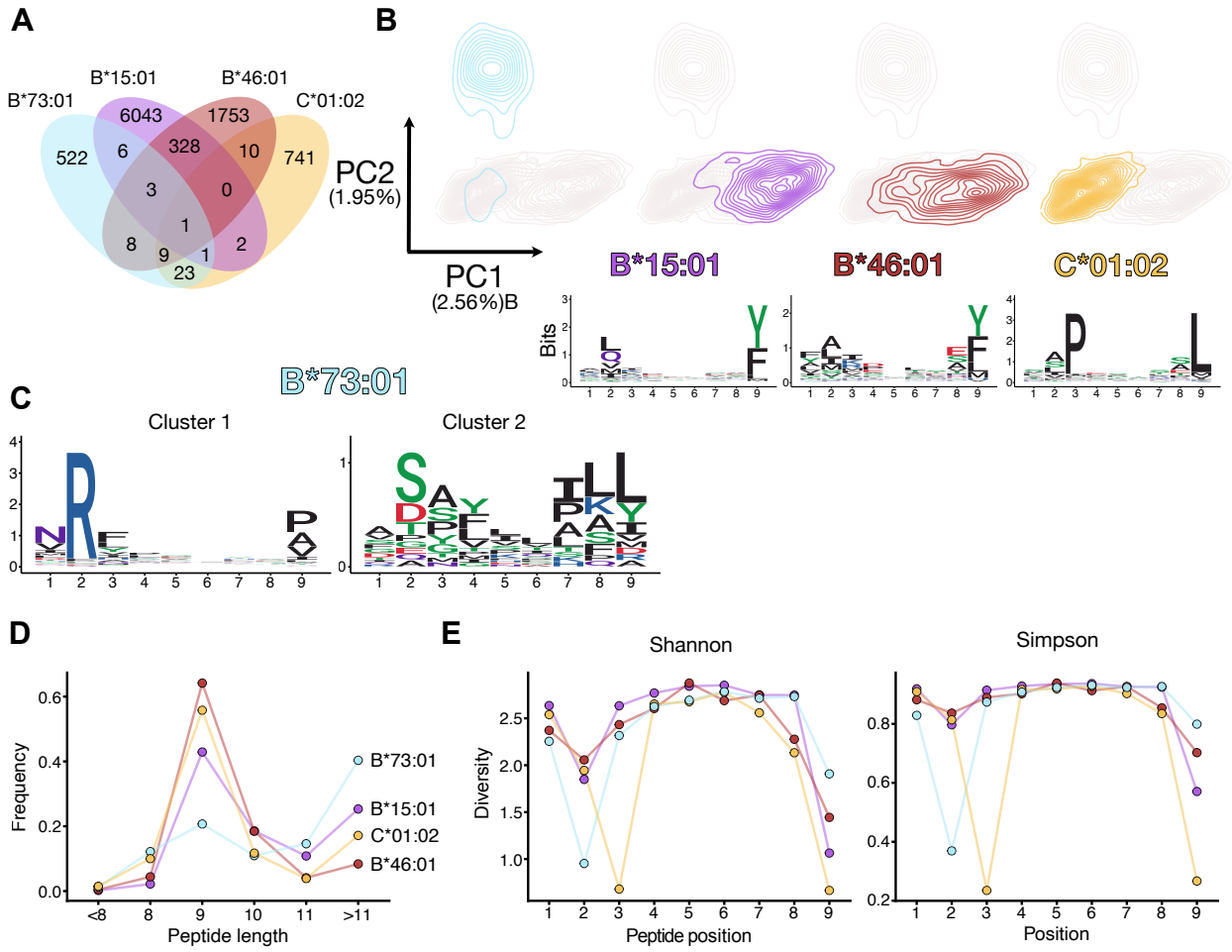

Figure 2.1 - HLA-B*73:01 independently evolved a rare peptide binding repertoire with an unusual length distribution.

A. Venn diagram showing the number of distinct peptides eluted from 4 HLA class I allotypes. **B.** Principal components analysis contour plots show the distribution of nonamer peptides bound by each HLA class I allotype. For comparison, in each plot the distribution of nonamer peptides from the other allotypes is shown in grey. The percentage of the variation in the dataset shown by the first (PC1) and second (PC2) principal components is listed to the left. **C.** Sequence logo of 9mers are shown for HLA-B*73:01, -B*15:01, -B*46:01, and -C*01:02. The total height of the column of amino acids represents the level of conservation at that position and the residue height denotes its relative frequency. The colors of each amino acid correspond to their biochemical characteristics: acidic (red), basic (blue), hydrophobic (black) and polar (green). **D.** For each allotype, the histogram divides the binding peptides according to length, given as the number of amino-acid residues. **E.** Shannon and Simpson diversity indices calculated for each position for each of 4 HLA class I allotypes using peptides eluted and sequenced by mass spectrometry in our study.

expressed, purified, and sequenced similarly in a previous study (**Figure 2.1A**) (31). With only

573 distinct peptides confidently identified, HLA-B*73:01 has the smallest of the four peptidomes (i.e. number of peptides) compared here and is an order of magnitude smaller than the greater than 6000 peptides bound by HLA-B*15:01 (**Figure 2.1A**). Consistent with the hypothesis that HLA-B*73:01 presents distinct peptides, fewer than 10% of the HLA-B*73:01 bound peptides were also identified among the peptides eluted from other alleles.

Visualization of the relative similarity of each peptidome by principal component analysis supports this finding, showing that most HLA-B*73:01-bound peptides have little spatial overlap with those of the other three alleles (**Figure 2.1B**), where the greatest contributor to PC2 is the hydropathy index or on average how hydrophobic a peptide is predicted to be (55%). A comparison of 9mer peptide binding motifs further illuminates this result, as HLA-B*73:01 has independently evolved a strong biochemical preference for an arginine at the second residue of the bound peptide (P2) (**Figure 2.1C**). Using an unsupervised alignment and clustering of peptide sequences (145,146), we identified 2 divergent peptide binding motifs for HLA-B*73:01 (**Figure 2.1C)**. While the peptides from cluster 1 are unique to HLA-B*73:01, the peptides from cluster 2 show a greater resemblance to peptides bound by HLA-B*15:01, HLA-B*46:01, and HLA-C*01:02. Cluster 1 peptides do also show a strong resemblance to peptides bound by the HLA-B*27 family of alleles, of which certain alleles have been shown to either protect individuals from infectious diseases or make them more susceptible to autoimmune conditions (147).

We observed that the length distribution of peptides bound by HLA-B*73:01 is evocative of those that bind HLA class II molecules, which, by virtue of their open-ended peptide binding groove, routinely bind longer peptides (148). While 9mers are the canonical length of HLA class I presented peptides, we found that only ~20% of HLA-B*73:01-presented peptides are 9 amino acids in length. For reference, ~40% of HLA-B*15:01-presented peptides are 9mers, the second lowest proportion of the alleles compared here (**Figure 2.1D**). Surprisingly, in comparison to presented 9mers, HLA-B*73:01 presents nearly double the frequency of peptides longer than 11 amino acids. To understand how HLA-B*73:01 is able to accommodate such long peptides, we stratified peptides by length and compared their peptide binding motifs. This investigation revealed that even as peptides get longer, they retain the P2 and PΩ anchors identified in HLA-B*73:01 cluster 1 9mers, suggesting that longer peptides likely bulge directly out from the peptide binding pocket without destabilizing the protein complex (**Figure S2.2A**). However, if accurate, this attribute would not be unique to HLA-B*73:01 (16).

Recent reports suggest that HLAs can be categorized as either ‘generalists’ or ‘specialists’ depending on the promiscuity of the peptides they can present (148). To quantify the diversity of peptides bound by HLA-B*73:01 we calculated diversity indices using either Shannon or Simpson measures at each position of presented 9mers (**Figure 2.1E**). Interestingly, while these analyses show that HLA-B*73:01 has relatively low sequence diversity at P1 and P2 (anchored within the A and B pocket, respectively), it has the highest sequence diversity of all 4 alleles at position 9 (anchored within the F pocket). In contrast, peptides presented by HLA-C*01:02 have low sequence diversity at P3 and P9, while surprisingly still presenting more peptides overall

(**Figure 2.1A**). Two things that may account for this are peptide preferences associated with the TAP machinery and the relative abundance of peptide fragments found in the human proteome that can bind different HLA alleles. In other words, peptides with a XXPXXXXXL motif may be more prevalent overall and possibly also within the ER than peptides with a XRXXXXXXP motif. Another contributing factor could be HLA-B*73:01's known interaction with TAPBPR, a second editor within the peptide-loading pathway, which, unlike tapasin, performs peptide exchange following release from the ER (149). Since the effect of TAPBPR on the peptides presented by HLA-C*01:02 and HLA-B*73:01 is currently unknown, future investigation will be needed to thoroughly test these hypotheses. Together, these data show that HLA-B*73:01 has evolved the ability to present peptides with a rare binding motif relative to other more common alleles. Additionally, these peptides can often be longer than typical HLA class I peptides and vary most at their C-terminal anchor, while being biased towards asparagine and arginine at the P1 position and P2 anchor position, respectively.

2.3 Peptides bound by HLA-B*73:01 are rarely presented by other class I molecules

While comparisons to alleles HLA-B*15:01, -B*46:01, and -C*01:02 can be insightful, we also sought to quantitatively compare HLA-B*73:01 to other common alleles found worldwide to better understand how unique HLA-B*73:01 truly is. Recent studies have massively expanded the amount of comprehensive and curated peptidome datasets for various HLA alleles (22,150–152). Thus we integrated our HLA-B*73:01 peptidome data with data from Sarkizova *et al.*, which includes greater than 185,000 total peptides from 95 alleles, chosen such that 95% of individuals worldwide have at least one of the HLA-A, HLA-B and HLA-C alleles included (22).

The peptidome for each allele was generated using the same parental cell line (BLCL 721.221) as was used in our study. To confirm that we could confidently combine these two datasets together, we compared the peptide binding motifs of eluted 9mers for HLA-B*15:01, -B*46:01, and -C*01:02. Reassuringly, we found that eluted 9mers were highly concordant between both datasets with Pearson correlation values of 0.93, 0.93, and 0.99 for HLA-B*15:01, -B*46:01, and -C*01:02, respectively (**Figure S2.1B**). This suggests that the peptidome of HLA-B*73:01

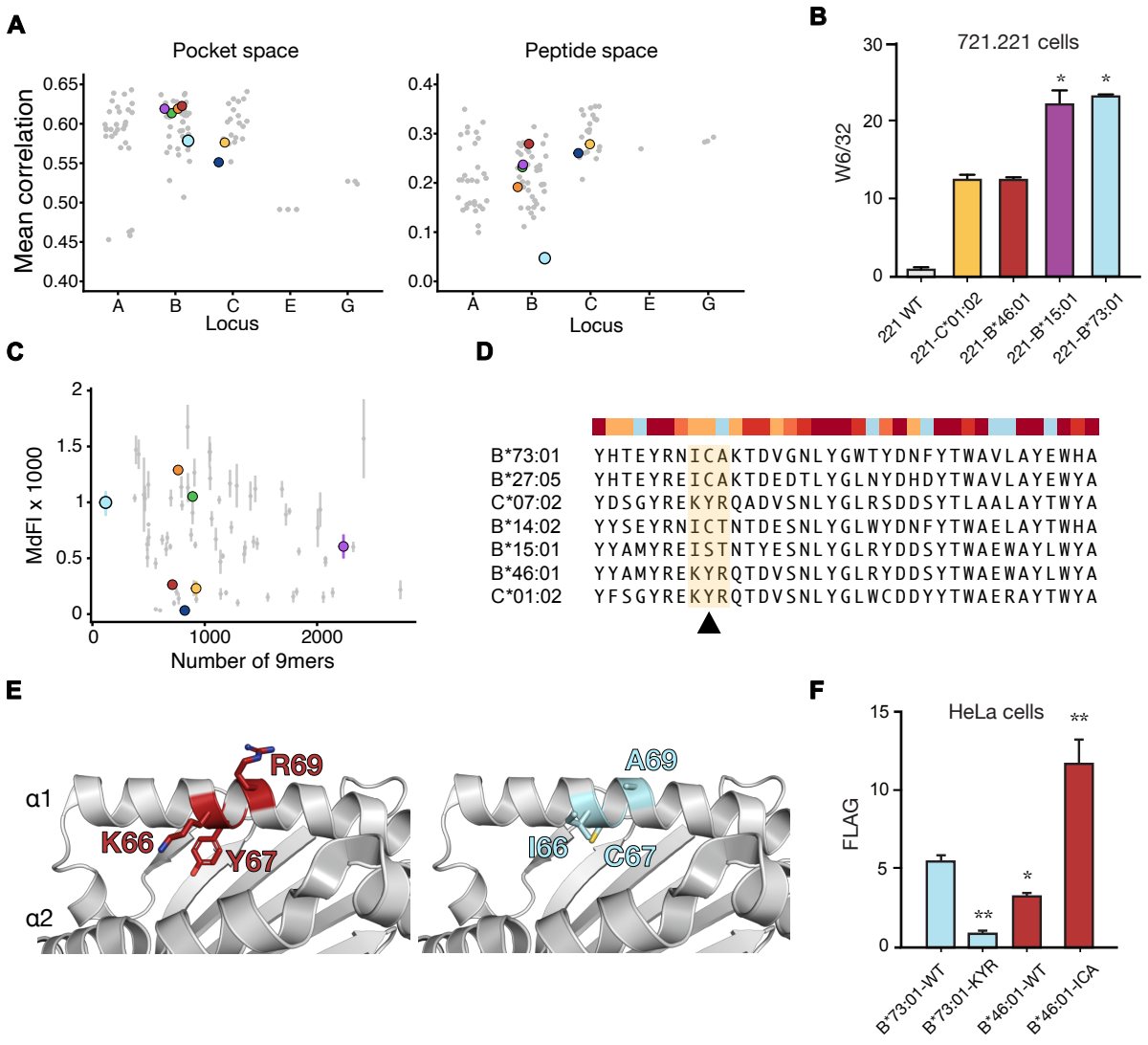

Figure 2.2 - HLA-B*73:01 combines a size-restricted and unique peptidome with above average cell-surface expression.

Figure 2.2 Continued
A. The mean correlation coefficient (Pearson) for each allele with every other allele in pocket space and peptide space stratified by HLA locus. Data includes the HLA-B*73:01 peptidome, peptides known to bind HLA-E (IEDB references), and peptides from Sarkizova *et al.* **B.** Cell-surface expression of 4 HLA class I allotypes as determined by W6/32 staining in transfected 721.221 cells (asterisks indicate expression that differs significantly from HLA-B*46:01 (*two-tailed t-test*, * indicates $p<0.005$). **C.** A scatterplot of total bound peptides by alleles in this study plotted across the cell-surface expression values of the same alleles as expressed on 721.221 cells taken from Bashirova et al. 2020. **D.** An alignment of HLA pseudo-sequences as defined in O'Donnell et al. 2020. These include residues 7, 9, 24, 45, 59, 62, 63, 66, 67, 69, 70, 73, 74, 76, 77, 80, 81, 84, 91, 95, 97, 99, 102, 114, 116, 118, 143, 147, 150, 152, 156, 158, 159, 163, 167, 171, 199 of the mature heavy chain. Highlighted in yellow is a motif known to affect the cell-surface expression of HLA-C alleles tested in panel E. **E.** A crystal structure of HLA-B*46:01 (4LCY) showing the positions and residues of the pseudo-sequence while highlighting the KYR motif in red or the ICA motif in cyan, following *in silico* mutagenesis in PyMol. **F.** Surface expression of flag-tagged wild-type (WT) HLA-B*73:01 and HLA-B*46:01 and position 66, 67 and 69 (KYR/ICA) swap mutants in HeLa cells as determined by anti-Flag Ab staining. (asterisks indicate expression that differs significantly from HLA-B*73:01-WT (*two-tailed t-test*, * indicates $p<0.05$, ** indicates $p<0.005$).

sequenced using our approach is comparable to previously published results, despite the use of over-expressed, soluble protein to generate our data (153).

To compare HLA-B*73:01 to other alleles, we then used a similar approach to Sarkizova *et al.* in which we compared HLA-B*73:01 to other alleles in both the peptides bound ("peptide space") and the variable pocket residues within the platform regions of HLA molecules ("pocket space"). These analyses revealed that the HLA-B*73:01 peptidome is most similar to the peptidomes of several other HLA class I alleles, namely HLA-B*27:05, -C*07:02, -B*14:02, -C*06:02, and -C*07:01, with Pearson correlation coefficients of 0.66, 0.47, 0.45, 0.45, and 0.41, respectively(**Figure S2.1A; left panel**). When comparing the peptide binding groove of HLA-B*73:01 to other HLA alleles, the previously mentioned alleles stood out with the addition of HLA-B*07:02, -B*07:04, -B*08:01, -B*13:02, -B*40:02, -B*40:06, -B*42:01, -B*54:01,

-B*55:01, -B*55:02, and -B*56:01 (**Figure S2.1B; right pane**l). A closer look at 9mers presented by alleles similar to HLA-B*73:01 in pocket space revealed that while they appear to have high similarity to the HLA-B*73:01 C-terminal anchor residue, the P2 anchor is what ultimately differentiates peptides bound by HLA-B*73:01. Thus, similar to comparisons made with only HLA-B*15:01, -B*46:01, and -C*01:02, the divergent nature of the HLA-B*73:01 peptidome stems from its strong preference for arginine at P2 and its interaction with the B pocket of HLA-B*73:01.

To quantify how unique HLA-B*73:01 is, we calculated the mean Pearson correlation coefficients for presented 9mers by comparing every allele to every other allele (**Figure 2.2A**). Interestingly, while HLA-B*73:01 was not the most different in pocket space, it was the most different in peptide space. These results are concordant with those found using only the 4 different alleles from our data – while the peptide binding groove of HLA-B*73:01 is an important factor for determining which peptides HLA-B*73:01 can present, other factors such as availability of peptides or distinct interactions with the peptide-loading complex (PLC) machinery may also play a role. In comparing the diversity of 9mers at each position more broadly, we noted that B2705 has even greater restriction at P2 than HLA-B*73:01 (**Figure S2.1D**). It is known that the KK10 peptide bound by B2705 of HIV origin shows signatures of escape mutations at the P2 position, mutating an arginine anchor to a lysine in order to escape B2705-mediated T-cell surveillance (154). More diversity in the HLA-B*73:01 P2 anchor position may allow HLA-B*73:01 to be more tolerant of such escape mutations relative to B2705. Indeed, our data suggest that 1% of HLA-B*73:01 presented peptides contain a lysine at

P2, whereas only 0.2% of such peptides bind to HLA-B*27:05. In summary, when comparing HLA-B*73:01 to other common alleles, HLA-B*73:01 remains unique in the peptides it can present, the relative diversity of peptides presented, and diversity of residues preferred at peptide anchor positions.

2.4 HLA-B*73:01 combines a size-restricted and unique peptidome with above average cell surface expression

Recent work suggests that peptide binding promiscuity of a given HLA class I peptidome is inversely correlated with its cell surface expression (155). Considering the small size of the HLA-B*73:01 peptide repertoire, we hypothesized that it would have comparably high cell surface expression levels relative to other alleles. In order to test this, we first compared the cell surface expression of HLA-B*73:01 relative to HLA-B*15:01, -B*46:01, and -C*01:02 expressed on 721.221 cells (which lack endogenous expression of HLA class I). Using just four alleles, we find that HLA-B*73:01 is the highest expressed allele, being expressed at a significantly higher level than both HLA-C*01:02 and -B*46:01, but at an approximately equal level to HLA-B*15:01 (**Figure 2.2B**). However, it was thought-provoking that we did not see the same inverse relationship previously described between peptidome promiscuity and cell surface expression.

In order to further investigate this unexpected finding, we incorporated cell surface expression data from a panel of 97 HLA alleles transfected into 721.221 cells (156). In agreement with comparisons to just HLA-B*15:01, -B*56:01, and -C*01:02, we found that the cell surface

expression level of HLA-B*73:01 was indeed above average (MFI of 989 relative to a mean 713.5; **Figure 2.2C**). To determine whether this was in line with expectations based on peptidome promiscuity, we decided to quantify the relationship between cell surface expression and the total number of presented 9mers by HLA-B*73:01 and alleles from Sarkizova et al. However, we found that although there was an inverse correlation between the two variables, the relationship was not significant ($p = 0.183$), suggesting that although peptidome promiscuity may contribute to cell surface expression levels, it is insufficient as a predictor by itself. Stratifying the data by HLA locus produced similar results. Thus, more work is required to dissect the complex relationship between cell surface expression and peptidome diversity in human HLA class I.

Based on these results, we considered other factors that may contribute to HLA-B*73:01 cell surface expression. The low cell surface expression of HLA-C*01:02 is consistent with previous studies showing that alleles of the HLA-C locus are expressed at a significantly lower level at the cell surface than alleles of the HLA-A or HLA-B loci (157,158). Contributing to this effect is the KYRV motif (**Figure 2.2D**), a four amino acid motif at positions 66, 67, 69 and 76 in the α1 helix that is present in all HLA-C alleles, but not in HLA-A or most HLA-B alleles (159). Because the gene conversion that created HLA-B*46:01 inserted the KYRV motif into the backbone of HLA-B*46:01, we hypothesized that this may contribute to the difference in expression between HLA-B*73:01 and HLA-B*46:01. To test the hypothesis, we examined the cell surface expression of swap mutants in HeLa cells where the residues at positions 66, 67 and 69 (both alleles encode Val76) present in HLA-B*46:01 were replaced with those found in HLA-

B*73:01 and vice versa (**Figure 2.2E**). Consistent with the results from 721.221 cells, we found that wild-type HLA-B*73:01 is expressed at a significantly higher level than HLA-B*46:01 (**Figure 2.2E**). Furthermore, we found that while the HLA-B*73:01-KYR mutant resulted in significantly less cell surface expression relative to WT, the HLA-B*46:01-ICA mutant resulted in significantly more cell surface expression (**Figure 2.2E**). These data are consistent with the hypothesis that the KYRV motif plays a significant role in determining the divergent levels of cell surface expression between these two C1+ HLA-B alleles in addition to peptidome promiscuity. Thus, while HLA-B*73:01 and HLA-B*46:01 both encode the C1 epitope and provide ligands for KIR2DL3, they differ considerably in their α1 domain sequence, which not only correlates with their distinctive peptidomes, but also with their divergent cell surface expression levels.

2.5 The crystal structure of HLA-B*73:01 bound to a 10mer reveals an unusual C-terminal bulged peptide extension

The exceptional number of polymorphisms between class I HLA alleles is what largely confers upon them their distinct antigen presenting abilities. With respect to HLA-B*73:01, we wondered how its molecular structure contributes to its restricted peptide binding repertoire. To address this question, we solved the crystal structure of HLA-B*73:01 in complex with KIR2DL2, loaded with a 10mer peptide, NRFAGFGIGL (KP1), to 2.9 Å resolution (**Table 1**). A comparison of crystal structures of HLA-C*07:02 (88,160) loaded with the same peptide that crystallized alone, in complex with KIR2DL2, or in complex with KIR2DL3 shows minimal differences in the peptide backbone and sidechains, implying that binding by KIR2DL2 does not

modify peptide positioning when presented by C0702. Therefore, we extend this observation to our analysis of HLA-B*73:01 as well; that is that KIR2DL2 binding does not induce a conformational change in KP1 presented by HLA-B*73:01 (**Figure S2.4**), although this assumption has not been experimentally verified.

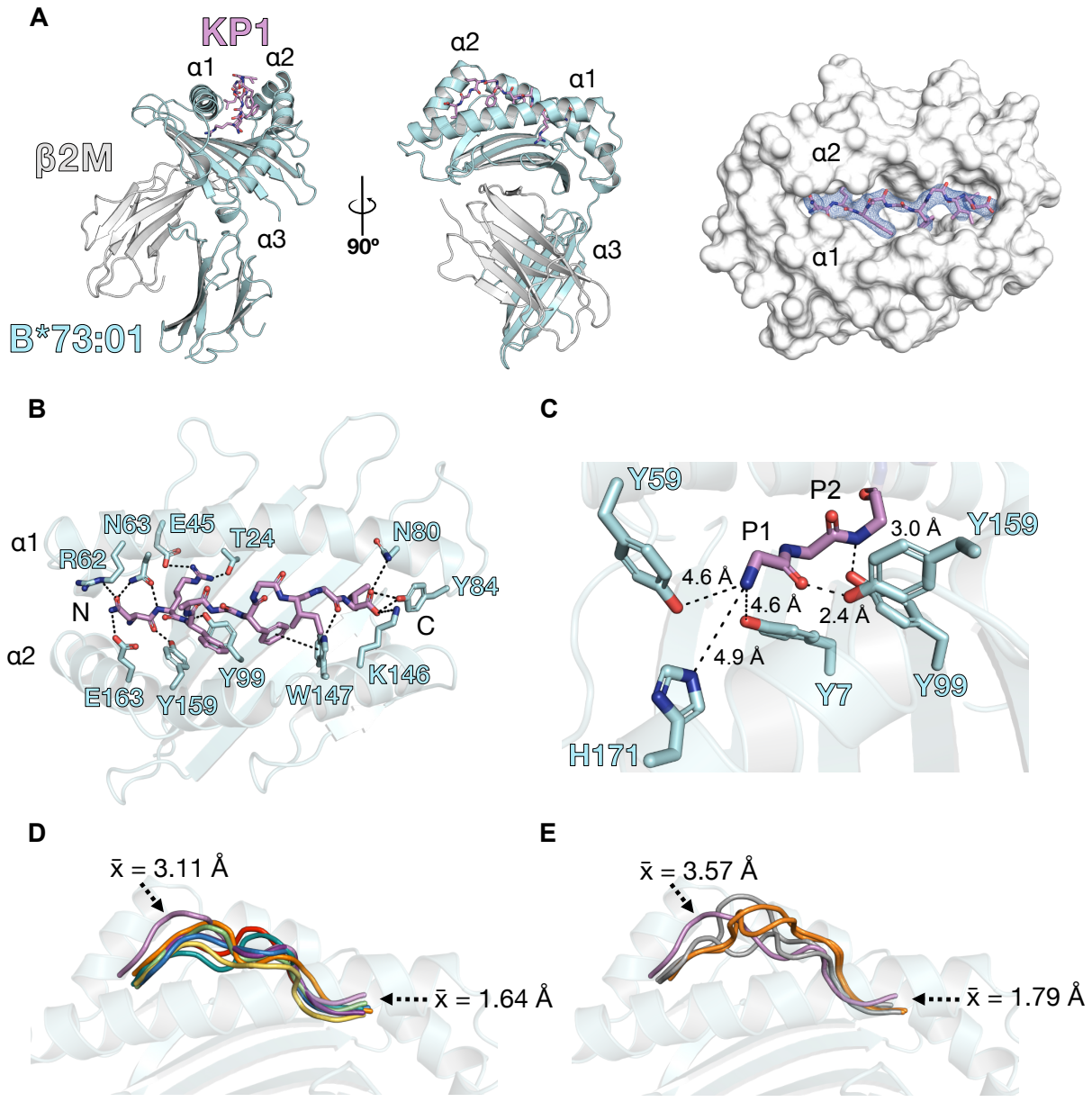

Figure 2.3 - The crystal structure of HLA-B*73:01 bound to a 10mer reveals an unusual C-terminal bulged peptide extension.

Figure 2.3 Continued
A. The crystal structure of HLA-B*73:01 was determined and resolved to 2.9 Å in complex with peptide KP1 (NRFAGFGIGL) and KIR2DL2 (not shown). The density within the peptide binding groove gives us high confidence that the position of KP1 is correct. **B.** The hydrogen bonding network between HLA-B*73:01 and its KP1 peptide ligand. Included as well are pi-interacting residues F6 of KP1 and W147 of HLA-B*73:01.**C.** A close up view of the backbone of KP1 at the first two amino acid residues and the distances from the amine group at the N-terminus of KP1 to other nearby polar atoms. Distances were calculated in PyMol. **D.** KP1 bulges out at the C-terminal end of the peptide over the F pocket of HLA-B*73:01 in a manner not seen in other alleles presenting a variety of 9mers. Colors correspond to allele colors as in Figure 1 and Figure S1. Distances were calculated in PyMol as the distance between alpha carbons at the second to last or first residue within each peptide ligand. **E.** Comparing even longer peptides, KP1 still stands out. Distances were calculated as in panel D.

Our model of HLA-B*73:01 shows that it adopts an overall conformation similar to that of other class I HLA alleles (**Figure 2.3A; left and middle panels**). Additionally, based on unoccupied electron density, we can confidently place the KP1 peptide and its sidechains (**Figure 2.3A; right panel**). The network of bonds that secure KP1 within the HLA-B*73:01 binding pocket is biased towards the N- and C-terminal ends of the peptide ligand — aside from a pi-stacking interaction between W147 of HLA-B*73:01 and F6 of KP1 — as is the case for most other class I HLA molecules (19) (**Figure 2.3B**). This bias is also consistent with trends seen in the peptide-binding motif where preferences are biased towards P1, P2, and PΩ (**Figure 2.1C**). A unique residue found in HLA-B*73:01 relative to other B alleles (aside from HLA-B*46:01) is N80, an essential residue of the C1 epitope that confers reactivity to KIR2DL2/3. Interestingly, N80 is also within close enough proximity (3.7Å) to the carboxyl end of KP1 to contribute a weak, electrostatic hydrogen bond.

Unlike most HLA alleles, HLA-B*73:01 has a slight preference for an asparagine at P1 of the peptide (**Figure 2.1C**). In our model, the asparagine at P1 of KP1 is able to utilize R62 and E163

Table 1 - Crystal structure data collection and refinement statistics

Statistics for the highest resolution shell are shown in parentheses

	HLA-B*73:01 with KP1 and KIR2DL2	B2M-Patr-F	B2M-Mamu-F
Wavelength			
Resolution range	54.3 - 2.9 (3.004 - 2.9)	55.74 - 2.81 (2.911 - 2.81)	39.61 - 2.11 (2.185 - 2.11)
Space group	P 43 21 2	C 2 2 21	I 1 2 1
Unit cell	92.69 92.69 200.98 90 90 90	90.797 111.477 85.272 90 90 90	40.87 92.6097 130.597 90 93.3853 90
Total reflections	545044 (20136)	119267 (8184)	180490 (5123)
Unique reflections	20180 (1973)	10756 (971)	25667 (2631)
Multiplicity	22.4 (16.8)	11.1 (5.7)	5.1 (3.5)
Completeness (%)	99.97 (100.00)	98.73 (90.99)	91.69 (94.50)
Mean I/sigma(I)	11.9 (0.9)	14.5 (1.7)	5.0 (0.6)
Wilson B-factor	68.74	65.48	27.14
R-merge	0.199 (3.260)	0.121 (0.798)	0.209 (2.198)
R-meas	0.203 (3.362)	0.126 (0.879)	0.233 (2.601)
R-pim	0.043 (0.813)	0.037 (0.359)	0.101 (1.347)
CC1/2	1.00 (0.411)	0.999 (0.817)	0.988 (0.287)
Reflections used in refinement	20177 (1973)	10736 (970)	25662 (2631)
Reflections used for R-free	999 (94)	508 (35)	1199 (120)
R-work	0.2225 (0.3146)	0.2315 (0.3626)	0.2439 (0.3282)
R-free	0.2551 (0.3436)	0.2950 (0.4615)	0.2958 (0.3922)
Number of non-hydrogen atoms	4764	3137	3060
Macromolecules	4645	3040	3060
Ligands	59	37	0
Solvent	60	60	0
Protein residues	588	382	382
RMS(bonds)	0.003	0.002	0.008
RMS(angles)	0.6	0.52	1.01
Ramachandran favored (%)	95.33	96.28	96.51
Ramachandran allowed (%)	4.67	3.46	3.23
Ramachandran outliers (%)	0	0.27	0.27
Rotamer outliers (%)	3.08	1.62	2.25
Clashscore	7.7	7.47	7.96
Average B-factor	70.86	65.34	37.52
Macromolecules	70.83	65.15	37.52
Ligands	80.78	84.85	
Solvent	63.27	62.96	
Number of TLS groups	15	18	10

on either side of the peptide binding groove in HLA-B*73:01 (**Figure 2.3B**), providing a potential explanation of this preference. Intriguingly, a closer investigation of the bonding network between the backbone of KP1 and HLA-B*73:01 at the N-terminus of the peptide revealed that two highly conserved tyrosines are altered in HLA-B*73:01 (**Figure 2.3C**). In most class I HLA molecules, tyrosines at positions 7, 99, 159, and 171 coordinate the amine group of the N-terminus of their peptide ligands (161). However, in HLA-B*73:01, tyrosine 171 is replaced by histidine, a feature shared with a subset of HLA-B alleles. The distance between the amine group at the N-terminus of KP1 and Y7 and H171 is likely too far (greater than 4Å) to form any relevant hydrogen bonds. Consequently, this shifts the N-terminus of KP1 up out of the pocket by an average of 1.64 Å relative to what has been observed for other 9mers (**Figure 2.3D; left panel**) and an average of 1.79 Å relative to other selected 10mers, an 11mer, and a 15mer (**Figure 2.3D; right panel**). This is in contrast to what is seen for alleles HLA-B*15:01, -C*01:02, and -B*27:05, where each allele has both Y7 and Y171 within reasonable distance to form hydrogen bonds with their respective peptide ligands (**Figure S2.3**).

HLA-B*14:02, which stood out as being very similar to HLA-B*73:01 in our analysis of HLA allele peptide and pocket space (**Figure S2.1**), also carries H171, and has a similar preference for asparagine or aspartic acid at P1 of its peptide ligands. Previous studies investigating peptide binding to three alleles of HLA-B*51 that carry either H171 or Y171 show that while the difference had minimal effect on peptide binding, they were able to generate a single cytotoxic T cell clone raised against HLA-B:51:01 (H171) that failed to lyse target cells expressing HLA-B*51:02 (Y171) (162,163) Thus, the A pocket of HLA-B*73:01 carries H171 instead of Y171,

which likely alters the conformation of bound peptides enough to potentially affect T cell responses, but not its peptide repertoire.

Now that we had a model of peptide-bound HLA-B*73:01, we decided to revisit our hypothesis that as HLA-B*73:01-presented peptides increase in length, they bulge out from the peptide binding groove (**Figure S2.2A**). When comparing our model to seven different class I HLA alleles, each in complex with a peptide 9mer, a bulge is apparent in KP1 near the C-terminus of the peptide, where the Cα of P8 is shifted upward out of the peptide binding groove by an average of 3.11Å (**Figure 2.3D**). When compared with structures of class I HLA complexed with five 10mers, an 11mer, and a 15mer, KP1 is shifted even further out of the peptide groove, on average 3.57Å. Together, these data suggest that the peptide binding mode of peptides presented by HLA-B*73:01 is unique relative to most other alleles in how it anchors its N-terminus within the A-pocket of the peptide binding groove and its ability to adapt peptide of longer lengths to its binding pocket by promoting a bulged conformation near the C-terminal end of the peptide.

2.6 The B and F pockets of HLA-B*73:01 illuminate constraints and flexibility in peptide binding

HLA-B*73:01 is part of an ancient allelic lineage of HLA-B alleles and is the only member of this lineage in humans (43). Interestingly, multiple alleles (HLA-B*27:05, -C*07:02, -C*06:02, and -B*14:02), including HLA-B*73:01, have convergently evolved to present similar binding motifs to. An alignment of MHC pseudosequences (**see Chapter 2.9 - Materials & Methods**) shows that the B pocket of HLA-B*73:01 uses almost the same exact residues as those of HLA-

B*27:05 and -B*14:02 (**Figure 2.4A**). An investigation of our structural model of HLA-B*73:01 shows that E45 and T24 are likely very important for anchoring the P2 arginine within the B pocket (**Figure 2.4B**). This is also similar to what is seen in structures of HLA-B*27:05 and -B*14:02 (**Figure S2.5A**). Notably, a previous study performing quantitative peptide binding studies tested dozens of different peptides and their ability to stabilize the HLA-B*73:01-β2M-peptide trimeric complex in solution (143). Surprisingly, although the majority of the peptides that bound most strongly contained an arginine at P2, the highest affinity binder actually contained a glutamic acid at P2. To reconcile this with our data, we noticed that HLA-B*73:01 also carries H9 and K70, two residues which are not involved in anchoring the peptides with a P2 arginine. HLA-B*40:02 is a B allele that strongly prefers a glutamic acid anchor at P2 and also carries H7 and N70, both of which form hydrogen bonds with the peptide anchor residue (**Figure S2.5A**). Thus, it is possible that some peptides containing a glutamic acid at P2 can stabilize HLA-B*73:01 as well. Indeed, our peptidome elution data also contains 11 peptides with a glutamic acid at P2, 8 of them between the lengths of 9-11 amino acids. However, it should be noted that while these residues are also found in HLA-B*27:05, to our knowledge there are no known peptide ligands that stabilize HLA-B*27:05 with a glutamic acid anchored at P2. Thus, HLA-B*73:01 may, in rare cases, be able to adopt its peptide binding groove to present peptides anchored at P2 with either strong basic or strong acidic properties.

In contrast, alleles HLA-C*07:02 and -C*06:02 have actually evolved a means of accommodating both arginines and tyrosines to be anchored from the P2 position, a feature likely not possible for HLA-B*73:01 (**Figure S2.1C**). Although we could identify 6 peptides with a

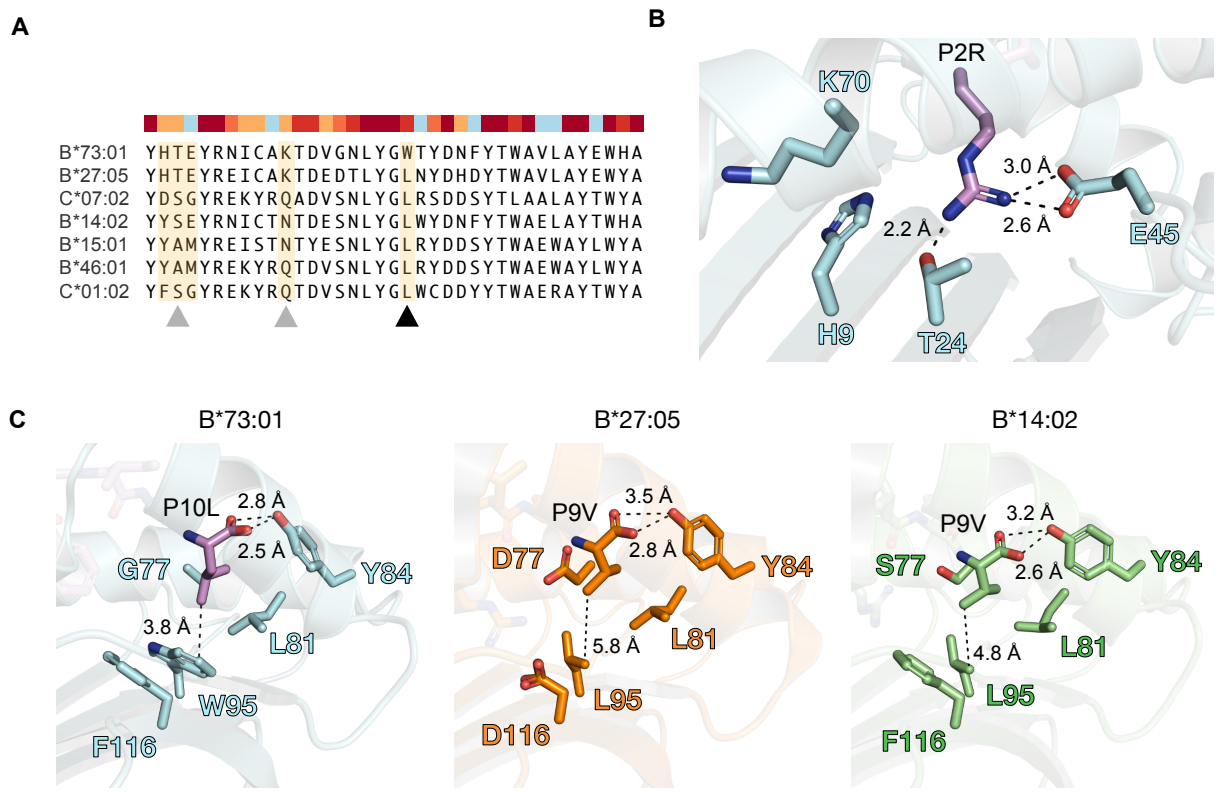

Figure 2.4 - The B and F pockets of B7301 illuminate constraints and flexibility in peptide binding.
A. An alignment of HLA pseudo-sequences as defined in O'Donnell et al. 2020 (174) to highlight residues of the B pocket (grey arrows) and residue 95 of the F pocket (black arrow). **B.** A close up view of the B pocket of HLA-B*73:01 shows that residues T24 and E45 are important for anchoring the arginine from P2 of KP1. H9 and K70 could play a role in anchoring more acidic side chains. **C.** The F pocket of HLA-B*73:01. contains a distinct W95 substitution not seen in many other alleles and explains why HLA-B*73:01. can only accommodate small hydrophobic C-terminal anchors.

tyrosine at P2, one is a 7mer, 3 are 8mers, one is a 16mer, and the other, a 35mer. A closer investigation of the structures of the B pockets of HLA-C*07:02 and -C*06:02 reveal that Y67 is likely very important in stabilizing a tyrosine anchor at P2 while the use of negatively charged D9 (in contrast to H9 in HLA-B*73:01) can stabilize an arginine anchor at P2 (**Figure S2.5B**).

While HLA-B*73:01 and -B*27:05 share identical B pocket architectures, they present few shared peptide sequences even when purified from the same cell. This may be due to the different architecture of their F pockets. In particular, HLA-B*73:01 carries a rare F pocket residue, W95, that only appears in roughly 12% of alleles in the set of alleles included in Sarkizova *et al.*, while roughly 43% contain a leucine at this position. W95, which sits directly beneath the C-terminal peptide anchor residue, likely contributes to the increased residue diversity seen at this position in HLA-B*73:01 (**Figure 2.1E**) by increasing surface area used by small hydrophobic residues to escape water molecules. HLA-B*27:05, on the other hand, has evolved a separate means of enhancing the diversity of its C-terminal peptide anchor, with a slightly negatively charged F pocket that can also accommodate positively charged anchors (89). This combination of relative constraint within the B pocket combined with hydrophobic plasticity within the F pocket represents a unique tradeoff for HLA-B*73:01 relative other alleles and contributes to its unique peptide repertoire, distinct from even HLA-B*27:05.

2.7 KIR2DL2 utilizes a unique docking angle to engage with HLA-B*73:01

HLA-B*73:01 and HLA-B*46:01 are distinct from all other human B alleles in that they carry the C1 epitope which confers reactivity to KIR2DL2/3. However, the only KIR2DL2/L3 complex structures determined to date were determined in complex with HLA-C, specifically HLA-C*03:04 and -C*07:02. Here, we determined the first complex structure of an HLA-B allele with KIR2DL2, thus giving us the first look at what differences exist relative to structures in complex with HLA-C. An overview of the complex shows that engagement of HLA-B*73:01 by KIR2DL2 is similar to other KIR2D complex structures (**Figure 2.5A**). KIR2DL2 binds

directly over the F pocket of the heavy chain, with its D1 and D2 domains positioned over the α2 and α1 helices, respectively, and makes no contact with the light chain.

Interestingly, when compared to other previously solved structures, the docking angle of KIR2DL2 onto HLA-B*73:01 is shifted ~15º towards the α2 helix (**Figure 2.5B**). This shift corresponds to a binding footprint that also appears biased towards the α2 helix (**Figure 2.5C** and **Figure 2.5D**). For example, whereas interactions between KIR2DL2 and HLA-C*07:02 along residues 72-84 of HLA-C*07:02 consist of 4 hydrogen bonds, interactions between KIR2DL2 and HLA-B*73:01 along the corresponding residues consists of only 2 hydrogen bonds. In contrast, interactions along residues 145-151 consist of a mix of hydrogen bonds and salt bridges that appear to be utilized in both structures. A comparison of HLA residues known to interact with KIR2D molecules reveals that the only difference between HLA-B*73:01 and HLA-B*46:01 and other HLA-C alleles is at residue 69, where HLA-B*73:01 encodes an alanine as opposed to an arginine. R69 has been shown to interact with E21 of KIR2DL molecules (93). However, E21A mutations in KIR2DL2 and KIR2DL3 did not impact binding to HLA-C*07:02 in a previous study (88). Thus, amino acid substitutions at residues known to directly interact with KIR2DL2 and KIR2DL3 likely do not explain the difference in how KIR2DL2 engages HLA-B*73:01 relative to previously solved complex structures.

A close up view of the buried surface area between KIR2DL2 and the KP1 peptide backbone shows that the partially positively charged amine group of the second to last residue within KP1 has also shifted towards the α2 helix of HLA-B*73:01 and that Q71 of KIR2DL2 tracks with it

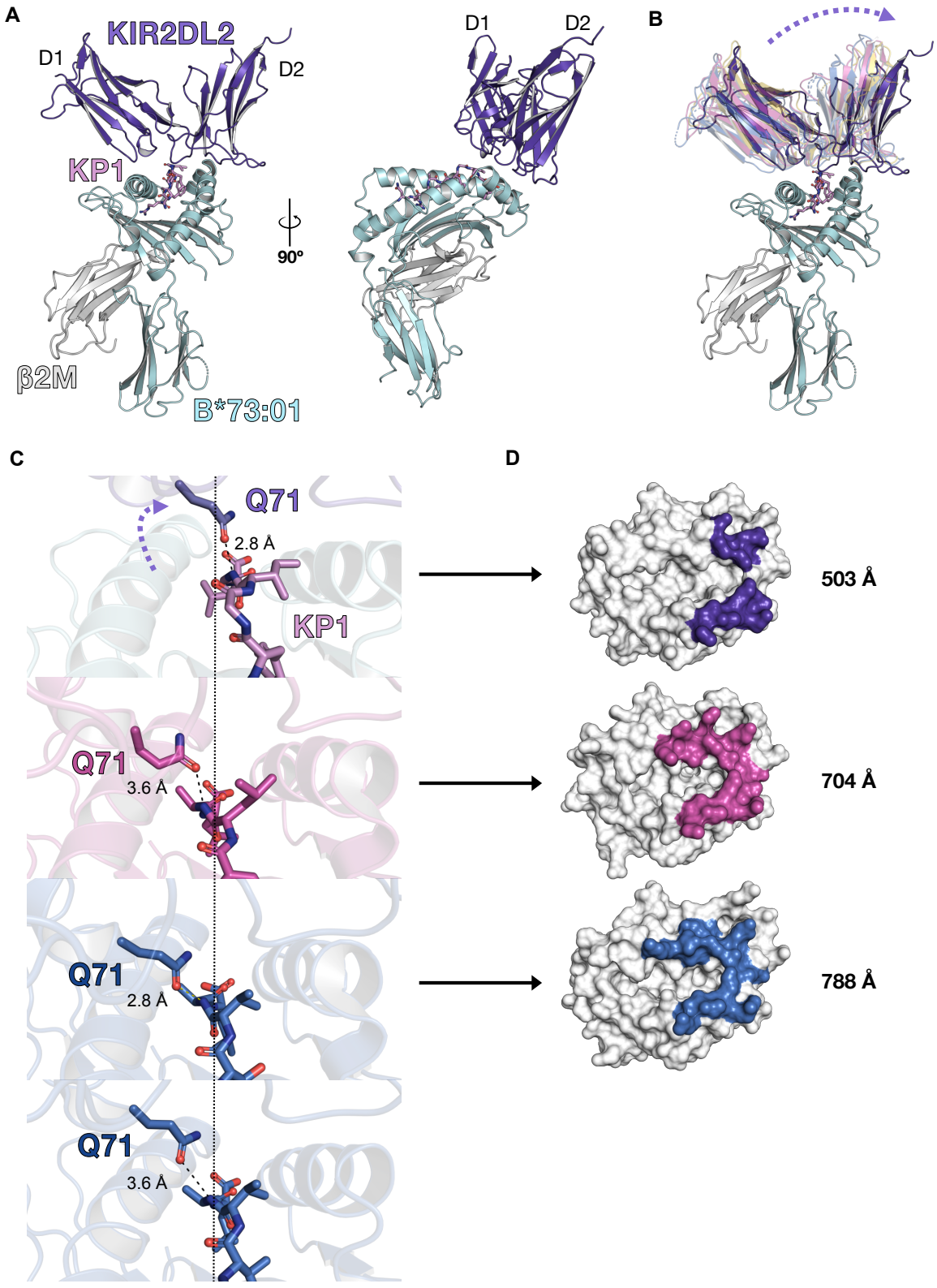

Figure 2.5 - The complex structure of B7301 and KIR2DL2 indicates the use of a unique docking angle.

Figure 2.5 Continued
A. A structural overview of the complex structure shows a similar binding mode to other solved KIR2DL2 complex structures directly over the F pocket of HLA-B*73:01. **B.** However, unlike previously solved structures, the angle of KIR2DL2 as it sits over HLA-B*73:01 is slightly tilted relative to other solved structures. **C.** Similar to previously solved KIR2DL2 complex structures, Q71of KIR2DL2 forms a hydrogen bond with the presented peptide backbone as P8 of the peptide backbone. This hydrogen bond could be the explanation for the altered docking angle of KIR2DL2 onto HLA-B*73:01. B7301 is shown on top, beneath it is HLA-C*03:04 in complex with KIR2DL2 in magenta, beneath that is HLA-C*07:02 in complex with KIR2DL2 in marine blue, and beneath that is HLA-C*07:02 in complex with KIR2DL3, also in marine blue. **D.** The KIR binding footprint shown as residues contributing significantly to buried surface area between KIR2DL2 and the corresponding HLA compared between solved structures is highly similar. The low affinity determined for KIR2DL2 and HLA-B*73:01. correlates with the relatively low buried surface area between them. The structures used correspond to the same structures in panel C.

(**Figure 2.5C**). In other complex structures, Q71 is closer to the α1 helices of their corresponding HLA ligands. Because the KP1 peptide is likely presented this way regardless of binding to KIR2DL2, the necessity of Q71 to track the backbone of KP1 to maintain an interaction with the pMHC complex may contribute to the unique docking angle of KIR2DL2 onto HLA-B*73:01.

The affinity of KIRs for peptide ligands depends on both the stability of the peptide-MHC complex and binding accessibility determined mainly by the PΩ-1 and PΩ-2 residues of presented peptides (90,164,165). Surprisingly, upon investigation of the KIR2DL2 interaction with HLA-B*73:01 by SPR, we found that the affinity of the complex is likely greater than 100 μM (**Figure S2.6**). It's unclear why the apparent affinity is so low; KP1 does not contain residues commonly considered to prevent KIR binding (I at P8, G at P9). However, it may be that peptide length influences KIR binding and therefore the KP1 peptide partially inhibits binding of HLA-B*73:01 to KIR2DL2. This hypothesis is consistent with the lengths of KIR2DL3-permissive peptides identified for HLA-B*46:01 (31). As previously detailed, longer peptides may result in

a protrusion above the F pocket, as is observed for KP1. This positioning overlaps with the KIR binding site, likely further impeding binding. However, with only a single model, we cannot with certainty whether these unique aspects of the HLA-B*73:01:2DL2 interface is due to HLA-B*73:01, the KP1 peptide, the low affinity of the interaction, or a combination of all three factors. Together, these results suggest that KIR2DL2 engages HLA-B*73:01 using a distinct docking angle, possibly in order to compensate for the excessive bulge in KP1 near the KIR2DL2 binding site. This allows Q71 of KIR2DL2 to form a hydrogen bond with the peptide backbone of KP1, an interaction conserved in all KIR2D structures determined to date, suggesting that this hydrogen bond is likely essential.

2.8 Discussion

This study characterizes the highest resolution and most comprehensive peptidome of HLA-B*73:01 to date. Compared to the most common worldwide HLA alleles, the HLA-B*73:01 peptidome is highly unique, differentiated by its preference for asparagine at P1, arginine at P2, and small hydrophobic residues at PΩ. HLA-B*73:01 also presents peptides with an unusual length distribution, facilitated by an ability to allow peptides to bulge out from the peptide binding groove while maintaining reasonable protein complex stability. The structure of HLA-B*73:01 helps explain some of its rare peptide presenting properties such as its unique A and F pockets. And finally, while the complex structure of HLA-B*73:01 with KIR2DL2 demonstrates overlapping binding footprints, the docking angle of KIR2DL2 onto HLA-B*73:01 appears to be unique compared to KIR2D complex structures solved with HLA-C*01:02, -C*03:04, and -C*07:02.

Highly compact peptidomes, like that of HLA-B*73:01, are thought to correlate with protection against one or a number of closely related pathogens (155). To our knowledge, there are no epidemiologic studies that have identified HLA-B*73:01 as a protective factor against the development or outcome of infectious disease, likely due to the low frequency of HLA-B*73:01 in the modern human population. However, HLA alleles that encode the C1 epitope, in combination with KIR2DL3, are known to provide protection against hepatitis C virus (141). Furthermore, the high cell surface expression of HLA-B*73:01 may correlate with protection against the progression of HIV infection (166). A recent study found that HLA-B*46:01, another C1+ HLA-B allele with significantly lower cell surface expression, is associated with rapid disease progression to AIDS in Asian HIV-positive cohorts (103). The authors suggest that this is likely due to an NK cell-mediated phenotype as carriers of HLA-B*46:01 possessed NK cells with distinct transcriptional profiles skewed towards immune activation. While the transcriptional phenotypes in NK cells of HLA-B*73:01 carriers remain unexplored, it could mean that HLA-B*73:01 carriers also have NK cells endowed with unique properties, some of which may have provided our ancestors with the ability to fight off local viral infections. Thus, the molecular characteristics of HLA-B*73:01 identified in this study may correlate with protection against at least two RNA viruses. Intriguingly, and in support of our hypothesis, a recent study suggests that RNA viruses were likely the major drivers of adaptive introgression of archaic alleles from Neanderthals to modern Europeans (167).

Alternatively, individuals that combine HLA alleles encoding the C1 epitope and KIR2DL3 are also more likely to fall victim to various forms of severe malaria (168–170). In fact, the high cell surface expression of HLA-B*73:01 determined by our study relative to other C1+ HLA-B and HLA-C alleles could make carriers of HLA-B*73:01 particularly susceptible to such diseases. While underlying immune mechanisms behind this outcome remain unknown, it could be that this outcome is mediated purely by the innate NK cell response to *Plasmodium* infection. However, KIR2DL2 has also been reported to play an important role in the adaptive immune response (142). Considering that clonally expanded, terminally differentiated CD8+ T cells are also known to express KIR molecules, the association between the C1 epitope, KIR2DL3, and severe malaria could also be T cell-mediated. Furthermore, the *Plasmodium falciparum* proteome is enriched with lysine residues, which are known to destabilize HLA-B*27:05 if at the P2 position. Considering a similar preference for arginine at P2 by HLA-B*73:01 and reports that HLA-B*27:05 has a latitude-dependent distribution negatively correlated with malaria prevalence (171), HLA-B*73:01 carriers might indeed be relatively defenseless against *Plasmodium* infection. It could also explain why HLA-B*73:01 was eventually lost in ancestral human populations that remained in Africa (43).

One important caveat to consider is that HLA-B*73:01 may not have been directly selected for, but rather that the increase in its frequency in certain parts of the world are due to genetic hitchhiking as part of adaptive haplotypes. The strong linkage disequilibrium between HLA-B*73:01 and -C*15:05, for example, might suggest that it was actually HLA-C*15:05 that was advantageous and HLA-B*73:01 simply came along for the ride. Interestingly, recent work

shows that escape mutations in an HIV antigen presented by HLA-C*15:05 allow the virus to escape T cell responses and lead to poor clinical outcomes (172). Thus, if HLA-C*15:05 was once advantageous, it appears not to be in the current day context of HIV infection.

In summary, our study shows that archaic HLA-B*73:01 has a highly unique peptidome with an unusual peptide length distribution. Further, we show that HLA-B*73:01 combines a highly restricted peptidome with high cell surface expression, characteristics that make it well-suited to combat one or a number of closely-related pathogens. Lastly, we use crystallography to highlight how HLA-B*73:01 is able to present these unique peptides relative to what appear to be highly similar alleles, and how its interface with KIR2DL2 is both conserved and distinct at the same time. These functional characteristics distinguish HLA-B*73:01 from other HLA class I alleles and likely provided early modern human migrants that inherited this allele from archaic humans with a selective advantage as they colonized Europe and Asia.

2.9 Materials & Methods

Peptidome sequencing and analysis

Isolation and analysis of peptides was performed as previously described (31). Briefly, 721.221 cells were transfected with constructs encoding a soluble form of HLA-B*73:01. The HLA-peptide complexes were affinity purified from the cell supernatant on an anti-HLA Class I (W6/32 antibody) sepharose column, eluted, and the bound peptides dissociated from the HLA by denaturation. Peptides were separated from the denatured proteins by ultrafiltration and separated into ~40 fractions by reverse-phase high-performance liquid chromatography (HPLC)

as described (31). Approximately 25% of each HPLC fraction was injected into a nano-scale reverse phase liquid chromatography Eksigent nano-LC-4000 (Sciex) system. Column specifications, mobile phase solvents, and the elution gradient were as described (31). Eluted fractions were ionized using a NanoSpray III (Sciex) ion source into a Sciex TripleTOF 5600 mass spectrometer.

Exact peptide sequences generated for HLA-B*73:01 (this study) and for HLA-B*15:01, HLA-B*46:01, and HLA-C*01:02 (31) were compared to detect shared peptides using Venny (https://bioinfogp.cnb.csic.es/tools/venny/). For PCA plots, four biochemical properties (molecular weight, hydropathy index, surface area and isoelectric point) were determined for each amino acid in each peptide resulting in 4 times X variables per peptide, where X equals the peptide length (e.g. 36 variables for a peptide of length 9). These variables were then used to calculate principal components for each peptide prior to plotting individual peptides in R, colored by allele of origin. Peptides of length 8-15 were extracted from the eluted ligand data for the four HLA class I alleles. A GibbsCluster analysis (https://services.healthtech.dtu.dk/service.php?GibbsCluster-2.0) was performed on each data set to identify binding motifs for distinct clusters and remove noise from the data.

Integration with and analysis of Sarkizova *et al.* data

Curated peptide lists were downloaded from Sarkizova *et al.* (22) and compared directly to peptides eluted from HLA-B*73:01 from this study. To calculate correlation coefficients in peptide space between all alleles, amino acid occurrences of all 20 amino acids were calculated

for each position of all 9mers generating a 9 by 20 matrix for each allele. These matrices were then flattened into vectors of 180 integers for each allele and used to directly calculate the Pearson correlation coefficients between alleles of presented 9mers. For plotting, alleles were clustered by hierarchical clustering implemented in the ggcorrplot function of the ggcorrplot package (https://github.com/kassambara/ggcorrplot) in R version 4.2.1 (2022-06-23) (173). The ggcorrplot function relies on the hclust R function, which by default implements the "complete" linkage method. This method defines the distance between two clusters as the maximum distance between their individual components.

To calculate correlation coefficients between alleles in pocket space, pseudosequences for each allele were used as described previously (174). Following a multiple sequence alignment, positions within the alignment that were 100% conserved between alleles were removed, leaving only variable aligned columns. These sequences were then used to calculate 61 different descriptors of their amino acid sequences using the aaDescriptors function of the Peptides package in R and default parameters. This generated a vector of 61 values for each allele which was then used to calculate Pearson correlation coefficients between all alleles. For plotting, alleles were clustered by hierarchical clustering implemented by the ggcorrplot function of the ggcorrplot package in R.

For plotting sequence logos, the R package ggseqlogo was used (175). For calculating Shannon and Simpson diversity indices, the diversity function from the R package vegan was used (176).

All code used to analyze these data can be found on https://github.com/biophilross/thesis.

Cell surface expression experiments

721.221 cells, which lack endogenous HLA class I expression, were transfected with constructs encoding full-length HLA-B*73:01, HLA-B*46:01, HLA-B*15:01 or HLA-C*01:02 and cultured as previously described (31). Cells expressing surface HLA class I were detected by flow cytometry (Accuri C6 cytometer, BD Biosciences). Expression levels of each allele were determined from the median fluorescence intensity (MFI) of the W6/32 antibody bound to each positive staining cell. Three independent transfections with at least 50,000 cells each were performed for each allele tested.

We examined the cell surface expression of wild-type and KYR/ICA mutant 3x FLAG-tagged HLA-B*73:011 and HLA-B*46:01 in HeLa cells. Recombinant cDNA encoding amino acids 1-338 of HLA-B*73:01 and HLA-B*46:01 with an N-terminal 3X FLAG-tag were manufactured by Genscript (Piscataway, NJ). Site-directed mutagenesis was performed with the QuikChange Kit (Stratagene), according to the manufacturer's instructions, to generate the two swap KYR/ICA mutants. HeLa cells were transfected with these constructs using the Fugene transfection reagent (Promega) and cultured as previously described (101). Cells expressing FLAG-tagged HLA class I were detected by flow cytometry (Accuri C6 cytometer, BD Biosciences). Expression levels of each allele or mutant were determined from the MFI of FITC-conjugated anti-FLAG antibody bound to each positive staining cell. Three independent transfections with at least 50,000 cells each were performed for each allele tested.

Crystal structure determination and analysis

Soluble HLA-B*73:01 and HLA-B*46:01 heavy chain constructs were initially codon optimized to improve bacterial expression. To codon optimize them, the first 60 bps (20 codons) of HLA-B*73:01 or HLA-B*46:01 were replaced with synonymous codons such that the GC content of the first 60 bps was lower than 50%. These GC optimized (GCO) constructs expressed several orders of magnitude better than WT and were used to generate protein-rich inclusion bodies.

GCO HLA-B*73:01 heavy chain (mature residues 1-276) and β2M light chain constructs were expressed in *E. coli* (BL21 DE3) for 4 hours following 1 mM IPTG induction. Cells were lysed using a triton-containing lysis buffer and incubated with lysozyme and DNase for one hour at 4ºC. Following incubation, cells were sonicated prior to centrifugation at 20,000xg to pellet isolate inclusion bodies. Inclusion bodies were denatured in 6M guanidinium chloride, 50 mM Tris-HCl pH 8.0 prior to refolding and either stored at -80ºC until use or used immediately. Refolding was carried out by rapid dilution into a buffer containing 400 mM L-arginine, 100 mM Tris-HCl, 2 mM EDTA, 5 mM oxidized glutathione, and 0.5 mM reduced glutathione adjusted to pH 8.3. 7 mgs of HLA-B*73:01 heavy chain, 1.75 mgs of β2M, and 1 mg of synthetic peptide purified to greater than 90% (Genscript) were added to a 50 mL reaction at the same time and mixed at 4ºC for 60 hours. Once the refolding reaction was complete, it was dialyzed into 20 mM Tris-HCL 50 mM NaCl pH 8.0 to dilute components of the refolding buffer at least 5000x. Samples were then 0.22 μm filtered prior to purification by anion exchange chromatography using a HiTrap QFF column. Following purification by QFF, monomeric fractions were pooled,

concentrated, and purified by size exclusion chromatography using a Superdex 200 increase 10/30 column (S200).

A soluble KIR2DL2*001 construct (residues 1-225 of the mature protein), covalently linked to an N-terminal 8x HisTag, separated by a 3C protease cleavage site, was expressed in BTI-Tn-5B1-4 insect cells (High Five) cells cultured in Insect-XPRESS Protein-free Insect Cell Medium (Lonza). BestBac 2.0 linearized DNA (Expression Systems) and a construct donor plasmid (pAcGP67a) were used to transfect Sf9 cells to generate a P1 viral stock. Virus was then amplified by shaking culture in Sf9 cells. Following Ni-NTA affinity purification, KIR2DL2 was treated for 2 hours at 37ºC with EndoF (made recombinantly in the Adams Lab) prior to another round of Ni-NTA affinity purification to remove the EndoF, which does not contain a HisTag. The 8x HisTag was then removed from KIR2DL2 using 3C protease (made recombinantly in the Adams Lab). Recombinant HLA-B*73:01 and KIR2DL2 were purified separately by size exclusion chromatography (SEC) into HBS (10 mM HEPES, 150 mM NaCl, pH 7.2), mixed at a 1:1 molar ratio, and concentrated to 6-10 mg/mL. Crystals initially grew in less than 24 hours in mother liquor containing Tris-HCl pH 8.5 and 25% PEG 8000. These crystals were then crushed using Seed Bead kits (Hampton Research) and used as microseeds to grow a larger crystal in mother liquor containing Tris-HCl pH 8.75 and 20% PEG 8000, which took 1-2 weeks to grow.

Diffraction data was collected at the Advanced Photon Source at Argonne National Laboratory on beamlines 23-ID-D. CCP4, Coot, and PHENIX software suites were used for molecular replacement and model refinement. PyMol and ChimeraX were used to make figures and

measure atomic distances. PDBe PISA v1.52 [20/10/2014] (https://www.ebi.ac.uk/pdbe/pisa/) was used to define interacting residues and the amount of buried surface area between polypeptide chains.

Surface plasmon resonance experiments

Recombinant B73, B46, KIR2DL2, and KIR2DL3 were purified as described above, but without the EndoF treatment step for the KIRs. Peptide-MHC constructs were designed with a C-terminal AviTag and biotinylated *in vitro* prior to final purification steps. Biotinylated MHCs were then immobilized to approximately 400 RUs using a Biotin CAPture Kit, Series S and a Biacore 8K (Cytiva). WT KIR2DL2 and KIR2DL3 analytes were ultimately purified into SPR buffer (10 mM HEPES buffered saline, 150 mM NaCl, 0.05% tween, adjusted to pH 7.4) flown over immobilized channels from concentrations of 0.625-80 μM.

2.10 Supplementary information

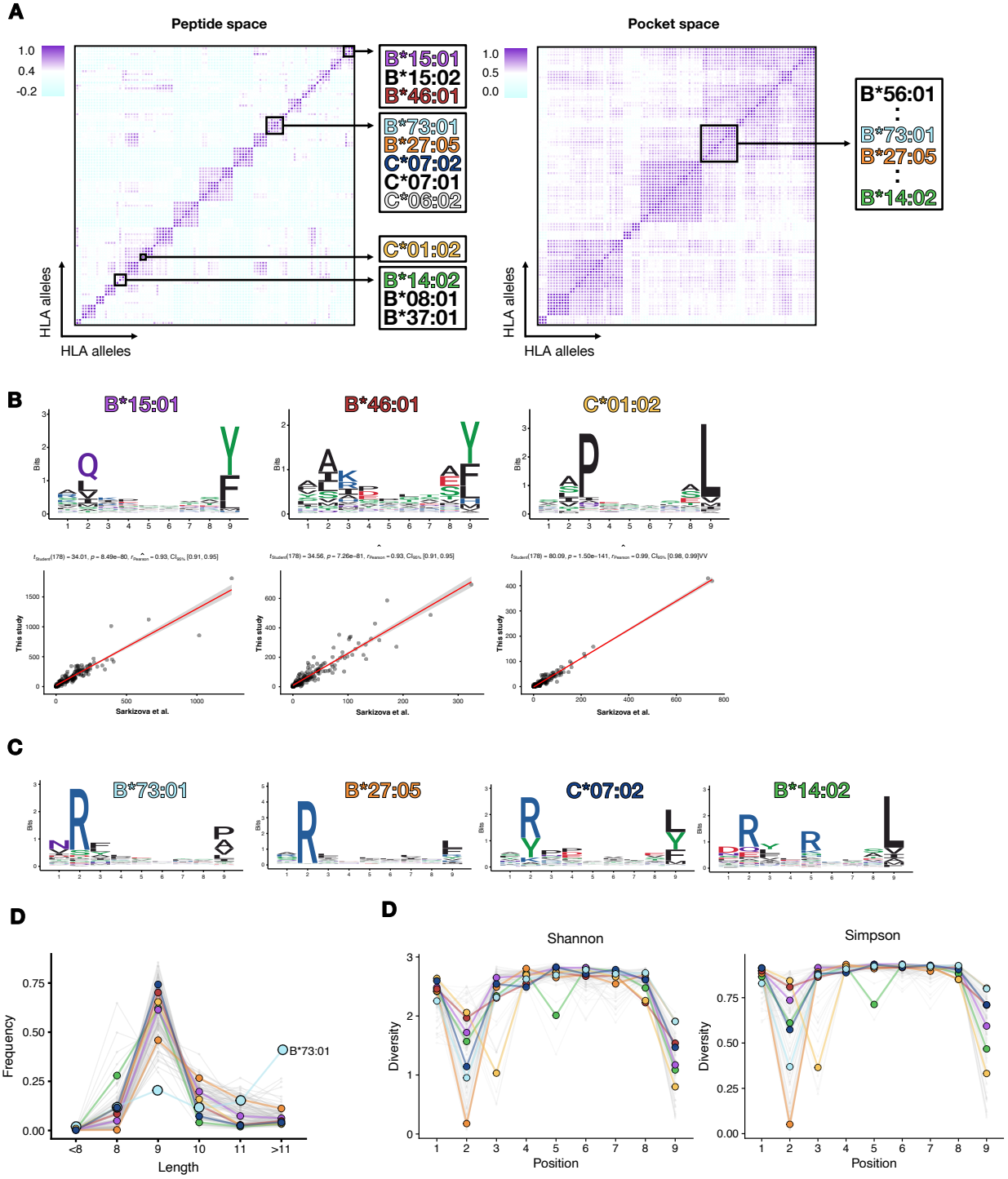

Figure S2.1 - Integration of the HLA-B*73:01 peptidome with data from Sarkizova et al.

Figure S2.1 Continuted
A. Heatmaps were generated using data from Sarkizova *et al.* was combined with eluted peptides from HLA-B*73:01 and representative peptides known to bind HLA-E (IEDB references) and then correlated (Pearson) with other peptidomes (left; peptide space) and other MHC pocket residues (right; pocket space), then clustered based on similarity to highly alleles with similar peptide binding motifs and/or similar binding grooves. Alleles with similar peptide repertoires and/or similar binding grooves are depicted on the right. **B.** The 9mer peptide binding motif profiles of HLA-B*15:01, -B*46:01, and -C*01:02 from Sarkizova *et al.* correlate positively with data generated in our previous study. **C.** Sequence logos of 9mer peptides bound by HLA-B*73:01, HLA-B*27:05, HLA-C*07:02, HLA-B*14:02, HLA-B*15:01, HLA-B*46:01, and HLA-C*01:02 are shown. The total height of the column of amino acids represents the level of conservation at that position and the residue height denotes its relative frequency. The colors of each amino acid correspond to their biochemical characteristics: acidic (red), basic (blue), hydrophobic (black) and polar (green). **D.** Length distributions of different alleles from Sarkizova *et al.* grouped into peptides less than 8, 8, 9, 10, 11, or more than 11 residues in length and including HLA-B*73:01.**E.** Shannon and Simpson diversity indices calculated for each position for each of all HLA class I allotypes included in Sarkizova *et al.* and HLA-B*73:01 from this study.

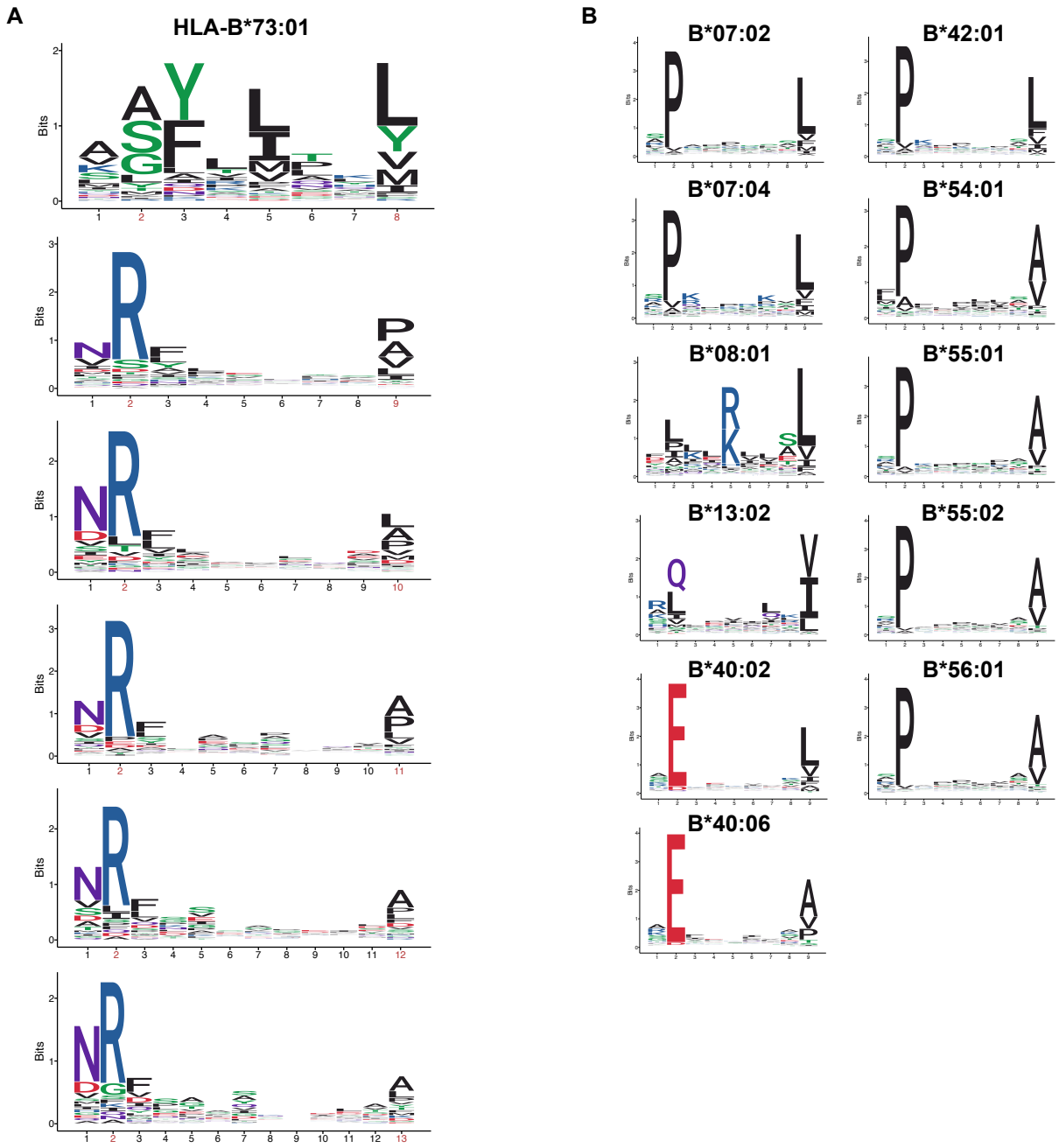

Figure S2.2 - HLA-B*73:01 peptide motif is similar at different lengths.
A. Sequence logos of all 8mers, 9mers, 10mers, 11mers, 12mers, and 13mers bound by HLA-B*73:01 derived from peptides eluted and sequenced by MS in our study. The total height of the column of amino acids represents the level of conservation at that position and the residue height denotes its relative frequency. The colors of each amino acid correspond to their biochemical characteristics: acidic (red), basic (blue), hydrophobic (black) and polar (green). **B.** 9mer binding motifs for alleles with similar binding grooves, but not necessarily similar peptides being bound.

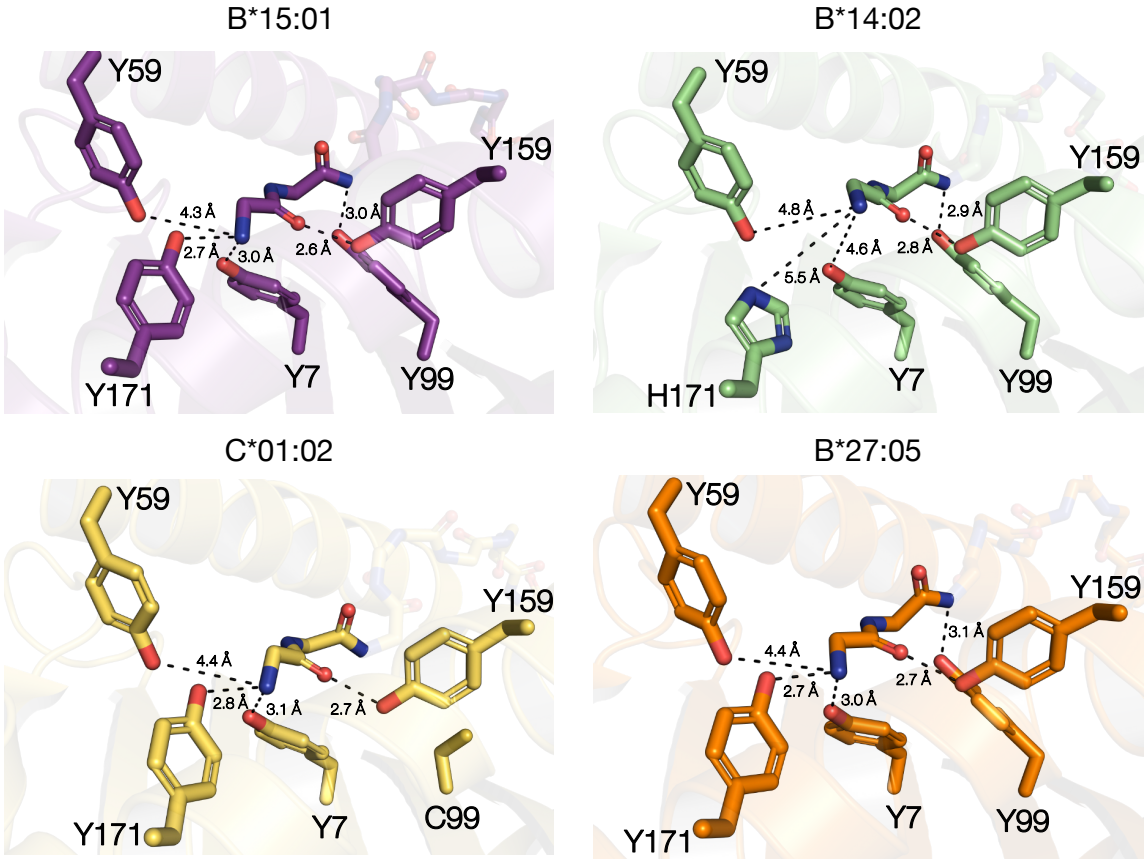

Figure S2.3 - Unique aspects of the A pocket of HLA-B*73:01 is likely due to Y171H substitution found in a subset of alleles.
A close up of the A pockets are shown for alleles HLA-B*15:01, -B*27:05, -C*01:02, and -B*14:02. Distances were calculated using PyMol.

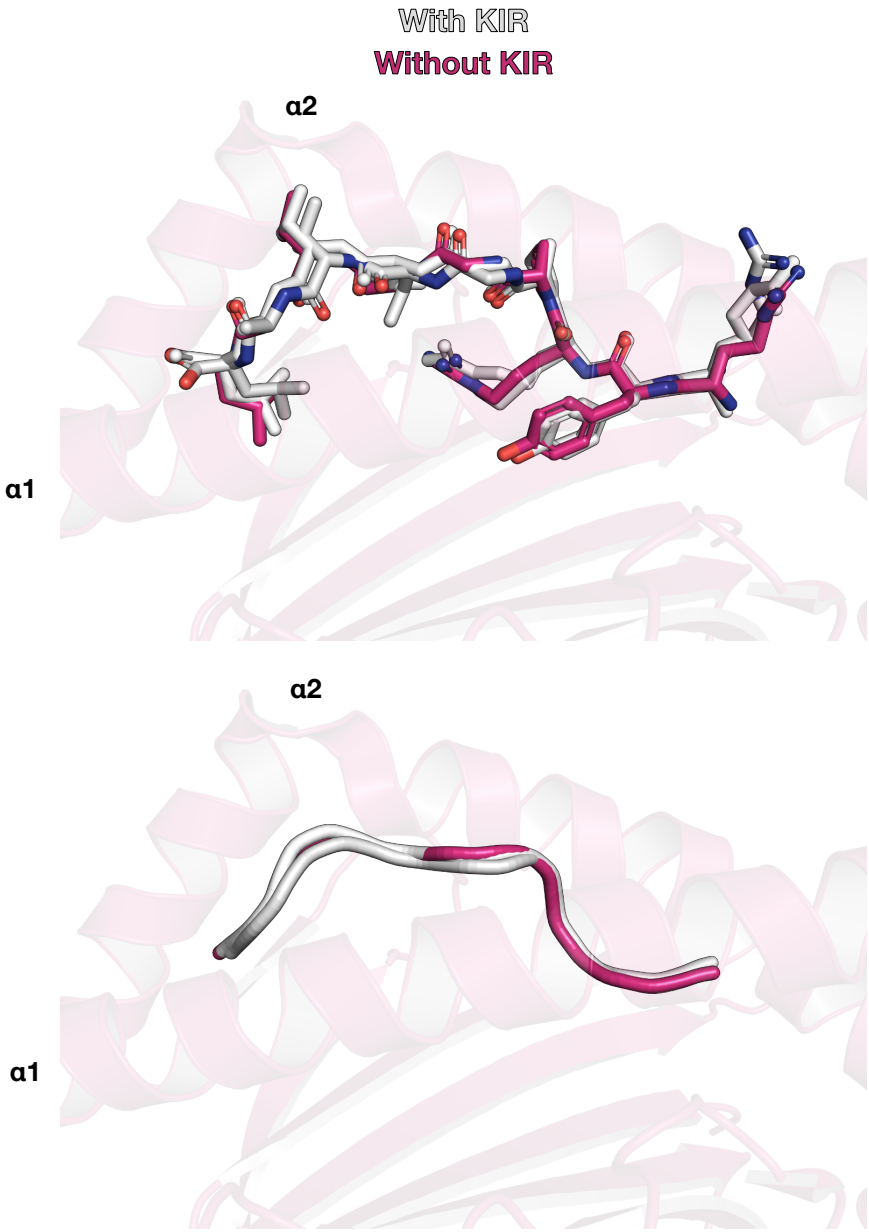

Figure S2.4 - KIR2DL2 and 2DL3 binding does not severely affect peptide presentation. Three structures of HLA-C*07:02 were aligned using β2M as a reference. In magenta is a structure of HLA-C*07:02 presenting RYRPGTVAL without a bound KIR. Two more structures in white are structures of HLA-C*07:02 presenting the same peptide, bound to KIR2DL2 or KIR2DL3.

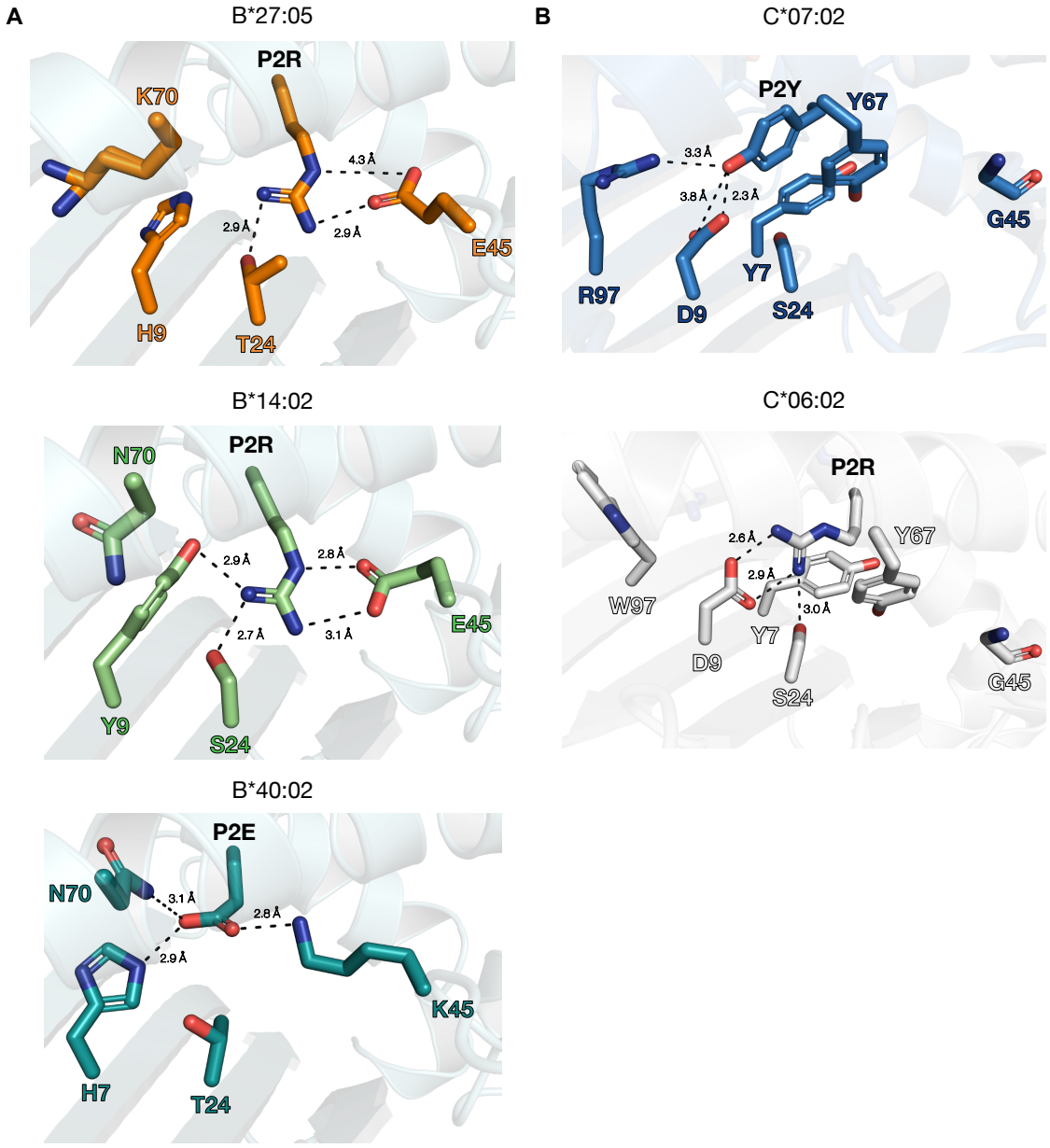

Figure S2.5 - HLA-B alleles display substantial plasticity in their B pocket.
A. A close up of B pockets of HLA-B*27:05, -B*14:02, and -B*40:02 show how other alleles use similar and different residues located on different sides of the binding pocket to accommodate positively or negatively charged anchors. **B.** HLA-C*07:02 and -C*06:02 have evolved a means of accommodating either arginine or tyrosines as P2 anchor residues, at positions that vary in HLA-B*73:01.

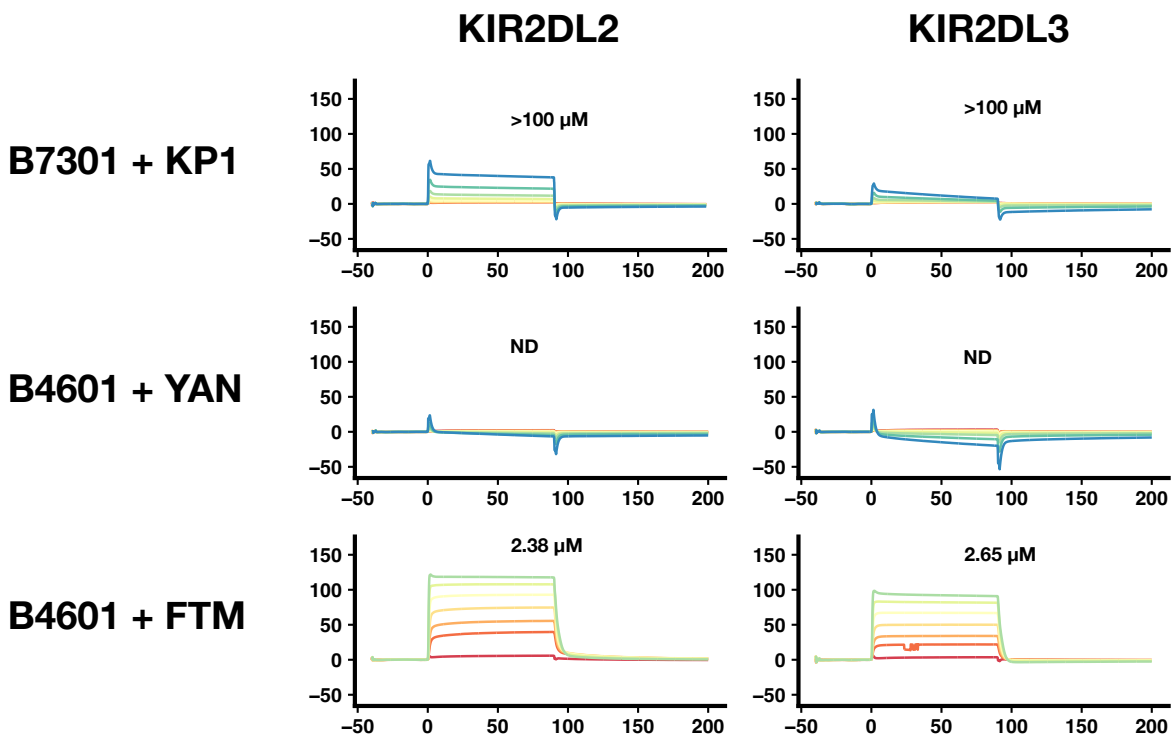

Figure S2.6 - SPR with KIR2DL2 and KIR2DL3 demonstrate an extremely low affinity complex.

Immobilized peptide MHC complexes are on the left while KIRs used are shown on top. Raw binding curves were plotted with response units on the Y-axis and time in seconds on the X-axis. The affinities for HLA-B*73:01-KP1 and -B*46:01-YAN peptide MHC complexes could not be accurately determined.

Chapter 3 – Presentation of different forms of the non-classical MHC, HLA-F and its non-human primate homologs

3.1 HLA-F can exist as both a peptide-loaded complex and "open conformer"

Of all class I HLA molecules, HLA-F is the least understood. Early attempts to characterize HLA-F led to the conclusion that unlike other HLA-I molecules, HLA-F is thought to be expressed at the cell surface without a peptide antigen and possibly without β2M, in a protein form known as its "open conformer" state (53,56,59,59). However, more recently, the Adams Lab and others have shown that some forms of HLA-F can, in fact, bind to peptides within its peptide binding groove (68,177,178). In its open conformer state, HLA-F has been shown to interact with various lineage II KIRs; most notably its interaction with KIR3DS1 leads to inhibition of HIV replication *in vitro* (58). On the other hand, when loaded with peptide, HLA-F no longer interacts with KIR3D molecules, but can bind to LIR1 (68). Therefore, understanding whether HLA-F can be expressed as both an open conformer and a peptide-loaded HLA complex *in vivo*, is crucial for deciphering its role in the various disease contexts in which its been implicated, including cancer (179), amyotrophic lateral sclerosis (180), and viral infection (181). Moving forward, the peptide-loaded forms of HLA-F protein will be referred to as pHLA-F, HLA-F complexed with β2M, but without peptide, as bHLA-F, and the heavy chain of HLA-F alone as ocHLA-F.

Up until 2017, HLA-F was the only HLA-I molecule for which a crystal structure had yet to be determined. However, in 2017, Dulberger *et al.* were able to determine the first crystal structure of HLA-F from protein made in insect cells, as opposed to refolded material. Overall, the

structure they determined showed that the structure of HLA-F is very similar to the structures of other class I molecules in regards to its platform domain, its interface with β2M, and α3 domain. Surprisingly, they also showed that the structure contained what appeared to be unoccupied electron density in the peptide binding groove, suggesting that it was likely loaded with peptide. Upon investigation of what these peptides looked like, they found another peculiarity: when expressed and purified from HEK293T (293T) cells, HLA-F was loaded with unusually long peptides. Based on the electron density found in the crystal structure and the sequences of bound peptides, it appeared that peptides were anchored into HLA-F at only the PΩ position within the F-pocket and that N-terminal peptide residues are likely not anchored at all. Consistent with the results of Dulberger *et al.* other groups have since published comparable data for HLA-F purified from immortalized human cell lines (177,178).

To explain the discrepancy in peptide binding relative to other HLA-I molecules, the authors noted that one residue was unique to HLA-F and ~0.01% of other class I alleles: a tryptophan at residue 62. Based on the structure, Dulberger *et al.* hypothesized that W62 likely blocks N-terminal peptide anchoring. Interestingly, the authors also noted that most non-human primate (NHP) orthologs of HLA-F had an arginine at this position (R62), aside from the orangutan, indicating that humans and orangutans had independently evolved W62, while the common ancestor of hominids and Old World Monkeys likely encoded for R62. Introducing this mutation into HLA-F (W62R) caused a reduction in peptide binding as measured by LC-MS/MS. This not only suggested that a single mutation transformed HLA-F from a primarily unloaded HLA molecule to one that could exist in either form, but provided an explanation for why previous

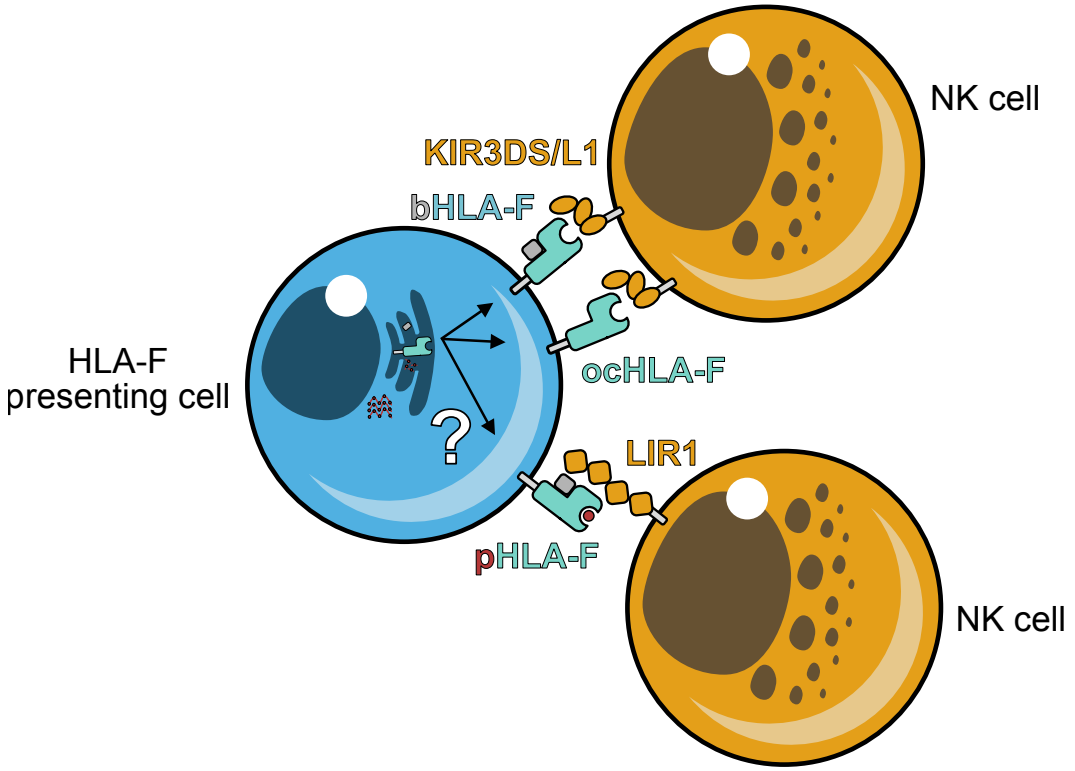

Figure 3.1 - A summary of past work on HLA-F presentation, regulation, and recognition. In this chapter, we addressed the question of HLA-F presentation and regulation, indicated by the white question mark. Which form is expressed on the cell surface under different conditions has important implications for which receptors HLA-F can interact with.

studies had concluded that HLA-F could be presented on the cell surface as ocHLA-F. For example, it could be that cells have a way of regulating which form of HLA-F they are presenting. Thus, Dulberger *et al.* proposed that peptide-presenting forms of HLA-F may be restricted to humans and orangutans, while HLA-F orthologs that contain R62 may prefer not to bind peptides and exist primarily in a peptide-free state.

Finally, the authors also showed that recombinantly expressed pHLA-F cannot signal through KIR3DS1 or KIR3DL2, while ocHLA-F could. However, pHLA-F could interact with LIR1. Thus, it follows that while ocHLA-F may function as a ligand for KIR3D molecules, pHLA-F

may function as a ligand for LIR1 *in vivo* (**Figure 3.1**). Naturally, these results raised many interesting follow-up questions. What distinguishes the structure of ocHLA-F from pHLA-F? Do other mutations in NHP orthologs compensate for the peptide-reducing capability of R62? And most importantly, if HLA-F can exist in multiple forms, which forms are relevant *in vivo*?

To address these questions, we generated conformation-specific reagents that recognize specific forms of HLA-F *in vitro* and on the cell surface. To our surprise, we found that contrary to previously proposed models, pHLA-F was the predominant form of HLA-F expressed on the surface of all cell lines tested. We also showed that β2M is essential for cell surface HLA-F expression and that cell surface expression also appears to be dependent on several components of the peptide-loading complex (PLC). We suggest that the differences between our results and those previously published are due to the fact that the previously used antibodies (including the 3D11) were generated against linear epitopes of HLA-F or refolded empty material. We show that 3D11 is not only biased towards ocHLA-F and bHLA-F, but also appears to bind unknown targets that are also dependent on β2M. Together, these data show that HLA-F is expressed and regulated on cells in a manner much more similar to other class I molecules than was previously thought.

3.2 Synthetic antibody (Fab) reagents can distinguish between different forms of HLA-F

We first set out to generate reagents specific for different forms of HLA-F. The pHLA-F and ocHLA-F used as targets for Fab generation were both derived from HLA-F expressed in insect cells. Using a novel purification strategy for the open conformer, we were able to purify both

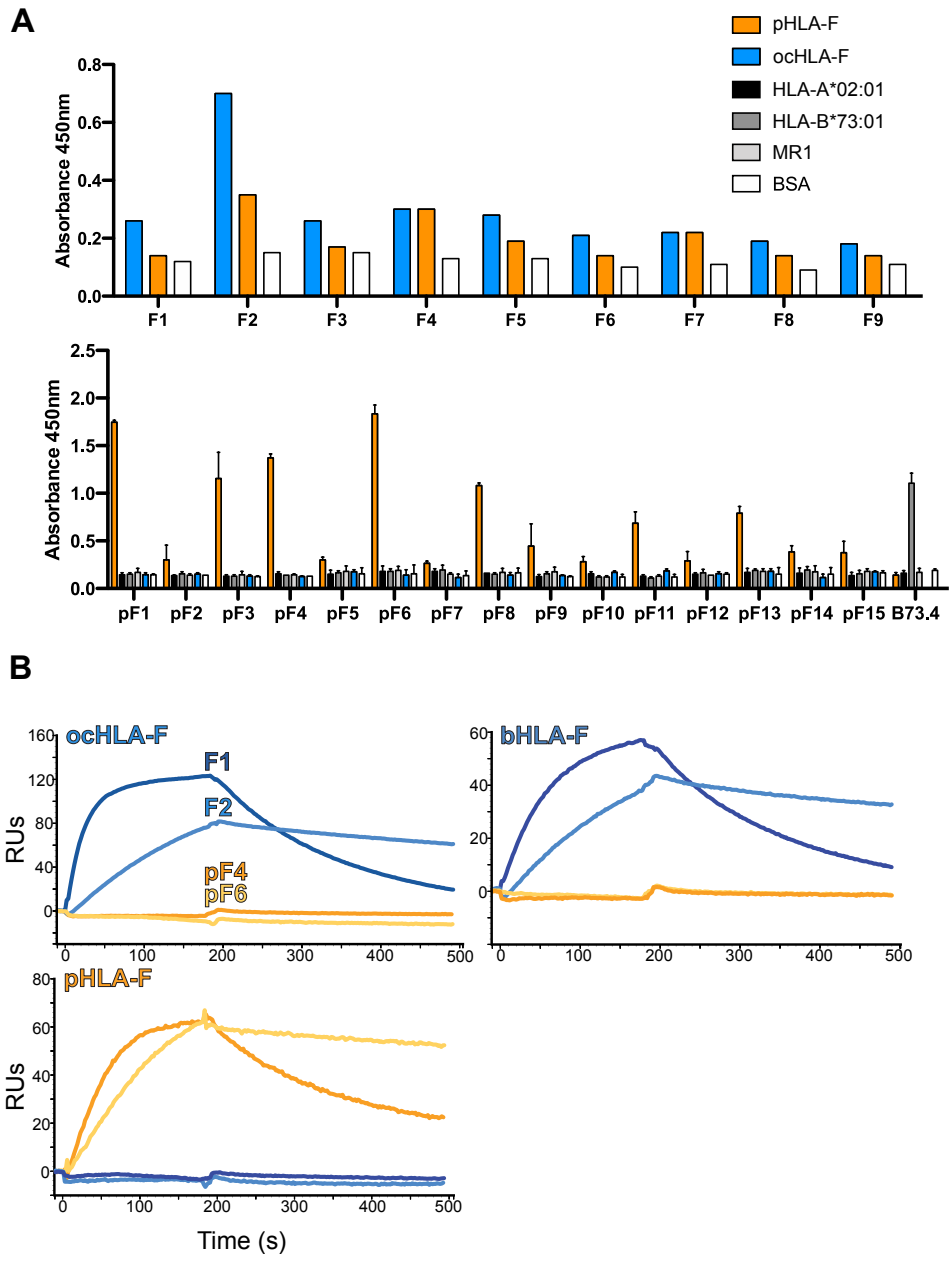

Figure 3.2 - *In vitro* validation of HLA-F conformation-specific fAbs.
A. A phage ELISA performed on purified phage populations expressing indicated fAbs and their reactivity with indicated MHC molecules or respective controls. Experiments only show minimal to no cross-reactivity with HLA-A*02:01, HLA-B*73:01, and MR1. **B.** Single concentration SPR curves indicate that while F1 and F2 only react with ocHLA-F and bHLA-F, pF4 and pF6 only react with pHLA-F.

ocHLA-F and pHLA-F from the same recombinant protein prep (**see Chapter 3.11 - Materials & Methods**) and use these as inputs for phage display-mediated selection (182). Following

several rounds of selection, we identified several promising Fabs specific to either ocHLA-F or pHLA-F as determined by phage ELISA (**Figure 3.2A**). For Fabs specific to pHLA-F, we were also able to confirm a lack of cross-reactivity with other common (HLA-A*02:01) and rare (HLA-B*73:01) class I molecules and the MHC-related molecule, MR1. Nanomolar affinities for both ocHLA-F and pHLA-F specific Fabs (from here on referred to collectively as ocF-Fabs and pF-Fabs, respectively) were observed as quantified by SPR, although pF-Fabs had an order of magnitude greater affinity than ocF-Fabs. Additionally, SPR was able to verify the specificity of our Fab reagents (**Figure 3.2B**). ocF-Fabs bound both bacterial and insect cell generated ocHLA-F material, while pF-Fabs bound only pHLA-F material. Additionally, we found that ocF-Fabs bound bacterial and insect cell-derived bHLA-F as well, indicating that they did not bind an epitope that would be inaccessible in the bHLA-F form. Thus, our reagents can distinguish between peptide-loaded (pHLA-F) and empty (ocHLA-F and bHLA-F) forms of HLA-F.

We were interested in determining which epitope the pF-Fabs recognized that allowed them to structurally distinguish empty HLA-F from pHLA-F. Using cryogenic electron microscopy (cryo-EM) we generated a 3.3 Å resolution map of pHLA-F bound to pF4 (data not shown). Interestingly, we found that pF4 recognized an epitope along the α1 helix near the A pocket. This region of the α1 helix is a common region bound by antibodies that prefer “empty” HLA molecules, specifically, HCA2, HC10, and 8G3 (183). Single concentration BLI experiments demonstrated that the W62R substitution alone drastically reduced the binding response to pF4 (**Figure 3.3C**). While the occurrence of W62 across class I HLA molecules is rare, it remains possible that pF4 could be cross-reactive with other HLA molecules other than HLA-F (**Figure**

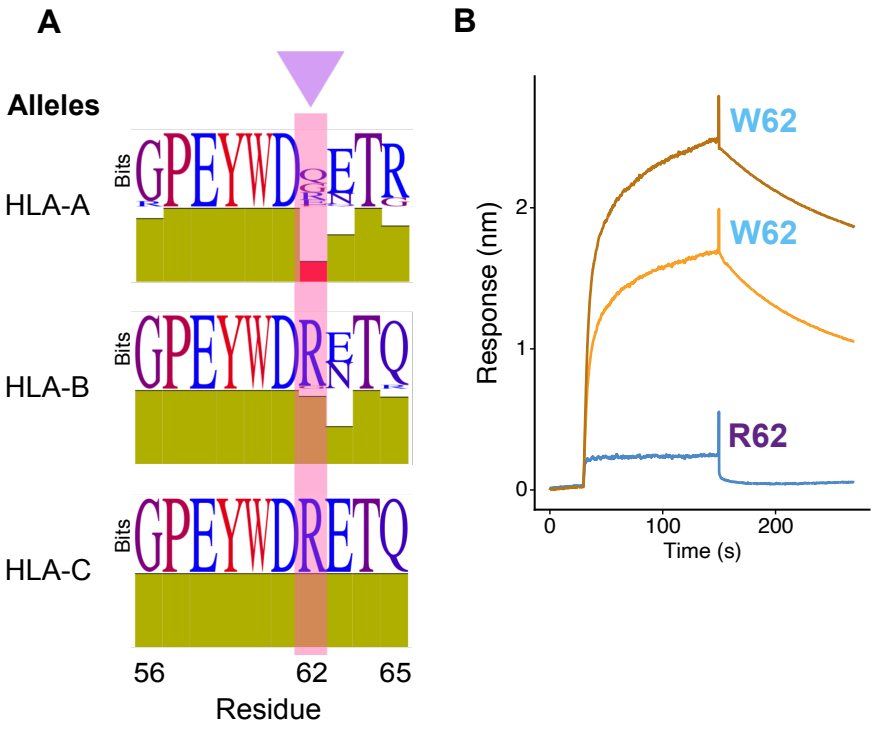

Figure 3.3 - Preliminary structural and binding data suggest that pF4 likely has minimal cross-reactivity with other class I molecules.
A. An alignment of HLA-A, HLA-B, and HLA-C alleles indicate that W62 rarely appears in HLAs outside of HLA-F and is often an arginine. Residues 56-65 encapsulate residues within the likely binding epitope of pF4 based on preliminary structural data (not shown). **B.** Single concentration bilayer interferometry binding curves show that immobilized pF4 can bind HLA-F and an orangutan HLA-F ortholog, but not a chimpanzee HLA-F ortholog. Human and orangutan molecules both encode W62, while the chimpanzee molecule encodes R62.

3.3B). Together, these data indicated that the conformation-specific Fab reagents recognize the α1 helix of HLA-F along an epitope previously reported to allow for differentiation between empty and loaded class I HLA molecules.

3.3 Fab reagents can be used to detect cell surface HLA-F

We next asked whether our Fab reagents could be used to detect cell surface expressed HLA-F by flow cytometry and whether they could do so with any specificity. Previous reports concluded that HLA-F can be expressed on the cell surface in its ocHLA-F or bHLA-F form, but these

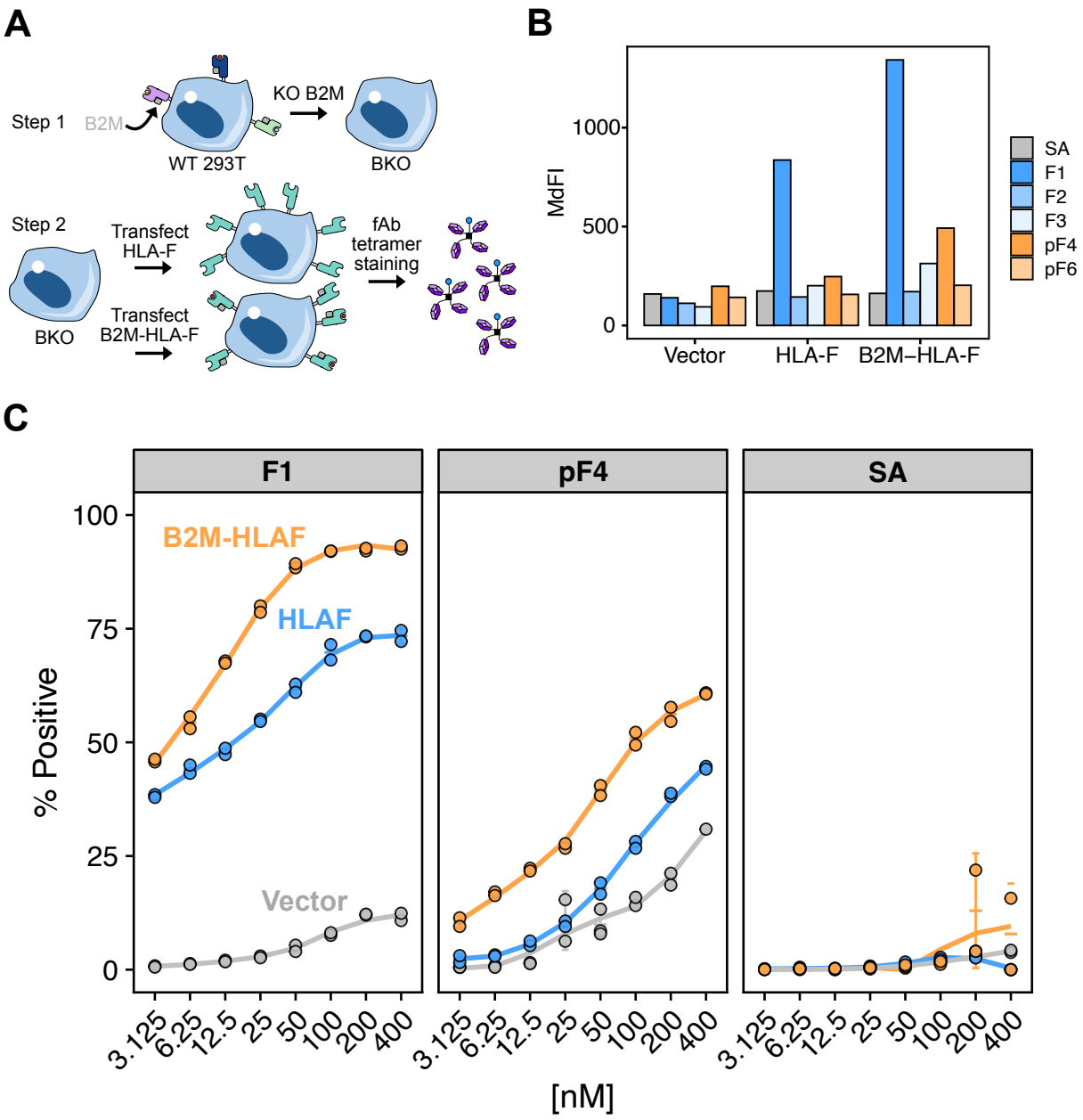

Figure 3.4 - Validating the use of labeled fAb tetramer reagents in recognizing cell-surface HLA-F.

A. A cartoon representation of our experimental design. First, β2M KO 293T cells were generated to ensure that they did not express complexed cell surface HLA molecules. Next, β2M KO lines (BKO) were used as transfectants for the HLA-F heavy chain alone or fused to β2M. Transfectants were then stained with Fab tetramers. **B.** A single replicate experiment testing different Fab reagents for their reactivity with BKO 293T cells transfected with the indicated constructs along the X-axis. **C.** Fab tetramer titrations of BKO 293T cells transfected with the indicated constructs. The tetramers used are displayed along the top while the colors here indicate the transfected construct. SA is the fluorophore-conjugated streptavidin used to make Fab tetramers mixed with buffer alone as a negative control.

conclusions were reached prior to Dulberger *et al.* showing that HLA-F could also present peptides. One challenge with such an experiment in finding a cell line that reliably expresses ocHLA-F or pHLA-F, but not both. To address this challenge, we engineered β2M knockout (KO) 293T cells using the CRISPR-Cas9 system (184). Using these β2M knockout 293T cells, we then transiently over-expressed full-length HLA-F heavy chain alone or N-terminally fused to β2M (β2M-HLAF). Our expectation was that by over-expressing the HLA-F heavy chain in the absence of β2M, any detected cell surface HLA-F would take the form of ocHLA-F, while over-expression of the β2M-HLA-F fusion construct would result in a mixture of cell surface bHLA-F and pHLA-F.

Using the 3D11 antibody, we observed that transiently transfected HLA-F and β2M-HLA-F were highly expressed on the cell surface, although the expression of β2M-HLA-F was higher. We then utilized this system to test whether fluorophore-conjugated Fab tetramers could detect cell surface HLA-F (**Figure 3.4A**). After testing various ocF-Fab and pF-Fab reagents, we found that the ocF-Fab #1 (F1) and pF4 Fabs were best able to detect HLA-F on the cell surface when comparing transfected cells to empty vector controls (**Figure 3.4B**). A titration series confirmed that while F1 stains both HLA-F and β2M-HLA-F transfected cells with minimal background, pF4 stained β2M-HLA-F transfected cells less well than F1, but confidently above both HLA-F and empty vector transfected cells (**Figure 3.4C**). Together, these data indicate that a combination of F1 and pF4 tetramers can distinguish between ocHLA-F and pHLA-F on the cell surface.

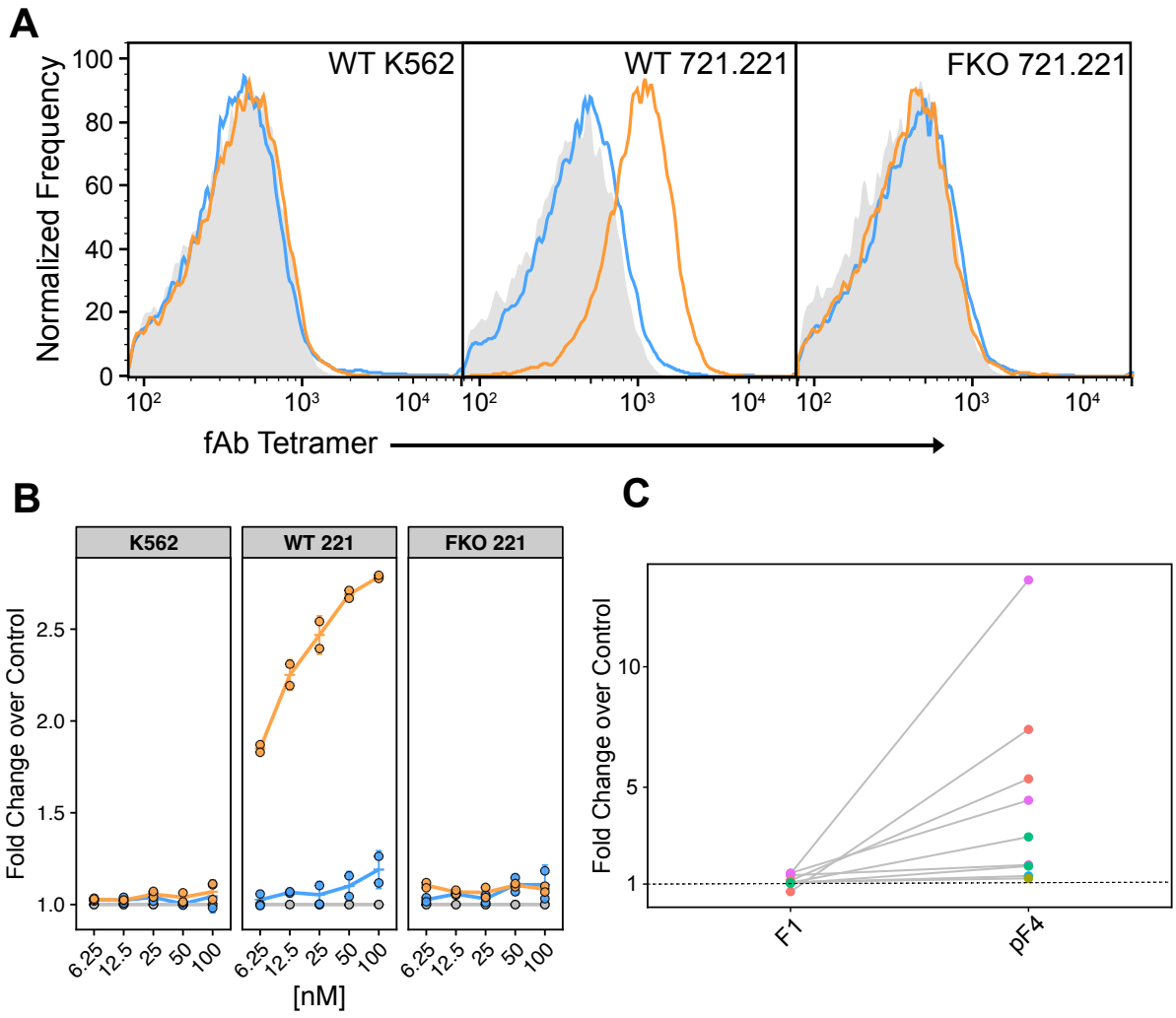

Figure 3.5 - Endogenously express HLA-F predominantly takes the form of pHLA-F. **A.** Representative histogram plots showing that on WT 721.221 cells, only pF4 stains cells noticeably above background. **B.** A titration of labeled streptavidin (SA) (grey), F1 (blue), and pF4 (orange) of WT K562 cells, a negative control, and WT and HLA-F KO 721.221 cells. **C.** A comparison of cell-surface expression of HLA-F as detected by F1 and pF4 in various cell lines. Shown are single-replicate values, but for several cell lines, this experiment has been repeated with similar qualitative results. Hepatocellular carcinoma (Huh7.5) and cervical carcinoma (HeLa) cell lines are also shown, but difficult to see because they did not display substantial cell-surface HLA-F by F1 or pF4.

3.4 Endogenous cell surface HLA-F expression is dominated by pHLA-F

After confirming that we could detect over-expressed HLA-F on the cell surface, we assayed for endogenously expressed HLA-F in a cell type known to supply relevant antigens to other

transfected and over-expressed classical HLA molecules (22). The 721.221 (221) cell line is known to be deficient in HLA-A and HLA-B, but endogenously expresses HLA-C (185), HLA-E (186), and HLA-F (53). In order to ensure that any binding by F1 or pF4 was due HLA-F recognition and not cross-reactivity with other HLA molecules, we used a similar strategy to engineer β2M KO and HLA-F KO 221 cell lines as we did with 293T cells. The K562 cell line has been previously shown to be devoid of any HLA-F protein by Western blot and was used as a negative control (187).

To our surprise, we found that only the pF4 tetramers stained WT 221 cells, while the F1 tetramers did not (**Figure 3.5A** and **Figure 3.5B**). This was in stark contrast to staining seen for transfected β2M KO 293T cells where F1 was able to stain cells with a much better signal-to-noise ratio (transfected cells relative to empty vector controls) than did pF4. As expected, pF4 fails to stain any of the four different HLA-F KO expanded clones, suggesting that pF4 is truly recognizing HLA-F and that it is unlikely that off-target effects of gene editing were responsible for our observations. Despite the lack of HLA-F protein expression detected by Western blot, we also observed that pF4 still stained the 221 β2M KO line, but to a lesser extent than WT cells (**Figure S3.2**). While the cause of this staining remains unclear, we suggest that these results are likely cell-type dependent because pF4 does not stain a similarly generated JY β2M KO clone (**Figure S3.3**). It is unlikely that pF4 is binding to ocHLA-F expressed on the surface due to the lack of β2M because we do not see compensatory staining with F1, nor is HLA-F detectable by Western blot in the 221 β2M KO cell line.

The surprising result that endogenous expression of cell surface HLA-F was only detectable by pF4 motivated us to test additional immortalized cell lines for cell surface HLA-F expression. Using trophoblast, endometrial stromal fibroblast, hepatocarcinoma, cervical carcinoma, and lymphoblastoid cell lines, we found that pF4 always stained cells to a greater degree than F1, which never stained cells above a negative control (fluorophore-conjugated streptavidin alone) (**Figure 3.5C**). Together these data indicated that endogenous expression of HLA-F on the cell surface of immortalized cell lines is loaded with peptides and takes the form of pHLA-F.

3.5 IFN-γ induction of HLA-F suggests that β2M is essential for cell surface expression

Thus far, we have shown that under basal conditions, endogenously expressed HLA-F mainly takes the form of pHLA-F. However, it remains possible that under certain conditions that lead to the upregulation of HLA-F transcript and protein or during limited availability of β2M, ocHLA-F becomes expressed on the cell surface. To test this hypothesis, we induced HLA-F expression by treatment with IFN-γ in cells with or without β2M. IFN-γ is a widely expressed cytokine expressed mainly by leukocytes in response to viral and microbial infection (188). Nearby cells that can sense IFN-γ respond by upregulating the expression of hundreds of genes, including genes that encode the PLC and some class I HLA molecules, such as HLA-F (189). For this experiment, we used WT and engineered 293T and Jurkat cells in which β2M or HLA-F had been knocked out by CRISPR (58).

To verify that IFN-γ induction was successful, we measured the surface expression of complexed class I HLA molecules using the W6/32 antibody, which should detect an increase in class I cell

surface expression. When using F1 and pF4 tetramers to test for cell surface HLA-F, we found

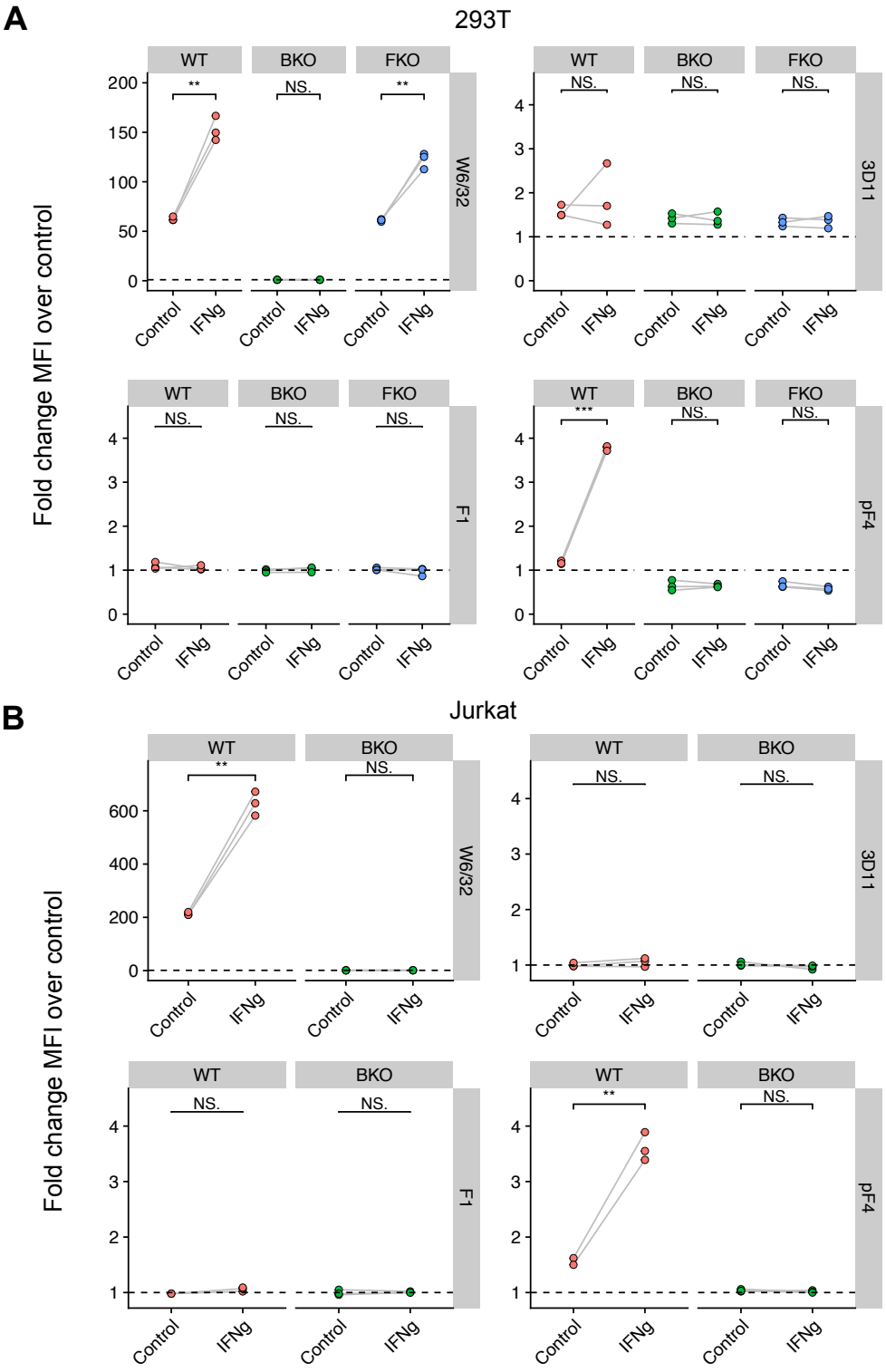

Figure 3.6 - IFNγ induction of engineered HEK293T cells and Jurkat cells show that β2M is essential for surface expression of pHLA-F, but does not effect expression of ocHLA-F.

Figure 3.6 Continued
A. WT, β2M KO, and HLA-F KO 293T cells were cultured with and without 10 ng/mL IFNγ for 2 days prior to staining with W6/32, 3D11, or F1 or pF4 tetramers. Y-axis is shown as fold-change MdFI of indicated antibody or tetramer relative to either an isotype control (W6/32 and 3D11) or labeled SA only control (F1 and pF4). Significance was measured using Welch's T-test. **<0.005 ***<0.0005. **B.** Same experiments as in A, but using WT and β2M KO Jurkat cells instead.

that in WT backgrounds, staining with pF4 was detected on 293T and Jurkat cells upon IFN-γ treatment while F1 failed to stain in any context (**Figure 3.6**). Interestingly, in the β2M KO background both pF4 and F1 failed to stain cells, whether treated with IFN-γ or not, suggesting that eliminating β2M does not lead to an increase in cell surface ocHLA-F. As expected, there was no staining by pF4 or F1 of HLA-F KO 293T cells with or without IFN-γ, indicating that pF4 is not likely cross-reacting with other HLA molecules. Together, these data suggest that even when induced by an external stimulus, HLA-F is expressed at the cell surface in the form of pHLA-F. Additionally, cell surface expression of pHLA-F requires β2M. More importantly, however, the loss of β2M does not lead to compensatory cell surface expression of ocHLA-F.

3.6 Cell surface pHLA-F may be dependent on several components of the PLC

HLA-F's dependence on components of the peptide-loading complex (PLC) remains ambiguous. One early study suggested that cell surface HLA-F was independent of TAP, which implied that surface presentation would most likely be in an open conformer form (53). If pHLA-F as detected by pF4 is truly loaded with peptides, then this form may also be sensitive to disruption of PLC components. In order to test whether disruption of PLC components modifies the abundance and/or form of HLA-F at the cell surface, we used CRISPR to knock out TAP1 and TPN, two important members of the PLC (9). Additionally, thanks to our conformation-specific

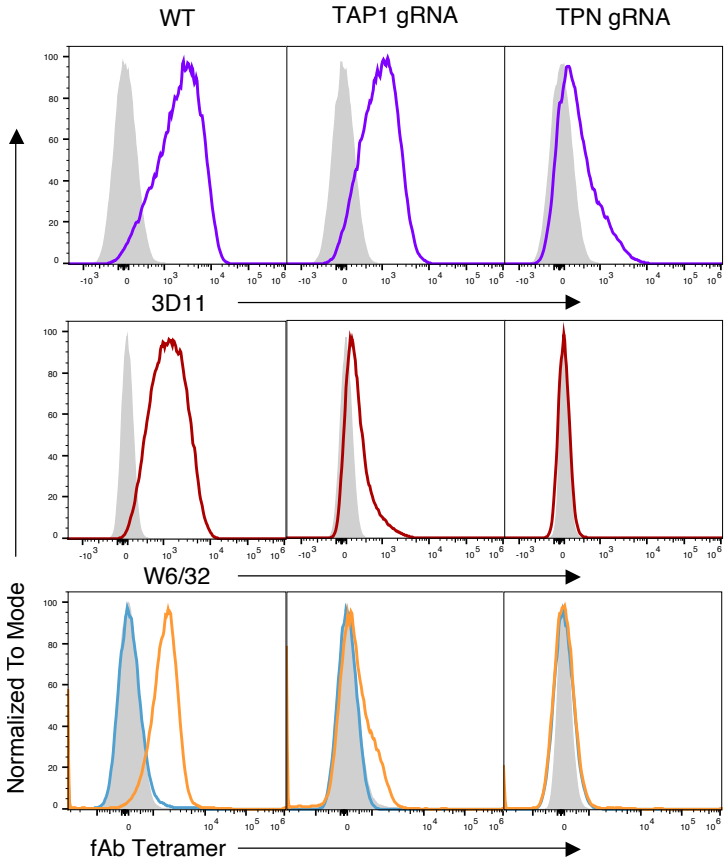

Figure 3.7 - Preliminary data suggests pHLA-F is dependent on TAP1 and TPN for cell surface expression.
Histograms of bulk-sorted 721.221 cells following nucleofection with either a TAP1 or TPN-targeting sgRNA. Cells were sorted on the top 1% of GFP expressing cells and expanded for 2-3 weeks prior to staining with indicated reagents.

Fabs, we can also test whether the disruption of any specific components leads to compensation at the cell surface by expression of ocHLA-F.

We were able to generate preliminary data for TAP1 and TPN KO cells. In line with a model suggesting regulation of pHLA-F by the canonical class I PLC pathway, pHLA-F expression was reduced in TAP1 and TPN KO lines as assessed by pF4 staining (**Figure 3.7**). Staining by W6/32 to assess cell surface expression HLA-C*01:02 in 221 cells follows a similar trend, consistent

with successful TAP1 and TPN KOs in these cells. Importantly, we did not observe compensatory expression of ocHLA-F as measured by F1 staining, suggesting that disruption of other canonical regulatory factors of HLA-I expression also do not lead to an increase of ocHLA-F cell surface, similar to our experiments using β2M KO cell lines. Nevertheless, definitive conclusions can only be reached once monoclonal lines have been expanded and KOs confirmed by Sanger sequencing, Western blot, and/or flow cytometry. However, these preliminary results are consistent with pHLA-F expression being regulated in a manner similar to other class I molecules.

3.7 ocHLA-F is expressed on the surface of activated lymphocytes

Since we had yet to observe ocHLA-F in an endogenous context, we sought to find a different system that would allow for the detection of ocHLA-F expression on the cell surface. Activated lymphocytes have long been recognized as producers of empty cell surface HLA molecules (190). Additionally, studies have shown that activated lymphocytes upregulate HLA-F protein expression (60). Thus, we hypothesized that perhaps ocHLA-F expression could be detected on the surface of activated lymphocytes. To test this hypothesis, we used primary frozen PBMCs from 2 healthy donors and cultured them for 5 days in the presence of IL-2 and a plate-bound anti-CD3 antibody. Cells were then harvested and assayed by flow cytometry. Our results showed that activated T cells expressed the highest abundance of ocHLA-F relative to any other cell type tested, although pHLA-F expression is still higher (**Figure 3.8**). These data confirm for the first time that ocHLA-F can be expressed endogenously on the surface of a primary cell type. Since these are peripheral blood-sourced T cells, these results are consistent with the model

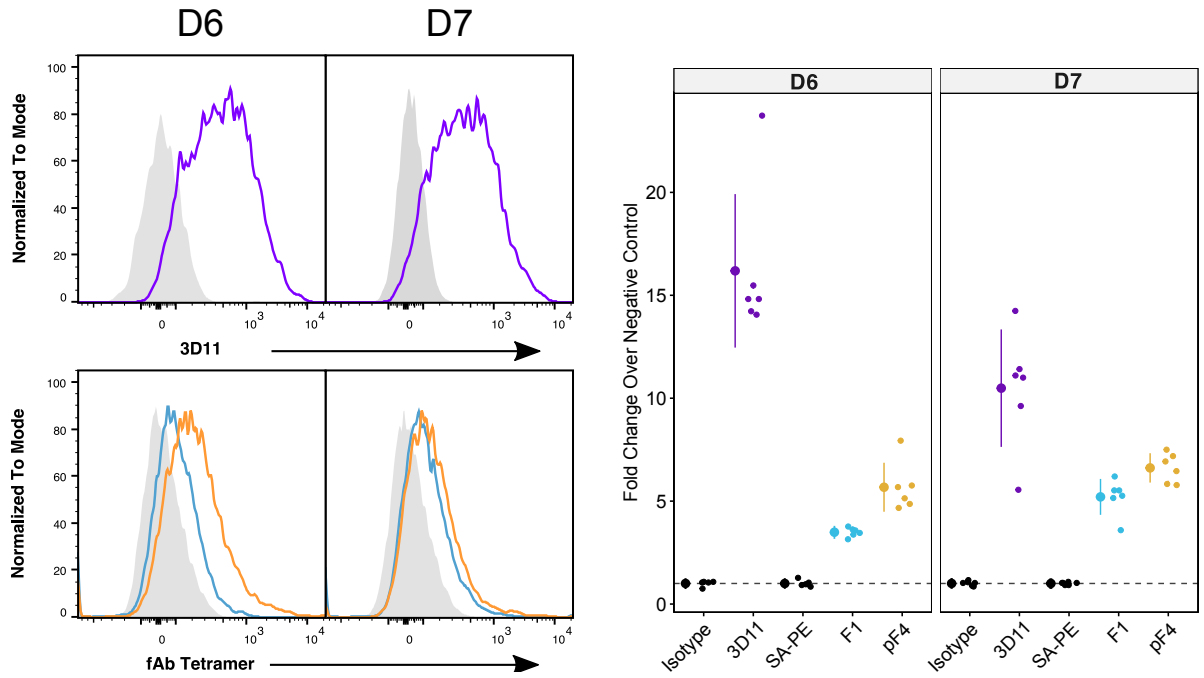

Figure 3.8 - Activated T cells express an almost equivalent amount of cell surface ocHLA-F and pHLA-F.
Representative flow cytometry plots (left) of live, CD3+ PBMCs from two donors (D6 and D7). Technical replicates shown as fold change over negative controls (right). For 3D11, the negative control is the indicated isotype while for F1 and pF4, the negative control is streptavidin-PE (SA-PE) only.

proposed by Garcia-Beltran *et al.* suggesting that HLA-F can be upregulated on HIV-infected T cells and signal through KIR3DS1 to inhibit HIV replication (58).

3.8 The 3D11 antibody is biased and binds to non-specific targets

To our knowledge, 3D11 is currently the only commercially available antibody for HLA-F that is validated for use in flow cytometry. However, it is not clear whether 3D11 can recognize all forms of HLA-F with comparable affinity. The 3D11 antibody, as well as other HLA-F antibodies, were originally designed to recognize empty refolded HLA-F material (52), which we now show is structurally different from pHLA-F. In order to investigate 3D11 specificity, we

tested its ability to stain the TAP1 and TPN KO cell lines (**Figure 3.7**). In this experiment, 3D11 staining was not depleted to the same extent in the knockouts as was observed for W6/32 and the pF4 reagents. Further, 3D11 did not detect pHLA-F induced by IFN-γ in 293T or Jurkat cells (**Figure 3.6**). One possible explanation for these results is that the 3D11 antibody is actually biased towards the empty forms of HLA-F and fails to effectively stain cells expressing only pHLA-F. Although not mutually exclusive, it could also be that 3D11 binds to one or several non-specific targets other than HLA-F. To test these hypotheses, we used a combination of HLA-F KO cell lines and purified protein reagents and compared their reactivity with the 3D11 antibody.

Using four validated 221 HLA-F KO lines, we observed that the 3D11 antibody was staining cells regardless of successful HLA-F KO (**Figure 3.9A**). HLA-F KO JY clones that have so far only been validated by flow cytometry (not by Western blot or Sanger sequencing) showed a similar result (**Figure 3.3**). Interestingly, although the KO of HLA-F in 221 and JY cells does not entirely abrogate 3D11 staining, the KO of β2M in both cell lines did, suggesting that 3D11 may bind a β2M-dependent protein aside from HLA-F (**Figure 3.9B**). Since it appeared as though 3D11 was binding non-specific targets, we sought to identify all 3D11 binding partners by immunoprecipitation followed by mass spectrometry. Although the experiment resulted in the specific pull-down of HLA-F from WT 221 cells and not from β2M or HLA-F KO 221 cells, there did not appear to be any obvious alternative targets expressed on the cell surface (**Table S1**). Additionally, no other HLA molecules were detected. Thus, future experiments will be

necessary to elucidate the cause of the non-specific binding of 3D11 to HLA-F KO cells observed here.

To test whether 3D11 is biased towards empty forms of HLA-F, we used biolayer interferometry (BLI) to measure the binding response of empty or peptide-loaded HLA-F made from insect cells for 3D11. Whether immobilizing 3D11 or HLA-F, the binding response was consistently higher between empty forms of HLA-F (ocHLA-F and bHLA-F) and 3D11 than the peptide-loaded form of HLA-F and 3D11 (**Figure 3.9C**). To corroborate these findings, we used protein-coated streptavidin-conjugated dynabeads in combination with flow cytometry. We expressed and purified biotinylated refolded ocHLA-F and bHLA-F and pHLA-F generated in insect cells (Hi5) and 293T cells. SA-conjugated dynabeads were then coated with biotinylated recombinant proteins, washed to remove excess/unbound protein, and then stained with a fluorophore-conjugated 3D11 antibody. Preliminary results suggest that indeed 3D11 binds 10.2 and 10.6-fold better to beads coated with ocHLA-F and bHLA-F, respectively, than beads coated with insect cell-derived pHLA-F and 1.5 and 1.6-fold better than beads coated with mammalian cell-derived pHLA-F (**Figure 3.9D**). This trend was also seen using the HCA2 antibody, known for its binding preference for empty HLA molecules (191). Thus, it appears as though not only is 3D11 binding non-specific targets, but that it is also biased towards empty forms of HLA-F. Although more replicates of each experiment need to be performed to account for biological and technical variation, if true, this would account for the discrepancy seen thus far between staining with 3D11 and F1 and pF4.

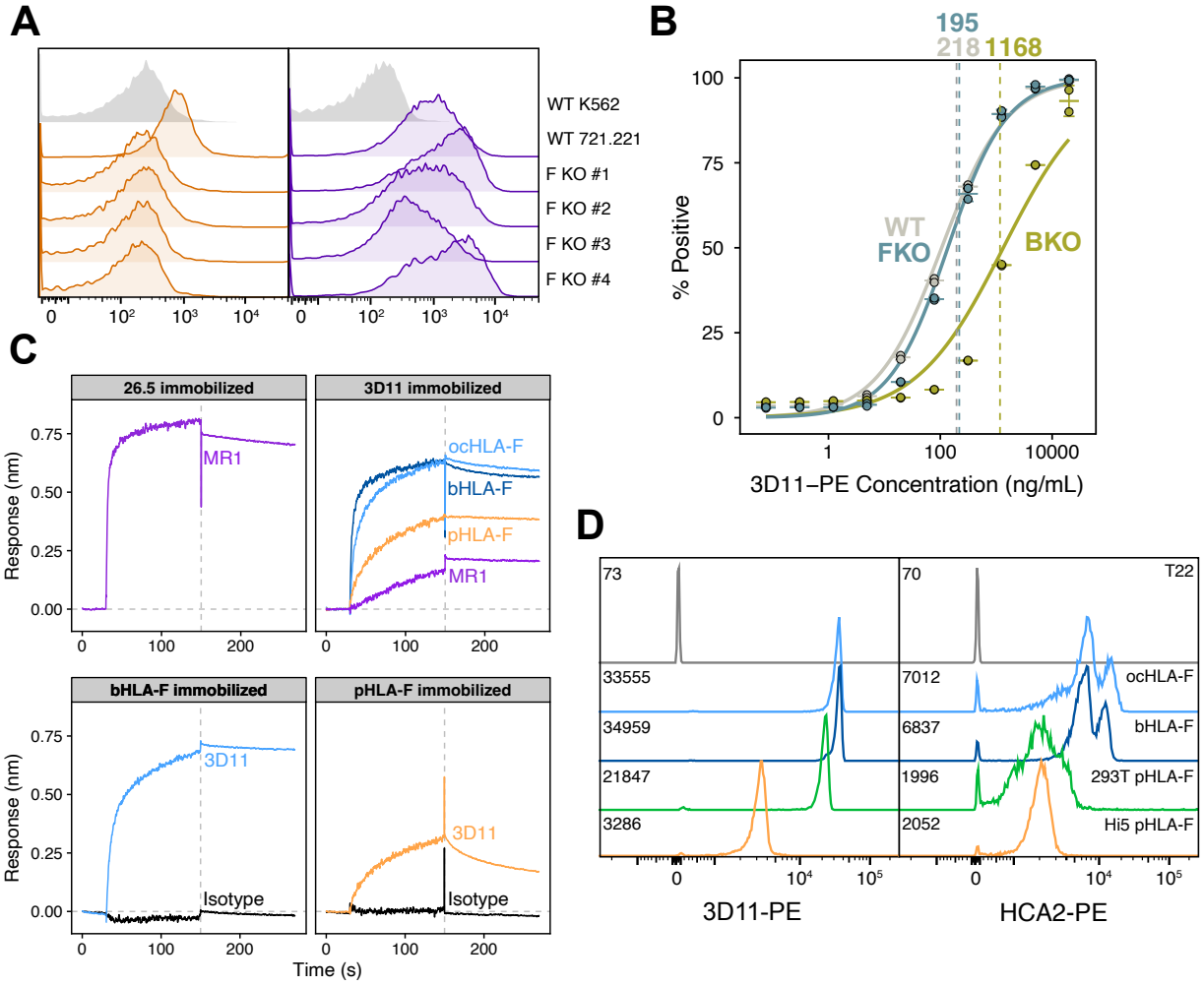

Figure 3.9 - The 3D11 antibody is likely cross-reactive and biased towards empty forms of HLA-F.

A. Flow cytometry results using pF4 (orange) or 3D11 (purple) to label cell-surface HLA-F in indicated cell lines. Sanger sequencing results of amplified genomic DNA from WT (1), HLA-F KO #1 (1,2), #2 (3,4), #3 (5,6), and #4 (7,8) 721.221 cells. **B.** A titration using the 3D11 antibody to label WT, β2M KO, and HLA-F KO 721.221 cell lines. **C.** Single concentration biolayer interferometry experiments using immobilized antibody or HLA-F indicated. 1 μM of MR1 and HLA-F analyte was used when antibodies were immobilized (top row). 833 nM of 3D11 or isotope antibody analyte was used when HLA-F was immobilized. **D.** Streptavidi-coated beads labeled with recombinant biotinylated proteins indicated on the right-hand side were appropriately washed to remove unbound protein and stained with either the 3D11 or the HCA2 antibody. MdFI values are indicated on the left.

3.9 NHP homologs of HLA-F present peptides regardless of W62 or R62

An important finding by Dulberger *et al.* was that the recombinant W62R HLA-F mutant was severely impaired in its ability to bind peptides as determined by total peptide abundance and number of peptides detected. These data raise an interesting evolutionary question about HLA-F and its orthologs: is there any association between W62 or R62 and the ability to present peptides by NHP ortholog (MHC-F) molecules? One hypothesis might be that MHC-F molecules with R62 are somehow worse peptide presenters than those with W62. Two ways in which this could manifest is either through the reduced ability of MHC-F molecules with R62 to bind peptides, similar to the W62R HLA-F mutant, or through the increased intrinsic stability of MHC-F with R62 without peptides. Importantly, these properties do not have to be mutually exclusive.

In order to test the hypothesis that MHC-F molecules with R62 cannot present peptides, we compared the stabilities and peptide-presenting abilities of recombinant forms of HLA-F with those of MHC-F molecules. To obtain the relevant sequences we cloned the mRNA of MHC-F using BLCLs derived from a chimpanzee (Patr-F), gorilla (Gogo-F), orangutan (Popy-F), gibbon (Hyla-F), and macaque (Mamu-F) kindly provided to us by the Parham lab at Stanford University and the Gilad lab at the University of Chicago. A phylogenetic tree that includes previously sequenced MHC-F genes from other species (68) along with those that we cloned matches expectations based on the species from which the BLCLs were derived (**Figure S3.4**).

In order to compare stabilities between different orthologs, we produced recombinant empty forms of each protein with and without β2M refolded from bacterial inclusion bodies. We also generated putatively peptide-loaded forms of each protein made in insect cells using a similar

scheme to what was previously done for HLA-F (68). When purifying HLA-F from bacteria without a peptide ligand, we consistently observe two states by SEC: an aggregated state that elutes early off the column – larger molecules elute earlier – and a monomeric state that elutes later at the expected volume based on its predicted molecular weight (**Figure 3.10A**). Additional verification by SDS-PAGE supports these observations. We have also observed that concentrating bacterially-derived HLA-F by centrifugal filtration will shift the equilibrium in solution towards the aggregated state. This, however, is not true of HLA-F expressed in insect cells, which elutes mainly as a monomer, and does not aggregate extensively when concentrated. Accordingly, we qualitatively assessed the stability of empty forms of MHC-F with respect to their ability to yield monomeric protein in the absence of ligand, their monomer-to-aggregate peak ratios, and their ability to remain monomeric at high concentrations.

Upon purification of recombinant empty MHC-F molecules, most orthologs were able to yield soluble monomers without the addition of a peptide ligand, similar to HLA-F (**Figure 310A,B**). SDS-PAGE confirms this for Patr-F and Hyla-F. Mamu-F was the only ortholog that did not refold (**Figure 3.11**). Concentrating either the chimpanzee-derived or orangutan-derived proteins resulted in a reduced monomer-to-aggregate peak ratio by SEC, also similar to what we observe for HLA-F (**Figure 3.10C**). The addition of β2M to the refolding reaction consistently increases the monomer-to-aggregate peak ratio, suggesting the presence of more stable monomers with β2M relative to just the heavy chain alone (**Figure 3.10D**). Upon production of recombinant MHC-Fs in insect cells, purification by SEC yielded more stable monomers with minimal to no aggregate peaks (**Figure 3.12**). Interestingly, Popy-F, derived from the orangutan, eluted as both

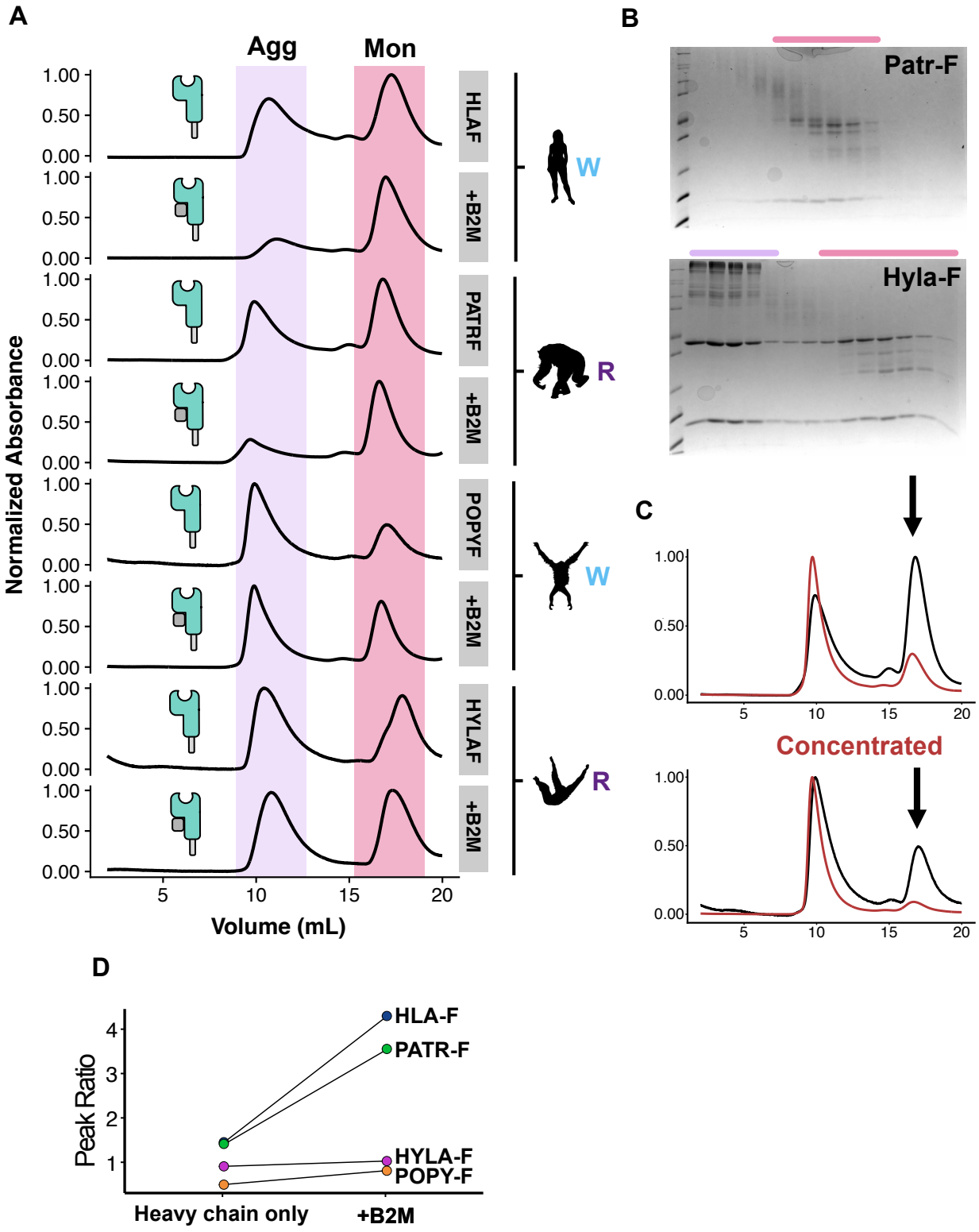

Figure 3.10 - Orthologs of HLA-F display similar stability when refolded without peptide ligands.

Figure 3.10 Continued
A. Size exclusion chromatography (SEC) traces of either HLA-F or an ortholog of HLA-F including chimpanzee, orangutan, or gibbon (ordered from top to bottom). Refolds were done either using the heavy chain alone or with β2M (+β2M). The violet highlighted volume indicates the aggregate peak (Agg) that we typically observe when refolding HLA-F without a peptide ligand. The pink highlighted volume indicates the monomer peak (mon) where we typically observe monodispered HLA-F eluting from the column. **B.** Non-reducing SDS-PAGE of SEC fractions from purifications of Patr-F (the chimpanzee ortholog) and Hyla-F (the gibbon ortholog). **C.** SEC traces before (black) and after (red) concentrating Patr-F (top row) and Popy-F (bottom row). **D.** Aggreate-to-monomer peak ratios calculated for refolds using just the heavy chain or heavy chain plus β2M.

a non-covalent dimer and monomer, a feature observed only for that ortholog. Based on these data, we observed no relationship between increased stability and empty MHC-F molecules with R62 compared to W62.

We hypothesized that peptide ligands account for the increased stability of MHC-F molecules expressed in insect cells, even those with R62. To test this, we determined the crystal structures of Patr-F and Mamu-F to 2.8Å and 2.1Å, respectively. Importantly, both Patr-F and Mamu-F encode R62. An alignment of Patr-F, Mamu-F, and HLA-F shows that all three structures are highly similar overall (**Figure S3.5**). Intriguingly, in both Patr-F and Mamu-F we observed clear electron densities for what appear to be peptide ligands within the peptide binding grooves (**Figure 3.12**). Similar to HLA-F, the density observed was incomplete or uncertain in places, making it difficult to build individual peptides into the model. However, in all structures including HLA-F, an anchor residue was clearly apparent near the middle of the peptide binding groove that likely corresponds to position 3, 4, or 5 of the peptide ligand. Density corresponding to an N-terminal anchor was only apparent in the structure of Mamu-F, perhaps because its structure is better resolved. Based on the electron density identified in Mamu-F, we hypothesize

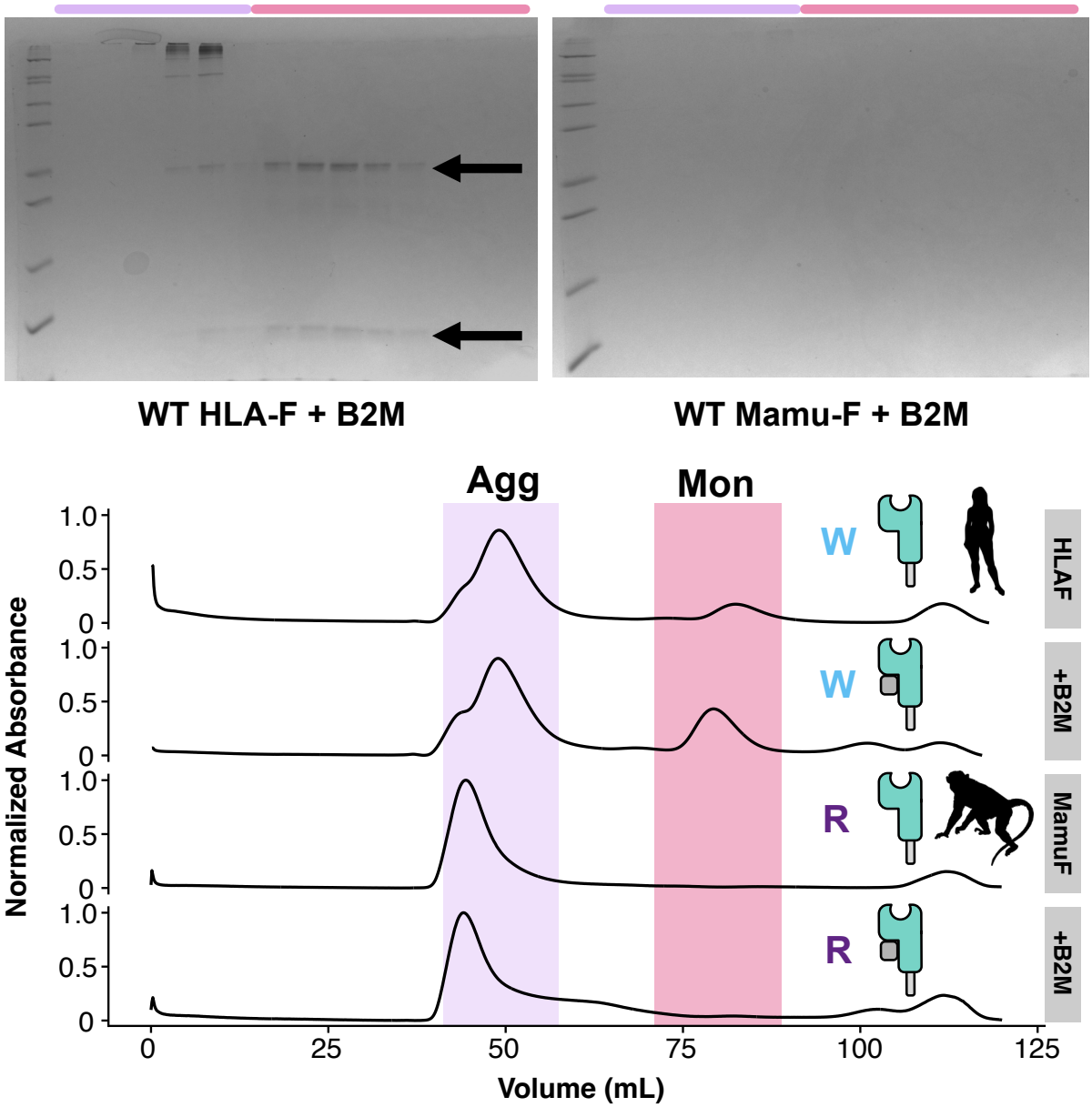

Figure 3.11 - Mamu-F, the macaque ortholog of HLA-F, does not refold at all without a ligand.
Non-reducing SDS-PAGE of SEC fractions portrayed below. While HLA-F refolds yield soluble monomer without the addition of a ligand, Mamu-F fails to yield any, even upon adding β2M.

that it may present peptides in three different states, each of which is bound differently at the N-terminus (**Figure 3.13**).

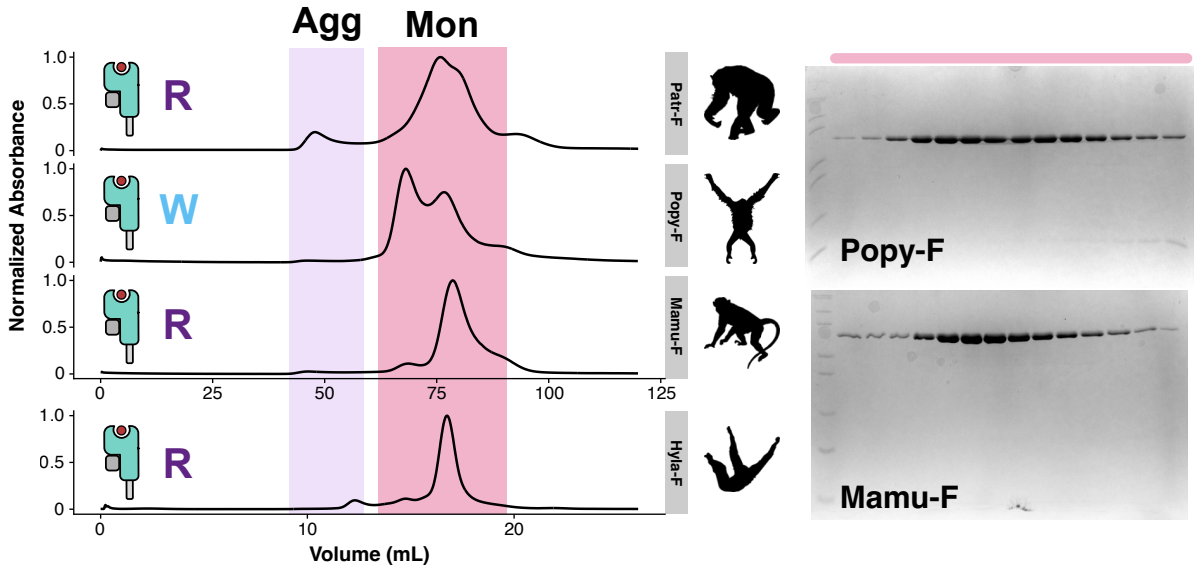

Figure 3.12 - MHC-F molecules display enhanced stability when expressed in insect cells. SEC traces show that when expressed in insect cells, Patr-F, Popy-F, Mamu-F, and Hyla-F (from top to bottom) exhibit minimal aggregation even once concentrated, unlike refolded material. Non-reducing SDS-PAGE shows clean, monomeric bands of the expected size. The second larger peak seen for Popy-F is a concentration-dependent non-covalent dimer.

To confirm the presence of peptides in these samples and determine what they look like, we eluted peptides from recombinant HLA-F, Patr-F, Popy-F, and Mamu-F expressed in insect cells. Dr. Samuel Weng of the UChicago Proteomics Core then processed the eluate by LC-MS/MS. In agreement with determined crystal structures, peptides were detected in each sample, some which were present in all species (**Table S3**). Immediately noticeable was that the length distribution of peptides eluted from HLA-F from Hi5 cells were on average shorter than peptides eluted from 293T cells by Dulberger *et al.* and from K562 cells by Hó *et al.* (**Figure 3.14A**). While the significance of these results are unknown, we have observed that recombinant HLA-F made in insect cells is more stable compared to 293T-derived material. Thus, this result might suggest that although HLA-F can present longer peptides, it is most stable when presenting shorter ones.

In order to analyze and compare peptide binding motifs, we used the GibbsCluster server to align and cluster peptides and visualize different peptide binding motifs within our data. Interestingly, upon analyzing datasets separately, some motifs appeared qualitatively similar to one another,

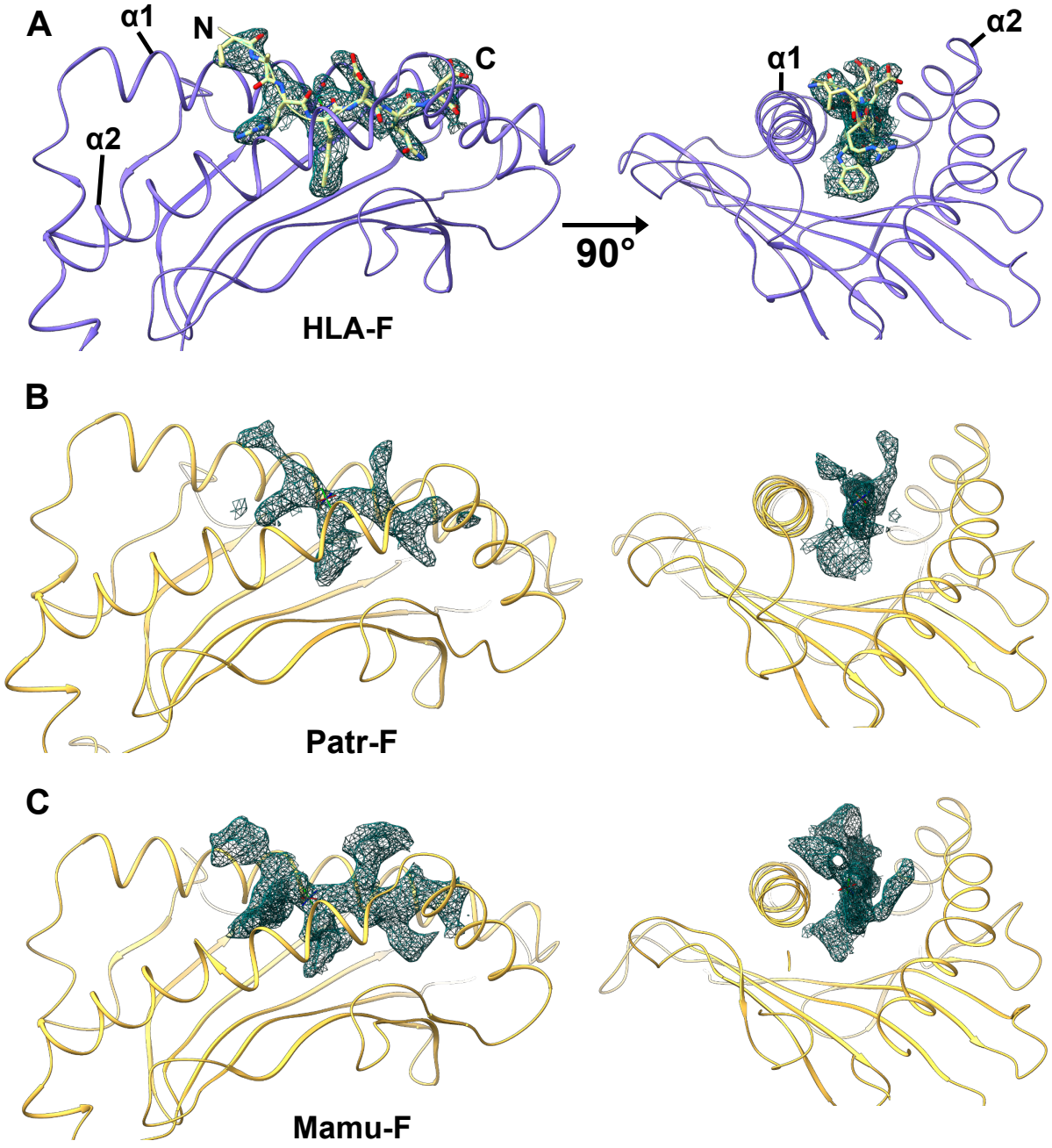

Figure 3.13 - Crystal structures of Patr-F and Mamu-F exhibit clear peptide density within their respective binding grooves.

Figure 3.13 Continued
A. The structure of HLA-F was determined by Dulberger *et al.* It was noted that its N-terminus is likely not anchored into the groove due to the lack of electron density within the binding pocket in that region. Furthermore, there was a prominent residue anchored in the center of the platform. **B,C.** The electron density found within the binding grooves of Patr-F and Mamu-F are nearly exactly the same. Each show weak electron density for anchors near the N-terminus, while showing prominent electron density for an anchor in the center of the platform. Because the structure of Mamu-F is better resolved, there is evidence of an N-terminal anchor as well.

suggesting that in each data set, certain peptides conformed to a distinct profile (**Figure 3.14B**). These motifs stood out not only for what appeared to be a charged anchor residue at PΩ, similar to what was reported previously (68), but also for a hydrophobic anchor in the middle of the peptide, concordant with ligand density in determined crystals structures of HLA-F, Patr-F, and Mamu-F (**Figure 3.12**). While peptides eluted from Hi5 and K562 cells had an enrichment for hydrophobic residues at P5, peptides eluted from 293T cells exhibited a similar enrichment at P3 (**Figure 3.14B**). Future experiments will need to confirm which of the identified peptides are true ligands for MHC-F molecules and how they are presented. Finally, a comparison of ligand intensities and the number of ligands across species revealed no relationships with residue 62 (**Figure 3.14C**). It should also be noted that while these peptides may represent those that fit within the biophysical restraints of MHC-F, they do not reflect the cellular context in which MHC-F naturally presents peptides, which may obscure more concrete binding motifs. Together, these data suggest that MHC-F molecules present peptides similarly to HLA-F and that R62 alone is not a meaningful predictor of the stability of empty forms of MHC-F nor of their ability to present peptides. However, it remains to be seen what peptides are most physiologically relevant within and across species

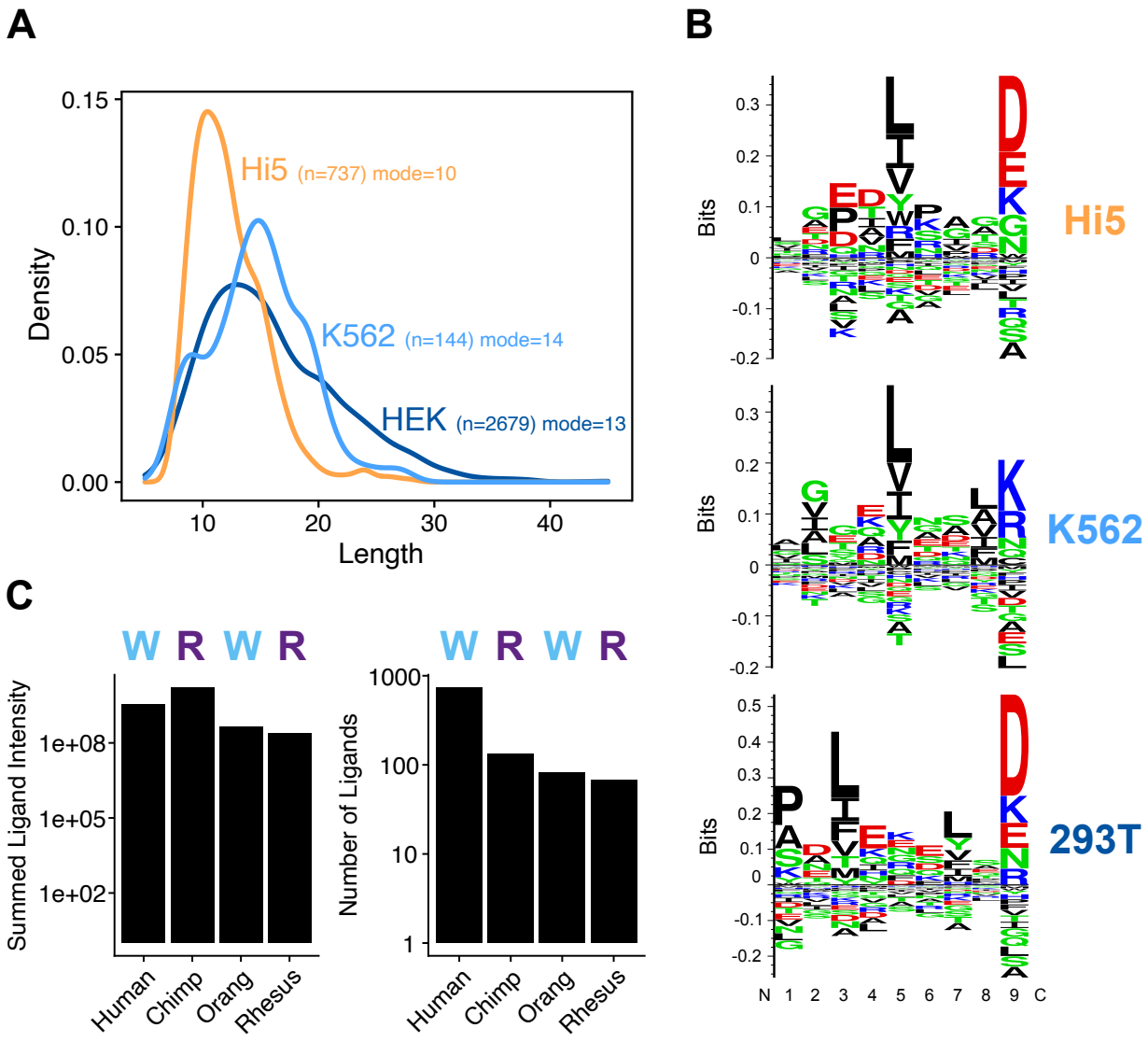

Figure 3.14 - LC-MS/MS confirms peptides bound by MHC-F molecules and highlights a distinct length distribution.
A. The length distributions of HLA-F peptides eluted from indicated cell lines. Hi5 cells represent data from this study. **B.** Selected clustered motifs output but GibbsCluster that appear qualitatively similar between datasets. **C.** Summed ligand intensities and total number of ligands identified from recombinant protein made in Hi5 cells.

3.10 The W62R HLA-F mutant is not able to signal through KIR3DS1

Our data suggested that MHC-F molecules natively containing R62 are able to present peptides in contrast to what was proposed by Dulberger *et al.* To further confirm our findings, we tested

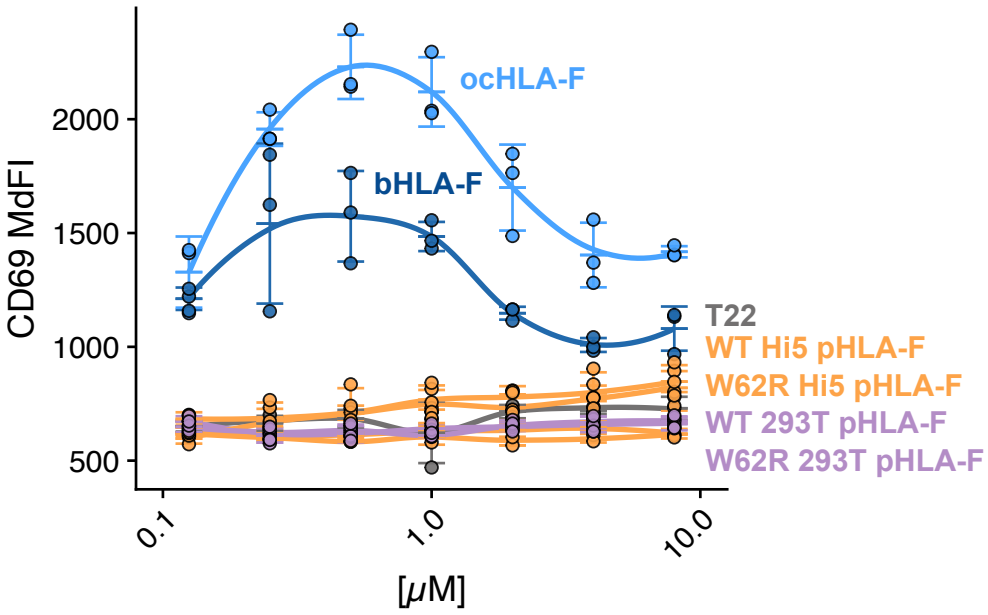

Figure 3.15 - W62R HLA-F mutants do not activate through KIR3DS1.
KIR3DS1:CD3ζ-engineered Jurkats were used for activation assays. Results show that only bacterially derived HLA-F is able to activate engineered Jurkats. Biotinylated monomers were immobilized onto traptavidin-coated plates at the indicated concentrations. Jurkats were incubated in wells overnight and cell-surface CD69 was used as a proxy for activation.

whether W62R HLA-F was able to signal through KIR3DS1, a known activating receptor of empty HLA-F, but not pHLA-F (68). We hypothesized that if W62R HLA-F is biased towards the unloaded state as opposed to an altered peptide repertoire, it should activate through KIR3DS1 to a greater degree than WT HLA-F, which is peptide-loaded when expressed in insect or mammalian cells. To detect a signaling event, we used Jurkat cells that express an engineered KIR3DS1 receptor containing the KIR3DS1 extracellular domain, a KIR3DL1 transmembrane domain, and a CD3ζ intracellular domain. In this system, if Jurkat cells expressing KIR3DS1 recognize extracellular ligands, signal transduction will induce upregulation of CD69 on the cell surface, which can then be assayed by flow cytometry. Ligand was presented in the form of immobilized recombinant protein in flat-well 96-well plates.

In line with our data suggesting that R62 does not bias HLA-F or MHC-F molecules towards an unloaded state, W62R HLA-F was not able to activate through KIR3DS1, whether expressed in insect or mammalian cells (**Figure 3.15**). As expected, WT pHLA-F expressed in either system failed to activate Jurkats, while refolded ocHLA-F and bHLA-F did so successfully. These data add to the evidence that HLA-F and MHC-F molecules encoding R62 are able to present peptides.

3.11 Discussion

HLA-F was long thought to be stably presented on the cell surface as an unloaded heavy chain with or without β2M. We successfully generated reagents that allow us to differentiate between ocHLA-F/bHLA-F and pHLA-F. In contrast to previously published results, using these reagents, we showed that pHLA-F is actually the dominant form of HLA-F expressed on the surface of cells and not ocHLA-F or bHLA-F. Using engineered cell lines, we also showed that β2M is necessary for surface expression of HLA-F and that HLA-F may be dependent on various components of the PLC important in peptide loading of class I MHC molecules. Finally, the fact that the 3D11 antibody appears to be biased and possibly cross-reactive also explains the differences seen in our data and previously published studies. All previous antibodies generated against HLA-F were generated against linear epitopes or refolded, empty material. Thus, a reasonable hypothesis is that reagents used to recognize HLA-F prior to those described here were biased towards detecting ocHLA-F, which could explain why HLA-F was thought to behave differently from other class I HLA molecules. While we could detect either ocHLA-F

and/or bHLA-F on the surface of activated T cells, previous studies have shown that this observation is not unique to HLA-F and has also been observed for other HLA molecules (192–194). This suggests that there is something unique about HLA expression on activated T cells that either allow for misfolded HLA-I molecules to escape regulation by the PLC or that they are more likely to lose their association with β2M and peptide. Of these two possible explanations for HLA-F, our results in IFN-γ induced β2M KO and HLA-F KO 293T and β2M KO Jurkat cells are more consistent with the latter.

Dulberger *et al.* hypothesized that MHC-F molecules may vary in their ability to present peptides based on whether they encode W62 or R62. Using crystal structures and LC-MS/MS we showed here that peptide presentation between HLA-F, Patr-F, and Mamu-F is relatively similar. All species tested presented peptides regardless of whether they encode W62 or R62. Furthermore, all orthologs exhibit similar or worse stability when refolded from bacterially-derived inclusion bodies without peptide. When expressed in insect cells and loaded with peptides, stability was greatly enhanced. Although crystal screens were conducted using refolded Patr-F and Hyla-F with and without β2M, no crystals were observed. Yet, a clear peptide binding profile could not be identified from eluted insect cell peptides. All of our attempts to refold HLA-F with a single synthesized peptide have failed. One possible explanation may be that overexpressed HLA-F picks up many low affinity peptides that are not amenable to refolding. In order to remedy this, one possible alternative would be to extract endogenously expressed HLA-F from the surface of cells using pF4. However, considering that HLA-F is expressed at least one order-of-magnitude lower than other class I molecules (**Figure 1.4**), this approach would be quite labor intensive.

The results presented here have several limitations. One is that only immortalized cell lines were engineered and investigated in this study. It remains to be determined whether primary cells express different patterns of cell surface HLA-F, either endogenously or upon induction by IFN-γ or viral infection. Another limitation is that our studies of the dependency of HLA-F on PLC components is preliminary. While transfected cells were sorted and expanded, single-cell clones need to be generated and validated prior to making any quantitative conclusions regarding the dependence of HLA-F dependence on the PLC. Lastly, our analysis of MHC-F molecules and their stability relative to HLA-F was largely qualitative. Experiments were not carried out in such a way that allowed quantitative statements. However, based on the available data, the major conclusion that residue 62 is a poor predictor of stability with or without peptide ligand is robust because each ortholog tested showed enhanced stability when expressed in insect cells, presumably because they were loaded with peptides. Thus, considering all available data, we hypothesize that the predominant form of HLA-F on the surface of cells if peptide-loaded under basal conditions, but that in different contexts, such as lymphocyte activation, ocHLA-F can also be expressed on the cell surface, opening up the possibility for interactions with KIR3D molecules.

3.12 Materials & Methods

Generating, expressing, and purifying conformation-specific Fab reagents

We designed a novel protein expression construct and purification strategy to yield ocHLA-F and pHLA-F made in Hi5 cells. HLA-F made in Hi5 cells had enhanced stability in comparison to protein expressed in *E. coli* and refolded from inclusion bodies. The construct design was similar to what has been reported previously (68) with one exception: a 3C protease cleavage site was added within the glycine-serine linker connecting β2M to HLA-F. Expression yield was comparable to the construct without the cleavage site. Following Ni-NTA affinity purification from clarified Hi5 supernatant, the elution was split in half. For pHLA-F, one half of this was specifically *in vitro* biotinylated at a C-terminal AviTag site, incubated overnight at 4ºC, and then purified over an S200 (see **Chapter 2 Materials & Methods**) by SEC into 10 mM HEPES 150 mM NaCl 0.02% azide (1X HBS). For ocHLA-F, one half was treated with 3C protease for at least 24 hours at 4ºC. Following 3C treatment, the protein solution was diluted into freshly prepared 8M urea and purified by reverse-phase chromatography over a C4 column using an acetonitrile gradient to separate the light chain from the heavy chain. Fractions corresponding to the HLA-F heavy chain were then dialyzed back into 1X HBS, *in vitro* biotinylated, and purified over S200.

Tomasz Ślężak of Dr. Tony Kossikoff's lab then used these biotinylated proteins to select for high affinity synthetic Fab reagents by phage display as described (195). Briefly, to generate F-Fabs, phage were first negatively selected using pHLA-F followed by positive selection with ocHLA-F. To generate pF-Fabs, phage were first negatively selected using ocHLA-F followed by positive selection with pHLA-F. After 4-5 rounds of negative and positive selection, a phage ELISA was used to assess the specificity and relative affinities of each pure Fab population.

Synthetic Fab heavy and light chains were then cloned into a single bacterial expression plasmid (either with or without a C-terminal AviTag) for periplasmic protein expression as described (195). 2xYT media was used to grow 1 L of each Fab and each expression was induced with 1 mM IPTG once the culture density reached OD_{600} ~0.6-1.0. After 4 hours of expression, cell pellets were collected by centrifugation and lysed by sonication. Following sonication, lysates were clarified and further purified by incubation at 65ºC for 30-45 minutes to remove heavy and light chain homodimers or other less stable protein products. Finally, lysate was clarified once more prior to purification by protein L column and eluted by 10% acetic acid directly into 1M Tris-HCl pH 8.0. Protein was dialyzed back into HBS and, for those Fabs with an AviTag, *in vitro* biotinylated with recombinant BirA ligase made in-house. Biotinylated Fabs were then purified once more over S200 by SEC. Biotinylated Fabs were used to make PE, APC, or BV421-labeled tetramers at a monomer concentration of 2 μM. For flow cytometry experiments, staining concentrations of 50 nM were used (1:40 dilution). Staining was performed in the dark, at room temperature, for 30 minutes, in 50 μL volumes.

Generation of CRISPR-Cas9 mediated knock out cell lines

Single guide RNAs (sgRNAs) were either utilized from previous publications (see **Table S4**) or designed to target HLA-F specifically and not other HLA genes using CRISPOR (196). Forward and reverse primers were then used to clone in the sgRNA of interest into p458X-GFP, which expressed the sgRNA of interest, an engineered SpCas9 enzyme, and a GFP reporter. pSpCas9(BB)-2A-GFP (PX458) was a gift from Feng Zhang (Addgene plasmid #48138 ; http://

n2t.net/addgene:48138). WT 293T/17 cells were transfected with p458X-GFP containing the β2M or HLA-F gRNAs (see **Table S3**) using PEI and two days later single-cell sorted on the top 1% of live, GFP positive cells into 96-well plates. Expanded clones were later validated as successful knock out lines by Sanger sequencing and cell surface staining using the W6/32 antibody (β2M KO lines) or by Sanger sequencing and cell surface staining using pF4 tetramers (HLA-F KO lines) (**Figure S3.1)**.

WT 721.221 and JY cells were nucleofected according to the nucleofection protocol found in the Lonza Cell Line Optimization Nucleofector Kit (Nucleofector 2b, Lonza) using the T-016 program with gRNAs targeting either β2M or HLA-F (see **Table S4**). After 24 hours, cells were bulk sorted on the top 1% of live, GFP-positive cells using the MACS Tyto FACS instrument. Cells were then expanded for 2-3 weeks, but only 721.221 and JY HLA-F and β2M KO lines were used to seed wells of 96-well plates by single-cell dilution cloning. Individual clones were then expanded and verified using a combination of Sanger sequencing, Western blot, and flow cytometry (**Figure S3.2**). Generation of polyclonal TAP1 and TPN KO lines were generated similarly, but single-cell clones were never generated.

Flow cytometry experiments

Antibodies used include W6/32 (Biolegend, anti-pan class I HLA complex) and 3D11 (BioLegend, anti-HLA-F). W6/32 was used at a dilution of 1:100 and 3D11 was used at a concentration of 1:20, unless otherwise noted. Tetramer staining using Fabs was done at 50 nM monomer concentrations. For staining, cells were aliquoted into 96-well plates and washed once

in 1X PBS + 3% FBS (flow buffer). Cells were then resuspended in 50 µL of prepared staining buffer containing the antibody/tetramer cocktail and incubated at room temperature for 30 minutes. Cells were then washed 1X in 200 µL flow buffer and resuspended in 200 µL flow buffer prior to data collection on the ThermoFisher Attune NxT flow cytometer. For live/dead staining, Zombie dyes were used prior to antibody staining by incubating cells in PBS for 30 minutes at room temperature, or DAPI was used post antibody staining at a 1:2000 dilution in 1X PBS for 5 minutes at room temperature. Negative controls include relevant isotype controls from Biolegend or PE, APC, or BV421-labeled streptavidin (SA) used to make tetramers of F1 and pF4.

For IFN-γ induction, cells were expanded to 60-80% confluency prior to exchanging spent media for fresh media that contained 10 ng/mL IFN-γ (treatment) or media only (control). Cells were then incubated for 48 hours at 37ºC prior to harvesting for antibody staining and flow cytometry.

Biolayer interferometry

Binding experiments were done by immobilizing biotinylated F1 or pF4 on a streptavidin (SA) tip to ~3 nm. Hi5 expressed and purified, β2M-fused HLA-F, Patr-F, or Popy-F concentrated to 1 µM were used as analytes and binding was measured relative to buffer blank controls (1X HBS).

For 3D11 binding experiments, unconjugated 3D11 or 26.5 (Biolegend) antibodies were immobilized onto Anti-mouse Fc capture (AMC) sensor tips to a response of ~1 nm. Hi5-derived

ocHLA-F, bHLA-F, pHLA-F, and bacterially-loaded MR1 was used at an analyte concentration of 1 μM. For the reverse experiment, bHLA-F or pHLA-F were immobilized onto Ni-NTA sensor tips to a response of ~3 nm. ocHLA-F was not used because it did not contain a multi-histidine tag. 3D11 and a corresponding isotype (Biolegend) antibody were used at an analyte concentration of 833 nM. All binding curves were buffer subtracted.

mRNA cloning of non-human primate homologs

BLCLs were generated of two chimpanzee (Peony and Razi), two gorilla (Oko and Shamba), two orangutan (Biji and Molec), two gibbon (Buddy and Vernette), and two macaques (Thian and Xuan) by the Parham (Stanford) and Gilad (University of Chicago) labs. Cell lines were cultured in RPMI 1640 + 10% FBS without antibiotics. 10e6 cells of each species were harvested, washed, and lysed using TRIzol and a Direct-zol RNA Miniprep kit kit to purify total RNA. 500 ng of total RNA was then used to generate cDNA libraries (Maxima First Strand cDNA Synthesis kit). 1 μL of each cDNA library was then used as a template for PCR using Q5 polymerase and primers designed by Rojo *et al.* (197) to amplify HLA-F homologs from chimpanzee, gorilla, orangutan, gibbon, and macaque. PCR products were gel purified and TA-cloned into a TOPO vector (Invitrogen TOPO TA Cloning Kit). White colonies were sequenced using M13 forward and reverse primers. Amino acid sequence alignments and and a phylogenetic tree using sequences of HLA-F homologs derived from Dulberger *et al.* were used to confirm that the homologs were truly of chimpanzee, gorilla, orangutan, gibbon, or macaque origin.

Cloning, expression, purification, and crystallization of non-human primate homologs

Soluble expression constructs of Peony-F (Patr-F), Oko-F (Gogo-F), Molec-F (Popy-F), Buddy-F (Hyla-F), and Thian-F (Mamu-F) were cloned into pET30a and pAcGP67a for bacterial and insect cell expression, respectively. All bacterial expression constructs were GC optimized, similar to what was described in the **Chapter 2 Materials & Methods** section for HLA-B*73:01. Insect cell expression vectors were N-terminally fused to β2M similar to HLA-F from Dulberger *et al.* (68) Bacterial and insect cell expression and purification was done as previously described for HLA-F (68). For crystallography, insect cell expressed protein was treated for 2 hours with EndoF enzyme expressed and purified in the Adams Lab prior to a second Ni-NTA purification to remove non-HisTagged EndoF. Hits for PeonyF were obtained and optimized in 2% PEG 400, 2M NaSO4, 0.1M Tris-HCl pH 8.0 and hits for Thian-F were obtained and optimized in 0.1M MES pH 6.0 and 20% PEG 4000.

Diffraction data was collected at the Advanced Photon Source at Argonne National Laboratory on beamlines 23-ID-D and 24-ID-E. CCP4, Coot, and PHENIX software suites were used for molecular replacement and model refinement. PyMol and ChimeraX were used to make figures and measure atomic distances.

In vitro T cell activation assay

Fresh (isolated same day) blood was used to enrich PBMCs by density gradient (Ficoll). PBMC aliquots were then frozen and stored in 100% FBS in liquid nitrogen storage until use. The day before thawing PBMCs, flat 96-well plates were coated with 100 µL of anti-CD3 (OKT3)

antibody diluted to 10 µg/mL in PBS and incubated at 4ºC overnight. For *in vitro* activation, PBMC aliquots were thawed and changed into media containing RPMI1640 + 10% FBS + 100 U/µL IL-2 + soluble anti-CD28 (CD28.2) antibody diluted to 2 µg/mL. Antibody solution was removed from coated plates and 200 µL of PBMCs at 1e6 cells/mL were aliquoted into appropriate wells. Cells were then incubated at 37°C, in 5% CO_2 for 5 days prior to harvesting for staining and data collection on the ThermoFisher Attune NxT flow cytometer.

KIR3DS1:CD3ζ Jurkat signaling assay

KIR3DS1:CD3ζ expressing Jurkat cells provided by the Altfeld lab were used to assay for signaling using protein-coated plates. Flat 96-well plates were coated with 2 µg/mL streptavidin (made in the Adams Lab) diluted in PBS and incubated at 4ºC overnight. Following coating with streptavidin, wells were washed once with 200 µL PBS and coated with C-terminally biotinylated ligands by incubating plates overnight at 4ºC. Plates are now ready to use and should be used within 1 week. Jurkats were expanded in RPMI1640 + 10% FBS. Once plates were ready, cells were harvested, resuspended in fresh media, and 50,000 cells in 200 µL were aliquoted into protein-coated wells. Plates were then incubated at 37ºC in 5% CO_2 for 16 hours prior to harvesting, staining with anti-CD3 (UCHT1) and anti-CD69 (FN50), and data collection on the ThermoFisher Attune NxT flow cytometer.

3D11 immunoprecipitation for mass spectrometry

Immunoprecipitation was performed largely as described (198). Briefly, ~25e6 WT, β2M KO, and HLA-F KO 721.221 cells were expanded, washed with PBS, and frozen at -80ºC until

needed. Cells were lysed using NP-40 based lysis buffer and mixed with 3D11-coupled or isotype control-coupled dynabeads for 18-24 hours at 4ºC while mixing. Beads were then washed and protein was eluted using 500 µL ammonium hydroxide. These samples were submitted for mass spectrometry where samples were concentrated by SpeedVac and trypsinized prior to injection.

HLA-F MHC-F peptide elution for LC-MS/MS

Proteins were expressed in Hi5 cells and purified as described (68). Once purified, at least 200 µg of protein per sample was mixed 1:1 with 20% acetic acid and boiled at 90ºC for 15-20 minutes. Samples were left to cool to room temperature for several minutes prior to filtering using a Centricon 3 kDa centrifugal filter. Samples were centrifuged for 45 minutes at 4000xg. Flow through was then frozen at -80ºC until ready for concentration by SpeedVac and downstream immunopeptidomics by mass spectrometry.

3.13 Supplementary materials

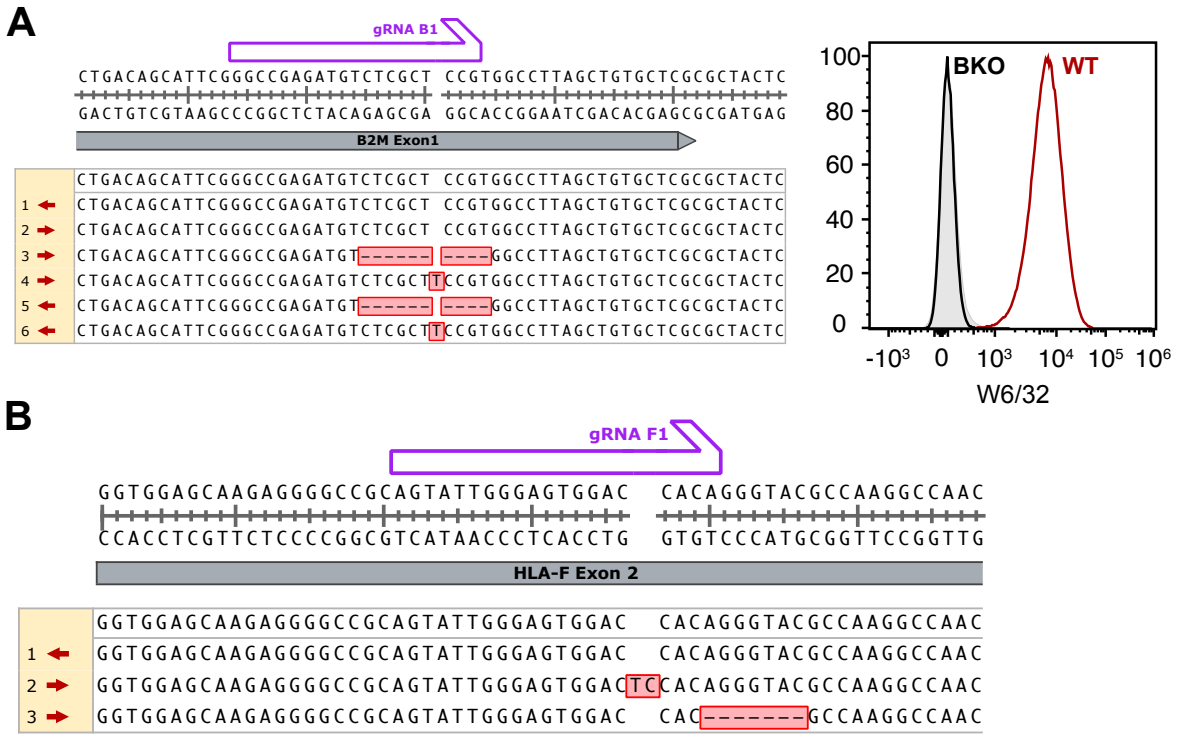

Figure S3.1 - Validation of 293T β2M and HLA-F knockout lines.
A. Sanger-sequencing data shows that while in the WT 293T cells, β2M remains intact (#1 and 2), in a β2M knockout clone, indels have disrupted the open reading frame (#3-6). Flow cytometry confirms that this clone does not express cell surface HLA class I as detected by W6/32 (right panel). **B.** Sanger-sequencing data shows that while in WT 293T cells, HLA-F remains intact (#1), in an HLA-F knockout clone, indels have disrupted the open reading

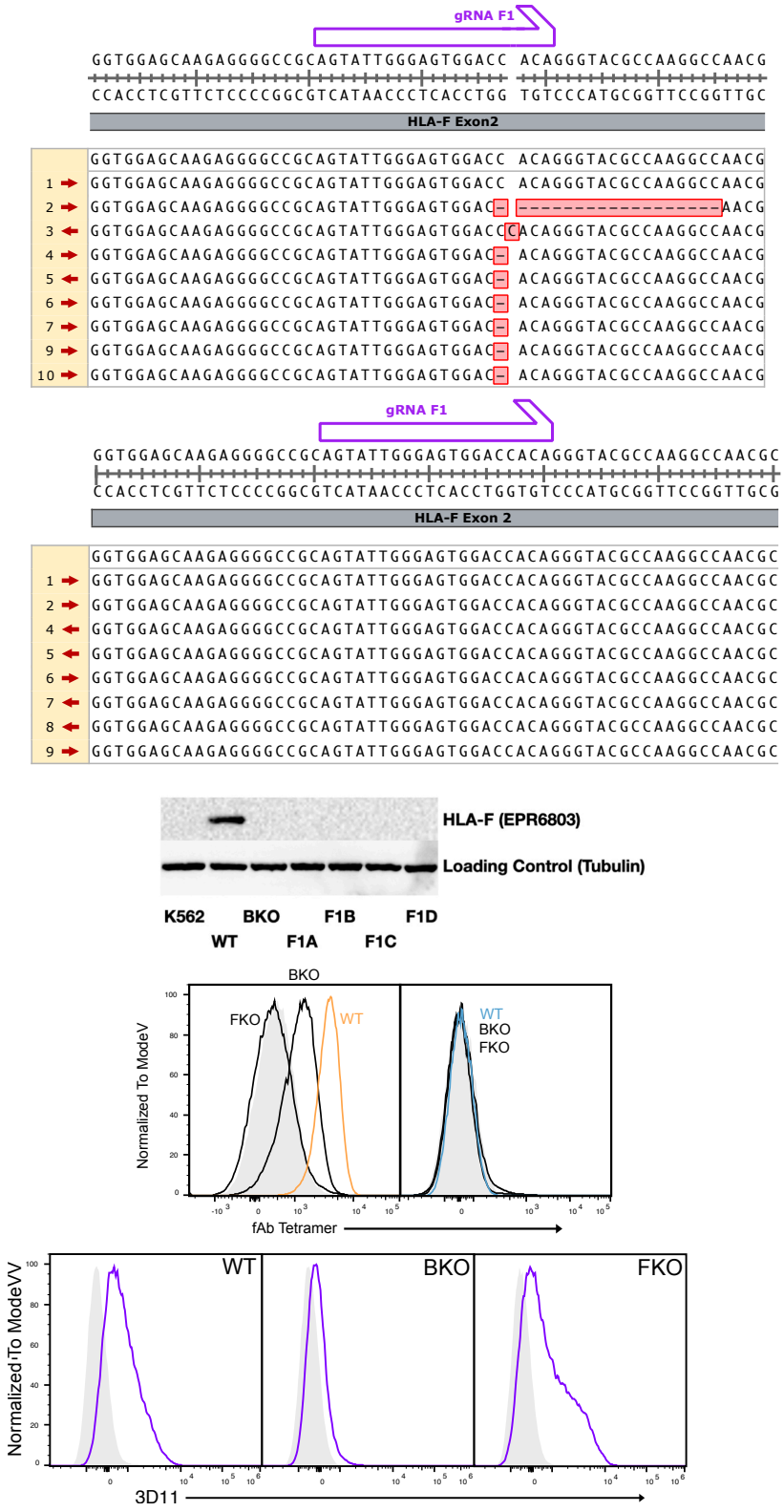

Figure S3.2 - Validation of 721.221 β2M and HLA-F knockout lines.

Figure S3.2 Continued

A. Sanger sequencing results of amplified genomic DNA from WT (1), HLA-F KO 1A (1,2), 1B (3,4), 1C (5,6), and 1D (7,8) 721.221 cells. Sanger-sequencing shows indels that disrupt the open reading frames of HLA-F only in HLA-F knockout clones (top row). Sanger-sequencing shows that in a β2M knockout clone of 721.221 cells, the HLA-F locus remains intact (middle row). Western blot for HLA-F shows that only in WT 721.221 cells is HLA-F protein still detected. Flow cytometry shows that while pF4 staining is completely ablated in a 721.221 HLA-F knockout clone, pF4 stains a β2M knockout clone, but with less intensity. The 3D11 antibody on the other hand continues to stain HLA-F knockout cells, but shows reduced staining of β2M knock out cells (bottom row).

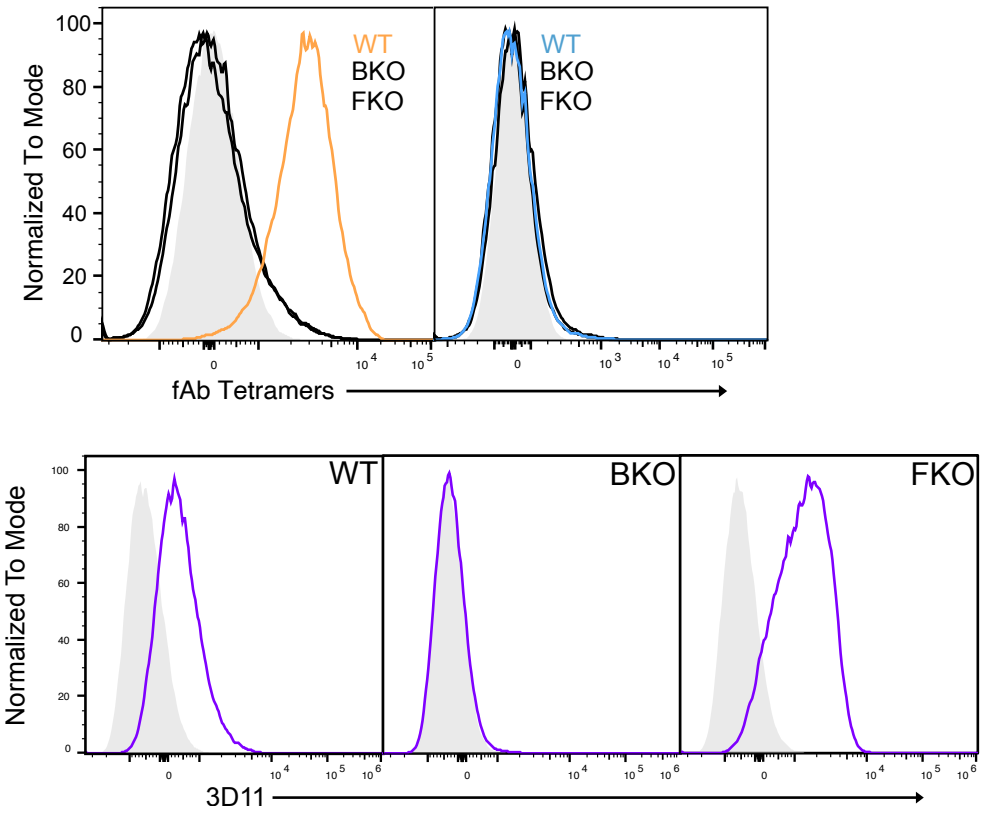

Figure S3.3 - Putative JY cell β2M and HLA-F knockout clones do not stain with F1 or pF4 fAbs.

Upon staining with F1 or pF4, only WT JY cells are positive for pF4. Once β2M or HLA-F have been knocked out, staining is completely ablated (top row). 3D11 shows a similar staining patter in JY cells as it does in 721.221 cells. While it stains WT and HLA-F knockout cells, it fails to stain cells without β2M.

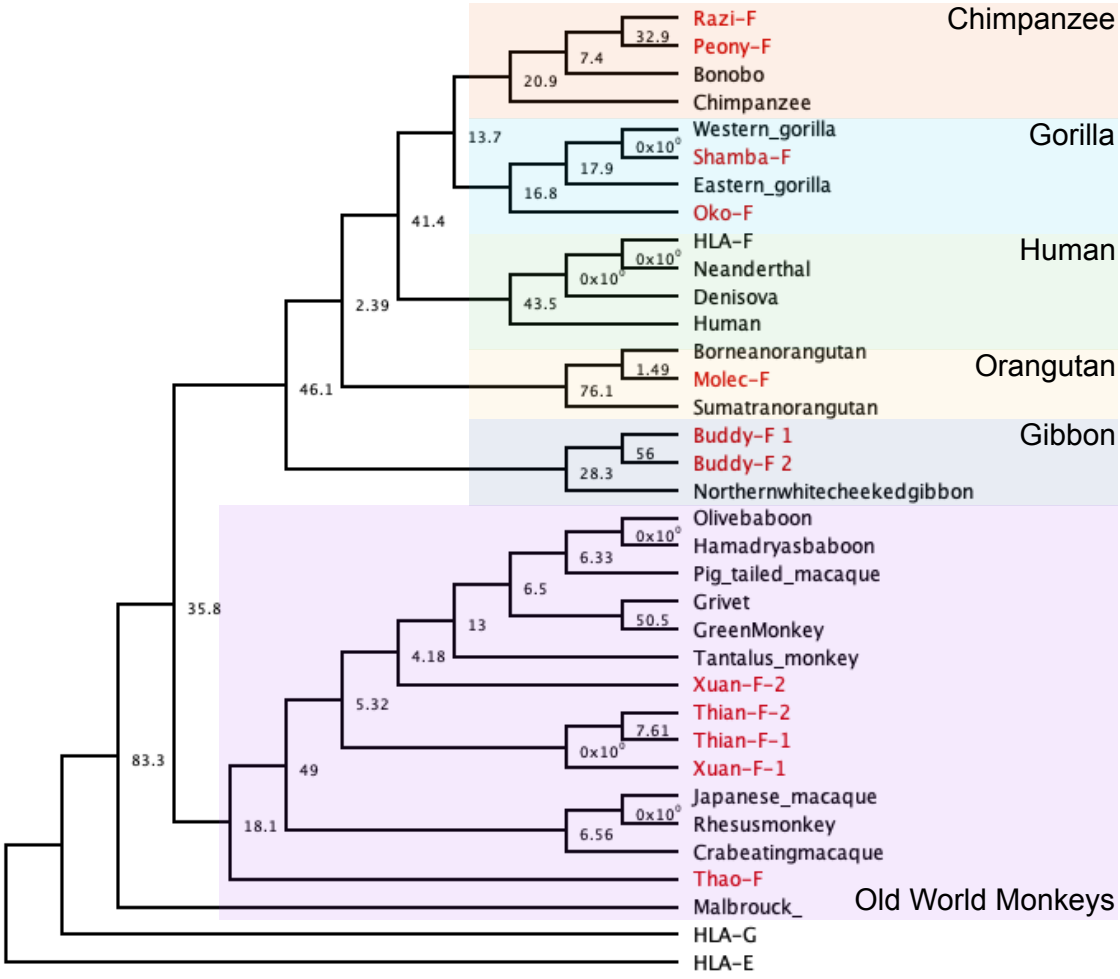

Figure S3.4 - Cladogram of HLA-F and its orthologs.
Full-length sequences from indicated sources were aligned using MUSCLE and phylogenetic tree generated using PAML. Numbers near ancestral nodes represent likelihood ratios. Tips labeled in red are sequences cloned in this study. Sequences in black are sequences that have been previously published (61).

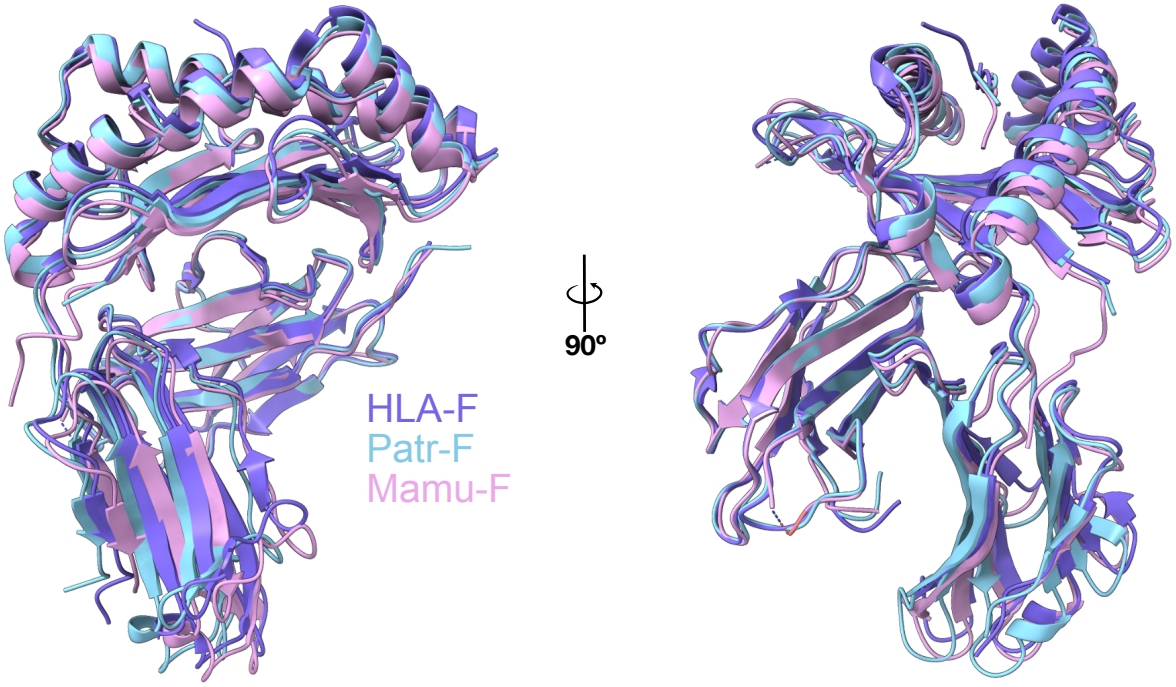

Figure S3.5 - Structural alignment of HLA-F, Patr-F, and Mamu-F.
A structural alignment of HLA-F, Patr-F, and Mamu-F show that their overall structures are highly similar. Structures were aligned using alpha carbons of the β2M backbones. Alignment RMSD between HLA-F and Patr-F was 0.32Å while between HLA-F and Mamu-F it was 3.77Å.

Chapter 4 - Regulation of HLA-F expression in human pregnancy

4.1 HLA-F as a novel regulator of human fecundability

The maternal-fetal interface is a composite of fetal and maternal cells where the fetal-derived placenta invades maternal tissue and anchors directly into the decidualized endometrium or decidua (105). The tips of placental villi become EVTs that invade deeply into the decidua and line maternal spiral arteries that provide the fetus with essential nutrients (105). During this process, EVTs communicate with dNK cells, maternal lymphocytes that play a crucial role in establishing fetomaternal tolerance, by expressing a unique combination of HLA ligands: HLA-C, -E, and -G (117,199). As mentioned in **Chapter 1.4.2,** HLA-F has also been implicated in playing an important role in pregnancy and its protein expression has been observed on both fetal and endometrial tissues. While its expression on EVTs remains unclear, evidence is starting to build that endometrial expression of HLA-F is associated with implantation, and subsequent pregnancy establishment, referred to as fecundability.

Studies of HLA in pregnancy have largely focused on EVTs because of data that supports their interactions with maternal lymphocytes as important drivers of pregnancy outcomes. More recently, studies have also focused on HLA expression in maternal tissues at the maternal fetal interface. In 2016, an important study by Burrows *et al.* test for associations between genome-wide quantitative trait loci (eQTLs) in mid-secretory phase endometrium and fecundability, measured as the time to a positive pregnancy test among women who were not using any contraceptive methods (127). They discovered several fecundability-associated eQTLs for *HLA-F*. Several of these eQTLs were located in the *HLA-F* promoter, but the most significant eQTL,

rs2523393, was located 10kb downstream of the *HLA-F* coding sequence, suggesting that this variant or a variant in strong LD with rs2523393 may reside in a sequence that regulates the transcription of *HLA-F*. Importantly, this effect was only significant for the maternal, not the paternal, genotype, indicating that only maternally-expressed *HLA-F* was relevant. The authors replicated this effect in an independent study of women (130). Together, these studies provide robust evidence for a link between *HLA-F* and pregnancy success.

Follow-up work confirmed that HLA-F protein is detectable in the endometrium by IHC (128) and flow cytometry (129) and that rs2523393 is located within a GATA2 transcription factor binding site (200). This GATA2 binding site was shown to be within a progesterone-responsive enhancer, and is responsive to progesterone signaling when transfected into endometrial stromal fibroblast (ESF) cell lines. Progesterone is a key hormone that is dynamically regulated throughout the menstrual cycle. Its abundance peaks during the secretory phase of the menstrual cycle and is essential in promoting the differentiation of the endometrium into the embryo-receptive decidua so that implantation can successfully occur (201,202). Consequently, *HLA-F* transcription was also shown to be regulated by GATA2 and progesterone receptor, suggesting that *HLA-F* expression also peaks during the "window of implantation" and could regulate maternal immune cells to facilitate successful implantation during this pivotal period.

Here, we investigated where HLA-F protein is being expressed, in what form HLA-F is being expressed, and how expression is regulated in the context of human pregnancy. To do this, we used a suite of protein reagents to detect HLA-F and a panel of cell line reagents to apply them

to. We showed that EVTs likely do not express *HLA-F* transcript or protein, whereas ESFs express cell surface pHLA-F. Furthermore, pHLA-F expression is regulated by hormones and cytokines known to play important roles during implantation. We then showed using HLA-F tetramers that decidual macrophages as the primary decidual immune cell that recognizes HLA-F expressed in the endometrium (108,109). Together, these findings lay the foundation for future work on understanding how HLA-F expressed in ESFs may successfully prime uterine leukocytes such that they can facilitate early pregnancy establishment and maintenance.

4.2 HLA-F is expressed on some immortalized trophoblast cell lines, but not primary trophoblasts

First we confirmed which cells at the maternal-fetal interface express HLA-F protein. Burrows *et al.* found that only the maternal genotype contributed to HLA-F's association with fecundability, suggesting that EVTs are not likely to be the cell type mediating this phenotype (127). However, several studies have reported HLA-F protein expression in EVTs by IHC and Western blot during the first trimester of pregnancy (52,120,121). One possible explanation for this is that EVTs could express a different form of HLA-F relative to maternal tissue and thus contribute to different phenotypes. Alternatively, it could also be that HLA-F expression in EVTs or other trophoblasts is not critical for implantation, but could be for later stages of pregnancy. To test this, we assayed for HLA-F protein expression by flow cytometry on trophoblast cell lines using our Fab reagents in combination with other previously described antibodies.

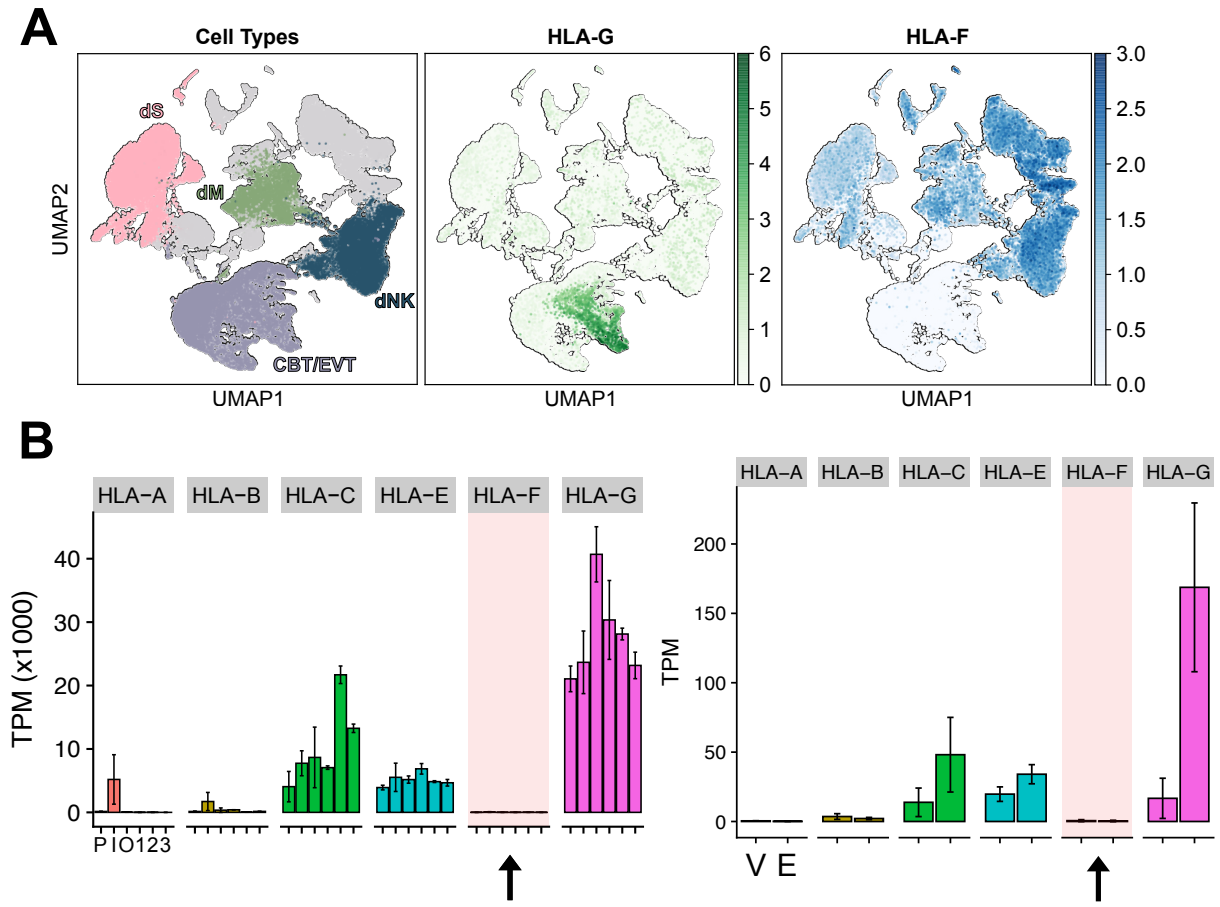

Figure 4.1 - HLA-F transcript is barely detectable in trophoblasts.
A. A reanalysis of data from Vento-Tormo *et al.* (112) shows that HLA-F transcript is expressed in nearly all cells except for trophoblasts. EVTs are marked by expression of HLA-G. Cells not marked by HLA-G expression are either villous cytotrophoblasts (CBTs) or syncytial trophoblasts (STs), which also do not appear to express HLA-F. dS = decidual stromal cells; dM = decidual macrophages; dNK = decidual natural killer cells; CBT/EVT = cytotrophoblasts/extravillous trophoblasts. **B.** Two bulk RNA-seq data sets (224,225) using isolated trophoblast populations show similarly that HLA-F transcript is barely above the limit of detection whereas HLA-C, HLA-E, and HLA-G transcript is abundantly expressed. P = primary EVTs, I = invasive EVTs, O = trophoblast organoids, 1,2, and 3 = differentiated trophoblast organoids, V = villous trophoblasts, EV = Extravillous trophoblasts.

First, we tested for HLA-F cell surface expression on several immortalized trophoblast cell lines. This included the choriocarcinoma cell line JEG3 (203), the large T antigen immortalized cell line, HTR8 (204), and the hTERT immortalized cell line, Swan71 (205). In HTR8 and Swan71 cells, staining was seen mainly by pF4, with little to no staining by 3D11 and F1 suggesting that

these cells are dominantly presenting pHLA-F (**Figure 4.1A**), similar to what was seen in **Chapter 3.4**. However, there was no staining by pF4, F1, or 3D11 in the JEG3 cell line, which best recapitulates the HLA expression pattern seen on primary trophoblasts (206). While these results suggest that HLA-F can be expressed on EVTs, primary trophoblast cells must be used to confirm these findings because trophoblast cell lines rarely recapitulate the HLA expression patterns of *ex vivo* cultured (primary) trophoblast lines (117,206,207). Thus, we also used primary trophoblasts to test for cell surface HLA-F expression. To do this, a collaborator at Cincinnati Children's hospital, Dr. Tamara Tilburgs, kindly provided several primary EVT lines derived from first trimester (elective terminations) and term (healthy labor) stages of pregnancy. These lines were generated as described in (207), but currently remain unpublished (**Table S5**).

Using 6 total primary EVT lines, 3 from derived from first trimester anchoring villi, 2 derived from term anchoring villi, and 1 derived from term chorioamniotic membranes distal to the implantation site (**Figure 4.1B)**, we used 3D11, F1, and pF4 staining reagents in combination with flow cytometry to assay for HLA-F cell surface expression. We also assayed HLA-G cell surface expression using the MEM-G/9 antibody (208) because it is known that primary EVT lines differentially express HLA-G and that this can be dependent on their time in culture. Staining of these primary cell lines showed that while they expressed cell surface HLA-G to varying degrees, we were not able to detect cell surface HLA-F expression using 3D11 (data not shown), F1, or pF4 staining reagents (**Figure 4.1C,D**). Thus, HLA-F expression on the HTR8 and Swan71 cell lines is likely an artifact of long-term *in vitro* culture and/or the immortalization process, an artifact seen also for HLA-A and HLA-B (206,209).

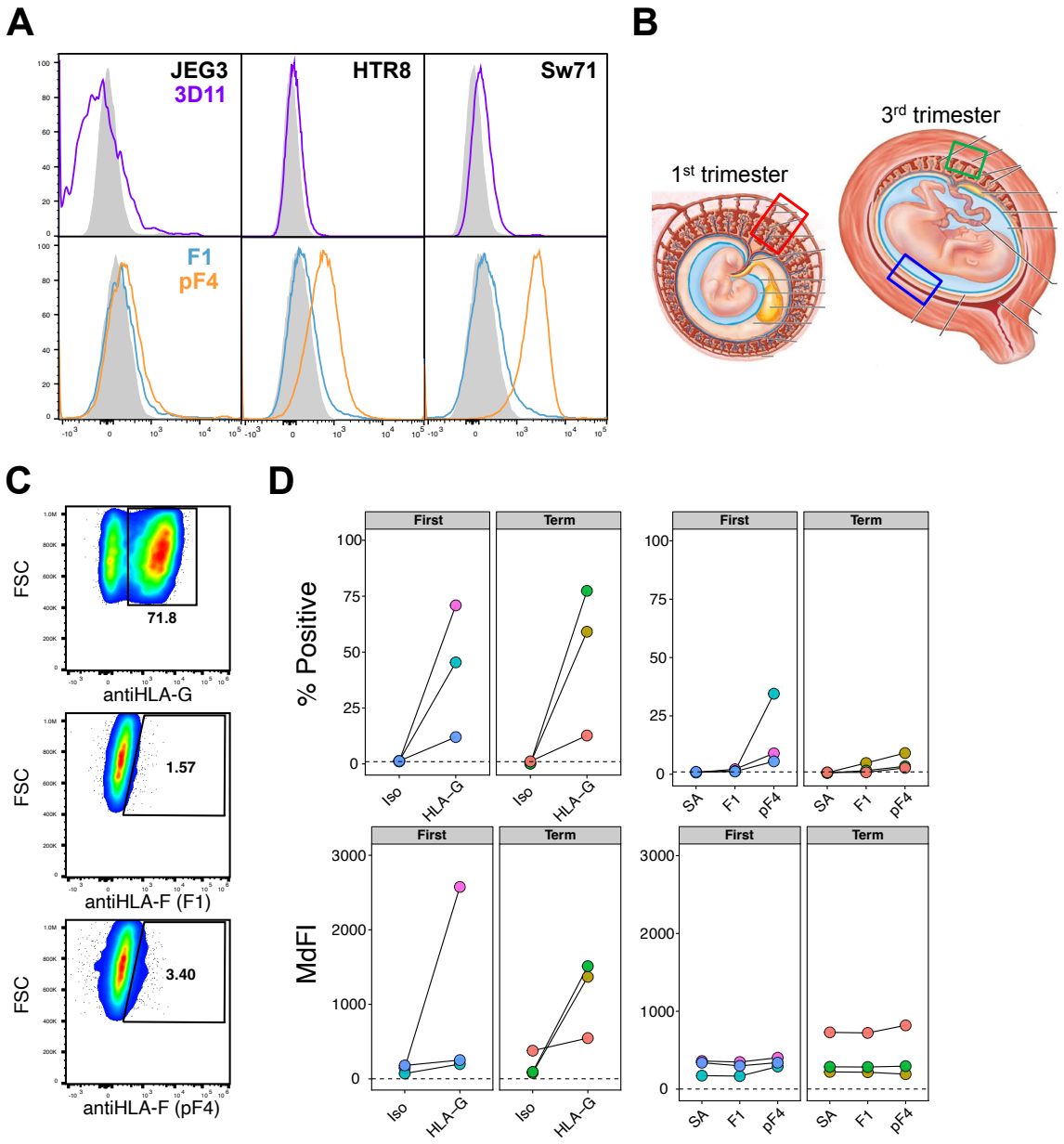

Figure 4.2 - Cell-surface HLA-F is not detectable on primary trophoblasts.
A. Flow cytometry shows cell-surface oHLA-F on several trophoblast cell lines. **B.** Anatomical locations of derived primary trophoblast cell lines are shown. **C.** Representative flow plots of primary trophoblast lines assaying for HLA-G (top), or for HLA-F by F1 (middle) or pF4 (bottom). **D.** Quantifications of flow cytometry experiments showed in panel C. The top row shows cells positively gated while the bottom row shows raw MdFI of controls (Iso or SA) or anti-HLA-G or anti-HLA-F reagents. The first column shows antiHLA-G staining while the second column shows antiHLA-F staining.

One explanation for the lack of HLA-F expression in primary EVTs is that HLA-F is sensitive to *in vitro* culture conditions, and may be downregulated during *in vitro* expansion, similar to what we observed for HLA-G. To complement our flow cytometry findings, we used recently published scRNA-seq data generated using tissue samples from women who underwent elective termination during their first trimester (112). Importantly, these data were obtained from fresh tissue samples and thus RNA expression profiles should more closely represent true *in vivo* expression profiles relative to *in vitro* cultured cell lines. However, one caveat is that these samples may have varied gestational ages that we could not control for. Additionally, the effects we are interested in occur around implantation while these samples were collected weeks afterwards. Nonetheless, in agreement with my data, *HLA-F* transcripts were detectable in nearly all cells except for trophoblasts (**Figure 4.2A**). In fact, HLA-F was lowly expressed or undetectable in all cells of fetal origin relative to cells of maternal origin (**Figure 4.2B**). These data suggest that HLA-F is not expressed on fetal tissue at the fetal-maternal interface during the first trimester, similar to the expression profiles observed for HLA-A and HLA-B.

A known limitation of scRNA-seq is in detecting low abundance transcripts (210). However, when analyzing data from bulk RNA-sequencing experiments, which sorted on different populations of trophoblasts, including EVTs, was detectable at low levels (transcripts per million ≥ 2 (211)) (**Figure 4.2C**) . In contrast, HLA-C, -E, and -G transcripts were expressed to degrees several fold greater than *HLA-F*, consistent with the observation that EVTs preferentially express this combination of HLA molecules. Even HLA-A and HLA-B, which are considered not to be expressed on trophoblasts (117), were expressed to a greater degree than *HLA-F*. Together, these

data point to a lack of HLA-F expression in placental cells. By utilizing newly generated RNA-seq data in combination with our reagents that we validated in **Chapter 3** as sensitive to HLA-F cell surface expression, we were able to more comprehensively test whether trophoblasts can express cell surface HLA-F. These results are in line with findings by Burrows *et al.* which indicated that paternally-expressed *HLA-F* does not contribute to successful pregnancy outcomes. Our data further shows that HLA-F is not expressed in fetal trophoblasts.

4.3 Endometrial stromal cells express cell surface pHLA-F and is regulated by cytokines and hormones known to be important during implantation

Next, we assayed for HLA-F cell surface expression in ESFs. HLA-F protein expression has been detected in endometrial tissue (128) by IHC and more recently by flow cytometry on ESFs (129). However, these data were generated using the 3D11 antibody, which we showed in **Chapter 3.7** was biased towards empty forms of HLA-F and likely binds non-specific targets. Thus, we utilized our new Fab staining reagents to confirm which form of HLA-F was expressed on the surface of two immortalized ESF cell lines. Considering that 3D11 is biased towards empty forms of HLA-F, we expected ESFs to stain with F1 and not pF4.

In contrast to these expectations, we found that immortalized ESFs express cell surface pHLA-F and not ocHLA-F (**Figure 4.3A**). Consistent with our model that 3D11 is biased towards ocHLA-F, we found that staining with 3D11 was barely detectable above background. Western blot was able to confirm that ESFs do indeed express HLA-F protein and glycosidase sensitivity assays confirmed that the majority of expressed HLA-F is found on the cell surface. It is well

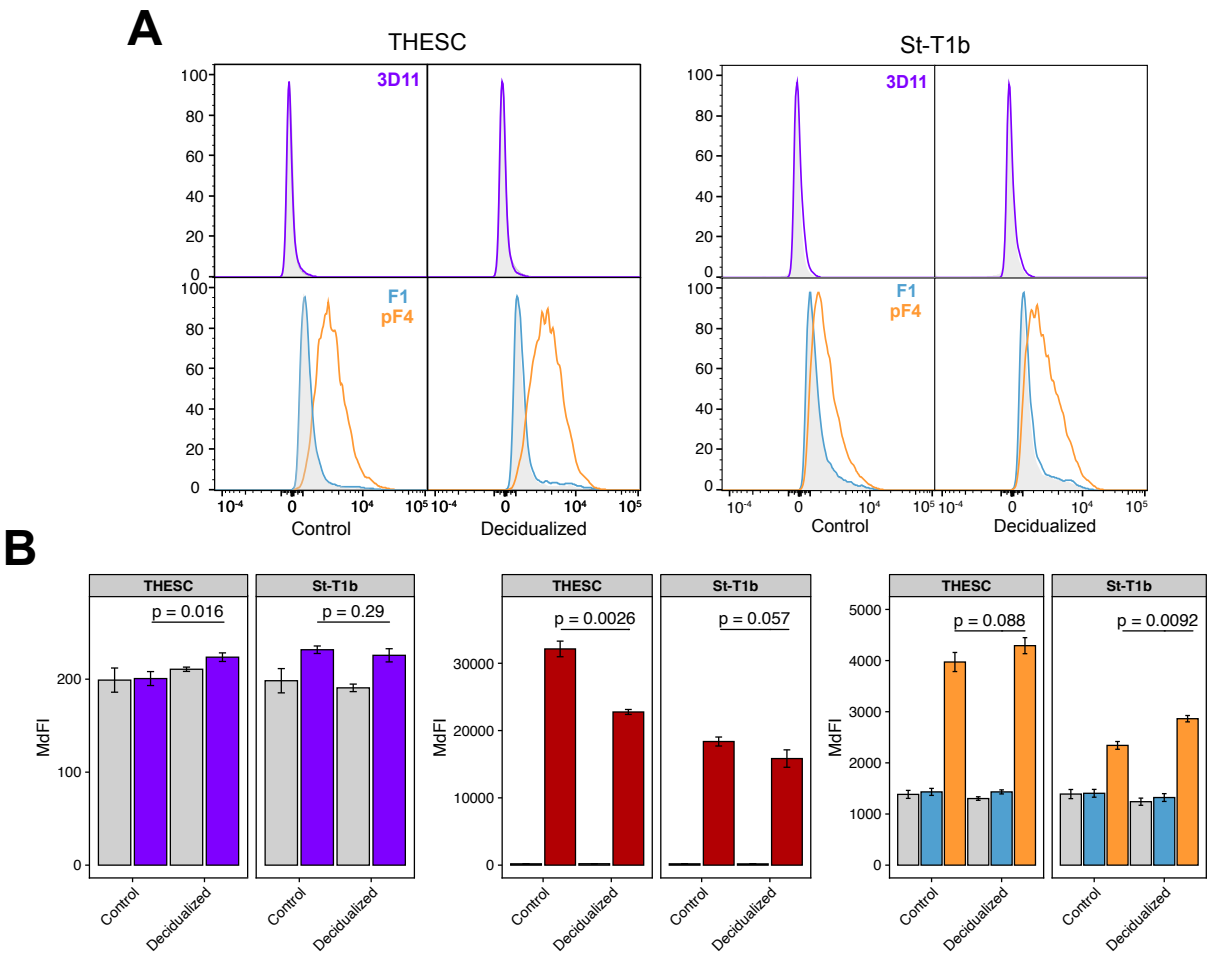

Figure 4.3 - Cell-surface pHLA-F is detected on immortalized ESF cell lines.
A. Representative flow cytometry histograms looking at THESC or St-T1b cell lines stained with the 3D11 antibody or F1 or pF4 tetramers. **B.** Quantified flow cytometry data is shown (n=3 separate cell flasks grown simultaneously). While no significant difference is noted for 3D11 staining, upon *in vitro* decidualization, pan class I HLA expression decreases in both cell lines, while pHLA-F expression increases. A two-tailed t-test was used to calculate p-values.

established that endoglycosidase H, Endo H, sensitivity can be used as a proxy for HLA cell surface expression (212). This is due to its preference for immature glycan modifications on HLAs that indicate that these molecules have not yet made it through Golgi-mediated transport to the plasma membrane. PNGase F, on the other hand, cleaves immature and mature glycan

modifications and serves as a positive control. These experiments therefore confirm that we are detecting true cell surface HLA-F expression.

Next we tested whether HLA-F protein expression on ESFs is upregulated during decidualization, similar to what has been reported for *HLA-F* transcript (213). During the latter half of the menstrual cycle (the secretory phase), progesterone and other hormones are responsible for differentiating (i.e. decidualizing) the endometrium in preparation for implantation. Studies have been able to successfully replicate these conditions *in vitro* using ESFs, which are responsive to stimuli that mimic progesterone and other less-well understood signals that promote decidualization (214). More details on this protocol can be found in **Chapter 4.6 - Materials & Methods**.

Using immortalized ESF cell lines (215,216), we were able to see a consistent increase in pHLA-F cell surface expression following *in vitro* decidualization (1.15 fold-increase in the THESC line; 1.37 fold-increase in the St-T1b line; **Figure 4.3B**). This is likely not due to off-target class we recognition by pF4 because staining using the W6/32 antibody shows a corresponding decrease in the expression of class I molecules in general. Although, the increase in HLA-F expression was only statistically significant in the St-T1b cell line, this is likely due to the small effect size and a larger sample size (i.e. more replicates) could mitigate any statistical uncertainty. Furthermore, an increase in HLA-F protein expression can be corroborated by Western blot (**Figure 4.4C**). Nevertheless, the effect size is quite small, and the fold-increase in

cell surface HLA-F is less than the fold-increase in HLA-F transcript published previously (213) and seen by Western blot.

Similar to the immortalized trophoblast cell lines, immortalized ESF lines may not faithfully recapitulate *in vivo* expression profiles. Thus, we further examined the cell surface HLA-F expression profile in primary ESFs. For this, we expanded first trimester ESF lines from four patients (see **Chapter 4.7 - Materials & Methods**) and assayed cell surface HLA-F expression by flow cytometry, similar to that performed for THESC and St-T1b lines. Similar to the immortalized cell lines, primary ESFs also expressed cell surface pHLA-F, but not ocHLA-F (**Figure 4.4A,B**). pHLA-F expression was increased upon *in vitro* decidualization, although statistical significance was not reached ($p = 0.17$), likely due to a small sample size. Additionally, the fold-increase in cell surface HLA-F following *in vitro* decidualization was similar to what is seen in immortalized ESFs. However, the same decrease in pan-class I surface expression detected by W6/32 in THESC and St-T1b lines was not seen in primary lines. This indicates that the artificial immortalization of ESFs may alter the expression profile of class I HLA relative to an *in vivo* setting, though HLA-F may not conform to this trend.

Interestingly, one ESF cell line derived from patient A63 did not express cell surface HLA-F whether decidualized or not (**Figure 4.5A**). However, it did express other class I molecules as indicated by W6/32 staining. To ensure that primary ESFs are responding as expected to *in vitro* decidualization media, we assayed for changes in the transcript abundance of two markers associated with ESF decidualization: prolactin (PRL) and the insulin-like growth factor binding

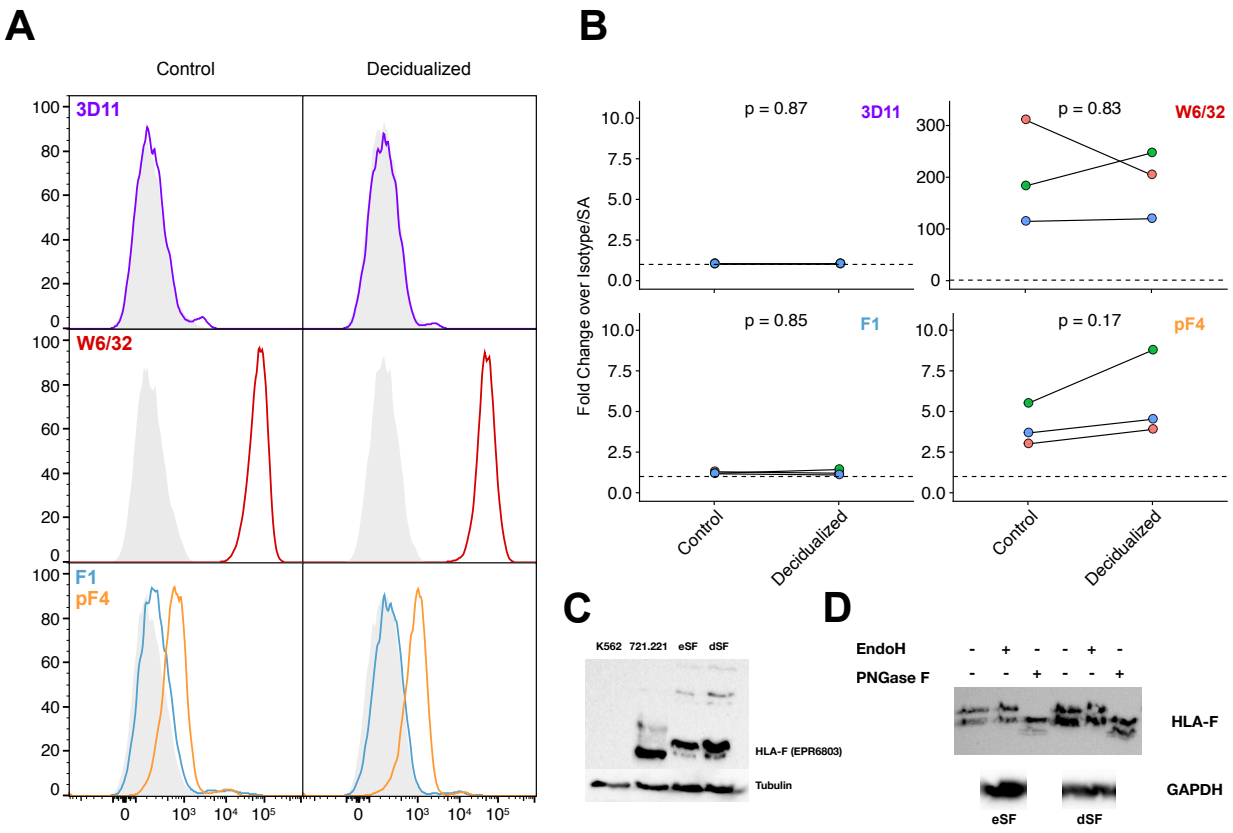

Figure 4.4 - Cell-surface pHLA-F is detected on primary ESF cell lines.
A. Representative flow cytometry histograms looking at the A13 cell line stained with the 3D11 or W6/32 antibody or F1 or pF4 tetramers. **B.** Quantified flow cytometry data is shown for the A13, B13, and B19 cell lines. Each color represents a different cell line derived from a different patient. P-values are derived from a paired sample t-test. **C.** A Western blot of the B19 cell line lysate cultured under control (eSF) or decidualization (dSF) conditions compared to K562 and 721.221 cell lysates. HLA-F expressed in dSF lysate is 1.88 fold greater than in eSF, once normalized by the tubulin loading control. **D.** THESC lysates were first treated with either EndoH or PNGase F to assess the glycosylation status of HLA-F prior to SDS-PAGE and blotting similar to panel C.

protein 1 (IGFBP1) genes (217). The three cell lines in which pHLA-F was detected significantly upregulated both PRL and IGFBP1 following culture in decidualization media, suggesting that these lines are responsive to decidualization stimuli and likely true ESFs (**Figure 4.5B**). However, the A63-derived line did not, suggesting either the presence of a contaminating cell type or a primary ESF line that has undergone genetic changes due to *in vitro* culture. This is not

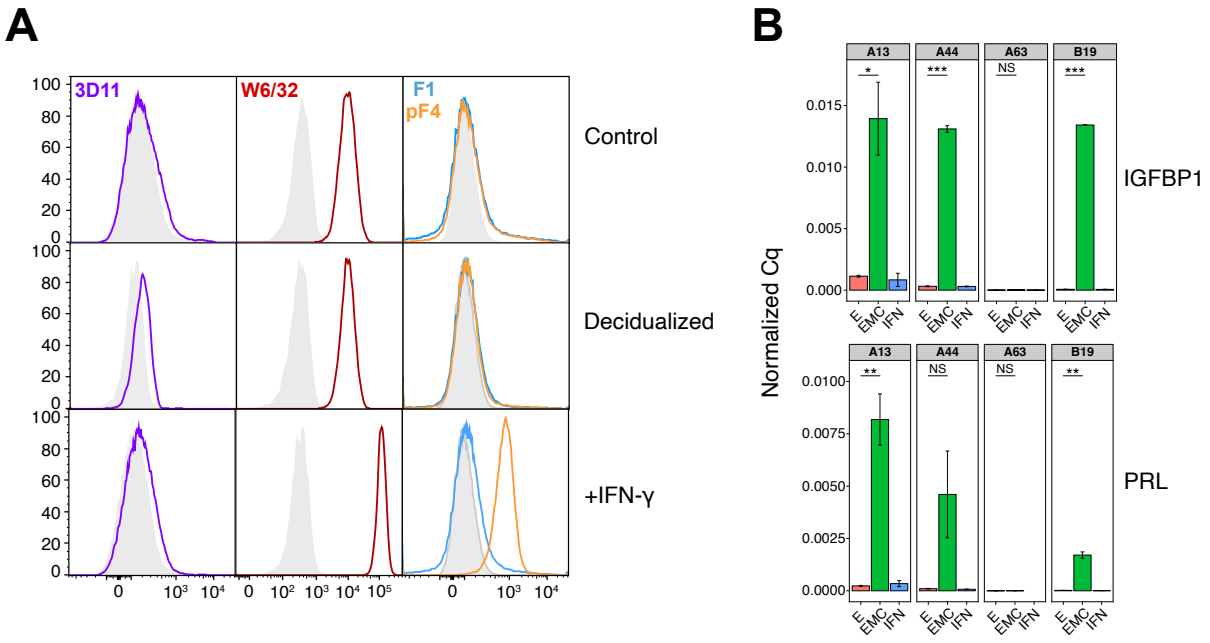

Figure 4.5 - The A63 cell line does not express HLA-F and is not responsive to *in vitro* decidualization stimuli.
A. Representative flow cytometry histograms looking at the A63 cell line stained with the 3D11 or W6/32 antibody or F1 or pF4 tetramers after culture in control media, decidualization media, or in control media with the addition of 10 ng/mL of IFN-γ. **B.** Transcript levels of IGFBP1 (top) and PRL (bottom) in primary expanded ESF lines cultured in control media (E), decidualization media (EMC) or control media with 10 ng/mL of IFN-γ (IFN). Cq values were normalized to L19 RNA, but were similar when normalized to GAPDH RNA (data not shown). A two-tailed t-test was used to determine statistical significance between technical replicates (n=3). * is ≤ 0.05, ** is ≤ 0.005, and *** is ≤ 0.0005.

surprising since the expansion protocol relies on the assumption that only ESFs will successfully expand under the culture conditions utilized. There is no explicit step to positively or negatively select for ESFs. Thus, lines can easily be contaminated by different cell types and it is important to test the purity of expanded cell lines by immunofluorescence and/or qPCR. A63 may actually be an expanded fetal stromal cell line due to its lack of HLA-F cell surface expression as is the case when looking at scRNA-seq data. It should also be noted that upon culture in IFN-γ, pHLA-F expression can be detected, suggesting that this cell line is capable of presenting cell surface

HLA-F, but does not do so under basal conditions. Further experiments will be needed to confirm the cellular identity of line A63.

Finally, we sought to test whether HLA-F expression was responsive to other relevant stimuli to more closely mimic the endometrial conditions during the window of implantation. There are now multiple lines of evidence to suggest that inflammatory cytokines are important in regulating the maternal immune system in the uterus during implantation (218) and that inflammation was actually essential for the evolution of embryo implantation as well (219). For example, studies in mice suggest that IFN-γ is important for appropriate arterial remodeling at the maternal-fetal interface (220). Thus, we tested whether the addition of other relevant cytokines in addition to MPA and cAMP (standard decidualization induction additives) would alter cell surface HLA-F expression. To do this, we tested the addition of interleukin-1-beta (IL1β) and human chorionic gonadotropin (hCG) (221), leukemia inhibitory factor (LIF) (200), and IFN-γ to MPA and cAMP separately or all together. In the two cell lines tested, the effects were consistently in the same direction and pHLA-F expression increased following the addition of cytokines (**Figure 4.6**). Although the greatest change in surface expression was seen in cells treated by IFN-γ, expression increased even further when all factors were included, indicating that the presence of other factors can further increase HLA-F expression beyond the effects of IFN-γ alone.

Together, these data show that cell surface HLA-F is expressed on ESFs and that immortalized ESF cell lines can reliably recapitulate HLA-F expression patterns. Furthermore, in line with

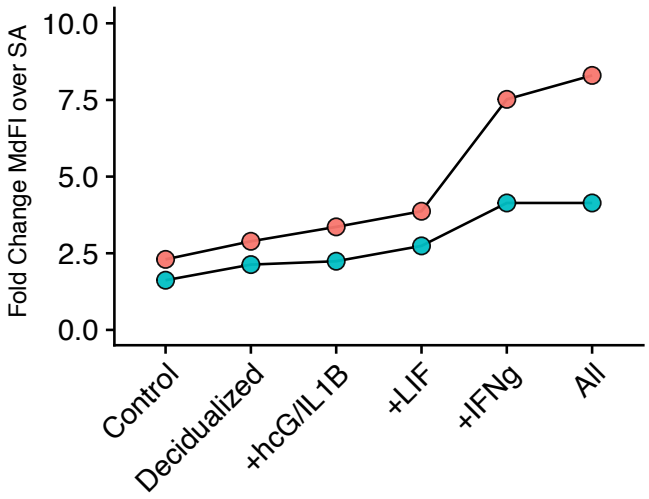

Figure 4.6 - *In vitro* decidualization combined with additional inflammatory cytokines increases cell-surface pHLA-F.
The B13 and B19 cell lines were cultured and subject to *in vitro* decidualization for 48 hours (see **Chapter 4.6 - Materials & Methods**). Other conditions were *in vitro* decidualization media plus the indicated additives. IL1B is interleukin-1-beta; hCG is human chorionic gonadotropin; LIF is leukemia inhibitory factor; IFNg is interferon gamma. Each indicated cytokine has been implicated in successful pregnancy establishment.

previously published data that show that HLA-F transcript is upregulated in ESFs upon *in vitro* decidualization (200), we show here that HLA-F protein is also increased, both within the cell and on the cell surface. However, the small effect size does raise the question of whether this increase is biologically relevant. Lastly, the detection of pHLA-F and not ocHLA-F on ESFs is also a key finding. Considering what we know about the reactivity of ocHLA-F and pHLA-F with various receptors, this data suggests that HLA-F is unlikely to engage with KIRs to mediate its fecundability-associated phenotype.

4.4 Tetramers of ocHLA-F and pHLA-F stain decidual leukocytes

Next we wanted to address what decidual leukocytes can recognize HLA-F. Cells that can recognize HLA-F may act as effector cells that mediate the fecundability-associated phenotype.

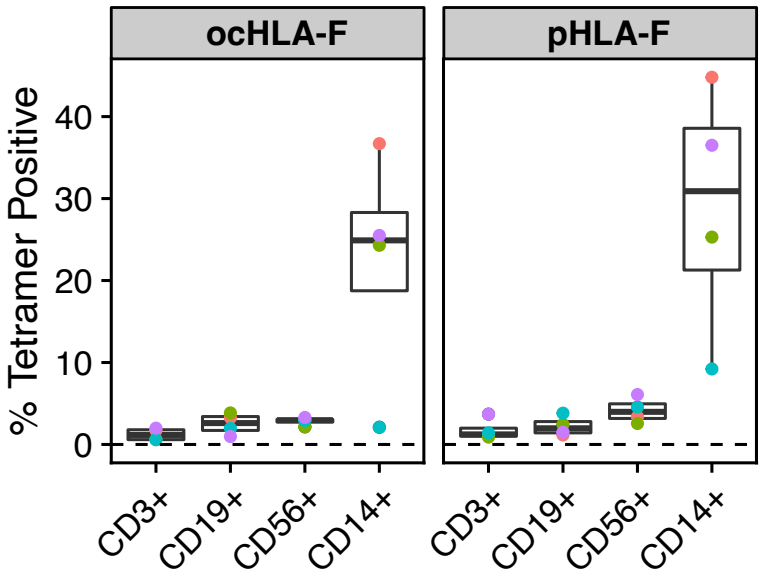

Figure 4.7 - Tetramer staining of decidual mononuclear cells.
ocHLA-F or HEK293T-derived pHLA-F tetramers were used to stain primary decidual mononuclear cells. Each color is a different patient (n=4). CD14+ labeled cells best recognized HLA-F whether using the ocHLA-F or pHLA-F tetramer.

We used a similar source of primary decidual cells as those that were used to expand primary ESF lines (see **Chapter 4.7 - Materials & Methods** for more details). It should be noted that although the influence of HLA-F on fecundability likely matters most during implantation, many hallmarks of decidual leukocyte maturation during the first trimester of pregnancy can already be seen during the secretory phase of the menstrual cycle, whether implantation occurs or not (223). Therefore, first trimester decidual leukocytes are a reasonable proxy for endometrial leukocyte phenotypes present during the window of implantation.

To test which cell types can recognize different forms of HLA-F, we made tetramer staining reagents of bacterially-derived ocHLA-F and HEK293T-derived pHLA-F. As a negative control tetramer, we used the nonclassical mouse MHC, T22, which is known to not present peptide

ligands (66). Upon staining of decidual mononuclear cells derived from four patients, decidual macrophages (C14+ cells) are stained best by both ocHLA-F and pHLA-F tetramers, followed by dNK (CD56+) cells and then decidual T (CD3+) cells (**Figure 4.7**). We hypothesize that this is mainly mediated by binding to LIR1 and LIR2 receptors, which are most highly expressed on macrophages and more lowly expressed on dNK and decidual T cells (112). Although LIR1 has a strong preference for complexed pHLA molecules, LIR2 is more agnostic and can also bind to open conformer HLA molecules (71). Considering that ocHLA-F is a known ligand of KIR3DS1, KIR3DL1, and KIR3DL2 (58), we were confused by the lack of staining of dNK cells by the ocHLA-F tetramer. While this could be explained by the low or lack of expression of these receptors in these patients, it is known that dNK cells can express KIR3DL1 and KIR3DL2 (108,109). These results suggest that decidual macrophages, which have been shown to play important roles in healthy pregnancy outcomes (213), are the most likely effector cells mediating the HLA-F associated fecundability phenotype.

4.5 Discussion

Non-classical class I HLAs are known to play important roles in establishing and maintaining healthy pregnancy. HLA-F has more recently been implicated in playing a role as well. Although it has since become clear that HLA-F expression levels are associated with fecundability, no studies have investigated which cell type or types express and recognize HLA-F. We show here that HLA-F is barely expressed, if at all, in primary EVTs. This is based on data from multiple studies that assayed transcriptome-wide transcript abundance by RNA sequencing (224,225) and my own flow cytometry experiments using immortalized and primary EVTs from first and third

trimester pregnancies. In contrast, ESFs express pHLA-F and hormones and cytokines known to undergo dynamic changes throughout the menstrual cycle regulate its cell surface expression levels. Using tetramers of HLA-F, we showed that decidual macrophages cells are particularly reactive to HLA-F, similar to what has been shown for peripheral blood monocytes and macrophages (57). To test whether HLA-F expression on ESFs has any effect on decidual macrophages, it will be crucial to use HLA-F KO ESF lines in combination with validated HLA-F blocking reagents and cell.

The influence of maternal HLA expression on pregnancy establishment and outcomes has not gotten the same amount of attention as the contributions of fetal HLA expression, especially on EVTs. However, more recent studies have shown that in mice, dNK education by maternal MHC-I is required for optimal fetal growth (125,126). Future studies will need to take advantage of genetic associations in humans, similar to what was done for HLA-F (127), to better understand how or if these findings in mice translate into humans.

One question that arises is why our results differ from previous studies showing that HLA-F is expressed on EVTs. As we showed in **Chapter 3**, the 3D11 antibody is biased and cross reactive. Since other antibodies were similarly generated against HLA-F, they could suffer from similar drawbacks, although that remains to be determined. Additionally, IHC is not as sensitive as flow cytometry and often requires fixing tissue prior to staining cells, which can change epitope availability and thus the reactivity profiles of antibodies. By flow cytometry, we were able to test for the expression of cell surface HLA-F on live cells, which presumably display cell surface

protein epitopes also displayed *in vivo*. Thus, it could be that the antibodies generated to date are not suitable for detecting HLA-F by IHC.

The Adams Lab has previously shown that recombinant pHLA-F fails to signal through KIR3DS1 or KIR3DL2, unlike ocHLA-F (68). Here, we further show that pHLA-F is the only form of HLA-F expressed on ESFs under basal or decidualization conditions. If previous results are any indication, pHLA-F is likely not to interact with KIR3D molecules unless cell type-specific peptides play a strong role. As of now, three studies have sought to elute peptides from HLA-F expressed in mammalian cells. Our work used HLA-F made in 293T cells and two other studies used 721.221 and K562 cell lines (68,177,178). In each study, peptides eluted showed similar length distributions and partly similar C-terminal anchor preferences. Future work will need to verify whether ESF-expressed HLA-F has similar or unique peptide binding preferences.

LIR1 is known to be an important regulator of pregnancy, mainly through its reactivity with HLA-G upon EVTs invasion into the endometrium (79). Thus, one possible mechanistic model of the association between HLA-F and fecundability is that HLA-F expression increases during the window of implantation in a genotype-specific manner and is able to react with LIR1 expressed on decidual macrophages to facilitate successful implantation in the higher expressing genotypes (**Figure 4.8**). However, this model will be difficult to test as LIR1 reacts with most class I molecules, which are also constitutively expressed on ESFs and other nucleated maternal cells. Alternatively, our group showed that the affinity of HLA-F for LIR1 is higher than other

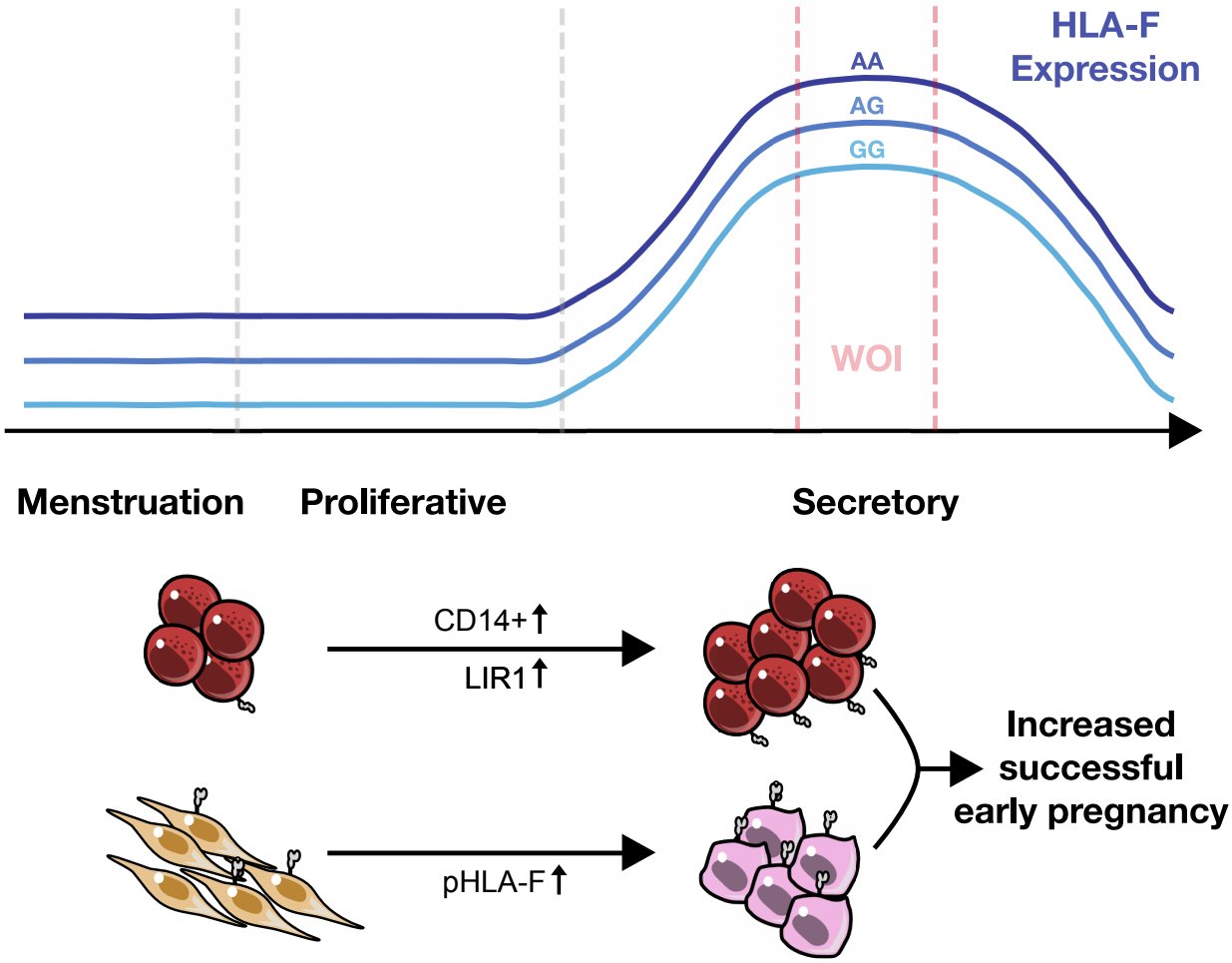

Figure 4.8 - A model for HLA-F's role in increasing fecundability.
pHLA-F is expressed at baseline in ESFs while LIR1 is expressed on CD14+ myeloid cells. Upon the start of the secretory cycle, decidualization stimuli increase both the abundance of leukocytes expressing LIR1 and pHLA-F expression on decidualized ESFs. Depending on your genotype at position rs2523393, HLA-F expression will either be higher or lower. We hypothesize that this interaction between pHLA-F and LIR1 increases the chances of successful implantation.

known class I molecules so it is possible that HLA-F plays an exceptional role in LIR1

recognition yet to be discovered.

4.6 Materials & Methods

Single-cell and bulk RNA-seq analysis

Single-cell RNA-seq data was downloaded from EMBL-EBI using the accession number E-MTAB-6701 (112). Processed counts were used as input into Seurat and ScanPy analysis software. Bulk RNA-seq data was downloaded from GEO using accession numbers GSE188352 (224) and GSE126530 (225) and used to assess the expression of class I HLA transcript as shown in **Figure 4.1**.

Primary and immortalized trophoblast culture

JEG3 choriocarcinoma and HTR8 and Swan71 immortalized cell lines were maintained in DMEM + 10% FBS (JEG3), RPMI 1640 + 10% FBS (HTR8), or DMEM:F12 + 10% FBS (Swan71). Primary trophoblasts were derived from first and third trimester abortive tissue samples from either the intervillous space at the placental implantation site or chorionic plate distal to the implantation site. Culture conditions were adapted from a previous study (207). For flow cytometry experiments, cells were detached using trypsin/EDTA, washed, and stained in 50 µL of staining buffer. Cells were then fixed in 100 µL of 2% PFA at room temperature for 15 minutes and resuspended in 200 µL of flow buffer (1X PBS + 3% FBS) prior to data collection on the ThermoFisher Attune NxT flow cytometer.

Culture of primary endometrial stromal cell lines

THESC and St-T1b cells were maintained in DMEM:F12 with charcoal-stripped FBS (CSFBS) added to 10%, sodium pyruvate added to 1 mM, Insulin-Selenium-Transferrin, L-glutamine added to 10 mM, and no antibiotics (THESC media). Primary ESF lines were generated by plating disassociated first trimester electively aborted tissue samples. Briefly, tissue was collected in basal RPMI 1640 and processed within 24 hours. Tissue was initially dissected to remove large blood coagulates and then digested for 30-45 minutes at 37ºC in HBSS mixed with collagenase, DNase, and hyaluronidase. Following this, erythrocytes and granulocytes were removed by Ficoll gradient and the mononuclear cell layer was extracted, washed, and frozen for long-term storage in liquid nitrogen. For the expansion of primary samples, cells were then thawed and plated into THESC media made with regular, heat-inactivated FBS instead of CSFBS (ESF expansion media) and with 1X Pen/Strep for 1 passage. After 1-3 days, cells were detached using TrypLE (Thermo Fisher) and replated in media without Pen/Strep. Cells were then either grown to 80% confluency with media changes every 2-3 days, or expanded with TrypLE and splitting cells 1:3. After 3-5 passages, cells were detached using TrypLE, counted, and frozen in ESF expansion media + 10% DMSO.

In vitro decidualization assays were carried out as previously described (200). Briefly, primary ESFs or ESF cell lines were thawed, cultured, and expanded in THESC maintenance media for 1-2 passages. Cells were then plated into T-25 flasks or 6-well plates and expanded until at least 90% confluent. Once cells were at least 90% confluent, media was exchanged into ESF priming media (THESC maintenance media with 2% CSFBS and 10 nM estradiol) overnight. The following day, media was changed again to ESF priming media (control) or ESF priming media

with added progestin (medroxyprogesterone acetate, MPA) and cyclic adenosine monophosphate (8-bromo-cAMP, cAMP) analogs (decidualization media). Cells were cultured for 48 hours prior to RNA extraction, protein extraction, or flow cytometry.

Making tetramers and staining cells

ocHLA-F tetramers were made by refolding bacterially expressed HLA-F from inclusion bodies as previously described (68). Purified protein was then *in vitro* biotinylated and tetramerized using SA-APC. pHLA-F monomers were made in 293T cells using a protocol adapted from Dulberger *et al.* (68). Briefly, β2M-fused HLA-F covalently linked to an N-terminal GFP tag was expressed in suspension-adapted 293T cells using the BacMam expression system (226) prior to purification as described in Dulberger *et al.* Purified protein was then *in vitro* biotinylated and tetramerized using APC conjugated SA (Agilent). Cells were stained with 400 nM tetramer in 1X PBS + 3% HI FBS for 45 minutes at room temperature. Decidual leukocyte populations were processed as described for primary ESF cell expansion. Cell aliquots were thawed and rested for 4 hours in RPMI 1640 + 10% FBS + 2 ng/mL IL-15 in a 50 mL conical at 37ºC. Cells were then collected, washed, and stained with Zombie live/dead stain, followed by tetramers, then antibody cocktails. After washing, cells were fixed as described for trophoblast cultures and data was collected on a ThermoFisher Attune NxT flow cytometer.

Chapter 5 - Conclusions and future directions

5.1 HLA-B*73:01 displays several unique molecular properties

Since its discovery nearly three decades ago, it was known that HLA-B*73:01 was extremely unusual (44). Its sequence indicated significant divergence from all characterized HLA alleles at the time, an observation that still holds true. Abi-Rached *et al.* proposed in 2011 that HLA-B*73:01 was likely reintroduced into the human gene pool through introgression with Denisovans and maintained for adaptive reasons that remain unclear. Due to its low frequency in human populations, we explored the molecular properties of HLA-B*73:01 *in vitro*.

We showed that in line with its divergent sequence, the peptidome of HLA-B*73:01 is equally distinct from other HLA-I molecules, that it presents very few self peptides, and that it is expressed at the cell surface at above-average levels. Using a structure of HLA-B*73:01 with KIR2DL2, we described its mode of peptide presentation, how KIR2DL2 structurally influences HLA-B*73:01 and its bound peptide, and how these observations compare with other alleles. Several aspects of the HLA-B*73:01 structure stood out: firstly, the similarity of HLA-B*73:01-presented self peptides to those presented by HLA-B*27:05 is due to a convergently-evolved B pocket. Secondly, it differentiates itself from HLA-B*27:05 and most other alleles by using a W95 in its F pocket to stabilize small hydrophobic residues at the PΩ position. Lastly, the KP1 peptide sits distinctly in the HLA-B*73:01 pocket in that its N terminus is shifted out of the pocket by greater than 1.6Å on average. The KP1 peptide bulges out of the pocket near where KIR2DL2 binds to HLA-B*73:01, potentially influencing its affinity. Together, these results

suggest that HLA-B*73:01 is divergent not only in sequence, but also in its peptidome and molecular signatures of peptide presentation.

The unique molecular properties of HLA-B*73:01 are consistent with its maintenance in the human population due to negative frequency-dependent selection (NFDS) (227). The NFDS model posits that microbial pathogens rapidly adapt to common HLA alleles, increasing the disease burden for individuals carrying common HLA alleles and giving the advantage to individuals carrying novel HLA alleles. Considering that HLA-B*73:01 presents peptides that are highly distinct from other HLA-I alleles and that pathogens are less likely to feel significant selection pressure from such a rare allele, individuals that carry HLA-B*73:01 likely faced less disease burden and were more likely to reproduce. However, it remains to be determined whether HLA-B*73:01 functions mainly as a ligand for KIR2DL2/L3 NK cell receptors or whether it can also provide significant T cell-mediated immunity as well.

Moving forward, it will also be important to identify CD8+ T cell epitopes presented by HLA-B*73:01. In on study investigating HLA alleles associated with HIV disease progression, HLA-B*73:01 was associated with slow disease progression (45). However, the effect was not statistically significant, likely due to a small sample size. Another way to test whether HLA-B*73:01 can provide T cell-mediated immunity against HIV would be to identify HLA-B*73:01+ HIV+ individuals and use tetramers of HLA-B*73:01 to find T cells specific for HLA-B*73:01 presenting HIV-derived peptide antigens. The peptidome of HLA-B*73:01 that we described can be used to predict high affinity HIV-derived peptide binders, test for T cell

reactivity, and generate tetramers to find T cells and HLA-B*73:01-restricted TCRs. Adding an understanding of HLA-B*73:01-restricted T cell-mediated immunity could add to the growing evolutionary story behind HLA-B*73:01's presence in the human gene pool.

5.2 Clarifying the role of pHLA-F *in vivo*

It took nearly three decades before the first crystal structure of HLA-F was determined (68). Prior to Dulberger *et al.*, it was widely assumed that HLA-F was expressed on the cell surface in its open conformer form. Studies showed that HLA-F remains expressed on the cell surface following acid treatment (58,60), that it may not associate with components of the PLC (53), and that cell-derived HLA-F does not present peptides (59). These traits made HLA-F different from all other HLA-I molecules, yet the molecular determinants of these unique properties remained unclear. Using HLA-F derived from insect cells, which is loaded with peptide and has superior stability relative to bacterial- and 293T-derived HLA-F material, we were able to generate Fab reagents that were more specific than previously described antibodies and possibly the first HLA-F staining reagent that can confidently detect cell surface pHLA-F.

Using these Fab reagents, we could address questions that were previously much more difficult to answer. We found that although ocHLA-F can be expressed on the surface of cells when overexpressed exogenously or on activated T cells, most cells that did express HLA-F, expressed higher levels of pHLA-F. Endogenously expressed HLA-F was dominated by cell surface pHLA-F and various cell lines engineered by CRISPR in combination with preliminary structural data suggest that our reagents are likely highly specific towards HLA-F recognition. Consistent with

these findings, we also determined that cells lacking β2M do not express endogenously-derived HLA-F on the cell surface and that HLA-F expression depends on TAP1 and TPN. Finally, we showed that even orthologs of HLA-F that encode for an arginine at residue 62, which was previously thought to ablate peptide binding, also present peptides, suggesting that peptide binding is the ancestral state of the hominid and old world monkey MHC-F common ancestor.

Interestingly, we found that the 3D11 antibody, the most commonly cited antibody used to detect cell surface HLA-F, is not only biased towards empty forms of HLA-F, but also likely cross-reactive. These results call into question conclusions that rely heavily on data generated using the 3D11 antibody. Furthermore, it appears that 3D11 was unable to detect HLA-F cell surface expression in many cell types, such as ESFs, which present HLA-F in the context of pregnancy and could be mediating HLA-F's association with early pregnancy success. Importantly, HLA-F has been associated with protection from several different viruses mediated by KIR3DS1+ NK cells thanks to reliance on the 3D11 antibody (58,61,62,181). While it is possible that virally infected cells upregulate the expression of ocHLA-F, which we did not test, a potential mechanism for switching between different forms of HLA-F is difficult to conceive of.

Considering that pHLA-F is a relevant form of HLA-F expressed *in vivo*, it will be crucial moving forward to determine structurally how HLA-F is able to present individual peptides, similar to what we were able to do with HLA-B*73:01. However, despite the availability of HLA-F-derived peptide elution data generated by LC-MS/MS, we have been unsuccessful in generating recombinant HLA-F bound to a single peptide antigen. One reason for this may be

that the refolding protocol that works for most HLA-I molecules is incompatible with HLA-F. For example, relative to HLA-I molecules, HLA class II molecules are notoriously difficult to refold with individual peptides hla refold (228–230). It could be that additional protocols need to be developed in order to purify homogeneously loaded HLA-F as well. Nevertheless, such a protocol would go a long way as it would allow access to structural data that has thus far proved elusive and for the investigation of HLA-F restricted T cells, whose TCRs are likely very specific for particular peptides.

Bibliography

1. Trowsdale J, Knight JC. Major Histocompatibility Complex Genomics and Human Disease. Annu Rev Genomics Hum Genet. 2013;14(1):301–23.

2. Stern LJ, Wiley DC. Antigenic peptide binding by class I and class II histocompatibility proteins. Structure. 1994 Apr 1;2(4):245–51.

3. Adams EJ, Luoma AM. The adaptable major histocompatibility complex (MHC) fold: structure and function of nonclassical and MHC class I-like molecules. Annu Rev Immunol. 2013 Jan;31:529–61.

4. Hennecke J, Wiley DC. T Cell Receptor–MHC Interactions up Close. Cell. 2001 Jan 12;104(1):1–4.

5. Pishesha N, Harmand TJ, Ploegh HL. A guide to antigen processing and presentation. Nat Rev Immunol. 2022 Apr 13;1–14.

6. Lang F, Schrörs B, Löwer M, Türeci Ö, Sahin U. Identification of neoantigens for individualized therapeutic cancer vaccines. Nat Rev Drug Discov. 2022 Apr;21(4):261–82.

7. Ross AB, Langer JD, Jovanovic M. Proteome Turnover in the Spotlight: Approaches, Applications, and Perspectives. Mol Cell Proteomics MCP. 2020 Dec 7;20:100016.

8. Grossmann N, Vakkasoglu AS, Hulpke S, Abele R, Gaudet R, Tampé R. Mechanistic determinants of the directionality and energetics of active export by a heterodimeric ABC transporter. Nat Commun. 2014 Nov 7;5(1):5419.

9. Cresswell P, Ackerman AL, Giodini A, Peaper DR, Wearsch PA. Mechanisms of MHC class I-restricted antigen processing and cross-presentation. Immunol Rev. 2005;207(1):145–57.

10. Raghavan M, Cid ND, Rizvi SM, Peters LR. MHC class I assembly: out and about. Trends Immunol. 2008 Sep 1;29(9):436–43.

11. Garbi N, Tan P, Diehl AD, Chambers BJ, Ljunggren HG, Momburg F, et al. Impaired immune responses and altered peptide repertoire in tapasin-deficient mice. Nat Immunol. 2000 Sep;1(3):234–8.

12. Blees A, Januliene D, Hofmann T, Koller N, Schmidt C, Trowitzsch S, et al. Structure of the human MHC-I peptide-loading complex. Nature. 2017 Nov;551(7681):525–8.

13. Jiang J, Taylor DK, Kim EJ, Boyd LF, Ahmad J, Mage MG, et al. Structural mechanism of tapasin-mediated MHC-I peptide loading in antigen presentation. Nat Commun. 2022 Sep 17;13(1):5470.

14. Neefjes J, Jongsma MLM, Paul P, Bakke O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat Rev Immunol. 2011 Dec;11(12):823–36.

15. An unstable beta 2-microglobulin: major histocompatibility complex class I heavy chain intermediate dissociates from calnexin and then is stabilized by binding peptide. J Exp Med. 1994 Dec 1;180(6):2163–71.

16. Guo HC, Jardetzky TS, Garrettt TPJ, Lane WS, Strominger JL, Wiley DC. Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. Nature. 1992 Nov;360(6402):364–6.

17. Bouvier M, Wiley DC. Importance of Peptide Amino and Carboxyl Termini to the Stability of MHC Class I Molecules. Science. 1994 Jul 15;265(5170):398–402.

18. Jh B, Ts J, Jc G, Lj S, Rg U, Jl S, et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature [Internet]. 1993 Jul 1 [cited 2022 Dec 24];364(6432). Available from: https://pubmed.ncbi.nlm.nih.gov/8316295/

19. Parker KC, Bednarek MA, Hull LK, Utz U, Cunningham B, Zweerink HJ, et al. Sequence motifs important for peptide binding to the human MHC class I molecule, HLA-A2. J Immunol. 1992 Dec 1;149(11):3580–7.

20. Madden D. The antigenic identity of peptide-MHC complexes: A comparison of the conformations of five viral peptides presented by HLA-A2. Cell. 1993 Nov 19;75(4):693–708.

21. Madden DR, Gorga JC, Strominger JL, Wiley DC. The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. Cell. 1992 Sep 18;70(6):1035–48.

22. Sarkizova S, Klaeger S, Le PM, Li LW, Oliveira G, Keshishian H, et al. A large peptidome dataset improves HLA class I epitope prediction across most of the human population. Nat Biotechnol. 2020 Feb;38(2):199–209.

23. Djaoud Z, Parham P. HLAs, TCRs, and KIRs, a Triumvirate of Human Cell-Mediated Immunity. Annu Rev Biochem. 2020 Jun 20;89:717–39.

24. Thammavongsa V, Raghuraman G, Filzen TM, Collins KL, Raghavan M. HLA-B44 Polymorphisms at Position 116 of the Heavy Chain Influence TAP Complex Binding via an Effect on Peptide Occupancy. J Immunol. 2006 Sep 1;177(5):3150–61.

25. Gangaev A, Ketelaars SLC, Isaeva OI, Patiwael S, Dopler A, Hoefakker K, et al. Identification and characterization of a SARS-CoV-2 specific CD8+ T cell response with immunodominant features. Nat Commun. 2021 May 10;12(1):2593.

26. Nagler A, Kalaora S, Barbolin C, Gangaev A, Ketelaars SLC, Alon M, et al. Identification of presented SARS-CoV-2 HLA class I and HLA class II peptides using HLA peptidomics. Cell Rep. 2021 Jun;35(13):109305.

27. Olafsdottir TA, Bjarnadottir K, Norddahl GL, Halldorsson GH, Melsted P, Gunnarsdottir K, et al. HLA alleles, disease severity, and age associate with T-cell responses following infection with SARS-CoV-2. Commun Biol. 2022 Sep 6;5(1):1–14.

28. Association of HLA class I with severe acute respiratory syndrome coronavirus infection | BMC Medical Genetics | Full Text [Internet]. [cited 2022 Oct 18]. Available from: https://bmcmedgenet.biomedcentral.com/articles/10.1186/1471-2350-4-9

29. Nguyen A, David JK, Maden SK, Wood MA, Weeder BR, Nellore A, et al. Human Leukocyte Antigen Susceptibility Map for Severe Acute Respiratory Syndrome Coronavirus 2. J Virol. 2020 Jun 16;94(13):e00510-20.

30. Stamatakis G, Samiotaki M, Mpakali A, Panayotou G, Stratikos E. Generation of SARS-CoV-2 S1 Spike Glycoprotein Putative Antigenic Epitopes in Vitro by Intracellular Aminopeptidases. J Proteome Res. 2020 Nov 6;19(11):4398–406.

31. Hilton HG, McMurtrey CP, Han AS, Djaoud Z, Guethlein LA, Blokhuis JH, et al. The Intergenic Recombinant HLA-B∗46:01 Has a Distinctive Peptidome that Includes KIR2DL3 Ligands. Cell Rep. 2017 May;19(7):1394–405.

32. Debebe BJ, Boelen L, Lee JC, IAVI Protocol C Investigators, Thio CL, Astemborski J, et al. Identifying the immune interactions underlying HLA class I disease associations. Davenport MP, Walczak AM, Lipsitch M, editors. eLife. 2020 Apr 2;9:e54558.

33. Neumann-Haefelin C. HLA-B27-mediated protection in HIV and hepatitis C virus infection and pathogenesis in spondyloarthritis: two sides of the same coin? Curr Opin Rheumatol. 2013 Jul;25(4):426–33.

34. Chen B, Li J, He C, Li D, Tong W, Zou Y, et al. Role of HLA-B27 in the pathogenesis of ankylosing spondylitis. Mol Med Rep. 2017 Apr;15(4):1943–51.

35. de MENTHON M, LAVALLEY MP, MALDINI C, GUILLEVIN L, MAHR A. HLA–B51/B5 and the Risk of Behçet's Disease: A Systematic Review and Meta-Analysis of Case–Control Genetic Association Studies. Arthritis Rheum. 2009 Oct 15;61(10):10.1002/art.24642.

36. Radwan J, Babik W, Kaufman J, Lenz TL, Winternitz J. Advances in the Evolutionary Understanding of MHC Polymorphism. Trends Genet. 2020 Apr 1;36(4):298–311.

37. Palmer WH, Telford M, Navarro A, Santpere G, Norman PJ. Human herpesvirus diversity is altered in HLA class I binding peptides. Proc Natl Acad Sci. 2022 May 3;119(18):e2123248119.

38. Immel A, Key FM, Szolek A, Barquera R, Robinson MK, Harrison GF, et al. Analysis of Genomic DNA from Medieval Plague Victims Suggests Long-Term Effect of Yersinia pestis on Human Immunity Genes. Mol Biol Evol. 2021 Oct 1;38(10):4059–76.

39. Parham P, Moffett A. Variable NK cell receptors and their MHC class I ligands in immunity, reproduction and human evolution. Nat Rev Immunol. 2013 Feb;13(2):133–44.

40. Phillips KP, Cable J, Mohammed RS, Herdegen-Radwan M, Raubic J, Przesmycka KJ, et al. Immunogenetic novelty confers a selective advantage in host–pathogen coevolution. Proc Natl Acad Sci U A. 2018 Jan;201708597.

41. Van Valen L. A New Evolutionary Law. Evol Theory. 1973 Jul;

42. Ejsmond MJ, Radwan J. Red Queen Processes Drive Positive Selection on Major Histocompatibility Complex (MHC) Genes. PLOS Comput Biol. 2015 Nov 24;11(11):e1004627.

43. Abi-Rached L, Jobin MJ, Kulkarni S, McWhinnie A, Dalva K, Gragert L, et al. The shaping of modern human immune systems by multiregional admixture with archaic humans. Science. 2011 Oct;334(6052):89–94.

44. Parham P, Arnett KL, Adams EJ, Barber LD, Domena JD, Stewart D, et al. The HLA-B73 antigen has a most unusual structure that defines a second lineage of HLA-B alleles. Tissue Antigens. 1994 May;43(5):302–13.

45. THE INTERNATIONAL HIV CONTROLLERS STUDY. The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation. Science. 2010 Dec 10;330(6010):1551–7.

46. Ombrello MJ, Kirino Y, de Bakker PIW, Gül A, Kastner DL, Remmers EF. Behçet disease-associated MHC class I residues implicate antigen binding and regulation of cell-mediated cytotoxicity. Proc Natl Acad Sci. 2014 Jun 17;111(24):8867–72.

47. Arnaiz-Villena A, Suarez-Trujillo F, Juarez I, Rodríguez-Sainz C, Palacio-Gruber J, Vaquero-Yuste C, et al. Evolution and molecular interactions of major histocompatibility complex (MHC)-G, -E and -F genes. Cell Mol Life Sci. 2022;79(8):464.

48. Otting N, de Groot NG, Bontrop RE. Evolution of HLA-F and its orthologues in primate species: a complex tale of conservation, diversification and inactivation. Immunogenetics. 2020 Dec;72(9–10):475–87.

49. Braud VM, Allan DS, Wilson D, McMichael AJ. TAP- and tapasin-dependent HLA-E surface expression correlates with the binding of an MHC class I leader peptide. Curr Biol CB. 1998 Jan 1;8(1):1–10.

50. Kaiser BK, Pizarro JC, Kerns J, Strong RK. Structural basis for NKG2A/CD94 recognition of HLA-E. Proc Natl Acad Sci U S A. 2008 May 6;105(18):6696–701.

51. Ferreira LMR, Meissner TB, Mikkelsen TS, Mallard W, O'Donnell CW, Tilburgs T, et al. A distant trophoblast-specific enhancer controls HLA-G expression at the maternal–fetal interface. Proc Natl Acad Sci U A. 2016 May;113(19):5364–9.

52. Ishitani A, Sageshima N, Lee N, Dorofeeva N, Hatake K, Marquardt H, et al. Protein expression and peptide binding suggest unique and interacting functional roles for HLA-E, F, and G in maternal-placental immune recognition. J Immunol. 2003 Aug;171(3):1376–84.

53. Lee N, Geraghty DE. HLA-F surface expression on B cell and monocyte cell lines is partially independent from tapasin and completely independent from TAP. J Immunol. 2003;

54. Parham P, Barnstable CJ, Bodmer WF. Use of a Monoclonal Antibody (W6/32) in Structural Studies of HLA-A,B,C Antigens. J Immunol. 1979 Jul 1;123(1):342–9.

55. Brodsky FM, Bodmer WF, Parham P. Characterization of a monoclonal anti-β2-microglobulin antibody and its use in the genetic and biochemical analysis of major histocompatibility antigens. Eur J Immunol. 1979;9(7):536–45.

56. Wainwright SD, Biro PA, Holmes CH. HLA-F is a predominantly empty, intracellular, TAP-associated MHC class Ib protein with a restricted expression pattern. J Immunol. 2000 Jan;164(1):319–28.

57. Lepin EJM, Bastin JM, Allan DSJ, Roncador G, Braud VM, Mason DY, et al. Functional characterization of HLA-F and binding of HLA-F tetramers to ILT2 and ILT4 receptors. Eur J Immunol. 2000;30(12):3552–61.

58. Garcia-Beltran WF, Hölzemer A, Martrus G, Chung AW, Pacheco Y, Simoneau CR, et al. Open conformers of HLA-F are high-affinity ligands of the activating NK-cell receptor KIR3DS1. Nat Immunol. 2016 Sep;17(9):1067–74.

59. Goodridge JP, Burian A, Lee N, Geraghty DE. HLA-F complex without peptide binds to MHC class I protein in the open conformer form. J Immunol. 2010 Jun;184(11):6199–208.

60. Lee N, Ishitani A, Geraghty DE. HLA-F is a surface marker on activated lymphocytes. Eur J Immunol. 2010 Aug;40(8):2308–18.

61. Koyro TF, Kraus E, Lunemann S, Hölzemer A, Wulf S, Jung J, et al. Upregulation of HLA-F expression by BK polyomavirus infection induces immune recognition by KIR3DS1-positive natural killer cells. Kidney Int. 2021 May;99(5):1140–8.

62. Lunemann S, Schöbel A, Kah J, Fittje P, Hölzemer A, Langeneckert AE, et al. Interactions Between KIR3DS1 and HLA-F Activate Natural Killer Cells to Control HCV Replication in Cell Culture. Gastroenterology. 2018 Nov;155(5):1366-1371.e3.

63. Jung JM, Ching W, Baumdick ME, Hofmann-Sieber H, Bosse JB, Koyro T, et al. KIR3DS1 directs NK cell-mediated protection against human adenovirus infections. Sci Immunol. 2021 Sep 17;6(63):eabe2942.

64. Müller S, Zocher G, Steinle A, Stehle T. Structure of the HCMV UL16-MICB Complex Elucidates Select Binding of a Viral Immunoevasin to Diverse NKG2D Ligands. PLOS Pathog. 2010 Jan 15;6(1):e1000723.

65. Li P, Willie ST, Bauer S, Morris DL, Spies T, Strong RK. Crystal Structure of the MHC Class I Homolog MIC-A, a γδ T Cell Ligand. Immunity. 1999 May 1;10(5):577–84.

66. Adams EJ, Chien YH, Garcia KC. Structure of a γδ T Cell Receptor in Complex with the Nonclassical MHC T22. Science. 2005 Apr 8;308(5719):227–31.

67. Kajikawa M, Ose T, Fukunaga Y, Okabe Y, Matsumoto N, Yonezawa K, et al. Structure of MHC class I-like MILL2 reveals heparan-sulfate binding and interdomain flexibility. Nat Commun. 2018 Oct 18;9(1):4330.

68. Dulberger CL, McMurtrey CP, Hölzemer A, Neu KE, Liu V, Steinbach AM, et al. Human Leukocyte Antigen F Presents Peptides and Regulates Immunity through Interactions with NK Cell Receptors. Immunity. 2017 Jun;46(6):1018-1029.e7.

69. Burian A, Wang KL, Finton KAK, Lee N, Ishitani A, Strong RK, et al. HLA-F and MHC-I Open Conformers Bind Natural Killer Cell Ig-Like Receptor KIR3DS1. PLOS ONE. 2016 Sep 20;11(9):e0163297.

70. Goodridge JP, Burian A, Lee N, Geraghty DE. HLA-F and MHC class I open conformers are ligands for NK cell Ig-like receptors. J Immunol. 2013 Oct;191(7):3553–62.

71. Shiroishi M, Kuroki K, Rasubala L, Tsumoto K, Kumagai I, Kurimoto E, et al. Structural basis for recognition of the nonclassical MHC molecule HLA-G by the leukocyte Ig-like receptor B2 (LILRB2/LIR2/ILT4/CD85d). Proc Natl Acad Sci U A. 2006 Oct;103(44):16412–7.

72. Liu F, Cocker ATH, Pugh JL, Djaoud Z, Parham P, Guethlein LA. Natural LILRB1 D1-D2 variants show frequency differences in populations and bind to HLA class I with various avidities. Immunogenetics. 2022 Dec 1;74(6):513–25.

73. Oliveira MLG de, Castelli EC, Veiga-Castelli LC, Pereira AE, Marcorin L, Carratto TMT, et al. Genetic diversity of the LILRB1 and LILRB2 coding regions in an admixed Brazilian population sample [Internet]. bioRxiv; 2021 [cited 2022 Nov 7]. p. 2021.04.16.440206. Available from: https://www.biorxiv.org/content/10.1101/2021.04.16.440206v1

74. Hirayasu K, Ohashi J, Tanaka H, Kashiwase K, Ogawa A, Takanashi M, et al. Evidence for natural selection on leukocyte immunoglobulin-like receptors for HLA class I in Northeast Asians. Am J Hum Genet. 2008 May;82(5):1075–83.

75. Barkal AA, Weiskopf K, Kao KS, Gordon SR, Rosental B, Yiu YY, et al. Engagement of MHC class I by the inhibitory receptor LILRB1 suppresses macrophages and is a target of cancer immunotherapy. Nat Immunol. 2018 Jan;19(1):76–84.

76. Saito F, Hirayasu K, Satoh T, Wang CW, Lusingu J, Arimori T, et al. Immune evasion of Plasmodium falciparum by RIFIN via inhibitory receptors. Nature. 2017 Nov;

77. Harrison TE, Mørch AM, Felce JH, Sakoguchi A, Reid AJ, Arase H, et al. Structural basis for RIFIN-mediated activation of LILRB1 in malaria. Nature. 2020 Nov;587(7833):309–12.

78. Gamliel M, Goldman-Wohl D, Isaacson B, Gur C, Stein N, Yamin R, et al. Trained Memory of Human Uterine NK Cells Enhances Their Function in Subsequent Pregnancies. Immunity. 2018 May;48(5):951-962.e5.

79. Fu B, Zhou Y, Ni X, Tong X, Xu X, Dong Z, et al. Natural Killer Cells Promote Fetal Development through the Secretion of Growth-Promoting Factors. Immunity. 2017 Dec;47(6):1100-1113.e6.

80. Yang Z, Bjorkman PJ. Structure of UL18, a peptide-binding viral MHC mimic, bound to a host inhibitory receptor. Proc Natl Acad Sci U A. 2008 Jul;105(29):10095–100.

81. Allan DSJ, Lepin EJM, Braud VM, O'Callaghan CA, McMichael AJ. Tetrameric complexes of HLA-E, HLA-F, and HLA-G. J Immunol Methods. 2002 Oct 1;268(1):43–50.

82. Long EO, Kim HS, Liu D, Peterson ME, Rajagopalan S. Controlling natural killer cell responses: integration of signals for activation and inhibition. Annu Rev Immunol. 2013;31:227–58.

83. Björkström NK, Béziat V, Cichocki F, Liu LL, Levine J, Larsson S, et al. CD8 T cells express randomly selected KIRs with distinct specificities compared with NK cells. Blood. 2012 Oct 25;120(17):3455–65.

84. Huard B, Karlsson L. KIR expression on self-reactive CD8+ T cells is controlled by T-cell receptor engagement. Nature. 2000 Jan;403(6767):325–8.

85. Li J, Zaslavsky M, Su Y, Guo J, Sikora MJ, van Unen V, et al. KIR+CD8+ T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19. Science. 2022 Mar 8;376(6590):eabi9591.

86. Guethlein LA, Norman PJ, Hilton HG, Parham P. Co-evolution of MHC class I and variable NK cell receptors in placental mammals. Immunol Rev. 2015 Sep;267(1):259–82.

87. Fan QR, Long EO, Wiley DC. Crystal structure of the human natural killer cell inhibitory receptor KIR2DL1-HLA-Cw4 complex. Nat Immunol. 2001 May;2(5):452–60.

88. Moradi S, Stankovic S, O'Connor GM, Pymm P, MacLachlan BJ, Faoro C, et al. Structural plasticity of KIR2DL2 and KIR2DL3 enables altered docking geometries atop HLA-C. Nat Commun. 2021 Apr 12;12(1):2173.

89. Stewart-Jones GBE, di Gleria K, Kollnberger S, McMichael AJ, Jones EY, Bowness P. Crystal structures and KIR3DL1 recognition of three immunodominant viral peptides complexed to HLA-B*2705. Eur J Immunol. 2005;35(2):341–51.

90. Sim MJW, Malaker SA, Khan A, Stowell JM, Shabanowitz J, Peterson ME, et al. Canonical and Cross-reactive Binding of NK Cell Inhibitory Receptors to HLA-C Allotypes Is Dictated by Peptides Bound to HLA-C. Front Immunol [Internet]. 2017 [cited 2022 Jul 28];8. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2017.00193

91. Naiyer MM, Cassidy SA, Magri A, Cowton V, Chen K, Mansour S, et al. KIR2DS2 recognizes conserved peptides derived from viral helicases in the context of HLA-C. Sci Immunol [Internet]. 2017 Sep 15 [cited 2022 Nov 7]; Available from: https://www.science.org/doi/10.1126/sciimmunol.aal5296

92. Sim MJW, Rajagopalan S, Altmann DM, Boyton RJ, Sun PD, Long EO. Human NK cell receptor KIR2DS4 detects a conserved bacterial epitope presented by HLA-C. Proc Natl Acad Sci U S A. 2019 Jun 25;116(26):12964–73.

93. Boyington JC, Motyka SA, Schuck P, Brooks AG, Sun PD. Crystal structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand. Nature. 2000 Jun;405(6786):537–43.

94. Liu J, Xiao Z, Ko HL, Shen M, Ren EC. Activating killer cell immunoglobulin-like receptor 2DS2 binds to HLA-A*11. Proc Natl Acad Sci U A. 2014 Feb;111(7):2662–7.

95. Yang Y, Bai H, Wu Y, Chen P, Zhou J, Lei J, et al. Activating receptor KIR2DS2 bound to HLA-C1 reveals the novel recognition features of activating receptor. Immunology. 2022 Mar;165(3):341–54.

96. Vivian JP, Duncan RC, Berry R, O'Connor GM, Reid HH, Beddoe T, et al. Killer cell immunoglobulin-like receptor 3DL1-mediated recognition of human leukocyte antigen B. Nature. 2011 Oct;479(7373):401–5.

97. Parham P, Guethlein LA. Genetics of Natural Killer Cells in Human Health, Disease, and Survival. Annu Rev Immunol. 2018 Feb;

98. Moesta AK, Norman PJ, Yawata M, Yawata N, Gleimer M, Parham P. Synergistic Polymorphism at Two Positions Distal to the Ligand-Binding Site Makes KIR2DL2 a Stronger Receptor for HLA-C Than KIR2DL3. J Immunol. 2008 Mar 15;180(6):3969–79.

99. Wroblewski EE, Parham P, Guethlein LA. Two to Tango: Co-evolution of Hominid Natural Killer Cell Receptors and MHC. Front Immunol. 2019 Feb 19;10:177.

100. Blokhuis JH, Hilton HG, Guethlein LA, Norman PJ, Nemat-Gorgani N, Nakimuli A, et al. KIR2DS5 allotypes that recognize the C2 epitope of HLA-C are common among Africans and absent from Europeans. Immun Inflamm Dis. 2017 Dec;5(4):461–8.

101. Hilton HG, Norman PJ, Nemat-Gorgani N, Goyos A, Hollenbach JA, Henn BM, et al. Loss and Gain of Natural Killer Cell Receptor Function in an African Hunter-Gatherer Population. PLOS Genet. 2015 Aug 20;11(8):e1005439.

102. Hilton HG, Moesta AK, Guethlein LA, Blokhuis J, Parham P, Norman PJ. The production of KIR-Fc fusion proteins and their use in a multiplex HLA class I binding assay. J Immunol Methods. 2015 Oct;425:79–87.

103. Li SS, Hickey A, Shangguan S, Ehrenberg PK, Geretz A, Butler L, et al. HLA-B∗46 associates with rapid HIV disease progression in Asian cohorts and prominent differences in NK cell phenotype. Cell Host Microbe. 2022 Aug 10;30(8):1173-1185.e8.

104. MEDAWAR, PB. Some immunological and endocrinological problems raised by the evolution of viviparity in vertebrates. Symp Soc Exp Biol. 1953;7:320–37.

105. Ferreira LMR, Meissner TB, Tilburgs T, Strominger JL. HLA-G: At the Interface of Maternal–Fetal Tolerance. Trends Immunol. 2017 Apr 1;38(4):272–86.

106. Benner M, Feyaerts D, García CC, Inci N, López SC, Fasse E, et al. Clusters of Tolerogenic B Cells Feature in the Dynamic Immunological Landscape of the Pregnant Uterus. Cell Rep. 2020 Sep 29;32(13):108204.

107. Hanna J, Goldman-Wohl D, Hamani Y, Avraham I, Greenfield C, Natanson-Yaron S, et al. Decidual NK cells regulate key developmental processes at the human fetal-maternal interface. Nat Med. 2006 Sep;12(9):1065–74.

108. Huhn O, Ivarsson MA, Gardner L, Hollinshead M, Stinchcombe JC, Chen P, et al. Distinctive phenotypes and functions of innate lymphoid cells in human decidua during early pregnancy. Nat Commun. 2020 Jan;11(1):381.

109. Koopman LA, Kopcow HD, Rybalov B, Boyson JE, Orange JS, Schatz F, et al. Human decidual natural killer cells are a unique NK cell subset with immunomodulatory potential. J Exp Med. 2003 Oct;198(8):1201–12.

110. Salvany-Celades M, van der Zwan A, Benner M, Setrajcic-Dragos V, Bougleux Gomes HA, Iyer V, et al. Three Types of Functional Regulatory T Cells Control T Cell Responses at the Human Maternal-Fetal Interface. Cell Rep. 2019 May 28;27(9):2537-2547.e5.

111. Strunz B, Bister J, Jönsson H, Filipovic I, Crona-Guterstam Y, Kvedaraite E, et al. Continuous human uterine NK cell differentiation in response to endometrial regeneration and pregnancy. Sci Immunol. 2021 Feb;6(56).

112. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, et al. Single-cell reconstruction of the early maternal–fetal interface in humans. Nature. 2018 Nov;563(7731):347–53.

113. Zdravkovic M, Aboagye-Mathiesen G, Zachar V, Tóth FD, Dalsgård AM, Hager H, et al. Immunosuppressive effects of human placental trophoblast interferon-beta on lymphocytes in vitro. Placenta. 1994 Sep;15(6):591–600.

114. King A, Kalra P, Loke YW. Human trophoblast cell resistance to decidual NK lysis is due to lack of NK target structure. Cell Immunol. 1990 May;127(2):230–7.

115. Zeldovich VB, Bakardjiev AI. Host defense and tolerance: unique challenges in the placenta. PLoS Pathog. 2012;8(8):e1002804.

116. Crespo ÂC, Strominger JL, Tilburgs T. Expression of KIR2DS1 by decidual natural killer cells increases their ability to control placental HCMV infection. Proc Natl Acad Sci U S A. 2016 Dec 27;113(52):15072–7.

117. Lee CQE, Gardner L, Turco M, Zhao N, Murray MJ, Coleman N, et al. What Is Trophoblast? A Combination of Criteria Define Human First-Trimester Trophoblast. Stem Cell Rep. 2016 Feb 9;6(2):257–72.

118. Moffett A, Colucci F. Uterine NK cells: active regulators at the maternal-fetal interface. J Clin Invest. 2014 May;124(5):1872–9.

119. Bulmer JN, Morrison L, Longfellow M, Ritson A, Pace D. Granulated lymphocytes in human endometrium: histochemical and immunohistochemical studies. Hum Reprod Oxf Engl. 1991 Jul;6(6):791–8.

120. Hackmon R, Pinnaduwage L, Zhang J, Lye SJ, Geraghty DE, Dunk CE. Definitive class I human leukocyte antigen expression in gestational placentation: HLA-F, HLA-E, HLA-C, and HLA-G in extravillous trophoblast invasion on placentation, pregnancy, and parturition. Am J Reprod Immunol. 2017;77(6).

121. Dunk CE, Bucher M, Zhang J, Hayder H, Geraghty DE, Lye SJ, et al. Human leukocyte antigen HLA-C, HLA-G, HLA-F, and HLA-E placental profiles are altered in

early severe preeclampsia and preterm birth with chorioamnionitis. Am J Obstet Gynecol. 2022 Oct;227(4):641.e1-641.e13.

122. Shobu T, Sageshima N, Tokui H, Omura M, Saito K, Nagatsuka Y, et al. The surface expression of HLA-F on decidual trophoblasts increases from mid to term gestation. J Reprod Immunol. 2006 Dec;72(1–2):18–32.

123. Nagamatsu T, Fujii T, Matsumoto J, Yamashita T, Kozuma S, Taketani Y. Human Leukocyte Antigen F Protein is Expressed in the Extra-Villous Trophoblasts but Not on the Cell Surface of them. Am J Reprod Immunol. 2006;56(3):172–7.

124. Sen Santara S, Crespo ÂC, Mulik S, Ovies C, Boulenouar S, Strominger JL, et al. Decidual NK cells kill Zika virus–infected trophoblasts. Proc Natl Acad Sci U S A. 2021 Nov 23;118(47):e2115410118.

125. Depierreux DM, Kieckbusch J, Shreeve N, Hawkes DA, Marsh B, Blelloch R, et al. Beyond Maternal Tolerance: Education of Uterine Natural Killer Cells by Maternal MHC Drives Fetal Growth. Front Immunol [Internet]. 2022 [cited 2022 Oct 8];13. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2022.808227

126. Kieckbusch J, Gaynor LM, Moffett A, Colucci F. MHC-dependent inhibition of uterine NK cells impedes fetal growth and decidual vascular remodelling. Nat Commun. 2014 Feb;5:3359.

127. Burrows CK, Kosova G, Herman C, Patterson K, Hartmann KE, Velez Edwards DR, et al. Expression Quantitative Trait Locus Mapping Studies in Mid-secretory Phase Endometrial Cells Identifies HLA-F and TAP2 as Fecundability-Associated Genes. PLoS Genet. 2016 Jul;12(7):e1005858.

128. Kofod L, Lindhard A, Bzorek M, Eriksen JO, Larsen LG, Hviid TVF. Endometrial immune markers are potential predictors of normal fertility and pregnancy after in vitro fertilization. Am J Reprod Immunol. 2017 Sep;78(3).

129. Papúchová H, Saxtorph MH, Hallager T, Jepsen IE, Eriksen JO, Persson G, et al. Endometrial HLA-F expression is influenced by genotypes and correlates differently with

immune cell infiltration in IVF and recurrent implantation failure patients. Hum Reprod Oxf Engl. 2022 Jul 30;37(8):1816–34.

130. Langkilde CH, Nilsson LL, Jørgensen N, Funck T, Perin TL, Hornstrup MB, et al. Variation in the HLA-F gene locus with functional impact is associated with pregnancy success and time-to-pregnancy after fertility treatment. Hum Reprod. 2020 Mar;35(3):705–17.

131. A Draft Sequence of the Neandertal Genome | Science [Internet]. [cited 2022 Aug 22]. Available from: https://www.science.org/doi/10.1126/science.1188021

132. A High-Coverage Genome Sequence from an Archaic Denisovan Individual | Science [Internet]. [cited 2022 Aug 22]. Available from: https://www.science.org/doi/10.1126/science.1224344

133. Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, et al. The complete genome sequence of a Neanderthal from the Altai Mountains. Nature. 2014 Jan;505(7481):43–9.

134. The Strength of Selection against Neanderthal Introgression | PLOS Genetics [Internet]. [cited 2022 Aug 22]. Available from: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006340

135. Dannemann M, Andrés AM, Kelso J. Introgression of Neandertal- and Denisovan-like Haplotypes Contributes to Adaptive Variation in Human Toll-like Receptors. Am J Hum Genet. 2016 Jan 7;98(1):22–33.

136. Resurrecting Surviving Neandertal Lineages from Modern Human Genomes | Science [Internet]. [cited 2022 Aug 22]. Available from: https://www.science.org/doi/10.1126/science.1245938

137. The great human expansion | PNAS [Internet]. [cited 2022 Aug 22]. Available from: https://www.pnas.org/doi/10.1073/pnas.1212380109

138. Hilton HG, Vago L, Aguilar AMO, Moesta AK, Graef T, Abi-Rached L, et al. Mutation at Positively Selected Positions in the Binding Site for HLA-C Shows That

KIR2DL1 Is a More Refined but Less Adaptable NK Cell Receptor Than KIR2DL3. J Immunol. 2012 Aug 1;189(3):1418–30.

139. Barber LD, Percival L, Valiante NM, Chen L, Lee C, Gumperz JE, et al. The inter-locus recombinant HLA-B*4601 has high selectivity in peptide binding and functions characteristic of HLA-C. J Exp Med. 1996 Aug 1;184(2):735–40.

140. Moesta A, Parham P. Diverse functionality among human NK cell receptors for the C1 epitope of HLA-C: KIR2DS2, KIR2DL2, and KIR2DL3. Front Immunol [Internet]. 2012 [cited 2022 Aug 31];3. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2012.00336

141. Khakoo SI, Thio CL, Martin MP, Brooks CR, Gao X, Astemborski J, et al. HLA and NK Cell Inhibitory Receptor Genes in Resolving Hepatitis C Virus Infection. Science. 2004 Aug 6;305(5685):872–4.

142. Basatena NKS al, MacNamara A, Vine AM, Thio CL, Astemborski J, Usuku K, et al. KIR2DL2 Enhances Protective and Detrimental HLA Class I-Mediated Immunity in Chronic Viral Infection. PLOS Pathog. 2011 Oct 13;7(10):e1002270.

143. Sylvester-Hvid C, Kristensen N, Blicher T, Ferré H, Lauemøller SL, Wolf XA, et al. Establishment of a quantitative ELISA capable of determining peptide - MHC class I interaction. Tissue Antigens. 2002 Apr;59(4):251–8.

144. Barber LD, Percival L, Parham P. Characterization of the peptide-binding specificity of HLA-B*7301. Tissue Antigens. 1996;47(6):472–7.

145. Andreatta M, Alvarez B, Nielsen M. GibbsCluster: unsupervised clustering and alignment of peptide sequences. Nucleic Acids Res. 2017 Jul 3;45(W1):W458–63.

146. Andreatta M, Lund O, Nielsen M. Simultaneous alignment and clustering of peptide data using a Gibbs sampling approach. Bioinforma Oxf Engl. 2013 Jan 1;29(1):8–14.

147. Sesma L, Montserrat V, Lamas JR, Marina A, Vázquez J, Castro JAL de. The Peptide Repertoires of HLA-B27 Subtypes Differentially Associated to

Spondyloarthropathy (B*2704 and B*2706) Differ by Specific Changes at Three Anchor Positions *. J Biol Chem. 2002 May 10;277(19):16744–9.

148. Kaufman J. Generalists and Specialists: A New View of How MHC Class I Molecules Fight Infectious Pathogens. Trends Immunol. 2018 May 1;39(5):367–79.

149. Uebel S, Kraas W, Kienle S, Wiesmüller KH, Jung G, Tampé R. Recognition principle of the TAP transporter disclosed by combinatorial peptide libraries. Proc Natl Acad Sci. 1997 Aug 19;94(17):8976–81.

150. Bulik-Sullivan B, Busby J, Palmer CD, Davis MJ, Murphy T, Clark A, et al. Deep learning using tumor HLA peptide mass spectrometry datasets improves neoantigen identification. Nat Biotechnol. 2019 Jan;37(1):55–63.

151. Faridi P, Li C, Ramarathinam SH, Vivian JP, Illing PT, Mifsud NA, et al. A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands. Sci Immunol. 2018 Oct 12;3(28):eaar3947.

152. Gfeller D, Guillaume P, Michaux J, Pak HS, Daniel RT, Racle J, et al. The Length Distribution and Multiple Specificity of Naturally Presented HLA-I Ligands. J Immunol. 2018 Dec 15;201(12):3705–16.

153. Scull KE, Dudek NL, Corbett AJ, Ramarathinam SH, Gorasia DG, Williamson NA, et al. Secreted HLA recapitulates the immunopeptidome and allows in-depth coverage of HLA A*02:01 ligands. Mol Immunol. 2012 Jun 1;51(2):136–42.

154. Goulder PJR, Brander C, Tang Y, Tremblay C, Colbert RA, Addo MM, et al. Evolution and transmission of stable CTL escape mutations in HIV infection. Nature. 2001 Jul;412(6844):334–8.

155. Chappell P, Meziane EK, Harrison M, Magiera \Lukasz, Hermann C, Mears L, et al. Expression levels of MHC class I molecules are inversely correlated with promiscuity of peptide binding. Elife. 2015 Apr;4:e05345.

156. Bashirova AA, Viard M, Naranbhai V, Grifoni A, Garcia-Beltran W, Akdag M, et al. HLA tapasin independence: broader peptide repertoire and HIV control. Proc Natl Acad Sci. 2020 Nov 10;117(45):28232–8.

157. McCutcheon JA, Gumperz J, Smith KD, Lutz CT, Parham P. Low HLA-C expression at cell surfaces correlates with increased turnover of heavy chain mRNA. J Exp Med. 1995 Jun 1;181(6):2085–95.

158. Reduced Cell Surface Expression of HLA-C Molecules Correlates with Restricted Peptide Binding and Stable TAP Interaction | The Journal of Immunology [Internet]. [cited 2022 Aug 23]. Available from: https://www.jimmunol.org/content/160/1/171.long

159. Sibilio L, Martayan A, Setini A, Monaco EL, Tremante E, Butler RH, et al. A Single Bottleneck in HLA-C Assembly*. J Biol Chem. 2008 Jan 18;283(3):1267–74.

160. Kaur G, Gras S, Mobbs JI, Vivian JP, Cortes A, Barber T, et al. Structural and regulatory diversity shape HLA-C protein expression levels. Nat Commun. 2017 Jun;8:15924.

161. Latron F, Pazmany L, Morrison J, Moots R, Saper MA, McMichael A, et al. A Critical Role for Conserved Residues in the Cleft of HLA-A2 in Presentation of a Nonapeptide to T Cells. Science. 1992 Aug 14;257(5072):964–7.

162. Kawaguchi G, Hildebrand WH, Hiraiwa M, Karaki S, Nagao T, Akiyama N, et al. Two subtypes of HLA-B51 differing by substitution at position 171 of the alpha 2 helix. Immunogenetics. 1992;37(1):57–63.

163. Kikuchi A, Sakaguchi T, Miwa K, Takamiya Y, Rammensee HG, Kaneko Y, et al. Binding of nonamer peptides to three HLA-B51 molecules which differ by a single amino acid substitution in the A-pocket. Immunogenetics. 1996 Sep 1;43(5):268–76.

164. Rajagopalan S, Long EO. The Direct Binding of a p58 Killer Cell Inhibitory Receptor to Human Histocompatibility Leukocyte Antigen (HLA)-Cw4 Exhibits Peptide Selectivity. J Exp Med. 1997 Apr 21;185(8):1523–8.

165. Fadda L, Borhis G, Ahmed P, Cheent K, Pageon SV, Cazaly A, et al. Peptide antagonism as a mechanism for NK cell activation. Proc Natl Acad Sci. 2010 Jun;107(22):10160–5.

166. Influence of HLA-C Expression Level on HIV Control | Science [Internet]. [cited 2022 Aug 23]. Available from: https://www.science.org/doi/10.1126/science.1232685

167. Enard D, Petrov DA. Evidence that RNA Viruses Drove Adaptive Introgression between Neanderthals and Modern Humans. Cell. 2018 Oct 4;175(2):360-371.e13.

168. Digitale JC, Callaway PC, Martin M, Nelson G, Viard M, Rek J, et al. Association of Inhibitory Killer Cell Immunoglobulin-like Receptor Ligands With Higher Plasmodium falciparum Parasite Prevalence. J Infect Dis. 2021 Jul 1;224(1):175–83.

169. Hirayasu K, Ohashi J, Kashiwase K, Hananantachai H, Naka I, Ogawa A, et al. Significant Association of KIR2DL3-HLA-C1 Combination with Cerebral Malaria and Implications for Co-evolution of KIR and HLA. PLOS Pathog. 2012 Mar 8;8(3):e1002565.

170. Prakash S, Ranjan P, Ghoshal U, Agrawal S. KIR-like activating natural killer cell receptors and their association with complicated malaria in north India. Acta Trop. 2018 Feb 1;178:55–60.

171. Mathieu A, Paladini F, Vacca A, Cauli A, Fiorillo MT, Sorrentino R. The interplay between the geographic distribution of HLA-B27 alleles and their role in infectious and autoimmune diseases: A unifying hypothesis. Autoimmun Rev. 2009 Mar 1;8(5):420–5.

172. Murakoshi H, Chikata T, Akahoshi T, Zou C, Borghan MA, Van Tran G, et al. Critical effect of Pol escape mutations associated with detrimental allele HLA-C∗15:05 on clinical outcome in HIV-1 subtype A/E infection. AIDS. 2021 Jan 1;35(1):33–43.

173. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2012. 2021.

174. O'Donnell TJ, Rubinsteyn A, Laserson U. MHCflurry 2.0: Improved Pan-Allele Prediction of MHC Class I-Presented Peptides by Incorporating Antigen Processing. Cell Syst. 2020 Jul 22;11(1):42-48.e7.

175. Wagih O. ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics. 2017 Nov 15;33(22):3645–7.

176. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. vegan: Community Ecology Package [Internet]. 2022 [cited 2022 Sep 18]. Available from: https://CRAN.R-project.org/package=vegan

177. Hò GGT, Heinen FJ, Huyton T, Blasczyk R, Bade-Döding C. HLA-F*01:01 presents peptides with N-terminal flexibility and a preferred length of 16 residues. Immunogenetics. 2019;71(5):353–60.

178. Hò GGT, Heinen FJ, Blasczyk R, Pich A, Bade-Doeding C. HLA-F Allele-Specific Peptide Restriction Represents an Exceptional Proteomic Footprint. Int J Mol Sci. 2019 Nov 8;20(22):5572.

179. Lin A, Zhang X, Ruan YY, Wang Q, Zhou WJ, Yan WH. HLA-F expression is a prognostic factor in patients with non-small-cell lung cancer. Lung Cancer. 2011 Dec;74(3):504–9.

180. Song S, Miranda CJ, Braun L, Meyer K, Frakes AE, Ferraiuolo L, et al. MHC class I protects motor neurons from astrocyte-induced toxicity in amyotrophic lateral sclerosis (ALS). Nat Med. 2016 Apr;22(4):397.

181. Jm J, W C, Me B, H HS, Jb B, T K, et al. KIR3DS1 directs NK cell-mediated protection against human adenovirus infections. Sci Immunol [Internet]. 2021 Sep 17 [cited 2022 Dec 23];6(63). Available from: https://pubmed.ncbi.nlm.nih.gov/34533978/

182. Paduch M, Kossiakoff AA. Generating Conformation and Complex-Specific Synthetic Antibodies. Methods Mol Biol Clifton NJ. 2017;1575:93–119.

183. McWilliam HEG, Mak JYW, Awad W, Zorkau M, Cruz-Gomez S, Lim HJ, et al. Endoplasmic reticulum chaperones stabilize ligand-receptive MR1 molecules for efficient presentation of metabolite antigens. Proc Natl Acad Sci. 2020 Oct 6;117(40):24974–85.

184. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012 Aug 17;337(6096):816–21.

185. Partridge T, Nicastri A, Kliszczak AE, Yindom LM, Kessler BM, Ternette N, et al. Discrimination Between Human Leukocyte Antigen Class I-Bound and Co-Purified HIV-Derived Peptides in Immunopeptidomics Workflows. Front Immunol. 2018;9:912.

186. Jiang H, Ware R, Stall A, Flaherty L, Chess L, Pernis B. Murine CD8+ T cells that specifically delete autologous CD4+ T cells expressing V beta 8 TCR: a role of the Qa-1 molecule. Immunity. 1995 Feb;2(2):185–94.

187. Burian A, Wang KL, Finton KAK, Lee N, Ishitani A, Strong RK, et al. HLA-F and MHC-I Open Conformers Bind Natural Killer Cell Ig-Like Receptor KIR3DS1. PLoS One. 2016 Sep;11(9):e0163297.

188. Schoenborn JR, Wilson CB. Regulation of Interferon-γ During Innate and Adaptive Immune Responses. In: Advances in Immunology [Internet]. Academic Press; 2007 [cited 2022 Dec 23]. p. 41–101. Available from: https://www.sciencedirect.com/science/article/pii/S0065277607960022

189. Gobin SJP, Zutphen M van, Woltman AM, Elsen PJ van den. Transactivation of Classical and Nonclassical HLA Class I Genes Through the IFN-Stimulated Response Element. J Immunol. 1999 Aug 1;163(3):1428–34.

190. Matko J, Bushkin Y, Wei T, Edidin M. Clustering of class I HLA molecules on the surfaces of activated and transformed human cells. J Immunol Baltim Md 1950. 1994 Apr 1;152(7):3353–60.

191. Seitz C, Uchanska-Ziegler B, Zank A, Ziegler A. The monoclonal antibody HCA2 recognises a broadly shared epitope on selected classical as well as several non-classical HLA class I molecules. Mol Immunol. 1998 Sep;35(13):819–27.

192. Schnabl E, Stockinger H, Majdic O, Gaugitsch H, Lindley IJ, Maurer D, et al. Activated human T lymphocytes express MHC class I heavy chains not associated with beta 2-microglobulin. J Exp Med. 1990 May 1;171(5):1431–42.

193. Santos SG, Powis SJ, Arosa FA. Misfolding of major histocompatibility complex class I molecules in activated T cells allows cis-interactions with receptors and signaling molecules and is associated with tyrosine phosphorylation. J Biol Chem. 2004 Dec 17;279(51):53062–70.

194. Arosa FA, de Jesus O, Porto G, Carmo AM, de Sousa M. Calreticulin is expressed on the cell surface of activated human peripheral blood T lymphocytes in association with major histocompatibility complex class I molecules. J Biol Chem. 1999 Jun 11;274(24):16917–22.

195. Dutka P, Mukherjee S, Gao X, Kang Y, de Waal PW, Wang L, et al. Development of “plug and play” fiducial marks for structural studies of GPCR signaling complexes by single particle cryo-EM. Struct Lond Engl 1993. 2019 Dec 3;27(12):1862-1874.e7.

196. Concordet JP, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res. 2018 Jul 2;46(W1):W242–5.

197. Rojo R, Castro MJ, Martinez-Laso J, Serrano-Vela JI, Morales P, Moscoso J, et al. MHC-F DNA sequences in bonobo, gorilla and orangutan. Tissue Antigens. 2005 Oct;66(4):277–83.

198. An optimized co-immunoprecipitation protocol for the analysis of endogenous protein-protein interactions in cell lines using mass spectrometry: STAR Protocols [Internet]. [cited 2022 Dec 17]. Available from: https://star-protocols.cell.com/protocols/1489

199. Hanna J, Goldman-Wohl D, Hamani Y, Avraham I, Greenfield C, Natanson-Yaron S, et al. Decidual NK cells regulate key developmental processes at the human fetal-maternal interface. Nat Med. 2006 Sep;12(9):1065–74.

200. Mika KM, Li X, DeMayo FJ, Lynch VJ. An ancient fecundability-associated polymorphism creates a new GATA2 binding site in a distal enhancer of HLA-F. bioRxiv. 2018.

201. Gellersen B, Brosens J. Cyclic AMP and progesterone receptor cross-talk in human endometrium: a decidualizing affair. J Endocrinol. 2003 Sep;178(3):357–72.

202. Gellersen B, Brosens IA, Brosens JJ. Decidualization of the human endometrium: mechanisms, functions, and clinical perspectives. Semin Reprod Med. 2007 Nov;25(6):445–53.

203. Kohler PO, Bridson WE. Isolation of hormone-producing clonal lines of human choriocarcinoma. J Clin Endocrinol Metab. 1971 May;32(5):683–7.

204. Graham CH, Hawley TS, Hawley RG, MacDougall JR, Kerbel RS, Khoo N, et al. Establishment and characterization of first trimester human trophoblast cells with extended lifespan. Exp Cell Res. 1993 Jun;206(2):204–11.

205. Straszewski-Chavez SL, Abrahams VM, Alvero AB, Aldo PB, Ma Y, Guller S, et al. The isolation and characterization of a novel telomerase immortalized first trimester trophoblast cell line, Swan 71. Placenta. 2009 Nov;30(11):939–48.

206. Apps R, Murphy SP, Fernando R, Gardner L, Ahad T, Moffett A. Human leucocyte antigen (HLA) expression of primary trophoblast cells and placental cell lines, determined using single antigen beads to characterize allotype specificities of anti-HLA antibodies. Immunology. 2009 May;127(1):26–39.

207. Okae H, Toh H, Sato T, Hiura H, Takahashi S, Shirane K, et al. Derivation of Human Trophoblast Stem Cells. Cell Stem Cell. 2018 Jan 4;22(1):50-63.e6.

208. Fournel S, Huc X, Aguerre-Girr M, Solier C, Legros M, Praud-Brethenou C, et al. Comparative reactivity of different HLA-G monoclonal antibodies to soluble HLA-G molecules. Tissue Antigens. 2000;55(6):510–8.

209. Pastuschek J, Nonn O, Gutiérrez-Samudio RN, Murrieta-Coxca JM, Müller J, Sanft J, et al. Molecular characteristics of established trophoblast-derived cell lines. Placenta. 2021 May 1;108:122–33.

210. Qiu P. Embracing the dropouts in single-cell RNA-seq analysis. Nat Commun. 2020 Mar 3;11(1):1169.

211. Mika K, Marinić M, Singh M, Muter J, Brosens JJ, Lynch VJ. Evolutionary transcriptomics implicates new genes and pathways in human pregnancy and adverse pregnancy outcomes. Rokas A, Perry GH, Stevens A, Wildman DE, Mesiano S, editors. eLife. 2021 Oct 8;10:e69584.

212. Owen MJ, Kissonerghis AM, Lodish HF, Crumpton MJ. Biosynthesis and maturation of HLA-DR antigens in vivo. J Biol Chem. 1981 Sep 10;256(17):8987–93.

213. Mika KM, Li X, DeMayo FJ, Lynch VJ. An Ancient Fecundability-Associated Polymorphism Creates a GATA2 Binding Site in a Distal Enhancer of HLA-F. Am J Hum Genet. 2018 Oct;103(4):509–21.

214. Tabanelli S, Tang B, Gurpide E. In vitro decidualization of human endometrial stromal cells. J Steroid Biochem Mol Biol. 1992 May;42(3–4):337–44.

215. Samalecos A, Reimann K, Wittmann S, Schulte HM, Brosens JJ, Bamberger AM, et al. Characterization of a novel telomerase-immortalized human endometrial stromal cell line, St-T1b. Reprod Biol Endocrinol RBE. 2009 Jul 20;7:76.

216. Krikun G, Mor G, Alvero A, Guller S, Schatz F, Sapi E, et al. A novel immortalized human endometrial stromal cell line with normal progestational response. Endocrinology. 2004 May;145(5):2291–6.

217. Sakabe NJ, Aneas I, Knoblauch N, Sobreira DR, Clark N, Paz C, et al. Transcriptome and regulatory maps of decidua-derived stromal cells inform gene discovery in preterm birth. Sci Adv. 2020 Dec 2;6(49):eabc8696.

218. Yockey LJ, Iwasaki A. Interferons and Proinflammatory Cytokines in Pregnancy and Fetal Development. Immunity. 2018 Sep 18;49(3):397–412.

219. Griffith OW, Chavan AR, Protopapas S, Maziarz J, Romero R, Wagner GP. Embryo implantation evolved from an ancestral inflammatory attachment reaction. Proc Natl Acad Sci U A. 2017 Aug;114(32):E6566–75.

220. Ashkar AliA, Croy BA. Interferon-γ Contributes to the Normalcy of Murine Pregnancy1. Biol Reprod. 1999 Aug 1;61(2):493–502.

221. Bourdiec A, Calvo E, Rao CV, Akoum A. Transcriptome Analysis Reveals New Insights into the Modulation of Endometrial Stromal Cell Receptive Phenotype by Embryo-Derived Signals Interleukin-1 and Human Chorionic Gonadotropin: Possible Involvement in Early Embryo Implantation. PLOS ONE. 2013 May 22;8(5):e64829.

222. Shuya LL, Menkhorst EM, Yap J, Li P, Lane N, Dimitriadis E. Leukemia Inhibitory Factor Enhances Endometrial Stromal Cell Decidualization in Humans and Mice. PLOS ONE. 2011 Sep 23;6(9):e25288.

223. Garcia-Alonso L, Handfield LF, Roberts K, Nikolakopoulou K, Fernando RC, Gardner L, et al. Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro. Nat Genet. 2021 Dec;53(12):1698–711.

224. Haider S, Lackner AI, Dietrich B, Kunihs V, Haslinger P, Meinhardt G, et al. Transforming growth factor-β signaling governs the differentiation program of extravillous trophoblasts in the developing human placenta. Proc Natl Acad Sci. 2022 Jul 12;119(28):e2120667119.

225. Vondra S, Kunihs V, Eberhart T, Eigner K, Bauer R, Haslinger P, et al. Metabolism of cholesterol and progesterone is differentially regulated in primary

trophoblastic subtypes and might be disturbed in recurrent miscarriages. J Lipid Res. 2019 Nov 1;60(11):1922–34.

226. Goehring A, Lee CH, Wang KH, Michel JC, Claxton DP, Baconguis I, et al. Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. Nat Protoc. 2014 Nov;9(11):2574–85.

227. Bodmer WF. Evolutionary Significance of the HL-A System. Nature. 1972 May;237(5351):139–45.

228. Stöckel J, Meinl E, Hahnel C, Malotka J, Seitz R, Drexler K, et al. Refolding of human class II major histocompatibility complex molecules isolated from Escherichia coli. Assembly of peptide-free heterodimers and increased refolding-yield in the presence of antigenic peptide. J Biol Chem. 1994 Nov 25;269(47):29571–8.

229. Altman JD, Reay PA, Davis MM. Formation of functional peptide complexes of class II major histocompatibility complex proteins from subunits produced in Escherichia coli. Proc Natl Acad Sci U S A. 1993 Nov 1;90(21):10330–4.

230. Cameron TO, Norris PJ, Patel A, Moulon C, Rosenberg ES, Mellins ED, et al. Labeling antigen-specific CD4(+) T cells with class II MHC oligomers. J Immunol Methods. 2002 Oct 1;268(1):51–69.